ATYPICAL EXIT EVENTS NEAR A REPELLING EQUILIBRIUM

YURI BAKHTIN AND HONG-BIN CHEN

ABSTRACT. We consider exit problems for small white noise perturbations of a dynamical system generated by a vector field, and a domain containing a critical point with all positive eigenvalues of linearization. We prove that, in the vanishing noise limit, the probability of exit through a generic set on the boundary is asymptotically polynomial in the noise strength, with exponent depending on the mutual position of the set and the flag of the invariant manifolds associated with the top eigenvalues. Furthermore, we compute the limiting exit distributions conditioned on atypical exit events of polynomially small probability and show that the limits are Radon–Nikodym equivalent to volume measures on certain manifolds that we construct. This situation is in sharp contrast with the large deviation picture where the limiting conditional distributions are point masses.

CONTENTS

1. INTRODUCTION

This paper is a part of our program on long-term behavior of dynamical systems with multiple unstable equilibria organized into heteroclinic networks, under small noisy perturbations. The existing work in this direction (see [SH90], [SA99], [ASK03] for early analysis with elements of heuristics and [Bak10], [Bak11], [AMB11], [BPG19], [BPG18], [BC19] for rigorous analysis) is a departure from the classical Freidlin–Wentzell (FW) theory of metastability. In FW, rare transitions can be described via large deviations

theory and happen at rates exponential in $-\epsilon^{-2}$ where ϵ is the strength of the perturbation:

$$(1.1) \qquad\qquad dX_t^\epsilon = b(X_t^\epsilon)dt + \epsilon\sigma(X_t^\epsilon)dW_t.$$

In [Kif81], it was shown that the exit from a neighborhood of an unstable critical point of b happens in time of the order of $\log\epsilon^{-1}$, in the most unstable direction, along the invariant manifold associated to the top eigenvalue of the linearization of the vector field b. In the case where the top eigenvalue is not simple, the limit of the exit location distribution was studied in [Eiz84].

In the case of a simple top eigenvalue, the results of [Kif81] were strengthened in [Bak10], [Bak11], [AMB11], where scaling limits for the distributions of exit locations were obtained and used to compute asymptotic probabilities of various pathways through the network. In particular, it turned out that there are interesting memory effects and in general the typical limiting behavior at logarithmic timescales is not simply a random walk on the graph of heteroclinic connections.

To study the dynamics over longer times, one has to study the rare events realizing unusual transitions that are improbable over logarithmic time scales, see the discussion of heteroclinic networks in [BPG18]. It was also understood in [BPG19] and [BPG18] that the leading contribution to these rare events is due to abnormally long stays in the neighborhood of the critical point. Asymptotic results on the decay of probabilities of these events were obtained for repelling equilibria in these papers for dimension 1 and in [BC19] for higher dimensions. The general results of [BC19] can be briefly summarized as follows. If all the eigenvalues of the linearization at the critical point are positive and simple and the leading one is $\lambda > 0$, then, for all $\alpha > 1/\lambda$ and initial conditions at distance of the order of ϵ to the critical point, it was shown that

$$\mathbb{P}\{\tau > \alpha\log\epsilon^{-1}\} = c\epsilon^\beta(1 + o(1)), \quad \epsilon \to 0,$$

where c and β were explicitly computed. Note that this is a more precise estimate than $\log\mathbb{P}\{\tau > \alpha\log\epsilon^{-1}\} = (\beta\log\epsilon)(1 + o(1))$ conjectured in [Mik95].

In the present paper, we extend the study of atypical exit times from [BC19] to the study of atypical exit locations in the same setting. We assume that the dynamics near the critical point (which we place at the origin in $\mathbb{R}^d$) admits a smooth conjugacy to the linear dynamics with simple characteristic exponents $\lambda_1 > \ldots > \lambda_d > 0$ and consider a neighborhood $\mathbf{D}$ of the origin, with smooth boundary $\partial\mathbf{D}$. For any subset A of $\partial\mathbf{D}$ possessing a certain regularity property (most relatively open subsets of $\partial\mathbf{D}$ fall into this category), we prove that

$$(1.2) \qquad\qquad \mathbb{P}\{X_\tau \in A\} = \epsilon^{\rho(A)}\mu(A)(1 + o(1)), \quad \epsilon \to 0,$$

where $\rho(A)$ and $\mu(A)$ are constants. The values that the exponent $\rho(A)$ can take belong to a discrete set of values $(\rho_i)_{i=1}^d$:

$$(1.3) \qquad\qquad \rho_i = \sum_{j<i}\left(\frac{\lambda_j}{\lambda_i} - 1\right), \quad i = 1, 2, \ldots, d.$$

Here and throughout this paper, the sum over an empty set is understood to be 0.

The relevant index $i = i(A)$ to be used in (1.3), i.e, such that $\rho(A) = \rho_{i(A)}$ in (1.2) is defined in the following way. For each $i = 1, \ldots, d$, there is a uniquely defined i-dimensional manifold M^i invariant under the flow generated by the drift vector field b, with tangent space at the origin spanned by eigenvectors associated with eigenvalues $\lambda_1, \ldots, \lambda_i$. These manifolds form a flag, i.e., $M^1 \subset M^2 \subset \ldots \subset M^d$, their traces on $\partial \mathbf{D}$ defined by $N^i = M^i \cap \partial \mathbf{D}$ also satisfy $N^1 \subset N^2 \subset \ldots \subset N^d$ and, additionally, $N^d \cap \partial \mathbf{D} = \partial \mathbf{D}$ and thus $i(A) = \min\{i \in \{1, \ldots, d\} : N^i \cap A \neq \emptyset\}$ is always well-defined, see Figure 1.

Our results mean that exits along manifolds of various dimensions have probabilities of different polynomial decay rates. Since $0 = \rho_1 < \rho_2 < \ldots \rho_d$, these probabilities are of the order of $\epsilon^{\rho_1}(= \epsilon^0 = 1) \gg \epsilon^{\rho_2} \gg \ldots \gg \epsilon^{\rho_d}$. The differences in the order of magnitude for these probabilities are due to a drastic distortion caused by exponential expansion with different rates in different eigendirections. One can say that the exit direction of the system is largely determined by its behavior in infinitesimal time which is then amplified by exponential growth with different rates in different directions.

In agreement with the results of [Kif81], exiting in the neighborhood of a two-point set N^1 (M^1 is a 1-dimensional manifold, i.e., a curve, associated to the most unstable direction) is a typical event which has asymptotic probability 1. Exiting away from it happens with probability of the order of ϵ^{ρ_2} and, conditioned on this polynomially rare event, the exit distribution concentrates on N_2. In general, exiting away from N_k is a rare event of probability of the order of $\epsilon^{\rho_{k+1}}$ and, conditioned on this rare event, the exit distribution concentrates on N_{k+1}. Moreover, these conditional distributions converge weakly, as $\epsilon \to 0$, to a limiting measure that is Radon–Nikodym equivalent to the k-dimensional volume on N_{k+1}, with a density that can be described explicitly. The basic case where the domain $\mathbf{D}$ is a cube and the vector field b is linear is at the heart of the analysis. It turns out that the limiting distributions of exit locations conditioned on exits through various faces of the cube show equidistribution properties that cannot be obtained through large deviation estimates and are surprising if one is used to the FW mindset.

We discuss the simple situation described above and build our intuition in Section 2. In Section 3 we give the general setting and our main results in detail. The proofs are given in Sections 4—6.

The techniques that we are using are primarily probabilistic. Most are based on the classical stochastic calculus tools and the key estimate is based on Malliavin calculus. In principle, exit problems can be addressed using PDE tools. For exits near unstable critical points, some elements of PDE-based analysis can be found in [Kif81], [Day95], and [CGLM13]. So it would be interesting to find a PDE approach to the problem solved in the present paper but we follow the path of probabilistic analysis using the basic approach similar to [Eiz84], [Day95], [Bak10], [Bak11], [AMB11], [Bak15], [BPG19], [BPG18], [BC19].

Concluding the introduction, let us briefly discuss two directions that will be natural continuations of the present work.

Although our new results and those on exit times from [BC19] are based on the same density estimates, we do not develop that connection further in this paper. In particular,
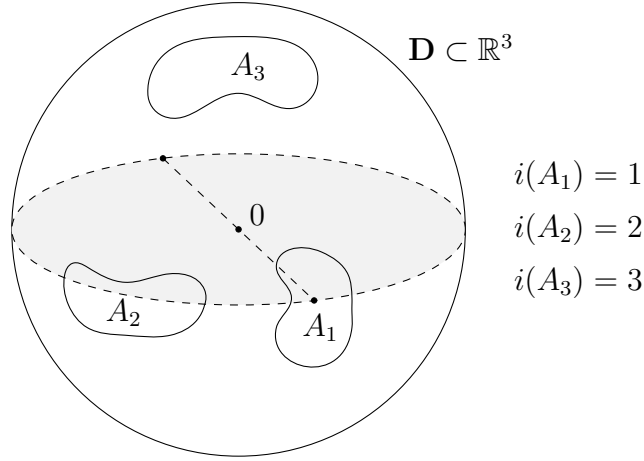
FIGURE 1. The dashed line segment and the shaded surface are the portions of M^1 and M^2 inside $\mathbf{D}$, respectively. The sets A_i, $i = 1, 2, 3$, lie on $\partial \mathbf{D}$. Probabilities of exit through these sets have different decay rates: there are constants c_1, c_2, c_3 such that $\mathbb{P}\{X_\tau \in A_1\} \to c_1$, $\mathbb{P}\{X_\tau \in A_2\} \sim c_2 \epsilon^{\rho_2}$, $\mathbb{P}\{X_\tau \in A_3\} \sim c_3 \epsilon^{\rho_3}$. Moreover, conditioned on $\{X_\tau \in A_2\}$, the distribution of X_τ has a limit concentrated on $A_2 \cap N_2 = A_2 \cap M_2$ and equivalent to the length measure on this curve; conditioned on $\{X_\tau \in A_3\}$, the distribution of X_τ has a limit equivalent to the area.

the detailed asymptotic analysis of the joint distribution of exit locations and exit times seems possible but harder, and we postpone it to another publication.

A more important question is the asymptotic behavior of exit distributions near hyperbolic critical points (saddles) of the driving vector field. Atypical events described in terms of the exit location are responsible for atypical transitions in heteroclinic networks. Similarly to the situation in this paper, their probability is expected to decay polynomially in ϵ leading to a hierarchy of transitions observable at various polynomial time scales, see the heuristic analysis in [BPG18]. The approach of the present paper based on Malliavin calculus density estimates from [BC19], will be an important ingredient in making this analysis rigorous in another forthcoming publication.

2. A HEURISTIC COMPUTATION FOR A SIMPLE CASE

Let us give a heuristic analysis of the simplest situation with exit distribution behavior that is counterintuitive from the point of view of the FW theory.

Suppose the diffusion $X = X^\epsilon$ in question is two-dimensional:

$$dX_t^1 = \lambda_1 X_t^1 dt + \epsilon dW_t^1,$$
$$dX_t^2 = \lambda_2 X_t^2 dt + \epsilon dW_t^2,$$

where $\lambda_1 > \lambda_2 > 0$ and $X_0^1 = X_0^2 = 0$, and W^1, W^2 are independent standard Wiener processes. We define $\tau = \inf\{t \geq 0 : X_t \in \partial \mathbf{D}\}$, where $\mathbf{D} = (-1,1)^2$ is a square and study the distribution of X_τ, the location of exit from $\mathbf{D}$.

When ϵ is small, it takes a long time to exit, and for large times t, the Duhamel principle gives

$$(2.1) \qquad X_t^k = \epsilon e^{\lambda_k t} \int_0^t e^{-\lambda_k s} dW_s^k \approx \epsilon e^{\lambda_k t} N_k,$$

where $N_k = \int_0^\infty e^{-\lambda_k s} dW_s^k$ is a centered Gaussian random variable with variance $1/(2\lambda_k)$. Denoting $\tau_k = \inf\{t \geq 0 : |X_t^k| = 1\}$, $k = 1,2$, we obtain from (2.1) that

$$(2.2) \qquad \tau_k \approx \frac{1}{\lambda_k} \log \frac{1}{\epsilon} + \frac{1}{\lambda_k} \log \frac{1}{|N_k|}.$$

Therefore, for small ϵ, typically we have $\tau_1 < \tau_2$. Moreover, plugging (2.2) with $k = 1$ into (2.1) with $k = 2$, we obtain

$$X_{\tau_1}^2 = \epsilon^{1-\frac{\lambda_2}{\lambda_1}} |N_1|^{-\frac{\lambda_2}{\lambda_1}} N_2 \to 0, \quad \epsilon \to 0,$$

so the typical random locations of exit X_τ will concentrate near points $q_\pm = (\pm 1, 0)$ where the invariant manifold associated with the leading eigenvalue λ_1 (i.e., the first axis) intersects $\partial \mathbf{D}$.

Let us now prohibit exits through the sides of $\mathbf{D}$ that contain $q_\pm$ and study the unlikely event B of exiting $\mathbf{D}$ through $[-1,1] \times \{-1,1\}$, i.e., we define $B = \{|X_\tau^2| = 1\}$. It turns out that $\mathbb{P}(B) = c\epsilon^{\frac{\lambda_1}{\lambda_2}-1}(1 + o(1))$ and the exit distribution conditioned on B is, somewhat surprisingly, asymptotically uniform on $[-1,1] \times \{-1,1\}$. Let us present a heuristic argument for this.

Introducing events

$$A_r = \left\{ |N_1| < r\epsilon^{\frac{\lambda_1}{\lambda_2}-1} |N_2|^{\frac{\lambda_1}{\lambda_2}} \right\}, \quad r > 0,$$

we obtain from (2.2) that

$$B = \{\tau_2 < \tau_1\} \approx A_1,$$

and, plugging (2.2) with $k = 2$ for t into (2.1) with $k = 1$, we obtain that

$$\{|X_{\tau_2}^1| \leq r\} \approx A_r, \quad r > 0.$$

Next,

$$\mathbb{P}(A_r) = \int_{\Sigma_{r,\epsilon}} g(x_1, x_2) dx_1 dx_2,$$

where $\Sigma_{r,\epsilon} = \{(x_1, x_2) : |x_1| < r\epsilon^{\frac{\lambda_1}{\lambda_2}-1} |x_2|^{\frac{\lambda_1}{\lambda_2}}\}$ and g is the joint Gaussian density of N_1 and N_2. As $\epsilon \to 0$, the domain $\Sigma_{r,\epsilon}$ shrinks to the axis $\{x_1 = 0\}$, so we can approximate $g(x_1, x_2)$ by $g(0, x_2)$ and conclude that

$$\mathbb{P}(A_r) = c\epsilon^{\frac{\lambda_1}{\lambda_2}-1} r(1 + o(1)), \quad \epsilon \to 0,$$

where

$$c = 2 \int_{\mathbb{R}} g(0, x) |x|^{\frac{\lambda_1}{\lambda_2}} dx.$$

Therefore,
$$\mathbb{P}(B) = \mathbb{P}\{\tau_2 < \tau_1\} = c\epsilon^{\frac{\lambda_1}{\lambda_2}-1}(1 + o(1)), \quad \epsilon \to 0,$$
and
$$\frac{\mathbb{P}\{|X_{\tau_2}^1| < r\}}{\mathbb{P}(B)} \to r, \quad \epsilon \to 0,$$
which, due to the symmetry of this example, implies that the limiting distribution is uniform.

Another implication of this calculation is that to realize B one needs typical values of N_2 and atypically small values of N_1. Due to (2.2), this translates into typical values of τ_2 and atypically large values of τ_1. One can say that the main effect of conditioning on B is conditioning X^1 to stay within $[-1, 1]$ for abnormally long times, with only a moderate effect on the evolution of X^2.

Let us now expand this example and consider a third coordinate evolving independently according to
$$dX_t^3 = \lambda_3 X_t^3 dt + \epsilon dW_t^3,$$
with $0 < \lambda_3 < \lambda_2$ and $X_0^3 = 0$. Now, the unlikely event of interest $B = \{|X_\tau^2| = 1\}$ corresponds to the exit through the union of two faces of the cube $\mathbf{D} = (-1, 1)^3$ given by $[-1, 1] \times \{-1, 1\} \times [-1, 1]$. From the analysis above we know that the exit will happen at time τ_2 corresponding to moderate values of N_2, i.e, near $\frac{1}{\lambda_k}\log\frac{1}{\epsilon}$. Plugging the expression for τ_2 from (2.2) into (2.1) for $k = 3$, we obtain that $X_{\tau_2}^3 = \epsilon^{1-\frac{\lambda_3}{\lambda_2}}|N_2|^{-\frac{\lambda_3}{\lambda_2}}N_3$. Hence, under conditioning on B, $X_{\tau_2}^3$ converges to 0. Combining this with our analysis of the two-dimensional situation above, we conclude that the exit distribution converges to the uniform distribution on $[-1, 1] \times \{-1, 1\} \times \{0\}$. This union of two one-dimensional segments should be viewed as the intersection of $\partial\mathbf{D}$ with the two-dimensional invariant manifold associated with λ_1 and λ_2, i.e., the x_1x_2-plane.

The goal of this paper may be described as to give a rigorous treatment of this example and its generalizations to higher dimensional nonlinear situations with space-dependent diffusion matrix and general domains.

3. SETTING AND THE MAIN RESULT

In $\mathbb{R}^d$, we consider an open simply connected set $\mathbf{D}$, a bounded vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ and the flow $(S_t)_{t\in\mathbb{R}}$ associated with b via the ODE

(3.1)
$$\begin{aligned} \tfrac{d}{dt}S_t x &= b(S_t x), \\ S_0 x &= x, \end{aligned}$$

satisfying the following conditions:
  — the origin $0 \in \mathbf{D}$;
  — $b(0) = 0$;
  — for all $x \in \mathbf{D} \setminus \{0\}$, the deterministic exit time

(3.2)
$$t(x) = \inf\{t \geq 0 : \ S_t x \in \partial\mathbf{D}\}$$

  satisfies $t(x) < \infty$. In particular, $x = 0$ is the only critical point of b in $\mathbf{D}$;
  — b is C^5 and $b(x) = \mathbf{a}x + q(x)$ where
    · $q(x) \leq C_q|x|^2$ for some $C_q > 0$,

· $\mathbf{a}$ is a $d \times d$ diagonal matrix with real entries $\lambda_1 > \lambda_2 > ... > \lambda_d > 0$;
— $\partial \mathbf{D}$ is C^1;
— b is transversal to $\partial \mathbf{D}$, i.e., $\langle \mathbf{n}(x), b(x) \rangle > 0$ for all $x \in \partial \mathbf{D}$, where $\mathbf{n}$ denotes the outer normal of $\partial \mathbf{D}$.

A more general situation where $\mathbf{a}$ is only assumed to have eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_d > 0$ can be reduced to this one by a diagonalizing linear transformation. Our results also hold for a broad class of domains with piecewise smooth boundaries but we restrict ourselves to domains with smooth boundaries for simplicity.

By the Hartman–Grobman Theorem (see, e.g., [KH95, Theorem 6.3.1]), there is an open neighborhood $\mathbf{U}$ of 0 and a homeomorphism $f : \mathbf{U} \to f(\mathbf{U})$ conjugating the flow S to the linear flow $\bar{S}$ generated by the vector field $x \mapsto \mathbf{a}x$ and given by $\bar{S}_t x = xe^{\lambda t} = (x^j e^{\lambda_j t})_{j=1}^d$, namely,

$$\frac{d}{dt} f(S_t x) = \mathbf{a} f(S_t x).$$

— in addition, we assume that f is a C^5 diffeomorphism.

Remark 3.1. Due to [Ste57], for this C^5 conjugacy condition to hold in our setting, it suffices to require (i) a smoothness condition: b is C^k for some $k \geq 5 \vee (\lambda_1/\lambda_n)$, and (ii) a no-resonanse condition:

$$\lambda_k \neq m_1 \lambda_1 + \ldots + m_d \lambda_d$$

for all $k = 1, \ldots, d$ and all nonnegative integer coefficients $m_1, \ldots, m_d$ satisfying $m_1 + \ldots + m_d \geq 2$.

The vector field $x \mapsto \mathbf{a}x$ is the pushforward of b under f, and since $\mathbf{a}$ is diagonal, f can be chosen to satisfy $f(0) = 0$ and $Df(0) = I$, the identity matrix.

We are interested in random perturbations of (3.1) given by (1.1), where
— $\epsilon \in (0, 1)$ is the noise amplitude parameter;
— $(W_t, \mathcal{F}_t)$ is a standard n-dimensional Wiener process with $n \geq d$;
— $\sigma = (\sigma_j^i)_{i=1,\ldots,d;\ j=1,\ldots,n}$ is a map from $\mathbb{R}^d$ into the space of $d \times n$ matrices satisfying
  · σ is C^3 (and , by adjustments outside $\mathbf{D}$, we may assume that σ has bounded derivatives in $\mathbb{R}^d$),
  · $\sigma(0) : \mathbb{R}^n \to \mathbb{R}^d$ is surjective.
We will study the solutions of (1.1) with initial data $X_0^\epsilon = \epsilon \xi_\epsilon \in \mathcal{F}_0$, where
— ξ_ϵ converges to some $\xi_0 \in \mathcal{F}_0$ in distribution as $\epsilon \to 0$;
— there are constants $C, c > 0$ independent of ϵ such that

(3.3) $$\mathbb{P}\{|\xi_\epsilon| > x\} \leq C e^{-|x|^c} \text{ for all } x \geq 0, \epsilon \in [0, 1).$$

To simplify notations, we often suppress the dependence on ϵ. In particular, we write X_t instead of X_t^ϵ. We introduce the first time for X_t to exit $\mathbf{D}$ as

(3.4) $$\tau = \tau_\epsilon = \inf\{t > 0 : X_t \notin \mathbf{D}\}.$$

Our main results concern the asymptotic properties of the distribution of X_τ, the location of exit of $X_\tau = X_{\tau_\epsilon}^\epsilon$ from $\mathbf{D}$. To state them, we need to introduce more definitions and notations.

— $\mathcal{B}_L = [-L, L]^d$, $L > 0$;
— $\mathbf{F}_{L\pm}^i = \mathcal{B}_L \cap \{x \in \mathbb{R}^d : x^i = \pm L\}$ is a face of $\mathcal{B}_L$, and $\mathbf{F}_L^i = \mathbf{F}_{L+}^i \cup \mathbf{F}_{L-}^i$;
— for $A \subset B \subset \mathbb{R}^d$, $\partial_B A$ and $\mathrm{int}_B A$ denote the boundary and the interior of A relative to B;
— $\partial = \partial_{\mathbb{R}^d}$, $\partial_L = \partial_{\partial \mathcal{B}_L}$, $\mathrm{int}_L = \mathrm{int}_{\partial \mathcal{B}_L}$;
— $\mathcal{H}^s$ denotes the s-dimensional Hausdorff measure.

For $k = 1, \ldots, d$, we define sets $\Lambda^k = \bigoplus_{i=1}^k \mathbb{R}e_i$, where $(e_i)_{i=1}^d$ is the standard basis for $\mathbb{R}^d$. The sets Λ^k are invariant manifolds for the linear flow $(\bar{S}_t)$ associated with top k exponents $\lambda_1, \ldots, \lambda_k$. Therefore, the sets

$$(3.5) \qquad M^k = \{x \in \mathbb{R}^d : S_t x \in f^{-1}(\Lambda^k) \text{ for some } t \in \mathbb{R}\}, \quad k = 1, \ldots, d,$$

are the k-dimensional invariant manifolds associated with top k exponents for the flow (S_t). Let us define the traces of these manifolds on the boundary by $N^k = M^k \cap \partial\mathbf{D}$ and note that due to our transversality assumptions, N^k is a $(k-1)$-dimensional C^1-manifold. In particular N^1 consists of two points, N^2 is a closed curve in $\partial\mathbf{D}$, and N^d coincides with $\partial\mathbf{D}$.

For any set $A \subset \partial\mathbf{D}$ we define the index of A to be

$$i(A) = \min\{k \in \{1, \ldots d\} : \overline{A} \cap N^k \neq \emptyset\},$$

see Figure 1. This notion is going to be useful because we will show that due to the presence of different exponential growth rates in different directions, the probabilities for the system to exit $\mathbf{D}$ near N^k have different orders of magnitude for different values of k. Thus, the index of A picks the manifold with the dominating contribution. However, this notion becomes truly meaningful and helps computing the asymptotics of exit probabilities only for sets with an additional regularity property which is compatible with the notion of weak convergence of probability measures, holds true for most relatively open subsets of $\partial\mathbf{D}$, and which we proceed to define.

Assuming $d \geq 2$, we say that a set $A \subset \partial\mathbf{D}$ is N-regular if it is Borel and satisfies

$$(3.6) \qquad\qquad \mathcal{H}^{i(A)-1}\{\partial_{\partial\mathbf{D}} A \cap N^{i(A)}\} = 0.$$

In the case of $d = 1$, all subsets of $\partial\mathbf{D}$ are considered to be N-regular.

We still need a few more elements of our construction.

There is a Euclidean ball O centered at 0, satisfying $\overline{f^{-1}(O)} \subset \mathbf{U}$, and such that the vector field $x \mapsto \mathbf{a}x$ is transversal to ∂O. Let us fix O and define

$$(3.7) \qquad\qquad L(O) = \sup\{L > 0 : \mathcal{B}_L \subset O\}.$$

For every $L \in (0, L(O))$, we can define $\psi_L : f^{-1}(\partial\mathcal{B}_L) \to \partial\mathbf{D}$ as the Poincare map along the flow (S_t):

$$(3.8) \qquad\qquad \psi_L(x) = S_{t(x)}x, \quad x \in f^{-1}(\partial\mathcal{B}_L),$$

where $t(\cdot)$ was introduced in (3.2). We can now define

$$(3.9) \qquad\qquad \zeta_L = f \circ \psi_L^{-1} : \partial\mathbf{D} \to \partial\mathcal{B}_L.$$

For $x, y \in \mathbb{R}^d$, we define

$$\tilde{\chi}^i(x^i, y) = \frac{1}{\sqrt{(2\pi)^d \det \mathcal{C}}} |x^i + y^i|^{\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \int_{\mathbb{R}^{d-i}} e^{-\frac{1}{2} x^\intercal \mathcal{C}_0^{-1} x} \Big|_{(x^1, \ldots, x^{i-1}) = -(y^1, \ldots, y^{i-1})} dx^{i+1} \ldots dx^d,$$

$$\chi_+^i(y) = \int_{[-y^i, \infty)} \tilde{\chi}^i(x^i, y) dx^i \quad \text{and} \quad \chi_-^i(y) = \int_{(-\infty, -y^i]} \tilde{\chi}^i(x^i, y) dx^i,$$

where

(3.10)
$$\mathcal{C}^{jk} = \sum_{l=1}^{n} \frac{\sigma_l^j(0) \sigma_l^k(0)}{\lambda_j + \lambda_k}.$$

For $L < L(O)$ and $i = 1, 2, \ldots, d$, we define the following measure on $\partial \mathbf{D}$:

(3.11)
$$\mu_L^i(A) = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \sum_{\bullet \in \{+, -\}} \mathbb{E} \chi_\bullet^i(\xi_0) \cdot \mathcal{H}^{i-1}(\zeta_L(A) \cap \mathbf{F}_{L\bullet}^i \cap \Lambda^i), \quad A \subset \partial \mathbf{D}.$$

In this definition, the set $\mathbf{F}_{L\bullet}^i \cap \Lambda^i$ is the union of two $(i-1)$-dimensional rectangles:

$$\mathbf{F}_{L\bullet}^i \cap \Lambda^i = [-L, L]^{i-1} \times \{-L, L\} \times \{0\}^{d-i},$$

and $\mathcal{H}^{i-1}(\cdot \cap \mathbf{F}_{L\bullet}^i \cap \Lambda^i)$ is simply the $(i-1)$-dimensional Euclidean volume (Lebesgue measure), so the measure μ_L^i is Radon–Nikodym equivalent to the volume measure on of $N^i \cap \zeta^{-1}(\mathbf{F}_L^i)$.

Recalling the definition of ρ_i in (1.3), we can now state our main result.

Theorem 3.2. *If A is an N-regular set with index i, then there is $L_A \in (0, L(O))$ such that for all $L \in (0, L_A)$*

(3.12)
$$\lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{X_\tau \in A\} = \mu_L^i(A).$$

Remark 3.3. The proof of this theorem also implies that the family of numbers (L_A) indexed by N-regular sets A can be chosen to satisfy $L_{A'} \geq L_A$ for $A' \subset A$.

Remark 3.4. Note that the scaling exponent ρ_i and the limiting constant $\mu_L^i(A)$ in (3.12) are defined explicitly. Thus, (3.12) provides a very precise approximation. Although the right-hand side of (3.12) seemingly involves L, in fact, it does not depend on $L \in (0, L_A)$. It is easy to see that $L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}}$ in the definition of $\mu_L^i(A)$ is the correct scaling factor compensating for distortions in directions $1, \ldots, i-1$ introduced by the linear flow that is a part of the definition of $\zeta_L(A)$. We also note that N-regular sets A such that $\mu_L^i(A) > 0$ (so Theorem 3.2 provides the truly leading term in the asymptotics) form a large class that includes, for example, ζ_L-preimages of small open balls with centers in $\mathbf{F}_{L\bullet}^i \cap \Lambda^i$.

According to Theorem 3.2, the decay rate of probability of exit is the same for all N-regular sets of the same index i. This, along with the fact that N-regular sets are specifically defined to be continuity sets for $\mathcal{H}^{i-1}$, allows us to state a corollary on the limiting behavior of conditional exit distributions.

If $\mathbb{P}\{X_\tau \in A\} \neq 0$, let ν_A^ϵ be the exit distribution of X conditioned on exiting from A:

$$\nu_A^\epsilon(\cdot) = \frac{\mathbb{P}\{X_\tau \in \cdot \cap A\}}{\mathbb{P}\{X_\tau \in A\}}.$$

We denote the weak convergence of finite positive Radon measures by "$\rightharpoonup$".

Theorem 3.5. *Let A be an N-regular set of index i and suppose that $\mu^i_L(A) > 0$. Then for all $L < L_A$ the following weak convergence holds as $\epsilon \to 0$:*

$$\nu^\epsilon_A \rightharpoonup \frac{\mu^i_L(\cdot \cap A)}{\mu^i_L(A)}.$$

Remark 3.6. The definition of μ^i_L in (3.11) together with the bi-Lipschitzness of ζ_L and (4.2) implies that $\mu^i_L(\cdot \cap A)$ is equivalent to the $(i-1)$-dimensional Hausdorff measure restricted to $\overline{A} \cap N^{i-1}$.

In the special case where $b(x) \equiv \mathbf{a}x$, $\mathbf{D} = \operatorname{int} \mathcal{B}_L$, $A = \mathbf{F}^i_{L\pm}$ for some $L > 0$ and $i = 1, \ldots, d$, the limiting measure in Theorem 3.5 is the uniform distribution on $\mathbf{F}^i_{L\pm} \cap \Lambda^i$. Thus Theorems 3.2 and 3.5 are natural generalizations of the simple 2- and 3-dimensional equidistribution examples discussed in Section 2.

It is important to stress that Theorem 3.5, where the limiting conditional distribution is equivalent to the volume measure on the manifold N^i, paints a picture drastically different from the typical large deviations picture where the limiting conditional distributions are often point masses concentrated at the minimizers of the large deviation rate function.

The unconditioned exit distribution was also shown to converge to an explicitly computed limit equivalent to the volume on a manifold of smaller dimension in [Eiz84]. In that paper, the eigenvalues of the linearization are not required to be simple but the assumptions on nonlinearity are fairly restrictive. At the core of the results of [Eiz84] and ours, is the fact that the transition probability over a small time interval is approximately Gaussian and this distribution is carried to the boundary almost deterministically by the flow, different directions being stretched with different rates. However, our results are more delicate since we have to zoom into the transition distribution studying its regularity at small scales with Malliavin calculus tools.

The plan of the proof is the following.

We are going to decompose the dynamics into two stages: (i) the evolution in the transformed coordinates until the exit from a small cube $\mathcal{B}_L$ (or, equivalently, from $f^{-1}(\mathcal{B}_L)$ in the original coordinates) and (ii) the evolution between exiting from $f^{-1}(\mathcal{B}_L)$ and exiting from $\mathbf{D}$. In the second stage, the process essentially follows the deterministic flow trajectory (S_t) and the associated Poincare map ψ_L, with error controlled by a FW large deviation estimate, so it is stage (i) that is central to the analysis. During stage (i), the evolution is well approximated by a Gaussian process due to approximate linearity of the drift, so to obtain the desired asymptotics we combine direct computations for this Gaussian process with estimates on the error of the Gaussian approximation based on Malliavin calculus bounds previously obtained in [BC19].

4. Proof of the main result

Theorem 3.2 will follow from two results that we give first. The first result helps to reduce the problem to considering only sets A with $\zeta_L(A)$ being a subset of the union of two faces of $\mathcal{B}_L$ associated with coordinate $i(A)$, and the second one computes the asymptotic probability of exit through such a set.

Proposition 4.1. *Let $A \subset \partial \mathbf{D}$ satisfy $i(A) = i$. The number L'_A defined by*

$$L'_A = \sup\{L \in (0, L(O)): \ \overline{\zeta_L(A)} \cap \mathbf{F}^j_L = \emptyset, \forall j < i; \mathcal{B}_L \subset O\}$$

is positive, (i.e., the set under the supremum is nonempty) and for all $L < L'_A$ we have

(1) $\overline{\zeta_L(A)} \cap \mathbf{F}^j_L = \emptyset$ *for all $j < i$;*
(2) *if $i < d$ and A is N-regular, then there are N-regular $A_0, A_1 \subset \partial \mathbf{D}$ such that*
  (a) $A_0 \subset A \subset A_0 \cup A_1$;
  (b) $i(A_0) = i$; $\quad \zeta_L(A) \cap \Lambda^i = \zeta_L(A_0) \cap \Lambda^i$; $\quad \zeta_L(A_0) \subset \mathrm{int}_{\partial \mathcal{B}_L} \mathbf{F}^i_L$;
  (c) $i(A_1) = i + 1$.

Proposition 4.2. *There is $L_0 > 0$ such that the following holds. Let $A \subset \partial \mathbf{D}$ be an arbitrary N-regular set with $i(A) = i$. If $L < L_0$ satisfies $\overline{\zeta_L(A)} \subset \mathrm{int}_{\partial \mathcal{B}_L} \mathbf{F}^i_L$, then*

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{X_\tau \in A\} = \mu^i_L(A).$$

Proposition 4.1 and Proposition 4.2 will be proved in Sections 6.1 and 6.2, respectively.

Proof of Theorem 3.2. Let $L_A = L'_A \wedge L_0$, where L_0 and L_A are defined in the propositions above. We immediately see that if $A' \subset A$, then $L_{A'} \geq L_A$, so Remark 3.3 is automatically justified.

The idea of the proof is to use Proposition 4.1 in order to approximate A by a union of regular sets of various indices such that Proposition 4.2 can be applied to each of them. More formally, we will use induction on $i(A)$, starting with the case $i(A) = d$.

Note that, by (1) of Proposition 4.1, $i(A) = d$ implies $\zeta_L(A) \subset \mathrm{int}_{\partial \mathcal{B}_L}(\mathbf{F}^d_L)$ for all $L < L_A$. Therefore, we can apply Proposition 4.2 to obtain

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_d} \mathbb{P}\{X_\tau \in A\} = \mu^d_L(A), \quad \text{for all } L < L_A,$$

which completes the proof of the induction basis.

For the induction step, let us assume that the desired result holds for all A with $i(A) = k$ for $i + 1 \leq k \leq d$. Let us show it is also true for A with $i(A) = i$.

Let us fix $L \in (0, L_A)$ arbitrarily. Since $i < d$ now, we can define A_0 and A_1 according to part (2) of Proposition 4.1. Since $A_0 \subset A$, Remark 3.3 implies $L < L_A \leq L_{A_0}$. Then, using part (2b) of Proposition 4.1, Proposition 4.2, and the definition of μ^i_L in (3.11), we obtain

$$(4.1) \qquad \lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{X_\tau \in A_0\} = \mu^i_L(A_0) = \mu^i_L(A).$$

By (2c) of Proposition 4.1, $i(A_1) = i + 1$, so by the induction hypothesis, for each $L' \leq L_{A_1}$,

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_{i+1}} \mathbb{P}\{X_\tau \in A_1\} = \mu^{i+1}_{L'}(A_1),$$

which implies that $\mathbb{P}\{X_\tau \in A_1\} = \mathcal{O}(\epsilon^{\rho_{i+1}}) = o(\epsilon^{\rho_i})$. Due to (2a),

$$|\mathbb{P}\{X_\tau \in A\} - \mathbb{P}\{X_\tau \in A_0\}| \leq \mathbb{P}\{X_\tau \in A_1\} = o(\epsilon^{\rho_i}).$$

Combining this with (4.1), we complete the induction step and the entire proof. $\qquad \square$

To prove Theorem 3.5, we need the following basic result.

Lemma 4.3. *Let $A \subset \partial \mathbf{D}$ be arbitrary. Then the following holds:*

(1) $i(A) = \min\left\{k \in \{1,\dots,d\} : \overline{\zeta_L(A)} \cap \Lambda^k \neq \emptyset\right\}$ for each $L < L(O)$;

(2) if A is Borel, then the following statements are equivalent:

    (a) A is N-regular,

    (b) $\mathcal{H}^{i(A)-1}\{\partial_L(\zeta_L(A)) \cap \Lambda^{i(A)}\} = 0$ for some $L < L(O)$,

    (c) $\mathcal{H}^{i(A)-1}\{\partial_L(\zeta_L(A)) \cap \Lambda^{i(A)}\} = 0$ for all $L < L(O)$.

Proof. First, ζ_L is a bi-Lipschitz homeomorphism, since it is a composition of a diffeomorphism f and the Poincaré map ψ_L^{-1} constructed in (3.8) from smooth flows transversal to locally smooth sections. Secondly, due to (3.5), the definition $N^k = M^k \cap \partial\mathbf{D}$, and the invariance of Λ^k under the linear flow $\bar{S}$, one can see that

$$(4.2) \qquad \zeta_L(\overline{A} \cap N^k) = \overline{\zeta_L(A)} \cap \Lambda^k \quad \text{and} \quad \zeta_L(\partial_{\partial\mathbf{D}} A \cap N^k) = \partial_L(\zeta_L(A)) \cap \Lambda^k,$$

which implies both parts (1) and (2) straightforwardly. $\qquad\square$

Proof of Theorem 3.5. We need to prove that

$$(4.3) \qquad \nu_A^\epsilon(B) \to \frac{\mu_L^i(B \cap A)}{\mu_L^i(A)}, \quad \epsilon \to 0,$$

for every continuity set B of the measure $\mu_L^i(\,\cdot\,\cap A)$ or, equivalently, by the definition (3.11), of $\mathcal{H}^{i-1}\big(\zeta_L(\,\cdot\,\cap A) \cap \mathbf{F}_{L\pm}^i \cap \Lambda^i\big)$ which is equal to $\mathcal{H}^{i-1}\big(\zeta_L(\,\cdot\,\cap A) \cap \Lambda^i\big)$ due to $L < L_A$. Using the inclusion $\partial_{\mathbf{D}}(B \cap A) \subset (\partial_{\mathbf{D}} B \cap \overline{A}) \cup (\overline{B} \cap \partial_{\mathbf{D}} A)$, the N-regularity of A (see (3.6)) and (2) of Lemma 4.3, we conclude that the continuity property of B implies that of $B \cap A$. Combining this with the fact that $\nu_A^\epsilon(B \cap A^c) = 0$ for all ϵ, we obtain that it is sufficient to check (4.3) for Borel subsets B of A with continuity property. For such a set B, either $i(B) = i(A)$, or $i(B) > i(A)$. In the first case, writing

$$\partial_{\mathbf{D}} B = \partial_{\mathbf{D}}(B \cap A) \subset (\partial_{\mathbf{D}} B \cap \overline{A}) \cup (\overline{B} \cap \partial_{\mathbf{D}} A),$$

using the continuity of B, part (2) of Lemma 4.3, and the N-regularity of A, we conclude that B is also N-regular, so (4.3) follows from Theorem 3.2. In the second case, part (1) of Lemma 4.3 implies $\overline{\zeta_L(B)} \cap \Lambda^i = \emptyset$. Therefore, $r = \operatorname{dist}(\zeta_L(B), \Lambda^i) > 0$, where

$$(4.4) \qquad \operatorname{dist}(C, D) = \inf\{|x - y| : \; x \in C, \; y \in D\} \wedge 1, \quad C, D \subset \mathbb{R}^d.$$

Since $\mathcal{H}^{i-1}\big(\zeta_L(B \cap A) \cap \mathbf{F}_{L\pm}^i \cap \Lambda^i\big) = 0$, it suffices to prove $\nu_A^\epsilon(B) \to 0$ to ensure (4.3). Let us define $B_L^i(r) = \{x \in \mathcal{B}_L : \operatorname{dist}(\{x\}, \Lambda^i) \geq r\}$. Since $\zeta_L(B) \subset B_L^i(r)$, and $\zeta_L^{-1}(B_L^i(r))$ is an N-regular set of index $i+1$ due to parts (1) and (2) of Lemma 4.3, we can apply Theorem 3.2 to $\zeta_L^{-1}(B_L^i(r))$ and conclude that $\mathbb{P}\{X_\tau \in B\} = o\big(\epsilon^{\rho_{i(A)}}\big) = o\big(\mathbb{P}\{X_\tau \in A\}\big)$, so (4.3) holds in this case as well. The proof is completed. $\qquad\square$

5. Exit from a box

Recall the definitions of $\mathcal{B}_L = [-L, L]^d$ and $\mathbf{F}_{L\pm}^i$ in Section 3. Let

$$\mathbf{F}_{L\pm,\delta}^i = \{x \in \mathbf{F}_{L\pm}^i : |x^j| \leq L - \delta \text{ for } j \neq i\},$$

$$\mathbf{F}_{L,\delta}^i = \mathbf{F}_{L+,\delta}^i \cup \mathbf{F}_{L-,\delta}^i.$$

Set

$$(5.1) \qquad \tau_L = \inf\{t > 0 : X_t \notin f^{-1}(\mathcal{B}_L)\},$$

where f is the linearizing conjugacy. The main result of this section gives the exit probability asymptotics for sets whose images under f are rectangles:

Proposition 5.1. *There exists $L_0 > 0$ such that for all positive $L \leq L_0$, we have*

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{X_{\tau_L} \in f^{-1}(A)\} = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \sum_{\bullet \in \{+,-\}} \mathbb{E}\chi_\bullet^i(\xi_0)\mathcal{H}^{i-1}(A \cap \mathbf{F}_{L\bullet}^i \cap \Lambda^i)$$

for all $i \in \{1, 2, \ldots, d\}$, and for all sets A with the following properties:

  — *A is a product of intervals which can be open, closed, or half-open.*
  — *$\overline{A} = [a^1, b^1] \times \ldots \times [a^{i-1}, b^{i-1}] \times \{\pm L\} \times [a^{i+1}, b^{i+1}] \times \ldots \times [a^d, b^d]$, where $[a^j, b^j] \subset (-L, L)$ for all $j \neq i$ and $a^j, b^j \neq 0$ for $j > i$.*

5.1. Derivation of Proposition 5.1 from auxiliary results.

From now on, we use standard summation convention over matching upper and lower indices. Let $Y_t = f(X_t)$, where f is the linearizing conjugacy. Using Itô's formula, we obtain

$$dY_t^i = \mathbf{a}_j^i Y_t^j dt + \epsilon \partial_k f^i(f^{-1}(Y_t))\sigma_j^k(f^{-1}(Y_t))dW_t^j$$

(5.2)
$$+ \frac{\epsilon^2}{2}\sum_{j,k=1}^d \partial_{j,k}^2 f^i(f^{-1}(Y_t))\langle \sigma^j(f^{-1}(Y_t)), \sigma^k(f^{-1}(Y_t))\rangle dt$$

$$= \lambda^i Y_t^i dt + \epsilon F_j^i(Y_t)dW_t^j + \epsilon^2 G^i(Y_t)dt,$$

where

  — $\lambda^i = \lambda_i$ to avoid summation in i;
  — F and G are C^3 (since f is C^5 and σ is C^3);
  — since $f(0) = 0$ and $Df(0) = I$, we have $F(0) = \sigma(0)$.

Since $F(0) = \sigma(0)$ is $d \times n$ with full rank and F is continuous, we can find $L_0 > 0$ small so that there is $c_0 > 0$ such that $\min_{|u|=1, u\in\mathbb{R}^d} |u^\intercal F(x)|^2 \geq c_0$ for all $x \in [-L_0, L_0]^d$. We shrink L_0 further, if necessary, to ensure $L_0 \leq L(O)$ as in (3.7). Since we will only care about exiting from a subset of $[-L_0, L_0]^d$, we modify F, G outside $[-L_0, L_0]^d$ so that

(5.3)
$$\min_{|u|=1, u\in\mathbb{R}^d} |u^\intercal F(x)|^2 \geq c_0, \text{ for all } x \in \mathbb{R}^d;$$
$$F, G \text{ and their derivatives are bounded.}$$

With this L_0 chosen, we will consider the following for the rest of this section, applying Duhamel's principle to (5.2) and setting $Y_0 = \epsilon y$,

(5.4)
$$Y_t^j = \epsilon e^{\lambda_j t} y^j + \epsilon e^{\lambda_j t}\left(\int_0^t e^{-\lambda_j s} F_l^j(Y_s)dW_s^l + \epsilon \int_0^t e^{-\lambda_j s} G^j(Y_s)ds\right)$$
$$= \epsilon e^{\lambda_j t}(y^j + M_t^j + \epsilon V_t^j) = \epsilon e^{\lambda_j t}(y^j + U_t^j),$$

where F, G are modified to ensure (5.3). We emphasize that M_t, V_t and U_t all depend on y and ϵ. We define $\mathbb{P}^{\epsilon y} = \mathbb{P}\{\cdot | Y_0 = \epsilon y\}$.

Let C_f be the Lipschitz constant of f and $c_f = C_f^{-1}$. Since $f(0) = 0$, we have $|\epsilon^{-1}f(\epsilon x)| \leq C_f|x|$ for all x and ϵ. In view of (3.3), we choose $\kappa > 0$ large so that for

(5.5)
$$K(\epsilon) = (\log \epsilon^{-1})^\kappa$$

we have

(5.6) $\qquad \mathbb{P}\{|\xi_\epsilon| > c_f K(\epsilon)\} \leq \epsilon^{\rho_d + \delta}$, for all $\epsilon \in (0, 1]$, for some $\delta > 0$.

By our definition of c_f, we have

$$\text{if } |x| \leq c_f K(\epsilon), \quad \text{then } |\epsilon^{-1} f(\epsilon x)| \leq K(\epsilon).$$

Remark 5.2. Later, when needed, κ in (5.5) will be adjusted to be even larger. This will not affect our results.

According to (5.1), $\tau_L = \inf\{t > 0 : Y_t \notin \mathcal{B}_L\}$. To prove Proposition 5.1, we first obtain asymptotics for Y_t exiting rectangular sets uniformly in $Y_0 = \epsilon y$ with $|y| \leq K(\epsilon)$. Recall that ρ_i is given in (1.3).

Proposition 5.3. *Consider Y_t defined by (5.4). If $L < L_0$, $i = 1, \ldots, d$, and A is a rectangle described in Proposition 5.1, then*

(5.7) $$\lim_{\epsilon \to 0} \sup_{|y| \leq K(\epsilon)} \left| \epsilon^{-\rho_i} \mathbb{P}^{\epsilon y}\{Y_{\tau_L} \in A\} - \chi_\pm^i\left(\epsilon^{-1} f^{-1}(\epsilon y)\right) c_A \right| = 0,$$

where

(5.8) $$c_A = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \prod_{j<i} (b^j - a^j) \prod_{j>i} \mathbb{1}_{\{0 \in (a^j, b^j)\}}.$$

Here and throughout the paper, the product over an empty set is understood to equal 1.

Derivation of Proposition 5.1 from Proposition 5.3. Consider (1.1) with $X_0 = \epsilon \xi_\epsilon$ described in (3.3) and observe that, by (5.6) and the above proposition,

(5.9) $$\epsilon^{-\rho_i} \mathbb{P}\{X_{\tau_L} \in f^{-1}(A)\} = \mathbb{E}\epsilon^{-\rho_i} \mathbb{P}^{f(\epsilon \xi_\epsilon)}\{Y_{\tau_L} \in A\} \mathbb{1}_{\{|\xi_\epsilon| \leq c_f K(\epsilon)\}} + o(1)$$

$$= c_A \mathbb{E}\chi_\pm^i(\xi_\epsilon) \mathbb{1}_{\{|\xi_\epsilon| \leq c_f K(\epsilon)\}} + o(1).$$

There is $C > 0$ such that

(5.10) $$|\chi_\pm^i(y)| \leq C(1 + |y|^{\sum_{j<i} \frac{\lambda_j}{\lambda_i}}), \quad x \in \mathbb{R}^d.$$

Due to the fast decay of the tail of ξ_ϵ imposed by (3.3), all positive moments of $|\chi_\pm^i(\xi_\epsilon)|$ are bounded uniformly in ϵ. Therefore, due to (5.6) and Hölder's inequality, we have $\mathbb{E}\chi_\pm^i(\xi_\epsilon)\mathbb{1}_{\{|\xi_\epsilon| > c_f K(\epsilon)\}} = o(1)$, which implies

(5.11) $$c_A \mathbb{E}\chi_\pm^i(\xi_\epsilon)\mathbb{1}_{\{|\xi_\epsilon| \leq c_f K(\epsilon)\}} = c_A \mathbb{E}\chi_\pm^i(\xi_\epsilon) + o(1).$$

This, along with the uniform tail bound on ξ_ϵ in (3.3), the polynomial bound on $\chi_\pm$ in (5.10), and continuity of $\chi_\pm$, implies

(5.12) $$\lim_{\epsilon \to 0} c_A \mathbb{E}\chi_\pm^i(\xi_\epsilon) = c_A \mathbb{E}\chi_\pm^i(\xi_0).$$

Combining (5.9), (5.11), and (5.12), we obtain

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{X_{\tau_L} \in f^{-1}(A)\} = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \mathbb{E}\chi_\pm^i(\xi_0) \mathcal{H}^{i-1}\{A \cap \Lambda^i\},$$

completing the proof. $\qquad \square$

5.2. **Proof of Proposition 5.3.** Let us fix any $L < L_0$ and $i \in \{1, 2, \ldots, d\}$. First, we remark that it suffices to consider A satisfying

$$(5.13) \qquad\qquad 0 \in (a^i,\ b^i), \quad \forall j > i.$$

In fact, if A does not satisfy (5.13), then we can find two rectangles A' and A'' such that

  (i) $\pi^{\leq i}(A) = \pi^{\leq i}(A') = \pi^{\leq i}(A'')$ where $\pi^{\leq i}$ is the projection onto the first i coordinates;
 (ii) $A' \subset A''$ and $A \subset A'' \setminus A'$;
(iii) A' and A'' satisfy (5.13).

By (5.8), we have $c_A = 0$ since A does not satisfy (5.13), and $c_{A'} = c_{A''}$ due to (i). These together with (ii) imply (we use $\eta_\pm = \chi^i_\pm(\epsilon^{-1}f^{-1}(\epsilon y))$):

$$\left|\epsilon^{-\rho_i}\mathbb{P}^{\epsilon y}\{Y_{\tau_L} \in A\} - \eta_\pm c_A\right| = \left|\epsilon^{-\rho_i}\mathbb{P}^{\epsilon y}\{Y_{\tau_L} \in A\}\right|$$
$$\leq \left|\epsilon^{-\rho_i}\mathbb{P}^{\epsilon y}\{Y_{\tau_L} \in A'\} - \eta_\pm c_{A'}\right| + \left|\epsilon^{-\rho_i}\mathbb{P}^{\epsilon y}\{Y_{\tau_L} \in A''\} - \eta_\pm c_{A''}\right|.$$

Finally, (iii) allows us to apply (5.7) to A' and A'', and thus (5.7) holds for A.

To avoid heavy notation, we also assume that A is closed. It can be readily checked that all our arguments are still valid if A is not closed.

Recall τ_L given in (5.1). Since L is fixed, for brevity, we write $\tau = \tau_L$ for the rest of the section. Here, we only study the case where $A \subset \mathbf{F}^i_{L+}$, which corresponds to $Y^i_\tau = L$. The case where $A \subset \mathbf{F}^i_{L-}$ (corresponding to $Y^i_\tau = -L$) can be considered in the same way.

We will need the following two statements.

Lemma 5.4. *Assume $0 \in (a^j, b^j)$ for all $j > i$. Let*

$$(5.14) \qquad\qquad T_0 = T_0(\epsilon) = \frac{1}{\lambda_i} \log \frac{L}{\epsilon(\log \epsilon^{-1})^{\kappa+1}}.$$

There are γ_j, $j = 1, \ldots, d$, satisfying

$$(5.15) \qquad\qquad 0 \vee \left(\frac{\lambda_j}{\lambda_i} - 1\right) < \gamma_j < \frac{\lambda_j}{\lambda_i} \quad j = 1, 2, \ldots, d,$$

such that

$$\mathbb{P}\{y + U_{T_0} \in B_-\} - o(\epsilon^{\rho_i}) \leq \mathbb{P}\{Y^j_\tau \in [a^j, b^j], \forall j \neq i; Y^i_\tau = L\} \leq \mathbb{P}\{y + U_{T_0} \in B_+\} + o(\epsilon^{\rho_i})$$

holds uniformly in $|y| \leq K(\epsilon)$, where

$$(5.16) \qquad \begin{aligned} B_\pm &= \cup_{x^i \in I_\pm} \left(B^{(x^i)}_{\pm,<i} \times \{x^i\} \times B^{(x^i)}_{\pm,>i}\right) \\ &= \cup_{x^i \in I_\pm} \left((J^{(x^i)}_{\pm,1} \times \ldots \times J^{(x^i)}_{\pm,i-1}) \times \{x^i\} \times (J^{(x^i)}_{\pm,i+1} \times \ldots \times J^{(x^i)}_{\pm,d})\right) \end{aligned}$$

with $I_\pm = \left(\mp \epsilon^{\gamma_i}, (\log \epsilon^{-1})^{\kappa+1} \pm \epsilon^{\gamma_i}\right]$, and for $j \neq i$

$$J^{(x^i)}_{\pm,j} = \left[a^j L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i}-1}(|x^i| \pm \epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}} \mp \epsilon^{\gamma_j}, b^j L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i}-1}(|x^i| \pm \epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}} \pm \epsilon^{\gamma_j}\right].$$

Note that due to (5.15), for small $\epsilon > 0$, the terms ϵ^{γ_j} are small compared to the leading order terms in the definition of $J^{(x^i)}_{\pm,j}$.

Lemma 5.5. *Let T_0 be given in (5.14) and $\mathcal{Z}$ be a centered Gaussian vector with covariance matrix $\mathcal{C}$ given in (3.10). Then for each $\upsilon \in (0,1)$, there is $\delta > 0$ such that*

$$\sup_{|y| \le \epsilon^{\upsilon-1}} \left| \mathbb{P}\{y + U_{T_0} \in B_\pm\} - \mathbb{P}\{y + \mathcal{Z} \in B_\pm\} \right| = o\left(\epsilon^{\rho_i+\delta}\right).$$

Let us define

(5.17) $$h_\epsilon(y) = \epsilon^{-\rho_i} \mathbb{P}\{y + \mathcal{Z} \in B_\pm\},$$

(5.18) $$h_0(y) = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \chi_\pm^i(y) \prod_{j<i}(b^j - a^j).$$

Note the dependence on $\pm$, which is suppressed to avoid heavy notation. Proposition 5.3 follows from Lemmas 5.4, 5.5 and the following estimate:

(5.19) $$\sup_{|y| \le K(\epsilon)} \left| h_\epsilon(y) - h_0(\epsilon^{-1} f^{-1}(\epsilon y)) \right| = o(1).$$

Our plan is to derive (5.19) in the remainder of this subsection and then prove Lemmas 5.4 and 5.5 in Subsection 5.3.

5.2.1. *Proof of* (5.19). We split (5.19) into estimating $|h_\epsilon(y) - h_0(y)|$ and $|h_0(y) - h_0(\epsilon^{-1} f^{-1}(\epsilon y))|$ separately. The techniques involved are elementary but the proof is tedious. We proceed in steps.

Step 1. We express h_ϵ and h_0 explicitly in the form of Gaussian integrals over some sets. For each $x \in \mathbb{R}^d$, let

$$x^{<i} = (x^1, ..., x^{i-1}), \quad x^{\ge i} = (x^i, ..., x^d), \quad x^{>i} = (x^{i+1}, ..., x^d), \quad \hat{x} = (x^{<i}, x^{>i});$$

(5.20) $$\hat{B}_\pm^{(x^i)} = B_{\pm,<i}^{(x^i)} \times B_{\pm,>i}^{(x^i)}$$

where $B_{\pm,<i}^{(x^i)}$ and $B_{\pm,>i}^{(x^i)}$ are given in (5.16). When y is fixed as in $Y_0 = \epsilon y$, we write $\tilde{x} = (-y^{<i}, x^{\ge i}) \in \mathbb{R}^d$ for each $x \in \mathbb{R}^d$.

Now, let us introduce

(5.21) $$g_\epsilon(x^i, y) = \frac{\epsilon^{-\rho_i}}{\sqrt{(2\pi)^d \det \mathcal{C}}} \int_{\hat{B}_\pm^{(x^i+y^i)} - \hat{y}} e^{-\frac{1}{2} x^\intercal \mathcal{C}^{-1} x} d\hat{x},$$

(5.22) $$g_0(x^i, y) = \frac{\prod_{j<i}(b^j - a^j) L^{-\frac{\lambda_j}{\lambda_i}} |x^i + y^i|^{\frac{\lambda_j}{\lambda_i}}}{\sqrt{(2\pi)^d \det \mathcal{C}}} \int_{\mathbb{R}^{d-i}} e^{-\frac{1}{2} \tilde{x}^\intercal \mathcal{C}_0^{-1} \tilde{x}} dx^{>i}.$$

Recall the definition of $I_\pm$ below (5.16). Additionally, we set

(5.23) $$\mathcal{I}_+(y^i) = [-y^i, \infty) \quad \text{and} \quad \mathcal{I}_-(y^i) = (-\infty, -y^i].$$

Using the definitions (5.17)–(5.18), we can see

$$h_\epsilon(y) = \int_{I_\pm - y^i} g_\epsilon(x^i, y) dx^i,$$

$$h_0(y) = \int_{\mathcal{I}_\pm(y^i)} g_0(x^i, y) dx^i.$$

Step 2. We record some useful estimates, which will be proved later. For simplicity of notation, we write

$$z_\epsilon(y) = \epsilon^{-1} f^{-1}(\epsilon y). \tag{5.24}$$

For convenience, we set $z_0(y) = y$. The following holds for all $\epsilon \in [0,1]$ and $x^i \in \mathbb{R}$,

$$\sup_{|y| \leq K(\epsilon)} |g_\epsilon(x^i, y)| \leq C(1 + K(\epsilon))^p e^{-c|x^i|^2}, \tag{5.25}$$

$$\sup_{|y| \leq K(\epsilon)} |g_\epsilon(x^i, y) - g_0(x^i, y)| \leq C\epsilon^q e^{-c|x^i|^2}, \tag{5.26}$$

$$\sup_{|y| \leq K(\epsilon)} |y - z_\epsilon(y)| \leq C\epsilon^q, \tag{5.27}$$

$$\sup_{|y| \leq K(\epsilon)} |g_0(x^i, y) - g_0(x^i, z_\epsilon(y))| \leq C\epsilon^q e^{-c|x^i|^2}, \tag{5.28}$$

for some $C, c, p, q > 0$.

Step 3. We estimate $|h_\epsilon(y) - h_0(y)|$ for $|y| \leq K(\epsilon)$. We shall only treat the case where $\pm$ is $+$ and $\mp$ is $-$. The other case is similar. We start by writing

$$|h_\epsilon(y) - h_0(y)| \leq \int_{I_+ - y^i} |g_\epsilon(x^i, y) - g_0(x^i, y)| dx^i + \int_{-\epsilon^{\gamma_i} - y^i}^{-y^i} |g_\epsilon(x^i, y)| dx^i$$
$$+ \int_{(\log \epsilon^{-1})^{\kappa+1} + \epsilon^{\gamma_i} - y^i}^{\infty} |g_0(x^i, y)| dx^i.$$

Using (5.25) and (5.26), we have, for some $q' > 0$,

$$\sup_{|y| \leq K(\epsilon)} |h_\epsilon(y) - h_0(y)| \leq \int_{\mathbb{R}} C\epsilon^q e^{-c|x^i|^2} dx^i \quad + C\epsilon^{\gamma_i}(1 + K(\epsilon))^p$$
$$+ \int_{(\log \epsilon^{-1})^{\kappa+1} + \epsilon^{\gamma_i} - K(\epsilon)}^{\infty} C(1 + K(\epsilon))^p e^{-c|x^i|^2} dx^i = o(\epsilon^{q'}). \tag{5.29}$$

Step 4. We estimate $|h_0(y) - h_0(z_\epsilon(y))|$ for $|y| \leq K(\epsilon)$. Recall the definition of $\mathcal{I}_\pm(y^i)$ in (5.23) and note that, due to (5.27),

$$\left| \mathcal{I}_\pm(y^i) \triangle \mathcal{I}_\pm(z_\epsilon^i(y)) \right| \leq |y^i - z_\epsilon^i(y)| \leq C\epsilon^q. \tag{5.30}$$

Here $\triangle$ denotes the symmetric difference of sets. By the formula for $h_0(y)$ in Step 1, we first write

$$|h_0(y) - h_0(z_\epsilon(y))|$$
$$\leq \int_{\mathcal{I}_\pm(z_\epsilon^i(y))} |g_0(x^i, y) - g_0(x^i, z_\epsilon(y))| dx^i + \int_{\mathcal{I}_\pm(y^i) \triangle \mathcal{I}_\pm(z_\epsilon^i(y))} |g_0(x^i, y)| + |g_0(x^i, z_\epsilon(y))| dx^i.$$

We can bound $|g_0(x^i, z_\epsilon(y))|$ by using (5.25) and (5.28). Apply this, (5.28) and (5.30) to see

$$\sup_{|y| \leq K(\epsilon)} |h_0(y) - h_0(z_\epsilon(y))| \leq \int_{\mathbb{R}} C\epsilon^q e^{-c|x^i|^2} dx^i + C\epsilon^q (1 + K(\epsilon))^p = o(\epsilon^{q''}) \tag{5.31}$$

for some $q'' > 0$.

In conclusion, (5.19) follows from (5.29) and (5.31).

It remains to prove estimates listed in Step 2. We prove them in the following order: (5.27), (5.25), (5.28), (5.26).

Proof of (5.27). Recall the definition of $z_\epsilon(y)$ in (5.24) and that the local diffeomorphism f satisfies $f(0) = 0$ and $Df(0) = I$. Since $|y| \le K(\epsilon)$, if ϵ is small, then $|\epsilon y|$ is uniformly close to 0. By expanding f^{-1} at 0, we can see that there are $C, q > 0$ such that

$$|y - z_\epsilon(y)| = |y - \epsilon^{-1}(f^{-1}(\epsilon y))| \le C\epsilon |y|^2 \le C\epsilon^q, \quad \forall |y| \le K(\epsilon).$$

This gives (5.27). □

Proof of (5.25). Since $\sigma(0)$ has full rank, by definition of $\mathcal{C}$ in (3.10), there is $c > 0$ such that

$$(5.32) \qquad\qquad e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x} \le e^{-c|x|^2}, \qquad \forall x \in \mathbb{R}^d.$$

Hence, there is $C > 0$ such that

$$|g_0(x^i, y)| \le C|x^i + y^i|^{\sum_{j<i} \frac{\lambda_j}{\lambda_i}} e^{-c|x^i|^2}.$$

Absorb polynomials of x^i into the exponential to see, for some $C, c, p > 0$,

$$\sup_{|y| \le K(\epsilon)} |g_0(x^i, y)| \le C(1 + K(\epsilon))^p e^{-c|x^i|^2}.$$

From this and (5.26), we obtain (5.25). □

Proof of (5.28). We simplify the expression (5.22) into

$$g_0(x^i, y) = C|x^i + y^i|^{p_0} \int_{\mathbb{R}^{d-i}} e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x}\Big|_{x^{<i}=-y^{<i}} dx^{>i},$$

for some $C, p_0 > 0$. Then, we have

$$|g_0(x^i, y) - g_0(x^i, z_\epsilon(y))|$$

$$(5.33) \qquad \le C\Big(|x^i + y^i|^{p_0} - |x^i + z_\epsilon^i(y)|^{p_0}\Big) \int_{\mathbb{R}^{d-i}} e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x}\Big|_{x^{<i}=-z_\epsilon^{<i}(y)} dx^{>i}$$

$$+ C|x^i + y^i|^{p_0} \int_{\mathbb{R}^{d-i}} \left| e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x}\Big|_{x^{<i}=-y^{<i}} - e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x}\Big|_{x^{<i}=-z_\epsilon^{<i}(y)} \right| dx^{>i}.$$

Let us estimate the terms on the right of (5.33). Using (5.27), we have, for some $C, p, q > 0$,

$$(5.34) \quad \left| |x^i + y^i|^{p_0} - |x^i + z_\epsilon^i(y)|^{p_0} \right| \le C(|x^i|^p + |y|^p)\epsilon^{q'} \le C\epsilon^q(|x^i|^p + 1), \quad \forall |y| \le K(\epsilon).$$

By (5.32), there are $C, c > 0$ such that

$$(5.35) \qquad\qquad \int_{\mathbb{R}^{d-i}} e^{-\frac{1}{2}x^\mathsf{T}\mathcal{C}^{-1}x}\Big|_{x^{<i}=-y^{<i}} dx^{>i} \le Ce^{-c|x^i|^2}.$$

To estimate the integrand of the last integral in (5.33), we need the following observation. Since $\mathcal{C}$ is symmetric and positive definite, there are $C, C' > 0$ such that, for all $w, z \in \mathbb{R}^d$,

$$\left| e^{-\frac{1}{2} w^\intercal \mathcal{C}^{-1} w} - e^{-\frac{1}{2} z^\intercal \mathcal{C}^{-1} z} \right| \leq C(e^{-c|w|^2} \vee e^{-c|z|^2})|w+z||w-z|$$

$$\leq C e^{-c|w|^2}(|2w| + |w-z|)|w-z| \mathbb{1}_{\{|w| \leq |z|\}}$$

$$+ C e^{-c|z|^2}(|2z| + |w-z|)|w-z| \mathbb{1}_{\{|w| > |z|\}}$$

$$(5.36) \qquad \leq C'(e^{-c'|w|^2} + e^{-c'|z|^2})(|w-z| + |w-z|^2).$$

Using this estimate and (5.27), we can obtain

$$\left| e^{-\frac{1}{2} x^\intercal \mathcal{C}^{-1} x} \Big|_{x^{<i} = -y^{<i}} - e^{-\frac{1}{2} x^\intercal \mathcal{C}^{-1} x} \Big|_{x^{<i} = -z_\epsilon^{<i}(y)} \right|$$

$$\leq e^{-c|x^{\geq i}|^2} \left(|y^{<i} - z_\epsilon^{<i}(y)| + |y^{<i} - z_\epsilon^{<i}(y)|^2 \right)$$

$$(5.37) \qquad \leq C \epsilon^q e^{-c|x^{\geq i}|^2}, \qquad \forall |y| \leq K(\epsilon),$$

Insert (5.34), (5.35), and (5.37) to the right hand side of (5.33) to see

$$|g_0(x^i, y) - g_0(x^i, z_\epsilon(y))|$$

$$\leq C \epsilon^q (|x^i|^p + 1) e^{-c|x^i|^2} + C(|x^i|^{p_0} + |K(\epsilon)|^{p_0}) \epsilon^q e^{-c|x^i|^2}$$

$$\leq C \epsilon^{q'} e^{-c'|x^i|^2}, \qquad \forall |y| \leq K(\epsilon),$$

for some $c', q' > 0$. This completes the proof.

$\square$

Proof of (5.26). Again, the techniques involved are elementary while the proof is tedious. Recall g_ϵ and g_0 in (5.21)–(5.22), and the notation in (5.20). To estimate the difference between g_ϵ and g_0, we introduce

$$\mathtt{I} = \frac{\epsilon^{-\rho_i}}{\sqrt{(2\pi)^d \det \mathcal{C}}} \int_{\hat{B}_\pm^{(x^i + y^i)} - \hat{y}} e^{-\frac{1}{2} \tilde{x}^\intercal \mathcal{C}^{-1} \tilde{x}} d\hat{x},$$

$$\mathtt{II} = \frac{\prod_{j<i}(b^j - a^j) L^{-\frac{\lambda_j}{\lambda_i}} |x^i + y^i|^{\frac{\lambda_j}{\lambda_i}}}{\sqrt{(2\pi)^d \det \mathcal{C}}} \int_{B_{\pm,>i}^{(x^i + y^i)} - y^{>i}} e^{-\frac{1}{2} \tilde{x}^\intercal \mathcal{C}^{-1} \tilde{x}} dx^{>i}.$$

Then, we write

$$(5.38) \qquad \left| g_\epsilon(x^i, y) - g_0(x^i, y) \right| \leq \left| g_\epsilon(x^i, y) - \mathtt{I} \right| + \left| \mathtt{I} - \mathtt{II} \right| + \left| \mathtt{II} - g_0(x^i, y) \right|.$$

We proceed in steps. In each step, we estimate one term on the right of the above display.

Step 1. We estimate $|g_\epsilon(x^i, y) - \mathtt{I}|$ for $|y| \leq K(\epsilon)$. We start by writing

$$|g_\epsilon(x^i, y) - \mathtt{I}| \leq C \epsilon^{-\rho_i} \int_{\hat{B}_\pm^{(x^i + y^i)} - \hat{y}} \left| e^{-\frac{1}{2} x^\intercal \mathcal{C}^{-1} x} - e^{-\frac{1}{2} \tilde{x}^\intercal \mathcal{C}^{-1} \tilde{x}} \right| d\hat{x}$$

Let us estimate the integrand. Recall the estimate (5.36). Using $\frac{\lambda_j}{\lambda_i} - 1 > 0$ for $j < i$, $|y| \leq K(\epsilon)$, and $e^{-c|x|^2}$ to absorb powers of $|x^i|$, we obtain that, for all x, y satisfying

$\hat{x} \in \hat{B}_{\pm}^{(x^i+y^i)} - \hat{y}$ and $|y| \leq K(\epsilon)$,

$$\left| e^{-\frac{1}{2}x^{\mathsf{T}}\mathcal{C}^{-1}x} - e^{-\frac{1}{2}\tilde{x}^{\mathsf{T}}\mathcal{C}^{-1}\tilde{x}} \right| \leq Ce^{-c_1|x^{\geq i}|^2}\left(|x^{<i} + y^{<i}| + |x^{<i} + y^{<i}|^2 \right)$$

$$\leq Ce^{-c_1|x^{\geq i}|^2} \sum_{j<i}(|v^j| + |v^j|^2)\Big|_{v^j=\left|\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(|x^i+y^i|+\epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}}+\epsilon^{\gamma_j}\right|} \leq C\epsilon^{q_1}e^{-c_2|x^{\geq i}|^2},$$

for some $C, c_1, c_2, q_1 > 0$. This, along with the definitions of $\hat{B}_{\pm}^{(x^i+y^i)}$ in (5.20), $B_{\pm,<i}^{(x^i+y^i)}$ in (5.16) and ρ_i in (1.3), implies

(5.39)
$$\begin{aligned}
|g_\epsilon(x^i, y) - \mathtt{I}| &\leq C\epsilon^{-\rho_i}\int_{\hat{B}_{\pm}^{(x^i+y^i)}-\hat{y}} \epsilon^{q_1}e^{-c_2|x^{\geq i}|^2}d\hat{x} \\
&= C\epsilon^{q_1-\rho_i}e^{-c_2|x^i|^2}\left(\int_{B_{\pm,<i}^{(x^i+y^i)}-y^{<i}} dx^{<i}\right)\left(\int_{B_{\pm,>i}^{(x^i+y^i)}-y^{>i}} e^{-c_2|x^{>i}|^2}dx^{>i}\right) \\
&\leq C\epsilon^{q_1-\rho_i}\left| B_{\pm,<i}^{(x^i+y^i)}\right| e^{-c_2|x^i|^2} \\
&\leq C\epsilon^{q_1-\rho_i}\epsilon^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}-1}(|x^i| + 1)^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c_2|x^i|^2} \leq C\epsilon^{q_1}e^{-c_3|x^i|^2}.
\end{aligned}$$

Here and henceforth we use $|B|$ to denote the Lebesgue measure of a set B.

Step 2. We estimate $|\mathtt{I} - \mathtt{II}|$. First note that, by integrating over the first $i - 1$ coordinates in $\mathtt{I}$ and the definition of $B_{\pm,<i}^{(x^i,y^i)}$ in (5.16), we have

$$\mathtt{I} = \frac{\prod_{j<i}\left((b^j - a^j)L^{-\frac{\lambda_j}{\lambda_i}}(|x^i + y^i| \pm \epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}} \pm 2\epsilon^{\gamma_j-(\frac{\lambda_j}{\lambda_i}-1)}\right)}{\prod_{j<i}(b^j - a^j)L^{-\frac{\lambda_j}{\lambda_i}}|x^i + y^i|^{\frac{\lambda_j}{\lambda_i}}}\mathtt{II}.$$

Also, $|\mathtt{II}| \leq C\prod_{j<i}|x^i + y^i|^{\frac{\lambda_j}{\lambda_i}}e^{-c|x^i|^2}$ for some $C > 0$. Hence, using $|y| \leq K(\epsilon)$ and $e^{-c|x^i|^2}$ to absorb powers of $|x^i|$, we can obtain, for some $C, c, q_2 > 0$,

(5.40)
$$\left| \mathtt{I} - \mathtt{II} \right| \leq C\epsilon^{q_2}e^{-c|x^i|^2}.$$

Step 3. We estimate $|\mathtt{II} - g_0(x^i, y)|$. Note that

$$\begin{aligned}
|\mathtt{II} - g_0(x^i, y)| &\leq C\int_{\mathbb{R}^{d-i}\setminus(B_{\pm,>i}^{(x^i+y^i)}-y^{>i})} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c|x^{\geq i}|^2}dx^{>i} \\
&\leq e^{-c|x^i|^2}\sum_{j>i}\int_{\mathbb{R}\setminus(J_{\pm,j}^{(x^i+y^i)}-y^j)} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c|x^j|^2}dx^j.
\end{aligned}$$

We split the integrals after the last inequality into

$$\begin{aligned}
&\int_{\mathbb{R}\setminus(J_{\pm,j}^{(x^i+y^i)}-y^j)} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c|x^j|^2}dx^j \\
&= \int_{b^jL^{-\frac{\lambda_j}{\lambda_i}}\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(|x^i+y^i|\pm\epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}}\pm\epsilon^{\gamma_j}-y^j}^{\infty} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c|x^j|^2}dx^j \\
&\quad + \int_{-\infty}^{a^jL^{-\frac{\lambda_j}{\lambda_i}}\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(|x^i+y^i|\pm\epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}}\mp\epsilon^{\gamma_j}-y^j} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}}e^{-c|x^j|^2}dx^j.
\end{aligned}$$

Choosing $q' > 0$ sufficiently small, we consider two cases. If $|x^i + y^i| \le \epsilon^{q'}$, then for some $q_3 > 0$, the above display is bounded by $C\epsilon^{q_3}$. For the case where $|x^i + y^i| > \epsilon^{q'}$, let us set $c_j = (|a^j| \wedge |b^j|) L^{-\frac{\lambda_j}{\lambda_i}}$ and recall that $\frac{\lambda_j}{\lambda_i} - 1 < 0$ for $j > i$. For some $q_4, q_5, p > 0$, the above display can be bounded by

$$2 \int_{c_j \epsilon^{\frac{\lambda_j}{\lambda_i}-1}(\epsilon^{q'}-\epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}}-\epsilon^{\gamma_j}-K(\epsilon)}^{\infty} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}} e^{-c|x^j|^2} dx^j$$

$$\le C\epsilon^{q_4} |x^i + y^i|^{\sum_{j<i}\frac{\lambda_j}{\lambda_i}} \le C\epsilon^{q_5}(|x^i|^p + 1) \qquad \forall |y| \le K(\epsilon).$$

In deriving the above inequality, we have used $\frac{\lambda_j}{\lambda_i} - 1 < 0$ and chosen $q' < \gamma_i$. Combining the above, we have

(5.41) $\qquad |\mathtt{II} - g_0(x^i, y)| \le C(\epsilon^{q_3} + \epsilon^{q_5}(1 + |x^i|^p))e^{-c|x^i|^2} \le C\epsilon^{q_6}e^{-c'|x^i|^2}.$

To conclude, we insert (5.39), (5.40), and (5.41) into (5.38). As a consequence, we obtain that, for some constants $q', C, c' > 0$, the following holds for ϵ sufficiently small,

$$\sup_{|y| \le K(\epsilon)} |g_\epsilon(x^i, y) - g_0(x^i, y)| \le C\epsilon^{q'} e^{-c'|x^i|^2},$$

as desired.

$\square$

5.3. **Proofs of Lemmas 5.4 and 5.5.**

Proof of Lemma 5.4. Let $\tau_j = \inf\{t > 0 : |Y_t^j| = L\}$. We recall (5.1) and the notation $\tau = \tau_L$. Hence, we have $\tau = \min_{j=1,2,\ldots,d}\{\tau_j\}$. First, we show the following.

Lemma 5.6. *If $[a^j, b^j] \subset (-L, L)$ for all $j \ne i$, then, with ρ_i defined in (1.3),*

$$\sup_{|y| \le K(\epsilon)} \left| \mathbb{P}\{Y_\tau^j \in [a^j, b^j], \forall j \ne i; Y_\tau^i = L\} - \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \ne i; Y_{\tau_i}^i = L\} \right| = o(\epsilon^{\rho_i}).$$

Proof. Since

$$\mathbb{P}\{Y_\tau^j \in [a^j, b^j], \forall j \ne i; Y_\tau^i = L\} = \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \ne i; Y_{\tau_i}^i = L; \tau = \tau_i\},$$

it remains to estimate the right-hand side of

$$\mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \ne i; Y_{\tau_i}^i = L\} - \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \ne i; Y_{\tau_i}^i = L; \tau = \tau_i\}$$
$$= \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \ne i; Y_{\tau_i}^i = L; \tau_i > \tau\}.$$

Using the strong Markov property and setting $c_j = |a^j| \vee |b^j|$, we can bound it by

$$\sum_{j \ne i} \mathbb{P}\{|Y_{\tau_i}^j| \le c_j, \tau_i > \tau_j\}$$

$$\le \sum_{j \ne i} \mathbb{E} \sum_{l=\pm L} \mathbb{1}_{\{Y_{\tau_j}^j = l\}} \mathbb{P}^{Y_{\tau_j}}\{|Y_{\tau_i}^j| \le c_j\} \le \sum_{j \ne i,\ l=\pm L} \mathbb{E}\mathbb{1}_{\{Y_{\tau_j}^j = l\}} \mathbb{P}^{Y_{\tau_j}}\{e^{\lambda_j \tau_i}|l + \epsilon U_{\tau_i}^j| \le c_j\}$$

$$\le \sum_{j \ne i} \mathbb{E}\mathbb{1}_{\{Y_{\tau_j}^j = l\}} \mathbb{P}^{Y_{\tau_j}}\{L - \epsilon|U_{\tau_i}^j| \le c_j\} \le \sum_{j \ne i} \epsilon^p (L - c_j)^{-p} \mathbb{E}|U_{\tau_i}^j|^p$$

for any $p > 0$. Let $p > \rho_i$. By (5.3), there is $C > 0$ such that, for all j, almost surely,

$$\text{(5.42)} \qquad \sup_{t \in [0,\infty)} \langle M^j \rangle_t \leq C \qquad \text{and} \qquad \sup_{t \in [0,\infty)} |V_t^j| \leq C.$$

This, along with BDG inequality, implies that $\mathbb{E}|U_{\tau_i}^j|^p$ is bounded uniformly in ϵ, and completes the proof. $\square$

We will approximate U_{τ_i} by U_{T_0}, where T_0 is given (5.14). By (5.4),

$$\text{(5.43)} \qquad L = |Y_{\tau_i}^i| = \epsilon e^{\lambda_i \tau_i} |y^i + U_{\tau_i}^i|, \qquad \text{or} \qquad \tau_i = \frac{1}{\lambda_i} \log \frac{L}{\epsilon |y^i + U_{\tau_i}^i|}.$$

Now (5.4) and (5.43) give

$$Y_{\tau_i}^j = L^{\frac{\lambda_j}{\lambda_i}} \epsilon^{1 - \frac{\lambda_j}{\lambda_i}} (y^j + U_{\tau_i}^j) |y^i + U_{\tau_i}^i|^{-\frac{\lambda_j}{\lambda_i}},$$

which implies

$$\text{(5.44)} \quad \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \neq i; \ Y_{\tau_i}^i = L\} = \mathbb{P}\{Y_{\tau_i}^j \in [a^j, b^j], \forall j \neq i; \ Y_{\tau_i}^i > 0\}$$

$$= \mathbb{P}\Big\{y^j + U_{\tau_i}^j \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1} |y^i + U_{\tau_i}^i|^{\frac{\lambda_j}{\lambda_i}} [a^j, b^j], \forall j \neq i; \ y^i + U_{\tau_i}^i > 0\Big\}.$$

Then, we compare τ_i with T_0 by showing that, for an appropriate choice of κ,

$$\text{(5.45)} \qquad \mathbb{P}\{\tau_i < T_0\} = \mathbb{P}\{|y^i + U_{\tau_i}^i| > (\log \epsilon^{-1})^{\kappa+1}\} = o(\epsilon^{\rho_i}).$$

By (5.42) and the exponential martingale inequality (see Problem 12.10 in [Bas11]), the following holds uniformly in $|y| \leq K(\epsilon)$ and ϵ sufficiently small,

$$\mathbb{P}\{\tau_i < T_0\} = \mathbb{P}\{|y^i + U_{\tau_i}^i| > (\log \epsilon^{-1})^{\kappa+1}; \tau_i < T_0\} \leq \mathbb{P}\{|y^i + U_{\tau_i \wedge T_0}^i| > (\log \epsilon^{-1})^{\kappa+1}\}$$

$$\leq \mathbb{P}\{|M_{\tau_i \wedge T_0}^i| > (\log \epsilon^{-1})^{\kappa+1} - (\log \epsilon^{-1})^{\kappa} - C\epsilon\}$$

$$\leq \mathbb{P}\{|M_{\tau_i \wedge T_0}^j| > \tfrac{1}{2}(\log \epsilon^{-1})^{\kappa}\}$$

$$\leq 2 \exp\big(-(8C)^{-1}(\log \epsilon^{-1})^{2\kappa}\big).$$

Therefore, it suffices to choose κ large enough (see Remark 5.2) to guarantee (5.45). So, with high probability, $\tau_i \geq T_0$. Let us choose δ to satisfy

$$0 < \delta < 2\frac{\lambda_d}{\lambda_i} = 2 \min_j \{\tfrac{\lambda_j}{\lambda_i}\} < 2.$$

Using the boundedness of F, and G, we can write for some $C_\delta > 0$:

$$\text{(5.46)} \quad \langle M^j \rangle_{\tau_i \vee T_0} - \langle M^j \rangle_{T_0} = \int_{T_0}^{\tau_i \vee T_0} e^{-2\lambda_j T_0} |F^j(Y_{s \wedge \tau})|^2 ds$$

$$\leq Ce^{-2\lambda_j T_0} \leq C\epsilon^{2\frac{\lambda_j}{\lambda_i}} (\log \epsilon^{-1})^{2\frac{\lambda_j}{\lambda_i}(\kappa+1)} \leq C_\delta \epsilon^{2\frac{\lambda_j}{\lambda_i} - \delta}$$

and

$$\text{(5.47)} \qquad |V_{\tau_i \vee T_0}^j - V_{T_0}^j| \leq Ce^{-\lambda_j T_0} \leq C_\delta \epsilon^{\frac{\lambda_j}{\lambda_i} - \frac{1}{2}\delta}.$$

Then we can choose $\gamma_j > 0$, $j \in \{1, 2, \ldots, d\}$ to satisfy, as anticipated in (5.15),

$$\text{(5.48)} \qquad 0 \vee \Big(\frac{\lambda_j}{\lambda_i} - 1\Big) < \gamma_j < \frac{\lambda_j}{\lambda_i} - \frac{1}{2}\delta.$$

By the exponential martingale inequality, estimates (5.46), (5.47), and the second inequality in (5.48), we have

$$(5.49) \quad \begin{aligned} \mathbb{P}\{|U^j_{\tau_i \vee T_0} - U^j_{T_0}| > \epsilon^{\gamma_j}\} &\leq \mathbb{P}\{|M^j_{\tau_i \vee T_0} - M^j_{T_0}| > \tfrac{1}{2}\epsilon^{\gamma_j}\} + \mathbb{P}\{\epsilon|V^j_{\tau_i \vee T_0} - V^j_{T_0}| > \tfrac{1}{2}\epsilon^{\gamma_j}\} \\ &\leq 2\exp\big(-\tfrac{1}{2C_\delta}\epsilon^{2\gamma_j - 2\frac{\lambda_j}{\lambda_i} + \delta}\big) + \mathbb{P}\{C_\delta \epsilon^{\frac{\lambda_j}{\lambda_i} - \frac{1}{2}\delta + 1} > \tfrac{1}{2}\epsilon^{\gamma_j}\} = o\big(\epsilon^{\rho_i}\big), \text{ for all } j. \end{aligned}$$

To see the upper bound in Lemma 5.4, observe that

$$\begin{aligned} \mathbb{P}\{Y^j_\tau \in [a^j, b^j], \forall j \neq i; Y^i_\tau = L\} &\leq \mathbb{P}\{Y^j_{\tau_i} \in [a^j, b^j], \forall j \neq i; Y^i_{\tau_i} = L\} + o\big(\epsilon^{\rho_i}\big) \\ &= \mathbb{P}\Big\{y^j + U^j_{\tau_i} \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1}|y^i + U^i_{\tau_i}|^{\frac{\lambda_j}{\lambda_i}}[a^j, b^j], \forall j \neq i; \ y^i + U^i_{\tau_i} > 0\Big\} + o\big(\epsilon^{\rho_i}\big) \\ &\leq \mathbb{P}\Big\{y^j + U^j_{\tau_i \vee T_0} \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1}|y^i + U^i_{\tau_i \vee T_0}|^{\frac{\lambda_j}{\lambda_i}}[a^j, b^j], \forall j \neq i; \\ &\qquad\qquad\qquad\qquad\qquad y^i + U^i_{\tau_i \vee T_0} \in \big(0, (\log \epsilon^{-1})^{\kappa+1}\big]\Big\} + o\big(\epsilon^{\rho_i}\big) \\ &\leq \mathbb{P}\{y + U_{T_0} \in B_+\} + o\big(\epsilon^{\rho_i}\big) \end{aligned}$$

where we used Lemma 5.6 in the first inequality, (5.44) in the identity, (5.45) in the second inequality, (5.49) in the third inequality.

For the lower bound, we have

$$\begin{aligned} &\mathbb{P}\{Y^j_\tau \in [a^j, b^j], \forall j \neq i; Y^i_\tau = L\} \\ &\geq \mathbb{P}\Big\{y^j + U^j_{\tau_i} \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1}|y^i + U^i_{\tau_i}|^{\frac{\lambda_j}{\lambda_i}}[a^j, b^j], \forall j \neq i; \quad y^i + U^i_{\tau_i} > 0\Big\} - o\big(\epsilon^{\rho_i}\big) \\ &\geq \mathbb{P}\Big\{y^j + U^j_{\tau_i \vee T_0} \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1}|y^i + U^i_{\tau_i \vee T_0}|^{\frac{\lambda_j}{\lambda_i}}[a^j, b^j], \forall j \neq i; \\ &\qquad\qquad\qquad\qquad\qquad y^i + U^i_{\tau_i \vee T_0} \in \big(0, (\log \epsilon^{-1})^{\kappa+1}\big]\Big\} - o\big(\epsilon^{\rho_i}\big) \\ &\geq \mathbb{P}\Big\{y^j + U^j_{\tau_i \vee T_0} \in L^{-\frac{\lambda_j}{\lambda_i}} \epsilon^{\frac{\lambda_j}{\lambda_i} - 1}\big(|y^i + U^i_{T_0}| - \epsilon^{\gamma_i}\big)^{\frac{\lambda_j}{\lambda_i}}[a^j, b^j], \forall j \neq i; \\ &\qquad\qquad\qquad\qquad\qquad y^i + U^i_{T_0} \in \big(\epsilon^{\gamma_i}, (\log \epsilon^{-1})^{\kappa+1} - \epsilon^{\gamma_i}\big]\Big\} - o\big(\epsilon^{\rho_i}\big) \\ &\geq \mathbb{P}\{y + U_{T_0} \in B_-\} - o\big(\epsilon^{\rho_i}\big) \end{aligned}$$

where we used Lemma 5.6 and (5.44) for the first inequality, (5.45) for the second inequality, (5.49) for the last two inequalities. We remark that in the penultimate inequality the factor $(|y^i + U^i_{T_0}| - \epsilon^{\gamma_i})^{\frac{\lambda_j}{\lambda_i}}$ is well-defined on the event we consider. This completes our proof of Lemma 5.4. $\qquad\square$

In order to prove Lemma 5.5, we recall the density estimates obtained in [BC19, Lemma 4.1] for the same setup and assumptions as in the present paper. For a random variable $\mathcal{X}$ with its value in $\mathbb{R}^d$, its density, if exists, is denoted as $\rho_{\mathcal{X}}$. Since U_t depends on y, we denote its density by $\rho^y_{U_t}$.

Lemma 5.7. *Consider* (5.4) *with* $Y_0 = \epsilon y$. *Let* $\mathtt{p}(x) = \sum_{j,k=1}^d x^{\frac{\lambda_j}{\lambda_k}}$ *for* $x \geq 0$ *and*

$$(5.50) \qquad Z_t^j = \int_0^t e^{-\lambda_j s} F_l^j(0) dW_s^l.$$

Then

(1) *there is a constant* $\theta > 0$ *such that for each* $\upsilon \in (0,1)$ *there are* $C, c, \delta > 0$ *such that, for* ϵ *sufficiently small,*

$$|\rho_{U_{T(\epsilon)}}^y(x) - \rho_{Z_{T(\epsilon)}}(x)| \leq C\epsilon^\delta \big(1 + \mathtt{p}(\epsilon^{1-\upsilon}|y|)\big)e^{-c|x|^2}, \quad x, y \in \mathbb{R}^d,$$

holds for all deterministic $T(\epsilon)$ *with* $1 \leq T(\epsilon) \leq \theta \log \epsilon^{-1}$;

(2) *for each* $\theta' > 0$, *there are constants* C', c', δ' *such that, for* ϵ *sufficiently small,*

$$|\rho_{Z_{T(\epsilon)}}(x) - \rho_{Z_\infty}(x)| \leq C'\epsilon^{\delta'} e^{-c'|x|^2}, \quad x \in \mathbb{R}^d,$$

holds for all deterministic $T(\epsilon)$ *with* $T(\epsilon) \geq \theta' \log \epsilon^{-1}$,

We will derive the following result from Lemma 5.7 and use it to prove Lemma 5.5.

Lemma 5.8. *For each* $\upsilon \in (0,1)$, *there is* $\delta > 0$ *such that*

$$\sup_{|y| \leq \epsilon^{\upsilon-1}} \big|\mathbb{P}\{y + U_{T_0} \in B_\pm\} - \mathbb{P}\{y + Z_{T_0} \in B_\pm\}\big| = o\big(\epsilon^{\rho_i+\delta}\big).$$

Let us derive Lemma 5.5 from these lemmas first and prove Lemma 5.8 after that.

Proof of Lemma 5.5. The definition (5.50) implies that Z_∞ is well-defined and has the same distribution as $\mathcal{Z}$. The definition of $B_\pm$ in (5.16) implies that there is $p > 0$ such that for small ϵ,

$$(5.51) \qquad B_\pm \subset \prod_{j=1}^d \Big(\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(\log \epsilon^{-1})^p[-1,1]\Big).$$

Since there is some $\theta' > 0$ such that $T_0 \geq \theta' \log \epsilon^{-1}$, by part (2) of Lemma 5.7 and (5.51), we obtain that, for ϵ sufficiently small,

$$\big|\mathbb{P}\{y + Z_{T_0} \in B_\pm\} - \mathbb{P}\{y + \mathcal{Z} \in B_\pm\}\big| = o\big(\epsilon^{\rho_i+\delta'}(\log \epsilon^{-1})^{pd}\big), \quad \forall y \in \mathbb{R}^d.$$

The above display and Lemma 5.8 together imply the result of Lemma 5.5. $\qquad\square$

To prove Lemma 5.8, we need some notation. For $\mathbf{v} \in \mathbb{R}^d$, $A \subset \mathbb{R}^d$ and $t \in \mathbb{R}$, we write $e^{\lambda t}\mathbf{v} = (e^{\lambda_j t}\mathbf{v}^j)_{j=1}^d \in \mathbb{R}^d$ and $e^{\lambda t}A = \{e^{\lambda t}x : x \in A\} \subset \mathbb{R}^d$.

Recalling $T_0 = T_0(\epsilon)$ from (5.14) and θ from the statement of Lemma 5.7, we set $N = \min\{n \in \mathbb{N} : \frac{T_0}{n} \leq \theta \log \epsilon^{-1}, \forall \epsilon \in (0, 1/2]\}$ and $t_k = \frac{k}{N}T_0$.

Lemma 5.8 is a specific case of the following result with $k = N$ and $w = 0$:

Lemma 5.9. *For each* $\upsilon \in (0,1)$, *there is a constant* υ' *and constants* $\epsilon_k, C_k, \delta_k$, $k = 1, 2, ..., N$ *such that*

(5.52)

$$\sup_{|y| \leq \epsilon^{\upsilon-1}} \sup_{|w| \leq \epsilon^{\upsilon'-1}} \big|\mathbb{P}^{\epsilon y}\{y + U_{t_k} + e^{-\lambda t_k}w \in B_\pm\} - \mathbb{P}\{y + Z_{t_k} + e^{-\lambda t_k}w \in B_\pm\}\big| \leq C_k\epsilon^{\rho_i+\delta_k},$$

holds for all $k = 1, 2, ..., N$ *and* $\epsilon \in (0, \epsilon_k]$.

Proof of Lemma 5.9. First, let us choose $v' \in (0,1)$ to satisfy

(5.53)
$$\frac{1}{N}\frac{\lambda_j}{\lambda_i} \geq \frac{1}{N}\frac{\lambda_d}{\lambda_i} > v', \quad \text{for all } j = 1, 2, ..., d.$$

For the case $k = 1$, Lemma 5.7 and (5.51) imply that

$$\sup_{|y|\leq\epsilon^{v-1}} \sup_{|w|\leq\epsilon^{v'-1}} \left| \mathbb{P}^{\epsilon y}\{y + U_{t_1} + e^{-\lambda t_1}w \in B_\pm\} - \mathbb{P}\{y + Z_{t_1} + e^{-\lambda t_1}w \in B_\pm\} \right|$$

$$\leq \sup_{|y|\leq\epsilon^{v-1}} \sup_{|w|\leq\epsilon^{v'-1}} \int_{\{x\in\mathbb{R}^d:y+x+e^{-\lambda t_1}w\in B_\pm\}} C\epsilon^\delta \left(1 + \mathsf{p}(\epsilon^{1-v}|y|)\right)e^{-c|x|^2}dx \leq C\epsilon^\delta|B_\pm| \leq C_1\epsilon^{\rho_i+\delta_1}$$

for some $C_1, \delta_1 > 0$.

We proceed by induction. Let $k \leq N$ and let us assume that (5.52) holds for $k-1$. Set $z(u) = e^{\lambda t_{k-1}}(y+u)$. The Markov property of Y_t implies that

$$\mathbb{P}^{\epsilon y}\{y + U_{t_k} + e^{-\lambda t_k}w \in B_\pm\} = \mathbb{P}^{\epsilon y}\{Y_{t_k} + \epsilon w \in \epsilon e^{\lambda t_k}B_\pm\}$$

$$= \mathbb{E}^{\epsilon y}\mathbb{P}^{Y_{t_{k-1}}}\{Y_{t_1} + \epsilon w \in \epsilon e^{\lambda t_k}B_\pm\}$$

$$= \mathbb{E}^{\epsilon y}\left(\mathbb{P}^{\epsilon z(u)}\{z(u) + U_{t_1} + e^{-\lambda t_1}w \in e^{\lambda t_{k-1}}B_\pm\}\big|_{u=U_{t_{k-1}}}\right).$$

To check (5.52) for k and complete the induction step, we must show that the error caused by replacing U_{t_1} by Z_{t_1} and $U_{t_{k-1}}$ by $Z_{t_{k-1}}$ in this expression is small. More precisely, (5.52) for k will follow immediately once we prove that there are $\epsilon_k, \delta', \delta'' > 0$ such that the following relations hold uniformly in $|y| \leq \epsilon^{v-1}$, $|w| \leq \epsilon^{v'-1}$ and $\epsilon \in (0, \epsilon_k]$:

(5.54)
$$|\mathbb{E}^{\epsilon y}A_\epsilon(U_{t_{k-1}}, w) - \mathbb{E}^{\epsilon y}B_\epsilon(U_{t_{k-1}}, w)| = o(\epsilon^{\rho_i+\delta'})$$

and

(5.55)
$$\left|\mathbb{E}^{\epsilon y}B_\epsilon(U_{t_{k-1}}, w) - C_\epsilon(y, w)\right| = o(\epsilon^{\rho_i+\delta''}),$$

where

$$A_\epsilon(u, w) = \mathbb{P}^{\epsilon z(u)}\{z(u) + U_{t_1} + e^{-\lambda t_1}w \in e^{\lambda t_{k-1}}B_\pm\},$$

$$B_\epsilon(u, w) = \mathbb{P}\{z(u) + Z_{t_1} + e^{-\lambda t_1}w \in e^{\lambda t_{k-1}}B_\pm\},$$

$$C_\epsilon(y, w) = \mathbb{P}\{y + Z_{t_k} + e^{-\lambda t_k}w \in B_\pm\}.$$

Let us derive (5.54). Due to part (1) of Lemma 5.7, there are $\delta', C', c' > 0$ such that

(5.56)
$$|A_\epsilon(u, w) - B_\epsilon(u, w)| \leq \int_{\{x\in\mathbb{R}^d:z(u)+x+e^{-\lambda t_1}w\in e^{\lambda t_{k-1}}B_\pm\}} C'\epsilon^{\delta'}\left(1 + \mathsf{p}(\epsilon^{1-v'}|z(u)|)\right)e^{-c'|x|^2}dx.$$

By (5.53), we have, for ϵ sufficiently small,

$$e^{\lambda_j t_{k-1}}\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(\log\epsilon^{-1})^p \leq e^{\lambda_j t_{N-1}}\epsilon^{\frac{\lambda_j}{\lambda_i}-1}(\log\epsilon^{-1})^p \leq \epsilon^{\frac{1}{N}\frac{\lambda_j}{\lambda_i}-1}(\log\epsilon^{-1})^p < \epsilon^{v'-1}.$$

This, together with (5.51), implies that there is a constant $C > 0$, such that if $z(u) + x + e^{-\lambda t_1}w \in e^{\lambda t_{k-1}}B_\pm$ and $|w| \leq \epsilon^{v'-1}$, then

(5.57)
$$\epsilon^{1-v'}|z(u)| \leq C + \epsilon^{1-v'}|x|.$$

Using $e^{-c'|x|^2}$ to absorb polynomials of $|x|$, from (5.56) and (5.57) we obtain, for some $C, c > 0$,

$$|A_\epsilon(u, w) - B_\epsilon(u, w)| \leq \epsilon^{\delta'} \int_{\{x \in \mathbb{R}^d : z(u) + x + e^{-\lambda t_1} w \in e^{\lambda t_{k-1}} B_\pm\}} Ce^{-c|x|^2} dx, \qquad |w| \leq \epsilon^{\upsilon'-1}.$$

Let $\mathcal{N}$ be a centered Gaussian with density proportional to $e^{-c|x|^2}$ and independent of $\mathcal{F}_{t_{k-1}}$. The above display implies that if $|w| \leq \epsilon^{\upsilon'-1}$, then

$$|\mathbb{E}^{\epsilon y} A_\epsilon(U_{t_{k-1}}, w) - \mathbb{E}^{\epsilon y} B_\epsilon(U_{t_{k-1}}, w)| \leq C\epsilon^{\delta'} \mathbb{P}^{\epsilon y}\{y + U_{t_{k-1}} + e^{-\lambda t_k} w + e^{-\lambda t_{k-1}} \mathcal{N} \in B_\pm\}.$$

Each entry of $e^{-\lambda t_1}$ decays like a small positive power of ϵ. So, for small ϵ,

(5.58) $|w| \leq \epsilon^{\upsilon'-1}$ implies $|e^{-\lambda t_1} w| + \log \epsilon^{-1} \leq \epsilon^{\upsilon'-1}.$

Therefore,

$$|\mathbb{E}^{\epsilon y} A_\epsilon(U_{t_{k-1}}, w) - \mathbb{E}^{\epsilon y} B_\epsilon(U_{t_{k-1}}, w)|$$
$$\leq C\epsilon^{\delta'} \mathbb{P}^{\epsilon y}\{y + U_{t_{k-1}} + e^{-\lambda t_{k-1}}(e^{-\lambda t_1} w + \mathcal{N}) \in B_\pm; \; |\mathcal{N}| \leq \log \epsilon^{-1}\} + o(\epsilon^{\rho_i + \delta'})$$
$$\leq C\epsilon^{\delta'} \mathbb{P}\{y + Z_{t_{k-1}} + e^{-\lambda t_{k-1}}(e^{-\lambda t_1} w + \mathcal{N}) \in B_\pm\} + o(\epsilon^{\rho_i + \delta_{k-1} + \delta'}) + o(\epsilon^{\rho_i + \delta'})$$
$$= o(\epsilon^{\rho_i + \delta'}),$$

uniformly in $|y| \leq \epsilon^{\upsilon-1}$ and $|w| \leq \epsilon^{\upsilon'-1}$. Here, in the second inequality we used the induction assumption allowed by (5.58), independence of $\mathcal{N}$, Fubini's theorem, and the superpolynomial decay of $\mathbb{P}\{|\mathcal{N}| > \log \epsilon^{-1}\}$. In the last line we used (5.51), the uniform boundedness of the density of $Z_{t_{k-1}}$, independence of $\mathcal{N}$ and Fubini's theorem. This completes the proof of (5.54).

Let us now prove (5.55). Let $\widetilde{Z}_{t_1}$ be a copy of Z_{t_1} independent of $\mathcal{F}_{t_{k-1}}$. The following holds uniformly in $|y| \leq \epsilon^{\upsilon-1}$ and $|w| \leq \epsilon^{\upsilon'-1}$:

$$\mathbb{E}^{\epsilon y} B_\epsilon(U_{t_{k-1}}, w)$$
$$= \mathbb{P}^{\epsilon y}\{y + U_{t_{k-1}} + e^{-\lambda t_{k-1}}(e^{-\lambda t_1} w + \widetilde{Z}_{t_1}) \in B_\pm\}$$
$$= \mathbb{P}^{\epsilon y}\{y + U_{t_{k-1}} + e^{-\lambda t_{k-1}}(e^{-\lambda t_1} w + \widetilde{Z}_{t_1}) \in B_\pm; \; |\widetilde{Z}_{t_1}| \leq \log \epsilon^{-1}\} + o(\epsilon^{\rho_i + \delta'})$$
$$= \mathbb{P}\{y + Z_{t_{k-1}} + e^{-\lambda t_{k-1}}(e^{-\lambda t_1} w + \widetilde{Z}_{t_1}) \in B_\pm; \; |\widetilde{Z}_{t_1}| \leq \log \epsilon^{-1}\} + o(\epsilon^{\rho_i + \delta_{k-1}}) + o(\epsilon^{\rho_i + \delta'})$$
$$= \mathbb{P}\{y + Z_{t_{k-1}} + e^{-\lambda t_{k-1}}\widetilde{Z}_{t_1} + e^{-\lambda t_k} w \in B_\pm\} + o(\epsilon^{\rho_i + \delta_{k-1} \wedge \delta'}),$$
$$= C_\epsilon(y, w) + o(\epsilon^{\rho_i + \delta_{k-1} \wedge \delta'}),$$

where we used the induction assumption in the third identity allowed by (5.58), independence of $\widetilde{Z}_{t_1}$ and Fubini's theorem. In the last line, we used the identity in distribution between $Z_{t_{k-1}} + e^{-\lambda t_{k-1}} \widetilde{Z}_{t_1}$ and Z_{t_k}. This proves (5.55) with $\delta'' = \delta_{k-1} \wedge \delta'$ completing the induction step and the entire proof. $\square$

6. EXTENSION TO A GENERAL DOMAIN

6.1. **Proof of Proposition 4.1.** We use the notation introduced in (3.5)–(3.9).

Let us first prove the inequality $L'_A > 0$ and part (1). The assumption $i(A) = i$ together with definitions (3.5) and (3.9) implies that for any $L < L(O)$, $\overline{\zeta_L(A)} \cap \Lambda^{i-1} = \emptyset$. Let us

fix any $L_0 < L(O)$. Since the set $\overline{\zeta_{L_0}(A)} \subset \partial\mathcal{B}_{L_0}$ is compact, we can find $r > 0$ such that $\max\{|y^m| : m > i - 1\}/|y| \geq r$ for all $y \in \overline{\zeta_{L_0}(A)}$. Let us choose $t_0 > 0$ such that for all $t > t_0$ the following holds: if $j \leq i - 1 \leq m$ and $|x^m| \leq |x^j|$, then $|e^{\lambda_m t} x^m|/|e^{\lambda_j t} x^j| < r$. Now if L is small enough to ensure that $\bar{S}_{t_0} \mathcal{B}_L \subset \mathcal{B}_{L_0}$, then for every $j \leq i - 1$ and every $x \in \mathbf{F}_L^j$, we are guaranteed that the orbit of x under $\bar{S}$ intersects $\partial\mathcal{B}_{L_0}$ at a point y satisfying $\max\{|y^m| : m > i - 1\}/|y| \leq \max\{|y^m| : m > i - 1\}/|y^j| < r$, so $y \notin \overline{\zeta_{L_0}(A)}$ and thus $x \notin \overline{\zeta_L(A)}$, which completes the proof of $L_A' > 0$ and part (1).

To prove part (2), we fix $L < L_A'$ arbitrarily and recall that $i(A) = i$. It suffices to define

$$
\begin{aligned}
B &= \{x \in \partial\mathcal{B}_L : \; |x_j| \leq L/2 \text{ for all } j > i\}, \\
C &= \{x \in \partial\mathcal{B}_L : \; |x_j| > L/2 \text{ for some } j > i\} = \partial\mathcal{B}_L \setminus B, \\
A_0 &= \zeta_L^{-1}(\zeta_L(A) \cap B) = A \cap \zeta_L^{-1}(B), \\
A_1 &= \zeta_L^{-1}(C).
\end{aligned}
$$

Property (2a) is obvious from the construction. The first of (2b) and (2c) hold due to (1) of Lemma 4.3 and the construction. The second item of (2b) follows from the construction, and the last one follows from part (1) and the definition of B. Lastly, the N-regularity of A_0 and A_1 can be verified through the construction and (2) of Lemma 4.3.

6.2. Proof of Proposition 4.2.
We recall the definitions of τ and τ_L given in (3.4) and (5.1). The goal is to show that the asymptotics of $\mathbb{P}\{X_\tau \in A\}$ is exactly captured by that of $\mathbb{P}\{X_{\tau_L} \in f^{-1} \circ \zeta_L(A)\}$ for suitable $A \subset \partial\mathbf{D}$, by which we are able to prove Proposition 4.2.

To this end, we need to approximate $\zeta_L(A)$. Simply taking a small neighborhood of that set makes it difficult to verify the continuity with respect to the measure $\mathcal{H}^{i-1}(\cdot \cap \mathbf{F}_{L,\delta}^i \cap \Lambda^i)$, which is required in Proposition 5.1. Hence, the following lemma is needed. We recall the definition of $\mathrm{dist}(\cdot, \cdot)$ in (4.4).

Lemma 6.1. *Let $d \geq 2$. Let $A \subset \partial\mathbf{D}$ be N-regular with $i(A) = i$. For $L < L_0$ defined in Proposition 5.1, let $B = \zeta_L(A)$ and assume that $\overline{B} \subset \mathrm{int}_L \mathbf{F}_L^i$. Then, there are two families of Borel sets $(B_\delta)_{\delta > 0}$ and $(B_{-\delta})_{\delta > 0}$ with the following properties:*
  (1) $B_{-\delta} \subset B \subset B_\delta \subset \mathbf{F}_{L,\delta_1}^i$ for some $\delta_1 > 0$ and all $\delta > 0$;
  (2) $\lim_{\delta \to 0} \mathcal{H}^{i-1}(B_{\pm\delta} \cap \Lambda^i) = \mathcal{H}^{i-1}(B \cap \Lambda^i)$ for all $\delta > 0$;
  (3) $B_{\pm\delta}$ are finite unions of rectangles described in Proposition 5.1, whose interiors are pairwise disjoint;
  (4) $\mathrm{dist}(\partial\mathcal{B}_L \setminus B_\delta, B) > 0$ and $\mathrm{dist}(\partial\mathcal{B}_L \setminus B, B_{-\delta}) > 0$ for all $\delta > 0$.

The next result shows that $\mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_{\pm\delta}\}$ is a very good approximation for $\mathbb{P}\{X_\tau \in A\}$.

Lemma 6.2. *Let $L < L_0$. For each $A \subset \partial\mathbf{D}$ as in Lemma 6.1, there is $\delta_0 > 0$ depending on A such that for each $\delta \in (0, \delta_0)$*

(6.1)
$$
\mathcal{O}(e^{-C_\delta \epsilon^{-2}}) + \mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_{-\delta}\} \leq \mathbb{P}\{X_\tau \in A\} \leq \mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_\delta\} + \mathcal{O}(e^{-C_\delta \epsilon^{-2}}),
$$

as $\epsilon \to 0$, for some $C_\delta > 0$ depending on δ.

To prove this lemma, we will need an estimate on the discrepancy between the deterministic path $S_t x$ and the perturbed one, i.e., the process X_t under $\mathbb{P}^x = \mathbb{P}\{\cdot \mid X_0 = x\}$. To that end, we will use the following consequence of Proposition 2.3 and Theorem 2.4 in [Aze80, Chapter III] which is an extension of the standard FW large deviation bound without an assumption of uniform ellipticity of σ. We state it here because we can use it directly for the case $d = 1$ in the proof of Proposition 4.2.

Lemma 6.3. *Let b and σ be Lipschitz and bounded. For all $\epsilon > 0$, let $(X_t^\epsilon)_{t \geq 0}$, be a solution of the Itô equation (1.1) with initial condition $X_0^\epsilon = x$, under a probability measure $\mathbb{P}$ and recall the definition of the flow (S_t) from (3.1). For each deterministic $T > 0$ and $\eta > 0$,*

$$\mathbb{P}\Big\{ \sup_{0 \leq t \leq T} |X_t^x - S_t x| > \eta \Big\} = \mathcal{O}(e^{-C\epsilon^{-2}})$$

holds uniformly in x, where C depends only on T, η and the Lipschitz constant of b.

Proof of Proposition 4.2. First, we consider $d \geq 2$. Splitting A into two sets if necessary, we can assume that $\overline{\zeta_L(A)} \subset \mathrm{int}_L\, \mathbf{F}_{L+}^i$ without loss of generality. Let δ_1 be defined by part (1) of Lemma 6.1. By compactness of $\overline{\zeta_L(A)}$, there is $\Delta \in (0, L_0 \wedge \delta_1)$ such that $\zeta_L(A) \subset \mathbf{F}_{L+,\Delta}^i$. We use (3) of Lemma 6.1 to represent $(\zeta_L(A))_{\pm\delta}$ as a finite union of rectangles with disjoint interiors. Applying Proposition 5.1 to these rectangles and noting that the contribution from (perhaps overlapping) boundaries of these rectangles is 0, we obtain

$$\lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_{\pm\delta}\} = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \mathbb{E}\chi_+^i(\xi_0) \mathcal{H}^{i-1}\{(\zeta_L(A))_{\pm\delta} \cap \Lambda^i\}.$$

Therefore, due to (2) of Lemma 6.1 ,

$$\lim_{\delta \to 0} \lim_{\epsilon \to 0} \epsilon^{-\rho_i} \mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_{\pm\delta}\} = L^{-\sum_{j<i} \frac{\lambda_j}{\lambda_i}} \mathbb{E}\chi_+^i(\xi_0) \mathcal{H}^{i-1}\{\zeta_L(A) \cap \Lambda^i\}.$$

Applying Lemma 6.2 we complete the proof for $d \geq 2$.

In the special case $d = 1$, we have $\partial \mathbf{D} = \{q_-, q_+\}$ with $q_- < 0 < q_+$, and $i = 1$. It suffices to study $\mathbb{P}\{X_\tau = q_\pm\}$. Note that $f^{-1}(\mathbf{F}_{L\pm,\delta}^1) = \{p_\pm\}$ where $p_\pm = f^{-1}(\pm L)$ satisfy $q_- < p_- < 0 < p_+ < q_+$. Proposition 5.1 implies that

$$\lim_{\epsilon \to 0} \mathbb{P}\{X_{\tau_L} = p_\pm\} = \mathbb{E}\chi_\pm^1(\xi_0).$$

Using Lemma 6.3 we conclude that $\mathbb{P}\{X_{\tau_L} = p_\pm;\ X_\tau \neq q_\pm\} = \mathcal{O}(e^{-C\epsilon^{-2}})$, which immediately implies that $\lim_{\epsilon \to 0} \mathbb{P}\{X_\tau = q_\pm\} = \mathbb{E}\chi_\pm^1(\xi_0)$ and completes the proof. $\square$

Proof of Lemma 6.1. Without loss of generality we may assume that $\overline{B} \subset \mathrm{int}_L\, \mathbf{F}_{L+}^i$. Let us choose $\delta_1 > 0$ such that

$$\overline{B} \subset \mathrm{int}_L\, \mathbf{F}_{L+,\delta_1}^i.$$

If $i = 1$, then $\mathbf{F}_{L+,\delta_1}^i \cap \Lambda^1 = B \cap \Lambda^1 = \{p\}$, where $p = (L, 0, \ldots, 0)$. Part (2) of Lemma 4.3 implies $\mathcal{H}^0(\partial_L B \cap \Lambda^1) = 0$, so $p \notin \partial_L B$ and thus $p \in \mathrm{int}_L B$. Hence, we can pick a closed rectangle R on $\mathbf{F}_{L+,\delta_1}^i$ such that $p \in R \subset \mathrm{int}_L B$. Setting $B_{-\delta} = R$ for all $\delta > 0$, we ensure properties (1)—(4) for $B_{-\delta}$. Choosing δ_1 sufficiently small and setting $B_\delta = \mathbf{F}_{L,\delta_1}^i$, for all $\delta > 0$, we ensure properties (1)—(4) for B_δ.

If $i = d$, then $\mathbf{F}^i_{L+,\delta_1} \cap \Lambda^d = \mathbf{F}^i_{L+,\delta_1}$. Since $\mathbf{F}^i_{L+,\delta_1}$ is $(d-1)$-dimensional and flat, the measure $\mathcal{H}^{d-1}(\cdot \cap \mathbf{F}^i_{L+,\delta_1} \cap \Lambda^d) = \mathcal{H}^{d-1}(\cdot \cap \mathbf{F}^i_{L+,\delta_1})$ can be viewed as the $(d-1)$-dimensional Lebesgue measure restricted on $\mathbf{F}^i_{L+,\delta_1}$. By the standard approximation arguments, we can choose $B_{-\delta}$ and B_δ to be two unions of finitely many rectangles, which satisfy (1) and (2). Slightly adjusting the rectangles, we can ensure (4).

For $1 < i < d$, we need an extended version of this construction. We construct the family $(B_\delta)_{\delta>0}$ first. Let us define a closed $(i-1)$-dimensional rectangle

$$Q = \mathbf{F}^i_{L+,\delta_1} \cap \Lambda^i = \{x \in \mathbb{R}^d : |x_1|, \ldots, |x_{i-1}| \leq L - \delta_1; \ x_i = L; \ x_{i+1} = \ldots = x_d = 0\}.$$

For every $\delta > 0$, using the compactness of $\overline{B} \cap Q$ and the fact that

$$(6.2) \qquad\qquad \mathcal{H}^{i-1}(\partial_L B \cap Q) = 0$$

(which follows from the regularity of A and Lemma 4.3), we can find a set G_δ satisfying the following:

$$(6.3) \qquad\qquad G_\delta \text{ is a finite union of open } (i-1)\text{-dimensional rectangles;}$$

$$(6.4) \qquad\qquad \overline{B} \cap Q \subset G_\delta \subset Q;$$

$$(6.5) \qquad\qquad \mathcal{H}^{i-1}(G_\delta \setminus (B \cap Q)) = \mathcal{H}^{i-1}(G_\delta \setminus (\overline{B} \cap Q)) < \delta.$$

Since $Q \setminus G_\delta$ and $\overline{B}$ are compact, we can adjust G_δ to additionally ensure that

$$(6.6) \qquad\qquad \text{dist}(Q \setminus G_\delta, \overline{B}) > 0.$$

Let π be the orthogonal projection onto Λ^i, namely

$$\pi : x \in \mathbb{R}^d \mapsto (x^1, x^2, ..., x^{i-1}, 0, \ldots, 0) \in \mathbb{R}^d.$$

Since $\mathbf{F}^i_{L+,\delta_1} \setminus \overline{B}$ is open and $Q \setminus G_\delta$ is closed in the relative topology of $\mathbf{F}^i_{L,\delta_1}$, (6.6) implies that there is some "thickness" $h(\delta) \in (0, \delta)$ such that

$$(6.7) \qquad\qquad K_\delta = \{x \in \mathbf{F}^i_{L,\delta_1} : \pi(x) \in Q \setminus G_\delta; \quad |x^j| < h(\delta), \ \forall j > i\}$$

satisfies

$$(6.8) \qquad\qquad \text{dist}(K_\delta, B) > 0.$$

Let us define $B_\delta = \mathbf{F}^i_{L+,\delta_1} \setminus K_\delta$. Parts (1) and (4) of the lemma now follow from (6.8).

Using (6.3) and subdividing rectangles if needed we can represent $\overline{G_\delta}$ as a finite union of $(i-1)$-dimensional closed rectangles with disjoint interiors. Part (3) follows now from (6.7) and the definition of B_δ.

Since

$$B_\delta \cap Q = B_\delta \cap \Lambda^i = G_\delta,$$

we have $\mathcal{H}^{i-1}(B_\delta \cap \Lambda^i) = \mathcal{H}^{i-1}(G_\delta)$. Thus,

$$0 \leq \mathcal{H}^{i-1}(B_\delta \cap \Lambda^i) - \mathcal{H}^{i-1}(B \cap \Lambda^i) = \mathcal{H}^{i-1}(G_\delta \setminus (B \cap Q)) < \delta,$$

by (6.4) and (6.5), so part (2) also follows.

To construct $B_{-\delta}$, we apply the same approach to the set $B_- = \mathbf{F}_{L+,\delta_1} \setminus B$ and note that due to the regularity of A, the set B_- satisfies a version of (6.2), namely,

$$\mathcal{H}^{i-1}(\partial_L B_- \cap Q) = 0,$$

so we can find a cover $G_{-\delta}$ of $\overline{B_-} \cap Q$ satisfying the versions of requirements (6.3)–(6.6) with B, G_δ replaced by $B_-, G_{-\delta}$. We can now define $K_{-\delta}$ via B_- and $G_{-\delta}$ similarly to (6.7)–(6.8), and check that properties (1)–(4) hold if we set $B_{-\delta} = K_{-\delta}$. $\qquad\square$

Proof of Lemma 6.2. To derive the upper bound in (6.1), we write

$$\mathbb{P}\{X_\tau \in A\} \leq \mathbb{P}\{f(X_{\tau_L}) \in (\zeta_L(A))_\delta\} + \mathtt{I},$$

where $\mathtt{I} = \mathbb{P}\{f(X_{\tau_L}) \notin (\zeta_L(A))_\delta,\ X_\tau \in A\}$ is the term we need to estimate.

Let $\Gamma = f^{-1}(\partial \mathcal{B}_L)$. Since b is transversal to both Γ and $\partial \mathbf{D}$, the inverse of the map ψ_L defined in (3.8) is Lipschitz on $\partial \mathbf{D}$.

Let us introduce $F_\delta = f^{-1}((\zeta_L(A))_\delta)$ and notice that $\gamma = \mathrm{dist}(\partial \mathbf{D} \setminus \psi_L(F_\delta), A) > 0$ due to the Lipschitz property of ψ_L^{-1} and (4) of Lemma 6.1.

Let $T_0 = \sup\{t(x):\ x \in \Gamma\}$, where $t(\cdot)$ was defined in (3.2), and $T_1 = T_0 + 1$. Due to the same transversality properties, by time T_1, all orbits under S originating from Γ exit $\mathbf{D}$ and end up at distance from $\partial \mathbf{D}$ that is bounded away from 0. Therefore there is $\eta > 0$ such that for every $x \in \Gamma$, and every continuous path $y : [0, T_1] \to \mathbb{R}^d$ such that $\sup_{t \in [0,T_1]} |y(t) - S_t x| \leq \eta$, the point $y_{\mathbf{D}}$ of the first intersection of the path y with $\partial \mathbf{D}$ is well-defined and satisfies $|y_{\mathbf{D}} - \psi_L(x)| < \gamma$.

We can now apply this statement along with Lemma 6.3 to see that

$$\mathtt{I} = \int_{\Gamma \setminus F_\delta} \mathbb{P}\{X_{\tau_L} \in dx\} \mathbb{P}^x\{X_\tau \in A\}$$

$$\leq \int_{\Gamma \setminus F_\delta} \mathbb{P}\{X_{\tau_L} \in dx\} \mathbb{P}^x \left\{ \sup_{0 \leq t \leq T_1} |X_t^x - S_t x| > \eta \right\} = \mathcal{O}(e^{-C\epsilon^{-2}})$$

for some $C = C(\delta) > 0$, which completes the proof of the upper bound in (6.1). The lower bound in (6.1) is derived similarly. $\qquad\square$

References

[AMB11]   Sergio Angel Almada Monter and Yuri Bakhtin. Normal forms approach to diffusion near hyperbolic equilibria. Nonlinearity, 24(6):1883–1907, 2011.

[ASK03]   Dieter Armbruster, Emily Stone, and Vivien Kirk. Noisy heteroclinic networks. Chaos, 13(1):71–86, 2003.

[Aze80]   R. Azencott. Grandes déviations et applications. In Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978), volume 774 of Lecture Notes in Math., pages 1–176. Springer, Berlin, 1980.

[Bak10]   Yuri Bakhtin. Small noise limit for diffusions near heteroclinic networks. Dyn. Syst., 25(3):413–431, 2010.

[Bak11]   Yuri Bakhtin. Noisy heteroclinic networks. Probability Theory and Related Fields, 150(1):1–42, Jun 2011.

[Bak15]   Yuri Bakhtin. On Gumbel limit for the length of reactive paths. Stochastics and Dynamics, 15(01):1550001, 2015.

[Bas11]   Richard F. Bass. Stochastic Processes. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2011.

[BC19]   Yuri Bakhtin and Hong-Bin Chen. Long exit times near a repelling equilibrium. Accepted at Annals Of Applied Probability; available at arXiv e-prints, page arXiv:1908.11840, Aug 2019.

[BPG18]   Yuri Bakhtin and Zsolt Pajor-Gyulai. Tails of exit times from unstable equilibria on the line. Accepted at Journal of Applied Probability; available at arXiv e-prints, page arXiv:1810.05341, Oct 2018.
[BPG19]   Yuri Bakhtin and Zsolt Pajor-Gyulai. Malliavin calculus approach to long exit times from an unstable equilibrium. Ann. Appl. Probab., 29(2):827–850, 04 2019.
[CGLM13] Frédéric Cérou, Arnaud Guyader, Tony Lelièvre, and Florent Malrieu. On the length of one-dimensional reactive paths. ALEA Lat. Am. J. Probab. Math. Stat., 10(1):359–389, 2013.
[Day95]   Martin V. Day. On the exit law from saddle points. Stochastic Process. Appl., 60(2):287–311, 1995.
[Eiz84]   Alexander Eizenberg. The exit distributions for small random perturbations of dynamical systems with a repulsive type stationary point. Stochastics, 12(3-4):251–275, 1984.
[KH95]   Anatole Katok and Boris Hasselblatt. Introduction to the modern theory of dynamical systems, volume 54 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.
[Kif81]   Yuri Kifer. The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point. Israel J. Math., 40(1):74–96, 1981.
[Mik95]   Toshio Mikami. Large deviations for the first exit time on small random perturbations of dynamical systems with a hyperbolic equilibrium point. Hokkaido Math. J., 24(3):491–525, 02 1995.
[SA99]   Emily Stone and Dieter Armbruster. Noise and $O(1)$ amplitude effects on heteroclinic cycles. Chaos: An Interdisciplinary Journal of Nonlinear Science, 9(2):499–506, 1999.
[SH90]   Emily Stone and Philip Holmes. Random perturbations of heteroclinic attractors. SIAM J. Appl. Math., 50(3):726–743, 1990.
[Ste57]   Shlomo Sternberg. Local contractions and a theorem of Poincaré. Amer. J. Math., 79:809–824, 1957.

Courant Institute of Mathematical Sciences, New York University, 251 Mercer St, New York, NY 10012

Email address: bakhtin@cims.nyu.edu, hbchen@cims.nyu.edu