

MIT Open Access Articles

Sampling Algebraic Varieties for Sum of Squares Programs

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Cifuentes, Diego and Pablo A. Parrilo. "Sampling Algebraic Varieties for Sum of Squares Programs." SIAM journal on optimization 27, no. 4 (2017): pp. 2381-2404.

As Published: 10.1137/15M1052548

Publisher: Society for Industrial & Applied Mathematics (SIAM)

Persistent URL: <https://hdl.handle.net/1721.1/121576>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



SAMPLING ALGEBRAIC VARIETIES FOR SUM OF SQUARES PROGRAMS*

DIEGO CIFUENTES[†] AND PABLO A. PARRILO[†]

Abstract. We study sum of squares (SOS) relaxations to optimize polynomial functions over a set $\mathcal{V} \cap \mathbb{R}^n$, where \mathcal{V} is a complex algebraic variety. We propose a new methodology that, rather than relying on some algebraic description, represents \mathcal{V} with a generic set of complex samples. This approach depends only on the geometry of \mathcal{V} , avoiding representation issues such as multiplicity and choice of generators. It also takes advantage of the coordinate ring structure to reduce the size of the corresponding semidefinite program (SDP). In addition, the input can be given as a straight-line program. Our methods are particularly appealing for varieties that are easy to sample from but for which the defining equations are complicated, such as $SO(n)$, Grassmannians, or rank k tensors. For arbitrary varieties, we can obtain the required samples by using the tools of numerical algebraic geometry. In this way we connect the areas of SOS optimization and numerical algebraic geometry.

Key words. sum of squares, sampling, SDP, algebraic variety, numerical algebraic geometry

AMS subject classifications. 90C22, 65H10

DOI. 10.1137/15M1052548

1. Introduction. Consider the ring $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ of multivariate polynomials and an algebraic variety $\mathcal{V} \subseteq \mathbb{C}^n$. For a given polynomial $p \in \mathbb{R}[x]$, we are interested in deciding whether

$$(1) \quad p(x) \geq 0 \quad \text{for all } x \in \mathcal{V} \cap \mathbb{R}^n.$$

More generally, we can consider the problem of finding lower bounds for a polynomial on a real variety. Recall that an algebraic variety can be given implicitly as the zero set of polynomial equations or, parametrically, as the image of \mathbb{C}^n under a rational map.

The decision problem in (1) is computationally hard, but there are simpler relaxations based on *sum of squares* (SOS) [23]. Recall that a polynomial $F \in \mathbb{R}[x]$ is SOS if it can be written in the form $F(x) = \sum_i f_i(x)^2$ for some $f_i \in \mathbb{R}[x]$. Given a bound $d \in \mathbb{N}$, a sufficient condition for (1) to hold is the existence of a polynomial $F \in \mathbb{R}[x]$ such that

$$(2) \quad p(z) = F(z) \text{ for all } z \in \mathcal{V} \quad (\text{i.e., } p \equiv F \bmod \mathbf{I}(\mathcal{V})); \quad F(x) \text{ is SOS}; \quad \deg(F) \leq 2d.$$

We refer to such an F as a d -SOS(\mathcal{V}) *certificate*. The main problem we address in this paper is the following.

PROBLEM. *Given a bound $d \in \mathbb{N}$, a polynomial $p(x)$, and a variety \mathcal{V} , find a d -SOS(\mathcal{V}) certificate (if it exists).*

It was shown in [22] that, given a Gröbner basis of the ideal $\mathbf{I}(\mathcal{V})$, the above problem reduces to a semidefinite program (SDP). To the best of our knowledge, this is the only known method to address this problem. This approach is quite effective for

*Received by the editors December 14, 2015; accepted for publication (in revised form) September 12, 2017; published electronically November 16, 2017.

<http://www.siam.org/journals/siopt/27-4/M105254.html>

[†]Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 (diegcif@mit.edu, parrilo@mit.edu).

varieties with simple Gröbner bases such as the hypercube $\{0, 1\}^n$ or hypersurfaces. Unfortunately, other than these simple cases, Gröbner bases computation is typically too expensive.

Given defining equations of the variety $\{h_j(x) = 0\}_j$, there is a *weaker* class of certificates based on writing $p(x)$ in the form $F(x) + \sum_j g_j(x)h_j(x)$ [23]. This approach is widely used in practice [4, 17], thanks to the convenience of allowing any set of defining equations. But this simplicity comes with a price, since the success of the relaxation now depends on the choice of a good set of equations $\{h_j\}_j$. Furthermore, the corresponding SDP is larger. Indeed, for fixed \mathcal{V} the number of unknowns is $O(d^{2n})$, whereas for (2) it is $O(d^{2 \dim \mathcal{V}})$; see Remark 2.4. We also point out that for several parametric varieties, notably secant varieties [15], the defining equations are not explicitly known, thus making this type of certificate unfeasible.

Sampling certificates. In this paper we propose an alternative geometric approach to compute $\text{SOS}(\mathcal{V})$ certificates. Rather than depending on an algebraic description of the variety, we rely on a generic set of samples $Z := \{z_1, \dots, z_S\} \subseteq \mathcal{V}$. By specializing the condition in (2) to such samples, we get the following.

DEFINITION 1.1. *Let $\mathcal{V} \subseteq \mathbb{C}^n$ be a variety, and let $p \in \mathbb{R}[x]$ be nonnegative on $\mathcal{V} \cap \mathbb{R}^n$. Given a bound $d \in \mathbb{N}$, a sampling d -SOS precertificate is a pair (F, Z) , where $F(x)$ is a polynomial and $Z = \{z_1, \dots, z_S\} \subseteq \mathcal{V}$ is a sample set such that*

$$(3) \quad p(z_s) = F(z_s) \text{ for } s = 1, \dots, S; \quad F(x) \text{ is SOS}; \quad \deg(F) \leq 2d.$$

The precertificate is correct if F is a d -SOS(\mathcal{V}) certificate, i.e., it satisfies (2).

Computing a sampling precertificate reduces to an SDP. We show that suitable genericity assumptions guarantee its correctness, thus giving us an $\text{SOS}(\mathcal{V})$ certificate. An interesting feature of our sampling methodology is that the only information needed about the variety is a *sampling oracle*, i.e., a procedure that generates generic samples. Note that sampling points is very simple when the variety has a known parametrization (e.g., for $SO(n)$, Grassmannians, or rank k tensors). For a general variety \mathcal{V} , the field of *numerical algebraic geometry* provides practical methods to sample generic points [32].

Contributions. This paper presents the following contributions.

- We introduce a new methodology to compute SOS certificates over an algebraic variety \mathcal{V} . This is a geometric formulation that represents \mathcal{V} with a generic set of complex samples, instead of relying on some algebraic description. In this way, we avoid algebraic issues such as multiplicity and dependence on the specific generators used. We analyze the correctness of our formulation, establishing sufficient conditions on the samples and the variety.
- Our methodology takes advantage of the coordinate ring structure to simplify the SDP. Moreover, it is the first such relaxation independent of Gröbner bases. This makes our methods appealing for many varieties that are easy to sample from but for which Gröbner bases computation is intractable. Examples of such varieties include $SO(n)$, Stiefel manifolds, Grassmannians, and secant varieties.
- We apply for the first time techniques from numerical algebraic geometry to SOS programs. In this way, we inherit some of the main strengths from this area. We highlight that these methods are trivially parallelizable, since they rely on homotopy continuation of many independent paths. They also allow

us to work with straight-line programs (i.e., polynomials do not need to be expanded).

Related work. Polynomial optimization problems have attracted major research in past years. SDP relaxations based on the SOS/moments method of [16, 23] constitute the major trend of study. The SOS literature is vast; we refer the reader to [4, 17, 18] for an overview.

Our sampling SOS methodology extends the ideas from Löfberg and Parrilo in [20], where they first consider sampling formulations for unconstrained SOS problems. They show that sampling formulations offer some numerical advantages over the standard approach. Most remarkably, the SDP has a low rank structure, which leads to a significant complexity improvement in interior point methods. In particular, low rank structure is exploited in the solvers SDPT3 and DSDP [3, 34]. Second, the SDP is usually better conditioned, as it relies on a set of orthogonal polynomials instead of a monomial basis. These properties make sampling formulations appealing, as seen in [19, 25, 26]. We will see that these properties are preserved in the variety case considered in this paper. We remark that our use of the samples differs from [20] in that for us, samples carry additional information about the underlying variety \mathcal{V} .

Different methods have been proposed to reduce complexity in SOS programs, in particular by exploiting symmetries, sparsity, and quotient ring structure; see [4, section 3.3], [18, section 8], and the references therein. This paper is only concerned with the last item, but we point out that all these techniques can be combined together. The Gröbner bases method to compute quotient ring SOS certificates was introduced in [22]; some further improvements were made in [24]. This is the default method for several varieties with simple Gröbner bases, particularly from combinatorial optimization [18, section 8]. Quotient ring methods have also been used for unconstrained optimization [21]. We point out that there was no “direct” method (without computing the radical \sqrt{I}) to obtain SOS certificates on the coordinate ring (i.e., $\text{SOS}(\mathcal{V})$ certificates).

Although the existence (or degree bounds) of SOS certificates is beyond the scope of this paper (see, e.g., [29]), we review some known results for completeness. In particular, $\text{SOS}(\mathcal{V})$ certificates exist if (i) \mathcal{V} is zero-dimensional, (ii) \mathcal{V} is one-dimensional and p is both strictly positive and bounded [31], (iii) \mathcal{V} is compact and p is strictly positive [30], and (iv) \mathcal{V} is a variety of minimal degree and p is quadratic [5]. For most varieties there exist nonnegative polynomials which are not SOS. Nonetheless, such instances can always be approximated by SOS polynomials (possibly of higher degree) [17, section 2.6].

Solution outline. Our approach to compute $d\text{-SOS}(\mathcal{V})$ certificates follows three main steps.

- (i) *Sampling:* Obtain a “good” set of samples Z on the variety. It will be sufficient for us to consider generic (random) samples on each component of the variety.
- (ii) *SDP:* Given a sample set Z , compute a sampling precertificate (F, Z) using an SDP.
- (iii) *Verification:* Check that the precertificate (F, Z) is correct. This reduces to the identity testing problem.

The structure of this paper is as follows. Section 2 presents some basic algebraic preliminaries. Afterwards, we approach each of the problems from above, although in a different order to simplify the exposition. Section 3 deals with (ii), section 4

with (iii), and section 5 with (i). Section 6 presents the complete sampling SOS methodology. Finally, section 7 shows several examples to illustrate our methods.

2. Preliminaries.

2.1. Algebraic geometry. Let \mathbb{K} denote a field which is either \mathbb{R} or \mathbb{C} , and let $\mathbb{K}[x] = \mathbb{K}[x_1, \dots, x_n]$ denote the ring of polynomials with coefficients in \mathbb{K} . The *ideal* generated by a set of polynomials $h = \{h_1, \dots, h_m\} \subseteq \mathbb{K}[x]$ is

$$I = \langle h \rangle := \{\sum_i g_i h_i : g_i \in \mathbb{K}[x]\}.$$

The *quotient ring* $\mathbb{K}[x]/I$ is the set of equivalence classes where $f \sim_I g$ if $f - g \in I$.

Given a set of polynomials $h \subseteq \mathbb{K}[x]$, its complex *algebraic variety* is

$$\mathcal{V} = \mathcal{V}_{\mathbb{C}}(h) := \{x \in \mathbb{C}^n : h_i(x) = 0 \text{ for } h_i \in h\}.$$

The corresponding real variety is $\mathcal{V} \cap \mathbb{R}^n$. Note that $\mathcal{V}_{\mathbb{C}}(h) = \mathcal{V}_{\mathbb{C}}(\langle h \rangle)$. In this paper we only consider complex varieties defined by real polynomials. It is easy to see that a complex variety \mathcal{V} can be defined by real polynomials if and only if it is *self-conjugate*, i.e., its complex conjugate $\bar{\mathcal{V}}$ is itself.

The *coordinate ring* of a variety $\mathcal{V} \subseteq \mathbb{C}^n$ is the quotient ring $\mathbb{K}[\mathcal{V}] := \mathbb{K}[x]/\mathbf{I}_{\mathbb{K}}(\mathcal{V})$, where $\mathbf{I}_{\mathbb{K}}(\mathcal{V})$ is the vanishing ideal

$$\mathbf{I}_{\mathbb{K}}(\mathcal{V}) := \{f \in \mathbb{K}[x] : f(x) = 0 \text{ for all } x \in \mathcal{V}\}.$$

Equivalently, $\mathbb{K}[\mathcal{V}]$ is the set of equivalence classes of polynomials where $f \sim_{\mathcal{V}} g$ if they define the same function on \mathcal{V} .

Remark 2.1. Hilbert's Nullstellensatz implies that $\mathbf{I}_{\mathbb{K}}(\mathcal{V}_{\mathbb{C}}(I)) = \sqrt{I}$ for any ideal $I \subseteq \mathbb{K}[x]$. It follows that $\mathbb{K}[\mathcal{V}]$ is equal to the quotient ring $\mathbb{K}[x]/I$ only if I is radical.

We say that a variety $\mathcal{V} \subseteq \mathbb{C}^n$ is *irreducible* if it is not the union of two proper varieties. Note that any variety parametrized by \mathbb{C}^n is irreducible. An arbitrary variety can be decomposed in a unique way in the form

$$\mathcal{V} = \mathcal{V}_1 \cup \dots \cup \mathcal{V}_r, \quad \text{where } \mathcal{V}_i \not\subseteq \mathcal{V}_j \text{ for } i \neq j.$$

The varieties \mathcal{V}_i are called the *irreducible components* of \mathcal{V} . If \mathcal{V} is self-conjugate, then either \mathcal{V}_i is also self-conjugate, or there is a pair $(\mathcal{V}_i, \mathcal{V}_j)$ of conjugate components.

2.2. Sampling varieties. Our technique requires a sampling oracle for the complex variety \mathcal{V} . More precisely, we need to sample generic (random) points in each irreducible component of \mathcal{V} . Observe that sampling points is easy whenever we have a parametrization. For instance, we can sample points from $SO(n)$ using the *Cayley parametrization*:

$$(4) \quad A \mapsto (\text{id}_n - A)(\text{id}_n + A)^{-1} \quad \text{for skew symmetric } A.$$

Other parametric varieties include Grassmannians, Stiefel manifolds, and secant varieties.

For a general variety \mathcal{V} , a practical way to compute sample points is through the tools of numerical algebraic geometry; we refer the reader to [2, 32] for an introduction. Homotopy continuation tools such as Bertini [1] and PHCpack [35] allow one to compute the irreducible decomposition of \mathcal{V} , and afterwards to sample an arbitrary number of points in any component. Typically the most expensive part is to produce

the decomposition; sampling points is relatively cheap. These numerical methods offer the following advantages with respect to symbolic methods such as Gröbner bases: they are trivially parallelizable (each path can be tracked independently), they allow for straight-line programs (polynomials do not need to be expanded), and they offer better numerical stability.

Remark 2.2 (complex samples). Even though we are only interested in real polynomials, our methods allow the sample points to be complex. This is an important feature, since computing real points on a variety is significantly harder than computing complex points.

Remark 2.3 (zero-dimensional case). The results from this paper are most useful for positive-dimensional varieties, particularly if the number of components is relatively small. The reason is that we treat each irreducible component separately. In particular, we take care of the zero-dimensional part of the variety exhaustively; i.e., we check for all such points that $p(x)$ is indeed nonnegative. If the whole variety is zero-dimensional, our algorithm reduces to a brute-force search.

2.3. SOS certificates on varieties. Consider a variety \mathcal{V} defined by equations $h = \{h_j\}_j$, and let $I = \langle h \rangle$ be the generated ideal. There are two traditional SOS methods to certify nonnegativity on $\mathcal{V} \cap \mathbb{R}^n$. An *equations d-SOS* certificate is a tuple of polynomials (F, g_1, \dots, g_m) such that

$$(5) \quad p(x) = F(x) + \sum_j g_j(x)h_j(x); \quad F(x) \text{ is SOS}; \quad \deg(F), \deg(g_j h_j) \leq 2d.$$

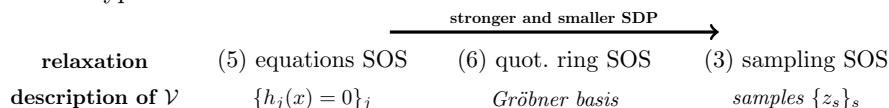
Finding such a certificate reduces to an SDP [23]. A *quotient ring d-SOS* certificate is a polynomial F such that

$$(6) \quad p - F \in I \quad (\text{i.e., } p \equiv F \bmod I); \quad F(x) \text{ is SOS}; \quad \deg(F) \leq 2d.$$

Given a Gröbner basis of I , the above reduces to an SDP [4, section 3.3.5]. For an introduction to Gröbner bases and quotient ring computations, we refer the reader to [9].

Quotient ring formulations are appealing for two main reasons. First, they are *stronger* than equations SOS (i.e., if (5) is feasible, then so is (6), but the converse is not true). Second, the size of the associated SDP is *smaller*, not only because of the absence of the equations g_j , but also since it takes into account the structure of the quotient ring. Consequently, quotient ring SOS has become the default approach for varieties with simple Gröbner bases (e.g., the hypercube $\{0, 1\}^n$). However, the expense of Gröbner bases computation limits its application to further cases.

Our sampling SOS methodology can be seen as a “better” quotient ring formulation because we work modulo the *radical* ideal $\sqrt{I} = \mathbf{I}(\mathcal{V})$, and thus the underlying space is the coordinate ring. The following diagram illustrates the relations among these three types of certificates:



Remark 2.4 (Hilbert function). For $k \in \mathbb{N}$, let $\mathcal{L}_k \subseteq \mathbb{R}[\mathcal{V}]$ be the linear space spanned by the polynomials of degree at most k . The function $H_{\mathcal{V}}(k) := \dim(\mathcal{L}_k)$ is known as the *Hilbert function*, and it plays an important role in sampling SOS (also

in quotient ring SOS). Indeed, the size of the positive semidefinite (PSD) matrix in the SDP is precisely $H_{\mathcal{V}}(d)$. We will also see in section 5 that the number of samples we require is given by $H_{\mathcal{V}}(2d)$. The Hilbert function can be bounded as follows [6]:

$$(7) \quad H_{\mathcal{V}}(k) \leq \binom{n+k}{k} \quad \text{and} \quad H_{\mathcal{V}}(k) \leq \deg \mathcal{V} \binom{\dim \mathcal{V} + k}{k} \quad \text{if } \mathcal{V} \text{ is equidimensional,}$$

where \deg, \dim denote the degree and dimension. The second bound implies that, for fixed \mathcal{V} , the size of the PSD matrix in sampling SOS is $O(d^{\dim \mathcal{V}})$. In contrast, for equations SOS we get $O(d^n)$.

3. Computing precertificates. In this section we show how, given a candidate sample set Z , computing sampling SOS precertificates reduces to an SDP. We will also determine what condition we need on the sample set in order for such precertificates to be correct. The answer will be given by the concept of *poisedness* from polynomial interpolation. Finally, we will show how to reduce the size of the SDP in order to take advantage of the coordinate ring structure.

3.1. Sampling SDP. For a degree bound d , let $u(x) \in \mathbb{R}[x]^N$ denote the vector with all $N = \binom{n+d}{d}$ monomials of degree at most d . Recall that a polynomial $F \in \mathbb{R}[x]$ is d -SOS if and only if

$$F(x) = Q \bullet u(x)u(x)^T$$

for some positive semidefinite matrix Q (denoted $Q \succeq 0$), where the notation \bullet represents trace inner product [23]. Computing a polynomial F satisfying (3) reduces to the following SDP:

$$(8) \quad \begin{array}{ll} \text{find} & Q \in \mathcal{S}^N, \quad Q \succeq 0 \\ \text{subject to} & p(z_s) = Q \bullet u(z_s)u(z_s)^T \quad \text{for } s = 1, \dots, S \end{array}$$

where \mathcal{S}^N denotes the space of $N \times N$ real symmetric matrices. Note that the matrix Q is real, whereas $p(z_s)$ and $u(z_s)$ are complex. Thus, each equality imposes a constraint on both the real and the imaginary part, i.e.,

$$\Re(p(z_s)) = Q \bullet \Re(u(z_s)u(z_s)^T), \quad \Im(p(z_s)) = Q \bullet \Im(u(z_s)u(z_s)^T).$$

The above SDP has two important features: the polynomial p can be given as a *straight-line program* (i.e., it does not need to be expanded) and the constraint matrices have *low rank*. Indeed, the rank of the constraint matrices $\Re(u(z_s)u(z_s)^T)$ and $\Im(u(z_s)u(z_s)^T)$ is at most two. This special rank structure can be exploited in interior point methods, as discussed in [3, 20, 26]. In particular, the Hessian assembly takes only $O(N^3)$ operations for low rank matrices, as opposed to $O(N^4)$ for unstructured matrices.

Observe that the monomial vector $u(x)$ can be replaced by any other polynomial set with the same linear span. In particular, we will see in section 3.3 that $u(x)$ can be chosen to be an orthogonal basis with respect to a natural inner product supported on the samples. Remarkably, this orthogonalization reduces complexity in the SDP by exploiting the algebraic *dependencies* of the coordinate ring $\mathbb{R}[\mathcal{V}]$. In addition, the conditioning of the problem might improve, as explained in [20] for the unconstrained case $\mathcal{V} = \mathbb{C}^n$.

Remark 3.1 (kernel/image form). The feasible set of (8) has the form $Q \succeq 0, Q \in \mathcal{Q}$, where

$$\mathcal{Q} = \{Q \in \mathcal{S}^N : Q \bullet A_i = b_i\}$$

is an affine subspace. We refer to the above representation of \mathcal{Q} as the *kernel form*. Alternatively, we can describe \mathcal{Q} explicitly by giving some generators, i.e.,

$$\mathcal{Q} = \left\{ Q_0 + \sum_j \lambda_j Q_j : \lambda_j \in \mathbb{R} \right\},$$

where $Q_0 \bullet A_i = b_i$ and $Q_j \bullet A_i = 0$. We refer to this representation as the *image form*. Depending on the problem, either of them might be more convenient. In particular, if the number of constraints is close to the dimension of \mathcal{S}^N , then the latter representation is more compact. This will be the case in the applications shown in sections 7.2 and 7.3. For a given problem, we can decide which representation is better by estimating the number of variables used in both of them, as discussed in [23].

3.2. Poisedness implies correctness. We just showed how to compute a sampling SOS precertificate for a given sample set. However, this precertificate might be incorrect unless we are cautious with the sample set, as illustrated in the next example.

Example 3.2 (incorrect precertificate). Let $\mathcal{V} \subseteq \mathbb{C}^2$ be the zero set of $h(x) := x_2^2 - 1$ that consists of two complex lines $\mathbb{C} \times \{1\}$ and $\mathbb{C} \times \{-1\}$. Let $p(x) := x_1^2 - x_2 + 1$, which is nonnegative on $\mathcal{V} \cap \mathbb{R}^2$. Let $Z := \{(k, 1)\}_{k=1}^5 \subseteq \mathcal{V}$ be a set of samples, and let $F(x) := x_1^2$. Observe that (F, Z) is a sampling SOS precertificate, but it is not correct because $p(0, -1) \neq F(0, -1)$. This example illustrates that a sample set, regardless of its size, might lead to incorrect precertificates if it does not capture well the geometry of the variety (in this case, Z misses one of the components of \mathcal{V}).

We now present a condition that guarantees the correctness of a precertificate. Let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$ be the coordinate ring of the variety, which is the space in which we will work. In particular, we will see the entries of the polynomial vector $u(x)$, as well as $p(x)$, as elements of \mathcal{R} . We need the following definition.

DEFINITION 3.3. Let $\mathcal{V} \subseteq \mathbb{C}^n$ be a self-conjugate variety, and let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$. Let $\mathcal{L} \subseteq \mathcal{R}$ be a linear subspace, and let $Z \subseteq \mathcal{V}$ be a set of samples. We say that (\mathcal{L}, Z) is *poised*¹ if the only polynomial $q \in \mathcal{L}$ such that $q(z) = 0$ for all $z \in Z$ is the zero polynomial.

Remark 3.4. For any finite-dimensional \mathcal{L} , there is a finite set Z such that (\mathcal{L}, Z) is poised.

Let $\mathcal{L}_d \subseteq \mathcal{R}$ be the linear space spanned by the entries of $u(x)$, and let $\mathcal{L}_{2d} \subseteq \mathcal{R}$ be spanned by the entries of $u(x)u(x)^T$. Note that $F(x) = Q \bullet u(x)u(x)^T \in \mathcal{L}_{2d}$. The following proposition tells us that poisedness guarantees the correctness of a sampling SOS precertificate. Thus, a *good set of samples* is one such that (\mathcal{L}_{2d}, Z) is poised.

PROPOSITION 3.5. Let $\mathcal{V} \subseteq \mathbb{C}^n$ be a self-conjugate variety, let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$, and let $p \in \mathcal{R}$ be nonnegative on $\mathcal{V} \cap \mathbb{R}^n$. Let (F, Z) be a sampling SOS precertificate, and let $\mathcal{L}_{2d} \subseteq \mathcal{R}$ be a linear subspace such that $p, F \in \mathcal{L}_{2d}$. If (\mathcal{L}_{2d}, Z) is poised, then (F, Z) is correct.

¹In polynomial interpolation it is usually further required that $|Z| = D$, where D is the dimension of \mathcal{L} [27]. We do not impose such a condition.

Proof. Let $g := p - F \in \mathcal{L}_{2d}$, and observe that $g(z) = 0$ for $z \in Z$. As (\mathcal{L}_{2d}, Z) is poised, this implies that $g = 0$ and thus $p = F \in \mathcal{R}$. \square

For the rest of this section we assume that the poisedness condition from above is satisfied. In section 5 we will discuss how to choose the samples in order to satisfy this requirement.

3.3. Reducing complexity. The size of the PSD matrix Q from (8) is $\binom{n+d}{d}$. We can reduce the size of this matrix by taking advantage of the coordinate ring structure. The size of the new matrix will be given by the Hilbert function $H_{\mathcal{V}}(d)$; see Remark 2.4. To do so, we simply need to find a basis of the linear subspace $\mathcal{L}_d \subseteq \mathcal{R}$ spanned by the entries of $u(x)$. We now explain how to get an orthogonal basis $u^o(x)$ with respect to the inner product given in the next proposition.

PROPOSITION 3.6. *Let $\mathcal{V} \subseteq \mathbb{C}^n$ be a self-conjugate variety, and let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$. Let $\mathcal{L}_d \subseteq \mathcal{R}$ be a linear subspace, and let $Z \subseteq \mathcal{V}$ be a set of samples. Let $\langle \cdot, \cdot \rangle_Z : \mathcal{L}_d \times \mathcal{L}_d \rightarrow \mathbb{R}$ be*

$$\langle f, g \rangle_Z = \sum_{z \in Z} (f(z)g(\bar{z}) + f(\bar{z})g(z)).$$

If (\mathcal{L}_d, Z) is poised, then $(\mathcal{L}_d, \langle \cdot, \cdot \rangle_Z)$ is a real inner product space.

Proof. It is clear that $\langle \cdot, \cdot \rangle_Z$ is bilinear and symmetric. Thus, we need only check positiveness. Observe that $\langle f, f \rangle_Z = \sum_{z \in Z} 2|f(z)|^2 \geq 0$, which is zero only if $f(z) = 0$ for all $z \in Z$. As $f \in \mathcal{L}_d$, the poisedness condition implies $f = 0$. \square

Remark 3.7. Note that if (\mathcal{L}_{2d}, Z) is poised, then (\mathcal{L}_d, Z) is also poised.

To find an orthogonal basis, we will operate on the evaluation matrix U with columns $u(z)$ for $z \in Z$. Consider the real matrix $W := [\Re(U) | \Im(U)]$. Observe that $u(x)$ is an orthogonal basis with respect to $\langle \cdot, \cdot \rangle_Z$ if and only if the rows of W are orthogonal with respect to the standard real inner product. Thus, we just need to orthogonalize the rows of W . Using an SVD (or rank revealing QR), we can obtain a decomposition $W = TW^o$, where W^o has orthogonal rows and T is a real full rank transformation matrix. Let U^o be such that $W^o = [\Re(U^o) | \Im(U^o)]$. The matrix U^o encodes the new vector of orthogonal polynomials $u^o(x)$. We note that directly orthogonalizing the matrix U does not work, as the transformation matrix T would be complex.

Algorithm 1 Orthogonal basis on the coordinate ring.

Input: Polynomial vector $u(x)$, samples Z of variety \mathcal{V}

Output: Orthogonal basis $u^o(x)$ and its evaluation matrix U^o

- 1: **procedure** ORTHBASIS($u(x), Z$)
 - 2: $U :=$ evaluation matrix with columns $u(z)$ for $z \in Z$
 - 3: $W := [\Re(U) | \Im(U)]$
 - 4: orthogonalize $W =: TW^o$, where $W^o(W^o)^T = \text{id}$
 - 5: let U^o be such that $W^o = [\Re(U^o) | \Im(U^o)]$
 - 6: let $u^o(x)$ be such that $u(x) = Tu^o(x)$
 - 7: **return** $u^o(x), U^o$
-

Example 3.8. Let \mathcal{V} be the complex variety of the set of rotation matrices $SO(2)$, i.e.,

$$(9) \quad \mathcal{V} = \{X \in \mathbb{C}^{2 \times 2} : X^T X = \text{id}_2, \det(X) = 1\}.$$

Let $p(X) = 4X_{21} - 2X_{11}X_{22} - 2X_{12}X_{21} + 3$, which is nonnegative on $\mathcal{V} \cap \mathbb{R}^{2 \times 2}$. We want to find a sampling SOS certificate. We can sample points on \mathcal{V} using the Cayley parametrization (4). Consider the following three complex samples:

$$z_1 = \begin{bmatrix} -0.6+0.8i & 1.2+0.4i \\ -1.2-0.4i & -0.6+0.8i \end{bmatrix}, \quad z_2 = \begin{bmatrix} -1.2+0.4i & 0.6+0.8i \\ -0.6-0.8i & -1.2+0.4i \end{bmatrix}, \quad z_3 = \begin{bmatrix} -0.75+0.25i & 0.75+0.25i \\ -0.75-0.25i & -0.75+0.25i \end{bmatrix}.$$

We fix the degree bound $d = 1$ and let $u(x) = (1, X_{11}, X_{12}, X_{21}, X_{22})$ be the monomials of degree at most d . The matrix of evaluations is

$$U = \begin{bmatrix} 1 & 1 & 1 \\ -0.6+0.8i & -1.2+0.4i & -0.75+0.25i \\ -1.2-0.4i & -0.6-0.8i & -0.75-0.25i \\ 1.2+0.4i & 0.6+0.8i & 0.75+0.25i \\ -0.6+0.8i & -1.2+0.4i & -0.75+0.25i \end{bmatrix}.$$

Using an SVD, we obtain the orthogonalized matrix U^o and the corresponding polynomial basis $u^o(x)$. Note that $u^o(x)$ has only three elements, as opposed to $u(x)$:

$$U^o = \begin{bmatrix} -0.5955+0.1005i & -0.5955-0.1005i & -0.5201 \\ 0.3058+0.6116i & -0.3058+0.6116i & 0.2548i \\ -0.0708+0.6411i & -0.0708-0.6411i & 0.4100 \end{bmatrix} \quad u^o(X) = (X_{21} + X_{22} - .8054, X_{21} - X_{22}, X_{21} + X_{22} + 2.4831).$$

The sampling SDP is

$$\begin{aligned} & \text{find} \quad Q \in \mathcal{S}^3, \quad Q \succeq 0 \\ & \text{subject to} \quad p(z_s) = Q \bullet u_s u_s^T \quad \text{for } s = 1, 2, 3, \end{aligned}$$

where u_s denotes the s th column of U^o . Solving the SDP, we obtain the sampling precertificate (F, Z) , where $F(X) = (2X_{21} + 1)^2$.

4. Verifying sampling precertificates. We now address the problem of testing the correctness of a precertificate (F, Z) . This problem is equivalent to determining whether the polynomial $f := p - F$ is identically zero on the variety \mathcal{V} , and it is known as the *identity testing* problem (see, e.g., [28] and the references therein). Note that the problem is nontrivial even if $\mathcal{V} = \mathbb{C}^n$, since f can be given as a straight-line program (such as a determinant). Nonetheless, there is a nice randomized algorithm to solve it, provided that we can efficiently sample the variety. Recall that generic samples can be obtained as explained in section 2.2. We now review the notion of genericity and show as well the solution to the identity testing problem.

4.1. Genericity. The notion of *genericity* is fundamental in algebraic geometry. Let $\mathcal{V} \subseteq \mathbb{C}^n$ be an irreducible variety of positive dimension. We say that a property holds *generically* on \mathcal{V} if there is a nonzero polynomial $q \in \mathbb{C}[\mathcal{V}]$ such that the property holds for any $z \in \mathcal{V}$ such that $q(z) \neq 0$. Informally, this means that the property holds outside of the small bad region given by $q(z) = 0$. In section 5.2 we will use a variation of this notion of genericity that is better suited for dealing with real polynomials.

Example 4.1. Let $\mathcal{V} = \mathbb{C}^{m \times m}$ be the space of $m \times m$ matrices. The property of being “nonsingular” is satisfied generically on \mathcal{V} , since a matrix $A \in \mathcal{V}$ is singular only if $\det(A) = 0$.

We often say that a sample point $z \in \mathcal{V}$ is *generic* if some property of interest (such as the conclusion of a theorem) is satisfied generically on \mathcal{V} . For instance, we may say “a generic $m \times m$ matrix is nonsingular.” A generic point can be understood as a random point on the variety.

PROPOSITION 4.2. *Let \mathcal{V} be an irreducible variety, let $f \in \mathbb{C}[\mathcal{V}]$ be a nonzero polynomial, and let $z \in \mathcal{V}$ be a generic sample. Then $f(z)$ is nonzero.*

Proof. The conclusion holds, except in the bad region defined by $f(z) = 0$. \square

4.2. Identity testing. Genericity allows us to derive *randomized* algorithms that succeed with *probability one* with respect to any distribution on \mathcal{V} with full support. In particular, Proposition 4.2 gives rise to Algorithm 2. This method efficiently solves the identity testing problem for an irreducible variety (provided that it can be sampled). Surprisingly, no efficient deterministic algorithm to this problem is known, and it is likely that finding such an algorithm is very hard [14].

Algorithm 2 Identity testing over \mathbb{C} .

Input: Polynomial $f \in \mathbb{C}[x]$, irreducible variety \mathcal{V}

Output: “True” if f is identically zero on \mathcal{V} . “False” otherwise.

```

1: procedure ISZERO( $f, \mathcal{V}$ )
2:    $z :=$  generic sample from  $\mathcal{V}$ 
3:   return True if  $f(z) = 0$  else False

```

Remark 4.3 (reducible varieties). If the variety \mathcal{V} is reducible, we can still solve the identity testing problem provided that we can sample each of its irreducible components. We simply need to apply Algorithm 2 to each component.

Remark 4.4 (probability one). Randomized algorithms derived from genericity statements provably work with probability one in exact arithmetic. However, in floating point arithmetic there is a nonzero probability of error, thus leading to Monte Carlo algorithms; for further discussion see [32, Chapter 4].

5. Selecting the samples. The missing step to complete our sampling SOS methodology is to describe how to obtain a good set of samples Z . Recall from section 3.2 that a good sample set must be such that (\mathcal{L}_{2d}, Z) is poised. Thus the question we address here is the following: given a linear space $\mathcal{L}_{2d} \subseteq \mathcal{R} = \mathbb{R}[\mathcal{V}]$, how can we get a sample set Z such that (\mathcal{L}_{2d}, Z) is poised? We will see in this section that this condition can be satisfied with a generic set of samples.

5.1. Poisedness again. Before proceeding to the selection of the samples, we first present an alternative characterization of poisedness. Let $\mathcal{L} \subseteq \mathcal{R}$ be a finite-dimensional subspace. Let $v(x) \in \mathcal{R}^N$ be a polynomial vector whose entries span \mathcal{L} . Let U be the matrix with columns $v(z)$ for $z \in Z$, and let $\hat{U} := [U | \bar{U}]$. We will refer to the (complex) rank of matrix \hat{U} as the *empirical dimension* of \mathcal{L} with respect to Z . It is easy to see that it does not depend on the choice of generators.

LEMMA 5.1. *(\mathcal{L}, Z) is poised if and only if the dimension of \mathcal{L} is equal to its empirical dimension.*

Proof. Let $D := \dim(\mathcal{L})$, and assume that $v(x) \in \mathcal{R}^D$ is a basis. Assume first that the $\text{rank}(\hat{U}) = D$. Note that any $q \in \mathcal{L}$ can be written uniquely in the form $q(x) = \mu^T v(x)$ for some $\mu \in \mathbb{R}^D$. The condition that $q(z) = 0$ for $z \in Z \cup \bar{Z}$ implies that $\mu^T \hat{U} = 0$. As \hat{U} has full row rank, then $\mu = 0$, and thus (\mathcal{L}, Z) is poised. Assume now that \hat{U} does not have full row rank. Then there is some nonzero $\lambda \in \mathbb{C}^D$ such that $\lambda^T \hat{U} = 0$. Observe that this implies that $\Re(\lambda)^T U = \Im(\lambda)^T U = 0$, where \Re, \Im denote the real/imaginary part. Thus, there is a nonzero $\mu \in \mathbb{R}^D$ such that $\mu^T U = 0$. Considering the polynomial $q(x) := \mu^T v(x)$, we conclude that (\mathcal{L}, Z) is not poised. \square

Remark 5.2. Since matrix \hat{U} has $2|Z|$ columns, it follows from the lemma that if (\mathcal{L}, Z) is poised, then $2|Z| \geq \dim \mathcal{L}$.

Example 5.3. Let $\mathcal{V} = \mathbb{C}$, and let $\mathcal{R} = \mathbb{R}[x]$ be the space of univariate polynomials. Let \mathcal{L} be the set of polynomials of degree less than N , and let $v(x) = (1, x, \dots, x^{N-1})$.

Let $Z \in \mathbb{C}^{N/2}$ be a tuple of complex samples, and let \hat{Z} be the concatenation of Z and \bar{Z} . The evaluation matrix \hat{U} in this case is the Vandermonde matrix of \hat{Z} , which is singular only if there are repeated elements in \hat{Z} . Therefore, (\mathcal{L}, Z) is poised if and only if the elements of \hat{Z} are all distinct.

5.2. How many samples. We return to the problem of finding a sample set Z such that (\mathcal{L}_{2d}, Z) is poised. As we decided that the samples will be random, the only missing point is to determine how many samples to take. Remark 5.2 tells us that we need at least $\lceil D/2 \rceil$ samples, where $D := \dim(\mathcal{L}_{2d})$. We wonder if this condition is *generically sufficient* to guarantee poisedness.

QUESTION. Let \mathcal{V} be a self-conjugate variety. Let $\mathcal{L}_{2d} \subseteq \mathbb{R}[\mathcal{V}]$ be a D -dimensional linear subspace, and let $Z \subseteq \mathcal{V}$ be a generic set of samples with $|Z| \geq D/2$. Is (\mathcal{L}_{2d}, Z) poised?

In order to make sense of the above question, we have to be more precise about the meaning of a generic set of samples. In section 4.1 we saw the definition of a generic sample of an irreducible variety. We have to extend this definition to multiple samples, taken possibly from a reducible variety. Below we formalize the exact notion of genericity that we use. It is slightly different from the one in section 4.1, as it includes the complex conjugates of the samples. The reason for including the conjugates is that it reflects the fact that we are working with real polynomials.

DEFINITION 5.4. Let $\mathcal{W} \subseteq \mathbb{C}^n$ be an irreducible variety, and let $Z = (z_1, \dots, z_S)$ be a tuple of S samples in \mathcal{W} . We say that Z satisfies a property *c-generically* (conjugate generically) if there is a polynomial $q \in \mathbb{C}[z_1, \dots, z_S, \bar{z}_1, \dots, \bar{z}_S]$ such that

- $q(z_1, \dots, z_S, \bar{z}_1, \dots, \bar{z}_S)$ is not identically zero when $z_1, \dots, z_S \in \mathcal{W}$;
- the property holds whenever $q(z_1, \dots, z_S, \bar{z}_1, \dots, \bar{z}_S) \neq 0$.

Let $\mathcal{W}_1, \dots, \mathcal{W}_r$ be irreducible varieties, and let $Z_1 \subseteq \mathcal{W}_1, \dots, Z_r \subseteq \mathcal{W}_r$ be tuples of samples. We say that (Z_1, \dots, Z_r) satisfies a property *c-generically* if there are polynomials $q_1 \in \mathbb{C}[Z_1, \bar{Z}_1], \dots, q_r \in \mathbb{C}[Z_r, \bar{Z}_r]$ such that

- $q_i(Z_i, \bar{Z}_i)$ is not identically zero on \mathcal{W}_i for $1 \leq i \leq r$;
- the property holds whenever $q_1(Z_1, \bar{Z}_1) \neq 0, \dots, q_r(Z_r, \bar{Z}_r) \neq 0$.

We say that Z (resp., Z_1, \dots, Z_r) is a *c-generic set of samples* if it satisfies some property of interest *c-generically*.

In the next section we will show that for an irreducible variety (or a conjugate pair of irreducible varieties) the answer to the above question is positive. However, for reducible varieties, we need to make sure that we have enough samples in each irreducible component, as will be discussed in section 5.5. Example 3.2 illustrates what might go wrong if we do not have enough samples in some component.

5.3. The irreducible case. Assume now that $\mathcal{V} = \mathcal{W} \cup \bar{\mathcal{W}}$, where $\mathcal{W} \subseteq \mathbb{C}^n$ is an irreducible variety. This means that \mathcal{V} is either a self-conjugate irreducible variety or a conjugate pair of irreducible varieties. In the latter case, note that we can assume, without loss of generality, that $Z \subseteq \mathcal{W}$ by possibly exchanging some samples with their complex conjugates. We show now that if the samples Z are *c-generic* and are at least as many as the dimensionality of the problem, then the poisedness property is satisfied.

THEOREM 5.5. Let $\mathcal{W} \subseteq \mathbb{C}^n$ be an irreducible variety, let $\mathcal{V} = \mathcal{W} \cup \bar{\mathcal{W}}$, and let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$. Let $\mathcal{L}_{2d} \subseteq \mathcal{R}$ be a linear subspace, and let $Z \subseteq \mathcal{W}$ be a *c-generic set of*

samples. If $|Z| \geq D/2$, where $D := \dim(\mathcal{L}_{2d})$, then (\mathcal{L}_{2d}, Z) is poised.²

Proof. Let $v(x) \in \mathcal{R}^D$ be a basis of \mathcal{L}_{2d} . Let $Z_j := \{z_1, \dots, z_j\}$, let $V_j \in \mathbb{C}^{D \times j}$ be the matrix with columns $\{v(z)\}_{z \in Z_j}$, and let $\hat{V}_j := [\Re(V_j) | \Im(V_j)] \in \mathbb{R}^{D \times 2j}$. Also denote $W_j := [\hat{V}_j | \Im(v(z_{j+1}))] \in \mathbb{R}^{D \times 2j+1}$. Because of Lemma 5.1, we just need to show that the matrix \hat{V}_S has rank D . To this end, we will show the following statements:

- If \hat{V}_{j-1} is full rank, then W_{j-1} is full rank c-generically.
- If W_{j-1} is full rank, then \hat{V}_j is full rank c-generically.

Clearly these statements imply that \hat{V}_S is full rank. Given the similarity between them, we only prove the latter.

Let $j \leq D/2$, and assume that W_{j-1} is full rank. We will show that there is a polynomial $Q \in \mathbb{C}[Z_j, \overline{Z_j}]$ which is not identically zero on \mathcal{W} and such that \hat{V}_j is full rank whenever $Q(Z_j, \overline{Z_j}) \neq 0$.

Assume that \hat{V}_j is not full rank. Then there must exist a vector $\lambda \in \mathbb{R}^{2j-1}$ such that

$$v(z_j) + v(\overline{z_j}) = 2\Re(v(z_j)) = W_{j-1}\lambda.$$

As W_{j-1} has less than D columns, there is some nonzero vector $\mu \in \mathbb{R}^D$ in its left kernel. Note that $\mu = \mu(Z_j, \overline{Z_j})$ can be parametrized as a rational function of $Z_j, \overline{Z_j}$, given that W_{j-1} is full rank. Let $q_\mu(x) := \mu^T v(x) \in \mathcal{R}$, which is nonzero due to the linear independence of $v(x)$. Observe that

$$q_\mu(z_j) + q_\mu(\overline{z_j}) = \mu^T W_{j-1} \lambda = 0.$$

As the coefficients of q_μ are rational functions on $Z_j, \overline{Z_j}$, we conclude that the samples satisfy a nonzero algebraic equation $Q \in \mathbb{C}[Z_j, \overline{Z_j}]$. \square

Remark 5.6. If the samples are real, it can be shown in a similar way that we need $|Z| \geq D$.

5.4. Verifying the number of samples. We just showed that, under genericity assumptions, the poisedness property is satisfied whenever we have as many samples as the dimension of the space. Concretely, we need to have $\lceil D/2 \rceil$ complex samples, where $D = \dim(\mathcal{L}_{2d})$. However, as the dimension D is not known a priori, it is uncertain how many samples to take. Therefore, we need some way to estimate such a dimension, and the natural quantity to consider is the empirical dimension D_e . The following corollary gives us a simple test that guarantees $D = D_e$.

COROLLARY 5.7. *Let $\mathcal{W} \subseteq \mathbb{C}^n$ be an irreducible variety, let $\mathcal{V} = \mathcal{W} \cup \overline{\mathcal{W}}$, and let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$. Let $\mathcal{L}_{2d} \subseteq \mathcal{R}$ be a linear subspace, and let $Z \subseteq \mathcal{W}$ be a c-generic set of samples. Let D be the dimension of \mathcal{L}_{2d} , and let D_e be its empirical dimension with respect to Z . If $D_e < 2|Z|$, then (\mathcal{L}_{2d}, Z) is poised (i.e., $D = D_e$).*

Proof. If $2|Z| < D$, it follows from the proof of Theorem 5.5 that $D_e = 2|Z|$. Therefore, we must have that $2|Z| \geq D$, and thus (\mathcal{L}_{2d}, Z) is poised because of Theorem 5.5. \square

The above corollary suggests a simple strategy that is summarized in Algorithm 3. We form the vector $u_2(x) = \text{vec}(u(x)u(x)^T)$, whose entries span \mathcal{L}_{2d} . Then we build the matrix of evaluations \hat{U}_2 with columns $u_2(z)$ for $z \in Z \cup \overline{Z}$. The rank of this

²This theorem is a special instance of the dimensionality problem in polynomial interpolation, and more elaborate versions can be found in the literature [8].

Algorithm 3 Test samples.**Input:** Polynomial vector $u(x)$, samples Z of a variety \mathcal{V} **Output:** “True” if generically we must have that (\mathcal{L}_{2d}, Z) is poised, where $\mathcal{L}_{2d} \subseteq \mathbb{R}[\mathcal{V}]$ is spanned by $u(x)u(x)^T$. “False” if we cannot guarantee it.

- 1: **procedure** GOODSAMPLES($u(x), Z$)
- 2: $\hat{U}_2 :=$ matrix with columns $\text{vec}(u(z)u(z)^T)$ for $z \in Z \cup \overline{Z}$
- 3: **return** False **if** \hat{U}_2 has full column rank **else** True

matrix is the empirical dimension D_e . If \hat{U}_2 does not have full column rank, the above corollary holds.

Remark 5.8. Consider the Hermitian matrix $\hat{U}_2^* \hat{U}_2$, where $*$ denotes the conjugate transpose. This matrix is often much smaller than \hat{U}_2 , and it can be constructed efficiently as

$$\hat{U}_2^* \hat{U}_2 = [\langle u(z_i), u(z_j) \rangle]_{z_i, z_j \in Z \cup \overline{Z}} = (\hat{U}^* \hat{U}) \circ (\hat{U}^* \hat{U}),$$

where \circ denotes the Hadamard product. Therefore, it is practical to use matrix $\hat{U}_2^* \hat{U}_2$ instead of \hat{U}_2 , given that they have the same rank.

Example 5.9. Consider the case of Example 3.8. We used $S = 3$ samples to compute the precertificate. To verify that the number of samples is sufficient, we construct the 6×6 matrix $\hat{U}_2^* \hat{U}_2$, obtaining the following:

$$\begin{bmatrix} 1.5581 & -0.2937+0.2562i & 0.1730-0.1158i & 0.0902+0.1118i & 0.0981 & -0.0676-0.0720i \\ -0.2937-0.2562i & 1.5581 & 0.1730+0.1158i & 0.0981 & 0.0902-0.1118i & -0.0676+0.0720i \\ 0.1730+0.1158i & 0.1730-0.1158i & 0.2535 & -0.0676-0.0720i & -0.0676+0.0720i & 0.1396 \\ 0.0902-0.1118i & 0.0981 & -0.0676+0.0720i & 1.5581 & -0.2937-0.2562i & 0.1730+0.1158i \\ 0.0981 & 0.0902+0.1118i & -0.0676-0.0720i & -0.2937+0.2562i & 1.5581 & 0.1730-0.1158i \\ -0.0676+0.0720i & -0.0676-0.0720i & 0.1396 & 0.1730-0.1158i & 0.1730+0.1158i & 0.2535 \end{bmatrix}$$

The rank of the above matrix is 5, and thus the condition from Corollary 5.7 is satisfied. Therefore, the number of samples is sufficient.

5.5. Reducible varieties. The analysis so far makes an irreducibility assumption on the variety \mathcal{V} . This assumption is satisfied for many varieties, in particular for any variety parametrized by \mathbb{C}^n . Even if \mathcal{V} is not irreducible, we can always work with each of its irreducible components independently. Indeed, note that $p \geq 0$ on some variety if and only if $p \geq 0$ on each irreducible component.

Nonetheless, there are circumstances in which we may not want to impose an irreducibility assumption, for example, if the variety has bad numerical properties and thus its irreducible components cannot be accurately estimated. In such situations, we can repeat the same analysis from before if we have some method that samples points from each irreducible component. For instance, if we intersect the variety \mathcal{V} with a generic hyperplane of complementary dimension, the intersection is a finite set that contains points in each irreducible component. Note that we do not know which component the samples belong to, but we are certain that there is at least one sample in each component.

The following corollary shows that if we have a sample set with enough points on each irreducible component, then (\mathcal{L}_{2d}, Z) is poised.

COROLLARY 5.10. *Let $\mathcal{W} \subseteq \mathbb{C}^n$ be a variety, let $\mathcal{V} = \mathcal{W} \cup \overline{\mathcal{W}}$, and let $\mathcal{R} = \mathbb{R}[\mathcal{V}]$. Let $\mathcal{L}_{2d} \subseteq \mathcal{R}$ be a linear subspace. Let $\mathcal{W} = \mathcal{W}_1 \cup \dots \cup \mathcal{W}_r$ be the irreducible decomposition, and let $Z_1 \subseteq \mathcal{W}_1, \dots, Z_r \subseteq \mathcal{W}_r$ be c -generic sets of samples. If $|Z_i| \geq D/2$ for all i , where $D := \dim(\mathcal{L}_{2d})$, then (\mathcal{L}_{2d}, Z) is poised.*

Proof. Let $f \in \mathcal{L}_{2d}$ be such that $f(z) = 0$ for all $z \in Z$. We want to show that f is the zero polynomial in $\mathbb{C}[\mathcal{V}]$. Let $\mathcal{V}_i := \mathcal{W}_i \cup \overline{\mathcal{W}_i}$, and let $\psi_i : \mathbb{C}[\mathcal{V}] \rightarrow \mathbb{C}[\mathcal{V}_i]$ be the restriction operator. It is clear that the dimension of $\psi_i(\mathcal{L}_{2d})$ is at most D . Thus, Theorem 5.5 says that $(\psi_i(\mathcal{L}_{2d}), Z_i)$ is poised whenever $q_i(Z_i, \overline{Z_i}) \neq 0$, for some polynomial q_i which is nonzero on \mathcal{W}_i . Note that $\psi_i(f)$ evaluates to zero on Z_i , and thus $\psi_i(f)$ must be the zero element in $\mathbb{C}[\mathcal{V}_i]$ whenever $q_i(Z_i, \overline{Z_i}) \neq 0$. Finally, observe that $(\psi_1 \times \cdots \times \psi_k) : \mathbb{C}[\mathcal{V}] \rightarrow \mathbb{C}[\mathcal{V}_1] \times \cdots \times \mathbb{C}[\mathcal{V}_k]$ is injective. We conclude that whenever $q_i(Z_i, \overline{Z_i}) \neq 0$, then $(\psi_1 \times \cdots \times \psi_k)(f)$ is zero and thus f must be zero. \square

6. Computing sampling certificates. We already developed all the tools needed to find a sampling SOS certificate, and we now put them together. Algorithm 4 summarizes our method for the case of an irreducible variety \mathcal{V} . Naturally, the most computationally expensive part is solving the SDP. Recall from Theorem 5.5 that the number of samples required is

$$(10) \quad S_{\min} := \lceil H_{\hat{\mathcal{V}}}(2d)/2 \rceil, \quad H_{\hat{\mathcal{V}}}(2d) := \dim(\mathcal{L}_{2d}) \leq \min \left\{ \binom{n+2d}{2d}, \deg \hat{V} \binom{\dim \hat{V} + 2d}{2d} \right\},$$

where $\hat{\mathcal{V}} = \mathcal{V} \cup \overline{\mathcal{V}}$, and where we used the bound from (7). Since $H_{\hat{\mathcal{V}}}(2d)$ is unknown in general, we use a simple search strategy in Algorithm 4. The algorithm terminates when the number of samples is at least S_{\min} .

In the case of a reducible variety, we might use Algorithm 4 for each of its irreducible components separately. If we cannot reliably identify such components, we need to take into account the considerations from section 5.5.

Remark 6.1 (zero-dimensional case). Note that a zero-dimensional variety is reducible, with each component consisting of a single point. Thus, in such a case our algorithm reduces to a brute-force enumeration over all solutions, and better strategies may exist. The main problem to address is that of producing small poised sets. We leave this as an open problem.

Algorithm 4 Sampling SOS.

Input: Polynomial $p \in \mathbb{R}[x]$ (given by an evaluation oracle), irreducible variety $\mathcal{V} \subseteq \mathbb{C}^n$ (given by a sampling oracle), degree bound $d \in \mathbb{N}$

Output: d -SOS(\mathcal{V}) certificate F , if it exists. “Null” if no certificate exists.

```

1: procedure SAMPLINGSOS( $p, \mathcal{V}, d$ )
2:    $u(x) :=$  vector with all monomials up to degree  $d$ 
3:    $S :=$  initial guess on the number of samples (an upper bound is given in (10))
4:    $Z :=$  generic set of  $S$  samples from  $\mathcal{V}$ 
5:    $u(x) := \text{ORTHBASIS}(u(x), Z)$  ▷ find basis of  $\mathbb{R}[\mathcal{V}]$ 
6:   if not GOODSAMPLES( $u(x), Z$ ) then ▷ check if there are enough samples
7:     increase  $S$  and go to 4
8:    $Q :=$  solution of SDP (8) (if none return Null) ▷ solve SDP
9:    $F(x) := Q \bullet u(x)u(x)^T$ 
10:  if not ISZERO( $p - F, \mathcal{V}$ ) then ▷ verify correctness
11:    the sample set was not generic enough; go to 4
12:  return  $F$ 
```

7. Examples. We now show several examples and numerical evaluations to illustrate our methodology. We implemented our algorithms in MATLAB using SDPT3 [34] to solve the SDPs. We also use Macaulay2 [12] for Gröbner bases and Bertini [1] for numerical algebraic geometry computations. The experiments are performed on an i7 PC with 8 cores of 3.40 GHz, 15.6 GB RAM, running Ubuntu 14.04.

We will compare our techniques with the following two methods: equations SOS (5) and the (Gröbner bases based) quotient ring SOS (6). For the convenience of the reader unfamiliar with these methods, Appendix A illustrates them with concrete, simple examples.

Remark 7.1 (polynomial optimization). Since some of the examples below are polynomial optimization problems, we recall that there are well-studied SOS relaxations [4, 17, 18]. Indeed, minimizing a polynomial $p(x)$ is equivalent to finding the largest γ such that $p(x) - \gamma$ is nonnegative, which can be relaxed to be a d -SOS polynomial. The solution γ^* obtained with the SDP gives a valid lower bound on $p(x)$. SOS lower bounds tend to be very good in practice, being the true minimum in certain applications [4, 17]. Moreover, if the minimizer x^* is unique and the dual matrix is rank one, then x^* might be recovered from the dual variables. In particular, $x^* = \Re(\sum_s y_s^* z_s)$ for the sampling SDP in (8), where y_s^* are the (complex) dual variables of the equality constraints.

7.1. Nilpotent matrices. Let $\mathcal{V} := \{X \in \mathbb{C}^{n \times n} : X^n = 0\}$ be the variety of nilpotent matrices. Let $p(X) := \det(X + \text{id}_n)$, which is nonnegative on \mathcal{V} (moreover, it is identically one). We compare different SOS methodologies to certify this.

First, consider the sampling approach. Let the degree bound $d = 1$, and let us take $S = \binom{n+2}{2}$ samples, which are always sufficient. Note that it is very easy to sample nilpotent matrices. For instance, we can generate a random triangular matrix with zero diagonal, and then apply a similarity transformation. For each sample $X_s \in \mathcal{V}$, we can efficiently evaluate $p(X_s)$ with Gaussian elimination. As $p(X_s) = 1$ for all samples X_s , we will obtain the trivial SOS decomposition $p(X) \equiv_{\mathcal{V}} (1)^2$.

Consider now the Gröbner bases approach. Let $h \subseteq \mathbb{R}[X]$ be the n^2 equations given by $X^n = 0$. We want to compute a Gröbner basis of h . Note, however, that the total number of terms in h is on the order of n^{n+1} , and the polynomials are all of degree n . Therefore, this Gröbner basis computation is extremely complicated.

If we are smarter, we can take a different set of defining equations of \mathcal{V} . Consider the polynomial $Q_X(t) := \det(t \text{id}_n - X) - t^n$, and let $h' \subseteq \mathbb{R}[X]$ be the equations given by the coefficients of $Q_X(t)$. It turns out that h' generates the radical ideal of $\langle h \rangle$, and moreover, it is a Gröbner basis [13, section 7]. However, h' has more than $n!$ terms. Once we have the Gröbner basis h' , we need to compute the normal form of p . To obtain this normal form we need to consider p as a dense polynomial. As both p and h' have on the order of $n!$ terms, performing this reduction is computationally intractable. If we are able to reduce it, we will conclude that $p(X) \equiv_{\mathcal{V}} 1$, as before.

Finally, note that equation SOS suffers from the same problems as the Gröbner bases approach. For this method there is an additional problem, which is that the monomial basis $u(X)$ will be very large in order to account for all the monomials in $p(X)$ and $h(X)$. This problem was avoided in the previous methods because of the quotient ring reductions.

This example illustrates two of the advantages of the sampling formulation: it avoids the algebraic problem of deciding which equations to use (e.g., h vs. h'), and it allows the use of straight-line programs (e.g., Gaussian elimination) for more efficient evaluations.

7.2. Weighted orthogonal Procrustes. We consider a family of optimization problems over varieties of orthogonal matrices. The Stiefel manifold $St(k, \mathbb{R}^n)$ is the set of orthonormal k -frames in \mathbb{R}^n . We identify it with the set of matrices $X \in \mathbb{R}^{n \times k}$ such that $X^T X = \text{id}_k$. Note that we can easily sample points from this variety, for instance, by using the Cayley parametrization. Alternatively, we can orthogonalize a random real matrix.

The weighted orthogonal Procrustes problem, also known as the Penrose regression problem, asks for a matrix $X \in St(k, \mathbb{R}^n)$ that minimizes $\|AXC - B\|$, for some matrices $A \in \mathbb{R}^{m_1 \times n}$, $B \in \mathbb{R}^{m_1 \times m_2}$, $C \in \mathbb{R}^{k \times m_2}$. There is no closed form solution for this problem, and several local optima may exist [7, 36].

Let $u(x)$ consist of all monomials up to some degree bound d . The sampling SDP is

$$\begin{aligned} & \max_{\gamma \in \mathbb{R}, Q \succeq 0} \quad \gamma \\ & \text{subject to} \quad \|AX_s C - B\|^2 - \gamma = Q \bullet u(X_s)u(X_s)^T \quad \text{for } s = 1, \dots, S, \\ & \quad \quad \quad X_s \in St(k, \mathbb{R}^n). \end{aligned}$$

Example 7.2 (see [7, Example 2]). Let $(n, k, m_1, m_2) = (4, 3, 5, 3)$, and consider the matrices

$$A^T = \begin{bmatrix} 0.2190 & 0.0470 & 0.6789 & 0.6793 & 0.9347 \\ 0.3835 & 0.5194 & 0.8310 & 0.0346 & 0.0535 \\ 0.5297 & 0.6711 & 0.0077 & 0.3834 & 0.0668 \\ 0.4175 & 0.6868 & 0.5890 & 0.9304 & 0.8462 \end{bmatrix}, B^T = \begin{bmatrix} 0.6526 & 0.2110 & 0.2229 & -0.4104 & -0.9381 \\ 0.6942 & 0.2204 & 0.2015 & 0.2994 & 1.0943 \\ 0.8299 & 1.1734 & -0.1727 & 0.0474 & -0.2351 \end{bmatrix}, C = \text{id}_3.$$

We consider the degree 1 SOS relaxation. Following Algorithm 4, we find out that $S = 43$ complex (or 85 real) samples are sufficient. More generally, the required number of samples is half of $H_V(2) = \binom{n+k+2}{2} - \binom{k+1}{2}$. By solving the above SDP we obtain a lower bound of 1.118147 on the minimum norm $\|AXC - B\|$. Furthermore, the dual SDP matrix has rank one, and thus we can recover a solution achieving such a lower bound as follows:

$$(X^*)^T = \begin{bmatrix} -0.0895 & 0.7472 & 0.2732 & -0.5992 \\ 0.7726 & -0.1843 & 0.6035 & -0.0702 \\ -0.5277 & 0.0163 & 0.7309 & 0.4324 \end{bmatrix}.$$

Table 1 compares different SDP formulations of the degree 1 SOS relaxation of the weighted orthogonal Procrustes problem. We consider the case where $m_1 = n$ and $m_2 = k$. The table shows the number of variables/constraints and the computation times for the equations SDP and the sampling SDP. The computation is performed on random instances, in which matrices A , B , C are generated from the standard normal distribution. For the sampling SDP we use the image form of the SDP (see section 3.1), given that it has low codimension. We remark that for the sampling SDP we include the preprocessing time, i.e., Algorithms 1 and 3.

We point out that the better performance of the sampling SDP is due to the fact that it makes use of the quotient ring structure. Although a similarly sized SDP could be derived using Gröbner bases, Table 1 shows that Gröbner bases computation is very expensive—much more expensive than solving the (larger) equations SDP. In particular, Macaulay2 ran out of memory for $n = 7$, $k = 5$.

7.3. Trace ratio problem. We now consider a problem on the Grassmannian manifold $Gr(k, \mathbb{R}^n)$, which is the set of all k -dimensional subspaces of \mathbb{R}^n . Note that we can easily sample points on $Gr(k, \mathbb{R}^n)$ by considering the subspace spanned by k random vectors. By identifying a subspace with the orthogonal projection onto it, we can view $Gr(k, \mathbb{R}^n)$ as the set of matrices $X \in \mathcal{S}^n$ satisfying $X^2 = X$ and $\text{tr}(X) = k$,

TABLE 1
Degree 1 SOS relaxations for the weighted orthogonal Procrustes problem.

n	k	Equations SDP			Sampling SDP			Gröbner bases time(s)
		variables	constraints	time(s)	variables	constraints	time(s)	
4	2	178	73	0.52	46	42	0.10	0.00
5	3	682	233	0.65	137	130	0.11	0.03
6	4	1970	576	1.18	326	315	0.15	9.94
7	5	4727	1207	3.56	667	651	0.31	out of memory
8	6	9954	2255	13.88	1226	1204	0.70	out of memory
9	7	19028	3873	42.14	2081	2052	2.11	out of memory
10	8	33762	6238	124.43	3322	3285	5.07	out of memory

so this is indeed a variety. The trace ratio problem looks for the maximizer of $\frac{\text{tr}(AX)}{\text{tr}(BX)}$ on $Gr(k, \mathbb{R}^n)$, for some given matrices $A, B \in \mathcal{S}^n$, $B \succ 0$. This problem arises in machine learning, and it can be efficiently solved by iterative methods, given that it has a unique local maximum [37]. We consider the following variation:

$$\max_{X \in Gr(k, \mathbb{R}^n)} \frac{\text{tr}(AX)}{\text{tr}(BX)} + \text{tr}(CX)$$

for some $A, B, C \in \mathcal{S}^n$, $B \succ 0$. This problem may have several local maxima, and thus local methods may not converge to the global optimum [38, 39].

To obtain an SOS relaxation, note that the problem is equivalent to minimizing γ such that $\text{tr}(BX)(\gamma - \text{tr}(CX)) - \text{tr}(AX)$ is nonnegative on $Gr(k, \mathbb{R}^n)$. Thus, the SDP to consider is

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, Q \succeq 0} \quad & \gamma \\ \text{subject to} \quad & \text{tr}(BX_s)(\gamma - \text{tr}(CX_s)) - \text{tr}(AX_s) = Q \bullet u(X_s)u(X_s)^T \quad \text{for } s = 1, \dots, S, \\ & X_s \in Gr(k, \mathbb{R}^n). \end{aligned}$$

Example 7.3 (see [39, Example 3.1]). Let $n = 3, k = 2$ and consider the matrices A, B, C from below. For the degree bound $d = 1$, Algorithm 4 gives that $S = 8$ complex (or 15 real) samples are sufficient. In general, the number of samples is half of $H_V(2) = \binom{\frac{1}{2}(n^2+n)}{2}$. Solving the above SDP, we get an upper bound of 28.692472. As the dual matrix has rank one, we can recover the optimal solution X^* .

$$A = \begin{bmatrix} 11 & 5 & 8 \\ 5 & 10 & 9 \\ 8 & 9 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & 7 & 7 \\ 7 & 10 & 8 \\ 7 & 8 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 15 & 10 & 9 \\ 10 & 7 & 6 \\ 9 & 6 & 6 \end{bmatrix}, \quad X^* = \begin{bmatrix} 0.61574 & 0.15424 & 0.46132 \\ 0.15424 & 0.93809 & -0.18517 \\ 0.46132 & -0.18517 & 0.44617 \end{bmatrix}.$$

As before, we compare the equations SDP and the sampling SDP of the degree 1 SOS relaxation. Table 2 shows the number of variables/constraints and the computation times on random instances for both methods. It also shows the computation time of Gröbner bases.

7.4. Low rank approximation. Consider the problem of finding the nearest rank k tensor. Let $\mathbb{C}^{n_1 \times \dots \times n_\ell}$ denote the set of tensors of order ℓ and dimensions (n_1, \dots, n_ℓ) , and let $\mathbb{C}_{\leq k}^{n_1 \times \dots \times n_\ell}$ be the closure of the space of tensors of rank at most k . Note that we can easily generate generic samples of rank k tensors. Given a real tensor $T \in \mathbb{R}^{n_1 \times \dots \times n_\ell}$, the rank k approximation problem asks for the nearest point $X \in \mathbb{R}_{\leq k}^{n_1 \times \dots \times n_\ell}$, i.e., the minimizer of $\|T - X\|^2$, where $\|\cdot\|$ denotes the norm of the vectorization.

TABLE 2
Degree 1 SOS relaxations for the trace ratio problem.

n	k	Equations SDP			Sampling SDP			Gröbner bases time(s)
		variables	constraints	time(s)	variables	constraints	time(s)	
4	2	342	188	0.47	56	45	0.10	0.00
5	3	897	393	0.71	121	105	0.11	0.02
6	4	2062	738	1.34	232	210	0.15	0.20
7	5	4265	1277	3.62	407	378	0.19	6.04
8	6	8106	2073	9.06	667	630	0.34	488.17
9	7	14387	3198	23.83	1036	990	0.61	out of memory
10	8	24142	4733	58.17	1541	1485	1.18	out of memory

Let $d := \lfloor k/2 \rfloor + 1$, and let $u(X)$ be the vector with all monomials of degree at most d . Denoting $\varsigma(X) := \|X\|^2$, we consider the following SDP relaxation:

$$\begin{aligned} & \max_{\gamma \in \mathbb{R}, Q \succeq 0} \quad \gamma \\ & \text{subject to} \quad (\|T - X_s\|^2 - \gamma) \varsigma(X_s)^{d-1} = Q \bullet u(X_s)u(X_s)^T \quad \text{for } s = 1, \dots, S, \\ & \quad \quad \quad X_s \in \mathbb{C}_{\leq k}^{n_1 \times \dots \times n_\ell}. \end{aligned}$$

We remark that computing the defining equations of the variety $\mathbb{C}_{\leq k}^{n_1 \times \dots \times n_\ell}$ is very complicated [15]. This means that using traditional SOS methods usually is not possible.

Example 7.4 (see [10, Example 3]). Let $T \in \mathbb{R}^{2 \times 2 \times 2 \times 2}$ be the tensor whose nonzero entries are

$$T_{1111} = 25.1, \quad T_{1121} = 0.3, \quad T_{1212} = 25.6, \quad T_{2111} = 0.3, \quad T_{2121} = 24.8, \quad T_{2222} = 23.$$

Consider the rank one approximation problem. Solving the above SDP ($d = 1, S = 49$) we obtain the lower bound 42.1216 on the minimum distance $\|T - X\|$. From the dual solution we recover the minimizer X^* , whose only nonzero entry is $X_{1212}^* = 25.6$.

Consider now the rank three approximation problem. The above SDP ($d = 2, S = 2422$) gives a lower bound of 23.0000. Again, we can recover the minimizer X^* , whose nonzero entries are

$$X_{1111}^* = 25.1, \quad X_{1121}^* = 0.3, \quad X_{1212}^* = 25.6, \quad X_{2111}^* = 0.3, \quad X_{2121}^* = 24.8.$$

To see that X^* is rank three, note that after removing the entry 25.6 we are left with a 2×2 matrix.

7.5. Certifying infeasibility. Given a complex variety $\mathcal{V} \subseteq \mathbb{C}^n$ consider the problem of certifying that $\mathcal{V} \cap \mathbb{R}^n$ is empty. A *Positivstellensatz* infeasibility certificate consists of showing that the constant polynomial -1 is SOS on the variety \mathcal{V} [23]. For instance, if $\mathcal{V} = \{i, -i\} \subseteq \mathbb{C}$, a Positivstellensatz certificate is that $-1 = x^2$ on the variety \mathcal{V} . We take an approach from numerical algebraic geometry, where we first compute a numerical irreducible decomposition of \mathcal{V} , and then use sampling SOS to obtain the infeasibility certificate. For a given vector $u(x)$ the SDP problem to solve is

$$\begin{aligned} & \text{find} \quad Q \succeq 0 \\ & \text{subject to} \quad -1 = Q \bullet u(z_s)u(z_s)^T \quad \text{for } s = 1, \dots, S, \\ & \quad \quad \quad z_s \in \mathcal{V}. \end{aligned}$$

Example 7.5. Let $\mathcal{V} \subseteq \mathbb{C}^9$ be the positive-dimensional part of the cyclic 9-roots problem. The cyclic 9-roots equations are

$$\begin{aligned} & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 \\ & x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_6 + x_6x_7 + x_7x_8 + x_8x_9 + x_9x_1 \\ & x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_5 + x_4x_5x_6 + x_5x_6x_7 + x_6x_7x_8 + x_7x_8x_9 + x_8x_9x_1 + x_9x_1x_2 \\ & \vdots \\ & x_1x_2x_3x_4x_5x_6x_7x_8 + x_2x_3x_4x_5x_6x_7x_8x_9 + \cdots + x_9x_1x_2x_3x_4x_5x_6x_7 \\ & x_1x_2x_3x_4x_5x_6x_7x_8x_9 - 1 \end{aligned}$$

The zero set of these equations consists of a two-dimensional variety \mathcal{V} of degree 18, and 6156 isolated solutions [11]. We remark that computing a Gröbner basis of these equations is very complicated unless its special structure is exploited. Indeed, Macaulay2 ran out of memory after five hours of computation.

We computed the irreducible decomposition of \mathcal{V} using Bertini; it took 2h, 45m with the default parameters. The variety \mathcal{V} decomposes into three pairs of conjugate irreducible varieties (each pair of degree 6). For each component we proceed to compute a sampling 2-SOS certificate. We require 31 complex samples on each component, which we obtained from Bertini in less than 1s. Note that the upper bound from (10) predicted $\frac{1}{2} \cdot \min\{\binom{13}{4}, 6\binom{6}{4}\} = 45$ samples. For each $j = 0, \dots, 5$, we solved the respective sampling SDP, obtaining an infeasibility certificate of the form

$$-1 = (R_j u(x))^T (R_j u(x)) \quad \text{for } x \in \mathcal{V}_j.$$

This allows us to conclude that each irreducible component of \mathcal{V} is purely complex. For instance, for the first irreducible component \mathcal{V}_0 it takes only 0.74s to obtain the certificate shown in Figure 1.

7.6. Amoeba membership. The (unlog) amoeba $\mathcal{A}_{\mathcal{V}} \subseteq \mathbb{R}_+^n$ of a variety $\mathcal{V} \subseteq \mathbb{C}^n$ is the image of \mathcal{V} under the absolute value function, i.e., $\mathcal{A}_{\mathcal{V}} = \{|z| : z \in \mathcal{V}\}$. The amoeba membership problem is to determine whether some point $\lambda \in \mathbb{R}_+^n$ belongs to $\mathcal{A}_{\mathcal{V}}$. Theobald and De Wolff recently proposed the use of Positivstellensatz certificates to prove that $\lambda \notin \mathcal{A}_{\mathcal{V}}$ [33]. We now briefly describe this approach.

For some $f \in \mathbb{C}[z]$, let $\Re(f), \Im(f) \in \mathbb{R}[x, y]$ be such that

$$f(x + iy) = \Re(f)(x, y) + i\Im(f)(x, y).$$

Consider the following sets of equations in $\mathbb{R}[x, y]$:

$$J_{\mathcal{V}} := \{\Re(f_j), \Im(f_j)\}_{j=1}^m, \quad h_{\lambda} := \{x_i^2 + y_i^2 - \lambda_i^2\}_{i=1}^n,$$

where f_j are the defining equations of \mathcal{V} . Theobald and De Wolff suggest computing a Gröbner basis of $J_{\mathcal{V}} \cup h_{\lambda}$ and then searching for a Positivstellensatz infeasibility certificate.

Consider the following approach based on a set of samples $Z \subseteq \mathcal{V}$. Let $\hat{\mathcal{V}} \in \mathbb{C}^{2n}$ be the zero set of $J_{\mathcal{V}} \subseteq \mathbb{R}[x, y]$. Note that if $z \in \mathcal{V}$, then $(\Re(z), \Im(z)) \in \hat{\mathcal{V}}$. Thus, given some monomial vectors $u(x, y)$ and $v(x, y)$, we can formulate the following SDP:

$$\begin{aligned} & \text{Find} \quad Q \succeq 0, \quad C \\ & \text{subject to} \quad -1 = Q \bullet u(x_s, y_s)u(x_s, y_s)^T + h_{\lambda}(x_s, y_s)^T C v(x_s, y_s) \quad \text{for } s = 1, \dots, S, \\ & \quad z_s = x_s + iy_s \in \mathcal{V}. \end{aligned}$$

$$R_0 = \begin{bmatrix} u(x) = (x_8^2, x_7x_9, x_6^2, x_5x_9, x_5x_7, x_4^2, x_3x_6, x_2x_7, x_2x_6, x_2^2, x_1x_3, x_1^2, x_8, x_7, x_6, x_5, x_3, x_1, 1) \\ \begin{array}{l} -0.9638686 \quad -0.3445318 \quad 0.8395791 \quad -1.9531033 \quad 0.6329543 \quad -0.0152284 \quad 0.0238164 \quad 0.4701138 \quad -1.9766327 \quad -0.8363703 \\ -0.3474835 \quad -0.3993919 \quad 0.5501348 \quad -1.2198730 \quad -0.2314149 \quad 0.0354563 \quad 1.0080575 \quad 0.4018444 \quad 1.0316339 \quad -0.6193326 \\ 0.0117704 \quad -0.5278490 \quad 0.6157589 \quad -0.3131173 \quad 0.2207819 \quad -0.0080541 \quad 0.4038186 \quad 0.1500184 \quad -0.2618475 \quad 0.3089739 \\ -0.0131866 \quad 0.1597228 \quad 0.1191077 \quad -0.1088218 \quad 0.0697348 \quad -0.1149430 \quad -0.5067092 \quad 0.1883695 \quad -0.5993569 \quad 0.0244521 \\ -0.4504113 \quad 0.0761266 \quad 0.0056933 \quad 0.1535964 \quad -0.0860039 \quad 0.0007534 \quad 0.1264270 \quad 0.0880389 \quad -0.0927822 \quad -0.1429983 \\ -0.0804265 \quad 0.1450405 \quad -0.0077285 \quad -0.1657304 \quad -0.3240087 \quad 0.0014097 \quad 0.0631496 \quad -0.4083965 \quad 0.0191162 \quad 0.0854950 \\ 0.0192110 \quad 0.1019831 \quad -0.1208989 \quad -0.1821975 \quad -0.1203214 \quad 0.0405222 \quad 0.0595267 \quad -0.1921851 \quad 0.0669972 \quad -0.1710978 \\ -0.1242984 \quad 0.1450764 \quad -0.2725352 \quad -0.1145423 \quad 0.0498037 \quad 0.0036466 \quad -0.0705293 \quad 0.2012444 \quad 0.0873671 \quad -0.2016367 \\ -0.0724047 \quad -0.0072012 \quad -0.0659910 \quad 0.1174698 \quad -0.0830511 \quad 0.0339559 \quad 0.1872153 \quad -0.0010648 \quad -0.1488432 \quad 0.1014557 \\ -0.1440253 \quad 0.0026597 \quad -0.1198142 \quad 0.0147434 \quad 0.1104305 \quad 0.0249783 \quad 0.0246079 \quad -0.0286915 \quad -0.0633959 \quad 0.0786306 \\ 0.0872131 \quad -0.0503333 \quad -0.0426310 \quad -0.0108485 \quad -0.1538320 \quad 0.0351373 \quad 0.0895051 \quad 0.0994606 \quad -0.0858176 \quad 0.0033518 \\ 0.0508126 \quad 0.0850471 \quad -0.0581353 \quad -0.0654513 \quad 0.0413446 \quad 0.0119532 \quad 0.0757765 \quad 0.0459333 \quad -0.0169990 \quad 0.1009677 \\ 0.0112157 \quad 0.0251923 \quad -0.0096306 \quad -0.0680984 \quad -0.0005761 \quad 0.0111481 \quad -0.0373533 \quad 0.0293210 \quad 0.0134771 \quad 0.1119457 \\ 0.0240035 \quad -0.1089833 \quad -0.0750114 \quad -0.0316807 \quad 0.0593823 \quad -0.0045386 \quad -0.0385441 \quad -0.0541272 \quad 0.0162211 \quad 0.0093965 \\ 0.0727416 \quad 0.0067146 \quad -0.0258618 \quad 0.0206007 \quad 0.0284529 \quad -0.0125337 \quad 0.0450957 \quad -0.0142651 \quad -0.0460162 \quad -0.0377493 \\ 0.0018262 \quad 0.0096039 \quad 0.0092749 \quad -0.0098153 \quad 0.0116513 \quad 0.0124708 \quad -0.0166840 \quad 0.0307406 \quad 0.0039079 \quad -0.0005090 \\ 0.0167276 \quad 0.0418916 \quad 0.0339853 \quad 0.0127189 \quad 0.0353915 \quad 0.0352643 \quad -0.0230590 \quad 0.0037635 \quad -0.0020096 \quad -0.0118329 \\ -0.0000118 \quad 0.0028062 \quad 0.0010184 \quad -0.0000054 \quad 0.0000008 \quad -0.0076840 \quad -0.0004971 \quad 0.0000191 \quad 0.0000175 \quad 0.0000097 \\ 0.0000403 \quad 0.0016845 \quad -0.0003547 \quad -0.0000443 \quad 0.0000077 \quad 0.0102089 \quad 0.0023837 \quad -0.0001760 \quad -0.0000343 \quad -0.0000479 \end{array} \end{bmatrix}$$

FIG. 1. *Positivstellensatz infeasibility certificate for the cyclic 9-roots problem.*

Example 7.6. Let $\mathcal{V} \subseteq \mathbb{C}^{nk}$ be the complex variety associated to the Stiefel manifold $St(k, \mathbb{R}^n)$. Let $\lambda = (1/n, 1/n, \dots, 1/n)$, and let us show that $\lambda \notin \mathcal{A}_{\mathcal{V}}$ using the SDP from above. We consider the degree 1 SOS relaxation for the case $n = 6, k = 4$. We require 1205 complex samples on \mathcal{V} , which we obtain using the Cayley parametrization. It takes only 0.79s to compute the Positivstellensatz certificate from below. On the other hand, Macaulay2 ran out of memory while computing a Gröbner basis of $J_{\mathcal{V}}$.

$$\begin{aligned} -1 &= (Ru(x, y))^T (Ru(x, y)) - 1.2 \sum_{i=1}^6 h_i(x, y) \quad \text{for } (x, y) \in \hat{\mathcal{V}}, \\ u(x, y) &= (y_6, y_5, y_4, y_3, y_2, y_1), \quad h_i(x, y) = x_i^2 + y_i^2 - 1/n^2, \\ R &= \begin{bmatrix} 0.1765714 & 0.8458754 & -0.3371163 & -1.0598462 & 0.0269367 & 0.6447252 \\ 0.2893688 & 0.1328983 & -1.4142041 & 0.4346374 & 0.1677938 & -0.2855976 \\ -0.4505154 & -0.6521358 & -0.3240160 & 0.2748310 & -0.0022626 & 1.2614402 \\ 1.0819066 & 0.4199281 & 0.3317461 & 0.7231132 & -0.3725210 & 0.5304889 \\ 0.8377745 & -1.0150421 & -0.1600336 & -0.6991182 & -0.3744590 & -0.1150085 \\ 0.4579696 & -0.1868200 & 0.2138378 & -0.0250102 & 1.4464173 & 0.1299494 \end{bmatrix}. \end{aligned}$$

Appendix A. Traditional SOS certificates. This appendix reviews two previously known methods to certify nonnegativity on a variety; we illustrate them with concrete, simple examples. No new results are presented.

A.1. Equations SOS. Let \mathcal{V} be a variety with defining equations $h = (h_1, \dots, h_m)$, and let $p \in \mathbb{R}[x]$ be nonnegative on $\mathcal{V} \cap \mathbb{R}^n$. The standard approach to certify this nonnegativity is to compute an SOS polynomial F and a tuple of polynomials $g = (g_1, \dots, g_m)$ that satisfy (5). Let $u(x) \in \mathbb{R}[x]^N$ consist of all monomials up to degree d . Computing an equation's d -SOS certificate (F, g) reduces to the following problem:

$$(11) \quad \begin{aligned} &\text{Find} \quad Q \in \mathcal{S}^N, \quad C \in \mathbb{R}^{m \times N}, \quad Q \succeq 0 \\ &\text{subject to} \quad p(x) = Q \bullet u(x)u(x)^T + h(x)^T C u(x), \end{aligned}$$

where $F(x) = Q \bullet u(x)u(x)^T$ and $g(x) = C u(x)$.

Example A.1. Let us revisit the case from Example 3.8. We want to certify that $p(X) = 4X_{21} - 2X_{11}X_{22} - 2X_{12}X_{21} + 3$ is nonnegative on the variety in (9). We fix a degree bound of $d = 1$ and let $u(X) = (1, X_{11}, X_{12}, X_{21}, X_{22})$. The dimensions of the matrices in the SDP (11) are $Q \in \mathcal{S}^5$, $C \in \mathbb{R}^{4 \times 5}$. Note that these dimensions are larger than in Example 3.8. Solving the SDP leads to the following SOS certificate:

$$\begin{aligned} p(X) &= F(X) - g_1 h_1 - g_2 h_2 + g_3 h_3 - g_4 h_4, \\ F(X) &= (X_{21} - X_{12} + 1)^2 + (X_{11} - X_{22})^2, \\ g_1 &= X_{12} + 1, \quad g_2 = X_{21} + 1, \quad g_3 = X_{11} + X_{22}, \quad g_4 = X_{21} + X_{12}, \\ h_1 &= X_{11}^2 + X_{21}^2 - 1, \quad h_2 = X_{12}^2 + X_{22}^2 - 1, \quad h_3 = X_{11}X_{12} + X_{21}X_{22}, \quad h_4 = \det(X) - 1. \end{aligned}$$

A.2. Quotient ring SOS. It is possible to take advantage of the quotient ring structure to obtain a simpler SDP. The standard approach requires a Gröbner basis of the ideal $I := \langle h \rangle$. We briefly explain the procedure. We refer the reader to [9] for an introduction to Gröbner bases.

Consider the quotient ring $\mathcal{R}' = \mathbb{R}[x]/I$, and let $\phi: \mathbb{R}[x] \rightarrow \mathcal{R}'$ be the canonical homomorphism. Since $\phi(g_j h_j) = 0$, when we view the SDP in (11) in the quotient

ring, we obtain the following:

$$(12) \quad \begin{array}{ll} \text{Find} & Q \in \mathbb{R}^{SN}, \quad Q \succeq 0 \\ \text{subject to} & \phi(p(x)) = Q \bullet \phi(u(x)u(x)^T). \end{array}$$

Note that the matrix C was eliminated. The above SDP requires methods to represent and compute the quotient ring \mathcal{R}' . Given a Gröbner basis gb of I , there is a simple way to achieve this. Concretely, any polynomial $f \in \mathbb{R}[x]$ can be written in *normal form*, denoted as $\phi_{gb}(f)$, with respect to gb . This normal form map ϕ_{gb} is effectively representing the quotient ring. In addition, there is a natural monomial basis $u(x)$ to use, given by the standard monomials with respect to gb .

Example A.2. Consider again the case from Example 3.8. Consider the Gröbner basis

$$gb = \{\underline{X_{11}} - X_{22}, \underline{X_{12}} + X_{21}, \underline{X_{21}^2} + X_{22}^2 - 1\},$$

where the leading monomials are underlined. Let $u(X) = (1, X_{21}, X_{22})$ be the standard monomials of degree at most $d = 1$. Computing the normal form $\phi_{gb}(u(X)u(X)^T)$, we obtain the following SDP:

$$\begin{array}{ll} \text{Find} & Q \in \mathcal{S}^3, \quad Q \succeq 0 \\ \text{subject to} & 4X_{21} - 4X_{22}^2 + 5 = Q \bullet \begin{bmatrix} 1 & X_{21} & X_{22} \\ X_{21} & 1 - X_{22}^2 & X_{21}X_{22} \\ X_{22} & X_{21}X_{22} & X_{22}^2 \end{bmatrix}. \end{array}$$

Solving the SDP leads to the SOS certificate $\phi_{gb}(p(X)) = \phi_{gb}((2X_{21} + 1)^2)$. Note that the certificate obtained, as well as the dimension of Q , agrees with that of Example 3.8.

Remark A.3 (sampling SOS). Our sampling approach can be seen as a quotient ring formulation. The difference is that we use the radical ideal $J = \sqrt{I}$ and the coordinate ring $\mathcal{R} = \mathbb{R}[x]/J$. Provided a sample set $Z = \{z_1, \dots, z_S\}$ our description of \mathcal{R} is given by the *evaluation map* $\phi_Z(f) := (f(z_1), \dots, f(z_S))$.

REFERENCES

- [1] D. J. BATES, J. D. HAUENSTEIN, A. J. SOMMESE, AND C. W. WAMPLER, *Bertini: Software for Numerical Algebraic Geometry*, www.nd.edu/~sommese/bertini, 2006.
- [2] D. J. BATES, J. D. HAUENSTEIN, A. J. SOMMESE, AND C. W. WAMPLER, *Numerically Solving Polynomial Systems with Bertini*, Software Environments Tools 25, SIAM, Philadelphia, 2013.
- [3] S. J. BENSON AND Y. YE, *DSDP5: Software for Semidefinite Programming*, Tech. report ANL/MCS-P1289-0905, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, 2005.
- [4] G. BLEKHERMAN, P. A. PARRILO, AND R. R. THOMAS, EDS., *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Ser. Optim. 13, SIAM, Philadelphia, 2013, <https://doi.org/10.1137/1.9781611972290>.
- [5] G. BLEKHERMAN, G. SMITH, AND M. VELASCO, *Sums of squares and varieties of minimal degree*, J. Amer. Math. Soc., 29 (2016), pp. 893–913.
- [6] M. CHARDIN, *Une majoration de la fonction de Hilbert et ses conséquences pour l'interpolation algébrique*, Bull. Soc. Math. France, 117 (1989), pp. 305–318.
- [7] M. T. CHU AND N. T. TRENDACILOV, *The orthogonally constrained regression revisited*, J. Comput. Graph. Statist., 10 (2001), pp. 746–771.
- [8] C. CILIBERTO, *Geometric aspects of polynomial interpolation in more variables and of Waring's problem*, in European Congress of Mathematics, Progr. Math. 201, Springer, New York, 2001, pp. 289–316.

- [9] D. A. COX, J. LITTLE, AND D. O'SHEA, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, New York, 2007.
- [10] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *On the best rank-1 and rank- (R_1, R_2, \dots, R_n) approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342, <https://doi.org/10.1137/S0895479898346995>.
- [11] J.-C. FAUGÈRE, *Finding all the solutions of Cyclic 9 using Gröbner basis techniques*, in Computer Mathematics—Proceedings of the Fifth Asian Symposium (ASCM 2001), vol. 9, World Scientific, Singapore, 2001, pp. 1–12.
- [12] D. R. GRAYSON AND M. E. STILLMAN, *Macaulay2, a Software System for Research in Algebraic Geometry*, <http://www.math.uiuc.edu/Macaulay2/>, 2002.
- [13] J. C. JANTZEN, *Nilpotent orbits in representation theory*, in Lie Theory, Springer, New York, 2004, pp. 1–211.
- [14] V. KABANETS AND R. IMPAGLIAZZO, *Derandomizing polynomial identity tests means proving circuit lower bounds*, Comput. Complex., 13 (2004), pp. 1–46.
- [15] J. M. LANDSBERG AND J. WEYMAN, *On the ideals and singularities of secant varieties of Segre varieties*, Bull. London Math. Soc., 39 (2007), pp. 685–697.
- [16] J. B. LASSERRE, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817, <https://doi.org/10.1137/S1052623400366802>.
- [17] J. B. LASSERRE, *Moments, Positive Polynomials and Their Applications*, Ser. Optim. Appl. 1, World Scientific, Singapore, 2009.
- [18] M. LAURENT, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging Applications of Algebraic Geometry, Springer, New York, 2009, pp. 157–270, <https://homepages.cwi.nl/~monique/files/moment-ima-update-new.pdf>.
- [19] Z. LIU AND L. VANDENBERGHE, *Low-rank structure in semidefinite programs derived from the KYP lemma*, in Proc. 46th IEEE Conference on Decision and Control, 2007, pp. 5652–5659.
- [20] J. LÖFBERG AND P. A. PARRILO, *From coefficients to samples: A new approach to SOS optimization*, in Proc. 43rd IEEE Conference on Decision and Control, vol. 3, 2004, pp. 3154–3159.
- [21] J. NIE, J. DEMMEL, AND B. STURMFELS, *Minimizing polynomials via sum of squares over the gradient ideal*, Math. Programming, 106 (2006), pp. 587–606.
- [22] P. A. PARRILO, *Exploiting structure in sum of squares programs*, in Proc. 42nd IEEE Conference on Decision and Control, vol. 5, 2003, pp. 4664–4669.
- [23] P. A. PARRILO, *Semidefinite programming relaxations for semialgebraic problems*, Math. Programming, 96 (2003), pp. 293–320.
- [24] F. PERMENTER AND P. A. PARRILO, *Selecting a monomial basis for sums of squares programming over a quotient ring*, in Proc. 51st IEEE Conference on Decision and Control, 2012, pp. 1871–1876.
- [25] T. ROH, B. DUMITRESCU, AND L. VANDENBERGHE, *Multidimensional FIR filter design via trigonometric sum-of-squares optimization*, IEEE J. Sel. Topics Signal Process., 1 (2007), pp. 641–650.
- [26] T. ROH AND L. VANDENBERGHE, *Discrete transforms, semidefinite programming, and sum-of-squares representations of nonnegative polynomials*, SIAM J. Optim., 16 (2006), pp. 939–964, <https://doi.org/10.1137/040612646>.
- [27] T. SAUER, *Polynomial interpolation in several variables: Lattices, differences, and ideals*, Stud. Comput. Math., 12 (2006), pp. 191–230.
- [28] N. SAXENA, *Progress on polynomial identity testing*, Bull. EATCS, 99 (2009), pp. 49–79.
- [29] C. SCHEIDERER, *Positivity and sums of squares: A guide to recent results*, in Emerging Applications of Algebraic Geometry, Springer, New York, 2009, pp. 271–324.
- [30] K. SCHMÜDGEN, *The k -moment problem for compact semi-algebraic sets*, Math. Ann., 289 (1991), pp. 203–206.
- [31] M. SCHWEIGHOFER, *Global optimization of polynomials using gradient tentacles and sums of squares*, SIAM J. Optim., 17 (2006), pp. 920–942, <https://doi.org/10.1137/050647098>.
- [32] A. J. SOMMESE AND C. W. WAMPLER II, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, Singapore, 2005.
- [33] T. THEOBALD AND T. DE WOLFF, *Approximating amoebas and coamoebas by sums of squares*, Math. Comp., 84 (2015), pp. 455–473.
- [34] R. H. TÜTÜNCÜ, K. C. TOH, AND M. J. TODD, *Solving semidefinite-quadratic-linear programs using SDPT3*, Math. Programming, 95 (2003), pp. 189–217.
- [35] J. VERSCHelde, *Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation*, ACM Trans. Math. Software (TOMS), 25 (1999), pp. 251–276.
- [36] T. VIKLANDS, *Algorithms for the Weighted Orthogonal Procrustes Problem and Other Least Squares Problems*, Ph.D. thesis, Umea University, Umea, Sweden, 2006.

- [37] H. WANG, S. YAN, D. XU, X. TANG, AND T. HUANG, *Trace ratio vs. ratio trace for dimensionality reduction*, in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–8.
- [38] L. H. ZHANG AND R. C. LI, *Maximization of the sum of the trace ratio on the Stiefel manifold*, I: *Theory*, Sci. China Math., 57 (2014), pp. 2495–2508.
- [39] L. H. ZHANG AND R. C. LI, *Maximization of the sum of the trace ratio on the Stiefel manifold*, II: *Computation*, Sci. China Math., 57 (2014), pp. 1–18.