LINEARLY SOLVABLE STOCHASTIC CONTROL LYAPUNOV FUNCTIONS[*]

YOKE PENG LEONG[†], MATANYA B. HOROWITZ[†], AND JOEL W. BURDICK[‡]

Abstract. This paper presents a new method for synthesizing stochastic control Lyapunov functions for a class of nonlinear stochastic control systems. The technique relies on a transformation of the classical nonlinear Hamilton–Jacobi–Bellman partial differential equation to a linear partial differential equation for a class of problems with a particular constraint on the stochastic forcing. This linear partial differential equation can then be relaxed to a linear differential inclusion, allowing for relaxed solutions to be generated using sum of squares programming. The resulting relaxed solutions are in fact viscosity super-/subsolutions, and by the maximum principle are pointwise upper and lower bounds to the underlying value function, even for coarse polynomial approximations. Furthermore, the pointwise upper bound is shown to be a stochastic control Lyapunov function, yielding a method for generating nonlinear controllers with pointwise bounded distance from the optimal cost when using the optimal controller. These approximate solutions may be computed with nonincreasing error via a hierarchy of semidefinite optimization problems. Finally, this paper develops a priori bounds on *trajectory suboptimality* when using these approximate value functions and demonstrates that these methods, and bounds, can be applied to a more general class of nonlinear systems not obeying the constraint on stochastic forcing. Simulated examples illustrate the methodology.

Key words. stochastic control Lyapunov function, sum of squares programming, Hamilton–Jacobi–Bellman equation, nonlinear systems, optimal control

AMS subject classifications. 93E15, 93E20

DOI. 10.1137/16M105767X

1. Introduction. The study of system stability is a central theme of control engineering. A primary tool for such studies is Lyapunov theory, wherein an energy-like function is used to show that some measure of distance from a stability point decays over time. The construction of Lyapunov functions that certify system stability advanced considerably with the introduction of sum of squares (SOS) programming, which has allowed for Lyapunov functions to be synthesized for both polynomial systems [28] and more general vector fields [27].

To address the more challenging problem of stabilization, rather than the analysis of an existing closed loop system, it is possible to generalize Lyapunov functions to incorporate control inputs. The existence of a control Lyapunov function (CLF) (see [20, 11, 33]) is sufficient for the construction of a stabilizing controller. However, the synthesis of a CLF for general systems remains an open question. Unfortunately, the SOS-based methods cannot be naively extended to the generation of CLF solutions, due to the bilinearity between the Lyapunov function and control input.

Due to the lack of a general CLF synthesis technique, an alternative is the use of receding horizon control (RHC), which allows for the incorporation of optimality criteria. Euler–Lagrange equations are used to construct a locally optimum trajectory [30], and stabilization is guaranteed by constraining the terminal cost in the RHC problem to be a CLF. Suboptimal CLFs have found extensive use, with applications

in legged locomotion [19] and distributed control [26]. Adding stochasticity to the governing dynamics compounds the difficulties of constructing Lyapunov functions [5, 10]. A complementary area in control engineering is the study of the Hamilton–Jacobi–Bellman (HJB) 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. [22]. The method is quite general, applicable to any system with polynomial nonlinearities.

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, Fleming [8], Holland [12], and other researchers thereafter [4, 7] have made connections between stochastic optimal control and reaction-diffusion equations through a logarithmic transformation. Recently, when studying stochastic control using the HJB equation, Kappen [17] and Todorov [37] discovered that particular assumptions on the structure of a dynamical system, given the name *linearly solvable* systems, allow a logarithmic transformation of the optimal control equation to a linear partial differential equation (PDE) form. The linearity of this class of problems has given rise to a growing body of research, with an overview available in [6]. Kappen's work focused on calculating solutions via *path integral* techniques. Todorov began with the analysis of particular Markov decision processes and showed the connection between the two paradigms. This work was built upon by Theodorou [35] into the path integral framework in use with dynamic motion primitives. These results have been developed in many compelling directions [34, 6, 38, 32].

This paper combines these previously disparate fields of linearly solvable optimal control and Lyapunov theory and provides a systematic way to construct stabilizing controllers with guaranteed performance. The result is a hierarchy of SOS programs that generates stochastic CLFs (SCLFs) for arbitrary *linearly solvable* systems. Such an approach has many benefits. First and foremost, this approach generates stabilizing controllers for an important class of nonlinear stochastic systems even when the optimal controller is not found. We prove that the approximate solutions generated by the SOS programs are pointwise upper and lower bounds to the true solutions. In fact, the upper bound solutions are SCLFs, which can be used to construct stabilizing controllers, and they bound the performance of the system when they are used to construct suboptimal controllers. Existing methods for the generation of SCLFs do not have such performance guarantees. Additionally, we demonstrate that, although the technique is based on linear solvability, it may be readily extended to more general systems, including deterministic systems, while inheriting the same performance guarantees.

A preliminary version of this work appeared in [13] and [15], where the use of sum of squares programming for solving the HJB was first considered. This paper builds on this recent body of research, studying the stabilization and optimality properties of the resulting solutions. These previous works focused on path planning, rather than stabilization, and did not include the stability analysis or suboptimality guarantees presented in this paper. A short version of this work appeared as [24], which included less detail and did not include the extension in section 5.

The rest of this paper is organized as follows. Section 2 reviews linearly solvable HJB equations, SCLFs, and SOS programming. Section 3 introduces a relaxed formulation of the HJB solutions which is efficiently computable using the SOS methodology. Section 4 analyzes the properties of the relaxed solutions, such as approximation errors relative to the exact solutions. This section shows that the relaxed solutions are SCLFs, and that the resulting controller is stabilizing. The upper bound solu-

tion is also shown to bound the performance when using the suboptimal controller. Section 5 summarizes an extension of the method to approximate optimal control problems which are not linearly solvable. Two examples are presented in section 6 to illustrate the optimization technique and its performance. Section 7 summarizes the findings of this work and discusses future research directions.

2. Background. This section briefly describes the paper's notation and reviews necessary background on the linear HJB equation, SCLFs, and SOS programming.

2.1. Notation. Table 1 summarizes the notation of different sets appearing in this paper.

TABLE 1
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 $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ 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

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

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 $x_{t:T}$.

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

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 $f(x_1, \ldots, x_n)$ is abbreviated as f when the arguments of the function are clear from the context.

2.2. Linear Hamilton–Jacobi–Bellman (HJB) equation. Consider the following affine nonlinear dynamical system:

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

where $x_t \in \Omega$ is the state at time t in a compact domain $\Omega \subset \mathbb{R}^n$, and $u_t \in \mathbb{R}^m$ is the control input; $f(x) \in \mathbb{R}^n[x]$, $G(x) \in \mathbb{R}^{n \times m}[x]$, and $B(x) \in \mathbb{R}^{n \times l}[x]$ 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 has independent increments with $\omega_t - \omega_s \sim \mathcal{N}(0, \Sigma_\epsilon(t - s))$, for $\mathcal{N}(\mu, \sigma^2)$, a normal distribution. The domain Ω is assumed to be a basic closed semialgebraic set defined as $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0 \ \forall i = 1, 2, \ldots, m\}$. 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:

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

subject to (1), where $\phi \in \mathbb{R}[x]$, $\phi : \Omega \to \mathbb{R}_+$ represents a state-dependent terminal cost, $q \in \mathbb{R}[x]$, $q : \Omega \to \mathbb{R}_+$ is state-dependent cost, and $R \in \mathbb{R}^{m \times m}$ is a positive definite matrix. The final time, 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 $\mathbb{E}_{\omega_t}$ is taken over all realizations of the noise ω_t. For stability of the resultant controller to the origin, $q(\cdot)$ and $\phi(\cdot)$ are also required to be positive definite functions.

The solution to this minimization problem is known as the *value function*, V : $\Omega \to \mathbb{R}_+$, where beginning from an initial point x_t at time t,

$$(3) \qquad V(x_t) = \min_{u_{t:T}} J(x_{t:T}, u_{t:T}).$$

Based on dynamic programming arguments [9, Chap. III.7], the associated HJB equation is a nonlinear second order PDE:

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

with boundary condition $V(x) = \phi(x)$. 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 that depends on the given application, to impose a large penalty for exiting the predefined domain. The optimal control effort, u^*, is given by

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

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

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

and using the logarithmic transformation

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

it is possible [36, 37, 16], after substitution and simplification, to obtain the following *linear* PDE from (4):

$$(8) \qquad 0 = -\frac{1}{\lambda} q \Psi + f^T (\nabla_x \Psi) + \frac{1}{2} Tr \left((\nabla_{xx} \Psi) \Sigma_t \right), \quad x \in \Omega,$$

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

This transformation of the value function has been deemed the *desirability* function [37, Tab. 1]. For brevity, define the expression

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

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

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

Condition (6) is trivially met for systems of the form $dx_t = f(x_t)\, dt + G(x_t)\,(u_t\, dt + d\omega_t)$, a pervasive assumption in the adaptive control literature [23]. This constraint 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. Additional discussion is given in [37, Supplementary Information, sec. 2.2].

2.3. Stochastic control Lyapunov functions. Before the SCLF is introduced, two forms of stability are defined, following the definitions in [18, Chap. 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 $\epsilon > 0$,*

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

where $X^{x,s}$ is the trajectory of (1) *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*

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

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

These notions of stability can be realized through the construction of SCLFs.

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

$$\mathcal{V}(0, t) = 0, \quad \mathcal{V}(x, t) \geq \mu(|x|) \quad \forall\, t > 0,$$
$$\exists\, u(x, t) \ \ s.t. \ \ L(\mathcal{V}(x, t)) \leq 0 \quad \forall\, (x, t) \in \mathcal{O}\backslash\{(0, t)\},$$

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

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

THEOREM 4 (Theorem 5.3 of [18]). *For system* (1), *assume that there exist an SCLF and a u satisfying 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 (see Theorem 5.5 and Corollary 5.1 of [18]). *For system* (1), *suppose that, in addition to the existence of an SCLF and a u satisfying Definition 3, u is time-invariant, and*

$$\mathcal{V}(x, t) \leq \mu'(|x|) \quad \forall\, t > 0,$$
$$L(\mathcal{V}(x, t)) < 0 \quad \forall\, (x, t) \in \mathcal{O}\backslash\{(0, t)\},$$

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

2.4. Sum of squares programming. SOS programming is the primary tool by which approximate solutions to the HJB equation are generated in this paper. In particular, we will show how the PDE that governs the HJB may be relaxed to a set of nonnegativity constraints. SOS methods will then allow for the construction of an optimization problem where these nonnegativity constraints may be enforced. A complete introduction to SOS programming is available in [28]. Here, we review the basic definition of SOS that is used throughout the paper.

DEFINITION 6. *A multivariate polynomial $f(x)$ is an SOS polynomial if there exist polynomials $f_0(x), \ldots, f_m(x)$ such that*

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

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

Accordingly, a sufficient condition for nonnegativity of a polynomial $f(x)$ is that $f(x) \in \mathbb{S}[x]$. Membership in the set $\mathbb{S}[x]$ may be tested as a convex problem [28].

THEOREM 7 (Theorem 3.3 of [28]). *The existence of an SOS decomposition of a polynomial in n variables of degree $2d$ can be decided by solving a semidefinite programming (SDP) feasibility problem. If the polynomial is dense (no sparsity), the dimension of the matrix inequality in the SDP is equal to $\binom{n+d}{d} \times \binom{n+d}{d}$.*

Hence, by adding SOS constraints to the set of all positive polynomials, testing nonnegativity of a polynomial becomes a tractable SDP. The converse question, is a nonnegative polynomial necessarily an SOS, is unfortunately false, indicating that this test is conservative [28]. Nonetheless, SOS feasibility is sufficiently powerful for our purposes. Theorem 7 guarantees a tractable procedure to determine whether a particular polynomial, possibly parameterized, is an SOS polynomial. Our method combines multiple polynomial constraints into an optimization formulation. To do so, we need to define the following polynomial sets.

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

$$(10) \qquad P(g_1, \ldots, g_m) = \left\{ \sum_{\nu \in \{0,1\}^m} s_\nu(x) g_1(x)^{\nu_1} \cdots g_m(x)^{\nu_m} \;\middle|\; s_\nu \in \mathbb{S}[x] \right\}.$$

The following proposition is trivial, but it is useful to incorporate the domain Ω into our optimization formulation later.

PROPOSITION 9. *Given $f(x) \in \mathbb{R}[x]$ and the domain*

$$\Omega = \{x \mid g_i(x) \in \mathbb{R}[x],\ g_i(x) \geq 0,\ i \in \{1, 2, \ldots, m\}\},$$

if $f(x) \in P(g_1, \ldots, g_m)$, then $f(x)$ is nonnegative on Ω. If there exists another polynomial $f'(x)$ such that $f'(x) \geq f(x)$ for all $x \in \Omega$, then $f'(x)$ is also nonnegative on Ω.

Proof. Because $g_i(x)$ and $s_i(x)$ are nonnegative, all functions in $P(\cdot)$ are nonnegative. The second statement is trivially true given the first statement. $\square$

Example. To illustrate an application of Proposition 9, consider a polynomial $f(x)$ defined on the domain $x \in [-1, 1]$. The bounded domain can be equivalently defined by polynomials with $g_1(x) = 1 + x$ 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 whether there exists $s_i(x)$ such that $f(x) - h(x) \in \mathbb{S}[x]$ becomes an SDP as given by Theorem 7.

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

Remark 10. Depending on the computational resources available, one may choose a subset of $P(\Omega)$ to reduce the size of the resulting SDP. However, the chances of finding a certificate are reduced as a consequence. This polynomial set is often used in the discussions of Schmüdgen's Positivstellensatz, which states that if $f(x)$ is positive on a compact domain Ω, then $f(x) \in P(\Omega)$ [22, 28].

3. SOS Relaxation of the HJB PDE. SOS programming has found many uses in combinatorial optimization, control theory, and other applications. This section now adds solving the linear HJB to this list. We would like to emphasize the following standing assumption, necessary in moment- and SOS-based methods [22, 28].

ASSUMPTION 11. *Assume that system* (1) *evolves on a compact domain $\Omega \subset \mathbb{R}^n$ and that Ω is a basic closed semialgebraic set such that $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, \ldots, k\}\}$ 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 the boundary.*

The following definitions formalize several operators that are useful in what follows.

DEFINITION 12. *Given a basic closed semialgebraic set $\Omega = \{x \mid g_i(x) \in \mathbb{R}[x], g_i(x) \geq 0, i \in \{1, \ldots, k\}\}$ 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 $s_\nu \in \mathcal{S}$ and $\mathcal{D}(\Omega, \mathcal{S}) \in P(\Omega)$.

DEFINITION 13. *Given a polynomial inequality $p(x) \geq 0$ defined on Ω, 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,*

$$\mathcal{T} = \{t_i(x) \mid t_i(x) \in \mathbb{R}[x], i \in \{1, \ldots, m\}\},$$

define the operator $\mathcal{B}$ as

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

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

3.1. Relaxation of the HJB equation. If the linear HJB (8) is not uniformly parabolic [3], a classical solution may not exist. The notion of *viscosity solutions* is developed to generalize the classical solution. We refer readers to [3] for a general discussion on viscosity solutions, and to [9] for a discussion on viscosity solutions related to Markov diffusion processes.

DEFINITION 14 (Definition 2.2 of [3]). *Given $\Omega \subset \mathbb{R}^N$ and a PDE*

$$(11) \qquad\qquad F(x, u, \nabla_x u, \nabla_{xx} u) = 0,$$

where $F : \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N) \to \mathbb{R}$, $\mathcal{S}(N)$ is the set of real symmetric $N \times N$ matrices, and F satisfies

$$F(x, r, p, X) \le F(x, s, p, Y) \quad \text{whenever } r \le s \text{ and } Y \le X,$$

then a viscosity subsolution *of* (11) *on Ω is a function $u \in USC(\Omega)$ such that*

$$F(x, u, \nabla_x u, \nabla_{xx} u) \le 0 \quad \forall\, x \in \Omega, \ (p, X) \in J_\Omega^{2,+} u(x).$$

Similarly, a viscosity supersolution *of* (11) *on Ω is a function $u \in LSC(\Omega)$ such that*

$$F(x, u, \nabla_x u, \nabla_{xx} u) \ge 0 \quad \forall\, x \in \Omega, \ (p, X) \in J_\Omega^{2,-} u(x).$$

Finally, u is a viscosity solution *of* (11) *on Ω if it is both a viscosity subsolution and a viscosity supersolution in Ω.*

The notations $USC(\Omega)$ and $LSC(\Omega)$ represent the sets of upper and lower semi-continuous functions on domain Ω, respectively, and $J_\Omega^{2,+} u(x)$ and $J_\Omega^{2,-} u(x)$ represent the second order "superjets" and "subjets" of u at x, respectively, a completely unrestrictive domain in our setting. For further details, readers may refer to [3]. For the remainder of this paper, we assume that unique nontrivial viscosity solutions to (4) and (8) exist (see [9, Chap. V]) and denote them as V^* and Ψ^*, respectively.

The equality constraints of (8) may be relaxed as follows:

$$(12a) \qquad\qquad \frac{1}{\lambda} q\Psi - \mathcal{L}(\Psi) \le (\ge)0,$$

$$(12b) \qquad\qquad \Psi(x) \le (\ge)\psi(x), \qquad x \in \partial\Omega.$$

Such a relaxation provides a pointwise bound to the solution Ψ^*, and this relaxation may be enforced via SOS programming. In particular, a solution to (12), denoted as Ψ_l (Ψ_u), is a lower (upper) bound on the solution Ψ^* over the entire problem domain.

THEOREM 15. *Given a smooth function Ψ_l (Ψ_u) that satisfies* (12), *then Ψ_l (Ψ_u) is a viscosity subsolution (supersolution), and $\Psi_l \le \Psi^*$ ($\Psi_u \ge \Psi^*$) for all $x \in \Omega$.*

Proof. By Definition 14, the solution Ψ_l is a viscosity subsolution where F in (11) is given by (12a). Note that Ψ^* is both a viscosity subsolution and a viscosity supersolution, and $\Psi_l \le \Psi^*$ on the boundary $\partial\Omega$. Hence, by the maximum principle [3, Thm. 3.3], $\Psi_l \le \Psi^*$ for all $x \in \Omega$. The proof is identical for Ψ_u. ☐

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 16. *If the solution to* (4) *is V^*, given solutions $V_u = -\lambda \log \Psi_l$ and $V_l = -\lambda \log \Psi_u$ from* (12), *then $V_u \ge V^*$ and $V_l \le V^*$.*

Proof. Recall that $V^* = -\lambda \log \Psi^*$. Apply Theorem 15, $V_u \geq V^*$, and $V_l \leq V^*$. $\square$

The solutions to (12) do not satisfy (8) exactly, but they provide pointwise bounds to the solution Ψ^*.

3.2. SOS program. Given that relaxation (12) results in a pointwise upper and lower bound to the exact solution of (8), we construct the following optimization problem that provides a suboptimal controller with bounded residual error:

$$(13) \qquad \min_{\Psi_l, \Psi_u} \quad \epsilon$$

$$\text{s.t.} \quad \frac{1}{\lambda} q \Psi_l - \mathcal{L}(\Psi_l) \leq 0, \quad x \in \Omega,$$

$$0 \leq \frac{1}{\lambda} q \Psi_u - \mathcal{L}(\Psi_u), \quad x \in \Omega,$$

$$\Psi_u - \Psi_l \leq \epsilon, \quad x \in \Omega,$$

$$0 \leq \Psi_l \leq \psi \leq \Psi_u, \quad x \in \partial\Omega,$$

$$\partial_{x^i} \Psi_l \leq 0, \quad x^i \geq 0,$$

$$\partial_{x^i} \Psi_l \geq 0, \quad x^i \leq 0,$$

$$\Psi_l(0) = 1,$$

where x^i is the ith component of $x \in \Omega$. As mentioned in subsection 3.1, 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 difference between the upper bound and lower bound solution is bounded, and the last three constraints ensure that the solution yields a stabilizing controller, as will be made clear in section 4. Note that in the optimization problem, Ψ_u and Ψ_l are polynomials whereby the coefficients and the degree for both are optimization variables. The term ϵ is related to the error of the approximation.

As discussed in the review of SOS techniques, a general optimization problem involving parameterized nonnegative polynomials is not necessarily tractable. In order to solve (13) using a polynomial-time algorithm, we restrict the polynomial inequalities such that they are SOS polynomials instead of nonnegative polynomials. We therefore apply Proposition 9 to relax optimization problem (13) into

$$(14) \qquad \min_{\Psi_l, \Psi_u, \mathcal{S}, \mathcal{T}} \quad \epsilon$$

$$\text{s.t.} \quad -\frac{1}{\lambda} q \Psi_l + \mathcal{L}(\Psi_l) - \mathcal{D}(\Omega, \mathcal{S}_1) \in \mathbb{S}[x],$$

$$\frac{1}{\lambda} q \Psi_u - \mathcal{L}(\Psi_u) - \mathcal{D}(\Omega, \mathcal{S}_2) \in \mathbb{S}[x],$$

$$\epsilon - (\Psi_u - \Psi_l) - \mathcal{D}(\Omega, \mathcal{S}_3) \in \mathbb{S}[x],$$

$$\mathcal{B}(\Psi_l, \partial\Omega, \mathcal{T}_1) \in \mathbb{S}[x],$$

$$\mathcal{B}(\psi - \Psi_l, \partial\Omega, \mathcal{T}_2) \in \mathbb{S}[x],$$

$$\mathcal{B}(\Psi_u - \psi, \partial\Omega, \mathcal{T}_3) \in \mathbb{S}[x],$$

$$-\partial_{x^i} \Psi_l - \mathcal{D}(\Omega \cap \{x^i \geq 0\}, \mathcal{S}_4) \in \mathbb{S}[x],$$

$$\partial_{x^i} \Psi_l - \mathcal{D}(\Omega \cap \{-x^i \geq 0\}, \mathcal{S}_5) \in \mathbb{S}[x],$$

$$\Psi_l(0) = 1,$$

where $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_4, \mathcal{S}_5)$, $\mathcal{S}_i \subseteq \mathbb{S}[x]$ is defined as in Definition 12, $\mathcal{T} = (\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3)$, and $\mathcal{T}_j \subseteq \mathbb{R}[x]$ is defined as in Definition 13. With a slight abuse of notation, $\mathcal{B}(\cdot) \in \mathbb{S}[x]$ implies that each polynomial in $\mathcal{B}(\cdot)$ is an SOS polynomial.

If the polynomial degrees are fixed, optimization problem (14) is convex and solvable using a semidefinite program via Theorem 7. The next section will discuss the systematic approach we used to solve the optimization problem. Henceforth, denote the solution to (14) as $(\Psi_u, \Psi_l, \mathcal{S}, \mathcal{T}, \epsilon)$.

Remark 17. By Definition 14, the viscosity solution is a continuous function. Consequently, the solution Ψ^* is a continuous function defined on a bounded domain. Hence, Ψ_u and Ψ_l can be made arbitrarily close to Ψ^* by the Stone–Weierstrass theorem [31] in (13). However, this guarantee is lost when Ψ_u and Ψ_l are restricted to being SOS polynomials. The feasible set of the optimization problem (14) is therefore not necessarily nonempty for a given polynomial degree. One would not expect feasibility for all instances of (14), as this would imply that there exists a linear stabilizing controller for any given system.

3.3. Controller synthesis. Let d be the maximum degree of Ψ_l, Ψ_u and polynomials in $\mathcal{S}$ and $\mathcal{T}$, and denote by $(\Psi_u^d, \Psi_l^d, \mathcal{S}^d, \mathcal{T}^d, \epsilon^d)$ a solution to (14) 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 (14).

In other words, if solutions exist for d and d' such that $d > d'$, then $\epsilon^d \leq \epsilon^{d'}$. Therefore, one could keep increasing the degree of polynomials in order to achieve tighter bounds on Ψ^* and, invariably, V^*. The use of such hierarchies has become commonplace in polynomial optimization [21, 28]. If at a certain degree, $\epsilon^d = 0$, the solution Ψ^* is found.

Once a satisfactory error is achieved or computational resources run out, the lower bound Ψ_l^d can be used to compute a suboptimal controller where d is the maximum degree computed. Recall that $u^* = -R^{-1}G^T\nabla_x V^*$ and $V^* = -\lambda \log \Psi^*$. The suboptimal controller u^ϵ for a given degree d and error ϵ^d is computed as $u^{\epsilon^d} = -R^{-1}G^T\nabla_x V_u^d$, where $V_u^d = -\lambda \log \Psi_l^d$. Even when ϵ^d is larger than a desired value, the solution Ψ_l^d still satisfies conditions in Definition 3 to yield a stabilizing suboptimal controller. The next section will analyze properties of the solutions and the suboptimal controller.

4. Analysis. This section establishes several properties of the solutions to the optimization problem (14) that are useful 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 relaxed solutions to the stochastic HJB equation are SCLFs, and the approximated solution leads to a stabilizing controller. Finally, we show that the costs of using the approximate solutions as controllers are bounded above by the approximated value functions.

4.1. Properties of approximated desirability functions. First, the approximation error of Ψ_l or Ψ_u obtained from (14) is computed relative to the true desirability function Ψ^*.

PROPOSITION 18. *Given a solution $(\Psi_u^d, \Psi_l^d, \mathcal{S}^d, \mathcal{T}^d, \epsilon^d)$ to (14) for a given degree d, the approximation error of the desirability function is bounded as $||\Psi^d - \Psi^*||_\infty \leq \epsilon^d$, where Ψ^d is either Ψ_u^d or Ψ_l^d.*

Proof. By Theorem 15, Ψ_l^d is the lower bound of Ψ^*, and Ψ_u^d is the upper bound of Ψ^*. So, $\epsilon^d \geq \Psi_u^d - \Psi_l^d \geq 0$ and $\Psi_u^d \geq \Psi^* \geq \Psi_l^d$. Combining both inequalities, one has $\Psi_u^d - \Psi^* \leq \epsilon^d$ and $\Psi^* - \Psi_l^d \leq \epsilon^d$. Therefore, $||\Psi^d - \Psi^*||_\infty \leq \epsilon^d$, where Ψ^d is either Ψ_u^d or Ψ_l^d.   □

PROPOSITION 19. *The hierarchy of SOS programs consisting of solutions to* (14) *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 larger than ϵ^d for all d. □

Thus, as the polynomial degree of the optimization problem is increased, the pointwise error ϵ is nonincreasing. Therefore, one could keep increasing the degree of polynomials in order to achieve tighter bounds on Ψ^* and, invariably, V^*. However, ϵ is only nonincreasing as the polynomial degree is increased, and a convergence of the bound ϵ to zero is not guaranteed because we restrict the approximating space to SOS. The possible lack of convergence to zero is the trade-off for an efficient algorithm.

Although the *bound* on the pointwise error is nonincreasing, the actual difference between Ψ and Ψ^* may increase between iterations.

COROLLARY 20. *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. By Proposition 19, $\epsilon^{d+1} \leq \epsilon^d$. Because $\gamma^{d+1} \leq \epsilon^{d+1}$, $\gamma^{d+1} \leq \epsilon^d$.   □

In other words, the approximation error of the desirability function for an SOS program using $d+1$ polynomial degree cannot increase such that it is larger than ϵ^d in each step of the hierarchy of SOS programs, which is nonincreasing.

4.2. Properties of approximated value functions. Up to this point, the analysis has focused on properties of the desirability solution. We now investigate the implications of these results for the value function, which is related to the desirability via the logarithmic transform (7). Henceforth, denote the solution to (4) as $V^*(x_t) = \min_{u[t:T]} \mathbb{E}_{\omega_t}[J(x_t)] = -\lambda \log \Psi^*(x_t)$, the solution to (14) for a fixed degree d as $(\Psi_u, \Psi_l, \mathcal{S}, \mathcal{T}, \epsilon)$, and the suboptimal value function computed from the solution of (14) as $V_u = -\lambda \log \Psi_l$. Only Ψ_l and V_u are considered henceforth, because Ψ_l, but not Ψ_u, gives an approximate value function that satisfies the properties of an SCLF in Definition 3, a fact shown in the next section.

THEOREM 21. *For all* $x \in \Omega$, V_u *is an upper bound of* V^* *such that*

$$(15) \qquad 0 \leq V_u - V^* \leq -\lambda \log\left(1 - \min\left\{1, \frac{\epsilon}{\eta}\right\}\right),$$

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

Proof. By Proposition 16, $V_u \geq V^*$ and hence $V_u - V^* \geq 0$. To prove the other inequality, by Proposition 18,

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

The last inequality holds because $\Psi^* \geq e^{-\frac{||V^*||_\infty}{\lambda}}$ 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 pointwise upper bound of V^*. $\qquad\square$

COROLLARY 22. *Let $V_u^d = -\lambda \log \Psi_l^d$ and $V_u^{d+1} = -\lambda \log \Psi_l^{d+1}$. If $\Psi_u^d - \Psi^* \leq \epsilon^d$ and $V_u^{d+1} - V^* = \gamma^{d+1}$, then $\gamma^{d+1} \leq -\lambda \log \left(1 - \min\left\{1, \frac{\epsilon^d}{\eta}\right\}\right)$.*

Proof. This result is given by Corollary 20 and Theorem 21. $\qquad\square$

At this point, we have shown that the lower bound of the desirability function yields an upper bound of the suboptimal cost. More importantly, the upper bound of the suboptimal cost is not increasing as the degree of polynomial increases.

4.3. Approximate HJB solutions are SCLFs. This section shows that the approximate value function derived from the approximation, Ψ_l, is an SCLF.

THEOREM 23. *V_u is an SCLF according to Definition 3.*

Proof. The constraint $\Psi_l(0) = 1$ in (14) 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 16, $V^u \geq V^*$. Hence, V^u is also a positive definite function. The second and third to last constraints in (14) ensure that Ψ_l is nonincreasing away from the origin. Hence, V_u is nondecreasing away from the origin satisfying $\mu(|x|) \leq V_u(x) \leq \mu'(|x|)$ for some $\mu, \mu' \in \mathcal{K}$.

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

$$(16) \qquad u^\epsilon = -R^{-1}G^T \nabla_x V_u,$$

the control law corresponding to V_u. 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, $u^\epsilon = \frac{\lambda}{\Psi_l}R^{-1}G^T \nabla_x \Psi_l$.

Then, from (9),

$$L(V_u) = -\frac{\lambda}{\Psi_l}(\nabla_x \Psi_l)^T \left(f + \frac{\lambda}{\Psi_l}GR^{-1}G^T \nabla_x \Psi_l \right)$$
$$+ \frac{1}{2}Tr\left(\left(\frac{\lambda}{\Psi_l^2}(\nabla_x \Psi_l)(\nabla_x \Psi_l)^T - \frac{\lambda}{\Psi_l}\nabla_{xx}\Psi_l \right) B\Sigma_\epsilon B \right),$$

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

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

From the first constraint in (14),

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

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

$$(17) \qquad L(V_u) \leq -q - \frac{\lambda}{2\Psi_l^2}(\nabla_x \Psi_l)^T \Sigma_t \nabla_x \Psi_l \leq 0,$$

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

COROLLARY 24. *The suboptimal controller $u^\epsilon = -R^{-1}G^T\nabla_x V_u$ is stabilizing in probability within the domain Ω.*

Proof. This corollary is a direct consequence of the constructive proof of Theorems 4 and 23. □

COROLLARY 25. *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 Theorems 5 and 23. In (17), $L(V_u) < 0$ for $x \in \Omega\backslash\{0\}$ if Σ_t is positive definite. Recall that q is positive definite in the problem formulation. □

4.4. 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 (16) can be bounded as follows.

THEOREM 26. *Given the control law $u^\epsilon = -R^{-1}G^T\nabla_x V_u$,*

$$(18) \qquad J_u \leq V_u \leq V^* - \lambda\log\left(1 - \min\left\{1, \frac{\epsilon}{\eta}\right\}\right),$$

where $J_u = \mathbb{E}_{\omega_t}[\phi_T(x_T) + \int_0^T q(x_t) + \frac{1}{2}u_t^T R u_t \ dt]$, the expected cost of the system when using the 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,

$$(19) \qquad V_u(x_t) = V_u(x_0, 0) + \int_0^t L(V_u)(x_s)ds + \int_0^t \nabla_x V_u(x_s)B(x_s)d\omega_s.$$

Given that V_u is derived from polynomial function Ψ_l, the integrals are well defined, and we can take the expectation of (19) to get

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

whereby the last term of (19) 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 (17),

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

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,

$$\mathbb{E}_{\omega_t}[V_u(x_T)] = V_u(x_0) + \mathbb{E}_{\omega_t}\left[\int_0^T L(V_u)(x_s)ds\right]$$

$$\leq V_u(x_0) - \mathbb{E}_{\omega_t}\left[\int_0^T q(x_s) + \frac{1}{2}(u_s^\epsilon)^T R u_s^\epsilon ds\right]$$

$$= V_u(x_0) - J(x_0, u^\epsilon) + \mathbb{E}_{\omega_t}[\phi(x_T)],$$

where the last equality is given by (2). 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 $V_u(x_0) \geq J(x_0, u^\epsilon)$. Theorem 21 gives the second inequality in the theorem. $\qquad\square$

5. Linearly solvable approximations. The approach presented in this paper would appear up to this point to be limited to systems that are linearly solvable, i.e., those that satisfy condition (6). However, the proposed methods may be extended to a system which does not satisfy these conditions by approximating the system with one that is linearly solvable. One example is to introduce stochastic forcing into an otherwise deterministic system.

We first construct a comparison theorem between HJB solutions and systems that share the same general dynamics but with differing noise covariance. This comparison allows for the approximated value function of one system to bound the value function for another, providing pointwise bounds, and indeed SCLFs, for those that do not satisfy (6).

PROPOSITION 27. *Suppose V^{a^*} is the solution to the HJB equation* (4) *with noise covariances Σ_a, and that V^b is a supersolution to* (4) *with identical parameters except the noise covariance Σ^b, where $\Sigma_b - \Sigma_a \succeq 0$. Then $V^b \geq V^{a^*}$ for all $x \in \Omega$.*

Proof. From [3, Def. 2.2], V is a viscosity supersolution to the HJB equation (4) with noise covariance Σ if it satisfies

$$(20) \quad 0 \leq -q - (\nabla_x V)^T f + \frac{1}{2}(\nabla_x V)^T GR^{-1}G^T(\nabla_x V) - \frac{1}{2}Tr\left((\nabla_{xx}V)B\Sigma B^T\right).$$

Since $\Sigma_b - \Sigma_a \succeq 0$, the following trace inequality holds:

$$Tr\left((\nabla_{xx}V^a)B\Sigma_b B^T\right) \geq Tr\left((\nabla_{xx}V^a)B\Sigma_a B^T\right).$$

Therefore, we have the inequality

$$0 \leq -q - (\nabla_x V^b)^T f + \frac{1}{2}(\nabla_x V^b)^T GR^{-1}G^T(\nabla_x V^b) - \frac{1}{2}Tr\left((\nabla_{xx}V^b)B\Sigma_b B^T\right)$$

$$\leq -q - (\nabla_x V^b)^T f + \frac{1}{2}(\nabla_x V^b)^T GR^{-1}G^T(\nabla_x V^b) - \frac{1}{2}Tr\left((\nabla_{xx}V^b)B\Sigma_a B^T\right),$$

which implies that V^b is in fact a viscosity supersolution to the system with noise covariance Σ^a (i.e., V^b satisfies (20) for Σ^a). As V^b is a supersolution to the system with parameter Σ^a, then $V^b \geq V^{a^*}$. $\qquad\square$

A particular class of such approximations arises from a deterministic HJB solution, which is not linearly solvable but is approximated by one that is linearly solvable. Consider a deterministic system of the form

$$(21) \qquad dx_t = (f(x_t) + G(x_t)u_t)\, dt$$

with cost function

$$(22) \qquad J(x, u) = \phi(x_T) + \int_0^T q(x_t) + \frac{1}{2}u_t R u_t\, dt,$$

where ϕ, q, R, f, G, and the state and input domains are defined as in the stochastic problem in subsection 2.2. Then, the HJB equation is given by

$$(23) \qquad 0 = q + (\nabla_x V)^T f - \frac{1}{2}(\nabla_x V)^T G R^{-1} G^T (\nabla_x V),$$

and the optimal control is given by $u^* = -R^{-1}G^T \nabla_x V$. In general, (23) is not a linear PDE.

COROLLARY 28. *Let V^* be the value function that solves* (23), *and V^u be the upper bound solution obtained from* (14) *where all parameters are the same as* (23) *and Σ_t is not zero. Then, V^u is an upper bound for V^* over the domain (i.e., $V^* \leq V^u$).*

Proof. A simple application of Proposition 27, where Σ_a takes the form of a zero matrix, gives $V^* \leq V^u$. □

Interestingly, using the solution from (14) and the transformation $V_u = -\lambda \log \Psi_l$, the suboptimal controller $u^\epsilon = -R^{-1}G^T \nabla_x V_u$ is a stabilizing controller for the deterministic system (21) if a simple condition is satisfied. This fact is shown using the Lyapunov theorem for deterministic systems introduced next [33].

DEFINITION 29. *Given the system* (21) *and cost function* (22), *a control Lyapunov function (CLF) is a proper positive definite function $\mathcal{V} \in \mathcal{C}^1$ on a compact domain $\Omega \cup \{0\}$ such that*

$$(24) \qquad \mathcal{V}(0) = 0, \ \mathcal{V}(x) \geq \mu(|x|) \quad \forall\, x \in \Omega \backslash \{0\},$$
$$\exists\, u(x) \ s.t. \ (\nabla_x \mathcal{V})^T(f + Gu) \leq 0 \quad \forall\, x \in \Omega \backslash \{0\},$$

where $\mu \in \mathcal{K}$.

THEOREM 30 (Theorem 2.5 of [33]). *Given a system* (21) *and cost function* (22), *if there exist a CLF V and a u satisfying Definition* 29, *then the controlled system is stable, and u is a stabilizing controller. Furthermore, if $(\nabla_x V)^T(f + Gu) < 0$ for all $x \in \Omega \backslash \{0\}$, the controlled system is asymptotically stable, and u is an asymptotically stabilizing controller.*

Verifying that the controller $u^\epsilon = -R^{-1}G^T \nabla_x V_u$ is in fact stabilizing and that V_u is a CLF may be approached as follows.

COROLLARY 31. *Given the controller $u^\epsilon = -R^{-1}G^T \nabla_x V_u$, if*

$$Tr\left((\nabla_{xx} V_u)\,\Sigma_t\right) \geq 0 \quad \forall\, x \in \Omega \backslash \{0\},$$

then u^ϵ is a stabilizing controller for (21). *If*

$$Tr\left((\nabla_{xx} V_u)\,\Sigma_t\right) > 0 \quad \forall\, x \in \Omega \backslash \{0\},$$

then u^ϵ is an asymptotically stabilizing controller for (21).

Proof. Recall that, from the proof of Theorem 23, all conditions in Definition 29 are satisfied by V_u except (24). To show that V_u satisfies (24), rearrange (4) to yield the following:

$$(\nabla_x V_u)^T (f + Gu^\epsilon) = (\nabla_x V_u)^T f - (\nabla_x V_u)^T GR^{-1}G^T(\nabla_x V_u)$$
$$\leq -q - \frac{1}{2}(\nabla_x V_u)^T GR^{-1}G^T(\nabla_x V_u) - \frac{1}{2}Tr\left((\nabla_{xx}V_u)\,\Sigma_t\right),$$

where $\Sigma_t = B\Sigma_\epsilon B^T$. Recall that q and R are positive definite. If $Tr\left((\nabla_{xx}V_u)\,\Sigma_t\right) \geq 0$ for all $x \in \Omega\backslash\{0\}$, then $(\nabla_x V_u)^T (f + Gu^\epsilon) \leq 0$, implying that V^u is a CLF and u^ϵ is a stabilizing controller by Theorem 30. Furthermore, if $Tr\left((\nabla_{xx}V_u)\,\Sigma_t\right) > 0$ for all $x \in \Omega\backslash\{0\}$, u^ϵ is an asymptotically stabilizing controller. □

In the deterministic case, Σ_t is free variable that can be chosen to be small according to the equality (6). Hence, (6) is no longer a constraint or an assumption, but it serves as a design principle for obtaining a CLF for system (21). Furthermore, given a Σ_t, the trace condition in Corollary 31 is easily enforced in (14) by adding one extra constraint in the optimization problem. Thus, the optimization problem (14) can also produce a CLF for the corresponding deterministic system, with analytical results from section 4, including a priori trajectory suboptimality bounds (Theorem 26), which are inherited as well.

6. Numeric examples. This section studies the computational characteristics of this method using two examples—a scalar system and a two-dimensional system. In the following problems, the optimization parser YALMIP [25] was used in conjunction with the semidefinite optimization package MOSEK [2]. In both examples, the continuous system is integrated numerically using Euler integration with a step size of 0.005 s during simulations.

6.1. Scalar unstable system. Consider the following scalar unstable nonlinear system:

$$(25) \qquad dx = \left(x^3 + 5x^2 + x + u\right)dt + d\omega$$

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. 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 $R = 1$. In the one-dimensional case, the origin, which is a boundary, divides the domain into two partitions, $x \leq 0$ and $x \geq 0$. Because of the natural division of the domain, the solutions for both domains can be represented by smooth polynomials and solved independently. The simulation is terminated when the trajectories enter the interval $[-0.005, 0.005]$ centered on the origin.

The desirability functions that result from solving (14) for varying polynomial degrees are shown in Figure 1. The true solution is computed by solving the HJB directly in Mathematica [39]. The kink at the origin is expected because the HJB PDE solution is not necessarily smooth at the boundary, and in this instance the origin is a zero-cost boundary.

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. Polynomial degrees below 14 are not feasible; hence this data is omitted from the plots. The suboptimal solution converges faster for $x > 0$
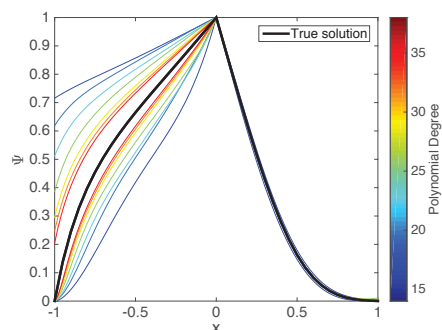
FIG. 1. *The desirability function of system* (25) *for varying polynomial degree. The true solution is the black curve.*
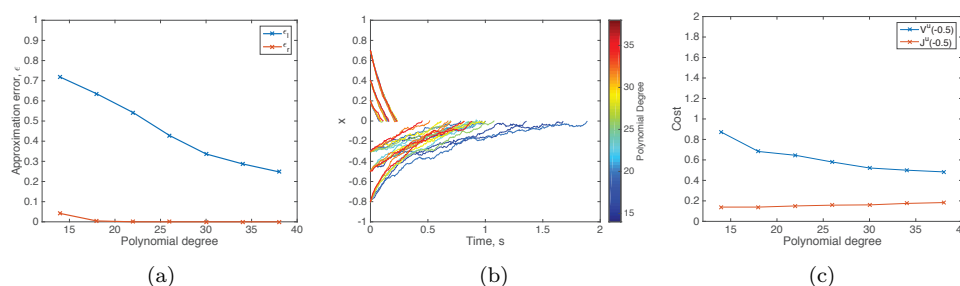

FIG. 2. *Computational results of system* (25). (a) *Convergence of the objective function of* (14) *as the degree of the polynomial increases. The approximation error for* $x \leq 0$ *is denoted by* ϵ_l, *and the approximation error for* $x \geq 0$ *is denoted by* ϵ_r. (b) *Sample trajectories using a controller computed from optimization problem* (14) *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* (14). *The initial condition is fixed at* $x_0 = -0.5$.

than for $x < 0$ when the degree of the polynomial increases, because the true solution for $x > 0$ has a simple quadratic-like shape that can be easily represented as a low degree SOS function.

Figure 2(b) shows sample trajectories using the controller computed from optimization problem (14) for different polynomial degrees. The controllers are stabilizing for six randomly chosen initial points. Unsurprisingly, the suboptimal solutions with low pointwise error result in the system converging more quickly towards the origin.

To compare J_u and V_u, a Monte Carlo experiment is illustrated in Figure 2(c). For each polynomial degree that is feasible, the controller obtained from Ψ_l in optimization problem (14) is implemented in 30 simulations of the system subject to random samples of Gaussian white noise with $\Sigma_\epsilon = 1$. The initial condition is fixed at $x_0 = -0.5$. In the figure, $V^u \geq J^u$ as expected, and the difference between the two decreases with increasing d.

6.2. Two-dimensional system. In the following example, we demonstrate the power of this technique on a two-dimensional system. Consider a nonlinear two-dimensional problem example with the following dynamics:

$$(26) \qquad \begin{bmatrix} dx \\ dy \end{bmatrix} = \left(2 \begin{bmatrix} x^5 - x^3 - x + xy^4 \\ y^5 - y^3 - y + yx^4 \end{bmatrix} + \begin{bmatrix} x \, u_1 \\ y \, u_2 \end{bmatrix} \right) dt + \begin{bmatrix} x \, d\omega_1 \\ y \, d\omega_2 \end{bmatrix}.$$

(a) Ψ, Degree $= 10$.   (b) Ψ, Degree $= 20$.   (c) V, Degree $= 10$.   (d) V, Degree $= 20$.
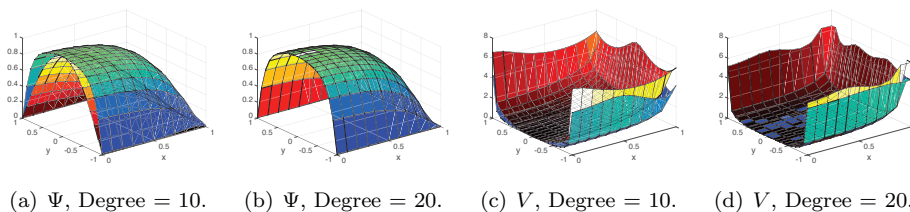
FIG. 3. *Approximated desirability functions and value functions for* (26) *when polynomial degrees are* 10 *and* 20. *In* (a) *and* (b), *the blue sheets are the upper bound solutions* Ψ_u, *and the red sheets are the lower bound solutions* Ψ_l. *The corresponding value functions are shown in* (c) *and* (d), *respectively.*

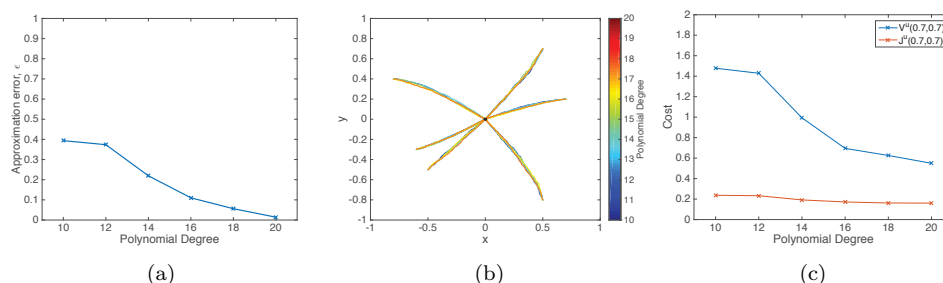(a)   (b)   (c)

FIG. 4. *Computational results of system* (26). (a) *Convergence of the variables in the objective function of* (14). (b) *Sample trajectories using the controller from optimization problem* (14) *with different polynomial degrees starting from six randomly chosen initial points.* (c) *The comparison between* J_u, *the expected cost, and* V_u, *the value function, for different polynomial degrees from optimization problem* (14). *The initial condition is fixed at* $x_0 = (0.7, 0.7)$.

The goal is to reach the origin at the boundary of the domain $\Omega = \{(x, y) \mid -1 \leq x \leq 1, -1 \leq y \leq 1\}$. The control penalty is $R = I_{2\times 2}$, and the state cost is $q(x) = x^2 + y^2$. The boundary conditions for the sides at $x = 1$, $x = -1$, $y = 1$, and $y = -1$ are set to $\phi(x, y) = 5$, while at the origin, the boundary has cost $\phi(0, 0) = 0$. The noise model considered is Gaussian white noise with zero mean and an identity covariance matrix.

The approximated desirability functions and their corresponding value functions are shown in Figure 3, with half of the domain $x \in [0, 1]$ shown in order to view the gaps between the upper and lower bound solutions. Figure 4(a) shows the convergence of the objective function of optimization problem (14) as the degree of the polynomial increases. There is no data below degree 10 because the optimization problem is not feasible in these cases. As shown in Figure 4(b), sample trajectories starting from six different initial points show that the controllers computed from Ψ_l for various degrees arrive at the origin. The trajectory is considered at the origin if it is within a distance of 0.01 from the origin.

Similar to the scalar example, a Monte Carlo experiment is performed to compare J_u and V_u. For each polynomial degree that is feasible, the controller obtained from Ψ_l in optimization problem (14) is implemented in 30 simulations of the system, subject to random samples of Gaussian white noise with $\Sigma_\epsilon = I_{2\times 2}$. The initial condition is fixed at $x_0 = (0.7, 0.7)$. Figure 4(c) shows 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 Ψ_l in optimization problem (14). As expected, $V^u \geq J^u$.

7. Conclusion. This paper has proposed a new method for approximating the solution to a class of optimal control problems for stochastic nonlinear systems via SOS 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 for a large class of stochastic nonlinear dynamical systems.

As is commonly seen when using SOS programming, the numerics of the semidefinite program may be cumbersome in practice. There are a number of avenues for future work aimed at improving the practical performance. First, the monomials of the polynomial approximation can be chosen strategically in order to decrease computation time while achieving high accuracy. A promising future direction is the synthesis of the work presented here with that of [14], wherein the curse of dimensionality is avoided via the strategic choice of basis functions. To improve the numerical conditioning of these optimization techniques, a domain partitioning technique is studied in [15], wherein the alternating direction method of multipliers is used to enable both parallelization and a solution representation that varies in resolution over the domain. In addition, there exists a growing body of literature devoted to increasing the numeric stability and scalability of SOS techniques [29, 1].

REFERENCES

[1] A. A. AHMADI AND P. PARRILO, *Towards scalable algorithms with formal guarantees for Lyapunov analysis of control systems via algebraic optimization*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2014, pp. 2272–2281.

[2] 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, New York, 2000, pp. 197–232.

[3] M. G. CRANDALL, H. ISHII, AND P.-L. LIONS, *Users guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., 27 (1992), pp. 1–67.

[4] P. DAI PRA, *A stochastic control approach to reciprocal diffusion processes*, Appl. Math. Optim., 23 (1991), pp. 313–329.

[5] H. DENG AND M. KRSTIĆ, *Stochastic nonlinear stabilization* I: *A backstepping design*, Systems Control Lett., 32 (1997), pp. 143–150.

[6] K. DVIJOTHAM AND E. TODOROV, *Linearly-solvable optimal control*, in Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, Wiley, New York, 2012, pp. 119–141.

[7] R. FILLIGER AND M.-O. HONGLER, *Relative entropy and efficiency measure for diffusion-mediated transport processes*, J. Phys. A, 38 (2005), p. 1247.

[8] W. H. FLEMING, *Logarithmic Transformations and Stochastic Control*, Springer, New York, 1982.

[9] W. H. FLEMING AND H. M. SONER, *Controlled Markov Processes and Viscosity Solutions*, Stoch. Model. Appl. Probab. 25, Springer, New York, 2006.

[10] P. FLORCHINGER, *Feedback stabilization of affine in the control stochastic differential systems by the control Lyapunov function method*, SIAM J. Control Optim., 35 (1997), pp. 500–511, https://doi.org/10.1137/S0363012995279961.

[11] R. A. FREEMAN AND J. A. PRIMBS, *Control Lyapunov functions: New ideas from an old source*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 1996, pp. 3926–3931.

[12] C. J. HOLLAND, *A new energy characterization of the smallest eigenvalue of the Schrödinger equation*, Comm. Pure Appl. Math., 30 (1977), pp. 755–765.

[13] M. HOROWITZ AND J. BURDICK, *Semidefinite relaxations for stochastic optimal control policies*, in Proceedings of the American Control Conference (ACC), 2014, pp. 3006–3012, https://doi.org/10.1109/ACC.2014.6859382.

[14] M. B. HOROWITZ, A. DAMLE, AND J. W. BURDICK, *Linear Hamilton-Jacobi-Bellman equations in high dimensions*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2014, pp. 5880–5887.

[15] M. B. Horowitz, I. Papusha, and J. W. Burdick, *Domain decomposition for stochastic optimal control*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2014, pp. 1866–1873.

[16] H. Kappen, *Linear theory for control of nonlinear stochastic systems*, Phys. Rev. Lett., 95 (2005), pp. 200201–200204.

[17] H. J. Kappen, *Path integrals and symmetry breaking for optimal control theory*, J. Statist. Mech., 2005 (2005), 11011.

[18] R. Khasminskii, *Stochastic Stability of Differential Equations*, Springer Science & Business Media, New York, 2011.

[19] S. Kolathaya and A. D. Ames, *Exponential convergence of a unified CLF controller for robotic systems under parameter uncertainty*, in Proceedings of the American Control Conference (ACC), 2014, pp. 3710–3715.

[20] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*, Wiley, New York, 1995.

[21] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817, https://doi.org/10.1137/S1052623400366802.

[22] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, *Nonlinear optimal control via occupation measures and LMI-relaxations*, SIAM J. Control Optim., 47 (2008), pp. 1643–1666, https://doi.org/10.1137/070685051.

[23] E. Lavretsky and K. Wise, *Robust and Adaptive Control: With Aerospace Applications*, Springer Science & Business Media, New York, 2012.

[24] Y. P. Leong, M. Horowitz, and J. Burdick, *Suboptimal stabilizing controllers for linearly solvable system*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2015, pp. 7157–7164.

[25] 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.

[26] P. Ogren, M. Egerstedt, and X. Hu, *A control Lyapunov function approach to multi-agent coordination*, in Proceedings of the 40th IEEE Conference on Decision and Control, 2001, pp. 1150–1155.

[27] A. Papachristodoulou and S. Prajna, *Analysis of non-polynomial systems using the sum of squares decomposition*, in Positive Polynomials in Control, Springer, New York, 2005, pp. 23–43.

[28] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.

[29] F. Permenter and P. Parrilo, *Basis selection for SOS programs via facial reduction and polyhedral approximations*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2014, pp. 6615–6620.

[30] J. A. Primbs, V. Nevistić, and J. C. Doyle, *Nonlinear optimal control: A control Lyapunov function and receding horizon perspective*, Asian J. Control, 1 (1999), pp. 14–24.

[31] W. Rudin, *Principles of Mathematical Analysis*, Internat. Ser. Pure Appl. Math., McGraw–Hill, New York, 1964.

[32] P. Rutquist, T. Wik, and C. Breitholtz, *Solving the Hamilton-Jacobi-Bellman equation for a stochastic system with state constraints*, in Proceedings of the IEEE International Conference on Decision and Control (CDC), 2014, pp. 1840–1845.

[33] E. D. Sontag, *A Lyapunov-like characterization of asymptotic controllability*, SIAM J. Control Optim., 21 (1983), pp. 462–471, https://doi.org/10.1137/0321028.

[34] F. Stulp, E. A. Theodorou, and S. Schaal, *Reinforcement learning with sequences of motion primitives for robust manipulation*, IEEE Trans. Robotics, 28 (2012), pp. 1360–1370.

[35] E. Theodorou, *Iterative Path Integral Stochastic Optimal Control: Theory and Applications to Motor Control*, Ph.D. thesis, Computer Science, University of Southern California, Los Angeles, 2011.

[36] E. Theodorou, F. Stulp, J. Buchli, and S. Schaal, *An iterative path integral stochastic optimal control approach for learning robotic tasks*, in Proceedings of the 18th IFAC World Congress, 2011, pp. 11594–11601.

[37] E. Todorov, *Efficient computation of optimal actions*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 11478–11483.

[38] W. Wiegerinck and B. Broek, *Stochastic optimal control in continuous spacetime multi-agent systems*, in Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, 2006, pp. 528–538.

[39] Wolfram Research, *Mathematica 10.0*, Wolfram Research, Champaign, IL, 2014.