Approximation of the joint spectral radius of a set of matrices using sum of squares

Pablo A. Parrilo[1] and Ali Jadbabaie[2]

[1] Laboratory for Information and Decision Systems
Massachusetts Institute of Technology, Cambridge, MA 02139, parrilo@mit.edu
[2] Dept. of Electrical and Systems Engineering and GRASP laboratory
University of Pennsylvania, Philadelphia, PA 19104, jadbabai@seas.upenn.edu

Abstract. We provide an asymptotically tight, computationally efficient approximation of the joint spectral radius of a set of matrices using sum of squares (SOS) programming. The approach is based on a search for a SOS polynomial that proves simultaneous contractibility of a finite set of matrices. We provide a bound on the quality of the approximation that unifies several earlier results and is independent of the number of matrices. Additionally, we present a comparison between our approximation scheme and a recent technique due to Blondel and Nesterov, based on lifting of matrices. Theoretical results and numerical investigations show that our approach yields tighter approximations.

1 Introduction

Stability of discrete linear inclusions has been a topic of major research over the past two decades. Such systems can be represented as a switched linear system of the form $x(k+1) = A_{\sigma(k)}x(k)$, where σ is a mapping from the integers to a given set of indices. The above model has been studied extensively across multiple disciplines, ranging from control theory, theory of non-negative matrices and Markov chains, wavelets, dynamical systems, etc. The fundamental question of interest is to determine whether $x(k)$ converges to a limit, or equivalently, whether the infinite matrix products chosen from the set of matrices converge [1–3]. The research on convergence of infinite products of matrices spans across four decades. A majority of results in this area has been provided in the special case of non-negative and/or stochastic matrices. A non-exhaustive list of related research providing several necessary and sufficient conditions for convergence of infinite products includes [4, 3, 5]. Despite the wealth of research in this area, finding algorithms that can determine the convergence remains elusive. Much of the difficulty of this problem stems from the hardness in computation or efficient approximation of the joint spectral radius of a finite set of matrices [6]. This is defined as

$$\rho(A_1, \dots, A_m) := \limsup_{k\to+\infty} \max_{\sigma\in\{1,\dots,m\}^k} ||A_{\sigma_k} \cdots A_{\sigma_2} A_{\sigma_1}||^{1/k}, \tag{1}$$

and represents the maximum growth (or decay) rate that can be obtained by taking arbitrary products of the matrices A_i, and its value is independent of the norm chosen. Daubechies and Lagarias [3] conjectured that the joint spectral radius is equal to a related quantity, the *generalized spectral radius*, which is defined in a similar way except for the fact that the norm of the product is replaced by the spectral radius. Later Berger and Wang [1] proved this conjecture to be true for finite set of matrices. It is well known that computing ρ is hard from a computational viewpoint, and even approximating it is difficult [7–9]. For rational matrices, the joint spectral radius is not a semialgebraic function of the data, thus ruling out a very large class of methods for its exact computation. When the matrices are non-negative and stochastic, the problem is decidable but PSPACE complete [10], a complexity class widely believed to be worse than NP.

It turns out that a necessary and sufficient condition for stability of a linear difference inclusion is for the corresponding matrices to have a subunit join spectral radius [4], i.e., $\rho(A_1, \ldots, A_m) < 1$. This, however, is a condition that is impossible to verify in general. A subunit joint spectral radius is on the other hand equivalent to existence of a common norm with respect of which all matrices in the set are contractive [11–13]. Unfortunately, the proof of this result is not constructive, as knowledge of the joint spectral radius is needed to construct such an *extremal norm* [13]. In fact a similar result, due to Dayawansa and Martin [14], holds for nonlinear systems that undergo switching. A common approach in trying to approximate the joint spectral radius or to prove that it is indeed subunit, has been to try to prove simultaneous contractibility (i.e., existence of a common norm with respect to which matrices are contractive), by searching for a common ellipsoidal norm, or equivalently, searching for a common quadratic Lyapunov function. The benefit of this approach is due to the fact that the search for a common ellipsoidal norm can be posed as a semidefinite program and solved efficiently using interior point techniques. However, it is not too difficult to generate examples where the discrete inclusion is *absolutely asymptotically stable*, i.e. asymptotically stable for all switching sequences, but a common quadratic Lyapunov function, (or equivalently a common ellipsoidal norm) does not exist.

Ando and Shih describe in [15] a constructive procedure for generating a set of m matrices for which the joint spectral radius is $\frac{1}{\sqrt{m}}$, but no quadratic Lyapunov function exists. They prove that the interval $[0, \frac{1}{\sqrt{m}})$ is effectively the "optimal" range for the joint spectral radius necessary to guarantee simultaneous contractibility under an ellipsoidal norm for a finite collection of m matrices. The range is denoted as optimal since it is the largest subset of $[0, 1)$ for which if the joint spectral radius is in this subset the collection of matrices is simultaneously contractible. Furthermore, they show that the optimal joint spectral radius range for a *bounded* set of $n \times n$ matrices is the interval $[0, \frac{1}{\sqrt{n}})$. The proof of this fact is based on John's ellipsoid theorem [15]. Roughly speaking, John's ellipsoid theorem implies that every convex body in the n dimensional Euclidean space that is symmetric around the origin can be approximated by an ellipsoid (from the inside and outside) up to a factor of $\frac{1}{\sqrt{n}}$. A major consequence of this result is that finding a common Lyapunov function becomes increasingly hard as the dimension goes up.

Recently, Blondel, Nesterov and Theys [16] showed a similar result (also based on John's ellipsoid theorem), that the best ellipsoidal norm approximation of the joint spectral radius provides a lower bound and an upper bound on the actual value. Given a set of matrices $\mathcal{M}$ with joint spectral radius ρ, and best ellipsoidal norm approximation $\hat{\rho}$, it is shown there that

$$\frac{1}{\sqrt{m}}\hat{\rho}(\mathcal{M}) \leq \rho(\mathcal{M}) \leq \hat{\rho}(\mathcal{M}) \tag{2}$$

Furthermore, in [17], Blondel and Nesterov proposed a scheme to approximate the joint spectral radius, by "lifting" the matrices using Kronecker products to provide better approximations. A common feature of these approaches is the appearance of convexity-based methods to provide certificates of the desired system properties.

In this paper, we develop a sum of squares (SOS) based scheme for the approximation of the joint spectral radius, and prove several results on the resulting quality of approximation. For this, we use two different techniques, one inspired by recent results of Barvinok [18] on approximation of norms by polynomials, and the other one based on a convergent iteration similar to that used for

Lyapunov inequalities. Our results provide a simple and unified derivation of most of the available bounds, including some new ones. As a consequence, we can use SOS polynomials to approximate the extremal norm that is equal to the joint spectral radius. We also show that this approximation is always tighter than the one provided by Blondel and Nesterov.

A description of the paper follows. In Section 2 we present a class of bounds on the joint spectral radius based on simultaneous contractivity with respect to a norm, followed by a sum of squares-based relaxation, and the corresponding suboptimality properties. In Section 3 we present some background material in multilinear algebra, necessary for our developments, and a derivation of a bound of the quality of the SOS relaxation. An alternative development is presented in Section 4, where a different bound on the performance of the SOS relaxation is given in terms of a very natural Lyapunov iteration, similar to the classical case. In Section 5 we make a comparison with earlier techniques and analyze a numerical example. Finally, in Section 6 we present our conclusions.

2 SOS norms

A natural way of bounding the joint spectral radius is to find a common norm under for which we can guarantee certain contractiveness properties for all the matrices. In this section, we first revisit this characterization, and introduce our method of using SOS relaxations to approximate this common norm.

2.1 Norms and the joint spectral radius

As we mentioned, there exists an intimate relationship between the spectral radius and the existence of a vector norm under which all the matrices are simultaneously contractive. This is summarized in the following theorem, a special case of Proposition 1 in [6] by Rota and Strang.

Theorem 1 ([6]). *Consider a finite set of matrices $\mathcal{A} = \{A_1, \dots, A_m\}$. For any $\epsilon > 0$, there exists a norm $\|\cdot\|$ in $\mathbb{R}^n$ (denoted as JSR norm hereafter) such that*

$$||A_i x|| \leq (\rho(\mathcal{A}) + \epsilon)||x||, \quad \forall x \in \mathbb{R}^n, \quad i = 1, \dots, m.$$

The theorem appears in this form, for instance, in Proposition 4 of [16]. The main idea in our approach is to replace the JSR norm that approximates the joint spectral radius with a homogeneous SOS polynomial $p(x)$ of degree $2d$. As we will see in the next sections, we can produce arbitrarily tight SOS approximations, while still being able to prove a bound on the resulting estimate.

2.2 Joint spectral radius and polynomials

As the results presented above indicate, the joint spectral radius can be characterized by finding a common norm under which all the maps are simultaneously contractive. As opposed to the unit ball of a norm, the level sets of a homogeneous polynomial are not necessarily convex (see for instance Figure 1). Nevertheless, as the following theorem suggests, we can still obtain upper bounds on the joint spectral radius by replacing norms with homogeneous polynomials.

Theorem 2. *Let $p(x)$ be a strictly positive homogeneous polynomial of degree $2d$ that satisfies*

$$p(A_i x) \leq \gamma^{2d} p(x), \qquad \forall x \in \mathbb{R}^n \quad i = 1, \dots, m.$$

Then, $\rho(A_1, \dots, A_m) \leq \gamma$.

Proof. If $p(x)$ is strictly positive, then by compactness of the unit ball in $\mathbb{R}^n$ and continuity of $p(x)$, there exist constants $0 < \alpha \leq \beta$, such that

$$\alpha\,||x||^{2d} \leq p(x) \leq \beta\,||x||^{2d} \qquad \forall x \in \mathbb{R}^n.$$

Then,

$$\begin{aligned}
||A_{\sigma_k} \dots A_{\sigma_1}|| &\leq \max_x \frac{||A_{\sigma_k} \dots A_{\sigma_1} x||}{||x||} \\
&\leq \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2d}} \max_x \frac{p(A_{\sigma_k} \dots A_{\sigma_1} x)^{\frac{1}{2d}}}{p(x)^{\frac{1}{2d}}} \\
&\leq \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2d}} \gamma^k.
\end{aligned}$$

From the definition of the joint spectral radius in equation (1), by taking kth roots and the limit $k \to \infty$ we immediately have the upper bound $\rho(A_1, \dots, A_m) \leq \gamma$.

2.3 Sums of Squares Programming

The condition in Theorem 2 involves positive polynomials, which are computationally hard to characterize. A useful scheme, introduced in [19, 20] and relatively well-known by now, relaxes the nonnegativity constraints to a much more tractable *sum of squares* condition, where $p(x)$ is required to have a decomposition as $p(x) = \sum_i p_i(x)^2$. The SOS condition can be equivalently expressed in terms of a semidefinite programming constraint, hence its tractability. In what follows, we briefly describe the basic SOS construction.

Consider a given multivariate polynomial for which we want to decide whether a sum of squares decomposition exists. This question is equivalent to a semidefinite programming (SDP) problem, because of the following result:

Theorem 3. *A homogeneous multivariate polynomial $p(x)$ of degree $2d$ is a sum of squares if and only if*

$$p(x) = (x^{[d]})^T Q x^{[d]}, \tag{3}$$

where $x^{[d]}$ is a vector whose entries are (possibly scaled) monomials of degree d in the x_i variables, and Q is a symmetric positive semidefinite matrix.

Since in general the entries of $x^{[d]}$ are not algebraically independent, the matrix Q in the representation (3) *is not unique*. In fact, there is an affine subspace of matrices Q that satisfy the equality, as can be easily seen by expanding the right-hand side and equating term by term. To obtain an SOS representation, we need to find a positive semidefinite matrix in this affine subspace. Therefore, the problem of checking if a polynomial can be decomposed as a sum of squares is *equivalent* to verifying whether a certain affine matrix subspace intersects the cone of positive definite matrices, and hence an SDP feasibility problem.

Example 1. Consider the quartic homogeneous polynomial in two variables described below, and define the vector of monomials as $[x^2, y^2, xy]^T$.

$$\begin{aligned} p(x, y) &= 2x^4 + 2x^3y - x^2y^2 + 5y^4 \\ &= \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix} \\ &= q_{11}x^4 + q_{22}y^4 + (q_{33} + 2q_{12})x^2y^2 + 2q_{13}x^3y + 2q_{23}xy^3 \end{aligned}$$

For the left- and right-hand sides to be identical, the following linear equations should hold:

$$q_{11} = 2, \quad q_{22} = 5, \quad q_{33} + 2q_{12} = -1, \quad 2q_{13} = 2, \quad 2q_{23} = 0. \tag{4}$$

A positive semidefinite Q that satisfies the linear equalities can then be found using SDP. A particular solution is given by:

$$Q = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = L^T L, \quad L = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix},$$

and therefore we have the sum of squares decomposition:

$$p(x, y) = \frac{1}{2}(2x^2 - 3y^2 + xy)^2 + \frac{1}{2}(y^2 + 3xy)^2.$$

□

2.4 Norms and SOS polynomials

The procedure described in the previous subsection can be easily modified to the case where the polynomial $p(x)$ is not fixed, but instead we search for an SOS polynomial in a given affine family (for instance, all polynomials of a given degree).

This line of thought immediately suggests the following SOS relaxation of the conditions in Theorem 2:

$$\rho_{SOS,2d} := \inf_{p(x) \in \mathbb{R}_{2d}[x], \gamma} \gamma \qquad \text{s.t.} \begin{cases} p(x) \text{ is SOS} \\ \gamma^{2d} p(x) - p(A_i x) \text{ is SOS} \end{cases} \tag{5}$$

where $\mathbb{R}_{2d}[x]$ is the set of homogeneous polynomials of degree $2d$.

For any fixed γ, the constraints in this problem are all of SOS type, and thus equivalent to semidefinite programming. Therefore, the computation of $\rho_{SOS,2d}$ is a quasiconvex problem, and can be easily solved with a standard SDP solver via bisection methods. By Theorem 2, the solution of this relaxation gives the bound

$$\rho(A_1, \ldots, A_m) \leq \rho_{SOS,2d}, \tag{6}$$

where $2d$ is the degree of the approximating polynomial.

2.5 Quality of approximation

What can be said about the quality of the bounds produced by the SOS relaxation? We present next some results to answer this question; a more complete characterization is developed in Section 3.1. A very helpful result in this direction is the following theorem of Barvinok, that quantifies how tightly SOS polynomials can approximate norms:

Theorem 4 ([18], p. 221). *Let $||\cdot||$ be a norm in $\mathbb{R}^n$. For any integer $d \geq 1$ there exists a homogeneous polynomial $p(x)$ in n variables of degree $2d$ such that*

1. *The polynomial $p(x)$ is a sum of squares.*
2. *For all $x \in \mathbb{R}^n$,*
$$p(x)^{\frac{1}{2d}} \leq ||x|| \leq k(n,d)\, p(x)^{\frac{1}{2d}},$$
where $k(n,d) := \binom{n+d-1}{d}^{\frac{1}{2d}}$.

For fixed state dimension n, by increasing the degree d of the approximating polynomials, the factor in the upper bound can be made arbitrarily close to one. In fact, for large d, we have the approximation
$$k(n,d) \;\approx\; 1 + \frac{n-1}{2}\frac{\log d}{d}.$$

With these preliminaries, we can now present the main result of this paper that proves a bound on the quality of SOS approximations to the joint spectral radius:

To apply these results to our problem, consider the following. If $\rho(A_1,\ldots,A_m) < \gamma$, by Theorem 1 (and sharper results in [11–13]) there exists a norm $\|\cdot\|$ such that
$$||A_i x|| \leq \gamma ||x||, \quad \forall x \in \mathbb{R}^n, i = 1,\ldots,m.$$

By Theorem 4, we can therefore approximate this norm with a homogeneous SOS polynomial $p(x)$ of degree $2d$ that will then satisfy
$$p(A_i x)^{\frac{1}{2d}} \leq ||A_i x|| \leq \gamma ||x|| \leq \gamma\, k(n,d)\, p(x)^{\frac{1}{2d}},$$

and thus we know that there exists a feasible solution of
$$\begin{cases} p(x) \text{ is SOS} \\ \alpha^{2d} p(x) - p(A_i x) \geq 0 \qquad i = 1,\ldots,m, \end{cases}$$

for $\alpha = k(n,d)\rho(A_1,\ldots,A_m)$.

Despite these appealing results, notice that in general we cannot directly conclude from this that the proposed SOS relaxation will always obtain a solution that is within $k(n,d)^{-1}$ from the true spectral radius. The reason is that even though we can prove the existence of a $p(x)$ that is SOS and for which $\alpha^{2d}p(x) - p(A_i x)$ are nonnegative for all i, it is unclear whether the last m expressions are actually SOS. We will show later in the paper that this is indeed the case. Before doing this, we concentrate first on two important cases of interest, where the approach described guarantees a good quality of approximation.

Planar systems The first case corresponds to two-dimensional (planar) systems, i.e., when $n = 2$. In this case, it always holds that nonnegative homogeneous bivariate polynomials are SOS (e.g., [21]). Thus, we have the following result:

Theorem 5. *Let $\{A_1, \ldots, A_m\} \subset \mathbb{R}^{2\times 2}$. Then, the SOS relaxation (5) always produces a solution satisfying:*

$$\tfrac{1}{2}\rho_{SOS,2d} \leq (d+1)^{-\frac{1}{2d}}\,\rho_{SOS,2d} \leq \rho(A_1, \ldots, A_m) \leq \rho_{SOS,2d}.$$

This result is independent *of the number m of matrices.*

Quadratic Lyapunov functions For the quadratic case (i.e., $2d = 2$), it is also true that nonnegative quadratic forms are sums of squares. Since

$$\binom{n+d-1}{d}^{\frac{1}{2d}} = \binom{n}{1}^{\frac{1}{2}} = \sqrt{n},$$

the inequality

$$\frac{1}{\sqrt{n}}\,\rho_{SOS,2} \leq \rho(A_1, \ldots, A_m) \leq \rho_{SOS,2} \tag{7}$$

follows. This bound exactly coincides with the results of Ando and Shih [15] or Blondel, Nesterov and Theys [17]. This is perhaps not surprising, since in this case both Ando and Shih's proof [15] and Barvinok's theorem rely on the use of John's ellipsoid to approximate the same underlying convex set.

Level sets and convexity Unlike the norms that appear in Theorem 1, an appealing feature of the SOS-based method is that we are not constrained to use polynomials with convex level sets. This enables in some cases much better bounds than what is promised by the theorems above, as illustrated in the following example.

Example 2. This is based on a construction by Ando and Shih [15]. Consider the problem of proving a bound on the joint spectral radius of the following matrices:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}.$$

For these matrices, it can be easily shown that $\rho(A_1, A_2) = 1$. Using a common quadratic Lyapunov function (i.e., the case $d = 2$), the upper bound on the joint spectral radius is equal to $\sqrt{2}$. However, even a simple quartic SOS Lyapunov function is enough to prove an upper bound of $1 + \epsilon$ for every $\epsilon > 0$, since the SOS polynomial

$$V(x) = (x_1^2 - x_2^2)^2 + \epsilon(x_1^2 + x_2^2)^2$$

satisfies

$$\begin{aligned}(1+\epsilon)V(x) - V(A_1 x) &= (x_2^2 - x_1^2 + \epsilon(x_1^2 + x_2^2))^2 \\ (1+\epsilon)V(x) - V(A_2 x) &= (x_1^2 - x_2^2 + \epsilon(x_1^2 + x_2^2))^2.\end{aligned}$$

The corresponding level sets of $V(x)$ are plotted in Figure 1, and are clearly non-convex.

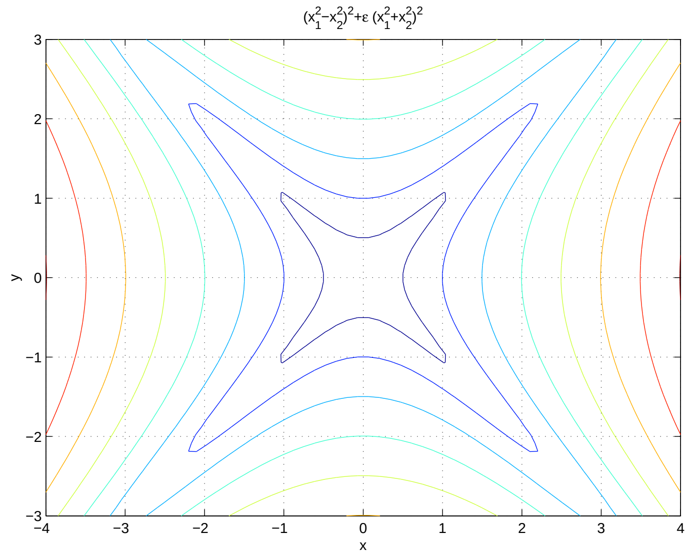

Fig. 1. Level sets of the quartic homogeneous polynomial $V(x_1, x_2)$. These define a Lyapunov function, under which both A_1 and A_2 are $(1+\epsilon)$-contractive. The value of ϵ is here equal to 0.01.

3 Symmetric algebra and induced matrices

In the upcoming sections, we present some further bounds on the quality of the SOS relaxation (5), either by a more refined analysis of the SOS polynomials in Barvinok's theorem or by explicitly producing an SOS Lyapunov function of guaranteed suboptimality properties. These constructions are quite natural, and parallel some lifting ideas as well as the classical iteration used in the solution of discrete-time Lyapunov inequalities.

Before proceeding further, we need to briefly revisit first some classical notions from multilinear algebra, namely the *symmetric algebra* of a vector space.

Consider a vector $x \in \mathbb{R}^n$, and an integer $d \geq 1$. We define its d-lift $x^{[d]}$ as a vector in $\mathbb{R}^N$, where $N := \binom{n+d-1}{d}$, with components $\{\sqrt{\binom{d}{\alpha}} x^\alpha\}_\alpha$, where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| := \sum_i \alpha_i = d$, and $\binom{d}{\alpha}$ denotes the multinomial coefficient $\binom{d}{\alpha_1, \alpha_2, \ldots, \alpha_n} = \frac{d!}{\alpha_1! \alpha_2! \ldots \alpha_n!}$. In other words, the components of the lifted vector are the monomials of degree d, scaled by the square root of the corresponding multinomial coefficients.

Example 3. Let $n = 2$, and $x = [u, v]^T$. Then, we have

$$\begin{bmatrix} u \\ v \end{bmatrix}^{[1]} = \begin{bmatrix} u \\ v \end{bmatrix}, \qquad \begin{bmatrix} u \\ v \end{bmatrix}^{[2]} = \begin{bmatrix} u^2 \\ \sqrt{2}uv \\ v^2 \end{bmatrix}, \qquad \begin{bmatrix} u \\ v \end{bmatrix}^{[3]} = \begin{bmatrix} u^3 \\ \sqrt{3}u^2v \\ \sqrt{3}uv^2 \\ v^3 \end{bmatrix}.$$

The main motivation for this specific scaling of the components, is to ensure that the lifting preserves (some of) the norm properties. In particular, if $||\cdot||$ denotes the standard Euclidean norm, it can be easily verified that $||x^{[d]}|| = ||x||^d$. Thus, the lifting operation provides a norm-preserving (up to power) embedding of $\mathbb{R}^n$ into $\mathbb{R}^N$ (in the projective case, this is the so-called *Veronese* embedding).

This concept can be directly extended from vectors to linear transformations. Consider a linear map in $\mathbb{R}^n$, and the associated $n \times n$ matrix A. Then, the lifting described above naturally induces an associated map in $\mathbb{R}^N$, that makes the corresponding diagram commute. The matrix representing this linear transformation is the *d-th induced matrix* of A, denoted by $A^{[d]}$, which is the unique $N \times N$ matrix that satisfies

$$A^{[d]} x^{[d]} = (Ax)^{[d]}$$

In systems and control, these classical constructions of multilinear algebra have been used under different names in several works, among them [22, 23] and (implicitly) [17]. Although not mentioned in the Control literature, there exists a simple explicit formula for the entries of these induced matrices; see [24, 25]. The d-th induced matrix $A^{[d]}$ has dimensions $N \times N$. Its entries are given by

$$A^{[d]}_{\alpha\beta} = \frac{\sqrt{\alpha!\beta!}}{d!} \operatorname{per} A_{\tilde{\alpha}\tilde{\beta}}, \tag{8}$$

where the indices α, β are multisets of $\{1, \ldots, n\}$ of cardinality d, and per indicates the *permanent*[3] of a square matrix. It can be shown that these operations define an algebra homomorphism, i.e., one that respects the structure of matrix multiplication. In particular, for any matrices A, B of compatible dimensions, the following identities hold:

$$(AB)^{[d]} = A^{[d]} B^{[d]}, \qquad (A^{-1})^{[d]} = (A^{[d]})^{-1}.$$

Furthermore, there is a simple and appealing relationship between the eigenvalues of $A^{[d]}$ and those of A. Concretely, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then the eigenvalues of $A^{[d]}$ are given by $\prod_{j \in S} \lambda_j$ where $S \subseteq \{1, \ldots, n\}, |S| = d$; there are exactly $\binom{n+d-1}{d}$ such multisets. A similar relationship holds for the corresponding eigenvectors. Essentially, as explained below in more detail, the induced matrices are the symmetry-reduced version of the d-fold Kronecker product.

As mentioned, the symmetric algebra and associated induced matrices are classical objects of multilinear algebra. Induced matrices, as defined above, as well as the more usual *compound matrices*, correspond to two specific isotypic components of the decomposition of the d-fold tensor product under the action of the symmetric group S^d (i.e., the *symmetric* and *skew-symmetric* algebras). Compound matrices are associated with the alternating character (hence their relationship with determinants), while induced matrices correspond instead to the trivial character, thus the connection with permanents. Similar constructions can be given for any other character of the symmetric group, by replacing the permanent in (8) with the suitable immanants; see [24] for additional details.

3.1 Bounds on the quality of $\rho_{SOS,2d}$

In this section we directly prove a bound on the approximation properties of the SOS approximation. As we will see, the techniques based on the lifting described will exactly yield the factor $k(n, d)^{-1}$ suggested by the application of Barvinok's theorem.

We first prove a preliminary result of the behavior of the joint spectral radius under lifting. The scaling properties described in the previous section can be applied to obtain the following:

---

[3] The permanent of a matrix $A \in \mathbb{R}^{n \times n}$ is defined as $\operatorname{per}(A) := \sum_{\sigma \in \Pi_n} \prod_{i=1}^{n} a_{i,\sigma(i)}$, where Π_n is the set of all permutations in n elements.

Lemma 1. *Given matrices* $\{A_1, \ldots, A_m\} \subset \mathbb{R}^{n\times n}$ *and an integer* $d \geq 1$*, the following identity holds:*

$$\rho(A_1^{[d]}, \ldots, A_m^{[d]}) = \rho(A_1, \ldots, A_m)^d.$$

The proof follows directly from the definition (1) and the two properties $(AB)^{[d]} = A^{[d]}B^{[d]}$, $||x^{[d]}|| = ||x||^d$, and it is thus omitted.

Combining all these inequalities, we obtain the main result of this paper:

Theorem 6. *The SOS relaxation (5) satisfies:*

$$\binom{n+d-1}{d}^{-\frac{1}{2d}} \rho_{SOS,2d} \leq \rho(A_1, \ldots, A_m) \leq \rho_{SOS,2d}. \tag{9}$$

Proof. From Lemma 1 and inequality (7), we have:

$$\binom{n+d-1}{d}^{-\frac{1}{2}} \rho_{SOS,2}(A_1^{[d]}, \ldots, A_m^{[d]}) \leq \rho(A_1, \ldots, A_m)^d \leq \rho_{SOS,2}(A_1^{[d]}, \ldots, A_m^{[d]}).$$

Combining this with the inequality (proven later in Theorem 9),

$$\rho_{SOS,2d}(A_1, \ldots, A_m)^d \leq \rho_{SOS,2}(A_1^{[d]}, \ldots, A_m^{[d]}),$$

the result follows.

4 Sum of squares Lyapunov iteration

We describe next an alternative approach to obtain bounds on the quality of the SOS approximation. As opposed to the results in the previous section, the bounds now explicitly depend on the number of matrices, but will usually be better in the case of small m.

Consider the iteration defined by

$$V_0(x) = 0, \qquad V_{k+1}(x) = Q(x) + \frac{1}{\beta}\sum_{i=1}^{m} V_k(A_i x), \tag{10}$$

where $Q(x)$ is a fixed n-variate homogeneous polynomial of degree $2d$ and $\beta > 0$. The iteration defines an affine map in the space of homogeneous polynomials of degree $2d$. As usual, the iteration will converge under certain assumptions on the spectral radius of this linear operator.

Theorem 7. *The iteration defined in (10) converges for arbitrary* $Q(x)$ *if* $\rho(A_1^{[2d]} + \cdots + A_m^{[2d]}) < \beta$*.*

Proof. Since the vector space of homogenous polynomials $\mathbb{R}_{2d}[x_1, \ldots, x_n]$ is isomorphic to the space of linear functionals on $(\mathbb{R}^n)^{[2d]}$, we can write $V_k(x) = \langle v_k, x^{[2d]}\rangle$, where $v_k \in \mathbb{R}^{\binom{n+2d-1}{2d}}$ is the vector of (scaled) coefficients of $V_k(x)$. Then, the iteration (10) can be simply expressed as:

$$v_{k+1} = q + \frac{1}{\beta}\sum_{i=1}^{m} A_i^{[2d]} v_k,$$

and it is well known that an affine iteration converges if the spectral radius of the linear term is less than one.

Theorem 8. *The following inequality holds:*

$$\rho_{SOS,2d} \leq \rho(A_1^{[2d]} + \cdots + A_m^{[2d]})^{\frac{1}{2d}}.$$

Proof. Choose a $Q(x)$ that is SOS, e.g., $Q(x) := (\sum_{i=1}^n x_i^2)^d$, and let $\beta = \rho(A_1^{[2d]} + \cdots + A_m^{[2d]}) + \epsilon$. The iteration (10) guarantees that V_{k+1} is SOS if V_k is. By induction, all the iterates V_k are SOS. By the choice of β and Theorem 7, the V_k converge to some homogeneous polynomial $V_\infty(x)$. By the closedness of the cone of SOS polynomials, the limit V_∞ is also SOS. Furthermore, we have

$$\beta V_\infty(x) - V_\infty(A_i x) = \beta Q(x) + \sum_{j \neq i} V_\infty(A_j x)$$

and therefore the expression on the right-hand side is SOS. This implies that $p(x) := V_\infty(x)$ is a feasible solution of the SOS relaxation (5). Taking $\epsilon \to 0$, the result follows.

Corollary 1. *For the SOS relaxation in (5), we always have*

$$\frac{1}{\sqrt[2d]{m}} \rho_{SOS,2d} \leq \rho(A_1, \ldots, A_m) \leq \rho_{SOS,2d}.$$

Proof. This follows directly from Theorem 8 and the inequalities (5.1) and (5.2) in [17].

The iteration (10) is the natural generalization of the Lyapunov recursion for the single matrix case, and of the construction by Ando and Shih in [15] for the quadratic case. Notice also that this corresponds exactly to the condition introduced in [17]. As a consequence, the SOS-based approach will *always* produce an upper bound at least as good as the one given by the Blondel-Nesterov procedure.

5 Comparison with earlier techniques

In this section we compare the $\rho_{SOS,2d}$ approach with some earlier bounds from the literature. We show that our bound is never weaker than those obtained by all the other procedures.

5.1 The Blondel-Nesterov technique

In [17], Blondel and Nesterov develop a method based on the calculation of the spectral radius of "lifted" matrices. They in fact present two different lifting procedures (the "Kronecker" and "semidefinite" lifting), and in Section 5 of their paper, they describe a family of bounds obtained by arbitrary combinations of these two liftings.

In contrast, the procedure presented in our paper relies on a single canonically defined lifting, that always dominates the Blondel-Nesterov contruction. Furthermore, it can be shown that our construction exactly corresponds to a fully symmetry-reduced version of their procedure, thus yielding the same (or better) bounds, but at a much smaller computational cost since the corresponding matrices are exponentially smaller.

5.2 Common quadratic Lyapunov functions

This method corresponds to finding a common quadratic Lyapunov function, either directly for the matrices A_i, or for the lifted matrices $A_i^{[d]}$. Specifically, let

$$\rho_{CQLF,2d} := \inf\{\gamma \mid \gamma^{2d} P - (A_i^{[d]})^T P A_i^{[d]} \succeq 0, \quad P \succ 0\}$$

This is essentially equivalent to what is discussed in Corollary 3 of [17], except that the matrices involved in our approach are exponentially smaller (of size $\binom{n+d-1}{d}$ rather than n^d), as all the symmetries have been taken out[4]. Notice also that, as a consequence of their definitions, we have

$$\rho_{CQLF,2d}(A_1, \ldots, A_m)^d = \rho_{SOS,2}(A_1^{[d]}, \ldots, A_m^{[d]}).$$

We can then collect most of these results in a single theorem:

Theorem 9. *The following inequalities between all the bounds hold:*

$$\rho(A_1, \ldots, A_m) \leq \rho_{SOS,2d} \leq \rho_{CQLF,2d} \leq \rho\left(\sum_{i=1}^{m} A_i^{[2d]}\right)^{\frac{1}{2d}}.$$

Proof. The left-most inequality is (6). The right-most inequality follows from a similar (but stronger) argument to the one given in Theorem 8 above, since the spectral radius condition $\rho(A_1^{[2d]} + \cdots + A_m^{[2d]}) < \beta$ actually implies the convergence of the matrix iteration in $\mathcal{S}^N$ given by

$$P_{k+1} = Q + \frac{1}{\beta}\sum_{i=1}^{m}(A_i^{[d]})^T P_k A_i^{[d]}, \qquad P_0 = I.$$

For the middle inequality, let $p(x) := (x^{[d]})^T P x^{[d]}$. Since $P \succ 0$, it follows that $p(x)$ is SOS. From $\gamma^{2d}P - (A_i^{[d]})^T P A_i^{[d]} \succeq 0$, left- and right-multiplying by $x^{[d]}$, we have that $\gamma^{2d}p(x) - p(A_i x)$ is also SOS, and thus $p(x)$ is a feasible solution of (5), from where the result directly follows.

Remark 1. We always have $\rho_{SOS,2} = \rho_{CQLF,2}$, since both correspond to the case of a common quadratic Lyapunov function for the matrices A_i.

5.3 Example

We present next a numerical example that compares the described techniques. In particular, we show that the bounds in Theorem 9 can all be strict.

Example 4. Consider the three 4×4 matrices

$$A_1 = \begin{bmatrix} 0 & 1 & 7 & 4 \\ 1 & 6 & -2 & -3 \\ -1 & -1 & -2 & -6 \\ 3 & 0 & 9 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -3 & 3 & 0 & -2 \\ -2 & 1 & 4 & 9 \\ 4 & -3 & 1 & 1 \\ 1 & -5 & -1 & -2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 4 & 5 & 10 \\ 0 & 5 & 1 & -4 \\ 0 & -1 & 4 & 6 \\ -1 & 5 & 0 & 1 \end{bmatrix}.$$

The value of the different approximations are presented in Table 1. A lower bound is $\rho(A_1A_3)^{\frac{1}{2}} \approx 8.9149$, which is extremely close (and perhaps exactly equal) to the upper bound $\rho_{SOS,4}$.

[4] There seems to be a typo in equation (7.4) of [17], as all the terms A_i^k should likely read $A_i^{\otimes k}$.

d	1	2	3
$\dim A_i^{[d]}$	4	10	20
$\dim A_i^{[2d]}$	10	35	84
$\rho_{SOS,2d}$	9.761	8.92	-
$\rho_{CQLF,2d}$	9.761	9.02	-
$\rho(\sum_i A_i^{[2d]})^{\frac{1}{2d}}$	12.519	9.887	9.3133

Table 1. Comparison of the different approximations for Example 4.

6 Conclusions

We provided an asymptotically tight, computationally efficient scheme for approximation of the joint spectral radius of a set of matrices using sum of squares programming. Utilizing the classical result of Rota and Strang on approximation of the joint spectral radius with a norm in conjunction with Barvinok's [18] result on approximation of any norm with a sum of squares polynomial, we provided an asymptotically tight estimate for the joint spectral radius which is independent of the number of matrices. Furthermore, we constructed an iterative scheme based on the generalization of a Lyapunov iteration which provides a bound on the joint spectral radius dependent on the number of matrices. Our results can be alternatively interpreted in a simpler way as providing a trajectory-preserving lifting to a polynomially-sized higher dimensional space, and proving contractiveness with respect to an ellipsoidal norm in that space. The results generalize earlier work of Ando and Shih [15], Blondel, Nesterov and Theys [16], and the lifting procedure of Blondel and Nesterov [17]. The good performance of our procedure was also verified using numerical examples.

References

1. Berger, M., Wang, Y.: Bounded semigroups of matrices. Linear Algebra and Applications **166** (1992) 21–27
2. Daubechies, I., Lagarias, J.C.: Sets of matrices all infinite products of which converge. Linear Algebra and Applications **161** (1992) 227–263
3. Daubechies, I., Lagarias, J.C.: Corrigendum/addendum to "Sets of matrices all infinite products of which converge". Linear Algebra and Applications **327** (2001) 69–83
4. Shih, M., Wu, J., Pang, C.T.: Asymptotic stability and generalized Gelfand spectral radius formula. Linear Algebra and Applications **251** (1997) 61–70
5. Leizarowitz, A.: On infinite products of stochastic matrices. Linear Algebra and Applications **168** (1992) 189–219
6. Rota, G.C., Strang, W.G.: A note on the joint spectral radius. Indag. Math. **22** (1960) 379–381
7. Blondel, V.D., Tsitsiklis, J.N.: A survey of computational complexity results in systems and control. Automatica **36** (2000) 1249–1274
8. Blondel, V.D., Tsitsiklis, J.N.: The boundedness of all products of a pair of matrices is undecidable. Systems and Control Letters **41** (2000) 135–140
9. Tsitsiklis, J.N., Blondel, V.: The Lyapunov exponent and joint spectral radius of pairs of matrices are hard- when not impossible- to compute and to approximate. Mathematics of Control, Signals, and Systems **10** (1997) 31–40
10. Hernek, D.: Random walks on colored graphs: Analysis and applications. PhD thesis, EECS Department, University of California at Berkeley (1995)

11. Barabanov, N.E.: Lyapunov indicators of discrete linear inclusions, parts I, II, and III. Translation from Avtomat. e. Telemekh. **2, 3 and 5** (1988) 40–46, 24–29, 17–44
12. Kozyakin, V.A.: Algebraic unsolvability of problem of absolute stability of desynchronized systems. Automation and Remote Control **51** (1990) 754–759
13. Wirth, F.: Joint spectral radius and extremal norms. Linear Algebra and Applications **251** (2002) 61–70
14. Dayawansa, W.P., Martin, C.F.: A converse Lyapunov theorem for a class of dynamical systems that undergo switching. IEEE Transactions on Automatic Control **44** (1999) 751–760
15. Ando, T., Shih, M.H.: Simultaneous contractibility. SIAM Journal on Matrix Analysis and Applications **19** (1998) 487–498
16. Blondel, V.D., Nesterov, Y., Theys, J.: On the accuracy of the ellipsoidal norm approximation of the joint spectral radius. Linear Algebra and Applications **394** (2005) 91–107
17. Blondel, V.D., Nesterov, Y.: Computationally efficient approximations of the joint spectral radius. SIAM J. Matrix Anal. Appl. **27**(1) (2005) 256–272
18. Barvinok, A.: A course in convexity. American Mathematical Society (2002)
19. Parrilo, P.A.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. PhD thesis, California Institute of Technology (2000) Available at `http://resolver.caltech.edu/CaltechETD:etd-05062004-055516`.
20. Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Prog. **96**(2, Ser. B) (2003) 293–320
21. Reznick, B.: Some concrete aspects of Hilbert's 17th problem. In: Contemporary Mathematics. Volume 253. American Mathematical Society (2000) 251–272
22. Brockett, R.: Lie algebras and Lie groups in control theory. In Mayne, D., Brockett, R., eds.: Geometric Methods in Systems Theory. D. Reidel Pub. Co. (1974) 17–56
23. Zelentsovsky, A.L.: Nonquadratic Lyapunov functions for robust stability analysis of linear uncertain systems. IEEE Trans. Automat. Control **39**(1) (1994) 135–138
24. Marcus, M.: Finite dimensional multilinear algebra. M. Dekker, New York (1973)
25. Marcus, M., Minc, H.: A survey of matrix theory and matrix inequalities. Dover Publications Inc., New York (1992) Reprint of the 1969 edition.