

Sum of squares approximations to energy functions

Hamza Adjerid
Department of Mathematics
Virginia Tech
Blacksburg, VA USA
ORCID 0009-0000-9025-0764

Jeff Borggaard
Department of Mathematics
Virginia Tech
Blacksburg, VA USA
ORCID 0000-0002-4023-7841

Abstract—Energy functions offer natural extensions of controllability and observability Gramians to nonlinear systems, enabling various applications such as computing reachable sets, optimizing actuator and sensor placement, performing balanced truncation, and designing feedback controllers. However, these extensions to nonlinear systems depend on solving Hamilton-Jacobi-Bellman (HJB) partial differential equations, which are infeasible for large-scale systems. Polynomial approximations are a viable alternative for modest-sized systems, but conventional polynomial approximations may yield negative values of the energy away from the origin. To address this issue, we explore polynomial approximations expressed as a sum of squares to ensure non-negative approximations. In this study, we focus on a reduced sum of squares polynomial where the coefficients are found through least-squares collocation—minimizing the HJB residual at sample points within a desired neighborhood of the origin. We validate the accuracy of these approximations through a case study with a closed-form solution and assess their effectiveness for controlling a ring of van der Pol oscillators with a Laplacian-like coupling term and discretized Burgers equation with source terms.

I. INTRODUCTION

Energy functions are pervasive in control theory. They were developed to extend controllability and observability Gramians to nonlinear systems for balancing [27] and have since been generalized to the H_∞ framework as the basis for nonlinear balanced truncation [25]. Since they can also serve as Lyapunov functions for closed-loop systems, other applications of energy functions include estimation of reachable sets and safe control algorithms [2], [30], sensor and actuator placement [11], and control design cf. [12]. The computational challenge to the use of energy functions are their reliance on solutions to the Hamilton-Jacobi-Bellman (HJB) equations. To avoid the curse-of-dimensionality, polynomial approximations have been introduced, e.g. [10], [14], [15]. However, these approximations often exhibit negative energy away from the origin.

Sum of squares (SOS) approximations in control theory also has a rich history that arose as a result of Parrilo [23] recasting many sum of squares problems as convex optimization problems and the development of efficient algorithms to solve them [7]. Their primary role in systems theory is to develop Lyapunov functions, which are guaranteed to be positive definite with negative definite derivatives along solutions and are useful for control synthesis [22], [24]. It is natural to consider SOS approximations for energy functions,

and this has been the subject of a number of studies, cf. [13] and have specifically been used for energy functions for a policy iteration control algorithm in recent work [21].

The remainder of this paper is organized as follows. In the next section, we provide a background on energy functions, a recent Kronecker product-based polynomial approximation that leads to efficient calculation of coefficients for modest systems, and a review of a typical sum of squares representation. Then we introduce three SOS formulations. The first embeds the Kronecker product polynomials within a higher degree polynomial to guarantee SOS. A second couples the Kronecker product polynomial with a nonlinear least-squares collocation procedure to determine the higher degree terms in a SOS. Finally, the third involves a modification of Parrilo's formulation and nonlinear least-squares collocation. To aid the optimization, we introduce a collocation strategy where we minimize residuals that are sampled over wider parameter sets. This allows us to warm-start the optimization to compute the SOS approximation in the region of interest. Numerical examples include a scalar test case where the analytic solution is available, a ring of van der Pol oscillators that are coupled through a Laplacian-like term, and an example generated by finite element discretization of Burgers equation. These show the better approximations to the energy functions away from the origin when using the modified Parrilo formulation and the warm-start strategy.

II. BACKGROUND

A. Energy functions

Consider the nonlinear, control affine input-output system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \quad (1)$$

where $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a smooth function, $\mathbf{B} \in \mathbb{R}^{n \times m}$, and $\mathbf{C} \in \mathbb{R}^{p \times n}$. The space of admissible controls are

$$\mathcal{U}^- = L_2(-\infty, 0; \mathbb{R}^m) \quad \text{and} \quad \mathcal{U}^+ = L_2(0, \infty; \mathbb{R}^m).$$

For this class of systems, energy functions are a natural extension of H_∞ control theory [28]. For any $\eta \leq 1$, we define

$$\mathcal{E}_\eta^-(\mathbf{x}_0) := \min_{\substack{\mathbf{u} \in \mathcal{U}^- \\ \mathbf{x}(-\infty)=\mathbf{0} \\ \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2} \int_{-\infty}^0 \eta \|\mathbf{y}(t)\|^2 + \|\mathbf{u}(t)\|^2 dt \quad (2)$$

as the *past energy*. The *future energy* is defined for three cases. For $0 < \eta \leq 1$,

$$\mathcal{E}_\eta^+(\mathbf{x}_0) := \min_{\substack{\mathbf{u} \in \mathcal{U}^+ \\ \mathbf{x}(0) = \mathbf{x}_0 \\ \mathbf{x}(\infty) = \mathbf{0}}} \frac{1}{2} \int_0^\infty \|\mathbf{y}(t)\|^2 + \frac{1}{\eta} \|\mathbf{u}(t)\|^2 dt; \quad (3)$$

for $\eta = 0$,

$$\mathcal{E}_\eta^+(\mathbf{x}_0) := \frac{1}{2} \int_0^\infty \|\mathbf{y}(t)\|^2 dt, \quad \mathbf{x}(0) = \mathbf{x}_0, \text{ and } \mathbf{u}(t) \equiv \mathbf{0}; \quad (4)$$

and for $\eta < 0$,

$$\mathcal{E}_\eta^+(\mathbf{x}_0) := \max_{\substack{\mathbf{u} \in \mathcal{U}^+ \\ \mathbf{x}(0) = \mathbf{x}_0 \\ \mathbf{x}(\infty) = \mathbf{0}}} \frac{1}{2} \int_0^\infty \|\mathbf{y}(t)\|^2 + \frac{1}{\eta} \|\mathbf{u}(t)\|^2 dt. \quad (5)$$

These are often defined using an L_2 -gain parameter γ_0 when $\eta = -\gamma_0^{-2}$ [28] or using an H_∞ -gain parameter $\gamma > 0$ when $\eta = 1 - \gamma^{-2}$ [25]. In the H_∞ setting, when we take the limit $\gamma \rightarrow 1$, $\eta = 0$, we recover that $\mathcal{E}_0^-$ is the controllability energy and $\mathcal{E}_0^+$ is the observability energy. Additionally, for the limit $\gamma \rightarrow \infty$, $\eta = 1$, we recover the energy functions used in HJB balancing [26]. Finally, when $\eta = -1$, the forward and past energy functions agree. These connections are detailed in [14].

These energy functions can be characterized by HJB equations, cf. [25]. If $\bar{\mathcal{E}}$ is a solution to

$$0 = \frac{\partial \bar{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) + \frac{1}{2} \frac{\partial \bar{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} \mathbf{B} \mathbf{B}^\top \frac{\partial \bar{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} - \frac{\eta}{2} \mathbf{x}^\top \mathbf{C}^\top \mathbf{C} \mathbf{x} \quad (6)$$

with $\bar{\mathcal{E}}(\mathbf{0}) = 0$ and $\mathbf{0}$ is an asymptotically stable fixed point of

$$\dot{\mathbf{x}} = - \left(\mathbf{f}(\mathbf{x}) + \mathbf{B} \mathbf{B}^\top \frac{\partial \bar{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} \right), \quad (7)$$

then $\bar{\mathcal{E}}(\mathbf{x})$ is the past energy function $\mathcal{E}_\eta^-(\mathbf{x})$ from (2).

Likewise, if $\tilde{\mathcal{E}}$ is a solution to

$$0 = \frac{\partial \tilde{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) - \frac{\eta}{2} \frac{\partial \tilde{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} \mathbf{B} \mathbf{B}^\top \frac{\partial \tilde{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}} + \frac{1}{2} \mathbf{x}^\top \mathbf{C}^\top \mathbf{C} \mathbf{x} \quad (8)$$

with $\tilde{\mathcal{E}}(\mathbf{0}) = 0$ and $\mathbf{0}$ is an asymptotically stable fixed point of

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) - \eta \mathbf{B} \mathbf{B}^\top \frac{\partial \tilde{\mathcal{E}}(\mathbf{x})}{\partial \mathbf{x}}, \quad (9)$$

then this solution $\tilde{\mathcal{E}}(\mathbf{x})$ is the future energy function $\mathcal{E}_\eta^+(\mathbf{x})$.

B. Polynomial approximations

Mesh-based approximations to Hamilton-Jacobi equations are prohibitive for systems with modest dimension. Therefore approximations using global basis functions, typically polynomials, were soon introduced [1], [8], [18], [20]. Polynomial approximations to energy functions in the forms defined above (2)–(5) were developed for nonlinear balancing applications as well [10]. In [14], the authors developed a scalable

approach to approximate energy functions by polynomials when they are written using Kronecker products:

$$\mathcal{E}_\eta^-(\mathbf{x}) \approx \frac{1}{2} (\mathbf{v}_2^\top \mathbf{x}^{\otimes 2} + \mathbf{v}_3^\top \mathbf{x}^{\otimes 3} + \cdots + \mathbf{v}_d^\top \mathbf{x}^{\otimes d}) \quad (10)$$

$$\mathcal{E}_\eta^+(\mathbf{x}) \approx \frac{1}{2} (\mathbf{w}_2^\top \mathbf{x}^{\otimes 2} + \mathbf{w}_3^\top \mathbf{x}^{\otimes 3} + \cdots + \mathbf{w}_d^\top \mathbf{x}^{\otimes d}), \quad (11)$$

where $\mathbf{x}^{\otimes d}$ is a Kronecker product involving d copies of the vector $\mathbf{x}$, e.g. $\mathbf{x}^{\otimes 3} = \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}$. The coefficients $\mathbf{v}_k$ and $\mathbf{w}_k$ involve structured linear systems that enable efficient computation. For $\mathbf{M} \in \mathbb{R}^{q \times n}$ we define the k -way *Lyapunov matrix* or a special *Kronecker sum* [4] matrix as

$$\mathcal{L}_k(\mathbf{M}) := \underbrace{\mathbf{M} \otimes \cdots \otimes \mathbf{I}_n}_{k \text{ times}} + \cdots + \underbrace{\mathbf{I}_n \otimes \cdots \otimes \mathbf{M}}_{k \text{ times}} \in \mathbb{R}^{n^{k-1}q \times n^k},$$

where $\mathbf{I}_n$ is the n -dimensional identity matrix. We present the results from [14] for calculating the coefficients $\mathbf{v}_k$ and $\mathbf{w}_k$ when system (1) is specialized to $\mathbf{f}(\mathbf{x}) \equiv \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{\otimes 2}$, and the pairs $(\mathbf{A}, \mathbf{B})$ and $(\mathbf{A}, \mathbf{C})$ are stabilizable and detectable.

Theorem 1: ([14, Th. 7]) Consider system (1) with the conditions stated above. Given $\eta \leq 1$, and past energy function $\mathcal{E}_\eta^-(\mathbf{x})$ expanded with the coefficients $\mathbf{v}_i$, $i = 2, \dots, d$ in (10). Then, $\mathbf{v}_2 = \text{vec}(\mathbf{V}_2)$, where $\mathbf{V}_2$ is the symmetric positive definite solution to the $\mathcal{H}_\infty$ Riccati equation

$$0 = \mathbf{A}^\top \mathbf{V}_2 + \mathbf{V}_2 \mathbf{A} - \eta \mathbf{C}^\top \mathbf{C} + \mathbf{V}_2 \mathbf{B} \mathbf{B}^\top \mathbf{V}_2. \quad (12)$$

Moreover, the coefficient vectors $\mathbf{v}_k = \text{vec}(\mathbf{V}_k) \in \mathbb{R}^{n^k}$ for $2 \leq k \leq d$ solve the linear systems

$$\begin{aligned} \mathcal{L}_k((\mathbf{A} + \mathbf{B} \mathbf{B}^\top \mathbf{V}_2)^\top) \mathbf{v}_k = \\ -\mathcal{L}_{k-1}(\mathbf{F}^\top) \mathbf{v}_{k-1} - \frac{\eta}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \text{vec}(\mathbf{V}_i^\top \mathbf{B} \mathbf{B}^\top \mathbf{V}_j). \end{aligned}$$

Theorem 2: ([14, Th. 6]) Consider system (1) with the conditions stated above. Given $\eta \leq 1$ and future energy function $\mathcal{E}_\eta^+(\mathbf{x})$ expanded with the coefficients $\mathbf{w}_i$, $i = 2, \dots, d$ in (11). Then, $\mathbf{w}_2 = \text{vec}(\mathbf{W}_2)$, where $\mathbf{W}_2$ is the symmetric positive definite solution to the $\mathcal{H}_\infty$ Riccati equation

$$0 = \mathbf{A}^\top \mathbf{W}_2 + \mathbf{W}_2 \mathbf{A} + \mathbf{C}^\top \mathbf{C} - \eta \mathbf{W}_2 \mathbf{B} \mathbf{B}^\top \mathbf{W}_2. \quad (13)$$

Moreover, the coefficient vectors $\mathbf{w}_k = \text{vec}(\mathbf{W}_k) \in \mathbb{R}^{n^k}$ for $2 \leq k \leq d$ solve the linear systems

$$\begin{aligned} \mathcal{L}_k((\mathbf{A} - \eta \mathbf{B} \mathbf{B}^\top \mathbf{W}_2)^\top) \mathbf{w}_k = \\ -\mathcal{L}_{k-1}(\mathbf{F}^\top) \mathbf{w}_{k-1} + \frac{\eta}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \text{vec}(\mathbf{W}_i^\top \mathbf{B} \mathbf{B}^\top \mathbf{W}_j). \end{aligned}$$

The proofs for both theorems can be found in [14]. While solving the systems grows exponentially in k and as a polynomial in n , the authors developed an efficient implementation taking advantage of the Kronecker structure of the systems based on the work in [5], [6], [9].

As shown in [14], polynomial approximations are accurate near the origin. The main issue with these approximations is negativity away from the origin. However, energy functions must be positive definite by definition. One way to overcome this issue is to propose function approximations that impose non-negative definiteness.

C. Sum of squares

A scalar valued function $f_{\text{sos}}(\mathbf{x})$ is a sum of squares if it can be written as

$$f_{\text{sos}}(\mathbf{x}) = \sum_{i=1}^N f_i^2(\mathbf{x}),$$

where $\{f_i(\mathbf{x})\}$ can be any collection of generic functions and N any integer. Polynomial sum of squares are much easier to work with

$$f_{\text{sos}}(\mathbf{x}) = \sum_{i=1}^N p_i^2(\mathbf{x}),$$

where each $p_i(\mathbf{x})$ is a polynomial.

Proposition: ([23]) A polynomial $p(\mathbf{x})$ of degree $2d$ is a SOS if and only if there exists a positive semi-definite matrix $\mathbf{Q}$ and vector $\mathbf{z}(\mathbf{x})$ that contains monomials in $\mathbf{x}$ of degree $\leq d$ such that:

$$p(\mathbf{x}) = \mathbf{z}(\mathbf{x})^\top \mathbf{Q} \mathbf{z}(\mathbf{x}). \quad (14)$$

For example [23], by defining $\mathbf{z}(\mathbf{x}) = [x_1^2 \quad x_1 x_2 \quad x_2^2]^\top$, we can rewrite

$$p(\mathbf{x}) = 2x_1^4 + 2x_1^3 x_2 - x_1^2 x_2^2 + 5x_2^4$$

in the form (14) with the positive semi-definite matrix

$$\mathbf{Q} = \begin{pmatrix} 2 & 1 & -3 \\ 1 & 5 & 0 \\ -3 & 0 & 5 \end{pmatrix} = \mathbf{L} \mathbf{L}^\top \quad \text{with} \quad \mathbf{L} = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 0 \\ 1 & 3 \\ -3 & 1 \end{pmatrix}.$$

The Cholesky factor $\mathbf{L}$ above allows us to specifically write p as the SOS

$$p(\mathbf{x}) = \frac{1}{2}(2x_1^2 + x_1 x_2 - 3x_2^2)^2 + \frac{1}{2}(x_2^2 + 3x_1 x_2)^2.$$

III. FORMULATION OF SOS POLYNOMIALS

The use of SOS programming in HJB problems have been addressed in prior works, e.g. [16]. In our work, we introduce three different SOS formulations:

- **Completing the sum of squares:** Adding higher degree terms to a given polynomial approximation (10) or (11) to make them sums of squares.
- **Completing the sum of squares with collocation:** Lower degree terms match a given polynomial approximation, and collocation is used to determine highest degree terms.
- **Collocation method:** Based on the Parrilo formulation of SOS (14) within a least-squares collocation method.

The approaches are illustrated below using the past energy function $\mathcal{E}_\eta^-(\mathbf{x})$, but the same procedure can be applied to the future energy function $\mathcal{E}_\eta^+(\mathbf{x})$ as well.

A. Completing the sum of squares

Given a degree d approximation to $\mathcal{E}_\eta^-(\mathbf{x})$ as in (10), we propose a sum of squares approximation $\mathcal{E}_{\text{sos}}^-(\mathbf{x})$ as:

$$\mathcal{E}_{\text{sos}}^-(\mathbf{x}) = (\tilde{\mathbf{v}}_1^\top \mathbf{x} + \tilde{\mathbf{v}}_2^\top \mathbf{x}^{\odot 2} + \cdots + \tilde{\mathbf{v}}_{d-1}^\top \mathbf{x}^{\odot(d-1)})^2.$$

In general, the terms $\tilde{\mathbf{v}}_k$ are matrices and thus, this is a sum of squares. The $d-1$ coefficients of $\mathcal{E}_{\text{sos}}^-(\mathbf{x})$ are found by matching the lowest degree terms in $\mathcal{E}_\eta^-(\mathbf{x})$. This is done without the involvement of the HJB equation. The HJB information is implicitly embedded in the polynomial approximation of $\mathcal{E}_\eta^-(\mathbf{x})$.

Example: Given $n = 2$, and a degree $d = 3$ polynomial approximation $\mathcal{E}_\eta^-(\mathbf{x})$, the degree of the SOS approximation is $2(d-1) = 4$. Therefore:

$$\mathcal{E}_{\text{sos}}^-(\mathbf{x}) = (\tilde{\mathbf{v}}_1^\top \mathbf{x} + \tilde{\mathbf{v}}_2^\top \mathbf{x}^{\odot 2})^2,$$

where $\tilde{\mathbf{v}}_1^\top \in \mathbb{R}^{2 \times 2}$ and $\tilde{\mathbf{v}}_2^\top \in \mathbb{R}^{2 \times 4}$. By expanding $\mathcal{E}_{\text{sos}}^-(\mathbf{x})$, we get:

$$\mathcal{E}_{\text{sos}}^-(\mathbf{x}) = \mathbf{x}^\top \tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^\top \mathbf{x} + 2\mathbf{x}^\top \tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_2^\top \mathbf{x}^{\odot 2} + (\mathbf{x}^{\odot 2})^\top \tilde{\mathbf{v}}_2 \tilde{\mathbf{v}}_2^\top \mathbf{x}^{\odot 2}$$

or

$$\mathcal{E}_{\text{sos}}^-(\mathbf{x}) = \text{vec}(\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^\top)^\top \mathbf{x}^{\odot 2} + 2\text{vec}(\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_2^\top)^\top \mathbf{x}^{\odot 3} + \text{vec}(\tilde{\mathbf{v}}_2 \tilde{\mathbf{v}}_2^\top)^\top \mathbf{x}^{\odot 4}$$

Now, matching the $O(\mathbf{x}^{\odot 2})$ and $O(\mathbf{x}^{\odot 3})$ terms leads to the following systems of equations:

$$\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^\top = \frac{1}{2} \mathbf{V}_2 \quad (\text{solve by Cholesky})$$

$$2\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_2^\top = \frac{1}{2} \mathbf{V}_3 \quad (\text{solve by backsubstitution})$$

where, $\mathbf{V}_2 = \text{reshape}(\mathbf{v}_2^\top, 2, 2)$, $\mathbf{V}_3 = \text{reshape}(\mathbf{v}_3^\top, 2, 4)$ and $\mathbf{v}_2^\top$ and $\mathbf{v}_3^\top$ are, respectively, the degree 2 and 3 coefficients of the polynomial approximation of $\mathcal{E}_\eta^-(\mathbf{x})$.

B. Completing the sum of squares with collocation

In this approach, we write $\mathcal{E}_{\text{sos},c}^-(\mathbf{x})$ in the same polynomial squared form with an additional term:

$$\mathcal{E}_{\text{sos},c}^-(\mathbf{x}) \equiv (\tilde{\mathbf{v}}_1^\top \mathbf{x} + \tilde{\mathbf{v}}_2^\top \mathbf{x}^{\odot 2} + \cdots + \tilde{\mathbf{v}}_{d-1}^\top \mathbf{x}^{\odot(d-1)} + \tilde{\mathbf{v}}_d^\top \mathbf{x}^{\odot d})^2$$

and we calculate for the first $d-1$ coefficients as above (matching to a degree $2d-1$ polynomial approximation).

To calculate $\tilde{\mathbf{v}}_d$ we introduce the residual of the HJB equation found by substituting $\mathcal{E}_{\text{sos},c}^-(\mathbf{x})$ into (8) at the point $\mathbf{x}$ which we denote by $\mathbf{R}(\mathbf{x})$. The least-squares collocation problem is then described by the functional

$$J(\tilde{\mathbf{v}}_d^\top) = \sum_{k=1}^N \|\mathbf{R}(\mathbf{x}^{(k)})\|^2,$$

that is formed using N sample points $\{\mathbf{x}^{(k)}\}$ in the approximation domain $\Omega \subset \mathbb{R}^n$. By using randomly selected sample points within Ω , this procedure has the potential to scale to larger dimensions. The missing coefficient in $\mathcal{E}_{\text{sos},c}^-(\mathbf{x})$ is then given by ($*$ in this work denotes optimal not an adjoint)

$$\tilde{\mathbf{v}}_d^{\top*} = \underset{\tilde{\mathbf{v}}_d^\top \in \mathbb{R}^{n^d}}{\text{argmin}} J(\tilde{\mathbf{v}}_d^\top).$$

C. Nonlinear least-squares collocation

Using Parrilo's formulation in (14), we write $\mathcal{E}_p^-(\mathbf{x})$ as:

$$\mathcal{E}_p^-(\mathbf{x}) = \mathbf{z}(\mathbf{x})^\top \mathbf{Q} \mathbf{z}(\mathbf{x}) \quad (15)$$

where $\mathbf{z}(\mathbf{x}) \in \mathbb{R}^\nu$ is a vector of all monomials of degree $\leq d$ and $\mathbf{Q} \in \mathbb{R}^{\nu \times \nu}$ is a positive definite matrix. The number of monomial terms, ν , in $\mathbf{z}(\mathbf{x})$ is defined as follows:

$$\nu = \sum_{i=1}^d \deg_i(n),$$

where $\deg_i(n)$ is the number of monomial terms of degree i in dimension n defined recursively by:

$$\deg_i(n) = \sum_{j=1}^n \deg_{i-1}(j) \quad \text{for } i \geq 2$$

with $\deg_1(n) = n$. The matrix $\mathbf{Q}$ can be factorized as $\mathbf{Q} = \mathbf{L}\mathbf{L}^\top$ and $\mathcal{E}_p^-(\mathbf{x})$ can be rewritten as:

$$\mathcal{E}_p^-(\mathbf{x}) = p_{\text{sos}}(\mathbf{x}) = \mathbf{z}(\mathbf{x})^\top \mathbf{L}\mathbf{L}^\top \mathbf{z}(\mathbf{x}) \quad (16)$$

where the entries of $\mathbf{L}$ are the decision variables of the optimization problem defined by the objective function

$$J(\mathbf{L}) = \sum_{k=1}^N \|\mathbf{R}(\mathbf{x}^{(k)})\|^2, \quad (17)$$

where now $\mathbf{R}(\mathbf{x})$ is the residual of the HJB equation at $\mathbf{x}$ when substituting (16) into (8). The optimal coefficients in (16) are then given by

$$\mathbf{L}^* = \underset{\mathbf{L} \in \mathbb{R}^{\nu \times \nu}}{\text{argmin}} J(\mathbf{L}).$$

D. Sequence of polynomial approximations

In these last two sections, we performed an optimization problem to compute unknown coefficients over the domain Ω . In the nonlinear least-squares collocation case especially, the entries of $\mathbf{L}$ strongly depend on the region of approximation. For approximation regions $\Omega \subset \mathbb{R}^n$ that lie within the unit hypercube, the lowest degree terms in the residual $\mathbf{R}$ tend have a more significant contribution than the higher degree terms. On the other hand, for approximation regions Ω where most sample points lie outside the unit hypercube, the higher degree terms have more contribution the residual $J(\mathbf{L})$. This motivates a sequence of polynomial approximations that increase as the approximation domain Ω increases, and has the effect of warm-starting the optimization algorithm.

Example: Consider the case $d = 4$ where the system has a cubic nonlinearity $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^\odot + \mathbf{B}\mathbf{u}$. The energy function takes the following form

$$\mathcal{E}_p(\mathbf{x}) = [\mathbf{x}^\top \mathbf{x}^\odot]^\top \begin{pmatrix} \mathbf{L}_{11} & 0 \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{L}_{11} & 0 \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{pmatrix}^\top [\mathbf{x}^\top \mathbf{x}^\odot]^\top.$$

We can expand this to find

$$\begin{aligned} \mathcal{E}_p(\mathbf{x}) = & \mathbf{x}^\top \mathbf{L}_{11} \mathbf{L}_{11}^\top \mathbf{x} + 2\mathbf{x}^\top \mathbf{L}_{11} \mathbf{L}_{21}^\top \mathbf{x}^\odot \\ & + \mathbf{x}^\odot \mathbf{L}_{21} \mathbf{L}_{21}^\top \mathbf{x} + \mathbf{L}_{22} \mathbf{L}_{22}^\top \mathbf{x}^\odot. \end{aligned}$$

We let $\mathbf{Q}_1 = \mathbf{L}_{11} \mathbf{L}_{11}^\top$, $\mathbf{Q}_2 = \mathbf{L}_{11} \mathbf{L}_{21}^\top$ and $\mathbf{Q}_3 = \mathbf{L}_{21} \mathbf{L}_{21}^\top + \mathbf{L}_{22} \mathbf{L}_{22}^\top$. Note that $\mathbf{Q}_3$ is the only matrix that depends on $\mathbf{L}_{22}$. Substituting $\mathcal{E}_p(\mathbf{x})$ into (8), the residual $\mathbf{R}(\mathbf{x})$ is a degree 6 polynomial, $\mathbf{Q}_3$, and hence $\mathbf{L}_{22}$, only appears in the degree 4, 5, and 6 terms, $r_4(\mathbf{L})$, $r_5(\mathbf{L})$ and $r_6(\mathbf{L})$, respectively. These coefficient functions are

$$\begin{aligned} r_4(\mathbf{L}) = & \text{vec}(8\mathbf{Q}_1 \mathbf{B} \mathbf{B}^\top \tilde{\mathbf{Q}}_3^\top + 2\mathbf{Q}_1 \mathbf{F} + 4\mathbf{A}^\top \tilde{\mathbf{Q}}_3^\top) \\ & + \text{vec}(18\tilde{\mathbf{Q}}_2 \mathbf{B} \mathbf{B}^\top \tilde{\mathbf{Q}}_2^\top) \\ r_5(\mathbf{L}) = & \text{vec}(24\tilde{\mathbf{Q}}_2 \mathbf{B} \mathbf{B}^\top \tilde{\mathbf{Q}}_3^\top + 6\tilde{\mathbf{Q}}_2 \mathbf{F}) \\ r_6(\mathbf{L}) = & \text{vec}(8\tilde{\mathbf{Q}}_3 \mathbf{B} \mathbf{B}^\top \tilde{\mathbf{Q}}_3^\top + 4\tilde{\mathbf{Q}}_3 \mathbf{F}), \end{aligned}$$

where $\tilde{\mathbf{Q}}_2$ and $\tilde{\mathbf{Q}}_3$ are reshaped versions of $\mathbf{Q}_2$ and $\mathbf{Q}_3$ to match the dimensions. When we attempt to minimize the residual with many sample points far outside the unit hypercube ($\|\mathbf{x}\| \gg 1$), the degree 6 term dominates and therefore the objective function is very sensitive to $\mathbf{L}_{22}$. Small changes to $\tilde{\mathbf{Q}}_3$ have such a large influence that the $r_6(\mathbf{L})$ term dominates the residual and this tends to drive $\tilde{\mathbf{Q}}_3$ to zero, but this also impacts $\mathbf{L}_{21}$ using the definition of $\mathbf{Q}_3$ above. Therefore, we elect to drop the $\mathbf{L}_{22}$ term in the case Ω is far outside the unit hypercube. In the case of more blocks in $\mathbf{L}$, it is desirable to drop the term representing the highest degree coefficients. For example, for degree 6 SOS approximations, we only keep the shaded portion of the matrix $\mathbf{L}$ illustrated below when Ω has more sample points outside the unit hypercube:

$$\begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix}$$

E. Windowing and issues for large regions Ω in $\mathbb{R}^n$

Performing the optimization of $J(\mathbf{L})$ for small regions around the origin is typically a convex optimization problem. To illustrate this, we take the scalar system from [14] and we write the degree 4 SOS approximation as in (16) for the past energy function $\mathcal{E}_p^-(\mathbf{x})$ with an approximation region $\Omega = [-1; 1] \subset \mathbb{R}$. Let $\mathbf{L}_{11}^*$ be the Cholesky factor of the solution of the Riccati equation (12). Now, if we fix the $\mathbf{L}_{11}$ entry of $\mathbf{L}$ at $\mathbf{L}_{11}^*$, Fig. 1 represents the objective function with respect to $\mathbf{L}_{12}$ and $\mathbf{L}_{22}$ (the *insensitivity* of J to $\mathbf{L}_{22}$ when most samples are *inside* the unit interval is clearly illustrated in this figure as well). If we consider larger domains, e.g. $\Omega = [-20; 20] \subset \mathbb{R}$, Figs. 2 and 3 show that the optimization problem quickly becomes non-convex.

This non-convexity can be partially overcome by introducing a *windowing* procedure that warm starts a sequence of optimization problems over progressively larger regions $\Omega_1 \subset \dots \subset \Omega_k \subset \dots \subset \Omega$. We start a collocation problem for Ω_1 , a small region around the origin, where the SOS approximation will easily capture the behaviour of the lower order terms of the polynomial approximation. Then, the optimization results for $\mathbf{L}$ are used as warm starts (initial guesses) for problems with successively larger domains Ω_k ,

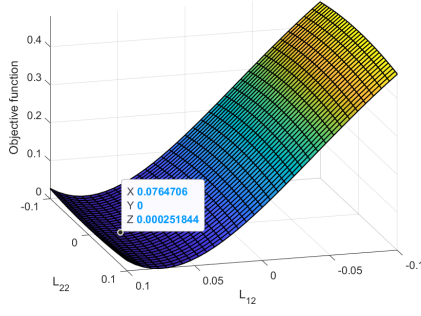


Fig. 1. Objective function when fixing $\mathbf{L}_{11}$ at $\mathbf{L}_{11}^*$ and choosing collocation points in the range $[-1, 1]$.

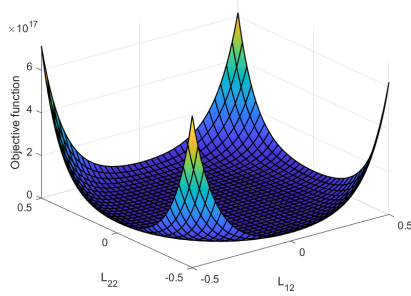


Fig. 2. Objective function when fixing $\mathbf{L}_{11}$ at $\mathbf{L}_{11}^*$ and choosing collocation points in the range $[-20, 20]$.

while we also increase the number of sample points to s_k . In this way, we achieve

$$\mathbf{L}_{\Omega_1}^{*(s_1)} \rightarrow \mathbf{L}_{\Omega_2}^{*(s_2)} \rightarrow \dots \rightarrow \mathbf{L}_{\Omega}^{*(N)}.$$

F. SOS nonlinear least-squares collocation algorithm

Algorithm 1 summarizes the procedure of approximating energy functions using the SOS collocation method described in Section III-C. We begin by solving the algebraic Riccati equation corresponding to the type of energy function we want to approximate, (12) or (13) to provide an initial guess for the $\mathbf{L}_{11}$ term. Then, we set the appropriate dimension ν of the $\mathbf{L}$ matrix: depending on the dimension of the system n , the degree of the approximation d , and the sequence of polynomial approximations (deciding on diagonal blocks of $\mathbf{L}$ to keep). The optimization problem is solved in a local region around the origin, to get the

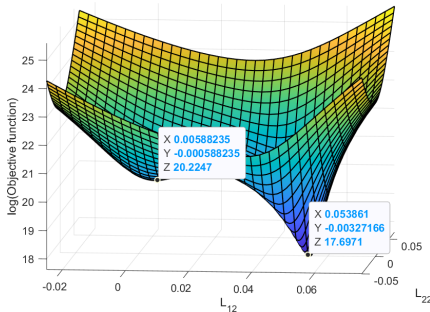


Fig. 3. Log of the objective function when fixing $\mathbf{L}_{11}$ at $\mathbf{L}_{11}^*$ and choosing collocation points in the range $[-20, 20]$.

first warm start on Ω_1 . We then proceed with windowing until the desired approximation region is reached. Finally, we construct the $\mathbf{Q} = \mathbf{L}\mathbf{L}^\top$ matrix that defines the energy function.

Algorithm 1: SOS approximation of energy functions.

Input: System matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{F}$, η , d the degree of SOS approximation, a nested set of approximation regions $\Omega_0 \subset \Omega_1 \dots \subset \Omega_r = \Omega$ and corresponding numbers of sample points $\{s_i\}_{i=0}^r$ ($s_r = N$).

Output: $\mathbf{Q}$, where $\mathbf{Q} = \mathbf{L}\mathbf{L}^\top$ in (15).

- 1) Solve the Algebraic Riccati equation (12) or (13), corresponding to the desired energy function, to get a warm start for the lowest degree term.
- 2) Calculate ν that depends on both n and d .
- 3) Calculate, ν_1 , the appropriate dimension that depends on the structure of the $\mathbf{L}$ resulting from the sequence of polynomial approximations.
- 4) To get a warm start $\mathbf{L}_0$ for the optimization, solve:

$$\mathbf{L}_0 = \underset{\mathbf{L} \in \mathbb{R}^{\nu \times \nu_1}}{\operatorname{argmin}} \sum_{k=1}^{s_0} \|\mathbf{R}(\mathbf{x}^{(k)})\|^2$$

where $\{\mathbf{x}^{(k)}\}_{k=1}^{s_0} \subset \Omega_0$, a small region containing $\mathbf{0}$.

- 5) For $i = 1, 2, \dots, r$

Solve:

$$\mathbf{L}_i = \underset{\mathbf{L} \in \mathbb{R}^{\nu \times \nu_1}}{\operatorname{argmin}} \sum_{k=1}^{s_i} \|\mathbf{R}(\mathbf{x}^{(k)})\|^2$$

with $\mathbf{L}_{i-1}$ as a warm start, and

s_i is the number of samples in Ω_i .

end

- 6) Finally, set $\mathbf{L} = \mathbf{L}_r$ and compute $\mathbf{Q} = \mathbf{L}\mathbf{L}^\top$.

IV. NUMERICAL RESULTS

We will perform numerical demonstrations of the proposed methods using a scalar example [14] to approximate the past energy $\mathcal{E}_\eta^-(\mathbf{x})$. Then we will test Algorithm 1 on a 6-dimensional ring of van der Pol oscillators where we approximate the future energy function $\mathcal{E}_\eta^+(\mathbf{x})$. Finally, we consider a 12-dimensional problem resulting from finite element approximation to Burgers equation. For the optimization, we used the IPOPT algorithm [29] along with the Julia package JuMP [17].

A. Scalar example

For our first study, we approximate the past energy function $\mathcal{E}_\eta^-(\mathbf{x})$ ($\gamma = \sqrt{2}$, $\eta = 0.5$) with degree 4, 6, and 8 SOS polynomials, for a scalar example found in [14] using $\Omega = [-8, 8]$. The system is given by

$$\dot{x}(t) = -2x(t) + x^2(t) + 2u(t) \quad \text{and} \quad y(t) = 2x(t).$$

We used symmetric windows starting at $[-1, 1]$ and doubling the size until we reached $[-8, 8]$. Since we can solve the HJB

in the 1D setting, the results of the polynomial approximations in Section II-B and III-F are compared to the analytical solution. The errors in the polynomial approximation, and the SOS approximation are presented in Figs. 4, 5, and 6.

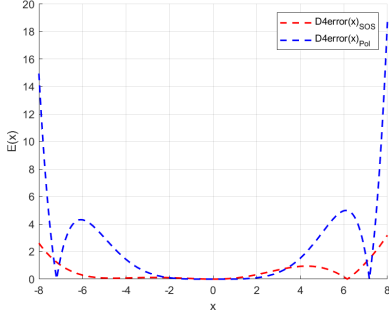


Fig. 4. Error of degree 4 approximations of energy functions $[-8, 8]$

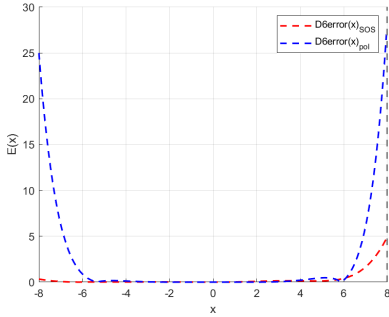


Fig. 5. Error of degree 6 approximations of energy functions $[-8, 8]$

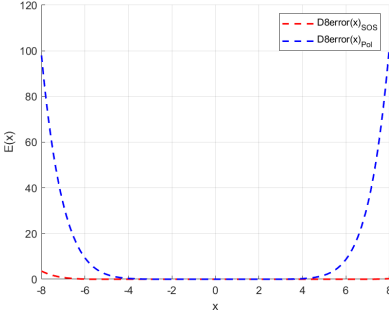


Fig. 6. Error of degree 8 approximations of energy functions $[-8, 8]$

As we increase the degree of the SOS approximation, the overall error gets smaller and is superior to the polynomial approximation in [14]. As expected, the polynomial approximation is good locally, but has a poor global behaviour that does not preserve the positivity and monotonicity properties that are expected for this example.

B. Ring of van der Pol equations

Consider the control of a ring of van der Pol oscillators,

$$\ddot{y}_i + (y_i^2 - 1)\dot{y}_i + y_i = y_{i-1} - 2y_i + y_{i+1} + b_i u_i(t), \quad (18)$$

for $i = 1, \dots, g$ with $y_i(0) = y^0$ and $\dot{y}_i(0) = 0$ (we identify $y_{g+1} = y_1$ and $y_g = y_0$ to close the ring). The coupling

TABLE I
STABILITY OF CLOSED-LOOP SYSTEMS FROM SAMPLED INITIAL GUESSES.

Windows in $\mathbb{R}^6$	Poly Approx	SOS Approx
$[-0.1, 0.1]^6$	Stable	Stable
$[-0.2, 0.2]^6$	Stable	Stable
$[-0.3, 0.3]^6$	Unstable 5/1000	Stable
$[-0.4, 0.4]^6$	Unstable 51/1000	Stable
$[-0.5, 0.5]^6$	Unstable 116/1000	Stable

terms on the right-hand-side resemble a discrete Laplacian. The stability of this system was studied in [19] and a related control problem considered in [3]. Here, we will consider the case when $g = 3$, with $b_1 = b_2 = 1$, $b_3 = 0$. Writing this as a first-order system places the system in the form (1) with $n = 6$ states and $m = 2$ control inputs. We define $\mathbf{C}$ using the positions of the 3 oscillators. We demonstrate our proposed SOS approximation on the future energy function $\mathcal{E}_\eta^+(\mathbf{x})$ using a degree 4 SOS approximation for the case $\eta = 1$.

Since we do not have an analytical solution available, we will evaluate the quality of the approximation by deriving the optimal control for the system using both the SOS approximation and the polynomial approximations

$$\begin{aligned} u_{sos}^*(t) &= -\mathbf{B}^\top \frac{\partial \mathcal{E}_{sos}^+(\mathbf{x})}{\partial \mathbf{x}} \\ u_{pol}^*(t) &= -\mathbf{B}^\top \frac{\partial \mathcal{E}_{pol}^+(\mathbf{x})}{\partial \mathbf{x}} \end{aligned} \quad (19)$$

running the closed-loop simulations of (18) and simultaneously approximating the integral term in the energy function given as

$$\mathcal{E}_\eta^+(\mathbf{x}_0) = \min_{\substack{u \in L_2[0, \infty) \\ \mathbf{x}(-\infty)=0 \quad \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2} \int_0^\infty \|y(t)\|^2 + \|u(t)\|^2 dt, \quad (20)$$

where x_0 is the initial condition for the system. We used 5 symmetric windows starting at $[-0.1, 0.1]^6$ and increasing the size by 0.1 in each direction until we reached $[-0.5, 0.5]^6$. To remove bias in our tests, we performed study using 1000 random initial conditions $\mathbf{x}_0$ in the approximation domains. The number of times the closed-loop system is unstable (or if was always stable) is reported in Table I. To show that the sum of squares polynomial approximation has comparable accuracy, we report the average relative errors in Table II. Note that we did not include the unstable results from the polynomial approximations in the reported averages (since the simulations didn't finish).

Table I shows that the SOS approximation provided a stable closed-loop solution for each of the 1000 random starting points in each approximation domain. On the other hand, as we increased the size of approximation region the polynomial approximation from [14] tends to give unstable solutions. This is to be expected since the polynomial approximations are designed to be very accurate near the origin, but without being an SOS polynomial, there is no

TABLE II

AVERAGE RELATIVE ERRORS IN $\mathcal{E}_1^+$ FROM 1000 INITIAL CONDITIONS.
THE ERRORS FROM THE UNSTABLE CASES WERE NOT INCLUDED.

Windows in $\mathbb{R}^6$	Poly Approx error	SOS Approx error
$[-0.1, 0.1]^6$	3.5465×10^{-4}	4.1133×10^{-3}
$[-0.2, 0.2]^6$	5.1609×10^{-4}	1.5224×10^{-2}
$[-0.3, 0.3]^6$	2.9679×10^{-3}	3.0859×10^{-2}
$[-0.4, 0.4]^6$	1.2120×10^{-2}	5.1790×10^{-2}
$[-0.5, 0.5]^6$	1.7436×10^{-2}	7.1467×10^{-2}

guarantee that positivity is maintained away from the origin. For example, starting at the initial condition,

$$\mathbf{x}_0^{\text{unstable}} = (-0.21 \ 0.08 \ 0.06 \ -0.35 \ 0.36 \ -0.47)^\top$$

leads to an unstable simulation. Table II shows that the average relative error, excluding the unstable initial conditions for the polynomial approximation, is acceptable for both the SOS and the polynomial approximation, with a clear advantage for the polynomial approximation in each of the approximation regions. The main advantage of the SOS approximations is that every initial condition in the neighborhood of the origin can be stabilized using (19).

C. Burgers equation

We also consider the control of the discretized 1D Burgers equation (using linear finite elements). Consider

$$z_t = -zz_x + \epsilon z_{xx} = \sum_{i=1}^m \chi_i^m(x) u_i(t), \quad x \in (0, 1), \quad t > 0, \quad (21)$$

with periodic boundary conditions and controlled outputs $y_i(t) = \int_0^1 \chi_i^p(x) z(t, x) dx$, $i = 1, \dots, p$. Here, $\chi_i^r(x)$ is the characteristic function over $[(i-1)/r, i/r]$ and $\epsilon = 5 \times 10^{-3}$. Approximating this PDE using 12 linear finite elements and choosing $m = p = 6$ creates a quadratic system of equations ($n = 12$) of the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{N}\mathbf{x}^\odot + \mathbf{B}u, \quad \text{and} \quad y = \mathbf{C}\mathbf{x}. \quad (22)$$

We demonstrate our proposed SOS approximation on the future energy function $\mathcal{E}_\eta^+(\mathbf{x})$ using a degree 4 SOS approximation for the case $\eta = 1$. As for the ring of van der Pol equations, we do not have an analytical solution. We will evaluate the quality of the approximation by deriving the optimal control for the discretized Burgers equation using both the SOS approximation and the polynomial approximation (19) and running the closed-loop simulations of (22) and simultaneously approximating the integral term in the energy function given by (20). For this example we ran the optimization for one window $[-0.1, 0.1]^{12}$, with 1000 random initial conditions $\mathbf{x}_0$ in the same window. To show the effectiveness of the approximation outside the optimization range, we also take 1000 random initial conditions $\mathbf{x}_0$ in $[-0.2, 0.2]^{12}$, $[-0.3, 0.3]^{12}$ and $[-0.4, 0.4]^{12}$ outside of the approximation window. We report the average relative errors in Table III. The results in Table III show that in this case both the SOS approximation and the polynomial approximation provide a

TABLE III

AVERAGE RELATIVE ERRORS IN $\mathcal{E}_1^+$ FROM 1000 INITIAL CONDITIONS.
THE ERRORS FROM THE UNSTABLE CASES WERE NOT INCLUDED.

Windows in $\mathbb{R}^{12}$	Poly. approx. error	SOS approx. error
$[-0.1, 0.1]^{12}$	7.3851×10^{-4}	3.4814×10^{-2}
$[-0.2, 0.2]^{12}$	8.1425×10^{-4}	4.5553×10^{-2}
$[-0.3, 0.3]^{12}$	1.0639×10^{-3}	6.0839×10^{-2}
$[-0.4, 0.4]^{12}$	1.8511×10^{-3}	7.5802×10^{-2}

good optimal control for the different random starting values both in the SOS approximation window and outside of it.

V. CONCLUSIONS

Three sum of squares polynomial approximation methods were described. In practice, the approach based on the form provided by [23] lead to the best results. The superiority of the SOS approximation over a Taylor series-based approximation was demonstrated using a scalar example with a known analytic formula. In a more challenging case, the results were mixed. The SOS approximation always provided stable closed-loop systems, but the Taylor polynomial approximations were more accurate on average when they produced stable closed-loop systems. However, away from the origin, the Taylor polynomial approximations failed to generate a stabilizing control in more than 10% of the randomly generated cases while the SOS polynomial approximations produced stable closed-loop systems in each of those same cases. The challenge of solving the nonlinear collocation problem limits the scalability of this approach to more modest sizes and higher degree SOS polynomials due to the large number of monomial terms that appear in the $\mathbf{L}$ matrices. Further work is underway to find better formulations of the optimization problems and test the performance on larger systems.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation under Grant CMMI-2130727.

REFERENCES

- [1] E. G. Albrekht. On the optimal stabilization of nonlinear systems. *Journal of Applied Mathematics and Mechanics*, 25:1254–1266, 1962.
- [2] Somil Bansal, Mo Chen, Sylvia Herbert, and Claire J. Tomlin. Hamilton-Jacobi reachability: A brief overview and recent advances, September 2017. arXiv:1709.07523 [cs, math].
- [3] Miguel Angel Barron. Stability of a ring of coupled van der Pol oscillators with non-uniform distribution of the coupling parameter. *Journal of Applied Research and Technology*, 14(1):62 – 66, 2016.
- [4] Michele Benzi and Valeria Simoncini. Approximation of functions of large matrices with Kronecker structure. *Numerische Mathematik*, 135(1):1–26, 2017.
- [5] Jeff Borggaard and Lizette Zietsman. The quadratic-quadratic regulator problem: Approximating feedback controls for quadratic-in-state nonlinear systems. In *2020 American Control Conference (ACC)*, pages 818–823. IEEE, 2020.
- [6] Jeff Borggaard and Lizette Zietsman. On approximating polynomial-quadratic regulator problems. *IFAC PaersOnLine*, 54(9):329–334, 2021.
- [7] Stephen P Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, 2004.

- [8] Tobias Breiten, Karl Kunisch, and Laurent Pfeiffer. Taylor expansions of the value function associated with a bilinear optimal control problem. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 36(5):1361–1399, 2019.
- [9] Minhong Chen and Daniel Kressner. Recursive blocked algorithms for linear systems with Kronecker product structure. *Numerical Algorithms*, 84(9):1199–1216, 2020.
- [10] Kenji Fujimoto and Daisuke Tsubakino. Computation of nonlinear balanced realization and model reduction based on Taylor series expansion. *Systems & Control Letters*, 57(4):283–289, 2008.
- [11] Didier Georges. The use of observability and controllability grammians or functions for optimal sensor and actuator location in finite-dimensional systems. In *Proceedings of 1995 34th IEEE Conference on Decision and Control*, volume 4, pages 3319–3324, New Orleans, LA, USA, 1995. IEEE.
- [12] Wassim M. Haddad, Tanmay Rajpurohit, and Xu Jin. Energy-based feedback control for stochastic port-controlled Hamiltonian systems. *Automatica*, 97:134–142, November 2018.
- [13] Zachary Jarvis-Wloszek, Ryan Feeley, Weehong Tan, Kunpeng Sun, and Andrew Packard. Controls applications of sum of squares programming. In *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Sciences*, pages 3–22. Springer, 2005.
- [14] Boris Kramer, Serkan Gugercin, Jeff Borggaard, and Linus Balecki. Nonlinear balanced truncation: Part 1—computing energy functions. 2023. arXiv:2209.07645 [math.OC].
- [15] Arthur J. Krener. Reduced order modeling of nonlinear control systems. In Alessandro Astolfi and Lorenzo Marconi, editors, *Analysis and Design of Nonlinear Control Systems*, pages 41–62. Springer Berlin Heidelberg, 2008.
- [16] Davide Lasagna, Deqing Huang, Owen R Tutty, and Sergei Chernyshenko. Sum-of-squares approach to feedback control of laminar wake flows. *Journal of Fluid Mechanics*, 809:628–663, 2016.
- [17] Miles Lubin, Oscar Dowson, Joaquim Dias Garcia, Joey Huchette, Benoît Legat, and Juan Pablo Vielma. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, 2023.
- [18] Dahlard L. Lukes. Optimal regulation of nonlinear dynamical systems. *SIAM Journal of Control*, 7:75–100, 1969.
- [19] Bonaventure Nana and Paul Wofo. Synchronization in a ring of four mutually coupled van der Pol oscillators: Theory and experiment. *Physical Review E*, 74(4):046213, 2006.
- [20] Carmeliza L. Navasca and Arthur J. Krener. Solution to Hamiltonian Jacobi Bellman equations. In *IEEE Conference on Decision and Control*, volume 39, pages 570–574, 2000.
- [21] Sajjad Pakkhesal and Saeed Shamaghdari. Sum-of-squares-based policy iteration for suboptimal control of polynomial time-varying systems. *Asian Journal of Control*, 24(6):3022–3031, 2022.
- [22] Antonis Papachristodoulou and Stephen Prajna. A tutorial on sum of squares techniques for systems analysis. Proc. 2005 American Control Conference, pages 2686–2700, 2005.
- [23] Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, Pasadena, CA, 2000.
- [24] Stephen Prajna, Antonis Papachristodoulou, and F Wu. Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach. Proc. 2004 5th Asian Control Conference, 2004.
- [25] Jacqueline M. Scherpen. H_∞ balancing for nonlinear systems. *International Journal of Robust and Nonlinear Control*, 6(7):645–668, 1996.
- [26] Jacqueline M. Scherpen and Arjan J. van der Schaft. Normalized coprime factorizations and balancing for unstable nonlinear systems. *International Journal of Control*, 60(6):1193–1222, 1994.
- [27] Jacqueline M. Scherpen. Balancing for nonlinear systems. *Systems & Control Letters*, 21:143–153, 1993.
- [28] Arjan J. van der Schaft. L_2 -gain analysis of nonlinear systems and nonlinear state-feedback H_∞ control. *IEEE Transactions on Automatic Control*, 37(6):770–784, 1992.
- [29] Andreas Wächter and Lorenz T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 2006.
- [30] Tianhao Wei and Changliu Liu. Safe control algorithms using energy functions: A unified framework, benchmark, and new directions, 2019. arXiv:1908.01883.