

Suboptimal Stabilizing Controllers for Linearly Solvable System

Yoke Peng Leong, Matanya B. Horowitz, and Joel W. Burdick

Abstract—This paper presents a novel method to synthesize stochastic control Lyapunov functions for a class of nonlinear, stochastic control systems. In this work, the classical nonlinear Hamilton-Jacobi-Bellman partial differential equation is transformed into a linear partial differential equation for a class of systems with a particular constraint on the stochastic disturbance. It is shown that this linear partial differential equation can be relaxed to a linear differential inclusion, allowing for approximating polynomial solutions to be generated using sum of squares programming. It is shown that the resulting solutions are stochastic control Lyapunov functions with a number of compelling properties. In particular, a-priori bounds on trajectory suboptimality are shown for these approximate value functions. The result is a technique whereby approximate solutions may be computed with non-increasing error via a hierarchy of semidefinite optimization problems.

I. INTRODUCTION

The stabilization of nonlinear systems is a central problem in control engineering. Lyapunov theory, wherein an energy-like function is used to show that some measure of distance from a stability point decays over time, is a critical tool for studying the convergence properties of a given system. Lyapunov theory may be generalized from analysis to synthesis of control systems using Control Lyapunov Function (CLF) [1]. However, the synthesis of a CLF for general systems remains a challenging open question, due to the bilinearity between the Lyapunov function and control input in the Lyapunov equation.

A complementary and related domain in control engineering is the study of the Hamilton-Jacobi-Bellman (HJB) equation, a partial differential equation that governs the optimal control of a system. Methods to calculate the solution to the HJB equation via semidefinite programming have been proposed previously by Lasserre et al. [2]. In this work, we propose an alternative line of study based on the linear structure of a particular form of the HJB equation. Since the late 1970s, researchers [3]–[6] have made connections between stochastic optimal control and reaction-diffusion equation through a logarithmic transformation. This line of research has recently been the subject of focused study by Kappen [7] and Todorov [8]. These results have been developed in a number of compelling directions [9]–[13].

This paper combines these previously disparate fields of dynamic programming and Lyapunov theory by considering the value function, the solution to a stochastic HJB equation,

as a Stochastic CLF (SCLF). The HJB solution is global, in that it incorporates all potential initial system states, and optimal. Here, we propose polynomial candidate approximate solutions to the HJB, extending recently developed tools in polynomial optimization to a new class of problems. It is already known that the solution to the deterministic HJB is in fact a CLF [14]. This paper shows that our approximated value function solutions are SCLFs as well.

A preliminary version of this work appeared in [15] and [16], where the use of semidefinite relaxations for solving the HJB were first considered. However, the stabilization properties of the resulting solutions were not investigated. Instead, these previous works focused on HJB solutions for path planning problems, and did not have guarantees on trajectory performance when using approximate solutions to the HJB.

The rest of this paper is organized as follows. Section II reviews the linearly solvable HJB equations, control Lyapunov functions, and sum of squares programming. Section III introduces a relaxed formulation of the HJB solutions which is efficiently computable using the sum of squares methodology. Section IV analyzes the properties of the relaxed solutions, such as approximation errors relative to the exact solutions. This section also shows that the relaxed solutions are SCLFs, and that the resulting controller is stabilizing. An example is presented in Section V to illustrate the optimization technique and its performance. Section VI summarizes the findings of this work and discusses future research directions.

II. BACKGROUND

This section briefly describes the notation and reviews necessary background on the linear HJB equation, SCLF, and SOS programming.

A. Notation

Table I summarizes the notation of different sets used in this work. A point on a trajectory, $x(t) \in \mathbb{R}^n$, at time t is denoted x_t , while the segment of this trajectory over the interval $[t, T]$ is denoted by $x_{[t, T]}$.

A compact domain in \mathbb{R}^n is denoted as Ω where $\Omega \subset \mathbb{R}^n$, and its boundary is denoted as $\partial\Omega$. A domain Ω is a *basic closed semialgebraic* set if there exists $g_i(x) \in \mathbb{R}[x]$ for $i = 1, 2, \dots, m$ such that $\Omega = \{x \mid g_i(x) \geq 0 \forall i = 1, 2, \dots, m\}$.

Given a polynomial $p(x)$, $p(x)$ is positive on domain Ω if $p(x) > 0 \forall x \in \Omega$, $p(x)$ is nonnegative on domain Ω if $p(x) \geq 0 \forall x \in \Omega$, and $p(x)$ is positive definite on domain Ω where $0 \in \Omega$, if $p(0) = 0$ and $p(x) > 0$ for all $x \in \Omega \setminus \{0\}$.

Y. P. Leong and M. B. Horowitz are with the Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA
ypleong@caltech.edu, mhorowitz@caltech.edu

J. W. Burdick is with the Mechanical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
jwb@robotics.caltech.edu

TABLE I
SET NOTATION

| Notation | Definition |
|------------------------------|---|
| \mathbb{Z}_+ | All positive integers |
| \mathbb{R} | All real numbers |
| \mathbb{R}_+ | All nonnegative real numbers |
| \mathbb{R}^n | All n -dimensional real vectors |
| $\mathbb{R}[x]$ | All real polynomial functions in x |
| $\mathbb{R}^{n \times m}$ | All $n \times m$ real matrices |
| $\mathbb{R}^{n \times m}[x]$ | All $M \in \mathbb{R}^{n \times m}$ such that $M_{i,j} \in \mathbb{R}[x] \forall i, j$ |
| \mathcal{K} | All continuous nondecreasing functions $\mu : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $\mu(0) = 0$, $\mu(r) > 0$ if $r > 0$, and $\mu(r) \geq \mu(r')$ if $r > r'$ |
| $\mathcal{C}^{k,k'}$ | All functions f such that f is k -differentiable with respect to the first argument and k' -differentiable with respect to the second argument |

If it exists, the infinity norm of a function is defined as $\|f\|_\infty = \sup_x |f(x)|$ for $x \in \Omega$. To improve readability, a function, $f(x_1, \dots, x_n)$, is abbreviated as f when the arguments of the function are clear from the context.

B. Linear Hamilton-Jacobi-Bellman (HJB) Equation

Consider the following affine nonlinear dynamical system,

$$dx_t = (f(x_t) + G(x_t)u_t) dt + B(x_t) d\omega_t \quad (1)$$

where $x_t \in \Omega$ is the state at time t in a compact state space domain $\Omega \subset \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ is the control input, $f(x) \in \mathbb{R}^n[x]$, $G(x) \in \mathbb{R}^{n \times m}[x]$, $B(x) \in \mathbb{R}^{n \times l}[x]$ are real polynomial functions of the state variables x , and $\omega_t \in \mathbb{R}^l$ is a vector consisting of Brownian motions with covariance Σ_ϵ , i.e., ω_t^i has independent increments with $\omega_t^i - \omega_s^i \sim \mathcal{N}(0, \Sigma_\epsilon(t-s))$, for $\mathcal{N}(\mu, \sigma^2)$ a normal distribution. The domain Ω is assumed to be a basic closed semialgebraic set defined as $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0 \forall i = 1, 2, \dots, m\}$. Without loss of generality, let $0 \in \Omega$ and $x = 0$ be the equilibrium point, whereby $f(0) = 0$, $G(0) = 0$ and $B(0) = 0$.

The goal is to minimize the following functional,

$$\mathbb{E}_{\omega_t}[J(x, u)] = \mathbb{E}_{\omega_t} \left[\phi(x_T) + \int_0^T q(x_t) + \frac{1}{2} u_t^T R u_t dt \right] \quad (2)$$

subject to (1), where $\phi \in \mathbb{R}[x]$, $\phi : \Omega \rightarrow \mathbb{R}_+$ represents a state-dependent terminal cost, $q \in \mathbb{R}[x]$, $q : \Omega \rightarrow \mathbb{R}_+$ is state dependent cost, and $R \in \mathbb{R}^{m \times m}$ is a positive definite matrix. T , unknown a priori, is the time at which the system reaches the domain boundary or the origin. This problem is generally called the *first exit* problem. The expectation \mathbb{E}_{ω_t} is taken over all realizations of the noise ω_t . For stability of the resultant controller to the origin, q and ϕ are also required to be positive definite functions. The solution to this minimization problem is known as the *value function*, $V : \Omega \rightarrow \mathbb{R}_+$, where beginning from an initial point x_t at time t

$$V(x_t) = \min_{u_{[t,T]}} \mathbb{E}_{\omega_t} [J(x_{[t,T]}, u_{[t,T]})]. \quad (3)$$

Based on dynamic programming arguments [17, Ch. III.7], the HJB equation associated with this problem is a nonlinear, second order partial differential equation (PDE)

$$0 = q + (\nabla_x V)^T f - \frac{1}{2} (\nabla_x V)^T G R^{-1} G^T (\nabla_x V) + \frac{1}{2} \text{Tr}((\nabla_{xx} V) B \Sigma_\epsilon B^T) \quad (4)$$

with boundary condition $V(x) = \phi(x)$ and the optimal control effort takes the form

$$u^* = -R^{-1} G^T \nabla_x V. \quad (5)$$

For the stabilization problem on a compact domain, it is appropriate to set the boundary condition to be $\phi(x) = 0$ for $x = 0$, indicating zero cost accrued for achieving the origin, and $\phi(x) > 0$ for $x \in \partial\Omega \setminus \{0\}$. In practice, $\phi(x)$ at the exterior boundary is usually chosen to be a large number depending on the applications to impose large penalty for exiting the predefined domain.

In general, (4) is difficult to solve due to its nonlinearity. However, with the assumption that there exists a $\lambda > 0$ and a control penalty cost R in (2) satisfying

$$\lambda G(x_t) R^{-1} G(x_t)^T = B(x_t) \Sigma_\epsilon B(x_t)^T \triangleq \Sigma(x_t) \triangleq \Sigma_t, \quad (6)$$

and using the logarithmic transformation

$$V = -\lambda \log \Psi, \quad (7)$$

it is possible [7], [8], after substitution and simplification, to obtain the following *linear* PDE from (4):

$$0 = -\frac{1}{\lambda} q \Psi + f^T (\nabla_x \Psi) + \frac{1}{2} \text{Tr}((\nabla_{xx} \Psi) \Sigma_t) \quad x \in \Omega$$

$$\Psi(x) = e^{-\frac{\phi(x)}{\lambda}} \quad x \in \partial\Omega. \quad (8)$$

This transformation of the value function has been deemed the *desirability* function [8]. For brevity, define the following expression

$$\mathcal{L}(\Psi) \triangleq f^T (\nabla_x \Psi) + \frac{1}{2} \text{Tr}((\nabla_{xx} \Psi) \Sigma_t)$$

and the function $\psi(x)$ at the boundary as

$$\psi(x) \triangleq e^{-\frac{\phi(x)}{\lambda}} \quad x \in \partial\Omega.$$

Condition (6) restricts the design of the control penalty R , such that control effort is highly penalized in subspaces with little noise, and lightly penalized in those with high noise. A specific case for which this condition is satisfied is for systems in which $B(x_t) = G(x_t)$. Additional discussion is given in [8].

C. Stochastic Control Lyapunov Functions (SCLF)

Before the stochastic control Lyapunov function (SCLF) is introduced, the definitions for two forms of stability are provided, following the definitions in [18, Ch. 5].

Definition 1. Given (1), the equilibrium point at $x = 0$ is stable in probability for $t \geq 0$ if for any $s \geq 0$ and $\epsilon > 0$,

$$\lim_{x \rightarrow 0} P \left\{ \sup_{t > s} |X^{x,s}(t)| > \epsilon \right\} = 0$$

where $X^{x,s}$ is the trajectory of (1) starting from x at time s .

Intuitively, Definition 1 is similar to the notion of stability for deterministic systems. The following is a stronger stability definition that is similar to the notion of asymptotic stability for deterministic systems.

Definition 2. Given (1), the equilibrium point at $x = 0$ is asymptotically stable in probability if it is stable in probability and

$$\lim_{x \rightarrow 0} P \left\{ \lim_{t \rightarrow \infty} |X^{x,s}(t)| = 0 \right\} = 1$$

where $X^{x,s}$ is the trajectory of (1) starting from x at time s .

For stochastic systems, the SCLF and Lyapunov theorems are defined as follows.

Definition 3. A stochastic control Lyapunov function (SCLF) for system (1) is a positive definite function $\mathcal{V} \in \mathcal{C}^{2,1}$ on a compact domain $\mathcal{O} = \Omega \cup \{0\} \times \{t > 0\}$ such that

$$\begin{aligned} \mathcal{V}(0, t) &= 0, \quad \mathcal{V}(x, t) \geq \mu(|x|) \quad \forall t \\ \exists u(x, t) \text{ s.t. } L(\mathcal{V}(x, t)) &\leq 0 \quad \forall (x, t) \in \mathcal{O} \setminus \{(0, t)\} \end{aligned}$$

where $\mu \in \mathcal{K}$, and

$$L(\mathcal{V}) = \partial_t \mathcal{V} + \nabla_x \mathcal{V}^T (f + Gu) + \frac{1}{2} \text{Tr}((\nabla_{xx} \mathcal{V}) B \Sigma_\epsilon B^T). \quad (9)$$

Theorem 4. [18, Thm. 5.3] For system (1), assume that there exists a SCLF and a u defined in Definition 3. Then, the equilibrium point $x = 0$ is stable in probability, and u is a stabilizing controller.

To achieve the stronger condition of asymptotic stability in probability, we have the following result.

Theorem 5. [18, Thm. 5.5 and Cor. 5.1] For system (1), suppose that in addition to the existence of a SCLF and a u defined in Definition 3, u is time-invariant,

$$\begin{aligned} \mathcal{V}(x, t) &\leq \mu'(|x|) \quad \forall t \\ L(\mathcal{V}(x, t)) &< 0 \quad \forall (x, t) \in \mathcal{O} \setminus \{(0, t)\} \end{aligned}$$

where $\mu' \in \mathcal{K}$. Then, the equilibrium point $x = 0$ is asymptotically stable in probability, and u is an asymptotically stabilizing controller.

D. Sum of Squares (SOS) Programming

This section provides a brief review of SOS programming, the tool by which we will use to generate approximate solutions to the HJB equation. A complete introduction to the subject of SOS programming is available in [19].

Definition 6. A multivariate polynomial $f(x)$ is a sum of squares (SOS) if there exist polynomials $f_0(x), \dots, f_m(x)$ such that

$$f(x) = \sum_{i=0}^m f_i^2(x).$$

The set of SOS polynomials in x is denoted as $\mathbb{S}[x]$.

A sufficient condition for non-negativity of a polynomial $f(x)$ is that $f(x) \in \mathbb{S}[x]$. This seemingly simple fact is compelling, as testing the membership of a polynomial in $\mathbb{S}[x]$ may be performed as a convex problem [19].

Theorem 7. [19, Thm. 3.3] The existence of a SOS decomposition of a polynomial in n variables of degree $2d$ can be decided by solving a semidefinite programming (SDP) feasibility problem.

Hence, by adding SOS constraints to the set of all positive polynomials, testing nonnegativity of a polynomial becomes a tractable SDP problem. The converse question, is a nonnegative polynomial necessarily a SOS, is unfortunately false, indicating that this test is conservative [19]. Nonetheless, SOS feasibility is sufficiently powerful for our purposes.

Theorem 7 guarantees a tractable procedure to determine whether a particular polynomial, possibly parameterized, is a SOS polynomial. Our method combines multiple polynomial constraints into an optimization formulation. To do so, we need to define the following polynomial set.

Definition 8. The preordering of polynomials $g_i(x) \in \mathbb{R}[x]$ for $i = 1, 2, \dots, m$ is the set

$$\begin{aligned} P(g_1, \dots, g_m) &= \left\{ \sum_{\nu \in \{0,1\}^m} s_\nu(x) g_1(x)^{\nu_1} \dots g_m(x)^{\nu_m} \mid s_\nu \in \mathbb{S}[x] \right\}. \end{aligned} \quad (10)$$

The following proposition is useful to incorporate the domain Ω in our optimization formulation later.

Proposition 9. Given $f(x) \in \mathbb{R}[x]$, if $f(x) \in P(g_1, \dots, g_m)$, on the domain $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, 2, \dots, m\}\}$, then $f(x)$ is nonnegative on Ω . If there exists another polynomial $f'(x)$ such that $f'(x) \geq f(x)$, then $f'(x)$ is also nonnegative on Ω .

To illustrate how this proposition applies, consider a polynomial $f(x)$ on a domain defined by $x \in [-1, 1]$. The bounded domain can be equivalently defined by polynomials $g_1(x) = 1 + x$ and $g_2(x) = 1 - x$. To certify that $f(x) \geq 0$ on the specified domain, construct a function $h(x) = s_1(x)(1 + x) + s_2(x)(1 - x) + s_3(x)(1 + x)(1 - x)$ where $s_i \in \mathbb{S}[x]$ and certify that $f(x) - h(x) \geq 0$. Notice that $h(x) \in P(1 + x, 1 - x)$, so $h(x) \geq 0$. If $f(x) - h(x) \geq 0$, then $f(x) \geq h(x) \geq 0$. Proposition 9 is applied here. Finding the correct $s_i(x)$ is not trivial in general. Nonetheless, as mentioned earlier, if we further impose that $f(x) - h(x) \in \mathbb{S}[x]$, then checking if there exists $s_i(x)$ such that $f(x) - h(x) \in \mathbb{S}[x]$ becomes a semidefinite feasibility program as given by Theorem 7. More concretely, the procedure may begin with a limited polynomial degree for $s_i(x)$, increasing the degree until a certificate is found (if one exists) or the computation resources are exhausted.

To simplify notation in later text, given a domain $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, 2, \dots, m\}\}$, we set the notation $P(\Omega) = P(g_1, \dots, g_m)$.

III. SUM-OF-SQUARES RELAXATION OF THE HJB PDE

This section demonstrates how SOS programming can be used to solve the linear HJB via an SOS relaxation. We would like to emphasize the following standing assumption, typical of moment and SOS-based methods [2], [19].

Assumption 10. *Assume that system (1) evolves on a compact domain $\Omega \subset \mathbb{R}^n$ that is also a basic closed semialgebraic set such that $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, \dots, k\}\}$ for some $k \geq 1$. Then, the boundary $\partial\Omega$ is polynomial representable. We use the notation $\partial\Omega = \{x \mid h_i(x) \in \mathbb{R}[x], \prod_{i=1}^m h_i(x) = 0\}$ for some $m \geq 1$ to describe this boundary.*

The following definitions formalize several operators that will be useful in later text.

Definition 11. *Given a basic closed semialgebraic set $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, \dots, k\}\}$ and a set of SOS polynomials, $\mathcal{S} = \{s_\nu(x) \mid s_\nu(x) \in \mathbb{S}[x], \nu \in \{0, 1\}^k\}$, define the operator \mathcal{D} as*

$$\mathcal{D}(\Omega, \mathcal{S}) = \sum_{\nu \in \{0, 1\}^k} s_\nu(x) g_1(x)^{\nu_1} \cdots g_k(x)^{\nu_k}$$

where $\mathcal{D}(\Omega, \mathcal{S}) \in P(\Omega)$.

Definition 12. *Given a polynomial inequality, $p(x) \geq 0$, the boundary of a compact set $\partial\Omega = \{x \mid h_i(x) \in \mathbb{R}[x], \prod_{i=1}^m h_i(x) = 0\}$ and a set of polynomials, $\mathcal{T} = \{t_i(x) \mid t_i(x) \in \mathbb{R}[x], i \in \{1, \dots, m\}\}$, define the operator \mathcal{B} as*

$$\mathcal{B}(p(x), \partial\Omega, \mathcal{T}) = \{p(x) - t_i(x)h_i(x) \mid i \in \{1, \dots, m\}\}$$

where \mathcal{B} returns a set of polynomials that is nonnegative on $\partial\Omega$.

A. Relaxation of the HJB equation

For the remainder of this paper, we assume that the unique solution to (4) and (8) exists in the viscosity solutions sense (see [17], Chapter V) and denote the unique solutions as V^* and Ψ^* respectively.

The equality constraints of (8) may be relaxed (in either direction) as follows

$$\begin{aligned} \frac{1}{\lambda}q\Psi - \mathcal{L}(\Psi) &\leq (\geq) 0 \\ \Psi(x) &\leq (\geq) \psi(x) \quad x \in \partial\Omega. \end{aligned} \quad (11)$$

This relaxation provides a point-wise bound to the true solution, and it may be enforced via SOS programming. In particular, a solution to (11), denoted as Ψ_l (Ψ_u), is a lower (upper) bound on the solution Ψ^* over the domain Ω .

Proposition 13. *Given a smooth function Ψ_l (Ψ_u) that satisfies (11), then Ψ_l (Ψ_u) is a viscosity subsolution (supersolution) and $\Psi_l \leq \Psi^*$ ($\Psi_u \geq \Psi^*$) for all $x \in \Omega$.*

Proof. By [20, Def. 2.2], the solution Ψ_l is a viscosity subsolution. Note that Ψ^* is both a viscosity subsolution and a viscosity supersolution, and $\Psi_l \leq \Psi^*$ on the boundary $\partial\Omega$. Hence, by the maximum principle for viscosity solutions [20,

Thm 3.3], $\Psi_l \leq \Psi^*$ for all $x \in \Omega$. Similar argument applies for Ψ_u . \square

Because the logarithmic transform (7) is monotonic, one can relate these bounds on the desirability function to bounds on the value function as follows:

Proposition 14. *If the solution to (4) is V^* , given solutions $V_u = -\lambda \log \Psi_l$ and $V_l = -\lambda \log \Psi_u$ from (11), then $V_u \geq V^*$ and $V_l \leq V^*$.*

B. Controller Synthesis

Given that relaxation (11) results in a point-wise upper and lower bound to the exact solution of (8), we construct the following optimization that provides a suboptimal controller with bounded residual error:

$$\begin{aligned} \min_{\Psi_l, \Psi_u} \quad & \epsilon & (12) \\ \text{s.t.} \quad & \frac{1}{\lambda}q\Psi_l - \mathcal{L}(\Psi_l) \leq 0 & x \in \Omega \\ & 0 \leq \frac{1}{\lambda}q\Psi_u - \mathcal{L}(\Psi_u) & x \in \Omega \\ & \Psi_u - \Psi_l \leq \epsilon & x \in \Omega \\ & 0 \leq \Psi_l \leq \psi \leq \Psi_u & x \in \partial\Omega \\ & \partial_{x^i}\Psi_l \leq 0 & x^i \geq 0 \\ & \partial_{x^i}\Psi_l \geq 0 & x^i \leq 0 \\ & \Psi_l(0) = 1 \end{aligned}$$

where x^i is the i -th component of $x \in \Omega$. As mentioned in Section III-A, the first two constraints result from the relaxations of the HJB equation, and the fourth constraint arises from the relaxation of the boundary conditions. The third constraint ensures that the solution error is bounded by ϵ , and the last three constraints ensure that the solution yields a stabilizing controller, as will be made clear in Section IV.

In order to solve (12) as a semidefinite optimization problem, we restrict the polynomial inequalities such that they are SOS polynomials instead of nonnegative polynomials. Therefore, after applying Proposition 9 to the domain constraints, the resulting optimization is

$$\begin{aligned} \min_{\Psi_l, \Psi_u, \mathcal{S}, \mathcal{T}} \quad & \epsilon & (13) \\ \text{s.t.} \quad & \frac{1}{\lambda}q\Psi_l + \mathcal{L}(\Psi_l) - \mathcal{D}(\Omega, \mathcal{S}_1) \in \mathbb{S}[x] \\ & \frac{1}{\lambda}q\Psi_u - \mathcal{L}(\Psi_u) - \mathcal{D}(\Omega, \mathcal{S}_2) \in \mathbb{S}[x] \\ & \epsilon - (\Psi_u - \Psi_l) - \mathcal{D}(\Omega, \mathcal{S}_3) \in \mathbb{S}[x] \\ & \mathcal{B}(\Psi_l, \partial\Omega, \mathcal{T}_1) \in \mathbb{S}[x] \\ & \mathcal{B}(\psi - \Psi_l, \partial\Omega, \mathcal{T}_2) \in \mathbb{S}[x] \\ & \mathcal{B}(\Psi_u - \psi, \partial\Omega, \mathcal{T}_3) \in \mathbb{S}[x] \\ & -\partial_{x^i}\Psi_l - \mathcal{D}(\Omega \cap \{x^i \geq 0\}, \mathcal{S}_4) \in \mathbb{S}[x] \\ & \partial_{x^i}\Psi_l - \mathcal{D}(\Omega \cap \{-x^i \geq 0\}, \mathcal{S}_5) \in \mathbb{S}[x] \\ & \Psi_l(0) = 1 \end{aligned}$$

where $\mathcal{S} = (\mathcal{S}_1, \dots, \mathcal{S}_5)$, $\mathcal{S}_i \subseteq \mathbb{S}[x]$ is defined as in Definition 11, $\mathcal{T} = (\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$, and $\mathcal{T}_j \subseteq \mathbb{R}[x]$ is defined as

in Definition 12. With a slight abuse of notation, $\mathcal{B}(\cdot) \in \mathbb{S}[x]$ implies that each polynomial in $\mathcal{B}(\cdot)$ is a SOS polynomial.

If the degrees of polynomials are fixed, optimization (13) is convex and may be solved as an SDP via Theorem 7. The next section will discuss the systematic approach we used to solve the optimization.

Remark 15. *By definition, the viscosity solution is a continuous function [20, Def. 2.2]. Consequently, the solution Ψ^* is a continuous function defined on a bounded domain. Hence, Ψ_u and Ψ_l can be made arbitrary close to Ψ^* by the Stone-Weierstrass Theorem [21] in (12). However, this guarantee is lost when Ψ_u and Ψ_l are restricted to be SOS polynomials. The feasible set of the optimization problem (13) is therefore not necessarily non-empty for a given polynomial degree.*

C. Hierarchy of SOS programs

Let d be the maximum degree of Ψ_l , Ψ_u and polynomials in \mathcal{S} and \mathcal{T} , and denote $(\Psi_u^d, \Psi_l^d, \mathcal{S}^d, \mathcal{T}^d, \epsilon^d)$ as a solution to (13) when the maximum polynomial degree is fixed at d . The hierarchy of SOS programs with increasing polynomial degree produces a sequence of possibly empty solutions $(\Psi_u^d, \Psi_l^d, \mathcal{S}^d, \mathcal{T}^d, \epsilon^d)_{d \in I}$, where $I \subset \mathbb{Z}_+$. This sequence will be shown in the next section to improve, under the metric of the objective in (13). The use of such hierarchies has become common in polynomial optimization [19], [22]. Once a satisfactory error is achieved or computational resources run out, the lower bound Ψ_l is used to compute the suboptimal controller. The suboptimal controller u^ϵ for a given error ϵ is computed as $u^\epsilon = -R^{-1}G^T \nabla_x V_u$ where $V_u = -\lambda \log \Psi_l$. The next section will analyze the properties of the solutions and the suboptimal controller.

IV. ANALYSIS

This section establishes appealing properties of the solutions to the optimization (13) that are relevant for feedback control. First, we show that the solutions in the SOS program hierarchy are uniformly bounded relative to the exact solutions. We next prove that the solutions to the relaxed stochastic HJB equation are SCLFs, and they yield stabilizing controllers. Finally, we show that the costs of using the approximate solutions as controllers are bounded above by the approximated value functions.

A. Properties of the Approximated Desirability Functions

First, compute the approximation error of the true desirability function Ψ_l or Ψ_u obtained from optimization (13).

Proposition 16. *Given a solution $(\Psi_u, \Psi_l, \mathcal{S}, \mathcal{T}, \epsilon)$ to (13) for a fixed degree d , the approximation error of a desirability function is bounded as $\|\Psi - \Psi^*\|_\infty \leq \epsilon$ where Ψ is either Ψ_u or Ψ_l .*

Proof. By Corollary 13, Ψ_l is the lower bound of Ψ^* , and Ψ_u is the upper bound of Ψ^* . So, $\epsilon \geq \Psi_u - \Psi_l \geq 0$ and $\Psi_u \geq \Psi^* \geq \Psi_l$. Combining both inequalities, one has $\Psi_u - \Psi^* \leq \epsilon$ and $\Psi^* - \Psi_l \leq \epsilon$. Therefore, $\|\Psi - \Psi^*\|_\infty \leq \epsilon$ where Ψ is either Ψ_u or Ψ_l . \square

Proposition 17. *The hierarchy of SOS programs consisting of solutions to (13) with increasing polynomial degree produces a sequence of solutions $(\Psi_u^d, \Psi_l^d, \mathcal{S}^d, \mathcal{T}^d, \epsilon^d)$ such that $\epsilon^{d+1} \leq \epsilon^d$ for all d .*

Proof. Polynomials of degree d form a subset of polynomials of degree $d+1$. Thus, at a higher polynomial degree $d+1$, a previous solution at a lower polynomial degree d is still a feasible solution when the coefficients for monomials with total degree $d+1$ is set to 0. Consequently, the optimal value ϵ^{d+1} cannot be smaller than ϵ^d for all d . \square

Although the bound on the pointwise error is non-increasing, the actual error may in fact increase between iterations. We bound this variation as follows.

Corollary 18. *Suppose $\|\Psi^d - \Psi^*\|_\infty \leq \epsilon^d$ and $\|\Psi^{d+1} - \Psi^*\|_\infty = \gamma^{d+1}$. Then, $\gamma^{d+1} \leq \epsilon^d$.*

Proof. From Proposition 17, $\gamma^{d+1} \leq \epsilon^{d+1} \leq \epsilon^d$. \square

Note that ϵ is only non-increasing as polynomial degree increases. Therefore, Proposition 17 and Corollary 18 does not guarantee a convergence of ϵ to zero.

B. Properties of the Approximated Value Function

We now investigate the implications of Corollary 18 upon the value function. Henceforth, denote the solution to (4) as $V^*(x_t) = \min_{u[t:T]} \mathbb{E}_{\omega_t} [J(x_t)] = -\lambda \log \Psi^*(x_t)$, and the suboptimal value function computed from the solution of (13) as $V_u = -\lambda \log \Psi_l$.

Theorem 19. *V_u is an upper bound of the optimal cost V^* such that*

$$0 \leq V_u - V^* \leq -\lambda \log \left(1 - \min \left\{ 1, \frac{\epsilon}{\eta} \right\} \right) \quad (14)$$

where $\eta = e^{-\frac{\|\Psi^*\|_\infty}{\lambda}}$.

Proof. By Proposition 14, $V_u \geq V^*$ and hence, $V_u - V^* \geq 0$. To prove the other inequality, by Proposition 16,

$$V_u - V^* = -\lambda \log \frac{\Psi_l}{\Psi^*} \leq -\lambda \log \frac{\Psi^* - \epsilon}{\Psi^*} \leq -\lambda \log \left(1 - \frac{\epsilon}{\eta} \right).$$

The last inequality holds because $\Psi^* \geq e^{-\frac{\|\Psi^*\|_\infty}{\lambda}}$ by definition in (7). Since Ψ_l is the lower bound of Ψ^* , the right hand side of the first equality is always a positive number. Therefore, V_u is a point-wise upper bound of V^* . \square

Corollary 20. *Let $V_u^d = -\lambda \log \Psi_l^d$ and $V_u^{d+1} = -\lambda \log \Psi_l^{d+1}$. If $V_u^d - V^* \leq \epsilon^d$ and $V_u^{d+1} - V^* = \gamma^{d+1}$, then $\gamma^{d+1} \leq -\lambda \log \left(1 - \min \left\{ 1, \frac{\epsilon^d}{\eta} \right\} \right)$.*

At this point, we have shown that the lower bound of the desirability function gives an upper bound of the suboptimal cost. More importantly, the upper bound of the suboptimal cost is non-increasing as the polynomial degree increases.

C. The Exact and Approximate HJB solutions are SCLFs

Here, we show that the approximate value function derived from the lower desirability approximation, Ψ_l , is a SCLF.

Theorem 21. V_u is a stochastic control Lyapunov function according to Definition 3.

Proof. The constraint $\Psi_l(0) = 1$ ensures that $V_u(0) = -\lambda \log \Psi_l(0) = 0$. Notice that all terms in $J(x, u)$ from (2) are positive definite, resulting in V^* being a positive definite function. In addition, by Proposition 14, $V^u \geq V^*$. Hence, V^u is also a positive definite function. The second and third to last constraints in (13) ensures that Ψ_l is nonincreasing. Hence, V_u is nondecreasing satisfying $\mu(|x|) \leq V_u(x) \leq \mu'(|x|)$ for some $\mu, \mu' \in \mathcal{K}$.

Next, show that there exists a u such that $L(V_u) \leq 0$. Following (5), let

$$u^\epsilon = -R^{-1}G^T \nabla_x V_u. \quad (15)$$

Notice that from the definition of V_u , $\nabla_x V_u = -\frac{\lambda}{\Psi_l} \nabla_x \Psi_l$ and $\nabla_{xx} V_u = \frac{\lambda}{\Psi_l^2} (\nabla_x \Psi_l) (\nabla_x \Psi_l)^T - \frac{\lambda}{\Psi_l} \nabla_{xx} \Psi_l$. So, $u = \frac{\lambda}{\Psi_l} R^{-1} G^T \nabla_x \Psi_l$. Then, from (9),

$$\begin{aligned} L(V_u) &= -\frac{\lambda}{\Psi_l} (\nabla_x \Psi_l)^T (f + \frac{\lambda}{\Psi_l} G R^{-1} G^T \nabla_x \Psi_l) \\ &\quad + \frac{1}{2} Tr \left(\left(\frac{\lambda}{\Psi_l^2} (\nabla_x \Psi_l) (\nabla_x \Psi_l)^T - \frac{\lambda}{\Psi_l} \nabla_{xx} \Psi_l \right) B \Sigma_\epsilon B \right) \end{aligned}$$

where $\partial_t V_u = 0$ because V_u is not a function of time. Applying the assumption in (6) and simplifying,

$$\begin{aligned} L(V_u) &= -\frac{\lambda}{\Psi_l} (\nabla_x \Psi_l)^T f - \frac{\lambda}{2\Psi_l^2} (\nabla_x \Psi_l)^T \Sigma_t \nabla_x \Psi_l \\ &\quad - \frac{\lambda}{2\Psi_l} Tr \left((\nabla_{xx} \Psi_l) \Sigma_t \right). \end{aligned}$$

From the first constraint in (13),

$$\begin{aligned} \frac{1}{\lambda} q \Psi_l - f^T (\nabla_x \Psi_l) - \frac{1}{2} Tr \left((\nabla_{xx} \Psi_l) \Sigma_t \right) \leq 0 &\implies \\ -\frac{\lambda}{\Psi_l} (\nabla_x \Psi_l)^T f \leq -q + \frac{\lambda}{2\Psi_l} Tr \left((\nabla_{xx} \Psi_l) \Sigma_t \right). \end{aligned}$$

Substituting this inequality into $L(V_u)$ and simplifying yields

$$L(V_u) \leq -q - \frac{\lambda}{2\Psi_l^2} (\nabla_x \Psi_l)^T \Sigma_t \nabla_x \Psi_l \leq 0 \quad (16)$$

because $q \geq 0$, $\lambda > 0$ and Σ_t is positive semidefinite by definition. Since V_u satisfies Definition 3, V_u is a SCLF. \square

Corollary 22. The suboptimal controller $u^\epsilon = -R^{-1}G^T \nabla_x V_u$ is stabilizing in probability within the domain Ω . If Σ_t is a positive definite matrix, the suboptimal controller $u^\epsilon = -R^{-1}G^T \nabla_x V_u$ is asymptotically stabilizing in probability within the domain Ω .

Proof. This corollary is a direct consequence of the constructive proof of Theorem 21 and Theorem 4. \square

D. Bound on the Total Trajectory Cost

We conclude this section by showing that the expected total trajectory cost incurred by the system while operating under the suboptimal controller of (15) is bounded.

Theorem 23. Given the control law $u^\epsilon = -R^{-1}G^T \nabla_x V_u$,

$$J_u \leq V_u \leq V^* - \lambda \log \left(1 - \min \left\{ 1, \frac{\epsilon}{\eta} \right\} \right) \quad (17)$$

where $J_u = \mathbb{E}_{\omega_t} [\phi_T(x_T) + \int_0^T r(x_t, u_t^\epsilon) dt]$, the expected cost of the system when using the given control law, u^ϵ .

Proof. By Itô's formula,

$$dV_u(x_t) = L(V_u)(x_t) dt + \nabla_x V_u(x_t) B(x_t) d\omega_t.$$

where $L(V)$ is defined in (9). Then,

$$\begin{aligned} V_u(x_t) &= V_u(x_0, 0) + \int_0^t L(V_u)(x_s) ds \\ &\quad + \int_0^t \nabla_x V_u(x_s) B(x_s) d\omega_s. \end{aligned} \quad (18)$$

Take the expectation of this equation to get

$$\mathbb{E}_{\omega_t} [V_u(x_t)] = V_u(x_0, 0) + \mathbb{E}_{\omega_t} \left[\int_0^t L(V_u)(x_s) ds \right]$$

whereby the last term of (18) drops out because the noise is assumed to have zero mean. The expectations of the other terms return the same terms because they are deterministic. From (16),

$$\begin{aligned} L(V_u) &\leq -q - \frac{\lambda}{2\Psi_l^2} (\nabla_x \Psi_l)^T \Sigma_t \nabla_x \Psi_l \\ &= -q - \frac{1}{2} (\nabla_x V_u)^T G R^{-1} G^T (\nabla_x V_u) \\ &= -q - \frac{1}{2} (u^\epsilon)^T R u^\epsilon \end{aligned}$$

where the first equality is given by the logarithmic transformation and the second equality is given by the control law $u^\epsilon = -R^{-1}G^T \nabla_x V_u$. Therefore,

$$\begin{aligned} \mathbb{E}_{\omega_t} [V_u(x_t)] &= V_u(x_0) + \mathbb{E}_{\omega_t} \left[\int_0^t L(V_u)(x_s) ds \right] \\ &\leq V_u(x_0) - \mathbb{E}_{\omega_t} \left[\int_0^t q(x_s) + \frac{1}{2} (u_s^\epsilon)^T R u_s^\epsilon ds \right] \\ &= V_u(x_0) - J(x_0, u^\epsilon) + \mathbb{E}_{\omega_t} [\phi(x_T)] \end{aligned}$$

Therefore, $V_u(x_0) - J(x_0, u^\epsilon) \geq \mathbb{E}_{\omega_t} [V_u(x_t) - \phi(x_t)]$. By definition, $V_u(x_T) \geq \phi(x_T)$ for all $x_T \in \Omega$. Thus, $\mathbb{E}_{\omega_t} [V_u(x_T) - \phi(x_T)] \geq 0$. Consequently, $V_u(x_0) - J(x_0, u^\epsilon) \geq 0$, and $V_u(x_0) \geq J(x_0, u^\epsilon)$. Lastly, Theorem 19 gives the second inequality in the theorem. \square

V. NUMERIC EXAMPLES

This section studies the computational characteristics of our method using a scalar unstable system. The optimization parser YALMIP [23] was used in conjunction with the semidefinite optimization package MOSEK [24] to solve the optimization problem (13).

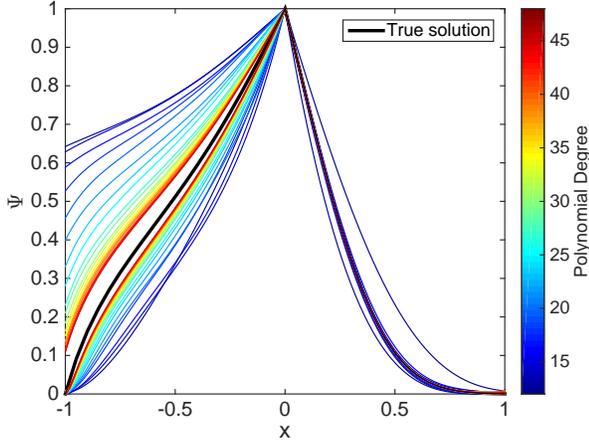


Fig. 1. The desirability function for varying polynomial degree. The true solution is the black curve.

Consider the following unstable scalar nonlinear system

$$dx = (-x^3 + 5x^2 + 3x + u) dt + d\omega \quad (19)$$

on the domain $x \in \Omega = \{x \mid -1 \leq x \leq 1\}$. The noise model considered is Gaussian white noise with zero mean and variance $\Sigma_\epsilon = 1$. The goal is to stabilize the system at the origin. Instead of zero, we choose the boundary at two ends of the domain to be $\Psi(-1) = 20e^{-10}$ and $\Psi(1) = 20e^{-10}$. At the origin, the boundary is set as $\Psi(0) = 1$. We set $q = x^2$, and $R = 1$. Because of the natural division of the domain, the solutions for both domains can be represented by smooth polynomials respectively, and solved independently.

The desirability functions that results from solving (13) for varying polynomial degrees are shown in Figure 1. The optimization problem is not feasible for polynomial degree below 12. The true solution is computed using Mathematica. The kink at the origin is expected because the HJB PDE solution is not necessarily smooth at the boundary, and in this situation the origin is itself a boundary between the two

domain halves. The approximation error ϵ for both partitions is shown in Figure 2(a) for increasing polynomial degree. As seen in the plots, the approximation improves as the polynomial degree increases.

To quantify the performance of the controller, a Monte Carlo experiment is performed. For each polynomial degree that is feasible, the controller obtained from Ψ_l in optimization (13) is implemented in 20 simulations of the system subject to random samples of Gaussian white noise with $\Sigma_\epsilon = 1$. The initial condition is fixed at $x_0 = -0.5$ and $t = 0$. The continuous system is integrated numerically using Euler integration with step size of 0.005s. The simulation is terminated when the trajectories enter the interval $[-0.005, 0.005]$ centered on the origin. Figure 2(c) shows the comparison between $J_u(x_0, t)$ and $V_u(x_0, t)$ for different polynomial degrees whereby J_u is the expected cost and V_u is the value function computed from Ψ_l in optimization (13). Figure 2(b) illustrates several sample trajectories. In general, the trajectories converge earlier when the polynomial degree is higher. This observation is expected because the approximation error is smaller as the polynomial degree increases.

VI. CONCLUSION

This paper proposes a novel method to solve the linear Hamilton Jacobi Equation of an optimal control problem with nonlinear, stochastic systems dynamics via sum of squares programming. Analytical results provide guarantees on the suboptimality of trajectories when using the approximate solutions for controller design. Consequently, one can synthesize a suboptimal stabilizing controller to nonlinear, stochastic dynamical systems.

To improve the algorithm, the monomials of the polynomial approximation can be chosen strategically in order to decrease computation time while achieving high accuracy. Thus, a promising future direction is the synthesis of the work presented here with that of [25], where HJB equations were solved in dimension twelve and higher. To improve

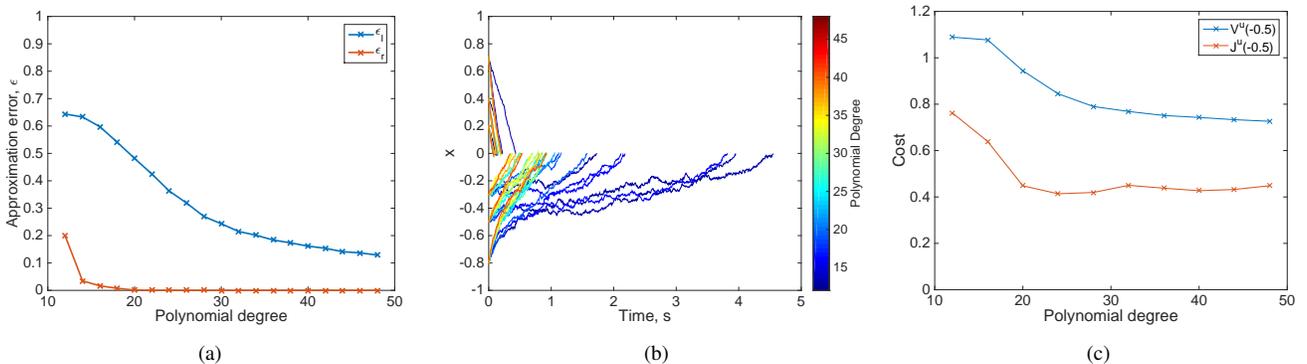


Fig. 2. Computational results of system (19). (a) Convergence of the objective function of (13) as the degree of polynomial increases. The approximation error for $x \leq 0$ is denoted as ϵ_l and the approximation error for $x \geq 0$ is denoted as ϵ_r . (b) Sample trajectories using controller computed from optimization problem (13) with different polynomial degrees starting from six randomly chosen initial points. (c) The comparison between J_u and V_u for different polynomial degrees whereby J_u is the expected cost and V_u is the value function computed from optimization problem (13). The initial condition is fixed at $x_0 = -0.5$.

the numerical conditioning of these optimization techniques, other numerical schemes are also under investigation [16].

There remains the question of the limitations placed by the structural constraint (6). A compelling research question is the suboptimality of controllers and trajectories when approximating systems that do not adhere to the constraint, such as deterministic systems or those with noise in states without a control channel.

REFERENCES

- [1] R. A. Freeman and J. A. Primbs, "Control Lyapunov functions: new ideas from an old source," in *Proceedings of the 35th IEEE Conference on Decision and Control*, vol. 4, 1996, pp. 3926–3931.
- [2] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, "Nonlinear optimal control via occupation measures and LMI-relaxations," *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [3] W. H. Fleming, *Logarithmic transformations and stochastic control*. Springer, 1982.
- [4] C. J. Holland, "A new energy characterization of the smallest eigenvalue of the schrödinger equation," *Communications on Pure and Applied Mathematics*, vol. 30, no. 6, pp. 755–765, 1977.
- [5] P. Dai Pra, "A stochastic control approach to reciprocal diffusion processes," *Applied mathematics and Optimization*, vol. 23, no. 1, pp. 313–329, 1991.
- [6] R. Filliger and M.-O. Hongler, "Relative entropy and efficiency measure for diffusion-mediated transport processes," *Journal of Physics A: Mathematical and General*, vol. 38, no. 6, p. 1247, 2005.
- [7] H. Kappen, "Linear theory for control of nonlinear stochastic systems," *Physical Review Letters*, vol. 95, no. 20, 2005.
- [8] E. Todorov, "Efficient computation of optimal actions," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 106, no. 28, pp. 11 478–11 483, 2009.
- [9] F. Stulp, E. A. Theodorou, and S. Schaal, "Reinforcement Learning With Sequences of Motion Primitives for Robust Manipulation," *IEEE Transactions on Robotics*, vol. 28, no. 6, pp. 1360–1370, 2012.
- [10] K. Dvijotham and E. Todorov, "Linearly-solvable optimal control," *Reinforcement learning and approximate dynamic programming for feedback control*, pp. 119–141, 2012.
- [11] B. Van Den Broek, W. Wiegerinck, and B. Kappen, "Graphical model inference in optimal control of stochastic multi-agent systems." *J. Artif. Intell. Res. (JAIR)*, vol. 32, pp. 95–122, 2008.
- [12] P. Rutquist, T. Wik, and C. Breitholtz, "Solving the Hamilton-Jacobi-Bellman equation for a stochastic system with state constraints," in *IEEE 53rd Conference on Decision and Control (CDC)*, 2014, pp. 1840–1845.
- [13] S. K. Shah and H. G. Tanner, "Dynamics-compatible potential fields using stochastic perturbations," in *23rd Mediterranean Conference on Control and Automation (MED)*, 2015, pp. 278–283.
- [14] J. A. Primbs, V. Nevistić, and J. C. Doyle, "Nonlinear optimal control: A control Lyapunov function and receding horizon perspective," *Asian Journal of Control*, vol. 1, no. 1, pp. 14–24, 1999.
- [15] M. Horowitz and J. Burdick, "Semidefinite relaxations for stochastic optimal control policies," in *American Controls Conf. (ACC)*, 2014, pp. 3006–3012.
- [16] M. B. Horowitz, I. Papusha, and J. W. Burdick, "Domain decomposition for stochastic optimal control," in *53rd Conference on Decision and Control (CDC)*, 2014, pp. 1866–1873.
- [17] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*. New York: Springer, 2006, vol. 25.
- [18] R. Khasminskii, *Stochastic Stability of Differential Equations*. Springer Science & Business Media, 2011, vol. 66.
- [19] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, vol. 96, no. 2, pp. 293–320, 2003.
- [20] M. G. Crandall, H. Ishii, and P.-L. Lions, "Users guide to viscosity solutions of second order partial differential equations," *Bulletin of the American Mathematical Society*, vol. 27, no. 1, pp. 1–67, 1992.
- [21] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1964, vol. 3.
- [22] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.
- [23] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB," in *IEEE International Symposium on Computer Aided Control Systems Design*, 2004, pp. 284–289.
- [24] E. D. Andersen and K. D. Andersen, "The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm," in *High performance optimization*. Springer, 2000, pp. 197–232.
- [25] M. B. Horowitz, A. Damle, and J. W. Burdick, "Linear Hamilton Jacobi Bellman Equations in High Dimensions," in *IEEE 53rd Conference on Decision and Control (CDC)*, 2014, pp. 5880–5887.