SIAM J. Control Optim.
Vol. 45, No. 4, pp. 1297–1313

ALMOST GLOBAL STOCHASTIC STABILITY*

RAMON VAN HANDEL†

Abstract. We develop a method to prove almost global stability of stochastic differential equations in the sense that almost every initial point (with respect to the Lebesgue measure) is asymptotically attracted to the origin with unit probability. The method can be viewed as a dual to Lyapunov's second method for stochastic differential equations and extends the deterministic result of [A. Rantzer, *Syst. Control Lett.*, 42 (2001), pp. 161–168]. The result can also be used in certain cases to find stabilizing controllers for stochastic nonlinear systems using convex optimization. The main technical tool is the theory of stochastic flows of diffeomorphisms.

Key words. stochastic stability, stochastic flows, nonlinear stochastic control

AMS subject classifications. 34F05, 60H10, 93C10, 93D15, 93E15

DOI. 10.1137/040618850

1. Introduction.

Lyapunov's second method or the method of Lyapunov functions, though developed in the late 19th century, remains one of the most important tools in the study of deterministic differential equations. The power of the method lies in the fact that an important qualitative property of a differential equation, the stability of an equilibrium point, can be proved without solving the equation explicitly. The theory was generalized to stochastic differential equations in the 1960s with fundamental contributions by Has'minskiĭ [10] and Kushner [15].

Lyapunov's method also underlies many important applications in the area of nonlinear control [11]. Finding optimal controls for nonlinear systems is generally an intractable problem, but often a solution can be found which stabilizes the system. Unlike in deterministic control theory, where nonlinear control is now a major field, there are very few results on stochastic nonlinear control. It is only recently that stochastic versions of the classical stabilization results of Jurdjevic-Quinn, Artstein, and Sontag were developed by Florchinger [7, 8, 9] and backstepping designs for stochastic strict-feedback systems were developed by Deng and Krstić [5] and Deng Krstić, and Williams [6].

In this paper we will not consider stochastic stability in the sense of Has'minskiĭ; rather, we ask the following question: for a given Itô stochastic differential equation on $\mathbb{R}^n$, can we prove that for almost every initial state (with respect to the Lebesgue measure on $\mathbb{R}^n$) the solution of the equation converges to the origin almost surely as $t \to \infty$, i.e., is the origin *almost* globally stable? This notion of stability is clearly weaker than global stability in the sense of Has'minskiĭ, but is of potential interest in many cases in which global stability may not be attained.

Our main result is a Lyapunov-type theorem that can be used to prove almost global stability of stochastic differential equations, extending the deterministic result of Rantzer [23]. The theorem has several remarkable properties. It can be viewed as a "dual" to Lyapunov's second method in the following sense: whereas the Lyapunov

*Received by the editors November 13, 2004; accepted for publication (in revised form) April 3, 2006; published electronically October 3, 2006. This work was supported by the ARO under grant DAAD19-03-1-0073.
http://www.siam.org/journals/sicon/45-4/61885.html

†Departments of Physics and Control & Dynamical Systems, California Institute of Technology 266-33, Pasadena, CA 91125 (ramon@its.caltech.edu).

condition reads $\mathscr{L}V < 0$, where $\mathscr{L}$ is the characteristic operator of the stochastic differential equation and V is the Lyapunov function, the condition that guarantees almost global stability reads $\mathscr{L}^*D < 0$, where $\mathscr{L}^*$ is the formal adjoint of $\mathscr{L}$ (also known as the Fokker–Planck operator). Hence the relation between the two theorems recalls the duality between densities and expectations which is prevalent throughout the theory of stochastic processes.

A further interesting property is the following convexity property. Suppose we are given an Itô equation of the form

$$x_t = x + \int_s^t (X_0(x_\tau) + u(x_\tau)Y(x_\tau))\,d\tau + \sum_{k=1}^m \int_s^t X_k(x_\tau)\,dW_\tau^k,$$

where $X_k(0) = 0$ for $k = 0, \ldots, m$ and $u(x)$ is a state feedback control. The goal is to design $u(x)$ such that the origin is almost globally stable. It is easily verified that the set of pairs of functions $(D(x), u(x)D(x))$ which satisfy $\mathscr{L}^*D < 0$ is convex. Note that the classical Lyapunov condition $\mathscr{L}V < 0$ is not convex.

The above convexity property was used in the deterministic case by Prajna, Parrilo, and Rantzer [22] to formulate the search for almost globally stabilizing controllers as a convex optimization problem, provided that X_k, Y, D, and u are rational functions. The method applies equally to the stochastic case and thus provides a tool for computer-aided design of stochastic nonlinear controllers.

It must be emphasized that almost global stability is a global property of the flow which places very few restrictions on the local behavior near the origin. In particular, local stability is not implied.[1] A very fruitful approach to studying the local dynamical behavior of stochastic differential equations (and more general random dynamical systems) is developed by Arnold [2]. First, the flow associated to the stochastic differential equation is linearized; then Oseledec's multiplicative ergodic theorem is used to provide a suitable "time-averaged" notion of the eigenvalues of the linearized flow. To prove almost global stability we do not linearize the flow, though the proofs still rely on the flow of diffeomorphisms generated by the stochastic equation. We refer to [1] for an introduction to the dynamical approach to stochastic analysis.

This paper is organized as follows. In section 2 we fix the notation that will be used in the remainder of the paper. In section 3 we reproduce the deterministic result of Rantzer [23] with a significantly different proof that generalizes to the stochastic case. In section 4 we review the theory of stochastic flows of diffeomorphisms generated by stochastic differential equations. Section 5 is devoted to the statement and proof of our main result for the case of globally Lipschitz continuous coefficients. In section 6 the main result is extended to cases in which the global Lipschitz condition does not necessarily hold. A few examples are given in section 7. Finally, in section 8 we discuss the application to control synthesis using convex optimization.

2. Notation.

Throughout this article we will consider (stochastic) differential equations in $\mathbb{R}^n$. The Lebesgue measure on $\mathbb{R}^n$ will be denoted by μ. $\mathbb{R}_+$ denotes the nonnegative real numbers and $\mathbb{Z}_+$ the nonnegative integers.

We remind the reader of the following definitions: For $0 < \alpha \leq 1$, a function $f : X \to Y$ from a normed space $(X, \|\cdot\|)$ to a normed space $(Y, \|\cdot\|_Y)$ is called

[1] The term "stability" seems a bit of a misnomer; despite that almost all points converge to the origin, a trajectory that starts close to the origin could move very far from the origin before converging to it. We have used the term that has been used in the deterministic literature, e.g., [18].

globally Hölder continuous of order α if there exists a positive constant C such that

$$\|f(x) - f(y)\|_Y \le C\|x - y\|^\alpha \qquad \forall x, y \in X. \tag{2.1}$$

f is locally Hölder continuous of order α if it satisfies the condition (2.1) on every bounded subset of X. f is called globally (locally) Lipschitz continuous if it is globally (locally) Hölder continuous of order 1. f is called a $C^{k,\alpha}$ function if it is k times continuously differentiable and the kth derivatives are locally Hölder continuous of order α for some $k \in \mathbb{Z}_+$ and $0 < \alpha \le 1$.

3. The deterministic case. In this section we give a new proof of Rantzer's theorem [23], which is a deterministic counterpart of our main result. Our proof demonstrates the main features of the proof of the stochastic result in the simpler deterministic case.

The following lemma is similar to Lemma A.1 in [23], and we omit the proof.

LEMMA 3.1. *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ be globally Lipschitz continuous, let $S \subset \mathbb{R}^n$ be an invariant set of $\dot{x}(t) = f(x(t))$, and let $Z \subset S$ be μ-measurable. Let $D \in C^1(S, \mathbb{R})$ be integrable on Z. Then*

$$\int_{\phi_t^{-1}(Z)} D(x)\,dx = \int_Z D(x)\,dx - \int_0^t \int_{\phi_\tau^{-1}(Z)} [\nabla \cdot (fD)](x)\,dx\,d\tau, \tag{3.1}$$

where $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ is the flow of f.

THEOREM 3.2. *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ be globally Lipschitz continuous and let $f(0) = 0$. Suppose there exists $D \in C^1(\mathbb{R}^n\backslash\{0\}, \mathbb{R}_+)$ such that D is integrable on $\{x \in \mathbb{R}^n : |x| > 1\}$ and $[\nabla \cdot (fD)](x) > 0$ for μ-almost all x. Then for μ-almost all initial states $x(0)$ the solution of $\dot{x}(t) = f(x(t))$ tends to the origin as $t \to \infty$.*

Proof. Let $S = \mathbb{R}^n\backslash\{0\}$, $\varepsilon > 0$ and $Z = \{x \in \mathbb{R}^n : |x| > \varepsilon\}$. Note that $\phi_t(x)$ is a diffeomorphism for every $t \in \mathbb{R}$; hence $\phi_t(x)$ is one-to-one, and as $\phi_t(0) = 0$, $t \in \mathbb{R}$, is a solution of $\dot{x}(t) = f(x(t))$ there can be no $x \in S$ such that $\phi_t(x) = 0$ for some $t \in \mathbb{R}$. We have thus verified the invariance of S under the flow $\phi_t(x)$. We now invoke Lemma 3.1. As $D(x)$ is nonnegative, expression (3.1) is also nonnegative. Furthermore, (3.1) is finite because D is integrable on Z, and is nonincreasing due to $[\nabla \cdot (fD)](x) \ge 0$. By monotone convergence the limit as $t \to \infty$ exists and is finite. Hence

$$\int_0^\infty \mathbb{D}(\phi_\tau^{-1}(Z))\,d\tau < \infty, \qquad \mathbb{D}(A) = \int_A [\nabla \cdot (fD)](x)\,dx.$$

Note that the assumption $[\nabla \cdot (fD)](x) \ge 0$ implies that $\mathbb{D}$ is a measure on S. The measure space $(S, \mathbb{D})$ is σ-finite as $\mathbb{D}(\{x \in S : \frac{1}{k} < |x| < k\}) < \infty$ for all $k > 1$ and $\bigcup_{n=2}^\infty \{x \in S : \frac{1}{k} < |x| < k\} = S$.

We now fix some $m \in \mathbb{N}$ and divide the halfline into bins $S_k^m = [(k-1)2^{-m}, k2^{-m}]$, $k \in \mathbb{N}$. From each bin we choose a time $t_k^m \in S_k^m$ such that

$$\mathbb{D}(\phi_{t_k^m}^{-1}(Z)) \le \inf_{t \in S_k^m} \mathbb{D}(\phi_t^{-1}(Z)) + 2^{-k}.$$

For fixed m, we denote this discrete grid by $T_m = \{t_k^m : k \in \mathbb{N}\}$. We now have

$$2^{-m} \sum_{k=1}^\infty \mathbb{D}(\phi_{t_k^m}^{-1}(Z)) \le 2^{-m} + \int_0^\infty \mathbb{D}(\phi_\tau^{-1}(Z))\,d\tau < \infty.$$

As $\mathbb{D}$ is σ-finite we can now apply the Borel–Cantelli lemma, which gives

$$\mathbb{D}\left(\limsup_{k\to\infty} \phi_{t_k^m}^{-1}(Z)\right) = \mu\left(\limsup_{k\to\infty} \phi_{t_k^m}^{-1}(Z)\right) = 0,$$

where the first equality follows as $[\nabla\cdot(fD)](x) > 0$ μ-a.e. implies $\mu \ll \mathbb{D}$. Consequently

$$\mu\left(\bigcup_{m=1}^{\infty} \limsup_{t\in T_m} \phi_t^{-1}(Z)\right) \le \sum_{m=1}^{\infty} \mu\left(\limsup_{t\in T_m} \phi_t^{-1}(Z)\right) = 0.$$

We have thus shown that the set of initial states x for which there are, for some m, infinitely many times $t \in T_m$ such that $\phi_t(x) \in Z$ has Lebesgue measure zero.

We now claim that if $\limsup_{t\to\infty} |\phi_t(x)| > \varepsilon$, then we can choose m so that there are infinitely many times t in T_m such that $\phi_t(x) \in Z$. The statement is trivial if also $\liminf_{t\to\infty} |\phi_t(x)| > \varepsilon$; let us thus assume that $\liminf_{t\to\infty} |\phi_t(x)| \le \varepsilon$. We will need the following result. Due to the global Lipschitz condition and $f(0) = 0$, we have

$$|\phi_t(x)| \le |\phi_s(x)| + \int_s^t |f(\phi_\sigma(x))|\, d\sigma \le |\phi_s(x)| + C\int_s^t |\phi_\sigma(x)|\, d\sigma \tag{3.2}$$

for some constant $C > 0$. Thus Gronwall's lemma gives $|\phi_t(x)| \le |\phi_s(x)|\, e^{C(t-s)}$. Now note that $\liminf_{t\to\infty} |\phi_t(x)| \le \varepsilon < \limsup_{t\to\infty} |\phi_t(x)|$ implies that there exist $\varepsilon'' > \varepsilon' > \varepsilon$ such that (i) there are infinitely many upcrossings of the curve $|\phi_t(x)|$ through ε', and (ii) $|\phi_t(x)|$ crosses ε'' infinitely often. Denote by t'' a time such that $|\phi_{t''}(x)| = \varepsilon''$ and by t' the latest time previous to t'' that $|\phi_{t'}(x)| = \varepsilon'$. Then clearly $t'' - t' \ge \frac{1}{C}\log\frac{\varepsilon''}{\varepsilon'}$. As this happens infinitely often, we conclude that $\phi_t(x)$ infinitely often spends a time in excess of $\frac{1}{C}\log\frac{\varepsilon''}{\varepsilon'}$ in Z. But then clearly m can be chosen large enough so that every such interval includes at least one of the $t_k^m \in T_m$.

We have now shown that for μ-almost all $x \in \mathbb{R}^n$ we have $\limsup_{t\to\infty} |\phi_t(x)| \le \varepsilon$, i.e., for μ-almost all $x \in \mathbb{R}^n$ $\exists t_e > 0$ such that $|\phi_t(x)| \le \varepsilon$ for $t \ge t_e$. But as this holds for any $\varepsilon > 0$ the trajectories must converge to the origin. □

4. Stochastic flows. The purpose of this section is to review, without proofs, some results of the theory of stochastic flows of diffeomorphisms generated by stochastic differential equations. A detailed exposition on the subject can be found in [13, 14], and shorter treatments are in [1, 2, 3, 12].

Throughout this article $(\Omega, \mathcal{F}, \mathbb{P})$ denotes the canonical Wiener space of the m-dimensional Brownian motion W_t with two-sided time $\mathbb{R}$. We also introduce the two-parameter filtration $\mathcal{F}_s^t = \sigma\{W_u^k - W_v^k : s \le v \le u \le t,\ 1 \le k \le m\}$. The extension to two-sided time is important in that it allows us to treat the Wiener process as a dynamical system [1, 2, 3].

THEOREM 4.1. *There exists a one-parameter group $\{\theta_t : t \in \mathbb{R}\}$ of measure-preserving transformations of $(\Omega, \mathcal{F}, \mathbb{P})$ such that $W_t(\theta_s\omega) = W_{t+s}(\omega) - W_s(\omega)$ for all $\omega \in \Omega$ and $s, t \in \mathbb{R}$.*

We will consider Itô stochastic differential equations of the form

$$x_t = x + \int_s^t X_0(x_\tau)\, d\tau + \sum_{k=1}^m \int_s^t X_k(x_\tau)\, dW_\tau^k \tag{4.1}$$

with the following assumptions:

1. $x \in \mathbb{R}^n$.

2. $X_k : \mathbb{R}^n \to \mathbb{R}^n$, $k = 0, \dots, m$, are globally Lipschitz continuous.

The global Lipschitz condition guarantees many nice properties of the solutions; we will assume it for the time being, and later relax this requirement somewhat in section 6.

Denote by $\xi_{s,t}(x,\omega)$ (or simply $\xi_{s,t}(x)$) the solution of (4.1) at time $t \geq s$ given the initial condition $x_s = x$. It is well known that in the case of globally Lipschitz continuous coefficients there exists a unique, nonexploding solution $\xi_{s,t}(x)$ which is an $\mathcal{F}_s^t$-semimartingale and is in L^p for any $p \geq 1$ (see, e.g., [13]).

THEOREM 4.2 (see [13, 2]). *Suppose X_k, $k = 0, \dots, m$, are globally Lipschitz continuous and let $s < t$. Then we have the following properties:*

1. *$\xi_{s,s}(x,\omega) = x$ for all s and ω.*
2. *For any u we have $\xi_{s,t}(\cdot, \theta_u\omega) = \xi_{s+u,t+u}(\cdot,\omega)$.*
3. *For almost all ω we have $\xi_{s,t}(\cdot,\omega) = \xi_{r,t}(\xi_{s,r}(\cdot,\omega),\omega)$ for all $s < r < t$.*
4. *$\xi_{s,t}(x)$ is $\mathbb{P}$-a.s. continuous in (s,t,x).*
5. *For almost all ω the map $\xi_{s,t}(\cdot,\omega) : \mathbb{R}^n \to \mathbb{R}^n$ is a homeomorphism for all $s < t$.*

The following result establishes that, under additional smoothness conditions, $\xi_{s,t}(x)$ is in fact a *stochastic flow of diffeomorphisms*.

THEOREM 4.3 (see [13]). *Suppose X_k, $k = 0, \dots, m$, are globally Lipschitz continuous and that they are $C^{p,\alpha}$ functions for some $p \geq 1$ and $0 < \alpha < 1$. Then for almost all ω the map $\xi_{s,t}(\cdot,\omega) : \mathbb{R}^n \to \mathbb{R}^n$ is a C^p diffeomorphism for any $s \leq t$, and*

$$\frac{\partial \xi_{s,t}(x)^i}{\partial x^j} = \delta^i_j + \sum_{\beta=1}^n \int_s^t \frac{\partial X_0^i}{\partial x^\beta}(\xi_{s,\tau}(x))\, \frac{\partial \xi_{s,\tau}(x)^\beta}{\partial x^j}\, d\tau + \sum_{k=1}^m \sum_{\beta=1}^n \int_s^t \frac{\partial X_k^i}{\partial x^\beta}(\xi_{s,\tau}(x))\, \frac{\partial \xi_{s,\tau}(x)^\beta}{\partial x^j}\, dW_\tau^k. \tag{4.2}$$

It will be convenient for our purposes to work with the inverse flow $\xi_{s,t}^{-1}(x)$, considered as a backward stochastic process in the time variable s (with t fixed). This will not give rise to ordinary Itô integrals as s behaves like a time-reversed variable, and hence the adaptedness of the process runs backward in time. The Itô backward integral is defined as [13]

$$\int_s^t f_\sigma \overleftarrow{dW_\sigma} \equiv \lim \text{ in prob} \sum_{k=0}^{n-1} f_{t_{k+1}}(W_{t_{k+1}} - W_{t_k}),$$

where f_s is a backward predictable process with $\int_s^t |f_u|^2\, du < \infty$ almost surely, and the formal construction of the integral from simple functions proceeds along the usual lines. The backward integral has similar properties to the forward integral; in particular, it is a backward $\mathcal{F}_s^t$-local martingale (for fixed t) and satisfies an Itô formula (see, e.g., [4, p. 124]), which is proved in the same way as its forward counterpart: given $\xi_s = \xi_t + \int_s^t a_\sigma\, d\sigma + \sum_k \int_s^t (b_\sigma)_k \overleftarrow{dW_\sigma^k}$ with backward predictable processes a_s, $(b_s)_k$ such that $\int_s^t a_\sigma\, d\sigma < \infty$ almost surely, $\int_s^t |(b_\sigma)_k|^2\, d\sigma < \infty$ almost surely, then for any C^2 function $F : \mathbb{R}^n \to \mathbb{R}$

$$F(\xi_s) = F(\xi_t) + \frac{1}{2} \sum_{k=1}^m \sum_{i,j=1}^n \int_s^t (b_\sigma)_k^i (b_\sigma)_k^j \frac{\partial^2 F}{\partial x^i \partial x^j}(\xi_\sigma)\, d\sigma + \sum_{i=1}^n \int_s^t a_\sigma^i \frac{\partial F}{\partial x^i}(\xi_\sigma)\, d\sigma + \sum_{k=1}^m \sum_{i=1}^n \int_s^t (b_\sigma)_k^i \frac{\partial F}{\partial x^i}(\xi_\sigma) \overleftarrow{dW_\sigma^k}. \tag{4.3}$$

We can now formulate the following result.

THEOREM 4.4 (see [13]). *Suppose X_k, $k = 0, \dots, m$, are globally Lipschitz continuous and that they are $C^{p,\alpha}$ functions for some $p \geq 2$ and $0 < \alpha < 1$. Then*

$$\xi_{s,t}^{-1}(x) = x - \int_s^t \tilde{X}_0(\xi_{\sigma,t}^{-1}(x))\, d\sigma - \sum_{k=1}^m \int_s^t X_k(\xi_{\sigma,t}^{-1}(x))\, \overleftarrow{dW_\sigma},$$

where we have defined

$$\tilde{X}_0(x) = X_0(x) - \sum_{k=1}^m \sum_{\beta=1}^n X_k^\beta(x) \frac{\partial}{\partial x^\beta} X_k(x).$$

This expression can be manipulated much in the same way as its forward counterpart. In particular, under the conditions of Theorem 4.4 and using (4.3), we obtain for any C^2 function $F : \mathbb{R}^n \to \mathbb{R}$ the backward Itô formula

$$\begin{aligned} F(\xi_{s,t}^{-1}(x)) = F(x) &+ \frac{1}{2} \sum_{k,i,j} \int_s^t X_k^i(\xi_{\sigma,t}^{-1}(x)) X_k^j(\xi_{\sigma,t}^{-1}(x)) \frac{\partial^2 F}{\partial x^i \partial x^j}(\xi_{\sigma,t}^{-1}(x))\, d\sigma \\ &- \sum_i \int_s^t \tilde{X}_0^i(\xi_{\sigma,t}^{-1}(x)) \frac{\partial F}{\partial x^i}(\xi_{\sigma,t}^{-1}(x))\, d\sigma - \sum_{k,i} \int_s^t X_k^i(\xi_{\sigma,t}^{-1}(x)) \frac{\partial F}{\partial x^i}(\xi_{\sigma,t}^{-1}(x))\, \overleftarrow{dW_\sigma^k}. \end{aligned} \tag{4.4}$$

Similarly we can differentiate the inverse flow, giving

$$\begin{aligned} \frac{\partial \xi_{s,t}^{-1}(x)^i}{\partial x^j} = \delta_j^i &- \sum_{\beta=1}^n \int_s^t \frac{\partial \tilde{X}_0^i}{\partial x^\beta}(\xi_{\sigma,t}^{-1}(x))\, \frac{\partial \xi_{\sigma,t}^{-1}(x)^\beta}{\partial x^j}\, d\sigma \\ &- \sum_{k=1}^m \sum_{\beta=1}^n \int_s^t \frac{\partial X_k^i}{\partial x^\beta}(\xi_{\sigma,t}^{-1}(x))\, \frac{\partial \xi_{\sigma,t}^{-1}(x)^\beta}{\partial x^j}\, \overleftarrow{dW_\sigma^k}. \end{aligned} \tag{4.5}$$

This expression is obtained, in the same way as its forward counterpart (4.2), by letting $y \to 0$ in the backward expression corresponding to [13, p. 219, eqn. (4)].

5. The main result.

We consider an Itô equation of the form (4.1). We write

$$\mathscr{L}^* f(x) = \frac{1}{2} \sum_{k=1}^m \sum_{i,j=1}^n \frac{\partial^2}{\partial x^i \partial x^j}(X_k^i(x) X_k^j(x) f(x)) - \sum_{i=1}^n \frac{\partial}{\partial x^i}(X_0^i(x) f(x)).$$

The following is our main result.

THEOREM 5.1. *Let X_k, $k = 0, \dots, m$, be globally Lipschitz continuous and $C^{2,\alpha}$ functions for some $\alpha > 0$, and let $X_k(0) = 0$. Suppose there exists $D \in C^2(\mathbb{R}^n \backslash \{0\}, \mathbb{R}_+)$ such that D is integrable on $\{x \in \mathbb{R}^n : |x| > 1\}$ and $\mathscr{L}^* D(x) < 0$ for μ-almost all x. Then for every initial time s and μ-a.e. initial state x the flow $\xi_{s,t}(x)$ tends to the origin as $t \to \infty$ $\mathbb{P}$-a.s.*

Before we prove the theorem, let us prove a stochastic version of Lemma 3.1.

LEMMA 5.2. *Suppose X_k, $k = 0, \dots, m$, are globally Lipschitz continuous and $C^{2,\alpha}$ functions for some $\alpha > 0$. Let $S_\ell \subset S_{\ell+1} \subset \mathbb{R}^n$ be an increasing sequence of open sets such that $\tau_\ell = \sup\{s < t : \xi_{s,t}^{-1}(x) \notin S_\ell\} \to -\infty$ as $\ell \to \infty$ $\mathbb{P}$-a.s. for every $x \in S = \bigcup_\ell S_\ell$. Suppose there is a $D \in C^2(S, \mathbb{R}_+)$ that is integrable on a* *measurable set $Z \subset S$, that obeys $\mathscr{L}^* D \le 0$ on S, and such that for each ℓ there is a $D_\ell \in C^2(\mathbb{R}^n, \mathbb{R}_+)$ that coincides with D on S_ℓ. Then*

$$0 \le \int_Z D(x)\,dx + \int_s^t \mathbb{E} \int_{\xi_{\sigma,t}^{-1}(Z)} \mathscr{L}^* D(x)\,dx\,d\sigma$$

for all $s \le t$, and in particular the limit as $s \to -\infty$ of this expression is well defined.

Proof. Denote by $\mathbf{J}_{s,t}(x)$ the matrix with elements $\mathbf{J}_{s,t}(x)^i_j = \partial \xi_{s,t}^{-1}(x)^i / \partial x^j$, i.e.,

$$\mathbf{J}_{s,t}(x)^i_j = \delta^i_j - \sum_\alpha \int_s^t \frac{\partial \tilde{X}_0^i}{\partial x^\alpha}(\xi_{\sigma,t}^{-1}(x)) \mathbf{J}_{\sigma,t}(x)^\alpha_j\,d\sigma - \sum_{k,\alpha} \int_s^t \frac{\partial X_k^i}{\partial x^\alpha}(\xi_{\sigma,t}^{-1}(x)) \mathbf{J}_{\sigma,t}(x)^\alpha_j\,\overleftarrow{dW_\sigma^k}$$

by (4.5). Denote by $|\mathbf{J}_{s,t}(x)|$ its determinant, i.e.,

$$|\mathbf{J}_{s,t}(x)| = \sum_{j_1 \cdots j_n = 1}^{n} \varepsilon_{j_1,\dots,j_n} \mathbf{J}_{s,t}(x)^1_{j_1} \mathbf{J}_{s,t}(x)^2_{j_2} \cdots \mathbf{J}_{s,t}(x)^n_{j_n},$$

where $\varepsilon_{j_1,\dots,j_n}$ is the antisymmetric tensor. Using Itô's rule and straightforward calculations we obtain

$$\begin{aligned}|\,\mathbf{J}_{s,t}(x)| = 1 &- \sum_i \int_s^t \frac{\partial \tilde{X}_0^i}{\partial x^i}(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,d\sigma - \sum_{k,i} \int_s^t \frac{\partial X_k^i}{\partial x^i}(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,\overleftarrow{dW_\sigma^k} \\ &+ \frac{1}{2} \sum_{k,i,j} \int_s^t \left[\frac{\partial X_k^i}{\partial x^i}(\xi_{\sigma,t}^{-1}(x)) \frac{\partial X_k^j}{\partial x^j}(\xi_{\sigma,t}^{-1}(x)) - \frac{\partial X_k^i}{\partial x^j}(\xi_{\sigma,t}^{-1}(x)) \frac{\partial X_k^j}{\partial x^i}(\xi_{\sigma,t}^{-1}(x)) \right] |\mathbf{J}_{\sigma,t}(x)|\,d\sigma.\end{aligned}$$

Note that as $\xi_{s,t}^{-1}(\cdot)$ is a diffeomorphism almost surely, its Jacobian $\mathbf{J}_{s,t}(\cdot)$ must almost surely be an invertible matrix; but as $|\mathbf{J}_{s,t}(x)|$ has almost surely continuous sample paths and $|\mathbf{J}_{t,t}(x)| = 1$, this implies that almost surely $|\mathbf{J}_{s,t}(x)| > 0$ for all $s < t$. Using (4.4) with $F = D_\ell$ and Itô's rule we obtain

$$\begin{aligned}0 \le D_\ell(\xi_{s,t}^{-1}(x)) |\mathbf{J}_{s,t}(x)| = D_\ell(x) &+ \int_s^t (\mathscr{L}^* D_\ell)(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,d\sigma \\ &- \sum_{k=1}^m \sum_{i=1}^n \int_s^t \frac{\partial X_k^i D_\ell}{\partial x^i}(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,\overleftarrow{dW_\sigma^k}.\end{aligned}$$

Now note that as D_ℓ coincides with D on S_ℓ, we can identify $(\mathscr{L}^* D_\ell)(\xi_{s\vee\tau_\ell,t}^{-1}(x)) = (\mathscr{L}^* D)(\xi_{s\vee\tau_\ell,t}^{-1}(x))$ for every ℓ. Moreover, as the last term in the expression above is a backward local martingale, there exists a sequence of stopping times $\tau'_p \searrow -\infty$ such that the stochastic integral stopped at τ'_p is a martingale. Replacing s by $s \vee \tau_\ell \vee \tau'_p$ in the expression above and taking the expectation gives

$$0 \le D(x) + \mathbb{E} \int_{s\vee\tau_\ell\vee\tau'_p}^t (\mathscr{L}^* D)(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,d\sigma.$$

We can now let $\ell, p \to \infty$ by monotone convergence. Integrating both sides gives

$$0 \le \int_Z D(x)\,dx + \int_s^t \mathbb{E} \int_Z (\mathscr{L}^* D)(\xi_{\sigma,t}^{-1}(x)) |\mathbf{J}_{\sigma,t}(x)|\,dx\,d\sigma,$$

where we have used Tonelli's theorem to change the order of integration. The result follows after a change of coordinates. □

The proof of Theorem 5.1 is similar to the proof of Theorem 3.2. The stochastic version of the argument following (3.2), however, is a little more subtle, as we do not have a pathwise upper bound on the rate of growth of sample paths. On the other hand, we can establish such a bound *in probability* which, together with the strong Markov property, is sufficient for our purposes; a similar argument was used in [16] to the same effect. For this purpose we give the following lemma, various versions of which appear in the literature (the result below is adapted from [6]).

LEMMA 5.3. *Let X_k, $k = 0, \ldots, m$, be locally Lipschitz continuous and let $\lambda > 0$. Then*

$$\mathbb{P}\left[\sup_{0\le\delta\le\Delta} |\xi_{s,s+\delta}(x) - x| \ge \lambda\right] \le K_1\Delta + K_2\Delta^2,$$

where $K_1, K_2 < \infty$ are constants that depend only on λ and $|x|$.

Proof. Let $\mathbf{W}_t$ be the m-vector with elements W_t^k and let $\mathbf{X}(\cdot)$ be the $n\times m$-matrix with entries $X_k^i(\cdot)$, $k = 1, \ldots, m$. For $r > 0$, define $B_r(x') = \{x \in \mathbb{R}^n : |x - x'| < r\}$, $B_r = B_r(0)$, and

$$\rho_0(r) = \sup_{|y|<r} |X_0(y)|, \qquad \rho_1(r) = \sup_{|y|<r} \|\mathbf{X}(y)\| = \sup_{|y|<r} \mathrm{tr}[\mathbf{X}(y)^T\mathbf{X}(y)]^{1/2}.$$

Let τ_r be the first exit time of $\xi_{s,t}(x)$ from B_r. In [6, p. 1240] it was established that

$$\mathbb{E}\left[\sup_{0\le\delta\le\Delta} |\xi_{s,(s+\delta)\wedge\tau_r}(x) - x|^2\right] \le 2\rho_0(r)^2\Delta^2 + 8\rho_1(r)^2\Delta.$$

Hence we have by Markov's inequality

$$\mathbb{P}\left[\sup_{0\le\delta\le\Delta} |\xi_{s,(s+\delta)\wedge\tau_r}(x) - x| \ge \lambda\right] \le \lambda^{-2}(2\rho_0(r)^2\Delta^2 + 8\rho_1(r)^2\Delta).$$

Now note that $B_\lambda(x)$ is strictly included in $B_{|x|+2\lambda}$, so that the first exit time from $B_\lambda(x)$ is no later than $\tau_{|x|+2\lambda}$. But then the events

$$\left\{\omega : \sup_{0\le\delta\le\Delta} |\xi_{s,(s+\delta)\wedge\tau_{|x|+2\lambda}}(x) - x| \ge \lambda\right\}, \quad \left\{\omega : \sup_{0\le\delta\le\Delta} |\xi_{s,s+\delta}(x) - x| \ge \lambda\right\}$$

are equivalent; after all, the events are equivalent on $\tau_{|x|+2\lambda} > s + \Delta$ by construction, whereas if $\tau_{|x|+2\lambda} \le s + \Delta$, both events must be true as $|\xi_{s,\tau_{|x|+2\lambda}}(x) - x| \ge \lambda$. Hence

$$\mathbb{P}\left[\sup_{0\le\delta\le\Delta} |\xi_{s,s+\delta}(x) - x| \ge \lambda\right] \le \lambda^{-2}(2\rho_0(|x| + 2\lambda)^2\Delta^2 + 8\rho_1(|x| + 2\lambda)^2\Delta),$$

where we have set $r = |x| + 2\lambda$. This completes the proof. □

We now turn to the proof of the main theorem.

Proof of Theorem 5.1. Let $\varepsilon > 0$ and $Z = \{x \in \mathbb{R}^n : |x| > \varepsilon\}$. We begin by applying Lemma 5.2. To this end, define $S_\ell = \{x \in \mathbb{R}^n : |x| > \ell^{-1}\}$, so $S = \bigcup_\ell S_\ell = \mathbb{R}^n\backslash\{0\}$. Clearly D is integrable on Z and there exists a $C^2(\mathbb{R}^n, \mathbb{R}_+)$-approximation D_ℓ of D for each ℓ. It remains to check that $\tau_\ell \to -\infty$. Suppose that this is not the case; then given $x \in S$ there must be a positive probability that $\xi_{s,t}^{-1}(x) = 0$ for some

$-\infty < s < t$. But $\xi_{s,t}^{-1}(0) = 0$ for all s and almost surely $\xi_{s,t}^{-1}(x)$ is one-to-one for all $s < t$, so this cannot happen. Hence all the conditions of Lemma 5.2 are satisfied, and we have

$$0 \le \int_Z D(x)\,dx + \int_s^t \mathbb{E} \int_{\xi_{\sigma,t}^{-1}(Z)} \mathscr{L}^* D(x)\,dx\,d\sigma. \tag{5.1}$$

Now note that (5.1) is nonincreasing with decreasing s due to $\mathscr{L}^* D \le 0$ and is finite because D is integrable on Z. By monotone convergence the limit as $s \to -\infty$ exists and is finite. Hence

$$\int_{-\infty}^t \mathbb{D}(\xi_{\sigma,t}^{-1}(Z))\,d\sigma < \infty, \qquad \mathbb{D}(A) = -\int_A \mathscr{L}^* D(x)\,(\mathbb{P}(d\omega) \times \mu(dx)),$$

where we have used Tonelli's theorem to convert the iterated integral to a single integral with respect to the product measure, and we slightly abuse our notation by writing $\mathbb{D}(\xi_{\sigma,t}^{-1}(Z)) = \mathbb{D}(\{(\omega, x) \in \Omega \times S : \xi_{\sigma,t}(x,\omega) \in Z\})$. Note that $\mathscr{L}^* D \le 0$ implies that $\mathbb{D}$ is a measure on $\Omega \times S$, and $\mathbb{D}$ is σ-finite as $\mathbb{D}(\Omega \times \{x \in S : \frac{1}{k} < |x| < k\}) < \infty$ for all $k > 1$ and $\bigcup_{n=2}^{\infty} (\Omega \times \{x \in S : \frac{1}{k} < |x| < k\}) = \Omega \times S$.

We now fix some $m \in \mathbb{N}$ and divide the halfline into bins $S_k^m = [(k-1)2^{-m}, k2^{-m}]$, $k \in \mathbb{N}$. From each bin we choose a time $t_k^m \in S_k^m$ such that

$$\mathbb{D}(\xi_{t-t_k^m,t}^{-1}(Z)) \le \inf_{s \in S_k^m} \mathbb{D}(\xi_{t-s,t}^{-1}(Z)) + 2^{-k}.$$

For fixed m, we denote this discrete grid by $T_m = \{t_k^m : k \in \mathbb{N}\}$. We now have

$$2^{-m} \sum_{k=1}^{\infty} \mathbb{D}(\xi_{t-t_k^m,t}^{-1}(Z)) \le 2^{-m} + \int_{-\infty}^t \mathbb{D}(\xi_{\sigma,t}^{-1}(Z))\,d\sigma < \infty.$$

Using the fact that the transformation θ_t of Theorem 4.1 is $\mathbb{P}$-preserving to shift the times t_k^m to the forward variable, we obtain

$$\sum_{k=1}^{\infty} \mathbb{D}(\xi_{s,s+t_k^m}^{-1}(Z)) = \sum_{k=1}^{\infty} \mathbb{D}(\xi_{t-t_k^m,t}^{-1}(Z)) < \infty.$$

As $\mathbb{D}$ is σ-finite we can now apply the Borel–Cantelli lemma, which gives

$$\mathbb{D}\left(\limsup_{k\to\infty} \xi_{s,s+t_k^m}^{-1}(Z)\right) = (\mathbb{P} \times \mu)\left(\limsup_{k\to\infty} \xi_{s,s+t_k^m}^{-1}(Z)\right) = 0,$$

where the first equality follows as $\mathscr{L}^* D(x) < 0$ μ-a.e. implies $\mathbb{P} \times \mu \ll \mathbb{D}$. Consequently

$$(\mathbb{P} \times \mu)\left(\bigcup_{m=1}^{\infty} \limsup_{t \in T_m} \xi_{s,s+t}^{-1}(Z)\right) \le \sum_{m=1}^{\infty} (\mathbb{P} \times \mu)\left(\limsup_{t \in T_m} \xi_{s,s+t}^{-1}(Z)\right) = 0.$$

We have thus shown that for all initial states x, except in a set $N \subset \mathbb{R}^n$ of Lebesgue measure zero, there is $\mathbb{P}$-a.s. for any m only a finite number of times t in the discrete grid T_m such that $\xi_{s,s+t}(x) \in Z$.

Let us fix an $x \notin N$. We now claim that the fact that $\mathbb{P}$-a.s. for any m there is only a finite number of times $t \in T_m$ such that $\xi_{s,s+t}(x) \in Z$ implies that $\mathbb{P}$-a.s. $\limsup_{t\to\infty} |\xi_{s,t}(x)| \le \varepsilon$. To see this, suppose $\mathbb{P}[\limsup_{t\to\infty} |\xi_{s,t}(x)| > \varepsilon] = \delta > 0$.

By monotone convergence $\mathbb{E}[\chi_{\limsup|\xi_{s,t}(x)|>\varepsilon'}] \nearrow \delta$ as $\varepsilon' \searrow \varepsilon$; hence there exists an $\varepsilon' > \varepsilon$ such that $\mathbb{P}[\limsup_{t\to\infty}|\xi_{s,t}(x)| > \varepsilon'] > 0$. We have already shown, however, that almost surely $|\xi_{s,t}(x)| \le \varepsilon$ for infinitely many times $t_n \nearrow \infty$. Hence

$$\mathbb{P}\left[\limsup_{t\to\infty}|\xi_{s,t}(x)| > \varepsilon'\right] > 0 \quad\Longrightarrow\quad \mathbb{P}[|\xi_{s,t}(x)| \text{ crosses } \varepsilon \text{ and } \varepsilon' \text{ infinitely often}] > 0.$$

Once we disprove the latter statement, the claim is proved by contradiction.

To this end, introduce the following sequence of predictable stopping times. Let $\sigma_0 = \inf\{t > s : |\xi_{s,t}(x)| \le \varepsilon\}$, $\tau_0 = \inf\{t > \sigma_0 : |\xi_{s,t}(x)| \ge \varepsilon'\}$, and for any $n > 0$ we set $\sigma_n = \inf\{t > \tau_{n-1} : |\xi_{s,t}(x)| \le \varepsilon\}$, $\tau_n = \inf\{t > \sigma_n : |\xi_{s,t}(x)| \ge \varepsilon'\}$. Define

$$\Omega_n(\Delta) = \{\omega \in \Omega : \tau_n < \infty,\ |\xi_{s,\tau_n+\delta}(x)| > \varepsilon\ \forall\, 0 \le \delta \le \Delta\}.$$

For any $\Delta > 0$, the set of $\omega \in \Omega$ such that $\omega \in \Omega_n(\Delta)$ for infinitely many n must be of $\mathbb{P}$-measure zero; after all, we can choose m sufficiently large so that every time interval of length Δ contains at least one point in T_m, and for points $t \in T_m$ we have $|\xi_{s,t}(x)| > \varepsilon$ only finitely often $\mathbb{P}$-a.s. Thus $\sum_n \chi_{\Omega_n(\Delta)} < \infty$ $\mathbb{P}$-a.s. To proceed, we use the following argument (see [17, pp. 398–399]). Introduce the discrete filtration $\mathcal{B}_k = \mathcal{F}_s^{\tau_{k+1}}$ and define $Z_k = X_k - Y_k$ with

$$X_k = \sum_{n=1}^{k} \chi_{\Omega_n(\Delta)}, \qquad Y_k = \sum_{n=1}^{k} \mathbb{E}[\chi_{\Omega_n(\Delta)}|\mathcal{B}_{n-1}].$$

As $\Omega_k(\Delta) \in \mathcal{B}_n$ for all $k \le n$, Z_k is a $\mathcal{B}_k$-martingale. Now define for $a > 0$ the stopping time $\kappa(a) = \inf\{n : Z_n > a\}$. As $|Z_k - Z_{k-1}| \le 1$ almost surely, the stopped process $Z_k' = Z_{k\wedge\kappa(a)}$ is a martingale that is bounded from above, and by the martingale convergence theorem Z_k' converges almost surely as $k \to \infty$ to a finite random variable Z_∞'. But as Z_k' and Z_k coincide on $\{\omega : \sup_n Z_n < a\}$ and $a > 0$ was chosen arbitrarily, we conclude that $Z_k \to Z_\infty < \infty$ on $\{\omega : \sup_n Z_n < \infty\}$ (modulo a null set). Note, however, that X_n and Y_n are both positive increasing processes and we have already established that $\sup_n X_n < \infty$ $\mathbb{P}$-a.s., so $\sup_n Z_n < \infty$ $\mathbb{P}$-a.s. But this implies that Z_k, and hence also Y_k, converges to a finite value $\mathbb{P}$-a.s. Thus we have established

$$\sum_{n=1}^{\infty} \mathbb{E}[\chi_{\Omega_n(\Delta)}|\mathcal{F}_s^{\tau_n}] < \infty \quad \mathbb{P}\text{-a.s.} \quad \text{for any } \Delta > 0.$$

Note that by the continuity of the sample paths, $|\xi_{s,\tau_n}(x)| = \varepsilon'$ on $\tau_n < \infty$. By Lemma 5.3, we can choose $\overline{\Delta} > 0$ sufficiently small such that

$$P(y) = \mathbb{P}\left[\sup_{0\le\delta\le\overline{\Delta}} |\xi_{s,s+\delta}(y) - y| < \frac{\varepsilon' - \varepsilon}{2}\right] \ge \frac{1}{2}$$

for all $|y| = \varepsilon'$. Using the strong Markov property, we can write

$$\infty > \sum_{n=1}^{\infty} \mathbb{E}[\chi_{\Omega_n(\overline{\Delta})}|\mathcal{F}_s^{\tau_n}] \ge \sum_{n=1}^{\infty} P(\xi_{s,\tau_n}(x))\,\chi_{\tau_n<\infty} \ge \frac{1}{2}\sum_{n=1}^{\infty} \chi_{\tau_n<\infty} \quad \mathbb{P}\text{-a.s.}$$

But this implies that $\tau_n < \infty$ finitely often $\mathbb{P}$-a.s., contradicting the assertion that $\mathbb{P}[|\xi_{s,t}(x)| \text{ crosses } \varepsilon \text{ and } \varepsilon' \text{ infinitely often}] > 0$. This is the desired result.

We have now shown that for μ-almost all $x \in \mathbb{R}^n$, $\mathbb{P}$-a.s. $\limsup_{t\to\infty} |\xi_{s,t}(x)| \le \varepsilon$, i.e., for μ-almost all $x \in \mathbb{R}^n$ $\mathbb{P}$-a.s., $\exists t_e > s$ such that $|\xi_{s,t}(x)| \le \varepsilon$ for $t \ge t_e$. But as this holds for any $\varepsilon > 0$ the flow must converge to the origin. ∎

The proof of Theorem 5.1 is readily extended to prove other assertions, such as the following instability theorem.

THEOREM 5.4. *Let X_k, $k = 0, \dots, m$, be globally Lipschitz continuous and $C^{2,\alpha}$ functions for some $\alpha > 0$. Suppose there exists a $D \in C^2(\mathbb{R}^n, \mathbb{R}_+)$ such that $\mathscr{L}^* D(x) < 0$ for μ-almost all x. Then for every initial time s and μ-a.e. initial state x the flow escapes to infinity, i.e., $|\xi_{s,t}(x)| \to \infty$ as $t \to \infty$ $\mathbb{P}$-a.s.*

Proof. Let $\varepsilon > 0$ and $Z' = \{x \in \mathbb{R}^n : |x| < \varepsilon\}$. Again we begin by applying Lemma 5.2. We can simply choose $S_\ell = S = \mathbb{R}^n$ for all ℓ; by nonexplosion $\tau_\ell = -\infty$ and the remaining conditions are evident. Hence

$$0 \le \int_{Z'} D(x)\, dx + \int_s^t \mathbb{E} \int_{\xi_{\sigma,t}^{-1}(Z')} \mathscr{L}^* D(x)\, dx\, d\sigma.$$

Proceeding in exactly the same way as in the proof of Theorem 5.1 we can now show that for μ-almost all $x \in \mathbb{R}^n$, $\mathbb{P}$-a.s. $\liminf_{t\to\infty} |\xi_{s,t}(x)| \ge \varepsilon$, i.e., for μ-almost all $x \in \mathbb{R}^n$ $\mathbb{P}$-a.s., $\exists t_e > s$ such that $|\xi_{s,t}(x)| \ge \varepsilon$ for $t \ge t_e$. But as this holds for any $\varepsilon > 0$ the flow must escape to infinity. ∎

Remark. At first sight the statements of Theorems 5.1 and 5.4 may seem contradictory, but this is not the case. The essential difference between the theorems is the region in $\mathbb{R}^n$ on which D is integrable. Roughly speaking, the idea behind the proofs of Theorems 5.1 and 5.4 is to show that if $\mathscr{L}^* D < 0$ μ-a.e., then the solution of the Itô equation can spend only a finite amount of time in any region on which D is integrable. Hence in Theorem 5.1 the solution will be attracted to the origin, whereas in Theorem 5.4 the solution is attracted to infinity.

If we try to satisfy the conditions of Theorems 5.1 and 5.4 simultaneously, we will run into problems. Suppose we have a nonnegative $D \in C^2(\mathbb{R}^n)$, as in Theorem 5.4, which is integrable as in Theorem 5.1. Then D is a normalizable density function, i.e., we could normalize D and interpret it as the density of the Itô equation at some point in time. But then $\mathscr{L}^* D < 0$ would imply that the associated Fokker–Planck equation does not preserve normalization of the density. Evidently Theorem 5.4 can only be satisfied if D is not integrable, whereas Theorem 5.1 requires D to have a singularity at the origin. See section 7 for examples.

6. Further results. In this section we extend the main result to cases in which the global Lipschitz condition is not necessarily satisfied. We first show that the result of Theorem 5.1 still holds if we can convert the coefficients of (4.1) to be globally Lipschitz continuous through a suitably chosen time transformation. In particular, this allows us to treat the case that X_k, $k = 0, \dots, m$, and their first derivatives are polynomially bounded, provided that some additional integrability conditions on D are satisfied. We also extend the main result to the case in which the flow is restricted to an invariant subset of $\mathbb{R}^n$ with compact closure.

THEOREM 6.1. *Let $X_k : \mathbb{R}^n \to \mathbb{R}^n$, $k = 0, \dots, m$, be measurable and let $X_k(0) = 0$. Suppose there is a strictly positive measurable map $c : \mathbb{R}^n \to (0, \infty)$ such that $c(x)$ and $c(x)^{-1}$ are locally bounded, and such that $c(x) X_0(x)$ and $\sqrt{c(x)}\, X_k(x)$, $k = 1, \dots, m$, are globally Lipschitz continuous and $C^{2,\alpha}$ functions for some $\alpha > 0$. Suppose there exists $D : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}_+$ such that $D(x)/c(x)$ is C^2 and is integrable on $\{x \in \mathbb{R}^n : |x| > 1\}$, and $\mathscr{L}^* D(x) < 0$ for μ-almost all x. Then for every initial* *time s and μ-a.e. initial state x the solution x_t of* (4.1) *tends to the origin as $t \to \infty$ $\mathbb{P}$-a.s.*

Proof. Consider the Itô equation

$$y_t = y_s + \int_s^t c(y_\tau) X_0(y_\tau)\, d\tau + \sum_{k=1}^m \int_s^t \sqrt{c(y_\tau)}\, X_k(y_\tau)\, dW_\tau^k. \tag{6.1}$$

We will write $Y_0(y) = c(y)X_0(y)$, $Y_k(y) = \sqrt{c(y)}\, X_k(y)$ $(k = 1, \dots, m)$, and $\tilde{D}(y) = D(y)/c(y)$. Note that by construction $\tilde{\mathscr{L}}^* \tilde{D}(y) = \mathscr{L}^* D(y)$, where $\tilde{\mathscr{L}}^*$ is the adjoint generator of (6.1). By our assumptions we can apply Theorem 5.1 to the Itô equation (6.1). Thus for all $y_s \in \mathbb{R}^n$, except in a set N with $\mu(N) = 0$, $y_t \to 0$ as $t \to \infty$ $\mathbb{P}$-a.s.

Now choose any $y_s \notin N$ and define

$$\beta_t = \int_s^t c(y_\tau)\, d\tau, \qquad \alpha_t = \inf\{s : \beta_s > t\}.$$

Note that α_τ is an $\mathcal{F}_s^t$-stopping time for each τ. We claim that $\beta_t \to \infty$ as $t \to \infty$ almost surely; indeed y_t almost surely spends an infinite amount of time in an arbitrarily small neighborhood of the origin, and as $c(x)^{-1}$ is locally bounded $c(x) \geq \delta > 0$ in any such neighborhood. Moreover, $\beta_t < \infty$ almost surely for any t as $c(x)$ is locally bounded, and hence $\alpha_t \to \infty$ as $t \to \infty$ almost surely. From [24, section V.26, p. 175], it follows that the time rescaled solution y_{α_t} is equivalent in law to the solution x_t of (4.1). But as almost all paths of the process y_t go to zero asymptotically and as $\alpha_t \to \infty$ almost surely, the result follows. ∎

COROLLARY 6.2. *Suppose that all the conditions of Theorem* 5.1 *are satisfied except the global Lipschitz condition. Suppose that additionally X_k, $k = 1, \dots, m$, satisfy $|X_k(x)| \leq C_k(1 + |x|^{p+1})$, $|\partial X_k(x)/\partial x^i| \leq C_k'(1 + |x|^p)$, and $|X_0(x)| \leq C_0(1 + |x|^{2p+1})$, $|\partial X_0(x)/\partial x^i| \leq C_0'(1+|x|^{2p})$ for some $p \geq 1$ and positive constants $C_k, C_k' < \infty$. If $(1+|x|^{2p})D(x)$ is integrable on $\{x \in \mathbb{R}^n : |x| > 1\}$, then Theorem* 5.1 *still holds.*

Proof. Let $c(x) = (1 + |x|^{2p})^{-1}$, and note that $c(x)$ is smooth, strictly positive and that $c(x)$ and $c(x)^{-1}$ are locally bounded. Let $Y_0(x) = c(x)X_0(x)$ and $Y_k(x) = \sqrt{c(x)}\, X_k(x)$, $k = 1, \dots, m$; as $\sqrt{c(x)}$ is smooth and X_k, $k = 0, \dots, m$, are $C^{2,\alpha}$, the coefficients Y_k are also $C^{2,\alpha}$. To show that Y_k are globally Lipschitz continuous it suffices to show that their first derivatives are bounded. Let us write $y_{k,i}(x) = |\partial Y_k/\partial x^i|(x)$. First consider the case $k = 0$. Then

$$y_{0,i}(x) \leq \frac{|\partial X_0(x)/\partial x^i|}{1 + |x|^{2p}} + \frac{2p|x|^{2p-2}|x^i|}{(1 + |x|^{2p})^2}\, |X_0(x)| \leq C_0' + C_0 \frac{2p|x|^{2p-2}|x^i|}{(1 + |x|^{2p})^2}(1 + |x|^{2p+1}),$$

which is bounded. Similarly, for $k \geq 1$

$$y_{k,i}(x) \leq C_k' \frac{1 + |x|^p}{\sqrt{1 + |x|^{2p}}} + C_k \frac{p|x|^{2p-2}|x^i|}{(1 + |x|^{2p})^{3/2}}(1 + |x|^{p+1})$$

is bounded. Finally, $D(x)/c(x) = (1 + |x|^{2p})D(x)$ is C^2, as $D(x)$ is C^2 and $c(x)^{-1}$ is smooth on $\mathbb{R}^n \backslash \{0\}$. Hence, provided $D(x)/c(x)$ is integrable on $\{x \in \mathbb{R}^n : |x| > 1\}$, we can apply Theorem 6.1. ∎

We now turn our attention to stochastic differential equations which evolve on an invariant set. The following notion of invariance is sufficient for our purposes.

DEFINITION 6.3. *A set K is called backward invariant with respect to the flow $\xi_{s,t}$ if $\xi_{s,t}^{-1}(K) \subset K$ almost surely for all $s < t$.*

We can now formulate the following result.

THEOREM 6.4. *Suppose the Itô equation* (4.1) *evolves on a backward invariant open set K with compact closure $\overline{K}$. Let X_k, $k = 0, \dots, m$, be $C^{2,\alpha}$ for some $\alpha > 0$ and let $X_k(0) = 0 \in \overline{K}$. Suppose there exists $D \in C^2(\overline{K}\backslash\{0\}, \mathbb{R}_+)$ such that $\mathscr{L}^* D(x) < 0$ for μ-almost all $x \in K$. Then for every initial time s and μ-a.e. initial state $x \in K$ the flow $\xi_{s,t}(x)$ tends to the origin as $t \to \infty$ $\mathbb{P}$-a.s.*

Proof. We will assume without loss of generality that X_k, $k = 0, \dots, m$, are globally Lipschitz continuous. Indeed, we can smoothly modify X_k outside K to have compact support without changing the properties of the flow in K, and as X_k are already locally Lipschitz continuous their modifications will be globally Lipschitz continuous.

Let $\varepsilon > 0$ and $Z = \{x \in K : |x| > \varepsilon\}$. D is integrable on Z, as D is bounded on Z and Z has compact closure. Let S_ℓ be an increasing sequence of open sets whose closure is strictly contained in $S = K\backslash\{0\}$, such that $\bigcup_\ell S_\ell = S$. Then there exists a $C^2(\mathbb{R}^n, \mathbb{R}_+)$-approximation D_ℓ of D for each ℓ, obtained by smoothly modifying D outside S_ℓ so that its support is contained in $\overline{K}$. That $\tau_\ell \to -\infty$ follows from backward invariance and from the one-to-one property of the flow. Hence all the conditions of Lemma 5.2 are satisfied, and we have

$$0 \leq \int_Z D(x)\,dx + \int_s^t \mathbb{E} \int_{\xi_{\sigma,t}^{-1}(Z)} \mathscr{L}^* D(x)\,dx\,d\sigma.$$

The remainder of the proof proceeds along the same lines as Theorem 5.1. □

7. Examples.

Example 1. Consider the Itô equation

$$\begin{aligned} dx_t &= (x_t^2 - 2x_t - z_t^2)\,dt + x_t\,dW_t, \\ dz_t &= 2z_t(x_t - 1)\,dt + z_t\,dW_t. \end{aligned} \tag{7.1}$$

Note that the line $z = 0$ is invariant under the flow of (7.1), where the solution $(x_t, 0)$ for an initial state $(x_0, 0)$ is given by

$$dx_t = (x_t^2 - 2x_t)\,dt + x_t\,dW_t.$$

This equation has an explicit solution (see also [2] for a detailed analysis of the dynamical behavior of this system):

$$x_t = \frac{x_0 e^{-2t} e^{W_t - \frac{1}{2}t}}{1 - x_0 \int_0^t e^{-2s} e^{W_s - \frac{1}{2}s} ds}.$$

Clearly $x_t(\omega)$, $\omega \in \Omega$, explodes in finite time if

$$x_0 > \left(\int_0^\infty e^{-2s} e^{W_s - \frac{1}{2}s} ds \right)^{-1} < \infty.$$

Hence the system (7.1) is certainly not globally stable.

Nonetheless, almost all points $(x_0, z_0) \in \mathbb{R}^2$ are attracted to the origin. To show this, apply Corollary 6.2 with

$$D(x, z) = \frac{1}{(x^2 + z^2)^2}, \qquad \mathscr{L}^* D(x, z) = -\frac{3}{(x^2 + z^2)^2} < 0.$$

Hence for almost every $(x_0, z_0) \in \mathbb{R}^2$, almost surely $(x_t, z_t) \to (0, 0)$ as $t \to \infty$.

Example 2. Consider the Itô equation

$$\begin{aligned} dx_t &= 12\,(2z_t - 1)x_t z_t\,dt - \tfrac{1}{2}x_t\,dt + (1 - 2z_t)x_t\,dW_t, \\ dy_t &= -\tfrac{1}{2}y_t\,dt + (1 - 2z_t)y_t\,dW_t, \\ dz_t &= -12 z_t x_t^2\,dt + 2(1 - z_t)z_t\,dW_t. \end{aligned} \tag{7.2}$$

Let $R_t = 2z_t - 2z_t^2 - x_t^2 - y_t^2$. By Itô's rule we have

$$dR_t = -4\,(1 - z_t)z_t R_t\,dt + 2\,(1 - 2z_t)R_t\,dW_t.$$

Evidently the ellipse $\{(x, y, z) \in \mathbb{R}^3 : 2z - 2z^2 - x^2 - y^2 = 0\}$ is invariant under (7.2). Local uniqueness of the solution implies that the interior of the ellipse is also invariant. Hence $K = \{(x, y, z) \in \mathbb{R}^3 : 2z - 2z^2 - x^2 - y^2 > 0\}$ is a (backward) invariant set for the system (7.2). Consider

$$D(x, y, z) = \frac{1}{z^2}, \qquad \mathscr{L}^* D(x, y, z) = -\frac{12x^2}{z^2} < 0 \quad \mu\text{-a.e.}$$

Hence by Theorem 6.4 for almost every $(x_0, y_0, z_0) \in K$, almost surely $(x_t, y_t, z_t) \to (0, 0, 0)$ as $t \to \infty$. Note that $(0, 0, 0)$ is certainly not globally stable: it is easily verified that any point with $x_0 = 0$ and $z_0 \neq 0$ is not attracted to $(0, 0, 0)$ almost surely, as in this case z_t has a constant nonzero mean.

Example 3. We consider again the system (7.2), but now we are interested in the behavior of points in the invariant set $K' = \{(x, y, z) \in \mathbb{R}^3 : 2z - 2z^2 - x^2 = 0,\ y = 0\}$. As K' is not an open set, we cannot apply Theorem 6.4 to study this case.

Define the transformation $(x, z) \mapsto p = x/z$. Note that p is the stereographic projection of $(x, y, z) \in K'$ which maps $(0, 0, 0) \mapsto \infty$. As the fixed point $(0, 0, 0)$ cannot be reached in finite time, we expect that the stereographic projection gives a well-defined dynamical system on $\mathbb{R}$. Using Itô's rule and $2z - 2z^2 - x^2 = 0$ we obtain the Itô equation

$$dp_t = \left(\frac{3}{2} + \frac{20}{2 + p^2}\right) p_t\,dt - p_t\,dW_t.$$

Note that this expression satisfies a global Lipschitz condition. Now consider

$$D(p) = \sqrt{2 + p^2}, \qquad \mathscr{L}^* D(p) = -\frac{42}{(2 + p^2)^{3/2}} < 0.$$

Hence by Theorem 5.4 for almost every $p_0 \in \mathbb{R}$, almost surely $p_t \to \infty$ as $t \to \infty$. This implies that the point $(x, y, z) = (0, 0, 0)$ is almost globally stable in K'.

8. Application to control synthesis. Consider an Itô equation of the form

$$x_t = x + \int_s^t (X_0(x_\tau) + u_\tau \tilde{X}_0(x_\tau))\,d\tau + \sum_{k=1}^m \int_s^t X_k(x_\tau)\,dW_\tau^k, \tag{8.1}$$

where $\tilde{X}_0$ and X_k, $k = 0, \ldots, m$, are $C^{2,\alpha}$ for some $\alpha > 0$, $X_k(0) = 0$ and u_t is a scalar control input. We consider instantaneous state feedback of the form $u_t = u(x_t)$ where $u(x)$ is $C^{2,\beta}$ for some $\beta > 0$ and $u(0) = 0$. Then by Theorem 5.1 or 6.4 or by

Corollary 6.2, $x_t \to 0$ as $t \to \infty$ almost surely for almost every x_0 if there exists a $D(x)$, with additional properties required by the appropriate theorem, such that

$$\begin{aligned} \mathscr{L}^* D(x) = \frac{1}{2} \sum_{k=1}^{m} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x^i \partial x^j} (X_k^i(x) X_k^j(x) D(x)) \\ - \sum_{i=1}^{n} \frac{\partial}{\partial x^i} (X_0^i(x) D(x) + \tilde{X}_0^i(x) u(x) D(x)) < 0 \qquad \mu\text{-a.e.} \end{aligned} \tag{8.2}$$

Note that this expression is affine in $D(x)$ and $u(x)D(x)$ and that the set of functions $(D(x), u(x)D(x))$ which satisfy (8.2) is convex. This fact has been used in the deterministic case by Prajna, Parrilo, and Rantzer [22] to search for "almost stabilizing" controllers for systems with polynomial vector fields using convex optimization. As the stochastic case enjoys the same convexity properties as the deterministic Theorem 3.2, this approach can also be applied to find stabilizing controllers for stochastic nonlinear systems. Note that that convex optimization cannot be used to search for globally stabilizing controllers using LaSalle's theorem [22] as LaSalle's convergence criterion [16, 6] is *not* convex.

The purpose of this section is to briefly outline the method of [22] for the synthesis of stabilizing controllers using convex optimization. We will also discuss a simple example.

Suppose that $\tilde{X}_0$ and X_k, $k = 0, \ldots, m$, are polynomial functions (the case of rational functions can be treated in a similar way). Consider $D(x)$ and $u(x)$ parametrized in the following way:

$$D(x) = \frac{a(x)}{b(x)^\gamma}, \qquad u(x) = \frac{c(x)}{a(x)}. \tag{8.3}$$

Here $b(x)$ is a nonnegative polynomial which vanishes only at the origin, $a(x)$ is a polynomial that is nonnegative in a neighborhood of the origin and is such that $u(x)$ is $C^{2,\beta}$, $c(x)$ is a polynomial that vanishes at the origin, and $\gamma > 0$ is a constant. The orders of the polynomials and γ can be chosen in such a way that $D(x)$ satisfies the integrability and growth requirements of Corollary 6.2. For fixed $b(x)$ and γ consider the expression

$$-b(x)^{\gamma+2} \mathscr{L}^* D(x) > 0 \qquad \mu\text{-a.e.} \tag{8.4}$$

with $D(x)$ and $u(x)$ given by (8.3) and $\mathscr{L}^*$ given by (8.2). Then (8.4) is a polynomial inequality that is linear in the polynomial coefficients of $a(x)$ and $c(x)$. Our goal is to formulate the search for these coefficients as a convex optimization problem.

Verifying whether (8.4) is satisfied comes down to testing nonnegativity of a polynomial (a nonnegative polynomial can only vanish on a finite set of points, and hence is positive μ-a.e.). This problem is known to be NP-hard in general; however, a powerful convex relaxation was suggested by Parrilo [20]. Instead of testing (8.4) directly we may ask whether the polynomial can be written as a *sum of squares*, i.e., whether $-b(x)^{\gamma+2} \mathscr{L}^* D(x) = \sum_i p_i(x)^2$ for a set of polynomials $p_i(x)$. The power of this relaxation comes from the fact that every sum of squares polynomial up to a specified order can be represented by a positive semidefinite matrix; hence the search for a sum of squares representation can be performed using semidefinite programming. As (8.4) is convex in $a(x)$ and $c(x)$ the following is a convex optimization problem:

Find polynomials $a(x)$, $c(x)$ such that $-b(x)^{\gamma+2} \mathscr{L}^* D(x)$ is a sum of squares.

This type of problem, known as a sum of squares program, can be solved in a highly efficient manner using the software SOSTOOLS [21]. We refer to [20, 21, 22] for further details on the computational technique.

Remark. Note that $a(0)$ and $c(0)$ depend only on the value of the constant coefficient of the polynomials $a(x)$ and $c(x)$. Thus $c(0) = 0$ can easily be enforced by fixing the constant coefficient of $c(x)$. To make sure $a(x)$ is nonnegative near the origin and $u(x)$ does not blow up, we can, for example, require $a(x)$ to be of the form $\lambda + p(x)$ with $\lambda > 0$ and $p(x)$ to be a sum of squares that vanishes at the origin.

Note that if the Itô equation (8.1) evolves on an invariant open set K with compact closure, then the sum of squares relaxation is overly restrictive. A related relaxation that guarantees only polynomial nonnegativity on K for the case that K is a semialgebraic set is considered in, e.g., [19].

Example. The following example is similar to an example in [22]. Consider the Itô equation

$$dx_t = (2x_t^3 + x_t^2 y_t - 6x_t y_t^2 + 5y_t^3)\,dt + (x_t^2 + y_t^2)\,dW_t,$$
$$dy_t = u_t\,dt - (x_t^2 + y_t^2)\,dW_t.$$

We choose $b(x, y) = x^2 + y^2$ and $\gamma = 2.5$. Using SOSTOOLS we find a solution with controller of order 3 and a constant $a(x)$. Note that these choices satisfy the integrability requirements of Corollary 6.2. We obtain a stabilizing controller

$$u(x) = -2.7\,x^3 + 4.6\,x^2 y - 6.7\,xy^2 - 3.4\,y^3,$$

where

$$-(x^2 + y^2)^{4.5} \mathscr{L}^*(x^2 + y^2)^{-2.5} = 0.35\,y^6 - 0.0015\,xy^5 + 0.6\,x^2y^4 + 0.0026\,x^3y^3 + 0.33\,x^4y^2 + 0.004\,x^5y + 0.13\,x^6$$

is a sum of squares polynomial.

Remark. A drawback of this method is that $b(x)$ and γ must be fixed at the outset. We have found that the method is very sensitive to the choice of $b(x)$ and γ even in the deterministic case; often an unfortunate choice will cause the search to be infeasible. Moreover it is not clear, even if there exists for polynomial $\tilde{X}_0$, X_k a rational u which almost globally stabilizes the system, that a rational D can always be found that satisfies $\mathscr{L}^* D < 0$. Nonetheless the method can be successful in cases where other methods fail, and as such could be a useful addition to the stochastic nonlinear control engineer's toolbox.

Acknowledgments. This work was performed in Hideo Mabuchi's group, and the author gratefully acknowledges his support. The author would like to thank Anders Rantzer, Luc Bouten, Stephen Prajna, Paige Randall, and especially Houman Owhadi for insightful discussions and comments. The author is particularly thankful to an anonymous referee for pointing out a gap in the proofs and for his careful reading of the manuscript, which has significantly improved the presentation.

REFERENCES

[1] L. Arnold, *The unfolding of dynamics in stochastic analysis*, Mat. Apl. Comput., 16 (1997), pp. 3–25.

[2] L. Arnold, *Random Dynamical Systems*, Springer-Verlag, Berlin, 1998.

[3] P. H. Baxendale, *Properties of stochastic flows of diffeomorphisms*, in Diffusion Processes and Related Problems in Analysis, Volume II, M. A. Pinsky and V. Wihstutz, eds., Progr. Probab. 27, Birkhäuser Boston, Boston, MA, 1992, pp. 3–35.

[4] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, UK, 1992.

[5] H. Deng and M. Krstić, *Stochastic nonlinear stabilization—part* I: *A backstepping design*, Syst. Control Lett., 32 (1997), pp. 143–150.

[6] H. Deng, M. Krstić, and R. J. Williams, *Stabilization of stochastic nonlinear systems driven by noise of unknown covariance*, IEEE Trans. Automat. Control, 46 (2001), pp. 1237–1253.

[7] P. Florchinger, *A universal formula for the stabilization of control stochastic differential equations*, Stochastic Anal. Appl., 11 (1993), pp. 155–162.

[8] P. Florchinger, *Lyapunov-like techniques for stochastic stability*, SIAM J. Control Optim., 33 (1995), pp. 1151–1169.

[9] P. Florchinger, *A stochastic Jurdjevic–Quinn theorem*, SIAM J. Control Optim., 41 (2002), pp. 83–88.

[10] R. Z. Has'minskiĭ, *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff, Germantown, MD, 1980.

[11] A. Isidori, *Nonlinear Control Systems*, 3rd ed., Springer-Verlag, Berlin, 1995.

[12] H. Kunita, *On the decomposition of solutions of stochastic differential equations*, in Stochastic Integrals, D. Williams, ed., Lecture Notes in Math. 851, Springer-Verlag, Berlin, 1981, pp. 213–255.

[13] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, in École d'Été de Probabilités de Saint-Flour XII, P. L. Hennequin, ed., Lecture Notes in Math. 1097, Springer-Verlag, Berlin, 1984, pp. 143–303.

[14] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, UK, 1990.

[15] H. J. Kushner, *Stochastic Stability and Control*, Academic Press, New York, 1967.

[16] H. J. Kushner, *Stochastic stability*, in Stability of Stochastic Dynamical Systems, R. F. Curtain, ed., Lecture Notes in Math. 294, Springer-Verlag, Berlin, 1972, pp. 97–124.

[17] M. Loève, *Probability Theory*, 3rd ed., Van Nostrand, Princeton, NJ, 1963.

[18] P. Monzón, *On necessary conditions for almost global stability*, IEEE Trans. Automat. Control, 48 (2003), pp. 631–634.

[19] A. Papachristodoulou and S. Prajna, *On the construction of Lyapunov functions using the sum of squares decomposition*, in Proceedings of the 41st IEEE Conference on Decision and Control, 2002, pp. 3482–3487.

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

[21] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, *SOSTOOLS: Sum of Squares Optimization Toolbox for MATLAB. User's Guide, Version* 2.00. Available from http://www.cds.caltech.edu/sostools, 2004.

[22] S. Prajna, P. A. Parrilo, and A. Rantzer, *Nonlinear control synthesis by convex optimization*, IEEE Trans. Automat. Control, 49 (2004), pp. 310–314.

[23] A. Rantzer, *A dual to Lyapunov's stability theorem*, Syst. Control Lett., 42 (2001), pp. 161–168.

[24] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Volume* 2: *Itô calculus*, 2nd ed., Cambridge University Press, Cambridge, UK, 2000.