

# Computing the Radius of Positive Semidefiniteness of a Multivariate Real Polynomial Via a Dual of Seidenberg's Method\*

Sharon Hutton  
Dept. of Mathematics,  
NCSU  
Raleigh, North Carolina  
27695-8205, USA  
sehutton@ncsu.edu

Erich L. Kaltofen  
Dept. of Mathematics, NCSU  
Raleigh, North Carolina  
27695-8205, USA  
kaltofen@math.ncsu.edu  
http://www.kaltofen.us

Lihong Zhi  
Key Laboratory of Mathematics  
Mechanization, AMSS  
Beijing 100190, China  
lzhi@mmrc.iss.ac.cn  
http://www.mmrc.iss.ac.cn/~lzhi

## ABSTRACT

We give a stability criterion for real polynomial inequalities with floating point or inexact scalars by estimating from below or computing the radius of semidefiniteness. That radius is the maximum deformation of the polynomial coefficient vector measured in a weighted Euclidean vector norm within which the inequality remains true. A large radius means that the inequalities may be considered numerically valid.

The radius of positive (or negative) semidefiniteness is the distance to the nearest polynomial with a real root, which has been thoroughly studied before. We solve this problem by parameterized Lagrangian multipliers and Karush-Kuhn-Tucker conditions. Our algorithms can compute the radius for several simultaneous inequalities including possibly additional linear coefficient constraints. Our distance measure is the weighted Euclidean coefficient norm, but we also discuss several formulas for the weighted infinity and 1-norms.

The computation of the nearest polynomial with a real root can be interpreted as a dual of Seidenberg's method that decides if a real hypersurface contains a real point. Sums-of-squares rational lower bound certificates for the radius of semidefiniteness provide a new approach to solving Seidenberg's problem, especially when the coefficients are numeric. They also offer a surprising alternative sum-of-squares proof for those polynomials that themselves cannot be represented by a polynomial sum-of-squares but that have a positive distance to the

nearest indefinite polynomial.

## Categories and Subject Descriptors

I.1.2 [Symbolic and Algebraic Manipulation]: Algorithms; G.1.6 [Numerical Analysis]: Global optimization

## General Terms

algorithms, experimentation

## Keywords

numeric polynomial inequality, nearest polynomial with a real root, sum-of-squares, approximate polynomial systems, semidefinite programming

## 1. INTRODUCTION

### 1.1 Motivation

Real polynomial or rational function global optimization is equivalent to establishing a polynomial inequality: the infimum  $\mu \in \mathbb{R}$  of a polynomial  $f \in \mathbb{R}[x_1, \dots, x_n]$  satisfies  $f(\xi_1, \dots, \xi_n) - \mu \geq 0$  for all  $x_i = \xi_i \in \mathbb{R}$ . In other words, the polynomial  $f - \mu$  is positive semidefinite. For univariate  $f$  ( $n = 1$ ) Sturm sequences [9] yield an efficient algorithm for deciding semidefiniteness. The bivariate case  $n = 2$  can be solved by Seidenberg's [26] algorithm (see also [9] and [11]), which is generalized to arbitrarily many variables via Lagrangian multipliers in [1, 25] or used in nonstandard decision methods [29]. Alternatively, one can use Artin's theorem of sum-of-squares and semidefinite programming (see, e.g., [12, 14]).

Here we consider the more difficult situation when the coefficients of  $f$  are not exactly known, which is the case when  $f$  is the result of an empirical measurement or a computation with floating point numbers. As a simple example consider Figure 1 below.

The middle polynomial  $f(x) = \frac{1}{3}(x-1)^2$  has a double real root at  $x = 1$ . In fact, it is the nearest polynomial with a real root to the polynomial  $x^2 + 1$  under the infinity norm [8]:  $\|f(x) - (x^2 + 1)\|_\infty = \frac{2}{3}$ ,

\*This material is based on work supported in part by the National Science Foundation under Grants CCF-0830347, CCF-0514585, DMS-0532140 and OISE-0913588 (Hutton and Kaltofen).

Lihong Zhi is supported by the Chinese National Natural Science Foundation under Grants 60821002/F02, 60911130369 and 10871194.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ISSAC 2010, 25–28 July 2010, Munich, Germany.

Copyright 2010 ACM 978-1-4503-0150-3/10/0007 ...\$10.00.

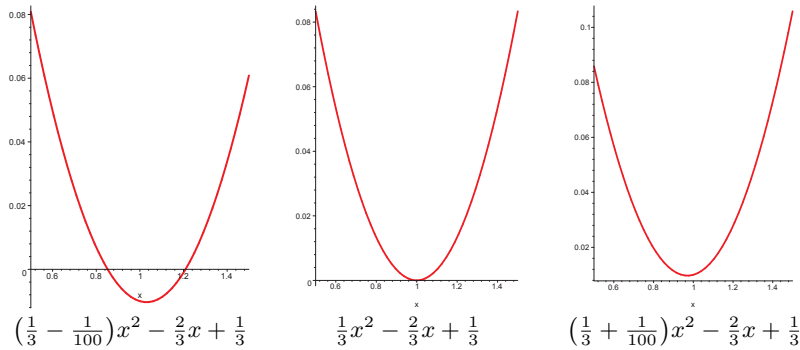


Figure 1: Root sensitivity

where for a polynomial  $g$  the norm  $\|g\|_\infty$  is the maximum of the absolute values of the coefficients of  $g$  (the height of  $g$ ). Small perturbations in the leading coefficient (one could also perturb the constant coefficient) make the polynomial  $f$  either indefinite (left polynomial in Figure 1, the polynomial changes sign) or positive definite (right polynomial). Therefore, the right polynomial, although positive definite, as an approximate polynomial is not numerically positive because a small change in its coefficients can make the polynomial indefinite. As in Kharitonov’s [16] interval polynomial stability criterion, we seek to compute by how much the coefficients in a polynomial can be deformed while still preserving nonnegativity. This distance is the coefficient vector norm distance to the nearest polynomial with a real root, which we shall call the *radius of positive semidefiniteness*. Note that there may not exist an affine optimizer—hence radius of positive semidefiniteness rather than distance to the nearest polynomial with a real root: the polynomial  $(1 - \epsilon^2)x^2 + y^2 - 2xy + 4$  attains for any  $\epsilon > 0$  negative values at  $x = y > 2/\epsilon$ . Thus the polynomial  $x^2 + y^2 - 2xy + 4$  has a radius of positive semidefiniteness = 0 although its global minimum is 4.

## 1.2 Results and Used Approach

We follow the approach by Karmarkar and Lakshman [15] (see also [3]) which first fixes a real root  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$  and gives a rational function  $\mathcal{N}(\alpha_1, \dots, \alpha_n)$  in the indeterminate  $\alpha$ ’s for the minimal distance from the given  $f$  to the nearest polynomial  $\tilde{f}$  with  $\tilde{f}(\alpha_1, \dots, \alpha_n) = 0$ . One then can compute the infimum of  $\mathcal{N}(\alpha_1, \dots, \alpha_n)$  over all real  $\alpha$ ’s. The case  $n \geq 2$  is from [27].

We rederive the multivariate formula for  $\mathcal{N}(\alpha_1, \dots, \alpha_n)$  in [27], for weighted  $\ell^2$  distance norms, by the method of Lagrangian multipliers. The weighted norms subsume the fixing of coefficients in [24] (see [4, Section 2.12.3.2.6] and Remark 5 below). Our approach also allows us to introduce linear constraints on the coefficients of  $\tilde{f}$ , as is done in [13] for the approximate GCD problem. Linear equality constraints generalize sparsity, which are equations of the form  $c_i = 0$ . Because the Jacobian of the Lagrange function remains linear in the control variables and multipliers, determinantal formulas parametric in the real root coordinates can be computed. Linear inequality constraints on the coefficients of  $\tilde{f}$ , for instance nonnegativity ( $c_i \geq 0$ ) can now be imposed via Karush-Kuhn-Tucker (KKT) conditions (see, e.g., [5]) and the

arising systems solved via linear programming, at least for a fixed real root. Parametric root coordinates or non-linear constraints necessitate non-linear techniques on the Lagrange and KKT conditions and are therefore in general of much higher computational complexity. Our approach allows multiple simultaneous  $f$ ’s and complex coefficients without modification.

Our solution can be interpreted as a dual of Seidenberg’s test whether a real hypersurface  $f$  has a real point. Seidenberg’s algorithm (and Safey El Din’s generalization) computes to a given real point in  $\mathbb{R}^n$  the nearest real point on  $f$  in terms of Euclidean distance. If  $f$  has no real solution the tangent equations have no real solutions. Our algorithm computes the nearest surface (in terms of coefficient norm) that has a real point. If  $f$  has a real point, the nearest surface is  $f$  itself. However, if a lower bound on the radius of semidefiniteness for any weight vector is  $> 0$ ,  $f$  has no real point, even when the coefficients of  $f$  are approximate. The latter can be certified by a sum-of-squares of rational functions, which leads to an entirely new verification that  $f$  is definite, i.e., has no real point, with possibly a very short certificate.

Polynomials with a radius of positive semidefiniteness  $> 0$  are quite special. Our Example 6 below demonstrates that a positive polynomial that is not a sum-of-squares of polynomials can have a lower bound certificate for the radius of positive semidefiniteness that is in fact a sum-of-squares of polynomials, which implies positive semidefiniteness of the polynomial itself. For such polynomials, sum-of-squares denominators in Artin-style certificates may never become necessary (see our conjecture at the end of Section 5).

## 1.3 Related Previous Results

Our method is conceptually that of hybrid symbolic-numeric computation such as computing approximate polynomial greatest common divisors and factorization.

Hitz and Kaltofen [7] derive Lakshman’s and Karmarkar’s formula for univariate  $f$  by a least square fit for the cofactor  $f(x)/(x - \alpha)$  and introduce linear equality constraints on the deformed coefficients. Zhi, Wu, Noda, Kai, Rezvan and Corless [31, 30, 23] generalize the formula to roots with given multiplicities. In [8]  $\ell^\infty$ -norm distances are introduced and Markus Hitz in the Summer of 1999 considered dual  $\ell^p$ -norms. Stetter [27] then generalized Lakshman and Karmarkar’s formula to

an arbitrary number of variables and dual  $\ell^p$  norm distances via Hölders inequality.

In [24, 21] Stetter's multivariate (complex) formula is applied to the important problem of computing the nearest consistent polynomial system, with zeros of a minimum given multiplicity, and a different proof via generalized Lagrangian interpolation is given. We observe that the  $\ell^\infty$ -norm formulas apply to the problem of consistent systems as well (see Theorem 7 below). In our setting, we determine the smallest deformation where all inequalities are simultaneously violated.

A related result [8] computes the nearest matrix in Frobenius norm that has a real eigenvalue. Sum-of-squares rational lower bound certificates were introduced in [12] to overcome the high algebraic degree in the exact real algebraic minima.

## 2. RADIUS OF POSITIVE SEMIDEFINITENESS

**Definition 1** Let  $w \in \mathbb{R}_{>0}^n$  be a vector of positive weights. For  $x = [x_1, \dots, x_n]^T \in \mathbb{R}^n$  the weighted  $\ell^2$  norm is

$$\|x\|_{2,w} = \sqrt{w_1 x_1^2 + \dots + w_n x_n^2}.$$

**Definition 2** Let  $\alpha = [\alpha_1, \dots, \alpha_n] \in \mathbb{R}^n$  be a prescribed real root and  $w \in \mathbb{R}_{>0}^n$  a weight vector. The distance to the nearest polynomial with a real root  $\alpha$  is defined as

$$\mathcal{N}_{2,w}^{[f]}(\alpha) = \left. \begin{array}{l} \inf_{\tilde{f} \in \mathbb{R}[x_1, \dots, x_n]} \|f - \tilde{f}\|_{2,w}^2 \\ \text{s. t. } \tilde{f}(\alpha) = 0, \\ \deg(\tilde{f}) \leq \deg(f). \end{array} \right\} (1)$$

If  $f$  and the used norm is clear from the context, we may write  $\mathcal{N}(\alpha)$  for the above infimum, which is actually a minimum (see Theorem 1 below).

**Theorem 1** Let  $f \in \mathbb{R}[x_1, \dots, x_n]$ ,

$$f(x_1, \dots, x_n) = \sum_{i_1 + \dots + i_n = 0}^d f_{i_1, \dots, i_n} x_1^{i_1} \dots x_n^{i_n}.$$

For  $\tau = [1, \alpha_1, \dots, \alpha_n, \dots, \alpha_1^{i_1} \dots \alpha_n^{i_n}, \dots]^T$ , the vector of possible term values in  $\tilde{f}$ , the distance to the nearest polynomial with a real root  $\alpha$  is

$$\mathcal{N}_{2,w}^{[f]}(\alpha) = \frac{f(\alpha)^2}{\tau^T D_w^{-1} \tau}. \quad (2)$$

Furthermore, the coefficient vector  $\vec{f}$ , for the polynomial  $\tilde{f}$  as in (1) is

$$\vec{f} = \tilde{f} - \frac{\tau^T \tilde{f}}{\tau^T D_w^{-1} \tau} D_w^{-1} \tau, \quad (3)$$

where  $\tilde{f}$  is the coefficient vector of  $f$  and  $D_w$  is a diagonal matrix of the weights. The polynomial  $\tilde{f}$  is the only polynomial that attains the infimum (2).

**Remark 1** The infimum

$$\rho_{2,w}(f) = \inf_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha) \quad (4)$$

is the unconstrained radius of positive semidefiniteness. Within any  $\epsilon$  of the radius (4) there is a polynomial that attains negative values: for any  $\epsilon > 0$  there is an  $\tilde{f}_\epsilon$  with a real root  $\alpha$  and  $\|f - \tilde{f}_\epsilon\|_{2,w}^2 < \rho_{2,w}(f) + \epsilon/2$ . Then  $(\tilde{f}_\epsilon - \delta)(\alpha) < 0$  for all  $\delta > 0$ , and in particular if  $w_1 \delta^2 < \epsilon/2$  we have

$$\|f - (\tilde{f}_\epsilon - \delta)\|_{2,w}^2 < \rho_{2,w}(f) + \epsilon.$$

In Section 4 we will permit constraints for the coefficients of  $\tilde{f}$ . Then a negative evaluation may be impossible: e.g.,  $\tilde{f}(x, y) = f_{2,0} x^2 + f_{0,2} y^2$  and  $f_{2,0} \geq 0$ ,  $f_{0,2} \geq 0$ . However, within less of the distance to the nearest polynomial with a real root a deformed  $\tilde{f}$  remains positive definite.

**Remark 2** If the weighted norm is the Euclidean norm then the formula becomes

$$\mathcal{N}_2(\alpha) = \frac{f(\alpha)^2}{\sum_{i_1 + \dots + i_n = 0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n}}. \quad (5)$$

**Remark 3** Different degree conditions in (5) give different rational functions. For example, if the individual variable degrees are bounded by  $d$ ,  $\deg_{x_j}(f) \leq d$  for all  $j$  with  $1 \leq j \leq n$ , then for the  $\ell^2$  norm

$$\mathcal{N}_2(\alpha) = \frac{f(\alpha)^2}{\sum_{i_1=0}^d \dots \sum_{i_n=0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n}}.$$

Comparing the denominators, we have

$$\sum_{i_1=0}^d \dots \sum_{i_n=0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n} \geq \sum_{i_1 + \dots + i_n = 0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n},$$

so

$$\begin{aligned} \inf \frac{f(\alpha)^2}{\sum_{i_1=0}^d \dots \sum_{i_n=0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n}} \\ \leq \inf \frac{f(\alpha)^2}{\sum_{i_1 + \dots + i_n = 0}^d \alpha_1^{2i_1} \dots \alpha_n^{2i_n}}, \end{aligned}$$

which must be since we optimize over a larger set of  $\tilde{f}$ .

**Remark 4** Theorem 1 can be generalized to a complex root  $\alpha$  and real/complex coefficients for  $f$ , which is the original setting of [3, 15, 27, 24]. Let  $f \in \mathbb{C}[x_1, \dots, x_n]$ ,  $\tau$  be as described in Theorem 1 then the distance to the nearest polynomial with root  $\alpha \in \mathbb{C}^n$  is

$$\mathcal{N}_2^{[f]}(\alpha) = \frac{(\bar{f}(\bar{\alpha}))(f(\alpha))}{\tau^H \tau}.$$

Here  $^H$  denotes the Hermitian transposed and  $\bar{\cdot}$  complex conjugation. Furthermore, the coefficient vector  $\vec{f}$ , for the polynomial  $\tilde{f}$  as in (1) is

$$\vec{f} = \tilde{f} - \frac{\tau^T \tilde{f}}{\tau^H \tau} \bar{\tau}.$$

Our proof and possible inclusion of weights for the real and imaginary parts is similar to the proof of Theorem 1.

**Remark 5** Theorem 1 is the real case of the theorems in [24] and [21] for complex roots. However, they use generalized Lagrangian interpolation for their proof. They also allow keeping selected coefficients of  $f$  as their input values and only deform the others in  $\tilde{f}$ , thus preserving sparsity or monicity, for instance. Our Theorem 1 has theirs as an immediate corollary by setting the weights of those coefficients to  $\infty$  in the limit. However, the problem may become ill-posed. If  $f$  has a nonzero constant coefficient which is fixed, and  $\alpha = 0$ , the set of  $\tilde{f}$  is empty. In Section 4 we generalize our approach to handle arbitrary linear constraints on the coefficients of  $\tilde{f}$ .

If a weight  $w_i \rightarrow 0$  in the limit then the corresponding coefficient in  $\tilde{f}$  becomes a “don’t care” deformation, i.e., any change in that coefficient is not taken into account in the distance measure. The “nearest” polynomial  $\tilde{f}$  with  $\alpha \in (\mathbb{R} \setminus \{0\})^n$  as a root then has distance 0, namely  $\tilde{f}(x) = f(x) - (f(\alpha)/\alpha^i) x^i$ , unless there are additional constraints on the coefficients of  $\tilde{f}$  in effect.

**Remark 6** The formulas in [7] and [15] use the weights  $w_i$  in the denominator of (2), not correctly their reciprocals  $1/w_i$ .

PROOF OF THEOREM 1. Let  $\vec{f}$ ,  $\tau$ , and  $f$  be as above. Denote the coefficients of  $\tilde{f}$  in (1) by

$$\tilde{f}(x_1, \dots, x_n) = \sum_{i_1 + \dots + i_n = 0}^d \tilde{f}_{i_1, \dots, i_n} x_1^{i_1} \cdots x_n^{i_n}.$$

Let  $\vec{f}$  be the coefficient vector of  $\tilde{f}$ . Also,  $\tilde{f}(\alpha_1, \dots, \alpha_n) = \tau^T \vec{f} = 0$ . We have  $\|f - \tilde{f}\|_{2,w} = (\vec{f} - \tilde{f})^T D_w (\vec{f} - \tilde{f})$ , the weighted  $\ell^2$  norm, where  $D_w$  is a diagonal matrix of the weights. Let  $\lambda$  be the Lagrange multiplier and

$$L(\alpha_1, \dots, \alpha_n, \lambda) = (\vec{f} - \tilde{f})^T D_w (\vec{f} - \tilde{f}) + \lambda \tau^T \vec{f}$$

the Lagrange function of our constrained optimization problem. We must check that  $\alpha$  is a regular point (i.e., the gradient of the constraint is not 0 at  $\alpha$ ). Since  $\nabla \tilde{f}(\alpha) \neq 0$  if  $\tau \neq 0$  then  $\alpha$  is a regular point as long as  $\alpha \neq 0$ . In the case  $\alpha = 0$  the constant coefficient of  $f$  is deformed to 0 and the formulas hold. The Jacobian of  $L$  w.r.t.  $\vec{f}$  and  $\lambda$  is

$$J_L = \begin{bmatrix} \vdots \\ \frac{\partial L}{\partial (\vec{f})_i} \\ \vdots \\ \frac{\partial L}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} -2D_w(\vec{f} - \tilde{f}) + \lambda\tau \\ \tau^T \vec{f} \end{bmatrix}.$$

Looking at the first block of the vector we have

$$-2D_w(\vec{f} - \tilde{f}) + \tau\lambda = -2D_w\vec{f} + 2D_w\tilde{f} + \tau\lambda = 0. \quad (6)$$

Multiplying by  $\tau^T D_w^{-1}$  we have

$$-2\tau^T D_w^{-1} D_w \vec{f} + 2\tau^T D_w^{-1} D_w \tilde{f} + \tau^T D_w^{-1} \tau \lambda = 0.$$

Recalling that  $\tilde{f}(\alpha_1, \dots, \alpha_n) = 0$  which means that  $\tau^T \tilde{f} = 0$  then we have

$$-2\tau^T D_w^{-1} D_w \vec{f} + 2\tau^T D_w^{-1} D_w \tilde{f} + \tau^T D_w^{-1} \tau \lambda = -2\tau^T \vec{f} + \tau^T D_w^{-1} \tau \lambda = 0.$$

Solving for  $\lambda$  we get  $\lambda = \frac{2\tau^T \vec{f}}{\tau^T D_w^{-1} \tau}$ . Looking at (6), we have  $\vec{f} - \tilde{f} = \frac{D_w^{-1} \tau \lambda}{2}$ . Substituting in for  $\lambda$  we obtain as the only solution  $\vec{f} - \tilde{f} = \frac{\tau^T \vec{f}}{\tau^T D_w^{-1} \tau} D_w^{-1} \tau$ . Finally,

$$\begin{aligned} \mathcal{N}_{2,w}^{[f]}(\alpha) &= \left( \frac{\tau^T \vec{f}}{\tau^T D_w^{-1} \tau} D_w^{-1} \tau \right)^T D_w \left( \frac{\tau^T \vec{f}}{\tau^T D_w^{-1} \tau} D_w^{-1} \tau \right) \\ &= \frac{\vec{f}^T \tau \tau^T D_w^{-1} \tau \tau^T \vec{f}}{\tau^T D_w^{-1} \tau \tau^T D_w^{-1} \tau}. \end{aligned}$$

So  $\mathcal{N}_{2,w}^{[f]}(\alpha) = \frac{f(\alpha)^2}{\tau^T D_w^{-1} \tau}$ .  $\square$

**Example 1** Here we give another example for the case that the infimum in (4) is not always attainable. Our first example was given at the end of Section 1.1. Consider the polynomial

$$f(x, y) = 1 - 2xy + x^2 y^2 + x^2 = (1 - xy)^2 + x^2.$$

We have that

$$\mathcal{N}_2(\alpha, \beta) = \frac{((1 - \alpha\beta)^2 + \alpha^2)^2}{\sum_{i+j=0}^4 \alpha^{2i} \beta^{2j}}.$$

Then  $\inf_{\alpha, \beta} \mathcal{N}_2(\alpha, \beta) = 0$ . Suppose now that there exists  $\alpha, \beta$  such that the numerator is 0. Then  $(1 - \alpha\beta) = 0$  and  $\alpha = 0$ . But if  $\alpha = 0$  then  $\alpha\beta = 0$ . Then  $1 - \alpha\beta \neq 0$ . Contradiction. Thus  $f$  does not have a real root and the infimum is not attainable.

We have

$$\mathcal{N}_2(\epsilon, \frac{1}{\epsilon}) = \frac{\epsilon^4}{\delta}, \delta = 3 + 2\epsilon^2 + \frac{2}{\epsilon^2} + 2\epsilon^4 + \frac{2}{\epsilon^4} + \epsilon^6 + \frac{1}{\epsilon^6} + \epsilon^8 + \frac{1}{\epsilon^8},$$

and the nearest polynomial to  $f$  with  $(\alpha, \beta) = (\epsilon, 1/\epsilon)$  as its root is

$$\begin{aligned} \tilde{f}(x, y) &= -\frac{\epsilon^6}{\delta} x^4 - \frac{\epsilon^4}{\delta} x^3 y + (1 - \frac{\epsilon^2}{\delta}) x^2 y^2 - \frac{1}{\delta} x y^3 - \frac{1}{\epsilon^2 \delta} y^4 \\ &\quad - \frac{\epsilon^5}{\delta} x^3 - \frac{\epsilon^3}{\delta} x^2 y - \frac{\epsilon}{\delta} x y^2 - \frac{1}{\epsilon \delta} y^3 + (1 - \frac{\epsilon^4}{\delta}) x^2 \\ &\quad - (2 + \frac{\epsilon^2}{\delta}) x y - \frac{1}{\delta} y^2 - \frac{\epsilon^3}{\delta} x - \frac{\epsilon}{\delta} y + 1 - \frac{\epsilon^2}{\delta}. \end{aligned}$$

Note that  $f(\epsilon, 1/\epsilon) - \epsilon^2 = 0$  has squared distance  $\epsilon^4$  from  $f$ , which is larger than  $\epsilon^4/3 > \epsilon^4/\delta$  for all  $\epsilon \neq 0$ .  $\square$

### 3. INFINITY AND ONE NORM

Theorem 1 can be generalized to include the  $\ell^1$ -norm,  $\ell^\infty$ -norm, and any  $\ell^p$ -norm. We discuss in more detail the results presented in [27, 4].

**Definition 3** We consider  $\mathbb{C}^n$  equipped with some norm  $\|\dots\|$ . The associated dual norm or operator norm  $\|\dots\|^*$  for the column vector  $v \in \mathbb{C}^n$  is defined by

$$\|v^T\|^* = \sup_{u \neq 0} \frac{|v^T u|}{\|u\|} = \sup_{\|u\|=1} |v^T u|.$$

Since we are taking the supremum over a compact domain, the maximum value is attained.

**Proposition 1 (Proposition 1 in [27])** For each  $u \in \mathbb{C}^n$ , with  $\|u\| = 1$ , there exist vectors  $v \in \mathbb{C}^n$ , with  $\|v^T\|^* = 1$ , such that  $|v^T u| = 1$ .  $\square$

It is well known that with  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $1 \leq p, q \leq \infty$ ,

$$\|\dots\| = \ell^p\text{-norm} \Leftrightarrow \|\dots\|^* = \ell^q\text{-norm}.$$

**Theorem 2** (see [27]) *Let  $\alpha = [\alpha_1, \dots, \alpha_n] \in \mathbb{C}^n$ ,  $\tau = [1, \alpha_1, \dots, \alpha_n, \dots, \alpha_1^{i_1} \dots \alpha_n^{i_n}, \dots]$ , the vector of possible term values of  $f$  with norm  $\|\dots\|$ , and  $\vec{f}, \tilde{f} \in \mathbb{C}^n$ , with dual norm  $\|\dots\|^*$ .  $\tilde{f}(\alpha) = \tau^T \vec{f} = 0$  requires*

$$\|\vec{f} - \tilde{f}\|^* \geq \frac{|f(\alpha)|}{\|\tau\|}.$$

Theorem 2 shows that Theorem 1 can be extended to any  $\ell^p$ -norm. We extend the results from Theorem 2 to the weighted  $\ell^1$  and  $\ell^\infty$ -norms. We prove Hölder's inequality for weighted  $\ell^1$ ,  $\ell^2$  and  $\ell^\infty$ -norms, which allows us to then follow the same proof as in [27] for Theorem 2. We further give an explicit formula for  $\tilde{f}$ .

**Theorem 3** *Let  $u, v \in \mathbb{C}^n$  and weights  $w_i$ . Then  $|v^T u| \leq \|u\|_{\infty, w} \|v\|_{1, 1/w}$ , where  $1/w$  is the vector of reciprocals of entries of  $w$ .*

PROOF. Looking at  $|v^T u|$ :

$$\begin{aligned} |v^T u| &= \left| \sum_i v_i u_i \right| = \left| \sum_i w_i v_i \frac{1}{w_i} u_i \right| \\ &\leq (\max_j w_j |v_j|) \sum_i \frac{1}{w_i} |u_i| = \|v\|_{\infty, w} \|u\|_{1, 1/w}. \quad \square \end{aligned}$$

**Corollary 1** *Let  $u, v \in \mathbb{C}^n$ , and weights  $w_i$ . Then*

$$|v^T u| \leq \|u\|_{1, w} \|v\|_{\infty, 1/w}.$$

**Theorem 4** *Let  $u, v \in \mathbb{C}^n$  and weights  $w_i$ . Then  $|v^T u| \leq \|u\|_{2, w} \|v\|_{2, 1/w}$*

PROOF OF THEOREM 4. Let  $\hat{u}_i = \sqrt{w_i} u_i$ ,  $\hat{v}_i = \frac{v_i}{\sqrt{w_i}}$ . Using the Cauchy-Schwartz inequality, we have:  
 $|v^T u| = |\hat{v}^T \hat{u}| \leq \left( \sum_i (\sqrt{w_i} u_i)^2 \right)^{1/2} \left( \sum_i \left( \frac{v_i}{\sqrt{w_i}} \right)^2 \right)^{1/2}$ .  
Therefore,  $|v^T u| \leq \|u\|_{2, w} \|v\|_{2, 1/w}$ .  $\square$

Now that we have proven Hölder's inequality for the weighted  $\ell^1$ ,  $\ell^2$  and  $\ell^\infty$ -norms, we can follow the proof of Theorem 2 to extend Theorem 1 to weighted  $\ell^1$  and  $\ell^\infty$ -norms. Theorem 4 would also yield an alternative proof of Theorem 1.

**Theorem 5** *For  $\tau, f$ , and  $\tilde{f}$  as described in Theorem 2,  $v = [1, \text{sgn}(\tau_i), \dots]$ , where  $v \in \mathbb{R}^\kappa$ ,  $\kappa$  is the dimension of  $f$  and  $\text{sgn}(\tau_i) = 1$  if  $\tau_i > 0$ ,  $\text{sgn}(\tau_i) = -1$  if  $\tau_i < 0$ , and  $\text{sgn}(\tau_i) = 0$  if  $\tau_i = 0$  with weighted  $\ell^\infty$ -norm and weights  $w_i$  then*

$$\mathcal{N}_{\infty, w}^{[f]}(\alpha) = \frac{|f(\alpha)|}{\|\tau\|_{1, 1/w}} \text{ and } \vec{f} = \tilde{f} - \frac{f(\alpha)}{\|\tau\|_{1, 1/w}} D_w^{-1} v.$$

PROOF OF THEOREM 5. From [27] and Theorem 3 we know that  $|f(\alpha)| = |(\vec{f} - \tilde{f})^T \tau| \leq \|\vec{f} - \tilde{f}\|_{\infty, w} \|\tau\|_{1, 1/w}$ . Therefore,  $\frac{|f(\alpha)|}{\|\tau\|_{1, 1/w}} \leq \|\vec{f} - \tilde{f}\|_{\infty, w}$ . For all  $j$  choose  $(\vec{f})_j$  such that  $\frac{f(\alpha)}{\|\tau\|_{1, 1/w}} = w_j (\vec{f} - \tilde{f})_j$ . From this we

get that  $w_j (\tilde{f})_j = w_j (\vec{f})_j - \frac{f(\alpha)}{\|\tau\|_{1, 1/w}}$ . Therefore,  $\tilde{f} = \vec{f} - \frac{f(\alpha)}{\|\tau\|_{1, 1/w}} D_w^{-1} v$ , which yields equality in the above inequality. This gives  $0 = \tilde{f}(\alpha) = \tau^T \vec{f} - \frac{f(\alpha)}{\|\tau\|_{1, 1/w}} \tau^T D_w^{-1} v$ .  $\square$

In the same way we obtain the following theorem.

**Theorem 6** *For  $\tau, f, \tilde{f}$ , and  $\text{sgn}(\tau_i)$  as described in Theorem 5 with weighted  $\ell^1$ -norm and weights  $w_i \geq 0$  then*

$$\mathcal{N}_{1, w}^{[f]}(\alpha) = \frac{|f(\alpha)|}{\|\tau\|_{\infty, 1/w}} \text{ and}$$

$$\vec{f}_i = \begin{cases} \tilde{f}_i & \text{for } i \neq j \\ \tilde{f}_i - \text{sgn}(\tau_i) \frac{f(\alpha)}{\|\tau\|_{\infty, 1/w}} \frac{1}{w_i} & \text{for } i = j \end{cases}$$

$$\text{where } \frac{|\tau_j|}{w_j} = \max_i \frac{|\tau_i|}{w_i}.$$

## 4. GENERALIZATIONS

Our method can be further generalized to include problems with linear constraints of the form  $H \vec{f} = p$ , where  $H \in \mathbb{R}^{t \times s}$ ,  $p \in \mathbb{R}^t$ , on the coefficient vector  $\vec{f}$  of  $\tilde{f}$ . We define

$$\mathcal{N}_{2, w}^{[f; H]}(\alpha) = \left. \begin{aligned} &\inf_{\tilde{f} \in \mathbb{R}[x_1, \dots, x_n]} \|f - \tilde{f}\|_{2, w}^2 \\ &\text{s. t. } \tilde{f}(\alpha) = 0, H \vec{f} = p, \\ &\quad \text{deg}(\tilde{f}) \leq \text{deg}(f). \end{aligned} \right\} \quad (7)$$

We note that the Jacobian of the Lagrange function corresponding to (7) constitutes a linear system in the unknown coefficients of  $\tilde{f}$  and the multipliers, hence a determinantal formula parameterized by the real root for the solution can be computed, which one can minimize.

**Example 2** Given a polynomial  $f(x, y) = x^2 + y^2 + 1$  find the nearest polynomial  $\tilde{f}(x, y) = \tilde{f}_{2,0} x^2 + \tilde{f}_{0,2} y^2 + \tilde{f}_{1,1} xy + \tilde{f}_{1,0} x + \tilde{f}_{0,1} y + \tilde{f}_{0,0}$  with  $\tilde{f}_{1,1} = \tilde{f}_{0,0}$  and  $\tilde{f}_{0,1} = 0$  and with the root  $(\alpha, \beta)$ . The Lagrangian function is

$$\begin{aligned} L(\alpha, \beta, \lambda) &= (\vec{f} - \tilde{f})^T (\vec{f} - \tilde{f}) \\ &\quad + \lambda_0 \tau^T \vec{f} + \lambda_1 (\tilde{f}_{1,1} - \tilde{f}_{0,0}) + \lambda_2 \tilde{f}_{0,1}, \end{aligned}$$

where the term vector  $\tau = [\alpha^2, \beta^2, \alpha\beta, \alpha, \beta, 1]$ . The Jacobian of  $L$  in  $\vec{f}$  and  $\lambda$  is zero for

$$\vec{f} = \begin{bmatrix} \tilde{f}_{2,0} \\ \tilde{f}_{0,2} \\ \tilde{f}_{1,1} \\ \tilde{f}_{1,0} \\ \tilde{f}_{0,1} \\ \tilde{f}_{0,0} \end{bmatrix} = \begin{bmatrix} -\frac{\alpha^2 - 2\beta^4 + \alpha^2 \beta^2 - 2\alpha\beta - 1 + \alpha^3 \beta}{2\alpha^2 + 2\beta^4 + 2\alpha^4 + \alpha^2 \beta^2 + 2\alpha\beta + 1} \\ \frac{2\alpha^2 + 2\alpha^4 - \alpha^2 \beta^2 + 2\alpha\beta + 1 - \beta^2 - \alpha\beta^3}{2\alpha^2 + 2\beta^4 + 2\alpha^4 + \alpha^2 \beta^2 + 2\alpha\beta + 1} \\ \frac{\beta^4 + \alpha^4 - \alpha\beta^3 - \alpha^3 \beta - \beta^2}{2\alpha^2 + 2\beta^4 + 2\alpha^4 + \alpha^2 \beta^2 + 2\alpha\beta + 1} \\ \frac{\alpha(1 + 2\beta^2 + \alpha\beta + 2\alpha^2)}{-2\alpha^2 + 2\beta^4 + 2\alpha^4 + \alpha^2 \beta^2 + 2\alpha\beta + 1} \\ 0 \\ \frac{\beta^4 + \alpha^4 - \alpha\beta^3 - \alpha^3 \beta - \beta^2}{2\alpha^2 + 2\beta^4 + 2\alpha^4 + \alpha^2 \beta^2 + 2\alpha\beta + 1} \end{bmatrix},$$

$$\lambda_0 = \frac{2(2\alpha^2 + \alpha\beta + 2\beta^2 + 1)}{2\alpha^4 + 2\alpha^2 + \alpha^2 \beta^2 + 2\alpha\beta + 1 + 2\beta^4},$$

$$\lambda_1 = \frac{-2(\alpha^4 + \alpha^3 \beta + \alpha^2 \beta^2 + \alpha\beta + \beta^3 \alpha + \beta^4 - \beta^2)}{2\alpha^4 + 2\alpha^2 + \alpha^2 \beta^2 + 2\alpha\beta + 1 + 2\beta^4},$$

$$\lambda_2 = \frac{-2(2\alpha^2 + \alpha\beta + 2\beta^2 + 1)\beta}{(2\alpha^4 + 2\alpha^2 + \alpha^2\beta^2 + 2\alpha\beta + 1 + 2\beta^4)}.$$

The minimum perturbation is

$$\mathcal{N}_2 = \frac{3\alpha^4 + 2\alpha^3\beta + 5\alpha^2\beta^2 + 3\alpha^2 + 2\alpha\beta + 2\alpha\beta^3 + 1 + 3\beta^4 + 2\beta^2}{2\alpha^2 + 2\alpha^4 + 2\beta^4 + \alpha^2\beta^2 + 2\alpha\beta + 1} \quad (8)$$

Running the `Minimize` procedure in Maple 13 we obtain  $\min_{(\alpha,\beta)} \mathcal{N}_2 = 1$  at the root  $(0, 0)$  and  $\tilde{f} = x^2 + y^2$ . That is the same deformed polynomial as for the unconstrained problem but derived from a different norm expression (8).

Note that before minimizing (8) one could restrict  $(\alpha, \beta)$  to lie on a parametric curve, thus constraining the variables rather than the coefficients, as is done in [7].  $\square$

Our method can be generalized even further to include inequalities,  $G\vec{f} \leq q$  with  $G \in \mathbb{R}^{m \times s}$ . Then

$$\left. \begin{aligned} \mathcal{N}_{2,w}^{[f;H;G]}(\alpha) = \\ \inf_{\tilde{f} \in \mathbb{R}[x_1, \dots, x_n]} \|f - \tilde{f}\|_{2,w}^2, \\ \text{s. t. } \tilde{f}(\alpha) = 0, G\vec{f} \leq q, H\vec{f} = p, \\ \deg(\tilde{f}) \leq \deg(f). \end{aligned} \right\} \quad (9)$$

Note that our constraint functions, being linear, are always convex. We can use the Karush-Kuhn-Tucker (KKT) conditions and the quantities as defined in Theorem 1. Using the KKT conditions in equation (9) with the Lagrange function

$$L = (\vec{f} - \tilde{f})^T D_w(\vec{f} - \tilde{f}) + \lambda_0 \tau^T \vec{f} + \lambda^T (H\vec{f} - p) + \mu^T (G\vec{f} - q).$$

The KKT conditions (for a regular point) are then

$$\left. \begin{aligned} \frac{\partial L}{\partial(\tilde{f})_i} = 0, \quad i = 1, \dots, s, \\ \tau^T \vec{f} = 0, \\ H\vec{f} = p, \\ G\vec{f} \leq q, \\ \mu_i \geq 0, \quad i = 1, \dots, m, \\ \mu^T (G\vec{f} - q) = 0. \end{aligned} \right\} \quad (10)$$

The last orthogonality conditions constitute branching:  $\mu_i = 0$  or  $(G\vec{f} - q)_i = 0$ , and (10) form linear programs.

**Example 3** Given a polynomial  $f(x, y) = x^2 + y^2 - 2y + 1$  and constraint  $\tilde{f}_{0,1} \geq 0$ , we determine the nearest polynomial  $\tilde{f}(x, y) = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0}$  with real root  $(0, 0)$ . The term vector for the root is  $\tau = [0, 0, 0, 0, 0, 1]$ . The Lagrangian function is  $L(\alpha, \beta, \lambda, \mu) = (\vec{f} - \tilde{f})^T (\vec{f} - \tilde{f}) + \lambda \tau^T \vec{f} + \mu(-\tilde{f}_{0,1})$ . We can formulate the KKT conditions as solving two linear programs:

$$\begin{aligned} &\text{Minimize} && 1 \\ &\text{subject to} && \partial L / \partial(\tilde{f})_i = 0, \quad i = 1, \dots, 6 \\ & && \tilde{f}_{0,0} = 0, \\ & && -\tilde{f}_{0,1} \leq 0, \\ & && \mu = 0, \end{aligned}$$

and

$$\begin{aligned} &\text{Minimize} && 1 \\ &\text{subject to} && \partial L / \partial(\tilde{f})_i = 0, \quad i = 1, \dots, 6, \\ & && \tilde{f}_{0,0} = 0, \\ & && -\tilde{f}_{0,1} = 0, \\ & && \mu \geq 0. \end{aligned}$$

The first linear program is infeasible; for the second linear program we obtain:

$$\tilde{f} = x^2 + y^2, \lambda = 2, \mu = 4, \text{ and } \mathcal{N}_2^{[f;G]} = 5.$$

The minimum perturbation can also be obtained by running the `Minimize` procedure in Maple 13 on the original optimization problem (9).  $\square$

The above result can be extended to systems. The distance to the nearest system with  $k$  equations and common root  $\alpha$  is defined as

$$\left. \begin{aligned} \inf_{\tilde{f}_1, \dots, \tilde{f}_k} \|f_1 - \tilde{f}_1\|_2^2 + \dots + \|f_k - \tilde{f}_k\|_2^2 \\ \text{s. t. } \tilde{f}_i(\alpha) = 0, \quad i = 1, \dots, k \\ f_i \in \mathbb{R}[x_1, \dots, x_n], \quad i = 1, \dots, k \\ \deg(\tilde{f}_i) \leq \deg(f_i), \quad i = 1, \dots, k \end{aligned} \right\} \quad (11)$$

Applying Theorem 1 and Theorem 5 to each individual  $\tilde{f}_k$  easily yields the following.

**Theorem 7** Let  $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_n]$ , with  $d_i = \deg(f_i)$ , The distance to the nearest system with a common root  $\alpha \in \mathbb{R}^n$  is in  $\ell^2$ -norm

$$\mathcal{N}_2^{\{f_1, \dots, f_k\}}(\alpha) = \frac{f_1(\alpha)^2}{\sum_{i_1 + \dots + i_n = 0}^{d_1} \alpha_1^{2i_1} \dots \alpha_n^{2i_n}} + \dots + \frac{f_k(\alpha)^2}{\sum_{i_1 + \dots + i_n = 0}^{d_k} \alpha_1^{2i_1} \dots \alpha_n^{2i_n}}, \quad (12)$$

and in  $\ell^\infty$ -norm

$$\mathcal{N}_{\infty,w}^{\{f_1, \dots, f_k\}}(\alpha) = \max_{1 \leq j \leq k} \frac{|f_j(\alpha)|}{\|\tau\|_{1,1/w}}.$$

The nearest polynomials, if they exist (see Example 1) are again determined by (3). Theorem 7 easily generalizes to include weighted norms. Linear equality and inequality constraints on the coefficients as described in (7) and (9) can also be applied.

**Example 4** Given polynomials

$$f_1(x, y) = x^4 + y^4 + 1 \text{ and } f_2(x, y) = x^2 + x^2 y^2 - 2xy + 1$$

we shall determine the minimum perturbation such that the deformed system of 2 equations has a real root.

For that, we compute the Gröbner basis of the numerators of the partial derivatives of (12) (cf. [2]). In Section 5 we present an alternative approach based on sum-of-squares certificates. The first equation in the obtained Gröbner basis is a polynomial in terms of  $\beta$  of degree 195. Next, we find all real roots of this polynomial and plug all 9 choices into a second polynomial in the Gröbner basis. We compute the norm of each possible point and select the minimum value. The minimum

perturbation obtained by solving the Gröbner basis of (12) in Maple is

$$\tilde{\mathcal{N}}_2 = 0.64597306998078277667 \quad (13)$$

for

$$(\alpha, \beta) = (-0.9138289555225176138, -1.1947071766554875688).$$

Note that for this example at least 25 mantissa digits must be used in Maple 13 in order to obtain the correct minimum.

We can then find the nearest polynomial system by plugging the root into equation (3) for each of the two polynomials:

$$\begin{aligned} \tilde{f}_1 = & 0.83448994938 + 0.15028000318x + \\ & 0.19773604528y - 0.17954059831xy - \\ & 0.13645140747x^2 - 0.23623667238y^2 + \\ & 0.12389530347x^3 + 0.21449844130xy^2 + \\ & 0.16301947576x^2y + 0.28223364788y^3 + \\ & 0.88750540206x^4 - 0.14801860821x^3y - \\ & 0.19476053763x^2y^2 - 0.25626282720xy^3 + \\ & 0.66281343538y^4, \end{aligned}$$

$$\begin{aligned} \tilde{f}_2 = & 0.96296934167 + 0.03362313909x + \\ & 0.04424079327y - 2.04016980557xy + \\ & 0.96947082410x^2 - 0.05285479322y^2 + \\ & 0.02771991571x^3 + 0.04799115499xy^2 + \\ & 0.03647342555x^2y + 0.06314600078y^3 - \\ & 0.02516916045x^4 - 0.03311718223x^3y + \\ & 0.95642493674x^2y^2 - 0.05733537729xy^3 - \\ & 0.07544098031y^4. \quad \square \end{aligned}$$

## 5. LOWER BOUND CERTIFICATES

The minimization of the rational function  $\mathcal{N}_{2,w}^{[f]} = \frac{f(\alpha)^2}{g(\alpha)^2}$  where  $g = \tau^T D_w^{-1} \tau$  defined in (2) can be reformulated as maximizing  $r$  such that  $f(\alpha)^2 - rg(\alpha)$  is nonnegative. We compute a lower bound of  $\inf_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha)$  by solving the SOS program [10, 19, 12, 14]:

$$\left. \begin{aligned} r^* := & \sup_{r \in \mathbb{R}, W} r \\ \text{s. t. } & f(\mathbf{X})^2 - rg(\mathbf{X}) = m_d(\mathbf{X})^T W m_d(\mathbf{X}) \\ & W \succeq 0, W^T = W \end{aligned} \right\} \quad (14)$$

where  $m_d(\mathbf{X})$  is the column vector of all terms in  $\mathbf{X}_1, \dots, \mathbf{X}_n$  up to degree  $d$ . The dimension of  $m_d(\mathbf{X})$  is  $\binom{n+d}{d}$ .

The SOS program (14) can be solved efficiently by algorithms in GloptiPoly [6], SOSTOOLS [22], YALMIP [17] and SeDuMi [28]. One can use GloptiPoly as described in [6] to extract the solutions  $\alpha$  which achieve the global minimum. However, since we are running fixed precision SDP solvers in Matlab, we can only obtain a numerical positive semidefinite matrix  $W$  and floating point number  $r^*$  which satisfy approximately

$$f(\mathbf{X})^2 - r^*g(\mathbf{X}) \approx m_d(\mathbf{X})^T \cdot W \cdot m_d(\mathbf{X}), \quad W \succeq 0. \quad (15)$$

So  $r^*$  is a lower bound of  $\inf_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha)$  approximately!

The lower bound  $\tilde{r}$  is certified if  $\tilde{r}$  and  $\tilde{W}$  hold the following conditions exactly:

$$f(\mathbf{X})^2 - \tilde{r}g(\mathbf{X}) = m_d(\mathbf{X})^T \cdot \tilde{W} \cdot m_d(\mathbf{X}), \quad \tilde{W} \succeq 0. \quad (16)$$

We can use Artin's theorem of sum-of-squares and semidefinite programming (see, e.g., [20, 12, 14]) to certify the computed minimum. We have done so for the minimum 1 of (8) of Example 2 and the rational lower bound  $\tilde{\mathcal{N}}_2 = 64597306998078108/10000000000000000$  of the real algebraic optimum (13) of Example 4.

**Example 5** ([29]) Given a polynomial

$$\begin{aligned} f &= x^2y^2 + x^2 - xy + y^4 - y^2 + 1 \\ &= (xy - 1/2)^2 + (y^2 - 1/2)^2 + x^2 + 1/2, \end{aligned}$$

decide the minimum perturbation such that the perturbed polynomial has a real root.

If we allow dense perturbations, after running `solvesos` in Matlab, we get the lower bound

$$\tilde{\mathcal{N}}_2 = 2.453484553428391600 \times 10^{-15}.$$

This is caused by the assumption that we can perturb  $f$  by any monomial terms with degree bounded by 4. In general, for  $f(x, y) - \epsilon x^4$  one has for  $x = y^2$  that  $g(y) = f(y^2, y) - \epsilon y^8$ . Notice that  $g(y)$  always has a real root, because  $g(0) = 1$  and  $g(\infty) = -\infty$ . We see that  $f$  has a radius of positive semidefiniteness that is 0. Hence, it would be more interesting to consider a weighted norm. For instance, if we only allow terms which appear in  $f$  to be perturbed, then the lower bound computed by `solvesos` in Matlab is  $\tilde{\mathcal{N}}_2 = 0.2469160193369205900$ . After applying the certification algorithm in [12, 14], we obtain the certified lower bound

$$\tilde{\mathcal{N}}_2 = 24691601933692029/100000000000000000. \quad \square$$

This means that  $f$  is positive since  $f(0, 0) = 1 > 0$ .

**Example 6** (see [18]) Consider the polynomial

$$f(x, y) = 2 - 3x^2y^2 + x^2y^4 + x^4y^2.$$

Notice that  $f$  is the result of adding one to the Motzkin polynomial. It is well-known that  $f$  is positive semidefinite but not an SOS, as seen in [18]. In fact  $f \geq 1$  for all  $x, y \in \mathbb{R}$ . First, we consider using a dense perturbation to obtain a lower bound for  $\mathcal{N}_2$ . We use Matlab to compute the approximate lower bound of  $\mathcal{N}_2$  and obtain 0 as the minimum, which is easily proven by considering  $f(x, y) - \epsilon x^5$ . Hence, we consider a weighted norm. We use infinite weights on the terms that have zero coefficients in  $f$ . Thus, we only allow the terms which appear in  $f$  to be perturbed (sparse deformation). The lower bound computed by `solvesos` in Matlab is

$$\tilde{\mathcal{N}}_2 = 0.1285480262594671800.$$

After applying the certification algorithm in [12, 14], we obtain the certified lower bound

$$\tilde{\mathcal{N}}_2 = 12854802625942833/100000000000000000.$$

We have computed an exact rational certificate (as in (16))  $f(x, y)^2 - 12854802625942833/100000000000000000 \times (1 + x^4y^8 + x^8y^4 + x^4y^4) = \text{SOS}$  (10 polynomial squares).

This means that the non-zero coefficients of  $f$  need to be perturbed (by at least 0.128 in  $\ell^2$ -norm squared) for  $f$  to have a real root. Since  $f(0, 0) = 2$ , we have

proven that  $f(x, y) > 0$  for all real  $x, y$  via a polynomial sum-of-squares certificate.  $\square$

Example 6 answers a question by one of the referees. In fact, we conjecture that such polynomial sums-of-squares always exist. More precisely, if for a real polynomial  $f(x_1, \dots, x_n)$  there exists a vector  $w$  of positive and infinite weights (excluding an infinite weight for the constant coefficient) such that  $\rho_{2,w}(f) > 0$  then in (14)  $r^* > 0$ . We have seen that  $\rho_{2,w}(f)$  easily is no larger than 0, provided  $f$  has a projective root at infinity, and the condition  $\rho_{2,w}(f) > 0$  makes  $f$  and  $w$  quite special.

**Acknowledgments:** Sharon Hutton thanks the EAPSI program staff at NSF and the staff at KLMM for their assistance with her 2009 Summer Institute in Beijing. We thank Mohab Safey El Din for his comments on Seidenberg's problem, and the reviewers for their remarks.

## 6. REFERENCES

- [1] AUBRY, P., ROULLIER, F., AND SAFEY EL DIN, M. Real solving for positive dimensional systems. *J. Symbolic Comput.* 34, 6 (Dec. 2002), 543–560. URL: <http://www.spiral.lip6.fr/~safey/Articles/RR-3992.ps.gz>.
- [2] BECKER, E., POWERS, V., AND WÖRMANN, T. Deciding positivity of real polynomials. In *Real Algebraic Geometry and Ordered Structures*, C. N. Delzell and J. J. Madden, Eds., vol. 253 of *Contemporary Math.* AMS, 2000, pp. 251–272.
- [3] CORLESS, R. M., GIANNI, P. M., TRAGER, B. M., AND WATT, S. M. The singular value decomposition for polynomial systems. In *Proc. 1995 Internat. Symp. Symbolic Algebraic Comput. ISSAC'95* (New York, N. Y., 1995), A. H. M. Levelt, Ed., ACM Press, pp. 96–103.
- [4] CORLESS, R. M., KALTOFEN, E., AND WATT, S. M. Hybrid methods. In *Computer Algebra Handbook*, J. Grabmeier, E. Kaltofen, and V. Weispfenning, Eds. Springer Verlag, Heidelberg, Germany, 2003, Section 2.12.3, pp. 112–125. URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/01/symnum.pdf>.
- [5] ECKER, J. G., AND KUPPERSCHMID, M. *Introduction to Operations Research*. John Wiley and Sons, 1988. 509 pp.
- [6] HENRION, D., AND LASSERRE, J.-B. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive polynomials in control*, D. Henrion and A. Garulli, Eds., vol. 312 of *Lecture Notes on Control and Information Sciences*. Springer Verlag, Heidelberg, Germany, 2005, pp. 293–310. URL: <http://homepages.laas.fr/henrion/Papers/extract.pdf>.
- [7] HITZ, M. A., AND KALTOFEN, E. Efficient algorithms for computing the nearest polynomial with constrained roots. In *Proc. 1998 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'98)* (New York, N. Y., 1998), O. Gloor, Ed., ACM Press, pp. 236–243. URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/98/HiKa98.pdf>.
- [8] HITZ, M. A., KALTOFEN, E., AND LAKSHMAN Y. N. Efficient algorithms for computing the nearest polynomial with a real root and related problems. In *Proc. 1999 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'99)* (New York, N. Y., 1999), S. Dooley, Ed., ACM Press, pp. 205–212. URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/99/HKL99.pdf>.
- [9] JACOBSON, N. *Basic Algebra I*. W. H. Freeman & Co., San Francisco, 1974.
- [10] JIBETEAN, D., AND DE KLERK, E. Global optimization of rational functions: a semidefinite programming approach. *Math. Program.* 106, 1 (2006), 93–109.
- [11] KALTOFEN, E. Computing the irreducible real factors and components of an algebraic curve. *Appl. Algebra Engin. Commun. Comput.* 1, 2 (1990), 135–148. URL: [http://www.math.ncsu.edu/~kaltofen/bibliography/90/Ka90\\_aaecc.pdf](http://www.math.ncsu.edu/~kaltofen/bibliography/90/Ka90_aaecc.pdf).
- [12] KALTOFEN, E., LI, B., YANG, Z., AND ZHI, L. Exact certification of global optimality of approximate factorizations via rationalizing sums-of-squares with floating point scalars. In *ISSAC 2008* (New York, N. Y., 2008), D. Jeffrey, Ed., ACM Press, pp. 155–163. URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/08/KLYZ08.pdf>.
- [13] KALTOFEN, E., YANG, Z., AND ZHI, L. Approximate greatest common divisors of several polynomials with linearly constrained coefficients and singular polynomials. In *ISSAC MMVI Proc. 2006 Internat. Symp. Symbolic Algebraic Comput.* (New York, N. Y., 2006), J.-G. Dumas, Ed., ACM Press, pp. 169–176. URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/06/KYZ06.pdf>.
- [14] KALTOFEN, E. L., LI, B., YANG, Z., AND ZHI, L. Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients, Jan. 2009. Accepted for publication in *J. Symbolic Comput.* URL: <http://www.math.ncsu.edu/~kaltofen/bibliography/09/KLYZ09.pdf>.
- [15] KARMARKAR, N. K., AND LAKSHMAN Y. N. On approximate GCDs of univariate polynomials. *J. Symbolic Comput.* 26, 6 (1998), 653–666. Special issue on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.
- [16] KHARITONOV, V. L. Asymptotic stability of an equilibrium of a family of systems of linear differential equations. *Differential Equations* 14 (1979), 1483–1485.
- [17] LÖFBERG, J. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. IEEE CCA/ISIC/CACSD Conf.* (Taipei, Taiwan, 2004). URL: <http://control.ee.ethz.ch/~joloef/yalmip.php>.
- [18] MARSHALL, M. *Positive Polynomials and Sums of Squares*. American Math. Soc., 2008. 187 pp.
- [19] NIE, J., DEMMEL, J., AND GU, M. Global minimization of rational functions and the nearest GCDs. *J. of Global Optimization* 40, 4 (2008), 697–718.
- [20] PEYRL, H., AND PARRILO, P. A. Computing sum of squares decompositions with rational coefficients. *Theoretical Comput. Sci.* 409 (2008), 269–281.
- [21] POPE, S., AND SZANTO, A. Nearest multivariate system with given root multiplicities. *Journal of Symbolic Computation* 44 (2009), 606–625. URL: <http://www4.ncsu.edu/~aszanto/Pope-Szanto-Preprint.pdf>.
- [22] PRAJNA, S., PAPACHRISTODOULOU, A., SELER, P., AND PARRILO, P. A. SOSTOOLS: Sum of squares optimization toolbox for MATLAB. Available from <http://www.cds.caltech.edu/sostools> and <http://www.mit.edu/~parrilo/sostools>, 2004.
- [23] REZVANI, N., AND CORLESS, R. M. The nearest singular polynomial with a given zero, revisited. *SIGSAM Bulletin* 39, 3 (Sept. 2005), 73–79.
- [24] RUATTA, O., SCIABICA, M., AND SZANTO, A. Over-constrained Weierstrass iteration and the nearest consistent system. Rapport de Recherche 5215, INRIA, [www.inria.fr](http://www.inria.fr), June 2004. Accepted in *Journal of Complexity*. URL: <ftp://ftp-sop.inria.fr/galaad/oruatta/RR-5215.pdf>.
- [25] SAFEY EL DIN, M. *Résolution réelle des systèmes polynomiaux en dimension positive*. Thèse de doctorat, Univ. Paris VI (Univ. Pierre et Marie Curie), Paris, France, 2001. URL: [http://www-spiral.lip6.fr/~safey/these\\_safey.ps.gz](http://www-spiral.lip6.fr/~safey/these_safey.ps.gz).
- [26] SEIDENBERG, A. A new decision method for elementary algebra. *Annals Math.* 60 (1954), 365–374.
- [27] STETTER, H. J. The nearest polynomial with a given zero, and similar problems. *SIGSAM Bulletin* 33, 4 (Dec. 1999), 2–4.
- [28] STURM, J. F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software* 11/12 (1999), 625–653.
- [29] ZENG, G. Nonstandard decision methods for the solvability of real polynomial equations. *Science in China Series-A Mathematics* 42, 12 (1999), 1251–1261.
- [30] ZHI, L., NODA, M.-T., KAI, H., AND WU, W. Hybrid method for computing the nearest singular polynomials. *Japan J. Industrial and Applied Math.* 21, 2 (June 2004), 149–162. URL: <http://www.mmrc.iss.ac.cn/~lzhi/Publications/zkzwjjiam04.pdf>.
- [31] ZHI, L., AND WU, W. Nearest singular polynomial. *J. Symbolic Comput.* 26, 6 (1998), 667–675. Special issue on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.