Large Deviations and Metastability Analysis for Heavy-Tailed Dynamical Systems

Xingyu Wang and Chang-Han Rhee

Industrial Engineering and Management Sciences, Northwestern University
Evanston, IL, 60613, USA

July 7, 2023

Abstract

This paper proposes a general framework that integrates the large deviations and metastability analysis of heavy-tailed stochastic dynamical systems. Employing this framework in the context of heavy-tailed stochastic difference/differential equations, we first establish a locally uniform sample path large deviations, and then translate the sample path large deviations into a sharp characterization of the joint distribution of the first exit time and exit location. As a result, we provide the heavy-tailed counterparts of the classical Freidlin-Wentzell and Eyring-Kramers theorems. Our findings also address an open question from [26], unveiling intricate phase transitions in the asymptotics of the first exit times under truncated heavy-tailed noises. Furthermore, we develop a set of machinery that elevates the first exit time analysis to the characterization of global dynamics of stochastic dynamical systems. With the machinery, we uncover the global behavior of the stochastic difference/differential equations and show that, after proper scaling, they converge to continuous-time Markov chains that only visit the widest minima of the potential function.

1 Introduction

The analysis of large deviations and metastability in stochastic dynamical systems has a rich history in probability theory and continues to be a vibrant field of research. For instance, the classical Freidlin-Wentzell theorem (see [51]) analyzed sample-path large deviations of Itô diffusions. Over the past few decades, the theory has seen numerous extensions, including the discrete-time version of Freidlin–Wentzell theorem (see, e.g., [39, 30]), large deviations for finite dimensional processes under relaxed assumptions (see, e.g., [13, 16, 15, 1, 17]), Freidlin–Wentzell-type bounds for infinite dimensional processes (see, e.g., [5, 6, 29]), and large deviations for stochastic partial differential equations (see, e.g., [50, 9, 46, 38]), to name a few. On the other hand, the exponential scaling and the pre-exponents in the asymptotics of first exit times under Brownian perturbations were characterized in the Eyring–Kramers law (see [19, 32]). There have been various theoretical advancements since this seminal work, such as the asymptotic characterization of the most likely exit path and the exit times for Brownian particles under more sophisticated gradient fields (see [35]), results for discrete-time processes (see, e.g., [31, 8]), and applications in queueing systems (see, e.g., [49]). For an alternative perspective on metastability based on potential theory, which diverges from the Freidlin-Wentzell theory, we refer the readers to [4].

In sharp contrast to the classical light-tailed analyses, stochastic dynamical systems exhibit fundamentally different large deviations and metastability behaviors under heavy-tailed perturbations. As shown in [24, 25, 26, 28], when the stochastic processes are driven by heavy-tailed noises, the exit events are typically caused by large perturbations of a small number of components rather than smooth tilting of the dynamics.

In this paper, we provide a general framework for heavy-tailed dynamical systems by developing a set of mathematical machinery that uncovers the interconnection between the large deviations, local stability, and global dynamics of stochastic processes. Building upon this unified framework, we characterize the sample-path large deviations and metastability of heavy-tailed stochastic difference and differential equations, thus offering the heavy-tailed counterparts of Freidlin–Wentzell and Eyring–Kramers theorems. Specifically, the main contributions of this work can be summarized as follows:

- For stochastic difference and differential equations with heavy-tailed increments, we establish a version of sample-path large deviations that is uniform w.r.t. the initial values. This is accomplished by rigorously characterizing a uniform version of the $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence introduced in [34]. In particular, this is the characterization of *catastrophe principle*, or *principle of big jumps*, which was confirmed in the special cases of Lévy processes and random walks in [45]. By extending the catastrophe principle to heavy-tailed stochastic difference/differential equations, we reveal a discrete hierarchy governing the causes and probabilities of rare events in a wide variety of heavy-tailed systems.

- We develop a framework for the first exit time analysis in heavy-tailed Markov processes. Central to the development is the concept of asymptotic atom where the process recurrently enters and (almost) regenerates. Leveraging the uniform version of sample-path large deviations over such asymptotic atoms, we obtain the sharp asymptotics of the joint law of (scaled) exit times and exit locations for heavy-tailed processes. Notably, we address an open question left in [26] and reveal intricate phase transitions in the first exit times of heavy-tailed dynamics under truncation.

- Additionally, we develop a machinery that establishes sample-path convergence to Markov jump processes building on the sharp first exit time and exit location analysis. As a direct application, we study the global dynamics of heavy-tailed stochastic difference and differential equations over a multi-well potential. Our findings systematically characterize an intriguing phenomena that, under truncation, the heavy-tailed processes closely resemble Markov jump processes that completely avoid narrow minima of the potential.

In this paper, we focus on the heavy-tailed phenomena captured by the notion of regular variation. Specifically, let $(Z_i)_{i\geq 1}$ be a sequence of iid random variables such that $\mathbf{E}Z_1 = 0$ and $\mathbf{P}(|Z_1| > x)$ is regularly varying with index $-\alpha$ as $x \to \infty$ for some $\alpha > 1$. That is, there exists some slowly varying function ϕ such that $\mathbf{P}(|Z_1| > x) = \phi(x)x^{-\alpha}$. For any $\eta > 0$ and $x \in \mathbb{R}$, let $\big(X^\eta_j(x)\big)_{j\geq 0}$ solves the stochastic difference equation

$$X^\eta_0(x) = x; \qquad X^\eta_j(x) = X^\eta_{j-1}(x) + \eta a\big(X^\eta_{j-1}(x)\big) + \eta\sigma\big(X^\eta_{j-1}(x)\big)Z_j \ \forall j \geq 1. \tag{1.1}$$

Throughout the paper, we adopt the convention that the subscript denotes time and the superscript η denotes the scaling parameter that tends to zero.

Another object of interest is the stochastic differential equation driven by heavy-tailed Lévy processes. Let L_t be a one-dimensional Lévy process with Lévy measure ν. Suppose that $\mathbf{E}L_1 = 0$ and $\nu\big((-\infty, -x) \cup (x, \infty)\big)$ is regularly varying with index $-\alpha$ as $x \to \infty$ for some $\alpha > 1$. For any $\eta > 0$ and $x \in \mathbb{R}$, we define $\bar{L}^\eta_t \triangleq \eta L_{t/\eta}$ as the scaled version of L_t, and let $Y^\eta_t(x)$ be the solution of the stochastic differential equation

$$Y^\eta_0(x) = x; \qquad dY^\eta_t(x) = a\big(Y^\eta_{t-}(x)\big)dt + \sigma\big(Y^\eta_{t-}(x)\big)d\bar{L}^\eta_t. \tag{1.2}$$

At the crux of this study, there is a fundamental difference between light-tailed and heavy-tailed stochastic dynamical systems. This difference lies in the mechanism through which system-wide rare events arise. In light-tailed systems, the rare events exhibits the conspiracy principle where everything goes wrong a little bit at any moment. In contrast, the catastrophe principle governs the most-likely cause of rare events in heavy-tailed systems, where a few catastrophes (i.e., extreme perturbations) drive the system-wide rare events, while the system's behavior is indistinguishable from the nominal behavior most of the time.

As a preliminary version of the catastrophe principle, the well-known principle of a single big jump characterizes the fact that extreme values in random walks and Lévy processes with regularly varying increments are usually caused by a single large perturbation. This line of investigation was initiated in [40, 41] and extended in [3, 14, 18, 21]. The principle of a single big jump has been derived at the functional level for random walks in [23] and established in a wide variety of stochastic dynamics with dependence structures; see, e.g., [7, 20, 22, 36, 37]. In contrast, the results for regularly varying Lévy processes and random walks developed in [45] embody the more general catastrophe principle at the sample-path level, addressing rare events that require multiple jumps to occur. See also [2] where similar large deviation results were obtained under different scaling.

In comparison,a the study of sample-path large deviations of stochastic dynamical systems, such as $X^\eta_j(x)$ and $Y^\eta_t(x)$ defined in (1.1) and (1.2) respectively, is still at an early stage. The only result we are aware of is [53]. Their key idea is to transfer the sample-path large deviations developed in [45] for Lévy processes onto stochastic differential equations through continuous mapping arguments. However, this approach does not work in general unless the diffusion coefficient $\sigma(\cdot)$ is held constant.

Compared to previous works, this paper establishes the catastrophe principle at much greater generality. In particular, we develop a uniform version of sample-path large deviations for heavy-tailed stochastic dynamical systems, which significantly enhances the subsequent metastability analysis. Take the stochastic difference equation $X^\eta_j(x)$ in (1.1) for example. Let $\boldsymbol{y}_t(x)$ be the solution to the ordinary differential equation (ODE) $d\boldsymbol{y}_t(x)/dt = a\big(\boldsymbol{y}_t(x)\big)$ with the initial condition $\boldsymbol{y}_0(x) = x$. Let $\boldsymbol{X}^\eta(x) \triangleq \big\{X^\eta_{\lfloor t/\eta \rfloor}(x) :\ t \in [0,1]\big\}$ be the time-scaled path of $X^\eta_j(x)$, and note that $\boldsymbol{X}^\eta(x)$ is a random element in $\mathbb{D}$, the space of RCLL functions over $[0,1]$. For a given compact set $A \subset \mathbb{R}$ and non-negative integer k, let $\mathbb{D}^{(k)}_A$ be the subset of $\mathbb{D}$ containing ODE paths $\boldsymbol{y}_\cdot(x)$ with exactly k perturbations and initial value $x \in A$. See Section 2.2 the rigorous definition of the concepts involved. Intuitively speaking, if $B \cap \mathbb{D}^{(k-1)}_A = \emptyset$ for some $B \in \mathbb{D}$, then it takes at least k perturbations for any ODE path $\boldsymbol{y}_\cdot(x)$ with $x \in A$ to enter set B. This index k plays a major role in our sample-path large deviations results. Indeed, for any Borel measurable $B \subseteq \mathbb{D}$ that is bounded away from (i.e., has a strictly positive distance from) $\mathbb{D}^{(k-1)}_A$ under Skorokhod J_1 metric, we obtain the following sharp asymptotics that is uniform w.r.t. any initial value over A:

$$\begin{aligned} \inf_{x\in A} \mathbf{C}^{(k)}\big(B^\circ; x\big) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x\in A} \mathbf{P}\Big(\boldsymbol{X}^\eta(x) \in B\Big)}{\big(\eta^{-1}\mathbf{P}(|Z_1| > \eta^{-1})\big)^k} \\ &\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x\in A} \mathbf{P}\Big(\boldsymbol{X}^\eta(x) \in B\Big)}{\big(\eta^{-1}\mathbf{P}(|Z_1| > \eta^{-1})\big)^k} \leq \sup_{x\in A} \mathbf{C}^{(k)}\big(B^-; x\big) < \infty. \end{aligned} \tag{1.3}$$

Here, $\mathbf{C}^{(k)}(\ \cdot\ ; x)$ is a Borel measure supported on $\mathbb{D}^{(k)}_A$, and B°, B^- are the interior and closure of B, respectively. See Theorem 2.2 for a formal statement of the results. As a manifestation of catastrophe principle, our results show that the index k—the minimum number of jumps needed to enter set B—dictates not only the most likely cause of events $\{\boldsymbol{X}^\eta(x) \in B\}$ (i.e., through at least k large perturbations in $\boldsymbol{X}^\eta(x)$) but also the polynomial rate of decay $(\eta^{-1}\mathbf{P}(|Z_1| > \eta^{-1}))^k$ of the probability $\mathbf{P}(\boldsymbol{X}^\eta(x) \in B)$. To establish uniform asymptotics of form (1.3), the key component is a uniform version of the $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence introduced in [34]. In Section 2.1, we develop the Portmanteau theorem for uniform convergence in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$, which is the backbone supporting our proofs of the uniform sample-path large deviations of form (1.3).

Furthermore, the uniform asymptotics described in (1.3) extend to stochastic dynamical systems beyond $X^\eta_j(x)$. For instance, let $\varphi_c(\cdot)$ be the projection operator from $\mathbb{R}$ onto $[-c, c]$. Let $b > 0$ be the truncation threshold, and define $\big(X^{\eta|b}_j(x)\big)_{j\geq 0}$ through the recursion

$$X^{\eta|b}_0(x) = x; \qquad X^{\eta|b}_j(x) = X^{\eta|b}_{j-1}(x) + \varphi_b\Big(\eta a\big(X^{\eta|b}_{j-1}(x)\big) + \eta\sigma\big(X^{\eta|b}_{j-1}(x)\big)Z_j\Big)\ \forall j \geq 1. \tag{1.4}$$

In other words, $X^{\eta|b}_j(x)$ is the modulated version of $X^{\eta}_j(x)$ where the distance traveled at each step is truncated under b. Theorem 2.3 presents the uniform sample-path large deviations for $X^{\eta|b}_j(x)$, which admits the same form as result (1.3). Similar results can be developed for the stochastic differential equation $Y^{\eta}_t(x)$ and its truncated counterpart, We collect the results in Section 2.2.3.

Next, we investigate the metastability of heavy-tailed stochastic dynamical systems with the drift coefficient set as $a(\cdot) = -U'(\cdot)$ in (1.1) and (1.2) for some potential function $U \in \mathcal{C}^1(\mathbb{R})$. Specifically, let $I = (s_{\text{left}}, s_{\text{right}})$ be some open interval containing the origin. Suppose that the entire domain I falls within the attraction field of the origin in the following sense: for the ODE path $d\boldsymbol{y}_t(x)/dt = -U'(\boldsymbol{y}_t(x))$ with initial condition $\boldsymbol{y}_0(x) = x$, the limit $\lim_{t\to\infty}\boldsymbol{y}_t(x) = 0$ holds for all $x \in I$. As a result, when initialized in I, the deterministic dynamical system will be attracted to and trapped at the origin. In contrast, under the presence of random perturbations, the escape from I becomes possible. The first exit time problem examines the law of the first time a stochastic dynamical system, such as $X^{\eta}_j(x)$ or $Y^{\eta}_t(x)$, exits from I due to the random perturbations. Of particular interest are the asymptotics of the first exit time as the noise magnitude decreases.

Originally motivated by the modeling of chemical reactions, the first exit time problem finds applications in numerous contexts, including physics [10, 11], extreme climate events [42], mathematical finance [48], and queueing systems [49]. The arguably best-known result in this field is the Eyring–Kramers law, which characterizes the exit time of Brownian particles. For references, see, e.g., [35]. Concerning Lévy-driven diffusions, [28, 24] derived the asymptotics of the first exit times under regularly varying noises, and [25] extended the results to the multi-dimensional settings. Furthermore, [26] investigated the case where the Lévy measure ν decays exponentially fast with speed $\nu\big((-\infty,-u]\cup[u,\infty)\big) \approx \exp(-u^{\alpha})$. The results revealed a surprising phase transition in the asymptotics of first exit times based on the index α. The hierarchy of exit times of Lévy-driven Langevin equations is summarized in [27].

Our approach to the first exit time problem relies on a general framework developed in Section 2.3.2. This framework uplifts the sample-path large deviations to first exit time analysis for general Markov chains. At the core of this framework lies the concept of asymptotic atoms, namely recurrent regions at which the process (almost) regenerates upon each visit. Our uniform sample-path large deviations then prove to be the right tool under this framework, empowering us to simultaneously characterize the behavior of the stochastic processes under any initial values over the asymptotic atoms. As an immediate application of the framework, we characterize the asymptotics of the joint law of first exit time and exit locations for a variety of heavy-tailed processes. In essence, under truncation threshold $b > 0$, it requires a minimum of $J^*_b = \lceil |s_{\text{left}}| \vee s_{\text{right}}/b \rceil$ jumps for the truncated dynamics $X^{\eta|b}_j(x)$ to exit from $I = (s_{\text{left}}, s_{\text{right}})$ when initialized at the origin. Theorem 2.6 then implies that for the first exit time $\tau^{\eta|b}(x) = \min\{j \geq 0 :\ X^{\eta}_j(x) \notin I\}$ and the exit location $X^{\eta|b}_{\tau^{\eta|b}(x)}(x)$, their joint law admits the limit (for all $x \in I$)

$$\Big(C^*_b \cdot \eta \cdot \big(\lambda(\eta)\big)^{J^*_b} \cdot \tau^{\eta|b}(x),\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\Big) \Rightarrow (E, V_b) \qquad \text{as } \eta \downarrow 0. \tag{1.5}$$

Here, E, V_b are two independent random variables, where E is Exponential with rate 1 and V_b is generated under some probability measure $C_b(\cdot)$ supported on I^c. C^*_b is a normalization constant, and the scale function $\lambda(\eta) = \eta^{-1} \cdot \mathbf{P}(|Z_1| > \eta^{-1})$ is roughly decaying at a polynomial rate $\eta^{\alpha-1}$ for small η with $\alpha > 1$ being the heavy-tailed index for noises Z_i. See Section 2.3.1 for definitions of the concepts involved. Notably, Theorem 2.6 presents an even stronger result where the asymptotics in (1.5) hold uniformly for initial values x over any compact set within I. Meanwhile, the first exit time analysis for $X^{\eta}_j(x)$ is obtained by sending the truncation threshold b to ∞. Similar first exit time analysis can be carried out for stochastic differential equations $Y^\eta_t(x)$ and the truncated counterparts, and we summarize the results in Section 2.3.3.

Compared to existing works [28, 24] in the regularly varying cases, our results allow for non-constant diffusion coefficient $\sigma(\cdot)$, analyze the impact of truncation, and eliminate the need for conditions such as $U \in \mathcal{C}^3(\mathbb{R})$ or non-degeneracy of $U''(\cdot)$ at the boundary of I. Additionally, our results address an open question that was partially explored in [26] regarding the impact of truncation on the first exit times. While [26] primarily focused on the truncation of Weibullian noises whose tail probability decays at rate $\exp(-u^\alpha)$ for some $\alpha \in (0,1)$, our work provides an important missing piece to the puzzle and unveil the effect of truncation in the regularly varying cases. In particular, we characterize an intricate phase transition in the asymptotics of $\tau^{\eta|b}(x)$ that was not observed in previous works. To be specific, by virtue of result (1.5) we find that the first exit time $\tau^{\eta|b}(x)$ is roughly of order $1/\eta^{1+J^*_b\cdot(\alpha-1)}$ for small η. In other words, the order of the first exit time $\tau^{\eta|b}(x)$ does not vary continuously with b; rather, it exhibits a discretized dependency on b through $J^*_b = \lceil |s_{\text{left}}| \vee s_{\text{right}}/b \rceil$, i.e., the minimum number of jumps required for the exit. This phase transition phenomenon further exemplifies the catastrophe principle under regularly varying noises, as the "cost" function J^*_b dictates not only the most likely cause (i.e., through J^*_b large noises) but also the rarity of the exit (i.e., occurring roughly once every $1/\eta^{1+J^*_b\cdot(\alpha-1)}$ steps).

In Section 2.4, we present a technical framework that connects the local stability and the global dynamics of stochastic processes. Specifically, the framework allows us to uplift the first exit time results to the sample-path convergence to jump processes. The power of this framework becomes evident when combined with the first exit time analysis for heavy-tailed dynamical systems. Indeed, consider a heavy-tailed stochastic process that traverses a multi-well potential U; see Figure 2.1 for an illustration of a potential U and the attraction fields therein. As a direct consequence of the framework, Theorem 2.9 shows the existence, under suitable conditions, of a CTMC $Y^{*|b}_t$ only visiting local minima in the widest attraction fields over U such that

$$\Big(X^{\eta|b}_{\lfloor t_1/\lambda^*_b(\eta)\rfloor}(x), \cdots, X^{\eta|b}_{\lfloor t_k/\lambda^*_b(\eta)\rfloor}(x)\Big) \Rightarrow \Big(Y^{*|b}_{t_1}, \cdots, Y^{*|b}_{t_k}\Big) \qquad \text{as } \eta \downarrow 0 \tag{1.6}$$

for all $k \geq 1$ and $0 < t_1 < \cdots < t_k$, under some time scaling of $\lambda^*_b(\eta)$. In particular, our result uncovers an intriguing phenomenon that, under truncations, the heavy-tailed dynamics (asymptotically) avoid any local minimum over U that is not wide enough. Regarding the concept of the widest attraction fields and the associated local minima, we note that the width is measured by the number of jumps (with sizes bounded by b) required to exit the attraction field, and we refer the readers to Section 2.4 for the rigorous definition.

Some of the results in Section 2.3 and Section 2.4 of this paper have been presented in a preliminary form at a conference [52]. The main focus of [52] was the connection between the metastability analysis of stochastic gradient descent (SGD) and its generalization performance in the context of machine learning. Compared to the ad-hoc approach in [52], this paper provides a systematic framework to study the global dynamics of significantly more general class of heavy-tailed dynamical systems. We also note that (i) by sending the truncation threshold b to ∞ in (1.6), we recover the global dynamics of $X^\eta_j(x)$ in Theorem 2.10; (ii) metastability analysis can be conducted analogously for stochastic differential equation $Y^\eta_t(x)$ and its corresponding truncated dynamics, which are summarized in Section 2.4.3.

The paper is structured as follows. Section 2.1 studies the uniform $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence, and Sections 2.2–2.4 present the main results. Specifically, Section 2.2 develops the sample-path large deviations, Section 2.3 carries out the first exit time analysis, and Section 2.4 presents the sample-path convergence of the global dynamics. Proofs are collected in Sections 3–5.

2 Main Results

This section presents the main results of this paper and discusses the implications. All the proofs are deferred to the later sections.

2.1 Preliminaries

We start with setting frequently used notations and reviewing the concept of $\mathbb{M}$-convergence introduced in [34]. Throughout the paper, we let $[n] \triangleq \{1, 2, \cdots, n\}$ for any positive integer n. Let $\mathbb{N} = \{0, 1, 2, \cdots\}$ be the set of non-negative integers. Let $(\mathbb{S}, \boldsymbol{d})$ be a complete and separable metric space with $\mathscr{S}_{\mathbb{S}}$ being the corresponding Borel σ-algebra. For any $E \subseteq \mathbb{S}$, let E° and E^- be the interior and closure of E, respectively. For any $r > 0$, let $E^r \triangleq \{y \in \mathbb{S} : \boldsymbol{d}(E, y) \leq r\}$ be the r-enlargement of a set E. Here for any set $A \subseteq \mathbb{S}$ and any $x \in \mathbb{S}$, we define $\boldsymbol{d}(A, x) \triangleq \inf\{\boldsymbol{d}(y, x) : y \in A\}$. Also, let $E_r \triangleq ((E^c)^r)^c$ be the r-shrinkage of E. Note that for any E, the enlargement E^r of E is closed, and the shrinkage E_r of E is open. We say that set $A \subseteq \mathbb{S}$ is bounded away from another set $B \subseteq \mathbb{S}$ if $\inf_{x \in A, y \in B} \boldsymbol{d}(x, y) > 0$. For any Borel measure μ on $(\mathbb{S}, \mathscr{S}_{\mathbb{S}})$, let the support of μ (denoted as $\text{supp}(\mu)$) be the smallest closed set C such that $\mu(\mathbb{S} \setminus C) = 0$. For any function $g : \mathbb{S} \to \mathbb{R}$, let $\text{supp}(g) \triangleq \big(\{x \in \mathbb{S} : g(x) \neq 0\}\big)^-$.

Given any Borel measurable subset $\mathbb{C} \subseteq \mathbb{S}$, let $\mathbb{S} \setminus \mathbb{C}$ be a subspace of $\mathbb{S}$ equipped with the relative topology with σ-algebra $\mathscr{S}_{\mathbb{S} \setminus \mathbb{C}} \triangleq \{A \in \mathscr{S}_{\mathbb{S}} : A \subseteq \mathbb{S} \setminus \mathbb{C}\}$. Let

$$\mathbb{M}(\mathbb{S} \setminus \mathbb{C}) \triangleq \Big\{\nu(\cdot) \text{ is a Borel measure on } \mathbb{S} \setminus \mathbb{C} : \ \nu(\mathbb{S} \setminus \mathbb{C}^r) < \infty \ \forall r > 0\Big\}.$$

$\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ can be topologized by the sub-basis constructed using sets of form $\{\nu \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C}) : \nu(f) \in G\}$, where $G \subseteq [0, \infty)$ is open, $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$, and $\mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ is the set of all real-valued, non-negative, bounded and continuous functions with support bounded away from $\mathbb{C}$ (i.e., $f(x) = 0 \ \forall x \in \mathbb{C}^r$ for some $r > 0$). Given a sequence $\mu_n \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ and some $\mu \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$, we say that μ_n converges to μ in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ as $n \to \infty$ if $\lim_{n \to \infty} |\mu_n(f) - \mu(f)| = 0$ for all $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$. See [34] for alternative definitions in the form of a Portmanteau Theorem. When the choice of $\mathbb{S}$ and $\mathbb{C}$ is clear from the context, we simply refer to it as $\mathbb{M}$-convergence. As demonstrated in [45], the sample path large deviations for heavy-tailed stochastic processes can be formulated as $\mathbb{M}$-convergence of scaled processes in the Skorokhod space. In this paper, we introduce a stronger version of $\mathbb{M}$-convergence, which facilitates the metastability analysis in the later sections.

Definition 2.1 (Uniform $\mathbb{M}$-convergence)**.** *Let Θ be a set of indices. Let $\mu^\eta_\theta, \mu_\theta \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ for each $\eta > 0$ and $\theta \in \Theta$. We say that μ^η_θ converges to μ_θ in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ uniformly in θ on Θ as $\eta \to 0$ if*

$$\lim_{\eta \downarrow 0} \sup_{\theta \in \Theta} |\mu^\eta_\theta(f) - \mu_\theta(f)| = 0 \qquad \forall f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C}).$$

If $\{\mu_\theta : \theta \in \Theta\}$ is sequentially compact, a Portmanteau-type theorem holds. The proof of this theorem is provided in Section 3.2.

Theorem 2.1 (Portmanteau theorem for uniform $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence)**.** *Let Θ be a set of indices. Let $\mu^\eta_\theta, \mu_\theta \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ for each $\eta > 0$ and $\theta \in \Theta$. Suppose that for any sequence of measures $(\mu_{\theta_n})_{n \geq 1}$, there exist a sub-sequence $(\mu_{\theta_{n_k}})_{k \geq 1}$ and some $\theta^* \in \Theta$ such that*

$$\lim_{k \to \infty} \mu_{\theta_{n_k}}(f) = \mu_{\theta^*}(f) \qquad \forall f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C}). \tag{2.1}$$

Then the next two statements are equivalent.

(i) μ^η_θ converges to μ_θ in $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ uniformly in θ on Θ as $\eta \downarrow 0$;

(ii) $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu^\eta_\theta(F) - \mu_\theta(F^\epsilon) \leq 0$ and $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu^\eta_\theta(G) - \mu_\theta(G_\epsilon) \geq 0$ for all $\epsilon > 0$, all closed $F \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$, and all open $G \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$.

Furthermore, claims (i) and (ii) both imply the following.

(iii) $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu^\eta_\theta(F) \leq \sup_{\theta \in \Theta} \mu_\theta(F)$ and $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu^\eta_\theta(G) \geq \inf_{\theta \in \Theta} \mu_\theta(G)$ for all closed $F \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$ and all open $G \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$.

2.2 Sample-Path Large Deviations

2.2.1 The Untruncated Case

Let $Z_1, Z_2, \ldots$ be the iid copies of some random variable Z and $\mathcal{F}$ be the σ-algebra generated by $(Z_j)_{j\geq 1}$. Let $\mathcal{F}_j$ be the σ-algebra generated by $Z_1, Z_2, \cdots, Z_j$ and $\mathcal{F}_0 \triangleq \{\emptyset, \Omega\}$. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ be a filtered probability space with filtration $\mathbb{F} = (\mathcal{F}_j)_{j\geq 0}$. The goal of this section is to study the sample-path large deviations for $\big\{X^\eta_j(x) : j \geq 0\big\}$, which is driven by the recursion

$$X^\eta_0(x) = x; \qquad X^\eta_j(x) = X^\eta_{j-1}(x) + \eta a\big(X^\eta_{j-1}(x)\big) + \eta\sigma\big(X^\eta_{j-1}(x)\big)Z_j, \quad \forall j \geq 1 \tag{2.2}$$

as $\eta \downarrow 0$. In particular, we are interested in the case where Z_i's are heavy-tailed. Heavy-tails are typically captured with the notion of regular variation. For any measurable function $\phi : (0,\infty) \to (0,\infty)$, we say that ϕ is regularly varying as $x \to \infty$ with index β (denoted as $\phi(x) \in \mathcal{RV}_\beta(x)$ as $x \to \infty$) if $\lim_{x\to\infty} \phi(tx)/\phi(x) = t^\beta$ for all $t > 0$. For details on the definition and properties of regularly varying functions, see, for example, chapter 2 of [44]. Throughout this paper, we say that a measurable function $\phi(\eta)$ is regularly varying as $\eta \downarrow 0$ with index β if $\lim_{\eta\downarrow 0} \phi(t\eta)/\phi(\eta) = t^\beta$ for any $t > 0$. We denote this as $\phi(\eta) \in \mathcal{RV}_\beta(\eta)$ as $\eta \downarrow 0$. Let

$$H^{(+)}(x) \triangleq \mathbf{P}(Z > x), \quad H^{(-)}(x) \triangleq \mathbf{P}(Z < -x), \quad H(x) \triangleq H^{(+)}(x) + H^{(-)}(x) = \mathbf{P}(|Z| > x). \tag{2.3}$$

We assume the following conditions regarding the law of the random variable Z:

Assumption 1 (Regularly Varying Noises)**.** $\mathbf{E}Z = 0$. *Besides, there exist* $\alpha > 1$ *and* $p^{(+)}, p^{(-)} \in (0,1)$ *with* $p^{(+)} + p^{(-)} = 1$ *such that*

$$H(x) \in \mathcal{RV}_{-\alpha}(x) \ \ \textit{as } x \to \infty; \quad \lim_{x\to\infty} \frac{H^{(+)}(x)}{H(x)} = p^{(+)}; \quad \lim_{x\to\infty} \frac{H^{(-)}(x)}{H(x)} = p^{(-)} = 1 - p^{(+)}.$$

Next, we introduce the following assumptions on the drift coefficient $a : \mathbb{R} \to \mathbb{R}$ and diffusion coefficient $\sigma : \mathbb{R} \to \mathbb{R}$. Note that the lower bounds for C and D in Assumption 2 and 3 are obviously not necessary. However we assume that $C \geq 1$ and $D \geq 1$ w.l.o.g. for the notation simplicity.

Assumption 2 (Lipschitz Continuity)**.** *There exists some* $D \in [1,\infty)$ *such that*

$$|\sigma(x) - \sigma(y)| \vee |a(x) - a(y)| \leq D|x - y| \quad \forall x, y \in \mathbb{R}.$$

Assumption 3 (Nondegeneracy)**.** $\sigma(x) > 0 \ \ \forall x \in \mathbb{R}$.

Assumption 4 (Boundedness)**.** *There exists some* $C \in [1,\infty)$ *such that*

$$|a(x)| \vee |\sigma(x)| \leq C \qquad \forall x \in \mathbb{R}.$$

To present the main results, we set a few notations. Let $(\mathbb{D}[0,T], \boldsymbol{d}_{J_1,[0,T]})$ be a metric space, where $\mathbb{D}[0,T]$ is the space of all RCLL functions on $[0,T]$ and $\boldsymbol{d}_{J_1,[0,T]}$ is the Skorodkhod J_1 metric

$$\boldsymbol{d}_{J_1,[0,T]}(x,y) \triangleq \inf_{\lambda\in\Lambda_T} \sup_{s\in[0,T]} |\lambda(s) - s| \vee |x(\lambda(s)) - y(s)|. \tag{2.4}$$

Here, Λ_T is the set of all homeomorphism on $[0,T]$. Given any $A \subseteq \mathbb{R}$, let $A^{k\uparrow} \triangleq \big\{(t_1, \cdots, t_k) \in A^k : \ t_1 < t_2 < \cdots < t_k\big\}$ be the set of sequences of increasing real numbers with length k on A. For any $k \in \mathbb{N}$ and $T > 0$, define mapping $h^{(k)}_{[0,T]} : \mathbb{R} \times \mathbb{R}^k \times (0,T]^{k\uparrow} \to \mathbb{D}[0,T]$ as follows. Given any $x_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$, and $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,T]^{k\uparrow}$, let $\xi = h^{(k)}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in \mathbb{D}[0,T]$ be the solution to

$$\xi_0 = x_0 \tag{2.5}$$

$$
\frac{d\xi_s}{ds} = a(\xi_s) \qquad \forall s \in [0,T],\ s \neq t_1, \cdots, t_k \tag{2.6}
$$

$$
\xi_s = \xi_{s-} + \sigma(\xi_{s-}) \cdot w_j \qquad \text{if } s = t_j \text{ for some } j \in [k]. \tag{2.7}
$$

Here for any $\xi \in \mathbb{D}[0,T]$ and $t \in (0,T]$, we use $\xi_{t-} = \lim_{s\uparrow t} \xi_s$ to denote the left limit of ξ at t, and we set $\xi_{0-} = \xi_0$. In essence, the mapping $h^{(k)}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t})$ produces the ODE path perturbed by jumps $w_1, \cdots, w_k$ (modulated by the drift coefficient $\sigma(\cdot)$) at times $t_1, \cdots, t_k$. We adopt the convention that $\xi = h^{(0)}_{[0,T]}(x_0)$ is the solution to the ODE $d\xi_s/ds = a(\xi_s)\ \forall s \in [0,T]$ under the initial condition $\xi_0 = x_0$. For any $\alpha > 1$, let ν_α be the (Borel) measure on $\mathbb{R}$ with

$$
\nu_\alpha[x,\infty) = p^{(+)} x^{-\alpha}, \quad \nu_\alpha(-\infty, -x] = p^{(-)} x^{-\alpha}, \quad \forall x > 0. \tag{2.8}
$$

where $p^{(+)}, p^{(-)}$ are the constants in Assumption 1. For any $t > 0$, let $\mathcal{L}_t$ be the Lebesgue measure restricted on $(0,t)$ and $\mathcal{L}^{k\uparrow}_t$ be the Lebesgue measure restricted on $(0,t)^{k\uparrow}$. Given any $T > 0$, $x \in \mathbb{R}$, and $k \geq 0$, let

$$
\mathbf{C}^{(k)}_{[0,T]}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{ h^{(k)}_{[0,T]}\big(x, \boldsymbol{w}, \boldsymbol{t}\big) \in \cdot \Big\} \nu^k_\alpha(d\boldsymbol{w}) \times \mathcal{L}^{k\uparrow}_T(d\boldsymbol{t}) \tag{2.9}
$$

where $\nu^k_\alpha(\cdot)$ is the k-fold product measure of ν_α. For $\big\{X^\eta_j(x) : j \geq 0\big\}$, we define the time-scaled version of the sample path as

$$
\boldsymbol{X}^\eta_{[0,T]}(x) \triangleq \big\{ X^\eta_{\lfloor t/\eta \rfloor}(x) :\ t \in [0,T] \big\}, \quad \forall T > 0 \tag{2.10}
$$

with $\lfloor x \rfloor \triangleq \max\{n \in \mathbb{Z} :\ n \leq x\}$ and $\lceil x \rceil \triangleq \min\{n \in \mathbb{Z} :\ n \geq x\}$. Note that $\boldsymbol{X}^\eta_{[0,T]}(x)$ is a $\mathbb{D}[0,T]$-valued random element. For any $k \in \mathbb{N}$ and $A \subseteq \mathbb{R}$, let

$$
\mathbb{D}^{(k)}_A[0,T] \triangleq h^{(k)}_{[0,T]}\big(A \times \mathbb{R}^k \times (0,T]^{k\uparrow}\big), \quad \forall T > 0 \tag{2.11}
$$

as the set that contains all ODE paths with k perturbations by time T. We adopt the convention that $\mathbb{D}^{(-1)}_A[0,T] \triangleq \emptyset$. Also, for any $\eta > 0$, let

$$
\lambda(\eta) \triangleq \eta^{-1} H(\eta^{-1}).
$$

From Assumption 1, one can see that $\lambda(\eta) \in \mathcal{RV}_{\alpha-1}(\eta)$ as $\eta \downarrow 0$. In case $T = 1$, we suppress the time horizon $[0,1]$ and write $\mathbb{D}$, $\boldsymbol{d}_{J_1}$, $h^{(k)}$, $\mathbf{C}^{(k)}$, $\mathbb{D}^{(k)}_A$, and $\boldsymbol{X}^\eta(x)$ to denote $\mathbb{D}[0,1]$, $\boldsymbol{d}_{J_1;[0,1]}$, $h^{(k)}_{[0,1]}$, $\mathbf{C}^{(k)}_{[0,1]}$, $\mathbb{D}^{(k)}_A[0,1]$, and $\boldsymbol{X}^\eta_{[0,1]}(x)$, respectively.

Now we are ready to state the main results. First, Theorem 2.2 establishes the uniform $\mathbb{M}$-convergence of (the law of) $\boldsymbol{X}^\eta_{[0,T]}(x)$ to $\mathbf{C}^{(k)}(\ \cdot\ ; x)$ and a uniform version of the sample-path large deviation for $\boldsymbol{X}^\eta_{[0,T]}(x)$. The proof is given in Section 3.3.

Theorem 2.2. *Under Assumptions 1, 2, 3, and 4, it holds for any $k \in \mathbb{N}$, $T > 0$, and any compact $A \subseteq \mathbb{R}$ that $\lambda^{-k}(\eta)\mathbf{P}\big(\boldsymbol{X}^\eta_{[0,T]}(x) \in\ \cdot\ \big) \to \mathbf{C}^{(k)}_{[0,T]}(\ \cdot\ ; x)$ in $\mathbb{M}\big(\mathbb{D}[0,T] \setminus \mathbb{D}^{(k-1)}_A[0,T]\big)$ uniformly in x on A as $\eta \to 0$. Furthermore, for any $B \in \mathscr{S}_{\mathbb{D}[0,T]}$ that is bounded away from $\mathbb{D}^{(k-1)}_A[0,T]$,*

$$
\begin{aligned}
\inf_{x \in A} \mathbf{C}^{(k)}_{[0,T]}\big(B^\circ; x\big) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}\big(\boldsymbol{X}^\eta_{[0,T]}(x) \in B\big)}{\lambda^k(\eta)} \\
&\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}\big(\boldsymbol{X}^\eta_{[0,T]}(x) \in B\big)}{\lambda^k(\eta)} \leq \sup_{x \in A} \mathbf{C}^{(k)}_{[0,T]}\big(B^-; x\big) < \infty.
\end{aligned} \tag{2.12}
$$

2.2.2 The Truncated Case

Interestingly enough, the sample-path large deviations for $X_j^\eta(x)$ are obtained by first studying its truncated counterpart of $X_j^\eta(x)$. Specifically, for any $x \in \mathbb{R}$, $b > 0$, and $\eta > 0$, on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, we define

$$X_0^{\eta|b}(x) = x, \qquad X_j^{\eta|b}(x) = X_{j-1}^{\eta|b}(x) + \varphi_b\Big(\eta a\big(X_{j-1}^{\eta|b}(x)\big) + \eta\sigma\big(X_{j-1}^{\eta|b}(x)\big)Z_j\Big) \quad \forall j \geq 1, \tag{2.13}$$

where the truncation operator $\varphi_\cdot(\cdot)$ is defined as

$$\varphi_c(w) \triangleq (w \wedge c) \vee (-c) \quad \forall w \in \mathbb{R}, c > 0. \tag{2.14}$$

Here $u \wedge v = \min\{u, v\}$ and $u \vee v = \max\{u, v\}$. For any $T, \eta, b > 0$, and $x \in \mathbb{R}$, let $\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \triangleq \big\{X_{\lfloor t/\eta \rfloor}^{\eta|b}(x) :\ t \in [0,T]\big\}$ be the time-scaled version of $X_j^{\eta|b}(x)$ embedded in the continuous-time space, and note that $\boldsymbol{X}_{[0,T]}^{\eta|b}(x)$ is a random element taking values in $\mathbb{D}[0,T]$.

For any $b, T \in (0,\infty)$ and $k \in \mathbb{N}$, define the mapping $h_{[0,T]}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0,T]^{k\uparrow} \to \mathbb{D}[0,T]$ as follows. Given any $x_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$, and $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,T]^{k\uparrow}$, let $\xi = h_{[0,T]}^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t})$ be the solution to

$$\xi_0 = x_0; \tag{2.15}$$

$$\frac{d\xi_s}{ds} = a(\xi_s) \quad \forall s \in [0,T],\ s \neq t_1, t_2, \cdots, t_k; \tag{2.16}$$

$$\xi_s = \xi_{s-} + \varphi_b\big(\sigma(\xi_{s-}) \cdot w_j\big) \quad \text{ if } s = t_j \text{ for some } j \in [k] \tag{2.17}$$

The mapping $h_{[0,T]}^{(k)|b}$ can be interpreted as a truncated analog of the mapping $h_{[0,T]}^{(k)}$ defined in (2.5)–(2.7). Here, $h_{[0,T]}^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t})$ also return an ODE path with k jumps, but the size of each jump is truncated under b. For any $b, T > 0$, $A \subseteq \mathbb{R}$ and $k = 0, 1, 2, \cdots$, let

$$\mathbb{D}_A^{(k)|b}[0,T] \triangleq h_{[0,T]}^{(k)|b}\big(A \times \mathbb{R}^k \times (0,T]^{k\uparrow}\big) \tag{2.18}$$

be the set of all ODE paths with k jumps, where the size of each jump is modulated by the drift coefficient $\sigma(\cdot)$ and then truncated under threshold b. We adopt the convention that $\mathbb{D}_A^{(-1)|b}[0,T] \triangleq \emptyset$. Given any $x \in \mathbb{R}$, $k \geq 0$, $b > 0$, and $T > 0$, let

$$\mathbf{C}_{[0,T]}^{(k)|b}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{h_{[0,T]}^{(k)|b}\big(x, \boldsymbol{w}, \boldsymbol{t}\big) \in \cdot\Big\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_T^{k\uparrow}(d\boldsymbol{t}). \tag{2.19}$$

Again, in case that $T = 1$, we set $\boldsymbol{X}^{\eta|b}(x) \triangleq \boldsymbol{X}_{[0,1]}^{\eta|b}(x)$, $h^{(k)|b} \triangleq h_{[0,1]}^{(k)|b}$, $\mathbb{D}_A^{(k)|b} \triangleq \mathbb{D}_A^{(k)|b}[0,1]$, and $\mathbf{C}^{(k)|b} \triangleq \mathbf{C}_{[0,1]}^{(k)|b}$. Now, we are ready to state the main result. See Section 3.3 for the proof.

Theorem 2.3. *Under Assumptions 1, 2, and 3, it holds for any $k \in \mathbb{N}$, any $b, T > 0$, and any compact $A \subseteq \mathbb{R}$ that $\lambda^{-k}(\eta)\mathbf{P}\big(\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \in\ \cdot\ \big) \to \mathbf{C}_{[0,T]}^{(k)|b}(\ \cdot\ ; x)$ in $\mathbb{M}\big(\mathbb{D}[0,T] \setminus \mathbb{D}_A^{(k-1)|b}[0,T]\big)$ uniformly in x on A as $\eta \to 0$. Furthermore, for any $B \in \mathscr{S}_{\mathbb{D}[0,T]}$ that is bounded away from $\mathbb{D}_A^{(k-1)|b}[0,T]$,*

$$\begin{aligned} \inf_{x \in A} \mathbf{C}_{[0,T]}^{(k)|b}\big(B^\circ; x\big) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \in B\big)}{\lambda^k(\eta)} \\ &\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \in B\big)}{\lambda^k(\eta)} \leq \sup_{x \in A} \mathbf{C}_{[0,T]}^{(k)|b}\big(B^-; x\big) < \infty. \end{aligned} \tag{2.20}$$

Here, we provide a high-level description of the proof strategy for Theorems 2.2 and 2.3. Specifically, the proof of Theorem 2.3 consists of two steps.

- First, we establish the asymptotic equivalence between $\boldsymbol{X}^{\eta|b}_{[0,T]}(x)$ and an ODE perturbed by the k "largest" noises in $(Z_j)_{j\leq T/\eta}$, in the sense that they admit the same limit in terms of $\mathbb{M}$-convergence as $\eta \downarrow 0$. The key technical tools are the concentration inequalities we developed in Lemma 3.4 that tightly control the fluctuations in $X^{\eta|b}_j(x)$ between any two "large" noises.
- Then it suffices to study the $\mathbb{M}$-convergence of this perturbed ODE. The foundation of this analysis is the asymptotic law of the top-k largest noises in $(Z_j)_{j\leq T/\eta}$ studied in Lemma 3.5.

See Section 3.3 for the detailed proof and the rigorous definitions of the concepts involved. Regarding Theorem 2.2, note that for b sufficiently large it is highly likely that $X^{\eta}_j(x)$ coincides with $X^{\eta|b}_j(x)$ for the entire period of $j \leq T/\eta$ (that is, the truncation operator φ_b did not come into effect for a long period due to the truncation threshold $b > 0$ being too large). By sending $b \to \infty$ and carefully analyzing the limits involved, we recover the results for $X^{\eta}_j(x)$ and prove Theorem 2.2.

2.2.3 Results for Stochastic Differential Equations

Lastly, we collect the results for the sample-path large deviations of stochastic differential equations driven by heavy-tailed Lévy processes. Recall that any one-dimensional Lévy process $\boldsymbol{L} = \{L_t :\ t \geq 0\}$ can be characterized by its generating triplet (c_L, σ_L, ν) where $c_L \in \mathbb{R}$ is the drift parameter, $\sigma_L \geq 0$ is the magnitude of the Brownian motion term in L_t, and ν is the Lévy measure of the Lévy process L_t characterizing the intensity of jumps in L_t. More precisely, we have the following representation

$$L_t \stackrel{d}{=} c_L t + \sigma_L B_t + \int_{|x|\leq 1} x\big[N([0,t]\times dx) - t\nu(dx)\big] + \int_{|x|>1} xN([0,t]\times dx) \tag{2.21}$$

where B is a standard Brownian motion, the measure ν satisfies $\int(|x|^2 \wedge 1)\nu(dx) < \infty$, and N is a Poisson random measure independent of B with intensity measure $\mathcal{L}_\infty \times \nu$. See chapter 4 of [47] for details. We impose the following assumption that characterizes the heavy-tailedness in the increments of L_t.

Assumption 5. $\mathbf{E}L_1 = 0$. *Besides, there exist $\alpha > 1$ and $p^{(-)}, p^{(+)} \in (0,1)$ such that for $H^{(+)}_L(x) \triangleq \nu(x,\infty)$, $H^{(-)}_L(x) \triangleq \nu(-\infty,-x)$ and $H_L(x) \triangleq \nu\Big((\infty,-x)\cup(x,\infty)\Big)$,*

- *$H_L(x) \in \mathcal{RV}_{-\alpha}(x)$ as $x \to \infty$;*
- *$\lim_{x\to\infty} H^{(+)}_L(x)/H_L(x) = p^{(+)},\ \lim_{x\to\infty} H^{(-)}_L(x)/H_L(x) = p^{(-)}$.*

Consider a filtered probability space $\big(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbf{P}\big)$ satisfying usual hypotheses stated in Chapter I, [43] and supporting the Lévy process L, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_t$ is the σ-algebra generated by $\{L_s : s \in [0,t]\}$. For $\eta \in (0,1]$, define the scaled process

$$\bar{\boldsymbol{L}}^{\eta} \triangleq \big\{\bar{L}^{\eta}_t = \eta L_{t/\eta} :\ t \in [0,1]\big\},$$

and let Y^{η}_t be the solution to SDE

$$Y^{\eta}_0(x) = x, \qquad dY^{\eta}_t(x) = a\big(Y^{\eta}_{t-}(x)\big)dt + \sigma\big(Y^{\eta}_{t-}(x)\big)d\bar{L}^{\eta}_t. \tag{2.22}$$

Recall the definitions of the mapping $h^{(k)}_{[0,T]}$ in (2.5)-(2.7) as well as the measure $\mathbf{C}^{(k)}_{[0,T]}(\ \cdot\ ; x)$ in (2.9). Also, recall the notion of uniform $\mathbb{M}$-convergence introduced in Definition 2.1. Define $\mathbf{Y}^{\eta}_{[0,T]}(x) = \{Y^{\eta}_t(x) :\ t \in [0,T]\}$ as a random element in $\mathbb{D}[0,T]$. In case that $T = 1$, we suppress $[0,1]$ and write $\boldsymbol{Y}^{\eta}(x)$. The next result characterizes the sample-path large deviations for $\mathbf{Y}^{\eta}_{[0,T]}(x)$ by establishing

$\mathbb{M}$-convergence that is uniform in the initial condition x. The proofs are almost identical to those of $X^\eta_j(x)$ and hence omitted to avoid repetition. Let

$$\lambda_L(\eta) \triangleq \eta^{-1}H_L(\eta^{-1}).$$

Theorem 2.4. *Under Assumptions 2, 3, 4, and 5, it holds for any $T>0$, $k \in \mathbb{N}$, and any compact set $A \subseteq \mathbb{R}$ that $\lambda_L^{-k}(\eta)\mathbf{P}\big(\boldsymbol{Y}^\eta_{[0,T]}(x) \in \ \cdot\ \big) \to \mathbf{C}^{(k)}_{[0,T]}(\ \cdot\ ;x)$ in $\mathbb{M}\big(\mathbb{D}[0,T] \setminus \mathbb{D}^{(k-1)}_A[0,T]\big)$ uniformly in x on A as $\eta \to 0$. Furthermore, for any $B \in \mathscr{S}_{\mathbb{D}[0,T]}$ that is bounded away from $\mathbb{D}^{(k-1)}_A[0,T]$,*

$$\begin{aligned}\inf_{x\in A}\mathbf{C}^{(k)}_{[0,T]}\big(B^\circ;x\big) \leq \liminf_{\eta\downarrow 0}\frac{\inf_{x\in A}\mathbf{P}\big(\boldsymbol{Y}^\eta_{[0,T]}(x)\in B\big)}{\lambda_L^k(\eta)} \\ \leq \limsup_{\eta\downarrow 0}\frac{\sup_{x\in A}\mathbf{P}\big(\boldsymbol{Y}^\eta_{[0,T]}(x)\in B\big)}{\lambda_L^k(\eta)} \leq \sup_{x\in A}\mathbf{C}^{(k)}_{[0,T]}\big(B^-;x\big) < \infty.\end{aligned}$$

Analogous to the truncated dynamics $X^{\eta|b}_j(x)$, we introduce processes $Y^{\eta|b}_t(x)$ that can be seen as a modulated version of $Y^\eta_t(x)$ where all jumps are truncated under the threshold value b. More generally, we consider the construction of a sequence of stochastic processes $\big(\boldsymbol{Y}^{\eta|b;(k)}_t(x;f,g)\big)_{k\geq 0}$ given any $f:\mathbb{R}\to\mathbb{R}$ and $g:\mathbb{R}\to\mathbb{R}$ that are Lipschitz continuous. First, for any $x \in \mathbb{R}$ and $t \geq 0$, let

$$dY^{\eta|b;(0)}_t(x;f,g) \triangleq f\big(Y^{\eta|b;(0)}_{t-}(x;f,g)\big)dt + g\big(Y^{\eta|b;(0)}_{t-}(x;f,g)\big)d\bar{L}^\eta_t \tag{2.23}$$

and set $\boldsymbol{Y}^{\eta|b;(0)}(x;f,g) \triangleq \big\{Y^{\eta|b;(0)}_t(x;f,g):\ t\in[0,1]\big\}$) for any $b>0$. As an immediate result of this construction, we have $Y^{\eta|b;(0)}_t(x;a,\sigma) = Y^\eta_t(x)$ and $\boldsymbol{Y}^{\eta|b;(0)}(x;a,\sigma) = \boldsymbol{Y}^\eta(x)$. Next, building upon the process $Y^{\eta|b;(0)}_t(x;f,g)$, we define

$$\tau^{\eta|b;(1)}_Y(x;f,g) \triangleq \min\Big\{t>0:\Big|g\big(Y^{\eta|b;(0)}_{t-}(x;f,g)\big)\cdot\Delta\bar{L}^\eta_t\Big| = \Big|\Delta Y^{\eta|b;(0)}_t(x;f,g)\Big| > b\Big\}, \tag{2.24}$$

$$W^{\eta|b;(1)}_Y(x;f,g) \triangleq \Delta Y^{\eta|b;(0)}_{\tau^{\eta|b;(1)}_Y(x;f,g)}(x;f,g) \tag{2.25}$$

as the arrival time and size of the first jump in $Y^{\eta|b;(0)}_t(x;f,g)$ that is larger than b. Furthermore, by proceeding recursively, we define (for any $k\geq 1$)

$$Y^{\eta|b;(k)}_{\tau^{\eta|b;(k)}_Y(x;f,g)}(x;f,g) \triangleq Y^{\eta|b;(k)}_{\tau^{\eta|b;(k)}_Y(x;f,g)-}(x;f,g) + \varphi_b\Big(W^{\eta|b;(k)}_Y(x;f,g)\Big), \tag{2.26}$$

$$dY^{\eta|b;(k)}_t(x;f,g) \triangleq f\big(Y^{\eta|b;(k)}_{t-}(x;f,g)\big)dt + g\big(Y^{\eta|b;(k)}_{t-}(x;f,g)\big)d\bar{L}^\eta_t \quad \forall t > \tau^{\eta|b;(k)}_Y(x;f,g), \tag{2.27}$$

$$\tau^{\eta|b;(k+1)}_Y(x;f,g) \triangleq \min\Big\{t>\tau^{\eta|b;(k)}_Y(x;f,g):\Big|g\big(Y^{\eta|b;(k)}_{t-}(x;f,g)\big)\cdot\Delta\bar{L}^\eta_t\Big| > b\Big\}, \tag{2.28}$$

$$W^{\eta|b;(k+1)}_Y(x;f,g) \triangleq \Delta Y^{\eta|b;(k)}_{\tau^{\eta|b;(k+1)}_Y(x;f,g)}(x;f,g) \tag{2.29}$$

Lastly, for any $t\geq 0, b>0$ and $x\in\mathbb{R}$, we define (under convention $\tau^{\eta|b}_{Y;f,g}(0;x)=0$)

$$Y^{\eta|b}_t(x) \triangleq \sum_{k\geq 0} Y^{\eta|b;(k)}_t(x;a,\sigma)\cdot\mathbb{I}\Big\{t\in\Big[\tau^{\eta|b;(k)}_Y(x;a,\sigma),\tau^{\eta|b;(k+1)}_Y(x;a,\sigma)\Big)\Big\} \tag{2.30}$$

and let $\mathbf{Y}^{\eta|b}_{[0,T]}(x) \triangleq \big\{Y^{\eta|b}_t(x):\ t\in[0,T]\big\}$. By definition, for any $t\geq 0, b>0, k\geq 0$ and $x\in\mathbb{R}$,

$$Y^{\eta|b}_t(x) = Y^{\eta|b;(k)}_t(x;a,\sigma) \qquad \Longleftrightarrow \qquad t\in\Big[\tau^{\eta|b;(k)}_Y(x;a,\sigma),\tau^{\eta|b;(k+1)}_Y(x;a,\sigma)\Big). \tag{2.31}$$

Again, in case that $T=1$ we suppress $[0,1]$ and write $\boldsymbol{Y}^{\eta|b}(x)$. The next result presents the sample-path large deviations for $Y^{\eta|b}_t(x)$. Once again, the proof is omitted as it closely resembles that of $X^{\eta|b}_j(x)$.

Theorem 2.5. *Under Assumptions 2, 3, and 5, it holds for any* $b, T > 0$, $k \in \mathbb{N}$, *and any compact set* $A \subseteq \mathbb{R}$ *that* $\lambda_L^{-k}(\eta)\mathbf{P}\big(\boldsymbol{Y}^{\eta|b}_{[0,T]}(x) \in \ \cdot\ \big) \to \mathbf{C}^{(k)|b}_{[0,T]}(\ \cdot\ ; x)$ *in* $\mathbb{M}\big(\mathbb{D}[0,T] \setminus \mathbb{D}^{(k-1)|b}_A[0,T]\big)$ *uniformly in* x *on* A *as* $\eta \to 0$. *Furthermore, for any* $B \in \mathscr{S}_{\mathbb{D}[0,T]}$ *that is bounded away from* $\mathbb{D}^{(k-1)|b}_A[0,T]$,

$$\inf_{x\in A} \mathbf{C}^{(k)|b}_{[0,T]}\big(B^\circ; x\big) \leq \liminf_{\eta\downarrow 0} \frac{\inf_{x\in A}\mathbf{P}\big(\boldsymbol{Y}^{\eta|b}_{[0,T]}(x) \in B\big)}{\lambda_L^k(\eta)}$$
$$\leq \limsup_{\eta\downarrow 0} \frac{\sup_{x\in A}\mathbf{P}\big(\boldsymbol{Y}^{\eta|b}_{[0,T]}(x) \in B\big)}{\lambda_L^k(\eta)} \leq \sup_{x\in A} \mathbf{C}^{(k)|b}_{[0,T]}\big(B^-; x\big) < \infty.$$

2.3 First Exit Time Analysis

2.3.1 Results for Stochastic Difference Equations

In this section, we address the first exit time analysis of $X^\eta_j(x)$ and $X^{\eta|b}_j(x)$, defined in (2.2) and (2.13), from an attraction field of some potential with a unique local minimum at the origin. Specifically, throughout Section 2.3.1, we fix an open interval $I \triangleq (s_{\text{left}}, s_{\text{right}})$ where $s_{\text{left}} < 0 < s_{\text{right}}$, and impose the following assumption on $a(\cdot)$.

Assumption 6. $a(0) = 0$. *Besides, it holds for all* $x \in I \setminus \{0\}$ *that* $a(x)x < 0$.

Consider the case where $a(\cdot) = -U'(\cdot)$ for some potential $U \in \mathcal{C}^1(\mathbb{R})$. Assumption 6 then implies that U has a unique local minimum at $x = 0$ over the domain I. Moreover, since $U'(x)x = -a(x)x > 0$ for all $x \in I \setminus \{0\}$, we know that the domain I is a subset of the attraction field of the origin in the following sense: the limit $\lim_{t\to\infty} \boldsymbol{y}_t(x) = 0$ holds for all $x \in I$ where $\boldsymbol{y}_t(x)$ is the solution of ODE

$$\boldsymbol{y}_0(x) = x, \qquad \frac{d\boldsymbol{y}_t(x)}{dt} = a\big(\boldsymbol{y}_t(x)\big) \quad \forall t \geq 0. \tag{2.32}$$

It is worth noticing that Assumption 6 is more flexible than standard assumptions in related works. For instance, [28, 24] required the second-order derivative $U''(\cdot)$ to be non-degenerate at the origin as well as the boundary points of I, with an extra condition of $U \in \mathcal{C}^3$ over a wide enough compact set, and held the drift coefficient $\sigma(\cdot)$ as constant. In contrast, we conduct a first exit time analysis with significantly relaxed assumptions.

Define

$$\tau^\eta(x) \triangleq \min\big\{j \geq 0 :\ X^\eta_j(x) \notin I\big\}, \qquad \tau^{\eta|b}(x) \triangleq \min\big\{j \geq 0 :\ X^{\eta|b}_j(x) \notin I\big\},$$

as the first exit time of $X^\eta_j(x)$ and $X^{\eta|b}_j(x)$ from I, respectively. To facilitate the presentation of the main results, we introduce a few concepts. Define $\breve{g}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0,\infty)^{k\uparrow} \to \mathbb{R}$ as the location of the perturbed ODE at the last jump time:

$$\breve{g}^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \triangleq h^{(k)|b}_{[0,t_k+1]}(x, \boldsymbol{w}, \boldsymbol{t})(t_k) \tag{2.33}$$

where $\boldsymbol{t} = (t_1, \ldots, t_k) \in (0,\infty)^{k\uparrow}$, $\boldsymbol{w} = (w_1, \ldots, w_k) \in \mathbb{R}^k$, and $h^{(k)|b}_{[0,T]} : \mathbb{R} \times \mathbb{R}^k \times (0,T]^{k\uparrow} \to \mathbb{D}[0,T]$ is as defined in (2.15)-(2.17). For $k = 0$, we adopt the convention that $\breve{g}^{(0)|b}(x) = x$. This allows us to define Borel measures (for each $k \geq 1$ and $b > 0$)

$$\breve{\mathbf{C}}^{(k)|b}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{\breve{g}^{(k-1)|b}\big(x + \varphi_b(\sigma(x)\cdot w_0), \boldsymbol{w}, \boldsymbol{t}\big) \in \ \cdot\ \Big\}\nu_\alpha(dw_0) \times \nu_\alpha^{k-1}(d\boldsymbol{w}) \times \mathcal{L}^{k-1\uparrow}_\infty(d\boldsymbol{t}) \tag{2.34}$$

with $\mathcal{L}^{k\uparrow}_\infty$ being the Lebesgue measure restricted on $\{(t_1,\cdots,t_k) \in (0,\infty)^k :\ 0 < t_1 < t_2 < \cdots < t_k\}$. Also, define

$$\breve{\mathbf{C}}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{x + \sigma(x)\cdot w \in \ \cdot\ \Big\}\nu_\alpha(dw). \tag{2.35}$$

In case that $x = 0$, we write $\check{\mathbf{C}}^{(k)|b}(\cdot) \triangleq \check{\mathbf{C}}^{(k)|b}(\ \cdot\ ; 0)$. and $\check{\mathbf{C}}(\cdot) \triangleq \check{\mathbf{C}}(\ \cdot\ ; 0)$. Also, let

$$r \triangleq |s_{\text{left}}| \wedge s_{\text{right}}, \qquad \mathcal{J}_b^* \triangleq \lceil r/b \rceil. \tag{2.36}$$

Here, r is the distance between the origin and I^c, and $\mathcal{J}_b^*$ is the number of jumps required to exit from I if the size of each jump is bounded by b.

Recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$ and $\lambda(\eta) = \eta^{-1}H(\eta^{-1})$. For any $k \geq 1$ we write $\lambda^k(\eta) = \big(\lambda(\eta)\big)^k$. As the main result of this section, Theorem 2.6 provides sharp asymptotics for the joint law of first exit times and exit locations in $X_j^{\eta|b}(x)$ and $X_j^\eta(x)$. The results are obtained through the general framework developed in Section 2.3.2. Specifically, the uniform sample-path large deviations developed in Section 2.2 prove to be the right tool in the first exit time analysis, allowing us to verify Condition 1 uniformly for all initial values over the asymptotic atoms $A(\epsilon) = (-\epsilon, \epsilon)$. See Section 2.3.2 for the general framework and Section 4.3 for the detailed proof of Theorem 2.6.

Theorem 2.6. *Let Assumptions 1, 2, 3, and 6 hold.*

(a) Let $b > 0$ be such that $s_{left}/b \notin \mathbb{Z}$ and $s_{right}/b \notin \mathbb{Z}$. For any $\epsilon > 0$, $t \geq 0$, and measurable set $B \subseteq I^c$,

$$\limsup_{\eta \downarrow 0} \sup_{x \in I_\epsilon} \mathbf{P}\bigg(C_b^* \eta \cdot \lambda^{\mathcal{J}_b^*}(\eta) \tau^{\eta|b}(x) > t;\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B \bigg) \leq \frac{\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^-)}{C_b^*} \cdot \exp(-t),$$

$$\liminf_{\eta \downarrow 0} \inf_{x \in I_\epsilon} \mathbf{P}\bigg(C_b^* \eta \cdot \lambda^{\mathcal{J}_b^*}(\eta) \tau^{\eta|b}(x) > t;\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B \bigg) \geq \frac{\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\circ)}{C_b^*} \cdot \exp(-t)$$

where $C_b^* \triangleq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(I^c)$.

(b) For any $t \geq 0$ and measurable set $B \subseteq I^c$,

$$\limsup_{\eta \downarrow 0} \sup_{x \in I_\epsilon} \mathbf{P}\bigg(C^* \eta \cdot \lambda(\eta) \tau^\eta(x) > t;\ X^\eta_{\tau^\eta(x)}(x) \in B \bigg) \leq \frac{\check{\mathbf{C}}(B^-)}{C^*} \cdot \exp(-t),$$

$$\liminf_{\eta \downarrow 0} \inf_{x \in I_\epsilon} \mathbf{P}\bigg(C^* \eta \cdot \lambda(\eta) \tau^\eta(x) > t;\ X^\eta_{\tau^\eta(x)}(x) \in B \bigg) \geq \frac{\check{\mathbf{C}}(B^\circ)}{C^*} \cdot \exp(-t)$$

where $C^* \triangleq \check{\mathbf{C}}(I^c)$.

2.3.2 General Framework

This section proposes a general framework that allows the analysis of the metastability and global dynamics of stochastic systems based on the sample path large deviations. Consider a general metric space $\mathbb{S}$ and a family of $\mathbb{S}$-valued Markov chains $\big\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\big\}$ parameterized by η, where $x \in \mathbb{S}$ denotes the initial state and j denotes the time index. We use the notation $\boldsymbol{V}^\eta_{[0,T]}(x) \triangleq \{V^\eta_{\lfloor t/\eta \rfloor}(x) : t \in [0,T]\}$ to denote the scaled version of $\{V_j^\eta(x) : j \geq 0\}$ as a $\mathbb{D}[0,T]$-valued random variable. For a given set E, let $\tau_E^\eta(x) \triangleq \min\{j \geq 0 : V_j^\eta(x) \in E\}$ denote $\{V_j^\eta(s) : j \geq 0\}$'s first hitting time of E. We consider an asymptotic domain of attraction $I \subseteq \mathbb{S}$, within which $\boldsymbol{V}^\eta_{[0,T]}(x)$ typically (i.e., as $\eta \downarrow 0$) stays throughout any given time horizon $[0,T]$ as far as $x \in I$. We will make these informal descriptions precise in Condition 1. In many cases, however, $V_\cdot^\eta(x)$ is bound to escape I eventually due to the stochasticity if we do not constrain the time horizon. The goal of this section is to establish an asymptotic limit of the joint distribution of the exit time $\tau^\eta_{I^c}(x)$ and the exit location $V^\eta_{\tau^\eta_{I^c}(x)}(x)$. Throughout this section, we will denote $V^\eta_{\tau^\eta_{I(\epsilon)^c}(x)}(x)$ and $V^\eta_{\tau^\eta_{I^c}(x)}(x)$ with $V^\eta_{\tau_\epsilon}(x)$ and $V^\eta_{\tau}(x)$, respectively, for notation simplicity.

We introduce the notion of asymptotic atom to facilitate the analyses. Let $\{I(\epsilon) \subseteq I : \epsilon > 0\}$ and $\{A(\epsilon) \subseteq \mathbb{S} : \epsilon > 0\}$ be collections of subsets of I such that $\bigcup_{\epsilon > 0} I(\epsilon) = I$ and $\bigcap_{\epsilon > 0} A(\epsilon) \neq \emptyset$. Let $C(\cdot)$ is a probability measure on $\mathbb{S} \setminus I$ satisfying $C(\partial I) = 0$.

Definition 2.2. *We say that* $\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\}$ *possesses an asymptotic atom* $\{A(\epsilon) \subseteq \mathbb{S} : \epsilon > 0\}$ *associated with the domain* I*, location measure* $C(\cdot)$*, scale* $\gamma : (0,\infty) \to (0,\infty)$*, and covering* $\{I(\epsilon) \subseteq I : \epsilon > 0\}$ *if the following holds: For each measurable set* $B \subseteq \mathbb{S}$*, there exist* $\delta_B : (0,\infty) \times (0,\infty) \to (0,\infty)$*,* $\epsilon_B > 0$*, and* $T_B > 0$ *such that*

$$C(B^\circ) - \delta_B(\epsilon, T) \leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) \leq T/\eta;\ V^\eta_{\tau_\epsilon}(x) \in B\big)}{\gamma(\eta) T/\eta} \tag{2.37}$$

$$\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) \leq T/\eta;\ V^\eta_{\tau_\epsilon}(x) \in B\big)}{\gamma(\eta) T/\eta} \leq C(B^-) + \delta_B(\epsilon, T) \tag{2.38}$$

$$\limsup_{\eta \downarrow 0} \frac{\sup_{x \in I(\epsilon)} \mathbf{P}\big(\tau^\eta_{(I(\epsilon)\setminus A(\epsilon))^c}(x) > T/\eta\big)}{\gamma(\eta) T/\eta} = 0 \tag{2.39}$$

$$\liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon)} \mathbf{P}\big(\tau^\eta_{A(\epsilon)}(x) \leq T/\eta\big) = 1 \tag{2.40}$$

for any $\epsilon \leq \epsilon_B$ *and* $T \geq T_B$*, where* $\gamma(\eta)/\eta \to 0$ *as* $\eta \downarrow 0$ *and* δ_B*'s are such that*

$$\lim_{\epsilon \to 0} \lim_{T \to \infty} \delta_B(\epsilon, T) = 0.$$

Condition 1. *A family* $\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\}$ *of Markov chains possesses an asymptotic atom* $\{A(\epsilon) \subseteq \mathbb{S} : \epsilon > 0\}$ *associated with the domain* I*, location measure* $C(\cdot)$*, scale* $\gamma : (0,\infty) \to (0,\infty)$*, and covering* $\{I(\epsilon) \subseteq I : \epsilon > 0\}$*.*

The following theorem is the key result of the general framework. See Section 4.1 for the proof of the theorem.

Theorem 2.7. *If Condition 1 holds, then the first exit time* $\tau^\eta_{I^c}(x)$ *scales as* $1/\gamma(\eta)$*, and the distribution of the location* $V^\eta_\tau(x)$ *at the first exit time converges to* $C(\cdot)$*. Moreover, the convergence is uniform over* $I(\epsilon)$ *for any* $\epsilon > 0$*. That is, for each* $\epsilon > 0$*, measurable* $B \subseteq I^c$*, and* $t \geq 0$*,*

$$C(B^\circ) \cdot e^{-t} \leq \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon)} \mathbf{P}\big(\gamma(\eta)\tau^\eta_{I^c}(x) > t,\ V^\eta_\tau(x) \in B\big)$$
$$\leq \limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon)} \mathbf{P}\big(\gamma(\eta)\tau^\eta_{I^c}(x) > t,\ V^\eta_\tau(x) \in B\big) \leq C(B^-) \cdot e^{-t}.$$

2.3.3 Results for Stochastic Differential Equations

Define stopping times

$$\tau^\eta_Y(x) \triangleq \inf\big\{t \geq 0 :\ Y^\eta_t(x) \notin I\big\}, \qquad \tau^{\eta|b}_Y(x) \triangleq \inf\big\{t \geq 0 :\ Y^{\eta|b}_t(x) \notin I\big\}.$$

as the first exit times of $Y^\eta_t(x)$ or $Y^{\eta|b}_t(x)$ from $I = (s_{\text{left}}, s_{\text{right}})$. Analogous to Theorem 2.6, the following result characterizes the asymptotic law of the first exit times $\tau^\eta_Y(x)$ and $\tau^{\eta|b}_Y(x)$ using the measures $\check{\mathbf{C}}^{(k)|b}(\cdot)$ defined in (2.34) and $\check{\mathbf{C}}(\cdot)$ defined in (2.35). We omit the proof due to its similarity to that of Theorem 2.6.

Theorem 2.8. *Let Assumptions 2, 3, 5, and 6 hold.*

(a) Let $b > 0$ *be such that* $s_{left}/b \notin \mathbb{Z}$ *and* $s_{right}/b \notin \mathbb{Z}$*. For any* $\epsilon > 0$*,* $t > 0$*, and measurable set* $B \subseteq I^c$*,*

$$\limsup_{\eta \downarrow 0} \sup_{x \in I_\epsilon} \mathbf{P}\left(C^*_b \lambda_L^{\mathcal{J}^*_b}(\eta) \tau^{\eta|b}_Y(x) > t;\ Y^{\eta|b}_{\tau^{\eta|b}_Y(x)}(x) \in B\right) \leq \frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}(B^-)}{C^*_b} \cdot \exp(-t),$$

$$\liminf_{\eta\downarrow 0}\inf_{x\in I_\epsilon}\mathbf{P}\bigg(C_b^*\lambda_L^{\mathcal{J}_b^*}(\eta)\tau_Y^{\eta|b}(x) > t;\ Y^{\eta|b}_{\tau_Y^{\eta|b}(x)}(x)\in B\bigg) \geq \frac{\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\circ)}{C_b^*}\cdot\exp(-t)$$

where $C_b^* \triangleq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(I^c)$.

(b) *For any* $t > 0$ *and measurable set* $B \subseteq I^c$,

$$\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\bigg(C^*\lambda_L(\eta)\tau_Y^{\eta}(x) > t;\ Y^{\eta}_{\tau_Y^{\eta}(x)}(x)\in B\bigg) \leq \frac{\check{\mathbf{C}}(B^-)}{C^*}\cdot\exp(-t),$$

$$\liminf_{\eta\downarrow 0}\inf_{x\in I_\epsilon}\mathbf{P}\bigg(C^*\lambda_L(\eta)\tau_Y^{\eta}(x) > t;\ Y^{\eta}_{\tau_Y^{\eta}(x)}(x)\in B\bigg) \geq \frac{\check{\mathbf{C}}(B^\circ)}{C^*}\cdot\exp(-t)$$

where $C^* \triangleq \check{\mathbf{C}}(I^c)$.

2.4 Sample-Path Convergence of Global Dynamics

2.4.1 Problem Setting and Main Results

Throughout Section 2.4, we set $a(\cdot) = -U'(\cdot)$ for some potential function $U : \mathbb{R} \to \mathbb{R}$ satisfying the following assumption.

Assumption 7. *Let* $U : \mathbb{R} \to \mathbb{R}$ *be a function in* $\mathcal{C}^1(\mathbb{R})$. *Besides, there exist a positive integer* $n_{min} \geq 2$ *and an ordered sequence of real numbers* $-\infty < m_1 < s_1 < m_2 < s_2 < \cdots < s_{n_{min}-1} < m_{n_{min}} < \infty$ *such that (under the convention* $s_0 = -\infty$ *and* $s_{n_{min}} = \infty$*)*

(i) $U'(x) = 0$ *iff* $x \in \{m_1, s_1, \cdots, s_{n_{min}-1}, m_{n_{min}}\}$;

(ii) $U'(x) < 0$ *for all* $x \in \bigcup_{j\in[n_{min}]}(s_{j-1}, m_j)$;

(iii) $U'(x) > 0$ *for all* $x \in \bigcup_{j\in[n_{min}]}(m_j, s_j)$.

See Figure 2.1 (Left) for an illustration of such function U with $n_{\min} = 3$. According to Assumption 7, the potential function U has finitely many local minima m_i. Meanwhile, the local maxima $s_1, \cdots, s_{n_{\min}-1}$ partition $\mathbb{R}$ into different regions $I_i \triangleq (s_{i-1}, s_i)$. Such regions are viewed as the attraction fields of the local minima m_i's: as the name suggests, any ODE $d\boldsymbol{y}_t(x)/dt = -U'(\boldsymbol{y}_t(x))$ with initial condition $\boldsymbol{y}_0(x) = x \in I_i$ admits the limit $\lim_{t\to\infty}\boldsymbol{y}_t(x) = m_i$. Building upon the first exit time analysis in Section 2.3, we characterize the global dynamics of $X^\eta_j(x)$ and $X^{\eta|b}_j(x)$. Note that we impose the condition $n_{\min} \geq 2$ simply to avoid the trivial case of $n_{\min} = 1$: in this case, no transition between different attraction fields will be observed due to the simple fact that there only exists one attraction field over potential U.

In order to present the main results, we introduce some concepts to help characterize the geometry of U. First, for each attraction field I_i, let

$$r_i \triangleq |m_i - s_{i-1}| \wedge |s_i - m_i| \tag{2.41}$$

be the effective radius of I_i, i.e., the minimum distance required to exit from I_i when starting from m_i. Next, for any $i \in [n_{\min}]$ and $j \in [n_{\min}]$ with $j \neq i$, let

$$\mathcal{J}_b^*(i) \triangleq \lceil r_i/b \rceil, \qquad \mathcal{J}_b^*(i,j) \triangleq \begin{cases} \lceil (s_{j-1} - m_i)/b \rceil & \text{if } j > i, \\ \lceil (m_i - s_j)/b \rceil & \text{if } j < i. \end{cases} \tag{2.42}$$

Here $\mathcal{J}_b^*(i)$ can be interpreted as the minimum number of jumps (with sizes bounded by b) required to escape from I_i, which also reflects the the width of I_i relative to the truncation threshold b. Besides, $\mathcal{J}_b^*(i,j)$ is the distance from m_i to I_j when measured against the truncation threshold $b > 0$. By definition, we must have $\mathcal{J}_b^*(i,j) \geq \mathcal{J}_b^*(i)$. Furthermore, the introduction of $\mathcal{J}_b^*(i)$ and $\mathcal{J}_b^*(i,j)$ allows us to formally develop the concept of typical transition graph.

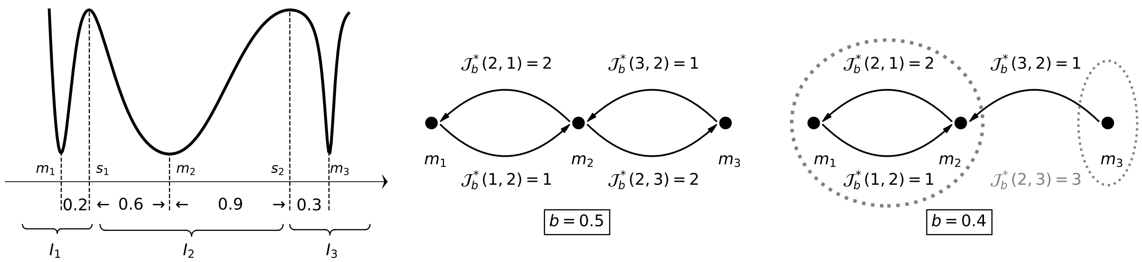

Figure 2.1: Typical transition graphs $\mathcal{G}_b$ under different gradient clipping thresholds b. **(Left)** The potential function U illustrated here has 3 attraction fields. For the second one $I_2 = (s_1, s_2)$, we have $s_2 - m_2 = 0.9, m_2 - s_1 = 0.6$. **(Middle)** The typical transition graph induced by $b = 0.5$. The entire graph $\mathcal{G}_b$ is irreducible since all nodes communicate with each other. **(Right)** The typical transition graph induced by $b = 0.4$. When $b = 0.4$, since $0.6 < 2b$ and $0.9 > 2b$, we have $\mathcal{J}_b^*(2,1) = 2$ and $\mathcal{J}_b^*(2,3) = 3$, and hence $\mathcal{J}_b^*(2) = 2 = \mathcal{J}_b^*(2,1) < \mathcal{J}_b^*(2,3)$. Therefore, the graph $\mathcal{G}_b$ does not contain the edge $m_2 \to m_3$ and there are two communication classes: $G_1 = \{m_1, m_2\}, G_2 = \{m_3\}$.

Definition 2.3 (Typical Transition Graph)**.** *Given a function U satisfying Assumption 7 and some $b > 0$, the b-typical transition graph is a directed graph $\mathcal{G}_b = (V, E_b)$ such that*

- *$V = \{m_1, \cdots, m_{n_{min}}\}$;*
- *An edge $(m_i \to m_j)$ is in E_b iff $\mathcal{J}_b^*(i,j) = \mathcal{J}_b^*(i)$.*

The graph $\mathcal{G}_b$ can be decomposed into different communication classes that are mutually exclusive. Specifically, for $m_i, m_j \in V$ with $i \neq j$, we say that m_i and m_j communicate if and only if there exists a path $(m_i \to m_{k_1} \to \cdots \to m_{k_n} \to m_j)$ as well as a path $(m_j \to m_{k'_1} \to \cdots \to m_{k'_{n'}} \to m_i)$ on $\mathcal{G}_b$. In this section we focus on the case where $\mathcal{G}_b$ is irreducible, i.e., all nodes communicate with each other on graph $\mathcal{G}_b$. See Figure 2.1 (Middle) and (Right) for the illustration of irreducible and reducible cases, respectively.

Now, we are ready to present Theorem 2.9 and show that, under proper time scaling, $X_j^{\eta|b}(x)$ converges (in terms of finite dimensional distributions) to a continuous-time Markov chain that only visits the widest attraction fields over U. Here, the width of each attraction field I_i is characterized by the relative width metric $\mathcal{J}_b^*(i)$ defined in (2.42). We use

$$\mathcal{J}_b^*(V) = \max_{i \in [n_{\min}]:\ m_i \in V} \mathcal{J}_b^*(i) \tag{2.43}$$

to denote the largest width (relative to the truncation threshold $b > 0$) among all attraction fields. Next, define

$$V_b^* \triangleq \{m_i :\ i \in [n_{\min}],\ \mathcal{J}_b^*(i) = \mathcal{J}_b^*(V)\} \tag{2.44}$$

as the set that contains all the widest local minima (when measured against the truncation threshold $b > 0$). Recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$ and $\lambda(\eta) = \eta^{-1} H(\eta^{-1}) \in \mathcal{RV}_{\alpha-1}(\eta)$. Define scale function

$$\lambda_b^*(\eta) \triangleq \eta \cdot \big(\lambda(\eta)\big)^{\mathcal{J}_b^*(V)} \in \mathcal{RV}_{\mathcal{J}_b^*(V)\cdot(\alpha-1)+1}(\eta). \tag{2.45}$$

We note that the condition $|s_j - m_i|/b \notin \mathbb{Z}\ \forall i \in [n_{\min}],\ j \in [n_{\min} - 1]$ in Theorem 2.9 is a mild one as it holds almost everywhere but countably many $b > 0$.

Theorem 2.9. *Let Assumptions 1, 2, 3, 4 and 7 hold. Let $b \in (0, \infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j \in [n_{min} - 1]$. Suppose that $\mathcal{G}_b$ is irreducible. There exist a continuous-time Markov chain $Y^{*|b}$ with state space V_b^*, as well as a random mapping π_b independent of $Y_t^{*|b}$ satisfying*

- $\pi_b(m) \equiv m$ *if* $m \in V_b^*$
- $\pi_b(m)$ *is a random variable that only takes value in* V_b^* *if* $m \notin V_b^*$

such that the following claim holds: given any $i \in [n_{min}]$*,* $x \in I_i$*, and* $0 < t_1 < t_2 < \cdots < t_k$*,*

$$\Big(X^{\eta|b}_{\lfloor t_1/\lambda_b^*(\eta)\rfloor}(x), \cdots, X^{\eta|b}_{\lfloor t_k/\lambda_b^*(\eta)\rfloor}(x)\Big) \Rightarrow \Big(Y^{*|b}_{t_1}(\pi(m_i)), \cdots, Y^{*|b}_{t_k}(\pi(m_i))\Big) \qquad \text{as } \eta \downarrow 0.$$

Theorem 2.9 will be established through an abstract framework developed in Section 2.4.2, which uplifts the first exit time analysis results to the sample-path convergence of global dynamics. The laws of $Y^{*|b}$ and π_b are specified in Section 5.1.

The next result studies the sample-path convergence of $X^\eta_j(x)$ (i.e., without truncation). The intuition is that, given any $T > 0$, there is a high chance that $X^\eta_j(x)$ coincides with the truncated $X^{\eta|b}_j(x)$ for all $j \leq T$, especially when the truncation threshold b is large. Therefore, by sending the truncation threshold b in $X^{\eta|b}_j(x)$ to ∞, we recover the results for $X^\eta_j(x)$. We specify the law of the limiting CTMC $Y^*_t(\cdot)$ in Section 5.1 and detail the proof in Section 5.2.

Theorem 2.10. *Let Assumptions 1, 2, 3, 4 and 7 hold. Given any* $i \in [n_{min}]$*,* $x \in I_i$*, and* $0 < t_1 < t_2 < \cdots < t_k$*,*

$$\Big(X^{\eta}_{\lfloor t_1/H(\eta^{-1})\rfloor}(x), \cdots, X^{\eta}_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\Big) \Rightarrow \Big(Y^{*}_{t_1}(m_i), \cdots, Y^{*}_{t_k}(m_i)\Big) \qquad \text{as } \eta \downarrow 0$$

where $H(\cdot) = \mathbf{P}(|Z_1| > \ \cdot\ )$ *and* $Y^*_t(\cdot)$ *is a CTMC with state space* $\{m_1, \cdots, m_{n_{min}}\}$*.*

Finally, we state a direct corollary of Theorem 2.9 that highlights the elimination of sharp minima under truncated heavy-tailed dynamics. Theorem 2.9 reveals that, under small η, the sample path of the truncated dynamics $X^{\eta|b}_j(x)$ closely resembles that of an CTMC that completely avoids all the narrower attraction fields of the potential U. Corollary 2.11 then further demonstrates that the fraction of time $X^{\eta|b}_j(x)$ spends around sharp minima converges in probability to 0 as $\eta \downarrow 0$, thus verifying the elimination effect under truncated heavy-tailed dynamics. See Section 5.5 for the proof.

Corollary 2.11. *Let Assumptions 1, 2, 3, 4 and 7 hold. Let* $b \in (0, \infty)$ *be such that* $|s_j - m_i|/b \notin \mathbb{Z}$ *for all* $i \in [n_{min}]$ *and* $j \in [n_{min} - 1]$*. Suppose that* $\mathcal{G}_b$ *is irreducible. Then given any* $i \in [n_{min}]$*,* $x \in I_i$*, and any* $T > 0$*,*

$$\frac{1}{T}\int_0^T \mathbb{I}\Big\{X^{\eta|b}_{\lfloor t/\lambda_b^*(\eta)\rfloor}(x) \in \bigcup_{j:\ m_j \notin V_b^*} I_j\Big\}dt \xrightarrow{\mathbf{P}} 0 \qquad \text{as } \eta \downarrow 0.$$

2.4.2 General Framework

Consider a general metric space $(\mathbb{S}, \boldsymbol{d})$. Let Y^η_t, $\hat{Y}^{\eta,\epsilon}_t$, and Y^*_t be $\mathbb{S}$-valued stochastic processes supported on the same probability space, Inspired by the approach in [28], we focus on the following condition that characterizes a type of asymptotic equivalence between processes Y^η_t and $\hat{Y}^{\eta,\epsilon}_t$.

Condition 2. *Given any* $0 < t_1 < t_2 < \cdots < t_k = t$*, the following claims hold for all* $\epsilon > 0$ *small enough:*

(i) $\Big(\hat{Y}^{\eta,\epsilon}_{t_1}, \hat{Y}^{\eta,\epsilon}_{t_2}, \cdots, \hat{Y}^{\eta,\epsilon}_{t_k}\Big) \Rightarrow \Big(Y^*_{t_1}, Y^*_{t_2}, \cdots, Y^*_{t_k}\Big)$ *as* $\eta \downarrow 0$*;*

(ii) For any $i \in [k]$*,* $\lim_{\eta\downarrow 0} \mathbf{P}\Big(\boldsymbol{d}(\hat{Y}^{\eta,\epsilon}_{t_i}, Y^\eta_{t_i}) \geq \epsilon\Big) = 0$*.*

As shown in Lemma 2.12 below, Condition 2 establishes a type of asymptotic equivalence between two families of stochastic processes Y^η_t and $\hat{Y}^{\eta,\epsilon}_t$ such that they admit the same limit Y^*_t in terms of finite dimensional distributions. See Section 5.3 for the proof of Lemma 2.12.

Lemma 2.12. *Suppose that Condition 2 holds. Given any $k \geq 1$ and $0 < t_1 < t_2 < \cdots < t_k$,*

$$\Big(Y^\eta_{t_1},\cdots,Y^\eta_{t_k}\Big) \Rightarrow \Big(Y^*_{t_1},\cdots,Y^*_{t_k}\Big) \qquad \textit{as } \eta \downarrow 0.$$

Naturally, the plan is to prove Theorem 2.9 via Lemma 2.12 by setting $Y^\eta_t = X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)$ and $Y^*_t = Y^{*|b}_t(\pi(m_i))$. To identify the right choice of process $\hat{Y}^{\eta,\epsilon}_t$ and facilitate the verification of sample-path convergence characterized in Condition 2, (i), we introduce the second key component of our framework, i.e., a technical tool for establishing the weak convergence at the sample-path level. Specifically, the following definition encapsulates the class of jumps processes considered in this paper.

Definition 2.4. *Let random variables $\Big((U_j)_{j\geq1},(V_j)_{j\geq1}\Big)$ be such that $V_j \in \mathbb{S}\ \forall j \geq 1$ for some general metric space $\mathbb{S}$, $U_j \in [0,\infty)$ for all $j \geq 1$, and $\lim_{i\to\infty}\mathbf{P}(\sum_{j\leq i}U_j > t) = 1\ \forall t > 0$. A continuous-time process Y_t on $\mathbb{R}$ is a $\Big((U_j)_{j\geq1},(V_j)_{j\geq1}\Big)$ jump process if (under the convention $V_0 \equiv 0$)*

$$Y_t = V_{\mathcal{J}(t)}\ \forall t \geq 0, \qquad \mathcal{J}(t) = \max\{J \geq 0:\ \sum_{j=1}^J U_j \leq t\}.$$

We add two remarks regarding this definition. First, $(U_j)_{j\geq1}$ and $(V_j)_{j\geq0}$ can be viewed as the inter-arrival times and destinations of jumps in Y_t, respectively. It is worth noticing that we allow for instantaneous jumps, i.e., $U_j = 0$. Nevertheless, the condition $\lim_{i\to\infty}\mathbf{P}(\sum_{j\leq i}U_j > t) = 1\ \forall t > 0$ prevents the concentration of infinitely many instantaneous jumps before any finite time $t \in (0,\infty)$, thus ensuring that the process $Y_t = V_{\mathcal{J}(t)}$ is almost surely well defined. In case that $U_j > 0\ \forall j$, the jump process Y_t admits the more standard expression $Y_t = V_i \iff t \in [\sum_{j=1}^i U_j, \sum_{j=1}^{i+1}U_j)$. Second, to account for the scenario where the process Y_t stays constant after a (possibly random) timestamp T, one can introduce dummy jumps that keep landing at the same location. For instance, suppose that after hitting $w \in \mathbb{S}$ the process Y_t is absorbed at w, then a representation compatible with Definition 2.4 is that, conditioning on $V_j = w$, we set U_k as iid Exp(1) RVs and $V_k \equiv w$ for all $k \geq j+1$.

As the second key component of the framework, Lemma 2.13 states that, in order to establish the convergence of jump processes, it suffices to verify the convergence of the inter-arrival times and destinations of jumps therein.

Lemma 2.13. *Let the metric space $(\mathbb{S},\boldsymbol{d})$ be separable. Let Y_t be a $\Big((U_j)_{j\geq1},(V_j)_{j\geq1}\Big)$ jumps process and, for each $n \geq 1$, Y^n_t be a $\Big((U^n_j)_{j\geq1},(V^n_j)_{j\geq1}\Big)$ jump process. Suppose that*

- *$(U^n_1,V^n_1,U^n_2,V^n_2,\cdots)$ converges in distribution to $(U_1,V_1,U_2,V_2,\cdots)$ as $n\to\infty$;*
- *For any $u > 0$ and any $j \geq 1$, $\mathbf{P}(U_1+\cdots+U_j = u) = 0$;*
- *For any $u > 0$, $\lim_{j\to\infty}\mathbf{P}(U_1+U_2+\cdots U_j > u) = 1$.*

Then for any $k \geq 1$ and $0 < t_1 < t_2 < \cdots < t_k < \infty$, $\Big(Y^n_{t_1},\cdots,Y^n_{t_k}\Big) \Rightarrow \Big(Y_{t_1},\cdots,Y_{t_k}\Big)$ as $n\to\infty$.

The verification of Condition 2, (i) hinges on the chioce of the approximator $\hat{Y}^{\eta,\epsilon}_t$. To this end, we construct a process $\hat{X}^{\eta,\epsilon|b}_t(x)$ as follows. Let (under convention $\hat{\tau}^{\eta,\epsilon|b}_0(x) \equiv 0$)

$$\hat{\tau}^{\eta,\epsilon|b}_1(x) \triangleq \min\Big\{j \geq 0:\ X^{\eta|b}_j(x) \in \bigcup_{i\in[n_{\min}]}(m_i-\epsilon,m_i+\epsilon)\Big\}, \tag{2.46}$$

$$\hat{\tau}^{\eta,\epsilon|b}_k(x) \triangleq \min\Big\{ j \geq \hat{\tau}^{\eta,\epsilon|b}_{k-1}(x): \ X^{\eta|b}_j(x) \in \bigcup_{i \neq \hat{\mathcal{I}}^{\eta,\epsilon|b}_{k-1}(x)} (m_i - \epsilon, m_i + \epsilon) \Big\} \qquad \forall k \geq 2. \tag{2.47}$$

Also, define $\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)$ by the rule

$$\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x) = i \iff X^{\eta|b}_{\hat{\tau}^{\eta,\epsilon|b}_k(x)}(x) \in I_i. \tag{2.48}$$

Essentially, $\hat{\tau}^{\eta,\epsilon|b}_k(x)$ records the k-th time $X^{\eta|b}_j(x)$ visits (the ϵ-neighborhood of) a local minimum and $\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)$ denotes the index of the visited local minimum. Let $\hat{X}^{\eta,\epsilon|b}_t(x)$ be the $\Big(\Big(\big(\hat{\tau}^{\eta,\epsilon|b}_k(x) - \hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)\big)\cdot \lambda^*_b(\eta)\Big)_{k\geq 1}, \big(m_{\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)}\big)_{k \geq 1}\Big)$ jump process. By definition, $\hat{X}^{\eta,\epsilon|b}_t(x)$ keeps track of how $X^{\eta|b}_j(x)$ traverses the potential U and makes transitions between the different local minima (under time scaling with $\lambda^*_b(\eta)$).

Using Lemma 2.13, the convergence of the jump process $\hat{X}^{\eta,\epsilon|b}_t(x)$ follows directly from the convergence of $\hat{\tau}^{\eta,\epsilon|b}_k(x) - \hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)$ and $m_{\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)}$, i.e., the inter-arrival times and destinations of the transitions in $X^{\eta|b}_j(x)$ between different attraction fields of U. This is exactly the content of the first exit time analysis. In particular, based on a straightforward adaptation of the first exit time analysis in Section 2.3.1 to the current setup, we obtain Proposition 2.14. The proof is detailed in Section 5.4.

Proposition 2.14 (Verifying Condition 2, (i))**.** *Let Assumptions 1, 2, 3, 4 and 7 hold. Let $b \in (0,\infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j \in [n_{min} - 1]$. Suppose that $\mathcal{G}_b$ is irreducible. Given any $\epsilon > 0$ small enough and any $i \in [n_{min}]$, $x \in I_i$,*

$$\Big(\hat{X}^{\eta,\epsilon|b}_{t_1}(x), \cdots, \hat{X}^{\eta,\epsilon|b}_{t_k}(x)\Big) \Rightarrow \Big(Y^{*|b}_{t_1}\big(\pi_b(m_i)\big), \cdots, Y^{*|b}_{t_k}\big(\pi_b(m_i)\big)\Big) \text{ as } \eta \downarrow 0 \qquad \forall k \geq 1,\ 0 < t_1 < \cdots < t_k.$$

Meanwhile, Proposition 2.15 verifies Condition 2, (ii) and confirms the equivalence between $\hat{X}^{\eta,\epsilon|b}_t(x)$ and $X^{\eta|b}_j(x)$ in the asymptotic sense. We give the proof in Section 5.4.

Proposition 2.15 (Verifying Condition 2, (ii))**.** *Let Assumptions 1, 2, 3, 4 and 7 hold. Let $b \in (0,\infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j \in [n_{min} - 1]$. Suppose that $\mathcal{G}_b$ is irreducible. Let $x \in \bigcup_{i \in [n_{min}]} I_i$. Given any $t > 0$, it holds for all $\epsilon > 0$ small enough that*

$$\lim_{\eta \downarrow 0} \mathbf{P}\bigg(\Big| X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta) \rfloor}(x) - \hat{X}^{\eta,\epsilon|b}_t(x) \Big| \geq \epsilon \bigg) = 0.$$

Now, we are ready to prove Theorem 2.9.

Proof of Theorem 2.9. Fix some $i \in [n_{\min}]$ and $x \in I_i$. Applying Propositions 2.14 and 2.15, we verify the conditions in Lemma 2.12 (under the choice of $Y^\eta_t = X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)$, $\hat{Y}^{\eta,\epsilon}_t = \hat{X}^{\eta,\epsilon|b}_t(x)$, and $Y^*_t = Y^{*|b}_t(\pi_b(m_i))$) and conclude the proof. □

2.4.3 Results for Stochastic Differential Equations

To conclude, we collect the sample-path convergence results for $Y^{\eta|b}_t(x)$ and $Y^\eta_t(x)$. We skip the proof as they are almost identical to those of $X^{\eta|b}_j(x)$ and $X^\eta_j(x)$. Recall the definition of V^*_b in (2.44) as the set that contains all the widest local minima m_i over U (when measured by the truncation threshold $b > 0$). Also, recall that $\lambda_L(\eta) = \eta^{-1}H_L(\eta^{-1})$ and $H_L(x) = H_L(x) = \nu\big((\infty, -x) \cup (x,\infty)\big)$, where ν is the Lévy measure of the Lévy process L_t. Define scale function

$$\lambda^*_{b;L}(\eta) \triangleq \big(\lambda_L(\eta)\big)^{\mathcal{J}^*_b(V)} \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)}(\eta). \tag{2.49}$$

Theorem 2.16. *Let Assumptions 2, 3, 4, 5, and 7 hold. Let $b \in (0,\infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j \in [n_{min} - 1]$. If $\mathcal{G}_b$ is irreducible, then given any $i \in [n_{min}]$, $x \in I_i$, and $0 < t_1 < t_2 < \cdots < t_k$,*

$$\Big(Y^{\eta|b}_{\lfloor t_1/\lambda^*_{b;L}(\eta)\rfloor}(x), \cdots, Y^{\eta|b}_{\lfloor t_k/\lambda^*_{b;L}(\eta)\rfloor}(x)\Big) \Rightarrow \Big(Y^{*|b}_{t_1}(\pi_b(m_i)), \cdots, Y^{*|b}_{t_k}(\pi_b(m_i))\Big) \qquad \text{as } \eta \downarrow 0$$

where the continuous-time Markov chain $Y^{*|b}_t$ and the random mapping $\pi_b(\cdot)$ are characterized in Theorem 2.9.

Theorem 2.17. *Let Assumptions 2, 3, 4, 5, and 7 hold. Given any $i \in [n_{min}]$, $x \in I_i$, and $0 < t_1 < t_2 < \cdots < t_k$,*

$$\Big(Y^{\eta}_{\lfloor t_1/\lambda_L(\eta)\rfloor}(x), \cdots, Y^{\eta}_{\lfloor t_k/\lambda_L(\eta)\rfloor}(x)\Big) \Rightarrow \Big(Y^{*}_{t_1}(m_i), \cdots, Y^{*}_{t_k}(m_i)\Big) \qquad \text{as } \eta \downarrow 0$$

where the continuous-time Markov chain Y^*_t is characterized in Theorem 2.10.

3 Uniform $\mathbb{M}$-Convergence and Sample Path Large Deviations

3.1 Technical Lemmas

Straightforward as they are, the proofs of the next two lemmas are provided for the sake of completeness.

Lemma 3.1. *Let $a : (0,\infty) \to (0,\infty)$, $b : (0,\infty) \to (0,\infty)$ be two functions such that $\lim_{\epsilon\downarrow 0} a(\epsilon) = 0, \lim_{\epsilon\downarrow 0} b(\epsilon) = 0$. Let $\{U(\epsilon) : \epsilon > 0\}$ be a family of geometric RVs with success rate $a(\epsilon)$, i.e. $\mathbf{P}(U(\epsilon) > k) = (1 - a(\epsilon))^k$ for $k \in \mathbb{N}$.*

(i) For any $c > 1$, there exists $\epsilon_0 > 0$ such that

$$\exp\Big(-\frac{c\cdot a(\epsilon)}{b(\epsilon)}\Big) \leq \mathbf{P}\Big(U(\epsilon) > \frac{1}{b(\epsilon)}\Big) \leq \exp\Big(-\frac{a(\epsilon)}{c\cdot b(\epsilon)}\Big) \quad \forall \epsilon \in (0,\epsilon_0).$$

(ii) Suppose that, in addition, $\lim_{\epsilon\downarrow 0} a(\epsilon)/b(\epsilon) = 0$. For any $c > 1$, there exists $\epsilon_0 > 0$ such that

$$\frac{a(\epsilon)}{c\cdot b(\epsilon)} \leq \mathbf{P}\Big(U(\epsilon) \leq \frac{1}{b(\epsilon)}\Big) \leq \frac{c\cdot a(\epsilon)}{b(\epsilon)} \quad \forall \epsilon \in (0,\epsilon_0).$$

Proof. (i) Note that $\mathbf{P}\Big(U(\epsilon) > \frac{1}{b(\epsilon)}\Big) = \Big(1 - a(\epsilon)\Big)^{\lfloor 1/b(\epsilon)\rfloor}$. By taking logarithm on both sides, we have

$$\ln \mathbf{P}\Big(U(\epsilon) > \frac{1}{b(\epsilon)}\Big) = \lfloor 1/b(\epsilon)\rfloor \ln\Big(1 - a(\epsilon)\Big) = \frac{\lfloor 1/b(\epsilon)\rfloor}{1/b(\epsilon)} \frac{\ln\Big(1 - a(\epsilon)\Big)}{-a(\epsilon)} \frac{-a(\epsilon)}{b(\epsilon)}.$$

Since $\lim_{x\to 0}\frac{\ln(1+x)}{x} = 1$, we know that for ϵ sufficiently small, we will have $-c\frac{a(\epsilon)}{b(\epsilon)} \leq \ln \mathbf{P}\Big(U(\epsilon) > \frac{1}{b(\epsilon)}\Big) \leq -\frac{a(\epsilon)}{c\cdot b(\epsilon)}$. By taking exponential on both sides, we conclude the proof.

(ii) To begin with, from the lower bound of part (i), we have

$$\mathbf{P}\Big(U(\epsilon) \leq \frac{1}{b(\epsilon)}\Big) = 1 - \mathbf{P}\Big(U(\epsilon) > \frac{1}{b(\epsilon)}\Big) \leq 1 - \exp\Big(-c\cdot\frac{a(\epsilon)}{b(\epsilon)}\Big) \leq \frac{c\cdot a(\epsilon)}{b(\epsilon)}$$

for sufficiently small $\epsilon > 0$. For the lower bound, recall that $1 - \exp(-x) \geq \frac{x}{\sqrt{c}}$ holds for $x > 0$ sufficiently close to 0. Since we assume $\lim_{\epsilon \downarrow 0} a(\epsilon)/b(\epsilon) = 0$, applying this bound with $x = \frac{a(\epsilon)}{\sqrt{c} \cdot b(\epsilon)}$ along with the upper bound of part (i), we get

$$\mathbf{P}\Big(U(\epsilon) \leq \frac{1}{b(\epsilon)}\Big) \geq 1 - \exp\Big(-\frac{1}{\sqrt{c}} \cdot \frac{a(\epsilon)}{b(\epsilon)}\Big) \geq \frac{a(\epsilon)}{c \cdot b(\epsilon)}$$

for sufficiently small ϵ. □

Lemma 3.2. *Suppose that a function $g : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous with $L < \infty$ such that $|g(x) - g(y)| \leq L|x - y|$ for all $x, y \in \mathbb{R}$. Given real numbers $x, \widetilde{x} \in \mathbb{R}$, and $\eta > 0$, and a sequence of real numbers $(z_i)_{i=1}^n$, let $\{x_k\}_{k=0,\ldots,n}$ and $\{\widetilde{x}_k\}_{k=0,\ldots,n}$ be constructed by*

$$x_0 = x, \quad x_k = x_{k-1} + \eta g(x_{k-1}) + \eta z_k \quad \textit{for} \;\; k = 1, 2, \cdots, n;$$
$$\widetilde{x}_0 = \widetilde{x}, \quad \widetilde{x}_k = \widetilde{x}_{k-1} + \eta g(\widetilde{x}_{k-1}) \quad \textit{for} \;\; k = 1, 2, \cdots, n.$$

If there exists some $\widetilde{c} \in (0, \infty)$ such that $\max_{k \leq n} \eta|z_1 + \cdots + z_k| + |x - \widetilde{x}| \leq \widetilde{c}$, then

$$\max_{k \leq n} |x_k - \widetilde{x}_k| \leq \widetilde{c} \cdot \exp(\eta L n).$$

Proof. Let $a_k \triangleq x_k - \widetilde{x}_k$ and note that $a_k = \eta \sum_{j=1}^k \big(g(\widetilde{x}_{j-1}) - g(x_{j-1})\big) + \eta(z_1 + \cdots + z_k) + x - \widetilde{x}$. Due to the Lipschitz continuity of $g(\cdot)$, this yields $|a_k| \leq \eta L(|a_0| + \cdots |a_{k-1}|) + \widetilde{c}$. It then follows from the discrete version of Gronwall's inequality (see, for example, Lemma A.3 of [33]) that $|a_k| \leq \tilde{c} \cdot \exp(\eta L k)$ for any $k = 0, 1, \cdots, n$. □

Let $\boldsymbol{x}_j^\eta(x)$ be the solution to

$$\boldsymbol{x}_0^\eta(x) = x, \qquad \boldsymbol{x}_j^\eta(x) = \boldsymbol{x}_{j-1}^\eta(x) + \eta a\big(\boldsymbol{x}_{j-1}^\eta(x)\big) \quad \forall j \geq 1. \tag{3.1}$$

After proper scaling of the time parameter, $\boldsymbol{x}_j^\eta$ approximates $\boldsymbol{y}_t$ with small η. In the next lemma, we bound the distance between $\boldsymbol{x}_{\lfloor t/\eta \rfloor}^\eta(x)$ and $\boldsymbol{y}_t(y)$.

Lemma 3.3. *Let Assumptions 2 and 4 hold. For any $\eta > 0, t > 0$ and $x, y \in \mathbb{R}$,*

$$\sup_{s \in [0,t]} |\boldsymbol{y}_s(y) - \boldsymbol{x}_{\lfloor s/\eta \rfloor}^\eta(x)| \leq (\eta C + |x - y|) \exp(Dt)$$

where $D, C \in [1, \infty)$ are the constants in Assumptions 2 and 4 respectively.

Proof. For any $s \geq 0$ that is not an integer, we write $\boldsymbol{x}_s^\eta(x) \triangleq \boldsymbol{x}_{\lfloor s \rfloor}^\eta(x)$. Also, we set $\boldsymbol{y}_s^\eta(y) \triangleq \boldsymbol{y}_{s\eta}(y)$ for any $s \geq 0$. Now observe that (for any $s \geq 0$)

$$\boldsymbol{y}_s^\eta(y) = \boldsymbol{y}_{\lfloor s \rfloor}^\eta(y) + \eta \int_{\lfloor s \rfloor}^s a(\boldsymbol{y}_u^\eta(y)) du$$
$$\boldsymbol{y}_{\lfloor s \rfloor}^\eta(y) = y + \eta \int_0^{\lfloor s \rfloor} a(\boldsymbol{y}_u^\eta(y)) du$$
$$\boldsymbol{x}_{\lfloor s \rfloor}^\eta(y) = x + \eta \int_0^{\lfloor s \rfloor} a(\boldsymbol{x}_u^\eta(y)) du.$$

Let $b(u) \triangleq \boldsymbol{y}_u^\eta(y) - \boldsymbol{x}_u^\eta(x)$. It suffices to show that $\sup_{u \in [0, t/\eta]} |b(u)| \leq (\eta C + |x - y|) \exp(Dt)$. To this end, we observe that (for any $s > 0$)

$$|b(s)| \leq |b(\lfloor s \rfloor)| + \big|\eta \int_{\lfloor s \rfloor}^s a\big(\boldsymbol{y}_u^\eta(y)\big) du\big| \leq |b(\lfloor s \rfloor)| + \eta C$$

$$\leq \eta \int_0^{\lfloor s \rfloor} \big| a\big(\boldsymbol{y}^\eta_u(y)\big) - a\big(\boldsymbol{x}^\eta_u(x)\big)\big| du + |x-y| + \eta C$$
$$\leq \eta D \int_0^s |b(u)| du + |x-y| + \eta C \qquad \text{due to Assumption 4.}$$

Apply Gronwall's inequality (see Theorem V.68 of [43]) to $b(\cdot)$ on interval $[0, t/\eta]$ and we conclude the proof. □

Our analysis hinges on the concept of the *large noises* among $(Z_j)_{j\geq 1}$, i.e., some Z_j large enough such that $\eta|Z_j|$ is larger than some prefixed threshold level $\delta > 0$. To be more concrete, for any $i \geq 1$ and $\eta, \delta > 0$, define the i^{th} arrival time of "large noises" and its size as

$$\tau^{>\delta}_i(\eta) \triangleq \min\{n > \tau^{>\delta}_{i-1}(\eta) :\ \eta|Z_n| > \delta\}, \quad \tau^{>\delta}_0(\eta) = 0 \tag{3.2}$$
$$W^{>\delta}_i(\eta) \triangleq Z_{\tau^{>\delta}_i(\eta)}. \tag{3.3}$$

For any $\delta > 0$ and $k = 1, 2, \cdots$, note that

$$\begin{aligned}\mathbf{P}\Big(\tau^{>\delta}_k(\eta) \leq \lfloor 1/\eta \rfloor\Big) &\leq \mathbf{P}\Big(\tau^{>\delta}_j(\eta) - \tau^{>\delta}_{j-1}(\eta) \leq \lfloor 1/\eta \rfloor \ \ \forall j \in [k]\Big) \\ &= \Big[\sum_{i=1}^{\lfloor 1/\eta \rfloor} \big(1 - H(\delta/\eta)\big)^{i-1} H(\delta/\eta)\Big]^k \leq \Big[\sum_{i=1}^{\lfloor 1/\eta \rfloor} H(\delta/\eta)\Big]^k \\ &\leq \Big[1/\eta \cdot H(\delta/\eta)\Big]^k. \end{aligned} \tag{3.4}$$

Recall the definition of filtration $\mathbb{F} = (\mathcal{F}_j)_{j\geq 0}$ where $\mathcal{F}_j$ is the σ-algebra generated by $Z_1, Z_2, \cdots, Z_j$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. In the next lemma, we establish a uniform asymptotic concentration bound for the weighted sum of Z_i's where the weights are adapted to the filtration $\mathbb{F}$. For any $M \in (0, \infty)$, let $\boldsymbol{\Gamma}_M$ denote the collection of families of random variables, over which we will prove the uniform asymptotics:

$$\boldsymbol{\Gamma}_M \triangleq \Big\{(V_j)_{j\geq 0} \text{ is adapted to } \mathbb{F} :\ |V_j| \leq M \ \forall j \geq 0 \text{ almost surely}\Big\}. \tag{3.5}$$

Let $\rho(t) \triangleq \exp(Dt)$ for any $t > 0$ where $D < \infty$ is the Lipschitz constant in Assumption 2.

Lemma 3.4. *Let Assumption 1 hold.*

(a) Given any $M > 0$, $N > 0$, $t > 0$, and $\epsilon > 0$, there exists $\delta_0 = \delta_0(\epsilon, M, N, t) > 0$ such that

$$\lim_{\eta \downarrow 0} \eta^{-N} \sup_{(V_i)_{i\geq 0} \in \boldsymbol{\Gamma}_M} \mathbf{P}\left(\max_{j \leq \lfloor t/\eta \rfloor \wedge \left(\tau^{>\delta}_1(\eta) - 1\right)} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i \Big| > \epsilon \right) = 0 \qquad \forall \delta \in (0, \delta_0).$$

(b) Furthermore, let Assumption 4 hold. For each i, define

$$A_i(\eta, b, \epsilon, \delta, x) \triangleq \Big\{ \max_{j \in I_i(\eta,\delta)} \eta \Big| \sum_{n = \tau^{>\delta}_{i-1}(\eta)+1}^{j} \sigma\big(X^{\eta|b}_{n-1}(x)\big) Z_n \Big| \leq \epsilon \Big\}; \tag{3.6}$$
$$I_i(\eta, \delta) \triangleq \big\{ j \in \mathbb{N} :\ \tau^{>\delta}_{i-1}(\eta) + 1 \leq j \leq \big(\tau^{>\delta}_i(\eta) - 1\big) \wedge \lfloor 1/\eta \rfloor \big\}. \tag{3.7}$$

Here we adopt the convention that (under $b = \infty$)

$$A_i(\eta, \infty, \epsilon, \delta, x) \triangleq \Big\{ \max_{j \in I_i(\eta,\delta)} \eta \Big| \sum_{n = \tau^{>\delta}_{i-1}(\eta)+1}^{j} \sigma\big(X^{\eta}_{n-1}(x)\big) Z_n \Big| \leq \epsilon \Big\}.$$

For any $k \geq 0$, $N > 0$, $\epsilon > 0$ and $b \in (0, \infty]$, there exists $\delta_0 = \delta_0(\epsilon, N) > 0$ such that

$$\lim_{\eta \downarrow 0} \eta^{-N} \sup_{x \in \mathbb{R}} \mathbf{P}\Big(\big(\bigcap_{i=1}^{k} A_i(\eta, b, \epsilon, \delta, x) \big)^c \Big) = 0 \qquad \forall \delta \in (0, \delta_0).$$

Proof. (a) Choose some β such that $\frac{1}{2 \wedge \alpha} < \beta < 1$. Let

$$Z_i^{(1)} \triangleq Z_i \mathbb{I}\Big\{ |Z_i| \leq \frac{1}{\eta^\beta} \Big\}, \quad \widehat{Z}_i^{(1)} \triangleq Z_i^{(1)} - \mathbf{E} Z_i^{(1)}, \quad Z_i^{(2)} \triangleq Z_i \mathbb{I}\Big\{ |Z_i| \in (\frac{1}{\eta^\beta}, \frac{\delta}{\eta}] \Big\} \ \ \forall i \geq 1.$$

Note that $\sum_{i=1}^j V_{i-1} Z_i = \sum_{i=1}^j V_{i-1} Z_i^{(1)} + \sum_{i=1}^j V_{i-1} Z_i^{(2)}$ on $j < \tau_1^{>\delta}(\eta)$, and hence,

$$\begin{aligned}
& \max_{j \leq \lfloor t/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1 \right)} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i \Big| \\
\leq & \max_{j \leq \lfloor t/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1 \right)} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i^{(1)} \Big| + \max_{j \leq \lfloor t/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1 \right)} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i^{(2)} \Big| \\
\leq & \max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i^{(1)} \Big| + \max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i^{(2)} \Big|. \\
\leq & \max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^j V_{i-1} \mathbf{E} Z_i^{(1)} \Big| + \max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^j V_{i-1} \widehat{Z}_i^{(1)} \Big| + \max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^j V_{i-1} Z_i^{(2)} \Big|.
\end{aligned}$$

Therefore, it suffices to show the existence of δ_0 such that for any $\delta \in (0, \delta_0)$,

$$\lim_{\eta \downarrow 0} \eta^{-N} \sup_{(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M} \mathbf{P}\Big(\max_{j \leq \lfloor t/\eta \rfloor} \eta \big| \sum_{i=1}^j V_{i-1} \mathbf{E} Z_i^{(1)} \big| > \frac{\epsilon}{3} \Big) = 0, \tag{3.8}$$

$$\lim_{\eta \downarrow 0} \eta^{-N} \sup_{(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M} \mathbf{P}\Big(\max_{j \leq \lfloor t/\eta \rfloor} \eta \big| \sum_{i=1}^j V_{i-1} \widehat{Z}_i^{(1)} \big| > \frac{\epsilon}{3} \Big) = 0, \tag{3.9}$$

$$\lim_{\eta \downarrow 0} \eta^{-N} \sup_{(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M} \mathbf{P}\Big(\max_{j \leq \lfloor t/\eta \rfloor} \eta \big| \sum_{i=1}^j V_{i-1} Z_i^{(2)} \big| > \frac{\epsilon}{3} \Big) = 0. \tag{3.10}$$

For (3.8), first recall that $\mathbf{E} Z_i = 0$, and hence,

$$\begin{aligned}
|\mathbf{E} Z_i^{(1)}| &= \big| \mathbf{E} Z_i \mathbb{I}\{ |Z_i| > 1/\eta^\beta \} \big| \leq \mathbf{E} |Z_i| \mathbb{I}\{ |Z_i| > 1/\eta^\beta \} \\
&= \mathbf{E}\Big[(|Z_i| - 1/\eta^\beta) \mathbb{I}\{ |Z_i| - 1/\eta^\beta > 0 \} \Big] + 1/\eta^\beta \cdot \mathbf{P}(|Z_i| > 1/\eta^\beta),
\end{aligned}$$

and since $(|Z_i| - 1/\eta^\beta)\mathbb{I}\{|Z_i| - 1/\eta^\beta > 0\}$ is non-negative,

$$\begin{aligned}
\mathbf{E}(|Z_i| - 1/\eta^\beta) \mathbb{I}\{ |Z_i| - 1/\eta^\beta > 0 \} &= \int_0^\infty \mathbf{P}\big((|Z_i| - 1/\eta^\beta) \mathbb{I}\{ |Z_i| - 1/\eta^\beta \} > x \big) dx \\
&= \int_0^\infty \mathbf{P}(|Z_i| - 1/\eta^\beta > x) dx = \int_{1/\eta^\beta}^\infty \mathbf{P}(|Z_1| > x) dx.
\end{aligned}$$

Recall that $H(x) = \mathbf{P}(|Z_1| > x) \in \mathcal{RV}_{-\alpha}$ as $x \to \infty$. Therefore, from Karamata's theorem,

$$|\mathbf{E} Z_i^{(1)}| \leq \int_{1/\eta^\beta}^\infty \mathbf{P}(|Z_1| > x) dx + 1/\eta^\beta \cdot \mathbf{P}(|Z_i| > 1/\eta^\beta) \in \mathcal{RV}_{(\alpha-1)\beta}(\eta) \tag{3.11}$$

as $\eta \downarrow 0$. Therefore, there exists some $\eta_0 > 0$ such that for any $(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M$ and $\eta \in (0, \eta_0)$,

$$\max_{j \leq \lfloor t/\eta \rfloor} \eta \left| \sum_{i=1}^{j} V_{i-1} \mathbf{E} Z_i^{(1)} \right| \leq tM \cdot \left| \mathbf{E} Z_i^{(1)} \right| < \epsilon/3,$$

from which we immediately get (3.8).

Next, for (3.9), fix a sufficiently large p satisfying

$$p \geq 1, \quad p > \frac{2N}{\beta}, \quad p > \frac{2N}{1-\beta}, \quad p > \frac{2N}{(\alpha-1)\beta} > \frac{2N}{(2\alpha-1)\beta}. \tag{3.12}$$

Note that for $(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M$ and $\eta > 0$, since $\{\eta V_{i-1} \widehat{Z}_i^{(i)} : i \geq 1\}$ is a martingale difference sequence,

$$\begin{aligned}
&\mathbf{E}\left[\left(\max_{j \leq \lfloor t/\eta \rfloor} \eta \Big| \sum_{i=1}^{j} V_{i-1} \widehat{Z}_i^{(1)} \Big| \right)^p \right] \\
&\leq c_1 \mathbf{E}\left[\left(\sum_{i=1}^{\lfloor t/\eta \rfloor} \left(\eta V_{i-1} \widehat{Z}_i^{(1)} \right)^2 \right)^{p/2} \right] \leq c_1 M^p \mathbf{E}\left[\left(\sum_{i=1}^{\lfloor t/\eta \rfloor} \left(\eta \widehat{Z}_i^{(1)} \right)^2 \right)^{p/2} \right] \\
&\leq c_1 c_2 M^p \mathbf{E}\left[\left(\max_{j \leq \lfloor t/\eta \rfloor} \Big| \sum_{i=1}^{j} \eta \widehat{Z}_i^{(1)} \Big| \right)^p \right] \leq c_1 c_2 \left(\frac{p}{p-1} \right)^p M^p \mathbf{E}\left[\left| \sum_{i=1}^{\lfloor t/\eta \rfloor} \eta \widehat{Z}_i^{(1)} \right|^p \right]
\end{aligned} \tag{3.13}$$

for some $c_1, c_2 > 0$ that only depend on p and won't vary with $(V_i)_{i \geq 0}$ and η. The first and third inequalities are from the uppper and lower bounds of Burkholder-Davis-Gundy inequality (Theorem 48, Chapter IV of [43]), respectively, and the fourth inequality is from Doob's maximal inequality. It then follows from Bernstein's inequality that for any $\eta > 0$ and any $s \in [0, t], y \geq 1$

$$\begin{aligned}
\mathbf{P}\Big(\Big| \sum_{j=1}^{\lfloor s/\eta \rfloor} \eta \widehat{Z}_j^{(1)} \Big|^p > \eta^{2N} y \Big) &= \mathbf{P}\Big(\Big| \sum_{j=1}^{\lfloor s/\eta \rfloor} \eta \widehat{Z}_j^{(1)} \Big| > \eta^{\frac{2N}{p}} y^{1/p} \Big) \\
&\leq 2 \exp\Big(- \frac{\frac{1}{2} \eta^{\frac{4N}{p}} \sqrt[p]{y^2}}{\frac{1}{3} \eta^{1-\beta+\frac{2N}{p}} \sqrt[p]{y} + \frac{t}{\eta} \cdot \eta^2 \cdot \mathbf{E}\big[(\widehat{Z}_1^{(1)})^2 \big]} \Big).
\end{aligned} \tag{3.14}$$

Our next goal is to show that $\frac{t}{\eta} \cdot \eta^2 \cdot \mathbf{E}\big[(\widehat{Z}_1^{(1)})^2\big] < \frac{1}{3} \eta^{1-\beta+\frac{2N}{p}}$ for any $\eta > 0$ small enough. First, due to $(a+b)^2 \leq 2a^2 + 2b^2$,

$$\mathbf{E}\big[(\widehat{Z}_1^{(1)})^2\big] = \mathbf{E}\big[\big(Z_i^{(1)} - \mathbf{E} Z_i^{(1)}\big)^2\big] \leq 2\mathbf{E}\big[\big(Z_i^{(1)}\big)^2\big] + 2\big[\mathbf{E} Z_i^{(1)}\big]^2 \leq 2\mathbf{E}\big[\big(Z_i^{(1)}\big)^2\big] + 2\big[\mathbf{E}|Z_i^{(1)}|\big]^2.$$

Also, it has been shown earlier that $\mathbf{E}|Z_i^{(1)}| \in \mathcal{RV}_{(\alpha-1)\beta}(\eta)$, and hence $\big[\mathbf{E}|Z_i^{(1)}|\big]^2 \in \mathcal{RV}_{2(\alpha-1)\beta}(\eta)$. From our choice of p in (3.12) that $p > \frac{2N}{(2\alpha-1)\beta}$, we have $1 + 2(\alpha-1)\beta > 1 - \beta + \frac{2N}{p}$, thus implying $\frac{t}{\eta} \cdot \eta^2 \cdot 2\big[\mathbf{E}|Z_i^{(1)}|\big]^2 < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$ for any $\eta > 0$ sufficiently small. Next, $\mathbf{E}\big[(Z_1^{(1)})^2\big] = \int_0^\infty 2x\mathbf{P}(|Z_1^{(1)}| > x)dx = \int_0^{1/\eta^\beta} 2x\mathbf{P}(|Z_1| > x)dx$. If $\alpha \in (1, 2]$, then Karamata's theorem implies $\int_0^{1/\eta^\beta} 2x\mathbf{P}(|Z_1| > x)dx \in \mathcal{RV}_{-(2-\alpha)\beta}(\eta)$ as $\eta \downarrow 0$. Given our choice of p in (3.12), one can see that $1-(2-\alpha)\beta > 1-\beta+\frac{2N}{p}$. As a result, for any $\eta > 0$ small enough we have $\frac{t}{\eta} \cdot \eta^2 \cdot 2\mathbf{E}\big[(Z_1^{(1)})^2\big] < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$. If $\alpha > 2$, then $\lim_{\eta \downarrow 0} \int_0^{1/\eta^\beta} 2x\mathbf{P}(|Z_1| > x)dx = \int_0^\infty 2x\mathbf{P}(|Z_1| > x)dx < \infty$. Also, (3.12) implies that $1 - \beta + \frac{2N}{p} < 1$. As a result, for any $\eta > 0$ small enough we have $\frac{t}{\eta} \cdot \eta^2 \cdot 2\mathbf{E}\big[(Z_1^{(1)})^2\big] < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$. In summary,

$$\frac{t}{\eta} \cdot \eta^2 \cdot \mathbf{E}\big[(\widehat{Z}_1^{(1)})^2\big] < \frac{1}{3}\eta^{1-\beta+\frac{2N}{p}} \tag{3.15}$$

holds for any $\eta > 0$ small enough. Along with (3.14), we yield that for any $\eta > 0$ small enough,

$$\mathbf{P}\Big(\Big|\sum_{j=1}^{\lfloor t/\eta \rfloor} \eta \widehat{Z}^{(1)}_j\Big|^p > \eta^{2N} y\Big) \leq 2\exp\Big(\frac{-\frac{1}{2}y^{1/p}}{\frac{2}{3}\eta^{1-\beta-\frac{2N}{p}}}\Big) \leq 2\exp\Big(-\frac{3}{4}y^{1/p}\Big) \qquad \forall y \geq 1,$$

where the last inequality is due to our choice of p in (3.12) that $1-\beta-\frac{2N}{p} > 0$. Moreover, since $\int_0^\infty \exp\big(-\frac{3}{4}y^{1/p}\big)dy < \infty$, one can see the existence of some $C^{(1)}_p < \infty$ such that $\mathbf{E}\Big|\sum_{j=1}^{\lfloor t/\eta \rfloor}\eta\widehat{Z}^{(1)}_j\Big|^p\Big/\eta^{2N} < C^{(1)}_p$ for all $\eta > 0$ small enough. Combining this bound, (3.13), and Markov inequality,

$$\begin{aligned}\mathbf{P}\Big(\max_{j\leq \lfloor t/\eta \rfloor}\Big|\sum_{i=1}^{j}\eta V_{i-1}\widehat{Z}^{(1)}_i\Big| > \frac{\epsilon}{3}\Big) &\leq \frac{\mathbf{E}\Big[\max_{j\leq\lfloor t/\eta\rfloor}\Big|\sum_{i=1}^{j}\eta V_{i-1}\widehat{Z}^{(1)}_i\Big|^p\Big]}{\epsilon^p/3^p} \\ &\leq \frac{c' M^p \mathbf{E}\Big|\sum_{j=1}^{\lfloor s/\eta\rfloor}\eta\widehat{Z}^{(1)}_i\Big|^p}{\epsilon^p/3^p} \leq \frac{c' M^p\cdot C^{(1)}_p}{\epsilon^p/3^p}\cdot\eta^{2N}\end{aligned}$$

for any $(V_i)_{i\geq 0} \in \mathbf{\Gamma}_M$ and all $\eta > 0$ sufficiently small. This proves (3.9).

Finally, for (3.10), recall that we have chosen β in such a way that $\alpha\beta - 1 > 0$. Fix a constant $J = \lceil \frac{N}{\alpha\beta-1}\rceil + 1$, and define $I(\eta) \triangleq \#\{i \leq \lfloor t/\eta \rfloor : Z^{(2)}_i \neq 0\}$. Besides, fix $\delta_0 = \frac{\epsilon}{3MJ}$. For any $\delta \in (0,\delta_0)$ and $(V_i)_{i\geq 0} \in \mathbf{\Gamma}_M$, note that on event $\{I(\eta) < J\}$, we must have $\max_{j\leq\lfloor t/\eta\rfloor}\eta\big|\sum_{i=1}^{j}V_{i-1}Z^{(2)}_i\big| < \eta\cdot M\cdot J\cdot\delta_0/\eta < MJ\delta_0 < \epsilon/3$. On the other hand,

$$\mathbf{P}\big(I(\eta)\geq J\big) \leq \binom{\lfloor t/\eta\rfloor}{J}\cdot\Big(H\big(1/\eta^\beta\big)\Big)^J \leq (t/\eta)^J\cdot\Big(H\big(1/\eta^\beta\big)\Big)^J \in \mathcal{RV}_{J(\alpha\beta-1)}(\eta) \text{ as } \eta\downarrow 0.$$

Lastly, the choice of $J = \lceil\frac{N}{\alpha\beta-1}\rceil + 1$ guarantees that $J(\alpha\beta-1) > N$, and hence,

$$\lim_{\eta\downarrow 0}\sup_{(V_i)_{i\geq 0}\in\mathbf{\Gamma}_M}\mathbf{P}\Big(\max_{j\leq\lfloor t/\eta\rfloor}\eta\big|\sum_{i=1}^{j}V_{i-1}Z^{(2)}_i\big| > \frac{\epsilon}{3}\Big)\Big/\eta^N \leq \lim_{\eta\downarrow 0}\sup_{(V_i)_{i\geq 0}\in\mathbf{\Gamma}_M}\mathbf{P}(I(\eta)\geq J)\Big/\eta^N = 0.$$

This concludes the proof of part (a).

(b) To ease notations, in this proof we write $X^{\eta|b} = X^\eta$ when $b = \infty$. Due to Assumption 4, it holds for any $x \in \mathbb{R}$ and any $\eta > 0, n \geq 0$ that $\sigma\big(X^{\eta|b}_n(x)\big) \leq C$. Therefore, $\{\sigma(X^{\eta|b}_i(x))\}_{i\geq 0} \in \mathbf{\Gamma}_C$. From the strong Markov property at stopping times $\big(\tau^{>\delta}_i(\eta)\big)_{j\geq 1}$,

$$\begin{aligned}\sup_{x\in\mathbb{R}}\mathbf{P}\Big(\big(\bigcap_{i=1}^{k}A_i(\eta,b,\epsilon,\delta,x)\big)^c\Big) &\leq \sum_{i=1}^{k}\sup_{x\in\mathbb{R}}\mathbf{P}\Big(\big(A_i(\eta,b,\epsilon,\delta,x)\big)^c\Big) \\ &\leq k\cdot\sup_{(V_i)_{i\geq 0}\in\mathbf{\Gamma}_C}\mathbf{P}\Big(\max_{j\leq\lfloor 1/\eta\rfloor\wedge\big(\tau^{>\delta}_1(\eta)-1\big)}\eta\big|\sum_{i=1}^{j}V_{i-1}Z_i\big| > \epsilon/2\Big)\end{aligned}$$

where $C < \infty$ is the constant in Assumption 4 and the set $\mathbf{\Gamma}_C$ is defined in (3.5). Thanks to part (a), one can find some $\delta_0 = \delta_0(\epsilon, C, N) \in (0,\bar{\delta})$ such that

$$\sup_{(V_i)_{i\geq 0}\in\mathbf{\Gamma}_C}\mathbf{P}\Big(\max_{j\leq\lfloor 1/\eta\rfloor\wedge\big(\tau^{>\delta}_1(\eta)-1\big)}\eta\big|\sum_{i=1}^{j}V_{i-1}Z_i\big| > \epsilon/2\Big) = \boldsymbol{o}(\eta^N)$$

(as $\eta\downarrow 0$) for any $\delta \in (0,\delta_0)$, which concludes the proof of part (b). $\square$

Next, for any $c > \delta > 0$, we study the law of $\big(\tau_j^{>\delta}(\eta)\big)_{j\geq 1}$ and $\big(W_j^{>\delta}(\eta)\big)_{j\geq 1}$ conditioned on event

$$E_{c,k}^{\delta}(\eta) \triangleq \Big\{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta|W_j^{>\delta}(\eta)| > c\ \ \forall j \in [k]\Big\}. \tag{3.16}$$

The intuition is that, on event $E_{c,k}^{\delta}(\eta)$, among the first $\lfloor 1/\eta \rfloor$ steps there are exactly k "large" jumps, all of which has size larger than c. Next, define random variable $W^*(c)$ with law

$$\mathbf{P}\big(W^*(c) > x\big) = p^{(+)}\Big(\frac{c}{x}\Big)^{\alpha}, \quad \mathbf{P}\big(-W^*(c) > x\big) = p^{(-)}\Big(\frac{c}{x}\Big)^{\alpha}\ \ \forall x > c, \tag{3.17}$$

and let $\big(W_j^*(c)\big)_{j\geq 1}$ be a sequence of iid copies of $W^*(c)$. Also, for $(U_j)_{j\geq 1}$, a sequence of iid copies of $\mathrm{Unif}(0,1)$ that is also independent of $\big(W_j^*(c)\big)_{j\geq 1}$, let $U_{(1;k)} \leq U_{(2;k)} \leq \cdots \leq U_{(k;k)}$ be the order statistics of $(U_j)_{j=1}^k$. For any random variable X and any Borel measureable set A, let $\mathscr{L}(X)$ be the law of X, and $\mathscr{L}(X|A)$ be the conditional law of X given event A.

Lemma 3.5. *Let Assumption 1 hold. For any $\delta > 0, c \geq \delta$ and $k \in \mathbb{Z}^+$,*

$$\lim_{\eta\downarrow 0}\frac{\mathbf{P}\big(E_{c,k}^{\delta}(\eta)\big)}{\lambda^k(\eta)} = \frac{1/c^{\alpha k}}{k!},$$

and

$$\begin{aligned}&\mathscr{L}\Big(\eta W_1^{>\delta}(\eta), \eta W_2^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta), \eta\tau_1^{>\delta}(\eta), \eta\tau_2^{>\delta}(\eta), \cdots, \eta\tau_k^{>\delta}(\eta)\Big| E_{c,k}^{\delta}(\eta)\Big)\\ \Rightarrow &\mathscr{L}\Big(W_1^*(c), W_2^*(c), \cdots, W_k^*(c), U_{(1;k)}, U_{(2;k)}, \cdots, U_{(k;k)}\Big) \text{ as } \eta \downarrow 0.\end{aligned}$$

Proof. Note that $\big(\tau_i^{>\delta}(\eta)\big)_{i\geq 1}$ is independent of $\big(W_i^{>\delta}(\eta)\big)_{i\geq 1}$. Therefore, $\mathbf{P}\big(E_{c,k}^{\delta}(\eta)\big) = \mathbf{P}\big(\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\big) \cdot \big(\mathbf{P}(\eta|W_1^{>\delta}(\eta)| > c)\big)^k$. Recall that $H(x) = \mathbf{P}(|Z_1| > x)$. Observe that

$$\begin{aligned}\mathbf{P}\Big(\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big) &= \mathbf{P}\Big(\#\big\{j \leq \lfloor 1/\eta \rfloor :\ \eta|Z_j| > \delta\big\} = k\Big)\\ &= \underbrace{\binom{\lfloor 1/\eta \rfloor}{k}}_{\triangleq q_1(\eta)}\underbrace{\Big(1 - H(\delta/\eta)\Big)^{\lfloor 1/\eta \rfloor - k}}_{\triangleq q_2(\eta)}\underbrace{\big(H(\delta/\eta)\big)^k}_{\triangleq q_3(\eta)}.\end{aligned} \tag{3.18}$$

For $q_1(\eta)$, note that

$$\lim_{\eta\downarrow 0}\frac{q_1(\eta)}{1/\eta^k} = \frac{\big(\lfloor 1/\eta \rfloor\big)\big(\lfloor 1/\eta \rfloor - 1\big)\cdots\big(\lfloor 1/\eta \rfloor - k + 1\big)/k!}{1/\eta^k} = \frac{1}{k!}. \tag{3.19}$$

Also, since $(\lfloor 1/\eta \rfloor - k) \cdot H(\delta/\eta) = \boldsymbol{o}(1)$ as $\eta \downarrow 0$, we have that $\lim_{\eta\downarrow 0} q_2(\eta) = 1$. Lastly, note that

$$\mathbf{P}\big(\eta|W_1^{>\delta}(\eta)| > c\big) = H(c/\eta)\Big/H(\delta/\eta),$$

and hence,

$$\lim_{\eta\downarrow 0}\frac{q_3(\eta)\cdot\Big(\mathbf{P}\big(\eta|W_1^{>\delta}(\eta)| > c\big)\Big)^k}{\big(H(1/\eta)\big)^k} = \lim_{\eta\downarrow 0}\frac{\big(H(\delta/\eta)\big)^k\cdot\Big(H(c/\eta)\Big/H(\delta/\eta)\Big)^k}{\big(H(1/\eta)\big)^k} = \lim_{\eta\downarrow 0}\frac{\big(H(c/\eta)\big)^k}{\big(H(1/\eta)\big)^k} = 1/c^{\alpha k} \tag{3.20}$$

Plugging (3.19) and (3.20) into (3.18), we yield

$$\lim_{\eta\downarrow 0}\frac{\mathbf{P}\big(E_{c,k}^{\delta}(\eta)\big)}{\lambda^k(\eta)} = \frac{q_1(\eta)\cdot q_2(\eta)\cdot q_3(\eta)\cdot\Big(\mathbf{P}\big(\eta|W_1^{>\delta}(\eta)| > c\big)\Big)^k}{1/\eta^k\big(H(1/\eta)\big)^k} = \frac{1/c^{\alpha k}}{k!}.$$

Next, we move onto the proof of the weak convergence. For any $x > c$,

$$\lim_{\eta \downarrow 0} \frac{\mathbf{P}\big(\eta W_1^{>\delta}(\eta) > x\big)}{\mathbf{P}\big(\eta |W_1^{>\delta}(\eta)| > c\big)} = p^{(+)}\Big(\frac{c}{x}\Big)^{\alpha}, \quad \lim_{\eta \downarrow 0} \frac{\mathbf{P}\big(\eta W_1^{>\delta}(\eta) < -x\big)}{\mathbf{P}\big(\eta |W_1^{>\delta}(\eta)| > c\big)} = p^{(-)}\Big(\frac{c}{x}\Big)^{\alpha}.$$

As a result, we must have $\mathscr{L}\Big(\eta W_1^{>\delta}(\eta), \eta W_2^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta) \Big| E^{\delta}_{c,k}(\eta)\Big) \Rightarrow \mathscr{L}\Big(W_1^*(c), \cdots, W_k^*(c)\Big)$. Moreover, notice that the sequences $\eta W_1^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta)$ and $\eta \tau_1^{>\delta}(\eta), \cdots, \eta \tau_k^{>\delta}(\eta)$ are conditionally independent on event $E^{\delta}_{c,k}(\eta)$. Indeed, for any $1 \leq i_1 < \cdots < i_k \leq \lfloor 1/\eta \rfloor$ and $c_1, \cdots, c_k > c$,

$$\begin{aligned}
&\frac{\mathbf{P}\Big(\tau_j^{>\delta}(\eta) = i_j \text{ and } \eta |W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k]\Big)}{\mathbf{P}\Big(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta |W_j^{>\delta}(\eta)| > c \ \ \forall j \in [k]\Big)} \\
&= \frac{\mathbf{P}\big(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1\big)\mathbf{P}\big(\eta |W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k]\big)}{\mathbf{P}\big(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\big)\mathbf{P}\big(\eta |W_j^{>\delta}(\eta)| > c \ \ \forall j \in [k]\big)} \\
&\text{due to the independence between } \big(\tau_i^{>\delta}(\eta)\big)_{i \geq 1} \text{ and } \big(W_i^{>\delta}(\eta)\big)_{i \geq 1} \\
&= \mathbf{P}\Big(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1 \ \Big|\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big) \cdot \mathbf{P}\Big(\eta |W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k] \ \Big|\ \eta |W_j^{>\delta}(\eta)| > c \ \ \forall j \in [k]\Big) \\
&= \mathbf{P}\Big(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1 \ \Big|\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta |W_j^{>\delta}(\eta)| > c \ \ \forall j \in [k]\Big) \\
&\quad \cdot \mathbf{P}\Big(\eta |W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k] \ \Big|\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta |W_j^{>\delta}(\eta)| > c \ \ \forall j \in [k]\Big).
\end{aligned}$$

Again, we applied the independence between $\big(\tau_i^{>\delta}(\eta)\big)_{i \geq 1}$ and $\big(W_i^{>\delta}(\eta)\big)_{i \geq 1}$. From the conditional independence between $\eta W_1^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta)$ and $\eta \tau_1^{>\delta}(\eta), \cdots, \eta \tau_k^{>\delta}(\eta)$ on event $E^{\delta}_{c,k}(\eta)$, we know that the limit of $\mathscr{L}\Big(\eta W_1^{>\delta}(\eta), \eta W_2^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta) \Big| E^{\delta}_{c,k}(\eta)\Big)$ is also independent from that of $\mathscr{L}\Big(\eta \tau_1^{>\delta}(\eta), \eta \tau_2^{>\delta}(\eta), \cdots, \eta \tau_k^{>\delta}(\eta) \Big| E^{\delta}_{c,k}(\eta)\Big)$. Therefore, it now only remains to show that

$$\mathscr{L}\Big(\eta \tau_1^{>\delta}(\eta), \eta \tau_2^{>\delta}(\eta), \cdots, \eta \tau_k^{>\delta}(\eta) \Big| E^{\delta}_{c,k}(\eta)\Big) \Rightarrow \mathscr{L}\Big(U_{(1;k)}, \cdots, U_{(k;k)}\Big).$$

Note that since both $\{\eta \tau_i^{>\delta}(\eta) : i = 1, \ldots, k\}$ and $\{U_{(i):k} : i = 1, \ldots, k\}$ are sorted in an ascending order, the joint CDFs are completely characterized by $\{t_i : i = 1, \ldots, k\}$'s such that $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \leq 1$. For any such $(t_1, \cdots, t_k) \in [0, t]^k$, note that

$$\begin{aligned}
&\mathbf{P}\Big(\eta \tau_1^{>\delta}(\eta) > t_1,\ \eta \tau_2^{>\delta}(\eta) > t_2,\ \cdots, \eta \tau_k^{>\delta}(\eta) > t_k \ \Big|\ E^{\delta}_{c,k}(\eta)\Big) \\
&= \mathbf{P}\Big(\eta \tau_1^{>\delta}(\eta) > t_1,\ \eta \tau_2^{>\delta}(\eta) > t_2,\ \cdots, \eta \tau_k^{>\delta}(\eta) > t_k \ \Big|\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big) \\
&= \frac{\mathbf{P}\Big(\eta \tau_1^{>\delta}(\eta) > t_1,\ \eta \tau_2^{>\delta}(\eta) > t_2,\ \cdots, \eta \tau_k^{>\delta}(\eta) > t_k;\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big)}{\mathbf{P}\Big(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big)}
\end{aligned}$$

and observe that

$$\begin{aligned}
&\frac{\mathbf{P}\Big(\eta \tau_1^{>\delta}(\eta) > t_1,\ \eta \tau_2^{>\delta}(\eta) > t_2,\ \cdots, \eta \tau_k^{>\delta}(\eta) > t_k;\ \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big)}{\mathbf{P}\Big(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big)} \\
&= \frac{\big|\boldsymbol{S}^{\eta}\big| \cdot q_2(\eta) q_3(\eta)}{q_1(\eta) q_2(\eta) q_3(\eta)} = \big|\boldsymbol{S}^{\eta}\big| \Big/ q_1(\eta)
\end{aligned}$$

where $\boldsymbol{S}^\eta \triangleq \Big\{(s_1,\cdots,s_k) \in \{1,2,\cdots,\lfloor 1/\eta\rfloor - 1\}^k:\ \eta s_j > t_j\ \forall j\in[k];\ s_1 < s_2 < \cdots < s_k\Big\}$. Note that

$$|\boldsymbol{S}^\eta| = \sum_{s_k=\lfloor \frac{t_k}{\eta}\rfloor+1}^{\lfloor 1/\eta\rfloor -1}\ \sum_{s_{k-1}=\lfloor \frac{t_{k-1}}{\eta}\rfloor+1}^{s_k-1}\ \sum_{s_{k-2}=\lfloor \frac{t_{k-2}}{\eta}\rfloor+1}^{s_{k-1}-1}\cdots \sum_{s_2=\lfloor \frac{t_2}{\eta}\rfloor+1}^{s_3-1}\ \sum_{s_1=\lfloor \frac{t_1}{\eta}\rfloor+1}^{s_2-1} 1.$$

Together with (3.19), we obtain

$$\begin{aligned}\lim_{\eta\downarrow 0}\big|\boldsymbol{S}^\eta\big|\Big/q_1(\eta) &= (k!)\cdot\lim_{\eta\downarrow 0}\frac{\big|\boldsymbol{S}^\eta\big|}{(1/\eta)^k} = (k!)\int_{t_k}^1\int_{t_{k-1}}^{s_k}\int_{t_{k-2}}^{s_{k-1}}\cdots\int_{t_2}^{s_3}\int_{t_1}^{s_2} ds_1 ds_2\cdots ds_k \\ &= \mathbf{P}\big(U_{(i;k)} > t_i\ \forall i\in[j]\big)\end{aligned}$$

and conclude the proof. $\square$

Recall the definitions of the sets $\mathbb{D}_A^{(k)}$ and $\mathbb{D}_A^{(k)|b}$ in (2.11) and (2.18) respectively. The next two results reveal useful properties on sets of form $\mathbb{D}_A^{(k)}$ and $\mathbb{D}_A^{(k)|b}$ when Assumptions 2 and 4 hold.

Lemma 3.6. *Let Assumptions 2 and 4 hold. Let $A \subseteq \mathbb{R}$ be compact and let $B \in \mathscr{S}_{\mathbb{D}}$. Let $k = 0,1,2,\cdots$. If B is bounded away from $\mathbb{D}_A^{(k-1)}$, then there exist $\bar\epsilon > 0$ and $\bar\delta > 0$ such that the following claims hold:*

(a) Given any $x\in A$, the condition $|w_j| > \bar\delta\ \forall j\in[k]$ must hold if $h^{(k)}(x,\boldsymbol{w},\boldsymbol{t}) \in B^{\bar\epsilon}$;

(b) $\boldsymbol{d}_{J_1}(B^{3\bar\epsilon},\mathbb{D}_A^{(k-1)}) > 0$.

Proof. The claims are trivial if A or B is an empty set. Also, the claims are trivially true if $k=0$; note that in (b) we have $\mathbb{D}_A^{(-1)} = \emptyset$. In this proof, therefore, we focus on the case where $A\neq\emptyset$, $B\neq\emptyset$, and $k\geq 1$.

Since B is bounded away from $\mathbb{D}_A^{(k-1)}$, there exists $\bar\epsilon>0$ such that $\boldsymbol{d}_{J_1}(B^{3\bar\epsilon},\mathbb{D}_A^{(k-1)})>0$ so that part (b) is satisfied. We will show that there exists a $\bar\delta$, which together with $\bar\epsilon$ satisfies (a) as well.Let $D\in[1,\infty)$ be the Lipschitz coefficient in Assumption 2. Besides, recall the constant $C\in(1,\infty)$ in Assumption 4 that satisfies $\sup_{x\in\mathbb{R}}|\sigma(x)|\leq C$. Let $\rho\triangleq\exp(D)$ and

$$\bar\delta \triangleq \frac{\bar\epsilon}{\rho C + 1}. \tag{3.21}$$

Note that $\bar\delta < \bar\epsilon$. To show that the claim (a) holds for such $\bar\epsilon$ and $\bar\delta$, we proceed with proof by contradiction. Suppose that there is some $\boldsymbol{t}=(t_1,\cdots,t_k)\in(0,1]^{k\uparrow}$, $\boldsymbol{w}=(w_1,\cdots,w_k)\in\mathbb{R}^k$, and $x_0\in A$ such that $\xi\triangleq h^{(k)}(x_0,\boldsymbol{w},\boldsymbol{t})\in B^{\bar\epsilon}$ yet $|w_j|\leq\bar\delta$ for some $j=1,2,\cdots,k$. We construct $\xi'\in\mathbb{D}_A^{(k-1)}$ such that $\boldsymbol{d}_{J_1}(\xi',\xi)<\bar\epsilon$. Let $J\triangleq\min\{j\in[k]:\ |w_j|<\bar\delta\}$. We focus on the case $J<k$, since the case $J=k$ is almost identical but only slightly simpler. Specifically, recall the definition of $h^{(0)}(\cdot)$ given below (2.7), and construct ξ' as

$$\xi'(s) \triangleq \begin{cases}\xi(s) & s\in[0,t_J)\\ h^{(0)}(\xi'(t_J-))(s-t_J) & s\in[t_J,t_{J+1})\\ \xi(s) & s\in[t_{J+1},t].\end{cases}$$

That is, ξ' is driven by the same ODE as ξ on $[t_J,t_{J+1})$, except that at the beginning of the intervals, ξ' starts from $\xi(t_J-)$ instead of $\xi(t_J)$. On the other hand, ξ' coincides with ξ outside of $[t_J,t_{J+1})$. To see how close ξ and ξ' are, note that from Assumption 4, we also have that $\big|\xi(t_J)-\xi(t_J-)\big| = \big|\sigma(\xi(t_J-))\cdot w_J\big|\leq C\bar\delta$. Then using Gronwall's inequality, we get

$$|\xi(s)-\xi'(s)| \leq \exp\big((t_{J+1}-t_J)D\big)\big|\xi(t_J)-\xi'(t_J-)\big|$$

$$\leq \rho\big|\xi(t_J) - \xi(t_J-)\big|$$
$$\leq \rho C\bar{\delta} < \bar{\epsilon}, \tag{3.22}$$

for all $s \in [t_J, t_{J+1})$. This implies that $\boldsymbol{d}_{J_1}(\xi,\xi') < \bar{\epsilon}$. However, this cannot be the case since $\xi \in B^{\bar{\epsilon}}$, $\xi' \in \mathbb{D}_A^{(k-1)}$, and we chose $\bar{\epsilon}$ such that $\boldsymbol{d}_{J_1}(B^{3\bar{\epsilon}}, \mathbb{D}_A^{(k-1)}) > 0$. This concludes the proof for the case with $J < k$. The proof for the case where $J = k$ is almost identical. The only difference is that ξ' is set to be $\xi'(s) = \xi(s)$ for all $s < t_k$, and $\xi'(s) = h^{(0)}\big(\xi'(t_k-)\big)(s - t_k)$ for all $s \in [t_k, 1]$, □

In some of the technical tools developed below, we will make use of the following uniform nondegeneracy assumption, which can be viewed as a stronger version of Assumption 3.

Assumption 8 (Uniform Nondegeneracy)**.** *There exists $c \in (0,1]$ such that $\inf_{x\in\mathbb{R}} \sigma(x) \geq c$.*

We make one observation related to Assumption 8 and the truncation operator φ_b defined in (2.14). For any $b, c > 0$, any $w \in \mathbb{R}$ and any $z \geq c$, note that for $\widetilde{w} \triangleq \varphi_{b/c}(w)$, we have $\varphi_b(z\cdot w) = \varphi_b(z\cdot\widetilde{w})$. Indeed, the claim is obviously true when $|w| \leq b/c$ (so $\widetilde{w} = w$); in case that $|w| > b/c$, we simply get $\varphi_b(z\cdot w) = \varphi_b(z\cdot\widetilde{w})$ with the value equal to b or $-b$. Combining this fact with $|\varphi_b(x) - \varphi_b(y)| \leq |x - y|\ \forall x, y \in \mathbb{R}$, we yield (for any $b, c > 0$, any $w_1, w_2 \in \mathbb{R}$, and any $z_1, z_2 \geq c$)

$$|\varphi_b(z_1\cdot w_1) - \varphi_b(z_2\cdot w_2)| \leq |z_1\widetilde{w}_1 - z_2\widetilde{w}_2| \qquad \text{where } \widetilde{w}_1 = \varphi_{b/c}(w_1),\ \widetilde{w}_2 = \varphi_{b/c}(w_2). \tag{3.23}$$

Now we are ready to develop a result for $\mathbb{D}_A^{(k)|b}$ that is analogous to Lemma 3.6.

Lemma 3.7. *Let Assumptions 2 and 4 hold. Let $A \subseteq \mathbb{R}$ be compact and let $B \in \mathscr{S}_{\mathbb{D}}$. Let $k = 0, 1, 2, \cdots$. If B is bounded away from $\mathbb{D}_A^{(k-1)|b}$, then there exist $\bar{\epsilon} > 0$ and $\bar{\delta} > 0$ such that the following claims hold:*

(a) Given any $x \in A$, the condition $|w_j| > \bar{\delta}\ \forall j \in [k]$ must hold if $h^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}}$;

(b) $\boldsymbol{d}_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b}) > 0$.

Furthermore, suppose that Assumption 8 holds, then there exist $\bar{\epsilon} > 0$ and $\bar{\delta} > 0$ such that

(c) Given any $x \in A$, the condition $|w_j| > \bar{\delta}\ \forall j \in [k]$ must hold if $h^{(k)|b+\bar{\epsilon}}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}}$,

(d) $\boldsymbol{d}_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}) > 0$.

Proof. The same arguments in Lemma 3.6 can be repeated here to identify some constants $\epsilon_0, \bar{\delta} > 0$ such that the following two claims hold:

- given any $x \in A$, the condition $|w_j| > \bar{\delta}\ \forall j \in [k]$ must hold if $h^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\epsilon_0}$;
- $\boldsymbol{d}_{J_1}(B, \mathbb{D}_A^{(k-1)|b}) > 3\epsilon_0$;

thus concluding the proof of (a),(b).

Let $\rho \triangleq \exp(D)$ with $D \in [1, \infty)$ being the Lipschitz coefficient in Assumption 2, $C \geq 1$ being the constant in Assumption 4, and $c \in (0, 1)$ being the constant in Assumption 8. We claim that

$$\xi = h^{(k)|b}\big(x, \boldsymbol{w}, \boldsymbol{t}\big),\ \xi' = h^{(k)|b+\epsilon}\big(x, \boldsymbol{w}, \boldsymbol{t}\big) \quad\Longrightarrow\quad \boldsymbol{d}_{J_1}(\xi, \xi') \leq \Big[2\rho\Big(1 + \frac{bD}{c}\Big)\Big]^k \epsilon \tag{3.24}$$

for any $\epsilon > 0$, $x \in \mathbb{R}$, $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1]^{k\uparrow}$, and $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$. Then we can pick some $\bar{\epsilon} > 0$ small enough such that $\big[2\rho\big(1 + \frac{bD}{c}\big)\big]^k \bar{\epsilon} < \epsilon_0/4$. First, for any $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1]^{k\uparrow}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$ and $x_0 \in A$ such that $h^{(k)|b+\bar{\epsilon}}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}}$, applying (3.24) we then get $h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon} + \frac{\epsilon_0}{2}} \subseteq B^{\epsilon_0}$ due to $\bar{\epsilon} < \epsilon_0/4$. Considering our choice of $\bar{\delta}$ in part (a), we must have $|w_j| > \bar{\delta}$ for all $j \in [k]$, thus concluding the proof of part (c).

Next, for part (d) we proceed with a proof by contradiction. Suppose that $\boldsymbol{d}_{J_1}(B^{\bar\epsilon}, \mathbb{D}_A^{(k-1)|b+\bar\epsilon}) = 0$. Then we can find some $\xi \in B$ and $\xi' = h^{(k)|b+\bar\epsilon}(x,\boldsymbol{w},\boldsymbol{t}) \in \mathbb{D}_A^{(k-1)|b+\bar\epsilon}$ such that $\boldsymbol{d}_{J_1}(\xi,\xi') < 2\bar\epsilon$. However, due to (3.24), it holds for $\hat\xi = h^{(k)|b}(x,\boldsymbol{w},\boldsymbol{t}) \in \mathbb{D}_A^{(k)|b}$ that $\boldsymbol{d}_{J_1}(\xi',\hat\xi) < \epsilon_0/2$, thus leading to the contradiction that $\boldsymbol{d}_{J_1}(B,\mathbb{D}_A^{(k)|b}) \leq \boldsymbol{d}_{J_1}(\xi,\hat\xi) \leq \boldsymbol{d}_{J_1}(\xi,\xi') + \boldsymbol{d}_{J_1}(\xi',\hat\xi) < 2\bar\epsilon + \frac{\epsilon_0}{2} < \epsilon_0$. This concludes the proof of part (d).

Now it only remains to prove (3.24). We fix some $x \in \mathbb{R}$, $\boldsymbol{t} = (t_1,\cdots,t_k) \in (0,1]^{k\uparrow}$, $\boldsymbol{w} = (w_1,\cdots,w_k) \in \mathbb{R}^k$. Also, let $t_0 = 0$, $t_{k+1} = 1$, $\xi = h^{(k)|b}\big(x,\boldsymbol{w},\boldsymbol{t}\big)$, $\xi' = h^{(k)|b+\epsilon}\big(x,\boldsymbol{w},\boldsymbol{t}\big)$ and $R_j \triangleq \sup_{t\in[0,t_j]}\big|\xi(t)-\xi'(t)\big|$. First of all, by definition of $h^{(k)|b}$, we get $R_1 = \big|\xi(t_1)-\xi'(t_1)\big| \leq \epsilon$. Now we proceed by induction and suppose that for some $j \in [k]$ we have $R_j \leq \big[2\rho\big(1+\frac{bD}{c}\big)\big]^{j-1}\epsilon$. On interval $t \in [t_j,t_{j+1})$, thanks to Assumption 2 we can apply Gronwall's inequality to get

$$\sup_{t\in[t_j,t_{j+1})}\big|\xi(t)-\xi'(t)\big| \leq \exp\big(D(t_{j+1}-t_j)\big)\big|\xi(t_j)-\xi'(t_j)\big| \leq \rho R_j. \tag{3.25}$$

Lastly, at $t = t_{j+1}$, if $j = k$ (so $t_{j+1} = 1$), the continuity of ξ,ξ' implies

$$|\xi(1)-\xi'(1)| = \lim_{t\to\infty}|\xi(t)-\xi'(t)| \leq \rho R_k \leq \rho\cdot\big[2\rho\big(1+\frac{bD}{c}\big)\big]^{k-1}\epsilon < \big[2\rho\big(1+\frac{bD}{c}\big)\big]^{k}\epsilon.$$

In case that $j \leq k-1$ so $t_{j+1} < 1$, the definition of $h^{(k)|b}$ implies (let $z_* \triangleq \xi(t_{j+1}-), z'_* \triangleq \xi'(t_{j+1}-)$)

$$\begin{aligned}
&\big|\xi(t_{j+1})-\xi'(t_{j+1})\big| \\
&= \big|z_* + \varphi_b\big(\sigma(z_*)w_{j+1}\big) - \big[z'_* + \varphi_{b+\epsilon}\big(\sigma(z'_*)w_{j+1}\big)\big]\big| \\
&\leq \big|z_*-z'_*\big| + \big|\varphi_b\big(\sigma(z_*)w_{j+1}\big) - \varphi_b\big(\sigma(z'_*)w_{j+1}\big)\big| + \big|\varphi_b\big(\sigma(z'_*)w_{j+1}\big) - \varphi_{b+\epsilon}\big(\sigma(z'_*)w_{j+1}\big)\big| \\
&\leq \big|z_*-z'_*\big| + \big|\varphi_b\big(\sigma(z_*)w_{j+1}\big) - \varphi_b\big(\sigma(z'_*)w_{j+1}\big)\big| + \epsilon \\
&\leq \big|z_*-z'_*\big| + |\sigma(z_*)-\sigma(z'_*)|\cdot|\varphi_{b/c}(w_{j+1})| + \epsilon \qquad \text{using (3.23)} \\
&\leq \big|z_*-z'_*\big| + D\cdot\big|z_*-z'_*\big|\cdot(b/c) + \epsilon \qquad \text{due to Lipschitz continuity of } \sigma\text{; see Assumption 2} \\
&= (1+\frac{bD}{c})|z_*-z'_*| + \epsilon \leq (1+\frac{bD}{c})\rho R_j + \epsilon \qquad \text{due to (3.25)} \\
&\leq \rho\big(1+\frac{bD}{c}\big)\cdot\big[2\rho\big(1+\frac{bD}{c}\big)\big]^{j-1}\epsilon + \epsilon \\
&\leq \big[2\rho\big(1+\frac{bD}{c}\big)\big]^{j}\epsilon.
\end{aligned}$$

The proof to (3.24) can be completed by arguing inductively for $j = 1,2,\cdots,k$. □

For any $\xi \in \mathbb{D}$, let $\|\xi\| \triangleq \sup_{t\in[0,1]}|\xi(t)|$. We present a result about the boundedness of all ξ in $\mathbb{D}_A^{(k)|b}$.

Lemma 3.8. *Let Assumptions 2 and 3 hold. Given an integer $k \geq 0$, some $-\infty < u \leq v < \infty$, and some $b > 0$, there exists $M = M(k,u,v,b) < \infty$ such that $\|\xi\| \leq M$ $\forall \xi \in \mathbb{D}_{[u,v]}^{(k)|b}$.*

Proof. Let $\xi^*(t) = \boldsymbol{y}_t(u)$. Let $N = |u-v| \vee b$ and $\rho = \exp(D) \geq 1$ where $D \in [1,\infty)$ is the Lipschitz coefficient in Assumption 2. Let $\xi = h^{(k)|b}(x,\boldsymbol{w},\boldsymbol{t})$ be an arbitrary element of $\mathbb{D}_A^{(k)|b}$ with $x \in A \subseteq [u,v]$, $\boldsymbol{w} = (w_1,\cdots,w_k) \in \mathbb{R}^k$, $\boldsymbol{t} = (t_1,\cdots,t_k) \in (0,1]^{k\uparrow}$. From Assumption 2 and Gronwall's inequality, we get $\sup_{t\in[0,t_1)}|\xi^*(t)-\xi(t)| \leq |x-u|\exp(Dt_1) \leq \rho|x-u| \leq \rho N$. Since $\xi^*(t)$ is continuous, and $|\xi(t_1)-\xi(t_1-)| \leq b$, we get $\sup_{t\in[0,t_1]}|\xi^*(t)-\xi(t)| \leq \rho N + b \leq 2\rho N$. Now proceed with induction. Adopt the convention that $t_{k+1} = 1$, and suppose that for some $j = 1,2,\cdots,k$,

$$\sup_{t\in[0,t_j]}|\xi^*(t)-\xi(t)| \leq \underbrace{(2\rho)^j N}_{\triangleq A_j}.$$

Then from Gronwall's inequality again, we get $|\xi^*(t) - \xi(t)| \leq \rho A_j$ for any $t \in [t_j, t_{j+1})$. Due to the continuity of ξ^* and the upper bound b on the jump size of ξ at t_{j+1}, we have

$$|\xi(t_{j+1}) - \xi^*(t_{j+1})| \leq \rho A_j + b \leq 2\rho A_j \leq A_{j+1}.$$

Therefore, $\sup_{t\in[0,t_{j+1}]} |\xi^*(t) - \xi(t)| \leq A_{j+1}$. By induction, we can conlude the proof with $M = A_{k+1} + \|\xi^*\| = (2\rho)^{k+1}N + \|\xi^*\|$. □

Next, we present a corollary that follows directly from the boundedness of $\mathbb{D}_A^{(k)|b}$ shown in Lemma 3.8. To facilitate the analysis, we consider the following "truncated" version of functions $a(\cdot), \sigma(\cdot)$. For any $M \geq 1$,

$$a_M(x) \triangleq \begin{cases} a(M) & \text{if } x > M, \\ a(-M) & \text{if } x < -M, \\ a(x) & \text{otherwise.} \end{cases} \qquad \sigma_M(x) \triangleq \begin{cases} \sigma(M) & \text{if } x > M, \\ \sigma(-M) & \text{if } x < -M, \\ \sigma(x) & \text{otherwise.} \end{cases} \tag{3.26}$$

Given any $a(\cdot), \sigma(\cdot)$ satisfying Assumptions 2 and 3, it is worth noticing that $a_M(\cdot), \sigma_M(\cdot)$ will satisfy Assumptions 2, 4, and 8. Similarly, recall the definition of the mapping $h^{(k)|b}$ in (2.15)-(2.17). We also consider its "truncated" counterpart by defining the mapping $h^{(k)|b}_{M\downarrow} : \mathbb{R} \times \mathbb{R}^k \times (0,1]^{k\uparrow} \to \mathbb{D}$ as follows. Given any $x_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$, $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1]^{k\uparrow}$, let $\xi = h^{(k)|b}_{M\downarrow}(x_0, \boldsymbol{w}, \boldsymbol{t})$ be the solution to

$$\xi_0 = x_0; \tag{3.27}$$

$$\frac{d\xi_t}{dt} = a_M(\xi_t) \qquad \forall t \in [0,1],\ t \neq t_1, t_2, \cdots, t_k; \tag{3.28}$$

$$\xi_t = \xi_{t-} + \varphi_b\big(\sigma_M(\xi_{t-})w_j\big) \qquad \text{if } t = t_j \text{ for some } j \in [k]. \tag{3.29}$$

Also, we let $\mathbb{D}^{(k)|b}_{A;M\downarrow} \triangleq h^{(k)|b}_{M\downarrow}\big(\mathbb{R} \times \mathbb{R}^k \times (0,1]^{k\uparrow}\big)$. One can see that the key difference between $h^{(k)|b}_{M\downarrow}$ and $h^{(k)|b}$ is that, when constructing $h^{(k)|b}_{M\downarrow}$, we use the truncated $a_M(\cdot), \sigma_M(\cdot)$ as the drift and diffusion coefficients instead of the vanilla $a(\cdot), \sigma(\cdot)$.

Corollary 3.9. *Let Assumptions 2 and 3 hold. Let $b > 0$, $k \geq 0$. Let $A \subseteq \mathbb{R}$ be compact. There exists $M_0 \in (0,\infty)$ such that for any $M \geq M_0$*

- $\sup_{t\leq 1} |\xi(t)| \leq M_0\ \forall \xi \in \mathbb{D}_A^{(k)|b} \cup \mathbb{D}^{(k)|b}_{A;M\downarrow}$;
- *For any $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1]^{k\uparrow}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$ and $x_0 \in A$,*

$$h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) = h^{(k)|b}_{M\downarrow}(x_0, \boldsymbol{w}, \boldsymbol{t}).$$

Proof. Let $-\infty < u < v < \infty$ be such that $A \subseteq [u,v]$. Given $x_0 \in A$, $\boldsymbol{w} \in \mathbb{R}^k$, and $\boldsymbol{t} \in (0,1]^{k\uparrow}$, let $\xi \triangleq h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in \mathbb{D}_A^{(k)|b} \subseteq \mathbb{D}^{(k)|b}_{[u,v]}$. Let $M_0 < \infty$ be the uniform upper bound associated with $\mathbb{D}^{(k)|b}_{[u,v]}$ in Lemma 3.8: i.e., $\sup_{t\in[0,1]} |\xi(t)| \leq M_0\ \forall \xi \in \mathbb{D}^{(k)|b}_{[u,v]}$. If $M \geq M_0$, then we must have $\xi = h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) = h^{(k)|b}_{M\downarrow}(x_0, \boldsymbol{w}, \boldsymbol{t})$ due to $\|\xi\| \leq M_0 \leq M$, and hence $\mathbb{D}^{(k)|b}_{A;M\downarrow} = \mathbb{D}_A^{(k)|b}$ This concludes the proof. □

Now we are ready to study the continuity of mappings $h^{(k)}$ defined in (2.5)-(2.7) and $h^{(k)|b}$ defined in (2.15)-(2.17).

Lemma 3.10. *Let Assumptions 2 and 3 hold. Given any $b, T > 0$ and any $k = 0, 1, 2, \cdots$, the mapping $h^{(k)|b}_{[0,T]}$ is continuous on $\mathbb{R} \times \mathbb{R}^k \times (0,T)^{k\uparrow}$.*

Proof. To ease notations we focus on the case where $T = 1$, but the proof is identical for any $T > 0$. Fix some $b > 0$ and $k = 0, 1, 2, \cdots$, some $x^* \in \mathbb{R}$, $\boldsymbol{w}^* = (w_1^*, \cdots, w_k^*) \in \mathbb{R}$ and $\boldsymbol{t}^* = (t_1^*, \cdots, t_k^*) \in (0,1)^{k\uparrow}$. Let $\xi^* = h^{(k)|b}(x^*, \boldsymbol{w}^*, \boldsymbol{t}^*)$. Also, fix some $\epsilon \in (0,1)$. It suffices to show the existence of some $\delta \in (0,1)$ such that $\boldsymbol{d}_{J_1}(\xi^*, \xi') < \epsilon$ for all $\xi' = h^{(k)|b}(x', \boldsymbol{w}', \boldsymbol{t}')$ with $x' \in \mathbb{R}$, $\boldsymbol{w}' = (w_1', \cdots, w_k') \in \mathbb{R}^k$, $\boldsymbol{t}' = (t_1', \cdots, t_k') \in (0,1)^{k\uparrow}$ satisfying

$$|x^* - x'| < \delta, \qquad |w_j' - w_j^*| \vee |t_j' - t_j| < \delta \ \forall j \in [k]. \tag{3.30}$$

In particular, by applying Corollary 3.9 onto $\mathbb{D}^{(k)|b}_{[x^*-1, x^*+1]}$, given any $M \in (0,\infty)$ large enough the claim $\|\xi^*\| + 1 < M$ and $\|\xi'\| + 1 < M$ holds for all $\xi' = h^{(k)|b}(x', \boldsymbol{w}', \boldsymbol{t}')$ satisfying (3.30). By picking an even larger M if necessary, we also ensure that $M \geq 1 + \max_{j\in[k]} |w_j^*|$. Let $a^* = a_M$, $\sigma^* = \sigma_M$ (see (3.26)). Let $C^* = \sup_{x\in[-M,M]} |a(x)| \vee \sigma(x) \vee 1$. Let $h^* = h^{(k)|b}_{M\downarrow}$, see (3.27)-(3.29). The choice of M implies that $\xi^* = h^*(x^*, \boldsymbol{w}^*, \boldsymbol{t}^*)$ and $\xi' = h^*(x', \boldsymbol{w}', \boldsymbol{t}')$.

Let $\rho \triangleq \exp(D) \geq 1$ where $D \in [1,\infty)$ is the Lipschitz coefficient in Assumption 2. We pick some $\tilde{\delta} > 0$ small enough such that

$$2\tilde{\delta} < 1 \wedge \epsilon; \qquad 2^k \rho^k (DM+1)^{k+1}(6C^* + \rho)\tilde{\delta} < \epsilon. \tag{3.31}$$

Also, by picking $\delta > 0$ small enough, it is guaranteed that (under convention $t_0^* = t_0' = 0$, $t_{k+1}^* = t_{k+1}' = 1$)

$$\delta < \widetilde{\delta} \vee 1; \qquad \max_{j\in[k]} \left| \frac{t_{j+1}^* - t_j^*}{t_{j+1}' - t_j'} - 1 \right| < \widetilde{\delta} \ \forall \boldsymbol{t}' = (t_1', \cdots, t_k') \in (0,1)^{k\uparrow}, \ \max_{j\in[k]} |t_j' - t_j^*| < \delta. \tag{3.32}$$

Now it only remains to show that, under the current the choice of δ, the bound $\boldsymbol{d}_{J_1}(\xi, \xi') < \epsilon$ follows from condition (3.30). To proceed, fix some ξ' satisfying condition (3.30). Define $\lambda : [0,1] \to [0,1]$ as

$$\lambda(u) = \begin{cases} 0 & \text{if } u = 0 \\ t_j^* + \frac{t_{j+1}^* - t_j^*}{t_{j+1}' - t_j'} \cdot (u - t_j') & \text{if } u \in (t_j', t_{j+1}'] \text{ for some } j = 0, 1, \cdots, k. \end{cases}$$

For any $u \in (0,1)$, let $j \in \{0, 1, \cdots, k\}$ be such that $u \in (t_j', t_{j+1}']$. Observe that

$$\begin{aligned} |\lambda(u) - u| &= \left| t_j^* + \frac{t_{j+1}^* - t_j^*}{t_{j+1}' - t_j'} \cdot (u - t_j') - u \right| = \left| t_j^* + \frac{t_{j+1}^* - t_j^*}{t_{j+1}' - t_j'} \cdot v - (v + t_j') \right| \quad \text{with } v \triangleq u - t_j' \\ &\leq |t_j^* - t_j'| + \left| \frac{t_{j+1}^* - t_j^*}{t_{j+1}' - t_j'} - 1 \right| \cdot v \\ &\leq \widetilde{\delta} + \widetilde{\delta} \cdot 1 < \epsilon. \end{aligned} \tag{3.33}$$

In summary, $\sup_{u\in[0,1]} |\lambda(u) - u| < \epsilon$. Moving on, we show $\sup_{u\in[0,1]} \left|\xi^*\big(\lambda(u)\big) - \xi'(u)\right| < \epsilon$. with an inductive argument. First, Assumption 2 allows us to apply Gronwall's inequality and get $\sup_{u\in(0, t_1^*\wedge t_1')} \big|\xi^*(u) - \xi'(u)\big| \leq \exp\big(D \cdot (t_1^* \wedge t_1')\big)|x^* - x'| \leq \rho\delta$. As a result, for any $u \in (0, t_1^* \wedge t_1')$,

$$\begin{aligned} \left|\xi^*\big(\lambda(u)\big) - \xi'(u)\right| &= \left|\xi^*\left(\frac{t_1^*}{t_1'} \cdot u\right) - \xi'(u)\right| \leq \left|\xi^*\left(\frac{t_1^*}{t_1'} \cdot u\right) - \xi^*(u)\right| + \left|\xi'(u) - \xi^*(u)\right| \\ &\leq \left|\xi^*\left(\frac{t_1^*}{t_1'} \cdot u\right) - \xi^*(u)\right| + \rho\delta \\ &\leq \sup_{x\in\mathbb{R}} |a^*(x)| \cdot \left|\frac{t_1^*}{t_1'} - 1\right| \cdot u + \rho\delta \qquad \text{due to } \xi^* = h^*(x^*, \boldsymbol{w}^*, \boldsymbol{t}^*) \\ &\leq C^*\widetilde{\delta} + \rho\widetilde{\delta} = (C^* + \rho)\widetilde{\delta} \qquad \text{due to (3.32)}. \end{aligned}$$

In case that $t'_1 \le t^*_1$, we already get $\sup_{u\in(0,t'_1)} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| < (4C^* + \rho)\widetilde{\delta}$. In case that $t^*_1 < t'_1$, due to $\xi' = h^*(x', \boldsymbol{w}', \boldsymbol{t}')$ for any $u \in [t^*_1, t'_1)$ as well as the properties (3.32)(3.33),

$$\big|\xi'(u) - \xi'(t^*_1)\big| \le \sup_{x\in\mathbb{R}} |a^*(x)| \cdot |u - t^*_1| < C^*\widetilde{\delta};$$
$$\big|\xi^*\big(\lambda(u)\big) - \xi^*\big(\lambda(t^*_1)\big)\big| \le \sup_{x\in\mathbb{R}} |a^*(x)| \cdot \big|\lambda(u) - \lambda(t^*_1)\big| < 2C^*\widetilde{\delta}.$$

As a result, $\sup_{u\in(0,t'_1)} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| < (4C^* + \rho)\widetilde{\delta}$. In addition, due to $|\varphi_b(x) - \varphi_b(y)| \le |x - y|$,

$$\begin{aligned}
&\Big|\xi^*\big(\lambda(t'_1)\big) - \xi'(t'_1)\Big| \\
&= \bigg|\xi^*\big(\lambda(t'_1-)\big) + \varphi_b\bigg(\sigma^*\Big(\xi^*\big(\lambda(t'_1-)\big)\Big)w^*_1\bigg) - \xi'(t'_1-) - \varphi_b\bigg(\sigma^*\Big(\xi'(t'_1-)\Big)w'_1\bigg)\bigg| \\
&\le \Big|\xi^*\big(\lambda(t'_1-)\big) - \xi'(t'_1-)\Big| + \Big|\sigma^*\Big(\xi^*\big(\lambda(t'_1-)\big)\Big)w^*_1 - \sigma^*\Big(\xi'(t'_1-)\Big)w'_1\Big| \\
&\le \Big|\xi^*\big(\lambda(t'_1-)\big) - \xi'(t'_1-)\Big| + \Big|\sigma^*\Big(\xi^*\big(\lambda(t'_1-)\big)\Big) - \sigma^*\Big(\xi'(t'_1-)\Big)\Big| \cdot |w^*_1| + \Big|\sigma^*\Big(\xi'(t'_1-)\Big)\Big| \cdot |w'_1 - w^*_1| \\
&< \Big|\xi^*\big(\lambda(t'_1-)\big) - \xi'(t'_1-)\Big| + \Big|\sigma^*\Big(\xi^*\big(\lambda(t'_1-)\big)\Big) - \sigma^*\Big(\xi'(t'_1-)\Big)\Big| \cdot M + C^*\delta \\
&\le (4C^* + \rho)\widetilde{\delta} + (4C^* + \rho)\widetilde{\delta} \cdot D \cdot M + C^*\delta \qquad \text{due to Assumption 2} \\
&= \big[(4C^* + \rho)(DM + 1) + C^*\big]\widetilde{\delta} \qquad \text{due to } \delta < \widetilde{\delta}.
\end{aligned}$$

In summary, $\sup_{u\in[0,t'_1]} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| \le \big[(4C^* + \rho)(DM + 1) + C^*\big]\widetilde{\delta} \le (DM + 1)(6C^* + \rho)\widetilde{\delta}$. Now we proceed inductively. Suppose that for some $j = 1, 2, \cdots, k$,

$$\sup_{u\in[0,t'_j]} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| \le \underbrace{2^{j-1}\rho^{j-1}(DM + 1)^j(6C^* + \rho)}_{\triangleq R_j}\widetilde{\delta}.$$

For any $v \in \big[0, (t'_{j+1} \wedge t^*_{j+1}) - t'_j\big)$,

$$\begin{aligned}
\big|\xi^*\big(\lambda(t'_j + v)\big) - \xi'(t'_j + v)\big| &\le \Big|\xi^*\big(\lambda(t'_j + v)\big) - \xi^*(t'_j + v)\Big| + \big|\xi^*(t'_j + v) - \xi'(t'_j + v)\big| \\
&\le \Big|\xi^*\big(\lambda(t'_j + v)\big) - \xi^*(t'_j + v)\Big| + \rho R_j\widetilde{\delta} \qquad \text{Using Gronwall's inequality} \\
&\le \sup_{x\in\mathbb{R}} |a^*(x)| \cdot |\lambda(t'_j + v) - (t'_j + v)| + \rho R_j\widetilde{\delta} \\
&\le 2C^*\widetilde{\delta} + \rho R_j\widetilde{\delta} \qquad \text{due to (3.33)}.
\end{aligned}$$

Again, in case that $t'_{j+1} \le t^*_{j+1}$, we already get $\sup_{u\in(0,t'_{j+1})} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| < \big(5C + \rho R_j\big)\widetilde{\delta}$. In case that $t^*_{j+1} < t'_{j+1}$, note that for any $u \in [t^*_{j+1}, t'_{j+1})$, one can apply properties (3.32)(3.33) to yield

$$\big|\xi'(u) - \xi'(t^*_{j+1})\big| \le \sup_{x\in\mathbb{R}} |a^*(x)| \cdot |u - t^*_{j+1}| < C^*\widetilde{\delta};$$
$$\big|\xi^*\big(\lambda(u)\big) - \xi^*\big(\lambda(t^*_{j+1})\big)\big| \le \sup_{x\in\mathbb{R}} |a^*(x)| \cdot \big|\lambda(u) - \lambda(t^*_{j+1})\big| < 2C^*\widetilde{\delta}.$$

In summary, we get $\sup_{u\in(0,t'_{j+1})} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| < \big(5C^* + \rho R_j\big)\widetilde{\delta}$. Lastly, in case that $j = k+1$ (so $t'_j = t'_{k+1} = t_j = t_{k+1} = 1$), we have $\Big|\xi^*(1) - \xi'(1)\Big| \le \limsup_{t\uparrow 1} \Big|\xi^*\big(\lambda(t)\big) - \xi'(t)\Big| \le \big(5C^* + \rho R_j\big)\widetilde{\delta} \le R_{j+1}\widetilde{\delta}$. In case that $j \le k$, using $|\varphi_b(x) - \varphi_b(y)| \le |x - y|$,

$$\Big|\xi^*\big(\lambda(t'_{j+1})\big) - \xi'(t'_{j+1})\Big|$$

$$
\begin{aligned}
&= \Big|\xi^*\big(\lambda(t'_{j+1}-)\big) + \varphi_b\Big(\sigma^*\Big(\xi^*\big(\lambda(t'_{j+1}-)\big)\Big)w^*_{j+1}\Big) - \xi'(t'_{j+1}-) - \varphi_b\Big(\sigma^*\Big(\xi'(t'_{j+1}-)\Big)w'_{j+1}\Big)\Big| \\
&\leq \Big|\xi^*\big(\lambda(t'_{j+1}-)\big) - \xi'(t'_{j+1}-)\Big| + \Big|\sigma^*\Big(\xi^*\big(\lambda(t'_{j+1}-)\big)\Big)w^*_{j+1} - \sigma^*\Big(\xi'(t'_{j+1}-)\Big)w'_{j+1}\Big| \\
&\leq \Big|\xi^*\big(\lambda(t'_{j+1}-)\big) - \xi'(t'_{j+1}-)\Big| + \Big|\sigma^*\Big(\xi^*\big(\lambda(t'_{j+1}-)\big)\Big) - \sigma^*\Big(\xi'(t'_{j+1}-)\Big)\Big| \cdot |w^*_{j+1}| \\
&\qquad + \Big|\sigma^*\Big(\xi'(t'_{j+1}-)\Big)\Big| \cdot |w'_{j+1} - w^*_{j+1}| \\
&< \Big|\xi^*\big(\lambda(t'_{j+1}-)\big) - \xi'(t'_{j+1}-)\Big| + \Big|\sigma^*\Big(\xi\big(\lambda(t'_{j+1}-)\big)\Big) - \sigma^*\Big(\xi'(t'_{j+1}-)\Big)\Big| \cdot M + C^*\delta \\
&\leq \big(5C^* + \rho R_j\big)\widetilde{\delta} + \big(5C^* + \rho R_j\big)\widetilde{\delta} \cdot D \cdot M + C^*\delta \qquad \text{because of Assumption 2} \\
&= \Big[\big(5C^* + \rho R_j\big)(DM+1) + C^*\Big]\widetilde{\delta} \leq \big(6C^* + \rho R_j\big)(DM+1)\widetilde{\delta} \\
&= 6C^*(DM+1)\widetilde{\delta} + \rho(DM+1)R_j\widetilde{\delta} \leq \rho(DM+1)R_j\widetilde{\delta} + \rho(DM+1)R_j\widetilde{\delta} \\
&= 2\rho(DM+1)R_j\widetilde{\delta} = 2^j\rho^j(DM+1)^{j+1}(6C^*+\rho)\widetilde{\delta} = R_{j+1}\widetilde{\delta},
\end{aligned}
$$

and hence $\sup_{u\in[0,t'_{j+1}]}\big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| \leq R_{j+1}\widetilde{\delta}$. By arguing inductively, we yield $\sup_{u\in[0,1]}\big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| \leq R_{k+1}\widetilde{\delta} < \epsilon$ due to our choice of $\widetilde{\delta}$ in (3.31). Combining this bound with (3.33), we get $\boldsymbol{d}_{J_1}(\xi^*, \xi') < \epsilon$ and conclude the proof. □

Lemma 3.11. *Let Assumptions 2, 3, and 4 hold. Given any $k = 0, 1, 2, \cdots$ and $T > 0$, the mapping $h^{(k)}_{[0,T]}$ is continuous on $\mathbb{R} \times \mathbb{R}^k \times (0,T)^{k\uparrow}$.*

Proof. To ease notations we focus on the case where $T = 1$, but the proof is identical for arbitrary $T > 0$. Fix some $k = 0, 1, 2, \cdots$, $x^* \in \mathbb{R}$, $\boldsymbol{w}^* = (w^*_1, \cdots, w^*_k) \in \mathbb{R}$ and $\boldsymbol{t}^* = (t^*_1, \cdots, t^*_k) \in (0,1)^{k\uparrow}$. We claim the existence of some $b = b(x^*, \boldsymbol{w}^*, \boldsymbol{t}^*) > 0$ such that for any $\delta \in (0,1)$, $x' \in \mathbb{R}$, $\boldsymbol{w}' \in \mathbb{R}^k$ and $\boldsymbol{t}' \in (0,1)^{k\uparrow}$ satisfying

$$
|x^* - x'| < \delta; \qquad |w'_j - w^*_j| \vee |t'_j - t^*_j| < \delta \ \forall j \in [k], \tag{3.34}
$$

we have $h^{(k)}(x', \boldsymbol{w}', \boldsymbol{t}') = h^{(k)|b}(x', \boldsymbol{w}', \boldsymbol{t}')$. Then the continuity of $h^{(k)}$ follows immediately from the continuity of $h^{(k)|b}$ established in Lemma 3.10. To find such $b > 0$, note that we can simply set $b = C \cdot \big(\max\{|w^*_j| : \ j \in [k]\} + 1\big)$ where $C \geq 1$ is the constant in Assumption 4 satisfying $\sup_{x\in\mathbb{R}}|\sigma(x)| \leq C$. Indeed, for any $\delta \in (0,1)$ and any $\delta \in (0,1)$, $x' \in \mathbb{R}$, $\boldsymbol{w}' \in \mathbb{R}^k$ and $\boldsymbol{t}' \in (0,1)^{k\uparrow}$ satisfying (3.34), for $\xi' = h^{(k)}(x', \boldsymbol{w}', \boldsymbol{t}')$ we have $|\xi'(t'_j-)w'_j| \leq C \cdot \big(\max\{|w^*_j| : \ j \in [k]\} + \delta\big) < b$ for all $j \in [k]$, thus implying $\xi' = h^{(k)|b}(x', \boldsymbol{w}', \boldsymbol{t}')$. This concludes the proof. □

As an important consequence of the previous discussion, we verify the sequential compactness condition (2.1) for measures $\mathbf{C}^{(k)}(\ \cdot\ ; x)$ and $\mathbf{C}^{(k)|b}(\ \cdot\ ; x)$ when we restrict x over a compact set A.

Lemma 3.12. *Let $T > 0$ and $k \geq 1$. Let $A \subseteq \mathbb{R}$ be compact.*

(a) Let Assumptions 2, 3, and 4 hold. For any sequence $x_n \in A$ and $x^* \in A$ such that $\lim_{n\to\infty} x_n = x^*$,

$$
\lim_{n\to\infty} \mathbf{C}^{(k)}(f; x_n) = \mathbf{C}^{(k)}(f; x^*) \qquad \forall f \in \mathcal{C}\big(\mathbb{D}[0,T] \setminus \mathbb{D}^{(k-1)}_A[0,T]\big).
$$

(b) Let Assumptions 2 and 3 hold. Let $b > 0$. For any sequence $x_n \in A$ and $x^* \in A$ such that $\lim_{n\to\infty} x_n = x^*$,

$$
\lim_{n\to\infty} \mathbf{C}^{(k)|b}(f; x_n) = \mathbf{C}^{(k)|b}(f; x^*) \qquad \forall f \in \mathcal{C}\big(\mathbb{D}[0,T] \setminus \mathbb{D}^{(k-1)|b}_A[0,T]\big).
$$

Proof. For convenience we consider the case $T = 1$, but the proof can easily extend for arbitrary $T > 0$.

(a) Pick some $f \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)})$. and let $\phi(x) = \phi_f(x) \triangleq \mathbf{C}^{(k)}(f; x)$. We argue that $\phi(x)$ is a continuous function using Dominated Convergence theorem. First, from the continuity of f and $h^{(k)}$ (see Lemma 3.11), for any sequence $y_m \in \mathbb{R}$ with $\lim_{m\to\infty} y_m = y \in \mathbb{R}$, we have

$$\lim_{m\to\infty} f\Big(h^{(k)}(y_m, \boldsymbol{w}, \boldsymbol{t})\Big) = f\Big(h^{(k)}(y, \boldsymbol{w}, \boldsymbol{t})\Big) \qquad \forall \boldsymbol{w} \in \mathbb{R}^k,\ \boldsymbol{t} \in (0,1)^{k\uparrow}.$$

Next, we apply Lemma 3.6 onto $B = \text{supp}(f)$, which is bounded away from $\mathbb{D}_A^{(k-1)}$, and find $\bar{\delta} > 0$ such that $h^{(k)}(x, \boldsymbol{w}, \boldsymbol{t}) \in B \Longrightarrow |w_j| > \bar{\delta}\ \forall j \in [k]$. As a result, $\big|f\big(h^{(k)}(x, \boldsymbol{w}, \boldsymbol{t})\big)\big| \leq \|f\| \cdot \mathbb{I}\big(|w_j| > \bar{\delta}\ \forall j \in [k]\big)$. Also, note that $\int \mathbb{I}\big(|w_j| > \bar{\delta}\ \forall j \in [k]\big)\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_1^{k\uparrow}(d\boldsymbol{t}) \leq 1/\bar{\delta}^{k\alpha} < \infty$. This allows us to apply Dominated Convergence theorem and establish the continuity of $\phi(x)$. This implies

$$\lim_{n\to\infty} \mathbf{C}^{(k)}(f; x_n) = \lim_{n\to\infty} \phi(x_n) = \phi(x^*) = \mathbf{C}^{(k)}(f; x^*).$$

Due to the arbitrariness of $f \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)})$ we conclude the proof of part (a).

(b) The proof is almost identical. The only differences are that we apply Lemma 3.10 (resp. Lemma 3.7) instead of Lemma 3.11 (resp. Lemma 3.6) so we omit the details. ☐

In the next lemma, we show that the image of $h^{(1)}$ (resp. $h^{(1)|b}$) provides good approximations of the sample path of X_j^η (resp. $X_j^{\eta|b}$) up until $\tau_1^{>\delta}(\eta)$, i.e. the arrival time of the first "large noise"; see (3.2),(3.3) for the definition of $\tau_i^{>\delta}(\eta), W_i^{>\delta}(\eta)$.

Lemma 3.13. *Let Assumptions 2 and 4 hold. Let $D, C \in [1, \infty)$ be the constants in Assumptions 2 and 4 respectively and let $\rho \triangleq \exp(D)$.*

(a) For any $\epsilon, \delta, \eta > 0$ and any $x, y \in \mathbb{R}$, it holds on the event

$$\Big\{ \max_{i \leq \lfloor 1/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1\right)} \eta \Big| \sum_{j=1}^{i} \sigma\big(X_{j-1}^\eta(x)\big) Z_j \Big| \leq \epsilon \Big\}$$

that

$$\sup_{t \in [0,1]:\ t < \eta\tau_1^{>\delta}(\eta)} \big|\xi_t - X_{\lfloor t/\eta \rfloor}^\eta(x)\big| \leq \rho \cdot \big(\epsilon + |x - y| + \eta C\big), \tag{3.35}$$

where

$$\xi = \begin{cases} h^{(1)}\big(y, \eta W_1^{>\delta}(\eta), \eta\tau_1^{>\delta}(\eta)\big) & \text{if } \eta\tau_1^{>\delta}(\eta) \leq 1, \\ h^{(0)}(y) & \text{if } \eta\tau_1^{>\delta}(\eta) > 1. \end{cases}$$

(b) Furthermore, suppose that Assumption 8 holds. For any $\epsilon, b > 0$, any $\delta \in (0, \frac{b}{2C})$, $\eta \in (0, \frac{b \wedge 1}{2C})$, and any $x, y \in \mathbb{R}$, it holds on event

$$\Big\{ \max_{i \leq \lfloor 1/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1\right)} \eta \Big| \sum_{j=1}^{i} \sigma\big(X_{j-1}^{\eta|b}(x)\big) Z_j \Big| \leq \epsilon \Big\}$$

that

$$\sup_{t \in [0,1]:\ t < \eta\tau_1^{>\delta}(\eta)} \big|\xi_t - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x)\big| \leq \rho \cdot \big(\epsilon + |x - y| + \eta C\big), \tag{3.36}$$

$$\sup_{t\in[0,1]:\ t\leq \eta\tau_1^{>\delta}(\eta)} \left|\xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\right| \leq \rho\cdot\Big(1+\frac{bD}{c}\Big)\big(\epsilon + |x-y| + 2\eta C\big) \tag{3.37}$$

where

$$\xi = \begin{cases} h^{(1)|b}\big(y, \eta W_1^{>\delta}(\eta), \eta\tau_1^{>\delta}(\eta)\big) & \text{if } \eta\tau_1^{>\delta}(\eta) \leq 1, \\ h^{(0)|b}(y) & \text{if } \eta\tau_1^{>\delta}(\eta) > 1. \end{cases}$$

Proof. (a) By definition of ξ, we have $\xi_t = \boldsymbol{y}_t(y) = h^{(0)}(y)(t)$ for any $t \in [0,1]$ with $t < \eta\tau_1^{>\delta}(\eta)$. Also, since $\tau_1^{>\delta}(\eta)$ only takes values in $\{1,2,\cdots\}$, we know that $\eta\tau_1^{>\delta}(\eta) \leq 1 \iff \tau_1^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor$ and $\eta\tau_1^{>\delta}(\eta) > 1 \iff \tau_1^{>\delta}(\eta) > \lfloor 1/\eta \rfloor$.

Let $A \triangleq \Big\{ \max_{i\leq \lfloor 1/\eta\rfloor \wedge (\tau_1^{>\delta}(\eta)-1)} \eta\Big|\sum_{j=1}^{i}\sigma\big(X^\eta_{j-1}(x)\big)Z_j\Big| \leq \epsilon\Big\}$. Recall the definition of the deterministic process $\boldsymbol{x}^\eta$ defined in (3.1). Applying Lemma 3.2, we know that on event A,

$$\Big|\boldsymbol{x}^\eta_j(x) - X^\eta_j(x)\Big| \leq \epsilon\cdot\exp(\eta D\cdot\lfloor 1/\eta\rfloor) \leq \rho\epsilon \qquad \forall j \leq \lfloor 1/\eta \rfloor \wedge \big(\tau_1^{>\delta}(\eta) - 1\big). \tag{3.38}$$

On the other hand, recall that $\boldsymbol{y}_t(y)$ is the solution to ODE $d\boldsymbol{y}_t(y)/dt = a\big(\boldsymbol{y}_t(y)\big)$ under initial condition $\boldsymbol{y}_0(y) = y$. Since $\xi_t = \boldsymbol{y}_t(y)$ on $t < \eta\tau_1^{>\delta}(\eta)$, by applying Lemma 3.3 we get

$$\sup_{t\in[0,1]:\ t<\eta\tau_1^{>\delta}(\eta)} \Big|\xi_t - \boldsymbol{x}^\eta_{\lfloor t/\eta\rfloor}(x)\Big| \leq \big(\eta C + |x-y|\big)\cdot\rho. \tag{3.39}$$

Therefore,

$$\sup_{t\in[0,1]:t<\eta\tau_1^{>\delta}(\eta)} \Big|\xi_t - X^\eta_{\lfloor t/\eta\rfloor}(x)\Big| \leq \rho\cdot\Big(\epsilon + |x-y| + \eta C\Big). \tag{3.40}$$

(b) Note that for any $x\in\mathbb{R}$ and any $t\in[0,1]$ with $t<\eta\tau_1^{>\delta}(\eta)$,

$$h^{(0)|b}(x)(t) = h^{(0)}(x)(t) = h^{(1)|b}\big(x,\eta W_1^{>\delta}(\eta),\eta\tau_1^{>\delta}(\eta)\big)(t) = h^{(1)}\big(x,\eta W_1^{>\delta}(\eta),\eta\tau_1^{>\delta}(\eta)\big)(t) = \boldsymbol{y}_t(x).$$

Also, for any w with $|w| \leq \delta < \frac{b}{2C}$, note that $\varphi_b\Big(\eta a(x) + \sigma(x)w\Big) = \eta a(x) + \sigma(x)w\ \forall x\in\mathbb{R}$ due to $\eta\sup_{x\in\mathbb{R}}|a(x)| \leq \eta C < \frac{b}{2}$ and $\sup_{x\in\mathbb{R}}\sigma(x)|w| \leq C|w| < b/2$ (recall our choice of $\eta C < \frac{b}{2}\wedge 1$). As a result, $X^\eta_j(x) = X^{\eta|b}_j(x)$ for all $x\in\mathbb{R}$ and $j < \tau_1^{>\delta}(\eta)$. It then follows directly from (3.40) that $\sup_{t\in[0,1]:t<\eta\tau_1^{>\delta}(\eta)}\big|\xi_t - X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\big| \leq \rho\cdot\big(\epsilon+|x-y|+\eta C\big)$. A direct consequence is (we write $\boldsymbol{y}(u;y) = \boldsymbol{y}_u(y)$, $\boldsymbol{y}(s-;y) = \lim_{u\uparrow s}\boldsymbol{y}_u(y)$, and $\xi(t) = \xi_t$ in this proof)

$$\Big|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big| \leq \rho\cdot\big(\epsilon+|x-y|+\eta C\big). \tag{3.41}$$

Therefore,

$$\begin{aligned}
&\Big|\xi\big(\eta\tau_1^{>\delta}(\eta)\big) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)}(x)\Big| \\
&= \Bigg|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) + \varphi_b\Big(\eta\sigma\Big(\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)\Big)W_1^{>\delta}(\eta)\Big) \\
&\qquad - \Big[X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x) + \varphi_b\Big(\eta a\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big) + \eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\Big)\Big]\Bigg| \\
&\leq \Big|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big|
\end{aligned}$$

$$+\underbrace{\left|\varphi_b\Big(\eta\sigma\Big(\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)\Big)W_1^{>\delta}(\eta)\Big)-\varphi_b\Big(\eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\Big)\right|}_{\triangleq I_1}$$

$$+\underbrace{\left|\varphi_b\Big(\eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\Big)-\varphi_b\Big(\eta a\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)+\eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\Big)\right|}_{\triangleq I_2}.$$

Based on observation (3.23), we get

$$\begin{aligned}I_1 &\leq \left|\varphi_{b/c}\big(\eta W_1^{>\delta}(\eta)\big)\right|\cdot\left|\sigma\Big(\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)\Big)-\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)\right| \\ &\leq \frac{b}{c}\cdot D\cdot\left|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)-X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\right|\leq\frac{bD}{c}\cdot\rho\cdot\big(\epsilon+|x-y|+\eta C\big)\end{aligned}$$

using Assumption 2 and the upper bound (3.41). On the other hand, from $|\varphi_b(x)-\varphi_b(y)|\leq|x-y|$ we get $I_2\leq\left|\eta a\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)\right|\leq\eta C$. In summary,

$$\begin{aligned}\sup_{t\in[0,1]:\ t\leq\eta\tau_1^{>\delta}(\eta)}\left|\xi_t-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right| &\leq \Big(1+\frac{bD}{c}\Big)\cdot\rho\cdot\big(\epsilon+|x-y|+\eta C\big)+\eta C \\ &\leq \Big(1+\frac{bD}{c}\Big)\cdot\rho\cdot\big(\epsilon+|x-y|+2\eta C\big).\end{aligned}$$

This concludes the proof of part (b). □

By applying Lemma 3.13 inductively, the next result illustrates how the image of the mapping $h^{(k)|b}$ approximates the path of $X^{\eta|b}_j(x)$.

Lemma 3.14. *Let Assumptions 2, 4, and 8 hold. Let $A_i(\eta,b,\epsilon,\delta,x)$ be defined as in (3.6). For any $k\geq 0$, $x\in\mathbb{R}$, $\epsilon,b>0$, $\delta\in(0,\frac{b}{2C})$, and $\eta\in(0,\frac{b\wedge\epsilon}{2C})$, it holds on event $\big(\bigcap_{i=1}^{k+1}A_i(\eta,b,\epsilon,\delta,x)\big)\cap\Big\{\tau_k^{>\delta}(\eta)\leq\lfloor 1/\eta\rfloor<\tau_{k+1}^{>\delta}(\eta)\Big\}$ that*

$$\sup_{t\in[0,1]}\left|\xi(t)-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right|<\Big[3\rho\cdot(1+\frac{bD}{c})\Big]^k\cdot 3\rho\epsilon.$$

where $\xi\triangleq h^{(k)|b}\big(x,\eta W_1^{>\delta}(\eta),\cdots,\eta W_k^{>\delta}(\eta),\eta\tau_1^{>\delta}(\eta),\cdots,\eta\tau_k^{>\delta}(\eta)\big)$, $\rho=\exp(D)\geq 1$, $D\in[1,\infty)$ is the Lipschitz coefficient in Assumption 2, $C\geq 1$ is the constant in Assumption 4, and $c\in(0,1)$ is the constant in Assumption 8.

Proof. First of all, on $A_1(\eta,b,\epsilon,\delta,x)$, one can apply (3.36) of Lemma 3.13 and obtain

$$\sup_{t\in[0,1]:\ t<\eta\tau_1^{>\delta}(\eta)}\left|\xi_t-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right|=\sup_{t\in[0,1]:\ t<\eta\tau_1^{>\delta}(\eta)}\left|\boldsymbol{y}_t(x)-X^{\eta}_{\lfloor t/\eta\rfloor}(x)\right|\leq\rho\cdot(\epsilon+\eta C)<2\rho\epsilon,$$

where we applied our choice of $\eta C<\epsilon/2$. In case that $k=0$, we can already conclude the proof. Henceforth in the proof, we focus on the case where $k\geq 1$. Now we can instead apply (3.37) of Lemma 3.13 to get

$$\sup_{t\in[0,\eta\tau_1^{>\delta}(\eta)]}\left|\xi_t-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right|\leq\rho\cdot\Big(1+\frac{bD}{c}\Big)\big(\epsilon+2\eta C\big)\leq 3\rho\cdot\Big(1+\frac{bD}{c}\Big)\epsilon$$

due to our choice of $2\eta C<\epsilon$. To proceed with an inductive argument, suppose that for some $j=1,2,\cdots,k-1$ we can show that

$$\sup_{t\in[0,1\wedge\eta\tau_j^{>\delta}(\eta)]}\left|\xi_t-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right|\leq\underbrace{\Big[3\rho\cdot(1+\frac{bD}{c})\Big]^j\epsilon}_{\triangleq R_j}.$$

To highlight the timestamp in the ODE $\boldsymbol{y}_t(y)$ we write $\boldsymbol{y}(t;y) = \boldsymbol{y}_t(y)$ in this proof. Note that for any $t \in \big[\eta\tau_j^{>\delta}(\eta), \eta\tau_{j+1}^{>\delta}(\eta)\big)$, we have $\xi_t = \boldsymbol{y}\Big(t - \eta\tau_j^{>\delta}(\eta); \xi_{\eta\tau_j^{>\delta}(\eta)}\Big)$. Therefore, by applying (3.37) of Lemma 3.13 again, we obtain

$$\begin{aligned}
\sup_{t \in \left[\eta\tau_j^{>\delta}(\eta), \eta\tau_{j+1}^{>\delta}(\eta)\right]} \left|\xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\right| &\leq \rho \cdot \Big(1 + \frac{bD}{c}\Big) \cdot (\epsilon + R_j + 2\eta C) \\
&\leq \rho \cdot \Big(1 + \frac{bD}{c}\Big) \cdot (2\epsilon + R_j) \qquad \text{due to } 2\eta C < \epsilon \\
&\leq 3\rho \cdot \Big(1 + \frac{bD}{c}\Big) R_j = R_{j+1} \qquad \text{due to } R_j > \epsilon.
\end{aligned}$$

Arguing inductively, we yield $\sup_{t \in [0, \eta\tau_k^{>\delta}(\eta)]} \left|\xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\right| \leq R_k = \big[3\rho \cdot (1 + \frac{bD}{c})\big]^k \epsilon$. Lastly, due to (3.35) of Lemma 3.13 and the fact that $\eta\tau_{k+1}^{>\delta}(\eta) > 1$,

$$\begin{aligned}
\sup_{t \in [\eta\tau_k^{>\delta}(\eta), 1]} \left|\xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\right| &\leq \rho \cdot (\epsilon + R_k + \eta C) \leq \rho \cdot (2\epsilon + R_k) \\
&\leq \rho \cdot 3R_k < \Big[3\rho \cdot (1 + \frac{bD}{c})\Big]^k \cdot 3\rho\epsilon
\end{aligned}$$

This concludes the proof. □

3.2 Proof of Theorem 2.1

We first provide the proof of Theorem 2.1, i.e., the Portmanteau theorem for the uniform $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence.

Proof of Theorem 2.1. We first prove $(i) \Rightarrow (ii)$ and proceed with a proof by contradiction. Suppose that the claim $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu^\eta_\theta(F) - \mu_\theta(F^\epsilon) \leq 0$ does not hold for some closed F bounded away from $\mathbb{C}$ and some $\epsilon > 0$. Then there exists some sequences $\eta_n \downarrow 0$ and $\theta_n \in \Theta$ and some $\delta > 0$ such that $\mu^{\eta_n}_{\theta_n}(F) - \mu_{\theta_n}(F^\epsilon) > \delta \ \forall n \geq 1$. Now, we make two observations. First, using Urysohn's lemma (see, e.g., lemma 2.3 of [34]), one can identify some $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ such that $\mathbb{I}_F \leq f \leq \mathbb{I}_{F^\epsilon}$, which leads to the bound $\mu^{\eta_n}_{\theta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq \mu^{\eta_n}_{\theta_n}(f) - \mu_{\theta_n}(f)$ for each n. Second, from statement (i) we get $\lim_{n\to\infty} \left|\mu^{\eta_n}_{\theta_n}(f) - \mu_{\theta_n}(f)\right| = 0$. In summary, we yield the contradiction

$$\begin{aligned}
\limsup_{n\to\infty} \mu^{\eta_n}_{\theta_n}(F) - \mu_{\theta_n}(F^\epsilon) &\leq \limsup_{n\to\infty} \mu^{\eta_n}_{\theta_n}(f) - \mu_{\theta_n}(f) \\
&\leq \lim_{n\to\infty} \left|\mu^{\eta_n}_{\theta_n}(f) - \mu_{\theta_n}(f)\right| = 0.
\end{aligned}$$

The case where claim $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu^\eta_\theta(G) - \mu_\theta(G^\epsilon) \geq 0$ does not hold for some open G bounded away from $\mathbb{C}$ and some $\epsilon > 0$ can be addressed analogously by applying Urysohn's lemma and constructing some $g \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ such that $\mathbb{I}_{G_\epsilon} \leq g \leq \mathbb{I}_G$. This concludes the proof of $(i) \Rightarrow (ii)$.

Next, we prove $(ii) \Rightarrow (i)$. Again, we consider a proof by contradiction. Suppose that the claim $\lim_{\eta \downarrow 0} \sup_{\theta \in \Theta} \left|\mu^\eta_\theta(g) - \mu_\theta(g)\right| = 0$ does not hold for some $g \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$. Then there exist some sequences $\eta_n \downarrow 0$, $\theta_n \in \Theta$ and some $\delta > 0$ such that

$$|\mu^{\eta_n}_{\theta_n}(g) - \mu_{\theta_n}(g)| > \delta \qquad \forall n \geq 1. \tag{3.42}$$

To proceed, we arbitrarily pick some closed $F \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$, and some open $G \subseteq \mathbb{S}$ that is bounded away from $\mathbb{C}$, and then make two observations. First, using claims in (ii), we have $\limsup_{n\to\infty} \mu^{\eta_n}_{\theta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq 0$ and $\liminf_{n\to\infty} \mu^{\eta_n}_{\theta_n}(G) - \mu_{\theta_n}(G_\epsilon) \geq 0$ for any $\epsilon > 0$. Next, due to the assumption (2.1), by picking a sub-sequence of θ_n if necessary we can find some μ_{θ^*} such that $\lim_{n\to\infty} \left|\mu_{\theta_n}(f) - \mu_{\theta^*}(f)\right| = 0$ for all $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$. By Portmanteau theorem for standard

$\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence (see theorem 2.1 of [34]), we yield $\limsup_{n\to\infty}\mu_{\theta_n}(F^\epsilon)\leq\mu_{\theta^*}(F^\epsilon)$ for the closed set F^ϵ and $\liminf_{n\to\infty}\mu_{\theta_n}(G_\epsilon)\geq\mu_{\theta^*}(G_\epsilon)$ for the open set G_ϵ. In summary, for any $\epsilon>0$,

$$\limsup_{n\to\infty}\mu^{\eta_n}_{\theta_n}(F)\leq\limsup_{n\to\infty}\mu_{\theta_n}(F^\epsilon)+\limsup_{n\to\infty}\mu^{\eta_n}_{\theta_n}(F)-\mu_{\theta_n}(F^\epsilon)\leq\mu_{\theta^*}(F^\epsilon),$$
$$\liminf_{n\to\infty}\mu^{\eta_n}_{\theta_n}(G)\geq\liminf_{n\to\infty}\mu_{\theta_n}(G_\epsilon)+\liminf_{n\to\infty}\mu^{\eta_n}_{\theta_n}(G)-\mu_{\theta_n}(G_\epsilon)\geq\mu_{\theta^*}(G_\epsilon).$$

Lastly, note that $\lim_{\epsilon\downarrow 0}\mu_{\theta^*}(F^\epsilon)=\mu_{\theta^*}(F)$ and $\lim_{\epsilon\downarrow 0}\mu_{\theta^*}(G_\epsilon)=\mu_{\theta^*}(G)$ due to continuity of measures and $\bigcap_{\epsilon>0}F^\epsilon=F$, $\bigcup_{\epsilon>0}G_\epsilon=G$. This allows us to apply Portmanteau theorem for standard $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence again and obtain that $\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(g)-\mu_{\theta^*}(g)\big|=0$ for the $g\in\mathcal{C}(\mathbb{S}\setminus\mathbb{C})$ fixed in (3.42). However, recall that we have already obtained $\lim_{n\to\infty}\big|\mu_{\theta_n}(g)-\mu_{\theta^*}(g)\big|=0$ using assumption (2.1). We hereby arrive at the contradiction

$$\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(g)-\mu_{\theta_n}(g)\big|\leq\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(g)-\mu_{\theta^*}(g)\big|+\lim_{n\to\infty}\big|\mu_{\theta^*}(g)-\mu_{\theta_n}(g)\big|=0$$

and conclude the proof of $(ii)\Rightarrow(i)$.

Due to the equivalence of (i) and (ii), it only remains to show that $(i)\Rightarrow(iii)$. Again, we consider a proof by contradiction. Suppose that the claim $\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F)\leq\sup_{\theta\in\Theta}\mu_\theta(F)$ in (iii) does not hold for some closed F bounded away from $\mathbb{C}$. Then we can find sequences $\eta_n\downarrow 0$, $\theta_n\in\Theta$ and some $\delta>0$ such that $\mu^{\eta_n}_{\theta_n}(F)>\sup_{\theta\in\Theta}\mu_\theta(F)+\delta\ \forall n\geq 1$. Next, due to the assumption (2.1), by picking a sub-sequence of θ_n if necessary we can find some μ_{θ^*} such that $\lim_{n\to\infty}\big|\mu_{\theta_n}(f)-\mu_{\theta^*}(f)\big|=0$ for all $f\in\mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. Meanwhile, (i) implies that $\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(f)-\mu_{\theta_n}(f)\big|=0$ for all $f\in\mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. Therefore,

$$\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(f)-\mu_{\theta^*}(f)\big|\leq\lim_{n\to\infty}\big|\mu^{\eta_n}_{\theta_n}(f)-\mu_{\theta_n}(f)\big|+\lim_{n\to\infty}\big|\mu_{\theta_n}(f)-\mu_{\theta^*}(f)\big|=0$$

for all $f\in\mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. By Portmanteau theorem for standard $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence, we yield the contradiction $\limsup_{n\to\infty}\mu^{\eta_n}_{\theta_n}(F)\leq\mu_{\theta^*}(F)\leq\sup_{\theta\in\Theta}\mu_\theta(F)$. In summary, we have established the claim $\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F)\leq\sup_{\theta\in\Theta}\mu_\theta(F)$ for all closed F bounded away from $\mathbb{C}$. The same approach can also be applied to show $\liminf_{\eta\downarrow 0}\inf_{\theta\in\Theta}\mu^\eta_\theta(G)\geq\inf_{\theta\in\Theta}\mu_\theta(G)$ for all open G bounded away from $\mathbb{C}$. This concludes the proof. □

To facilitate the application of Theorem 2.1, we introduce the concept of asymptotic equivalence between two families of random objects. Specifically, we consider a generalized version of asymptotic equivalence over $\mathbb{S}\setminus\mathbb{C}$, which is equivalent to definition 2.9 in [12].

Definition 3.1. *Let X_n and Y_n be random elements taking values in a complete separable metric space $(\mathbb{S},\boldsymbol{d})$. Let ϵ_n be a sequence of positive real numbers. Let $\mathbb{C}\subseteq\mathbb{S}$ be Borel measurable. X_n is said to be* ***asymptotically equivalent to** Y_n **in** $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ **with respect to** ϵ_n if for any $\Delta>0$ and any $B\in\mathscr{S}_\mathbb{S}$ bounded away from $\mathbb{C}$,*

$$\lim_{n\to\infty}\epsilon_n^{-1}\mathbf{P}\Big(\boldsymbol{d}\big(X_n,Y_n\big)\mathbb{I}\big(X_n\in B\ \text{or}\ Y_n\in B\big)>\Delta\Big)=0.$$

In case that $\mathbb{C}=\emptyset$, Definition 3.1 simply degenerates to the standard notion of asymptotic equivalence; see definition 1 of [45]. The following lemma demonstrates the application of the asymptotic equivalence and is plays an important role in our analysis below.

Lemma 3.15 (Lemma 2.11 of [12])**.** *Let X_n and Y_n be random elements taking values in a complete separable metric space $(\mathbb{S},\boldsymbol{d})$ and let $\mathbb{C}\subseteq\mathbb{S}$ be Borel measurable. Suppose that $\epsilon_n^{-1}\mathbf{P}(X_n\in\cdot)\to\mu(\cdot)$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ for some sequence of positive real numbers ϵ_n. If X_n is asymptotically equivalent to Y_n when bounded away from $\mathbb{C}$ with respect to ϵ_n, then $\epsilon_n^{-1}\mathbf{P}(Y_n\in\cdot)\to\mu(\cdot)$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$.*

3.3 Proof of Theorems 2.2 and 2.3

In the proofs of Theorems 2.2 and 2.3 below, without loss of generality we focus on the case where $T = 1$. But we note that the proof for the cases with arbitrary $T > 0$ is identical.

Recall the notion of uniform $\mathbb{M}$-convergence introduced in Definition 2.1. At first glance, the uniform version of $\mathbb{M}$-convergence stated in Theorem 2.2 and 2.3 is stronger than the standard $\mathbb{M}$-convergence introduced in [34]. Nevertheless, under the conditions provided in Theorem 2.2 or 2.3 regarding the initial conditions of $\boldsymbol{X}^\eta$ or $\boldsymbol{X}^{\eta|b}$, we can show that it suffices to prove the standard notion of $\mathbb{M}$-convergence. In particular, the proofs to Theorem 2.2 and 2.3 hinge on the following key result for $\boldsymbol{X}^{\eta|b}$.

Proposition 3.16. *Let η_n be a sequence of strictly positive real numbers with $\lim_{n\to\infty}\eta_n = 0$. Let compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ be such that $\lim_{n\to\infty} x_n = x^*$. Under Assumptions 1, 2, and 3, it holds for any $k = 0, 1, 2, \cdots$ and $b > 0$ that*

$$\mathbf{P}\big(\boldsymbol{X}^{\eta_n|b}(x_n) \in \ \cdot\ \big)/\lambda^k(\eta_n) \to \mathbf{C}^{(k)|b}\big(\ \cdot\ ; x^*\big) \text{ in } \mathbb{M}\big(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)|b}\big) \text{ as } n\to\infty.$$

As the first application of Proposition 3.16, we prepare a similar result for the unclipped dynamics $\boldsymbol{X}^\eta$ defined in (2.10).

Proposition 3.17. *Let η_n be a sequence of strictly positive real numbers with $\lim_{n\to\infty}\eta_n = 0$. Let compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ be such that $\lim_{n\to\infty} x_n = x^*$. Under Assumptions 1, 2, 3, and 4, it holds for any $k = 0, 1, 2, \cdots$ that*

$$\mathbf{P}\big(\boldsymbol{X}^{\eta_n}(x_n) \in \ \cdot\ \big)/\lambda^k(\eta_n) \to \mathbf{C}^{(k)}\big(\ \cdot\ ; x^*\big) \text{ in } \mathbb{M}\big(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)}\big) \text{ as } n\to\infty.$$

Proof. Fix some $k = 0, 1, 2, \cdots$ and some $g \in \mathcal{C}\big(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)}\big)$. By virtue of Portmanteau theorem for $\mathbb{M}$-convergence (see theorem 2.1 of [34]), it suffices to show that

$$\lim_{n\to\infty} \mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n}(x_n)\big)\big]/\lambda^k(\eta_n) = \mathbf{C}^{(k)}(g; x^*).$$

To this end, we first set $B \triangleq \text{supp}(g)$ and observe that for any $n \geq 1$ and any $\delta, b > 0$,

$$\begin{aligned}
&\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n}(x_n)\big)\big] \\
&= \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\mathbb{I}\big(\boldsymbol{X}^{\eta_n}(x_n) \in B\big)\big] \\
&= \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\mathbb{I}\big(\tau^{>\delta}_{k+1}(\eta_n) < \lfloor 1/\eta_n \rfloor;\ \boldsymbol{X}^{\eta_n}(x_n) \in B\big)\big] \\
&+ \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\mathbb{I}\big(\tau^{>\delta}_{k}(\eta_n) > \lfloor 1/\eta_n \rfloor;\ \boldsymbol{X}^{\eta_n}(x_n) \in B\big)\big] \\
&+ \mathbf{E}\Big[g(\boldsymbol{X}^{\eta_n}(x_n))\mathbb{I}\Big(\tau^{>\delta}_{k}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau^{>\delta}_{k+1}(\eta_n);\ \eta_n|W^{>\delta}_j(\eta_n)| > \frac{b}{2C} \text{ for some } j \in [k];\ \boldsymbol{X}^{\eta_n}(x_n) \in B\Big)\Big] \\
&+ \underbrace{\mathbf{E}\Big[g(\boldsymbol{X}^{\eta_n}(x_n))\mathbb{I}\Big(\tau^{>\delta}_{k}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau^{>\delta}_{k+1}(\eta_n);\ \eta_n|W^{>\delta}_j(\eta_n)| \leq \frac{b}{2C}\ \forall j \in [k];\ \boldsymbol{X}^{\eta_n}(x_n) \in B\Big)\Big]}_{\triangleq I_*(n,b,\delta)}
\end{aligned}$$

where $C \geq 1$ is the constant in Assumption 4 such that $|a(x)| \vee \sigma(x) \leq C$ for any $x \in \mathbb{R}$. Now we focus on term $I_*(n,b,\delta)$ and let

$$\widetilde{A}(n,b,\delta) \triangleq \Big\{\tau^{>\delta}_{k}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau^{>\delta}_{k+1}(\eta_n);\ \eta_n|W^{>\delta}_j(\eta_n)| \leq \frac{b}{2C}\ \forall j \in [k];\ \boldsymbol{X}^{\eta_n}(x_n) \in B\Big\}.$$

For any n large enough, we have $\eta_n \cdot \sup_{x\in\mathbb{R}}|a(x)| \leq \eta_n C \leq b/2$. As a result,for such n and any $\delta \in (0, \frac{b}{2C})$, on event $\widetilde{A}(n,b,\delta)$ the step-size (before truncation) $\eta a\big(X^{\eta|b}_{j-1}(x)\big) + \eta\sigma\big(X^{\eta|b}_{j-1}(x)\big)Z_j$ of

$X_j^{\eta|b}$ is less than b for each $j \leq \lfloor 1/\eta_n \rfloor$, and hence $\boldsymbol{X}^{\eta_n}(x_n) = \boldsymbol{X}^{\eta_n|b}(x_n)$. This observation leads to the following upper bound: Given any $b > 0$ and $\delta \in (0, \frac{b}{2C})$, it holds for any n large enough that

$$\begin{aligned}
\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\big] \leq \|g\| & \underbrace{\mathbf{P}\big(\tau_{k+1}^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor\big)}_{\triangleq p_1(n,\delta)} \\
& + \|g\| \underbrace{\mathbf{P}\big(\tau_k^{>\delta}(\eta_n) > \lfloor 1/\eta_n \rfloor;\ \boldsymbol{X}^{\eta_n}(x_n) \in B\big)}_{\triangleq p_2(n,\delta)} \\
& + \|g\| \underbrace{\mathbf{P}\Big(\tau_k^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau_{k+1}^{>\delta}(\eta_n);\ \eta_n |W_j^{>\delta}(\eta_n)| > \frac{b}{2C} \text{ for some } j \in [k]\Big)}_{\triangleq p_3(n,b,\delta)} \\
& + \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\big].
\end{aligned}$$

Meanwhile, given any n large enough, any $b > 0$ and any $\delta \in (0, \frac{b}{2C})$, we obtain the lower bound

$$\begin{aligned}
\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\big] & \geq \mathbf{E}[I_*(n,b,\delta)] \\
& = \mathbf{E}\Big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\mathbb{I}\Big(\widetilde{A}(n,b,\delta)\Big)\Big] \quad \text{due to } \boldsymbol{X}^{\eta_n}(x_n) = \boldsymbol{X}^{\eta_n|b}(x_n) \text{ on } \widetilde{A}(n,b,\delta) \\
& \geq \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\big] - \|g\| \, \mathbf{P}\Big(\big(\widetilde{A}(n,b,\delta)\big)^c\Big) \\
& \geq \mathbf{E}\big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\big] - \|g\| \cdot \big[p_1(n,\delta) + p_2(n,\delta) + p_3(n,b,\delta)\big].
\end{aligned}$$

Suppose we can find some $\delta > 0$ satisfying

$$\lim_{n\to\infty} p_1(n,\delta)/\lambda^k(\eta_n) = 0, \tag{3.43}$$

$$\lim_{n\to\infty} p_2(n,\delta)/\lambda^k(\eta_n) = 0. \tag{3.44}$$

Fix such δ. Furthermore, we claim that for any $b > 0$,

$$\limsup_{n\to\infty} p_3(n,b,\delta)/\lambda^k(\eta_n) \leq \psi_\delta(b) \triangleq \frac{k}{\delta^{\alpha k}} \cdot \big(\frac{\delta}{2C}\big)^\alpha \cdot \frac{1}{b^\alpha}. \tag{3.45}$$

Note that $\lim_{b\to\infty} \psi_\delta(b) = 0$. Lastly, we claim that

$$\lim_{b\to\infty} \mathbf{C}^{(k)|b}(g;x^*) = \mathbf{C}^{(k)}(g;x^*). \tag{3.46}$$

Then by combining (3.43)–(3.44) with the upper and lower bounds for $\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\big]$ established earlier, we see that for any b large enough (such that $\frac{b}{2C} > \delta$),

$$\begin{aligned}
& \lim_{n\to\infty} \frac{\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\big]}{\lambda^k(\eta_n)} - \|g\|\,\psi_\delta(b) \leq \lim_{n\to\infty} \frac{\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\big]}{\lambda^k(\eta_n)} \leq \lim_{n\to\infty} \frac{\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n|b}(x_n))\big]}{\lambda^k(\eta_n)} + \|g\|\,\psi_\delta(b), \\
\Longrightarrow & - \|g\|\,\psi_\delta(b) + \mathbf{C}^{(k)|b}(g;x^*) \leq \lim_{n\to\infty} \frac{\mathbf{E}\big[g(\boldsymbol{X}^{\eta_n}(x_n))\big]}{\lambda^k(\eta_n)} \leq \|g\|\,\psi_\delta(b) + \mathbf{C}^{(k)|b}(g;x^*).
\end{aligned}$$

In the last line of the display, we applied Proposition 3.16. Letting b tend to ∞ and applying the limit (3.46), we conclude the proof. Now it only remains to prove (3.43) (3.44) (3.45) (3.46).

Proof of Claim (3.43):

Applying (3.4), we see that $p_1(n,\delta) \leq \big(H(\frac{\delta}{\eta_n})/\eta_n\big)^{k+1}$ holds for any $\delta > 0$ and any $n \geq 1$. Due to the regularly varying nature of $H(\cdot)$, we then yield $\limsup_{n\to\infty} \frac{p_1(n,\delta)}{\lambda^{k+1}(\eta_n)} \leq 1/\delta^{\alpha(k+1)} < \infty$. To show that claim (3.43) holds for any $\delta > 0$ we only need to note that

$$\limsup_{n\to\infty} \frac{p_1(n,\delta)}{\lambda^k(\eta_n)} \leq \limsup_{n\to\infty} \frac{p_1(n,\delta)}{\lambda^{k+1}(\eta_n)} \cdot \lim_{n\to\infty} \lambda(\eta_n) \leq \frac{1}{\delta^{\alpha(k+1)}} \cdot \lim_{n\to\infty} \frac{H(1/\eta_n)}{\eta_n} = 0$$

due to $\frac{H(1/\eta)}{\eta} = \lambda(\eta) \in \mathcal{RV}_{\alpha-1}(\eta)$ as $\eta \downarrow 0$ and $\alpha > 1$.

Proof of Claim (3.44):

We claim the existence of some $\epsilon > 0$ such that

$$\Big\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor;\ \boldsymbol{X}^\eta(x) \in B\Big\} \cap \big(\bigcap_{i=1}^{k+1} A_i(\eta,\infty,\epsilon,\delta,x)\big) = \emptyset \qquad \forall x \in A,\ \delta > 0,\ \eta \in (0, \frac{\epsilon}{C\rho}) \tag{3.47}$$

where $D, C \in [1,\infty)$ are the constants in Assumptions 2 and 4 respectively, $\rho \triangleq \exp(D)$, and event $A_i(\eta,b,\epsilon,\delta,x)$ is defined in (3.6). Then for any $\delta > 0$, we yield

$$\limsup_{n\to\infty} p_2(n,\delta)\big/\lambda^k(\eta_n) \leq \limsup_{n\to\infty}\sup_{x\in A}\mathbf{P}\Big(\big(\bigcap_{i=1}^{k+1} A_i(\eta_n,\infty,\epsilon,\delta,x)\big)^c\Big)\Big/\lambda^k(\eta_n).$$

Applying Lemma 3.4 (b) with some $N > k(\alpha-1)$, we conclude that claim (3.44) holds for all $\delta > 0$ small enough. Now it only remains to find $\epsilon > 0$ that satisfies condition (3.47). To this end, we first note that the set $B = \text{supp}(g)$ is bounded away from $\mathbb{D}_A^{(k-1)}$. By applying Lemma 3.6 one can find $\bar\epsilon > 0$ such that $\boldsymbol{d}_{J_1}(B^{\bar\epsilon}, \mathbb{D}_A^{(k-1)}) > \bar\epsilon$. Now we show that (3.47) holds for any $\epsilon > 0$ small enough with $(\rho+1)\epsilon < \bar\epsilon$. To see why, we fix such ϵ as well as some $x \in A$, $\delta > 0$ and $\eta \in (0, \frac{\epsilon}{C\rho})$. Next, define process $\breve{\boldsymbol{X}}^{\eta,\delta}(x) \triangleq \big\{\breve X_t^{\eta,\delta}(x):\ t \in [0,1]\big\}$ as the solution to (under initial condition $\breve X_0^{\eta,\delta}(x) = x$)

$$\begin{aligned}
\frac{d\breve X_t^{\eta,\delta}(x)}{dt} &= a\big(\breve X_t^{\eta,\delta}(x)\big) \qquad \forall t \geq 0,\ t \notin \{\eta\tau_j^{>\delta}(\eta):\ j \geq 1\},\\
\breve X^{\eta,\delta}_{\eta\tau_i^{>\delta}(\eta)}(x) &= X^\eta_{\tau_i^{>\delta}(\eta)}(x) \qquad \forall j \geq 1.
\end{aligned}$$

On event $\big(\cap_{i=1}^{k+1} A_i(\eta,\infty,\epsilon,\delta,x)\big) \cap \{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$, observe that

$$\begin{aligned}
&\boldsymbol{d}_{J_1}\big(\breve{\boldsymbol{X}}^{\eta,\delta}(x), \boldsymbol{X}^\eta(x)\big)\\
&\leq \sup_{t \in \big[0,\eta\tau_1^{>\delta}(\eta)\big)\cup\big[\eta\tau_1^{>\delta}(\eta),\eta\tau_2^{>\delta}(\eta)\big)\cup\cdots\cup\big[\eta\tau_k^{>\delta}(\eta),\eta\tau_{k+1}^{>\delta}(\eta)\big)} \Big|\breve X_t^{\eta,\delta}(x) - X^\eta_{\lfloor t/\eta \rfloor}(x)\Big|\\
&\leq \rho\cdot\big(\epsilon + \eta C\big) \leq \rho\epsilon + \epsilon < \bar\epsilon \qquad \text{because of (3.35) of Lemma 3.13.}
\end{aligned}$$

In the last line of the display above, we applied $\eta < \frac{\epsilon}{C\rho}$ and our choice of $(\rho+1)\epsilon < \bar\epsilon$. However, on $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ we have $\breve{\boldsymbol{X}}^{\eta,\delta}(x) \in \mathbb{D}_A^{(k-1)}$. As a result, on event $\big(\cap_{i=1}^{k+1} A_i(\eta,\infty,\epsilon,\delta,x)\big) \cap \{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ we must have $\boldsymbol{d}_{J_1}\big(\mathbb{D}_A^{(k-1)}, \boldsymbol{X}^\eta(x)\big) < \bar\epsilon$, and hence $\boldsymbol{X}^\eta(x) \notin B$ due to the fact that $\boldsymbol{d}_{J_1}(B^{\bar\epsilon}, \mathbb{D}_A^{(k-1)}) > \bar\epsilon$. This establishes (3.47).

Proof of Claim (3.45):

Due to the independence between $\big(\tau_i^{>\delta}(\eta) - \tau_{j-1}^\eta(\delta)\big)_{j\geq 1}$ and $\big(W_i^{>\delta}(\eta)\big)_{j\geq 1}$,

$$\begin{aligned}
p_3(n,b,\delta) &= \mathbf{P}\Big(\tau_k^{>\delta}(\eta_n) < \lfloor 1/\eta_n \rfloor < \tau_{k+1}^{>\delta}(\eta_n)\Big)\mathbf{P}\Big(\eta_n|W_j^{>\delta}(\eta_n)| > \frac{b}{2C} \text{ for some } j \in [k]\Big)\\
&\leq \mathbf{P}\Big(\tau_k^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor\Big)\cdot\sum_{j=1}^k \mathbf{P}\Big(\eta_n|W_j^{>\delta}(\eta_n)| > \frac{b}{2C}\Big)\\
&\leq \Big(\frac{H(\delta/\eta_n)}{\eta_n}\Big)^k\cdot k\cdot\frac{H\big(\frac{b}{2C}\cdot\frac{1}{\eta_n}\big)}{H\big(\delta\cdot\frac{1}{\eta_n}\big)}.
\end{aligned}$$

Due to $H(x) \in \mathcal{RV}_{-\alpha}(x)$ as $x \to \infty$, we conclude that $\limsup_{n\to\infty}\frac{p_4(n,b,\delta)}{\lambda^k(\eta_n)} \leq \frac{k}{\delta^{\alpha k}}\cdot\big(\frac{\delta}{2C}\big)^\alpha\cdot\frac{1}{b^\alpha} = \psi_\delta(b)$.

Proof of Claim (3.46):

The proof relies on the following claim: for any $S \in \mathscr{S}_{\mathbb{D}}$ that is bounded away from $\mathbb{D}_A^{(k-1)}$,

$$\lim_{b\to\infty} \mathbf{C}^{(k)|b}(S;x^*) = \mathbf{C}^{(k)}(S;x^*). \tag{3.48}$$

Then for $g \in \mathcal{C}(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)})$ fixed at the beginning of the proof, we know that $B = \text{supp}(g)$ is bounded away from $\mathbb{D}_A^{(k-1)}$. Also, for an arbitrarily selected $\Delta > 0$, an approximation to g using simple functions implies the existence of some $N \in \mathbb{N}$, some sequence of real numbers $\big(c_g^{(i)}\big)_{i=1}^N$, some sequence $\big(B_g^{(i)}\big)_{i=1}^N$ of Borel measurable sets on $\mathbb{D}$ that are bounded away from $\mathbb{D}_A^{(k-1)}$ such that the following claims hold for $g^\Delta(\cdot) \triangleq \sum_{i=1}^N c_g^{(i)}\mathbb{I}\big(\ \cdot\ \in B_g^{(i)}\big)$:

$$B_g^{(i)} \subseteq B \ \ \forall i \in [N]; \qquad \big|g^\Delta(\xi) - g(\xi)\big| < \Delta \ \ \forall \xi \in \mathbb{D}.$$

Now observe that

$$\begin{aligned}\limsup_{b\to\infty}\Big|\mathbf{C}^{(k)|b}(g;x^*) - \mathbf{C}^{(k)}(g;x^*)\Big| &\leq \limsup_{b\to\infty}\Big|\mathbf{C}^{(k)|b}(g;x^*) - \mathbf{C}^{(k)|b}(g^\Delta;x^*)\Big| \\ &\quad + \limsup_{b\to\infty}\Big|\mathbf{C}^{(k)|b}(g^\Delta;x^*) - \mathbf{C}^{(k)}(g^\Delta;x^*)\Big| \\ &\quad + \limsup_{b\to\infty}\Big|\mathbf{C}^{(k)}(g^\Delta;x^*) - \mathbf{C}^{(k)}(g;x^*)\Big|\end{aligned}$$

First, note that $\mathbf{C}^{(k)|b}(g^\Delta;x^*) = \sum_{i=1}^N c_g^{(i)}\mathbf{C}^{(k)|b}(B_g^{(i)};x^*)$ and $\mathbf{C}^{(k)}(g^\Delta;x^*) = \sum_{i=1}^N c_g^{(i)}\mathbf{C}^{(k)}(B_g^{(i)};x^*)$. Therefore, applying (3.48), we get $\limsup_{b\to\infty}\Big|\mathbf{C}^{(k)|b}(g^\Delta;x^*) - \mathbf{C}^{(k)}(g^\Delta;x^*)\Big| = 0$. Next, note that $\Big|\mathbf{C}^{(k)|b}(g^\Delta;x^*) - \mathbf{C}^{(k)|b}(g;x^*)\Big| \leq \Delta\cdot\mathbf{C}^{(k)|b}(B;x^*)$ and $\Big|\mathbf{C}^{(k)}(g^\Delta;x^*) - \mathbf{C}^{(k)}(g;x^*)\Big| \leq \Delta\cdot\mathbf{C}^{(k)}(B;x^*)$. Thanks to (3.48) again, we get $\limsup_{b\to\infty}\Big|\mathbf{C}^{(k)|b}(g;x^*) - \mathbf{C}^{(k)}(g;x^*)\Big| \leq 2\Delta\cdot\mathbf{C}^{(k)}(B;x^*)$. The arbitrariness of $\Delta > 0$ allows us to conclude the proof of (3.45).

We prove (3.48) by applying Dominated Convergence theorem. From the definition in (2.19),

$$\mathbf{C}^{(k)|b}(S;x^*) \triangleq \int \mathbb{I}\Big\{h^{(k)|b}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big) \in S\Big\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_1^{k\uparrow}(d\boldsymbol{t})$$

where $S \in \mathscr{S}_{\mathbb{D}}$ is bounded away from $\mathbb{D}_A^{(k-1)}$. First, for any $\boldsymbol{w} \in \mathbb{R}^k$, $\boldsymbol{t} \in (0,1)^{k\uparrow}$ and $x_0 \in \mathbb{R}$, let $M \triangleq \max_{j\in[k]}|w_j|$. For any $b > MC$ where $C \geq 1$ is the constant satisfying such that $\sup_{x\in\mathbb{R}}|a(x)| \vee \sigma(x) \leq C$ (see Assumption 4), by comparing the definition of $h^{(k)}$ and $h^{(k)|b}$ it is easy to see that $h^{(k)|b}(x^*,\boldsymbol{w},\boldsymbol{t}) = h^{(k)}(x^*,\boldsymbol{w},\boldsymbol{t})$. This implies $\lim_{b\to\infty}\mathbb{I}\big\{h^{(k)|b}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big) \in S\big\} = \mathbb{I}\big\{h^{(k)}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big) \in S\big\}$ for all $\boldsymbol{w} \in \mathbb{R}^k$ and $\boldsymbol{t} \in (0,1)^{k\uparrow}$. In order to apply Dominated Convergence theorem and conclude the proof of (3.48), it suffices to find an integrable function that dominates $\mathbb{I}\big\{h^{(k)|b}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big) \in S\big\}$. Specifically, since S is bounded away from $\mathbb{D}_A^{(k-1)}$, we can find some $\bar\epsilon > 0$ such that $\boldsymbol{d}_{J_1}(S,\mathbb{D}_A^{(k-1)}) > \bar\epsilon$. Also, let $\rho = \exp(D)$ where $D \in [1,\infty)$ is the Lipschitz coefficient in Assumption 2. Fix some $\bar\delta < \frac{\bar\epsilon}{\rho C}$. We claim that

$$\mathbb{I}\Big\{h^{(k)|b}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big) \in S\Big\} \leq \mathbb{I}\Big\{|w_j| > \bar\delta\ \forall j \in [k]\Big\} \qquad \forall b > 0,\ \boldsymbol{w} \in \mathbb{R}^k,\ \boldsymbol{t} \in (0,1)^{k\uparrow}. \tag{3.49}$$

From $\int \mathbb{I}\{|w_j| > \bar\delta\ \forall j \in [k]\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_1^{k\uparrow}(d\boldsymbol{t}) \leq 1/\bar\delta^{k\alpha} < \infty$ we conclude the proof. Now it only remains to prove (3.49). Fix some $\boldsymbol{w} = (w_1,\cdots,w_k) \in \mathbb{R}^k$, $\boldsymbol{t} = (t_1,\cdots,t_k) \in (0,1)^{k\uparrow}$, and $b > 0$. Let $\xi_b = h^{(k)|b}(x^*,\boldsymbol{w},\boldsymbol{t})$. Suppose there is some $J \in [k]$ such that $|w_J| \leq \bar\delta$. It suffices to show that $\xi_b \notin S$. To this end, define $\xi \in \mathbb{D}$ as (recall that $\boldsymbol{y}_\cdot(x)$ is the ODE defined in (2.32))

$$\xi(s) \triangleq \begin{cases}\xi_b(s) & s \in [0,t_J) \\ \boldsymbol{y}_{s-t_J}(\xi(t_J-)) & s \in [t_J,t_{J+1}) \\ \xi_b(s) & s \in [t_{J+1},t].\end{cases}$$

Note that $\xi \in \mathbb{D}_A^{(k-1)}$ and $|\xi(t_J) - \xi_b(t_J)| = |\Delta\xi_b(t_J)| = \big|\sigma\big(\xi_b(t_J-)\big)\cdot w_J\big|$. Applying Gronwall's inequality, we then yield that for all $s \in [t_J, t_{J-1})$,

$$\begin{aligned}
|\xi_b(s) - \xi(s)| &\leq \exp\big(D(s-t_J)\big)\cdot\big|\sigma\big(\xi_b(t_J-)\big)\cdot w_J\big| \\
&\leq \rho\cdot\big|\sigma\big(\xi_b(t_J-)\big)\cdot w_J\big| \qquad \text{where } \rho = \exp(D) \\
&\leq \rho C|w_J| \qquad \text{due to } \sup_{x\in\mathbb{R}}|\sigma(x)| \leq C \text{ , see Assumption 4} \\
&\leq \rho C\bar{\delta} < \bar{\epsilon} \qquad \text{due to our choice of } \bar{\delta} < \frac{\bar{\epsilon}}{\rho C},
\end{aligned}$$

which implies $\boldsymbol{d}_{J_1}(\xi,\xi_b) < \bar{\epsilon}$. However, due to $\xi \in \mathbb{D}_A^{(k-1)}$ and $\boldsymbol{d}_{J_1}(S, \mathbb{D}_A^{(k-1)}) > \bar{\epsilon}$, we must have $\xi_b \notin S$. This concludes the proof of (3.49). □

With Proposition 3.17 in our arsenal, we prove Theorem 2.2.

Proof of Theorem 2.2. For simplicity of notations we focus on the case where $T = 1$, but the proof below can be easily generalized for arbitrary $T > 0$.

We first prove the uniform $\mathbb{M}$-convergence. Specifically, we proceed with a proof by contradiction. Fix some $k = 0, 1, \cdots$ and suppose that there is some $f \in \mathcal{C}(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)})$, some sequence $\eta_n > 0$ with limit $\lim_{n\to\infty}\eta_n = 0$, some sequence $x_n \in A$, and $\epsilon > 0$ such that $\big|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x_n)\big| > \epsilon\ \forall n \geq 1$ where $\mu_n^{(k)}(\cdot) \triangleq \mathbf{P}\big(\boldsymbol{X}^{\eta_n}(x_n) \in \ \cdot\ \big)/\lambda^k(\eta_n)$. Since A is compact, by picking a proper subsequence we can assume w.l.o.g. that $\lim_{n\to\infty} x_n = x^*$ for some $x^* \in A$. This allows us to apply Proposition 3.17 and yield $\lim_{n\to\infty}\big|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x^*)\big| = 0$. On the other hand, using part (a) of Lemma 3.12, we get $\lim_{n\to\infty}\big|\mathbf{C}^{(k)}(f;x_n) - \mathbf{C}^{(k)}(f;x^*)\big| = 0$. Therefore, we arrive at the contradiction

$$\lim_{n\to\infty}\big|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x_n)\big| \leq \lim_{n\to\infty}\big|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x^*)\big| + \lim_{n\to\infty}\big|\mathbf{C}^{(k)}(f;x^*) - \mathbf{C}^{(k)}(f;x_n)\big| = 0$$

and conclude the proof of the uniform $\mathbb{M}$-convergence claim.

Next, we prove the uniform sample-path large deviations stated in (2.12). Part (a) of Lemma 3.12 verifies the compactness condition (2.1) for measures $\mathbf{C}^{(k)}(\ \cdot\ ;x)$ with $x \in A$. In light of the Portmanteau theorem for uniform $\mathbb{M}$-convergence (i.e., Theorem 2.1), most claims follow directly from Theorem 2.2 and it only remains to verify that $\sup_{x\in A}\mathbf{C}^{(k)}\big(B^-;x\big) < \infty$.

Note that B^- is bounded away from $\mathbb{D}_A^{(k-1)}$. This allows us to apply Lemma 3.6 and find $\bar{\epsilon} > 0$ and $\bar{\delta} > 0$ such that

- Given any $x \in A$, $h^{(k)}(x,\boldsymbol{w},\boldsymbol{t}) \in B^{\bar{\epsilon}} \Longrightarrow |w_j| > \bar{\delta}\ \forall j \in [k]$,
- $B^{\bar{\epsilon}} \cap \mathbb{D}_A^{(k-1)} = \emptyset$.

Then by the definition of $\mathbf{C}^{(k)|b}$ in (2.9),

$$\begin{aligned}
\sup_{x\in A}\mathbf{C}^{(k)}(B^-;x) &= \sup_{x\in A}\int \mathbb{I}\Big\{h^{(k)}\big(x,\boldsymbol{w},\boldsymbol{t}\big) \in B^- \cap \mathbb{D}_A^{(k)|b}\Big\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_1^{k\uparrow}(d\boldsymbol{t}) \\
&\leq \int \mathbb{I}\Big\{|w_j| > \bar{\delta}\ \forall j \in [k]\Big\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_1^{k\uparrow}(d\boldsymbol{t}) \leq 1/\bar{\delta}^{k\alpha} < \infty.
\end{aligned}$$

This concludes the proof. □

Similarly, building upon Proposition 3.16, we provide the proof to Theorem 2.3.

Proof of Theorem 2.3. The proof-by-contradiction approach in Theorem 2.2 can be applied here to establish the uniform $\mathbb{M}$-convergence. The only difference is that we apply Proposition 3.16 (resp., part (b) of Lemma 3.12) instead of Proposition 3.17 (resp., part (a) of Lemma 3.12). Similarly, the proof to the uniform sample-path large deviations stated in (2.20) is almost identical to that of (2.12) in Theorem 2.2. In particular, the only differences are that we apply part (b) of Lemma 3.12 (resp., Lemma 3.7) instead of part (a) of Lemma 3.12 (resp., Lemma 3.6). To avoid repetition we omit the details. □

3.3.1 Proof of Proposition 3.16

As has been demonstrated earlier, Proposition 3.16 lays the foundation for the sample-path LDPs of heavy-tailed stochastic difference equations. To disentangle the technicalities involved, the first step we will take is to provide further reduction to the assumptions in Proposition 3.16. Specifically, we show that it suffices to prove the seemingly more restrictive results stated below, where we impose the the boundedness condition in Assumption 4 and the stronger uniform nondegeneracy condition in Assumption 8.

Proposition 3.18. *Let η_n be a sequence of strictly positive real numbers with $\lim_{n\to\infty}\eta_n = 0$. Let compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ be such that $\lim_{n\to\infty} x_n = x^*$. Under Assumptions 1, 2, 4, and 8, it holds for any $k = 0, 1, 2, \cdots$ and $b > 0$ that*

$$\mathbf{P}\big(\boldsymbol{X}^{\eta_n|b}(x_n) \in \ \cdot\ \big)/\lambda^k(\eta_n) \to \mathbf{C}^{(k)|b}\big(\ \cdot\ ; x^*\big) \text{ in } \mathbb{M}\big(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)|b}\big) \text{ as } n \to \infty.$$

Proof of Proposition 3.16. Fix some $b > 0, k \geq 0$, as well as some $g \in \mathcal{C}\big(\mathbb{D}\setminus\mathbb{D}_A^{(k-1)|b}\big)$ that is also uniformly continuous on $\mathbb{D}$. Thanks to the Portmanteau theorem for $\mathbb{M}$-convergence (see theorem 2.1 of [34]), it suffices to show that $\lim_{n\to\infty}\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\big]/\lambda^k(\eta_n) = \mathbf{C}^{(k)|b}(g; x^*)$. Let $B \triangleq \text{supp}(g)$. Note that B is bounded away from $\mathbb{D}_A^{(k-1)|b}$. Applying Corollary 3.9, we can fix some M_0 such that the following claim holds for any $M \geq M_0$: for any $\xi = h_{M\downarrow}^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t})$ with $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0, 1]^{k\uparrow}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$ and $x_0 \in A$,

$$\xi = h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) = h_{M\downarrow}^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}); \quad \sup_{t\in[0,1]}|\xi(t)| \leq M_0. \tag{3.50}$$

Here the mapping $h_{M\downarrow}^{(k)|b}$ is defined in (3.27)-(3.29). Now fix some $M \geq M_0+1$ and recall the definitions of a_M, σ_M in (3.26). Also, define stochastic processes $\widetilde{\boldsymbol{X}}^{\eta|b}(x) \triangleq \big\{\widetilde{X}_{\lfloor t/\eta\rfloor}^{\eta|b}(x) : \ t \in [0, 1]\big\}$ as

$$\widetilde{X}_j^{\eta|b}(x) = \widetilde{X}_{j-1}^{\eta|b}(x) + \varphi_b\Big(\eta a_M\big(\widetilde{X}_{j-1}^{\eta|b}(x)\big) + \eta\sigma_M\big(\widetilde{X}_{j-1}^{\eta|b}(x)\big)Z_j\Big) \quad \forall j \geq 1$$

under initial condition $\widetilde{X}_0^{\eta|b}(x) = x$. In particular, by comparing the definition of $\widetilde{X}_j^{\eta|b}(x)$ with that of $X_j^{\eta|b}(x)$ in (2.13), we must have (for any $x \in \mathbb{R}, \eta > 0$)

$$\sup_{t\in[0,1]}\big|\widetilde{X}_{\lfloor t/\eta\rfloor}^{\eta|b}(x)\big| > M \iff \sup_{t\in[0,1]}\big|X_{\lfloor t/\eta\rfloor}^{\eta|b}(x)\big| > M, \tag{3.51}$$

$$\sup_{t\in[0,1]}\big|X_{\lfloor t/\eta\rfloor}^{\eta|b}(x)\big| \leq M \implies \boldsymbol{X}^{\eta|b}(x) = \widetilde{\boldsymbol{X}}^{\eta|b}(x). \tag{3.52}$$

Now observe that for any $n \geq 1$ (recall that $B = \text{supp}(g)$)

$$\begin{aligned}\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\big] &= \mathbf{E}\Big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\mathbb{I}\Big\{\boldsymbol{X}^{\eta_n|b}(x_n) \in B;\ \sup_{t\in[0,1]}\big|X_{\lfloor t/\eta\rfloor}^{\eta_n|b}(x_n)\big| \leq M\Big\}\Big] \\ &\quad + \mathbf{E}\Big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\mathbb{I}\Big\{\boldsymbol{X}^{\eta_n|b}(x_n) \in B;\ \sup_{t\in[0,1]}\big|X_{\lfloor t/\eta\rfloor}^{\eta_n|b}(x_n)\big| > M\Big\}\Big].\end{aligned} \tag{3.53}$$

An upper bound then follows immediately from (3.51) and (3.52):

$$\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\big] \leq \mathbf{E}\Big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\Big] + \|g\|\,\mathbf{P}\Big(\sup_{t\in[0,1]}\big|\widetilde{X}_{\lfloor t/\eta\rfloor}^{\eta_n|b}(x_n)\big| > M\Big).$$

Similarly, by bounding the first term on the R.H.S. of (3.53) using (3.51) and (3.52), we obtain

$$\begin{aligned}\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\big] &\geq \mathbf{E}\Big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\mathbb{I}\Big\{\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\in B;\ \sup_{t\in[0,1]}\big|\widetilde{X}^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big|\leq M\Big\}\Big] \\ &\geq \mathbf{E}\Big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\Big] - \|g\|\,\mathbf{P}\Big(\sup_{t\in[0,1]}\big|\widetilde{X}^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big| > M\Big).\end{aligned}$$

To conclude the proof, it only remains to show that

$$\lim_{n\to\infty}\lambda^{-k}(\eta_n)\mathbf{E}\big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\big] = \mathbf{C}^{(k)|b}(g;x^*), \tag{3.54}$$

$$\lim_{n\to\infty}\lambda^{-k}(\eta_n)\mathbf{P}\Big(\sup_{t\in[0,1]}\big|\widetilde{X}^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big| > M\Big) = 0. \tag{3.55}$$

Proof of Claim (3.54):

Under Assumption 3, one can easily see that a_M,σ_M would satisfy Assumption 4 and 8. This allows us to apply Proposition 3.18 and obtain $\lim_{n\to\infty}\lambda^{-k}(\eta_n)\mathbf{E}\big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\big] = \widetilde{\mathbf{C}}^{(k)|b}(g;x^*)$ where

$$\widetilde{\mathbf{C}}^{(k)|b}(\ \cdot\ ;x) \triangleq \int \mathbb{I}\Big\{h^{(k)|b}_{M\downarrow}\big(x,\boldsymbol{w},\boldsymbol{t}\big)\in\ \cdot\ \Big\}\nu^k_\alpha(d\boldsymbol{w})\times\mathcal{L}^{k\uparrow}_1(d\boldsymbol{t}).$$

Given (3.50) and the fact that $x^*\in A$, we immediately get $\widetilde{\mathbf{C}}^{(k)|b}(\ \cdot\ ;x^*) = \mathbf{C}^{(k)|b}(\ \cdot\ ;x^*)$ and conclude the proof of (3.54).

Proof of Claim (3.55):

Let $E \triangleq \{\xi\in\mathbb{D}:\ \sup_{t\in[0,1]}|\xi(t)| > M\}$. Suppose we can show that E is bounded away from $\mathbb{D}^{(k)|b}_A$, then by applying Proposition 3.18 again we get $\limsup_{n\to\infty}\mathbf{P}\Big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\in E\Big)\Big/\lambda^{k+1}(\eta_n) < \infty$, which then implies (3.55). To see why E is bounded away from $\mathbb{D}^{(k)|b}_A$, note that it follows directly from (3.50) that

$$\xi\in\mathbb{D}^{(k)|b}_A \Longrightarrow \sup_{t\in[0,1]}|\xi(t)|\leq M_0\leq M-1$$

due to our choice of $M\geq M_0+1$ at the beginning. Therefore, we yield $\boldsymbol{d}_{J_1}\big(\mathbb{D}^{(k)|b}_A, E\big)\geq 1$ and conclude the proof. □

The rest of Section 3.3 is devoted to establishing Proposition 3.18. In light of Lemma 3.15, a natural approach to the $\mathbb{M}$-convergence claim in Proposition 3.18 is to construct some process $\hat{\boldsymbol{X}}^{\eta|b;(k)}$ that is not only asymptotically equivalent to $\boldsymbol{X}^{\eta|b}$ (as $\eta\downarrow 0$) but also (under the right scaling) approaches to $\mathbf{C}^{(k)}_b$ in the sense of $\mathbb{M}$-convergence. To properly introduce the process $\hat{\boldsymbol{X}}^{\eta|b;(k)}$, a few new definitions are in order. For any $j\geq 1$ and $n\geq j$ let

$$\mathcal{J}_Z(c,n) \triangleq \#\{i\in[n]:\ |Z_i|\geq c\}\ \ \forall c\geq 0;\quad \boldsymbol{Z}^{(j)}(\eta)\triangleq\max\Big\{c\geq 0:\ \mathcal{J}_Z(c,\lfloor 1/\eta\rfloor)\geq j\Big\}. \tag{3.56}$$

In other words, $\mathcal{J}_Z(c,n)$ counts the number of elements in $\{|Z_i|: i\in[n]\}$ that are larger than c, and $\boldsymbol{Z}^{(j)}(\eta)$ identifies the value of the j^{th} largest element in $\{|Z_i|: i\leq\lfloor 1/\eta\rfloor\}$. Moreover, let

$$\tau^{(j)}_i(\eta)\triangleq\min\big\{k>\tau^{(j)}_{i-1}(\eta):\ |Z_k|\geq\boldsymbol{Z}^{(j)}(\eta)\big\},\quad W^{(j)}_i(\eta)\triangleq Z_{\tau^{(j)}_i(\eta)}\ \ \forall i=1,2,\cdots,j \tag{3.57}$$

with the convention that $\tau^{(j)}_0(\eta)=0$. Note that $\big(\tau^{(j)}_i(\eta),W^{(j)}_i(\eta)\big)_{i\in[j]}$ record the arrival time and size of the top j elements (in terms of absolute value) of $\{|Z_i|: i\in[n]\}$. In case that there are ties between the values of $\{|Z_i|: i\leq\lfloor 1/\eta\rfloor\}$, under our definition we always pick the first j elements. Now for any $j \geq 1$ and any $\eta, b > 0, x \in \mathbb{R}$, we are able to define $\hat{\boldsymbol{X}}^{\eta|b;(j)}(x) \triangleq \big\{\hat{X}_t^{\eta|b;(j)}(x) : \ t \in [0,1]\big\}$ as the solution to

$$\frac{d\hat{X}_t^{\eta|b;(j)}(x)}{dt} = a\big(\hat{X}_t^{\eta|b;(j)}(x)\big) \quad \forall t \in [0,1],\ t \notin \big\{\eta\tau_i^{(j)}(\eta) :\ i \in [j]\big\}, \tag{3.58}$$

$$\hat{X}_t^{\eta|b;(j)}(x) = \hat{X}_{t-}^{\eta|b;(j)}(x) + \varphi_b\Big(\eta\sigma\big(\hat{X}_{t-}^{\eta|b;(j)}(x)\big)W_i^{(j)}(\eta)\Big) \quad \text{if } t = \eta\tau_i^{(j)}(\eta) \text{ for some } i \in [j]. \tag{3.59}$$

with initial condition $\hat{X}_0^{\eta|b;(j)}(x) = x$. For the case $j = 0$, we adopt the convention that

$$d\hat{X}_t^{\eta|b;(0)}(x)/dt = a\big(\hat{X}_t^{\eta|b;(0)}(x)\big)\ \forall t \in [0,1]$$

with $\hat{X}_0^{\eta|b;(0)}(x) = x$. The key observation is that, by definition of $\hat{\boldsymbol{X}}^{\eta|b;(k)}$, it holds for any $\eta, b > 0$, $k \geq 0$, and $x \in \mathbb{R}$ that

$$\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta) \qquad \Longrightarrow \qquad \hat{\boldsymbol{X}}^{\eta|b;(k)}(x) = h^{(k)|b}\big(x, \eta\boldsymbol{W}^{>\delta}(\eta), \eta\boldsymbol{\tau}^{>\delta}(\eta)\big) \tag{3.60}$$

with $\boldsymbol{W}^{>\delta}(\eta) = (W_1^{>\delta}(\eta), \cdots, W_k^{>\delta}(\eta))$ and $\boldsymbol{\tau}^{>\delta}(\eta) = (\tau_1^{>\delta}(\eta), \cdots, \tau_k^{>\delta}(\eta))$. The following two results allow us to apply Lemma 3.15, thus bridging the gap between $\boldsymbol{X}^{\eta|b}$ and the limiting measure $\mathbf{C}^{(k)|b}$ in the sense of $\mathbb{M}$-convergence.

Proposition 3.19. *Let η_n be a sequence of strictly positive real numbers such that $\lim_{n\to\infty} \eta_n = 0$. Let compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ be such that $\lim_{n\to\infty} x_n = x^*$. Under Assumptions 1, 2, 4, and 8, it holds for any $k = 0, 1, 2, \cdots$ and $b > 0$ that $\boldsymbol{X}^{\eta_n|b}(x_n)$ is asymptotically equivalent to $\hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n)$ (as $n \to \infty$) w.r.t. $\lambda^k(\eta_n)$ when bounded away from $\mathbb{D}_A^{(k-1)|b}$.*

Proposition 3.20. *Let η_n be a sequence of strictly positive real numbers with $\lim_{n\to\infty} \eta_n = 0$. Let compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ be such that $\lim_{n\to\infty} x_n = x^*$. Under Assumptions 1, 2, 4, and 8, it holds for any $k = 0, 1, 2, \cdots$ and $b > 0$ that*

$$\mathbf{P}\Big(\hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n) \in \ \cdot\ \Big)\Big/\lambda^k(\eta_n) \to \mathbf{C}^{(k)|b}\big(\ \cdot\ ; x^*\big) \ \text{ in } \mathbb{M}\big(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)|b}\big) \ \text{ as } n \to \infty$$

where the measure $\mathbf{C}^{(k)|b}$ is defined in (2.19).

Proof of Proposition 3.18. In light of Lemma 3.15, it is a direct corollary of Propositions 3.19 and 3.20. □

Now it only remains to prove Propositions 3.19 and 3.20.

Proof of Proposition 3.19. Fix some $b > 0, k \geq 0$, and some sequence of strictly positive real numbers η_n with $\lim_{n\to\infty} \eta_n = 0$. Also, fix a compact set $A \subseteq \mathbb{R}$ and $x_n, x^* \in A$ such that $\lim_{n\to\infty} x_n = x^*$. Meanwhile, arbitrarily pick some $\Delta > 0$ and some $B \in \mathscr{S}_{\mathbb{D}}$ that is bounded away from $\mathbb{D}_A^{(k-1)|b}$. It suffices to show that

$$\lim_{n\to\infty} \mathbf{P}\Big(\boldsymbol{d}_{J_1}\big(\boldsymbol{X}^{\eta_n|b}(x_n), \hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n)\big)\mathbb{I}\big(\boldsymbol{X}^{\eta_n|b}(x_n) \text{ or } \hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n) \in B\big) > \Delta\Big)\Big/\lambda^k(\eta_n) = 0. \tag{3.61}$$

Applying Lemma 3.7, we can fix some $\bar{\epsilon} > 0$ and $\bar{\delta} \in (0, \frac{b}{3C})$ such that for any $x \in A$, $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1]^{k\uparrow}$, and $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$,

$$h^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \text{ or } h^{(k)|b+\bar{\epsilon}}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \qquad \Longrightarrow \qquad |w_i| > 3C\bar{\delta}/c\ \forall i \in [k]. \tag{3.62}$$

$$\boldsymbol{d}_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}) > \bar{\epsilon} \tag{3.63}$$

where $C \geq 1$ and $0 < c \leq 1$ are the constants in Assumptions 4 and 8, respectively. Meanwhile, let

$$B_0 \triangleq \{\boldsymbol{X}^{\eta|b}(x) \in B \text{ or } \hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \in B;\ \boldsymbol{d}_{J_1}\big(\boldsymbol{X}^{\eta|b}(x), \hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big) > \Delta\},$$

$$B_1 \triangleq \{\tau^{>\delta}_{k+1}(\eta) > \lfloor 1/\eta \rfloor\},$$
$$B_2 \triangleq \{\tau^{>\delta}_{k}(\eta) \leq \lfloor 1/\eta \rfloor\},$$
$$B_3 \triangleq \big\{\eta|W^{>\delta}_i(\eta)| > \bar{\delta} \text{ for all } i \in [k]\big\}.$$

Note that

$$B_0 = (B_0 \cap B_1^c) \cup (B_0 \cap B_1 \cap B_2^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3). \tag{3.64}$$

To proceed, set $\rho^{(k)} \triangleq \Big[3\rho\cdot(1+\frac{bD}{c})\Big]^k \cdot 3\rho$ where $\rho = \exp(D)$ and $D \in [1,\infty)$ is the Lipschitz coefficient in Assumption 2. For any $\epsilon > 0$ small enough so that

$$\rho^{(k)}\sqrt{\epsilon} < \Delta, \quad \epsilon < \frac{\bar{\delta}}{2\rho}, \quad \epsilon < \bar{\epsilon}/2, \quad \epsilon \in (0,1),$$

we claim that

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_0 \cap B_1^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.65}$$
$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_0 \cap B_1 \cap B_2^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.66}$$
$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_0 \cap B_1 \cap B_2 \cap B_3^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.67}$$
$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_0 \cap B_1 \cap B_2 \cap B_3\Big)\Big/\lambda^k(\eta) = 0 \tag{3.68}$$

if we pick $\delta > 0$ sufficiently small. Now fix such δ. Combining these claims with the decomposition of event B_0 in (3.64), we establish (3.61). Now we conclude the proof of this proposition with the proofs of claims (3.65)–(3.68).

Proof of (3.65):

For any $\delta > 0$, note that (3.4) implies that $\sup_{x\in A}\mathbf{P}(B_0 \cap B_1^c) \leq \mathbf{P}(B_1^c) \leq \big(\eta^{-1}H(\delta\eta^{-1})\big)^{k+1} = \boldsymbol{o}\big(\lambda^k(\eta)\big)$, from which the claim follows.

Proof of (3.66):

It suffices to find $\delta > 0$ such that

$$\lim_{\eta\downarrow 0}\mathbf{P}\Big(\underbrace{B_0 \cap \{\tau^{>\delta}_k(\eta) > \lfloor 1/\eta \rfloor\}}_{\triangleq \widetilde{B}}\Big)\Big/\lambda^k(\eta) = 0$$

In particular, we focus on $\delta \in (0, \bar{\delta} \wedge \frac{b}{2})$ with $\bar{\delta}$ characterized in (3.62). By definition, $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) = h^{(k)|b}\big(x, \eta\tau^{(k)}_1(\eta), \cdots, \eta\tau^{(k)}_k(\eta), \eta W^{(k)}_1(\eta), \cdots, \eta W^{(k)}_k(\eta)\big)$. Moreover, on $\{\tau^{>\delta}_k(\eta) > \lfloor 1/\eta \rfloor\}$ we must have $\#\big\{i \in \big[\lfloor 1/\eta \rfloor\big] : \ \eta|Z_i| > \delta\big\} < k$. From the definition of $\boldsymbol{Z}^{(k)}(\eta)$ in (3.56), we then have that $\min_{i\in[k]}\eta|W^{(k)}_i(\eta)| \leq \delta$. In light of (3.62), we yield $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \notin B^{\bar{\epsilon}}$ on $\{\tau^{>\delta}_k(\eta) > \lfloor 1/\eta \rfloor\}$, and hence

$$\widetilde{B} \subseteq \{\boldsymbol{X}^{\eta|b}(x) \in B\} \cap \{\tau^{>\delta}_k(\eta) > \lfloor 1/\eta \rfloor\}.$$

Let event $A_i(\eta, b, \epsilon, \delta, x)$ be defined as in (3.6). Suppose that

$$\{\boldsymbol{X}^{\eta|b}(x) \in B\} \cap \{\tau^{>\delta}_k(\eta) > \lfloor 1/\eta \rfloor\} \cap \big(\cap_{i=1}^k A_i(\eta, b, \epsilon, \delta, x)\big) = \emptyset \tag{3.69}$$

holds for all $\eta > 0$ small enough with $\eta < \min\{\frac{b\wedge 1}{2C}, \frac{\epsilon}{C}\}$, any $\delta \in (0, \frac{b}{2C})$, and any $x \in A$. Then

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}(\widetilde{B})\Big/\lambda^k(\eta) \leq \limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(\big(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\big)^c\Big)\Big/\lambda^k(\eta).$$

To conclude the proof, one only need to apply Lemma 3.4 (b) with some $N > k(\alpha-1)$.

Now it only remains to prove claim (3.69). To proceed, let process $\breve{X}^{\eta|b;\delta}_t(x)$ be the solution to

$$\frac{d\breve{X}^{\eta|b;\delta}_t(x)}{dt} = a\big(\breve{X}^{\eta|b;\delta}_t(x)\big) \qquad \forall t\in[0,\infty)\setminus\{\eta\tau^{>\delta}_j(\eta):\ j\geq 1\}, \tag{3.70}$$

$$\breve{X}^{\eta|b;\delta}_{\eta\tau^{>\delta}_j(\eta)}(x) = X^{\eta|b}_{\tau^{>\delta}_j(\eta)}(x) \qquad \forall j\geq 1 \tag{3.71}$$

under the initial condition $\breve{X}^{\eta|b;\delta}_0(x)=x$. Let $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x)\triangleq\big\{\breve{X}^{\eta|b;\delta}_t(x):\ t\in[0,1]\big\}$. For any $j\geq 1$, observe that on event $\big(\cap_{i=1}^j A_i(\eta,b,\epsilon,\delta,x)\big)\cap\{\tau^{>\delta}_j(\eta)>\lfloor 1/\eta\rfloor\}$,

$$\begin{aligned}
&\boldsymbol{d}_{J_1}\big(\breve{\boldsymbol{X}}^{\eta|b;\delta}(x),\boldsymbol{X}^{\eta|b}(x)\big)\\
&\leq \sup_{t\in\left[0,\eta\tau^{>\delta}_1(\eta)\right)\cup\left[\eta\tau^{>\delta}_1(\eta),\eta\tau^{>\delta}_2(\eta)\right)\cup\cdots\cup\left[\eta\tau^{>\delta}_{j-1}(\eta),\eta\tau^{>\delta}_j(\eta)\right)}\left|\breve{X}^{\eta|b;\delta}_t(x)-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right|\\
&\leq \rho\cdot\big(\epsilon+\eta C\big)\leq 2\rho\epsilon<\bar\epsilon \qquad \text{due to (3.36) of Lemma 3.13.}
\end{aligned}\tag{3.72}$$

Therefore, on event $\big(\cap_{i=1}^k A_i(\eta,b,\epsilon,\delta,x)\big)\cap\{\tau^{>\delta}_k(\eta)>\lfloor 1/\eta\rfloor\}$, it holds for any $j\in[k-1]$ with $\eta\tau^{>\delta}_j(\eta)\leq 1$ that

$$\begin{aligned}
\left|\Delta\breve{X}^{\eta|b;\delta}_{\eta\tau^{>\delta}_j(\eta)}(x)\right| &= \left|\breve{X}^{\eta|b;\delta}_{\eta\tau^{>\delta}_j(\eta)-}(x)-X^{\eta|b}_{\tau^{>\delta}_j(\eta)}(x)\right| \qquad \text{see (3.71)}\\
&\leq \left|\breve{X}^{\eta|b;\delta}_{\eta\tau^{>\delta}_j(\eta)-}(x)-X^{\eta|b}_{\tau^{>\delta}_j(\eta)-1}(x)\right|+\left|X^{\eta|b}_{\tau^{>\delta}_j(\eta)-1}(x)-X^{\eta|b}_{\tau^{>\delta}_j(\eta)}(x)\right|\\
&<\bar\epsilon+b.
\end{aligned}\tag{3.73}$$

As a result, on event $\big(\cap_{i=1}^k A_i(\eta,b,\epsilon,\delta,x)\big)\cap\{\tau^{>\delta}_k(\eta)>\lfloor 1/\eta\rfloor\}$, we have $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x)\in\mathbb{D}^{(k-1)|b+\bar\epsilon}_A$. Considering the facts that $\mathbb{D}^{(k-1)|b+\bar\epsilon}_A$ is bounded away from $B^{\bar\epsilon}$ (see (3.63)) as well as $\boldsymbol{d}_{J_1}\big(\breve{\boldsymbol{X}}^{\eta|b;\delta}(x),\boldsymbol{X}^{\eta|b}(x)\big)<\bar\epsilon$ shown in (3.72), we have just established that $\boldsymbol{X}^{\eta|b}(x)\notin B$, thus establishing (3.69).

Proof of (3.67):

On event $B_1\cap B_2=\{\tau^{>\delta}_k(\eta)\leq\lfloor 1/\eta\rfloor<\tau^{>\delta}_{k+1}(\eta)\}$, it follows from (3.60) that $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)=h^{(k)|b}\big(x,\eta W^{>\delta}_1(\eta),\cdots,\eta W^{>\delta}_k(\eta),\eta\tau^{>\delta}_1(\eta),\cdots,\eta\tau^{>\delta}_k(\eta)\big)$. Furthermore, on B_3^c, there is some $i\in[k]$ with $|\eta W^{>\delta}_i(\eta)|\leq\bar\delta$. Considering the choice of $\bar\delta$ in (3.62), on event $B_1\cap B_2\cap B_3^c$ we have $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\notin B$, and hence

$$B_0\cap B_1\cap B_2\cap B_3^c\subseteq\{\boldsymbol{X}^{\eta|b}(x)\in B\}\cap\big\{\tau^{>\delta}_k(\eta)\leq\lfloor 1/\eta\rfloor<\tau^{>\delta}_{k+1}(\eta);\ \eta|W^{>\delta}_i(\eta)|\leq\bar\delta\text{ for some }i\in[k]\big\}.$$

Furthermore, we claim that for any $x\in A$, any $\delta\in(0,\bar\delta\wedge\frac{b}{2C})$ and any $\eta>0$ satisfying $\eta<\min\{\frac{b\wedge 1}{2C},\bar\delta\}$,

$$\begin{aligned}
&\{\boldsymbol{X}^{\eta|b}(x)\in B\}\cap\big\{\tau^{>\delta}_k(\eta)\leq\lfloor 1/\eta\rfloor<\tau^{>\delta}_{k+1}(\eta);\ \eta|W^{>\delta}_i(\eta)|\leq\bar\delta\text{ for some }i\in[k]\big\}\\
&\qquad\qquad\cap\Big(\bigcap_{i=1}^{k+1}A_i(\eta,b,\epsilon,\delta,x)\Big)=\emptyset.
\end{aligned}\tag{3.74}$$

Then it follows immediately that for any $\delta\in(0,\bar\delta\wedge\frac{b}{2})$,

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_0\cap B_1\cap B_2\cap B_3^c\Big)\Big/\lambda^k(\eta)\leq\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(\big(\bigcap_{i=1}^{k+1}A_i(\eta,b,\epsilon,\delta,x)\big)^c\Big)\Big/\lambda^k(\eta).$$

Applying Lemma 3.4 (b) with some $N>k(\alpha-1)$, the conclusion of the proof follows.

We are left with proving the claim (3.74). First, note that on this event, there exists some $J \in [k]$ such that $\eta|W_J^{>\delta}(\eta)| \leq \bar{\delta}$. Next, recall the definition of $\breve{X}_t^{\eta|b;\delta}(x)$ in (3.70)-(3.71), and note that it has been shown in (3.72) (with $j = k+1$) that

$$\sup_{t\in[0,1]} \left|\breve{X}_t^{\eta|b;\delta}(x) - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\right| < 2\rho\epsilon < \bar{\epsilon}. \tag{3.75}$$

If we can show that $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$, then (3.75) immediately leads to $\boldsymbol{X}^{\eta|b}(x) \notin B$, thus proving claim (3.74). To proceed, first note that

$$\begin{aligned}
\left|\Delta \breve{X}^{\eta|b;\delta}_{\eta\tau_J^{>\delta}(\eta)}(x)\right| &\leq \left|\breve{X}^{\eta|b;\delta}_{\eta\tau_J^{>\delta}(\eta)-}(x) - X^{\eta|b}_{\tau_J^{>\delta}(\eta)-1}(x)\right| + \left|X^{\eta|b}_{\tau_J^{>\delta}(\eta)-1}(x) - X^{\eta|b}_{\tau_J^{>\delta}(\eta)}(x)\right| \qquad \text{see (3.71)} \\
&\leq 2\rho\epsilon + \eta\left|a\big(X^{\eta|b}_{\tau_J^{>\delta}(\eta)-1}(x)\big) + \sigma\big(X^{\eta|b}_{\tau_J^{>\delta}(\eta)-1}(x)\big)W_J^{>\delta}(\eta)\right| \qquad \text{using (3.75)} \\
&\leq 2\rho\epsilon + \eta C + C\bar{\delta} < 3C\bar{\delta} \qquad \text{due to } 2\rho\epsilon < \bar{\delta},\ \eta < \bar{\delta},\text{ and } C \geq 1.
\end{aligned}$$

Meanwhile, the calculations in (3.73) can be repeated to show that $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) \in \mathbb{D}_A^{(k)|b+\bar{\epsilon}}$, and hence $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) = h^{(k)|b+\bar{\epsilon}}\big(x, \widetilde{w}_1, \cdots, \widetilde{w}_k, \eta\tau_1^{>\delta}(\eta), \cdots, \eta\tau_k^{>\delta}(\eta)\big)$ for some $(\widetilde{w}_1, \cdots, \widetilde{w}_k) \in \mathbb{R}^k$. Due to $0 < c \leq \sigma(y) \leq C\ \forall y \in \mathbb{R}$ (see Assumptions 4 and 8),

$$3C\bar{\delta} > \left|\Delta \breve{X}^{\eta|b;\delta}_{\eta\tau_J^{>\delta}(\eta)}(x)\right| = \varphi_{b+\bar{\epsilon}}\left(\left|\sigma\Big(\breve{X}^{\eta|b;\delta}_{\eta\tau_J^{>\delta}(\eta)-}(x)\Big)\cdot \widetilde{w}_J\right|\right) \geq c\cdot|\widetilde{w}_J|,$$

which implies $|\widetilde{w}_J| < 3C\bar{\delta}/c$. In light of our choice of $\bar{\delta}$ in (3.62), we yield $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$ and conclude the proof.

Proof of (3.68):

We focus on $\delta \in (0, \bar{\delta} \wedge \frac{b}{2C})$. On event $B_1 \cap B_2 = \{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\}$, $\hat{\boldsymbol{X}}^{\eta|b;(k)}$ admits the expression in (3.60). This allows us to apply Lemma 3.14 and show that, for any $x \in A$ and any $\eta \in (0, \frac{\epsilon\wedge b}{2C})$, the inequality

$$\boldsymbol{d}_{J_1}\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x), \boldsymbol{X}^{\eta|b}(x)\big) \leq \sup_{t\in[0,1]} |\hat{X}_t^{\eta|b;(k)}(x) - X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)| < \rho^{(k)}\epsilon$$

holds on event $\big(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\big) \cap B_1 \cap B_2 \cap B_3 \cap B_0$. Due to our choice of $\rho^{(k)}\epsilon < \Delta$ at the beginning of the proof, we get $\big(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\big) \cap B_1 \cap B_2 \cap B_3 \cap B_0 = \emptyset$. Therefore,

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(B_1 \cap B_2 \cap B_3 \cap B_0\Big)\Big/\lambda^k(\eta) \leq \limsup_{\eta\downarrow 0}\sup_{x\in A}\mathbf{P}\Big(\big(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\big)^c\Big)\Big/\lambda^k(\eta).$$

Again, by applying Lemma 3.4 (b) with some $N > k(\alpha - 1)$, we conclude the proof. □

In order to prove Proposition 3.20, we first prepare a lemma regarding a weak convergence claim on event $E^\delta_{c,k}(\eta) \triangleq \Big\{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta\rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta|W_j^{>\delta}(\eta)| > c\ \ \forall j \in [k]\Big\}$ defined in (3.16).

Lemma 3.21. *Let Assumption 1 hold. Let $A \subseteq \mathbb{R}$ be a compact set. Let bounded function $\Phi : \mathbb{R}\times\mathbb{R}^k\times(0,1]^{k\uparrow} \to \mathbb{R}$ be continuous on $\mathbb{R}\times\mathbb{R}^k\times(0,1)^{k\uparrow}$. For any $\delta > 0, c > \delta$ and $k = 0, 1, 2, \cdots$,*

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\left|\frac{\mathbf{E}\Big[\Phi\big(x, \eta W_1^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta), \eta\tau_1^{>\delta}(\eta), \cdots, \eta\tau_k^{>\delta}(\eta)\big)\mathbb{I}_{E^\delta_{c,k}(\eta)}\Big]}{\lambda^k(\eta)} - \frac{(1/c^{\alpha k})\phi_{c,k}(x)}{k!}\right| = 0$$

where $\phi_{c,k}(x) \triangleq \mathbf{E}\Big[\Phi\big(x, W_1^*(c), \cdots, W_k^*(c), U_{(1;k)}, \cdots, U_{(k;k)}\big)\Big]$.

Proof. Fix some $\delta > 0, c > \delta$ and $k = 0, 1, \cdots$. We proceed with a proof by contradiction. Suppose there exist some $\epsilon > 0$, some sequence $x_n \in A$, and some sequence $\eta_n > 0$ such that

$$\left|\lambda^{-k}(\eta_n)\mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\mathbb{I}_{E^{\delta}_{c,k}(\eta_n)}\Big] - (1/k!)\cdot c^{-\alpha k}\cdot \phi_{c,k}(x_n)\right| > \epsilon \quad \forall n \geq 1 \tag{3.76}$$

where $\boldsymbol{W}^{\eta} \triangleq (\eta W_1^{>\delta}(\eta), \cdots, \eta W_k^{>\delta}(\eta))$; $\boldsymbol{\tau}^{\eta} \triangleq (\eta\tau_1^{>\delta}(\eta), \cdots, \eta\tau_k^{>\delta}(\eta))$. Since A is compact, we can always pick a converging subsequence x_{n_k} such that $x_{n_k} \to x^*$ for some $x^* \in A$. To ease the notation complexity, let's assume (w.l.o.g.) that $x_n \to x^*$. Now observe that

$$\begin{aligned}
&\lim_{n\to\infty} \lambda^{-k}(\eta_n)\mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\mathbb{I}_{E^{\delta}_{c,k}(\eta_n)}\Big] \\
&= \Big[\lim_{n\to\infty} \lambda^{-k}(\eta_n)\mathbf{P}\big(E^{\delta}_{c,k}(\eta_n)\big)\Big] \cdot \lim_{n\to\infty} \mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\Big| E^{\delta}_{c,k}(\eta_n)\Big] \\
&= (1/k!)\cdot c^{-\alpha k}\cdot \mathbf{E}\Big[\Phi\big(x^*, \boldsymbol{W}^*, \boldsymbol{U}^*\big)\Big] = (1/k!)\cdot c^{-\alpha k}\cdot \phi_{c,k}(x^*) \qquad \text{due to Lemma 3.5}
\end{aligned}$$

where $\boldsymbol{W}^* \triangleq \big(W_j^*(c)\big)_{j=1}^k$, $\boldsymbol{U}^* \triangleq \big(U_{(j;k)}\big)_{j=1}^k$. However, by Bounded Convergence theorem, we see that $\phi_{c,k}$ is also continuous, and hence $\phi_{c,k}(x_n) \to \phi_{c,k}(x^*)$. This leads to a contradiction with (3.76) and allows us to conclude the proof. □

We are now ready to prove Proposition 3.20.

Proof of Proposition 3.20. Fix some $b > 0$, some $k = 0, 1, 2, \cdots$ and $g \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)|b})$ (i.e., $g : \mathbb{D} \to [0, \infty)$ is continuous and bounded with support $B \triangleq \text{supp}(g)$ bounded away from $\mathbb{D}_A^{(k-1)|b}$). First of all, from Lemma 3.7 we can fix some $\bar{\delta} > 0$ such that the following claim holds for any $x_0 \in A$ and any $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1)^{k\uparrow}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$:

$$h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \Longrightarrow |w_j| > \bar{\delta} \;\; \forall j \in [k]. \tag{3.77}$$

Fix some $\delta \in (0, \bar{\delta} \wedge \frac{b}{2})$, and observe that for any $\eta > 0$ and $x \in A$,

$$\begin{aligned}
g\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big) &= \underbrace{g\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big)\mathbb{I}\big\{\tau^{>\delta}_{k+1}(\eta) \leq \lfloor 1/\eta \rfloor\big\}}_{\triangleq I_1(\eta,x)} + \underbrace{g\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big)\mathbb{I}\big\{\tau^{>\delta}_{k}(\eta) > \lfloor 1/\eta \rfloor\big\}}_{\triangleq I_2(\eta,x)} \\
&\quad + \underbrace{g\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big)\mathbb{I}\big\{\tau^{>\delta}_{k}(\eta) \leq \lfloor 1/\eta \rfloor < \tau^{>\delta}_{k+1}(\eta);\ |\eta W^{>\delta}_j(\eta)| \leq \bar{\delta} \text{ for some } j \in [k]\big\}}_{\triangleq I_3(\eta,x)} \\
&\quad + \underbrace{g\big(\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big)\mathbb{I}\big(E^{\delta}_{\bar{\delta},k}(\eta)\big)}_{\triangleq I_4(\eta,x)}.
\end{aligned}$$

For term $I_1(\eta, x)$, it follows from (3.4) that $\sup_{x\in\mathbb{R}} \mathbf{E}[I_1(\eta, x)] \leq \|g\| \cdot \Big[\frac{1}{\eta_n} \cdot H(\delta/\eta_n)\Big]^{k+1}$. Therefore, $\lim_{\eta\downarrow 0} \sup_{x\in A} \mathbf{E}[I_1(\eta, x)] \Big/ \big(\eta^{-1}H(\eta^{-1})\big)^k \leq \frac{\|g\|}{\delta^{\alpha(k+1)}} \cdot \lim_{n\to\infty} \frac{H(1/\eta)}{\eta} = 0$.

Next, by definition, $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) = h^{(k)|b}\big(x, \eta\tau_1^{(k)}(\eta), \cdots, \eta\tau_k^{(k)}(\eta), \eta W_1^{(k)}(\eta), \cdots, \eta W_k^{(k)}(\eta)\big)$. Moreover, on event $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$, we must have $\#\big\{i \in \big[\lfloor 1/\eta \rfloor\big] :\ \eta|Z_i| > \delta\big\} < k$. From the definition of $\boldsymbol{Z}^{(k)}(\eta)$ in (3.56), we then have that $\min_{i\in[k]} \eta|W_i^{(k)}(\eta)| \leq \delta$. In light of (3.77) and our choice of $\delta < \bar{\delta}$, for any $x \in A$ and any $\eta > 0$, on event $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ we have $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \notin B$ for $B = \text{supp}(g)$, thus implying $I_2(\eta, x) = 0$ for any $x \in A$ and $\eta > 0$.

Moving onto term $I_3(\eta, x)$, on event $\{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\}$ the process $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)$ admits the expression in (3.60), which implies $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \notin B$. due to (3.77) and our choice of $\delta < \bar{\delta}$. In summary, we get $I_3(\eta, x) = 0$.

Lastly, on event $E^{\delta}_{\bar{\delta},k}(\eta)$, the process $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)$ would again admits the expression in (3.60). As a result, for any $\eta > 0$ and $x \in A$, we have

$$\mathbf{E}[I_4(\eta,x)] = \mathbf{E}\big[\Phi\big(x, \eta\boldsymbol{W}^{>\delta}(\eta), \eta\boldsymbol{\tau}^{>\delta}(\eta)\big)\mathbb{I}\big(E^{\delta}_{\bar{\delta},k}(\eta)\big)\big]$$

where $\boldsymbol{W}^{>\delta}(\eta) = (W^{>\delta}_1(\eta), \cdots, W^{>\delta}_k(\eta))$, $\boldsymbol{\tau}^{>\delta}(\eta) = (\tau^{>\delta}_1(\eta), \cdots, \tau^{>\delta}_k(\eta))$, and $\Phi : \mathbb{R}\times\mathbb{R}^k\times(0,1)^{k\uparrow} \to \mathbb{R}$ is defined as $\Phi(x_0, \boldsymbol{w}, \boldsymbol{t}) \triangleq g\big(h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t})\big)$. Meanwhile, let $\phi(x) \triangleq \mathbf{E}\Big[\Phi\big(x, W^*_1(\bar{\delta}), \cdots, W^*_k(\bar{\delta}), U_{(1;k)}, \cdots, U_{(k;k)}\big)\Big]$. First, the continuity of mapping Φ on $\mathbb{R}\times\mathbb{R}^k\times(0,1)^{k\uparrow}$ follows directly from the continuity of g and $h^{(k)|b}$ (see Lemma 3.10). Besides, $\|\Phi\| \leq \|g\| < \infty$. It then follows from the continuity of Φ and Bounded Convergence Theorem that ϕ is also continuous. Also, $\|\phi\| \leq \|\Phi\| \leq \|g\| < \infty$. Now observe that

$$\limsup_{\eta\downarrow 0}\sup_{x\in A}\Big|\lambda^{-k}(\eta)\mathbf{E}\Big[\Phi\big(x, \eta\boldsymbol{W}^{>\delta}(\eta), \eta\boldsymbol{\tau}^{>\delta}(\eta)\big)\mathbb{I}\big(E^{\delta}_{\bar{\delta},k}(\eta)\big)\Big] - (1/k!)\cdot c^{-\alpha k}\cdot\phi_{c,k}(x)\Big| = 0$$

due to Lemma 3.21. Meanwhile, due to continuity of $\phi(\cdot)$, for any $x_n, x^* \in A$ with $\lim_{n\to\infty} x_n = x^*$, we have $\lim_{n\to\infty}\phi(x_n) = \phi(x^*)$. To conclude the proof, we only need to show that $\frac{(1/\bar{\delta}^{\alpha k})\phi(x^*)}{k!} = \mathbf{C}^{(k)|b}(g;x^*)$. In particular, note that

$$\phi(x^*) = \int g\big(h^{(k)|b}(x^*, w_1, \cdots, w_k, t_1, \cdots, t_k)\big)\mathbb{I}\big\{|w_j| > \bar{\delta}\ \forall j \in [k]\big\}$$
$$\mathbf{P}\Big(U_{(1;k)} \in dt_1, \cdots, U_{(k;k)} \in dt_k\Big) \times \Big(\prod_{j=1}^{k}\bar{\delta}^{\alpha}\cdot\nu_\alpha(dw_j)\Big).$$

First, using (3.77), we must have $g\big(h^{(k)}(x^*, w_1, \cdots, w_k, \boldsymbol{t})\big) = 0$ if there is some $j \in [k]$ with $|w_j| \leq \bar{\delta}$. Next, $\mathbf{P}\Big(U_{(1;k)} \in dt_1, \cdots, U_{(k;k)} \in dt_k\Big) = k!\cdot\mathbb{I}\{0 < t_1 < t_2 < \cdots < t_k < 1\}\mathcal{L}^{k\uparrow}_1(dt_1, \cdots, dt_k)$ where $\mathcal{L}^{k\uparrow}_1$ is the Lebesgue measure restricted on $(0,1)^{k\uparrow}$. The conclusion of the proof then follows from

$$\phi(x^*) = k!\cdot\bar{\delta}^{\alpha k}\int g\big(h^{(k)|b}(x^*, \boldsymbol{w}, \boldsymbol{t})\big)\nu^k_\alpha(d\boldsymbol{w})\times\mathcal{L}^{k\uparrow}_1(d\boldsymbol{t}) = k!\cdot\bar{\delta}^{\alpha k}\cdot\mathbf{C}^{(k)}_b\big(g;x^*\big),$$

where we appealed to the definition in (2.19) in the last equality. □

4 First Exit Time Analysis

4.1 Proof of Theorem 2.7

Our proof of Theorem 2.7 hinges on the following proposition.

Proposition 4.1. *Suppose that Condition 1 holds. For each measurable set $B \subseteq \mathbb{S}$ and $t \geq 0$, there exists $\delta_{t,B}(\epsilon)$ such that*

$$C(B^\circ)\cdot e^{-t} - \delta_{t,B}(\epsilon) \leq \liminf_{\eta\downarrow 0}\inf_{x\in A(\epsilon)}\mathbf{P}\big(\gamma(\eta)\tau^{\eta}_{I(\epsilon)^c}(x) > t;\ V^{\eta}_{\tau_\epsilon}(x) \in B\big)$$
$$\leq \limsup_{\eta\downarrow 0}\sup_{x\in A(\epsilon)}\mathbf{P}\big(\gamma(\eta)\tau^{\eta}_{I(\epsilon)^c}(x) > t;\ V^{\eta}_{\tau_\epsilon}(x) \in B\big) \leq C(B^-)\cdot e^{-t} + \delta_{t,B}(\epsilon).$$

for all sufficiently small $\epsilon > 0$, where $\delta_{t,B}(\epsilon) \to 0$ as $\epsilon \to 0$.

Proof. Fix some measurable $B \subseteq \mathbb{S}$ and $t \geq 0$. Henceforth in the proof, given any choice of $0 < r < R$, we only consider $\epsilon \in (0, \epsilon_B)$ and T sufficiently large such that Condition 1 holds with T replaced with $\frac{1-r}{2}T$, $\frac{2-r}{2}T$, rT, and RT. Let

$$\rho^{\eta}_i(x) \triangleq \inf\Big\{j \geq \rho^{\eta}_{i-1}(x) + \lfloor rT/\eta\rfloor : V^{\eta}_j(x) \in A(\epsilon)\Big\}$$

where $\rho^\eta_0(x) = 0$. One can interpret these as the i^{th} asymptotic regeneration times after cooling period rT/η. We start with the following two observations: For any $0 < r < R$,

$$\begin{aligned}
\mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \in \big(RT/\eta,\, \rho^\eta_1(y)\big]\Big) &\leq \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \wedge \rho^\eta_1(y) > RT/\eta\Big) \\
&\leq \mathbf{P}\Big(V^\eta_j(y) \in I(\epsilon)\setminus A(\epsilon) \quad \forall j \in \big[\lfloor rT/\eta\rfloor,\, RT/\eta\big]\Big) \\
&\leq \sup_{z\in I(\epsilon)\setminus A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{(I(\epsilon)\setminus A(\epsilon))^c}(z) > \frac{R-r}{2}T/\eta\Big) \\
&= \gamma(\eta)T/\eta \cdot o(1) \qquad (4.1)
\end{aligned}$$

where the last equality is from (2.39) of Condition 1, and

$$\begin{aligned}
&\sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\; \tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big) \\
&\leq \sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\; \tau^\eta_{I(\epsilon)^c}(y) \leq RT/\eta\Big) + \sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \in \big(RT/\eta,\, \rho^\eta_1(y)\big]\Big) \\
&\leq \sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\; \tau^\eta_{I(\epsilon)^c}(y) \leq RT/\eta\Big) + \gamma(\eta)T/\eta \cdot o(1) \\
&\leq \big(C(B^-) + \delta_B(\epsilon, RT) + o(1)\big)\cdot \gamma(\eta)RT/\eta \qquad (4.2)
\end{aligned}$$

where the second inequaility is from (4.1) and the last equality is from (2.38) of Condition 1.

We work with different choices of R and r for the lower and upper bounds. For the lower bound, we work with $R > r > 1$ and set $K = \left\lceil \frac{t/\gamma(\eta)}{T/\eta} \right\rceil$. Note that for $\eta \in \big(0, (r-1)T\big)$, we have $\lfloor rT/\eta\rfloor \geq T/\eta$ and hence $\rho^\eta_K(x) \geq K\lfloor rT/\eta\rfloor \geq t/\gamma(\eta)$. Note also that from the Markov property conditioning on $\mathcal{F}_{\rho^\eta_j(x)}$,

$$\begin{aligned}
&\inf_{x\in A(\epsilon)} \mathbf{P}\big(\gamma(\eta)\tau^\eta_{I(\epsilon)^c}(x) > t;\; V^\eta_{\tau_\epsilon}(x) \in B\big) \\
&\geq \inf_{x\in A(\epsilon)} \mathbf{P}(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_K(x);\; V^\eta_{\tau_\epsilon}(x) \in B) = \inf_{x\in A(\epsilon)} \sum_{j=K}^\infty \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) \in \big(\rho^\eta_j(x),\, \rho^\eta_{j+1}(x)\big];\; V^\eta_{\tau_\epsilon}(x) \in B\Big) \\
&\geq \inf_{x\in A(\epsilon)} \sum_{j=K}^\infty \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) \in \big(\rho^\eta_j(x),\, \rho^\eta_j(x) + T/\eta\big];\; V^\eta_{\tau_\epsilon}(x) \in B\Big) \\
&\geq \inf_{x\in A(\epsilon)} \sum_{j=K}^\infty \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq T/\eta;\; V^\eta_{\tau_\epsilon}(y) \in B\Big)\cdot \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big). \\
&\geq \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq T/\eta;\; V^\eta_{\tau_\epsilon}(y) \in B\Big)\cdot \sum_{j=K}^\infty \inf_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big). \qquad (4.3)
\end{aligned}$$

From the Markov property conditioning on $\mathcal{F}_{\rho^\eta_j(x)}$, the second term can be bounded as follows:

$$\begin{aligned}
&\sum_{j=K}^\infty \inf_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big) \\
&\geq \sum_{j=0}^\infty \bigg(\inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) > \rho^\eta_1(y)\Big)\bigg)^{K+j} = \sum_{j=0}^\infty \bigg(1 - \sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)\bigg)^{K+j} \\
&= \frac{1}{\sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)} \cdot \bigg(1 - \sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)\bigg)^{\lceil \frac{t/\gamma(\eta)}{T/\eta} \rceil}
\end{aligned}$$

$$\geq \frac{1}{\big(1+\delta_{\mathbb{S}}(\epsilon,RT)+o(1)\big)\cdot\gamma(\eta)RT/\eta}\cdot\bigg(1-\big(1+\delta_{\mathbb{S}}(\epsilon,RT)+o(1)\big)\cdot\gamma(\eta)RT/\eta\bigg)^{\frac{t/\gamma(\eta)}{T/\eta}+1}. \tag{4.4}$$

where the last inequality is from (4.2) with $B=\mathbb{S}$. From (4.3), (4.4), and (2.37) of Condition 1, we have

$$\begin{aligned}
&\liminf_{\eta\downarrow 0}\inf_{x\in A(\epsilon)}\mathbf{P}\big(\gamma(\eta)\tau^{\eta}_{I(\epsilon)^c}(x)>t;\ V^{\eta}_{\tau_\epsilon}(x)\in B\big)\\
&\geq \liminf_{\eta\downarrow 0}\frac{C(B^\circ)-\delta_B(\epsilon,T)+o(1)}{\big(1+\delta_{\mathbb{S}}(\epsilon,RT)+o(1)\big)\cdot R}\cdot\bigg(1-\big(1+\delta_{\mathbb{S}}(\epsilon,RT)+o(1)\big)\cdot\gamma(\eta)RT/\eta\bigg)^{\frac{R\cdot t}{\gamma(\eta)RT/\eta}+1}.\\
&\geq \frac{C(B^\circ)-\delta_B(\epsilon,T)}{1+\delta_{\mathbb{S}}(\epsilon,RT)}\cdot\exp\Big(-\big(1+\delta_{\mathbb{S}}(\epsilon,RT)\big)\cdot R\cdot t\Big).
\end{aligned}$$

By taking limit $T\to\infty$ and then considering an R arbitrarily close to 1, it is straightforward to check that the desired lower bound holds.

Moving on to the upper bound, we set $R=1$ and fix an arbitrary $r\in(0,1)$. Set $k=\Big\lfloor\frac{t/\gamma(\eta)}{T/\eta}\Big\rfloor$ and note that

$$\begin{aligned}
\sup_{x\in A(\epsilon)}\mathbf{P}\big(\gamma(\eta)\tau^{\eta}_{I(\epsilon)^c}(x)>t;\ V^{\eta}_{\tau_\epsilon}(x)\in B\big) &= \sup_{x\in A(\epsilon)}\mathbf{P}\big(\tau^{\eta}_{I(\epsilon)^c}(x)>t/\gamma(\eta);\ V^{\eta}_{\tau_\epsilon}(x)\in B\big)\\
&= \underbrace{\sup_{x\in A(\epsilon)}\mathbf{P}\big(\tau^{\eta}_{I(\epsilon)^c}(x)>t/\gamma(\eta)\geq\rho^{\eta}_k(x);\ V^{\eta}_{\tau_\epsilon}(x)\in B\big)}_{\text{(I)}}\\
&\quad+\underbrace{\sup_{x\in A(\epsilon)}\mathbf{P}\big(\tau^{\eta}_{I(\epsilon)^c}(x)>t/\gamma(\eta);\ \rho^{\eta}_k(x)>t/\gamma(\eta);\ V^{\eta}_{\tau_\epsilon}(x)\in B\big)}_{\text{(II)}}
\end{aligned}$$

We first show that (II) vanishes as $\eta\to 0$. Our proof hinges on the following claim:

$$\big\{\tau^{\eta}_{I(\epsilon)^c}(x)>t/\gamma(\eta);\ \rho^{\eta}_k(x)>t/\gamma(\eta)\big\}\ \subseteq\ \bigcup_{j=1}^{k}\big\{\tau^{\eta}_{I(\epsilon)^c}(x)\wedge\rho^{\eta}_j(x)-\rho^{\eta}_{j-1}(x)\geq T/\eta\big\}$$

Proof of the claim: Suppose that $\tau^{\eta}_{I(\epsilon)^c}(x)>t/\gamma(\eta)$ and $\rho^{\eta}_k(x)>t/\gamma(\eta)$. Let $k^*\triangleq\max\{j\geq 1:\rho^{\eta}_j(x)\leq t/\gamma(\eta)\}$. Note that $k^*<k$. We consider two cases separately: (i) $\rho^{\eta}_{k^*}(x)/k^*>(t/\gamma(\eta)-T/\eta)/k^*$ and (ii) $\rho^{\eta}_{k^*}(x)\leq t/\gamma(\eta)-T/\eta$. In case of (i), since $\rho^{\eta}_{k^*}(x)/k^*$ is the average of $\{\rho^{\eta}_j(x)-\rho^{\eta}_{j-1}(x):j=1,\ldots,k^*\}$, there exists $j^*\leq k^*$ such that

$$\rho^{\eta}_{j^*}(x)-\rho^{\eta}_{j^*-1}(x)>\frac{t/\gamma(\eta)-T/\eta}{k^*}\geq\frac{kT/\eta-T/\eta}{k-1}=T/\eta$$

Note that since $\rho^{\eta}_{j^*}(x)\leq\rho^{\eta}_{k^*}(x)\leq t/\gamma(\eta)\leq\tau^{\eta}_{I(\epsilon)^c}(x)$, this proves the claim for case (i). For case (ii), note that

$$\rho^{\eta}_{k^*+1}(x)\wedge\tau^{\eta}_{I(\epsilon)^c}(x)-\rho^{\eta}_{k^*}(x)\geq t/\gamma(\eta)-(t/\gamma(\eta)-T/\eta)=T/\eta,$$

which proves the claim.

Now, with the claim in hand, we have that

$$\begin{aligned}
\text{(II)}&\leq\sum_{j=1}^{k}\sup_{x\in A(\epsilon)}\mathbf{P}\big(\tau^{\eta}_{I(\epsilon)^c}(x)\wedge\rho^{\eta}_j(x)-\rho^{\eta}_{j-1}(x)\geq T/\eta\big)\\
&=\sum_{j=1}^{k}\sup_{x\in A(\epsilon)}\mathbf{E}\Big[\mathbf{P}\big(\tau^{\eta}_{I(\epsilon)^c}(x)\wedge\rho^{\eta}_j(x)-\rho^{\eta}_{j-1}(x)\geq T/\eta\big|\mathcal{F}_{\rho^{\eta}_{j-1}(x)}\big)\Big]
\end{aligned}$$

$$\leq \sum_{j=1}^{k} \sup_{y\in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(y) \wedge \rho^\eta_1(y) \geq T/\eta\big)$$
$$\leq \frac{t}{\gamma(\eta)T/\eta} \cdot \gamma(\eta)T/\eta \cdot o(1) = o(1)$$

for sufficiently large T's, where the last inequality is from the definition of k and (4.1). We are now left with bounding (I) from above.

$$\begin{aligned}
\text{(I)} &= \sup_{x\in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) > t/\gamma(\eta) \geq \rho^\eta_K(x);\ V^\eta_{\tau_\epsilon}(x) \in B\big) \leq \sup_{x\in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_K(x);\ V^\eta_{\tau_\epsilon}(x) \in B\big)\\
&= \sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) \in \big(\rho^\eta_j(x),\, \rho^\eta_{j+1}(x)\big];\ V^\eta_{\tau_\epsilon}(x) \in B\Big)\\
&= \sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{E}\Big[\mathbf{E}\Big[\mathbb{I}\big\{V^\eta_{\tau_\epsilon}(x) \in B\big\} \cdot \mathbb{I}\big\{\tau^\eta_{I(\epsilon)^c}(x) \leq \rho^\eta_{j+1}(x)\big\}\Big|\mathcal{F}_{\rho^\eta_j(x)}\Big] \cdot \mathbb{I}\big\{\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\big\}\Big]\\
&\leq \sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{E}\Big[\sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\ \tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big) \cdot \mathbb{I}\big\{\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\big\}\Big]\\
&= \sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\ \tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big) \cdot \sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big)
\end{aligned}$$

The first term can be bounded via (4.2) with $R = 1$:

$$\begin{aligned}
&\sup_{y\in A(\epsilon)} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(y) \in B;\ \tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)\\
&\leq \big(C(B^-) + \delta_B(\epsilon, T) + o(1)\big) \cdot \gamma(\eta)T/\eta + \frac{1-r}{2} \cdot \gamma(\eta)T/\eta \cdot o(1)
\end{aligned}$$

whereas the second term is bounded via (2.37) of Condition 1 as follows:

$$\begin{aligned}
&\sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big)\\
&\leq \sum_{j=0}^{\infty} \bigg(\sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) > \lfloor rT/\eta \rfloor\Big)\bigg)^{k+j} = \sum_{j=0}^{\infty} \bigg(1 - \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq rT/\eta\Big)\bigg)^{k+j}\\
&\leq \frac{1}{\inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq rT/\eta\Big)} \cdot \bigg(1 - \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq rT/\eta\Big)\bigg)^{\frac{t/\gamma(\eta)}{T/\eta}-1}\\
&= \frac{1}{r \cdot \big(1 - \delta_B(\epsilon, rT) + o(1)\big) \cdot \gamma(\eta)T/\eta} \cdot \Big(1 - r \cdot \big(1 - \delta_B(\epsilon, rT) + o(1)\big) \cdot \gamma(\eta)T/\eta\Big)^{\frac{t}{\gamma(\eta)T/\eta}-1}
\end{aligned}$$

Therefore,

$$\limsup_{\eta\downarrow 0} \sup_{x\in A(\epsilon)} \mathbf{P}\big(\gamma(\eta)\tau^\eta_{I(\epsilon)^c}(x) > t;\ V^\eta_{\tau_\epsilon}(x) \in B\big) \leq \frac{C(B^-) + \delta_B(\epsilon, T)}{r \cdot (1 - \delta_B(\epsilon, rT))} \cdot \exp\Big(-r \cdot \big(1 - \delta_B(\epsilon, rT)\big) \cdot t\Big).$$

Again, taking $T \to \infty$ and considering r arbitrarily close to 1, we can check that the desired upper bound holds. □

Now we are ready to prove Theorem 2.7.

Proof of Theorem 2.7. We first claim that for any $\epsilon, \epsilon' > 0$, $t \geq 0$, and measurable $B \subseteq \mathbb{S}$,

$$\begin{aligned} C(B^\circ) \cdot e^{-t} - \delta_{t,B}(\epsilon) &\leq \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I(\epsilon)^c}(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B\Big) \\ &\leq \limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I(\epsilon)^c}(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B\Big) \leq C(B^-) \cdot e^{-t} + \delta_{t,B}(\epsilon) \end{aligned} \tag{4.5}$$

where $\delta_{t,B}(\epsilon)$ is characterized in Proposition 4.1 such that $\delta_{t,B}(\epsilon) \to 0$ as $\epsilon \to 0$. Now, note that for any measurable $B \subseteq I^c$,

$$\begin{aligned} &\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I^c}(x) > t,\, V^\eta_\tau(x) \in B\Big) \\ &= \underbrace{\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I^c}(x) > t,\, V^\eta_\tau(x) \in B,\, V^\eta_{\tau_\epsilon}(x) \in I\Big)}_{\text{(I)}} + \underbrace{\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I^c}(x) > t,\, V^\eta_\tau(x) \in B,\, V^\eta_{\tau_\epsilon}(x) \notin I\Big)}_{\text{(II)}} \end{aligned}$$

and since

$$\text{(I)} \leq \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(x) \in I\Big) \qquad \text{and} \qquad \text{(II)} = \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B \setminus I\Big),$$

we have that

$$\begin{aligned} \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I^c}(x) > t,\, V^\eta_\tau(x) \in B\Big) &\geq \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B \setminus I\Big) \\ &\geq C\big((B \setminus I)^\circ\big) \cdot e^{-t} - \delta_{t,B\setminus I}(\epsilon) \\ &= C(B^\circ) \cdot e^{-t} - \delta_{t,B\setminus I}(\epsilon) \end{aligned}$$

due to $B \subseteq I^c$, and

$$\begin{aligned} &\limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_{I^c}(x) > t,\, V^\eta_\tau(x) \in B\Big) \\ &\leq \limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B \setminus I\Big) + \limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon')} \mathbf{P}\Big(V^\eta_{\tau_\epsilon}(x) \in I\Big) \\ &\leq C\big((B \setminus I)^-\big) \cdot e^{-t} + \delta_{t,B\setminus I}(\epsilon) + C(I^-) + \delta_{0,I}(\epsilon) \\ &= C(B^-) \cdot e^{-t} + \delta_{t,B\setminus I}(\epsilon) + \delta_{0,I}(\epsilon). \end{aligned}$$

Taking $\epsilon \to 0$, we arrive at the desired lower and upper bounds of the theorem. Now we are left with the proof of the claim 4.5 is true. Note that for any $x \in I$,

$$\begin{aligned} &\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B\Big) \\ &= \mathbf{E}\bigg[\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B \Big| \mathcal{F}_{\tau^\eta_{A(\epsilon)}(x)}\Big) \cdot \Big(\mathbb{I}\big\{\tau^\eta_{A(\epsilon)}(x) \leq T/\eta\big\} + \mathbb{I}\big\{\tau^\eta_{A(\epsilon)}(x) > T/\eta\big\}\Big)\bigg] \end{aligned} \tag{4.6}$$

Fix an arbitrary $s > 0$, and note that from the Markov property,

$$\begin{aligned} &\mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(x) > t,\, V^\eta_{\tau_\epsilon}(x) \in B\Big) \\ &\leq \mathbf{E}\bigg[\sup_{y \in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_\epsilon(y) > t/\gamma(\eta) - T/\eta,\, V^\eta_{\tau_\epsilon}(y) \in B\Big) \cdot \mathbb{I}\big\{\tau^\eta_{A(\epsilon)}(x) \leq T/\eta\big\}\bigg] + \mathbf{P}\Big(\tau^\eta_{A(\epsilon)}(x) > T/\eta\Big) \\ &\leq \sup_{y \in A(\epsilon)} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^\eta_\epsilon(y) > t - s,\, V^\eta_{\tau_\epsilon}(y) \in B\Big) + \mathbf{P}\Big(\tau^\eta_{A(\epsilon)}(x) > T/\eta\Big) \end{aligned}$$

for sufficiently small η's; here, we applied $\gamma(\eta)/\eta \to 0$ as $\eta \downarrow 0$ in the last inequality. In light of (2.40) of Condition 1, by taking $\eta \to 0$ uniformly over $x \in I(\epsilon')$ and then $T \to \infty$ we yield

$$\limsup_{\eta \downarrow 0} \sup_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^{\eta}_{\epsilon}(x) > t,\, V^{\eta}_{\tau_\epsilon}(x) \in B\Big) \leq C(B^-) \cdot e^{-(t-s)} + \delta_{t,B}(\epsilon)$$

Considering an arbitrarily small $s > 0$, we get the upper bound of the claim (4.5). For the loswer bound, again from (4.6) and the Markov property,

$$\begin{aligned}
&\liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon')} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^{\eta}_{\epsilon}(x) > t,\, V^{\eta}_{\tau_\epsilon}(x) \in B\Big) \\
&\geq \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon')} \mathbf{E}\bigg[\inf_{y \in A(\epsilon)} \mathbf{P}\Big(\tau^{\eta}_{\epsilon}(y) > t/\gamma(\eta),\, V^{\eta}_{\tau_\epsilon}(y) \in B\Big) \cdot \mathbb{I}\big\{\tau^{\eta}_{A(\epsilon)}(x) \leq T/\eta\big\}\bigg] \\
&\geq \liminf_{\eta \downarrow 0} \inf_{y \in A(\epsilon)} \mathbf{P}\Big(\gamma(\eta) \cdot \tau^{\eta}_{\epsilon}(y) > t,\, V^{\eta}_{\tau_\epsilon}(y) \in B\Big) \cdot \inf_{x \in I(\epsilon')} \mathbf{P}\big(\tau^{\eta}_{A(\epsilon)}(x) \leq T/\eta\big) \\
&\geq C(B^\circ) - \delta_{t,B}(\epsilon),
\end{aligned}$$

which is the desired lower bound of the claim 4.5. This concludes the proof. □

4.2 Technical Lemmas for measures $\check{\mathbf{C}}^{(k)|b}(\,\cdot\,)$

In order to prove Theorem 2.6, in Section 4.2 we first prepare several technical lemmas that reveal important properties of measure $\check{\mathbf{C}}^{(k)|b}(\cdot)$ defined in (2.34). Throughout this section, we impose Assumptions 1, 2, 3, and 6 for all results derived below. Besides, we fix a few useful constants. For the sake of notation simplicity, for the majority of this section we fix some $b > 0$ such that $s_{\text{left}}/b \notin \mathbb{Z}$ and $s_{\text{right}}/b \notin \mathbb{Z}$. With this, for $r = |s_{\text{left}}| \wedge s_{\text{right}}$ we have $r > (\mathcal{J}^*_b - 1)b$. This allows us to fix, throughout this section, some $\bar{\epsilon} > 0$ small enough such that

$$\bar{\epsilon} \in (0,1), \qquad r > (\mathcal{J}^*_b - 1)b + 3\bar{\epsilon}. \tag{4.7}$$

Next, for any $\epsilon \in (0, \bar{\epsilon})$, let

$$\boldsymbol{t}(\epsilon) \triangleq \min\big\{t \geq 0 :\ \boldsymbol{y}_t(s_{\text{left}} + \epsilon) \in [-\epsilon, \epsilon] \text{ and } \boldsymbol{y}_t(s_{\text{right}} - \epsilon) \in [-\epsilon, \epsilon]\big\} \tag{4.8}$$

for the ODE $\boldsymbol{y}_t(x)$ defined in (2.32), and recall that $I_\epsilon \triangleq (s_{\text{left}} + \epsilon, s_{\text{right}} - \epsilon)$ is the r-shrinkage of set I. Also, we use $I^-_\epsilon = [s_{\text{left}} + \epsilon, s_{\text{right}} - \epsilon]$ to denote the closure of I_ϵ. Then, the definition of $\boldsymbol{t}(\cdot)$ immediately implies

$$\boldsymbol{y}_t(y) \in [-\epsilon, \epsilon] \qquad \forall y \in I^-_\epsilon,\ \forall t \geq \boldsymbol{t}(\epsilon). \tag{4.9}$$

Recall that $I^- = [s_{\text{left}}, s_{\text{right}}]$. The following lemma collects useful properties of the mapping $h^{(k)|b}_{[0,T]}$ defined in (2.15)-(2.17).

Lemma 4.2. *Let Assumptions 2 and 6 hold. Let $\bar{\epsilon} > 0$ be the constant characterized in (4.7). Furthermore, suppose that $\sup_{x \in I^-} |a(x)| \vee |\sigma(x)| \leq C$ for some $C \geq 1$ and $\inf_{x \in I^-} \sigma(x) \geq c$ for some $c \in (0,1]$. (We adopt the convention that $t_0 = 0$.)*

(a) Suppose that $\mathcal{J}^*_b \geq 2$. It holds for all $T > 0$, $x_0 \in [-b-\bar{\epsilon}, b+\bar{\epsilon}]$, $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}^*_b - 2}) \in \mathbb{R}^{\mathcal{J}^*_b - 2}$, and $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}^*_b - 2}) \in (0,T]^{\mathcal{J}^*_b - 2\uparrow}$ that

$$\sup_{t \in [0,T]} |\xi(t)| \leq (\mathcal{J}^*_b - 1)b + \bar{\epsilon} < r - 2\bar{\epsilon} \text{ where } \xi \triangleq h^{(\mathcal{J}^*_b - 2)|b}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t}).$$

(b) *It holds for all* $T > 0$*,* $x_0 \in [-\bar{\epsilon}, \bar{\epsilon}]$*,* $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*-1}) \in \mathbb{R}^{\mathcal{J}_b^*-1}$*, and* $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*-1}) \in (0,T]^{\mathcal{J}_b^*-1\uparrow}$ *that*

$$\sup_{t\in[0,T]} |\xi(t)| \leq (\mathcal{J}_b^* - 1)b + \bar{\epsilon} < r - 2\bar{\epsilon} \text{ where } \xi \triangleq h^{(\mathcal{J}_b^*-1)|b}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t}).$$

(c) *There exist* $\bar{\delta} > 0$ *and* $\bar{t} > 0$ *such that the following claim holds. Let* $T > 0$*,* $x_0 \in [-\bar{\epsilon}, \bar{\epsilon}]$*,* $w_0 \in \mathbb{R}$*,* $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*-1}) \in \mathbb{R}^{\mathcal{J}_b^*-1}$*, and* $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*-1}) \in (0,T]^{\mathcal{J}_b^*-1\uparrow}$*. If*

$$\sup_{t\in[0,T]} |\xi(t)| \geq r - \bar{\epsilon} \text{ where } \xi \triangleq h^{(\mathcal{J}_b^*-1)|b}_{[0,T]}\Big(x_0 + \varphi_b\big(\sigma(x_0)\cdot w_0\big), \boldsymbol{w}, \boldsymbol{t}\Big),$$

then

- $\sup_{t\in[0,t_{\mathcal{J}_b^*-1})} |\xi(t)| \leq (\mathcal{J}_b^* - 1)b + \bar{\epsilon} < r - 2\bar{\epsilon}$*;*
- $|\xi(t_{\mathcal{J}_b^*-1})| \geq r - \bar{\epsilon}$*;*
- $\inf_{t\in[0,t_{\mathcal{J}_b^*-1}]} |\xi(t)| \geq \bar{\epsilon}$*;*
- $|w_j| > \bar{\delta}$ *for all* $j = 0, 1, \cdots, \mathcal{J}_b^* - 1$*;*
- $t_{\mathcal{J}_b^*-1} < \bar{t}$*.*

(d) *Let* $T > 0$*,* $x \in \mathbb{R}, \boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*}) \in \mathbb{R}^{\mathcal{J}_b^*}, \boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*}) \in (0,T]^{\mathcal{J}_b^*\uparrow}$ *and* $\epsilon \in (0, \bar{\epsilon})$*. If* $|\xi(t_1-)| < \epsilon$ *for* $\xi = h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, \boldsymbol{t})$*, then*

$$|\xi(t_{\mathcal{J}_b^*}) - \hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| < \Big[2\exp\big(D(T-t_1)\big)\cdot\Big(1 + \frac{bD}{c}\Big)\Big]^{\mathcal{J}_b^*+1}\cdot\epsilon$$

where $\hat{\xi} = h^{(\mathcal{J}_b^*-1)|b}_{[0,T-t_1]}\big(\varphi_b(\sigma(0)\cdot w_1), (w_2, \cdots, w_{\mathcal{J}_b^*}), (t_2 - t_1, t_3 - t_1, \cdots, t_{\mathcal{J}_b^*} - t_1)\big)$ *and* $D \geq 1$ *is the constant in Assumption 2.*

(e) *Given* $\Delta > 0$*, there exists* $\epsilon_0 = \epsilon_0(\Delta) \in (0, \bar{\epsilon})$ *such that for any* $T > 0$*,* $\theta > r - \bar{\epsilon}$*,* $x \in [-\epsilon_0, \epsilon_0]$*,* $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*}) \in \mathbb{R}^{\mathcal{J}_b^*}$*, and* $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*}) \in (0,T]^{\mathcal{J}_b^*\uparrow}$*,*

$$|\xi(t_{\mathcal{J}_b^*})| \vee |\hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| > \theta \qquad \Longrightarrow \qquad |\hat{\xi}(t_{\mathcal{J}_b^*} - t_1) - \xi(t_{\mathcal{J}_b^*})| < \Delta$$

where $\xi = h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, \boldsymbol{t})$ *and* $\hat{\xi} = h^{(\mathcal{J}_b^*-1)|b}_{[0,T-t_1]}\big(\varphi_b(\sigma(0)\cdot w_1), (w_2, \cdots, w_{\mathcal{J}_b^*}), (t_2 - t_1, t_3 - t_1, \cdots, t_{\mathcal{J}_b^*} - t_1)\big)$*.*

Proof. Before the proof of the claims, we stress that the validity of all claims do not depend on the value of $\sigma(\cdot)$ and $a(\cdot)$ outside of I^-. Take part (a) as an example. Suppose that we can prove part (a) under the stronger assumption that $\sup_{x\in\mathbb{R}} |a(x)| \wedge \sigma(x) \leq C$ for some $C \in [1, \infty)$ and $\inf_{x\in\mathbb{R}} \sigma(x) \geq c$ for some $c \in (0, 1]$. Then due to $\sup_{t\in[0,T]} |\xi(t)| < r = |s_{\text{left}}| \wedge s_{\text{right}}$ for $\xi = h^{(\mathcal{J}_b^*-2)|b}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t})$, we have $\xi(t) \in I^-$ for all $t \in [0, T]$. This implies that part (a) is still valid even if we only have $\sup_{x\in I^-} |a(x)| \wedge \sigma(x) \leq C$ and $\inf_{x\in I^-} \sigma(x) \geq c$. The same applies to all the other claims. Therefore, in the proof below we assume w.l.o.g. that the strong assumptions $\sup_{x\in\mathbb{R}} |a(x)| \wedge \sigma(x) \leq C$ for some $C \in [1, \infty)$ and $\inf_{x\in\mathbb{R}} \sigma(x) \geq c$ for some $c \in (0, 1]$ hold.

(a) The proof hinges on the following observation. For any $j \geq 0, T > 0, x_0 \in \mathbb{R}, \boldsymbol{w} = (w_1, \cdots, w_j) \in \mathbb{R}^j$ and $\boldsymbol{t} = (t_1, \cdots, t_j) \in (0,T]^{j\uparrow}$, let $\xi = h^{(j)|b}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t})$. The condition $a(x)x \leq 0\ \forall x \in (-\gamma, \gamma)$ implies that

$$\frac{d|\xi(t)|}{dt} = -\big|a\big(\xi(t)\big)\big| \quad \forall t \in [0,T] \setminus \{t_1, \cdots, t_j\} \tag{4.10}$$

Specifically, suppose that $\mathcal{J}^*_b \geq 2$. For all $T > 0, x_0 \in [-b-\bar{\epsilon}, b+\bar{\epsilon}], \boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}^*_b - 2}) \in \mathbb{R}^{\mathcal{J}^*_b - 2}$ and $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}^*_b - 2}) \in (0,T]^{\mathcal{J}^*_b - 2\uparrow}$, it holds for $\xi \triangleq h^{(\mathcal{J}^*_b - 2)|b}_{[0,T]}(x_0, \boldsymbol{w}, \boldsymbol{t})$ that $d|\xi(t)|/dt \leq 0$ for any $t \in [0,T] \setminus \{t_1, \cdots, t_{\mathcal{J}^*_b - 2}\}$, thus leading to

$$\begin{aligned}
\sup_{t\in[0,T]} |\xi(t)| &\leq |\xi(0)| + \sum_{t \leq T} |\Delta\xi(t)| \\
&\leq |\xi(0)| + (\mathcal{J}^*_b - 2)b \qquad \text{due to truncation operators } \varphi_b \text{ in } h^{(\mathcal{J}^*_b - 2)|b}_{[0,T]} \\
&\leq b + \bar{\epsilon} + (\mathcal{J}^*_b - 2)b \\
&= (\mathcal{J}^*_b - 1)b + \bar{\epsilon} < r - 2\bar{\epsilon} \qquad \text{due to (4.7).}
\end{aligned}$$

This concludes the proof of part (a).

(b) The proof is almost identical to that of part (a). In particular, it follows from (4.10) that $d|\xi(t)|/dt \leq 0$ for any $t \in [0,T] \setminus \{t_1, \cdots, t_{\mathcal{J}^*_b - 1}\}$. Therefore,

$$\begin{aligned}
\sup_{t\in[0,T]} |\xi(t)| &\leq |\xi(0)| + \sum_{t \leq T} |\Delta\xi(t)| \\
&\leq |\xi(0)| + (\mathcal{J}^*_b - 1)b \quad \text{due to truncation operators } \varphi_b \text{ in } h^{(\mathcal{J}^*_b - 1)|b}_{[0,T]} \\
&\leq \bar{\epsilon} + (\mathcal{J}^*_b - 1)b < r - 2\bar{\epsilon} \qquad \text{due to (4.7).}
\end{aligned}$$

(c) We start from the claim that $\sup_{t\in[0, t_{\mathcal{J}^*_b - 1})} |\xi(t)| < r - 2\bar{\epsilon}$. The case with $\mathcal{J}^*_b = 1$ is trivial since $[0, t_{\mathcal{J}^*_b - 1}) = [0, 0)$ is an empty set. Now consider the case where $\mathcal{J}^*_b \geq 2$. For $\hat{x}_0 \triangleq x_0 + \varphi_b\big(\sigma(x_0)\cdot w_0\big)$, we have $|\hat{x}_0| \leq \bar{\epsilon} + b$. By setting $\hat{\boldsymbol{w}} = (w_1, \cdots, w_{\mathcal{J}^*_b - 2}), \hat{\boldsymbol{t}} = (t_1, \cdots, t_{\mathcal{J}^*_b - 2})$ and $\hat{\xi} = h^{(\mathcal{J}^*_b - 2)|b}_{[0,T]}(\hat{x}_0, \hat{\boldsymbol{w}}, \hat{\boldsymbol{t}})$, we get $\xi(t) = \hat{\xi}(t)$ for all $t \in [0, t_{\mathcal{J}^*_b - 1})$. It then follows directly from results in part (a) that $\sup_{t\in[0, t_{\mathcal{J}^*_b - 1})} |\xi(t)| = \sup_{t\in[0, t_{\mathcal{J}^*_b - 1})} |\hat{\xi}(t)| \leq (\mathcal{J}^*_b - 1)b + \bar{\epsilon} < r - 2\bar{\epsilon}$.

Next, to see why the claim $|\xi(t_{\mathcal{J}^*_b - 1})| \geq r - \bar{\epsilon}$ is true, note that we already know $\sup_{t\in[0,T]} |\xi(t)| \geq r - \bar{\epsilon}$, and we have just shown that $\sup_{t\in[0, t_{\mathcal{J}^*_b - 1})} |\xi(t)| < r - 2\bar{\epsilon}$. Now consider the following proof by contradiction. Suppose that $|\xi(t_{\mathcal{J}^*_b - 1})| < r - \bar{\epsilon}$. Then by definition of the mapping $h^{(\mathcal{J}^*_b - 1)|b}_{[0,T]}$, we know that $\xi(t)$ is continuous on $t \in [t_{\mathcal{J}^*_b - 1}, T]$. Given observation (4.10), we yield the contradiction that $\sup_{t\in[t_{\mathcal{J}^*_b - 1}, T]} |\xi(t)| \leq |\xi(t_{\mathcal{J}^*_b - 1})| \wedge \big(\sup_{t\in[0, t_{\mathcal{J}^*_b - 1})} |\xi(t)|\big) < r - \bar{\epsilon}$. This concludes the proof.

Similarly, to show the claim $\inf_{t\in[0, t_{\mathcal{J}^*_b - 1}]} |\xi(t)| \geq \bar{\epsilon}$ we proceed with a proof by contradiction. Suppose there is some $t \in [0, t_{\mathcal{J}^*_b - 1}]$ such that $|\xi(t)| < \bar{\epsilon}$. Then observation (4.10) implies that

$$\begin{aligned}
|\xi(t_{\mathcal{J}^*_b - 1})| &\leq |\xi(t)| + \sum_{s\in(t, t_{\mathcal{J}^*_b - 1}]} |\Delta\xi(s)| \\
&\leq \bar{\epsilon} + (\mathcal{J}^*_b - 1)b \qquad \text{due to truncation operators } \varphi_b \text{ in } h^{(\mathcal{J}^*_b - 1)|b}_{[0,T]} \\
&< r - 2\bar{\epsilon} \qquad \text{due to (4.7).}
\end{aligned}$$

However, we have just shown that $|\xi(t_{\mathcal{J}^*_b - 1})| \geq r - \bar{\epsilon}$ must hold. With this contradiction established we conclude the proof.

Recall that $C \geq 1$ be the constant satisfying $\sup_{x\in\mathbb{R}} |\sigma(x)| \leq C$. We show that for any $\bar{\delta} > 0$ small enough such that

$$(\mathcal{J}^*_b - 1)b + 3\bar{\epsilon} + C\bar{\delta} < r,$$

we have $|w_j| > \bar{\delta}$ for all $j = 0, 1, \cdots, \mathcal{J}^*_b - 1$. Again, suppose that the claim does not hold. Then there is some $j^* = 0, 1, \cdots, \mathcal{J}^*_b - 1$ with $|w_{j^*}| \leq \bar{\delta}$. From observation (4.10), we get

$$|\xi(t_{\mathcal{J}^*_b - 1})| \leq |\xi(0)| + \sum_{t\in[0, t_{\mathcal{J}^*_b - 1}]} |\Delta\xi(t)|$$

$$\leq |x_0| + \varphi_b\Big(\Big|\sigma(x_0)\cdot w_0\Big|\Big) + \sum_{j=1}^{\mathcal{J}_b^*-1}\varphi_b\Big(\Big|\sigma\big(\xi(t_j-)\big)\cdot w_j\Big|\Big)$$

$$\leq \bar{\epsilon} + (\mathcal{J}_b^* - 1)b + C\bar{\delta} \qquad \text{due to } |x_0|\leq\bar{\epsilon}, |w_{j^*}|\leq\bar{\delta} \text{ and } |\sigma(y)|\leq C \text{ for all } y\in\mathbb{R}$$

$$< r - 2\bar{\epsilon} \qquad \text{due to our choice of } \bar{\delta}.$$

This contradiction with the fact $|\xi(t_{\mathcal{J}_b^*-1})| \geq r - \bar{\epsilon}$ allows us to conclude the proof.

Lastly, we move onto the claim $t_{\mathcal{J}_b^*-1} < \bar{t}$. If $\mathcal{J}_b^* = 1$, then due to $t_0 = 0$ the claim is trivially true for any $\bar{t} > 0$. Now we focus on the case where $\mathcal{J}_b^* \geq 2$ and start by specifying the constant $\bar{t}$. From the continuity of $a(\cdot)$ (see Assumption 2) and the fact that $a(y) \neq 0\ \forall y \in (-r,0)\cup(0,r)$, we can find some $c_{\bar{\epsilon}} > 0$ such that $|a(y)| \geq c_{\bar{\epsilon}}$ for all $y \in [-r+\bar{\epsilon}, -\bar{\epsilon}]\cup[\bar{\epsilon}, r-\bar{\epsilon}]$. Now we pick some

$$t_{\bar{\epsilon}} \triangleq r/c_{\bar{\epsilon}}, \qquad \bar{t} = (\mathcal{J}_b^*-1)\cdot t_{\bar{\epsilon}}.$$

We proceed with a proof by contradiction. Suppose that $t_{\mathcal{J}_b^*-1} \geq \bar{t} = (\mathcal{J}_b^*-1)\cdot t_{\bar{\epsilon}}$, then we can find some $j^* = 1,2,\cdots,\mathcal{J}_b^*-1$ such that $t_{j^*} - t_{j^*-1} \geq t_{\bar{\epsilon}}$. First, recall that we have shown that $|\xi(t_{j^*-1})| < r - \bar{\epsilon}$. Next, note that we must have $|\xi(t)| < \bar{\epsilon}$ for some $t \in [t_{j^*-1}, t_{j^*})$. Indeed, suppose that $|\xi(t)| \geq \bar{\epsilon}$ for all $t \in [t_{j^*-1}, t_{j^*})$. Then from observation (4.10) and the fact that $|a(y)| \geq c_{\bar{\epsilon}}$ for all $y \in [-\gamma, -\bar{\epsilon}]\cup[\bar{\epsilon},\gamma]$, we yield

$$|\xi(t_{j^*}-)| \leq |\xi(t_{j^*-1})| - c_{\bar{\epsilon}}\cdot t_{\bar{\epsilon}} \leq r - c_{\bar{\epsilon}}\cdot\frac{r}{c_{\bar{\epsilon}}} = 0.$$

The continuity of $\xi(t)$ on $t \in [t_{j^*-1}, t_{j^*})$ then implies that for any $t \in [t_{j^*-1}, t_{j^*})$ close enough to t_{j^*}, we have $|\xi(t)| < \bar{\epsilon}$. However, note that we have shown that $\inf_{t\in[0,t_{\mathcal{J}_b^*-1}]}|\xi(t)| \geq \bar{\epsilon}$. With this contradiction established, we conclude the proof.

(d) Let $R_j \triangleq |\xi(t_j) - \hat{\xi}(t_j - t_1)|$ for any $j \in [\mathcal{J}_b^*]$ and $R_0 \triangleq |\xi(t_1) - \hat{\xi}(0)|$. We start by analyzing R_0. First, note that $\xi(t_1) = \xi(t_1-) + \varphi_b\Big(\sigma\big(\xi(t_1-)\big)\cdot w_1\Big)$ and $\hat{\xi}(0) = \varphi_b\big(\sigma(0)\cdot w_1\big)$. Using (4.10), we get $|\xi(t_1-)| \leq |x_0| \leq \epsilon$. As a result,

$$\begin{aligned}
R_0 &\leq \epsilon_0 + \Big|\varphi_b\Big(\sigma\big(\xi(t_1-)\big)\cdot w_1\Big) - \varphi_b\big(\sigma(0)\cdot w_1\big)\Big| \\
&\leq \epsilon + \Big|\sigma\big(\xi(t_1-)\big) - \sigma(0)\Big|\cdot\big|\varphi_{b/c}(w_1)\big| \qquad \text{using (3.23)} \\
&\leq \epsilon + D\epsilon\cdot\frac{b}{c} = (1 + \frac{bD}{c})\cdot\epsilon \qquad \text{because of Assumption 2.}
\end{aligned}$$

We proceed with an induction argument. Suppose that for some $j = 0,1,\cdots,\mathcal{J}_b^*-1$, we have $R_j \leq \rho^{j+1}\cdot\epsilon$ with

$$\rho \triangleq \exp(DT)\cdot\Big(1 + \frac{bD}{c}\Big).$$

Then by applying Gronwall's inequality for $u \in [t_j, t_{j+1})$, we get

$$\sup_{u\in[t_j,t_{j+1})}|\xi(u) - \hat{\xi}(u - t_1)| \leq R_j\cdot\exp\big(D(t_{j+1}-t_j)\big) \leq \exp(DT)R_j.$$

Then at $t = t_{j+1}$ we have (set $\hat{t}_{j+1} \triangleq t_{j+1} - t_1$)

$$\begin{aligned}
R_{j+1} &= |\hat{\xi}(\hat{t}_{j+1}) - \xi(t_{j+1})| \\
&= \Big|\hat{\xi}(\hat{t}_{j+1}-) + \varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big)\cdot w_{j+1}\Big) - \Big[\xi(t_{j+1}-) + \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big)\cdot w_{j+1}\Big)\Big]\Big| \\
&\leq \Big|\hat{\xi}(\hat{t}_{j+1}-) - \xi(t_{j+1}-)\Big| + \Big|\varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big)\cdot w_{j+1}\Big) - \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big)\cdot w_{j+1}\Big)\Big|
\end{aligned}$$

$$
\begin{aligned}
&\le \exp(DT)R_j + \Big|\varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big)\cdot w_{j+1}\Big) - \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big)\cdot w_{j+1}\Big)\Big| \\
&\le \exp(DT)R_j + \Big|\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big) - \sigma\big(\xi(t_{j+1}-)\big)\Big|\cdot\big|\varphi_{b/c}(w_{j+1})\big| \qquad \text{using (3.23)} \\
&\le \exp(DT)R_j + D\big|\hat{\xi}(t_{j+1}-) - \xi(t_{j+1}-)\big|\cdot\frac{b}{c} \qquad \text{due to Assumption 2} \\
&\le \exp(DT)R_j + \frac{bD}{c}\cdot\exp(DT)R_j = \Big(1+\frac{bD}{c}\Big)\exp(DT)R_j \le \rho^{j+2}\cdot\epsilon.
\end{aligned}
$$

By arguing inductively we conclude the proof.

(e) Note that the statement is not affected by the values of ξ outside of $[0, t_{\mathcal{J}_b^*}]$ or the values of $\hat{\xi}$ outside of $[0, t_{\mathcal{J}_b^*} - t_1]$. Therefore, without loss of generality we set $T = t_{\mathcal{J}_b^*} + 1$. Suppose we can show that

$$
|\xi(t_{\mathcal{J}_b^*}) - \hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| < \underbrace{\Big[2\exp\big(D(\bar{t}+1)\big)\cdot\Big(1+\frac{bD}{c}\Big)\Big]^{\mathcal{J}_b^*+1}}_{\triangleq \rho^*}\cdot\epsilon_0 \qquad \forall \epsilon_0 \in (0,\bar{\epsilon}]. \tag{4.11}
$$

Then one can see that part (e) holds for any $\epsilon_0 \in (0,\bar{\epsilon})$ small enough such that $\rho^*\epsilon_0 < \Delta$.

Now it remains to prove claim (4.11). From observation (4.10), we get $|\xi(t_1-)| \le |\xi(0)| = |x| \le \epsilon_0$. This allows us to apply results in part (d) and get (recall our choice of $T = t_{\mathcal{J}_b^*} + 1$)

$$
|\xi(t_{\mathcal{J}_b^*}) - \hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| < \Big[2\exp\big(D(t_{\mathcal{J}_b^*} - t_1 + 1)\big)\cdot\Big(1+\frac{bD}{c}\Big)\Big]^{\mathcal{J}_b^*+1}\cdot\epsilon_0.
$$

Lastly, note that if $|\hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| > \theta > r - \bar{\epsilon}$, then $t_{\mathcal{J}_b^*} - t_1 < \bar{t}$ from part (c). Likewise, from part (c), $|\xi(t_{\mathcal{J}_b^*})| > \theta > r - \bar{\epsilon}$, then $t_{\mathcal{J}_b^*} < \bar{t}$. Therefore, in either case, $t_{\mathcal{J}_b^*} - t_1 + 1 \le \bar{t} + 1$. This concludes the proof. ☐

The following lemmas reveal important properties of the measure $\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}$.

Lemma 4.3. *For any $|\gamma| > (\mathcal{J}_b^* - 1)b + \bar{\epsilon}$ such that $\gamma/b \notin \mathbb{Z}$,*

$$
\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\{\gamma\}) = 0.
$$

Proof. First, consider the case where $\mathcal{J}_b^* = 1$. $\breve{\mathbf{C}}^{(1)|b}(\{\gamma\}) = \nu_\alpha\Big(\big\{w:\ \varphi_b\big(\sigma(0)\cdot w\big) = \gamma\big\}\Big)$. Since $\gamma \ne b$, we know that $\big\{w:\ \varphi_b\big(\sigma(0)\cdot w\big) = \gamma\big\} \subseteq \{\frac{\gamma}{\sigma(0)}\}$. The absolute continuity of ν_α (w.r.t the Lebesgue measure) then implies that $\breve{\mathbf{C}}^{(1)|b}(\{\gamma\}) = 0$.

Now we focus on the case where $\mathcal{J}_b^* \ge 2$. Observe that

$$
\begin{aligned}
&\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\{\gamma\}) \\
&= \int \mathbb{I}\Big(\mathbb{I}\Big\{g^{(\mathcal{J}_b^*-1)|b}\big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*-2}, w^*, t_1, \cdots, t_{\mathcal{J}_b^*-2}, t_{\mathcal{J}_b^*-2} + t^*\big) = \gamma\Big\} \\
&\qquad\qquad \times \nu_\alpha(dw^*)\mathcal{L}(dt^*)\Big)\nu_\alpha^{\mathcal{J}_b^*-1}(dw_1, \cdots, dw_{\mathcal{J}_b^*-1}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-2\uparrow}(dt_1, \cdots, dt_{\mathcal{J}_b^*-2}) \\
&= \int \Big(\int_{(t^*,w^*)\in E(\boldsymbol{w},\boldsymbol{t})} \nu_\alpha(dw^*)\mathcal{L}(dt^*)\Big)\nu_\alpha^{\mathcal{J}_b^*-1}(d\boldsymbol{w}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-2\uparrow}(d\boldsymbol{t})
\end{aligned}
$$

where

$$
\begin{aligned}
E(\boldsymbol{w},\boldsymbol{t}) &= \Big\{(w,t)\in\mathbb{R}\times(0,\infty):\ \varphi_b\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) + \sigma\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\Big)w\Big) = \gamma\Big\}, \\
\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t}) &= \breve{g}^{(\mathcal{J}_b^*-2)|b}\big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*-2}, t_1, \cdots, t_{\mathcal{J}_b^*-2}\big).
\end{aligned}
$$

Here $\boldsymbol{y}_t(x)$ is the ODE defined in (2.32). Furthermore, we claim that for any $\boldsymbol{w}, \boldsymbol{t}$, there exist some continuous function $w^* : (0,\infty) \to \mathbb{R}$ and some $t^* \in (0,\infty)$ such that

$$E(\boldsymbol{w}, \boldsymbol{t}) \subseteq \big\{(w,t) \in \mathbb{R} \times (0,\infty) : w = w^*(t) \text{ or } t = t^*\big\}. \tag{4.12}$$

Then set $E(\boldsymbol{w}, \boldsymbol{t})$ charges zero mass under Lebesgues measure on $\mathbb{R} \times (0,\infty)$. From the absolute continuity of $\nu_\alpha \times \mathcal{L}$ (w.r.t. Lebesgues measure on $\mathbb{R} \times (0,\infty)$) we get $\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\{\gamma\}) = 0$.

Now it only remains to prove claim (4.12). Henceforth in this proof we fix some $\boldsymbol{w} \in \mathbb{R}^{\mathcal{J}_b^*-1}$ and $\boldsymbol{t} \in (0,\infty)^{\mathcal{J}_b^*-2\uparrow}$. We first note that due to $|\gamma| > (\mathcal{J}_b^*-1)b + \bar{\epsilon}$, it follows from part (a) of Lemma 4.2 that $|\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})| \leq (\mathcal{J}_b^*-1)b + \bar{\epsilon} < \gamma$. If $\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t}) = 0$, then $a(0) = 0$ implies that $\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) = 0$ for all $t \geq 0$. Due to the assumption that $\gamma \neq b$, in this case we have $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) - \gamma| = \gamma \neq b$ for all $t \geq 0$. Otherwise, Assumption 6 implies that $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)|$ is monotone decreasing w.r.t. t. Since $|\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})| < \gamma$, we must also have $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)| < \gamma$ for all $t \geq 0$. As a result, for $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) - \gamma| = b$ to hold, we need $\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) = y$ for some $|y| < \gamma$, $|y-\gamma| = b$. There exists at most one y that satisfies this condition: that is, $y = \gamma - b$ if $\gamma > b$, and no solution if $\gamma < b$. Due to the strict monotonicity of $\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)$ w.r.t. t, there exists at most one $t^* = t^*(\boldsymbol{w},\boldsymbol{t})$ such that $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) - \gamma| = b$.

Now for any $t > 0, t \neq t^*$, we know that $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) - \gamma| \neq b$. If there is some $w \in \mathbb{R}$ such that $\varphi_b\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) + \sigma\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\Big)w\Big) = \gamma$, then from the fact that $|\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) - \gamma| \neq b$, the only possible choice for w is $w = \frac{\gamma - \boldsymbol{y}_t\left(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\right)}{\sigma\left(\boldsymbol{y}_t\left(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\right)\right)}$. (Note that this quantity is well-defined due to $\sigma(x) > 0\ \forall x \in \mathbb{R}$; see Assumption 3.) By setting $w^*(t) \triangleq \frac{\gamma - \boldsymbol{y}_t\left(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\right)}{\sigma\left(\boldsymbol{y}_t\left(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\right)\right)}$ we conclude the proof. □

Lemma 4.4. $\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(I^c) \in (0,\infty)$.

Proof. Let $\bar{t}, \bar{\delta}$ be the constants characterized in Lemma 4.2. We start with the proof of finiteness. Recall that $r = |s_{\text{left}}| \wedge s_{\text{right}}$, and observe

$$\begin{aligned}
&\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((-\infty, s_{\text{left}}] \cup [s_{\text{right}}, \infty)\Big) \\
&\leq \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(\mathbb{R} \setminus \big[-(r-\bar{\epsilon}), r-\bar{\epsilon}\big]\Big) \\
&= \int \mathbb{I}\Big\{\Big|\breve{g}^{(\mathcal{J}_b^*-1)|b}\big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), (w_1,\cdots,w_{\mathcal{J}_b^*-1}), (t_1,\cdots,t_{\mathcal{J}_b^*-1})\big)\Big| > \gamma - \bar{\epsilon}\Big\} \\
&\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1,\cdots,dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(dt_1,\cdots,dt_{\mathcal{J}_b^*-1}) \\
&= \int \mathbb{I}\Big\{\Big|h^{(\mathcal{J}_b^*-1)|b}_{[0,1+t_{\mathcal{J}_b^*-1}]}\big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), (w_1,\cdots,w_{\mathcal{J}_b^*-1}), (t_1,\cdots,t_{\mathcal{J}_b^*-1})\big)(t_{\mathcal{J}_b^*-1})\Big| > \gamma - \bar{\epsilon}\Big\} \\
&\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1,\cdots,dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(dt_1,\cdots,dt_{\mathcal{J}_b^*-1}) \\
&\leq \int \mathbb{I}\Big(|w_j| > \bar{\delta}\ \forall j \in [\mathcal{J}_b^*];\ t_{\mathcal{J}_b^*-1} < \bar{t}\Big)\nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(d\boldsymbol{t}) \qquad \text{using part (c) of Lemma 4.2} \\
&\leq \bar{t}^{\mathcal{J}_b^*-1}/\bar{\delta}^{\alpha\mathcal{J}_b^*} < \infty.
\end{aligned}$$

Next, we move onto the proof of the strict positivity. Without loss of generality, assume that $s_{\text{right}} \leq |s_{\text{left}}|$. Then due to $r/b \notin \mathbb{Z}$, we have $(\mathcal{J}_b^*-1)b < s_{\text{right}} < \mathcal{J}_b^* b$. First, consider the case where $\mathcal{J}_b^* = 1$. Then for all $w \geq \frac{b}{\sigma(0)}$ we have $\varphi_b\big(\sigma(0)\cdot w\big) = b > s_{\text{right}}$. Therefore,

$$\breve{\mathbf{C}}^{(1)|b}\Big([s_{\text{right}}, \infty)\Big) = \int \mathbb{I}\big\{\varphi_b\big(\sigma(0)\cdot w\big) \geq s_{\text{right}}\big\}\nu_\alpha(dw) \geq \int_{w \in [\frac{b}{\sigma(0)},\infty)} \nu_\alpha(dw) = \Big(\frac{\sigma(0)}{b}\Big)^\alpha > 0.$$

Now consider the case where $\mathcal{J}_b^* \geq 2$. In particular, we claim the existence of some $(w_1, \cdots, w_{\mathcal{J}_b^*}) \in \mathbb{R}^{\mathcal{J}_b^*}$ and $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*-1}) \in (0,\infty)^{\mathcal{J}_b^*-1\uparrow}$ such that

$$\begin{aligned} &\breve{g}^{(\mathcal{J}_b^*)|b}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1, \cdots, w_{\mathcal{J}_b^*-1}, \boldsymbol{t}\Big) \\ &= h^{(\mathcal{J}_b^*-1)|b}_{[0,t_{\mathcal{J}_b^*-1}+1]}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1, \cdots, w_{\mathcal{J}_b^*-1}, \boldsymbol{t}\Big)(t_{\mathcal{J}_b^*-1}) > s_{\text{right}}. \end{aligned} \tag{4.13}$$

Then from the continuity of mapping $h^{(\mathcal{J}_b^*-1)|b}_{[0,t_{\mathcal{J}_b^*-1}+1]}$ (see Lemma 3.10), we can fix some $\Delta > 0$ such that for all w_j' such that $|w_j' - w_j| < \Delta$ and $|t_j' - t_j| < \Delta$, we have

$$\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b(\sigma(0)\cdot w'_{\mathcal{J}_b^*}), w_1', \cdots, w'_{\mathcal{J}_b^*-1}, t_1', \cdots, t'_{\mathcal{J}_b^*-1}\Big) > s_{\text{right}}.$$

Then we can conclude the proof with

$$\begin{aligned} &\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big([s_{\text{right}},\infty)\Big) \\ &\geq \int \mathbb{I}\Big\{|w_j' - w_j| < \Delta\ \forall j \in [\mathcal{J}_b^*];\ |t_j' - t_j| < \Delta\ \forall j \in [\mathcal{J}_b^*-1]\Big\} \\ &\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1', \cdots, dw'_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1}(dt_1', \cdots, dt'_{\mathcal{J}_b^*-1}) \\ &> 0. \end{aligned}$$

It only remains to show (4.13). By Assumptions 2 and 3, we can fix some $C_0 > 0$ such that $|a(x)| \leq C_0$ for all $x \in [s_{\text{left}}, s_{\text{right}}]$, as well as some $c > 0$ such that $\inf_{x\in[s_{\text{left}},s_{\text{right}}]}\sigma(x) \geq c$. Now set $w_1 = \cdots = w_{\mathcal{J}_b^*} = b/c$, Also, pick some $\Delta > 0$ and set $t_k = k\Delta$ (with convention $t_0 \triangleq 0$). For $\xi \triangleq h^{(\mathcal{J}_b^*-1)|b}_{[0,t_{\mathcal{J}_b^*-1}+1]}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1, \cdots, w_{\mathcal{J}_b^*-1}, t_1, \cdots, t_{\mathcal{J}_b^*-1}\Big)$, note that part (c) of Lemma 4.2 implies $\sup_{t\in[0,t_{\mathcal{J}_b^*-1})}|\xi(t)| \leq (\mathcal{J}_b^*-1)b + \bar{\epsilon}$, so we must have $\xi(t) \in [s_{\text{left}}, s_{\text{right}}]$ for all $t < t_{\mathcal{J}_b^*-1}$. This implies $\big|a\big(\xi(t)\big)\big| \leq C_0$ for all $t < t_{\mathcal{J}_b^*-1}$. Now we make a few observations. First, at $t_0 = 0$ we have $\xi(0) = \varphi_b\big(\sigma(0)\cdot w_{\mathcal{J}_b^*}\big) = b$ due to $\sigma(0)\cdot w_{\mathcal{J}^*} \geq c \cdot \frac{b}{c} = b$. Also, note that (for any $j = 1, 2, \cdots, \mathcal{J}_b^*-1$)

$$\begin{aligned} \xi(t_j) &= \xi(t_{j-1}) + \int_{s\in[t_{j-1},t_j)} a\big(\xi(s)\big)ds + \varphi_b\big(\sigma(\xi(t_j-)\cdot w_j\big) \\ &= \xi(t_{j-1}) + \int_{s\in[t_{j-1},t_j)} a\big(\xi(s)\big)ds + b \quad \text{ due to } \sigma\big(\xi(t_j-)\big)\cdot w_j \geq c\cdot\frac{b}{c} = b \\ &\geq \xi(t_{j-1}) - C_0\cdot(t_j - t_{j-1}) + b \quad \text{because of } a(x)x \leq 0 \text{ (see Assumption 6) and } \big|a\big(\xi(t)\big)\big| \leq C_0 \\ &= \xi(t_{j-1}) - C_0\Delta + b. \end{aligned}$$

By arguing inductively, we get

$$\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1, \cdots, w_{\mathcal{J}_b^*-1}, \boldsymbol{t}\Big) = \xi(t_{\mathcal{J}_b^*-1}) \geq \mathcal{J}_b^* b - (\mathcal{J}_b^*-1)C_0\Delta.$$

Due to $\mathcal{J}_b^* b > s_{\text{right}}$, it then holds for all $\Delta > 0$ small enough that $\mathcal{J}_b^* b - (\mathcal{J}_b^*-1)C_0\Delta > s_{\text{right}}$. This concludes the proof. □

Lemma 4.5. *Let $\bar{t}, \bar{\delta}$ be the constants characterized in Lemma 4.2. Given $\Delta \in (0, \bar{\epsilon}/2)$, there exists $\epsilon_0 > 0$ such that for any $\epsilon \in (0, \epsilon_0]$, $T \geq \bar{t}$, and measurable $B \subseteq (I_{\bar{\epsilon}/2})^c$,*

$$\begin{aligned} (T - \bar{t})\cdot\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B_\Delta\Big) &\leq \inf_{x\in[-\epsilon,\epsilon]} \mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\bigg(\Big(\breve{E}(\epsilon,B,T)\Big)^\circ;\ x\bigg) \\ &\leq \sup_{x\in[-\epsilon,\epsilon]} \mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\bigg(\Big(\breve{E}(\epsilon,B,T)\Big)^-;\ x\bigg) \leq T\cdot\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^\Delta\Big) + \big(\bar{t}/\bar{\delta}^\alpha\big)^{\mathcal{J}_b^*}. \end{aligned}$$

where $\breve{E}(\epsilon,B,T) \triangleq \Big\{\xi \in \mathbb{D}[0,T]:\ \exists t \leq T \text{ s.t. } \xi(t) \in B \text{ and } \xi(s) \in I_\epsilon\ \forall s \in [0,t)\Big\}$.

Proof. Using part (e) of Lemma 4.2, for the fixed $\Delta > 0$ we can fix some $\epsilon_0 \in (0, \Delta/2)$ such that the following claim holds (recall that $r = |s_{\text{left}}| \wedge s_{\text{right}}$): For any $T > 0$, $x \in [-\epsilon_0, \epsilon_0]$, $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*}) \in \mathbb{R}^{\mathcal{J}_b^*}$, and $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}_b^*}) \in (0, T]^{\mathcal{J}_b^* \uparrow}$,

$$|\xi(t_{\mathcal{J}_b^*})| \vee |\hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| > r - \bar{\epsilon} \qquad \Longrightarrow \qquad |\hat{\xi}(t_{\mathcal{J}_b^*} - t_1) - \xi(t_{\mathcal{J}_b^*})| < \Delta/2 \tag{4.14}$$

where $\xi = h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, \boldsymbol{t})$ and $\breve{g}^{(\mathcal{J}_b^*-1)|b}\big(\varphi_b(\sigma(0) \cdot w_1), w_2, \cdots, w_{\mathcal{J}_b^*}, t_2 - t_1, t_3 - t_1, \cdots, t_{\mathcal{J}_b^*} - t_1\big)$.

Henceforth in the proof we fix some $\epsilon \in (0, \epsilon_0]$ and $B \subseteq (I_{\bar{\epsilon}/2})^c$. To prove the upper bound, we start with the following observation. For any $\xi \in \breve{E}(\epsilon, B, T)$ and any ξ' such that $\boldsymbol{d}_{J_1,[0,T]}(\xi, \xi') < \epsilon$, due to $\epsilon \leq \epsilon_0 < \Delta/2$, we can find some $t' \in [0, T]$ such that $\xi'(t') \in B^{\Delta/2}$. This implies

$$\big(\breve{E}(\epsilon, B, T)\big)^- \subseteq \big(\breve{E}(\epsilon, B, T)\big)^\epsilon \subseteq \Big\{\xi \in \mathbb{D}[0, T] : \ \xi(t) \in B^{\Delta/2} \text{ for some } t \in [0, T]\Big\}.$$

By definition of the measure $\mathbf{C}^{(k)|b}_{[0,T]}$ in (2.9),

$$\begin{aligned}
&\mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\Big(\big(\breve{E}(\epsilon, B, T)\big)^-; \ x\Big) \\
&\leq \int \mathbb{I}\Big\{\exists t \in [0, T] \text{ s.t. } h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, \boldsymbol{t})(t) \in B^{\Delta/2}\Big\} \ \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_T^{\mathcal{J}_b^* \uparrow}(d\boldsymbol{t}) \\
&\text{(by setting } u_j \triangleq t_j - t_1 \text{ for all } j = 2, 3, \cdots, \mathcal{J}_b^*) \\
&= \int \bigg(\int \mathbb{I}\Big\{\exists t \in [0, T] \text{ s.t. } h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(t) \in B^{\Delta/2}\Big\} \\
&\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \bigg) \mathcal{L}(dt_1) \\
&= \int \phi_B(t_1, x) \mathcal{L}_T(dt_1)
\end{aligned} \tag{4.15}$$

where

$$\begin{aligned}
\phi_B(t_1, x) = \int \mathbb{I}\Big\{&\exists t \in [0, T] \text{ s.t. } h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(t) \in B^{\Delta/2}\Big\} \\
&\times \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}).
\end{aligned}$$

For any $x \in [-\epsilon_0, \epsilon_0]$, note that $\boldsymbol{y}_t(x) \in [-\epsilon_0, \epsilon_0]$ $\forall t \geq 0$. Also, note that due to $B \subseteq (I_{\bar{\epsilon}/2})^c$, we have $\inf_{w \in B} |w| \geq r - \bar{\epsilon}/2$. Because of our choice of $\Delta \in (0, \bar{\epsilon}/2)$, we then have $\inf_{w \in B^{\Delta/2}} |w| > r - \bar{\epsilon}$. Using property (4.14), for all $t_1 \in (0, T]$ and $x \in [-\epsilon_0, \epsilon_0]$ we have the upper bound

$$\begin{aligned}
\phi_B(t_1, x) \leq \int \mathbb{I}\Big\{&\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0) \cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B^\Delta\Big\} \\
&\times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*})
\end{aligned} \tag{4.16}$$

due to part (c) of Lemma 4.2. In particular, if we only consider $t_1 \in (0, T - \bar{t})$, then for any $x \in [-\epsilon_0, \epsilon_0]$ it follows from (4.16) that

$$\begin{aligned}
\phi_B(t_1, x) &\leq \int \mathbb{I}\Big\{\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0) \cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B^\Delta\Big\} \\
&\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\text{due to } T - t_1 > \bar{t} \text{ (from } t_1 \in (0, T - \bar{t})) \text{ and } u_{\mathcal{J}^*} < \bar{t} \text{ (see part (c) of Lemma 4.2)} \\
&= \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\Delta).
\end{aligned}$$

On the other hand, for all $t_1 \in [T-\bar{t}, T)$ and $x \in [-\epsilon_0, \epsilon_0]$, from (4.16) we get

$$\begin{aligned}
\phi_B(t_1, x) &\leq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*} \Big) \in B^{\Delta} \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\leq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*} \Big) \in B^{\Delta} \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to } T - t_1 \leq \bar{t} \\
&\leq \int \mathbb{I}\Big\{ \Big| \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*} \Big) \Big| > r - \bar{\epsilon} \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to } \Delta < \bar{\epsilon} \text{ and recall } r = |s_{\text{left}}| \wedge s_{\text{right}} \\
&\leq \int \mathbb{I}\Big\{ |w_j| > \bar{\delta}\ \forall j \in [\mathcal{J}_b^*] \Big\} \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to part (c) of Lemma 4.2} \\
&\leq (1/\bar{\delta})^{\alpha \mathcal{J}_b^*} \cdot \bar{t}^{\mathcal{J}_b^*-1}. \qquad (4.17)
\end{aligned}$$

Therefore, in (4.15) we obtain (for all $x \in [-\epsilon_0, \epsilon_0]$)

$$\begin{aligned}
\int \phi_B(t_1, x)\mathcal{L}_T(dt_1) &= \int_{t_1 \in (0, T-\bar{t})} \phi_B(t_1, x)\mathcal{L}_T(dt_1) + \int_{t_1 \in [T-\bar{t}, T)} \phi_B(t_1, x)\mathcal{L}_T(dt_1) \\
&\leq (T - \bar{t}) \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^{\Delta}) + \bar{t} \cdot (1/\bar{\delta}^{\alpha})^{\mathcal{J}_b^*} \cdot \bar{t}^{\mathcal{J}_b^*-1} \\
&\leq T \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^{\Delta}) + (\bar{t}/\bar{\delta}^{\alpha})^{\mathcal{J}_b^*}
\end{aligned}$$

and conclude the proof of the upper bound.

The proof of the lower bound is almost identical. Specifically, let $\widetilde{E} = \big\{ \xi \in \mathbb{D}[0,T] :\ \exists t \in [0,T] \text{ s.t. } \xi(t) \in B_{\Delta/2},\ \xi(s) \in I_{2\epsilon}\ \forall s \in [0,t) \big\}$. For any $\xi \in \widetilde{E}$ and any ξ' with $\boldsymbol{d}_{J_1,[0,T]}(\xi, \xi') < \epsilon$, due to $\epsilon \leq \epsilon_0 < \Delta/2$ there must be some $t' \in [0,T]$ such that $\xi'(t') \in B$ and $\xi'(s) \in I_\epsilon\ \forall s \in [0, t')$. This implies

$$\Big\{ \xi \in \mathbb{D}[0,T] :\ \exists t \in [0,T] \text{ s.t. } \xi(t) \in B_{\Delta/2},\ \xi(s) \in I_{2\epsilon}\ \forall s \in [0,t) \Big\} \subseteq \big(\breve{E}(\epsilon, B, T) \big)_\epsilon \subseteq \big(\breve{E}(\epsilon, B, T) \big)^{\circ}.$$

As a result,

$$\begin{aligned}
&\mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\Big(\big(\breve{E}(\epsilon, B, T) \big)^{\circ};\ x \Big) \\
&\geq \int \mathbb{I}\Big\{ \exists t \in [0,T] \text{ s.t. } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, \boldsymbol{t})(t) \in B_{\Delta/2} \text{ and } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, \boldsymbol{t})(s) \in I_{2\epsilon}\ \forall s \in [0,t) \Big\}\ \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_T^{\mathcal{J}_b^*\uparrow}(d\boldsymbol{t}) \\
&= \int \widetilde{\phi}_B(t_1, x)\mathcal{L}_T(dt_1)
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\phi}_B(t_1, x) = \int \mathbb{I}\Big\{ &\exists t \in [0,T] \text{ s.t. } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(t) \in B_{\Delta/2} \\
&\text{and } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(s) \in I_{2\epsilon}\ \forall s \in [0,t) \Big\}
\end{aligned}$$

$$\times\, \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2,\cdots,du_{\mathcal{J}_b^*}).$$

Analogous to the argument for (4.16), using property (4.14) we yield that for all $t \in (0, T-\bar{t})$:

$$\begin{aligned}\phi_B(t_1,x) &\geq \int \mathbb{I}\Big\{\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2,\cdots,w_{\mathcal{J}_b^*}, u_2,\cdots,u_{\mathcal{J}_b^*}\Big) \in B_\Delta\Big\} \\ &\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1,\cdots,dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2,\cdots,du_{\mathcal{J}_b^*}) \\ &= \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B_\Delta).\end{aligned}$$

due to part (c) of Lemma 4.2 again. To avoid repetitions we omit the details here. □

Lemma 4.6. *Let $\bar{\epsilon} \in (0,b)$ be defined as in (4.7). Let positive integer k, open interval $S \subseteq \mathbb{R}$, and $b > 0$ be such that $d_S \geq k$ and $r_S - (d_S - 1)\cdot b > \bar{\epsilon}$ where*

$$r_S \triangleq \inf\{|x| : \ x \in S\}, \qquad d_S \triangleq \lceil r_S/b \rceil.$$

Then

$$\breve{\mathbf{C}}^{(k)|b}(S) > 0 \qquad \Longleftrightarrow \qquad d_S = k.$$

Proof. We first prove that $\breve{\mathbf{C}}^{(k)|b}(S) > 0 \Longrightarrow d_S = k$. By definition of $\breve{\mathbf{C}}^{(k)|b}$ in (2.34), there must be some $w_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1,\cdots,w_{k-1}) \in \mathbb{R}^{k-1}$, and $\boldsymbol{t} = (t_1,\cdots,t_{k-1}) \in (0,\infty)^{(k-1)\uparrow}$ such that (let $T = t_{k-1}+1$)

$$h_{[0,T]}^{(k-1)|b}\big(\varphi_b(\sigma(0)\cdot w_0), \boldsymbol{w}, \boldsymbol{t}\big)(t_{k-1}) \in S. \tag{4.18}$$

However, part (a) of Lemma 4.2 implies that $\big|h_{[0,T]}^{(k-1)|b}\big(\varphi_b(\sigma(0)\cdot w_0), \boldsymbol{w}, \boldsymbol{t}\big)(t)\big| < (k-1)\cdot b + \bar{\epsilon}$ for all $t \in [0, t_{k-1})$. Therefore,

$$\begin{aligned} r_S \leq \big|h_{[0,T]}^{(k-1)|b}\big(\varphi_b(\sigma(0)\cdot w_0), \boldsymbol{w}, \boldsymbol{t}\big)(t_{k-1})\big| &\leq \big|h_{[0,T]}^{(k-1)|b}\big(\varphi_b(\sigma(0)\cdot w_0), \boldsymbol{w}, \boldsymbol{t}\big)(t_{k-1}-)\big| + b \\ &\leq k\cdot b + \bar{\epsilon}.\end{aligned}$$

This leads to $r_S/b < k+1$, and hence $d_S = k$ or $k+1$. Furthermore, suppose that $d_S = k+1$. Then $r_S \leq k\cdot b + \bar{\epsilon}$ immediately contradicts the assumption $r_S - (d_S-1)\cdot b = r_S - k\cdot b > \bar{\epsilon}$. This concludes the proof of $d_S = k$.

Next, we prove that $d_S = k \Longrightarrow \breve{\mathbf{C}}^{(k)|b}(S) > 0$. In particular, suppose that we can find some $w_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1,\cdots,w_{k-1}) \in \mathbb{R}^{k-1}$, and $\boldsymbol{t} = (t_1,\cdots,t_{k-1}) \in (0,\infty)^{(k-1)\uparrow}$ such that (4.18) holds under the choice of $T = t_{k-1}+1$. Then from the continuity of mapping $h_{[0,T]}^{(k)|b}$ (see Lemma 3.10), one can find some $\Delta > 0$ small enough such that

$$S \supseteq \Big\{(w_0', \boldsymbol{w}', \boldsymbol{t}') \in \mathbb{R}\times\mathbb{R}^{k-1}\times(0,T)^k : \ |w_0' - w_0| < \Delta;\ \max_{i\in[k-1]} |w_i' - w_i| \vee |t_i' - t_i| < \Delta\Big\}.$$

Note that for $\Delta > 0$ small enough, we can ensure that $\boldsymbol{t}' = (t_1',\cdots,t_{k-1}') \in (0,T)^{(k-1)\uparrow}$ if $\max_{i\in[k-1]}|t_i' - t_i| < \Delta$ (that is, $\boldsymbol{t}'$ is still strictly increasing). Therefore, $\breve{\mathbf{C}}^{(k)|b}(S) \geq \Big(\prod_{i\in[k-1]}\int_{(t_i-\Delta,t_i+\Delta)}\mathcal{L}(dt)\Big)\cdot\Big(\prod_{i=0,1,\cdots,k-1}\int_{(w_i-\Delta,w_i+\Delta)}\nu_\alpha(dw)\Big) > 0$.

Now, it suffices to find some $w_0 \in \mathbb{R}$, $\boldsymbol{w} = (w_1,\cdots,w_{k-1}) \in \mathbb{R}^{k-1}$, and $\boldsymbol{t} = (t_1,\cdots,t_{k-1}) \in (0,\infty)^{(k-1)\uparrow}$ such that (4.18) holds. Due to $r_S - (d_S-1)\cdot b > \bar{\epsilon}$ we know that $r_S > 0$, which implies $0 \notin S$. W.l.o.g. we assume that the open interval S is on the R.H.S. of the origin. First, due to $d_S = k$, we can find some $\delta > 0$ and $x \in S$ such that $x < kb + \delta$. Next, let $t_i = \Delta\cdot i$ for some $\Delta > 0$.

By Assumption 3, we can fix some constant $c > 0$ such that $\inf_{x \in [s_{\text{left}}, s_{\text{right}}]} \sigma(x) \geq c$. Also, we set $w_i = b/c$ for all $i = 0, 1, \cdots, k-2$. By picking $\Delta > 0$ small enough we can ensure that

$$x_{k-1} \triangleq h^{(k-1)|b}_{[0,T]}(\varphi_b(\sigma(0) \cdot w_0), \boldsymbol{w}, \boldsymbol{t})(t_{k-1}-) > (k-1) \cdot b - \delta.$$

Lastly, note that $h^{(k-1)|b}_{[0,T]}(\varphi_b(\sigma(0) \cdot w_0), \boldsymbol{w}, \boldsymbol{t})(t_{k-1}) = x_{k-1} + \varphi_b(\sigma(x_{k-1}) \cdot w_{k-1})$, and $x - x_{k-1} < b$ due to $x_{k-1} > (k-1) \cdot b - \delta$ and $x < kb - \delta$. By setting $w_{k-1} = (x - x_{k-1})/\sigma(x_{k-1})$, we yield $h^{(k-1)|b}_{[0,T]}(\varphi_b(\sigma(0) \cdot w_0), \boldsymbol{w}, \boldsymbol{t})(t_{k-1}) = x \in S$ and conclude the proof. □

4.3 Proof of Theorem 2.6

In this section, we apply the framework developed in Section 2.3.2. and prove Theorem 2.6. Analogous to Section 4.2, we impose Assumptions 1, 2, 3, and 6 and adopt the choices of $\bar{\epsilon} > 0$, $r > 0$, and $\boldsymbol{t}(\epsilon)$ in (4.7), (4.8), and (4.8) throughout this section.

Let us consider a specialized version of Condition 1 where $\mathbb{S} = \mathbb{R}$, $A(\epsilon) = (-\epsilon, \epsilon)$, $I = (s_{\text{left}}, s_{\text{right}})$, and $I(\epsilon)$ is set to be $I_\epsilon = (s_{\text{left}} + \epsilon, s_{\text{right}} - \epsilon)$. Let $V^\eta_j(x) = X^{\eta|b}_j(x)$. Meanwhile, recall that $C^*_b = \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\big(I^c\big)$, and it has been established in Lemma 4.4 that $C^*_b \in (0, \infty)$. Now, recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$ and $\lambda(\eta) = \eta^{-1} H(\eta^{-1})$, and set

$$C(\ \cdot\ ) \triangleq \frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}(\ \cdot\ \setminus I)}{C^*_b}, \qquad \gamma(\eta) \triangleq C^*_b \cdot \eta \cdot \big(\lambda(\eta)\big)^{\mathcal{J}^*_b}.$$

Note that $\partial I = \{s_{\text{left}}, s_{\text{right}}\}$ and recall our assumption $s_{\text{left}}/b \notin \mathbb{Z}$ and $s_{\text{right}}/b \notin \mathbb{Z}$. Also, our choice of constant $\bar{\epsilon}$ in (4.7) ensures that $|s_{\text{left}}| \wedge s_{\text{right}} > (\mathcal{J}^*_b - 1) \cdot b + \bar{\epsilon}$. Lemma 4.3 then verifies $\check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}(\partial I) = 0$ and hence $C(\partial I) = 0$. Besides, note that $\gamma(\eta) T/\eta = C^*_b T \cdot \big(\lambda(\eta)\big)^{\mathcal{J}^*_b}$.

We start by establishing conditions (2.37) and (2.38). First, given any $B \subseteq \mathbb{R}$ we specify the choice of function $\boldsymbol{\delta}_B(\epsilon, T)$ in Condition 1. From the continuity of measures, we get $\lim_{\Delta \downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B^\Delta \cap I^c) \setminus (B^- \cap I^c)\Big) = 0$ and $\lim_{\Delta \downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B^\circ \cap I^c) \setminus (B_\Delta \cap I^c)\Big) = 0$. This allows us to fix a sequence $(\Delta^{(n)})_{n \geq 1}$ such that $\Delta^{(n+1)} \in (0, \Delta^{(n)}/2)$ and

$$\check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B^\Delta \cap I^c) \setminus (B^- \cap I^c)\Big) \vee \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B^\circ \cap I^c) \setminus (B_\Delta \cap I^c)\Big) \leq 1/2^n \tag{4.19}$$

for each $n \geq 1$. Next, recall the definition of set $\breve{E}(\epsilon, B, T)$ in Lemma 4.5, and let $\widetilde{B}(\epsilon) \triangleq B \setminus I_\epsilon$. Using Lemma 4.5, we are able to fix another sequence $(\epsilon^{(n)})_{n \geq 1}$ such that $\epsilon^{(n)} \in (0, \bar{\epsilon}]\ \forall n \geq 1$ and for any $n \geq 1$, $\epsilon \in (0, \epsilon^{(n)}]$, we have

$$\sup_{x \in [-\epsilon, \epsilon]} \mathbf{C}^{(\mathcal{J}^*_b)|b}_{[0,T]}\Big(\big(\breve{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^-;\ x\Big) \leq T \cdot \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B \setminus I_\epsilon)^{\Delta^{(n)}}\Big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}^*_b}, \tag{4.20}$$

$$\inf_{x \in [-\epsilon, \epsilon]} \mathbf{C}^{(\mathcal{J}^*_b)|b}_{[0,T]}\Big(\big(\breve{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^\circ;\ x\Big) \geq (T - \bar{t}) \cdot \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((B \setminus I_\epsilon)_{\Delta^{(n)}}\Big). \tag{4.21}$$

Given any $\epsilon \in (0, \epsilon^{(1)}]$, there uniquely exists some $n = n_\epsilon \geq 1$ such that $\epsilon \in (\epsilon^{(n+1)}, \epsilon^{(n)}]$. This allows us to set

$$\begin{aligned}\breve{\boldsymbol{\delta}}_B(\epsilon, T) = T \cdot \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big(B^{\Delta^{(n)}} \setminus B^-\Big) \vee \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big(B^\circ \setminus B_{\Delta^{(n)}}\Big) \vee \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((\partial I)^{\epsilon + \Delta^{(n)}}\Big) \\ + \bar{t} \cdot \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\big(B^\circ \setminus I\big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}^*_b}\end{aligned} \tag{4.22}$$

and $\boldsymbol{\delta}_B(\epsilon, T) = \breve{\boldsymbol{\delta}}_B(\epsilon, T)/(C^*_b \cdot T)$. First, due to (4.19), we get

$$\lim_{T \to \infty} \boldsymbol{\delta}_B(\epsilon, T) \leq \frac{1}{C^*_b} \cdot \Big[\frac{1}{2^{n_\epsilon}} \vee \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}\Big((\partial I)^{\epsilon + \Delta^{(n_\epsilon)}}\Big)\Big]$$

where n_ϵ is the unique positive integer satisfying $\epsilon \in (\epsilon^{(n+1)}, \epsilon^{(n)}]$. Moreover, as $\epsilon \downarrow 0$ we get $n_\epsilon \to \infty$. Since ∂I is closed, we get $\cap_{r>0}(\partial I)^r = \partial I$, which then implies $\lim_{r\downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^r\Big) = \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(\partial I\Big) = 0$ due to continuity of measures. In summary, we have verified that $\lim_{\epsilon\downarrow 0}\lim_{T\to\infty} \boldsymbol{\delta}_B(\epsilon,T) = 0$.

Now, we are ready to verify conditions (2.37) and (2.38). Specifically, we introduce stopping times

$$\tau_\epsilon^{\eta|b}(x) \triangleq \min\big\{j \geq 0 :\ X_j^{\eta|b}(x) \notin I_\epsilon\big\}. \tag{4.23}$$

Lemma 4.7 (Verifying conditions (2.37) and (2.38))**.** *Let $\bar{t}$ be characterized as in Lemma 4.2. Given any measurable $B \subseteq \mathbb{R}$, any $\epsilon > 0$ small enough, and any $T > \bar{t}$,*

$$\begin{aligned} C(B^\circ) - \boldsymbol{\delta}_B(\epsilon,T) &\leq \liminf_{\eta\downarrow 0}\ \inf_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\gamma(\eta)T/\eta} \\ &\leq \limsup_{\eta\downarrow 0}\ \sup_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\gamma(\eta)T/\eta} \leq C(B^-) + \boldsymbol{\delta}_B(\epsilon,T). \end{aligned}$$

Proof. Recall that $\gamma(\eta)T/\eta = C_b^* T\cdot\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}$, $C(\cdot) = \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\,\cdot\setminus I)/C_b^*$, and $\boldsymbol{\delta}_B(\epsilon,T) = \check{\boldsymbol{\delta}}_B(\epsilon,T)/(C_b^*\cdot T)$. By rearranging the terms, it suffices to show that

$$\limsup_{\eta\downarrow 0}\ \sup_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} \leq T\cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^-\setminus I) + \check{\boldsymbol{\delta}}_B(\epsilon,T), \tag{4.24}$$

$$\liminf_{\eta\downarrow 0}\ \inf_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} \geq T\cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\circ\setminus I) - \check{\boldsymbol{\delta}}_B(\epsilon,T). \tag{4.25}$$

To proceed, recall the definition of set $\check{E}(\epsilon,B,T)$ in Lemma 4.5. Let $\widetilde{B}(\epsilon) \triangleq B\setminus I_\epsilon$. Note that

$$\Big\{\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big\} = \Big\{\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in \widetilde{B}(\epsilon)\Big\} = \Big\{\boldsymbol{X}^{\eta|b}_{[0,T]}(x) \in \check{E}(\epsilon,\widetilde{B}(\epsilon),T)\Big\}.$$

For any $\epsilon \in (0,\bar{\epsilon})$ and $\xi \in \check{E}(\epsilon,\widetilde{B}(\epsilon),T)$, there is $t \in [0,T]$ such that $\xi(t) \notin I_\epsilon$ and hence $|\xi(t)| \geq r-\epsilon > r-\bar{\epsilon}$. On the other hand, using part (b) of Lemma 4.2, it holds for all $\xi \in \mathbb{D}^{(\mathcal{J}_b^*-1)|b}_{[-\epsilon,\epsilon]}[0,T]$ that $\sup_{t\in[0,T]}|\xi(t)| < r - 2\bar{\epsilon}$. In summary, we have established that

$$\boldsymbol{d}_{J_1,[0,T]}\Big(\check{E}(\epsilon,\widetilde{B}(\epsilon),T),\ \mathbb{D}^{(\mathcal{J}_b^*-1)|b}_{[-\epsilon,\epsilon]}[0,T]\Big) > 0$$

for all $\epsilon > 0$ small enough. Now let $n = n_\epsilon$ be the unique positive integer such that $\epsilon \in (\epsilon^{(n+1)}, \epsilon^{(n)}]$. It follows from Theorem 2.3 that

$$\begin{aligned} \limsup_{\eta\downarrow 0}\ \sup_{x\in[-\epsilon,\epsilon]} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} &\leq \sup_{x\in[-\epsilon,\epsilon]} \mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\Big(\big(\check{E}(\epsilon,\widetilde{B}(\epsilon),T)\big)^-; x\Big) \\ &\leq T\cdot\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B\setminus I_\epsilon)^{\Delta^{(n)}}\Big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}_b^*}; \end{aligned} \tag{4.26}$$

here the last inequality we applied property (4.20). Furthermore,

$$\begin{aligned} \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B\setminus I_\epsilon)^{\Delta^{(n)}}\Big) &\leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}}\Big) \qquad \text{due to } (E\cup F)^\Delta \subseteq E^\Delta \cup F^\Delta \\ &= \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}} \cap I^c\Big) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}} \cap I\Big) \end{aligned}$$

$$
\begin{aligned}
&\leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \setminus I\Big) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((I_\epsilon^c)^{\Delta^{(n)}} \cap I\Big) \\
&\leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \setminus I\Big) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\epsilon+\Delta^{(n)}}\Big) \\
&\leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{-} \setminus I\Big) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^{\Delta^{(n)}} \cap I^c) \setminus (B^- \cap I^c)\Big) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\epsilon+\Delta^{(n)}}\Big)
\end{aligned}
$$

Considering the definition of $\check{\boldsymbol{\delta}}_B$ in (4.22), one can plug this bound back into (4.26) and yield the upper bound (4.24). Similarly, by applying Theorem 2.3 and property (4.21), we obtain

$$
\begin{aligned}
\liminf_{\eta \downarrow 0} \inf_{x \in [-\epsilon,\epsilon]} \frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x) \leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} &\geq \inf_{x \in [-\epsilon,\epsilon]} \mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\Big(\big(\check{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^\circ; x\Big) \\
&\geq (T - \bar{t}) \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I_\epsilon)_{\Delta^{(n)}}\Big).
\end{aligned} \tag{4.27}
$$

Furthermore, from the preliminary bound $(E \cap F)_\Delta \supseteq E_\Delta \cap F_\Delta$ we get

$$
\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I_\epsilon)_{\Delta^{(n)}}\Big) \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I)_{\Delta^{(n)}}\Big) \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B_{\Delta^{(n)}} \cap I^c_{\Delta^{(n)}}\Big).
$$

Together with the fact that $B_\Delta \setminus I = B_\Delta \cap I^c \subseteq \Big(B_\Delta \cap (I^c)_\Delta\Big) \cup \Big(I^c \setminus (I^c)_\Delta\Big)$, we yield

$$
\begin{aligned}
\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I_\epsilon)_{\Delta^{(n)}}\Big) &\geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B_{\Delta^{(n)}} \setminus I\Big) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(I^c \setminus I^c_{\Delta^{(n)}}\Big) \\
&\geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B_{\Delta^{(n)}} \setminus I\Big) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\Delta^{(n)}}\Big) \\
&\geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^\circ \setminus I\Big) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^\circ \cap I^c) \setminus (B_{\Delta^{(n)}} \cap I^c)\Big) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\Delta^{(n)}}\Big).
\end{aligned}
$$

Plugging this bound back into (4.27), we establish the lower bound (4.25) and conclude the proof. □

The next two results verify conditions (2.39) and (2.40). Let

$$
R_\epsilon^{\eta|b}(x) \triangleq \min\big\{j \geq 0:\ X_j^{\eta|b}(x) \in (-\epsilon, \epsilon)\big\} \tag{4.28}
$$

be the first time $X_j^{\eta|b}(x)$ returned to the ϵ-neighborhood of the origin. Under our choice of $A(\epsilon) = (-\epsilon, \epsilon)$ and $I(\epsilon) = I_\epsilon = (s_\text{left}+\epsilon, s_\text{right}-\epsilon)$, the event $\{\tau^\eta_{(I(\epsilon)\setminus A(\epsilon))^c}(x) > T/\eta\}$ in condition (2.39) means that $X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon)$ for all $j \leq T/\eta$. Also, recall that $\gamma(\eta)T/\eta = C_b^* T \cdot \big(\lambda(\eta)\big)^{\mathcal{J}_b^*}$. Therefore, to verify condition (2.39), it suffices to prove the following result.

Lemma 4.8 (Verifying condition (2.39))**.** *Given any $k \geq 1$ and $\epsilon \in (0, \bar{\epsilon})$, it holds for all $T \geq k \cdot \boldsymbol{t}(\epsilon/2)$ that*

$$
\lim_{\eta \downarrow 0} \sup_{x \in I_\epsilon^-} \frac{1}{\lambda^{k-1}(\eta)} \mathbf{P}\Big(X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon) \quad \forall j \leq T/\eta\Big) = 0.
$$

Proof. First, $\Big\{X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon)\ \ \forall j \leq T/\eta\Big\} = \big\{\boldsymbol{X}^{\eta|b}_{[0,T]}(x) \in E(\epsilon)\big\}$ where

$$
E(\epsilon) \triangleq \big\{\xi \in \mathbb{D}[0,T]:\ \xi(t) \in I_\epsilon \setminus (-\epsilon, \epsilon)\ \ \forall t \in [0,T]\big\}.
$$

Recall the definition of $\mathbb{D}_A^{(k)|b}[0,T]$ in (2.18). We claim that $E(\epsilon)$ is bounded away from $\mathbb{D}^{(k-1)|b}_{I_\epsilon^-}[0,T]$. This allows us to apply Theorem 2.3 and conclude that

$$
\sup_{x \in I_\epsilon^-} \mathbf{P}\Big(\boldsymbol{X}^{\eta|b}_{[0,T]}(x) \in E(\epsilon)\Big) = \boldsymbol{O}\big(\lambda^k(\eta)\big) = \boldsymbol{o}\big(\lambda^{k-1}(\eta)\big) \quad \text{as } \eta \downarrow 0.
$$

Now it only remains to verify that $E(\epsilon)$ is bounded away from $\mathbb{D}^{(k-1)|b}_{I^-_\epsilon}[0,T]$, which can be established if we show that for any $\xi \in \mathbb{D}^{(k-1)|b}_{I^-_\epsilon}[0,T]$ and $\xi' \in E(\epsilon)$,

$$\boldsymbol{d}_{J_1,[0,T]}(\xi,\xi') \geq \frac{\epsilon}{2}. \tag{4.29}$$

First, if $\xi(t) \notin I_{\epsilon/2}$ for some $t \leq T$, then by definition of $E(\epsilon)$ we get $\boldsymbol{d}_{J_1,[0,T]}(\xi,\xi') \geq \frac{\epsilon}{2}$. Now suppose that $\xi(t) \in I_{\epsilon/2}$ for all $t \leq T$. Let $x_0 \in I^-_\epsilon$, $(w_1,\cdots,w_{k-1}) \in \mathbb{R}^{k-1}$, and $(t_1,\cdots,t_{k-1}) \in (0,T]^{k-1\uparrow}$ be such that $\xi = h^{(k-1)|b}_{[0,T]}(x_0,w_1,\cdots,w_{k-1},t_1,\cdots,t_{k-1})$. With the convention that $t_0 = 0$ and $t_k = T$, we have

$$\xi(t) = \boldsymbol{y}_{t-t_{j-1}}\big(\xi(t_{j-1})\big) \qquad \forall t \in [t_{j-1},t_j). \tag{4.30}$$

for each $j \in [k]$. Here $\boldsymbol{y}_\cdot(x)$ is the ODE defined in (2.32). Also, note that due to the assumption $T \geq k \cdot \boldsymbol{t}(\epsilon/2)$, there exists some $j \in [k]$ such that $t_j - t_{j-1} \geq \boldsymbol{t}(\epsilon/2)$. However, note that we have assumed that $\xi(t_{j-1}) \in I_{\epsilon/2}$. Combining (4.30) along with property (4.9), we get $\lim_{t\uparrow t_j}\xi(t) \in [-\frac{\epsilon}{2},\frac{\epsilon}{2}]$. On the other hand, $\xi'(t) \notin (-\epsilon,\epsilon)$ for all $t \in [0,T]$, which implies that $\boldsymbol{d}_{J_1,[0,T]}(\xi,\xi') \geq \frac{\epsilon}{2}$. This concludes the proof. □

Lastly, we establish condition (2.40). Note that the first visit time $\tau^\eta_{A(\epsilon)}(x)$ therein coincides with $R^{\eta|b}_\epsilon(x)$ defined in (4.28) due to our choice of $A(\epsilon) = (-\epsilon,\epsilon)$.

Lemma 4.9 (Verifying condition (2.40))**.** *Let $\boldsymbol{t}(\cdot)$ be defined as in (4.8) and*

$$E(\eta,\epsilon,x) \triangleq \Big\{R^{\eta|b}_\epsilon(x) \leq \frac{\boldsymbol{t}(\epsilon/2)}{\eta};\ X^{\eta|b}_j(x) \in I_{\epsilon/2}\ \forall j \leq R^{\eta|b}_\epsilon(x)\Big\}.$$

For each $\epsilon \in (0,\bar{\epsilon})$ we have $\lim_{\eta\downarrow 0}\sup_{x\in I^-_\epsilon}\mathbf{P}\Big(\big(E(\eta,\epsilon,x)\big)^c\Big) = 0$.

Proof. First, note that $\big(E(\eta,\epsilon,x)\big)^c \subseteq \big\{\boldsymbol{X}^{\eta|b}_{[0,\boldsymbol{t}(\epsilon/2)]}(x) \in E^*_1(\epsilon)\cup E^*_2(\epsilon)\cup E^*_3(\epsilon)\big\}$ where

$$E^*_1(\epsilon) \triangleq \big\{\xi \in \mathbb{D}[0,\boldsymbol{t}(\epsilon/2)]:\ \xi(t) \notin (-\epsilon,\epsilon)\ \forall t \in [0,\boldsymbol{t}(\epsilon/2)]\big\},$$
$$E^*_2(\epsilon) \triangleq \big\{\xi \in \mathbb{D}[0,\boldsymbol{t}(\epsilon/2)]:\ \exists 0 \leq s \leq t \leq \boldsymbol{t}(\epsilon/2)\ s.t.\ \xi(t) \in (-\epsilon,\epsilon),\ \xi(s) \notin I_{\epsilon/2}\big\}.$$

Recall the definition of $\mathbb{D}^{(k)|b}_A[0,T]$ in (2.18). We claim that both $E^*_1(\epsilon)$ and $E^*_2(\epsilon)$ are bounded away from

$$\mathbb{D}^{(0)|b}_{I^-_\epsilon}[0,\boldsymbol{t}(\epsilon/2)] = \Big\{\{\boldsymbol{y}_t(x):\ t \in [0,\boldsymbol{t}(\epsilon/2)]\}:\ x \in I^-_\epsilon\Big\}.$$

To see why, note that from Assumption 6 and property (4.9), we get $\boldsymbol{y}_{\boldsymbol{t}(\epsilon/2)}(x) \in [-\frac{\epsilon}{2},\frac{\epsilon}{2}]$ and $\boldsymbol{y}_t(x) \in I_\epsilon$, $|\boldsymbol{y}_t(x)| \leq |x|$ for all t and x such that $t \in [0,\boldsymbol{t}(\epsilon/2)]$ and $x \in I^-_\epsilon$. Therefore,

$$\boldsymbol{d}_{J_1,[0,\boldsymbol{t}(\epsilon/2)]}\Big(\mathbb{D}^{(0)|b}_{I^-_\epsilon}[0,\boldsymbol{t}(\epsilon/2)],\ E^*_1(\epsilon)\Big) \geq \frac{\epsilon}{2} > 0, \tag{4.31}$$
$$\boldsymbol{d}_{J_1,[0,\boldsymbol{t}(\epsilon/2)]}\Big(\mathbb{D}^{(0)|b}_{I^-_\epsilon}[0,\boldsymbol{t}(\epsilon/2)],\ E^*_2(\epsilon)\Big) \geq \frac{\epsilon}{2} > 0. \tag{4.32}$$

This allows us to apply Theorem 2.3 and obtain $\sup_{x\in I^-_\epsilon}\mathbf{P}\Big(\big(E(\eta,\epsilon,x)\big)^c\Big) \leq \sup_{x\in I^-_\epsilon}\mathbf{P}\big(\boldsymbol{X}^{\eta|b}_{[0,\boldsymbol{t}(\epsilon/2)]}(x) \in E^*_1(\epsilon)\cup E^*_2(\epsilon)\big) = \boldsymbol{O}\big(\lambda(\eta)\big)$ as $\eta \downarrow 0$. To conclude the proof, one only needs to note that $\lambda(\eta) \in \mathcal{RV}_{\alpha-1}(\eta)$ (with $\alpha > 1$) and hence $\lim_{\eta\downarrow 0}\lambda(\eta) = 0$. □

Now we are ready to provide the proof of Theorem 2.6.

Proof of Theorem 2.6. (a) Since Lemmas 4.7–4.9 have verified Condition 1, part (a) of Theorem 2.6 follows immediately from Theorem 2.7.

(b) Note that the value of $\sigma(\cdot)$ and $a(\cdot)$ outside of $I^- = [s_{\text{left}}, s_{\text{right}}]$ has no impact on the first exit time problem. Therefore, by modifying the value of $\sigma(\cdot)$ and $a(\cdot)$ outside of I^-, we can assume w.l.o.g. that there is some $C > 0$ such that $0 \leq \sigma(x) \leq C$ and $|a(x)| \leq C$ for all $x \in \mathbb{R}$. We start with a few observations. First, note that under any $\eta \in (0, \frac{b}{2C})$, on the event $\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq t\}$ the step-size (before truncation) $\eta a\big(X^{\eta|b}_{j-1}(x)\big) + \eta\sigma\big(X^{\eta|b}_{j-1}(x)\big)Z_j$ of $X^{\eta|b}_j$ is less than b for each $j \leq t$. Therefore, $X^{\eta|b}_j(x)$ and $X^{\eta}_j(x)$ coincide for such j's. In other words, for any $\eta \in (0, \frac{b}{2C})$, on event $\big\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq t\big\}$ we have

$$X^{\eta|b}_j(x) = X^{\eta}_j(x)\ \forall j \leq t. \tag{4.33}$$

Second, note that for any $b > |s_{\text{left}}| \vee s_{\text{right}}$ we have $\mathcal{J}^*_b = 1$. More importantly, given any measurable $A \subseteq \mathbb{R}$ such that $r_A = \inf\{|x| :\ x \in A\} > 0$, we claim that

$$\lim_{b\to\infty} \check{\mathbf{C}}^{(1)|b}(A) = \check{\mathbf{C}}(A). \tag{4.34}$$

This claim follows from a simple application of the dominated convergence theorem. Indeed, by definition of $\check{\mathbf{C}}^{(1)|b}$, we get $\check{\mathbf{C}}^{(1)|b}(A) = \int \mathbb{I}\big\{\varphi_b\big(\sigma(0)\cdot w\big) \in A\big\}\nu_\alpha(dw)$. For $f_b(w) \triangleq \mathbb{I}\big\{\varphi_b\big(\sigma(0)\cdot w\big)\big)$, we first note that given $w \in \mathbb{R}$, we have $f_b(w) = f(w) \triangleq \mathbb{I}\big\{\sigma(0)\cdot w\big)$ for all $b > |w|\cdot\sigma(0)$. Therefore, $\lim_{b\to\infty} f_b(w) = f(w)$ holds for all $w \in \mathbb{R}$. Next, due to $r_A > 0$, we have $f_b(w) \leq \mathbb{I}\big\{|w| \geq r_A/\sigma(0)\big\}$ for all $b > 0$ and $w \in \mathbb{R}$. Meanwhile, note that $\int \mathbb{I}\big\{|w| \geq r_A/\sigma(0)\big\}\nu_\alpha(dw) = (\sigma(0)/r_A)^\alpha < \infty$. This allows us to apply dominated convergence theorem and establish (4.34). Similarly, for all $b > |s_{\text{left}}| \vee s_{\text{right}}$, we have

$$C^*_b = \check{\mathbf{C}}^{(1)|b}(I^c) = \int \mathbb{I}\Big\{\varphi_b\big(\sigma(0)\cdot w\big) \in I^c\Big\}\nu_\alpha(dw) = \int \mathbb{I}\Big\{\sigma(0)\cdot w \in I^c\Big\}\nu_\alpha(dw) = \check{\mathbf{C}}(I^c) \triangleq C^*. \tag{4.35}$$

To see why, it suffices to notice that for such b,

$$\varphi_b\big(\sigma(0)\cdot w\big) \notin I \qquad \Longleftrightarrow \qquad \sigma(0)\cdot w \notin I.$$

Now, we fix $t \geq 0$ and $B \subseteq I^c$. Also, henceforth in the proof we only consider $b > |s_{\text{left}}| \vee s_{\text{right}}$ large enough such that $C^* = C^*_b$. An immediate consequence of this choice of b is that $\mathcal{J}^*_b = \lceil r/b \rceil = 1$. First, note that $\lambda(\eta) = \eta^{-1}\cdot H(\eta^{-1})$ and hence $\eta\cdot\lambda(\eta) = H(\eta^{-1})$. To analyze the probability of event $A(\eta, x) = \{C^* H(\eta^{-1})\tau^\eta(x) > t,\ X^\eta_{\tau^\eta(x)}(x) \in B\}$, we arbitrarily pick some $T > t$ and observe that

$$A(\eta, x) = \underbrace{\Big\{C^* H(\eta^{-1})\tau^\eta(x) \in (t, T],\ X^\eta_{\tau^\eta(x)}(x) \in B\Big\}}_{\triangleq A_1(\eta,x,T)} \cup \underbrace{\Big\{C^* H(\eta^{-1})\tau^\eta(x) > T,\ X^\eta_{\tau^\eta(x)}(x) \in B\Big\}}_{\triangleq A_2(\eta,x,T)}. \tag{4.36}$$

Let $E_b(\eta, T) \triangleq \big\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq \frac{T}{C^* H(\eta^{-1})}\big\}$. To analyze the probability of $A_1(\eta, x, T)$, we further decompose the event as $A_1(\eta, x, T) = \big(A_1(\eta, x, T) \cap E_b(\eta, T)\big) \cup \big(A_1(\eta, x, T) \setminus E_b(\eta, T)\big)$. First, for all $\eta \in (0, \frac{b}{2C})$,

$$\begin{aligned}
&\mathbf{P}\Big(A_1(\eta, x, T) \cap E_b(\eta, T)\Big) \\
&= \mathbf{P}\bigg(\Big\{C^*_b\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) \in (t, T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B\Big\} \cap E_b(\eta, T)\bigg) \qquad \text{due to (4.33) and (4.35)} \\
&\leq \mathbf{P}\bigg(C^*_b\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) \in (t, T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B\bigg)
\end{aligned}$$

$$= \mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) > t,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big) - \mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) > T,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big).$$

Using part (a) of Theorem 2.6 and observation (4.35), we get

$$\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\Big(A_1(\eta,x,T)\cap E_b(\eta,T)\Big) \le \frac{\check{\mathbf{C}}^{(1)|b}(B^-)}{C^*}\cdot\exp(-t) - \frac{\check{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*}\cdot\exp(-T). \tag{4.37}$$

On the other hand, $\sup_{x\in I_\epsilon}\mathbf{P}\big(A_1(\eta,x,T)\setminus E_b(\eta,T)\big) \le \mathbf{P}\big((E_b(\eta,T))^c\big) = \mathbf{P}\big(\eta|Z_j| > \frac{b}{2C}$ for some $j \le \frac{T}{C^*H(\eta^{-1})}\big)$. Applying Lemma 3.1 (i), we get

$$\begin{aligned}\limsup_{\eta\downarrow 0}\mathbf{P}\Big(\eta|Z_j| > \frac{b}{2C}\text{ for some } j\le\frac{T}{C^*H(\eta^{-1})}\Big) &= 1 - \liminf_{\eta\downarrow 0}\mathbf{P}\Big(\mathrm{Geom}\Big(H(\frac{b}{\eta\cdot 2C})\Big) > \frac{T}{C^*H(\eta^{-1})}\Big)\\ &\le 1 - \lim_{\eta\downarrow 0}\exp\Big(-\frac{T\cdot H(\eta^{-1}\cdot\frac{b}{2C})}{C^*H(\eta^{-1})}\Big)\\ &= 1-\exp\Big(-\frac{T}{C^*}\cdot\Big(\frac{2C}{b}\Big)^{\alpha}\Big).\end{aligned}\tag{4.38}$$

Similarly,

$$\begin{aligned}A_2(\eta,x,T) &\subseteq \Big\{C^*H(\eta^{-1})\tau^\eta(x) > T\Big\}\\ &= \Big(\Big\{C^*H(\eta^{-1})\tau^\eta(x) > T\Big\}\cap E_b(\eta,T)\Big)\cup\Big(\Big\{C^*H(\eta^{-1})\tau^\eta(x) > T\Big\}\setminus E_b(\eta,T)\Big).\end{aligned}$$

On $\{C^*H(\eta^{-1})\tau^\eta(x) > T\}\cap E_b(\eta,T)$, due to (4.33) we have $\tau^\eta(x) = \tau^{\eta|b}(x)$. Also, from (4.35) we get $C^* = C_b^*$. Using part (a) of Theorem 2.6 again, we get

$$\limsup_{\eta\downarrow 0}\mathbf{P}\Big(\Big\{C^*H(\eta^{-1})\tau^\eta(x) > T\Big\}\cap E_b(\eta,T)\Big) \le \limsup_{\eta\downarrow 0}\mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) > T\Big) \le \exp(-T). \tag{4.39}$$

Meanwhile, the limit of $\sup_{x\in I_\epsilon}\mathbf{P}\big(C^*H(\eta^{-1})\tau^\eta(x) > T\}\cap E_b(\eta,T)\big)$ as $\eta\downarrow 0$ is again bounded by (4.38). Collecting (4.37), (4.38), and (4.39), we have shown that for all $b > 0$ large enough and all $T > t$,

$$\begin{aligned}\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\Big(A(\eta,x)\Big) \le \frac{\check{\mathbf{C}}^{(1)|b}(B^-)}{C^*}\cdot\exp(-t) - \frac{\check{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*}\cdot\exp(-T) + \exp(-T)\\ + 2\cdot\Big[1-\exp\Big(-\frac{T}{C^*}\cdot\Big(\frac{2C}{b}\Big)^{\alpha}\Big)\Big].\end{aligned}$$

In light of claim (4.34), we can drive $b\to\infty$ and obtain $\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\big(A(\eta,x)\big) \le \frac{\check{\mathbf{C}}(B^-)}{C^*}\cdot\exp(-t) - \frac{\check{\mathbf{C}}(B^\circ)}{C^*}\cdot\exp(-T)+\exp(-T)$. Letting T tend to ∞, we conclude the proof of the upper bound.

The lower bound can be established analogously. In particular, from the decomposition in (4.36), we get

$$\begin{aligned}&\inf_{x\in I_\epsilon}\mathbf{P}\big(A(\eta,x)\big)\\ &\ge \inf_{x\in I_\epsilon}\mathbf{P}\big(A_1(\eta,x,T)\big) \ge \inf_{x\in I_\epsilon}\mathbf{P}\big(A_1(\eta,x,T)\cap E_b(\eta,T)\big)\\ &= \inf_{x\in I_\epsilon}\mathbf{P}\Big(\Big\{C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)\in(t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big\}\cap E_b(\eta,T)\Big) \qquad \text{due to (4.33) and (4.35)}\end{aligned}$$

$$\geq \inf_{x \in I_\epsilon} \mathbf{P}\Big(C^*_b\eta \cdot \lambda(\eta)\tau^{\eta|b}(x) \in (t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B \Big) - \mathbf{P}\Big(\big(E_b(\eta,T)\big)^c \Big)$$

$$\geq \inf_{x \in I_\epsilon} \mathbf{P}\Big(C^*_b\eta \cdot \lambda(\eta)\tau^{\eta|b}(x) > t,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B \Big) - \sup_{x \in I_\epsilon} \mathbf{P}\Big(C^*_b\eta \cdot \lambda(\eta)\tau^{\eta|b}(x) > T,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B \Big)$$
$$- \mathbf{P}\Big(\big(E_b(\eta,T)\big)^c \Big).$$

Using part (a) of Theorem 2.6 and the limit in (4.38), we yield (for all $b > 0$ large enough and all $T > t$)

$$\liminf_{\eta \downarrow 0} \inf_{x \in I_\epsilon} \mathbf{P}\Big(A(\eta,x)\Big) \leq \frac{\check{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*} \cdot \exp(-t) - \frac{\check{\mathbf{C}}^{(1)|b}(B^-)}{C^*} \cdot \exp(-T) - \Bigg[1 - \exp\Bigg(-\frac{T}{C^*} \cdot \Big(\frac{2C}{b}\Big)^\alpha\Bigg)\Bigg].$$

Sending $b \to \infty$ and then $T \to \infty$, we conclude the proof of the lower bound. □

5 Sample-Path Convergence of Global Dynamics

5.1 Law of the Limiting Markov Chains in Theorems 2.9 and 2.10

Consider some $b \in (0,\infty)$ such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{\min}]$ and $j \in [n_{\min} - 1]$. This allows us to fix some $\bar{\epsilon} \in (0, 1 \wedge b)$ such that

$$r_i > \big(\mathcal{J}^*_b(i) - 1\big)b + 3\bar{\epsilon}, \qquad [m_i - \bar{\epsilon}, m_i + \bar{\epsilon}] \subseteq [s_{i-1} + \bar{\epsilon}, s_i - \bar{\epsilon}] \qquad \forall i \in [n_{\min}] \tag{5.1}$$

with r_i and $\mathcal{J}^*_b(i)$ defined in (2.41) and (2.42), respectively. Recall the definition of $\check{\mathbf{C}}^{(k)|b}(\ \cdot\ ; x)$ in (2.34), and define (for $i, j \in [n_{\min}]$ with $i \neq j$)

$$q_b(i,j) \triangleq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j; m_i), \qquad q_b(i) \triangleq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_i; m_i). \tag{5.2}$$

First, note that $q_b(i) = \sum_{j \in [n_{\min}]:\ j \neq i} q_b(i,j) + \sum_{j \in [n_{\min}-1]} \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(\{s_j\}; m_i)$. From (5.1), we have $|s_j - m_i| > (\mathcal{J}^*_b(i) - 1) \cdot b$. By assumption $|s_j - m_i|/b \notin \mathbb{Z}$ for all $j \in [n_{\min} - 1]$, one can then apply Lemma 4.3 to show that $\sum_{j \in [n_{\min}-1]} \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(\{s_j\}; m_i) = 0$. Together with Lemma 4.4, we yield that $q_b(i) = \sum_{j \in [n_{\min}]:\ j \neq i} q_b(i,j) \in (0,\infty)$. Furthermore, based on our choice of $\bar{\epsilon}$ of (5.1) in (5.1), we can Lemma 4.6 and yield

$$q_b(i,j) > 0 \qquad \Longleftrightarrow \qquad \mathcal{J}^*_b(i,j) = \mathcal{J}^*_b(i).$$

First, we detail the law of $Y^{*|b}$ and π_b. Given any $m_{\text{init}} \in V$, consider a $\Big((U_j)_{j\geq 1}, (V_j)_{j \geq 1}\Big)$ jump process $Y^{*|b}_t(m_{\text{init}})$ defined as follows. Set $V_1 = m_{\text{init}}$ and $U_1 = 0$, and (for any $t > 0$, $l \geq 1$, and $i, j \in [n_{\text{init}}]$ with $i \neq j$)

$$\begin{aligned} &\mathbf{P}\Big(U_{l+1} < t,\ V_{l+1} = m_j\ \Big|\ V_l = m_i, (V_j)^{l-1}_{j=1},\ (U_j)^l_{j=1}\Big) = \mathbf{P}\Big(U_{l+1} < t,\ V_{l+1} = m_j\ \Big|\ V_l = m_i\Big) \\ &= \begin{cases} \frac{q_b(i,j)}{q_b(i)} & \text{if } m_i \notin V^*_b, \\ \frac{q_b(i,j)}{q_b(i)} \cdot \Big(1 - \exp\big(-q_b(i)t\big)\Big) & \text{if } m_i \in V^*_b. \end{cases} \end{aligned} \tag{5.3}$$

In other words, conditioning on $V_l = m_i$, the law of U_{l+1} and V_{l+1} are independent: we have $V_{l+1} = m_j$ with probability $q_b(i,j)/q_b(i)$; as for U_{l+1}, we set $U_{l+1} \equiv 0$ if $m_i \notin V^*_b$ (i.e., the current value m_i is not a widest minimum), and set U_{l+1} as an Exponential RV with rate $q_b(i)$ otherwise.

We make two observations. First, if $m_{\text{init}} \in V^*_b$, then $Y^{*|b}_t(m_{\text{init}})$ is a continuous-time Markov chain that only visits V^*_b due to the fact that any jump at $m \notin V^*_b$ is instantaneous. Second, if

$m_{\text{init}} \notin V_b^*$, then a series of instantaneous jumps will immediately send $Y_t^{*|b}(m_{\text{init}})$ to some $m \in V_b^*$ at time $t = 0$. This procedure is summarized by the random mapping π_b. In particular, consider a discrete-time Markov chain $\widetilde{Y}_k^{*|b}(\cdot)$ on V where the one-step transition probability from m_i to m_j is $q_b(i,j)/q_b(i)$. Let

$$\tau_{\text{DTMC}}^{*|b}(m) \triangleq \min\{k \geq 0 : \ \widetilde{Y}_k^{*|b}(m) \in V_b^*\}, \qquad p_b(i,j) \triangleq \mathbf{P}\Big(\widetilde{Y}^{*|b}_{\tau^{*|b}_{\text{DTMC}}(m_i)}(m_i) = m_j\Big). \tag{5.4}$$

Here, we can view any $m \in V_b^*$ as an absorbing state, and $p_b(i,j)$ is the absorption probability at m_j under initial condition m_i. Returning to the continuous-time jump process $Y_t^{*|b}$, one can see that $Y_0^{*|b}(m_i) = m_j$ with probability $p_b(i,j)$. Therefore, under the definition

$$\pi_b(m_i) = m_j \qquad \text{w.p. } p_b(i,j), \tag{5.5}$$

the inclusion of random mapping π_b in the initial value of $Y_t^{*|b}\big(\pi_b(m_i)\big)$ only serves to capture all the instantaneous jumps at $t = 0$ before the first visit to some $m \in V_b^*$.

Recall the definition of measure $\check{\mathbf{C}}(\ \cdot\ ; x)$ in (2.35). For $i, j \in [n_{\min}]$ with $i \neq j$, let

$$q(i,j) \triangleq \check{\mathbf{C}}(I_j; m_i). \tag{5.6}$$

Let $Y_t^*(\cdot)$ be a continuous-time Markov chain over states $\{m_i : \ i \in [n_{\min}]\}$ with generator characterized by $q(i,j)$ and initial condition $Y_0^*(m) = m$.

5.2 Proof of Theorem 2.10

Proof of Theorem 2.10. For any $b > \max_{i \in [n_{\min}],\ j \in [n_{\min}-1]} |m_i - s_j|$, by definitions in (2.42) we have $\mathcal{J}_b^*(i,j) = \mathcal{J}_b^*(i) = 1$ for all $i \in [n_{\min}]$ and $j \in [n_{\min} - 1]$. Therefore, for such $b > 0$ large enough, we also have $\lambda_b^*(\eta) = \eta \cdot \lambda(\eta) = H(\eta^{-1})$. Henceforth in this proof, we only consider such large b.

Fix some $i \in [n_{\min}]$, $x \in I_i$, and $0 < t_1 < t_2 < \cdots < t_k$. Also, pick some closed set $A \subseteq \mathbb{R}^k$. Observe that

$$\begin{aligned}
&\mathbf{P}\bigg(\Big(X^\eta_{\lfloor t_1/H(\eta^{-1})\rfloor}(x), \cdots, X^\eta_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\Big) \in A\bigg) \qquad (5.7)\\
&\leq \mathbf{P}\bigg(\Big(X^{\eta|b}_{\lfloor t_1/H(\eta^{-1})\rfloor}(x), \cdots, X^{\eta|b}_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\Big) \in A,\ X_j^{\eta|b}(x) = X_j^\eta(x)\ \forall j \leq \lfloor t_k/H(\eta^{-1})\rfloor\bigg)\\
&\qquad + \mathbf{P}\bigg(\Big(X^\eta_{\lfloor t_1/H(\eta^{-1})\rfloor}(x), \cdots, X^\eta_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\Big) \in A, X_j^{\eta|b}(x) \neq X_j^\eta(x) \text{ for some } j \leq \lfloor t_k/H(\eta^{-1})\rfloor\bigg)\\
&\leq \underbrace{\mathbf{P}\bigg(\Big(X^{\eta|b}_{\lfloor t_1/H(\eta^{-1})\rfloor}(x), \cdots, X^{\eta|b}_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\Big) \in A\bigg)}_{\text{(I)}} + \underbrace{\mathbf{P}\bigg(X_j^{\eta|b}(x) \neq X_j^\eta(x) \text{ for some } j \leq \lfloor t_k/H(\eta^{-1})\rfloor\bigg)}_{\text{(II)}}.
\end{aligned}$$

For term (I), it follows from Theorem 2.9 that $\limsup_{\eta \downarrow 0} \text{(I)} \leq \mathbf{P}\Big(\big(Y_{t_1}^{*|b}(m_i), \cdots, Y_{t_k}^{*|b}(m_i)\big) \in A\Big)$. Here, the process $Y_t^{*|b}(m_i)$ defined in Section 5.1 is simply an irreducible continuous-time Markov chain with generator $q_b(i,j)$. Indeed, any pair of nodes m_i and m_j would communicate with each other on the b-typical transition graph $\mathcal{G}_b$ (see Definition 2.3) due to $\mathcal{J}_b^*(i,j) = \mathcal{J}_b^*(i) = 1$ for all $i \in [n_{\min}]$ and $j \in [n_{\min} - 1]$, which implies $m_i \in V_b^*$ for all $i \in [n_{\min}]$.

For term (II), we make two observations. First, recall that $C \in [1, \infty)$ is the constant in Assumption 4 such that $\sup_{x \in \mathbb{R}} |a(x)| \vee \sigma(x) \leq C$. Under any $\eta \in (0, \frac{b}{2C})$, on the event $\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq \lfloor t_k/H(\eta^{-1})\rfloor\}$ the step-size (before truncation) $\eta a\big(X_{j-1}^{\eta|b}(x)\big) + \eta\sigma\big(X_{j-1}^{\eta|b}(x)\big)Z_j$ of $X_j^{\eta|b}$ is less than b for each $j \leq \lfloor t_k/H(\eta^{-1})\rfloor$. Therefore, $X_j^{\eta|b}(x)$ and $X_j^\eta(x)$ coincide for such j's. In other words, for any $\eta \in (0, \frac{b}{2C})$, we have $\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq \lfloor t_k/H(\eta^{-1})\rfloor\} \subseteq \{X^{\eta|b}_j(x) = X^\eta_j(x)\ \forall j \leq \lfloor t_k/H(\eta^{-1})\rfloor\}$. which leads to (recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$)

$$\limsup_{\eta\downarrow 0} \text{(II)} \leq \limsup_{\eta\downarrow 0} \mathbf{P}\Big(\exists j \leq \lfloor t_k/H(\eta^{-1})\rfloor\ s.t.\ \eta|Z_j| > \frac{b}{2C}\Big)$$
$$\leq \limsup_{\eta\downarrow 0} \frac{t_k}{H(\eta^{-1})}\cdot H(\eta^{-1}\cdot\frac{b}{2C}) = t_k\cdot\Big(\frac{2C}{b}\Big)^\alpha$$

due to $H(x) \in \mathcal{RV}_{-\alpha}(x)$ as $x \to \infty$. In summary,

$$\limsup_{\eta\downarrow 0} \mathbf{P}\Big(\big(X^\eta_{\lfloor t_1/H(\eta^{-1})\rfloor}(x),\cdots,X^\eta_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\big)\in A\Big) \leq \mathbf{P}\Big(\big(Y^{*|b}_{t_1}(m_i),\cdots,Y^{*|b}_{t_k}(m_i)\big)\in A\Big)+t_k\cdot\Big(\frac{2C}{b}\Big)^\alpha.$$

Furthermore, note that for all b large enough, we have $q_b(i,j) = q(i,j)$ for all $i,j \in [n_{\min}]$ with $i \neq j$. To see why, we fix some $i,j \in [n_{\min}]$ with $i \neq j$. For all b large enough, we have $\mathcal{J}^*_b(i,j) = 1$, and hence (see (5.2) and (5.6) for definitions of $q_b(i,j)$ and $q(i,j)$)

$$q(i,j) = \nu_\alpha\Big(\big\{w \in \mathbb{R}:\ m_i + \sigma(m_i)\cdot w \in I_j\big\}\Big),\quad q_b(i,j) = \nu_\alpha\Big(\big\{w \in \mathbb{R}:\ m_i + \varphi_b\big(\sigma(m_i)\cdot w\big) \in I_j\big\}\Big).$$

Suppose that I_j has bounded support (i.e., $j = 2,3,\cdots,n_{\min}-1$ so that I_j is not the leftmost or the rightmost attraction field), then it holds for all b large enough that $m_i - b \notin I_j$ and $m_i + b \notin I_j$. Under such large b, for $m_i + \varphi_b\big(\sigma(m_i)\cdot w\big) \in I_j$ to hold we must have $|\sigma(m_i)\cdot w| < b$, thus implying $m_i + \varphi_b\big(\sigma(m_i)\cdot w\big) = m_i + \sigma(m_i)\cdot w$ and hence $q_b(i,j) = q(i,j)$. Next, consider the case where $j = 1$ so $I_j = I_1 = (-\infty, s_1)$ is the leftmost attraction field. For any b large enough we must have $m_i - z \in (-\infty, s_1) = I_1$ for all $z \geq b$. This also implies $m_i + \varphi_b\big(\sigma(m_i)\cdot w\big) \in I_1 \iff m_i + \sigma(m_i)\cdot w \in I_1$. The same argument can be applied to the case with $j = n_{\min}$ (that is, $I_j = (s_{n_{\min}-1},\infty)$ is the rightmost attraction field).

Now that we know $q_b(i,j) = q(i,j)$ for all b large enough, the claim $Y^{*|b}_t(m_i) = Y^*_t(m_i)\ \forall t \geq 0$ must hold for all b large enough as both CTMCs have the same generator. Therefore, for the closed set $A \subseteq \mathbb{R}^k$, $\lim_{b\to\infty}\mathbf{P}\Big(\big(Y^{*|b}_{t_1}(m_i),\cdots,Y^{*|b}_{t_k}(m_i)\big)\in A\Big) = \mathbf{P}\Big(\big(Y^*_{t_1}(m_i),\cdots,Y^*_{t_k}(m_i)\big)\in A\Big)$. Together with the fact that $\lim_{b\to\infty}\big(\frac{2C}{b}\big)^\alpha = 0$, in (5.7) we obtain $\limsup_{\eta\downarrow 0}\mathbf{P}\Big(\big(X^\eta_{\lfloor t_1/H(\eta^{-1})\rfloor}(x),\cdots,X^\eta_{\lfloor t_k/H(\eta^{-1})\rfloor}(x)\big) \in A\Big) \leq \mathbf{P}\Big(\big(Y^*_{t_1}(m_i),\cdots,Y^*_{t_k}(m_i)\big)\in A\Big)$. From the arbitrariness of the closed set A, we conclude the proof with Portmanteau theorem. □

5.3 Proof of Lemmas 2.12 and 2.13

Proof of Lemma 2.12. Fix some $k \geq 1$ and $0 < t_1 < t_2 < \cdots < t_k$. Pick some open set $G \subseteq \mathbb{S}^k$ where $\mathbb{S}^k$ is the k-fold product space of $\mathbb{S}$ with uniform metric $\boldsymbol{d}^{(k)}\big((x_1,\cdots,x_k),(y_1,\cdots,y_k)\big) = \max_{i\in[k]}\boldsymbol{d}(x_i,y_i)$. By Portmanteau theorem, it suffices to show that $\liminf_{\eta\downarrow 0}\mathbf{P}\Big((Y^\eta_{t_1},\cdots,Y^\eta_{t_k}) \in G\Big) \geq \mathbf{P}\Big((Y^*_{t_1},\cdots,Y^*_{t_k})\in G\Big)$.

By part (ii) of Condition 2, $\lim_{n\to\infty}\mathbf{P}\Big(\boldsymbol{d}^{(k)}(\hat{\boldsymbol{Y}}^{\eta,\epsilon},\boldsymbol{Y}^\eta)\geq\epsilon\Big) = 0$ holds for all $\epsilon > 0$ small enough where $\hat{\boldsymbol{Y}}^{\eta,\epsilon} = (\hat{Y}^{\eta,\epsilon}_{t_1},\cdots,\hat{Y}^{\eta,\epsilon}_{t_k})$ and $\boldsymbol{Y}^\eta = (Y^\eta_{t_1},\cdots,Y^\eta_{t_k})$. Meanwhile,

$$\mathbf{P}\Big((Y^\eta_{t_1},\cdots,Y^\eta_{t_k})\in G\Big) \geq \mathbf{P}\Big((Y^\eta_{t_1},\cdots,Y^\eta_{t_k})\in G,\ \boldsymbol{d}^{(k)}(\hat{\boldsymbol{Y}}^{\eta,\epsilon},\boldsymbol{Y}^\eta)<\epsilon\Big)$$
$$\geq \mathbf{P}\Big((\hat{Y}^{\eta,\epsilon}_{t_1},\cdots,\hat{Y}^{\eta,\epsilon}_{t_k})\in G_\epsilon,\ \boldsymbol{d}^{(k)}(\hat{\boldsymbol{Y}}^{\eta,\epsilon},\boldsymbol{Y}^\eta)<\epsilon\Big)$$
$$\geq \mathbf{P}\Big((\hat{Y}^{\eta,\epsilon}_{t_1},\cdots,\hat{Y}^{\eta,\epsilon}_{t_k})\in G_\epsilon\Big) - \mathbf{P}\Big(\boldsymbol{d}^{(k)}(\hat{\boldsymbol{Y}}^{\eta,\epsilon},\boldsymbol{Y}^\eta)\geq\epsilon\Big).$$

Here, G_ϵ is the ϵ-shrinkage of set G and note that G_ϵ is open. By part (i) of Condition 2, we get $\liminf_{\eta\downarrow 0}\mathbf{P}\Big((\hat{Y}^{\eta,\epsilon}_{t_1},\cdots,\hat{Y}^{\eta,\epsilon}_{t_k})\in G_\epsilon\Big)\geq \mathbf{P}\Big((Y^*_{t_1},\cdots,Y^*_{t_k})\in G_\epsilon)$. In summary, we have shown that $\liminf_{\eta\downarrow 0}\mathbf{P}\Big((Y^\eta_{t_1},\cdots,Y^\eta_{t_k})\in G\Big)\geq \mathbf{P}\Big((Y^*_{t_1},\cdots,Y^*_{t_k})\in G_\epsilon\Big)$. Let $\epsilon\downarrow 0$ and we conclude the proof due to continuity of measures and $\cup_{\epsilon>0}G_\epsilon = G$ for the open set G. □

Proof of Lemma 2.13. Fix some $k\in\mathbb{N}$ and $0<t_1<t_2<\cdots<t_k<\infty$. Set $t=t_k$. Pick some $\epsilon>0$. By assumption, one can fix some $J(\epsilon)>0$ such that $\mathbf{P}(\sum_{j=1}^{J(\epsilon)}U_j\leq t)<\epsilon$ as well as $N(\epsilon)$ such that $\mathbf{P}(\sum_{j=1}^{J(\epsilon)}U^n_j\leq t)<\epsilon$ for all $n\geq N(\epsilon)$. Also, we can fix $\Delta(\epsilon)>0$ such that $\mathbf{P}\Big(\sum_{i=1}^{j}U_i\in\bigcup_{l\in[k]}[t_l-\Delta(\epsilon),t_l+\Delta(\epsilon)]\text{ for some }j\leq J(\epsilon)\Big)<\epsilon$. Throughout the proof, we may abuse the notation slightly and write $N=N(\epsilon)$, $J=J(\epsilon)$ and $\Delta=\Delta(\epsilon)$ when there is no ambiguity.

For any probability measure μ, let $\mathscr{L}_\mu(X)$ be the law of the random element X under μ. Due to $\mathbb{S}$ being separable, we can apply Skorokhod's representation theorem and construct a probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},\mathbf{Q})$ that supports random variables $(\widetilde{U}^n_1,\widetilde{V}^n_1,\widetilde{U}^n_2,\widetilde{V}^n_2\cdots)_{n\geq 1}$ and $(\widetilde{U}_1,\widetilde{V}_1,\widetilde{U}_2,\widetilde{V}_2,\cdots)$ such that the following conditions hold:

- $\mathscr{L}_{\mathbf{P}}(U^n_1,V^n_1,U^n_2,V^n_2,\cdots)=\mathscr{L}_{\mathbf{Q}}(\widetilde{U}^n_1,\widetilde{V}^n_1,\widetilde{U}^n_2,\widetilde{V}^n_2\cdots)$ for all $n\geq 1$;
- $\mathscr{L}_{\mathbf{P}}(U_1,V_1,U_2,V_2,\cdots)=\mathscr{L}_{\mathbf{Q}}(\widetilde{U}_1,\widetilde{V}_1,\widetilde{U}_2,\widetilde{V}_2,\cdots)$;
- $\widetilde{U}^n_j\xrightarrow{\mathbf{Q}-a.s.}\widetilde{U}_j$ and $\widetilde{V}^n_j\xrightarrow{\mathbf{Q}-a.s.}\widetilde{V}_j$ as $n\to\infty$ for all $j\in[J]$.

This allows us to construct a coupling between processes Y_t and Y^n_t on $(\widetilde{\Omega},\widetilde{\mathcal{F}},\mathbf{Q})$ by setting Y_t as the $\Big((\widetilde{U}_j)_{j\geq 1},(\widetilde{V}_j)_{j\geq 1}\Big)$ jump process and (for each $n\geq 1$) Y^n_t as the $\Big((\widetilde{U}^n_j)_{j\geq 1},(\widetilde{V}^n_j)_{j\geq 1}\Big)$ jump process. Furthermore, define processes

$$Y^{n,\downarrow J}_s=Y^n_{s\wedge\sum_{j=1}^J\widetilde{U}^n_j},\qquad Y^{\downarrow J}_s=Y_{s\wedge\sum_{j=1}^J\widetilde{U}_j}.$$

We make a few observations under $\mathbf{Q}$. First, on event $\{\sum_{j=1}^J\widetilde{U}_j>t,\ \sum_{j=1}^J\widetilde{U}^n_j>t\}$, we have $Y^n_s=Y^{n,\downarrow J}_s$ and $Y_s=Y^{\downarrow J}_s$ for all $s\in[0,t]$. Next, for each $i\in[k]$ we define

$$\mathcal{I}^{\leftarrow}_i(\Delta)=\max\{j\geq 0:\ \widetilde{U}_1+\cdots\widetilde{U}_j\leq t_i-\Delta\},\qquad \mathcal{I}^{\rightarrow}_i(\Delta)=\min\{j\geq 0:\ \widetilde{U}_1+\cdots\widetilde{U}_j\geq t_i+\Delta\}.$$

On event $A_n(\Delta)=\{\sum_{i=1}^j\widetilde{U}_i\notin\bigcup_{l\in[k]}[t_l-\Delta,t_l+\Delta]\ \forall j\leq J\}\cap\{\sum_{j=1}^J\widetilde{U}_j>t,\ \sum_{j=1}^J\widetilde{U}^n_j>t\}$, we have $\mathcal{I}^{\rightarrow}_i(\Delta)=\mathcal{I}^{\leftarrow}_i(\Delta)+1$ for all $i\in[k]$. Therefore, on this event it holds $\mathbf{Q}$-a.s. that (for all $i\in[k]$)

$$\lim_{n\to\infty}\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)}\widetilde{U}^n_j=\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)}\widetilde{U}_j\leq t_i-\Delta,\qquad \lim_{n\to\infty}\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)+1}\widetilde{U}^n_j=\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)+1}\widetilde{U}_j\geq t_i+\Delta,$$
$$\lim_{n\to\infty}\widetilde{V}^n_{I^{\leftarrow}_i(\Delta)}=\widetilde{V}_{I^{\leftarrow}_i(\Delta)}.$$

Therefore, on this event it holds $\mathbf{Q}$-a.s. that $\lim_{n\to\infty}Y^n_{t_i}=\lim_{n\to\infty}\widetilde{V}^n_{I^{\leftarrow}_i(\Delta)}=\widetilde{V}_{I^{\leftarrow}_i(\Delta)}=Y_{t_i}$ for all $i\in[k]$. As a result, for any $g:\mathbb{S}^k\to\mathbb{R}$ that is bounded and continuous, note that (let $\boldsymbol{Y}^n=(Y^n_{t_1},\cdots,Y^n_{t_k})$, $\boldsymbol{Y}=(Y_{t_1},\cdots,Y_{t_k})$, and $\|g\|=\sup_{\boldsymbol{y}\in\mathbb{S}^k}|g(\boldsymbol{y})|$)

$$\begin{aligned}\limsup_{n\to\infty}\Big|\mathbf{E}g(\boldsymbol{Y}^n)-\mathbf{E}g(\boldsymbol{Y})\Big| &\leq \limsup_{n\to\infty}\mathbf{E}_{\mathbf{Q}}\Big|g(\boldsymbol{Y}^n)-g(\boldsymbol{Y})\Big| \\ &=\limsup_{n\to\infty}\mathbf{E}_{\mathbf{Q}}\Big|g(\boldsymbol{Y}^n)-g(\boldsymbol{Y})\Big|\mathbb{I}_{A_n(\Delta)}+\limsup_{n\to\infty}\mathbf{E}_{\mathbf{Q}}\Big|g(\boldsymbol{Y}^n)-g(\boldsymbol{Y})\Big|\mathbb{I}_{(A_n(\Delta))^c} \\ &\leq 0+\|g\|\limsup_{n\to\infty}\mathbf{Q}\Big(\big(A_n(\Delta)\big)^c\Big)\qquad\text{due to }\boldsymbol{Y}^n\xrightarrow{\mathbf{Q}-a.s.}\boldsymbol{Y}\text{ on }A_n(\Delta)\end{aligned}$$

$$\leq \|g\| \cdot \bigg(\limsup_{n \to \infty} \mathbf{Q}(\sum_{i=1}^{J} \widetilde{U}_j \leq t) + \limsup_{n \to \infty} \mathbf{Q}(\sum_{i=1}^{J} \widetilde{U}_j^n \leq t)$$

$$+ \limsup_{n \to \infty} \mathbf{Q}\Big(\sum_{i=1}^{j} \widetilde{U}_i \in \bigcup_{l \in [k]} [t_l - \Delta, t_l + \Delta] \text{ for some } j \leq J \Big) \bigg)$$

$$\leq \|g\| \cdot 3\epsilon.$$

The last inequality follows from our choice of $J = J(\epsilon)$, $N = N(\epsilon)$, and $\Delta = \Delta(\epsilon)$ at the beginning of the proof. From the arbitrariness of the mapping g and $\epsilon > 0$, we conclude the proof with Portmanteau theorem. □

5.4 Proof of Propositions 2.14 and 2.15

In this section, we fix some $b \in (0, \infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{\text{min}}]$ and $j \in [n_{\text{min}} - 1]$. This allows us to fix some $\bar{\epsilon} \in (0, 1 \wedge b)$ such that (5.1) holds.

In our proof of Propositions 2.14 and 2.15, the key tools are the first exit analysis results, i.e., Theorem 2.6 and technical lemmas developed in Section 4.3. Note that Theorem 2.6 is applied on some open interval I with bounded support. Returning to the potential U characterized in Assumption 7, while for all $i = 2, \cdots, n_{\text{min}}$ the attraction field I_i does have bounded support, for $i = 1$ or n_{min} (that is, the leftmost or the rightmost attraction field) note that $I_1 = (-\infty, s_1)$ and $I_{n_{\text{min}}} = (s_{n_{\text{min}}-1}, \infty)$ are not bounded. Besides, for technical reasons, in our analysis below we will bound the probability that the heavy-tailed dynamics visit $S(\delta) \triangleq \bigcup_{i \in [n_{\text{min}}-1]} [s_i - \delta, s_i + \delta]$ (i.e., the union of the δ-neighborhood of any the boundary point s_i) and show that $X_j^{\eta|b}(x)$ is almost always outside of $S(\delta)$. As a result, we will frequently apply results such as Theorem 2.6 onto sets of form

$$I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M, M) = (I_i)_\delta \cap (-M, M)$$

for some $\delta, M > 0$. For any $M > 0$ large enough such that $-M < m_1 < s_1 < \cdots < s_{n_{\text{min}}-1} < m_{n_{\text{min}}} < M$, we have $I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M, M) = (s_{i-1} + \delta, s_i - \delta)$ for all $i = 2, 3, \cdots, n_{\text{min}} - 1$ (i.e., any attraction field that is not the leftmost or the rightmost one); also, we have $I_{1;\delta,M} = (s_0 + \delta, s_1 - \delta) \cap (-M, M) = (-M, s_1 - \delta)$ (due to $s_0 = -\infty$) and $I_{n_{\text{min}};\delta,M} = (s_{n_{\text{min}}-1} + \delta, s_{n_{\text{min}}} - \delta) \cap (-M, M) = (s_{n_{\text{min}}-1} + \delta, M)$ (due to $s_{n_{\text{min}}} = \infty$).

We first prepare a technical lemma and show that, during any transition between the attraction fields, $X_j^{\eta|b}(x)$ is unlikely to get too close to any of the boundary points s_i's or exit a wide enough compact set. Let

$$\sigma_{i;\epsilon}^{\eta|b}(x) \triangleq \min\Big\{ j \geq 0 : \ X_j^{\eta|b}(x) \in \bigcup_{l \neq i} (m_l - \epsilon, m_l + \epsilon) \Big\}, \tag{5.8}$$

$$\tau_{i;\delta,M}^{\eta|b}(x) \triangleq \min\Big\{ j \geq 0 : \ X_j^{\eta|b}(x) \notin I_{i;\delta,M} \Big\}. \tag{5.9}$$

Lemma 5.1. *Let Assumptions 1, 2, 3, 4 and 7 hold. Let $b \in (0, \infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j = 0, 1, \cdots, n_{min}$. There exists $M > 0$ such that*

$$\max_{i \in [n_{min}]} \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big((-M, M)^c; m_i \big) = 0, \tag{5.10}$$

Furthermore, given any $\Delta > 0$, it holds for all $\delta > 0$ small enough and all $\epsilon > 0$ small enough that

$$\limsup_{\eta \downarrow 0} \max_{i \in [n_{min}]} \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\Big(\exists j < \sigma_{i;\epsilon}^{\eta|b}(x) \ s.t. \ X_j^{\eta|b}(x) \in S(\delta) \ or \ \big| X_j^{\eta|b}(x) \big| \geq M + 1 \Big) < \Delta. \tag{5.11}$$

Proof. In light of Lemma 4.6, it holds for all $M > 0$ large enough such that

$$\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big((-M,M)^c; m_i\big) = 0 \qquad \forall i \in [n_{\min}].$$

This concludes the proof of (5.10).

Henceforth in this proof, we fix such large M satisfying $|M - m_i|/b \notin \mathbb{Z}\ \forall i \in [n_{\min}]$ and $M > \max_{i\in[n_{\min}]}(\mathcal{J}_b^*(i) - 1)b + \bar{\epsilon}$, where $\bar{\epsilon} > 0$ is the constant in (5.1). Also, we fix some $\epsilon \in (0, \bar{\epsilon})$, and show that (5.11) holds for such ϵ.

Recall the definition of $\tau^{\eta|b}_{i;\delta,M}(x)$ in (5.9) and $I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M, M)$. We make a few observations regarding the stopping time $\tau^{\eta|b}_{i;2\delta,M}(x) = \min\{j \geq 0 :\ X^{\eta|b}_j(x) \notin I_{i;2\delta,M}\}$ with $\delta > 0$. First, due to $I_{i;2\delta,M} \subseteq I_{i;\delta,M}$, we must have $\tau^{\eta|b}_{i;2\delta,M}(x) \leq \tau^{\eta|b}_{i;\delta,M}(x) \leq \sigma^{\eta|b}_{i;\epsilon}(x)$ and $X^{\eta|b}_j(x) \notin S(\delta)$, $|X^{\eta|b}_j(x)| < M$ for all $j < \tau^{\eta|b}_{i;2\delta,M}(x)$. Next, on event

$$A_0(\eta,\delta,x) \triangleq \Big\{X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in (-M,M);\ X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \notin S(2\delta)\Big\},$$

there exists some $j \in [n_{\min}],\ j \neq i$ such that $X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in I_{j;2\delta,M}$. Now define

$$A_1(\eta,\delta,x) \triangleq \Big\{\exists j < \sigma^{\eta|b}_{i;\epsilon}(x)\ s.t.\ X^{\eta|b}_j(x) \in S(\delta)\Big\},\ A_2(\eta,x) \triangleq \Big\{\exists j < \sigma^{\eta|b}_{i;\epsilon}(x)\ s.t.\ |X^{\eta|b}_j(x)| \geq M+1\Big\}.$$

Let $R^{\eta|b}_{j;\epsilon}(x) \triangleq \min\{k \geq 0 :\ X^{\eta|b}_k(x) \in (m_j - \epsilon, m_j + \epsilon)\}$. From the strong Markov property at $\tau^{\eta|b}_{i;2\delta,M}(x)$,

$$\begin{aligned}
&\max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\bigg(\Big(A_1(\eta,\delta,x) \cup A_2(\eta,x)\Big) \cap A_0(\eta,\delta,x)\bigg) \\
&\leq \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\bigg(A_1(\eta,\delta,x) \cup A_2(\eta,x)\ \bigg|\ A_0(\eta,\delta,x)\bigg) \\
&\leq \max_{j\in[n_{\min}]} \underbrace{\sup_{y\in[s_{j-1}+2\delta,s_j-2\delta]\cap(-M,M)} \mathbf{P}\bigg(\Big\{X^{\eta|b}_k(x) \in [s_{j-1}+\delta, s_j-\delta] \cap (-M-1, M+1)\ \forall\exists k < R^{\eta|b}_{j;\epsilon}(x)\Big\}^c\bigg)}_{p_j(\eta)}.
\end{aligned}$$

For any $j \in [n_{\min}]$ and any $\delta > 0$ small enough, b y applying Lemma 4.9 onto $I_j \cap (-M-1, M+1)$ (with parameter ϵ therein set as 2δ) we get $\lim_{\eta\downarrow 0} p_j(\eta) = 0$. In summary, we have shown that $\limsup_{\eta\downarrow 0}\max_{i\in[n_{\min}]}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\big(A_1(\eta,\delta,x)\cup A_2(\eta,x)\big)\cap A_0(\eta,\delta,x)\Big) = 0$. Meanwhile, to establish (5.11) it only remains to show that $\limsup_{\eta\downarrow 0}\max_{i\in[n_{\min}]}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\big(A_0(\eta,\delta,x)\big)^c\Big) < \Delta$. This can be proved if we show that for all $\delta > 0$ small enough,

$$\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \notin (-M,M)\Big) = 0,$$

$$\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in S(2\delta)\Big) < \Delta.$$

To proceed, we fix some $i \in [n_{\min}]$. Note that $I_{i;\delta,M} \subset I_i$ and hence $I^c_{i;\delta,M} \supset I^c_i$. First, our choice of M at the beginning ensures $\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}((-M,M)^c; m_i) = 0$. By applying part (a) of Theorem 2.6 onto $I_{i;\delta,M}$, we obtain

$$\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \notin (-M,M)\Big) = 0.$$

Next, recall the assumption $|s_j - m_i|/b \notin \mathbb{Z}$ for all $j \in [n_{\min} - 1]$. From Lemma 4.3, we get $\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(\{s_1, \cdots, s_{n_{\min}}\}; m_i\big) = 0$, which then implies $\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(S(2\delta); m_i\big) < q_b(i) \cdot \Delta$ for all $\delta > 0$ small enough. Meanwhile, due to $I_{i;2\delta,M} \subset I_i$, we have $I^c_{i;2\delta,M} \supset I^c_i$ and hence $\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\Big(I^c_{i;2\delta,M}; m_i\Big) \geq q_b(i)$. Applying part (a) of Theorem 2.6 again, we yield (for all $\delta > 0$ small enough)

$$\limsup_{\eta \downarrow 0} \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in S(2\delta) \bigg) \leq \frac{\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\Big(S(2\delta); m_i\Big)}{\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\Big(I^c_{i;2\delta,M}; m_i\Big)} < \Delta.$$

This concludes the proof of (5.11). □

The next result is an adaptation of the first exit time analysis in Section 2.3 to the current setup.

Proposition 5.2. *Let Assumptions 1, 2, 3, 4 and 7 hold. Let $b \in (0, \infty)$ be such that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j = 0, 1, \cdots, n_{min}$. There exists $\bar{\epsilon} > 0$ such that the following claims hold.*

(i) Let $R^{\eta|b}_{i;\epsilon}(x) \triangleq \min\{j \geq 0 :\ X^{\eta|b}_j(x) \in (m_i - \epsilon, m_i + \epsilon)\}$. For any $\epsilon \in (0, \bar{\epsilon})$, $t > 0$ and $i \in [n_{min}]$,

$$\liminf_{\eta \downarrow 0} \inf_{x \in [s_{i-1} + \epsilon, s_i - \epsilon]} \mathbf{P}\bigg(R^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_j(x) \in I_i\ \forall j \leq R^{\eta|b}_{i;\epsilon}(x) \bigg) = 1.$$

(ii) Let $i, j \in [n_{min}]$ be such that $i \neq j$. Let $\sigma^{\eta|b}_{i;\epsilon}(x) \triangleq \min\{j \geq 0 :\ X^{\eta|b}_j(x) \in \bigcup_{l \neq i}(m_l - \epsilon, m_l + \epsilon)\}$. If $m_i \in V^*_b$, then for any $\epsilon \in (0, \bar{\epsilon})$ and any $t \geq 0$,

$$\liminf_{\eta \downarrow 0} \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \geq \exp\big(- q_b(i) \cdot t \big) \cdot \frac{q_b(i,j)}{q_b(i)},$$
$$\limsup_{\eta \downarrow 0} \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \leq \exp\big(- q_b(i) \cdot t \big) \cdot \frac{q_b(i,j)}{q_b(i)}.$$

If $m_i \notin V^*_b$, then for any $\epsilon \in (0, \bar{\epsilon})$ and any $t \geq 0$,

$$\begin{aligned} \frac{q_b(i,j)}{q_b(i)} &\leq \liminf_{\eta \downarrow 0} \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \\ &\leq \limsup_{\eta \downarrow 0} \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \leq \frac{q_b(i,j)}{q_b(i)}. \end{aligned}$$

Proof. Throughout this proof, the constant $\bar{\epsilon} \in (0, 1 \wedge b)$ is as specified in (5.1).

(i) Fix some $\epsilon \in (0, \bar{\epsilon})$. Pick some $M > 0$ large enough such that $|M| > \sup_{x \in [s_{i-1}+\epsilon, s_i - \epsilon]} |x| + \bar{\epsilon}$. Let

$$\boldsymbol{t}_i(x, \epsilon) \triangleq \inf\{t \geq 0 :\ \boldsymbol{y}_t(x) \in (m_i - \epsilon, m_i + \epsilon)\} \tag{5.12}$$

where $\boldsymbol{y}_\cdot(x)$ solves the ODE $d\boldsymbol{y}_t(x)/dt = -U'(\boldsymbol{y}_t(x))$ and initial condition $\boldsymbol{y}_0(x) = x$. Set

$$T = \sup\Big\{ \boldsymbol{t}_i(x, \frac{\epsilon}{2}) :\ x \in [-M + \epsilon, M - \epsilon] \cap [s_{i-1} + \epsilon, s_i - \epsilon] \Big\}$$

By Assumption 7, we have $\boldsymbol{t}_i(x, \frac{\epsilon}{2}) < \infty$ for all $x \in [-M + \epsilon, M - \epsilon] \cap [s_{i-1} + \epsilon, s_i - \epsilon]$, with $\boldsymbol{t}_i(\ \cdot\ , \frac{\epsilon}{2})$ being continuous over $x \in [-M + \epsilon, M - \epsilon] \cap [s_{i-1} + \epsilon, s_i - \epsilon]$. This implies $T < \infty$. Next, recall that $\lambda^*_b(\eta) \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)+1}(\eta)$ as $\eta \downarrow 0$. Due to $\mathcal{J}^*_b(V) \geq 1$, we have $\mathcal{J}^*_b(V) \cdot (\alpha - 1) + 1 > 1$. This implies $\frac{t}{\lambda^*_b(\eta)} > \frac{T}{\eta}$ for all $\eta > 0$ sufficiently small, and hence (for such small η)

$$\inf_{x \in [s_{i-1} + \epsilon, s_i - \epsilon]} \mathbf{P}\bigg(R^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_j(x) \in I_i\ \forall j \leq R^{\eta|b}_{i;\epsilon}(x) \bigg)$$

$$\geq \inf_{x \in [s_{i-1}+\epsilon, s_i - \epsilon]} \mathbf{P}\bigg(R^{\eta|b}_{i;\epsilon}(x) \leq T/\eta,\ X^{\eta|b}_j(x) \in I_i\ \forall j \leq R^{\eta|b}_{i;\epsilon}(x) \bigg).$$

By applying Lemma 4.9 onto $(-M,M)\cap I_j$, we conclude the proof of part (i).

(ii) Recall that $\lambda^*_{i;b}(\eta) \triangleq \eta \cdot \lambda^{\mathcal{J}^*_b(i)}(\eta)$. To prove claims in part (ii), It suffices to establish the following upper and lower bounds: for all $i,j \in [n_{\min}]$ such that $i \neq j$, all $\epsilon \in (0,\bar{\epsilon})$, and all $t \geq 0$,

$$\liminf_{\eta \downarrow 0} \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \geq \exp\big(- q_b(i) \cdot t \big) \cdot \frac{q_b(i,j)}{q_b(i)}, \tag{5.13}$$

$$\limsup_{\eta \downarrow 0} \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \leq \exp\big(- q_b(i) \cdot t \big) \cdot \frac{q_b(i,j)}{q_b(i)}. \tag{5.14}$$

To see why, we first consider the case with $m_i \in V^*_b$, which implies $\mathcal{J}^*_b(i) = \mathcal{J}^*_b(V)$; see (2.44) for the definition of $\mathcal{J}^*_b(V)$. As a result, we have $\lambda^*_{i;b}(\eta) = \lambda^*_b(\eta) \triangleq \eta \cdot \lambda^{\mathcal{J}^*_b(V)}(\eta)$ and the upper and lower bounds in part (ii) follow immediately from the (5.13) and (5.14).

Next, in case that $m_i \notin V^*_b$, we have $\mathcal{J}^*_b(i) < \mathcal{J}^*_b(V)$, and hence $\frac{\lambda^*_{i;b}(\eta)}{\lambda^*_b(\eta)} \to \infty$ as $\eta \downarrow 0$. If $t = 0$, then the upper and lower bounds in part (ii) are still immediate consequences of (5.13) and (5.14). Now, we focus on the case where $t > 0$ and start from the lower bound. Given any $T > 0$, we have $t \cdot \frac{\lambda^*_{i;b}(\eta)}{\lambda^*_b(\eta)} > T$ eventually for all η small enough. Therefore,

$$\begin{aligned}
&\inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \\
&\leq \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_{i;b}(\eta) \leq T,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \\
&= \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) - \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_{i;b}(\eta) > T,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg).
\end{aligned}$$

Applying (5.13) and (5.14), we get

$$\liminf_{\eta \downarrow 0} \inf_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \geq \frac{q_b(i,j)}{q_b(i)} \cdot \Big(1 - \exp\big(- q_b(i) \cdot T \big) \Big).$$

Let T tend to ∞, and we conclude the proof of the lower bound in part (ii) for the case where $m_i \notin V^*_b$. As for the upper bound, note that

$$\sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg) \leq \sup_{x \in [m_i - \epsilon, m_i + \epsilon]} \mathbf{P}\bigg(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg).$$

Applying (5.14) with $t = 0$, we conclude the proof of the upper bound in the case where $m_i \notin V^*_b$.

The rest of this proof is devoted to establishing (5.13) and (5.14). Here, we state one fact that will be applied in the analysis below. By assumption $|s_j - m_i|/b \notin \mathbb{Z}$ for all $j \in [n_{\min} - 1]$, one can apply Lemma 4.3 and obtain $\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(\{s_1, \cdots, s_{n_{\min}-1}\}; m_i\big) = 0$. Due to $I_j = (s_{j-1}, s_j)$, we then have

$$\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I_j; m_i \Big) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^-_j; m_i \Big) \qquad \forall i,j \in [n_{\min}] \text{ with } i \neq j. \tag{5.15}$$

Proof of Lower Bound (5.13).

Recall that $I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M,M)$ and $\tau^{\eta|b}_{i;\delta,M}(x) = \min\{k \geq 0 :\ X^{\eta|b}_k(x) \notin I_{i;\delta,M}\}$. Now, observe that (for any $\delta > 0$)

$$\bigg\{ \sigma^{\eta|b}_{i;\epsilon}(x) \cdot \lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \bigg\}$$

$$\supseteq \underbrace{\Big\{ \tau^{\eta|b}_{i;\delta,M}(x)\cdot \lambda^*_{i;b}(\eta) > t;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{j;\delta,M+1} \Big\}}_{\text{(I)}} \cap \underbrace{\Big\{ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \Big\}}_{\text{(II)}}.$$

Given any $T > 0$, strong Markov property at $\tau^{\eta|b}_{i;\delta,M}(x)$ implies that (for all $\eta > 0$ small enough)

$$\begin{aligned}
\inf_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(\text{(II)}\ \Big|\ \text{(I)}\Big) &\geq \inf_{y \in I_{j;\delta,M+1}} \mathbf{P}\Big(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(y)}(y) \in I_j \Big) \\
&\geq \inf_{y \in I_{j;\delta,M+1}} \mathbf{P}\Big(R^{\eta|b}_{j;\epsilon}(y) \leq T/\eta;\ X^{\eta|b}_k(y) \in I_j\ \forall k \leq R^{\eta|b}_{j;\epsilon}(y)\Big).
\end{aligned}$$

Recall the definition of $\boldsymbol{t}_j(x,\epsilon)$ in (5.12), and set $T = \sup\Big\{\boldsymbol{t}_j(x,\frac{\epsilon}{2}):\ x \in [-M-1,M+1]\cap[s_{j-1}+\delta, s_j-\delta]\Big\} < \infty$. By applying Lemma 4.9 again, we yield

$$\begin{aligned}
&\liminf_{\eta\downarrow 0} \inf_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(\text{(II)}\ \Big|\ \text{(I)}\Big) \\
&\geq \liminf_{\eta\downarrow 0} \inf_{y \in I_{j;\delta,M+1}} \mathbf{P}\Big(R^{\eta|b}_{j;\epsilon}(y) \leq T/\eta;\ X^{\eta|b}_k(y) \in I_j\ \forall k \leq R^{\eta|b}_{j;\epsilon}(y)\Big) = 1.
\end{aligned} \tag{5.16}$$

Next, we move onto the analysis of event (I). Let $M \in (0,\infty)$ be such that the claim (5.10) of Lemma 5.1 holds. Fix some $\Delta > 0$. Meanwhile, by assumption $|s_j - m_i|/b \notin \mathbb{Z}$ for all $j \in [n_{\min}-1]$, one can apply Lemma 4.3 and obtain $\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(\{s_1,\cdots,s_{n_{\min}-1}\};m_i\big) = 0$. Due to the continuity of measures, it then holds for all $\delta > 0$ small enough that (recall that $q_b(i) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_i^c;m_i)$; see (5.2))

$$\begin{aligned}
&\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big((s_{i-1}-\delta, s_i+\delta)^c; m_i\Big) \\
&\leq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I_i^c;m_i\big) + \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big([s_{i-1}-\delta,s_{i-1}+\delta]\cup[s_i-\delta,s_i+\delta];m_i\Big) \\
&< (1+\Delta)\cdot q_b(i).
\end{aligned}$$

Therefore, for the set $I^c_{i;\delta,M} = \big((-M,M)\cap(s_{i-1}+\delta,s_i-\delta)\big)^c$, it holds for all $\delta > 0$ small enough that

$$\begin{aligned}
\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^c_{i;\delta,M};m_i\Big) &\leq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big((-M,M)^c;m_i\Big) + \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big((s_{i-1}+\delta,s_i-\delta)^c;m_i\Big) \\
&= 0 + \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big((s_{i-1}+\delta,s_i-\delta)^c;m_i\Big) \qquad \text{using (5.10)} \\
&< (1+\Delta)\cdot q_b(i).
\end{aligned} \tag{5.17}$$

On the other hand, recall the definition of $q_b(i,j) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j;m_i)$ in (5.2), and note that

$$\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I^c_{i;\delta,M};m_i\big) \geq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I^c_i;m_i\big) = q_b(i) \tag{5.18}$$

due to $I_{i;\delta,M} \subseteq I_i$ and hence $I^c_{i;\delta,M} \supseteq I^c_i$. Next, observe that

$$\text{(I)} = \underbrace{\Big\{ \tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x))}(x) \in I_j \Big\}}_{\text{(III)}} \cap \underbrace{\Big\{ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{j;\delta,M+1} \Big\}}_{\text{(IV)}}.$$

By applying part (a) of Theorem 2.6 onto $I_{i;\delta,M}$, we yield (for any δ small enough)

$$\liminf_{\eta\downarrow 0} \inf_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(\text{(III)}\Big) \geq \exp\Big(- \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^c_{i;\delta,M};m_i\Big)\cdot t\Big) \cdot \frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I_j;m_i\Big)}{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^c_{i;\delta,M};m_i\Big)}$$

$$> \frac{\exp\Big(-(1+\Delta)q_b(i)\cdot t\Big)}{1+\Delta}\cdot\frac{q_b(i,j)}{q_b(i)}.$$

In the last inequality, we applied (5.17) and (5.18). On the other hand, due to $\tau^{\eta|b}_{i;\delta,M}(x) \leq \sigma^{\eta|b}_{i;\epsilon}(x)$, we obtain $\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\text{IV})^c\Big) < \Delta$ for all $\delta > 0$ small enough by applying (5.11) of Lemma 5.1. In summary, for all $\delta > 0$ small enough,

$$\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\text{I})\Big) \geq \frac{\exp\Big(-(1+\Delta)q_b(i)\cdot t\Big)}{1+\Delta}\cdot\frac{q_b(i,j)}{q_b(i)} - \Delta. \tag{5.19}$$

Combining (5.16) and (5.19), we yield $\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\Big) \geq \frac{\exp(-(1+\Delta)q_b(i)\cdot t)}{1+\Delta}\cdot\frac{q_b(i,j)}{q_b(i)} - \Delta$. Let $\Delta \downarrow 0$ and we conclude the proof of the lower bound.

Proof of Upper Bound (5.14).

Let $(\text{I}) = \Big\{\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\Big\}$. Given some $M > 0$ and $\delta > 0$, define event $(\text{II}) = \Big\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in (-M-1,M+1)\setminus S(\delta)\Big\}$. We start from the decomposition $(\text{I}) = \big((\text{I})\setminus(\text{II})\big)\cup\big((\text{I})\cap(\text{II})\big)$. First, arbitrarily pick some $\Delta > 0$, and let $M \in (0,\infty)$ be such that the claim (5.10) of Lemma 5.1 holds. The claim

$$\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\text{I})\setminus(\text{II})\Big) \leq \limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\text{II})^c\Big) < \Delta \tag{5.20}$$

for all $\delta > 0$ small enough follows directly from (5.11) of Lemma 5.1. Next, on event $(\text{I})\cap(\text{II})$, there exists some $K \in [n_{\min}]$, $K \neq i$ such that $X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in (-M-1,M+1)\cap(s_{K-1}+\delta, s_K-\delta) = I_{K;\delta,M+1}$. For each $k \in [n_{\min}]$ with $k \neq i$, define event

$$(k) = (\text{I})\cap(\text{II})\cap\Big\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{k;\delta,M+1}\Big\}$$

and note that $\bigcup_{k\in[n_{\min}]:\ k\neq i}(k) = (\text{I})\cap(\text{II})$. To proceed, consider the following decomposition

$$(k) = \underbrace{\bigg((k)\cap\Big\{\Big(\sigma^{\eta|b}_{i;\epsilon}(x)-\tau^{\eta|b}_{i;\delta,M}(x)\Big)\cdot\lambda^*_{i;b}(\eta) > \Delta\Big\}\bigg)}_{(k,1)}\cup\underbrace{\bigg((k)\cap\Big\{\Big(\sigma^{\eta|b}_{i;\epsilon}(x)-\tau^{\eta|b}_{i;\delta,M}(x)\Big)\cdot\lambda^*_{i;b}(\eta) \leq \Delta\Big\}\bigg)}_{(k,2)}.$$

We fix some $k \in [n_{\min}]$ with $k \neq i$ and analyze the probability of events $(k,1)$ and $(k,2)$ separately. First, recall that $\lambda^*_{i;b}(\eta) = \eta\cdot\lambda^{\mathcal{J}^*_b(i)}(\eta) \in \mathcal{RV}_{\mathcal{J}^*_b(i)\cdot(\alpha-1)+1}(\eta)$, so given any $T > 0$ it holds for all $\eta > 0$ small enough that $\frac{\Delta}{\lambda^*_{i;b}(\eta)} > \frac{T}{\eta}$. Now, we pick $T = \sup\big\{\boldsymbol{t}_k(x,\frac{\epsilon}{2}):\ x \in I^-_{k;\delta,M+1}\big\}$ with $\boldsymbol{t}_k(\cdot,\cdot)$ defined in (5.12), and observe that

$$\begin{aligned}
&\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((k,1)\Big)\\
&\leq \limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((k)\cap\Big\{\sigma^{\eta|b}_{i;\epsilon}(x)-\tau^{\eta|b}_{i;\delta,M}(x) > T/\eta\Big\}\Big)\\
&\leq \limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{k;\delta,M+1};\ \sigma^{\eta|b}_{i;\epsilon}(x)-\tau^{\eta|b}_{i;\delta,M}(x) > T/\eta\Big)\\
&\leq \limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M+1}}\mathbf{P}\Big(\sigma^{\eta|b}_{i;\epsilon}(y) > T/\eta\Big) \qquad \text{due to strong Markov property at } \tau^{\eta|b}_{i;\delta,M}(x)
\end{aligned}$$

$$\le \limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M+1}} \mathbf{P}\Big(X^{\eta|b}_j(y)\notin(m_k-\epsilon,m_k+\epsilon)\ \forall j\le T/\eta\Big)$$
$$= 0 \qquad \text{due to Lemma 4.9.} \tag{5.21}$$

Next, for all $k\neq i$,

$$\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((k,2)\Big)$$
$$\le \sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta)>t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_{k;\delta,M+1};\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x)\in I_j\Big)$$
$$\le \sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta)>t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_k\Big)$$
$$\cdot\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x)\in I_j\ \Big|\ \tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta)>t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_{k;\delta,M+1}\Big)$$
$$\le \sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\underbrace{\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta)>t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_k}_{(k,\mathrm{I})}\Big)\cdot\sup_{y\in I_{k;\delta,M+1}}\mathbf{P}\Big(\underbrace{X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(y)}(y)\in I_j}_{(k,\mathrm{II})}\Big).$$

In the last inequality we applied the strong Markov property at $\tau^{\eta|b}_{i;\delta,M}(x)$. Applying part (a) of Theorem 2.6 onto $I_{i;\delta,M}$ and the bound (5.18), we yield (for any δ small enough)

$$\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((k,\mathrm{I})\Big)\le\exp\Big(-\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^c_{i;\delta,M};m_i\Big)\cdot(t-\Delta)\Big)\cdot\frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^-_k;m_i\Big)}{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\Big(I^c_{i;\delta,M};m_i\Big)}$$
$$\le\exp\big(-q_b(i)\cdot(t-\Delta)\big)\cdot\frac{q_b(i,k)}{q_b(i)}. \tag{5.22}$$

Here, we also applied (5.15) to show that $\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I^-_k;m_i\big)=\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I_k;m_i\big)=q_b(i,k)$. Next, we analyze the probability of event (k,II). If $k=j$, we apply the trivial upper bound $\mathbf{P}((k,\mathrm{II}))\le 1$. If $k\neq j$, recall that $R^{\eta|b}_{k;\epsilon}(x)=\min\{n\ge 0:\ X^{\eta|b}_n(x)\in(m_k-\epsilon,m_k+\epsilon)\}$ is the first time $X^{\eta|b}_n(x)$ visits the ϵ-neighborhood of m_k, and note that

$$\sup_{y\in I_{k;\delta,M+1}}\mathbf{P}((k,\mathrm{II}))\le\sup_{y\in I_{k;\delta,M+1}}\mathbf{P}\Big(\exists n<R^{\eta|b}_{k;\epsilon}(y)\ s.t.\ X^{\eta|b}_n(y)\notin I_{k;\frac{\delta}{2},M+2}\Big).$$

Indeed, on event (k,II), the first local minimum visited by $X^{\eta|b}_n(y)$ is m_j even though the initial value $X^{\eta|b}_0(y)=y$ belongs to $I_{k;\delta,M+1}\subset I_k$; this implies that $X^{\eta|b}_n(y)$ must have left I_k (and hence $I_{k;\frac{\delta}{2},M+2}$) before visiting the neighborhood of m_k. Applying Lemma 4.9 onto $I_k\cap(-M-2,M+2)$ (with the parameter ϵ therein set as δ), we obtain $\limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M+1}}\mathbf{P}((k,\mathrm{II}))=0$ for all $\delta>0$ small enough in the case of $k\neq j$. Combining this result with (5.20), (5.21), and (5.22), we yield that

$$\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta)>t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x)\in I_j\Big)$$
$$\le\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\mathrm{II})^c\Big)$$
$$+\sum_{k\in[n_{\min}]:\ k\neq i}\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((k,\mathrm{I})\Big)\cdot\limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M}}\mathbf{P}\Big((k,\mathrm{II})\Big)$$
$$\le\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((\mathrm{II})^c\Big)+\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big((j,\mathrm{I})\Big)$$

$$\leq \Delta + \exp\big(-q_b(i)\cdot(t-\Delta)\big)\cdot\frac{q_b(i,j)}{q_b(i)}.$$

Let $\Delta \downarrow 0$ and we conclude the proof of the upper bound. $\square$

Now, we are ready to prove Proposition 2.14.

Proof of Proposition 2.14. Recall that $Y^{*|b}_t(m_i)$ is a $\Big((U_j)_{j\geq 1},(V_j)_{j\geq 1}\Big)$ jump process with $U_1=0$, $V_i=m_i$, and the law of $\Big((U_j)_{j\geq 2},(V_j)_{j\geq 2}\Big)$ specified in (5.3). To prove the weak convergence claim in terms of finite dimensional distributions, it suffices to verify the conditions in Lemma 2.13. In particular, since $Y^{*|b}_t(m_i)$ is a continuous-time Markov chain that is irreducible with finitely many states, the only condition we need to check is the following: Given $\epsilon>0$ and $\eta>0$, let $U^\eta_k = \Big(\big(\hat{\tau}^{\eta,\epsilon|b}_k(x)-\hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)\big)\cdot\lambda^*_b(\eta)$ and $V^\eta_k = m_{\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)}$ (for definitions, see (2.46)–(2.48)); it holds for all $\epsilon>0$ small enough that $(U^\eta_1,V^\eta_1,U^\eta_2,V^\eta_2,\cdots)$ converges in distribution to $(U_1,V_1,U_2,V_2,\cdots)$ as $\eta\downarrow 0$. This is equivalent to proving that, for each $N\geq 1$, $(U^\eta_1,V^\eta_1,\cdots,U^\eta_N,V^\eta_N)$ converges in distribution to $(U_1,V_1,\cdots,U_N,V_N)$ as $\eta\downarrow 0$.

Fix some $N=1,2,\cdots$. First, from part (i) of Proposition 5.2, we get $\big(U^\eta_1,V^\eta_1\big)\Rightarrow(0,m_i)=(U_1,V_1)$ as $\eta\downarrow 0$. Next, for any $n\geq 1$, any $t_l\in(0,\infty)$, any $v_l\in\{m_i:\ i\in[n_{\min}]\}$, and $t>0$, $i,j\in[n_{\min}]$ with $i\neq j$, it follows directly from part (ii) of Proposition 5.2 that

$$\begin{aligned}&\lim_{\eta\downarrow 0}\mathbf{P}\bigg(U^\eta_{n+1}\leq t,\ V^\eta_{n+1}=m_j\ \bigg|\ V^\eta_n=m_i,\ V^\eta_l=v_l\ \forall l\in[n-1],\ U^\eta_l\leq t_l\ \forall l\in[n]\bigg)\\ &=\begin{cases}\frac{q_b(i,j)}{q_b(i)} & \text{if } m_i\notin V^*_b,\\ \frac{q_b(i,j)}{q_b(i)}\cdot\Big(1-\exp\big(-q_b(i)t\big)\Big) & \text{if } m_i\in V^*_b.\end{cases}\end{aligned}$$

This coincides with the conditional law of $\mathbf{P}\Big(U_{n+1}<t,\ V_{n+1}=m_j\ \Big|\ V_n=m_i,\ (V_j)_{j=1}^{n-1},\ (U_j)_{j=1}^n\Big)$ specified in (5.3). By arguing inductively, we conclude the proof. $\square$

Lastly, we give the proof of Proposition 2.15.

Proof of Proposition 2.15. If $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)\in\bigcup_{l\in[n_{\min}]}(m_l-\epsilon,m_l+\epsilon)$, then due to the definition of $\hat{X}^{\eta,\epsilon|b}_t(x)$ as the marker of the *last visited local minimum* (under time scaling of $\lambda^*_b(\eta)$; see (2.46)-(2.48) for the definition of the process $\hat{X}^{\eta,\epsilon|b}_t(x)$), we must have $|X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)-\hat{X}^{\eta,\epsilon|b}_t(x)|<\epsilon$. Therefore, it suffices to show that for any $\epsilon\in(0,\bar{\epsilon})$ (with $\bar{\epsilon}$ specified in (5.1))

$$\lim_{\eta\downarrow 0}\mathbf{P}\Big(X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)\in\bigcup_{l\in[n_{\min}]}(m_l-\epsilon,m_l+\epsilon)\Big)=1.$$

To proceed, pick some $\delta_t\in(0,\frac{t}{3})$, $\delta>0$, and $M>0$. Recall that $H(\cdot)=\mathbf{P}(|Z_1|>\cdot)$ and $S(\delta)=\bigcup_{i\in[n_{\min}-1]}[s_i-\delta,s_i+\delta]$. Define event

$$(\mathrm{I})=\Big\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor-\lfloor 2\delta_t/H(\eta^{-1})\rfloor}(x)\in(-M,M)\backslash S(\delta)\Big\}.$$

Let $t_1(\eta)=\lfloor t/\lambda^*_b(\eta)\rfloor-\lfloor 2\delta_t/H(\eta^{-1})\rfloor$. On event (I), let $R^\eta\triangleq\min\{j\geq t_1(\eta):\ X^{\eta|b}_j(x)\in\bigcup_{l\in[n_{\min}]}(m_l-\frac{\epsilon}{2},m_l+\frac{\epsilon}{2})\}$ and set $\hat{\mathcal{I}}^\eta$ by the rule $\hat{\mathcal{I}}^\eta=j\iff X^{\eta|b}_{R^\eta}(x)\in I_j$. Now we can define event

$$(\mathrm{II})=\Big\{R^\eta-t_1(\eta)\leq\delta_t/H(\eta^{-1})\Big\}.$$

On event (I) ∩ (II) we have $\lfloor t/\lambda^*_b(\eta)\rfloor - \lfloor 2\delta_t/H(\eta^{-1})\rfloor \leq R^\eta \leq \lfloor t/\lambda^*_b(\eta)\rfloor$. Let $\tau^\eta \triangleq \min\{j \geq R^\eta : X^{\eta|b}_j(x) \notin (m_{\hat{\mathcal{I}}^\eta} - \epsilon, m_{\hat{\mathcal{I}}^\eta} + \epsilon)\}$, and define event

$$\text{(III)} = \Big\{\tau^\eta - R^\eta > 2\delta_t/H(\eta^{-1})\Big\}.$$

On event (I) ∩ (II) ∩ (III), we have $\tau^\eta > \lfloor t/\lambda^*_b(\eta)\rfloor \geq R^\eta$, and hence $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x) \in \bigcup_{l\in[n_{\min}]}(m_l - \epsilon, m_l + \epsilon)$. Furthermore, we claim that for any $\Delta > 0$ there exist $\delta_t \in (0, \frac{t}{3})$, $\delta > 0$, and $M > 0$ such that

$$\liminf_{\eta\downarrow 0} \mathbf{P}\Big(\text{(I)}\Big) \geq 1 - \Delta, \tag{5.23}$$

$$\liminf_{\eta\downarrow 0} \mathbf{P}\Big(\text{(II)} \;\Big|\; \text{(I)}\Big) \geq 1, \tag{5.24}$$

$$\liminf_{\eta\downarrow 0} \mathbf{P}\Big(\text{(III)} \;\Big|\; \text{(I)} \cap \text{(II)}\Big) \geq 1 - \Delta. \tag{5.25}$$

An immediate consequence is that $\liminf_{\eta\downarrow 0} \mathbf{P}(\text{(I)}\cap\text{(II)}\cap\text{(III)}) \geq (1-\Delta)^2$. Let $\Delta \downarrow 0$ and we conclude the proof. Now it only remains to establish (5.23) (5.24) (5.25). Throughout the remainer of this proof, we fix some $\epsilon \in (0, \bar{\epsilon})$ and $\Delta > 0$.

Proof of (5.23).

Let $I_{M,\delta} = (-M, M)\backslash S(\delta)$. Recall the definition of $\hat{\tau}^{\eta,\epsilon|b}_j(x)$ in (2.46)(2.47). For any $N \in \mathbb{Z}_+$, on event

$$\bigg(\bigcap_{k=1}^{N-1} \underbrace{\Big\{X^{\eta|b}_j(x) \in I_{M,\delta}\ \forall j \in \big[\hat{\tau}^{\eta,\epsilon|b}_k(x), \hat{\tau}^{\eta,\epsilon|b}_{k+1}(x)\big]\Big\}}_{A_k(\eta)}\bigg) \cap \underbrace{\Big\{\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq t_1(\eta)\Big\}}_{B_1(\eta)} \cap \underbrace{\Big\{\tau^{\eta,\epsilon|b}_N > \lfloor t/\lambda^*_b(\eta)\rfloor\Big\}}_{B_2(\eta)}$$

we have $X^{\eta|b}_j(x) \in I_{M,\delta}$ for all $j \in [\hat{\tau}^{\eta,\epsilon|b}_1(x), \hat{\tau}^{\eta,\epsilon|b}_N(x)]$ and $\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq t_1(\eta) < \hat{\tau}^{\eta,\epsilon|b}_N(x)$, thus implying $X^{\eta|b}_{t_1(\eta)}(x) \in I_{M,\delta}$. Therefore, it suffices to show the existence of some M, N, and δ such that

$$\limsup_{\eta\downarrow 0}\bigg[\mathbf{P}\Big(B^c_1(\eta)\Big) + \mathbf{P}\Big(B^c_2(\eta)\Big) + \sum_{k=1}^{N-1}\mathbf{P}\Big(A^c_k(\eta)\Big)\bigg] < \Delta \qquad \forall \delta_t \in (0, t/3). \tag{5.26}$$

Recall that $t_1(\eta) = \lfloor t/\lambda^*_b(\eta)\rfloor - \lfloor 2\delta_t/H(\eta^{-1})\rfloor$. Fix some $u \in (0, t/3)$. First, due to $\lambda^*_b(\eta) = \eta \cdot \big(\eta^{-1}H(\eta^{-1})\big)^{\mathcal{J}^*_b(V)}$ and $\mathcal{J}^*_b(V) \geq 1$, if $\delta_t \in (0, t/3)$ then it holds eventually for all η small enough that $t_1(\eta) > u/\lambda^*_b(\eta)$. Let $i \in [n_{\min}]$ be such that $x \in I_i$ and let $R^{\eta|b}_{i;\epsilon}(x) = \min\{j \geq 0 :\ X^{\eta|b}_j(x) \in [m_i - \epsilon, m_i + \epsilon]\}$. Since $\hat{\tau}^{\eta,\epsilon|b}_1(x)$ is the first visit time to $\bigcup_{l\in[n_{\min}]}(m_l - \epsilon, m_l + \epsilon)$ (i.e., the ϵ-neighborhood of any local minima m_l), we have $\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq R^{\eta|b}_{i;\epsilon}(x)$, and hence

$$\begin{aligned}\limsup_{\eta\downarrow 0}\mathbf{P}\Big(B^c_1(\eta)\Big) &\leq \limsup_{\eta\downarrow 0}\mathbf{P}\Big(\hat{\tau}^{\eta,\epsilon|b}_1(x) > u/\lambda^*_b(\eta)\Big) \leq \limsup_{\eta\downarrow 0}\mathbf{P}\Big(\lambda^*_b(\eta)\cdot R^{\eta|b}_{i;\epsilon}(x) > u\Big)\\ &= 0 \qquad \text{using Proposition 5.2 (i).}\end{aligned} \tag{5.27}$$

We move onto the analysis of event $B_2(\eta)$ and the choice of N. Recall that $Y^{*|b}_t(x)$ is the irreducible, continuous-time Markov chain over V^*_b with law specified in (5.3). In particular, we can fix some N large enough such that $\mathbf{P}(U_1 + \cdots + U_N \leq t) < \Delta/2$. Then from the weak convergence stated in Proposition 2.14, we get

$$\begin{aligned}\limsup_{\eta\downarrow 0}\mathbf{P}\Big(B^c_2(\eta)\Big) &\leq \limsup_{\eta\downarrow 0}\mathbf{P}\bigg(\sum_{n=1}^{N}\big(\tau^{\eta,\epsilon|b}_n(x) - \tau^{\eta,\epsilon|b}_{n-1}(x)\big)\cdot\lambda^*_b(\eta) \leq t\bigg)\\ &\leq \mathbf{P}(U_1 + \cdots + U_N \leq t) < \Delta/2.\end{aligned} \tag{5.28}$$

Meanwhile, recall that $\sigma^{\eta|b}_{k;\epsilon}(x) = \min\{j \geq 0 :\ X^{\eta|b}_j(x) \in \bigcup_{l\neq k}(m_l - \epsilon, m_l + \epsilon)\}$ (i.e., the first time $X^{\eta|b}_j(x)$ visits the ϵ-neighborhood of some m_l that is different from m_k); also, for all $k \geq 2$, $\hat{\tau}^{\eta,\epsilon|b}_k(x)$ is the first time since $\hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)$ that $X^{\eta|b}_j(x)$ visits the ϵ-neighborhood of some m_l that is different from the one visited at $\hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)$. From the strong Markov property at $\hat{\tau}^{\eta,\epsilon|b}_k(x)$, we then get

$$\sup_{k\geq 1}\mathbf{P}\Big(A^c_k(\eta)\Big) \leq \max_{l\in[n_{\min}]}\ \sup_{y\in[m_l-\epsilon,m_l+\epsilon]}\mathbf{P}\Big(\exists j < \sigma^{\eta|b}_{l;\epsilon}(y)\ s.t.\ X^{\eta|b}_j(y) \in S(\delta) \text{ or } \big|X^{\eta|b}_j(y)\big| \geq M\Big).$$

Applying Lemma 5.1, we are able to fix some $M > 0$ and $\delta \in (0, \epsilon/2)$ such that $\limsup_{\eta\downarrow 0}\mathbf{P}\big(A^c_k(\eta)\big) \leq \frac{\Delta}{2N}\ \forall k \in [N-1]$. Combining this bound with (5.27) and (5.28), we finish the proof of (5.26). As a concluding remark, note that our proof of claim (5.23) relies on the specific choices of M and δ but allows for arbitrary $\delta_t \in (0, t/2)$. In the proof of claims (5.24) and (5.25) below, we adopt the same choice of M and δ so these two parameters will be fixed henceforth in this proof.

Proof of (5.24).

We show that the claim holds for all $\delta_t \in (0, t/3)$. Due to $H(x) \in \mathcal{RV}_{-\alpha}(x)$ and $\alpha > 1$, given any $T > 0$ we have $T/\eta < \delta_t/H(\eta^{-1})$ eventually for all η small enough. Recall that $I_{j;\delta,M} = (s_{j-1}+\delta, s_j - \delta) \cap (-M, M)$. By Markov property at $t_1(\eta)$, for any $T > 0$ it holds for all $\eta > 0$ small enough that

$$\begin{aligned}\mathbf{P}\Big((\text{II})^c\ \Big|\ (\text{I})\Big) &\leq \max_{k\in[n_{\min}]}\ \sup_{y\in I_{k;\delta,M}}\mathbf{P}\Big(X^{\eta|b}_j(y) \notin \bigcup_{l\in[n_{\min}]}(m_l - \frac{\epsilon}{2}, m_l + \frac{\epsilon}{2})\ \forall j \leq \delta_t/H(\eta^{-1})\Big)\\ &\leq \max_{k\in[n_{\min}]}\ \sup_{y\in I_{k;\delta,M}}\mathbf{P}\Big(R^{\eta|b}_{k;\epsilon/2}(y) > \delta_t/H(\eta^{-1})\Big)\\ &\leq \max_{k\in[n_{\min}]}\ \sup_{y\in I_{k;\delta,M}}\mathbf{P}\Big(R^{\eta|b}_{k;\epsilon/2}(y) > T/\eta\Big)\end{aligned}$$

where $R^{\eta|b}_{k;\epsilon/2}(y) = \min\{j \geq 0 :\ X^{\eta|b}_j(y) \in (m_k - \frac{\epsilon}{2}, m_k + \frac{\epsilon}{2})\}$.

Let $\boldsymbol{t}_k(x,\epsilon) \triangleq \inf\{t \geq 0 :\ \boldsymbol{y}_t(x) \in (m_k - \epsilon, m_k + \epsilon)\}$. By Assumption 7, $\boldsymbol{t}_k(x, \frac{\epsilon}{4}) < \infty$ for all $x \in [-M-1, M+1]\cap[s_{k-1}+\frac{\delta}{2}, s_k-\frac{\delta}{2}]$, with $\boldsymbol{t}_k(\ \cdot\ , \frac{\epsilon}{4})$ being continuous over $[-M-1, M+1]\cap[s_{k-1}+\frac{\delta}{2}, s_k-\frac{\delta}{2}]$. As a result, we can fix $T \in (0,\infty)$ large enough such that

$$T > \sup\Big\{\boldsymbol{t}_k(x, \frac{\epsilon}{4}) :\ x \in [-M-1, M+1] \cap [s_{k-1} + \frac{\delta}{2}, s_k - \frac{\delta}{2}]\Big\} \qquad \forall k \in [n_{\min}].$$

For each $k \in [n_{\min}]$, by applying Lemma 4.9 onto $(-M-1, M+1) \cap (s_{k-1}, s_k)$, we are able to show that $\limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M}}\mathbf{P}\Big(R^{\eta|b}_{k;\epsilon/2}(y) > T/\eta\Big) = 0$. This concludes the proof of claim (5.24).

Proof of (5.25).

We show that claim (5.25) holds for all $\delta_t \in (0, t/3)$ small enough. By strong Markov property at R^η,

$$\mathbf{P}\Big((\text{III})^c\ \Big|\ (\text{I}) \cap (\text{II})\Big) \leq \max_{k\in[n_{\min}]}\ \sup_{y\in[m_k-\epsilon/2, m_k+\epsilon/2]}\mathbf{P}\Big(\exists j \leq \frac{2\delta_t}{H(\eta^{-1})}\ s.t.\ X^{\eta|b}_j(y) \notin (m_k - \epsilon, m_k + \epsilon)\Big).$$

Also, note that $\epsilon < \bar{\epsilon} < b$; see (5.1). For each $k \in [n_{\min}]$, by applying part (a) of Theorem 2.6 onto $(m_k - \epsilon, m_k + \epsilon)$, we obtain some $c_{k,\epsilon} \in (0,\infty)$ such that for any $u > 0$,

$$\limsup_{\eta\downarrow 0}\ \sup_{y\in[m_k-\epsilon/2, m_k+\epsilon/2]}\mathbf{P}\Big(\exists j \leq \frac{u}{H(\eta^{-1})}\ s.t.\ X^{\eta|b}_j(y) \notin (m_k - \epsilon, m_k + \epsilon)\Big) \leq 1 - \exp(-c_{k,\epsilon}\cdot u).$$

By picking δ_t small enough, we ensure that $1 - \exp(-c_{k,\epsilon}\cdot 2\delta_t) < \Delta$ for all $k \in [n_{\min}]$, thus completing the proof of claim (5.25). □

5.5 Proof of Corollary 2.11

Proof of Corollary 2.11. For simplicity of notations we focus on the case where $T = 1$, but the proof below can be easily generalized for arbitrary $T > 0$.

Fix some $b > 0$ such that $\mathcal{G}_b$ is irreducible and $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{\min}]$ and $j \in [n_{\min} - 1]$. Also, fix some $x \in \bigcup_{i \in [n_{\min}]} I_i$ and let $T^{\eta}_{\text{narrow}} \triangleq \int_0^1 \mathbb{I}\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta) \rfloor}(x) \in \bigcup_{j:\ j \notin V^*_b} I_j\}dt$. The goal is to show that $T^{\eta}_{\text{narrow}} \xrightarrow{\mathbf{P}} 0$ as $\eta \downarrow 0$. To proceed, let $K^{\eta}_N \triangleq \sum_{n=1}^{N-1} \mathcal{I}^{\eta}_N(n)$ where $I^{\eta}_N(n) \triangleq \mathbb{I}\Big\{\exists t \in (\frac{n}{N}, \frac{n+1}{N}]$ s.t. $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta) \rfloor} \in \bigcup_{j:\ j \notin V^*_b} I_j\Big\}$, and note that

$$T^{\eta}_{\text{narrow}} = \sum_{n=0}^{N-1} \int_{n/N}^{(n+1)/N} \mathbb{I}\Big\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta) \rfloor}(x) \in \bigcup_{j:\ j \notin V^*_b} I_j\Big\}dt \leq \frac{1}{N} + \sum_{n=1}^{N-1} \frac{1}{N} \cdot \mathcal{I}^{\eta}_N(n) = \frac{1 + K^{\eta}_N}{N}.$$

The proof hinges on the following claims: there exist some $C \in (0, \infty)$ and a family of events A^{η}_N such that

(i) for all positive integer N large enough, $\lim_{\eta \downarrow 0} \mathbf{P}(A^{\eta}_N) = 1$;

(ii) for all positive integer N large enough, there exists $\bar{\eta} = \bar{\eta}(N) > 0$ such that under any $\eta \in (0, \bar{\eta})$,

$$\mathbf{P}(K^{\eta}_N \geq j \mid A^{\eta}_N) \leq \mathbf{P}\Big(\text{Binom}(N, \frac{2C}{N}) \geq j\Big) \qquad \forall j = 1, 2, \cdots, N.$$

Here, $\text{Binom}(n, p)$ is the Binomial RV representing the number of successful trials among n Bernoulli trials with success rate p. Then given any N large enough, $\eta \in (0, \bar{\eta}(N))$ and any $\beta \in (0, 1)$,

$$\begin{aligned}
\mathbf{P}\Big(T^{\eta}_{\text{narrow}} \geq \underbrace{\frac{1 + 2C + \sqrt{N^{\beta}}}{N}}_{\triangleq \delta(N.\beta)}\Big) &\leq \mathbf{P}(K^{\eta}_N \geq 2C + \sqrt{N^{\beta}}) \\
&= \mathbf{P}\big(\{K^{\eta}_N \geq 2C + \sqrt{N^{\beta}}\} \cap A^{\eta}_N\big) + \mathbf{P}\big(\{K^{\eta}_N \geq 2C + \sqrt{N^{\beta}}\} \setminus A^{\eta}_N\big) \\
&\leq \mathbf{P}\Big(\text{Binom}(N, \frac{2C}{N}) \geq 2C + \sqrt{N^{\beta}}\Big) + \mathbf{P}\big((A^{\eta}_N)^c\big) \qquad \text{by claim (ii)} \\
&\leq \mathbf{P}\Big(\Big(\text{Binom}(N, \frac{2C}{N}) - 2C\Big)^2 \geq N^{\beta}\Big) + \mathbf{P}\big((A^{\eta}_N)^c\big) \\
&\leq \frac{\text{var}\Big[\text{Binom}(N, \frac{2C}{N})\Big]}{N^{\beta}} + \mathbf{P}\big((A^{\eta}_N)^c\big) \qquad \text{by Markov's inequality} \\
&\leq \frac{2C}{N^{\beta}} + \mathbf{P}\big((A^{\eta}_N)^c\big).
\end{aligned}$$

Using claim (i) and by driving $\eta \downarrow 0$, we get $\limsup_{\eta \downarrow 0} \mathbf{P}\big(T^{\eta}_{\text{narrow}} \geq \delta(N, \beta)\big) \leq 2C/N^{\beta}$ for all N large enough. Lastly, note that $C/N^{\beta} \to 0$ as $N \to \infty$; also, under our choice of $\beta \in (0, 1)$ we have $\lim_{N \to \infty} \delta(N, \beta) = 0$. This implies $T^{\eta}_{\text{narrow}} \xrightarrow{\mathbf{P}} 0$ as $\eta \downarrow 0$.

Now, it only remains to verify claims (i) and (ii). First, we specify the choice of events A^{η}_N. Let $t_N(n) = n/N$. For some $\epsilon > 0$, let

$$A^{\eta}_N(n) \triangleq \Big\{X^{\eta|b}_{\lfloor t_N(k)/\lambda^*_b(\eta) \rfloor}(x) \in \bigcup_{i:\ m_i \in V^*_b} (m_i - \epsilon, m_i + \epsilon)\ \forall k \in [n]\Big\}$$

and let $A^{\eta}_N = A^{\eta}_N(N)$. Note that $A^{\eta}_N(1) \supseteq A^{\eta}_N(2) \supseteq \cdots \supseteq A^{\eta}_N(N) = A^{\eta}_N$. Furthermore, in Theorem 2.9 note that the limiting CTMC $Y^{*|b}_t$ only visits V^*_b. As a result, given any positive integer N, one can find $\epsilon = \epsilon(N) > 0$ small enough such that $\lim_{\eta \downarrow 0} \mathbf{P}(A^{\eta}_N) = 1$.

Next, let $\big(\widetilde{\mathcal{I}}^\eta_N(n)\big)_{n\in[N-1]}$ be a random vector with law $\mathcal{L}\Big(\big(\mathcal{I}^\eta_N(n)\big)_{n\in[N-1]}\ \Big|\ A^\eta_N\Big)$. Suppose there exists some $C\in(0,\infty)$ such that for all N large enough, there is $\bar\eta=\bar\eta(N)>0$ for the following claim to hold: Given any $n\in[N-1]$ and sequence $i_j\in\{0,1\}\ \forall j\in[n-1]$,

$$\mathbf{P}\Big(\widetilde{\mathcal{I}}^\eta_N(n)=1\ \Big|\ \widetilde{\mathcal{I}}^\eta_N(j)=i_j\ \forall j\in[n-1]\Big)<2C/N\qquad\forall\eta\in(0,\bar\eta).\tag{5.29}$$

Then given any N sufficiently large and any $\eta\in(0,\bar\eta(N))$, there exists a coupling between iid Bernoulli RVs $(\mathcal{I}_N(n))_{n\in[N-1]}$ with success rate $2C/N$ and $(\widetilde{\mathcal{I}}^\eta_N(n))_{n\in[N-1]}$ such that $\widetilde{\mathcal{I}}^\eta_N(n)\leq\mathcal{I}_N(n)\ \forall n\in[N-1]$ almost surely. This coupling between $(\mathcal{I}_N(n))_{n\in[N-1]}$ and $(\widetilde{\mathcal{I}}^\eta_N(n))_{n\in[N-1]}$ immediately verifies claim (ii).

Lastly, we prove condition (5.29) under the choice of $C>\max_{i:\ m_i\in V^*_b}q_b(i)$ with $q_b(i)$ defined in Section 5.1. Besides, due to $\lim_{x\to 0}\frac{1-\exp(-x)}{x}=1$, it holds for all N large enough that

$$1-\exp\Big(-C\cdot\frac{1}{N}\Big)<\sqrt{2}C/N.\tag{5.30}$$

Henceforth in this proof, we fix such large N. Now, given any $n\in[N-1]$ and sequence $i_j\in\{0,1\}\ \forall j\in[n-1]$, observe that

$$\begin{aligned}
&\mathbf{P}\Big(\widetilde{\mathcal{I}}^\eta_N(n)=1\ \Big|\ \widetilde{\mathcal{I}}^\eta_N(j)=i_j\ \forall j\in[n-1]\Big)\\
&=\frac{\mathbf{P}\Big(\widetilde{\mathcal{I}}^\eta_N(n)=1;\ \widetilde{\mathcal{I}}^\eta_N(j)=i_j\ \forall j\in[n-1]\Big)}{\mathbf{P}\Big(\widetilde{\mathcal{I}}^\eta_N(j)=i_j\ \forall j\in[n-1]\Big)}\\
&=\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(n)=1;\ \mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)\Big/\mathbf{P}(A^\eta_N)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)\Big/\mathbf{P}(A^\eta_N)}\\
&\leq\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(n)=1;\ \mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)}\qquad\text{due to }A^\eta_N(n)\supseteq A^\eta_N\\
&=\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(n)=1;\ \mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}\cdot\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)}\\
&=\underbrace{\mathbf{P}\Big(\mathcal{I}^\eta_N(n)=1\ \Big|\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}_{\triangleq p^\eta_1}\cdot\underbrace{\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)}}_{\triangleq p^\eta_2}.
\end{aligned}$$

For term p^η_1, by Markov property of $X^{\eta|b}_j(x)$ at $j=\lfloor t_N(n)/\lambda^*_b(\eta)\rfloor$,

$$\begin{aligned}
p^\eta_1&\leq\sup_{y\in\bigcup_{i:\ m_i\in V^*_b}(m_i-\epsilon,m_i+\epsilon)}\mathbf{P}\bigg(X^{\eta|b}_j(y)\notin\bigcup_{i:\ m_i\in V^*_b}(m_i-\epsilon,m_i+\epsilon)\text{ for some }j\leq\lfloor\frac{1/N}{\lambda^*_b(\eta)}\rfloor\bigg)\\
&\leq\max_{i:\ m_i\in V^*_b}\ \sup_{y\in(m_i-\epsilon,m_i+\epsilon)}\mathbf{P}\bigg(X^{\eta|b}_j(y)\notin I_i\text{ for some }j\leq\lfloor\frac{1/N}{\lambda^*_b(\eta)}\rfloor\bigg).
\end{aligned}$$

By part (ii) of Proposition 5.2, there is some $\bar\eta=\bar\eta(N)>0$ such that for all $\eta\in(0,\bar\eta)$, we have $p^\eta_1<1-\exp(-C\cdot 1/N)<\sqrt{2}\cdot C/N$ due to our choice of $C>\max_{i:\ m_i\in V^*_b}q_b(i)$ and the choice of N in (5.30). As for term p^η_2, note that for any event B, we have

$$\frac{\mathbf{P}(B\cap A^\eta_N(n))}{\mathbf{P}(B\cap A^\eta_N)}\leq\frac{\mathbf{P}(B)}{\mathbf{P}(B)-\mathbf{P}\big((A^\eta_N)^c\big)}\to 1\qquad\text{as }\eta\downarrow 1\text{ due to }\lim_{\eta\downarrow 0}\mathbf{P}(A^\eta_N)=1.\tag{5.31}$$

In the definition of p_2^η, note that there are only finitely many choices of $n \in [N-1]$ and finitely many combinations for $i_j \in \{0,1\}\ \forall j \in [n-1]$. By considering each of the finitely many choices for $B = \{\mathcal{I}_N^\eta(j) = i_j\ \forall j \in [n-1]\}$ in (5.31), we can find some $\bar{\eta} = \bar{\eta}(N)$ such that $p_2^\eta < \sqrt{2}\ \forall \eta \in (0,\bar{\eta})$ uniformly for all the choices of $n \in [N-1]$ and sequence i_j. Combining the bounds $p_1^\eta < \sqrt{2}C/N$ and $p_2^\eta < \sqrt{2}$ (for all η small enough), we verify condition (5.29) and conclude the proof. □

References

[1] P. Baldi and L. Caramellino. General freidlin–wentzell large deviations and positive diffusions. *Statistics & Probability Letters*, 81(8):1218–1229, 2011.

[2] H. Bernhard, B. Das, et al. Heavy-tailed random walks, buffered queues and hidden large deviations. *Bernoulli*, 26(1):61–92, 2020.

[3] A. A. Borovkov and K. A. Borovkov. *Asymptotic Analysis of Random Walks: Heavy-Tailed Distributions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2008.

[4] A. Bovier and F. Den Hollander. *Metastability: a potential-theoretic approach*, volume 351. Springer, 2016.

[5] A. Budhiraja and P. Dupuis. A variational representation for positive functionals of infinite dimensional brownian motion. *Probability and mathematical statistics-Wroclaw University*, 20(1):39–61, 2000.

[6] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability*, 36(4):1390 – 1420, 2008.

[7] D. Buraczewski, E. Damek, T. Mikosch, and J. Zienkiewicz. Large deviations for solutions to stochastic recurrence equations under Kesten's condition. *The Annals of Probability*, 41(4):2755 – 2790, 2013.

[8] O. Catoni and R. Cerf. The exit path of a markov chain with rare transitions. *ESAIM: Probability and Statistics*, 1:95–144, 1997.

[9] S. Cerrai and M. Röckner. Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-Lipshitz reaction term. *The Annals of Probability*, 32(1B):1100 – 1139, 2004.

[10] A. V. Chechkin, V. Y. Gonchar, J. Klafter, and R. Metzler. Barrier crossing of a lévy flight. *Europhysics Letters*, 72(3):348, 2005.

[11] A. V. Chechkin, O. Y. Sliusarenko, R. Metzler, and J. Klafter. Barrier crossing driven by lévy noise: Universality and the role of noise intensity. *Physical Review E*, 75(4):041101, 2007.

[12] B. Chen, C.-H. Rhee, and B. Zwart. Sample-path large deviations for a class of heavy-tailed markov additive processes. *arXiv preprint arXiv:2010.10751*, 2023.

[13] P. Chigansky and R. Liptser. *The Freidlin-Wentzell LDP with Rapidly Growing Coeffcients*, pages 197–219. World Scientific, 2007.

[14] D. Denisov, A. B. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: The big-jump domain. *The Annals of Probability*, 36(5):1946 – 1991, 2008.

[15] C. Donati-Martin. Large deviations for wishart processes. *Probability and Mathematical Statistics*, 28, 2008.

[16] C. Donati-Martin, A. Rouault, M. Yor, and M. Zani. Large deviations for squares of bessel and ornstein-uhlenbeck processes. *Probability Theory and Related Fields*, 129:261–289, 2004.

[17] P. Dupuis and R. S. Ellis. *A weak convergence approach to the theory of large deviations*. John Wiley & Sons, 2011.

[18] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media, 2013.

[19] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[20] S. Foss, T. Konstantopoulos, and S. Zachary. Discrete and continuous time modulated random walks with heavy-tailed increments. *Journal of Theoretical Probability*, 20:581–612, 2007.

[21] S. Foss, D. Korshunov, S. Zachary, et al. *An introduction to heavy-tailed and subexponential distributions*, volume 6. Springer, 2011.

[22] H. Hult and F. Lindskog. Extremal behavior of stochastic integrals driven by regularly varying Lévy processes. *The Annals of Probability*, 35(1):309 – 339, 2007.

[23] H. Hult, F. Lindskog, T. Mikosch, and G. Samorodnitsky. Functional large deviations for multivariate regularly varying random walks. *The Annals of Applied Probability*, 15(4):2651 – 2680, 2005.

[24] P. Imkeller and I. Pavlyukevich. First exit times of sdes driven by stable lévy processes. *Stochastic Processes and their Applications*, 116(4):611–642, 2006.

[25] P. Imkeller, I. Pavlyukevich, and M. Stauch. First exit times of non-linear dynamical systems in $\mathbb{R}^d$ perturbed by multifractal Lévy noise. *Journal of Statistical Physics*, 141(1):94–119, 2010.

[26] P. Imkeller, I. Pavlyukevich, and T. Wetzel. First exit times for Lévy-driven diffusions with exponentially light jumps. *The Annals of Probability*, 37(2):530 – 564, 2009.

[27] P. Imkeller, I. Pavlyukevich, and T. Wetzel. The hierarchy of exit times of Lévy-driven Langevin equations. *The European Physical Journal Special Topics*, 191(1):211–222, 2010.

[28] Imkeller, Peter and Pavlyukevich, Ilya. Metastable behaviour of small noise lévy-driven diffusions. *ESAIM: PS*, 12:412–437, 2008.

[29] G. Kallianpur and J. Xiong. Stochastic differential equations in infinite dimensional spaces. *Lecture Notes-Monograph Series*, 26:iii–342, 1995.

[30] Y. Kifer. A Discrete-Time Version of the Wentzell-Friedlin Theory. *The Annals of Probability*, 18(4):1676 – 1692, 1990.

[31] F. Klebaner and O. Zeitouni. The exit problem for a class of period doubling systems. *Ann. Appl. Probab*, 4:1188–1205, 1994.

[32] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[33] R. Kruse. *Strong and weak approximation of semilinear stochastic evolution equations*. Springer, 2014.

[34] F. Lindskog, S. I. Resnick, J. Roy, et al. Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probability Surveys*, 11:270–314, 2014.

[35] A. D. W. M. I. Freidlin. *Random Perturbations of Dynamical Systems*. Springer New York, NY, 1998.

[36] T. Mikosch and O. Wintenberger. Precise large deviations for dependent regularly varying sequences. *Probability Theory and Related Fields*, 156(3-4):851–887, 2013.

[37] T. Mikosch and O. Wintenberger. A large deviations approach to limit theory for heavy-tailed time series. *Probability Theory and Related Fields*, 166:233–269, 2016.

[38] M. T. Mohan. Wentzell–freidlin large deviation principle for stochastic convective brinkman–forchheimer equations. *Journal of Mathematical Fluid Mechanics*, 23(3):62, 2021.

[39] G. J. Morrow and S. Sawyer. Large deviation results for a class of markov chains arising from population genetics. *The Annals of Probability*, 17(3):1124–1146, 1989.

[40] A. Nagaev. Limit theorems for large deviations where cramér's conditions are violated. *Fiz-Mat. Nauk*, 7:17–22, 1969.

[41] A. V. Nagaev. On a property of sums of independent random variables. *Theory of Probability & Its Applications*, 22(2):326–338, 1978.

[42] C. Penland and B. D. Ewald. On modelling physical systems with stochastic models: diffusion versus lévy processes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2455–2474, 2008.

[43] P. E. Protter. *Stochastic integration and differential equations*. Springer, 2005.

[44] S. I. Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media, 2007.

[45] C.-H. Rhee, J. Blanchet, B. Zwart, et al. Sample path large deviations for lévy processes and random walks with regularly varying increments. *The Annals of Probability*, 47(6):3551–3605, 2019.

[46] M. Röckner, T. Zhang, and X. Zhang. Large deviations for stochastic tamed 3d navier-stokes equations. *Applied Mathematics and Optimization*, 61:267–285, 2010.

[47] K.-i. Sato, S. Ken-Iti, and A. Katok. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge university press, 1999.

[48] E. Scalas, R. Gorenflo, and F. Mainardi. Fractional calculus and continuous-time finance. *Physica A: Statistical Mechanics and its Applications*, 284(1-4):376–384, 2000.

[49] A. Shwartz and A. Weiss. *Large deviations for performance analysis: queues, communication and computing*, volume 5. CRC Press, 1995.

[50] R. B. Sowers. Large deviations for a reaction-diffusion equation with non-gaussian perturbations. *The Annals of Probability*, 20(1):504–537, 1992.

[51] A. D. Ventsel' and M. I. Freidlin. On small random perturbations of dynamical systems. *Russian Mathematical Surveys*, 25(1):1, feb 1970.

[52] X. Wang, S. Oh, and C.-H. Rhee. Eliminating sharp minima from SGD with truncated heavy-tailed noise. In *International Conference on Learning Representations*, 2022.

[53] W. Wei, Q. Huang, and J. Duan. Large deviations for sde driven by heavy-tailed l\'evy processes. *arXiv preprint arXiv:2101.03856*, 2021.