A functional law of the iterated logarithm for weakly hypoelliptic diffusions at time zero[☆]

Marco Carfagnini[a], Juraj Földes[b,*], David P. Herzog[c]

[a] *Department of Mathematics, University of Connecticut, 341 Mansfield Road U1009, Storrs, CT 06269-1009, United States of America*

[b] *Department of Mathematics, University of Virginia, 322 Kerchof Hall, Charlottesville, VA 22904-4137, United States of America*

[c] *Department of Mathematics, Iowa State University, 474 Carver, 411 Morrill Rd., Ames, IA 50011-2104, United States of America*

Abstract

We study the almost sure behavior of solutions of stochastic differential equations (SDEs) as time goes to zero. Our main general result establishes a functional law of the iterated logarithm (LIL) that applies in the setting of SDEs with degenerate noise satisfying the *weak Hörmander condition* but not the *strong Hörmander condition*. That is, SDEs in which the drift terms must be used in order to conclude hypoellipticity. As a corollary of this result, we obtain the almost sure behavior as time goes to zero of a given direction in the equation, even if noise is not present explicitly in that direction. The techniques used to prove the main results are based on large deviations applied to a non-trivial rescaling of the original system. In concrete examples, we show how to find the proper rescaling to obtain the functional LIL. Furthermore, we apply the main results to the problem of identifying regular points for hypoelliptic diffusions. Consequently, we obtain a control-theoretic criteria for a given point to be regular for the process.
© 2022 Elsevier B.V. All rights reserved.

1. Introduction

The law of the iterated logarithm (LIL) for an i.i.d. sequence of random variables $X_1, X_2, \ldots, X_n, \ldots$ with mean zero and unit variance reads

$$\limsup_{n \to \infty} \frac{\pm S_n}{\sqrt{2n \log \log n}} = 1 \quad \text{a.s.}, \tag{1.1}$$

where $S_n := X_1 + X_2 + \cdots + X_n$ and log denotes the natural logarithm. The formula (1.1) was first established by Hartman and Whintner in 1941 [18] as a generalization of earlier works of Khinchin in 1924 [24] and Kolmogorov in 1929 [28]. Analogously, an LIL holds for a standard, real-valued Brownian motion W_t as $t \to \infty$ by replacing n by t and S_n by W_t in (1.1). Furthermore, one can use (1.1) for the Brownian motion at time infinity to obtain an LIL at time zero

$$\limsup_{t \to 0^+} \frac{\pm W_t}{\sqrt{2t \log \log t^{-1}}} = 1 \quad \text{a.s.} \tag{1.2}$$

by Brownian inversion.

Note that LILs provide an asymptotic window, for example $[-\sqrt{2t \log \log t^{-1}}, \sqrt{2t \log \log t^{-1}}]$ for the process W_t as $t \to 0^+$, complementing the usual central limit scaling. More precisely, it follows that the set of limit points of the scaled processes $S_n / \sqrt{2n \log \log n}$ as $n \to \infty$ or $W_t / \sqrt{2t \log \log t^{-1}}$ as $t \to 0^+$ is the interval $[-1, 1]$ [18]. A further generalization of this limit set analysis is due to Strassen [36], which for W_t at $t = 0$ establishes that, for almost every ω, the set of limit points (in the space of continuous paths $\mathscr{C}([0, 1]; \mathbf{R})$) of the family

$$Y_t^\varepsilon(\omega) := \frac{W_{\varepsilon t}(\omega)}{\sqrt{2\varepsilon \log \log \varepsilon^{-1}}}, \quad t \in [0, 1], \tag{1.3}$$

as $\varepsilon \to 0^+$ is the set of functions $f \in \mathscr{C}^0([0, 1]; \mathbf{R})$ with $f_0 = 0$ and $\int_0^1 |\dot{f}_s|^2 \, ds \leq 1$. Here,

$$\mathscr{C}^0([0, 1]; \mathbf{R}^k) := \{ f \in \mathscr{C}([0, 1]; \mathbf{R}^k) : \dot{f} \in L^2([0, 1]) \}.$$

Observe that, by the fundamental theorem of calculus and Jensen's inequality, the condition $\int_0^1 |\dot{f}_s|^2 \, ds \leq 1$ implies that $|f_1| \leq 1$. By choosing $f_s = \pm s$, the extremal values ± 1 of the pointwise limit set $[-1, 1]$ are attained, and, by setting $t = 1$ in (1.3), we obtain (1.2) as a corollary of Strassen's result.

The goal of this paper is to provide a framework for establishing Strassen-type LILs at time zero that applies in the setting of *weakly hypoellptic* diffusions. To formulate our results, we fix positive integers $d, k \in \mathbf{N}$, a non-empty, open set $U \subset \mathbf{R}^d$, and consider an Itô stochastic differential equation (SDE) on U of the form

$$\begin{cases} dx_t = \tilde{b}(x_t) \, dt + \tilde{\sigma}(x_t) \, dB_t, \\ x_0 = x \in U, \end{cases} \tag{1.4}$$

where B_t is a standard, k-dimensional Brownian motion defined on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$. In (1.4), we assume $\tilde{b} : U \to \mathbf{R}^d$ and $\tilde{\sigma} : U \to M_{d \times k}$ are *locally Lipschitz* on U; that is, Lipschitz continuous on every compact subset of U, while $M_{d \times k}$ denotes the set of $d \times k$ matrices with real-valued entries. Under these assumptions, the solution (1.4) can be defined pathwise until the first time $\tau(x_\cdot)$ at which x_t exits U. If $\tau(x_\cdot)$ is finite with positive probability, we fix a *death state* $\Delta \notin U$ and set $x_t = \Delta$ for $t \geq \tau(x_\cdot)$.

In this paper, we focus on degenerate diffusions, or equivalently on matrices $\tilde{\sigma}(x)$ that have rank strictly less than d at the initial condition x, so that the dynamics x_t defined by (1.4) is not trivially dominated at time zero in every direction by the process $x + \tilde{\sigma}(x)B_t$. Otherwise, a functional LIL can be readily obtained by rescaling Eq. (1.4) according to the LIL Brownian scaling (1.3) and passing to a (functional) limit using the theory of large deviations [2–4]. To see why, suppose for simplicity that $\tilde{\sigma}$ is a constant $d \times k$ matrix on U and all expressions below are well-defined. Then, the rescaled process

$$y_t^\varepsilon := x + \frac{x_{\varepsilon t} - x}{\sqrt{2\varepsilon \log \log \varepsilon^{-1}}}, \quad t \in [0, 1], \ \varepsilon > 0, \tag{1.5}$$

satisfies the integral equation

$$y_t^\varepsilon = x + \int_0^t \frac{\sqrt{\varepsilon}}{\sqrt{2 \log \log \varepsilon^{-1}}} b(x_{\varepsilon s}) \, ds + \frac{\tilde{\sigma} B_{\varepsilon t}}{\sqrt{2\varepsilon \log \log \varepsilon^{-1}}}. \tag{1.6}$$

Since the integral in (1.6) is small in ε, one then expects, and indeed it can be proved that, for almost all ω, the set of limit points of $y_\cdot^\varepsilon(\omega)$ in $\mathscr{C}([0, 1]; \mathbf{R}^k)$ as $\varepsilon \to 0$ is precisely the set of $g \in \mathscr{C}^0([0, 1]; \mathbf{R}^d)$ of the form

$$g_t = x + \tilde{\sigma} f_t \tag{1.7}$$

for some $f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k)$ with $f_0 = 0$ and $\int_0^1 |\dot{f}_s|^2 \, ds \leq 1$. Thus if $\text{rank}(\tilde{\sigma}(x)) = d$, then the asymptotic behavior of every component of y_t^ε, and consequently every component of x_t, can be readily characterized.

On the other hand, if $\text{rank}(\tilde{\sigma}(x)) < d$, then the same rescaling (1.5) is valid but less informative. In particular, for the directions that are in the range of $\tilde{\sigma}(x)$, the argument above gives correct asymptotic behavior. However, for the directions in $\mathbf{R}^d$ perpendicular to the range of $\tilde{\sigma}(x)$, the limiting trajectories are constant almost surely, meaning that the dynamics restricted to this subspace is finer and thus a different scaling is required. Even if one can heuristically estimate the right scale in the perpendicular directions, more work is needed to establish Strassen's law for the corresponding process.

In this direction, more general Strassen-type LILs for stochastic differential Eqs. (1.4) can be found in the pioneering works of Baldi [3] (at time infinity) and later Caramellino [10] (at time zero). Both works are limited by the permissible scaling transformations (called a *sequence of contractions*) applied to the diffusion to obtain the LIL. In particular, such scalings do not allow for *weakly hypoelliptic* diffusions; that is, diffusions that satisfy the *weak Hörmander* condition but not the *strong Hörmander* condition (see Section 3 for further information). Intuitively, weakly hypoelliptic diffusions are those in which the noise must spread through the drift so that the process does not live on a lower-dimensional manifold of U in small times. On the other hand, strongly hypoelliptic diffusions are those in which noise only spreads through the diffusion matrix, i.e. the drift is not needed, in order to reach all directions of the phase space. Weakly hypoelliptic diffusions arise in a number of natural settings, from finite-dimensional stochastic models in turbulence, see [5–7,13,14,16,22,34,39], to canonical models in statistical mechanics and machine learning, see [9,12,17,21,30,31], where local almost sure behavior at time zero is not understood precisely but is nevertheless important. Usually, one can estimate the behavior at time zero by finding the support of the process using control theory via the Support Theorems [37,38]. This usually gives the behavior of the process in small times with positive, usually very small, probability [20]. The LILs deduced here provide a refinement of the support of the diffusion in small times.

In this paper, we improve upon the main results in [3,10] by showing more general scaling transformations are permitted to obtain the functional LIL using the theory of large deviations. Specifically, our methods allow for different scalings in each component which need not be functions of the standard elliptic/Brownian scaling $\sqrt{\varepsilon \log \log \varepsilon^{-1}}$ as in [3,10], but can be general regularly varying functions in ε as $\varepsilon \to 0^+$. See Definition 2.1 in conjunction with Definition 2.2. In addition, in the setting of SDEs with additive noise, our transformations can in fact be time dependent (see Example 3.2), which is needed for degenerate diffusions with non-vanishing drift at the initial condition. There, because of the time dependence, the rescaled problem has a more complicated structure, which is why we restrict to the case of additive noise. Furthermore, we apply the main results to several nontrivial examples, e.g. the Iterated Kolmogorov equation in general dimensions and a stochastic Lorenz '96 model, capturing the a.s. behavior in each of these equations in each direction as time goes to zero started from the origin.

Our original motivation for LILs in the setting of hypoelliptic diffusions was to derive a probabilistic method for determining whether certain boundary points are *regular* or *irregular* (cf. Section 7). Such information is crucial when solving second-order linear hypoelliptic boundary-value problems, for example the Dirichlet or Poisson problems, in a domain [33]. Understanding when a particular point is regular or irregular is a long-standing open question and we refer the reader to the work of Kogoj [26,27] which employs analytical methods from PDEs to provide sufficient conditions for classical solvability of the Dirichlet problem and Harnack-type estimates. In this paper, we use the LIL to derive a control theoretic condition for a given boundary point to be regular or irregular (see Section 7).

The organization of this paper is as follows. In Section 2, we outline notation and state our main general results while in Section 3 we apply these results to concrete examples. We recommend the reader not familiar with the methods to first loosely read Section 3 to obtain some ideas of how to arrive at an LIL in the weakly hypoelliptic setting before reading Section 2. Section 4 outlines the needed results from the theory of large deviations to establish the main general results pertaining to the LILs, which are proved in Section 5 and Section 6. In Section 7, we derive our criteria for a point on the boundary to be regular or irregular. There, we also discuss applications of this criteria to the design of piecewise C^1 boundaries on which all points are regular.

2. Setting, notation and main results

The setup in this section is similar to that in Baldi [3] and Caramellino [10], but with several differences. First, our setting is slightly more general, which allows for more general transformations of the original process; that is, the process that satisfies (1.4). Second, we work primarily in the space of *explosive paths*, as opposed to continuous paths, defined below, and consequently, we employ notation and results from Azencott [2]. Note that the space of explosive paths was considered in [10] as well, though many of the proofs are carried out in the space of continuous paths. Importantly, the space of explosive paths allows us to work around issues of finite-time explosion in both the SDE and its limiting ODE. In particular, we can remove Assumption (A) (iii) of [10], although in many examples this condition is satisfied.

2.1. Laws of the iterated logarithm and large deviations

The crucial ingredient in the proof of the functional LIL is a change of coordinates, i.e. a rescaling, for the system (1.4), allowing one to reformulate the problem using the theory of large

deviations. Unless the diffusion (1.4) is uniformly elliptic at the initial state x, the change of coordinates varies depending on the dynamics. To see how to construct such a transformation, we provide concrete examples in Section 3.

Unlike in [3,10], instead of assuming a particular transformation, we simply associate to Eq. (1.4) a small parameter $\varepsilon_* > 0$ and family of processes $\{y^\varepsilon\}_{\varepsilon \in (0,\varepsilon_*]}$ satisfying an SDE of the form

$$\begin{cases} dy_t^\varepsilon = b_\varepsilon(y_t^\varepsilon)\,dt + \frac{1}{\sqrt{r(\varepsilon)}}\sigma_\varepsilon(y_t^\varepsilon)\,dB_t^\varepsilon, \\ y_0^\varepsilon = x, \end{cases} \tag{2.1}$$

where for each $\varepsilon \in (0, \varepsilon_*]$, we set $B_t^\varepsilon := \frac{1}{\sqrt{\varepsilon}}B_{\varepsilon t}$, where B_t is the standard, k-dimensional Brownian motion introduced in (1.4), and let

$$r(\varepsilon) := \log\log\varepsilon^{-1}. \tag{2.2}$$

We furthermore assume that the coefficients $b_\varepsilon, \sigma_\varepsilon$ satisfy the following conditions.

Assumption 1. There exist $\varepsilon_* > 0$, a non-empty open set $U^* \subset \mathbf{R}^d$ (not necessarily the same as U) and locally Lipschitz functions $b : U^* \to \mathbf{R}^d$ and $\sigma : U^* \to M_{d \times k}$ such that the following properties hold.

(i) For every $\varepsilon \in (0, \varepsilon_*]$, the coefficients $b_\varepsilon : U^* \to \mathbf{R}^d$, $\sigma_\varepsilon : U^* \to M_{d \times k}$ are locally Lipschitz on U^*.

(ii) For every compact $K \subset U^*$,

$$\lim_{\varepsilon \to 0^+}\sup_{y \in K}|b_\varepsilon(y) - b(y)| = 0,$$

$$\lim_{\varepsilon \to 0^+}\sup_{y \in K}\|\sigma_\varepsilon(y) - \sigma(y)\| = 0,$$

where $\|\cdot\|$ denotes a matrix norm.

In Fig. 1, we have provided a sketch of all of some of the objects introduced thus far, with a generic "mapping" Φ_ε relating the two processes, x_t and y_t^ε.

Remark 2.1. Note that b and σ in Assumption 1 are not the same as $\tilde{b}$ and $\tilde{\sigma}$ in Eq. (1.4). One should think of (2.1) as a rescaled version of (1.4), where b_ε and σ_ε depend on $\tilde{b}$ and $\tilde{\sigma}$.

Remark 2.2. When comparing the noise terms in (1.6) and in (2.1), note that the $\sqrt{\varepsilon}$ in the denominator in (1.6) is included in B_t^ε in (2.1) while the $\sqrt{2}$ is included in σ_ε.

2.2. The space of explosive trajectories $\mathscr{E}([0, t]; V)$

In order to treat possible finite-time blow-up of either y^ε in (2.1) or g_t (see (2.9) below), we work in the space of *explosive trajectories*. That is, for any $m \in \mathbf{N}$, $t \in (0, 1]$ and open $V \subset \mathbf{R}^m$, we again, offering a slight abuse of notation, fix a death state $\Delta \notin V$ and let $V \cup \{\Delta\}$ be the Alexandroff compactification of V. Then, $\mathscr{E}([0, t]; V)$ denotes the space of continuous mappings
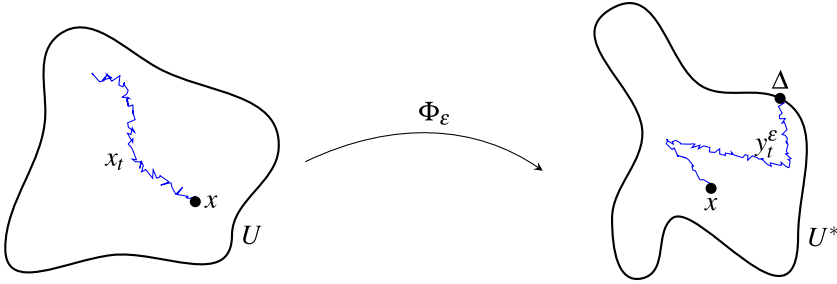
$$g : [0, t] \to V \cup \{\Delta\} \tag{2.3}$$

Fig. 1. A diagram representing a cartoon relationship between the process x_t and y_t^ε via the generic mapping Φ_ε, which should be thought of as a rescaling of Eq. (1.4).

such that if $g_{t_0} = \Delta$ for some $t_0 \in [0, t]$, then $g_s = \Delta$ for all $s \in [t_0, t]$. For any $x \in V$ and $t \in (0, 1]$, we define $\mathscr{E}_x([0, t]; V) := \{g \in \mathscr{E}([0, t]; V) : g_0 = x\}$. If $g \in \mathscr{E}([0, t]; V)$, let

$$\tau_t(g) = \inf\{s \in [0, t] : g_s = \Delta\} \tag{2.4}$$

denote the *time of explosion* of g, where we set $\inf \emptyset = \infty$. Define

$$\mathscr{E}^0([0, t]; V) = \{g \in \mathscr{E}([0, t]; V) : \dot{g} \in L^2([0, s]) \text{ for any } s < \tau_t(g), s \leq t\}, \tag{2.5}$$

where $\dot{g}$ denotes the time derivative of g. In other words, $g \in \mathscr{E}^0([0, t]; V)$ means that g belongs locally to the Sobolev space $H^1([0, \tau_t(g)) \cap [0, t]; V)$. We denote by $\mathscr{C}([0, t]; V)$ and $\mathscr{C}_x([0, t]; V)$, $x \in V$, $t \in (0, 1]$, respectively the space of continuous $g : [0, t] \to V$ and continuous $g : [0, t] \to V$ with $g_0 = x$. In particular, $g \in \mathscr{C}([0, t]; V)$ implies $\tau_t(g) = \infty$. Let

$$\mathscr{C}^0([0, t]; V) = \mathscr{C}([0, t]; V) \cap \mathscr{E}^0([0, t]; V) \tag{2.6}$$

and observe that $\mathscr{C}^0([0, t]; V)$ coincides with the Sobolev space $H^1([0, t], V)$.

It is important to equip the space $\mathscr{E}([0, t]; V)$ with a topology compatible with the topology of $\mathscr{C}([0, t]; V)$. As in [2], we define the closed sets in $\mathscr{E}([0, t]; V)$ by specifying convergent sequences. That is, we say that a sequence $g_n \in \mathscr{E}([0, t]; V)$ *converges* to $g \in \mathscr{E}([0, t]; V)$ as $n \to \infty$ if g_n converges uniformly to g on compact subsets of $[0, \tau_t(g)) \cap [0, t]$, or equivalently, if for any $s \in [0, \tau_t(g)), s \leq t$, there exists $N \in \mathbf{N}$ such that $\{g_n\}_{n \geq N} \subset \mathscr{C}([0, s]; V)$ and $g_n \to g$ as $n \to \infty$ in the space $\mathscr{C}([0, s]; V)$. The topologies on $\mathscr{C}^0([0, t]; V)$ and $\mathscr{E}^0([0, t]; V)$ are then induced by the topologies on, respectively, $\mathscr{C}([0, t]; V)$ and $\mathscr{E}([0, t]; V)$ intersected with $H^1_{\text{loc}}([0, \tau_t(g)) \cap [0, t]; V)$.

For $g, h \in \mathscr{E}([0, t]; V)$ and $s \leq t$, we define

$$d_s(g, h) = \begin{cases} \sup_{u \in [0, s]} |g_u - h_u|, & \text{if } s < \tau_t(g) \wedge \tau_t(h) \\ \infty & \text{otherwise.} \end{cases} \tag{2.7}$$

Clearly, $d_s(g, g) = 0$ if $s < \tau_t(g)$ and $d_s(g, h) = d_s(h, g)$, where both sides are either infinite or finite and equal. Also note that for any $f, g, h \in \mathscr{E}([0, t]; V)$ it follows that

$$d_s(f, g) \leq d_s(f, h) + d_s(h, g) \tag{2.8}$$

since the above reduces to the usual triangle inequality in $\mathscr{C}([0, s]; V)$ if $s < \tau_t(f) \wedge \tau_t(g) \wedge \tau_t(h)$. On the other hand if $s \geq \tau_t(f) \wedge \tau_t(g) \wedge \tau_t(h)$, then the righthand side is always infinite, so the inequality (2.8) is trivially satisfied.

Remark 2.3. Note that d_s is not a metric on $\mathscr{E}([0, t]; V)$ since, for example, $d_s(g, h) = \infty$ if g, h are identically equal to Δ. However, we do not need d to be a metric below.

For any $A \subset \mathscr{C}([0, t]; V)$ and any $g \in \mathscr{E}([0, t]; V)$ we set

$$d_t(g, A) = \inf\{d_t(g, h) \,:\, h \in A\}.$$

Remark 2.4. If $g \in \mathscr{E}([0, 1]; U)$, then its restriction to $[0, t]$, $t \in (0, 1]$, also belongs to $\mathscr{E}([0, t]; U)$. Below, we slightly abuse notation by denoting this restriction g as well.

2.3. Statement of the main general results

To employ a large deviation principle for y_t^ε solving (2.1), the following family of deterministic ODEs on the open set U^* is of particular importance:

$$\begin{cases} \dot{g}_t = b(g_t) + \sigma(g_t)\dot{f}_t \,, \\ g_0 = x \,, \end{cases} \tag{2.9}$$

where b, σ, U^* are as in Assumption 1, $f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k)$, and $x \in U^*$. By classical results for equations with locally Lipschitz coefficients (see, for example, [2, Proposition 2.3, p. 75]), if $\tau(g)$ denotes the first exit time of the solution of (2.9) from U^*, Assumption 1 ensures that for any $f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k)$ and $x \in U^*$, Eq. (2.9) has a unique solution $g \in \mathscr{E}^0([0, 1]; U^*)$ provided we set $g(t) = \Delta$ for any $t \geq \tau(g)$. We also assume that (2.9) is always satisfied for any $t \geq \tau(g)$ and we let $g = S_x(f)$. Here, $S : U^* \times \mathscr{C}^0([0, 1]; \mathbf{R}^k) \to \mathscr{E}^0([0, 1]; U^*)$ is the mapping given by

$$(x, f) \mapsto g = S_x(f). \tag{2.10}$$

In general, the mapping S is not continuous. However, for every $a \geq 0$ the restriction of S to $U^* \times \mathscr{C}_a$ where

$$\mathscr{C}_a := \{f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k) \,:\, \textstyle\int_0^1 |\dot{f}_t|^2 \, dt \leq a\}, \tag{2.11}$$

is continuous [2, Proposition 2.8, p.75]. Furthermore, by standard arguments (considering the equation for the difference of solutions and estimating using Gronwall's inequality) Assumption 1 implies that for any compact sets $K, L \subset U^*$ with $K \subset \text{interior}(L)$ and for any $a > 0$, there exist constants $T > 0, C > 0$ such that

(p1) For all $x \in K$ and all $f \in \mathscr{C}_a$, $g = S_x(f)$ has $\tau_1(g) > T$ and $g([0, T]) \subset L$.
(p2) For all $x, y \in K$, $f \in \mathscr{C}_a$, and $s \leq T$

$$d_s(S_x(f), S_y(f)) \leq e^{Cs}|x - y|.$$

(p3) If $\{f_n\} \subset \mathscr{C}_a$ converges in $H^1([0, 1])$ to $f \in \mathscr{C}_a$, then $S_x(f_n) \to S_x(f)$ in $\mathscr{C}^0([0, T]; U^*)$.

An essential notion for us is the *Cramer transform* $\lambda : \mathscr{E}([0, 1]; U^*) \to [0, \infty]$ defined by

$$\lambda(g) = \inf\{\tfrac{1}{2}\textstyle\int_0^1 |\dot{f}_t|^2 \, dt \,:\, f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k) \text{ and } S_{g_0}(f) = g\}, \tag{2.12}$$

where we set $\inf \emptyset = \infty$ and $\dot{f}_s = 0$ for any $s \geq \tau_1(g)$, or equivalently $S_x(f)(s) = g_s$ is satisfied for any f if $s \geq \tau_1(g)$. We remark that we can choose $\dot{f}_t = 0$ for $t > \tau_1(g)$ since in this time range Eq. (2.9) is satisfied by definition for any f. Note that our definition is equivalent to one introduced in [2, (4), Chapter IV] by [2, Proposition 2.10, Chapter IV]. Now,

[2, Proposition 2.10, Chapter IV] yields that if $\lambda(g) < \infty$, then the infimum in (2.12) is attained, $g \mapsto \lambda(g)$ is lower semi-continuous, and for any compact set $K \subset U^*$ and any $a \geq 0$, the set

$$\{g \in \mathscr{E}([0, 1]; U^*) : g_0 \in K, \lambda(g) \leq a\} \tag{2.13}$$

is compact in $\mathscr{E}([0, 1]; U^*)$. It thus follows that for every $x \in U^*$ the set

$$\mathscr{K}_x = \{g \in \mathscr{E}_x([0, 1]; \mathbf{R}^d) : \lambda(g) \leq 1\} \tag{2.14}$$

is also compact by choosing $K = \{x\}$ and $a = 1$ in (2.13).

By property (p1) and the relation (2.12), for any compact set $L \subset U^*$ and $x \in interior(L)$ there exists a time

$$t_* = t_*(x) \in (0, 1] \tag{2.15}$$

such that for all $g \in \mathscr{K}_x$, $g([0, t_*]) \subset interior(L)$. Let $\mathscr{K}_x(t_*)$ be the set of functions in $\mathscr{K}_x$ restricted to $\mathscr{E}([0, t_*]; U^*)$; that is, $\mathscr{K}_x(t_*) := \{g_{|[0,t_*]}, \ g \in \mathscr{K}_x\}$. Our goal is to show that, under Assumptions 1 and 2 (defined below) $\mathscr{K}_x(t_*)$ is the set of limit points, almost surely, of y^ε solving (2.1) as $\varepsilon \to 0$ (when restricted to $[0, t_*]$). Then, we show how one can relate y^ε back to the original process x_t solving (1.4) to obtain the desired functional LIL.

Assumption 2. Consider the compact set $L \subset U^*$ and constant $t_* > 0$ defined in (2.15). The process y^ε with $x \in interior(L)$ solving (2.1) satisfies the following properties:

(i) For every $\delta > 0$, there exists $c_0 \in (0, 1)$ such that for any $c \in (c_0, 1)$ there is a **P**-almost surely finite random variable $J = J(\omega, c) \in \mathbf{N}$ such that $j \geq J$ and $\varepsilon \in [c^{j+1}, c^j]$ implies $y_t^\varepsilon \in L$ for all $t \in [0, t_*]$ and

$$d_{t_*}(y^{c^j}, y^\varepsilon) < \delta. \tag{2.16}$$

(ii) For $\varepsilon_* > 0$ as in Assumption 1, the mapping $\varepsilon \mapsto y^\varepsilon : (0, \varepsilon_*) \to \mathscr{E}_x([0, t_*]; U^*)$ is continuous, **P**-almost surely.

Remark 2.5. Assumption 2 essentially allows one to reduce the proof of the main result (Theorem 2.6) to the countable sequence $\{y^{c^j}\}$ instead of $\{y^\varepsilon\}$.

Theorem 2.6. *Suppose that Assumptions 1 and 2 are both satisfied for some $\varepsilon_* > 0$ and non-empty open U^*. Fix compact $L \subset U^*$, $x \in interior(L)$ and $t_* = t_*(x) \in (0, 1]$ as in (2.15). Then, for **P**-almost every ω, we have the following conclusions:*

(i) *The set $\mathscr{Y}(\omega) := \{y^\varepsilon(\omega)\}_{\varepsilon \in (0, \varepsilon_*]}$ is relatively compact in $\mathscr{E}([0, t_*]; U^*)$.*

(ii) *$d_{t_*}(y^\varepsilon(\omega), \mathscr{K}_x(t_*)) \to 0$ as $\varepsilon \to 0$.*

(iii) *For every $h \in \mathscr{K}_x(t_*)$, $\{y^\varepsilon(\omega)\}_{\varepsilon \in (0, \varepsilon_*]}$ has a subsequence $\{y^{\varepsilon_j(\omega, h)}(\omega)\}_{j=1}^\infty$ with $\varepsilon_j(\omega, h) \downarrow 0$ as $j \to \infty$ such that*

$$d_{t_*}(y^{\varepsilon_j(\omega, h)}(\omega), h) \to 0 \quad as \quad j \to \infty.$$

In Section 3 we make heavy use of the following corollary of Theorem 2.6, which is a basic topological consequence of relative compactness and continuity.

Corollary 2.7. *Under the hypotheses of Theorem 2.6, let X be a Hausdorff topological space and suppose that $F : \mathscr{E}([0, t_*]; U^*) \to X$ is continuous. Then for **P**-almost every ω, $F(\mathscr{Y}(\omega))$ is relatively compact in X and $F(\mathscr{K}_x(t_*))$ is the limit set of $F(\mathscr{Y}(\omega))$ as $\varepsilon \to 0$.*

Proof of Corollary 2.7. The second conclusion follows immediately by continuity. Since $\mathscr{Y}(\omega)$ is relatively compact almost surely, then $\overline{\mathscr{Y}(\omega)}$ is compact, almost surely. For **P**-almost every ω, $F(\overline{\mathscr{Y}(\omega)})$ is compact since F is continuous, and therefore closed in the Hausdorff topological space X. Hence, $F(\mathscr{Y}(\omega)) \subset F(\overline{\mathscr{Y}(\omega)})$ implies $\overline{F(\mathscr{Y}(\omega))} \subset F(\overline{\mathscr{Y}(\omega)})$. Since a closed subset of a compact set is compact, the assertion follows. $\square$

In practice, it is relatively straightforward to check Assumption 1. However, Assumption 2, especially part (i), requires more work to validate. We next explore verifiable conditions under which Assumption 2 holds.

2.4. Sufficient conditions for Assumption 2

Heuristically, if Assumption 1 is satisfied and the system (2.1) arises from the original equation (1.4) under a *reasonable* change of coordinates, then Assumption 2 also holds. However, even more is true if the noise is additive. That is, if $\sigma_\varepsilon(x) \equiv \sigma_\varepsilon$ is a constant matrix for every $\varepsilon \in (0, \varepsilon_*]$ and Assumption 1 holds, then with an additional marginal continuity hypothesis, Assumption 2 holds for (2.1) independent of any relationship to the original equation (1.4).

To introduce an allowable change of coordinates that maps (1.4) to (2.1), we need further notation. We denote by $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $\beta = (\beta_1, \ldots, \beta_d)$ multiindices taking values in $\mathbf{R}^d$, and we write $\alpha \succeq \beta$ (respectively $\alpha \succ \beta$) if $\alpha_i \geq \beta_i$ (respectively $\alpha_i > \beta_i$) for all i. Note that this is equivalent to the partial ordering on the positive cone. When the context is clear, we use 0 and 1 to denote the multiindices $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$, respectively. For multiindices α, β, we let the product $\alpha\beta$ denote the multiindex $(\alpha_1\beta_1, \ldots, \alpha_d\beta_d)$ and if $\alpha \succ 0$, we define the multiindex $\alpha^{-1} = (\alpha_1^{-1}, \alpha_2^{-1}, \ldots, \alpha_d^{-1})$. Finally, for any multi-index α, we define

$$|\alpha| = \sqrt{\sum_i \alpha_i^2}. \tag{2.17}$$

Definition 2.1. Suppose that, for every multiindex $\alpha \succ 0$, $\Phi_\alpha : U \to U_\alpha$ is a C^2-bijection, where $U_\alpha \subset \mathbf{R}^d$ is open. We call $\{\Phi_\alpha\}_{\alpha \succ 0}$ *a family of weak contractions centered at* $x \in U$ if the following conditions are met:

(i) For every multiindex $\alpha \succ 0$, $\Phi_\alpha(x) = x$;
(ii) For all multiindices $\alpha \succeq \beta \succ 0$ we have

$$|\Phi_\alpha(y) - \Phi_\alpha(z)| \leq |\Phi_\beta(y) - \Phi_\beta(z)|$$

for all $y, z \in U$.
(iii) There exist $\kappa > 0$ and an open set $U^* \subset \mathbf{R}^d$ such that $x \in U^*$, $U^* \subset U_\alpha$ for each $|\alpha| < \kappa$, and for any compact set $K \subset U^*$ and any $\varepsilon > 0$ there exists $\delta > 0$ such that $|\alpha\beta^{-1} - 1| < \delta$ and $|\alpha| < \kappa$, $|\beta| < \kappa$ imply

$$|\Phi_\alpha \circ \Phi_\beta^{-1}(y) - y| < \varepsilon$$

for all $y \in K$.

Example 2.1. Let U be an open neighborhood of $0 \in \mathbf{R}^d$, and for any multiindex $\alpha \succ 0$, let $\Phi_\alpha^1 : U \to \mathbf{R}^d$ be given by

$$\Phi_\alpha^1(y) := (y_1\alpha_1^{-1}, y_2\alpha_2^{-1}, \ldots, y_d\alpha_d^{-1}). \tag{2.18}$$

Define the open set $U_\alpha = \Phi_\alpha^1(U)$. Then, $\{\Phi_\alpha^1\}_{\alpha>0}$ defines a family of weak contractions centered at $0 \in U$. Also, by *shifting* everything above, if U now denotes an open neighborhood of $x \in \mathbf{R}^d$, the family $\{\Phi_\alpha\}_{\alpha>0}$ defined by

$$\Phi_\alpha(y) = x - \Phi_\alpha^1(x) + \Phi_\alpha^1(y) = x + \Phi_\alpha^1(y - x), \quad y \in U,$$

is a family of weak contractions centered at x by setting $U_\alpha = \Phi_\alpha(U)$.

Remark 2.8. Compared with Baldi [3] and Caramellino [10], the index α in Definition 2.1 is allowed to be a multiindex rather than a positive real parameter. With some minor additional structure (see Definition 2.2), this affords more general transformations of (1.4) rather than functions of $\sqrt{\varepsilon \log \log \varepsilon^{-1}}$ alone (see Section 3). It is expected that a similar condition can be used to deduce LILs for diffusions at time infinity as well.

In order to specify a change of coordinates from x_t to y_t^ε, we need to impose assumptions on the dependence of the multiindex α on ε. Below, this dependence is determined using a heuristic scaling argument which in turn dictates the asymptotic behavior of x_t at time $t = 0$. The conditions outlined in the next definition are natural and satisfied in the examples in which we are interested.

Definition 2.2. Fix $\varepsilon_0 > 0$. We call $\psi : [0, \varepsilon_0] \to [0, \infty)^d$ an *asymptotic index* if all of the following conditions are met:

  (i) ψ is continuous.
 (ii) $\psi(0) = 0$ and $\psi(u) \prec \psi(v)$ as a multiindex for any $0 < u < v \leq \varepsilon_0$.
(iii) For any $\varepsilon > 0$, there exists $\delta \in (0, 1)$ such that for all $c \in (1 - \delta, 1)$ there exists $J \in \mathbf{N}$ such that for any $j \geq J$:

$$\delta_1, \delta_2 \in [c^{j+1}, c^j] \quad \text{implies} \quad |\psi(\delta_1)\psi(\delta_2)^{-1} - 1| < \varepsilon.$$

Example 2.2. For any positive integers $\ell, k \in \mathbf{N}$, and $\varepsilon_*(\ell, k) > 0$ small enough, let $\psi_{\ell,k} : [0, \varepsilon_*(\ell, k)] \to [0, \infty)$ be given by

$$\psi_{\ell,k}(\varepsilon) = \begin{cases} \sqrt{\varepsilon^\ell (\log\log\varepsilon^{-1})^k} & \text{for } \varepsilon \in (0, \varepsilon_*(\ell, k)], \\ 0 & \text{for } \varepsilon = 0. \end{cases} \tag{2.19}$$

Then, for any $(\ell_1, k_1), \ldots, (\ell_d, k_d) \in \mathbf{N} \times \mathbf{N}$ and $\varepsilon_* := \min\{\varepsilon_*(\ell_i, k_i) : i = 1, \ldots, d\} > 0$, $\psi : [0, \varepsilon_*] \to [0, \infty)^d$ defined as

$$\psi(\varepsilon) = (\psi_{\ell_1,k_1}(\varepsilon), \ldots, \psi_{\ell_d,k_d}(\varepsilon))$$

is an asymptotic index. Indeed, by standard calculations for $\varepsilon \in (0, \varepsilon_*]$ and $\varepsilon_* > 0$ small enough we have

$$\frac{d}{d\varepsilon}(\psi_{\ell,k}(\varepsilon))^2 = \varepsilon^{\ell-1}(\log\log\varepsilon^{-1})^{k-1}\left(\ell \log\log\varepsilon^{-1} - \frac{k}{\log\varepsilon^{-1}}\right) > 0, \tag{2.20}$$

which implies (ii). Also, by just proved monotonicity, if $c^{j+1} \leq \delta_2 \leq \delta_1 \leq c^j$

$$1 \leq \frac{\psi_{\ell,k}(\delta_1)}{\psi_{\ell,k}(\delta_2)} = \left(\frac{\delta_1}{\delta_2}\right)^{\frac{\ell}{2}}\left(\frac{\log\log\delta_1^{-1}}{\log\log\delta_2^{-1}}\right)^{\frac{k}{2}} \leq \frac{1}{c^{\frac{\ell}{2}}}\left(\frac{\log(j\log c^{-1})}{\log((j+1)\log c^{-1})}\right)^{\frac{k}{2}}. \tag{2.21}$$

Then, for any small $\varepsilon > 0$ there exists $\delta > 0$ such that $c^{\frac{\ell}{2}} > 1 - \frac{\varepsilon}{3}$ for any $c \in (1 - \delta, 1)$ and by choosing J large, we obtain that for any $j \geq J$

$$\left(\frac{\log(j \log c^{-1})}{\log((j+1) \log c^{-1})} \right)^{\frac{k}{2}} < 1 + \frac{\varepsilon}{3}. \tag{2.22}$$

Thus, if $\varepsilon \in (0, 1)$

$$0 \leq \frac{\psi_{\ell,k}(\delta_1)}{\psi_{\ell,k}(\delta_2)} - 1 \leq \frac{1 + \frac{\varepsilon}{3}}{1 - \frac{\varepsilon}{3}} - 1 < \varepsilon. \tag{2.23}$$

Hence property (iii) follows. The case $c^{j+1} \leq \delta_1 \leq \delta_2 \leq c^j$ follows analogously.

Example 2.3. In Example 2.2, instead of choosing each coordinate of the form (2.19), one could replace $\psi_{\ell,k}$ by a continuous, strictly increasing function $\varphi : [0, \varepsilon_*] \to [0, \infty)$ with $\varphi(0) = 0$ and φ regularly varying at 0.

Using the previous two concepts, we now connect Eqs. (1.4) and (2.1). Suppose that $\{\Phi_\alpha\}_{\alpha>0}$ is a family of weak contractions centered at $x \in U$ and $\psi : [0, \varepsilon_*] \to \mathbf{R}^d$ is an asymptotic index. We can choose ε_* small enough so that $|\psi(\varepsilon)| < \kappa$, where κ is given in Definition 2.1(iii). Observe that for any $\varepsilon \in (0, \varepsilon_*]$ and $t < \varepsilon^{-1}\tau(x.)$ (see (2.4)), the family of processes

$$y_t^\varepsilon := \Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) \tag{2.24}$$

satisfies, by Itô's formula, an SDE on U^* of the form (2.1) with $b_\varepsilon : U^* \to \mathbf{R}^d, \sigma_\varepsilon : U^* \to M_{d \times k}$ given by

$$b_\varepsilon(y) = \varepsilon \tilde{L} \, \Phi_{\psi(\varepsilon)}(\Phi_{\psi(\varepsilon)}^{-1}(y)), \tag{2.25}$$

$$\sigma_\varepsilon(y) = \sqrt{\varepsilon r(\varepsilon)} D \, \Phi_{\psi(\varepsilon)}(\Phi_{\psi(\varepsilon)}^{-1}(y)) \tilde{\sigma}(\Phi_{\psi(\varepsilon)}^{-1}(y)), \tag{2.26}$$

where

$$\tilde{L} = \sum_{i=1}^{d} \tilde{b}_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{d} (\tilde{\sigma}(x)\tilde{\sigma}(x)^T)_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \tag{2.27}$$

with $\tilde{b}$ and $\tilde{\sigma}$ as in (1.4), and U^* as in Assumption 1.

Lemma 2.9. *Suppose that $\{\Phi_\alpha\}_{\alpha>0}$ is a family of weak contractions centered at $x \in U$ and $\psi : [0, \varepsilon_*] \to \mathbf{R}^d$ is an asymptotic index and suppose $|\psi(\varepsilon_*)| < \kappa$, where κ is as in Definition 2.1(iii) for the appropriate U^*. If $b_\varepsilon, \sigma_\varepsilon$ given by (2.25)–(2.26) satisfy Assumption 1 with already fixed $\varepsilon_* > 0$ and U^*, then the family of processes $\{y^\varepsilon\}_{\varepsilon \in (0, \varepsilon_*]}$ given by (2.24) satisfies Assumption 2.*

Lemma 2.9 is proved in Section 6 along with the following corollary.

Corollary 2.10. *Consider the family of processes $\{y^\varepsilon\}_{\varepsilon \in (0, \varepsilon_*]}$ defined by relation (2.1) and suppose that $\sigma_\varepsilon(x) \equiv \sigma_\varepsilon$ is a family of constant $d \times k$ matrices. If ε_*, U^*, $b_\varepsilon, \sigma_\varepsilon$ satisfy Assumption 1 and for every $\delta \in (0, \varepsilon_*]$ we have that $\sigma_\varepsilon \to \sigma_\delta$ and $b_\varepsilon \to b_\delta$ as $\varepsilon \to \delta$ uniformly on compact subsets of U^*, then $\{y^\varepsilon\}_{\varepsilon \in (0, \varepsilon_*]}$ satisfies Assumption 2.*

Remark 2.11. Perhaps the most surprising consequence of Lemma 2.9 is that it is used to prove Corollary 2.10, even though there is no reference to an underlying mapping from x_t to y_t^ε.

3. Law of the iterated logarithm examples in the weakly hypoelliptic setting

3.1. Weakly hypoelliptic diffusions

Since Theorem 2.6 is a generalization of the main result in [10], all applications discussed there also follow from Theorem 2.6. We therefore refer the reader to [10] to see how to obtain a functional LIL at time zero for d-dimensional Brownian motion, elliptic SDEs as well as some iterated stochastic integrals. In particular, by the example in [10, Section 5], we see that Theorem 2.6 can also be applied when the noise is state-dependent. Here, we provide examples not covered by [10] which follow from Theorem 2.6.

All of the SDEs discussed below fall within the class of *weakly hypoelliptic* diffusions with additive noise; that is, each SDE below is of the form (1.4), where $\tilde{b} \in C^\infty(U)$ and $\tilde{\sigma}(x) \equiv \tilde{\sigma}$ is a $d \times k$ constant matrix such that the range of $\tilde{\sigma}$, denoted by $\mathscr{R}(\tilde{\sigma})$, has dimension strictly less than d, but *Hörmander's condition* is satisfied. That is, we say that the columns $\tilde{\sigma}^1, \tilde{\sigma}^2, \ldots, \tilde{\sigma}^k$ of $\tilde{\sigma}$ and $\tilde{\sigma}^0(x) := \tilde{b}(x)$, viewed as vector fields on U, satisfy *Hörmander's condition* on U if the list

$$
\begin{array}{ll}
\tilde{\sigma}^{\ell_1}(x) & \ell_1 = 1, 2, \ldots, k \\
[\tilde{\sigma}^{\ell_1}, \tilde{\sigma}^{\ell_2}](x) & \ell_1, \ell_2 = 0, 1, \ldots, k \\
[\tilde{\sigma}^{\ell_1}, [\tilde{\sigma}^{\ell_2}, \tilde{\sigma}^{\ell_3}]](x) & \ell_1, \ell_2, \ell_3 = 0, 1, \ldots, k \\
\qquad \vdots \qquad \vdots &
\end{array}
\tag{H}
$$

spans the tangent space at all points $x \in U$. In the above, $[X, Y]$ denotes the *commutator* of the vector fields X and Y; that is, if $X = (X_j(x))$ and $Y = (Y_j(x))$, then

$$
[X, Y](x) := \sum_{j=1}^d \sum_{i=1}^d \left\{ X_i(x) \frac{\partial Y_j(x)}{\partial x_i} - Y_i(x) \frac{\partial X_j(x)}{\partial x_i} \right\} \frac{\partial}{\partial x_j}.
$$

A celebrated theorem of Hörmander [23] shows that if condition (H) is satisfied, then the operators $\tilde{L}, \tilde{L}^*, \partial_t \pm \tilde{L}, \partial_t \pm \tilde{L}^*$ are all hypoelliptic on the respective domains $U, U, (0, \infty) \times U$, $(0, \infty) \times U$, where $\tilde{L}$ is as in (2.27) and $\tilde{L}^*$ denotes the formal $L^2(dx)$-adjoint of $\tilde{L}$. As a consequence, the distribution of the solution process x_t restricted to Borel subsets of U is absolutely continuous with respect to Lebesgue measure with transition density $q_t(x, y)$. Furthermore, $(t, x, y) \mapsto q_t(x, y) \in C^\infty((0, \infty) \times U \times U)$. If the process x_t exits U in finite time, then the law of x_t has a singular component on ∂U. However, this component is not present prior to exiting.

One interpretation of Hörmander's theorem is that condition (H) ensures that x_t is not locally restricted to a lower-dimensional submanifold of U. Indeed, the noise is either acting explicitly in directions $\tilde{\sigma}^1, \ldots, \tilde{\sigma}^k$, or it propagates implicitly through the drift term $\tilde{\sigma}^0 = \tilde{b}$, as represented by the commutators in (H). However, condition (H) does not guarantee that the process reaches all points in a small neighborhood in short times. For example, the process may be restricted to a cone (still satisfying Hörmander condition) as opposed to a ball [20, Example 3.4]. The goal of our LILs deduced below is to provide further insight into the a.s., small time behavior.

3.2. Examples

We now consider several concrete examples, starting with the so-called *Iterated Kolmogorov* diffusion.

Example 3.1 (*Iterated Kolmogorov*). Consider the following SDE on $\mathbf{R}^d$

$$
\begin{aligned}
dx_1 &= x_2 \, dt \,, \\
dx_2 &= x_3 \, dt \,, \\
&\vdots \quad \vdots \\
dx_{d-1} &= x_d \, dt \,, \\
dx_d &= dW_t \,,
\end{aligned}
\tag{3.1}
$$

where W_t is a standard one-dimensional Brownian motion defined on $(\Omega, \mathscr{F}, \mathbf{P})$ and the process

$$
x_t := (x_1(t), \ldots, x_d(t))
$$

has initial condition $x_0 = (0, 0, \ldots, 0)$. Our goal is to establish an LIL for the first coordinate x_1, which is given by the iterated time integral of the Brownian motion

$$
x_1(t) = \int_0^t \int_0^{t_2} \cdots \int_0^{t_d} dW_s \, dt_d \, dt_{d-1} \ldots dt_2.
$$

This was the one of the main goals of the paper [29] by Lachal. Historically, the case $d = 2$ in (3.1) is the first known example of a *hypoelliptic* diffusion, as discovered by Kolmogorov. For further information, see the discussion in the introduction of [23].

For any multiindex $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \succ 0$, define $\Phi_\alpha : \mathbf{R}^d \to \mathbf{R}^d$ by

$$
\Phi_\alpha(y) = (\alpha_1^{-1} y_1, \alpha_2^{-1} y_2, \ldots, \alpha_d^{-1} y_d)
$$

and note by Example 2.1, $\{\Phi_\alpha\}_{\alpha>0}$ is a family of weak contractions centered at the origin in $\mathbf{R}^d$. Furthermore, by Example 2.2, for $\varepsilon_* > 0$ small enough, $\psi : [0, \varepsilon_*] \to [0, \infty)^d$ given by

$$
\psi(\varepsilon) = \left(\sqrt{\varepsilon^{2d-1} \log\log \varepsilon^{-1}}, \sqrt{\varepsilon^{2d-3} \log\log \varepsilon^{-1}}, \ldots, \sqrt{\varepsilon^3 \log\log \varepsilon^{-1}}, \sqrt{\varepsilon \log\log \varepsilon^{-1}} \right)
$$

is an asymptotic index. To see that Assumption 1 is satisfied for the transformed diffusion y_t^ε defined by

$$
y_t^\varepsilon = \Phi_{\psi(\varepsilon)}(x_{\varepsilon t}),
\tag{3.2}
$$

we observe that, by construction, y_t^ε solves the following SDE

$$
\begin{aligned}
dy_1 &= y_2 \, dt \,, \\
&\vdots \quad \vdots \\
dy_{d-1} &= y_d \, dt \,, \\
dy_d &= \frac{dB_\varepsilon(t)}{\sqrt{r(\varepsilon)}} \,,
\end{aligned}
\tag{3.3}
$$

with $y_0 = (y_1(0), \ldots, y_d(0)) = 0$ and $B_\varepsilon(t) = \varepsilon^{-1/2} W_{\varepsilon t}$ being a standard Brownian motion on $\mathbf{R}$, and $r(\varepsilon) = \log\log \varepsilon^{-1}$. Since, $b_\varepsilon \equiv b$ is a linear function on $\mathbf{R}^d$ and $\sigma_\varepsilon \equiv \sigma$ is a constant matrix, Assumption 1 is satisfied. In addition, by Lemma 2.9, Assumption 2 holds true for y_t^ε. Furthermore, since the process is non-explosive, we may set $t_* = 1$ in the statements of Theorem 2.6 and Corollary 2.7.

For the projection $\pi : \mathbf{R}^d \to \mathbf{R}$ onto the first coordinate, consider the continuous map $F : \mathscr{E}([0, 1]; \mathbf{R}^d) \to \mathbf{R} \cup \{\Delta\}$ given by $F(g) = \pi g_1$ if $g_1 \in \mathbf{R}^d$ and $F(g) = \Delta$ if $g_1 = \Delta$,

where $g_1 = g_{t=1}$. Observe that $\mathbf{R} \cup \{\Delta\}$ is a Hausdorff space. Then, Corollary 2.7 implies that $\{F(y^\varepsilon)\}_{0 < \varepsilon < \varepsilon_*}$ is relatively compact in $\mathbf{R} \cup \{\Delta\}$. Furthermore, for $f \in \mathscr{C}^0([0, 1]; \mathbf{R})$ let

$$J_1(f) = \int_0^1 \int_0^{t_2} \cdots \int_0^{t_{d-1}} f_{t_d} \, dt_d \ldots dt_2 . \tag{3.4}$$

Then, the almost sure limit set of $(y_1^\varepsilon)_1$ as $\varepsilon \to 0$ is given by

$$F(\mathscr{K}_0) = \{J_1(f) : \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \leq 1\}.$$

The embedding $H^1([0, 1]) \hookrightarrow L^\infty([0, 1])$ implies that the constants

$$M = \sup\{J_1(f) : \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \leq 1\},$$

$$m = \inf\{J_1(f) : \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \leq 1\},$$

are finite and by choosing $f_s = \pm s$, one has $M, m \neq 0$, and therefore almost surely

$$\limsup_{\varepsilon \to 0} \frac{x_1(\varepsilon)}{\sqrt{\varepsilon^{2d-1} \log \log \varepsilon^{-1}}} = M > 0,$$

$$\liminf_{\varepsilon \to 0} \frac{x_1(\varepsilon)}{\sqrt{\varepsilon^{2d-1} \log \log \varepsilon^{-1}}} = m.$$

Moreover, $m = -M$ since J is an odd function of f.

Example 3.2. This example shows the utility of Corollary 2.10. Consider again the same system as in Example 3.1, but with $d = 2$ and the process $x_t = (x_1(t), x_2(t))$ starting from a general initial condition $x_0 = (x_1(0), x_2(0)) \in \mathbf{R}^2$. Equivalently, we can consider the process solving

$$dx_1 = (x_2 + c) \, dt, \qquad dx_2 = dW_t \tag{3.5}$$

for some c, with $x_0 = 0$. Again, our goal is to obtain a LIL for the first coordinate x_1. If we define

$$y_1^\varepsilon(t) = \frac{x_1(\varepsilon t) - x_1(0) - t\varepsilon x_2(0)}{\sqrt{\varepsilon^3 \log \log \varepsilon^{-1}}} + x_1(0) + t x_2(0), \tag{3.6}$$

$$y_2^\varepsilon(t) = \frac{x_2(t\varepsilon) - x_2(0)}{\sqrt{\varepsilon \log \log \varepsilon^{-1}}} + x_2(0),$$

then $y_t^\varepsilon := (y_1^\varepsilon(t), y_2^\varepsilon(t))$ satisfies

$$dy_1^\varepsilon = y_2^\varepsilon \, dt,$$

$$dy_2^\varepsilon = \tfrac{1}{\sqrt{r(\varepsilon)}} dW_t^\varepsilon,$$

where $(y_1^\varepsilon(0), y_2^\varepsilon(0)) = (x_1(0), x_2(0))$ and $W_t^\varepsilon = \tfrac{1}{\sqrt{\varepsilon}} W_{\varepsilon t}$. Note that Assumption 1 is clearly satisfied for the process y_t^ε. Due to the explicit dependence on time, the mapping $x \mapsto y^\varepsilon$ in (3.6) does not satisfy the assumptions of Lemma 2.9 if $x_2(0) \neq 0$. However, Corollary 2.10 ensures that Assumption 2 is satisfied. Also, the solution of the associated deterministic problem

$$\dot{y}_1 = y_2, \qquad \dot{y}_2 = \dot{f} \tag{3.7}$$

is given by

$$y_2(t) = x_2(0) + f(t), \qquad y_1(t) = x_1(0) + x_2(0)t + \int_0^t f(s)ds, \qquad (3.8)$$

where $f(0) = 0$. If $J_1 : \mathscr{C}^0([0, 1]; \mathbf{R}) \to \mathbf{R}$ is given by

$$J_1(f) = \int_0^1 f_s \, ds \qquad \text{and} \qquad M = \sup\{J_1(f) : \tfrac{1}{2}\int_0^1 (\dot{f}_s)^2 \, ds \le 1\}, \qquad (3.9)$$

then as in Example 3.1, one has $0 < M < \infty$ and, **P**-almost surely,

$$\limsup_{\varepsilon \to 0} \left(\frac{x_1(\varepsilon) - x_1(0) - \varepsilon x_2(0)}{\sqrt{\varepsilon^3 \log \log \varepsilon^{-1}}} \right) = M, \qquad (3.10)$$

$$\liminf_{\varepsilon \to 0} \left(\frac{x_1(\varepsilon) - x_1(0) - \varepsilon x_2(0)}{\sqrt{\varepsilon^3 \log \log \varepsilon^{-1}}} \right) = -M. \qquad (3.11)$$

Note that (3.10) implies that for any $\delta > 0$ there exists a (random) sequence $\varepsilon_n = \varepsilon_n(\omega) > 0$ such that $\varepsilon_n \to 0$ as $n \to \infty$ and

$$x_1(\varepsilon_n) \ge x_1(0) + \varepsilon_n x_2(0) + \sqrt{\varepsilon_n^3 \log \log \varepsilon_n^{-1}} (M - \delta).$$

Similarly, using (3.11); for any $\delta > 0$ there exists $\bar{\varepsilon}_n = \bar{\varepsilon}_n(\omega) > 0$ such that $\bar{\varepsilon}_n \to 0$ as $n \to \infty$ and

$$x_1(\bar{\varepsilon}_n) \le x_1(0) + \bar{\varepsilon}_n x_2(0) + \sqrt{\bar{\varepsilon}_n^3 \log \log \bar{\varepsilon}_n^{-1}} (-M + \delta).$$

Now suppose $x_2(0) > 0$ as the case $x_2(0) < 0$ is treated similarly. Since $\sqrt{\varepsilon^3 \log \log \varepsilon^{-1}} \ll \varepsilon$ for small ε, we obtain that $x_1(t) \ge x_1(0)$ for all small times t. Intuitively, if we rewrite our system as (3.5), then since $x_0 = 0$ one has $x_2 < c = x_2(0)$ for all small times. Thus, x_1 is increasing for small times as our analysis shows. Observe that if there was a noise in the x_1 coordinate, then it would change sign on the time scale $\sqrt{\varepsilon} \gg \varepsilon$, and therefore $x_1(t) - x_1(0)$ would change sign as well.

Example 3.3. As our next example, we consider the following diffusion on $\mathbf{R}^2$

$$\begin{aligned} dx_1 &= (x_1^2 - x_2^2)\, dt, \\ dx_2 &= 2x_1 x_2 \, dt + dB_t, \end{aligned} \qquad (3.12)$$

where $x_t = (x_1(t), x_2(t))$ has initial condition $x_0 = (x_1(0), x_2(0)) = (0, 0)$ and B_t is a standard, one-dimensional Brownian motion on $(\Omega, \mathscr{F}, \mathbf{P})$. This particular diffusion has been extensively studied (see [1,8,16,19]). In particular, one of the main results in these works is that the diffusion defined by (3.12) is non-explosive for all initial conditions in $\mathbf{R}^2$. This is true despite the fact that the associated deterministic dynamics (obtained by deleting dB_t from (3.12)) explodes in finite time when started from $(s, 0)$ with $s > 0$. Here, we study the behavior at time zero of the first coordinate, x_1 (the behavior of the second coordinate is trivial).

First, define the family $\{\Phi_\alpha\}_{\alpha>0}$ of weak contractions centered at $(0, 0) \in \mathbf{R}^2$ by

$$\Phi_\alpha(y_1, y_2) = \left(y_1 \alpha_1^{-1}, y_2 \alpha_2^{-1} \right).$$

We also define, for $\varepsilon_* > 0$ small enough, the asymptotic index $\psi : [0, \varepsilon_*] \to [0, \infty)^2$ as

$$\psi(\varepsilon) = (\varepsilon^2 \log \log \varepsilon^{-1}, \sqrt{\varepsilon \log \log \varepsilon^{-1}}).$$

Then, by (2.25) and (2.26) the process

$$y_t^\varepsilon = \Phi_{\psi(\varepsilon)}(x_{\varepsilon t})$$

satisfies the SDE

$$dy_1 = \varepsilon^3 \log \log \varepsilon^{-1} y_1^2 \, dt - y_2^2 \, dt,$$
$$dy_2 = 2\varepsilon^3 \log \log \varepsilon^{-1} y_1 y_2 \, dt + \frac{1}{\sqrt{r(\varepsilon)}} \, dB_\varepsilon,$$

where $(y_1(0), y_2(0)) = (0, 0)$ and B_ε is a standard, one-dimensional Brownian motion. Note that

$$b_\varepsilon(y_1, y_2) = (\varepsilon^3 \log \log \varepsilon^{-1} y_1^2 - y_2^2, \ 2\varepsilon^3 \log \log \varepsilon^{-1} y_1 y_2)$$

and $\sigma_\varepsilon \equiv \sigma$ is a constant matrix. Since

$$b_\varepsilon(y_1, y_2) \to b(y_1, y_2) = (-y_2^2, 0)$$

uniformly on compact subsets of $\mathbf{R}^2$ as $\varepsilon \to 0$, it follows that Assumption 1 is satisfied and by Lemma 2.9, Assumption 2 is also satisfied. Thus, Theorem 2.6 and Corollary 2.7 apply for any sufficiently small $t_* \in (0, 1]$. However, because the limiting ODE

$$\dot{y}_1 = -y_2^2,$$
$$\dot{y}_2 = \dot{f},$$

is well-defined for all times $t \in [0, 1]$ for any $f \in \mathscr{C}^0([0, 1]; \mathbf{R}^2)$ (simply integrate it), we may take $t_* = 1$.

With a slight abuse of notation, we let $\pi : \mathbf{R}^2 \to \mathbf{R}$ denote the projection onto the first coordinate, and $F : \mathscr{E}([0, 1]; \mathbf{R}^2) \to \mathbf{R} \cup \{\Delta\}$ be such that $F(g) = \pi g_1$ if $g_1 \in \mathbf{R}^2$ and Δ otherwise (cf. Example 3.1). Corollary 2.7 implies that $\{F(y^\varepsilon)\}_{\varepsilon \in (0, \varepsilon_*]}$ is relatively compact in $\mathbf{R} \cup \{\Delta\}$. Moreover, if $J_2 : \mathscr{C}_0([0, 1]; \mathbf{R}) \to (-\infty, 0]$ is given by

$$J_2(f) = -\int_0^1 f_s^2 \, ds,$$

then the a.s. limit set of $F(y^\varepsilon)$ as $\varepsilon \to 0$ is given by

$$F(\mathscr{K}_0) = \{J_2(f) : \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \le 1\}.$$

Note that

$$-M := \inf\{J_2(f) : \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \le 1\} \in (-\infty, 0),$$

since $H^1([0, 1]) \hookrightarrow L^\infty([0, 1])$, and consequently almost surely

$$\liminf_{\varepsilon \to 0} \frac{x_1(\varepsilon)}{\varepsilon^3 \log \log \varepsilon^{-1}} = -M, \qquad \limsup_{\varepsilon \to 0} \frac{x_1(\varepsilon)}{\varepsilon^3 \log \log \varepsilon^{-1}} = 0.$$

Example 3.4. Next, we consider the following Lorenz 96 model with $d = 5$

$$\begin{aligned}
dx_1 &= (x_2 - x_4)x_5 \, dt - x_1 \, dt + dB_1, \\
dx_2 &= (x_3 - x_5)x_1 \, dt - x_2 \, dt + dB_2, \\
dx_3 &= (x_4 - x_1)x_2 \, dt - x_3 \, dt, \\
dx_4 &= (x_5 - x_2)x_3 \, dt - x_4 \, dt, \\
dx_5 &= (x_1 - x_3)x_4 \, dt - x_5 \, dt,
\end{aligned} \tag{3.13}$$

where the process $x_t = (x_1(t), x_2(t), \ldots, x_5(t))$ above is assumed to evolve on $\mathbf{R}^5$ starting initially at $x_0 = 0$, and B_i, $i = 1, 2$, are independent, standard Brownian motions defined on $(\Omega, \mathscr{F}, \mathbf{P})$. Using nearly identical computations, one can also treat the general Lorenz 96 model in d dimensions, but for clarity we prefer the concrete scenario above. Our interest stems from the fact that the nonlinearities mimic pairwise interactions in the Navier–Stokes equation. Here we will analyze the small-time behavior of x_5.

In this example, again the family $\{\Phi_\alpha\}_{\alpha>0}$ of weak contractions centered at $0 \in \mathbf{R}^5$ is the same as above

$$\Phi_\alpha(y) = (y_1 \alpha_1^{-1}, y_2 \alpha_2^{-1}, \ldots, y_5 \alpha_5^{-1}).$$

For $\varepsilon_* > 0$ small enough, we define $\psi : [0, \varepsilon] \to [0, \infty)^5$ by

$$\psi(\varepsilon) = (\sqrt{\varepsilon \log \log \varepsilon^{-1}}, \sqrt{\varepsilon \log \log \varepsilon^{-1}}, \varepsilon^2 \log \log \varepsilon^{-1}, \sqrt{\varepsilon^7 (\log \log \varepsilon)^3}, \varepsilon^5 (\log \log \varepsilon^{-1})^2)$$

and note that ψ is an asymptotic index for (3.13).

Let y^ε be given by $y_t^\varepsilon = \Phi_{\psi(\varepsilon)}(x_{\varepsilon t})$ and note that y^ε solves the following SDE

$$dy_1 = b_{1,\varepsilon}(y_1, y_2, y_4, y_5) \, dt + \frac{1}{\sqrt{r(\varepsilon)}} \, dB_{1,\varepsilon} \,,$$

$$dy_2 = b_{2,\varepsilon}(y_1, y_2, y_3, y_5) + \frac{1}{\sqrt{r(\varepsilon)}} \, dB_{2,\varepsilon} \,,$$

$$dy_3 = -y_1 y_2 \, dt + b_{3,\varepsilon}(y_2, y_3, y_4) \, dt \,,$$

$$dy_4 = -y_2 y_3 \, dt + b_{4,\varepsilon}(y_3, y_4, y_5) \, dt \,,$$

$$dy_5 = y_1 y_4 \, dt + b_{5,\varepsilon}(y_3, y_4, y_5) \, dt \,,$$

where $b_{i,\varepsilon} \to 0$, $i = 1, 2, 3, 4, 5$, as $\varepsilon \to 0$ uniformly on compact subsets in $\mathbf{R}^5$. Furthermore, one can check that the $b_{i,\varepsilon}$ are locally Lipschitz on $\mathbf{R}^5$, so that Assumption 1 is satisfied. Thus Theorem 2.6 and Corollary 2.7 both apply. In this case, the limiting ODE is

$$\dot{y}_1 = \dot{f}_1 \,, \qquad \dot{y}_2 = \dot{f}_2 \,, \qquad \dot{y}_3 = -y_1 y_2 \,, \qquad \dot{y}_4 = -y_2 y_3 \,, \qquad \dot{y}_5 = y_1 y_4 \,, \qquad (3.14)$$

with $y(0) = 0$.

Note that we can solve (3.14) explicitly and take $t_* = 1$ again. Indeed, let $\pi_5 : \mathbf{R}^5 \to \mathbf{R}$ denote the projection onto the fifth coordinate and, slightly abusing notation again, define $F : \mathscr{E}([0, 1]; \mathbf{R}^5) \to \mathbf{R} \cup \{\Delta\}$ by $F(g) = \pi_5 g_1$ if $g_1 \in \mathbf{R}^5$ and $F(g) = \Delta$ otherwise. Applying Corollary 2.7, we note that $\{F(y^\varepsilon)\}_{0<\varepsilon\le\varepsilon_*}$ is relatively compact in $\mathbf{R} \cup \{\Delta\}$. Furthermore, define $J_3 : \mathscr{C}_0([0, 1]; \mathbf{R}^2) \to \mathbf{R}$ as y_5 in (3.14)

$$J_3(f_1, f_2) = \int_0^1 f_1(t) \left(\int_0^t f_2(s) \left(\int_0^s f_1(r) f_2(r) \, dr \right) ds \right) dt,$$

where $f_1(0) = f_2(0) = 0$. Let

$$M = \sup\{J_3(f_1, f_2) \,:\, \tfrac{1}{2} \int_0^1 |(\dot{f}_1, \dot{f}_2)|^2 \le 1\}$$

$$m = \inf\{J_3(f_1, f_2) \,:\, \tfrac{1}{2} \int_0^1 |(\dot{f}_1, \dot{f}_2)|^2 \le 1\} \,.$$

Setting $\dot{f} = \dot{g} = 1$ we see that J_3 clearly attains positive values. However, seeing that J_3 can realize negative is not immediately obvious. Nevertheless, by choosing $f_1(t) = \sin(5t)$ and

$f_2(t) = \sin(t)$ one has $J_3(f_1, f_2) \approx -0.00605$. Thus by a proper rescaling to guarantee the constraint on $\dot{f}_1$ and $\dot{f}_2$, we obtain then (almost surely)

$$\limsup_{\varepsilon \to 0} \frac{x_5(\varepsilon)}{\varepsilon^5 (\log \log \varepsilon^{-1})^2} = M,$$

$$\liminf_{\varepsilon \to 0} \frac{x_5(\varepsilon)}{\varepsilon^5 (\log \log \varepsilon^{-1})^2} = m,$$

for some $M > 0$ and $m < 0$.

Remark 3.1. Because the LIL at time zero is local phenomena, there is nothing important about the SDEs above being defined on all of $\mathbf{R}^d$. One can consider the same equations (or different ones) defined in a neighborhood U of the initial condition x, with scalings taking the same forms. One then applies the general results of this paper, but with U^* in place of all of $\mathbf{R}^d$. See also Example 2.1.

Remark 3.2. Let W_t be a standard, real-valued Brownian motion defined on $(\Omega, \mathscr{F}, \mathbf{P})$. Chung's LIL [11] for W_t at time zero states that, $\mathbf{P}$-almost surely,

$$\liminf_{\varepsilon \to 0} \left\{ \sqrt{\frac{\log \log \varepsilon^{-1}}{\varepsilon}} \max_{0 \leqslant t \leqslant \varepsilon} |W_t| \right\} = \frac{\pi}{\sqrt{8}}. \tag{3.15}$$

A natural question is whether Theorem 2.6 and Corollary 2.7 can be used to deduce a similar results for SDEs in the examples above? The short answer is that the results do apply, but do not obviously capture the precise asymptotic behavior that one would expect for the $\liminf$ of the running maximum. To see why in more detail, we have provided the next example.

Example 3.5. Consider (3.1) for $d = 2$ and define the maximum process for the first coordinate $x_1^*(t) := \sup_{s \in [0,t]} |x_1(s)|$. Using the scaling property [25, equation (2.1)]

$$x_1^*(t) \overset{d}{=} t^{3/2} x_1^*(1) \qquad \forall t \geq 0$$

along with inversion, one of the main results in [25] states that, $\mathbf{P}$-almost surely,

$$\liminf_{\varepsilon \to 0} \frac{x_1^*(\varepsilon)}{\phi(\varepsilon)} = c, \qquad \phi(\varepsilon) := \frac{\varepsilon^{3/2}}{(\log \log \varepsilon^{-1})^{3/2}}, \tag{3.16}$$

for some deterministic constant $c \in (0, \infty)$. To investigate properties of x^* in our framework, let $\pi : \mathbf{R}^2 \to \mathbf{R}$ be the projection onto the first coordinate, and let $F : \mathscr{E}([0, 1]; \mathbf{R}^2) \to \mathbf{R} \cup \{\Delta\}$ be a continuous mapping given by $F(g) = \sup_{s \in [0,1]} |\pi g_s|$ if $g_1 \in \mathbf{R}^2$ and $F(g) = \Delta$ if $g_1 = \Delta$. Following Example 3.1, if $J : \mathscr{C}^0([0, 1]; \mathbf{R}) \to \mathbf{R}$ is given by

$$J(f) = \sup_{0 \leq t \leq 1} \left| \int_0^t f_s \, ds \right|, \tag{3.17}$$

then the almost sure limit set of $F(y^\varepsilon)$ as $\varepsilon \to 0$ is

$$F(\mathscr{K}_0) = \{ J(f) \, : \, \tfrac{1}{2} \int_0^1 (\dot{f}_s)^2 \, ds \leq 1 \}.$$

Here, we recall that y^ε is as in (3.2) with $d = 2$. From this we deduce that, almost surely,

$$\limsup_{\varepsilon \to 0} \frac{x_1^*(\varepsilon)}{\varepsilon^{3/2} \sqrt{\log \log \varepsilon^{-1}}} = M, \tag{3.18}$$

$$\liminf_{\varepsilon \to 0} \frac{x_1^*(\varepsilon)}{\varepsilon^{3/2}\sqrt{\log\log\varepsilon^{-1}}} = 0, \tag{3.19}$$

for some constant $M \in (0, \infty)$. Note that while relation (3.18) provides precise asymptotics, the second relation (3.19), though consistent with (3.16), does not establish (3.16). This is because the lim inf of x^* acts on a smaller scale which is not captured by our scaling.

4. Large deviations

In this section, we outline two key results that are used to prove Theorem 2.6. Both results follow almost immediately from the existing literature. Here, we only provide slight adjustments, if needed, to connect with our setting.

The first result stated below is a consequence of [4, Theorem 1.1], which is an improvement of [2, 2.4 Théorème, Chapitre III]. The only difference here is that we are not assuming [4, (A.3)], but our proof follows nearly identical localization procedure for an open set U^* as opposed to $\mathbf{R}^d$ in [4].

Recall that $(x, f) \mapsto S_x(f)$ is the solution operator of (2.9), $r(\varepsilon) = \log\log\varepsilon^{-1}$, and the distance-like function d_t is defined in (2.7). We also recall that for every $\varepsilon > 0$, $B_t^\varepsilon = \varepsilon^{-1/2}B_{\varepsilon t}$, where B_t denotes a standard Brownian motion on $\mathbf{R}^k$, and that $\varepsilon_* > 0$ and U^* were fixed in Assumption 1.

Theorem 4.1. *Suppose that Assumption 1 is satisfied and let $K \subset U^*$ be a compact set and $t \in (0, 1]$. For every $\rho > 0, R > 0, a > 0$, there exist $\varepsilon_0 \in (0, \varepsilon_*], \alpha > 0$ such that for all $x \in K, f \in \mathscr{C}^0([0, t]; \mathbf{R}^k)$ and $g = S_x(f)$ with*

$$\int_0^t |\dot{f}_s|^2 \, ds \leq a \qquad and \qquad g([0, t]) \subset K,$$

we have the following estimate

$$\mathbf{P}\left\{d_t\left(\tfrac{1}{\sqrt{r(\varepsilon)}}B^\varepsilon, f\right) \leq \alpha \quad and \quad d_t(y^\varepsilon, g) > \rho\right\} \leq e^{-Rr(\varepsilon)}$$

for all $0 < \varepsilon \leq \varepsilon_0$.

Proof. Since K is compact, there is $\rho_0 \in (0, \rho]$ and a compact set $K' \subset U^*$ containing a $2\rho_0$ neighborhood of K. Let $V \subset \mathbf{R}^d$ be a bounded, open set such that $K' \subset V \subset U^*$ and let $\varphi : \mathbf{R}^d \to [0, 1]$ be C^∞ function with $\varphi = 1$ on K' and 0 on V^c. For any function p on U^* we define p_φ on $\mathbf{R}^d$ by

$$p_\varphi(y) = \begin{cases} p(y)\varphi(y) & y \in U^*, \\ 0 & y \notin U^*. \end{cases}$$

Thus, to $b, \sigma, b_\varepsilon, \sigma_\varepsilon$ we associate respectively $b_\varphi, \sigma_\varphi, b_{\varphi,\varepsilon}, \sigma_{\varphi,\varepsilon}$. Then, by construction and Assumption 1, $b_\varphi, \sigma_\varphi, b_{\varphi,\varepsilon}, \sigma_{\varphi,\varepsilon}$ are bounded and globally Lipschitz on $\mathbf{R}^d$ and $b_{\varphi,\varepsilon} \to b_\varphi$, $\sigma_{\varphi,\varepsilon} \to \sigma_\varphi$ uniformly on $\mathbf{R}^d$ as $\varepsilon \to 0$. Let y_φ^ε be the unique solution of the Itô SDE

$$\begin{cases} dy_t = b_{\varphi,\varepsilon}(y_t)\,dt + \frac{1}{\sqrt{r(\varepsilon)}}\sigma_{\varphi,\varepsilon}(y_t)\,dB_t^\varepsilon, \\ y_0 = x. \end{cases}$$

By standard arguments, the solution above belongs almost surely to $\mathscr{C}_x([0, t]; U)$. Then [4, Theorem 1.1] with $h(\varepsilon) \equiv 1$ and ε^{-2} replaced by $r(\varepsilon) < \varepsilon^{-2}$, provides the existence of

$\varepsilon_0 > 0, \alpha > 0$ such that for all $x \in K$, $f \in \mathscr{C}^0([0, t]; \mathbf{R}^k)$ and $g = S_x(f)$ with

$$\int_0^t |\dot{f}_s|^2 \, ds \leq a \qquad \text{and} \qquad g([0, t]) \subset K$$

we have, for every $\varepsilon \in (0, \varepsilon_0]$

$$\mathbf{P}\left\{d_t\left(\tfrac{1}{\sqrt{r(\varepsilon)}} B^\varepsilon, f\right) \leq \alpha \ \text{ and } \ d_t(y_\varphi^\varepsilon, g) > \rho_0\right\} \leq e^{-Rr(\varepsilon)}.$$

However, y_φ^ε and y^ε coincide until the first time both y_φ^ε and y^ε exit K'. Since $g([0, t]) \subset K$ and $\rho > \rho_0$

$$\mathbf{P}\left\{d_t\left(\tfrac{1}{\sqrt{r(\varepsilon)}} B^\varepsilon, f\right) \leq \alpha \ \text{ and } \ d_t(y^\varepsilon, g) > \rho\right\}$$

$$\leq \mathbf{P}\left\{d_t\left(\tfrac{1}{\sqrt{r(\varepsilon)}} B^\varepsilon, f\right) \leq \alpha \ \text{ and } \ d_t(y^\varepsilon, g) > \rho_0\right\}$$

$$\leq \mathbf{P}\left\{d_t\left(\tfrac{1}{\sqrt{r(\varepsilon)}} B^\varepsilon, f\right) \leq \alpha \ \text{ and } \ d_t(y_\varphi^\varepsilon, g) > \rho_0\right\}$$

$$\leq e^{-Rr(\varepsilon)}$$

and the proof is finished. $\square$

Recall the Cramer transform $\lambda : \mathscr{E}([0, 1]; U^*) \to [0, +\infty]$ introduced in (2.12), and for any $A \subset \mathscr{E}_x([0, 1]; U^*)$ Borel set define

$$\Lambda(A) = \inf_{g \in A} \lambda(g). \tag{4.1}$$

Theorem 4.2. *Suppose that Assumption 1 is satisfied. For any Borel set $A \subset \mathscr{E}_x([0, 1]; U^*)$,*

$$-\Lambda(interior(A)) \leq \liminf_{\varepsilon \to 0} \frac{1}{r(\varepsilon)} \log \mathbf{P}\{y^\varepsilon \in A\} \leq \limsup_{\varepsilon \to 0} \frac{1}{r(\varepsilon)} \log \mathbf{P}\{y^\varepsilon \in A\} \leq -\Lambda(\overline{A}), \tag{4.2}$$

where interior(A) and $\overline{A}$ respectively denote the interior and closure of A.

For the proof of Theorem 4.2, we refer to [2, proof of 2.13 Théorème, Chapitre III] with ε^{-2} replaced by $r(\varepsilon)$, which works in our setting as one merely needs σ to be locally Lipschitz on U^* rather than C^1.

Given the previous two results, we are now prepared to prove Theorem 2.6.

5. Proof of Theorem 2.6

The proof of Theorem 2.6 is similar to the proof of the main result in Baldi [3] and also Caramellino [10], but with a different topology and set of assumptions.

We first need some auxiliary results.

Proposition 5.1. *Suppose that Assumption 1 is satisfied and $x \in U^*$, t_*, and $\mathscr{K}_x(t_*)$ are as in the statement of Theorem 2.6. Then the following assertions hold.*

(i) *For any $c \in (0, 1)$:*

$$\mathbf{P}\left\{\lim_{j \to \infty} d_{t_*}(y^{c^j}, \mathscr{K}_x(t_*)) = 0\right\} = 1.$$

(ii) *If Assumption 2(i) is furthermore satisfied, then*

$$\mathbf{P}\left\{ \lim_{\varepsilon \to 0} d_{t_*}(y^\varepsilon, \mathscr{K}_x(t_*)) = 0 \right\} = 1.$$

Proof. To show (i), fix $\delta > 0$ and consider the set

$$\mathscr{K}_{x,\delta} = \{g \in \mathscr{E}_x([0, 1]; U^*) : d_{t_*}(g, \mathscr{K}_x(t_*)) \geq \delta\}. \tag{5.1}$$

First we claim that $\mathscr{K}_{x,\delta}$ is closed in $\mathscr{E}([0, 1]; U^*)$. Indeed, if $g_n \in \mathscr{K}_{x,\delta}$ converges to g in $\mathscr{E}([0, 1]; U^*)$ as $n \to \infty$ and $g(t_*) = \Delta$, then $d_{t_*}(g, \mathscr{K}_x(t_*)) = \infty$ since, by the definition of t^*, $h(t^*) \neq \Delta$ for any $h \in \mathscr{K}_x(t_*)$. Hence, $g \in \mathscr{K}_{x,\delta}$. If on the other hand $g(t_*) \in U^*$, then by the continuity of g, $\tau_1(g) > t_*$ and by the definition of convergence in $\mathscr{E}([0, 1]; U^*)$ one has that $\tau_1(g_n) > t_*$ for any sufficiently large n. Then, the triangle inequality (2.8) and the definition of d_{t_*} imply for any $h \in \mathscr{K}_x(t_*)$

$$d_{t_*}(g_n, g) + d_{t_*}(g, h) \geq d_{t_*}(g_n, h) \geq \delta. \tag{5.2}$$

Passing $n \to \infty$, we obtain $g \in \mathscr{K}_{x,\delta}$.

We now claim that there exists $\delta' > 0$ for which $\Lambda(\mathscr{K}_{x,\delta}) > 1 + \delta'$. Suppose to the contrary that $\Lambda(\mathscr{K}_{x,\delta}) \leq 1$. By definition, there exists a sequence $g_n \in \mathscr{K}_{x,\delta}$ such that

$$\lim_{n \to \infty} \lambda(g_n) = \Lambda(\mathscr{K}_{x,\delta}) \leq 1.$$

Thus, for all n large enough, $g_n \in \mathscr{M} := \{g \in \mathscr{E}_x([0, 1]; U^*) : \lambda(g) \leq 2\}$. The set $\mathscr{M}$ is sequentially compact since $\mathscr{M}$ is the image of the sequentially compact set (the compact Sobolev embedding $H^1 \hookrightarrow \mathscr{C}$)

$$\mathscr{C}_2 = \{f \in \mathscr{C}([0, 1]; \mathbf{R}^k) : \tfrac{1}{2}\int_0^1 |\dot{f}_s|^2 \, ds \leq 2\}$$

under the continuous mapping $S_x : \mathscr{C}_2 \to \mathscr{E}_x([0, 1]; U^*)$ given by $S_x(f)$. Hence, the sequence $\{g_n\}$ has a convergence subsequence $\{g_{n_k}\}$ converging to some $g \in \mathscr{K}_{x,\delta}$. The lower semicontinuity of λ then implies

$$1 \geq \liminf_{k \to \infty} \lambda(g_{n_k}) \geq \lambda(g).$$

In particular, $g \in \mathscr{K}_x$, contradicting closedness and the definition of $\mathscr{K}_{x,\delta}$. Thus we have shown that there is $\delta' > 0$ so that $\Lambda(\mathscr{K}_{x,\delta}) > 1 + \delta'$.

By Theorem 4.2 and the fact that $\mathscr{K}_{x,\delta}$ is closed

$$\limsup_{\varepsilon \to 0} \frac{1}{r(\varepsilon)} \log \mathbf{P}\{y^\varepsilon \in \mathscr{K}_{x,\delta}\} \leq -(1 + \delta').$$

Hence, using the definition of $r(\varepsilon)$ we have for all j large and $c \in (0, 1)$

$$\mathbf{P}\{y^{c^j} \in \mathscr{K}_{x,\delta}\} \leq \frac{C}{j^{1+\frac{\delta'}{2}}} \tag{5.3}$$

for some constant $C = C(c) > 0$. The Borel–Cantelli lemma then implies

$$\mathbf{P}\left\{ \limsup_{j \to \infty} d_{t_*}(y^{c^j}, \mathscr{K}_x(t_*)) \geq \delta \right\} = 0.$$

Since $\delta > 0$ was arbitrary, $\lim_{j \to \infty} d_{t_*}(y^{c^j}, \mathscr{K}_x(t_*)) = 0$ almost surely, finishing the proof of part (i).

To establish part (ii), fix $\delta > 0$ and by Assumption 2(i) choose a constant $c \in (0, 1)$ and a (random) index $J_0 = J_0(\omega, c) > 0$ such that for all $j \geq J_0$ and $\varepsilon \in [c^{j+1}, c^j]$

$$y_{t_*}^\varepsilon \in U^* \qquad \text{and} \qquad d_{t_*}(y^{c^j}, y^\varepsilon) < \frac{\delta}{2}. \tag{5.4}$$

By (5.3), we can increase J_0 if necessary so that $j \geq J_0$ implies

$$d_{t_*}(y^{c^j}, \mathscr{K}_x(t_*)) < \frac{\delta}{2}.$$

Using (5.4), the triangle inequality (2.8), for any $j \geq J_0$ and $\varepsilon \in [c^{j+1}, c^j]$ one has

$$d_{t_*}(y^\varepsilon, \mathscr{K}_x(t_*)) < \delta,$$

and part (ii) follows. $\square$

Proposition 5.2. *Suppose that Assumption 1 is satisfied. Let $g \in \mathscr{K}_x$ be such that $\lambda(g) < 1$. Then, for all $\varepsilon > 0$ and $c \in (0, 1)$ we have*

$$\mathbf{P}\{d_{t_*}(y^{c^j}, g) < \varepsilon \text{ for infinitely many } j\} = 1.$$

Proof. In the proof, we abbreviate *for infinitely many j* as *i.o. j*. Fix $g \in \mathscr{K}_x$ with $a := \lambda(g) < 1$ and fix $\varepsilon > 0$ and $c \in (0, 1)$. Since the infimum in the definition of λ (see (2.12)) is attained, there exists $f \in \mathscr{C}_a([0, 1]; \mathbf{R}^k)$ so that $g = S_x(f)$. By shifting f by a constant value, we may assume without loss of generality that $f_0 = 0$, as the time derivative is invariant under this shift. For $a_* > 0$ and $c \in (0, 1)$ define events

$$F_j = \left\{ d_{t_*}\left(\frac{1}{\sqrt{r(c^j)}} B^{c^j}, f \right) < a_* \right\} \qquad \text{and} \qquad H_j = \{d_{t_*}(y^{c^j}, g) < \varepsilon\}.$$

Then, Theorem 4.1 implies that there exist $a_* > 0$ and $J > 0$ such that $j \geq J$ implies

$$\mathbf{P}\{F_j \cap H_j^c\} \leq \exp(-2r(c^j)) \leq \frac{C}{j^2}$$

for some constant $C > 0$. The Borel–Cantelli lemma then implies $\mathbf{P}\{F_j \cap H_j^c \text{ i.o.} j\} = 0$. Now, by Mueller [32] or Gantert [15] we have $1 = \mathbf{P}\{F_j \text{ i.o. } j\}$. Thus,

$$1 = \mathbf{P}\{F_j \text{ i.o. } j\} \leq \mathbf{P}\{F_j \cap H_j \text{ i.o. } j\} + \mathbf{P}\{F_j \cap H_j^c \text{ i.o. } j\} \leq \mathbf{P}\{H_j \text{ i.o. } j\},$$

as desired. $\square$

We will also need the following topological result.

Lemma 5.3. *Let $t \in (0, 1]$, $V \subset \mathbf{R}^d$ be open, $x \in V$ and suppose $K \subset \mathscr{C}_x([0, t]; V)$. Then, K is compact in $\mathscr{C}_x([0, t]; V)$ if and only if K is compact in $\mathscr{E}_x([0, t]; V)$.*

Proof. Let $K \subset \mathscr{C}_x([0, t]; V)$ be compact in $\mathscr{E}_x([0, t]; V)$. To show that K is compact in $\mathscr{C}_x([0, t]; V)$, it suffices to prove that if $\mathscr{V}$ is open in $\mathscr{C}_x([0, t]; V)$, then $\mathscr{V}$ is open in $\mathscr{E}_x([0, t]; V)$. Equivalently, we show that $\mathscr{W} = \mathscr{E}_x([0, t]; V) \setminus \mathscr{V}$ is closed in $\mathscr{E}_x([0, t]; V)$. Let $g_n \in \mathscr{W}$ be such that $g_n \to g \in \mathscr{E}_x([0, t]; V)$. If $g(t) = \Delta$, then clearly $g \notin \mathscr{V} \subset \mathscr{C}_x([0, t]; V)$, and therefore $g \in \mathscr{W}$. On the other hand, if $g(t) \in V$, then $g \in \mathscr{C}_x([0, t]; V)$ by definition of $\mathscr{E}_x([0, t]; V)$. By definition of the topology on $\mathscr{E}_x([0, t]; V)$, $g_n \in \mathscr{W} \cap \mathscr{C}_x([0, t]; V)$ for all $n \geq N$, $N > 0$ large enough. Since $\mathscr{W} \cap \mathscr{C}_x([0, t]; V)$ is closed and $g_n \to g$ in the topology of $\mathscr{C}_x([0, t]; V)$, then $g \in \mathscr{W}$. Thus, $\mathscr{W}$ is closed in $\mathscr{E}_x([0, t]; V)$.

Conversely, suppose $K \subset \mathscr{C}_x([0,t];V)$ is compact in $\mathscr{C}_x([0,t];V)$. To prove that K is compact in $\mathscr{E}_x([0,t];V)$, it is enough to show that if $\mathscr{V}$ is open in $\mathscr{E}_x([0,t];V)$, then $\mathscr{V} \cap \mathscr{C}_x([0,t];V)$ is open in $\mathscr{C}_x([0,t];V)$. Note that this follows immediately from the fact that if $g_n \to g$ in $\mathscr{C}_x([0,t];V)$, then $g_n \to g$ in $\mathscr{E}_x([0,t];V)$. $\square$

Given the previous three results, we next prove the main general result, Theorem 2.6.

Proof of Theorem 2.6. We have already established part (ii) in Proposition 5.1(ii). To prove part (i), we need to show that $\mathscr{Y}(\omega)$ is relatively compact, almost surely. By Proposition 5.1(ii), for any $\delta > 0$ there exists $\varepsilon_0 = \varepsilon_0(\omega, \delta) \in (0, \varepsilon_*]$ such that $d_{t_*}(y^\varepsilon, \mathscr{K}_x(t_*)) \le \delta$ for all $\varepsilon \in (0, \varepsilon_0]$. In particular, for any small $\delta > 0$, $y^\varepsilon(t_*) \ne \Delta$ a.s. for any $\varepsilon \in [0, \varepsilon_0]$. Hence, $y^\varepsilon \in \mathscr{C}_x([0, t_*]; U^*)$ for any $\varepsilon \in (0, \varepsilon_0]$, and therefore the closure of $\{y^\varepsilon\}_{0 < \varepsilon \le \varepsilon_0}$ in $\mathscr{E}_x([0, t_*]; U^*)$ is the same as the closure in $\mathscr{C}_x([0, t_*]; U^*)$.

We claim that the closure of $\{y^\varepsilon\}_{0 < \varepsilon \le \varepsilon_0}$ in $\mathscr{E}_x([0, t_*]; U^*)$ in $\mathscr{C}_x([0, t_*]; U^*)$ is compact in $\mathscr{C}_x([0, t_*]; U^*)$, almost surely. Consequently, the closure of $\{y^\varepsilon\}_{0 < \varepsilon \le \varepsilon_0}$ is compact in $\mathscr{E}_x([0, t_*]; U^*)$ almost surely by Lemma 5.3 finishing the proof of part (i) of the result. To prove the claim, first note that by Assumption 2(ii), there exists a subset $\Omega_1 \subset \Omega$ with $\mathbf{P}(\Omega_1) = 1$ such that $t \mapsto y_t^\varepsilon(\omega)$ is continuous up to the time of explosion, for every $\omega \in \Omega_1$ and every $\varepsilon \in (0, \varepsilon_*)$. Let $\Omega_2 := \{\omega : \lim_{\varepsilon \to 0} d_{t_*}(y^\varepsilon(\omega), \mathscr{K}_x(t_*)) = 0\}$. Then by Proposition 5.1(ii), we have that $\mathbf{P}(\Omega_2) = 1$. Hence, setting $\Omega_{1,2} := \Omega_1 \cap \Omega_2$, we also have $\mathbf{P}(\Omega_{1,2}) = 1$. Note that we have that $t \mapsto y_t^\varepsilon(\omega)$ is continuous for every $\omega \in \Omega_{1,2}$ and every $\varepsilon \in (0, \varepsilon_0]$. Let us now fix a sequence $\{\varepsilon_n\} \subset (0, \varepsilon_0]$. By passing to a subsequence, we can suppose that $\varepsilon_n \to \varepsilon_\infty \in [0, \varepsilon_0]$. If $\varepsilon_\infty > 0$, then by Assumption 2(ii), $y^{\varepsilon_n}(\omega) \to y^{\varepsilon_\infty}(\omega)$ in $\mathscr{E}_x([0, t_*]; U^*)$, for all $\omega \in \Omega_{1,2}$. Note that this convergence happens in $\mathscr{C}_x([0, t_*]; U^*)$ for all $\omega \in \Omega_{1,2}$ by definition of ε_0. If $\varepsilon_\infty = 0$, then by Proposition 5.1(ii), $y^{\varepsilon_n}(\omega) \to \mathscr{K}_x(t_*)$ in $\mathscr{C}_x([0, t_*]; U^*)$ for all $\omega \in \Omega_{1,2}$. However, since $\mathscr{K}_x(t_*)$ is compact, there exists a subsequence of $\{\varepsilon_n\}$, again denoted by $\{\varepsilon_n\}$, such that y^{ε_n} converges to $g \in \mathscr{K}_x(t_*)$, and therefore g belongs to the closure of $\{y^\varepsilon\}_{0 < \varepsilon < \varepsilon_0}$. Our claim now follows. We next prove part (iii) of the statement. Note that by Proposition 5.1(ii), for any $\omega \in \Omega_{1,2}$ the limit set $\mathscr{K}'$ of $\mathscr{Y}$ as $\varepsilon \to 0$ is contained in $\mathscr{K}_x(t_*)$. To see the other inclusion, first note that $\mathscr{K}_x(t_*)$ is compact and hence it contains a countable dense subset, say $\mathscr{D} = \{g_m\}_{m=1}^\infty$ with $\lambda(g_m) \le 1$ for all $m \ge 1$. Note that we can find a countable dense subset $\mathscr{D}'$ of $\mathscr{K}_x(t_*)$ such that $\lambda(g) < 1$ for all $g \in \mathscr{D}'$. Indeed, for every $g_m \in \mathscr{D}$, the infimum in (2.12) is attained and thus there exists $f_m \in \mathscr{C}^0([0, 1]; \mathbf{R}^k)$ with

$$\frac{1}{2} \int_0^1 |\dot{f}_{m_t}|^2 dt \le 1 \quad \text{and} \quad S_x(f_m) = g_m.$$

For each $n > 1$ define $f_m^n = (1 - 1/n) f_m$ and let $g_m^n := S_x(f_m^n)$. Clearly, $\lambda(g_m^n) < 1$ and $\mathscr{D}' := \{g_m^n\}_{m,n=1}^\infty$ is the desired countable dense subset of $\mathscr{K}_x(t_*)$. By Proposition 5.2, for every $k, n, m \ge 1$, and $c \in (0, 1)$ the set

$$\Omega_{n,m,k} := \left\{\omega : d_{t_*}(y^{c^j}(\omega), g_m^n) < \frac{1}{2k} \text{ for infinitely many } j\right\}$$

satisfies $\mathbf{P}(\Omega_{n,m,k}) = 1$. Now, set $\Omega' := \bigcap_{n,m,k=1}^\infty \Omega_{n,m,k}$ and note that $\mathbf{P}(\Omega') = 1$. For any $g \in \mathscr{K}_x(t_*)$ with $\lambda(g) \le 1$ and for every $k \ge 1$ there exist $n(k), m(k) \ge 1$ and $g_{m(k)}^{n(k)} \in \mathscr{D}'$ such

that $d_{t_*}(g, g_{m(k)}^{n(k)}) < \frac{1}{2k}$. Thus, by the triangle inequality, for every $k \geq 1$

$$\Omega' \subset \left\{ \omega : d_{t_*}\left(y^{c^j}(\omega), g_{m(k)}^{n(k)} \right) < \frac{1}{2k} \text{ for infinitely many } j \right\}$$

$$\subset \left\{ \omega : d_{t_*}(y^{c^j}(\omega), g) < \frac{1}{2k} + d_{t_*}\left(g, g_{m(k)}^{n(k)} \right) \text{ for infinitely many } j \right\}$$

$$\subset \left\{ \omega : d_{t_*}(y^{c^j}(\omega), g) < \frac{1}{k} \text{ for infinitely many } j \right\}.$$

Hence,

$$\Omega' \subset \bigcap_{k=1}^{\infty} \left\{ \omega : d_{t_*}(y^{c^j}(\omega), g) < \frac{1}{k} \text{ for infinitely many } j \right\}.$$

Thus, we proved that there exists a set Ω' with $\mathbf{P}(\Omega') = 1$ such that any $g \in \mathcal{K}_x(t_*)$ with $\lambda(g) \leq 1$ belongs to $\mathcal{K}'$ for all $\omega \in \Omega'$. $\quad\square$

6. Proof of Lemma 2.9 and Corollary 2.10

In this section, we prove Lemma 2.9 and Corollary 2.10, which give basic criteria for the family $\{y^\varepsilon\}_{0<\varepsilon\leq\varepsilon_*}$ to satisfy Assumption 2. We begin with the:

Proof of Lemma 2.9. We first prove Assumption 2(i). Recall the definition of $t_* = t_*(x) \in (0, 1]$ and $L \subset U^*$ in (2.15) and recall that $\mathcal{K}_x(t_*)$ is compact in both $\mathcal{C}_x([0, t_*]; U^*)$ and $\mathcal{E}_x([0, t_*]; U^*)$ by Lemma 5.3. From (2.15), it follows that $g([0, t_*]) \subset L \subset U^*$ for all $g \in \mathcal{K}_x(t_*)$. Choose a compact set $L' \subset U^*$ with $L \subset \text{interior}(L')$ and fix $\delta > 0$ such that $2\delta < \text{dist}(L', \partial U^*)$. Using compactness of $\mathcal{K}_x(t_*)$ in $\mathcal{C}_x([0, t_*]; U^*)$, the Arzelà–Ascoli theorem implies that the set $\mathcal{K}_x(t_*)$ is equicontinuous, and therefore for any $c \in (0, 1)$ sufficiently close to 1 one has

$$\sup_{h \in \mathcal{K}_x(t_*)} \sup_{\substack{t \in [0, t_*] \\ s \in [ct, t]}} |h(t) - h(s)| < \frac{\delta}{3}. \tag{6.1}$$

By properties (i) and (ii) in the definition of an asymptotic index (Definition 2.2), we can decrease ε_0 if necessary such that $|\psi(u)| < \kappa$ for any $u \in [0, \varepsilon_0]$, where κ is as in Definition 2.1(iii). Then, since $\{\Phi_\alpha\}$ is a family of weak contractions centered at x, for all $c \in (0, 1)$ close enough to 1, there exists $J_1 > 0$ (deterministic) such that for all $j \geq J_1$ and all $\varepsilon \in [c^{j+1}, c^j]$ one has $U^* \subset U_{\psi(c^j)}$ and

$$|\Phi_{\psi(\varepsilon)} \circ \Phi_{\psi(c^j)}^{-1}(y) - y| + 2|\Phi_{\psi(c^{j+1})} \circ \Phi_{\psi(c^j)}^{-1}(y) - y| < \frac{\delta}{3} \tag{6.2}$$

for all $y \in L'$. Fix $c \in (0, 1)$ sufficiently close to 1 such that (6.1) and (6.2) are satisfied.

By Proposition 5.1(i), for almost every ω we can choose a finite $J_0 = J_0(\omega, c) \geq J_1$ such that $j \geq J_0$ implies $y_s^{c^j} \in L'$ for any $s \in [0, t_*]$. Also, from (2.8) for all $j \geq J_0$ and $\varepsilon \in [c^{j+1}, c^j]$ it follows that

$$d_{t_*}(y^\varepsilon, y^{c^j}) \leq d_{t_*}\left(\Phi_{\psi(\varepsilon)} \circ \Phi_{\psi(c^j)}^{-1}(y^{c^j}), y^{c^j} \right) + d_{t_*}(\Phi_{\psi(\varepsilon)}(x_{\varepsilon\cdot}), \Phi_{\psi(\varepsilon)}(x_{c^j\cdot})). \tag{6.3}$$

Let us prove that the second term on the right hand side is finite. First, observe that

$$\Phi_{\psi(\varepsilon)}(x_{c^j t}) = \Phi_{\psi(\varepsilon)} \circ \Phi_{\psi(c^j)}^{-1}(y_t^{c^j}).$$

Thus, since $j \geqslant J_0$ we have $y_s^{c^j} \in L'$ for any $s \in [0, t_*]$, and one has by (6.2)

$$d_{t_*}(\Phi_{\psi(\varepsilon)} \circ \Phi_{\psi(c^j)}^{-1}(y_{\cdot}^{c^j}), y_{\cdot}^{c^j}) < \frac{\delta}{3}. \tag{6.4}$$

In particular, since $2\delta < \mathrm{dist}(L', \partial U^*)$, $\Phi_{\psi(\varepsilon)}(x_{c^j t}) \in U^*$ for any $t \in [0, t^*]$.

Next, since ψ is an asymptotic index and $\{\Phi_\alpha\}_{\alpha > 0}$ is a family of weak contractions centered at x, one has for any $t \leq t^*$, $t < \tau_{t_*}(y^\varepsilon)$

$$\begin{aligned}
|\Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) - \Phi_{\psi(\varepsilon)}(x_{c^j t})| &\leq |\Phi_{\psi(c^{j+1})}(x_{\varepsilon t}) - \Phi_{\psi(c^{j+1})}(x_{c^j t})| \\
&\leq \sup_{s \in [ct, t]} |\Phi_{\psi(c^{j+1})}(x_{c^j s}) - \Phi_{\psi(c^{j+1})}(x_{c^j t})| \\
&\leq \sup_{s \in [ct, t]} \left\{ |\Phi_{\psi(c^{j+1})}(x_{c^j s}) - y_s^{c^j}| \right. \\
&\qquad\quad \left. + |y_t^{c^j} - \Phi_{\psi(c^{j+1})}(x_{c^j t})| + |y_t^{c^j} - y_s^{c^j}| \right\} \\
&\leq 2 d_{t_*}(\Phi_{\psi(c^{j+1})} \circ \Phi_{\psi(c^j)}^{-1}(y_{\cdot}^{c^j}), y_{\cdot}^{c^j}) + \sup_{\substack{t \in [0, t_*] \\ s \in [ct, t]}} |y_t^{c^j} - y_s^{c^j}|.
\end{aligned}$$

Consequently, from (6.1) and (6.2) for any $g \in \mathscr{K}_x(t_*)$ and $t \leq t_*$ it follows

$$\begin{aligned}
|\Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) - \Phi_{\psi(\varepsilon)}(x_{c^j t})| &\leq 2 d_{t_*}(\Phi_{\psi(c^{j+1})} \circ \Phi_{\psi(c^j)}^{-1}(y_{\cdot}^{c^j}), y_{\cdot}^{c^j}) + 2 d_{t_*}(y^{c^j}, g) \\
&\quad + \sup_{h \in \mathscr{K}_x(t_*)} \sup_{\substack{t \in [0, t_*] \\ s \in [ct, t]}} |h(t) - h(s)| \\
&\leq \frac{2\delta}{3} + 2 d_{t_*}(y^{c^j}, g).
\end{aligned}$$

Since the left hand side is independent of g, we can take the infimum with respect to $g \in \mathscr{K}_x(t_*)$ and obtain

$$|\Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) - \Phi_{\psi(\varepsilon)}(x_{c^j t})| \leq \frac{2\delta}{3} + 2 d_{t_*}(y^{c^j}, \mathscr{K}_x(t_*)). \tag{6.5}$$

Thus, by Proposition 5.1, by increasing J_2 if needed, for any $j \geq J_2$ and $t \in [0, t_*]$ we have

$$|\Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) - \Phi_{\psi(\varepsilon)}(x_{c^j t})| \leq \delta. \tag{6.6}$$

Consequently, since $2\delta < \mathrm{dist}(L', \partial U^*)$, employing (6.4) we obtain that $y_t^\varepsilon \in U^*$ for all $t \in [0, t_*]$.

Returning to (6.3) and using (6.2), definition of d_{t_*}, and (6.6), we obtain for any $j \geq J_2$

$$d_{t_*}(y^\varepsilon, y^{c^j}) \leq d_{t_*}(\Phi_{\psi(\varepsilon)} \circ \Phi_{\psi(c^j)}^{-1}(y_{\cdot}^{c^j}), y_{\cdot}^{c^j}) + \sup_{t \in [0, t_*]} |\Phi_{\psi(\varepsilon)}(x_{\varepsilon t}) - \Phi_{\psi(\varepsilon)}(x_{c^j t})| \leq \frac{4\delta}{3} \tag{6.7}$$

and Assumption 2(i) follows.

To prove Assumption 2(ii), fix $\varepsilon_\infty \in (0, \varepsilon_*)$, let ω be a realization of the noise, and $t_1 \in [0, \tau_{t_*}(y^{\varepsilon_\infty}(\omega)))$ with $t_1 \leq t_*$. Similar to the proof of Theorem 2.6, let Ω_1 be the subset of Ω with $\mathbf{P}(\Omega_1) = 1$ such that $y_{\cdot}^\varepsilon(\omega)$ is continuous for every $\omega \in \Omega_1$ and every $\varepsilon \in (0, \varepsilon_0]$. The assertion follows once we prove that for all $\omega \in \Omega_1$, for any sequence ε_n with $\varepsilon_n \to \varepsilon_\infty$ as $n \to \infty$ one has $y^{\varepsilon_n}(\omega) \to y^{\varepsilon_\infty}(\omega)$ as $n \to \infty$ in the space $\mathscr{C}_x([0, t_1]; U^*)$. To simplify the notation, we will often drop the argument ω below. Since $t_1 < \tau_{t_*}(y^{\varepsilon_\infty})$, then $\delta_0 := \inf_{t \in [0, t_1]} \mathrm{dist}(y_t^{\varepsilon_\infty}, \partial U^*) > 0$. In particular $y^{\varepsilon_\infty}([0, t_1]) \subset K$ for some compact $K = K(\omega)$

with $K \subset U^*$ almost surely. Fix $\delta \in (0, \delta_0]$ and large enough n such that $\varepsilon_n > \varepsilon_\infty/2$ and

$$|\Phi_{\psi(\varepsilon_n)} \circ \Phi_{\psi(\varepsilon_\infty)}^{-1}(y) - y| < \frac{\delta}{3} \tag{6.8}$$

for any $y \in K$. Note that such n exists since $\varepsilon_n \to \varepsilon_\infty$, $\{\Phi_\alpha\}_{\alpha > 0}$ is a sequence of weak contractions and ψ an asymptotic index.

Then, from (6.8) and the fact that $y \mapsto \Phi_{\psi(\varepsilon_\infty/2)}(y) \in C^2(U)$, we obtain for any $t \leq t_1$ with $t < \tau_{t^*}(y^{\varepsilon_n})$

$$|y_t^{\varepsilon_n} - y_t^{\varepsilon_\infty}| \leq |\Phi_{\psi(\varepsilon_n)}(x_{\varepsilon_n t}) - \Phi_{\psi(\varepsilon_n)}(x_{\varepsilon_\infty t})| + |\Phi_{\psi(\varepsilon_n)} \circ \Phi_{\psi(\varepsilon_\infty)}^{-1}(y_t^{\varepsilon_\infty}) - y_t^{\varepsilon_\infty}|$$

$$\leq |\Phi_{\psi(\varepsilon_\infty/2)}(x_{\varepsilon_n t}) - \Phi_{\psi(\varepsilon_\infty/2)}(x_{\varepsilon_\infty t})| + \frac{\delta}{3}. \tag{6.9}$$

Using the fact that

$$s \mapsto x_s(\omega) : [0, t] \to \mathscr{E}_x([0, t]; U^*) \tag{6.10}$$

is continuous for all $\omega \in \Omega_1$, we obtain that for any sufficiently large n, $\text{dist}(y_t^{\varepsilon_n}, \partial U^*) \geq \frac{\delta_0}{2}$ for any $t \leq t_1$, $t < \tau_{t^*}(y^{\varepsilon_n})$. Again, a standard extension argument implies that $\text{dist}(y_t^{\varepsilon_n}, \partial U^*) \geq \frac{\delta_0}{2}$ for any $t \leq t_1$, and in particular $t_1 < \tau_{t^*}(y^{\varepsilon_n})$. Finally, passing $n \to \infty$ using (6.10), and since $\delta > 0$ was arbitrary, the assertion follows. $\quad\square$

We now turn our attention to the proof of Corollary 2.10.

Proof of Corollary 2.10. Fix $c \in (0, 1)$, $\delta > 0$, and $\varepsilon \in [c^{j+1}, c^j]$. By Proposition 5.1(i) there is an almost surely finite random variable $J_0 = J_0(c, \omega)$ such that

$$y^{c^j}([0, t_*]) \subset \text{interior}(L) \quad \text{for all} \quad j \geq J_0,$$

where t_* and L are as in (2.15) (see also Definition 2.2). Define $T_\varepsilon = T_\varepsilon(\omega) = \inf\{t \geq 0 : y_t^\varepsilon \notin \text{interior}(L)\}$ and for every $t \geq 0$ set $T_\varepsilon(t) = t \wedge T_\varepsilon$. Then, for any $j \geq J_0$ and $t \in [0, t_*]$ we obtain

$$\sup_{s \leq T_\varepsilon(t)} |y_s^\varepsilon - y_s^{c^j}| \leq \int_0^{T_\varepsilon(t)} |b_\varepsilon(y_s^\varepsilon) - b_{c^j}(y_s^{c^j})| \, ds + \sup_{s \leq t_*} \left| \frac{\sigma_\varepsilon}{\sqrt{\varepsilon r(\varepsilon)}} B_{\varepsilon s} - \frac{\sigma_{c^j}}{\sqrt{c^j r(c^j)}} B_{c^j s} \right|$$

$$=: S_1 + S_2(j, t_*). \tag{6.11}$$

To estimate S_1, note that since both y^ε, y^{c^j} map $[0, T_\varepsilon(t_*)]$ to the compact set $L \subset U^*$ for $j \geq J_0$, we have by Assumption 1

$$S_1 \leq \int_0^{T_\varepsilon(t)} |b_\varepsilon(y_s^\varepsilon) - b(y_s^\varepsilon)| \, ds + \int_0^{T_\varepsilon(t)} |b_{c^j}(y_s^{c^j}) - b(y_s^{c^j})| \, ds$$

$$+ \int_0^{T_\varepsilon(t)} |b(y_s^\varepsilon) - b(y_s^{c^j})| \, ds$$

$$\leq 2 C_j(L) t_* + C_L \int_0^t \sup_{v \leq T_\varepsilon(s)} |y_v^\varepsilon - y_v^{c^j}| \, ds$$

where $t \leq t_*$, $C_j(L), C_L > 0$ are deterministic constants and $C_j(L) \to 0$ as $j \to \infty$. Combining the previous estimate with (6.11) and using Gronwall's inequality gives

$$\sup_{s \leq T_\varepsilon(t_*)} |y_s^\varepsilon - y_s^{c^j}| \leq (2 C_j(L) t_* + S_2(j, t_*)) e^{C_L t_*}$$

for any $j \geq J_0$.

In order to estimate $S_2(j, t_*)$, for any multiindex $\alpha > 0$ belonging to $\mathbf{R}^k$, let $\Phi_\alpha : \mathbf{R}^k \to \mathbf{R}^k$ and $\psi : [0, \varepsilon_*] \to \mathbf{R}^k$ be given by

$$\Phi_\alpha(y) = (y_1 \alpha_1^{-1}, \ldots, y_k \alpha_k^{-1}) \qquad \text{and} \qquad \psi(\varepsilon) = (\sqrt{\varepsilon \log \log \varepsilon^{-1}}, \ldots, \sqrt{\varepsilon \log \log \varepsilon^{-1}}),$$

where $\varepsilon_* \in (0, e^{-1})$. By Example 2.1, $\{\Phi_\alpha\}_{\alpha>0}$ is a family of weak contractions centered at 0 in $\mathbf{R}^k$ while by Example 2.2, for $\varepsilon_* > 0$ sufficiently small, $\psi : [0, \varepsilon_*] \to [0, \infty)^k$ is an asymptotic index. We can then estimate $S_2(j, t_*)$ as follows

$$S_2(j, t_*) \leq \sup_{t \leq t_*} \left| \sigma_\varepsilon \big(\Phi_{\psi(\varepsilon)}(B_{\varepsilon t}) - \Phi_{\psi(c^j)}(B_{c^j t}) \big) \right| + \sup_{t \leq t_*} |(\sigma_\varepsilon - \sigma_{c^j}) \Phi_{\psi(c^j)}(B_{t c^j})|$$

$$\leq \|\sigma_\varepsilon\| d_{t_*}(\Phi_{\psi(\varepsilon)}(B_{\varepsilon \cdot}), \Phi_{\psi(c^j)}(B_{c^j \cdot})) + D_j \sup_{t \leq t_*} \frac{|B_{t c^j}|}{\sqrt{c^j \log \log c^{-j}}},$$

where D_j is a deterministic constant with $D_j \to 0$ as $j \to \infty$ and $\| \cdot \|$ denotes the matrix norm. Now, the assumptions of Lemma 2.9 are satisfied with $x_t = B_t$ solving (1.4) with $\tilde{b} = 0$ and $\tilde{\sigma}$ being the $d \times d$ identity matrix. For any $M > 0$, by Lemma 2.9 there is $c \in (0, 1)$ and $J_1 = J_1(\omega, c, M) > 0$ such that for any $j \geq J_1$ and $\varepsilon \in [c^{j+1}, c^j]$ we obtain

$$d_{t_*}(\Phi_{\psi(\varepsilon)}(B_{\varepsilon \cdot}), \Phi_{\psi(c^j)}(B_{c^j \cdot})) \leq \frac{1}{M(\|\sigma\| + 1)}.$$

By the standard LIL for Brownian motion, for any $c \in (0, 1)$, there exists $J_2 = J_2(\omega, c) > 0$ such that for any $j \geq J_2$ it follows that, almost surely,

$$\sup_{t \leq t_*} \frac{|B_{t c^j}|}{\sqrt{c^j \log \log c^{-j}}} \leq 2.$$

Overall, for any $M > 0$, there exists $\delta' > 0$ such that for any $c \in (1 - \delta', 1)$ there is $J_3 = J_3(\omega, c, M)$ such that for all $j \geq J_3$ one has

$$\sup_{s \leq T_\varepsilon(t_*)} |y_s^\varepsilon - y_s^{c^j}| \leq (2C_j(L)t_* + M^{-1} \|\sigma^\varepsilon\| D_j)e^{C_L t_*} < \delta$$

for all $\varepsilon \in [c^{j+1}, c^j]$. By increasing M and J_3 if necessary, Proposition 5.1 part (i) ensures $T_\varepsilon > t_*$, and therefore $T_\varepsilon(t_*) = t_*$, so that Assumption 2(i) is satisfied.

In order to establish Assumption 2 part (ii), fix $\varepsilon_0 \in (0, \varepsilon_*]$ and ω in a subset of Ω of full measure specified below and let $t < \tau_{t_*}(y_\cdot^{\varepsilon_0}(\omega))$ or $t = t^*$ if $y_{t_*}^{\varepsilon_0}(\omega) \neq \Delta$. Since $s \mapsto y_s^{\varepsilon_0}(\omega) : [0, t] \to U^*$ is continuous, there is a compact set $K \subset U^*$ such that the image of $[0, t]$ under the map $s \mapsto y_s^{\varepsilon_0}(\omega)$ is contained in K.

Fix any $\varepsilon \in (0, \varepsilon_*]$ and set $S_\varepsilon(\omega) := \inf\{t \geq 0 : y_t^\varepsilon \notin K\}$ and $S_\varepsilon(\omega, s) = S_\varepsilon(\omega) \wedge s$ for any $s > 0$. To simplify the notation, we drop the explicit dependence on ω and proceed as above to find that for any $s \leq t \leq t_*$

$$\sup_{v \leq S_\varepsilon(s)} |y_v^\varepsilon - y_v^{\varepsilon_0}| \leq \int_0^{S_\varepsilon(s)} |b_{\varepsilon_0}(y_v^{\varepsilon_0}) - b_{\varepsilon_0}(y_v^\varepsilon)| \, dv + \int_0^{S_\varepsilon(s)} |b_{\varepsilon_0}(y_v^\varepsilon) - b_\varepsilon(y_v^\varepsilon)| \, dv$$

$$+ \sup_{v \leq t} \left| \frac{\sigma_{\varepsilon_0}}{\sqrt{\varepsilon_0 r(\varepsilon_0)}} B_{v \varepsilon_0} - \frac{\sigma_\varepsilon}{\sqrt{\varepsilon r(\varepsilon)}} B_{v \varepsilon} \right|$$

$$\leq C_{\varepsilon, \varepsilon_0}(K)t_* + C_{\varepsilon_0}(K) \int_0^{S_\varepsilon(s)} \sup_{w \leq S_\varepsilon(v)} |y_w^\varepsilon - y_w^{\varepsilon_0}| \, dv$$

$$+ \sup_{v \leq t} \left| \frac{\sigma_{\varepsilon_0}}{\sqrt{\varepsilon_0 r(\varepsilon_0)}} B_{v \varepsilon_0} - \frac{\sigma_\varepsilon}{\sqrt{\varepsilon r(\varepsilon)}} B_{v \varepsilon} \right|,$$

where $C_{\varepsilon, \varepsilon_0}(K)$, $C_{\varepsilon_0}(K)$ are deterministic constants with $C_{\varepsilon, \varepsilon_0}(K) \to 0$ as $\varepsilon \to \varepsilon_0$. Gronwall's inequality, $t_* \leq 1$, and the almost sure path continuity of Brownian motion then imply

$$\sup_{v \leq S_\varepsilon(t)} |y_v^\varepsilon - y_v^{\varepsilon_0}| \leq C_{\varepsilon, \varepsilon_0}'(K) e^{C_{\varepsilon_0}(K)t}$$

for some $C_{\varepsilon, \varepsilon_0}'$ such that $C_{\varepsilon, \varepsilon_0}' \to 0$ as $\varepsilon \to \varepsilon_0$. By continuity, and $t < \tau_{t_*}(y_\cdot^{\varepsilon_0}(\omega))$ we obtain for ε sufficiently close to ε_0 that $S_\varepsilon(t) = t$, and therefore

$$\sup_{v \leq t} |y_v^\varepsilon - y_v^{\varepsilon_0}| \leq C_{\varepsilon, \varepsilon_0}(K) e^{C_{\varepsilon_0}(K)t}.$$

for any ε sufficiently close to ε_0. Passing $\varepsilon \to \varepsilon_0$ and using the definition of convergence in $\mathscr{E}_x([0, t_*]; U^*)$, we obtain the desired result. $\square$

7. Application: Criteria for regular points on boundary of a bounded domain in $\mathbf{R}^d$

Throughout this section, for simplicity we suppose $U = U^* = \mathbf{R}^d$ and $V \subset \mathbf{R}^d$ is a non-empty, open set. We moreover suppose that $\partial V := \overline{V} \setminus V \subset \mathbf{R}^d$ is non-empty, where $\overline{V}$ denotes the closure of V in $\mathbf{R}^d$.

For $x \in \partial V$, our goal is to use Theorem 2.6 to deduce criteria for the diffusion x_t solving (1.4) to be *regular* at x. Specifically, we say that $x \in \partial V$ is *regular for* (x_t, V) if

$$\mathbf{P}_x\{\tau_{\overline{V}} > 0\} = 0,$$

where

$$\tau_{\overline{V}} = \inf\{t > 0 : x_t \notin \overline{V}\}. \tag{7.1}$$

We call x *irregular for* (x_t, V) otherwise.

Remark 7.1. Note that $x \in \partial V$ irregular for (x_t, V) means that x_t spends, with positive probability, a positive amount of time in $\overline{V}$ before exiting $\overline{V}$. Because the event $\{\tau_{\overline{V}} > 0\}$ belongs to the germ σ-field $\bigcap_{t>0} \mathscr{F}_t$, Blumenthal's $0-1$ law implies that this event either has probability 0 or 1. Thus $x \in \partial V$ is irregular for (x_t, V) if and only if $\mathbf{P}_x\{\tau_{\overline{V}} > 0\} = 1$.

In order to state the main result of this section we recall (cf. (2.9)) the deterministic system associated to (2.1):

$$\begin{cases} \dot{g}_t = b(g_t) + \sigma(g_t)\dot{f}_t, \\ g_0 = x. \end{cases} \tag{7.2}$$

In this section, we view (7.2) as a control problem, with controls f belonging to the class (cf. (2.11))

$$\mathscr{C}_1 = \{f \in \mathscr{C}^0([0, 1]; \mathbf{R}^k) : \tfrac{1}{2} \int_0^1 |\dot{f}_s|^2 \, ds \leq 1\}. \tag{7.3}$$

Let L be a compact set containing a neighborhood of $x \in \partial V$ and let $t_* > 0$ be as in (2.15). For any $t \in [0, t_*]$ and any $f \in \mathscr{C}_1$, let $S_x^t(f) \in L$ denote the solution of (7.2) at time t. For $t \in [0, t_*]$, let

$$\mathscr{A}(x, t) = \{y \in U : S_x^t(f) = y \text{ for some } f \in \mathscr{C}_1\}, \tag{7.4}$$

where $\mathscr{C}_1$ is as in (7.3) and

$$\mathscr{A}(x, \leq t_*) = \bigcup_{t \in [0, t_*]} \mathscr{A}(x, t). \tag{7.5}$$

Consider the processes x_t and y_t^ε defined by (1.4) and (2.1), respectively, both having initial condition $x \in \partial V$.

Definition 7.1. We say that a point $z \in \overline{V}^c$ is *asymptotically invariant at* x if there exists $\delta > 0$ such that the following statement holds almost surely: whenever $y_t^\varepsilon \in B_\delta(z)$ (y^ε solves (2.1)) for some $t \in (0, t_*]$ and some $\varepsilon \in (0, \varepsilon_*]$, then $x_{\varepsilon t} \in \overline{V}^c$.

Let $\mathscr{I}_x \subset \overline{V}^c$ be the set of asymptotically invariant points at x.

Example 7.1. Let $x = 0 \in \mathbf{R}^d$ and

$$V = \{y \in \mathbf{R}^d \,:\, y_d < 0\}.$$

Suppose that $\Phi_\alpha : \mathbf{R}^d \to \mathbf{R}^d$, $\alpha > 0$, is of the form

$$\Phi_\alpha(y) = (y_1 \alpha_1^{-1}, y_2 \alpha_2^{-1}, \dots, y_d \alpha_d^{-1})$$

and $\psi : [0, \varepsilon_*] \to [0, \infty)^d$ is an asymptotic index. If $y_t^\varepsilon := \Phi_{\psi(\varepsilon)}(x_{\varepsilon t})$, then any $z \in \overline{V}^c$ is asymptotically invariant at x. Indeed, choose any $\delta > 0$ such that $B_\delta(z) \subset V^c$ and use that $\psi_d(\varepsilon) > 0$.

We have the following result, which is a consequence of Theorem 2.6.

Theorem 7.2. *Suppose that Assumptions 1 and 2 are satisfied. Then, $x \in \partial V$ is regular for (x_t, V) if*

$$\mathscr{A}(x, \leq t_*) \cap \mathscr{I}_x \neq \emptyset. \tag{7.6}$$

Proof. Suppose $z \in \mathscr{A}(x, \leq t_*) \cap \mathscr{I}_x$. Since z is asymptotically invariant, there exists $\delta > 0$ such that whenever $y_t^\varepsilon \in B_\delta(z)$ we have $x_{\varepsilon t} \in \overline{V}^c$, almost surely. Since $z \in \mathscr{A}(x, \leq t_*)$, there exist $t \in (0, t_*]$ and an $f \in \mathscr{C}_1$ such that $S_x^t(f) = z \in \overline{V}^c$. Let $g = S_x(f) \in \mathscr{C}^0([0, t_*]; \mathbf{R}^d)$ and note that by definition, $g \in \mathscr{K}_x(t_*)$. If $\lambda(g) = 1$, then as in the proof of Theorem 2.6, we can find g^* with $\lambda(g^*) < 1$ and $\sup_{t \in [0, t^*]} |g_t - g_t^*| < \delta/2$. If, on the other hand, $\lambda(g) < 1$ we simply set $g^* = g$. By Proposition 5.2, there exists a deterministic sequence $\varepsilon_n > \varepsilon_{n+1} > 0$ with $\varepsilon_n \to 0$ such that

$$\mathbf{P}\left\{d_{t_*}(y^{\varepsilon_n}, g^*) < \frac{\delta}{2} \text{ for infinitely many } n\right\} = 1,$$

and consequently

$$\mathbf{P}\left\{d_{t_*}(y^{\varepsilon_n}, g) < \delta \text{ for infinitely many } n\right\} = 1.$$

Since $g_t = z$, then

$$\mathbf{P}\left\{|y_t^{\varepsilon_n} - z| < \delta \text{ for infinitely many } n\right\} = 1.$$

Thus by the asymptotic invariance of z, $x_{\varepsilon_n t} \in \overline{V}^c$ for infinitely many n, almost surely. Hence,

$$\mathbf{P}_x\{\tau_{\overline{V}} = 0\} \geq \mathbf{P}\{d_{t_*}(y^{\varepsilon_n}, g) < \delta \text{ for infinitely many } n\} = 1$$

and the proof is finished. $\square$

The next two results provide sufficient conditions on the noise that guarantee a given boundary point is regular. Before proceeding, we let $\mathscr{R}(A)$ denote the range of the matrix A, or equivalently the space spanned by the columns of A.

Proposition 7.3. *Let $x \in \partial V$ and suppose that $\tilde{b} \in C^\infty(\mathbf{R}^d; \mathbf{R}^d)$ and $\tilde{\sigma} \in C^\infty(\mathbf{R}^d; M_{d \times k})$. Assume there exists $v \in \mathscr{R}(\tilde{\sigma}(x))$ satisfying the following two properties:*

(qi) $x + \lambda v \in \overline{V}^c$ for all $\lambda \in (0, 1]$.
(qii) For all $\lambda \in (0, 1]$, there exists $\delta_\lambda > 0$ such that if $\delta_{\varepsilon, \lambda} = \sqrt{\varepsilon \log \log \varepsilon^{-1}} \delta_\lambda$ and $\lambda_\varepsilon = \lambda \sqrt{\varepsilon \log \log \varepsilon^{-1}}$, then

$$B_{\delta_{\varepsilon,\lambda}}(x + \lambda_\varepsilon v) \subset \overline{V}^c$$

for all $\varepsilon > 0$ small enough.

Then, x is regular for (x_t, V).

Proof. For any multiindex $\alpha > 0$, let $\Phi_\alpha^1, \Phi_\alpha : \mathbf{R}^d \to \mathbf{R}^d$ be given by

$$\Phi_\alpha^1(y) = (y_1 \alpha_1^{-1}, \ldots, y_d \alpha_d^{-1}) \qquad \text{and} \qquad \Phi_\alpha(y) = \Phi_\alpha^1(y - x) + x.$$

By Example 2.1, $\{\Phi_\alpha\}_{\alpha>0}$ is a family of weak contractions centered at $x \in \partial V \subset \mathbf{R}^d$. If we define, for $\varepsilon_* > 0$ small enough, $\psi : [0, \varepsilon_*] \to [0, \infty)^d$ as

$$\psi(\varepsilon) = (\sqrt{\varepsilon \log \log \varepsilon^{-1}}, \ldots, \sqrt{\varepsilon \log \log \varepsilon^{-1}}),$$

then by Example 2.2, ψ is an asymptotic index. For $\varepsilon \in (0, \varepsilon_*]$, define

$$y_t^\varepsilon = \Phi_{\alpha(\varepsilon)}(x_{\varepsilon t}). \tag{7.7}$$

We thus see that for all $t < \tau_1(x.)\varepsilon^{-1}$ (see (2.4) for the definition of $\tau_1(x.)$), y_t^ε satisfies an SDE of the form (2.1) (c.f. (2.24)–(2.26)) and it is easy to check that

$$b_\varepsilon(y) \to 0 \qquad \text{and} \qquad \sigma_\varepsilon(y) \to \tilde{\sigma}(x) \tag{7.8}$$

as $\varepsilon \to 0$ for every $y \in \mathbf{R}^d$ with the convergence above uniform on compact subsets of $\mathbf{R}^d$. Furthermore, $b_\varepsilon, \sigma_\varepsilon$ are locally Lipschitz on $\mathbf{R}^d$ for every $\varepsilon \in (0, \varepsilon_*]$. Thus Assumption 1 and, by Lemma 2.9, Assumption 2 are both satisfied.

The associated deterministic system is

$$\begin{cases} \dot{g}_t = \tilde{\sigma}(x) \dot{f}_t, \\ g_0 = x, \end{cases} \tag{7.9}$$

where $f \in \mathscr{C}_1 = \{h \in \mathscr{C}^0([0, 1]; \mathbf{R}^k) : \frac{1}{2} \int_0^1 |\dot{f}_s|^2 ds \leq 1\}$. Note that (7.9) has constant coefficients as $x \in \partial V$ is the initial condition, which is fixed. Let $v \in \mathscr{R}(\tilde{\sigma}(x))$ satisfy (qi) and (qii). In particular, $v = \tilde{\sigma}(x)w$ for some $w \in \mathbf{R}^k$. Hence, for any small enough $\lambda \in (0, 1]$, $f_t := \lambda t w \in \mathscr{C}_1$ and then $g_t = x + \lambda v t$. Also, there is $\lambda^0 > 0$ such that all points $z \in \mathbf{R}^d$ of the form

$$z = x + \lambda v \qquad \lambda \in (0, \lambda^0]$$

belong to $\mathscr{A}(x, \leq 1)$. Since $\overline{V}^c$ is open, there exists $\delta_{\lambda^0} > 0$ such that $B_{\delta_{\lambda^0}}(x + \lambda^0 v) \subset \overline{V}^c$, and then by property (qii), $B_{\delta_{\varepsilon,\lambda^0}}(x + \lambda_\varepsilon^0 v) \subset \overline{V}^c$ for all $\varepsilon > 0$ small enough. But, almost surely, $y_t^\varepsilon \in B_{\delta_{\lambda^0}}(x + \lambda^0 v)$ if and only if $x_{\varepsilon t} \in B_{\delta_{\varepsilon,\lambda^0}}(x + \lambda_\varepsilon^0 v) \subset \overline{V}^c$. Hence,

$$x + \lambda^0 v \in \mathscr{A}(x, \leq 1) \cap \mathscr{I}_x \neq \emptyset$$

and the proof follows from Theorem 7.2. $\quad\square$

Corollary 7.4. *Let $x \in \partial V$ and suppose that $\tilde{b} \in C^\infty(\mathbf{R}^d; \mathbf{R}^d)$ and $\tilde{\sigma} \in C^\infty(\mathbf{R}^d; M_{d \times k})$. Assume that there exist a unit vector $n(x)$ and $\delta > 0$ such that $B_\delta(x + \delta n(x))$ is tangent to ∂V at x and is contained in $\overline{V}^c$ (also known as the exterior sphere condition). If there is $w \in \mathscr{R}(\tilde{\sigma}(x))$ with $w \cdot n(x) > 0$, then x is regular for (x_t, V).*

Proof. Without loss of generality, we can rotate and shift the set V so that $x = 0$ and $n(0) = e_d$. In what follows, we thus assume $B_\delta(\delta e_d) \subset \overline{V}^c$ is tangent to ∂V at 0 and there exists a vector $w \in \mathscr{R}(\tilde{\sigma}(0))$ such that $w \cdot n(0) = w_d > 0$. Since $w_d > 0$, there exists $\lambda_0 > 0$

$$\lambda w \in B_\delta(\delta e_d) \subset \overline{V}^c \qquad \text{for any} \quad \lambda \in (0, \lambda_0].$$

Hence, for the choice of $v = \lambda_0 w$, Proposition 7.3 part (qi) is satisfied because $v \in \mathscr{R}(\tilde{\sigma}(0))$. Since $B_\delta(\delta e_d)$ is open, there exists $\delta' > 0$ such that $B_{\delta'}(v) \subset B_\delta(\delta e_d)$. By convexity, whenever $v' \in B_{\delta'}(v)$ then $\lambda v' \in B_\delta(\delta e_d)$ for all $\lambda \in (0, 1]$. Since $v' \in B_{\delta'}(v)$ if and only if $\sqrt{\varepsilon \log \log \varepsilon^{-1}} v' \in B_{\delta_\varepsilon}(v_\varepsilon)$ for any small $\varepsilon > 0$, property (qii) of Proposition 7.3 follows. An application of Proposition 7.3 finishes the proof. $\square$

Using a nearly identical proof, we can also obtain the following result which is the so-called *exterior cone condition*.

Corollary 7.5. *Let $x \in \partial V$ and suppose that $\tilde{b} \in C^\infty(\mathbf{R}^d; \mathbf{R}^d)$ and $\tilde{\sigma} \in C^\infty(\mathbf{R}^d; M_{d \times k})$. Suppose that V satisfies the exterior cone condition at x; that is, there exists a basis $\{x_1, x_2, \ldots, x_d\}$ of $\mathbf{R}^d$ such that*

$$Cone(x; x_1, \ldots, x_d) := \{x + \lambda_1 x_1 + \cdots + \lambda_d x_d : \lambda_i \in (0, 1)\} \subset \overline{V}^c.$$

If the column space of $\tilde{\sigma}(x)$ contains a vector w such that $x + w \in Cone(x; x_1, \ldots, x_d)$, then x is regular for (x_t, V).

If $\mathscr{R}(\tilde{\sigma}(x))$ is not all of $\mathbf{R}^d$, then identifying regular points for (x_t, V) on ∂V can be complicated, because one has to know the almost sure dynamics near the boundary point. Moreover, the method used in the proof of Proposition 7.3 is not sufficient to characterize all points. However, as the next examples illustrate, the techniques developed here can be still useful.

Example 7.2. As in Example 3.2, we consider the iterated Kolmogorov equation in dimension $d = 2$ and assume $V = B_r(0)$ for a given $r > 0$. Corollary 7.4 implies that all points on $x_0 = (x_1(0), x_2(0)) \in \partial V$ with $x_2(0) \neq 0$ have normal vector with non-zero second component, and therefore are regular for (x_t, V). On the other hand, if $x_2(0) = 0$, then $n(x_0) = (\pm 1, 0)$ and relations (3.10) and (3.11) with $x_2(0) = 0$ imply that points $(\pm r, 0)$ are also regular for (x_t, V). Note that the same result holds if $V = B_r(0)^c$.

Example 7.3. Next, consider the same problem as in Examples 7.2 and 3.2 but with $V \subset \mathbf{R}^2$ assumed to be a general bounded open set with C^∞ boundary ∂V. Then, by Corollary 7.4, all points $x_0 \in \partial V$, where the outward unit normal $n(x_0) = (n_1(x_0), n_2(x_0))$ to ∂V has $n_2(x) \neq 0$ are regular, on the other hand, $n(x_0) = (\pm 1, 0)$ for some $x_0 \in \partial V$, then (3.10) and (3.11)

imply that x_0 is regular if and only if $x_2(0) \geq 0$ and $n(x_0) = (+1, 0)$, or $x_2(0) \leq 0$ and $n(x_0) = (-1, 0)$.

7.1. Modification of the boundary

If the domain V is not apriori specified, the idea of this section is to slightly modify V so that all points on the boundary are regular. We do this using polygonal approximations under the assumption that there is noise in a uniform direction on the boundary and V is convex and bounded with *non-flat C^1* boundary ∂V; that is, ∂V is C^1 and for each $x \in \partial V$ and any $r > 0$, the set $\partial V \cap B(x, r)$ is not a subset of a hyperplane. The latter condition is satisfied, for example, if V is strictly convex, or if at each $x \in \partial V$ there is at least one non-zero principal curvature.

The construction of our polygonal approximations makes use of convex hulls of randomly chosen points on the boundary. There are many different ways to do the selection of points, but here we do it according to Hausdorff measure π on ∂V. That is, we will choose sufficiently many vertices independently and according to law of π. Intuitively, for a large number of vertices, the convex hull of these points should be *close* to V. The main result in Schütt and Werner [35] makes this precise on a set of high probability. It turns out that such a resulting polygon cannot have, almost surely, any face parallel to the uniform direction in which the noise acts on the boundary. Thus, we can then apply Corollary 7.5.

Remark 7.6. Polygonal approximation of smooth domains is used, for example, in finite element method (FEM) to numerically solve differential equations.

Theorem 7.7. *Let $d \geq 2$, $v \in \mathbf{R}^d$ be a unit vector, and suppose that $V \subset \mathbf{R}^d$ is convex, bounded and non-flat C^1 boundary ∂V. Consider a probability space $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \mathbf{Q})$ such that $\xi_1, \xi_2, \xi_3 \ldots$ are i.i.d. random variables defined on $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \mathbf{Q})$ with distribution π. Then, for every $\varepsilon > 0$, there exist $n(\varepsilon) > 0$ and a set S_ε with $\mathbf{Q}(S_\varepsilon) > 1 - \varepsilon$ such that the following properties are satisfied:*

(i) *For every collection of d-points in $\{\xi_1, \xi_2, \ldots, \xi_{n(\varepsilon)}\}$ there exists a unique hyperplane H that contains those d points. Furthermore, H is not parallel to v.*
(ii) *If $[\xi_1, \xi_2, \ldots, \xi_{n(\varepsilon)}]$ denotes the closed convex hull of $\xi_1, \xi_2, \ldots, \xi_{n(\varepsilon)}$, then*

$$|\overline{V}| - |[\xi_1, \xi_2, \ldots, \xi_{n(\varepsilon)}]| < \varepsilon,$$

where $|A|$ denotes the Lebesgue measure of A in $\mathbf{R}^d$.

In order to prove Theorem 7.7, we first establish the following auxiliary result.

Lemma 7.8. *Suppose $d \geq 2$ and that $V \subset \mathbf{R}^d$ is convex with C^1 non-flat boundary ∂V. Then, any hyperplane in $\mathbf{R}^d$ intersects ∂V only on a set of π-measure zero.*

Remark 7.9. Note that in this result, we may drop the hypothesis that V is bounded.

Proof of Lemma 7.8. Without loss of generality, assume that the hyperplane $\mathscr{P}$ has normal vector $e_1 = (1, 0, \ldots, 0)$ and $\mathscr{P} \cap \partial V \neq \emptyset$. Define

$$\ell_+ = \sup\{\alpha \in \mathbf{R} : \partial V \cap \{x : x_1 = \alpha\} \neq \emptyset\},$$
$$\ell_- = \inf\{\alpha \in \mathbf{R} : \partial V \cap \{x : x_1 = \alpha\} \neq \emptyset\}.$$

If the set $\{\alpha \in \mathbf{R} \ : \ \partial V \cap \{x \ : \ x_1 = \alpha\}\}$ is not bounded above (respectively below), we set $\ell_+ = \infty$ (respectively $\ell_- = -\infty$). Note that if $\ell_+ \in \mathbf{R}$, ℓ_+ corresponds to the first coordinate of the *rightmost* point on ∂V. Similarly, if $\ell_- \in \mathbf{R}$, ℓ_- corresponds to the first coordinate of the *leftmost* point. Observe that $\mathscr{P} \cap \partial V \neq \emptyset$ implies $\mathscr{P} = \{x \ : \ x_1 = \alpha\}$ for some α with $\alpha \in [\ell_-, \ell_+]$, $|\alpha| < \infty$.

First assume $\alpha = \ell_+ < \infty$. Without loss of generality, we may assume $\ell_+ = 0$ and $0 \in \mathscr{P} \cap \partial V$, otherwise we shift both sets. To every $X \in \mathscr{P} \cap \partial V$, we associate its position vector $x = \overrightarrow{0X}$. Let k be largest number of points $\{X_1, X_2, \ldots, X_k\} \subset \mathscr{P} \cap \partial V$ such that the corresponding set of position vectors $\{x_1, x_2, \ldots, x_k\}$ is linearly independent. If $k < d - 1$, then $\mathscr{L} := \text{span} \{x_1, \ldots, x_k\}$ is $k \leq d - 2$ dimensional linear space, and therefore $\partial V \cap \mathscr{P} \subset \mathscr{L} \cap \mathscr{P}$ is at most $(d - 2)$-dimensional. Thus, $\pi(\partial V \cap \mathscr{P}) = 0$ as desired. If $k = d - 1$, then by the convexity of V and the fact that $\alpha = \ell_+$, any convex combination of $\{x_1, x_2, \ldots, x_{d-1}\}$ belongs to $\partial V \cap \mathscr{P}$. However, this implies that ∂V is locally a hyperplane about some point on ∂V, a contradiction to the fact that ∂V is non-flat.

A similar argument can be applied in the case when $\alpha = \ell_- > -\infty$.

Finally, suppose $\alpha \in (\ell_-, \ell_+)$. Fix $x \in \partial V \cap \mathscr{P}$ and note that the normal vector to ∂V at x is not parallel to e_1. Indeed, otherwise $\mathscr{P}$ is tangent to V and by convexity, V lies on one side of $\mathscr{P}$, a contradiction to $\alpha \in (\ell_-, \ell_+)$.

As above, we can without loss of generality assume $\alpha = 0$ and $x = 0$. We parametrize ∂V as $\partial V = \{y \in \mathbf{R}^d \ : \ \Phi(y) = 0\}$ for some C^1-function $\Phi : \mathbf{R}^d \to \mathbf{R}$ with $\Phi(0) = 0$ and, by the claim, $\nabla \Phi(0) \neq \lambda e_1$ for any $\lambda \in \mathbf{R}$. Then,

$$\partial V \cap \mathscr{P} \subset \{y = (0, y_2, \ldots, y_d) : \Phi(0, y_2, \ldots, y_d) = 0\}. \tag{7.10}$$

To solve $\Phi(0, y_2, \ldots, y_d) = 0$ we note that $\Phi(0) = 0$ and there exists $j \geq 2$ such that $\partial_{x_j} \Phi(0) \neq 0$. By the implicit function theorem, $y_j = \phi(y_2, \ldots, y_{j-1}, y_{j+1}, \ldots, y_{d-1})$ locally for some C^1 function ϕ, and therefore $\partial V \cap \mathscr{P}$ is locally a $(d - 2)$-dimensional manifold. In particular $\pi(\partial V \cap \mathscr{P} \cap B_\rho) = 0$ for some $\rho > 0$.

Since the countable union of set of zero measure is also set of zero measure, the result follows. $\square$

We are now ready to prove Theorem 7.7.

Proof of Theorem 7.7. For any measurable $A_i \subset \partial V$ and any n, we have

$$\mathbf{Q}\{\xi_1 \in A_1, \ldots, \xi_n \in A_n\} = \pi(A_1)\pi(A_2)\ldots\pi(A_n).$$

For any collection of points $y_1, \ldots, y_k$ denote the set of vectors $\{\overrightarrow{y_1 y_j}, j > 1\}$ by $\langle y_1, \ldots, y_k \rangle$. Observe that $\langle y_1, \ldots, y_k \rangle$ depends on the arrangement of points, but the span of the vectors $\langle y_1, \ldots, y_k \rangle$ does not. Note that z belongs to the affine space defined by $y_1, \ldots, y_k$ if an only if $z \in y_1 + \text{span}\{\langle y_1, \ldots, y_k \rangle\} =: \mathscr{A}(y_1, \ldots, y_k)$. Also, we denote $y_1, \ldots \hat{y}_j, \ldots, y_k$ the sequence of points (or similarly vectors), where the point y_j is omitted from the list. For $i_1, i_2, \ldots, i_\ell \in \mathbf{N}$ all distinct with $\ell \leq d + 1$, let

$$D_{i_1 i_2 \ldots i_\ell} = \{\omega \in \widetilde{\Omega} \ : \ \langle \xi_{i_1}(\omega), \xi_{i_2}(\omega), \ldots, \xi_{i_\ell}(\omega) \rangle \text{ are linearly independent}\}. \tag{7.11}$$

We claim that $\mathbf{Q}(D_{i_1 i_2 \ldots i_\ell}) = 1$. Indeed, since $i_1, i_2, \ldots, i_\ell \in \mathbf{N}$ are all distinct with $\ell \leq d+1$, then by Lemma 7.8

$$\mathbf{Q}(D^c_{i_1 i_2 \ldots i_\ell}) \leq \sum_{j=2}^{\ell} \mathbf{Q}\{\overrightarrow{\xi_{i_1}\xi_{i_j}} \in \mathrm{span}\{\langle \xi_{i_1}, \ldots, \hat{\xi}_{i_j}, \ldots, \xi_{i_\ell}\rangle\}\}$$

$$= \sum_{j=2}^{\ell} \int_{x_{i_j} \in \mathscr{A}(x_{i_1}, \ldots, \hat{x}_{i_j}, \ldots, x_{i_\ell}) \cap \partial V} \pi(dx_{i_j}) \pi^{\ell-1}(dx_{i_1} \ldots d\hat{x}_{i_j} \ldots dx_{i_\ell}) = 0,$$

since $\mathscr{A}(x_{i_1}, \ldots, \hat{x}_{i_j}, \ldots, x_{i_\ell})$ is at most $(\ell-1)$-dimensional affine space which intersects ∂V on a set of π-measure 0 (see Lemma 7.8).

Fix a vector $v \in \mathbf{R}^d$. If $\langle x_1, x_2, \ldots, x_{d-1}\rangle$ are independent vectors, define

$$A_v(x_1, x_2, \ldots, x_{d-1}) = \{x_d \in \partial V : v \in \mathrm{span}\langle x_1, \ldots, x_d\rangle\}$$

and if $\langle x_1, \ldots, x_{d-1}\rangle$ are dependent define $A_v(x_1, x_2, \ldots, x_{d-1}) = \emptyset$. Note that if $v \notin \langle x_1, x_2, \ldots, x_d\rangle$ and $\langle x_1, x_2, \ldots, x_d\rangle$ are independent vectors, then $A_v(x_1, x_2, \ldots, x_d)$ is the intersection of ∂V and the affine $(d-1)$-dimensional space that contains points $x_1, x_2, \ldots, x_d$ and is parallel to v. For any distinct collection $i_1, i_2, \ldots, i_d \in \mathbf{N}$, $D_{i_1 i_2 \ldots i_d}$ is a set of full measure, and therefore

$$\mathbf{Q}(F_{i_1 i_2 \ldots i_d}) := \mathbf{Q}(\{\omega \in \widetilde{\Omega} : \mathscr{A}(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \ldots, \xi_{i_d}(\omega)) \text{ is parallel to } v\})$$

$$= \mathbf{Q}(\{\omega \in D_{i_1 i_2 \ldots i_d} : \mathscr{A}(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \ldots, \xi_{i_d}(\omega)) \text{ is parallel to } v\}).$$

Hence, by the Fubini–Tonelli theorem

$$\mathbf{Q}(F_{i_1 i_2 \ldots i_d}) = \mathbf{Q}(\{\omega \in D_{i_1 i_2 \ldots i_d} : \xi_{i_j}(\omega) \in A_v(\xi_{i_1}(\omega), \ldots, \hat{\xi}_{i_j}(\omega), \ldots, \xi_{i_d}(\omega))\})$$

$$= \int_{\{x_j \in A_v(x_1, \ldots, \hat{x}_j, \ldots, x_d)\}} \pi(dx_j) \pi^{d-1}(dx_1 \ldots \hat{x}_j \ldots dx_d) = 0,$$

where in the last equality we used Lemma 7.8 and the fact that $A_v(x_1, \ldots, \hat{x}_j, \ldots, x_d)$ is at most $(d-1)$-dimensional affine space, and has π-measure zero when intersected with ∂V.

By taking finite unions,

$$F := \bigcup_{\substack{i_1, i_2, \ldots, i_d \in \{1, \ldots, N\} \\ \text{all distinct}}} F_{i_1 i_2 \ldots i_d}, \tag{7.12}$$

we have $\mathbf{Q}(F) = 0$. Thus for any $N \geq d$ and any $\omega \in F^c$, with $\mathbf{Q}(F^c) = 1$ the realization

$$\xi_1(\omega), \xi_2(\omega), \ldots, \xi_N(\omega)$$

satisfies (i).

To obtain (ii), fix $\varepsilon > 0$ and by [35, Theorem 1.1] there is a sufficiently large $N \geq d+1$ such that

$$|V| - \mathbf{E_Q}|[\xi_1, \xi_2, \ldots, \xi_N]| = |V| - \mathbf{E_Q} \mathbf{1}_{F^c}|[\xi_1, \xi_2, \ldots, \xi_N]| < \varepsilon^2.$$

Since convexity of V implies $|[\xi_1, \xi_2, \ldots, \xi_N]| \leq |V|$, then by Chebyshev's inequality

$$\mathbf{Q}\{|V| - |[\xi_1, \xi_2, \ldots, \xi_N]| > \varepsilon\} \leq \frac{|V| - \mathbf{E_Q} \mathbf{1}_{F^c}|[\xi_1, \xi_2, \ldots, \xi_N]|}{\varepsilon} < \varepsilon, \tag{7.13}$$

which concludes the proof. $\square$

An almost immediate consequence of the previous result is the following corollary.

Corollary 7.10. *Suppose that $V \subset \mathbf{R}^d$, $d \geq 2$ is a non-empty, bounded, convex, non-flat domain with C^1 boundary ∂V. Assume $\tilde{b} \in C^\infty(\mathbf{R}^d; \mathbf{R}^d)$ and $\tilde{\sigma} \in C^\infty(\mathbf{R}^d; M_{d \times k})$ and there exists a unit vector $v \in \mathbf{R}^d$ such that $v \in \mathcal{R}(\sigma(x))$ for every $x \in \partial V$. Then, for every $\varepsilon > 0$, there exists a non-empty, open convex domain $D_\varepsilon \subset V$ with piecewise smooth boundary ∂D_ε such that every $y \in \partial D_\varepsilon$ is regular for (x_t, D_ε) and $|V| - |D_\varepsilon| < \varepsilon$.*

Proof. Fix $\varepsilon > 0$ and select points $x_1, x_2, \ldots, x_{n(\varepsilon)} \in \partial V$ satisfying both (i) and (ii) of Theorem 7.7. Define

$$D_\varepsilon = \mathrm{interior}([x_1, x_2, \ldots, x_{n(\varepsilon)}]).$$

and note that $|V| - |D_\varepsilon| < \varepsilon$ and ∂D_ε is piecewise C^2. Since no face on the boundary ∂D_ε is parallel to v, Corollary 7.5 implies the assertion. $\square$

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Athreya, T. Kolba, J.C. Mattingly, Propagating Lyapunov functions to prove noise-induced stabilization, Electron. J. Probab. 17 (2012) 1–38.

[2] R. Azencott, Grandes déviations et applications, in: Ecole d'Été de Probabilités de Saint-Flour VIII-1978, Springer, 1980, pp. 1–176.

[3] P. Baldi, Large deviations and functional iterated logarithm law for diffusion processes, Probab. Theory Related Fields 71 (3) (1986) 435–453.

[4] P. Baldi, M. Chaleyat-Maurel, An extension of Ventsel-Freidlin estimates, in: Stochastic Analysis and Related Topics, Springer, 1988, pp. 305–327.

[5] J. Bec, A. Celani, M. Cencini, S. Musacchio, Clustering and collisions of heavy particles in random smooth flows, Phys. Fluids 17 (7) (2005) 073301.

[6] J. Bec, M. Cencini, R. Hillerbrand, Clustering of heavy particles in random self-similar flow, Phys. Rev. E 75 (2) (2007) 025301.

[7] J. Bedrossian, K. Liss, Quantitative spectral gaps and uniform lower bounds in the small noise limit for Markov semigroups generated by hypoelliptic stochastic differential equations, 2020, arXiv preprint arXiv:2007.13297.

[8] J. Birrell, D.P. Herzog, J. Wehr, The transition from ergodic to explosive behavior in a family of stochastic differential equations, Stochastic Process. Appl. 122 (4) (2012) 1519–1539.

[9] E. Camrud, D.P. Herzog, G. Stoltz, M. Gordina, Weighted L^2-contractivity of langevin dynamics with singular potentials, 2021, arXiv preprint arXiv:2104.10574.

[10] L. Caramellino, Strassen's law of the iterated logarithm for diffusion processes for small time, Stochastic Process. Appl. 74 (1) (1998) 1–19.

[11] K.L. Chung, On the maximum partial sums of sequences of independent random variables, Trans. Amer. Math. Soc. 64 (2) (1948) 205–233.

[12] N. Cuneo, J.P. Eckmann, M. Hairer, L. Rey-Bellet, Non-equilibrium steady states for networks of oscillators, Electron. J. Probab. 23 (2018) 1–28.

[13] W. E, J.C. Mattingly, Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation, Comm. Pure Appl. Math. 54 (11) (2001) 1386–1402.

[14] J. Foldes, N.E. Glatt-Holtz, D.P. Herzog, Sensitivity of steady states in a degenerately-damped stochastic Lorenz system, 2020, arXiv preprint arXiv:2009.08429.

[15] N. Gantert, An inversion of Strassen's law of the iterated logarithm for small time, Ann. Probab. (1993) 1045–1049.

[16] K. Gawędzki, D.P. Herzog, J. Wehr, Ergodic properties of a model for turbulent dispersion of inertial particles, Comm. Math. Phys. 308 (1) (2011) 49–80.

[17] G. Giacomin, S. Olla, E. Saada, H. Spohn, G. Stoltz, Stochastic Dynamics Out of Equilibrium, Springer, 2019.
[18] P. Hartman, A. Wintner, On the law of the iterated logarithm, Amer. J. Math. 63 (1) (1941) 169–176.
[19] D.P. Herzog, J.C. Mattingly, Noise-induced stabilization of planar flows I, Electron. J. Probab. 20 (2015) 1–43.
[20] D.P. Herzog, J.C. Mattingly, A practical criterion for positivity of transition densities, Nonlinearity 28 (8) (2015) 2823–2845.
[21] D.P. Herzog, J.C. Mattingly, Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials, Comm. Pure Appl. Math. 72 (10) (2019) 2231–2255.
[22] D.P. Herzog, H.D. Nguyen, Stability and invariant measure asymptotics in a model for heavy particles in rough turbulent flows, 2021, arXiv preprint arXiv:2104.08629.
[23] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1) (1967) 147–171.
[24] A. Khintchine, Über einen satz der wahrscheinlichkeitsrechnung, Fundam. Math. 6 (1) (1924) 9–20.
[25] D. Khoshnevisan, Z. Shi, Chung's law for integrated Brownian motion, Trans. Amer. Math. Soc. (10) (1998) 4253–4264.
[26] A.E. Kogoj, A control condition for a weak Harnack inequality, Nonlinear Anal. 75 (11) (2012) 4198–4204.
[27] A.E. Kogoj, On the Dirichlet problem for hypoelliptic evolution equations: Perron–Wiener solution and a cone-type criterion, J. Differential Equations 262 (3) (2017) 1524–1539.
[28] A. Kolmogoroff, Über das gesetz des iterierten logarithmus, Math. Ann. 101 (1) (1929) 126–135.
[29] T. Lachal, Local asymptotic classes for the successive primitives of Brownian motion, Ann. Probab. (1997) 1712–1734.
[30] B. Leimkuhler, M. Sachs, G. Stoltz, Hypocoercivity properties of adaptive Langevin dynamics, SIAM J. Appl. Math. 80 (3) (2020) 1197–1222.
[31] T. Leliévre, G. Stoltz, M. Rousset, Free Energy Computations: A Mathematical Perspective, World Scientific, 2010.
[32] C. Mueller, A unification of Strassen's law and Levy's modulus of continuity, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 56 (2) (1981) 163–179.
[33] B. Øksendal, Stochastic Differential Equations, Stochastic Differential Equations, Springer, 2003.
[34] M. Romito, Ergodicity of the finite dimensional approximation of the 3d Navier–Stokes equations forced by a degenerate noise, J. Stat. Phys. 114 (1) (2004) 155–177.
[35] C. Schütt, E. Werner, Polytopes with vertices chosen randomly from the boundary of a convex body, in: Geometric Aspects of Functional Analysis, Springer, 2003, pp. 241–422.
[36] V. Strassen, An invariance principle for the law of the iterated logarithm, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 3 (3) (1964) 211–226.
[37] D. Stroock, S.R.S. Varadhan, On degenerate elliptic–parabolic operators of second order and their associated diffusions, Comm. Pure Appl. Math. 25 (6) (1972) 651–713.
[38] D.W. Stroock, S.R.S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 3, Univ. California, Berkeley, Calif., 1972, pp. 333–359, 1970/1971.
[39] M.C. Zelati, M. Hairer, A noise-induced transition in the Lorenz system, Comm. Math. Phys. 383 (3) (2021) 2243–2274.