Bounds on the Pythagoras number of the sum of square magnitudes of Laurent polynomials

Thanh Hieu Le [*†]
Department of Mathematics, Quy Nhon University, Vietnam

Marc Van Barel [‡§]
Department of Computer Science, KU Leuven, Belgium

August 17, 2015

Abstract

This paper presents a lower and upper bound of the Pythagoras number of sum of square magnitudes of Laurent polynomials (*sosm-polynomials*). To prove these bounds, properties of the corresponding system of quadratic polynomial equations are used. Applying this method, a new proof for the best (known until now) upper bound of the Pythagoras number of real polynomials is also presented.

Keyworks: Pythagoras number, sum of squares, sum of square magnitudes of polynomials

[*] *Email*: `lethanhhieu@qne.edu.vn`

[†] The author is partially supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.01-2014.30.

[‡] *Email*: `marc.vanbarel@cs.kuleuven.be`

[§] The research was partially supported by the Research Council KU Leuven, project OT/10/038 (Multi-parameter model order reduction and its applications), PF/10/002 Optimization in Engineering Centre (OPTEC), by the Fund for Scientific Research–Flanders (Belgium), G.0828.14N (Multivariate polynomial and rational interpolation and approximation), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization).

1 Introduction

Non-negative (complex Laurent or real) polynomials play a fundamental role in several applications, see, e.g., [8, 7, 14]. Several optimization problems can be reformulated over the cone of non-negative polynomials. Each real non-negative polynomial (i.e., its coefficients are real and it is in real variables) can be approximated well by a sum of squares polynomial under the 1-norm [12]. A polynomial which is a sum of real squares (say, *sos-polynomial*) or which is a sum of square magnitudes of complex polynomials (say, *sosm-polynomial*) might have several sos- or sosm-representations. It is useful in practical computations when parametrizing the corresponding sets of sos- or sosm-polynomials, to find a representation with the minimum number of sos(m)-terms.

Let $\mathbb{R}[x]_{n,d}$ be the set of all real-valued n-variable polynomials of degree at most d. Let $\Sigma(n,d)$ denote the set of sum of squares (*sos*) polynomials, i.e.,

$$\Sigma(n,d) \triangleq \left\{ f \in \mathbb{R}[x]_{n,2d} : \ f(x) = \sum_{i=1}^{r} v_i(x)^2, \forall x \in \mathbb{R}^n; v_i \in \mathbb{R}[x]_{n,d}, \forall i = 1,\ldots,r \right\}.$$

The positive integer number

$$\pi(f) \triangleq \min \{ r \in \mathbb{N} : f \text{ is sum of } r \text{ squares} \}$$

is called the *Pythagoras number* or *the length* of f [4, 6, 19]. It is well-known that a polynomial f is an sos-polynomial if and only if there exists a positive semidefinite real symmetric matrix F such that f can be expressed as

$$f(x) = \mathbf{v}_d(x)^T F \mathbf{v}_d(x), \forall x \in \mathbb{R}^n$$

where $\mathbf{v}_d(x)$ is the column vector of all possible monomials $x^\alpha \triangleq x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ in $\mathbb{R}[x]_{n,d}$.

To define sum of square magnitude (*sosm*) polynomials, we need the following notations. Let $\mathbb{C}[x]_{n,d}$ denote the set of all complex-valued n-variable polynomials of degree at most d. In this case the polynomials are defined on the $n-$torus

$$\mathbb{T}^n \triangleq \{ z \in \mathbb{C}^n : |z_i| = 1, \forall i = 1,\ldots,n \}.$$

The set of sum of square magnitude (*sosm*) polynomials in n variables of degree d

$\Sigma^{\Im}(n,d)$ is defined as

$$\Sigma^{\Im}(n,d) \triangleq \left\{ g(x) : \ g(x) = \sum_{i=1}^{r} q_i(x)^2, \forall x \in \mathbb{T}^n; q_i \in \mathbb{C}[x]_{n,d}, \forall i = 1, \ldots, r \right\}.$$

Analogously, for each $g \in \Sigma^{\Im}(n,d)$, the Pythagoras number of g is defined as

$$\pi(g) \triangleq \min \{ r \in \mathbb{N} : g \text{ is sum of } r \text{ square magnitudes of polynomials} \}.$$

A Laurent polynomial is a sum of square magnitudes of polynomials in $\Sigma^{\Im}(n,d)$ if and only if there is a positive semidefinite Hermitian matrix G such that

$$g(z) = \mathbf{v}_d(z)^H G \mathbf{v}_d(z), \forall z \in \mathbb{T}^n,$$

where $\mathbf{v}_d(z)$ denotes the column vector of monomials $z^\alpha = z_1^{\alpha_1} \ldots z_n^{\alpha_n}$ in $\mathbb{C}[z]_{n,d}$.

The sets of possible exponents of polynomials in $\mathbb{R}[x]_{n,d}$, $\Sigma(n,2d)$ and $\Sigma^{\Im}(n,d)$ are defined, respectively, as follows:

$$\begin{aligned}
\Omega(n,d) &\triangleq \left\{ \alpha \in \mathbb{N}^n : |\alpha| \triangleq \sum_{j=1}^{n} \alpha_j \leq d \right\}, \\
\Gamma(n,d) &\triangleq \Omega(n,d) + \Omega(n,d) \\
&= \left\{ \gamma \in \mathbb{N}^n : |\gamma| \triangleq \sum_{j=1}^{n} \gamma_j \leq 2d \right\}, \\
\Gamma^{\Im}(n,d) &\triangleq \Omega(n,d) - \Omega(n,d).
\end{aligned}$$

Denote

$$e(n,d) \triangleq \frac{(n+d)!}{n!d!}, \quad a(n,d) \triangleq \frac{(n+2d)!}{n!(2d)!},$$

then the cardinalities of $\Omega(n,d)$ and $\Gamma(n,d)$ are (see [6])

$$|\Omega(n,d)| = e(n,d) \triangleq \hat{e}, \ |\Gamma(n,d)| = a(n,d) \triangleq \hat{a}.$$

We note that there has not been a formula for the cardinality of $\Gamma^{\Im}(n,d) \triangleq \hat{k}$ in the literature. To formulate the following theorems on lower and upper bounds for the

Pythagoras number of $\Sigma(n,d)$ and $\Sigma^{\Im}(n,d)$, we need the following notation:

$$\begin{aligned} L(n,d) &\triangleq \frac{2\hat{e}+1-\sqrt{[2\hat{e}+1]^2-8\hat{a}}}{2}, \\ U(n,d) &\triangleq \frac{\sqrt{1+8\hat{a}}-1}{2}, \\ L^{\Im}(n,d) &\triangleq \hat{e}-\sqrt{\hat{e}^2-\hat{k}}, \\ U^{\Im}(n,d) &\triangleq \frac{\sqrt{8\hat{k}+1}-1}{2}. \end{aligned}$$

Lower and upper bounds for the Pythagoras number of either sos-polynomials or sosm-polynomials are given by the following theorems.

Theorem 1. *[6] For any $f \in \Sigma(n,d)$, we have*

$$L(n,d) \leq \pi(f) \leq U(n,d) \leq \hat{e}.$$

Theorem 2. *For any $g \in \Sigma^{\Im}(n,d)$, we have*

$$L^{\Im}(n,d) \leq \pi(g) \leq \min\left\{U^{\Im}(n,d), \hat{e}\right\}.$$

The paper is organized as follows. Section 2 summarizes some important properties of the cones of positive semidefinite real symmetric and complex Hermitian matrices which will be used in subsequent sections. The bounds in Theorem 1 were given in [6] with the corresponding proof. Section 3 presents a new proof for the upper bound $U(n,d)$. The key result to give the proofs of the upper bound of either sos- or sosm-polynomials, Proposition 3, is also shown in this section. Section 4 deals with the proof of Theorem 2. In Section 5, a formula is derived for $\hat{k} \triangleq \Gamma^{\Im}(n,d)$ with $n = 2,3,\ldots,6$. Also, some examples for different values of n and d are presented showing that $U^{\Im}(n,d)$ can be less than $\hat{e} \triangleq e(n,d)$ for certain values of n and d and vice versa. Section 6 gives the conclusions.

2 Cones of positive semidefinite matrices

This section summarizes some properties of the cones $\mathbb{S}_+^\mu$ and $\mathbb{H}_+^\mu$ of real symmetric and complex Hermitian positive semidefinite $\mu \times \mu$ matrices, respectively. The results in this section are well-know in the literature. They are listed here without proofs.

Proposition 1. [5, 10] *The cones $\mathbb{S}_+^\mu, \mathbb{H}_+^\mu$ are*

- *proper, i.e., they are closed, convex, have nonempty interior (solid) and contain no line (pointed);*
- *self-dual.*

The following proposition summarizes the fact that either the space $\mathbb{S}^\mu$ of real symmetric matrices or the space $\mathbb{H}^\mu$ of complex Hermitian matrices can be identified with an adapted real Hilbert space. Throughout this chapter, unless otherwise stated, $\langle ., . \rangle$ denotes either the "trace" inner product $\langle A, B \rangle = \text{Trace}(A^H B)$ of matrices or the standard inner product in $\mathbb{C}^r$. We use the notation $(x_{ij})_{\text{condition on indices } i \text{and } j}$ to denote a vector containing the elements x_{ij} where the index j is varied faster than the index i.

Proposition 2. i) (See, e.g., [9]) *Suppose $\mathbb{S}^\mu$ is endowed with the "trace" inner product $\langle ., . \rangle$ and $\mathbb{R}^{\mu(\mu+1)/2}$ is endowed with the inner product*

$$\langle x, y \rangle_D = x^T D y, \forall x = (x_{ij})_{1 \le i \le j \le \mu}, y = (y_{ij})_{1 \le i \le j \le \mu} \in \mathbb{R}^{\mu(\mu+1)/2}$$

where $D = \text{diag}(d_{11}, d_{12}, \ldots, d_{1\mu}, \ldots, d_{\mu\mu})$ is a diagonal matrix with $d_{ii} = 1$ and $d_{ij} = 2$ for $1 \le i < j \le \mu$. Then the space $\mathbb{S}^\mu$ is isometrically isomorphic to $\mathbb{R}^{\mu(\mu+1)/2}$ under the map

$$\mathbb{S}^\mu \ni X = [x_{ij}]_{i,j=1,\ldots,\mu} \mapsto \tilde{x} = (x_{ij})_{1 \le i \le j \le \mu} \in \mathbb{R}^{\mu(\mu+1)/2}.$$

ii) (See, e.g., [10]) *The space $\mathbb{H}^\mu$ is isometrically isomorphic to $\mathbb{R}^{\mu^2}$, endowed with the standard inner product, under the map that maps each matrix $[a_{ij}] \in \mathbb{H}^\mu$ to the following vector in $\mathbb{R}^{\mu^2}$*

$$(a_{11}, \sqrt{2}\text{Re}(a_{12}), \sqrt{2}\text{Im}(a_{12}), \ldots, a_{22}, \sqrt{2}\text{Re}(a_{23}), \sqrt{2}\text{Im}(a_{23}), \ldots, a_{\mu\mu})^T.$$

For any $A = X + \imath Y, B = U + \imath V \in \mathbb{H}^\mu$, with $X, U \in \mathbb{S}^\mu$, $Y, V \in \mathbb{R}^{\mu \times \mu}$, $Y^T = -Y, V^T = -V$, the inner product on $\mathbb{H}^\mu$ can be expressed as

$$\langle A, B \rangle = \langle X, U \rangle + \langle Y, V \rangle = \langle \tilde{x}, \tilde{u} \rangle_D + \langle \tilde{y}, \tilde{v} \rangle_D,$$

where the corresponding vector $\tilde{y}, \tilde{v}$ of the skew-symmetric matrices Y, V is defined, respectively, via the map

$$W = [w_{ij}]_{i,j=1,\ldots,\mu} \mapsto \tilde{w} = (w_{ij})_{1 \le i \le j \le \mu}.$$

3 Upper bounds on the Pythagoras number of sos-polynomials

In this section, we give a new proof for the upper bound of the Pythagoras number of sos-polynomials given in Theorem 1. Some lower and upper bounds of the Pythagoras number of such polynomials were also presented in [19, 4]. The upper bound $U(n,d)$ in Theorem 1 is the sharpest (by now) and given in [6]. The authors proved such bound by using the "method of cages". This method is based on the Newton polytope of the sets of exponents $\Omega(n,d)$ and $\Gamma(n,d)$. A polynomial $f \in \mathbb{R}[x]_{n,2d}$ can always be represented as a linear combination of monomials. Moreover, if it is sos, $f = \sum f_i^2$, $f_i \in \mathbb{R}[x]_{n,d}$, then its coefficients can be represented as a quadratic polynomial of the coefficients of the f_i's. Each of these quadratic polynomials is called a "vectorial quadratic form" [3]. Given the coefficients of f, determining the coefficients of the polynomials f_i is equivalent to solving a system of quadratic equations.

Theorem 1 is a direct consequence of Theorems 4.4 and 6.1 in [6]. It says that every sum of squares polynomial can be expressed as a sum of at most $\lfloor U(n,d) \rfloor$ squares, where $\lfloor . \rfloor$ denotes the integer part of a real number. We now prove the upper bound using the theory of systems of "vectorial quadratic form" equations.

3.1 A new proof for the upper bound $U(n,d)$ of Theorem 1

We first recall some facts of vectorial quadratic forms from [3]. One can view each $m \times r$ real matrix H as an $m-$tuple of column vectors in $\mathbb{R}^r$, i.e.,

$$H \triangleq [h_1, \ldots, h_m]^T \in \mathbb{R}^{m \times r}.$$

A (real) *vectorial quadratic form* corresponding to the real symmetric matrix $Q = [q_{ij}] \in \mathbb{S}^m$ is a map $q : \mathbb{R}^{m \times r} \longrightarrow \mathbb{R}$ defined by

$$\begin{aligned} q(H) &= \sum_{i=1}^{m} q_{ii} \langle h_i, h_i \rangle + \sum_{i \neq j} q_{ij} \langle h_i, h_j \rangle \\ &= \sum_{i=1}^{m} q_{ii} \langle h_i, h_i \rangle + 2 \sum_{1 \leq i < j \leq m} q_{ij} \langle h_i, h_j \rangle. \end{aligned}$$

Then it is easy to see that

$$q(H) = \langle Q, HH^T \rangle.$$

Notice that the $(i,j)-$entry of HH^T is $\langle h_i, h_j\rangle$ for all $i,j = 1,\dots m$. Before giving our proof, we list the following result from [3].

Proposition 3. *Suppose $Q_1,\dots,Q_l$ are symmetric matrices of order n and $a_1,\dots,a_l$ are real numbers. If a positive semidefinite matrix X exists such that*

$$\langle Q_i, X\rangle = a_i, \forall i = 1,\dots,l,$$

then there exists a positive semidefinite matrix X_0 satisfying the l equations above and

$$\operatorname{rank}(X_0) \leq \left\lfloor \frac{\sqrt{8l+1}-1}{2} \right\rfloor.$$

Now, suppose f is a sos-polynomial in n real variables and of degree $2d$, say

$$f(x) = \sum_{i=1}^{r} p_i(x)^2, \;\; p_i(x) \in \mathbb{R}[x]_{n,d}, \;\; x \in \mathbb{R}^n, \;\; \forall i = 1,\dots,r.$$

Suppose furthermore that f is expressed in the classical basis as

$$f(x) = \sum_{\gamma\in\Gamma(n,d)} f_\gamma x^\gamma, \quad x^\gamma \triangleq x_1^{\gamma_1}\dots,x_n^{\gamma_n}. \tag{1}$$

Let V be the matrix whose columns are the column vectors of coefficients of $p_i's$. Then

$$f(x) = \mathbf{v}_d(x)^T(VV^T)\mathbf{v}_d(x). \tag{2}$$

Identifying the coefficients of f in the two expressions (1) and (2), we have

$$\sum_{\substack{\beta+\alpha=\gamma\\ \alpha,\beta\in\Omega(n,d)}} \Big(\sum_{i=1}^{r} p_{\alpha i}p_{\beta i}\Big) = f_\gamma, \forall\gamma\in\Gamma(n,d), \tag{3}$$

where $p_i = [p_{\alpha i}]_{\alpha\in\Omega(n,d)}$ is the column vector of coefficients of the polynomial $p_i(x)$. This gives us a system of $\hat{a} = |\Gamma(n,d)|$ equations of quadratic polynomials in $\hat{e}r$ variables $(p_{\alpha i}),\; i = 1,\dots,r,\; \alpha\in\Omega(n,d)$.

To apply Proposition 3, we define a vectorial quadratic form as follows. For each

$\gamma \in \Gamma(n,d)$, denote by $Q_\gamma = [Q^\gamma_{\alpha\beta}]_{\alpha,\beta\in\Omega(n,d)}$ the symmetric matrix defined by

$$Q^\gamma_{\alpha\beta} = \begin{cases} 1 & \text{if} \quad \alpha+\beta=\gamma, \alpha=\beta, \\ \frac{1}{2} & \text{if} \quad \alpha+\beta=\gamma, \alpha\neq\beta, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The corresponding vectorial quadratic form $q_\gamma : \mathbb{R}^{\hat{e}\times r} \longrightarrow \mathbb{R}$ is defined by

$$q_\gamma(H) = \langle Q_\gamma, HH^T\rangle = \sum_{\substack{\beta+\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \left(\sum_{i=1}^{r} h_{\alpha i}h_{\beta i}\right), \ H=(h_{\alpha i}) \in \mathbb{R}^{\hat{e}\times r}. \tag{5}$$

From (3), (4) and (5), it follows that the associated matrix VV^T of f satisfies Proposition 3, and hence a positive semidefinite matrix X_0 exists such that

$$\operatorname{rank}(X_0) \leq \left\lfloor \frac{\sqrt{8\hat{a}+1}-1}{2} \right\rfloor = \lfloor U(n,d) \rfloor .$$

The conclusion is obtained from the fact that

$$\pi(f) = \min\left\{\operatorname{rank}(A) : A \in \mathbb{S}^{\hat{e}}_+, \mathbf{v}_d(x)^T A \mathbf{v}_d(x) = f(x), \forall x \in \mathbb{R}^n\right\}.$$

3.2 Remarks

Remark 1. Results on the facial structure of linear programs and semidefinite programs [17, 16, 18] also give an upper bound on the Pythagoras number of real polynomials but it is not as sharp as $U(n,d)$. Several nice properties of faces of the cone of positive semidefinite matrices can be found in [20, 18, 17, 2, 10, 1, 16]. This unsharp upper bound is derived by considering the following primal and dual semidefinite programs, respectively,

$$\begin{array}{ll} \min_X \langle C, X\rangle & \\ \text{subject to} & X \in \mathbb{S}^{\hat{e}}_+, \\ & \langle Q_\gamma, X\rangle = f_\gamma, \forall \gamma \in \Gamma(n,d), \end{array} \tag{6}$$

and

$$\begin{array}{ll} \max_z \sum_{\gamma\in\Gamma(n,d)} f_\gamma z_\gamma & \\ \text{subject to} & z = (z_\gamma)_{\gamma\in\Gamma(n,d)} \in \mathbb{R}^{\hat{a}}, \\ & \left(C - \sum_{\gamma\in\Gamma(n,d)} z_\gamma Q_\gamma\right) \succeq 0, \end{array} \tag{7}$$

where $C \in \mathbb{S}^{\hat{e}}$. Pataki [16, 17, 18] proved that for any feasible point X (with rank r) of the primal semidefinite program (6), the following rank inequality holds

$$\frac{r(r+1)}{2} \leq \hat{a} + \dim \mathcal{F},$$

where $\mathcal{F}$ is the smallest face of the feasible set containing X. This certainly gives a weaker upper bound than the one given in Theorem 1 because

$$\frac{U(n,d)[U(n,d)+1]}{2} \leq \hat{a}.$$

Remark 2. In [3] it is shown that there always exists a positive definite matrix C for which the following inequality holds

$$\operatorname{rank}\left(C - \sum_{\gamma \in \Gamma(n,d)} x_\gamma Q_\gamma\right) \geq \hat{e} - \left\lfloor \frac{\sqrt{8\hat{a}+1}-1}{2} \right\rfloor \tag{8}$$

for all $\{x_\gamma\}_{\gamma \in \Gamma(n,d)} \subset \mathbb{R}$. A consequence when such a matrix C exists is that both primal and dual semidefinite programs (6) and (7) have an optimal solution. The key is that the matrices $\{Q_\gamma\}_{\gamma \in \Gamma(n,d)}$ are linearly independent.

Proposition 4. *The matrices $\{Q_\gamma\}_{\gamma \in \Gamma(n,d)}$ defined in (4) are linearly independent.*

Proof. Notice that for $\gamma, \gamma' \in \Gamma(n,d)$, if $\gamma \neq \gamma'$ then for any $\alpha, \alpha', \beta, \beta' \in \Omega(n,d)$ such that $\alpha + \beta = \gamma$ and $\alpha' + \beta' = \gamma'$ we have $(\alpha, \beta) \neq (\alpha', \beta')$. This implies that any nonzero entry of the matrix Q_γ does not appear at the same position as the nonzero ones of $Q_{\gamma'}$. This gives us the conclusion of the proposition. □

4 Bounds on the Pythagoras number of sosm-polynomials

We start this section by stating the following proposition which allows us to consider only the polynomials being sums of square magnitudes of linearly independent polynomials. In a sosm-representation of a polynomial, if a sosm-term polynomial is a linear combination of the polynomials of other terms then its square magnitude is not necessarily a linear combination of the other square magnitudes.

Proposition 5. *If $g(z)$ is a sum of r square magnitudes of polynomials $q_i(z) \in \mathbb{C}[z]_{n,d}, i = 1, \ldots, r$, and the column vectors of coefficients of the polynomials $q_i(z)$ are*

linearly dependent then it can be expressed as a sum of at most $\hat{e}$ square magnitudes of linearly independent polynomials.

Proof. Suppose $g(z) = \sum_{i=1}^{r} |q_i(z)|^2, \forall z \in \mathbb{T}^n$ where $q_i(z) \in \mathbb{C}[z]_{n,d}, 1 = 1, \ldots, r$. Then it has a matrix representation (see, e.g., [13])

$$g(z) = \sum_{i=1}^{r} |q_i(z)|^2 = \mathbf{v}_d(z)^H \Big(\sum_{i=1}^{r} \bar{\mathbf{q}}_k \mathbf{q}_k^T \Big) \mathbf{v}_d(z) = \mathbf{v}_d(z)^H (\overline{G} G^T) \mathbf{v}_d(z), \tag{9}$$

where $\mathbf{q}_i$ denotes the column vector of coefficients of the polynomial $q_i(z)$, $G = [\mathbf{q}_1, \ldots, \mathbf{q}_r]$, $\overline{G}$ is the element-wise conjugate of G. Since the sosm-term polynomials $q_i(z)$ are linearly dependent, $\text{rank}(G) = s < r$. Applying the Cholesky factorization we have $\overline{G} G^T = \overline{L} L^T$ where $L \in \mathbb{C}^{\hat{e} \times s}$ is lower triangular of rank s. We obtain the new representation of the polynomial

$$g(\tilde{z}) = \mathbf{v}_d(\tilde{z})^H (\overline{L} L^T) \mathbf{v}_d(\tilde{z}).$$

This implies that the new polynomial is a sum of $\hat{e}$ square magnitudes of polynomials, and the sosm-term polynomials are linearly independent. □

Because of Proposition 5, one can assume in the rest of this chapter that the sosm-term polynomials of a sosm-polynomial are linearly independent.

Now, suppose

$$g(z) = \sum_{\gamma \in \Gamma^{\Im}(n,d)} g_\gamma z^\gamma, \quad \deg(g) = d, \tag{10}$$

is a sum of r square magnitudes of polynomials with a matrix representation as in (9). Note that its sosm-term polynomials are linearly independent. For each $i = 1, \ldots, r$, by $q_{\alpha i}$ denote the αth coefficient of the polynomial $q_i(z)$. Identifying the coefficients of $g(z)$ in the matrix and the canonical-basis representations as in (9) and (10), respectively, we get

$$\sum_{\substack{\beta - \alpha = \gamma \\ \alpha, \beta \in \Omega(n,d)}} \Big(\sum_{i=1}^{r} \bar{q}_{\alpha i} q_{\beta i} \Big) = g_\gamma, \ \forall \gamma \in \Gamma^{\Im}(n,d). \tag{11}$$

So if the Laurent polynomial $g(z)$ is sosm on the $n-$torus $\mathbb{T}^n$, then $g_\gamma = \bar{g}_{-\gamma}$ for all

$\gamma \in \Gamma^{\Im}(n,d),\ \gamma \geqq 0$ (componentwise). So, the equations in (11) are reduced to

$$\sum_{\substack{\beta-\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \Big(\sum_{i=1}^{r} \bar{q}_{\alpha i} q_{\beta i}\Big) = g_\gamma,\ \forall \gamma \in \Gamma^{\Im}(n,d), \gamma \geqq 0. \tag{12}$$

One also notices from the matrix representation of sosm-polynomials that the Pythagoras number of sosm-polynomials is bounded above by $\hat{e}$, i.e., $r \leq \hat{e}$. We now give another upper bound for the Pythagoras number of such polynomials in the next subsection.

4.1 The upper bound

In this subsection, we convert the system of complex quadratic equations (12) to one of real quadratic equations. Then we apply Proposition 3 to obtain an upper bound for sosm-polynomials.

Firstly, in (12), set

$$\begin{aligned} q_{\alpha i} &= x_{\alpha i} + \imath y_{\alpha i}, \quad x_{\alpha i}, y_{\alpha i} \in \mathbb{R},\ \forall \alpha \in \Omega(n,d), \quad \forall i = 1,\ldots,r, \\ g_\gamma &= u_\gamma + \imath v_\gamma, \quad u_\gamma, v_\gamma \in \mathbb{R},\ \forall \gamma \in \Gamma^{\Im}(n,d), \quad \gamma \geqq 0. \end{aligned}$$

The system (12) is then equivalent to the system of $k = |\Gamma^{\Im}(n,d)|$ real quadratic equations

$$\begin{aligned} \sum_{\substack{\beta-\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \sum_{i=1}^{r} \Big(x_{\alpha i} x_{\beta i} + y_{\alpha i} y_{\beta i}\Big) &= u_\gamma, \forall \gamma \in \Gamma^{\Im}(n,d), \gamma \geqq 0, \\ \sum_{\substack{\beta-\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \sum_{i=1}^{r} \Big(x_{\alpha i} y_{\beta i} - y_{\alpha i} x_{\beta i}\Big) &= v_\gamma, \forall \gamma \in \Gamma^{\Im}(n,d), \gamma \gneqq 0. \end{aligned} \tag{13}$$

On the other hand, using the Cholesky factorization of Prop. 5, one can assume the

matrix G to have the lower triangular form

$$G = \begin{pmatrix} * & & \\ \vdots & \ddots & \\ * & \dots & * \\ \vdots & \ddots & \vdots \\ * & \dots & * \end{pmatrix} \in \mathbb{C}^{\hat{e}\times r}$$

where $*$ denotes a possible nonzero entry, and each diagonal element on any of its columns is real and nonzero. The matrix G can then be expressed as $G = X + \imath Y, X, Y \in \mathbb{R}^{\hat{e}\times r}$ where X contains $\hat{e}r - \frac{r(r-1)}{2}$ possible nonzero entries and Y contains $\hat{e}r - \frac{r(r+1)}{2}$ possible nonzero entries. Let $X = (x_{\alpha i})_{\alpha,i}, Y = (y_{\alpha i})_{\alpha,i} \in \mathbb{R}^{\hat{e}\times r}$. From (13), consider $\hat{k} \triangleq |\Gamma^{\Im}(n,d)|$ vectorial quadratic forms defined on $\mathbb{R}^{2\hat{e}\times r}$

$$\begin{aligned} q_\gamma^{\Re}(X;Y) &\triangleq \sum_{\substack{\beta-\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \sum_{i=1}^{r} \left(x_{\alpha i}x_{\beta i} + y_{\alpha i}y_{\beta i}\right), \quad \gamma \geqq 0, \\ q_\gamma^{\Im}(X;Y) &\triangleq \sum_{\substack{\beta-\alpha=\gamma \\ \alpha,\beta\in\Omega(n,d)}} \sum_{i=1}^{r} \left(x_{\alpha i}y_{\beta i} - y_{\alpha i}x_{\beta i}\right), \quad \gamma \gneqq 0. \end{aligned} \tag{14}$$

Note that in (14), some of the parameters $y_{\alpha i}$ and $x_{\alpha i}$ are zero (corresponding to the strictly upper triangular part of G). Set

$$\mathrm{sym}(X,Y) = \begin{pmatrix} XX^T & XY^T \\ YX^T & YY^T \end{pmatrix} \in \mathbb{S}^{2\hat{e}}.$$

It is clear that

$$\mathrm{sym}(X,Y) = \begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X^T & Y^T \end{pmatrix} \succeq 0.$$

Then there exist symmetric matrices $Q_\gamma^{\Re}, Q_\gamma^{\Im} \in \mathbb{S}^{2\hat{e}}$ such that

$$q_\gamma^{\Re}(X;Y) = \langle Q_\gamma^{\Re}, \mathrm{sym}(X,Y)\rangle, \quad q_\gamma^{\Im}(X;Y) = \langle Q_\gamma^{\Im}, \mathrm{sym}(X,Y)\rangle.$$

Thus, for a given sosm-polynomial $g(z) = \sum_{i=1}^{r} |q_i(z)|^2 = \mathbf{v}_d(z)^H(\overline{G}G^T)\mathbf{v}_d(z)$ with $G = X + \imath Y$ as above, there exist $\hat{k}$ symmetric matrices $Q_\gamma^{\Re}, Q_\gamma^{\Im} \in \mathbb{S}^{2\hat{e}}$ and a positive

semidefinite matrix $A = \mathrm{sym}(X, Y)$ such that

$$\langle Q_\gamma^{\Re}, A\rangle = u_\gamma, \ \ \forall \gamma \geqq 0 \text{ and } \langle Q_\gamma^{\Im}, A\rangle = v_\gamma, \ \ \forall \gamma \gneqq 0.$$

By Proposition 3, there is a matrix $A_0 \in \mathbb{S}_+^{2\hat{e}}$ satisfying the conditions above and

$$\mathrm{rank}(A_0) \leq \frac{\sqrt{8\hat{k}+1}-1}{2}.$$

Finally, let $A_0 = A_1 A_1^T$ be the Cholesky decomposition of A_0. Then $A_1 \in \mathbb{R}^{2\hat{e}\times s}$ with $\mathrm{rank}(A_0) = \mathrm{rank}(A_1) = s$. Let X_0 and Y_0 be the matrices taken from the first and last $\hat{e}$ rows of A_1, respectively. Then $A_0 = \mathrm{sym}(X_0, Y_0)$. The matrix $G_0 = X_0 + \imath Y_0 \in \mathbb{C}^{\hat{e}\times s}$ is also an associated matrix of the polynomial $g(z)$, and if $s \leq r$ then we have

$$\pi(g) \leq \mathrm{rank}(G_0) \leq \mathrm{rank}\begin{pmatrix} X_0 \\ \imath Y_0 \end{pmatrix} = \mathrm{rank}\begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = \mathrm{rank}(A_0) \leq U^{\Im}(n,d).$$

The second inequality can be found in [15].

Since $\pi(g) \leq \hat{e}$, we have the following.

Proposition 6. *If the Laurent polynomial $g(z)$ of degree d is sosm on the $n-$torus $\mathbb{T}^n$ then*

$$\pi(g) \leq \min\{\hat{e}, U^{\Im}(n,d)\}.$$

4.2 The lower bound

To give a proof for the lower bound we need a result from function theory and dimension theory. In particular, one concerns the dimension under polynomial mappings. The "dimension" here stands for the dimension of topological spaces, see, e.g., [11, 6]. More precisely, we say that a subset of $\mathbb{R}^\mu$ has dimension μ if its interior is nonempty.

Proposition 7. [6] *A polynomial mapping $\Phi : \mathbb{R}^\mu \to \mathbb{R}^\nu$, i.e., with coordinate functions that are polynomials, always satisfies the dimensional inequality*

$$\dim(\mathrm{Im}(\Phi)) \leq \mu.$$

Below we prove that the set of sosm-polynomials can be embedded in the range space of a polynomial mapping. Indeed, consider the polynomial mapping

$$\Phi \triangleq (q_\gamma^{\Re}, q_\gamma^{\Im})_{\gamma \geqq 0} : \mathbb{R}^{2\hat{e}r - r^2} \equiv \mathbb{R}^{\hat{e}r - \frac{r(r-1)}{2}} \times \mathbb{R}^{\hat{e}r - \frac{r(r+1)}{2}} \to \mathbb{R}^k$$

defined by (14). Since for any $g(z) \in \Sigma^{\Im}(n,d)$ there exist two real matrices X, Y so that (13) is satisfied, $\Sigma^{\Im}(n,d)$ is isomorphic to a subset of $\mathrm{Im}(\Phi)$. We will prove that $\mathrm{int}(\Sigma^{\Im}(n,d)) \neq \emptyset$ then so is $\mathrm{int}(\mathrm{Im}(\Phi))$. We then apply Proposition 7 to give the lower bound in Theorem 2. For any $\varepsilon \in \mathbb{C}$ and any $\alpha, \beta \in \Omega(n,d)$, the polynomial

$$(\varepsilon z^{\alpha} + z^{\beta})\overline{(\varepsilon z^{\alpha} + z^{\beta})} = |\varepsilon|^2 + 1 + \varepsilon z^{\alpha-\beta} + \bar{\varepsilon} z^{\beta-\alpha}, z \in \mathbb{T}^n,$$

is sosm. Let $f(z) = |\Gamma^{\Im}(n,d)| + 1, \forall z \in \mathbb{T}^n$. It is certain that f is sosm. We prove that $f \in \mathrm{int}(\Sigma^{\Im}(n,d))$. Indeed, for any

$$h(z) = \sum_{\substack{\gamma \in \Gamma^{\Im}(n,d) \\ \gamma \gneqq 0}} (\varepsilon_\gamma z^{\gamma} + \bar{\varepsilon}_\gamma z^{-\gamma}) + \varepsilon_0 \in \Sigma^{\Im}(n,d), \sum_{\gamma \geqq 0} |\varepsilon_\gamma|^2 \leq 1,$$

then h belongs to the unit ball in $\mathbb{R}^k$, and the polynomial

$$f + h = \sum_{\gamma \gneqq 0} \left(\varepsilon_\gamma z^{\gamma} + \bar{\varepsilon}_\gamma z^{-\gamma} + |\varepsilon_\gamma|^2 + 1\right) + (1 - \sum_{\gamma \gneqq 0} |\varepsilon_\gamma|^2)$$

is sosm. This means that f is an interior point of $\Sigma^{\Im}(n,d)$. So

$$\dim \mathrm{Im}(\Phi) = \dim \Sigma^{\Im}(n,d) = k.$$

Proposition 7 implies that

$$r^2 - 2\hat{e}r + k \leq 0.$$

Moreover, if $g(z)$ is sosm then from $g(z) = \mathbf{v}_d(z)^H A \mathbf{v}_d(z), A = (a_{\alpha\beta})$ we have

$$g_\gamma = \sum_{\beta - \alpha = \gamma} a_{\alpha\beta}, \forall \gamma \in \Gamma^{\Im}(n,d).$$

There are $\hat{e}^2$ entries $a_{\alpha\beta}$, and if $\gamma \neq \gamma'$ then for $\alpha - \beta = \gamma, \alpha' - \beta' = \gamma'$, we have $(\alpha, \beta) \neq (\alpha', \beta')$. This implies

$$k \leq \hat{e}^2,$$

and hence the inequality $r^2 - 2\hat{e}r + k \leq 0$ is equivalent to

$$\hat{e} - \sqrt{\hat{e}^2 - k} \leq r \leq \hat{e} + \sqrt{\hat{e}^2 - k}.$$

We thus have the following.

Table 1: Values of $p_n(d)$ with respect to some pairs of (n, d).

n	d	$\lvert\Gamma^{\Im}(n,d)\rvert = p_n(d)$	n	d	$\lvert\Gamma^{\Im}(n,d)\rvert = p_n(d)$
3	1	13	5	1	31
	2	55		2	271
	3	147		3	1281
	4	309		4	4251
				5	11253
				6	25493
4	1	21	6	1	43
	2	131		2	505
	3	471		3	3067
	4	1251		4	12559
	5	2751		5	39733
				6	104959
				7	242845

Proposition 8. *For any sosm-polynomial $g \in \Sigma^{\Im}(n, d)$, we have*

$$\pi(g) \geq L^{\Im}(n, d).$$

5 The cardinality of $\Gamma^{\Im}(n, d)$ and examples

In this section, we give some examples demonstrating why the upper bound in Theorem 2 should be taken as $\min\{U^{\Im}(n, d), \hat{e}\}$. First of all we derive a formula for $\Gamma^{\Im}(n, d)$ for $n = 2, 3, \ldots, 6$.

For a fixed number of variables n, one can see that $|\Gamma^{\Im}(n, d)|$ is a polynomial $p_n(d)$ in d of degree n. So, if one knows $n + 1$ values of $p_n(d)$ with respect to $n + 1$ values of d then by using Lagrange interpolation one obtains an explicit formula of $p_n(d)$. Table 5 shows such values of $p_n(d)$, which are numerically determined, with respect to several values of n. Based on this interpolation data, we have the following.

Proposition 9. *A formula for $|\Gamma^{\Im}(n, d)|$ for $n = 2, 3, \ldots, 6$ is given by*

i) $|\Gamma^{\Im}(2, d)| = (2d + 1)^2 - d(d + 1)$.

ii) $|\Gamma^{\Im}(3, d)| = (2d + 1)^3 - \frac{7}{3}d(d + 1)(2d + 1)$.

iii) $|\Gamma^{\Im}(4, d)| = (2d + 1)^4 - \frac{1}{12}d(d + 1)(157d^2 + 157d + 46)$.

Table 2: Values of $L^{\Im}(n,d)$, $\hat{e} \triangleq e(n,d)$ and $U^{\Im}(n,d)$ with respect to values of (n,d).

n	d	$\lceil L^{\Im}(n,d) \rceil$	$\hat{e} \triangleq e(n,d)$	$\lfloor U^{\Im}(n,d) \rfloor$
4	1	3	5	6
	2	6	15	15
	3	8	35	30
	4	10	70	49
	5	12	126	73
5	1	4	6	7
	2	8	21	22
	3	13	56	50
	4	19	126	91
	5	24	252	149

iv) $|\Gamma^{\Im}(5,d)| = \frac{1}{120}(252d^5 + 630d^4 + 1120d^3 + 1050d^2 + 548d + 120)$.

v) $|\Gamma^{\Im}(6,d)| = \frac{1}{720}(924d^6 + 2772d^5 + 6720d^4 + 8820d^3 + 7476d^2 + 3528d + 720)$.

Based on the above proposition, we can give several examples showing that $U^{\Im}(n,d)$ is not always less than or equal to $\hat{e}$.

By the Fejér-Riesz Theorem, the Pythagoras number of sosm-polynomials in one variable is one. In case of two variables, one can prove that $\lfloor U^{\Im}(2,d) \rfloor \leq e(2,d)$, $\forall d \geq 2$. The same estimation for $n = 3$, one also obtains $\lfloor U^{\Im}(3,d) \rfloor \leq e(3,d)$, $\forall d \geq 2$. Table 5 shows some values of n and small d for which $e(n,d) < \lfloor U^{\Im}(n,d) \rfloor$.

Note that the conjectured formula of the Pythagoras number of sosm-polynomials given in [13] satisfies the bounds of Theorem 2. The upper bound turns out to be a sharp one.

6 Conclusion

A lower and sharp upper bound for the Pythagoras number of sosm-polynomials were presented. These bounds are new and could be useful in practice, leading to a reduction in computational complexity when problems are considered over the cone of such polynomials. A new proof for the known upper bound of the Pythagoras number of sos-polynomials has also been presented.

References

[1] G. P. Barker. The lattice of faces of a finite dimensional cone. *Linear Algebra and its Applications*, 7:71–82, 1973.

[2] G. P. Barker. Theory of cones. *Linear Algebra and its Applications*, 39:263–291, 1981.

[3] A. I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13:189–202, 1995.

[4] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*. Springer-Verlag, Berlin, Heidelberg, 1998.

[5] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[6] M.-D. Choi, T. Y. Lam, and B. Reznick. Sums of squares of real polynomials. In *Proceedings of Symposia in Pure Mathematics*, volume 58, pages 103–126, 1995.

[7] Y. V. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. Convex optimization over positive polynomials and filter design. In *Proceedings UKACC Int. Conf. Control 2000*, 2000. Paper SS41.

[8] J. S. Geronimo and M.-J. Lai. Factorization of multivariate positive Laurent polynomials. *Journal of Approximation Theory*, 139(1-2):327–345, April 2006.

[9] O. Güler. Barrier function in interior point methods. *Mathematics of Operations Research*, 21(4):860–885, 1996.

[10] R. D. Hill and S. R. Waters. On the cone of positive semidefinite matrices. *Linear Algebra and its Applications*, 90:81–88, 1987.

[11] W. Hurewicz and H. Wallman. *Dimension theory*. Princeton University Press, 1948.

[12] J. B. Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM Review*, 49(4):651–669, 2007.

[13] T. H. Le, L. Sorber, and M. Van Barel. The Pythagoras number of real sum of squares polynomials and sum of square magnitudes of polynomials. *Calcolo*, 50(4):283–303, 2013.

[14] T. H. Le and M. Van Barel. A convex optimization method to solve a filter design problem. *Journal of Computational and Applied Mathematics*, 255:183–192, 2014.

[15] G. Marsaglia and G. P. H. Styan. When does $\text{rank}(A+B) = \text{rank}(A)+\text{rank}(B)$? *Canadian Mathematical Bulletin*, 15(3):451–452, 1972.

[16] G. Pataki. Cone-LP's and semidefinite programs: Geometry and a simplex-type method. In *Integer Programming and Combinatorial Optimization*, volume 1084 of *Lecture Notes in Computer Science*, pages 162–174, 1996.

[17] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research*, 23(2):339–358, May 1998.

[18] G. Pataki. The Geometry of Semidefinite Programming. In H Wolkowicz, R Saigal, and L Vandenberghe, editors, *Handbook of semidefinite programming: theory, algorithms, and applications*, International series in operations research and management science. Kluwer Academic Publishers, 2000.

[19] A. Prestel and C. N Delzell. *Positive polynomials*. Springer monographs in mathematics. Springer, 2001.

[20] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.