2020 American Control Conference
Denver, CO, USA, July 1-3, 2020

Lyapunov-Like Conditions for Tight Exit Probability Bounds through Comparison Theorems for SDEs

Petter Nilsson and Aaron D. Ames

***Abstract*—Computing upper bounds on *exit probabilities*—the probability that a system reaches certain "bad" sets—may assist decision-making in control of stochastic systems. Existing analytical bounds for systems described by stochastic differential equations are quite loose, especially for low-probability events, which limits their applicability in practical situations. In this paper we analyze why existing bounds are loose, and conclude that it is a fundamental issue with the underlying techniques based on martingale inequalities. As an alternative, we give comparison results for stochastic differential equations that via a Lyapunov-like function allow exit probabilities of an n-dimensional system to be upper-bounded by an exit probability of a one-dimensional Ornstein-Uhlenbeck process. Even though no closed-form expression is known for the latter, it depends on three or four parameters and can be a priori tabulated for applications. We extend these ideas to the controlled setting and state a stochastic analogue of control barrier functions. The bounds are illustrated on numerical examples and are shown to be much tighter than those based on martingale inequalities.**

I. Introduction

Lyapunov techniques have arguably been among the most influential tools in the history of control theory. Classic Lyapunov results allow us to infer stability of a point from a Lyapunov function that satisfies certain properties, and many modifications have been proposed to infer other properties such as invariance or stability of a set. In this paper we refer to such methods as *certificate*-driven, since knowledge of a Lyapunov-like function is a certificate for the satisfaction of some property. While traditional Lyapunov techniques generalize to systems that are subject to disturbance, they are rooted in the robust tradition of only giving a binary yes/no answer to questions of stability or invariance. However, many components of modern systems are by nature stochastic, and it is oftentimes desirable to make decisions that involve trade-offs between multiple goals. In these cases a binary answer to the question of whether a property holds under certain conditions is less informative than knowing its probability.

While stochastic models can be approximated with robust counterparts that capture a large proportion of the possible behaviors, a property may be impossible to prove for a robust model due to simultaneous inclusion of behaviors that are very unlikely. If a robust model that includes 95% of all behaviors satisfies a property, then the property is satisfied with at least 95% probability, but it is difficult to build such a model without simultaneously including rare worst-case behaviors that may render the property false.

The authors are with the Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA, 91125 USA. {pettni,ames}@caltech.edu.

To exemplify, consider a system $\dot{x} = f(x, u, d)$, and the question of how to make assumptions on the disturbance signal d. The disturbance may represent model shortcomings such as unmodeled fast control loops for actuators, sensing noise injected into the aggregate system, or external disturbances that fluctuate around zero. A common property among these signals is that their expected long-term average is zero, but the signals may deviate from zero for short periods of time. It can be assumed that *adverse* signals with long (in time) and large (in magnitude) deviations are comparatively less likely than *benign* signals with short and/or small deviations. However, it becomes difficult to distinguish between these two types of signals in a robust model. The most common type of robust assumption is on the form $d \in D$ where D is a compact set containing zero, which takes into account spatial but not temporal characteristics of the signal. In order to include the benign signals in such a model, D must include typical extreme values, but there is no way of doing so without simultaneously including the adverse signals, and any distinction between the two classes disappears. Thus, from the robust model viewpoint there is no difference between a controller that can handle the benign signals but not the adverse ones, and a controller that can handle neither.

The shortcoming described above is partially mitigated by considering more sophisticated robust models such as integral constraints that account for both temporal and spatial characteristics, but the yes/no dichotomy remains an issue since there is no differentiation between the likelihood of different signals within the model: relative probabilities of the behaviors are masked. It is likely impossible to build meaningful autonomous systems with a 0% failure rate, which poses a problem for verification in the robust framework: *under this assumption successful robust verification is by definition only possible with respect to models that do not capture all possible behaviors.*

In stochastic models these issues are mitigated since the noise models quantify the likelihood of different behaviors. For a stable stochastic differential equation adverse signals (as described above) are significantly less likely than benign signals, although both are accounted for. Verification in a stochastic framework could additionally serve as a tool to explore the design space, in particular how different assumptions regarding uncertain components affect the probability of various properties being satisfied. While verification of finite stochastic systems via model checking has received considerable attention [1], [2], there has been less work on continuous certificate-driven verification. Kushner [3] developed a theory

978-1-5386-8266-1/$31.00 ©2020 AACC

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

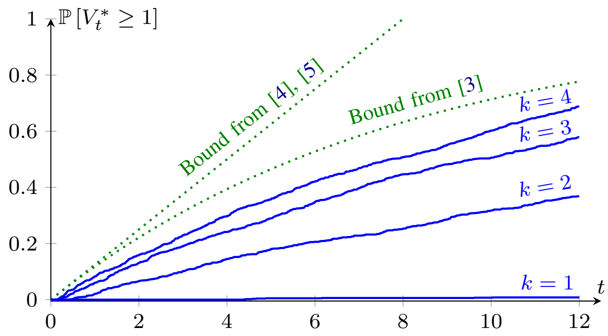

Fig. 1. Comparison of Euler–Maruyama Monte Carlo estimates ($N = 1,000$) of the probability $\mathbb{P}[V_t^* \geq 1]$ for the process $\mathrm{d}V_t = \left(-2\mu V_t + \frac{\sigma^2}{2}\right)\mathrm{d}t + k\sigma\sqrt{2V_t}\mathrm{d}W_t$ for $k = 1, 2, 3, 4$ and $\sigma = 0.5$, $\mu = 1$, with analytical bounds from the literature. Since the analytical bounds hold for all values of k the bounds become loose for the case of interest $k = 1$.

of stochastic stability, later applied to verification by Prajna, Jadbabaie, and Pappas [4] and by Steinhardt and Tedrake [5]. These works all rely on similar (super)martingale arguments to infer bounds on the probability that properties are violated, but the martingale property is only concerned with expectation and not with variance. A similar problem as described above therefore arises: just like robust models fail to distinguish between benign and adverse signals, a martingale viewpoint does not distinguish between martingales with different variance magnitudes. Thus, such bounds produce the same values for all martingales, irrespective of noise magnitude. As the following example shows, this often leads to loose bounds for processes with low variance.

Example 1. *Consider the one-dimensional stochastic system*

$$\mathrm{d}X_t = -\mu X_t \mathrm{d}t + \sigma \mathrm{d}W_t \tag{1}$$

with stable deterministic dynamics, and where W_t is a Wiener process. The time evolution of a Lyapunov candidate $V_t = V(X_t) = X_t^2/2$ is

$$\mathrm{d}V_t = \left(-2\mu V_t + \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\sqrt{2V_t}\mathrm{d}W_t. \tag{2}$$

Suppose we are interested in bounding the probability that V_t passes above a level λ on the interval $[0, T]$, i.e. upper-bound

$$\mathbb{P}\left[\sup_{0\leq s\leq t} V_s \geq \lambda\right]. \tag{3}$$

Existing results in the literature (summarized in Section II-C below) upper-bound this probability via martingale arguments based on inequalities for the deterministic part $(-2\mu V_t + \sigma^2/2)\mathrm{d}t$, but disregard the stochastic part $\sigma\sqrt{2V_t}\mathrm{d}W_t$. As a result, these bounds hold for any process that has the same deterministic dynamics. But, as Fig. 1 shows, the probability (3) is heavily dependent on the stochastic part of the dynamics and disregarding it results in loose bounds when it is small.

Based on these observations it seems necessary to move beyond the martingale approach to obtain useful probability estimates, especially for low-probability phenomena. However, it is difficult to calculate the density of the running maximum for all but the most basic stochastic processes—the problem can be phrased as solving a parabolic partial differential equation for which analytical solutions are unknown.

In this paper we instead propose to express the bounds as canonical quantities that represent exit probabilities for one-dimensional processes. Our first result considers the Ornstein-Uhlenbeck process $\mathrm{d}X_t = -\alpha X_t \mathrm{d}t + \sigma \mathrm{d}W_t$, which is the canonical "stable" stochastic process, in the sense that it converges to a stationary probability distribution. The results we present are parameterized by three resp. four quantities, so although no explicit formula is available, it is feasible to simply tabulate solutions for relevant parameter ranges. To this end, we show that the property exhibited in Fig. 1—that larger variance leads to higher probabilities—holds in quite general settings. It may seem intuitive that more noise leads to a higher probability of reaching a given level, but as we show with a counterexample the reality is slightly more suble. One may view these results as a stochastic analogue of the comparison theorems that are typically employed in traditional Lyapunov proofs: in the deterministic setting comparison theorems are used to show that if $\mathrm{d}V \leq -\alpha V \mathrm{d}t$, then the system $\mathrm{d}x = -\alpha x \mathrm{d}t$ serves as an upper bound, which is the limit in $\sigma \to 0$ of an Ornstein-Uhlenbeck process. As we illustrate with examples, this approach sharpens bounds in the stochastic case, and generalizes robust methods in the sense that when the noise magnitude goes to zero we retrieve known results from the robust paradigm.

In the following section we review existing results from the robust and stochastic settings. We then present our main results in Section III, show two numerical examples in Section IV, and conclude the paper in Section V.

II. Background

Consider a stochastic differential equation (SDE)

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{4}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times p}$, and W_t is p-dimensional vector of independent Wiener processes. We let

$$\mathbb{P}_x[\cdot] = \mathbb{P}[\cdot \mid X_0 = x] \tag{5}$$

denote the probability measure conditioned on an initial state. In this paper we are concerned with *exit probabilities*, which are probabilities of the form $\mathbb{P}_x[\exists t \in [0, T] : X_t \notin S]$ for some set S. We remark that exit probabilities are well defined even if the stochastic process is only defined in a neighborhood of S, since the process can be stopped on exiting S without affecting the exit probability.

The well-known Itô Lemma from stochastic calculus states that the time evolution of a function $V(t, X_t)$ is governed by

$$\mathrm{d}V(t, X_t) = \mathcal{L}V(t, X_t)\mathrm{d}t + (\nabla_x V)(t, X_t)\sigma(X_t)\mathrm{d}W_t, \tag{6}$$

where the infinitesimal generator $\mathcal{L}$ of the process is

$$\begin{aligned}\mathcal{L}V(t, x) = \frac{\partial V(t, x)}{\partial t} &+ (\nabla_x V)(t, x) f(x) \\ &+ \frac{1}{2}\left\langle \sigma(x), (\nabla_x^2 V)(t, x)\sigma(x)\right\rangle_F,\end{aligned} \tag{7}$$

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

where $\langle \cdot, \cdot \rangle_F$ denotes the Frobenius matrix inner product [6].

There is a multitude of results in the literature regarding implications from satisfaction of an inequality of the form

$$\mathcal{L}V(t, X_t) \leq -\alpha V(t, X_t) + c_t. \tag{8}$$

In Lyapunov theory for deterministic systems, satisfaction of this type of inequality with $c_t = 0$ is of course a necessary and sufficient condition for exponential stability, but that is not the only result of its kind. Below we survey existing results on input-to-state stability, deterministic barrier functions, as well as stochastic stability, which all derive certain system properties starting from the inequality (8).

A. Input-to-State Stability

Equation (8) is a special case of the storage function condition for *input-to-state stability*, which is equivalent to an upper bound on the magnitude of $V(x)$ [7, Theorem 3.4]. Input-to-state stability is a powerful theoretical concept, and there are strong converse results, but it can be cumbersome to utilize in practical situtations as the upper bounds are only implicitly given. Input-to-state stability has also been extended to hybrid stochastic systems [8], but only for inequalities in *expectation*, which are less informative than bounds on exit probabilities.

B. Deterministic Barrier Functions

In contrast to input-to-state stability, *barrier functions* are very practical in that the desired bounds are stated explicitly. The usual *barrier function condition* (e.g. [9]) is stated for a smooth function $h(t, x)$ as

$$\mathcal{L}h(t, x) + \alpha h(t, x) \geq 0 \tag{9}$$

for invariance of the (here assumed to be compact) set $\{t, x : h(t, x) \geq 0\}$. By defining the non-negative function $\tilde{h}(t, x) = \bar{h}(t) - h(t, x)$ for $\bar{h}(t) = \sup_x h(t, x)$, it follows that $\{t, x : h(t, x) \geq 0\} = \{t, x : \tilde{h}(t, x) \leq \bar{h}(t)\}$ and the barrier function condition (9) becomes

$$\mathcal{L}\tilde{h}(t, x) \leq -\alpha\tilde{h}(t, x) + \alpha\bar{h}(t) + \frac{\mathrm{d}\bar{h}(t)}{\mathrm{d}t}, \tag{10}$$

which is of the form (8).

C. Martingale Results for Stochastic Stability

We next summarize existing results in the stochastic setting, which are the bounds shown in Fig. 1. The sharpest bound in the stochastic setting that we are aware of is due to Kushner [3, Theorem III.1], which states that if (8) holds for a constant $c_t = c$ (Kusher considers the more general case of a deterministic time-dependent c_t, but for space reasons we present this less general result) and a continuous non-negative function $V(t, x)$, then

$$\begin{aligned}\mathbb{P}_x&\left[\sup_{0 \leq t \leq T} V(t, X_t) \geq \lambda\right] \\ &\leq \frac{V(0, x) + \left(e^{T\min(\alpha, c/\lambda)} - 1\right)\max(\lambda, c/\alpha)}{\lambda e^{T\min(\alpha, c/\lambda)}}.\end{aligned} \tag{11}$$

This can be proven by introducing an auxiliary function $W(t, x) = e^{\gamma ct}V(t, x) + (e^{\gamma cT} - e^{\gamma ct})/\gamma$ for an appropriate $\gamma < \alpha/c$, showing that it is a supermartingale, and applying Doob's martingale inequality [6, Theorem 3.2.4]. A weaker upper bound $(V(0, x) + cT)/\lambda$ can be obtained via the same method by considering the supermartingale $W(t, x) = V(t, x) + c(T - t)$, which yields the bound in [4] and [5].

While the stochastic dynamics of the state X_t are implicitly accounted for in (8) (the second-order term in the Itô formula affects whether (8) is satisfied, so more noise makes it more difficult to satisfy (8)), the stochastic part of (6) is disregarded. Therefore classical deterministic results are not retrieved when applying e.g. inequality (11) to a deterministic system.

III. Main Result

In this work we seek a unifying theory for the stochastic and deterministic settings. As opposed to the stochastic works surveyed above that are based solely on inequality (8), we also consider the stochastic dynamics of a certificate function V. This allows us to obtain exit probability bounds that are tighter, especially for low-probability events. Furthermore, as the stochastic effect goes to zero we retrieve the deterministic results discussed in Section II-B as a special case.

A. Comparison Theorem for Ornstein-Uhlenbeck-Like Processes

Our main results are based on the following theorem that gives conditions for when more noise in a 1-dimensional process is associated with a higher probability of the process reaching a certain level. Theorem 1 can be seen as a monotonicity result in the sense that more noise increases the probability of the supremum attaining a given level.

Theorem 1. *Let X_t^1 and X_t^2 be two one-dimensional Itô processes with $X_0^1 = X_0^2 = x$ and dynamics*

$$\mathrm{d}X_t^1 = -\theta X_t^1 \mathrm{d}t + Y_t \mathrm{d}W_t^1, \tag{12a}$$

$$\mathrm{d}X_t^2 = -\theta X_t^2 \mathrm{d}t + \sigma \mathrm{d}W_t^2. \tag{12b}$$

where W_t^1, W_t^2 are Wiener processes, and Y_t is X_t^1-measurable. If $|Y_t| \leq \sigma$ for all $t \in [0, T]$, then for $\lambda > 0$

$$\mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t^1 \geq \lambda\right] \leq \mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t^2 \geq \lambda\right]. \tag{13}$$

Proof. We show the result with an argument rooted in the theory of optimal control. Consider the problem of maximizing the probability that the set $\{x : x \geq \lambda\}$ is reached before time T for the system

$$\mathrm{d}X_t = -\theta X_t \mathrm{d}t + Y_t \mathrm{d}W_t, \tag{14}$$

where Y_t is considered a *control input* that is constrained to $[-\sigma, \sigma]$. The optimal control problem we are interested in is the stochastic reachability problem

$$\begin{aligned}&\sup_{Y_t}\ \mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t \geq \lambda\right], \\ &\text{s.t. } X_t \text{ follows (14)}, |Y_t| \leq \sigma.\end{aligned} \tag{15}$$

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

If we can show that the optimal control input is equal to σ the statement of the theorem follows.

Let $\Phi^Y(t,x)$ denote the value function of the problem for a fixed control input Y_t, i.e. the probability that the level set $X_t = \lambda$ is reached when starting at x at time t using input Y_t. From the Itô formula and by reversing time it can be shown that Φ^Y satisfies the Hamilton-Jacobi PDE

$$\frac{\partial \Phi^Y}{\partial t} = \frac{Y_t^2}{2}\frac{\partial^2 \Phi^Y}{\partial x^2} - \theta x \frac{\partial \Phi^Y}{\partial x} \tag{16}$$

on $(t,x) \in [0,T] \times [-\infty, \lambda]$ (e.g., [10]). The maximum principle gives that the optimal input Y^* satisfies

$$Y^*(t,x) \in \arg\sup_Y \frac{Y^2}{2}\frac{\partial^2 \Phi^*(t,x)}{\partial x^2}, \tag{17}$$

where Φ^* is the optimal value function. Thus, if we can show that (17) is satisfied for $Y(t) \equiv \sigma$ and the corresponding value function the result follows. To this end we study the PDE

$$\begin{cases} \frac{\partial v \Phi}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2 \Phi}{\partial x^2} - \theta x \frac{\partial \Phi}{\partial x} & (t,x) \in [0,T] \times [-\infty, \lambda], \\ \Phi(0,x) = 0 & x \in [-\infty, \lambda], \\ \Phi(t,\lambda) = 1 & t \in [0,T], \end{cases} \tag{18}$$

where we have added appropriate initial and boundary conditions for the reachability problem.

If we can show that $\frac{\partial^2 \Phi}{\partial x^2} > 0$ for all times then (17) holds and the result follows. To establish this we invoke a result on preservation of convexity by Janson and Tysk [11, Theorem 10.2] stating conditions under which the parabolic PDE

$$\begin{cases} \frac{\partial \phi}{\partial t} = a\frac{\partial^2 \phi}{\partial x^2} + b\frac{\partial \phi}{\partial x} + c, & (t,x) \in [0,T] \times [0,1], \\ \phi(0,x) = \phi_0(x), & x \in [0,1], \\ \phi(t,0) = \phi(t,1) = 0, & t \in [0,T], \end{cases} \tag{19}$$

preserves convexity in the sense that convexity of $\phi_0(x)$ implies convexity in x of $\phi(t,x)$. This holds given smoothness assumptions on a, b, c and the following conditions: i) $b(0,t) \geq 0, b(1,t) \leq 0$, ii) $2c + \frac{\partial b}{\partial x}$ is a function of only t, and iii) c is concave in x.

It is fairly straightforward to pose a smoothened version of (18) on this form: let $\tilde{\Phi} = \Phi - 1$ to homogenize the boundary condition. This affects the initial condition, but we can approximate the resulting initial condition $\mathbf{1}_{x \geq \lambda} - 1$ (convex on $[-\infty, \lambda]$) arbitrarily well with a smooth convex function[1]. Furthermore, let $x \leftarrow (x - \underline{x})/(\lambda - \underline{x})$ for an arbitrarily small lower bound $\underline{x}$ to rescale the spatial domain to $[0,1]$. Conditions ii) and iii) are trivially satisfied, and condition i) is satisfied if and only if $\lambda \geq 0$. Thus [11, Theorem 10.2] applies, so $\Phi(t,x)$ is convex in x for all t, hence $Y_t \equiv \sigma$ is optimal for (15), and the result follows. □

Remark 1. As can be seen in the proof, the condition $\lambda > 0$ is necessary for the proof strategy to work. Interestingly, numerically solving (18) for values $\lambda < 0$ reveals that the result indeed does not hold in this case. Fig. 2 illustrates that for the case $\lambda < 0$ the reachability probability is *not* monotonic

[1] $\mathbf{1}_{x \geq \lambda}$ denotes the indicator function equal to 1 on the set $\{x \geq \lambda\}$.

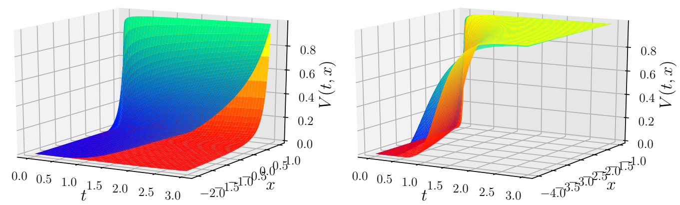

Fig. 2. Numerical solutions to (18) obtained with the FiPy [12] library for $\sigma = 1$ (blue/green) and $\sigma = 0.5$ (red/yellow). The left plot shows the solution for $\lambda = 1$ on the spatial domain $[-2, 1]$, and the right plot shows the solution for $\lambda = -1$ on the spatial domain $[-4, -1]$. For the case $\lambda = 1$ larger noise implies uniformly higher values of $V(t,x)$, corresponding to higher probability of reaching the set $\{x \geq 1\}$, but this is not true for the case $\lambda = -1$ where in some regions a larger noise level corresponds to lower reachability probability.

in the noise level: a smaller noise level is sometimes associated with a higher reachability probability. This phenomenon is best understood with a simple example: consider a system $\mathrm{d}X_t = -X_t \mathrm{d}t + \sigma \mathrm{d}W_t$ starting at $X_0 = -2$. For $\sigma = 0$ the probability of reaching the level -1 within time $\log 2 + \epsilon$ is equal to 1. However, for $|\sigma| > 0$ the probability is strictly smaller than 1. The distinction between the case $\lambda > 0$ and $\lambda < 0$ is that for $\lambda < 0$ the noise might counteract the drift dynamics, whereas this does not happen for $\lambda > 0$ where the drift dynamics flow "away" from λ. Exactly how the "optimal" noise is characterized in the case $\lambda < 0$ is an interesting question for future research.

The right-hand side exit probability in (13) depends on five parameters: x, θ, σ, T and λ. However, with the transformations $U_t = \frac{\sqrt{\theta}}{\sigma} X_t^2$ and $s = \theta t$ we can consolidate parameters and arrive at the following corollary which provides an upper bound in terms of a quantity that is a function of just three parameters.

Corollary 1. *Let X_t^1 follow (12a) and let U_t be a standard Ornstein-Uhlenbeck process ((12b) with $\theta = \sigma = 1$). Then*

$$\mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t^1 \geq \lambda\right] \leq \mathbb{P}_{x\frac{\sqrt{\theta}}{\sigma}}\left[\sup_{0 \leq s \leq \theta T} U_s \geq \lambda \frac{\sqrt{\theta}}{\sigma}\right]. \tag{20}$$

Furthermore, the same proof technique can be used to obtain a result with affine drift and more general diffusion terms.

Corollary 2. *Let X_t^1 and X_t^2 be two one-dimensional non-negative Itô processes with $X_0^1 = X_0^2 = x$ and dynamics*

$$\mathrm{d}X_t^1 = \theta(\mu - X_t^1)\mathrm{d}t + Y_t (X_t^1)^k \mathrm{d}W_t^1, \tag{21a}$$

$$\mathrm{d}X_t^2 = \theta(\mu - X_t^2)\mathrm{d}t + \sigma (X_t^2)^k \mathrm{d}W_t^2, \tag{21b}$$

where $\mu > 0$, $k > 0$ is a positive exponent, and W_t^1, W_t^2 are Wiener processes, and Y_t is X_t^1-measurable. If $|Y_t| \leq \sigma$ for all $t \in [0,T]$, then for $\lambda > 0$

$$\mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t^1 \geq \mu + \lambda\right] \leq \mathbb{P}_x\left[\sup_{0 \leq t \leq T} X_t^2 \geq \mu + \lambda\right]. \tag{22}$$

Proof. The proof of Theorem 1 applies when the analysis is confined to the spatial domain $[0, \lambda]$. □

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

B. Stochastic Stability

We next leverage the result in the previous subsection to obtain results on stochastic stability for multi-dimensional systems. Consider a system

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + e(X_t)\mathrm{d}W_t \tag{23}$$

for a multi-dimensional Wiener process W_t, and assume that there is a Lyapunov-like function $V(x)$ such that

$$\mathcal{L}V(x) \le -\alpha V(x) + c, \tag{24a}$$

$$\|\nabla V(x)e(x)\|_2 \le \sigma. \tag{24b}$$

That is, V satisfies the same stability condition discussed in Section II, but also a condition that bounds the influence of noise on its time evolution.

Theorem 2. *Consider the system (23) and a function V satisfying (24). Then it holds that*

$$\begin{aligned}&\mathbb{P}_x\left[\sup_{0\le t\le T} V(x_t) \ge \frac{c}{\alpha} + \lambda\right] \\ &\qquad \le \Omega\left(\frac{\sqrt{\alpha}}{\sigma}\left(V(x) - \frac{c}{\alpha}\right), \alpha T, \frac{\sqrt{\alpha}}{\sigma}\lambda\right),\end{aligned} \tag{25}$$

where

$$\Omega(u, T, \lambda) = \mathbb{P}_u\left[\sup_{0\le t\le T} U_t \ge \lambda\right] \tag{26}$$

is the λ-exit probability on $[0, T]$ for a standard one-dimensional Ornstein-Uhlenbeck process starting at $U_0 = u$.

Proof. By the Itô formula the time evolution of V satisfies

$$\begin{aligned}\mathrm{d}V(x_t) &= \mathcal{L}V(x_t)\mathrm{d}t + \nabla V(x_t)e(x_t)\mathrm{d}W_t \\ &\le (-\alpha V(x_t) + c)\,\mathrm{d}t + \nabla V(x_t)e(x_t)\mathrm{d}W_t.\end{aligned}$$

By a drift comparison theorem for stochastic processes [13] we know that $V(x_t) \le \tilde{V}_t$ for the stochastic process $\tilde{V}_t$ with $\tilde{V}_0 = V(x_0)$ and

$$\mathrm{d}\tilde{V}_t = -\alpha\left(\tilde{V}_t - \frac{c}{\alpha}\right)\mathrm{d}t + \nabla V(x_t)e(x_t)\mathrm{d}W_t.$$

The transformation $\hat{V} = \tilde{V} - \frac{c}{\alpha}$ and equation (24b) then give the result via Corollary 1. □

Computing or bounding $\Omega(u, T, \lambda)$ is a difficult problem that is subject to ongoing research [10]. There are results in the literature that give closed-form bounds on $\mathbb{E}\left[\sup_{0\le t\le T}|U_t|\right]$ [14], that can be combined with Markov's inequality for a bound on Ω, but the result is not very tight. But absent an analytic expression, Ω is a quantity that only depends on three parameters. We therefore argue that for practical purposes it is feasible to maintain a database of values of Ω and interpolate between them in order to evaluate desired probability bounds. These values can be obtained via Monte Carlo simulations, by solving the PDE as above for different values of λ, or through more sophisticated numerical techniques [10]. Lookup tables are not without precedent in control: tabulation is often done in model-predictive control to avoid solving optimization problems on-line, and is necessary in for example HJB level set methods that rely on numerical solutions of PDEs.

Our second main result is obtained by using Corollary 2 instead of Theorem 1, and is practical for the case of quadratic Lyapunov functions.

Theorem 3. *Consider the system (23) and a non-negative function V satisfying (24a) and*

$$\|\nabla V(x)e(x)\|_2 \le \sigma\sqrt{V(x)}, \tag{27}$$

Then it holds that

$$\mathbb{P}_x\left[\sup_{0\le t\le T} V(x_t) \ge \frac{c}{\alpha} + \lambda\right] \le \mathbb{P}_z\left[\sup_{0\le t\le \alpha T} Z_t \ge 1 + \frac{\lambda\alpha}{c}\right], \tag{28}$$

where Z_t is the process

$$\mathrm{d}Z_t = (1 - Z_t)\,\mathrm{d}t + \frac{\sigma\sqrt{Z_t}}{\sqrt{c}}\mathrm{d}W_t, \quad Z_0 = z = \frac{\alpha}{c}V(x_0).$$

As can be seen, the right-hand side exit probability in (28) depends on four lumped parameters: time horizon αT, level set limit $\lambda\alpha/c$, initial condition $\alpha V(x_0)/c$, and noise magnitude $\sigma/\sqrt{c}$.

Proof. We follow the same initial steps as in the proof of Theorem 2. Due to monotonicity and (27) we can consider the behavior of the system

$$\mathrm{d}\tilde{V}_t = \left(c - \alpha\tilde{V}_t\right)\mathrm{d}t + \sigma\sqrt{\tilde{V}_t}\mathrm{d}W_t, \quad \tilde{V}_0 = V(x_0)$$

as an upper bound. It holds that $\tilde{V}_t = cZ_t/\alpha$ for the system

$$\mathrm{d}Z_t = \alpha(1 - Z_t)\mathrm{d}t + \frac{\sigma\sqrt{\alpha}}{\sqrt{c}}\sqrt{Z_t}\mathrm{d}W_t.$$

Rescaling time $t \leftarrow t\alpha$ results in the statement. □

C. Controlled Stochastic Stability

We next generalize the results above to the controlled setting. Consider a controlled diffusion

$$\mathrm{d}X_t = (f(X_t) + g(X_t)u_t)\,\mathrm{d}t + e(X_t)\mathrm{d}W_t, \tag{29}$$

where $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ models the effect of a controlled input $u_t \in U \subset \mathbb{R}^m$ on the drift dynamics. For this system we define the operator

$$\mathcal{L}_g V(t, x) = \nabla_x V(t, x) g(x) \tag{30}$$

and let $\mathcal{L}V$ be the expression in (7) that depends on f but not on g. Then the following result holds which can be seen as a stochastic analogue of invariance results for control barrier functions [9][2]:

Theorem 4. *Consider the system (29) and a function $V : \mathbb{R}^n \to \mathbb{R}$ satisfying (24b), and assume that the set*

$$U_{CBF}(x) = \{u : \mathcal{L}V(x) + \mathcal{L}_g V(x)u \le -\alpha V(x) + c\} \tag{31}$$

[2]Results in [9] allow for α to be a $\mathcal{K}_\infty$ function rather than a scalar, which enables converse results to be obtained.

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

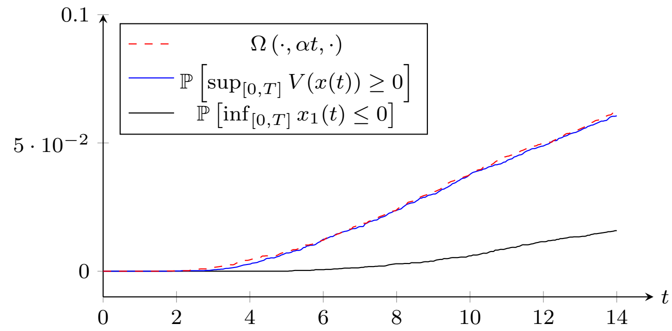

Fig. 3. Empirical exit probabilities for the example in Section IV-A. The dashed red curve is the theoretical upper bound, the blue solid curve is the *certificate* exit probability, and the solid blue line is the exit probability of x_1 becoming negative. Results in this paper imply that $\mathbb{P}\left[\inf_{[0,T]} x_1(t) \leq 0\right] \leq \mathbb{P}\left[\sup_{[0,T]} V(x(t)) \geq 0\right] \leq \Omega(\cdot, \alpha t, \cdot)$, and in this particular case the second inequality is in fact an equality.

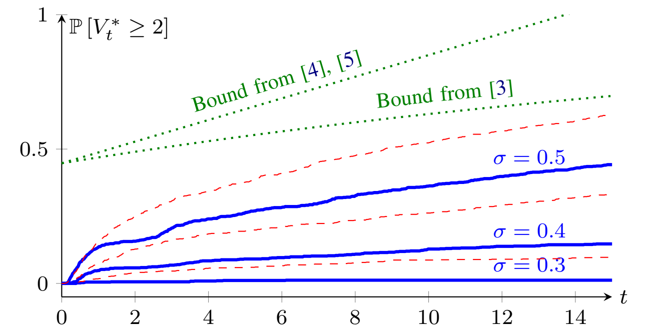

Fig. 4. Illustration of the empirical exit probabilities for different values of σ (solid blue lines), and the corresponding empirical bounds (dashed red). Also shown are the analytical bounds described in Section II that do not depend on σ (dotted green).

is non-empty for all x such that $V(x) < c/\alpha + \lambda$. Then, for any controller such that $u_t \in U_{CBF}(X_t)$ it holds that

$$\mathbb{P}_x\left[\sup_{0\leq t\leq T} V(X_t) \geq \frac{c}{\alpha} + \lambda\right] \leq \Omega\left(\frac{\sqrt{\alpha}}{\sigma}\left(V(x) - \frac{c}{\alpha}\right), \alpha T, \frac{\sqrt{\alpha}}{\sigma}\lambda\right). \quad (32)$$

In addition, an analogous result in the controlled setting under assumption (27) holds, resulting in the inequality (28), but for space reasons we do not state it here.

As $\sigma \to 0$ the Ornstein-Uhlenbeck process in (12b) becomes the deterministic system $\dot{x} = -\theta x$. Thus, in the noise-free limit we retrieve invariance results associated with control barrier functions.

IV. Examples

We show two numerical examples where the results are applied. In the first example Theorem 4 is applied to a linear system with a linear certificate. The second example utilizes Theorem 3 via a quadratic certificate. The numerical results have been obtained by simulating SDEs using the Matlab Financial Toolbox.

A. Linear System with Linear Certificate

The first example illustrates how a controller can be designed to enforce an upper bound on exit probabilities. Take the system

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= u dt + \sigma dW_t, \end{aligned} \quad (33)$$

and the function

$$V(x_1, x_2) = -x_1 - \beta x_2 \quad (34)$$

that is less than zero for $x_1 + \beta x_2 \geq 0$, i.e. for states where the risk of crossing the line $x_1 = 0$ is low. In particular, if $V(x) \leq 0$ along a trajectory, then $x_1 \geq 0$ everywhere on the trajectory.

We implement a controller u such that

$$\mathcal{L}V(x_1, x_2) + \mathcal{L}_g V(x_1, x_2) u = -\alpha V(x_1, x_2) + \alpha M, \quad (35)$$

which is always possible since $\mathcal{L}_g V(x_1, x_2) = -\beta \neq 0$. This construction is analogous to an established method for constructing control barrier functions for systems with relative degree greater than one [15]. We consider the probability of crossing the level set $V(x_1, x_2) = 0$ starting at an initial point (x_1^0, x_2^0) and utilizing the controller described above. We have

$$\|\nabla V \cdot E\|_2 = \beta\sigma, \quad (36)$$

so Theorem 4 applies, giving that the probability of crossing into $V(x_1, x_2) \geq 0$ within time T is upper bounded by

$$\Omega\left(\frac{\sqrt{\alpha}}{\sigma}\left(V(x_1^0, x_2^0) - M\right), \alpha T, -\frac{\sqrt{\alpha}}{\sigma}M\right). \quad (37)$$

As an example we consider the following parameters: $\alpha = 0.1, \beta = 1, \sigma = 0.2, M = -1$, and $(x_1^0, x_2^0) = (4, -3)$. Fig. 3 shows the probability (37), the probability of crossing the line $V(x_1, x_2) = 0$, and the probability of crossing the line $x_1 = 0$, all estimated from 10,000 Monte-Carlo simulations. In this particular case the upper bound is in fact achieved; equation (37) corresponds to the exact probability that V becomes positive. There is however a gap between this probability and the probability of reaching $x_1 < 0$, since it is possible to "recover" from $V > 0$ without crossing into the region $x_1 < 0$.

Note that the controller (35) is implemented as an *equality* to exactly achieve a certain deterministic convergence rate. A more traditional barrier function application would be to design what we might call an *exit probability filter*, or a *risk filter*, that enforces the *inequality* $\mathcal{L}V \leq -\alpha V + \alpha M$ by potentially overriding a nominal controller [9].

This example is somewhat atypical in that (35) holds on a non-bounded set. On a compact set it would not be possible to have it hold for a negative M. From a practical viewpoint the example can be thought of as the local dynamics of a finite-time avoidance maneuver where the risk of a collision is to be estimated.

B. Linear System with Quadratic Certificate

We next apply the results to linear systems coupled with quadratic Luapunov functions. Consider the system

$$dX_t = (AX_t + Bu_t)\,dt + E dW_t \quad (38)$$

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.

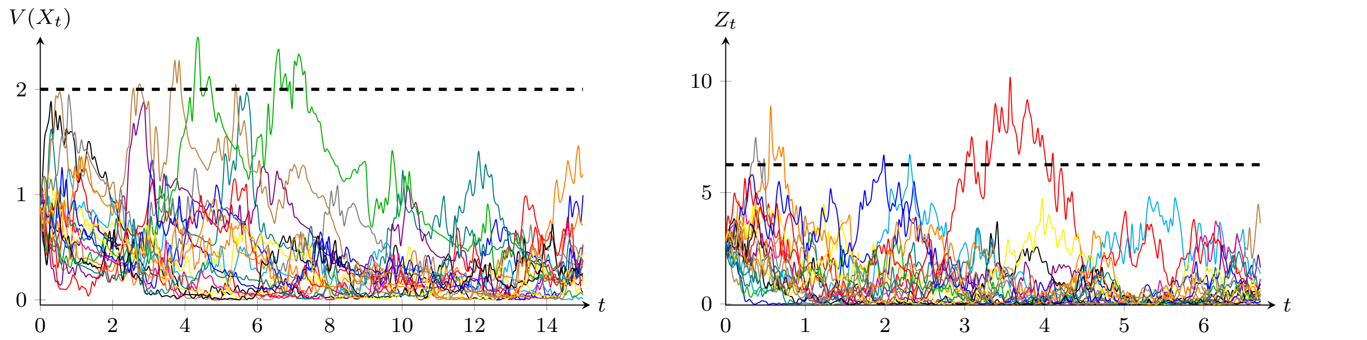

Fig. 5. Sample trajectories for the case $\sigma = 0.4$. Left: sample trajectories of $V(X_t)$ over the horizon $[0, 15]$ and the bound $\lambda = 2$. Right: trajectories of Z_t over the horizon $[0, 15\alpha]$ and the bound $\alpha\lambda/c = 6.25$.

where A, B, E are matrices of appropriate dimensions. For a Lyapunov candidate $V(x) = x^T P x$ with $P \succ 0$ we have that if P solves the algebraic Ricatti equation

$$A^T P + PA - PBB^T P = -Q, \tag{39}$$

then the control input $u_t = -\frac{1}{2} B^T P X_t$ yields

$$\mathcal{L}V(x) = -x^T Q x + \langle E, PE \rangle_F \leq -\alpha V + \langle E, PE \rangle_F \,, \tag{40}$$

where α is the maximal α such that $Q \succeq \alpha P$ (which can be found by solving an LMI), and

$$\nabla V(x) e(x) = 2x^T P E. \tag{41}$$

Thus Theorem 3 applies with

$$c = \langle E, PE \rangle_F \,, \quad \sigma = 2\|L^T E\|, \tag{42}$$

where $P = LL^T$ is the Cholesky decomposition of P.

We illustrate this numerically with the system

$$\mathrm{d}X_t = \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} X_t + u_t \right) \mathrm{d}t + \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \mathrm{d}W_t. \tag{43}$$

Although this is a very simple system, it serves to illustrate the difference in bound tightness. Setting $Q = 0.2I_2$ yields $P = 0.447I_2$ and the parameters $\alpha = 0.4472$, $c = 0.1431$, and $\sigma = 0.7566$. Sample trajectories of $V(X_t)$ and of Z_t from Theorem 3 simulated over 3,000 discrete steps, are shown in Fig. 5. Estimated exit probabilities calculated from 1,000 sample trajectories are shown in Fig. 4 together with analytic bounds from the literature.

V. Conclusions

In this paper we presented a comparison theorem for one-dimensional SDEs, and applied it to upper-bound exit probabilities for multi-dimensional SDEs in terms of an exit probability of a one-dimensional process. As opposed to known closed-form bounds in the literature, these bounds take the noise magnitude into account and are therefore much tighter for small-probability events. Furthermore, this is a generalization of known results, in the sense that as the noise magnitude goes to zero, we retrieve invariance results from the barrier function literature.

The bounds are not on closed form, but depend on three resp. four parameters and can therefore feasibly be tabulated for an application. Closed-form bounds would however be more desirable; it is conceivable that such bounds can be obtained via approximate solutions to the PDE (18), but we leave this question for the future. Another direction for future work is to use sums-of-squares programming [16] to automatically search for certificate functions V that satisfy the conditions in Theorem 2.

References

[1] Christian Dehnert, Sebastian Junges, Joost-Pieter Katoen, and Matthias Volk. A Storm is Coming: A Modern Probabilistic Model Checker. In *Proc. CAV*, pages 592–600, 2017.

[2] Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of Probabilistic Real-Time Systems. In *Proc. CAV*, volume 6806, pages 585–591, 2011.

[3] Harold J. Kushner. *Stochastic Stability and Control*. Academic Press Inc., 1967.

[4] Stephen Prajna, Ali Jadbabaie, and George J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, 52(8):1415–1428, 2007.

[5] Jacob Steinhardt and Russ Tedrake. Finite-time regional verification of stochastic non-linear systems. *International Journal of Robotics Research*, 31(7):901–923, 2012.

[6] Bernt Øksendal. *Stochastic Differential Equations*. Springer Berlin Heidelberg, 6 edition, 2003.

[7] Eduardo D. Sontag. *Input to state stability: Basic concepts and results*, pages 163–220. Springer, 2006.

[8] Majid Zamani, Matthias Rungger, and Peyman Mohajerin Esfahani. Approximations of Stochastic Hybrid Systems: A Compositional Approach. *IEEE Transactions on Automatic Control*, 62(6):2838–2853, 2017.

[9] Aaron D. Ames, Xiangru Xu, Jessy W. Grizzle, and Paulo Tabuada. Control Barrier Function Based Quadratic Programs for Safety Critical Systems. *IEEE Trans. Autom. Control*, 62(8):3861–3876, 2017.

[10] Alexander Lipton and Vadim Kaushansky. On the First Hitting Time Density of an Ornstein-Uhlenbeck Process. 2018. https://arxiv.org/abs/1810.02390.

[11] Svante Janson and Johan Tysk. Preservation of convexity of solutions to parabolic equations. *Journal of Differential Equations*, 206(1):182–226, 2004.

[12] Jonathan E. Guyer, Daniel Wheeler, and James A. Warren. FiPy: Partial differential equations with Python. *Computing in Science & Engineering*, 11(3):6–15, 2009.

[13] Shige Peng and Xuehong Zhu. Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations. *Stochastic Processes and their Applications*, 116(3):370–380, 2006.

[14] S E Graversen and G Peskir. Maximal inequalities for the Ornstein-Uhlenbeck process. *Proc. Amer. Math. Soc.*, 128(10):3035–3042, 2000.

[15] Quan Nguyen and Koushil Sreenath. Exponential Control Barrier Functions for enforcing high relative-degree safety-critical constraints. In *Proc. ACC*, pages 322–328, 2016.

[16] Jean Bernard Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2010.

Authorized licensed use limited to: CALIFORNIA INSTITUTE OF TECHNOLOGY. Downloaded on September 23,2020 at 21:58:55 UTC from IEEE Xplore. Restrictions apply.