# 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

May 27, 2024

**Abstract**

This paper introduces a novel framework for large deviations and metastability analysis in heavy-tailed stochastic dynamical systems. Employing this framework in the context of stochastic difference equations $X^\eta_{j+1}(x) = X^\eta_j(x) + \eta a\big(X^\eta_j(x)\big) + \eta\sigma\big(X^\eta_j(x)\big)Z_{j+1}$ and its variation with a truncated dynamics, we first establish locally uniform sample path large deviations and then translate these large-deviation asymptotics into a sharp characterization of the joint distribution of the first exit time and exit location. As a result, we obtain the heavy-tailed counterparts of the classical Freidlin-Wentzell and Eyring-Kramers theorems. Furthermore, our exit time and exit location analysis paves the way to a sharp characterization of the global dynamics of a popular variant of heavy-tailed stochastic gradient descent (SGD) algorithm: we show that, after proper scaling, heavy-tailed SGDs with gradient clipping converge to continuous-time Markov chains that only visit the widest minima of the potential function. In the appendix, we also present the corresponding results for the Lévy driven SDEs.

## Contents

# 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, and over the past few decades, the theory has seen numerous extensions, including the discrete-time version of Freidlin–Wentzell theorem (see, e.g., [38, 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., [6, 7, 28]), and large deviations for stochastic partial differential equations (see, e.g., [50, 9, 46, 37]), 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 these seminal works, such as the asymptotic characterization of the most likely exit path and the exit times for Brownian particles under more sophisticated gradient fields (see [36]), 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 [5].

While such developments provide powerful means to understand rare events and metastability of classical light-tailed systems, they often fail to provide useful bounds when it comes to the heavy-tailed systems. As shown in [23, 25, 26, 24], 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. This is in sharp contrast to the light-tailed counterparts where rare events typically arise via smooth tilting of the nominal dynamics. Due to such a stark difference in the mechanism through which rare events arise, heavy-tailed systems exhibit a fundamentally different large deviations and metastability behaviors and call for a different set of technical tools for successful analysis.

In this paper, we build a general framework for asymptotic analysis of heavy-tailed dynamical systems by developing a set of machinery that uncovers the interconnection between the large deviations, local stability, and global dynamics of stochastic processes. Building upon this framework, we characterize the sample-path large deviations and metastability of heavy-tailed stochastic difference equations (and stochastic differential equations in the appendix), thus offering the heavy-tailed counterparts of Freidlin–Wentzell and Eyring–Kramers theory. Furthermore, we show that our framework is powerful enough to elegantly characterize the global behavior of the heavy-tailed systems. More precisely, the main contributions of this article can be summarized as follows:

- **Heavy-tailed Large Deviations:** We establish sample-path large deviations for heavy-tailed dynamical systems. We propose a new heavy-tailed large deviations formulation that is locally uniform w.r.t. the initial values. We accomplish this by formulating a uniform version of $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence [35, 45]. Our large deviations characterize the *catastrophe principle* (also known as the *principle of big jumps*), which reveals a discrete hierarchy governing the causes and probabilities of a wide variety of rare events associated with heavy-tailed stochastic difference/differential equations. Moreover, this new formulation of the heavy-tailed large deviations paves the way to the analysis of local stability and global dynamics.
- **Local Stability Analysis:** We establish a scaling limit of the exit-time and exit-location for stochastic difference equations. We accomplish this by developing a machinery for local stability analysis of general (heavy-tailed) Markov processes. Central to the development is the concept of asymptotic atoms, where the process recurrently enters and asymptotically regenerates. Leveraging the locally uniform version of sample-path large deviations over such asymptotic atoms, we obtain sharp asymptotics of the joint law of the (scaled) exit-times and exit-locations for heavy-tailed processes. Notably, this complements the investigation of the exit times under the truncated dynamics, which was first analyzed in [26] in the context of Weibull tails.
- **Characterization of Global Dynamics:** Building on the exit-time and exit-location analysis, we establish a scaling limit of the heavy-tailed dynamical systems over a multi-well potential at the process level. The scaling limit is a Markov jump process whose state space consists of the local minima of the potential. In particular, our findings systematically characterize a curious phenomena that the truncated heavy-tailed processes avoid narrow local minima altoghether in the limit. As a direct application, we prove an ergodic theorem, which shows that the fraction of time such processes spend in the narrow attraction field converges to zero as the step-size tends to zero. Precise characterization of such phenomena is of fundamental importance in order to understand and control the effectiveness of the stochastic gradient descent (SGD) algorithms in training deep neural networks.

Specifically, we focus on the class of heavy-tailed phenomena captured by the notion of regular variation. 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 $\phi$ such that $\mathbf{P}(|Z_1| > x) = \phi(x)x^{-\alpha}$. For any $\eta > 0$ and $x \in \mathbb{R}$, let $\big(X^\eta_j(x)\big)_{j\geq 0}$ be the solution of the following stochastic difference equation

$$X^\eta_0(x) = x; \qquad X^\eta_{j+1}(x) = X^\eta_j(x) + \eta a\big(X^\eta_j(x)\big) + \eta\sigma\big(X^\eta_j(x)\big)Z_{j+1} \quad \forall j \geq 0. \tag{1.1}$$

Throughout this paper, we adopt the convention that the subscript denotes the time, and the superscript $\eta$ denotes the scaling parameter that tends to zero. Furthermore, we also consider a truncated variation of $X^\eta_{j+1}(x)$ which is arguably more relevant when $Z_i$'s are heavy-tailed. Specifically, let $\varphi_b(\cdot) : x \mapsto \frac{x}{|x|}\max\{b, |x|\}$ be the projection operator from $\mathbb{R}$ onto $[-b, b]$, where $b > 0$ is a truncation threshold. Define $\big(X^{\eta|b}_j(x)\big)_{j\geq 0}$ with the following recursion:

$$X^{\eta|b}_0(x) = x; \qquad X^{\eta|b}_{j+1}(x) = X^{\eta|b}_j(x) + \varphi_b\Big(\eta a\big(X^{\eta|b}_j(x)\big) + \eta\sigma\big(X^{\eta|b}_j(x)\big)Z_{j+1}\Big) \quad \forall j \geq 0. \tag{1.2}$$

In other words, $X^{\eta|b}_j(x)$ is a modulated version of $X^\eta_j(x)$ where the distance traveled at each step is truncated at $b$. Such dynamical systems arise in the training algorithms for deep neural networks, and their global dynamics has a close connection to the curious ability of SGDs to regularize the deep neural networks algorithmically. See, for example, [52] and the references therein for more details. Note that (1.1) and (1.2) can be viewed as discretizations of small noise SDEs driven by Lévy processes. All the results we establish for (1.1) and (1.2) in this paper can also be established for the stochastic differential equations driven by regularly-varying Lévy processes through a straightforward adaptation of the machinery we develop in this paper. The results for Lévy-driven SDEs are summarized in Appendix A.

At the crux of this study 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 system-wide rare events are characterized by the *conspiracy principle*: the system deviates from its nominal behavior because the entire system behaves subtly differently from the norm, as if it has conspired. In contrast, *the catastrophe principle* governs the rare events in heavy-tailed systems: catastrophic failures (i.e., extremely large deviations from the average behavior) in a small number of components drive the system-wide rare events, and the behavior of the rest of the system is indistinguishable from the nominal behavior.

The principle of a single big jump, a special case of the catastrophe principle, has been discussed in the heavy-tail and extreme value theory literature for a long time. That is, in many heavy-tailed systems, the system-wide rare events arise due to exactly one catastrophe. This line of investigation was initiated in the classical works [39, 40]. The summary of the subsequent developments in the context of processes with independent increments can be found in, for example, [4, 14, 18, 21]. The principle of a single big jump has been rigorously confirmed for random walks in the form of heavy-tailed large deviations at the sample-path level in [22]. More recently, [45] established a fully general catastrophe principle, which goes beyond the principle of a single big jump and characterizes the rare events driven by any number of catastrophes for regularly varying Lévy processes and random walks. For example, let $\mathbb{D}$ denote the space of càdlàg functions over $[0,1]$, let $S_j \triangleq Z_1 + \cdots + Z_j$ denote a mean-zero random walk, and let $\boldsymbol{S}^n \triangleq \{S^n_{\lfloor nt \rfloor}/n : t \in [0,1]\}$ denote a scaled version of $S_j$. Suppose that $Z_i$'s have a regularly varying tail with index $\alpha$ as above. Then, the sample path large deviations established in [45] takes the following form: for "general" $B \in \mathbb{D}$,

$$
\begin{aligned}
0 < \mathbf{C}_k(B^\circ) &\leq \liminf_{n\to\infty} \frac{\mathbf{P}(\boldsymbol{S}^n \in B)}{(n\mathbf{P}(|Z_1| > n))^k} \\
&\leq \limsup_{n\to\infty} \frac{\mathbf{P}(\boldsymbol{S}^n \in B)}{(n\mathbf{P}(|Z_1| > n))^k} \leq \mathbf{C}_k(B^-) < \infty,
\end{aligned}
\tag{1.3}
$$

where $k$ is the minimal number of jumps that a step function must possess in order to belong to $B$, $\mathbf{C}_k(\cdot)$ is a measure on $\mathbb{D}$ supported on the set of step functions with $k$ or less jumps, and $B^\circ$ and $B^-$ are the interior and closure of $B$, respectively. Here, $k$, as a function of $B$, plays the role of the infimum of rate function over $B$ in the classical light-tailed large deviation principle (LDP) formulation. Note that in contrast to the standard log-asymptotics in light-tailed counterparts, (1.3) provides exact asymptotics. See also [2] where similar asymptotic bounds were obtained for random walks under more general scaling. Below, we describe the three main contributions of this paper in more detail.

**Large Deviations for Heavy-Tailed Dynamical Systems.** The first contribution of this paper is to characterize the catastrophe principle for a general class of heavy-tailed stochastic dynamical systems in the form of a "locally uniform" heavy-tailed large deviations at the sample-path level. This turns out to be the right large deviations formulation for the purpose of the subsequent analysis of local stability and global dynamics. To be specific, let $\boldsymbol{X}^\eta(x) \triangleq \{X^\eta_{\lfloor t/\eta \rfloor}(x) : \ t \in [0,1]\}$ be the time-scaled version of $X^\eta_j(x)$ embedded in the continuous time, and note that $\boldsymbol{X}^\eta(x)$ is a random element in $\mathbb{D}$. As $\eta$ decreases, $\boldsymbol{X}^\eta(x)$ converges to a deterministic limit $\{\boldsymbol{y}_t(x) : t \in [0,1]\}$, where $d\boldsymbol{y}_t(x)/dt = a(\boldsymbol{y}_t(x))$ with initial value $\boldsymbol{y}_0(x) = x$. Let $B \subseteq \mathbb{D}$ be a Borel set w.r.t. the $J_1$ topology and $A \subset \mathbb{R}$ be a compact set. We establish the following asymptotic bound for each $k$:

$$\begin{aligned}\inf_{x \in A} \mathbf{C}^{(k)}(B^\circ; x) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}\big(\boldsymbol{X}^\eta(x) \in B\big)}{\big(\eta^{-1}\mathbf{P}(|Z_1| > \eta^{-1})\big)^k} \\ &\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}\big(\boldsymbol{X}^\eta(x) \in B\big)}{\big(\eta^{-1}\mathbf{P}(|Z_1| > \eta^{-1})\big)^k} \leq \sup_{x \in A} \mathbf{C}^{(k)}(B^-; x).\end{aligned} \tag{1.4}$$

The precise statement and the definition of $\mathbf{C}^{(k)}$ can be found in Theorem 2.3 and Section 2.2.1, but here we just point out that the index $k$ that leads to non-degenerate upper and lower bounds in (1.4) is the minimum number of jumps that needs to be added to the path of $\boldsymbol{y}_t(x)$ for it to enter the set $B$ given $x \in A$. Such a $k$ dictates the precise polynomial decay rate of the rare-event probability and corresponds to the infimum of rate function of the classical large deviations framework. Note also that as the set $A$ shrinks to an atom, the upper and lower bounds in (1.4) become tighter, and hence, (1.4) is a *locally uniform* version of the large deviations formulation in (1.3). An important implication of (1.4) is a sharp characterization of the catastrophe principle. Specifically, Section 2.2.2 proves that the conditional distribution of $\boldsymbol{X}^\eta(x)$ given the rare event of interest converges to the distribution of a piecewise deterministic random function $\boldsymbol{X}^*_{|B}(x)$ with precisely $k$ random jumps whose sizes are bounded from below:

$$\mathcal{L}\big(\boldsymbol{X}^\eta(x)\big|\boldsymbol{X}^\eta(x) \in B\big) \to \mathcal{L}\big(\boldsymbol{X}^*_{|B}(x)\big).$$

Note that the perturbation associated with $Z_i$ is modulated by $\eta\sigma(X^\eta_{i-1}(x))$. Hence, the jump size associated with $Z_i$ being bounded from below implies that $Z_i$ is of order $1/\eta$. This confirms that the rare event $\{\boldsymbol{X}^\eta(x) \in B\}$ arises almost always because of $k$ catastrophically large $Z_i$'s, whereas the rest of the system is indistinguishable from its nominal behavior.

The notion of $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence, introduced in [35] and further developed in [45], was a key technical tool behind (1.3). In this paper, we introduce a uniform version of the $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$-convergence to establish the uniform asymptotics in (1.4) and prove an associated Portmanteau theorem (Theorem 2.2); see Section 2.1. These developments form the backbone that supports our proofs of the uniform sample-path large deviations in (1.4). Furthermore, we also establish the locally uniform asymptotics for $\boldsymbol{X}^{\eta|b}(x)$ in Theorem 2.4. As Section 2.3 and Section 2.4 elaborate, such large deviations of $\boldsymbol{X}^{\eta|b}(x)$ leads to exit times and scaled sample paths with structurally different asymptotic limits compared to those associated with $\boldsymbol{X}^\eta(x)$.

**Exit Time Analysis.** The second contribution of this paper is the local stability analysis for heavy-tailed systems. The first exit time problem finds applications in numerous contexts, including chemical reactions [32], physics [10, 11], extreme climate events [42], mathematical finance [48], and queueing systems [49]. A classical result in this literature is the Eyring-Kramers law [20, 33], which characterizes the exit time of Brownian particles; see also [36].

Unlike the classical light-tailed context where the dynamical systems are driven by Brownian noise, the exit times of the heavy-tailed Lévy-driven SDEs exhibit fundamentally different characteristics, and their successful analysis is a relatively recent development [23, 24]. These results were extended to the multi-dimensional settings in [25] and, later, stochastic difference equations driven by $\alpha$-stable noises in [41] as well. The exit times characterized in this line of research is a manifestation of the principle of a single big jump in the context of the exit times of the stochastic dynamical systems. In contrast, our focus in this paper is to build a systematic tool that facilitates the analysis of the exit times even when they are driven by multiple big jump events as in the case of $\boldsymbol{X}^{\eta|b}(x)$. Indeed, we characterize the asymptotics of the joint law of the first exit time and the exit location for heavy-tailed processes.

We consider (1.1) with drift coefficients $a(\cdot) = -U'(\cdot)$ 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 $\boldsymbol{y}_0(x) = x$, it holds that $\lim_{t\to\infty} \boldsymbol{y}_t(x) = 0$ for all $x \in I$. As a result, when initialized within $I$, the deterministic process will be attracted to and be trapped around the origin. In contrast, under the presence of random perturbations, although $X^{\eta}_j(x)$ and $X^{\eta|b}_j(x)$ are attracted to the origin most of the times, they will eventually escape from $I$ if one waits long enough. Of particular interest are the asymptotics of the first exit time as $\eta \to \infty$. Theorem 2.6 establishes that the joint law of the first exit time $\tau^{\eta|b}(x) = \min\{j \geq 0 : \ X^{\eta|b}_j(x) \notin I\}$ and the exit location $X^{\eta|b}_{\tau}(x) \triangleq X^{\eta|b}_{\tau^{\eta|b}(x)}(x)$ admits the following limit (for all $x \in I$):

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

with some (deterministic) time-scaling function $\lambda^I_b(\eta)$. Here, $E$ is an exponential random variable with the rate parameter 1, and $V_b$ is some random element independent of $E$ and supported on $I^c$. The exact law of $V_b$ and the definition of $\lambda^I_b(\eta)$ are provided in Section 2.3.1. Here, we note that $\lambda^I_b(\eta)$ is regularly varying with index $-[1 + \mathcal{J}^*_b(\alpha - 1)]$, where $\mathcal{J}^*_b$ is the "discretized width" of domain $I$ relative to the truncation threshold $b$; see (2.26) for the precise definition. Intuitively speaking, $\mathcal{J}^*_b$ is the minimal number of jumps of size $b$ to escape from $I$, and hence, the wider the domain $I$ is, the longer the exit time $\tau^{\eta|b}(x)$ will be asymptotically. Theorem 2.6 also obtains the first exit time analysis for $X^{\eta}_j(x)$ by considering an arbitrarily large truncation threshold $b \approx \infty$.

Our approach hinges on a general machinery we develop in Section 2.3.2. At the core of this development lies the concept of asymptotic atoms, namely, nested regions of recurrence at which the process asymptotically regenerates upon each visit. Our locally uniform sample-path large deviations then prove to be the right tool in this framework, empowering us to simultaneously characterize the behavior of the stochastic processes under all the initial values over the asymptotic atoms. We also point out that our results make weaker assumptions than the previous results, allowing for non-constant diffusion coefficient $\sigma(\cdot)$ and eliminating the need for regularity conditions such as $U \in \mathcal{C}^3(\mathbb{R})$ and non-degeneracy of $U''(\cdot)$ at the boundary of $I$.

It should be noted that [26] also investigated the exit events driven by multiple jumps. However, the mechanism through which multiple jumps arise in their context is through a different tail behavior of the increment distribution that is lighter than any polynomial rate—more precisely, a Weibull tail—and it is fundamentally different from that of the regularly varying case. Along with the aforementioned results [23, 24, 25] for regularly varying SDEs, [26] paints interesting picture of the hierarchy in the asymptotics of the first exit times. See [27] for the summary of such hierarchy. Our results complement the picture and provide a missing piece of the puzzle by unveiling the precise effect of truncation in the regularly varying cases. In particular, we characterize a discrete structure of phase transitions in (1.5), where we find that the first exit time $\tau^{\eta|b}(x)$ is (roughly) of order $1/\eta^{1+\mathcal{J}^*_b\cdot(\alpha-1)}$ for small $\eta$. This means that the order of the first exit time $\tau^{\eta|b}(x)$ does not vary continuously with $b$; rather, it exhibits a discrete dependence on $b$ through $\mathcal{J}^*_b$.

**Characterization of the Global Dynamics:** The third contribution of this paper is to characterize the global dynamics of heavy-tailed stochastic dynamical systems. 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 its attraction fields. Under suitable conditions, Theorem 2.10 establishes that the stochastic process $X^{\eta|b}_j(x)$ converges to a Markov jump process that visits only the widest local minima with proper scaling. By considering an arbitrarily large truncation threshold $b \approx \infty$, we also recover the sample-path convergence of the untruncated dynamics $X^{\eta}_j(x)$ in Theorem 2.11. The modes of convergence are in finite dimensional distributions and weakly w.r.t. the $L^p$ norm in $\mathbb{D}[0, \infty)$. See Section 2.4.2 for precise definitions and statements.

As a consequence of the sharp characterization of the global dynamics in Theorem 2.10, we also obtain Corollary 2.12, which proves an ergodic theorem for the fraction of the times $X^{\eta|b}_j(x)$ spends in narrow attraction fields: roughly speaking,

$$\frac{1}{T\cdot\lambda^*_b(\eta)}\sum_{i=1}^{T\cdot\lambda^*_b(\eta)}\mathbb{I}\Big\{X^{\eta|b}_i(x)\in\bigcup_{j:\ m_j\in\text{ wide minima}}(m_j-\epsilon,m_j+\epsilon)\Big\}\xrightarrow{p}1\qquad\text{as }\eta\downarrow 0$$

where $\lambda^*_b(\eta)$ is a scaling function of $\eta$ that is regularly varying with index $\mathcal{J}^*_b(V)\cdot(\alpha-1)+1$. Here $\mathcal{J}^*_b(V)$ is the maximal relative width of $U$'s attraction fields. This uncovers an intriguing phenomenon: combined with truncation, the heavy-tailed processes will avoid any local minimum of $U$ that is not the widest. The precise definitions of the widest attraction fields and the associated local minima are given in Section 2.4, but here we note that the width is measured by the number of jumps (with sizes bounded by $b$) required to exit the attraction field.

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 training deep neural networks. Compared to the brute force approach in [52], the current paper provides a systematic framework to characterize the global dynamics for significantly more general class of heavy-tailed dynamical systems.

The rest of the paper is organized as follows. Section 2 presents the main results of this paper. Section 3, Section 4, and Section 5 provide the proofs of Sections 2.1 and 2.2, Section 2.3, and Section 2.4, respectively. The results for SDEs driven by Lévy processes with regularly varying increments are collected in Appendix A.

# 2 Main Results

This section presents the main results of this paper and discusses their implications. Section 2.1 introduces the uniform version of $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence and presents an associated portmanteau theorem. 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. All the proofs are deferred to the later sections.

Before presenting the main results, we set frequently used notations. 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 metric space with $\mathscr{S}_\mathbb{S}$ being the corresponding Borel $\sigma$-algebra. For any $E\subseteq\mathbb{S}$, let $E^\circ$ 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 $\mu$ on $(\mathbb{S},\mathscr{S}_\mathbb{S})$, let the support of $\mu$ (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 $\text{supp}(g)\triangleq\big(\{x\in\mathbb{S}:\ g(x)\neq 0\}\big)^-$.

## 2.1 Uniform $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-Convergence

This section extends the notion of $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence [35, 45] to a uniform version and prove an associated portmanteau theorem. Such developments pave the way to the locally uniform heavy-tailed sample-path large deviations.

Specifically, in this section we consider some metric space $(\mathbb{S},\boldsymbol{d})$ that is complete and separable. 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 $\sigma$-algebra $\mathscr{S}_{\mathbb{S}\setminus\mathbb{C}}\triangleq\{A\in\mathscr{S}_\mathbb{S}:\ A\subseteq\mathbb{S}\setminus\mathbb{C}\}$. Let

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

$\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ 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 $\mu_n$ converges to $\mu$ 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 [35] 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 in terms of $\mathbb{M}$-convergence of the scaled process in the Skorokhod space. In this paper, we introduce a stronger version of $\mathbb{M}$-convergence, which facilitates the analysis of the local stability and global dynamics in the later sections.

**Definition 2.1** (Uniform $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence)**.** *Let $\Theta$ 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 $\mu^\eta_\theta$ converges to $\mu_\theta$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ uniformly in $\theta$ on $\Theta$ 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 is provided in Section 3.1.

**Theorem 2.2** (Portmanteau theorem for uniform $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence)**.** *Let $\Theta$ be a set of indices. Let $\mu^\eta_\theta,\ \mu_\theta \in \mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ for each $\eta > 0$ and $\theta \in \Theta$. If, 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 three statements are equivalent:*

*(i)* *$\mu^\eta_\theta$ converges to $\mu_\theta$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$ uniformly in $\theta$ on $\Theta$ as $\eta \downarrow 0$;*

*(ii)* *$\lim_{\eta\downarrow 0}\sup_{\theta\in\Theta}|\mu^\eta_\theta(f) - \mu_\theta(f)| = 0$ for each $f \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$ that is also uniformly continuous on $\mathbb{S}$;*

*(iii)* *$\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F) - \mu_\theta(F^\epsilon) \leq 0$ 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, any of the claims (i)–(iii) implies the following.*

*(iv)* *$\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F) \leq \sup_{\theta\in\Theta}\mu_\theta(F)$ 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 $\mathbb{C}$.*

To conclude, we provide two additional remarks regarding Theorem 2.2. First, it is not possible to strengthen statement (*iii*) and assert that

$$\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F) - \mu_\theta(F) \leq 0, \qquad \liminf_{\eta\downarrow 0}\inf_{\theta\in\Theta}\mu^\eta_\theta(G) - \mu_\theta(G) \geq 0 \tag{2.2}$$

for all closed $F \subseteq \mathbb{S}$ bounded away from $\mathbb{C}$ and all open $G \subseteq \mathbb{S}$ bounded away from $\mathbb{C}$. In other words, in statement (*iii*) the $\epsilon$-fattening in $F^\epsilon$ and $\epsilon$-shrinking in $G_\epsilon$ are indispensable. Indeed, we demonstrate through a counterexample that, due to the infinite cardinality of the collections of measures $\{\mu^\eta_\theta : \theta \in \Theta\}$ and $\{\mu_\theta :\ \theta \in \Theta\}$, the claims in (2.2) can easily fall apart while statements (*i*)–(*iii*) hold true. Specifically, by setting $\mathbb{C} = \emptyset$ and $\mathbb{S} = \mathbb{R}$, the $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence degenerates to the weak convergence of Borel measures on $\mathbb{R}$. Set $\Theta = [-1, 1]$ and

$$\mu^\eta_\theta \triangleq \boldsymbol{\delta}_{\theta-\eta}, \qquad \mu_\theta \triangleq \boldsymbol{\delta}_\theta,$$

where $\boldsymbol{\delta}_x$ is the Dirac measure at $x$. For closed set $F = [-1,0]$ and any $\eta \in (0,2)$,

$$\sup_{\theta\in\Theta} \mu^\eta_\theta(F) - \mu_\theta(F) \geq \boldsymbol{\delta}_{-\eta/2}\big([-1,0]\big) - \boldsymbol{\delta}_{\eta/2}\big([-1,0]\big) \qquad \text{by picking } \theta = \eta/2$$
$$= \mathbb{I}\bigg\{\frac{-\eta}{2} \in [-1,0]\bigg\} - \mathbb{I}\bigg\{\frac{\eta}{2} \in [-1,0]\bigg\} = 1,$$

thus implying $\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu^\eta_\theta(F) - \mu_\theta(F) \geq 1$.

Secondly, while statement $(iv)$ holds as the key component when establishing the sample-path large deviation results, it is indeed strictly weaker than the other claims for one obvious reason: unlike statements $(i)$–$(iii)$, the content of statement $(iv)$ does not require $\mu^\eta_\theta$ to converge to $\mu_\theta$ for any $\theta \in \Theta$. To illustrate that $(iv)$ does not imply $(i)$–$(iii)$, it suffices to examine the following case where $\mathbb{C} = \emptyset$, $\mathbb{S} = \mathbb{R}$, $\Theta = [-1,1]$, $\mu^\eta_\theta = \boldsymbol{\delta}_{-\theta}$, and $\mu_\theta = \boldsymbol{\delta}_\theta$.

## 2.2 Heavy-Tailed Large Deviations

In Section 2.2.1, we study the sample-path large deviations for stochastic difference equations with heavy-tailed increments. Section 2.2.2 then characterizes the catastrophe principle of heavy-tailed systems by presenting the conditional limit theorems that reveal a discrete hierarchy of the most likely scenarios and probabilities of rare events in heavy-tailed stochastic difference equations.

### 2.2.1 Sample-Path Large Deviations

Let $Z_1, Z_2, \ldots$ be the iid copies of some random variable $Z$ and $\mathcal{F}$ be the $\sigma$-algebra generated by $(Z_j)_{j\geq 1}$. Let $\mathcal{F}_j$ be the $\sigma$-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

$$X^\eta_0(x) = x; \qquad X^\eta_j(x) = X^\eta_{j-1}(x) + \eta a\big(X^\eta_{j-1}(x)\big) + \eta\sigma\big(X^\eta_{j-1}(x)\big)Z_j, \quad \forall j \geq 1 \tag{2.3}$$

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 $\phi$ is regularly varying as $x \to \infty$ with index $\beta$ (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 of 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 $\beta$ 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

$$H^{(+)}(x) \triangleq \mathbf{P}(Z > x), \quad H^{(-)}(x) \triangleq \mathbf{P}(Z < -x), \quad H(x) \triangleq H^{(+)}(x) + H^{(-)}(x) = \mathbf{P}(|Z| > x). \tag{2.4}$$

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*

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

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 notational 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)**.** $\sigma(x) > 0 \quad \forall x \in \mathbb{R}$.

**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 càdlàg functions on $[0,T]$ and $\boldsymbol{d}_{J_1}^{[0,T]}$ is the Skorodkhod $J_1$ metric

$$\boldsymbol{d}_{J_1}^{[0,T]}(x,y) \triangleq \inf_{\lambda \in \Lambda_T} \sup_{t \in [0,T]} |\lambda(t) - t| \vee |x(\lambda(t)) - y(t)|. \tag{2.5}$$

Here, $\Lambda_T$ is the set of all homeomorphism on $[0,T]$. Given any $A \subseteq \mathbb{R}$, let $A^{k\uparrow} \triangleq \{(t_1, \cdots, t_k) \in A^k : \ t_1 < t_2 < \cdots < t_k\}$ 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 $h_{[0,T]}^{(k)} : \mathbb{R} \times \mathbb{R}^k \times (0,T]^{k\uparrow} \to \mathbb{D}[0,T]$ 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 $\xi = h_{[0,T]}^{(k)}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in \mathbb{D}[0,T]$ be the solution to

$$\xi_0 = x_0 \tag{2.6}$$

$$\frac{d\xi_s}{ds} = a(\xi_s) \qquad \forall s \in [0,T], \ s \neq t_1, \cdots, t_k \tag{2.7}$$

$$\xi_s = \xi_{s-} + \sigma(\xi_{s-}) \cdot w_j \qquad \text{if } s = t_j \text{ for some } j \in [k]. \tag{2.8}$$

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 $\xi$ at $t$, and we set $\xi_{0-} = \xi_0$. In essence, the mapping $h_{[0,T]}^{(k)}(x_0, \boldsymbol{w}, \boldsymbol{t})$ 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 $\xi = h_{[0,T]}^{(0)}(x_0)$ 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 $\nu_\alpha$ be the (Borel) measure on $\mathbb{R}$ with

$$\nu_\alpha[x, \infty) = p^{(+)} x^{-\alpha}, \quad \nu_\alpha(-\infty, -x] = p^{(-)} x^{-\alpha}, \quad \forall x > 0. \tag{2.9}$$

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 $\mathcal{L}_t^{k\uparrow}$ be the Lebesgue measure restricted on $(0,t)^{k\uparrow}$. Given any $T > 0$, $x \in \mathbb{R}$, and $k \geq 0$, let

$$\mathbf{C}_{[0,T]}^{(k)}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{ h_{[0,T]}^{(k)}\big(x, \boldsymbol{w}, \boldsymbol{t}\big) \in \ \cdot \ \Big\} \nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_T^{k\uparrow}(d\boldsymbol{t}) \tag{2.10}$$

where $\nu_\alpha^k(\cdot)$ is the $k$-fold product measure of $\nu_\alpha$. For $\{X_j^\eta(x) : j \geq 0\}$, we define the time-scaled version of the sample path as

$$\boldsymbol{X}_{[0,T]}^\eta(x) \triangleq \big\{ X_{\lfloor t/\eta \rfloor}^\eta(x) : \ t \in [0,T] \big\}, \quad \forall T > 0 \tag{2.11}$$

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 $\boldsymbol{X}_{[0,T]}^\eta(x)$ is a $\mathbb{D}[0,T]$-valued random element. For any $k \in \mathbb{N}$ and $A \subseteq \mathbb{R}$, let

$$\mathbb{D}_A^{(k)}[0,T] \triangleq h_{[0,T]}^{(k)}\big(A \times \mathbb{R}^k \times (0,T]^{k\uparrow}\big), \quad \forall T > 0 \tag{2.12}$$

as the set that contains all ODE paths with $k$ perturbations by time $T$. We adopt the convention that $\mathbb{D}_A^{(-1)}[0,T] \triangleq \emptyset$. 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 $\mathbb{D}$, $\boldsymbol{d}_{J_1}$, $h^{(k)}$, $\mathbf{C}^{(k)}$, $\mathbb{D}_A^{(k)}$, and $\boldsymbol{X}^\eta(x)$ to denote $\mathbb{D}[0,1]$, $\boldsymbol{d}_{J_1}^{[0,1]}$, $h_{[0,1]}^{(k)}$, $\mathbf{C}_{[0,1]}^{(k)}$, $\mathbb{D}_A^{(k)}[0,1]$, and $\boldsymbol{X}_{[0,1]}^\eta(x)$, respectively.

Now, we are ready to state Theorem 2.3, which establishes the uniform $\mathbb{M}$-convergence of (the law of) $\boldsymbol{X}_{[0,T]}^\eta(x)$ to $\mathbf{C}^{(k)}(\ \cdot\ ; x)$ and a uniform version of the sample-path large deviations for $\boldsymbol{X}_{[0,T]}^\eta(x)$.

**Theorem 2.3.** *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 $\lambda^{-k}(\eta)\mathbf{P}\big(\boldsymbol{X}_{[0,T]}^\eta(x) \in\ \cdot\ \big) \to \mathbf{C}_{[0,T]}^{(k)}(\ \cdot\ ; x)$ in $\mathbb{M}\big(\mathbb{D}[0,T] \setminus \mathbb{D}_A^{(k-1)}[0,T]\big)$ 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 $\mathbb{D}_A^{(k-1)}[0,T]$,*

$$
\begin{aligned}
\inf_{x \in A} \mathbf{C}_{[0,T]}^{(k)}(B^\circ; x) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^\eta(x) \in B\big)}{\lambda^k(\eta)} \\
&\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^\eta(x) \in B\big)}{\lambda^k(\eta)} \leq \sup_{x \in A} \mathbf{C}_{[0,T]}^{(k)}(B^-; x) < \infty.
\end{aligned}
\tag{2.13}
$$

**Remark 1.** *We add a remark on the connection between (2.13) and the classical LDP framework. Given any measurable $B \subseteq \mathbb{D}[0,T]$, there is a particular $k$ that plays the role of the rate function. Specifically, let $\mathcal{J}_{A;T}(B) \triangleq \min\{k \in \mathbb{N} :\ B \cap \mathbb{D}_A^{(k)}[0,T] \neq \emptyset\}$. In great generality, this coincides with the unique value of $k \in \mathbb{N}$ for which the lower bound $\inf_{x \in A} \mathbf{C}_{[0,T]}^{(k)}(B^\circ; x)$ in (2.13) is strictly positive, and $\lambda^{\mathcal{J}_{A;T}(B)}(\eta)$ characterizes the exact rate of decay for both $\inf_{x \in A} \mathbf{P}(\boldsymbol{X}_{[0,T]}^\eta(x) \in B)$ and $\sup_{x \in A} \mathbf{P}(\boldsymbol{X}_{[0,T]}^\eta(x) \in B)$ as $\eta \downarrow 0$. It should be noted these results are exact asymptotics as opposed to the log asymptotics in classical LDP framework. In case that the set $A$ is a singleton, $T = 1$, $a \equiv 0$, and $\sigma \equiv 1$, the process $\boldsymbol{X}_{[0,T]}^\eta(x)$ will degenerate to a Lévy process, and $\mathcal{J}_{A;T}(\cdot)$ will reduce to $\mathcal{J}(\cdot)$ defined in equation (3.3) of [45]. This confirms that Theorem 2.3 is a proper generalization of the heavy-tailed large deviations for Lévy processes and random walks in [45].*

The proof of Theorem 2.3 will be given in Section 3.3. Interestingly enough, the results are obtained by first studying its truncated counterpart. 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

$$
X_0^{\eta|b}(x) = x, \qquad X_j^{\eta|b}(x) = X_{j-1}^{\eta|b}(x) + \varphi_b\Big(\eta a\big(X_{j-1}^{\eta|b}(x)\big) + \eta\sigma\big(X_{j-1}^{\eta|b}(x)\big)Z_j\Big) \quad \forall j \geq 1, \tag{2.14}
$$

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

$$
\varphi_c(w) \triangleq (w \wedge c) \vee (-c) \quad \forall w \in \mathbb{R}, c > 0. \tag{2.15}
$$

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 $\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \triangleq \{X_{\lfloor t/\eta \rfloor}^{\eta|b}(x) :\ t \in [0,T]\}$ be the time-scaled version of $X_j^{\eta|b}(x)$ embedded 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

$$
\xi_0 = x_0; \tag{2.16}
$$

$$
\frac{d\xi_s}{ds} = a(\xi_s) \quad \forall s \in [0,T],\ s \neq t_1, t_2, \cdots, t_k; \tag{2.17}
$$

$$
\xi_s = \xi_{s-} + \varphi_b\big(\sigma(\xi_{s-}) \cdot w_j\big) \quad \text{ if } s = t_j \text{ for some } j \in [k] \tag{2.18}
$$

The mapping $h_{[0,T]}^{(k)|b}$ can be interpreted as a truncated analog of the mapping $h_{[0,T]}^{(k)}$ defined in (2.6)–(2.8). In other words, $h_{[0,T]}^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t})$ returns an ODE path perturbed by $k$ jumps, but the size of each jump is truncated under $b$. For any $b$, $T > 0$, $A \subseteq \mathbb{R}$ and $k = 0, 1, 2, \ldots$, let

$$\mathbb{D}_A^{(k)|b}[0,T] \triangleq h_{[0,T]}^{(k)|b}\big(A \times \mathbb{R}^k \times (0,T]^{k\uparrow}\big) \tag{2.19}$$

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 $\mathbb{D}_A^{(-1)|b}[0,T] \triangleq \emptyset$. We collect and establish useful properties of mappings $h_{[0,T]}^{(k)}$, $h_{[0,T]}^{(k)|b}$ and sets $\mathbb{D}_A^{(k)}[0,T]$, $\mathbb{D}_A^{(k)|b}[0,T]$ in Section B.

Given any $x \in \mathbb{R}$, $k \geq 0$, $b > 0$, and $T > 0$, let

$$\mathbf{C}_{[0,T]}^{(k)|b}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{h_{[0,T]}^{(k)|b}\big(x, \boldsymbol{w}, \boldsymbol{t}\big) \in \ \cdot\ \Big\}\nu_\alpha^k(d\boldsymbol{w}) \times \mathcal{L}_T^{k\uparrow}(d\boldsymbol{t}). \tag{2.20}$$

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 2.4.** *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 $\mathbb{D}_A^{(k-1)|b}[0,T]$,*

$$\begin{aligned}
\inf_{x \in A} \mathbf{C}_{[0,T]}^{(k)|b}\big(B^\circ; x\big) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \in B\big)}{\lambda^k(\eta)} \\
&\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}\big(\boldsymbol{X}_{[0,T]}^{\eta|b}(x) \in B\big)}{\lambda^k(\eta)} \leq \sup_{x \in A} \mathbf{C}_{[0,T]}^{(k)|b}\big(B^-; x\big) < \infty.
\end{aligned} \tag{2.21}$$

Here, we provide a high-level description of the proof strategy for Theorems 2.3 and 2.4. Specifically, the proof of Theorem 2.4 consists of two steps.

- First, we establish the asymptotic equivalence between $\boldsymbol{X}_{[0,T]}^{\eta|b}(x)$ 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 3.3 that tightly control the fluctuations in $X_j^{\eta|b}(x)$ 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 3.4.

See Section 3.3 for the detailed proof and the rigorous definitions of the concepts involved. Regarding Theorem 2.3, note that for any $b$ sufficiently large, it is highly likely that $X_j^\eta(x)$ coincides with $X_j^{\eta|b}(x)$ for the entire period of $j \leq T/\eta$ (that is, the truncation operator $\varphi_b$ did not come into effect for a long period due to the truncation threshold $b > 0$ being large). By sending $b \to \infty$ and carefully analyzing the limits involved, we recover the sample-path large deviations for $X_j^\eta(x)$ and prove Theorem 2.3.

### 2.2.2 Catastrophe Principle

Perhaps the most important implication of large deviations bounds is the identification of conditional distributions of the stochastic processes given the rare events of interest. This section precisely identifies the distributional limits of the conditional laws of $\boldsymbol{X}^\eta(x)$ and $\boldsymbol{X}^{\eta|b}(x)$, respectively, given the rare events.

More precisely, the conditional limit theorem below follows immediately from the sample-path large deviations established above, i.e., (2.13) and (2.21). While all the results in Section 2.2.2 can be easily extended to $\mathbb{D}[0,T]$ with arbitrary $T \in (0,\infty)$, we focus on $\mathbb{D} = \mathbb{D}[0,1]$ for the sake of clarity of presentation.

**Corollary 2.5.** *Let Assumptions 1, 2, and 3 hold.*

*(i) For some $b > 0$, $k \in \mathbb{N}$, $x \in \mathbb{R}$, and measurable $B \subseteq \mathbb{D}$, suppose that $B$ is bounded away from $\mathbb{D}^{(k-1)|b}_{\{x\}}$, $B \cap \mathbb{D}^{(k)|b}_{\{x\}} \neq \emptyset$, and $\mathbf{C}^{(k)|b}(B^\circ) = \mathbf{C}^{(k)|b}(B^-) > 0$. Then*

$$\mathbf{P}\big(\boldsymbol{X}^{\eta|b}(x) \in \cdot \mid \boldsymbol{X}^{\eta|b}(x) \in B\big) \Rightarrow \frac{\mathbf{C}^{(k)|b}(\,\cdot \cap B; x)}{\mathbf{C}^{(k)|b}(B; x)} \qquad \text{as } \eta \downarrow 0.$$

*(ii) Furthermore, suppose that Assumption 4 holds. For some $k \in \mathbb{N}$, $x \in \mathbb{R}$, and measurable $B \subseteq \mathbb{D}$, suppose that $B$ is bounded away from $\mathbb{D}^{(k-1)}_{\{x\}}$, $B \cap \mathbb{D}^{(k)}_{\{x\}} \neq \emptyset$, and $\mathbf{C}^{(k)}(B^\circ) = \mathbf{C}^{(k)}(B^-) > 0$. Then*

$$\mathbf{P}\big(\boldsymbol{X}^{\eta}(x) \in \cdot \,\big|\, \boldsymbol{X}^{\eta}(x) \in B\big) \Rightarrow \frac{\mathbf{C}^{(k)}(\cdot \cap B; x)}{\mathbf{C}^{(k)}(B; x)} \qquad \text{as } \eta \downarrow 0.$$

**Remark 2.** *Note that Corollary 2.5 is a sharp characterization of* catastrophe principle *for $\boldsymbol{X}^{\eta|b}(x)$ and $\boldsymbol{X}^{\eta}(x)$. By definition of measures $\mathbf{C}^{(k)|b}$, its support is on the set of paths of the form*

$$h^{(k)|b}\big(x, (w_1, \cdots, w_k), (u_1, \cdots, u_k)\big),$$

*where the mapping $h^{(k)|b}$ is defined in (2.16)–(2.18), and $w_i$'s are bounded from below; see, for instance, Lemma 3.5 and 3.6. This is a clear manifestation of the catastrophe principle: whenever the rare event arises, the conditional distribution resembles the nominal path (i.e., the solution of the associated ODE) perturbed by precisely $k$ jumps. In fact, the definition of $\mathbf{C}^{(k)|b}$ also implies that the the jump sizes are Pareto (modulated by $\sigma(\cdot)$) and the jump times are uniform, conditional on the perturbed path belonging to $B$. Similar interpretation applies to $\boldsymbol{X}^{\eta}(x)$ in part (ii).*

## 2.3 Local Stability Analysis

This section analyzes the local stability of $X^{\eta}_j(x)$ and $X^{\eta|b}_j(x)$. Section 2.3.1 establishes the scaling limits of their exit times. Section 2.3.2 introduces a framework that facilitates such analysis for general Markov chains.

### 2.3.1 First Exit Times

In this section, we analyze the first exit times of $X^{\eta}_j(x)$ and $X^{\eta|b}_j(x)$ 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 $s_{\text{left}} < 0 < s_{\text{right}}$ w.l.o.g. We impose the following assumption on $a(\cdot)$.

**Assumption 5.** *$a(0) = 0$. Besides, it holds for all $x \in I \setminus \{0\}$ that $a(x)x < 0$.*

Of particular interest is the case where $a(\cdot) = -U'(\cdot)$ for some potential $U \in \mathcal{C}^1(\mathbb{R})$. Assumption 5 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

$$\boldsymbol{y}_0(x) = x, \qquad \frac{d\boldsymbol{y}_t(x)}{dt} = a\big(\boldsymbol{y}_t(x)\big) \ \ \forall t \geq 0. \tag{2.22}$$

It is worth noticing that Assumption 5 is more flexible than the assumptions in other related works in the literature. For instance, [24, 23] 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, Assumption 5 is close to minimum assumption required for $I$ to be an attraction field associated with the origin.

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 $\check{g}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0,\infty)^{k\uparrow} \to \mathbb{R}$ as the location of the perturbed ODE at the last jump time:

$$\check{g}^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \triangleq h^{(k)|b}_{[0,t_k+1]}(x, \boldsymbol{w}, \boldsymbol{t})(t_k) \tag{2.23}$$

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 (2.16)-(2.18). For $k = 0$, we adopt the convention that $\check{g}^{(0)|b}(x) = x$. This allows us to define Borel measures (for each $k \geq 1$ and $b > 0$)

$$\check{\mathbf{C}}^{(k)|b}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{\check{g}^{(k-1)|b}\big(x + \varphi_b\big(\sigma(x)\cdot w_0\big), \boldsymbol{w}, \boldsymbol{t}\big) \in\ \cdot\ \Big\}\nu_\alpha(dw_0) \times \nu_\alpha^{k-1}(d\boldsymbol{w}) \times \mathcal{L}^{k-1\uparrow}_\infty(d\boldsymbol{t}) \tag{2.24}$$

with $\mathcal{L}^{k\uparrow}_\infty$ being the Lebesgue measure restricted on $\{(t_1, \cdots, t_k) \in (0,\infty)^k :\ 0 < t_1 < t_2 < \cdots < t_k\}$. Section C collects useful properties of the measure $\check{\mathbf{C}}^{(k)|b}$. Also, define

$$\check{\mathbf{C}}(\ \cdot\ ; x) \triangleq \int \mathbb{I}\Big\{x + \sigma(x)\cdot w \in\ \cdot\ \Big\}\nu_\alpha(dw). \tag{2.25}$$

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

$$l \triangleq \inf_{x \in I^c} |x| = |s_{\text{left}}| \wedge s_{\text{right}}, \qquad \mathcal{J}^*_b \triangleq \lceil l/b \rceil. \tag{2.26}$$

Here, $l$ is the distance between the origin and $I^c$, and $\mathcal{J}^*_b$ is the smallest 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 of $X^{\eta|b}_j(x)$ and $X^{\eta}_j(x)$. The results are obtained through a general machinery we develop in Section 2.3.2. The proof of Theorem 2.6 is provided in Section 4.2.

**Theorem 2.6.** *Let Assumptions 1, 2, 3, and 5 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 $B \subseteq I^c$,*

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

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

*where $C^*_b \triangleq \check{\mathbf{C}}^{(\mathcal{J}^*_b)|b}(I^c) \in (0,\infty)$.*

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

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

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

*where $C^* \triangleq \check{\mathbf{C}}(I^c) \in (0,\infty)$.*

### 2.3.2 General Framework

This section proposes a general framework that enables sharp characterization of exit times and exit locations of Markov chains. The new heavy-tailed large deviations formulation introduced in Section 2.2 is conducive to this framework.

Consider a general metric space $(\mathbb{S}, \boldsymbol{d})$ and a family of $\mathbb{S}$-valued Markov chains $\big\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\big\}$ parameterized by $\eta$, where $x \in \mathbb{S}$ denotes the initial state and $j$ denotes the time index. We use $\boldsymbol{V}_{[0,T]}^\eta(x) \triangleq \{V_{\lfloor t/\eta \rfloor}^\eta(x) : t \in [0,T]\}$ to denote the scaled version of $\{V_j^\eta(x) : j \geq 0\}$ as a $\mathbb{D}[0,T]$-valued random element. 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 $\boldsymbol{V}_{[0,T]}^\eta(x)$ typically (i.e., as $\eta \downarrow 0$) stays within $I$ throughout any fixed time horizon $[0,T]$ as far as the initial state $x$ is in $I$. However, if one considers an infinite time horizon, $V_\cdot^\eta(x)$ is typically bound to escape $I$ eventually due to the stochasticity. The goal of this section is to establish an asymptotic limit of the joint distribution of the exit time $\tau_{I^c}^\eta(x)$ and the exit location $V_{\tau_{I^c}^\eta(x)}^\eta(x)$. Throughout this section, we will denote $V_{\tau_{I(\epsilon)^c}^\eta(x)}^\eta(x)$ and $V_{\tau_{I^c}^\eta(x)}^\eta(x)$ with $V_{\tau_\epsilon}^\eta(x)$ and $V_\tau^\eta(x)$, respectively, for notation simplicity.

We introduce the notion of asymptotic atoms 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 2.7.** *$\big\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\big\}$ 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*

$$C(B^\circ) - \delta_B(\epsilon, T) \leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A(\epsilon)} \mathbf{P}\big(\tau_{I(\epsilon)^c}^\eta(x) \leq T/\eta;\ V_{\tau_\epsilon}^\eta(x) \in B\big)}{\gamma(\eta)T/\eta} \tag{2.27}$$

$$\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A(\epsilon)} \mathbf{P}\big(\tau_{I(\epsilon)^c}^\eta(x) \leq T/\eta;\ V_{\tau_\epsilon}^\eta(x) \in B\big)}{\gamma(\eta)T/\eta} \leq C(B^-) + \delta_B(\epsilon, T) \tag{2.28}$$

$$\limsup_{\eta \downarrow 0} \frac{\sup_{x \in I(\epsilon)} \mathbf{P}\big(\tau_{(I(\epsilon)\setminus A(\epsilon))^c}^\eta(x) > T/\eta\big)}{\gamma(\eta)T/\eta} = 0 \tag{2.29}$$

$$\liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon)} \mathbf{P}\big(\tau_{A(\epsilon)}^\eta(x) \leq T/\eta\big) = 1 \tag{2.30}$$

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

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

To see how Definition 2.7 asymptotically characterize the atoms in $V_\cdot^\eta(x)$ for the first exit analysis from domain $I$, note that the condition (2.30) requires the process to efficiently return to the asymptotic atoms $A(\epsilon)$. The conditions (2.27) and (2.28) then state that, upon hitting the asymptotic atoms $A(\epsilon)$, the process almost regenerates in terms of the law of the exit time $\tau_{I(\epsilon)^c}^\eta(x)$ and exit locations $V_{\tau_\epsilon}^\eta(x)$. Furthermore, the condition (2.29) prevents the process $V_\cdot^\eta(x)$ from spending a long time without either returning to the asymptotic atoms $A(\epsilon)$ or exiting from $I(\epsilon)$, which covers the domain $I$ as $\epsilon$ tends to 0.

The existence of an asymptotic atom is a sufficient condition for characterization of exit time and location asymptotics as in Theorem 2.6. To minimize repetition, we refer to the existence of an asymptotic atom—with specific domain, location measure, scale, and covering—Condition 1 throughout the paper.

**Condition 1.** *A family $\big\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\big\}$ 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 this section. See Section 4.1 for the proof of the theorem.

**Theorem 2.8.** *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 $t \geq 0$,*

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

## 2.4 Characterization of Global Dynamics

In this section, we consider the case where $a(\cdot) = -U'(\cdot)$ for some general multimodal potential function $U : \mathbb{R} \to \mathbb{R}$ and characterize the global behavior of $X^{\eta}_j(x)$ and $X^{\eta|b}_j(x)$. We show that, after proper scaling, their sample paths converge to those of Markov jump processes whose state spaces consist of local minima. Curiously, the state space of the limit process associated with $X^{\eta|b}_j(x)$ consists of only the widest local minima.

### 2.4.1 Problem Setting

We consider a multimodal potential function with local minima $\{m_1, m_2, \ldots, m_{n_{min}}\}$. More precisely, we make the following assumption throughout this section.

**Assumption 6.** *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 $n_{\min} = 3$. Note that the local maxima $s_1, \cdots, s_{n_{\min}-1}$ divide $\mathbb{R}$ into different regions $I_i \triangleq (s_{i-1}, s_i)$ for $i = 0, \ldots, n_{min}$. Such regions can be viewed as the *attraction fields* of the local minima $m_i$'s. That is, the ODE $\boldsymbol{y}_t(x)$ defined in (2.22) (with $a = -U'$) admits the limit $\lim_{t\to\infty} \boldsymbol{y}_t(x) = m_i$ if $x \in I_i$.

Building upon the first exit time analysis in Section 2.3, we characterize the global dynamics of $X^{\eta}_j(x)$ and $X^{\eta|b}_j(x)$ in this section. Note that we impose the condition $n_{\min} \geq 2$ simply to avoid the trivial case of $n_{\min} = 1$, in which case there exists only one attraction field. In order to present the main results, we introduce some definitions to facilitate the characterization of the geometry of $U$. First, for each attraction field $I_i$, let

$$l_i \triangleq \inf_{x \in I_i^c} |x - m_i| = |m_i - s_{i-1}| \wedge |s_i - m_i| \tag{2.31}$$

be the effective "width" of $I_i$, i.e., the minimum distance between $m_i$ and the outside of $I_i$. Next, for any $i \in [n_{\min}]$ and $j \in [n_{\min}]$ with $j \neq i$, let

$$\mathcal{J}^*_b(i) \triangleq \lceil l_i/b \rceil, \qquad l_{i,j} \triangleq \inf_{x \in I_j} |x - m_i| = \begin{cases} s_{j-1} - m_i & \text{if } j > i \\ m_i - s_j & \text{if } j < i \end{cases}, \qquad \mathcal{J}^*_b(i,j) \triangleq \lceil l_{i,j}/b \rceil. \tag{2.32}$$

Recall that $b$ is the truncation threshold for $X^{\eta|b}_j(x)$. Here, $\mathcal{J}^*_b(i)$ can be interpreted as the *discretized width* of $I_i$ w.r.t. the *resolution* $b$, in the sense that it is the minimum number of jumps (with sizes bounded by $b$) required to escape from $I_i$. Furthermore, $\mathcal{J}^*_b(i,j)$ is the minimal number of steps required to travel from $m_i$ to $I_j$ under the truncation threshold $b$. By definition, we must have $\mathcal{J}^*_b(i,j) \geq \mathcal{J}^*_b(i)$. With $\mathcal{J}^*_b(i)$ and $\mathcal{J}^*_b(i,j)$, we define the *typical transition graph* as follows.

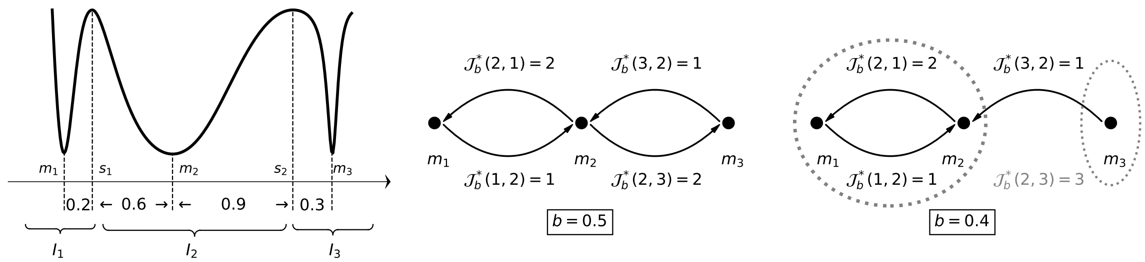


Figure 2.1: Typical transition graphs $\mathcal{G}_b$ associated with different gradient clipping thresholds $b$. **(Left)** The potential function $U$ illustrated here has 3 attraction fields. For the second one $I_2 = (s_1, s_2)$, we have $s_2 - m_2 = 0.9, m_2 - s_1 = 0.6$. **(Middle)** The typical transition graph associated with $b = 0.5$. The entire graph $\mathcal{G}_b$ is irreducible since all nodes communicate with each other. **(Right)** The typical transition graph associated with $b = 0.4$. 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 2.9** (Typical Transition Graph)**.** *Given a function $U$ satisfying Assumption 6 and some $b > 0$, the typical transition graph associated with threshold $b$ 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. 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. Specifically, we impose the following assumption on the truncation threshold $b$. We note that the second condition is a mild one, as it holds for almost every $b > 0$ except for countably many.

**Assumption 7.** *$b \in (0, \infty)$ is such that*

- *$\mathcal{G}_b$ is irreducible,*
- *$|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{min}]$ and $j \in [n_{min} - 1]$.*

### 2.4.2 Sample Path Convergence

We are now ready to present the main result of this section. Theorem 2.10 establishes that, under a proper time scaling, the sample path of $X_j^{\eta|b}(x)$ converges to that of a Markov jump process, which only visits the widest local minima of $U$. Here, the width of each attraction field $I_i$ is characterized by $\mathcal{J}_b^*(i)$ defined in (2.32). We use

$$\mathcal{J}_b^*(V) \triangleq \max_{i:\ m_i \in V} \mathcal{J}_b^*(i) \tag{2.33}$$

to denote the largest width—when discretized w.r.t. the resolution $b$—among all attraction fields. Next, define

$$V_b^* \triangleq \{m_i :\ \mathcal{J}_b^*(i) = \mathcal{J}_b^*(V)\} \tag{2.34}$$

as the set containing all the widest local minima.

In Theorem 2.10, the scaling limit of the sample path of $X^{\eta|b}_j(x)$ will be characterized in terms of the following two modes of convergence. First, we say that $\{S^\eta_t : t > 0\}$ converges to $\{S^*_t : t > 0\}$ *in finite-dimensional distributions* (f.d.d.) if we have $\big(S^\eta_{t_1}, \cdots, S^\eta_{t_k}\big) \Rightarrow \big(S^*_{t_1}, \cdots, S^*_{t_k}\big)$ as $\eta \downarrow 0$ for any $k \geq 1$ and $0 < t_1 < t_2 < \cdots < t_k < \infty$. We also denote this as $\{S^\eta_t : t > 0\} \stackrel{f.d.d.}{\to} \{S^*_t : t > 0\}$.

**Remark 3.** *We consider the convergence in f.d.d. only on $(0, \infty)$, thus excluding $t = 0$. This is because in Theorem 2.10, under the proper time scaling, the value of $X^{\eta|b}_{\lfloor t \rfloor}(x)$ at $t$ close to 0 quickly converges to that of $Y^{*|b}_0$ (i.e., the initial value of the limit Markov jump process), but not exactly at $t = 0$. The same applies to Theorem 2.11.*

Next, we recall the convergence w.r.t. the $L_p$ topology in $\mathbb{D}[0, \infty)$. For any $p \in [1, \infty)$ and $T \in (0, \infty)$, let

$$\boldsymbol{d}^{[0,T]}_{L_p}(x, y) \triangleq \left( \int_0^T |x_t - y_t|^p dt \right)^{1/p} \qquad \forall x, y \in \mathbb{D}[0, T] \tag{2.35}$$

be the $L_p$ metric on $\mathbb{D}[0, T]$. For any $T > 0$, define the projection $\pi_T : \mathbb{D}[0, \infty) \to \mathbb{D}[0, T]$ such that

$$\pi_T(\xi)_t = \xi_t \qquad \forall t \in [0, T]. \tag{2.36}$$

Now, we define

$$\boldsymbol{d}^{[0,\infty)}_{L_p}(x, y) \triangleq \sum_{k \geq 1} \frac{1 \wedge \boldsymbol{d}^{[0,k]}_{L_p}\big(\pi_k(x), \pi_k(y)\big)}{2^k} \qquad \forall x, y \in \mathbb{D}[0, \infty) \tag{2.37}$$

and note that $\boldsymbol{d}^{[0,\infty)}_{L_p}$ is a metric on $\mathbb{D}[0, \infty)$. Hereafter in this paper, the continuity of a functional $f : \mathbb{D}[0, \infty) \to \mathbb{R}$ is understood w.r.t. the topology induced by $\boldsymbol{d}^{[0,\infty)}_{L_p}$. We say that the sequence of càdlàg processes $\{S^\eta_t : \ t \geq 0\}$ converges in distribution to $\{S^*_t : \ t \geq 0\}$ *w.r.t. the $L_p$ topology in* $\mathbb{D}[0, \infty)$ as $\eta \downarrow 0$ if $\lim_{\eta \downarrow 0} \mathbf{E} f(S^\eta_\cdot) = \mathbf{E} f(S^*_\cdot)$ for all $f : \mathbb{D}[0, \infty) \to \mathbb{R}$ that is bounded and continuous. We denote this with $S^\eta_\cdot \Rightarrow S^*_\cdot$ in $(\mathbb{D}[0, \infty), \boldsymbol{d}^{[0,\infty)}_{L_p})$ or $\{S^\eta_t : t \geq 0\} \Rightarrow \{S^*_t : t \geq 0\}$ in $(\mathbb{D}[0, \infty), \boldsymbol{d}^{[0,\infty)}_{L_p})$.

Recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$ and $\lambda(\eta) = \eta^{-1} H(\eta^{-1}) \in \mathcal{RV}_{\alpha-1}(\eta)$. Define a scaling function

$$\lambda^*_b(\eta) \triangleq \eta \cdot \big(\lambda(\eta)\big)^{\mathcal{J}^*_b(V)} \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)+1}(\eta) \qquad \text{as } \eta \downarrow 0. \tag{2.38}$$

We are now ready to state the main result.

**Theorem 2.10.** *Let Assumptions 1, 2, 3, 6, and 7 hold. Let $p \in [1, \infty)$, $i \in [n_{min}]$, and $x \in I_i$. As $\eta \downarrow 0$,*

$$\{X^{\eta|b}_{\lfloor \cdot / \lambda^*_b(\eta) \rfloor}(x) : \ t > 0\} \stackrel{f.d.d.}{\to} \{Y^{*|b}_t : \ t > 0\} \quad \textit{and} \quad X^{\eta|b}_{\lfloor \cdot / \lambda^*_b(\eta) \rfloor}(x) \Rightarrow Y^{*|b}_\cdot \ \textit{in} \ (\mathbb{D}[0, \infty), \boldsymbol{d}^{[0,\infty)}_{L_p}),$$

*where $Y^{*|b}_t$ is a continuous-time Markov chain with a finite state space $V^*_b$, initial distribution (see (2.41) for the definition of $\theta_b$),*

$$\mathbf{P}(Y^{*|b}_0 = m_j) = \theta_b(m_j | m_i) \qquad \forall m_j \in V^*_b, \tag{2.39}$$

*and infinitesimal generator (see (2.40) for the definition of $q_b$)*

$$Q^{*|b}(i, j) = \sum_{j' \in [n_{min}]: \ j' \neq i} q_b(i, j')\theta_b(m_j | m_{j'}) \qquad \forall m_i, \ m_j \in V^*_b \ \textit{with} \ m_i \neq m_j,$$
$$Q^{*|b}(i, i) = -\sum_{m_j \in V^*_b: \ j \neq i} Q^{*|b}(i, j) \qquad \forall m_i \in V^*_b.$$

We provide the proof of Theorem 2.10 in Section 5.2. Here, we specify the law of limiting Markov jump process $Y^{*|b}$. Recall the definition of $\check{\mathbf{C}}^{(k)|b}(\ \cdot\ ; x)$ in (2.24). Let

$$q_b(i,j) \triangleq \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}(I_j; m_i), \qquad q_b(i) \triangleq \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}(I_i^c; m_i). \tag{2.40}$$

Note that $\sum_{j\in[n_{\min}]:\ j\neq i} q_b(i,j) = q_b(i) \in (0,\infty)$ for any $i \in [n_{\min}]$; see (5.14). This allows us to define a discrete-time Markov chain $(S_n)_{n\geq 0}$ over state space $V$, with any state $v \in V_b^*$ being an absorbing state, such that the one-step transition kernel $\mathbf{P}(S_{n+1} = m_j | S_n = m_i) = q_b(i,j)/q_b(i)$ holds for any $m_i \in V \setminus V_b^*$ and any $m_j \in V$. Next, define

$$\theta_b(m_j|m_i) \triangleq \mathbf{P}(S_n = m_j \text{ for some } n \geq 0 \mid S_0 = m_i) \tag{2.41}$$

for any $m_i \in V$ and any $m_j \in V_b^*$ as the probability of being absorbed at $m_j$ when starting from $m_i$. For any $m_i \in V_b^*$, by definition of $\theta_b(\cdot|m_i)$, we see that $\theta_b(m_i|m_i) = 1$. In case that $m_i \in V \setminus V_b^*$, the evaluation of $\theta_b(m_j|m_i)$ is straightforward using the fundamental matrix of the Markov chain; see, for instance, Chapter 3.3 of [29]. Lastly, given the generator of $Y^{*|b}$, we have

$$\mathbf{P}(Y^{*|b}_{t+h} = m_j \mid Y^{*|b}_t = m_i) = h \cdot \sum_{j'\in[n_{\min}]:\ j'\neq i} q_b(i,j')\theta_b(m_j|m_{j'}) + \boldsymbol{o}(h) \qquad \text{as } h \downarrow 0$$

for any $m_i,\ m_j \in V_b^*$ with $m_i \neq m_j$.

Moving onto the untruncated process $X_j^\eta(x)$, Theorem 2.11 establishes a sample-path level convergence of $X_j^\eta(x)$ by sending $b \to \infty$ in Theorem 2.10. In particular, given any $T > 0$, there is a high chance that $X_j^\eta(x)$ coincides with the truncated dynamics $X_j^{\eta|b}(x)$ for all $j \leq T$ if the truncation threshold $b$ is large. Therefore, as the truncation threshold $b$ of $X_j^{\eta|b}(x)$ tends to $\infty$ in Theorem 2.10, we recover the results for $X_j^\eta(x)$. More precisely, recall the definition of measure $\check{\mathbf{C}}(\ \cdot\ ; x)$ in (2.25). For $i,j \in [n_{\min}]$ with $i \neq j$, let

$$q(i,j) \triangleq \check{\mathbf{C}}(I_j; m_i), \qquad q(i) \triangleq \sum_{j\in[n_{\min}]:\ j\neq i} q(i,j). \tag{2.42}$$

Recall that $H(\cdot) = \mathbf{P}(|Z| > \cdot)$. Theorem 2.11 shows that, under time scaling $1/H(\eta^{-1})$, the process $X_j^\eta$ converges in distribution to a Markov jump process at the sample-path level. The proof is given in Section 5.3.

**Theorem 2.11.** *Let Assumptions 1, 2, 3, 4, and 6 hold. Let $p \in [1,\infty)$, $i \in [n_{min}]$, and $x \in I_i$. As $\eta \downarrow 0$,*

$$\{X^\eta_{\lfloor t/H(\eta^{-1})\rfloor}(x) :\ t > 0\} \overset{f.d.d.}{\to} \{Y_t^* :\ t > 0\} \quad \textit{and} \quad X^\eta_{\lfloor \cdot/H(\eta^{-1})\rfloor}(x) \Rightarrow Y^*_\cdot \ \textit{in}\ (\mathbb{D}[0,\infty), \boldsymbol{d}_{L_p}^{[0,\infty)})$$

*where $Y_t^*$ is a continuous-time Markov chain with a finite state space $V$, initial value $Y_0^* = m_i$, and infinitesimal generator*

$$Q^*(i,j) = q(i,j) \qquad \forall m_i,\ m_j \in V \ \textit{with}\ m_i \neq m_j,$$
$$Q^*(i,i) = -\sum_{j\in[n_{min}]:\ j\neq i} Q^*(i,j) = -q(i) \qquad \forall m_i \in V.$$

Finally, we state a direct corollary of Theorem 2.10 that highlights the elimination of sharp minima under truncated heavy-tailed dynamics. Theorem 2.10 reveals that, under small $\eta$, the sample path of the truncated dynamics $X_j^{\eta|b}(x)$ closely resembles that of a Markov jump process that completely avoids all the narrower attraction fields of the potential $U$. Corollary 2.12 then further demonstrates that the fraction of time $X_j^{\eta|b}(x)$ spends around sharp minima converges in probability to 0 as $\eta \downarrow 0$.

This result follows directly from Theorem 2.10 and the continuous mapping argument. In particular, given any $\epsilon,\ T > 0$ and mapping

$$f(\xi) = \frac{1}{T}\int_0^T \mathbb{I}\Big\{\xi_t \in \bigcup_{j:\ m_j \in V_b^*} (m_j - \epsilon, m_j + \epsilon)\Big\}dt,$$

one can see that $f : \mathbb{D}[0,\infty) \to \mathbb{R}$ is continuous at any $\xi$ that, over $[0,T]$, only takes values in $V_b^*$ and only makes finitely many jumps.

**Corollary 2.12.** *Let Assumptions 1, 2, 3, 6, and 7 hold. For any $i \in [n_{min}]$, $x \in I_i$, $T > 0$, and $\epsilon > 0$,*

$$\frac{1}{T \cdot \lambda_b^*(\eta)} \sum_{i=1}^{T\cdot\lambda_b^*(\eta)} \mathbb{I}\Big\{X_i^{\eta|b}(x) \in \bigcup_{j:\ m_j \in V_b^*} (m_j - \epsilon, m_j + \epsilon)\Big\} \xrightarrow{p} 1 \qquad \text{as } \eta \downarrow 0$$

*where $\xrightarrow{p}$ stands for convergence in probability.*

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

Here, we collect the proofs for Sections 2.1 and 2.2. Specifically, Section 3.1 provides the proof of Theorem 2.2, i.e., the Portmanteau theorem for the uniform $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence. Section 3.2 further develops a set of technical tools, which will then be applied to establish the sample-path large deviations results (i.e., Theorems 2.3 and 2.4) in Section 3.3.

## 3.1 Proof of Theorem 2.2

*Proof of Theorem 2.2.* **Proof of** $(i) \Rightarrow (ii)$. It follows directly from Definition 2.1.

**Proof of** $(ii) \Rightarrow (iii)$. We consider a proof by contradiction. Suppose that the upper bound $\limsup_{\eta\downarrow 0}\sup_{\theta\in\Theta}\mu_\theta^\eta(F) - \mu_\theta(F^\epsilon) \le 0$ does not hold for some closed $F$ bounded away from $\mathbb{C}$ and some $\epsilon > 0$. Then there exist a sequence $\eta_n \downarrow 0$, a sequence $\theta_n \in \Theta$, and some $\delta > 0$ such that $\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) > \delta\ \forall n \ge 1$. Now, we make two observations. First, using Urysohn's lemma (see, e.g., lemma 2.3 of [35]), one can identify some $f \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$, which is also uniformly continuous on $\mathbb{S}$, such that $\mathbb{I}_F \le f \le \mathbb{I}_{F^\epsilon}$. This leads to the bound $\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \le \mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)$ for each $n$. Secondly, from statement $(ii)$ we get $\lim_{n\to\infty}\left|\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)\right| = 0$. In summary, we yield the contradiction

$$\limsup_{n\to\infty}\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \le \limsup_{n\to\infty}\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f) \le \lim_{n\to\infty}\left|\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)\right| = 0.$$

Analogously, if the claim $\liminf_{\eta\downarrow 0}\inf_{\theta\in\Theta}\mu_\theta^\eta(G) - \mu_\theta(G^\epsilon) \ge 0$, supposedly, does not hold for some open $G$ bounded away from $\mathbb{C}$ and some $\epsilon > 0$, then we can yield a similar contradiction by applying Urysohn's lemma and constructing some uniformly continuous $g \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$ such that $\mathbb{I}_{G_\epsilon} \le g \le \mathbb{I}_G$. This concludes the proof of $(ii) \Rightarrow (iii)$.

**Proof of** $(iii) \Rightarrow (i)$**.** Again, we proceed with a proof by contradiction. Suppose that the claim $\lim_{\eta\downarrow 0}\sup_{\theta\in\Theta}\left|\mu_\theta^\eta(g) - \mu_\theta(g)\right| = 0$ 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

$$|\mu_{\theta_n}^{\eta_n}(g) - \mu_{\theta_n}(g)| > \delta \qquad \forall n \ge 1. \tag{3.1}$$

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 $\mathbb{C}$. First, using claims in $(iii)$, we get $\limsup_{n\to\infty}\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \le 0$

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 condition (2.1), by picking a sub-sequence of $\theta_n$ if necessary we can find some $\mu_{\theta^*}$ 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 [35]), we yield $\limsup_{n\to\infty} \mu_{\theta_n}(F^\epsilon) \leq \mu_{\theta^*}(F^\epsilon)$ and $\liminf_{n\to\infty} \mu_{\theta_n}(G_\epsilon) \geq \mu_{\theta^*}(G_\epsilon)$. 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 $\lim_{n\to\infty} \left|\mu^{\eta_n}_{\theta_n}(g) - \mu_{\theta^*}(g)\right| = 0$ for the $g \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$ fixed in (3.1). However, recall that we have already obtained $\lim_{n\to\infty} \left|\mu_{\theta_n}(g) - \mu_{\theta^*}(g)\right| = 0$ using assumption (2.1). We now arrive at the contradiction

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

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

**Proof of** $(i) \Rightarrow (iv)$**.** Due to the equivalence of $(i)$, $(ii)$, and $(iii)$, it only remains to show that $(i) \Rightarrow (iv)$. Suppose, for the sake of contradiction, that the claim $\limsup_{\eta\downarrow 0} \sup_{\theta\in\Theta} \mu^\eta_\theta(F) \leq \sup_{\theta\in\Theta} \mu_\theta(F)$ in $(iv)$ 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 $\theta_n$ if necessary we can find some $\mu_{\theta^*}$ such that $\lim_{n\to\infty} \left|\mu_{\theta_n}(f) - \mu_{\theta^*}(f)\right| = 0$ for all $f \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. Meanwhile, $(i)$ implies that $\lim_{n\to\infty} \left|\mu^{\eta_n}_{\theta_n}(f) - \mu_{\theta_n}(f)\right| = 0$ for all $f \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. Therefore,

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

for all $f \in \mathcal{C}(\mathbb{S}\setminus\mathbb{C})$. By Portmanteau theorem for standard $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$-convergence, 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. $\square$

To facilitate the application of Theorem 2.2, 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 $\epsilon_n$ 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*** $\epsilon_n$ *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 3.2** (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 $\epsilon_n$. If $X_n$ is asymptotically equivalent to $Y_n$ when bounded away from $\mathbb{C}$ with respect to $\epsilon_n$, then $\epsilon_n^{-1}\mathbf{P}(Y_n \in \cdot) \to \mu(\cdot)$ in $\mathbb{M}(\mathbb{S}\setminus\mathbb{C})$.*

### 3.2 Technical Lemmas for Theorems 2.3 and 2.4

Our analysis hinges on the concepts 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 $i^{\text{th}}$ arrival time of "large noises" and its size as

$$\tau_i^{>\delta}(\eta) \triangleq \min\{n > \tau_{i-1}^{>\delta}(\eta) :\ \eta|Z_n| > \delta\}, \quad \tau_0^{>\delta}(\eta) = 0 \tag{3.2}$$

$$W_i^{>\delta}(\eta) \triangleq Z_{\tau_i^{>\delta}(\eta)}. \tag{3.3}$$

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

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

Recall the definition of filtration $\mathbb{F} = (\mathcal{F}_j)_{j\geq 0}$ where $\mathcal{F}_j$ is the $\sigma$-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 $\mathbf{\Gamma}_M$ denote the collection of families of random variables, over which we will prove the uniform asymptotics:

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

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

**Lemma 3.3.** *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*

$$\lim_{\eta\downarrow 0} \eta^{-N} \sup_{(V_i)_{i\geq 0}\in\mathbf{\Gamma}_M} \mathbf{P}\left(\max_{j \leq \lfloor t/\eta \rfloor \wedge \left(\tau_1^{>\delta}(\eta) - 1\right)} \eta\Big|\sum_{i=1}^{j} V_{i-1}Z_i\Big| > \epsilon\right) = 0 \qquad \forall \delta \in (0, \delta_0).$$

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

$$A_i(\eta, b, \epsilon, \delta, x) \triangleq \Big\{\max_{j\in I_i(\eta,\delta)} \eta\Big|\sum_{n=\tau_{i-1}^{>\delta}(\eta)+1}^{j} \sigma\big(X_{n-1}^{\eta|b}(x)\big)Z_n\Big| \leq \epsilon\Big\}; \tag{3.6}$$

$$I_i(\eta, \delta) \triangleq \big\{j \in \mathbb{N} :\ \tau_{i-1}^{>\delta}(\eta) + 1 \leq j \leq \big(\tau_i^{>\delta}(\eta) - 1\big) \wedge \lfloor 1/\eta \rfloor\big\}. \tag{3.7}$$

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

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

*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 $\beta$ 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 $\delta_0$ 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 $\eta \in (0, \eta_0)$,

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

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,

$$\begin{aligned}
&\mathbf{E}\left[\left(\max_{j\leq \lfloor t/\eta \rfloor} \eta\Big|\sum_{i=1}^{j} V_{i-1}\widehat{Z}^{(1)}_i\Big|\right)^p\right] \\
&\leq c_1\mathbf{E}\left[\left(\sum_{i=1}^{\lfloor t/\eta \rfloor}\left(\eta V_{i-1}\widehat{Z}^{(1)}_i\right)^2\right)^{p/2}\right] \leq c_1 M^p\mathbf{E}\left[\left(\sum_{i=1}^{\lfloor t/\eta \rfloor}\left(\eta\widehat{Z}^{(1)}_i\right)^2\right)^{p/2}\right] \\
&\leq c_1c_2M^p\mathbf{E}\left[\left(\max_{j\leq \lfloor t/\eta \rfloor}\Big|\sum_{i=1}^{j}\eta\widehat{Z}^{(1)}_i\Big|\right)^p\right] \leq c_1c_2\left(\frac{p}{p-1}\right)^p M^p\mathbf{E}\left[\left|\sum_{i=1}^{\lfloor t/\eta \rfloor}\eta\widehat{Z}^{(1)}_i\right|^p\right]
\end{aligned} \tag{3.13}$$

for some $c_1, c_2 > 0$ that only depend on $p$ and won't vary with $(V_i)_{i\geq 0}$ and $\eta$. 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 $s \in [0, t], y \geq 1$

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

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 $(a+b)^2 \leq 2a^2 + 2b^2$,

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

Also, it has been shown earlier that $\mathbf{E}|Z^{(1)}_i| \in \mathcal{RV}_{(\alpha-1)\beta}(\eta)$, and hence $\big[\mathbf{E}|Z^{(1)}_i|\big]^2 \in \mathcal{RV}_{2(\alpha-1)\beta}(\eta)$. 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 $\frac{t}{\eta}\cdot\eta^2\cdot 2\big[\mathbf{E}|Z^{(1)}_i|\big]^2 < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$ 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 $\beta$ in such a way that $\alpha\beta-1>0$. Fix a constant $J=\lceil\frac{N}{\alpha\beta-1}\rceil+1$, and define $I(\eta)\triangleq\#\big\{i\leq\lfloor t/\eta\rfloor:\ Z^{(2)}_i\neq 0\big\}$. 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 $\big(\tau^{>\delta}_i(\eta)\big)_{j\geq 1}$,

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

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

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

(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 $\big(\tau^{>\delta}_j(\eta)\big)_{j\geq 1}$ and $\big(W^{>\delta}_j(\eta)\big)_{j\geq 1}$ conditioned on event

$$E^\delta_{c,k}(\eta)\triangleq\Big\{\tau^{>\delta}_k(\eta)\leq\lfloor 1/\eta\rfloor<\tau^{>\delta}_{k+1}(\eta);\ \eta|W^{>\delta}_j(\eta)|>c\ \ \forall j\in[k]\Big\}.\tag{3.16}$$

The intuition is that, on event $E^{\delta}_{c,k}(\eta)$, 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 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 3.4.** *Let Assumption 1 hold. For any $\delta > 0, c \geq \delta$ and $k \in \mathbb{Z}^+$,*

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

*and*

$$\begin{aligned}&\mathscr{L}\Big(\eta W^{>\delta}_1(\eta), \eta W^{>\delta}_2(\eta), \cdots, \eta W^{>\delta}_k(\eta), \eta\tau^{>\delta}_1(\eta), \eta\tau^{>\delta}_2(\eta), \cdots, \eta\tau^{>\delta}_k(\eta)\Big| E^{\delta}_{c,k}(\eta)\Big)\\ \Rightarrow&\mathscr{L}\Big(W^*_1(c), W^*_2(c), \cdots, W^*_k(c), U_{(1;k)}, U_{(2;k)}, \cdots, U_{(k;k)}\Big) \text{ as } \eta \downarrow 0.\end{aligned}$$

*Proof.* Note that $\big(\tau^{>\delta}_i(\eta)\big)_{i\geq 1}$ is independent of $\big(W^{>\delta}_i(\eta)\big)_{i\geq 1}$. Therefore, $\mathbf{P}\big(E^{\delta}_{c,k}(\eta)\big) = \mathbf{P}\big(\tau^{>\delta}_k(\eta) \leq \lfloor 1/\eta \rfloor < \tau^{>\delta}_{k+1}(\eta)\big) \cdot \big(\mathbf{P}(\eta\big|W^{>\delta}_1(\eta)\big| > c)\big)^k$. Recall that $H(x) = \mathbf{P}(|Z_1| > x)$. Observe that

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

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

$$\mathbf{P}\big(\eta\big|W^{>\delta}_1(\eta)\big| > c\big) = H(c/\eta)\Big/H(\delta/\eta),$$

and hence,

$$\lim_{\eta\downarrow 0} \frac{q_3(\eta) \cdot \Big(\mathbf{P}\big(\eta\big|W^{>\delta}_1(\eta)\big| > c\big)\Big)^k}{\big(H(1/\eta)\big)^k} = \lim_{\eta\downarrow 0} \frac{\big(H(\delta/\eta)\big)^k \cdot \Big(H(c/\eta)\Big/H(\delta/\eta)\Big)^k}{\big(H(1/\eta)\big)^k} = \lim_{\eta\downarrow 0} \frac{\big(H(c/\eta)\big)^k}{\big(H(1/\eta)\big)^k} = 1/c^{\alpha k} \tag{3.20}$$

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

$$\lim_{\eta\downarrow 0} \frac{\mathbf{P}\big(E^{\delta}_{c,k}(\eta)\big)}{\lambda^k(\eta)} = \frac{q_1(\eta) \cdot q_2(\eta) \cdot q_3(\eta) \cdot \Big(\mathbf{P}\big(\eta\big|W^{>\delta}_1(\eta)\big| > c\big)\Big)^k}{1/\eta^k\big(H(1/\eta)\big)^k} = \frac{1/c^{\alpha k}}{k!}.$$

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. □

Next, we present several results about mappings $h^{(k)}_{[0,T]}$ defined in (2.6)–(2.8) and $h^{(k)|b}_{[0,T]}$ defined in (2.16)–(2.18). These results will serve as crucial tools when establishing Theorems 2.3 and 2.4. As the proofs rely on similar arguments and calculations, independent of the arguments in other sections of our analyses, we collect the proofs of Lemmas 3.5–3.8 in Section B.

Recall the definitions of the sets $\mathbb{D}^{(k)}_A$ and $\mathbb{D}^{(k)|b}_A$ in (2.12) and (2.19), respectively, which are the images of mappings $h^{(k)}$ and $h^{(k)|b}$. The next two results reveal useful properties of $\mathbb{D}^{(k)}_A$ and $\mathbb{D}^{(k)|b}_A$ when Assumptions 2 and 4 hold.

**Lemma 3.5.** *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 $\mathbb{D}^{(k-1)}_A$, 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) $\boldsymbol{d}_{J_1}(B^{3\bar\epsilon}, \mathbb{D}^{(k-1)}_A) > 0$.*

As an intermediate step of the proof, 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 $\inf_{x \in \mathbb{R}} \sigma(x) \geq c$.*

Now, we state a result for $\mathbb{D}^{(k)|b}_A$ that is analogous to Lemma 3.5.

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

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

Lemma 3.7 reveals that the image of $h^{(1)}$ (resp. $h^{(1)|b}$) provides good approximations of the sample path of $X^\eta_j$ (resp. $X^{\eta|b}_j$) up until $\tau^{>\delta}_1(\eta)$, i.e. the arrival time of the first "large noise"; see (3.2),(3.3) for the definition of $\tau^{>\delta}_i(\eta), W^{>\delta}_i(\eta)$.

**Lemma 3.7.** *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*

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

*that*

$$\sup_{t \in [0,1]:\ t < \eta \tau_1^{>\delta}(\eta)} \big| \xi_t - X^\eta_{\lfloor t/\eta \rfloor}(x) \big| \le \rho \cdot \big(\epsilon + |x-y| + \eta C\big), \tag{3.21}$$

*where*

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

(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*

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

*that*

$$\sup_{t \in [0,1]:\ t < \eta \tau_1^{>\delta}(\eta)} \big| \xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x) \big| \le \rho \cdot \big(\epsilon + |x-y| + \eta C\big), \tag{3.22}$$

$$\sup_{t \in [0,1]:\ t \le \eta \tau_1^{>\delta}(\eta)} \big| \xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x) \big| \le \rho \cdot \Big(1 + \frac{bD}{c}\Big)\big(\epsilon + |x-y| + 2\eta C\big) \tag{3.23}$$

*where*

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

By applying Lemma 3.7 inductively, the next result illustrates how the image of the mapping $h^{(k)|b}$ approximates the path of $X^{\eta|b}_j(x)$.

**Lemma 3.8.** *Let Assumptions 2, 4, and 8 hold. Let $A_i(\eta, b, \epsilon, \delta, x)$ be defined as in (3.6). For any $k \ge 0$, $x \in \mathbb{R}$, $\epsilon, b > 0$, $\delta \in (0, \frac{b}{2C})$, and $\eta \in (0, \frac{b \wedge \epsilon}{2C})$, it holds on event $\big(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\big) \cap \Big\{\tau_k^{>\delta}(\eta) \le \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\Big\}$ that*

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

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

To conclude, Lemma 3.9 provides tools for verifying 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 3.9.** *Let $T > 0$ and $k \geq 1$. Let $A \subseteq \mathbb{R}$ be compact.*

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

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

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

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

*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 B.4), 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 3.5 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 B.3 (resp. Lemma 3.6) instead of Lemma B.4 (resp. Lemma 3.5) so we omit the details. □

### 3.3 Proofs of Theorems 2.3 and 2.4

In the proofs of Theorems 2.3 and 2.4 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 2.3 and 2.4 is stronger than the standard $\mathbb{M}$-convergence introduced in [35]. Nevertheless, under the conditions provided in Theorem 2.3 or 2.4 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 2.3 and 2.4 hinge on the following key result for $\boldsymbol{X}^{\eta|b}$.

**Proposition 3.10.** *Let $\eta_n$ 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*

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

As the first application of Proposition 3.10, we prepare a similar result for the unclipped dynamics $\boldsymbol{X}^\eta$ defined in (2.11).

**Proposition 3.11.** *Let $\eta_n$ 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*

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

*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 [35]), 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 $\delta, b > 0$,

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

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 $X^{\eta|b}_j$ 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

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

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

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

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

$$\lim_{n\to\infty} p_1(n, \delta)/\lambda^k(\eta_n) = 0, \tag{3.24}$$

$$\lim_{n\to\infty} p_2(n,\delta)/\lambda^k(\eta_n) = 0. \tag{3.25}$$

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

$$\limsup_{n\to\infty} p_3(n,b,\delta)/\lambda^k(\eta_n) \leq \psi_\delta(b) \triangleq \frac{k}{\delta^{\alpha k}} \cdot \big(\frac{\delta}{2C}\big)^\alpha \cdot \frac{1}{b^\alpha}. \tag{3.26}$$

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

$$\lim_{b\to\infty} \mathbf{C}^{(k)|b}(g;x^*) = \mathbf{C}^{(k)}(g;x^*). \tag{3.27}$$

Then by combining (3.24)–(3.26) 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 3.10. Letting $b$ tend to $\infty$ and applying the limit (3.27), we conclude the proof. Now it only remains to prove (3.24) (3.25) (3.26) (3.27).

**Proof of Claim** (3.24):

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 (3.24) 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** (3.25):

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

$$\Big\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor;\ \boldsymbol{X}^\eta(x) \in B\Big\} \cap \big(\bigcap_{i=1}^{k+1} A_i(\eta,\infty,\epsilon,\delta,x)\big) = \emptyset \qquad \forall x \in A,\ \ \delta > 0,\ \ \eta \in (0,\frac{\epsilon}{C\rho}) \tag{3.28}$$

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

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

Applying Lemma 3.3 (b) with some $N > k(\alpha-1)$, we conclude that claim (3.25) holds for all $\delta > 0$ small enough. Now it only remains to find $\epsilon > 0$ that satisfies condition (3.28). 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 3.5 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 (3.28) holds for any $\epsilon > 0$ small enough with $(\rho+1)\epsilon < \bar{\epsilon}$. To see why, we fix such $\epsilon$ 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 $\breve{X}_0^{\eta,\delta}(x) = x$)

$$\frac{d\breve{X}_t^{\eta,\delta}(x)}{dt} = a\big(\breve{X}_t^{\eta,\delta}(x)\big) \qquad \forall t \geq 0,\ t \notin \{\eta\tau_j^{>\delta}(\eta):\ j \geq 1\},$$

$$\breve{X}^{\eta,\delta}_{\eta\tau_i^{>\delta}(\eta)}(x) = X^{\eta}_{\tau_i^{>\delta}(\eta)}(x) \qquad \forall j \geq 1.$$

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

$$\begin{aligned}
&\boldsymbol{d}_{J_1}\big(\breve{\boldsymbol{X}}^{\eta,\delta}(x), \boldsymbol{X}^{\eta}(x)\big) \\
&\leq \sup_{t \in \left[0,\eta\tau_1^{>\delta}(\eta)\right) \cup \left[\eta\tau_1^{>\delta}(\eta),\eta\tau_2^{>\delta}(\eta)\right) \cup \cdots \cup \left[\eta\tau_k^{>\delta}(\eta),\eta\tau_{k+1}^{>\delta}(\eta)\right)} \left|\breve{X}^{\eta,\delta}_t(x) - X^{\eta}_{\lfloor t/\eta \rfloor}(x)\right| \\
&\leq \rho\cdot\big(\epsilon + \eta C\big) \leq \rho\epsilon + \epsilon < \bar{\epsilon} \qquad \text{because of (3.21) of Lemma 3.7.}
\end{aligned}$$

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 (3.28).

**Proof of Claim** (3.26):

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** (3.27):

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

$$\lim_{b\to\infty} \mathbf{C}^{(k)|b}(S;x^*) = \mathbf{C}^{(k)}(S;x^*). \tag{3.29}$$

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 (3.29), we get $\limsup_{b\to\infty}\left|\mathbf{C}^{(k)|b}(g^{\Delta};x^*) - \mathbf{C}^{(k)}(g^{\Delta};x^*)\right| = 0$. Next, note that $\left|\mathbf{C}^{(k)|b}(g^{\Delta};x^*) - \mathbf{C}^{(k)|b}(g;x^*)\right| \leq \Delta\cdot\mathbf{C}^{(k)|b}(B;x^*)$ and $\left|\mathbf{C}^{(k)}(g^{\Delta};x^*) - \mathbf{C}^{(k)}(g;x^*)\right| \leq \Delta\cdot\mathbf{C}^{(k)}(B;x^*)$. Thanks to (3.29) again, we get $\limsup_{b\to\infty}\left|\mathbf{C}^{(k)|b}(g;x^*) - \mathbf{C}^{(k)}(g;x^*)\right| \leq 2\Delta\cdot\mathbf{C}^{(k)}(B;x^*)$. The arbitrariness of $\Delta > 0$ allows us to conclude the proof of (3.26).

We prove (3.29) by applying Dominated Convergence theorem. From the definition in (2.20),

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

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 (3.29), 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

$$\mathbb{I}\Big\{h^{(k)|b}\big(x^*,\boldsymbol{w},\boldsymbol{t}\big)\in S\Big\} \leq \mathbb{I}\Big\{|w_j| > \bar\delta\ \forall j\in[k]\Big\} \qquad \forall b>0,\ \boldsymbol{w}\in\mathbb{R}^k,\ \boldsymbol{t}\in(0,1)^{k\uparrow}. \tag{3.30}$$

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 (3.30). 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 (2.22))

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

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 (3.30). □

With Proposition 3.11 in our arsenal, we prove Theorem 2.3.

*Proof of Theorem 2.3.* 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 $\left|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x_n)\right| > \epsilon\ \forall n \geq 1$ 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 3.11 and yield $\lim_{n\to\infty}\left|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f;x^*)\right| = 0$. On the other hand, using part (a) of Lemma 3.9, we get $\lim_{n\to\infty}\left|\mathbf{C}^{(k)}(f;x_n) - \mathbf{C}^{(k)}(f;x^*)\right| = 0$. Therefore, we arrive at the contradiction

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

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

Next, we prove the uniform sample-path large deviations stated in (2.13). Part (a) of Lemma 3.9 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 2.2), most claims follow directly from Theorem 2.3 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 3.5 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 (2.10),

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

This concludes the proof. □

Similarly, building upon Proposition 3.10, we provide the proof to Theorem 2.4.

*Proof of Theorem 2.4.* The proof-by-contradiction approach in Theorem 2.3 can be applied here to establish the uniform $\mathbb{M}$-convergence. The only difference is that we apply Proposition 3.10 (resp., part (b) of Lemma 3.9) instead of Proposition 3.11 (resp., part (a) of Lemma 3.9). Similarly, the proof to the uniform sample-path large deviations stated in (2.21) is almost identical to that of (2.13) in Theorem 2.3. In particular, the only differences are that we apply part (b) of Lemma 3.9 (resp., Lemma 3.6) instead of part (a) of Lemma 3.9 (resp., Lemma 3.5). To avoid repetition we omit the details. □

### 3.3.1 Proof of Proposition 3.10

To facilitate the analysis, we consider the following "truncated" version of functions $a(\cdot), \sigma(\cdot)$. For any $M \geq 1$,

$$a_M(x) \triangleq \begin{cases} a(M) & \text{if } x > M, \\ a(-M) & \text{if } x < -M, \\ a(x) & \text{otherwise.}\end{cases} \qquad \sigma_M(x) \triangleq \begin{cases} \sigma(M) & \text{if } x > M, \\ \sigma(-M) & \text{if } x < -M, \\ \sigma(x) & \text{otherwise.}\end{cases} \tag{3.31}$$

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 (2.16)-(2.18). We also consider its "truncated" counterpart by defining the mapping $h_{M\downarrow}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0,1]^{k\uparrow} \to \mathbb{D}$ 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

$$\xi_0 = x_0; \tag{3.32}$$

$$\frac{d\xi_t}{dt} = a_M(\xi_t) \qquad \forall t \in [0,1],\ t \neq t_1, t_2, \cdots, t_k; \tag{3.33}$$

$$\xi_t = \xi_{t-} + \varphi_b\big(\sigma_M(\xi_{t-})w_j\big) \qquad \text{if } t = t_j \text{ for some } j \in [k]. \tag{3.34}$$

Also, we let

$$\mathbb{D}^{(k)|b}_{A;M\downarrow} \triangleq h^{(k)|b}_{M\downarrow}\big(\mathbb{R} \times \mathbb{R}^k \times (0,1]^{k\uparrow}\big). \tag{3.35}$$

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 $a(\cdot), \sigma(\cdot)$.

As has been demonstrated earlier, Proposition 3.10 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 3.10. 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 3.12.** *Let $\eta_n$ 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*

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

*Proof of Proposition 3.10.* Fix some $b > 0, k \geq 0$, as well as some $g \in \mathcal{C}\big(\mathbb{D} \setminus \mathbb{D}^{(k-1)|b}_A\big)$ that is also uniformly continuous on $\mathbb{D}$. Thanks to the Portmanteau theorem for $\mathbb{M}$-convergence (see theorem 2.1 of [35]), 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 $\mathbb{D}^{(k-1)|b}_A$. Applying Corollary B.2, we can fix some $M_0$ such that the following claim holds for any $M \geq M_0$ : for any $\xi = h^{(k)|b}_{M\downarrow}(x_0, \boldsymbol{w}, \boldsymbol{t})$ 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 $x_0 \in A$,

$$\xi = h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) = h^{(k)|b}_{M\downarrow}(x_0, \boldsymbol{w}, \boldsymbol{t}); \quad \sup_{t\in[0,1]} |\xi(t)| \leq M_0. \tag{3.36}$$

Here the mapping $h^{(k)|b}_{M\downarrow}$ is defined in (3.32)–(3.34). Now fix some $M \geq M_0 + 1$ and recall the definitions of $a_M, \sigma_M$ in (3.31). Also, define stochastic processes $\widetilde{\boldsymbol{X}}^{\eta|b}(x) \triangleq \big\{\widetilde{X}^{\eta|b}_{\lfloor t/\eta\rfloor}(x) :\ t \in [0,1]\big\}$ as

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

under initial condition $\widetilde{X}^{\eta|b}_0(x) = x$. In particular, by comparing the definition of $\widetilde{X}^{\eta|b}_j(x)$ with that of $X^{\eta|b}_j(x)$ in (2.14), we must have (for any $x \in \mathbb{R}, \eta > 0$)

$$\sup_{t\in[0,1]} \big|\widetilde{X}^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\big| > M \iff \sup_{t\in[0,1]} \big|X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\big| > M, \tag{3.37}$$

$$\sup_{t\in[0,1]} \big|X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\big| \leq M \implies \boldsymbol{X}^{\eta|b}(x) = \widetilde{\boldsymbol{X}}^{\eta|b}(x). \tag{3.38}$$

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

$$\begin{aligned}\mathbf{E}\big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\big] = {} & \mathbf{E}\Big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\mathbb{I}\Big\{\boldsymbol{X}^{\eta_n|b}(x_n) \in B;\ \sup_{t\in[0,1]}\big|X^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big| \leq M\Big\}\Big] \\ & + \mathbf{E}\Big[g\big(\boldsymbol{X}^{\eta_n|b}(x_n)\big)\mathbb{I}\Big\{\boldsymbol{X}^{\eta_n|b}(x_n) \in B;\ \sup_{t\in[0,1]}\big|X^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big| > M\Big\}\Big].\end{aligned} \tag{3.39}$$

An upper bound then follows immediately from (3.37) and (3.38):

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

Similarly, by bounding the first term on the R.H.S. of (3.39) using (3.37) and (3.38), 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

$$\lim_{n\to\infty}\lambda^{-k}(\eta_n)\mathbf{E}\big[g\big(\widetilde{\boldsymbol{X}}^{\eta_n|b}(x_n)\big)\big] = \mathbf{C}^{(k)|b}(g;x^*), \tag{3.40}$$

$$\lim_{n\to\infty}\lambda^{-k}(\eta_n)\mathbf{P}\Big(\sup_{t\in[0,1]}\big|\widetilde{X}^{\eta_n|b}_{\lfloor t/\eta\rfloor}(x_n)\big| > M\Big) = 0. \tag{3.41}$$

**Proof of Claim** (3.40):

Under Assumption 3, one can easily see that $a_M, \sigma_M$ would satisfy Assumption 4 and 8. This allows us to apply Proposition 3.12 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

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

Given (3.36) 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 (3.40).

**Proof of Claim** (3.41):

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 3.12 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 (3.41). To see why $E$ is bounded away from $\mathbb{D}^{(k)|b}_A$, note that it follows directly from (3.36) 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 3.12. In light of Lemma 3.2, a natural approach to the $\mathbb{M}$-convergence claim in Proposition 3.12 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

$$\mathcal{J}_Z(c,n) \triangleq \#\{i \in [n]:\ |Z_i| \geq c\}\ \ \forall c \geq 0; \quad \boldsymbol{Z}^{(j)}(\eta) \triangleq \max\Big\{c \geq 0:\ \mathcal{J}_Z(c,\lfloor 1/\eta\rfloor) \geq j\Big\}. \tag{3.42}$$

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 $j^{\text{th}}$ largest element in $\{|Z_i| : i \leq \lfloor 1/\eta \rfloor\}$. Moreover, let

$$\tau_i^{(j)}(\eta) \triangleq \min\left\{k > \tau_{i-1}^{(j)}(\eta) : \ |Z_k| \geq \boldsymbol{Z}^{(j)}(\eta)\right\}, \quad W_i^{(j)}(\eta) \triangleq Z_{\tau_i^{(j)}(\eta)} \ \ \forall i = 1,2,\cdots,j \tag{3.43}$$

with the convention that $\tau_0^{(j)}(\eta) = 0$. Note that $\big(\tau_i^{(j)}(\eta), W_i^{(j)}(\eta)\big)_{i\in[j]}$ 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

$$\frac{d\hat{X}_t^{\eta|b;(j)}(x)}{dt} = a\big(\hat{X}_t^{\eta|b;(j)}(x)\big) \quad \forall t \in [0,1], \ t \notin \big\{\eta\tau_i^{(j)}(\eta) : \ i \in [j]\big\}, \tag{3.44}$$

$$\hat{X}_t^{\eta|b;(j)}(x) = \hat{X}_{t-}^{\eta|b;(j)}(x) + \varphi_b\Big(\eta\sigma\big(\hat{X}_{t-}^{\eta|b;(j)}(x)\big)W_i^{(j)}(\eta)\Big) \quad \text{if } t = \eta\tau_i^{(j)}(\eta) \text{ for some } i \in [j]. \tag{3.45}$$

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

$$\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta) \qquad \Longrightarrow \qquad \hat{\boldsymbol{X}}^{\eta|b;(k)}(x) = h^{(k)|b}\big(x, \eta\boldsymbol{W}^{>\delta}(\eta), \eta\boldsymbol{\tau}^{>\delta}(\eta)\big) \tag{3.46}$$

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 3.2, 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 3.13.** *Let $\eta_n$ 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 3.14.** *Let $\eta_n$ 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*

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

*where the measure $\mathbf{C}^{(k)|b}$ is defined in (2.20).*

*Proof of Proposition 3.12.* In light of Lemma 3.2, it is a direct corollary of Propositions 3.13 and 3.14. □

Now it only remains to prove Propositions 3.13 and 3.14.

*Proof of Proposition 3.13.* Fix some $b > 0, k \geq 0$, and some sequence of strictly positive real numbers $\eta_n$ 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

$$\lim_{n\to\infty} \mathbf{P}\Big(\boldsymbol{d}_{J_1}\big(\boldsymbol{X}^{\eta_n|b}(x_n), \hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n)\big)\mathbb{I}\big(\boldsymbol{X}^{\eta_n|b}(x_n) \text{ or } \hat{\boldsymbol{X}}^{\eta_n|b;(k)}(x_n) \in B\big) > \Delta\Big)\Big/\lambda^k(\eta_n) = 0. \tag{3.47}$$

Applying Lemma 3.6, 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 $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$,

$$h^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \text{ or } h^{(k)|b+\bar{\epsilon}}(x, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \Longrightarrow |w_i| > 3C\bar{\delta}/c \ \forall i \in [k]. \tag{3.48}$$

$$\boldsymbol{d}_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}) > \bar{\epsilon} \tag{3.49}$$

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

$$B_0 \triangleq \{\boldsymbol{X}^{\eta|b}(x) \in B \text{ or } \hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \in B;\ \boldsymbol{d}_{J_1}\big(\boldsymbol{X}^{\eta|b}(x), \hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\big) > \Delta\},$$
$$B_1 \triangleq \{\tau_{k+1}^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\},$$
$$B_2 \triangleq \{\tau_{k}^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor\},$$
$$B_3 \triangleq \big\{\eta|W_i^{>\delta}(\eta)| > \bar{\delta} \text{ for all } i \in [k]\big\}.$$

Note that

$$B_0 = (B_0 \cap B_1^c) \cup (B_0 \cap B_1 \cap B_2^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3). \tag{3.50}$$

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

$$\limsup_{\eta \downarrow 0} \sup_{x \in A} \mathbf{P}\Big(B_0 \cap B_1^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.51}$$

$$\limsup_{\eta \downarrow 0} \sup_{x \in A} \mathbf{P}\Big(B_0 \cap B_1 \cap B_2^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.52}$$

$$\limsup_{\eta \downarrow 0} \sup_{x \in A} \mathbf{P}\Big(B_0 \cap B_1 \cap B_2 \cap B_3^c\Big)\Big/\lambda^k(\eta) = 0, \tag{3.53}$$

$$\limsup_{\eta \downarrow 0} \sup_{x \in A} \mathbf{P}\Big(B_0 \cap B_1 \cap B_2 \cap B_3\Big)\Big/\lambda^k(\eta) = 0 \tag{3.54}$$

if we pick $\delta > 0$ sufficiently small. Now fix such $\delta$. Combining these claims with the decomposition of event $B_0$ in (3.50), we establish (3.47). Now we conclude the proof of this proposition with the proofs of claims (3.51)–(3.54).

**Proof of** (3.51):

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** (3.52):

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

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

In particular, we focus on $\delta \in (0, \bar{\delta} \wedge \frac{b}{2})$ with $\bar{\delta}$ characterized in (3.48). 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 $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ 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 (3.42), we then have that $\min_{i \in [k]} \eta|W_i^{(k)}(\eta)| \leq \delta$. In light of (3.48), we yield $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \notin B^{\bar{\epsilon}}$ on $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$, and hence

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

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

$$\{\boldsymbol{X}^{\eta|b}(x)\in B\}\cap\{\tau_k^{>\delta}(\eta)>\lfloor 1/\eta\rfloor\}\cap\big(\cap_{i=1}^k A_i(\eta,b,\epsilon,\delta,x)\big)=\emptyset \tag{3.55}$$

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 3.3 (b) with some $N>k(\alpha-1)$.

Now it only remains to prove claim (3.55). To proceed, let process $\breve{X}_t^{\eta|b;\delta}(x)$ be the solution to

$$\frac{d\breve{X}_t^{\eta|b;\delta}(x)}{dt}=a\big(\breve{X}_t^{\eta|b;\delta}(x)\big)\qquad\forall t\in[0,\infty)\setminus\{\eta\tau_j^{>\delta}(\eta):\ j\geq 1\}, \tag{3.56}$$

$$\breve{X}^{\eta|b;\delta}_{\eta\tau_j^{>\delta}(\eta)}(x)=X^{\eta|b}_{\tau_j^{>\delta}(\eta)}(x)\qquad\forall j\geq 1 \tag{3.57}$$

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

$$\begin{aligned}
&\boldsymbol{d}_{J_1}(\breve{\boldsymbol{X}}^{\eta|b;\delta}(x),\boldsymbol{X}^{\eta|b}(x))\\
&\leq\sup_{t\in\big[0,\eta\tau_1^{>\delta}(\eta)\big)\cup\big[\eta\tau_1^{>\delta}(\eta),\eta\tau_2^{>\delta}(\eta)\big)\cup\cdots\cup\big[\eta\tau_{j-1}^{>\delta}(\eta),\eta\tau_j^{>\delta}(\eta)\big)}\Big|\breve{X}_t^{\eta|b;\delta}(x)-X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\Big|\\
&\leq\rho\cdot\big(\epsilon+\eta C\big)\leq 2\rho\epsilon<\bar{\epsilon}\qquad\text{due to (3.22) of Lemma 3.7.}
\end{aligned}\tag{3.58}$$

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

$$\begin{aligned}
\Big|\Delta\breve{X}^{\eta|b;\delta}_{\eta\tau_j^{>\delta}(\eta)}(x)\Big|&=\Big|\breve{X}^{\eta|b;\delta}_{\eta\tau_j^{>\delta}(\eta)-}(x)-X^{\eta|b}_{\tau_j^{>\delta}(\eta)}(x)\Big|\qquad\text{see (3.57)}\\
&\leq\Big|\breve{X}^{\eta|b;\delta}_{\eta\tau_j^{>\delta}(\eta)-}(x)-X^{\eta|b}_{\tau_j^{>\delta}(\eta)-1}(x)\Big|+\Big|X^{\eta|b}_{\tau_j^{>\delta}(\eta)-1}(x)-X^{\eta|b}_{\tau_j^{>\delta}(\eta)}(x)\Big|\\
&<\bar{\epsilon}+b.
\end{aligned}\tag{3.59}$$

As a result, on event $\big(\cap_{i=1}^k A_i(\eta,b,\epsilon,\delta,x)\big)\cap\{\tau_k^{>\delta}(\eta)>\lfloor 1/\eta\rfloor\}$, we have $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x)\in\mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}$. Considering the facts that $\mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}$ is bounded away from $B^{\bar{\epsilon}}$ (see (3.49)) as well as $\boldsymbol{d}_{J_1}(\breve{\boldsymbol{X}}^{\eta|b;\delta}(x),\boldsymbol{X}^{\eta|b}(x))<\bar{\epsilon}$ shown in (3.58), we have just established that $\boldsymbol{X}^{\eta|b}(x)\notin B$, thus establishing (3.55).

**Proof of** (3.53):

On event $B_1\cap B_2=\{\tau_k^{>\delta}(\eta)\leq\lfloor 1/\eta\rfloor<\tau_{k+1}^{>\delta}(\eta)\}$, it follows from (3.46) that $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)=h^{(k)|b}\big(x,\eta W_1^{>\delta}(\eta),\cdots,\eta W_k^{>\delta}(\eta),\eta\tau_1^{>\delta}(\eta),\cdots,\eta\tau_k^{>\delta}(\eta)\big)$. Furthermore, on $B_3^c$, there is some $i\in[k]$ with $|\eta W_i^{>\delta}(\eta)|\leq\bar{\delta}$. Considering the choice of $\bar{\delta}$ in (3.48), on event $B_1\cap B_2\cap B_3^c$ we have $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x)\notin B$, and hence

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

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

$$\begin{aligned}
&\{\boldsymbol{X}^{\eta|b}(x)\in B\}\cap\big\{\tau_k^{>\delta}(\eta)\leq\lfloor 1/\eta\rfloor<\tau_{k+1}^{>\delta}(\eta);\ \eta|W_i^{>\delta}(\eta)|\leq\bar{\delta}\text{ for some }i\in[k]\big\}\\
&\qquad\cap\big(\bigcap_{i=1}^{k+1}A_i(\eta,b,\epsilon,\delta,x)\big)=\emptyset.
\end{aligned}\tag{3.60}$$

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

We are left with proving the claim (3.60). 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 (3.56)-(3.57), and note that it has been shown in (3.58) (with $j = k+1$) that

$$\sup_{t \in [0,1]} \Big| \breve{X}_t^{\eta|b;\delta}(x) - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x) \Big| < 2\rho\epsilon < \bar{\epsilon}. \tag{3.61}$$

If we can show that $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$, then (3.61) immediately leads to $\boldsymbol{X}^{\eta|b}(x) \notin B$, thus proving claim (3.60). To proceed, first note that

$$\begin{aligned}
\Big|\Delta \breve{X}_{\eta\tau_J^{>\delta}(\eta)}^{\eta|b;\delta}(x)\Big| &\leq \Big|\breve{X}_{\eta\tau_J^{>\delta}(\eta)-}^{\eta|b;\delta}(x) - X_{\tau_J^{>\delta}(\eta)-1}^{\eta|b}(x)\Big| + \Big|X_{\tau_J^{>\delta}(\eta)-1}^{\eta|b}(x) - X_{\tau_J^{>\delta}(\eta)}^{\eta|b}(x)\Big| \qquad \text{see (3.57)} \\
&\leq 2\rho\epsilon + \eta\Big|a\big(X_{\tau_J^{>\delta}(\eta)-1}^{\eta|b}(x)\big) + \sigma\big(X_{\tau_J^{>\delta}(\eta)-1}^{\eta|b}(x)\big)W_J^{>\delta}(\eta)\Big| \qquad \text{using (3.61)} \\
&\leq 2\rho\epsilon + \eta C + C\bar{\delta} < 3C\bar{\delta} \qquad \text{due to } 2\rho\epsilon < \bar{\delta},\ \eta < \bar{\delta}, \text{ and } C \geq 1.
\end{aligned}$$

Meanwhile, the calculations in (3.59) 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),

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

which implies $|\widetilde{w}_J| < 3C\bar{\delta}/c$. In light of our choice of $\bar{\delta}$ in (3.48), we yield $\breve{\boldsymbol{X}}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$ and conclude the proof.

**Proof of** (3.54):

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 (3.46). This allows us to apply Lemma 3.8 and show that, for any $x \in A$ and any $\eta \in (0, \frac{\epsilon \wedge b}{2C})$, the inequality

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

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 3.3 (b) with some $N > k(\alpha - 1)$, we conclude the proof. $\square$

In order to prove Proposition 3.14, we first prepare a lemma regarding a weak convergence claim on event $E_{c,k}^{\delta}(\eta) \triangleq \Big\{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta);\ \eta|W_j^{>\delta}(\eta)| > c\ \ \forall j \in [k]\Big\}$ defined in (3.16).

**Lemma 3.15.** *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 $k = 0,1,2,\cdots$,*

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

*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

$$\left| \lambda^{-k}(\eta_n)\mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\mathbb{I}_{E^\delta_{c,k}(\eta_n)}\Big] - (1/k!) \cdot c^{-\alpha k} \cdot \phi_{c,k}(x_n) \right| > \epsilon \quad \forall n \geq 1 \tag{3.62}$$

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

$$\begin{aligned}
&\lim_{n\to\infty} \lambda^{-k}(\eta_n)\mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\mathbb{I}_{E^\delta_{c,k}(\eta_n)}\Big] \\
&= \Big[\lim_{n\to\infty} \lambda^{-k}(\eta_n)\mathbf{P}\big(E^\delta_{c,k}(\eta_n)\big)\Big] \cdot \lim_{n\to\infty} \mathbf{E}\Big[\Phi\big(x_n, \boldsymbol{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}\big)\Big| E^\delta_{c,k}(\eta_n)\Big] \\
&= (1/k!) \cdot c^{-\alpha k} \cdot \mathbf{E}\Big[\Phi\big(x^*, \boldsymbol{W}^*, \boldsymbol{U}^*\big)\Big] = (1/k!) \cdot c^{-\alpha k} \cdot \phi_{c,k}(x^*) \qquad \text{due to Lemma 3.4}
\end{aligned}$$

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 (3.62) and allows us to conclude the proof. $\square$

We are now ready to prove Proposition 3.14.

*Proof of Proposition 3.14.* 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 3.6 we can fix some $\bar{\delta} > 0$ such that the following claim holds for any $x_0 \in A$ and any $\boldsymbol{t} = (t_1, \cdots, t_k) \in (0,1)^{k\uparrow}$, $\boldsymbol{w} = (w_1, \cdots, w_k) \in \mathbb{R}^k$:

$$h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in B^{\bar{\epsilon}} \Longrightarrow |w_j| > \bar{\delta} \quad \forall j \in [k]. \tag{3.63}$$

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 (3.42), we then have that $\min_{i\in[k]} \eta|W_i^{(k)}(\eta)| \leq \delta$. In light of (3.63) 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 (3.46), which implies $\hat{\boldsymbol{X}}^{\eta|b;(k)}(x) \notin B$. due to (3.63) 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 (3.46). 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 $\boldsymbol{W}^{>\delta}(\eta) = (W_1^{>\delta}(\eta), \cdots, W_k^{>\delta}(\eta))$, $\boldsymbol{\tau}^{>\delta}(\eta) = (\tau_1^{>\delta}(\eta), \cdots, \tau_k^{>\delta}(\eta))$, 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 $\Phi$ 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 B.3). Besides, $\|\Phi\| \leq \|g\| < \infty$. It then follows from the continuity of $\Phi$ and Bounded Convergence Theorem that $\phi$ 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 3.15. 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 (3.63), 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, $\mathbf{P}\Big(U_{(1;k)} \in dt_1, \cdots, U_{(k;k)} \in dt_k\Big) = k! \cdot \mathbb{I}\{0 < t_1 < t_2 < \cdots < t_k < 1\}\mathcal{L}_1^{k\uparrow}(dt_1, \cdots, dt_k)$ where $\mathcal{L}_1^{k\uparrow}$ is the Lebesgue measure restricted on $(0,1)^{k\uparrow}$. The conclusion of the proof then follows from

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

where we appealed to the definition in (2.20) in the last equality. □

# 4 First Exit Time Analysis

In this section, we collect the proofs for Section 2.3. Specifically, Section 4.1 develops the general framework for first exit analysis of Markov processes by establishing Theorem 2.8. Section 4.2 then applies the framework in the context of heavy-tailed stochastic difference equations and proves Theorem 2.6.

## 4.1 Proof of Theorem 2.8

Our proof of Theorem 2.8 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*

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

*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 $i^{\text{th}}$ asymptotic regeneration times after cooling period $rT/\eta$. We start with the following two observations: For any $0 < r < R$,

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

where the last equality is from (2.29) of Condition 1, and

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

where the second inequaility is from (4.1) and the last equality is from (2.28) 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 $K = \Big\lceil \frac{t/\gamma(\eta)}{T/\eta} \Big\rceil$. 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 $\mathcal{F}_{\rho^\eta_j(x)}$,

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

$$\geq \inf_{x\in A(\epsilon)} \sum_{j=K}^{\infty} \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq T/\eta;\ V^\eta_{\tau_\epsilon}(y) \in B\Big) \cdot \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big).$$

$$\geq \inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq T/\eta;\ V^\eta_{\tau_\epsilon}(y) \in B\Big) \cdot \sum_{j=K}^{\infty} \inf_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big). \tag{4.3}$$

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

$$\begin{aligned}
&\sum_{j=K}^{\infty} \inf_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big) \\
&\geq \sum_{j=0}^{\infty} \bigg(\inf_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) > \rho^\eta_1(y)\Big)\bigg)^{K+j} = \sum_{j=0}^{\infty} \bigg(1 - \sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)\bigg)^{K+j} \\
&= \frac{1}{\sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)} \cdot \bigg(1 - \sup_{y\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(y) \leq \rho^\eta_1(y)\Big)\bigg)^{\lceil \frac{t/\gamma(\eta)}{T/\eta} \rceil} \\
&\geq \frac{1}{\big(1+\delta_{\mathbb{S}}(\epsilon, RT) + o(1)\big)\cdot \gamma(\eta)RT/\eta} \cdot \bigg(1 - \big(1+\delta_{\mathbb{S}}(\epsilon, RT) + o(1)\big)\cdot \gamma(\eta)RT/\eta\bigg)^{\frac{t/\gamma(\eta)}{T/\eta}+1}.
\end{aligned} \tag{4.4}$$

where the last inequality is from (4.2) with $B = \mathbb{S}$. From (4.3), (4.4), and (2.27) 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

$$\begin{aligned}
\text{(II)} &\leq \sum_{j=1}^{k} \sup_{x\in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) \wedge \rho^\eta_j(x) - \rho^\eta_{j-1}(x) \geq T/\eta\big) \\
&= \sum_{j=1}^{k} \sup_{x\in A(\epsilon)} \mathbf{E}\Big[\mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(x) \wedge \rho^\eta_j(x) - \rho^\eta_{j-1}(x) \geq T/\eta \big| \mathcal{F}_{\rho^\eta_{j-1}(x)}\big)\Big] \\
&\leq \sum_{j=1}^{k} \sup_{y\in A(\epsilon)} \mathbf{P}\big(\tau^\eta_{I(\epsilon)^c}(y) \wedge \rho^\eta_1(y) \geq T/\eta\big) \\
&\leq \frac{t}{\gamma(\eta)T/\eta} \cdot \gamma(\eta)T/\eta \cdot o(1) = o(1)
\end{aligned}$$

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.

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

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 (2.27) of Condition 1 as follows:

$$\sum_{j=k}^{\infty} \sup_{x\in A(\epsilon)} \mathbf{P}\Big(\tau^\eta_{I(\epsilon)^c}(x) > \rho^\eta_j(x)\Big)$$

$$\leq \sum_{j=0}^{\infty}\bigg(\sup_{y\in A(\epsilon)}\mathbf{P}\Big(\tau^{\eta}_{I(\epsilon)^c}(y) > \lfloor rT/\eta\rfloor\Big)\bigg)^{k+j} = \sum_{j=0}^{\infty}\bigg(1 - \inf_{y\in A(\epsilon)}\mathbf{P}\Big(\tau^{\eta}_{I(\epsilon)^c}(y) \leq rT/\eta\Big)\bigg)^{k+j}$$

$$\leq \frac{1}{\inf_{y\in A(\epsilon)}\mathbf{P}\Big(\tau^{\eta}_{I(\epsilon)^c}(y) \leq rT/\eta\Big)} \cdot \bigg(1 - \inf_{y\in A(\epsilon)}\mathbf{P}\Big(\tau^{\eta}_{I(\epsilon)^c}(y) \leq rT/\eta\Big)\bigg)^{\frac{t/\gamma(\eta)}{T/\eta}-1}$$

$$= \frac{1}{r\cdot\big(1-\delta_B(\epsilon,rT)+o(1)\big)\cdot\gamma(\eta)T/\eta}\cdot\Big(1 - r\cdot\big(1-\delta_B(\epsilon,rT)+o(1)\big)\cdot\gamma(\eta)T/\eta\Big)^{\frac{t}{\gamma(\eta)T/\eta}-1}$$

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 2.8.

*Proof of Theorem 2.8.* 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

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

$$= C(B^-) \cdot e^{-t} + \delta_{t,B\setminus I}(\epsilon) + \delta_{0,I}(\epsilon).$$

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 $x \in I$,

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

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

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

for sufficiently small $\eta$'s; here, we applied $\gamma(\eta)/\eta \to 0$ as $\eta \downarrow 0$ in the last inequality. In light of (2.30) 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,

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

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

## 4.2 Proof of Theorem 2.6

In this section, we apply the framework developed in Section 2.3.2 and prove Theorem 2.6. Throughout this section, we impose Assumptions 1, 2, 3, and 5. Besides, we fix a few useful constants. For the majority of this section we fix the truncation threshold $b > 0$ such that $s_{\text{left}}/b \notin \mathbb{Z}$ and $s_{\text{right}}/b \notin \mathbb{Z}$. With this, for $l = \inf_{x \in I^c} |x| = |s_{\text{left}}| \wedge s_{\text{right}}$ we have $l > (\mathcal{J}^*_b - 1)b$. This allows us to fix, throughout this section, some $\bar{\epsilon} > 0$ small enough such that

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

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 (2.22). Also, recall that $I_\epsilon \triangleq (s_{\text{left}} + \epsilon, s_{\text{right}} - \epsilon)$ is the $\epsilon$-shrinkage of set $I$. We use $I^-_\epsilon = [s_{\text{left}} + \epsilon, s_{\text{right}} - \epsilon]$ to denote the closure of $I_\epsilon$. Then, the definition of $\boldsymbol{t}(\cdot)$ immediately implies

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

In our analysis below, we make use of the following inequality. We collect the proof in Section C, together with the proofs of several other useful properties regarding measures $\breve{\mathbf{C}}^{(k)|b}$.

**Lemma 4.2.** *The following claim holds for $\bar{\delta} > 0$ small enough and $\bar{t} > 0$ large enough: 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 $B \subseteq (I_{\bar{\epsilon}/2})^c$,*

$$\begin{aligned}(T - \bar{t}) \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B_\Delta) &\leq \inf_{x \in [-\epsilon,\epsilon]} \mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\bigg(\Big(\breve{E}(\epsilon, B, T)\Big)^\circ;\ x\bigg) \\ &\leq \sup_{x \in [-\epsilon,\epsilon]} \mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\bigg(\Big(\breve{E}(\epsilon, B, T)\Big)^-;\ x\bigg) \leq T \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\Delta) + \big(\bar{t}/\bar{\delta}^\alpha\big)^{\mathcal{J}_b^*}.\end{aligned}$$

*where $\breve{E}(\epsilon, B, T) \triangleq \Big\{\xi \in \mathbb{D}[0,T]:\ \exists t \leq T \text{ s.t. } \xi(t) \in B \text{ and } \xi(s) \in I_\epsilon\ \forall s \in [0,t)\Big\}$.*

To see how to apply the framework developed in Section 2.3.2, 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 $V_j^\eta(x) = X_j^{\eta|b}(x)$. Meanwhile, recall that $C_b^* = \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\big(I^c\big)$, and note that it is established in Lemma C.3 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

$$C(\ \cdot\ ) \triangleq \frac{\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\ \cdot\ \setminus I)}{C_b^*}, \qquad \gamma(\eta) \triangleq C_b^* \cdot \eta \cdot \big(\lambda(\eta)\big)^{\mathcal{J}_b^*}. \tag{4.10}$$

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 $l = |s_{\text{left}}| \wedge s_{\text{right}} > (\mathcal{J}_b^* - 1) \cdot b + \bar{\epsilon}$. Lemma C.2 then verifies $\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\partial I) = 0$ 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 (2.27) and (2.28). 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 $\lim_{\Delta \downarrow 0} \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^\Delta \cap I^c) \setminus (B^- \cap I^c)\Big) = 0$ and $\lim_{\Delta \downarrow 0} \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^\circ \cap I^c) \setminus (B_\Delta \cap I^c)\Big) = 0$. This allows us to fix a sequence $(\Delta^{(n)})_{n \geq 1}$ such that $\Delta^{(n+1)} \in (0, \Delta^{(n)}/2)$ and

$$\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^\Delta \cap I^c) \setminus (B^- \cap I^c)\Big) \vee \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B^\circ \cap I^c) \setminus (B_\Delta \cap I^c)\Big) \leq 1/2^n \tag{4.11}$$

for each $n \geq 1$. Next, recall the definition of set $\breve{E}(\epsilon, B, T)$ in Lemma 4.2, and let $\widetilde{B}(\epsilon) \triangleq B \setminus I_\epsilon$. Using Lemma 4.2, 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

$$\sup_{x \in [-\epsilon,\epsilon]} \mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\Big(\big(\breve{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^-;\ x\Big) \leq T \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I_\epsilon)^{\Delta^{(n)}}\Big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}_b^*}, \tag{4.12}$$

$$\inf_{x \in [-\epsilon,\epsilon]} \mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\Big(\big(\breve{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^\circ;\ x\Big) \geq (T - \bar{t}) \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B \setminus I_\epsilon)_{\Delta^{(n)}}\Big). \tag{4.13}$$

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

$$\begin{aligned}\breve{\boldsymbol{\delta}}_B(\epsilon, T) = T \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^{\Delta^{(n)}} \setminus B^-\Big) \vee \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(B^\circ \setminus B_{\Delta^{(n)}}\Big) \vee \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\epsilon + \Delta^{(n)}}\Big) \\ + \bar{t} \cdot \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\big(B^\circ \setminus I\big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}_b^*}\end{aligned} \tag{4.14}$$

and $\boldsymbol{\delta}_B(\epsilon, T) = \breve{\boldsymbol{\delta}}_B(\epsilon, T)/(C_b^* \cdot T)$. First, due to (4.11), we get

$$\lim_{T \to \infty} \boldsymbol{\delta}_B(\epsilon, T) \leq \frac{1}{C_b^*} \cdot \Big[\frac{1}{2^{n_\epsilon}} \vee \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((\partial I)^{\epsilon + \Delta^{(n_\epsilon)}}\Big)\Big]$$

where $n_\epsilon$ 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 $\partial I$ 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 (2.27) and (2.28). Specifically, we introduce stopping times

$$\tau^{\eta|b}_\epsilon(x) \triangleq \min\big\{j \geq 0 : \ X^{\eta|b}_j(x) \notin I_\epsilon\big\}. \tag{4.15}$$

**Lemma 4.3** (Verifying conditions (2.27) and (2.28))**.** *Let $\bar{t}$ be characterized as in Lemma C.1. Given any measurable $B \subseteq \mathbb{R}$, any $\epsilon > 0$ small enough, and any $T > \bar{t}$,*

$$\begin{aligned}
C(B^\circ) - \boldsymbol{\delta}_B(\epsilon, T) &\leq \liminf_{\eta\downarrow 0}\inf_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big)}{\gamma(\eta)T/\eta} \\
&\leq \limsup_{\eta\downarrow 0}\sup_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big)}{\gamma(\eta)T/\eta} \leq C(B^-) + \boldsymbol{\delta}_B(\epsilon, T).
\end{aligned}$$

*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

$$\limsup_{\eta\downarrow 0}\sup_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} \leq T\cdot\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^-\setminus I) + \check{\boldsymbol{\delta}}_B(\epsilon, T), \tag{4.16}$$

$$\liminf_{\eta\downarrow 0}\inf_{x\in(-\epsilon,\epsilon)} \frac{\mathbf{P}\Big(\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} \geq T\cdot\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\circ\setminus I) - \check{\boldsymbol{\delta}}_B(\epsilon, T). \tag{4.17}$$

To proceed, recall the definition of set $\check{E}(\epsilon, B, T)$ in Lemma 4.2. Let $\widetilde{B}(\epsilon) \triangleq B \setminus I_\epsilon$. Note that

$$\Big\{\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big\} = \Big\{\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in \widetilde{B}(\epsilon)\Big\} = \Big\{\boldsymbol{X}^{\eta|b}_{[0,T]}(x) \in \check{E}(\epsilon, \widetilde{B}(\epsilon), T)\Big\}.$$

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 $|\xi(t)| \geq l - \epsilon > l - \bar{\epsilon}$. On the other hand, using part $(b)$ of Lemma C.1, it holds for all $\xi \in \mathbb{D}^{(\mathcal{J}_b^*-1)|b}_{[-\epsilon,\epsilon]}[0, T]$ that $\sup_{t\in[0,T]}|\xi(t)| < l - -2\bar{\epsilon}$. In summary, we have established that

$$\boldsymbol{d}^{[0,T]}_{J_1}\Big(\check{E}(\epsilon, \widetilde{B}(\epsilon), T),\ \mathbb{D}^{(\mathcal{J}_b^*-1)|b}_{[-\epsilon,\epsilon]}[0, T]\Big) > 0$$

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 2.4 that

$$\begin{aligned}
\limsup_{\eta\downarrow 0}\sup_{x\in[-\epsilon,\epsilon]} \frac{\mathbf{P}\Big(\tau^{\eta|b}_\epsilon(x) \leq T/\eta;\ X^{\eta|b}_{\tau^{\eta|b}_\epsilon(x)}(x) \in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} &\leq \sup_{x\in[-\epsilon,\epsilon]} \mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\Big(\big(\check{E}(\epsilon, \widetilde{B}(\epsilon), T)\big)^-; x\Big) \\
&\leq T\cdot\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B\setminus I_\epsilon)^{\Delta^{(n)}}\Big) + (\bar{t}/\bar{\delta}^\alpha)^{\mathcal{J}_b^*};
\end{aligned} \tag{4.18}$$

here the last inequality we applied property (4.12). 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 (4.14), one can plug this bound back into (4.18) and yield the upper bound (4.16). Similarly, by applying Theorem 2.4 and property (4.13), we obtain

$$
\begin{aligned}
\liminf_{\eta\downarrow 0}\inf_{x\in[-\epsilon,\epsilon]}\frac{\mathbf{P}\Big(\tau_\epsilon^{\eta|b}(x)\leq T/\eta;\ X^{\eta|b}_{\tau_\epsilon^{\eta|b}(x)}(x)\in B\Big)}{\big(\lambda(\eta)\big)^{\mathcal{J}_b^*}} &\geq \inf_{x\in[-\epsilon,\epsilon]}\mathbf{C}^{(\mathcal{J}_b^*)|b}_{[0,T]}\Big(\big(\check{E}(\epsilon,\widetilde{B}(\epsilon),T)\big)^\circ;x\Big)\\
&\geq (T-\bar{t})\cdot\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((B\setminus I_\epsilon)_{\Delta^{(n)}}\Big).
\end{aligned}
\tag{4.19}
$$

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 (4.19), we establish the lower bound (4.17) and conclude the proof. □

The next two results verify conditions (2.29) and (2.30). Let

$$
R_\epsilon^{\eta|b}(x)\triangleq\min\big\{j\geq 0:\ X_j^{\eta|b}(x)\in(-\epsilon,\epsilon)\big\}
\tag{4.20}
$$

be the first time $X_j^{\eta|b}(x)$ returned to the $\epsilon$-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 (2.29) 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 (2.29), it suffices to prove the following result.

**Lemma 4.4** (Verifying condition (2.29))**.** *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 (2.19). We claim that $E(\epsilon)$ is bounded away from $\mathbb{D}^{(k-1)|b}_{I_\epsilon^-}[0,T]$. This allows us to apply Theorem 2.4 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 $\xi' \in E(\epsilon)$,

$$\boldsymbol{d}^{[0,T]}_{J_1}(\xi,\xi') \geq \frac{\epsilon}{2}. \tag{4.21}$$

First, if $\xi(t) \notin I_{\epsilon/2}$ for some $t \leq T$, then by definition of $E(\epsilon)$ we get $\boldsymbol{d}^{[0,T]}_{J_1}(\xi,\xi') \geq \frac{\epsilon}{2}$. 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

$$\xi(t) = \boldsymbol{y}_{t-t_{j-1}}\big(\xi(t_{j-1})\big) \qquad \forall t \in [t_{j-1},t_j). \tag{4.22}$$

for each $j \in [k]$. Here $\boldsymbol{y}_\cdot(x)$ is the ODE defined in (2.22). 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 (4.22) 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 $\boldsymbol{d}^{[0,T]}_{J_1}(\xi,\xi') \geq \frac{\epsilon}{2}$. This concludes the proof. □

Lastly, we establish condition (2.30). Note that the first visit time $\tau^\eta_{A(\epsilon)}(x)$ therein coincides with $R^{\eta|b}_\epsilon(x)$ defined in (4.20) due to our choice of $A(\epsilon) = (-\epsilon,\epsilon)$.

**Lemma 4.5** (Verifying condition (2.30))**.** *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 (2.19). 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 5 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,

$$\boldsymbol{d}^{[0,\boldsymbol{t}(\epsilon/2)]}_{J_1}\Big(\mathbb{D}^{(0)|b}_{I^-_\epsilon}[0,\boldsymbol{t}(\epsilon/2)],\ E^*_1(\epsilon)\Big) \geq \frac{\epsilon}{2} > 0, \tag{4.23}$$
$$\boldsymbol{d}^{[0,\boldsymbol{t}(\epsilon/2)]}_{J_1}\Big(\mathbb{D}^{(0)|b}_{I^-_\epsilon}[0,\boldsymbol{t}(\epsilon/2)],\ E^*_2(\epsilon)\Big) \geq \frac{\epsilon}{2} > 0. \tag{4.24}$$

This allows us to apply Theorem 2.4 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 $\lim_{\eta\downarrow 0}\lambda(\eta) = 0$. □

We prepare some tools that will not be directly applied in the proof of Theorem 2.6 but will be useful for the subsequent analysis in Section 5. First, we provide some straightforward bounds for the law of geometric random variables.

**Lemma 4.6.** *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) In addition, suppose that $\lim_{\epsilon\downarrow 0} a(\epsilon)/b(\epsilon) = 0$. 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 $\mathbf{P}(U(\epsilon) > \frac{1}{b(\epsilon)}) = \big(1-a(\epsilon)\big)^{\lfloor 1/b(\epsilon)\rfloor}$. 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 $\epsilon$ 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 $\epsilon$. □

Next, recall $\tau_1^{>\delta}(\eta)$ defined in (3.2) as the first arrival time of a noise $Z_n$ with size larger than $\delta/\eta$. Lemma 4.7 shows that it is very unlikely for $X_j^{\eta|b}$ to deviate far from the local minimum before $\tau_1^{>\delta}(\eta)$ (that is, before the arrival of any "large" noise).

**Lemma 4.7.** *Given any $\epsilon \in (0,\bar{\epsilon})$ and positive integer $N$, there is some $\bar{\delta} > 0$ such that*

$$\lim_{\eta\downarrow 0}\sup_{x\in(-\frac{\epsilon}{2},\frac{\epsilon}{2})} \mathbf{P}\Big(X_j^{\eta|b}(x) \notin (-\epsilon,\epsilon) \text{ for some } j < \tau_1^{>\delta}(\eta)\Big)\Big/\eta^N = 0 \qquad \forall \delta \in (0,\bar{\delta}).$$

*Proof.* Note that the values of $a(\cdot)$ and $\sigma(\cdot)$ outside of $[-\epsilon,\epsilon] \subseteq [-\bar{\epsilon},\bar{\epsilon}]$ has no impact on the first exit time from $(-\epsilon,\epsilon)$ when starting from $(-\epsilon/2,\epsilon/2)$. In light of Assumption 3, by modifying the values of $a(\cdot)$ and $\sigma(\cdot)$ outside of $[-\bar{\epsilon},\bar{\epsilon}]$ we can assume w.l.o.g. the existence of some $0 < c \leq C < \infty$ that

$$\inf_{x\in\mathbb{R}} \sigma(x) \geq c, \qquad \sup_{x\in\mathbb{R}} \sigma(x) \vee |a(x)| \leq C. \tag{4.25}$$

That is, in this proof we assume w.l.o.g. that Assumptions 4 and 8 hold.

For any $r > 0$, let $T^\eta_r(x) \triangleq \min\{j \geq 0 : \ X^{\eta|b}_j(x) \notin (-r,r)\}$. Due to the monotonicity in $\tau^{>\delta'}_1(\eta) \leq \tau^{>\delta}_1(\eta)$ for any $0 < \delta' < \delta$, it suffices to show that for any positive integer $N$ and any small enough $\epsilon > 0$, there is some $\delta = \delta(N,\epsilon) > 0$ such that

$$\limsup_{\eta \downarrow 0} \sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\big(T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta)\big)/\eta^N = 0. \tag{4.26}$$

Fix some $\beta > \alpha$ where $\alpha > 1$ is specified in Assumption 1. Also, pick some $\theta \in (0, \beta - \alpha)$. Applying Lemma 4.6 (i), we see that the claim $\mathbf{P}(\tau^{>\delta}_1(\eta) > 1/\eta^\beta) = \boldsymbol{o}\big(\exp(-1/\eta^\theta)\big)$ (as $\eta \downarrow 0$) holds for any $\delta > 0$. Also, note that $\tau^{>\delta}_1(\eta)$ only takes integer values, and observe that

$$\big\{T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta)\big\} \subseteq \big\{T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta) \leq 1/\eta^\beta\big\} \cup \big\{\tau^{>\delta}_1(\eta) > 1/\eta^\beta\big\}.$$

Therefore, to prove (4.26) we only need to find some $\delta > 0$ such that

$$\sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\big(T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta) \leq \lfloor 1/\eta^\beta \rfloor\big) = \boldsymbol{o}(\eta^N) \quad \text{ as } \eta \downarrow 0. \tag{4.27}$$

Recall the definition of $\boldsymbol{t}(\epsilon)$ in (4.8). Let $t \triangleq \boldsymbol{t}(\epsilon/2) < \infty$ and $K(\eta) \triangleq \lceil \frac{\lfloor 1/\eta^\beta \rfloor}{\lfloor t/\eta \rfloor} \rceil$. Note that $K(\eta) = \boldsymbol{O}(1/\eta^{\beta-1})$. Next, we fix some $\widetilde{\epsilon} > 0$ small enough such that $2\exp(tD)\widetilde{\epsilon} < \epsilon/2$, with $D < \infty$ being the Lipschitz constant in Assumption 2. Define events

$$\widetilde{A}_k(\eta,x) \triangleq \Bigg\{ \max_{(k-1)\lfloor \frac{t}{\eta} \rfloor + 1 \leq j \leq k \lfloor \frac{t}{\eta} \rfloor \wedge \big(\tau^{>\delta}_1(\eta) - 1\big)} \eta \Bigg| \sum_{i = (k-1)\lfloor \frac{t}{\eta} \rfloor + 1}^{j} \sigma\big(X^{\eta|b}_{i-1}(x)\big) Z_i \Bigg| \leq \widetilde{\epsilon} \Bigg\}.$$

For any $x \in (-\epsilon,\epsilon)$, any $\delta \in (0, \frac{b}{2C})$ and any $\eta \in (0, \frac{\widetilde{\epsilon}}{C} \wedge \frac{b \wedge 1}{2C})$ (where $C$ is specified in (4.25)), on event $\widetilde{A}_1(\eta,x)$ we observe the following facts. First, from (3.22) in part $(b)$ of Lemma 3.7 (in particular, the consideration of $s \leq t/\eta$ instead of $s \leq 1/\eta$ necessitates a rescaling of $\rho = \exp(Dt)$ in Lemma 3.7),

$$\sup_{s \leq \frac{t}{\eta} \wedge \big(\tau^{>\delta}_1(\eta) - 1\big)} \Big| \boldsymbol{y}_{\eta s}(x) - X^{\eta|b}_{\lfloor s \rfloor}(x) \Big| < \exp(tD)\widetilde{\epsilon} + \exp(tD)\eta C < 2\exp(tD)\widetilde{\epsilon} < \epsilon/2$$

due to our choice of $\eta$ and $\widetilde{\epsilon}$ above. Next, by Assumption 5, we have $\boldsymbol{y}_s(x) \in (-\epsilon,\epsilon)\ \forall s \geq 0$ and $\boldsymbol{y}_t(x) \in (-\epsilon/2, \epsilon/2)$. Combining these two facts, we get that $X^{\eta|b}_s(x)$ for all $s \leq \lfloor t/\eta \rfloor \wedge \big(\tau^{>\delta}_1(\eta) - 1\big)$ and, in case that $\tau^{>\delta}_1(\eta) > \lfloor t/\eta \rfloor$, we must have $X^{\eta|b}_{\lfloor t/\eta \rfloor}(x) \in (-\epsilon,\epsilon)$. By repeating this argument inductively for $k = 2, 3, \cdots, K(\eta)$, we can see that for any $x \in (-\epsilon,\epsilon)$, any $\delta \in (0, \frac{b}{2C})$, and any $\eta \in (0, \frac{\widetilde{\epsilon}}{C} \wedge \frac{b \wedge 1}{2C})$, it holds on event $\bigcap_{k=1}^{K(\eta)} \widetilde{A}_k(\eta,x)$ that

$$X^{\eta|b}_j(x) \in (-2\epsilon, 2\epsilon) \qquad \forall j \leq \lfloor 1/\eta^\beta \rfloor \wedge \big(\tau^{>\delta}_1(\eta) - 1\big) \leq K(\eta)\lfloor t/\eta \rfloor \wedge \big(\tau^{>\delta}_1(\eta) - 1\big).$$

As a result, for any $x \in (-\epsilon,\epsilon)$, any $\delta \in (0, \frac{b}{2C})$, and any $\eta \in (0, \frac{\widetilde{\epsilon}}{C} \wedge \frac{b \wedge 1}{2C})$,

$$\sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\Big(T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta)\Big) \leq \sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\Big( \bigcup_{k=1}^{K(\eta)} \Big(\widetilde{A}_k(\eta,x)\Big)^c \Big).$$

Lastly, due to part $(a)$ of Lemma 3.3, the claim $\sup_{k \in [K(\eta)]} \sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\Big(\big(\widetilde{A}_k(\eta,x)\big)^c\Big) = \boldsymbol{o}(\eta^{N+\beta-1})$ holds for all $\delta > 0$ small enough, which leads to

$$\sup_{x \in (-\epsilon,\epsilon)} \mathbf{P}\Big(T^\eta_{2\epsilon}(x) < \tau^{>\delta}_1(\eta)\Big) \leq K(\eta) \cdot \boldsymbol{o}(\eta^{N+\beta-1}) \leq \boldsymbol{O}(1/\eta^{\beta-1}) \cdot \boldsymbol{o}(\eta^{N+\beta-1}) = \boldsymbol{o}(\eta^N).$$

This verifies claim (4.27) and concludes the proof. □

We conclude this section with the proof of Theorem 2.6.

*Proof of Theorem 2.6.* (a) Since Lemmas 4.3–4.5 have verified Condition 1, part (a) of Theorem 2.6 follows immediately from Theorem 2.8. Note that it is established in Lemma C.3 that $C_b^* \in (0,\infty)$

(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

$$X^{\eta|b}_j(x) = X^{\eta}_j(x)\ \forall j \leq t. \tag{4.28}$$

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

$$\lim_{b\to\infty} \check{\mathbf{C}}^{(1)|b}(A) = \check{\mathbf{C}}(A). \tag{4.29}$$

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 (4.29). Similarly, for all $b > |s_{\text{left}}| \vee s_{\text{right}}$, we have

$$C^*_b = \check{\mathbf{C}}^{(1)|b}(I^c) = \int \mathbb{I}\Big\{\varphi_b\big(\sigma(0)\cdot w\big) \in I^c\Big\}\nu_\alpha(dw) = \int \mathbb{I}\Big\{\sigma(0)\cdot w \in I^c\Big\}\nu_\alpha(dw) = \check{\mathbf{C}}(I^c) \triangleq C^*. \tag{4.30}$$

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 $\mathcal{J}^*_b = \lceil l/b \rceil = 1$. 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

$$A(\eta,x) = \underbrace{\Big\{C^*H(\eta^{-1})\tau^\eta(x) \in (t,T],\ X^\eta_{\tau^\eta(x)}(x) \in B\Big\}}_{\triangleq A_1(\eta,x,T)} \cup \underbrace{\Big\{C^*H(\eta^{-1})\tau^\eta(x) > T,\ X^\eta_{\tau^\eta(x)}(x) \in B\Big\}}_{\triangleq A_2(\eta,x,T)}. \tag{4.31}$$

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 $\eta \in (0, \frac{b}{2C})$,

$$\begin{aligned}
&\mathbf{P}\Big(A_1(\eta,x,T) \cap E_b(\eta,T)\Big) \\
&= \mathbf{P}\bigg(\Big\{C^*_b\eta\cdot\lambda(\eta)\tau^{\eta|b}(x) \in (t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x) \in B\Big\} \cap E_b(\eta,T)\bigg) \qquad \text{due to (4.28) and (4.30)}
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)\in(t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big)\\
&= \mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)>t,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big)-\mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)>T,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big).
\end{aligned}
$$

Using part (a) of Theorem 2.6 and observation (4.30), we get

$$
\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\Big(A_1(\eta,x,T)\cap E_b(\eta,T)\Big)\leq \frac{\check{\mathbf{C}}^{(1)|b}(B^-)}{C^*}\cdot\exp(-t)-\frac{\check{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*}\cdot\exp(-T). \tag{4.32}
$$

On the other hand, $\sup_{x\in I_\epsilon}\mathbf{P}\big(A_1(\eta,x,T)\setminus E_b(\eta,T)\big)\leq\mathbf{P}\big((E_b(\eta,T))^c\big)=\mathbf{P}\big(\eta|Z_j|>\frac{b}{2C}\text{ for some }j\leq\frac{T}{C^*H(\eta^{-1})}\big)$. Applying Lemma 4.6 (i), we get

$$
\begin{aligned}
\limsup_{\eta\downarrow 0}\mathbf{P}\bigg(\eta|Z_j|>\frac{b}{2C}\text{ for some }j\leq\frac{T}{C^*H(\eta^{-1})}\bigg)&=1-\liminf_{\eta\downarrow 0}\mathbf{P}\bigg(\mathrm{Geom}\Big(H\big(\frac{b}{\eta\cdot 2C}\big)\Big)>\frac{T}{C^*H(\eta^{-1})}\bigg)\\
&\leq 1-\lim_{\eta\downarrow 0}\exp\bigg(-\frac{T\cdot H(\eta^{-1}\cdot\frac{b}{2C})}{C^*H(\eta^{-1})}\bigg)\\
&=1-\exp\bigg(-\frac{T}{C^*}\cdot\Big(\frac{2C}{b}\Big)^\alpha\bigg).
\end{aligned}\tag{4.33}
$$

Similarly,

$$
\begin{aligned}
A_2(\eta,x,T)&\subseteq\Big\{C^*H(\eta^{-1})\tau^\eta(x)>T\Big\}\\
&=\bigg(\Big\{C^*H(\eta^{-1})\tau^\eta(x)>T\Big\}\cap E_b(\eta,T)\bigg)\cup\bigg(\Big\{C^*H(\eta^{-1})\tau^\eta(x)>T\Big\}\setminus E_b(\eta,T)\bigg).
\end{aligned}
$$

On $\{C^*H(\eta^{-1})\tau^\eta(x)>T\}\cap E_b(\eta,T)$, due to (4.28) we have $\tau^\eta(x)=\tau^{\eta|b}(x)$. Also, from (4.30) we get $C^*=C_b^*$. Using part (a) of Theorem 2.6 again, we get

$$
\limsup_{\eta\downarrow 0}\mathbf{P}\bigg(\Big\{C^*H(\eta^{-1})\tau^\eta(x)>T\Big\}\cap E_b(\eta,T)\bigg)\leq\limsup_{\eta\downarrow 0}\mathbf{P}\Big(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)>T\Big)\leq\exp(-T). \tag{4.34}
$$

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 (4.33). Collecting (4.32), (4.33), and (4.34), we have shown that for all $b>0$ large enough and all $T>t$,

$$
\begin{aligned}
\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\Big(A(\eta,x)\Big)\leq&\frac{\check{\mathbf{C}}^{(1)|b}(B^-)}{C^*}\cdot\exp(-t)-\frac{\check{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*}\cdot\exp(-T)+\exp(-T)\\
&+2\cdot\bigg[1-\exp\bigg(-\frac{T}{C^*}\cdot\Big(\frac{2C}{b}\Big)^\alpha\bigg)\bigg].
\end{aligned}
$$

In light of claim (4.29), we can drive $b\to\infty$ and obtain $\limsup_{\eta\downarrow 0}\sup_{x\in I_\epsilon}\mathbf{P}\big(A(\eta,x)\big)\leq\frac{\check{\mathbf{C}}(B^-)}{C^*}\cdot\exp(-t)-\frac{\check{\mathbf{C}}(B^\circ)}{C^*}\cdot\exp(-T)+\exp(-T)$. Letting $T$ tend to $\infty$, we conclude the proof of the upper bound.

The lower bound can be established analogously. In particular, from the decomposition in (4.31), we get

$$
\inf_{x\in I_\epsilon}\mathbf{P}\big(A(\eta,x)\big)
$$

$$
\begin{aligned}
&\geq \inf_{x\in I_\epsilon} \mathbf{P}\big(A_1(\eta,x,T)\big) \geq \inf_{x\in I_\epsilon} \mathbf{P}\big(A_1(\eta,x,T)\cap E_b(\eta,T)\big) \\
&= \inf_{x\in I_\epsilon} \mathbf{P}\bigg(\Big\{C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)\in(t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\Big\}\cap E_b(\eta,T)\bigg) \qquad \text{due to (4.28) and (4.30)} \\
&\geq \inf_{x\in I_\epsilon} \mathbf{P}\bigg(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)\in(t,T],\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\bigg) - \mathbf{P}\Big(\big(E_b(\eta,T)\big)^c\Big) \\
&\geq \inf_{x\in I_\epsilon} \mathbf{P}\bigg(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)> t,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\bigg) - \sup_{x\in I_\epsilon}\mathbf{P}\bigg(C_b^*\eta\cdot\lambda(\eta)\tau^{\eta|b}(x)> T,\ X^{\eta|b}_{\tau^{\eta|b}(x)}(x)\in B\bigg) \\
&\qquad - \mathbf{P}\Big(\big(E_b(\eta,T)\big)^c\Big).
\end{aligned}
$$

Using part (a) of Theorem 2.6 and the limit in (4.33), we yield (for all $b>0$ large enough and all $T>t$)

$$
\liminf_{\eta\downarrow 0}\inf_{x\in I_\epsilon}\mathbf{P}\Big(A(\eta,x)\Big) \leq \frac{\breve{\mathbf{C}}^{(1)|b}(B^\circ)}{C^*}\cdot\exp(-t) - \frac{\breve{\mathbf{C}}^{(1)|b}(B^-)}{C^*}\cdot\exp(-T) - \bigg[1-\exp\bigg(-\frac{T}{C^*}\cdot\Big(\frac{2C}{b}\Big)^\alpha\bigg)\bigg].
$$

Sending $b\to\infty$ and then $T\to\infty$, we conclude the proof of the lower bound. □

# 5 Sample-Path Convergence of Global Dynamics

Here, we collect the proofs for Section 2.4. Specifically, Section 5.1 develops a set of technical tools for verifying the convergence of jump processes in f.d.d. and $(\mathbb{D}[0,\infty),\boldsymbol{d}^{[0,\infty)}_{L_p})$. Applying this framework, Section 5.2 provides the proof of Theorem 2.10, based on which we give the proof of Theorem 2.11 in Section 5.3.

## 5.1 Technical Lemmas for Theorem 2.10

Let $Y^\eta_\cdot$ and $Y^*_\cdot$ be random elements in $\mathbb{D}[0,\infty)$, i.e., $\mathbb{R}$-valued càdlàg processes. We start by discussing a few properties of the weak convergence in $(\mathbb{D}[0,\infty),\boldsymbol{d}^{[0,\infty)}_{L_p})$. In particular, a similar mode of convergence in $(\mathbb{D}[0,T],\boldsymbol{d}^{[0,T]}_{L_p})$ can be defined analogously for any $T\in(0,\infty)$. Recall the projection mapping $\pi_T$ defined in (2.36). We say that $Y^\eta_\cdot \Rightarrow Y^*_\cdot$ in $(\mathbb{D}[0,T],\boldsymbol{d}^{[0,T]}_{L_p})$ if

$$
\lim_{\eta\downarrow 0}\mathbf{E}f\big(\pi_T(S^\eta_\cdot)\big) = \mathbf{E}f\big(\pi_T(S^*_\cdot)\big) \qquad \forall f:\mathbb{D}[0,T]\to\mathbb{R}\text{ continuous and bounded;}
$$

see (2.35) for the definition of $\boldsymbol{d}^{[0,T]}_{L_p}$. More precisely, the $L_p$ norm $\boldsymbol{d}^{[0,T]}_{L_p}$ induces a metric over a quotient space $\mathbb{D}[0,T]/\mathcal{N}$. In particular, since we are dealing with the càdlàg space $\mathbb{D}[0,T]$, we set $\mathcal{N}=\{\xi\in\mathbb{D}[0,T]:\ \xi_t\equiv 0\ \forall t\in[0,T)\}$, which is the set containing all paths in $\mathbb{D}[0,T]$ that is constant zero except for the endpoint.

First, Lemma 5.1 shows that the convergence in $(\mathbb{D}[0,\infty),\boldsymbol{d}^{[0,\infty)}_{L_p})$ follows from the convergence in $(\mathbb{D}[0,T],\boldsymbol{d}^{[0,T]}_{L_p})$.

**Lemma 5.1.** *Let $p\in[1,\infty)$. If $Y^\eta_\cdot\Rightarrow Y^*_\cdot$ in $(\mathbb{D}[0,T],\boldsymbol{d}^{[0,T]}_{L_p})$ as $\eta\downarrow 0$ for any positive integer $T$, then $Y^\eta_\cdot\Rightarrow Y^*_\cdot$ in $(\mathbb{D}[0,\infty),\boldsymbol{d}^{[0,\infty)}_{L_p})$ as $\eta\downarrow 0$.*

*Proof.* By Portmanteau Theorem, it suffices to show that $\lim_{\eta\downarrow 0}\mathbf{E}f(Y^\eta_\cdot)=\mathbf{E}f(Y^*_\cdot)$ holds for any $f:\mathbb{D}[0,\infty)\to\mathbb{R}$ that is bounded and uniformly continuous. To proceed, we arbitrarily pick one such $f$ and some $\epsilon>0$. By virtue of the uniform continuity of $f$, there exists some $\delta>0$ such that $|f(x)-f(y)|<\epsilon$ whenever $\boldsymbol{d}^{[0,\infty)}_{L_p}(x,y)<\delta$. By definition of $\boldsymbol{d}^{[0,\infty)}_{L_p}$ in (2.37), we must have

$\boldsymbol{d}_{L_p}^{[0,\infty)}(x,y) < 1/2^{\lfloor T \rfloor - 1}$ if $x_t = y_t$ for all $t \in [0,T)$. Now, we fix some positive integer $T$ large enough such that $1/2^{T-1} < \delta$. Define $\widetilde{\pi}_T : \mathbb{D}[0,\infty) \to \mathbb{D}[0,\infty)$ by

$$\widetilde{\pi}_T(\xi)_t \triangleq \begin{cases} \xi_t & \text{if } t \in [0,T) \\ 0 & \text{if } t \geq T \end{cases}$$

and set $\widetilde{f}_T(\xi) \triangleq f\big(\widetilde{\pi}_T(\xi)\big)$. We now have $\boldsymbol{d}_{L_p}^{[0,\infty)}\big(\xi, \widetilde{\pi}_T(\xi)\big) < \delta$ and hence $|f(\xi) - \widetilde{f}_T(\xi)| < \epsilon$ for any $\xi \in \mathbb{D}[0,\infty)$. As a result,

$$\limsup_{\eta \downarrow 0} |\mathbf{E} f(Y^\eta_\cdot) - \mathbf{E}\widetilde{f}_T(Y^\eta_\cdot)| < \epsilon, \qquad |\mathbf{E} f(Y^*_\cdot) - \mathbf{E}\widetilde{f}_T(Y^*_\cdot)| < \epsilon. \tag{5.1}$$

Furthermore, let $\pi_T^\dagger : \mathbb{D}[0,T] \to \mathbb{D}[0,\infty)$ be defined as

$$\pi^\dagger(\xi')_t \triangleq \begin{cases} \xi'_t & \text{if } t \in [0,T) \\ 0 & \text{if } t \geq T \end{cases},$$

which, at an intuitive level, is interpreted as a "pseudo inverse" of the projection mapping $\pi_T$ defined in (2.36). Also, define functional $f_T : \mathbb{D}[0,T] \to \mathbb{R}$ by $f_T(\cdot) \triangleq f\big(\pi_T^\dagger(\cdot)\big)$. It is easy to see that $(i)$ $f_T$ is continuous due to the continuity of $f$ and $\pi_T^\dagger$, and $(ii)$ for any $\xi \in \mathbb{D}[0,\infty)$, we have $\widetilde{f}_T(\xi) = f_T\big(\pi_T(\xi)\big)$. Due to the assumption $Y^\eta_\cdot \Rightarrow Y^*_\cdot$ in $(\mathbb{D}[0,T], \boldsymbol{d}_{L_p}^{[0,T]})$, we now yield

$$\lim_{\eta \downarrow 0} |\mathbf{E}\widetilde{f}_T(Y^\eta_\cdot) - \mathbf{E}\widetilde{f}_T(Y^*_\cdot)| = 0. \tag{5.2}$$

Combining (5.1) and (5.2), we get $\limsup_{\eta \downarrow 0} |\mathbf{E} f(Y^\eta_\cdot) - \mathbf{E} f(Y^*_\cdot)| < 2\epsilon$. Driving $\epsilon \to 0$, we conclude the proof. □

Lemma 5.2 then provides a Prohorov-type argument where weak convergence in $(\mathbb{D}[0,T], \boldsymbol{d}_{L_p}^{[0,T]})$ can be established using the convergence in f.d.d. and a tightness condition. The proof is a straightforward adaptation of its $J_1$ counterparts. For the sake of clarity, the next proof will, w.l.o.g., focus on the case where $T = 1$, but we stress that the arguments can be easily extended to $\mathbb{D}[0,T]$ with arbitrary $T \in (0,\infty)$. Recall that we write $\mathbb{D} = \mathbb{D}[0,1]$.

**Lemma 5.2.** *Let $T \in (0,\infty)$, $p \in [1,\infty)$, and $\mathcal{T}$ be a dense subset of $(0,T)$. Suppose that the laws of $Y^{\eta_n}_\cdot$ are tight in $(\mathbb{D}[0,T], \boldsymbol{d}_{L_p}^{[0,T]})$ for any sequence $\eta_n > 0$ with $\lim_n \eta_n = 0$, and*

$$(Y^\eta_{t_1}, \cdots, Y^\eta_{t_k}) \Rightarrow (Y^*_{t_1}, \cdots, Y^*_{t_k}) \text{ as } \eta \downarrow 0 \qquad \forall k = 1, 2, \cdots,\ \forall (t_1, \cdots, t_k) \in \mathcal{T}^{k\uparrow}. \tag{5.3}$$

*Then $Y^\eta_\cdot \Rightarrow Y^*_\cdot$ in $(\mathbb{D}[0,T], \boldsymbol{d}_{L_p}^{[0,T]})$ as $\eta \downarrow 0$.*

*Proof.* The arguments below are adapted from the standard proofs in [3] for $J_1$ topology. For any $0 \leq t_1 < t_2 < \cdots < t_k \leq 1$, let $\pi_{(t_1,\cdots,t_k)} : \mathbb{D} \to \mathbb{R}^k$ be the projection mapping, i.e., $\pi_{(t_1,\cdots,t_k)}(\xi) = (\xi_{t_1}, \xi_{t_2}, \cdots, \xi_{t_k})$. Let $\mathcal{R}^k$ be the Borel $\sigma$-algebra for $\mathbb{R}^k$. Let $p[\pi_{\boldsymbol{t}} : \boldsymbol{t} \in \mathcal{T}]$ be the collection of all sets of form $\pi^{-1}_{(t_1,\cdots,t_k)} H$, where $k \geq 1$, $H \in \mathcal{R}^k$, and $t_1 < \cdots < t_k$ with $t_i \in \mathcal{T}$ for each $i \in [k]$. It suffices to show that (recall that $\boldsymbol{d}_{L_p} = \boldsymbol{d}_{L_p}^{[0,1]}$ and let $\mathcal{D}_p$ be the Borel $\sigma$-algebra of $(\mathbb{D}, \boldsymbol{d}_{L_p})$)

$$p[\pi_{\boldsymbol{t}} : \boldsymbol{t} \in \mathcal{T}] \text{ is a separating class for } (\mathbb{D}, \boldsymbol{d}_{L_p}). \tag{5.4}$$

In other words, any two Borel probability measures $\mu$ and $\nu$ over $(\mathbb{D}, \boldsymbol{d}_{L_p})$ would coincide (i.e., $\mu(A) = \nu(A)\ \forall A \in \mathcal{D}_p$) if $\mu(A) = \nu(A)\ \forall A \in p[\pi_{\boldsymbol{t}} : \boldsymbol{t} \in \mathcal{T}]$. To see why claim (5.4) is a sufficient condition, note that the tightness condition in Lemma 5.2 implies that the sequence $Y^{\eta_n}_\cdot$ has a converging subsequence, while the claim (5.4) and assumption (5.3) dictates that the limiting distribution must be that of $Y^*_\cdot$.

The remainder of this proof is devoted to establishing claim (5.4). First, we show that the projection mapping of form $\pi_{(t_1,\cdots,t_k)}:\mathbb{D}\to\mathbb{R}^k$ is $\mathcal{D}_p/\mathcal{R}^k$ measurable when $0\leq t_1<\cdots<t_k<1$, which immediately confirms that $p[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]\subseteq\mathcal{D}_p$. To do so, it suffices to prove that $\pi_{(t)}$ is measurable for any given $t\in[0,1)$. Define $h_\epsilon(x):\mathbb{D}\to\mathbb{R}$ by $h_\epsilon(x)=\epsilon^{-1}\int_t^{t+\epsilon}x_s ds$. W.l.o.g. we only consider $\epsilon$ small enough such that $t+\epsilon\leq 1$. For any $x,\ y\in\mathbb{D}$ and $\Delta\in(0,1)$,

$$\begin{aligned}|h_\epsilon(x)-h_\epsilon(y)| &\leq \epsilon^{-1}\int_t^{t+\epsilon}|x_s-y_s|ds\\ &=\epsilon^{-1}\int_t^{t+\epsilon}|x_s-y_s|\mathbb{I}\{|x_s-y_s|>\Delta\}ds+\epsilon^{-1}\int_t^{t+\epsilon}|x_s-y_s|\mathbb{I}\{|x_s-y_s|\leq\Delta\}ds\\ &\leq \epsilon^{-1}\int_t^{t+\epsilon}\frac{|x_s-y_s|^p}{|\Delta|^p}ds+\Delta.\end{aligned}$$

Therefore, for any sequence $y^{(n)}\in\mathbb{D}$ such that $\boldsymbol{d}_{L_p}(y^{(n)},x)\to 0$, we have $\limsup_{n\to\infty}|h_\epsilon(x)-h_\epsilon(y^{(n)})|\leq\Delta$. Driving $\Delta\downarrow 0$, we see that $h_\epsilon(\cdot)$ is a continuous mapping. On the other hand, the right continuity of all paths in $\mathbb{D}$ implies that $h_\epsilon(x)\to\pi_{(t)}(x)$ as $\epsilon\to 0$ for all $x\in\mathbb{D}$. As a result, the limiting mapping $\pi_{(t)}$ must be $\mathcal{D}_p/\mathcal{R}$ measurable.

Let $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]$ be the $\sigma$-algebra generated by $p[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]$. Note that we have verified $p[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]\subseteq\mathcal{D}_p$, which implies $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]\subseteq\mathcal{D}_p$ since $\mathcal{D}_p$ is also a $\sigma$-algebra. Furthermore, suppose that we can show

$$\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]\supseteq\mathcal{D}_p\qquad(\text{and hence }\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]=\mathcal{D}_p),\tag{5.5}$$

then we can confirm claim (5.4) using $\pi-\lambda$ Theorem. Indeed, for any Borel probability measures $\mu$ and $\nu$ over $(\mathbb{D},\boldsymbol{d}_{L_p})$, note that $\mathcal{L}\triangleq\{A\in\mathcal{D}_p:\ \mu(A)=\nu(A)\}$ is a $\lambda$-system. Whenever $p[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]\subseteq\mathcal{L}$, by applying $\pi-\lambda$ Theorem we then get $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]=\mathcal{D}_p\subseteq\mathcal{L}$. This concludes that $p[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]$ is a separating class.

Now, it only remains to prove claim (5.5). Since $\mathcal{T}$ is a dense subset of $(0,T)$, for each $m\geq 1$ we can pick some positive integer $k$ and some $0<s_1<\cdots<s_k<1$, with $s_i\in\mathcal{T}$, such that $\max_{i\in[k+1]}|s_{i+1}-s_i|<m^{-1}$, under the convention that $s_0=0$ and $s_{k+1}=1$. Now, construct a mapping $V_m:\mathbb{R}^k\to\mathbb{D}$ as follows: for each $\boldsymbol{\alpha}=(\alpha_1,\cdots,\alpha_k)\in\mathbb{R}^k$, define $\xi=V_m(\boldsymbol{\alpha})$ by setting $\xi_t=\alpha_i$ if $t\in[s_i,s_{i+1})$ for each $i\in[k+1]$ (with the convention that $\alpha_0=0$) and $\xi_1=\alpha_k$. It is easy to see that $V_m$ is continuous, and hence $\mathcal{R}^k/\mathcal{D}_p$ measurable. Besides, we have shown that $\pi_{(t_1,\cdots,t_k)}$ is $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]/\mathcal{R}^k$ measurable. As a result, the composition $V_m^*\triangleq V_m\pi_{s_1,\cdots,s_k}:\mathbb{D}\to\mathbb{D}$ is $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]/\mathcal{D}_p$ measurable.

To proceed, fix some $\epsilon>0$. For any $x\in\mathbb{D}$, define $x'\in\mathbb{D}$ such that $x'_t=x_t$ for all $t\in[\epsilon,1-\epsilon)$ and $x'_t=0$ otherwise. The boundedness of any path in $\mathbb{D}$ implies the existence of some $M_x\in(0,\infty)$ such that $\sup_t|x_t|\leq M_x$. Next, note that

$$\boldsymbol{d}_{L_p}(V_m^*x,x)\leq\underbrace{\boldsymbol{d}_{L_p}(V_m^*x',x')}_{\text{(I)}}+\underbrace{\boldsymbol{d}_{L_p}(V_m^*x',V_m^*x)}_{\text{(II)}}+\underbrace{\boldsymbol{d}_{L_p}(x',x)}_{\text{(III)}}.$$

First, it was shown in Theorem 12.5 of [3] that $\lim_{m\to\infty}\boldsymbol{d}_{J_1}(V_m^*x',x')=0$. This immediately implies that $\lim_{m\to\infty}\boldsymbol{d}_{L_p}(V_m^*x',x')=0$. Next, by definition of $x'$, we have $\limsup_{m\to\infty}\big[\text{(II)}\big]^p\leq(2M_x)^p\cdot 2\epsilon$ and $\limsup_{m\to\infty}\big[\text{(III)}\big]^p\leq(2M_x)^p\cdot 2\epsilon$. Driving $\epsilon\downarrow 0$, we obtain that $\lim_{m\to\infty}\boldsymbol{d}_{L_p}(V_m^*x,x)=0$ for all $x\in\mathbb{D}$. This implies that the identity mapping $\mathbf{I}(\xi)=\xi$ is also $\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]/\mathcal{D}_p$ measurable, which leads to $\mathcal{D}_p\subseteq\sigma[\pi_{\boldsymbol{t}}:\boldsymbol{t}\in\mathcal{T}]$ and concludes the proof. $\square$

Moreover, consider a family of $\mathbb{R}$-valued càdlàg processes $\hat{Y}_t^{\eta,\epsilon}$, supported on the same underlying probability space with process $Y_t^\eta$, that satisfies the following condition.

**Condition 2.** *For any $T \in (0,\infty)$ and $p \in [1,\infty)$, the following claims hold for all $\epsilon > 0$ small enough:*

*(i)* $\{\hat{Y}^{\eta,\epsilon}_t : \ t > 0\} \stackrel{f.d.d.}{\to} \{Y^*_t : \ t > 0\}$ *and* $\hat{Y}^{\eta,\epsilon}_{\cdot} \Rightarrow Y^*_{\cdot}$ *in* $(\mathbb{D}[0,T], \boldsymbol{d}^{[0,T]}_{L_p})$ *as* $\eta \downarrow 0$*;*

*(ii)* $\lim_{\eta\to 0}\mathbf{P}(|\hat{Y}^{\eta,\epsilon}_T - Y^\eta_T| \geq \epsilon) = 0$ *and* $\lim_{\eta\downarrow 0}\mathbf{P}\Big(\boldsymbol{d}^{[0,T]}_{L_p}(\hat{Y}^{\eta,\epsilon}_{\cdot}, Y^\eta_{\cdot}) \geq 2\epsilon\Big) = 0.$

As the first component of our framework, Lemma 5.3 shows that, under Condition 2, both $Y^\eta_t$ and $\hat{Y}^{\eta,\epsilon}_t$ admit the same limit $Y^*_t$.

**Lemma 5.3.** *If Condition 2 holds, then* $\{Y^\eta_t : \ t > 0\} \stackrel{f.d.d.}{\to} \{Y^*_t : \ t > 0\}$ *and, for any* $T > 0$, $Y^\eta_{\cdot} \Rightarrow Y^*_{\cdot}$ *in* $(\mathbb{D}[0,T], \boldsymbol{d}^{[0,T]}_{L_p})$ *as* $\eta \downarrow 0$.

*Proof.* We start with the $L_p$ convergence. By Portmanteau Theorem, it suffices to show that $\liminf_{\eta\downarrow 0}\mathbf{P}(Y^\eta_{\cdot} \in G) \geq \mathbf{P}(Y^*_{\cdot} \in G)$ for any open set $G$ in the $L_p$ topology of $\mathbb{D}[0,T]$. Next, (recall that $G_\epsilon$ is the $\epsilon$-shrinkage of $G$, and $G_\epsilon$ is also an open set)

$$\begin{aligned}\mathbf{P}(Y^\eta_{\cdot} \in G) &\geq \mathbf{P}(Y^\eta_{\cdot} \in G,\ \boldsymbol{d}^{[0,T]}_{L_p}(\hat{Y}^{\eta,\epsilon}_{\cdot}, Y^\eta_{\cdot}) < 2\epsilon) \geq \mathbf{P}(\hat{Y}^{\eta,\epsilon}_{\cdot} \in G_{2\epsilon},\ \boldsymbol{d}^{[0,T]}_{L_p}(\hat{Y}^{\eta,\epsilon}_{\cdot}, Y^\eta_{\cdot}) < 2\epsilon) \\ &\geq \mathbf{P}(\hat{Y}^{\eta,\epsilon}_{\cdot} \in G_{2\epsilon}) - \mathbf{P}(\boldsymbol{d}^{[0,T]}_{L_p}(\hat{Y}^{\eta,\epsilon}_{\cdot}, Y^\eta_{\cdot}) \geq 2\epsilon).\end{aligned}$$

For small enough $\epsilon > 0$, using part $(i)$ of Condition 2 we get $\liminf_{\eta\downarrow 0}\mathbf{P}(\hat{Y}^{\eta,\epsilon}_{\cdot} \in G_{2\epsilon}) \geq \mathbf{P}(Y^*_{\cdot} \in G_{2\epsilon})$, and by part $(ii)$ of Condition 2 we have $\lim_{\eta\downarrow 0}\mathbf{P}(\boldsymbol{d}^{[0,T]}_{L_p}(\hat{Y}^{\eta,\epsilon}_{\cdot}, Y^\eta_{\cdot}) \geq 2\epsilon) = 0$. Therefore, $\liminf_{\eta\downarrow 0}\mathbf{P}(Y^\eta_{\cdot} \in G) \geq \mathbf{P}(Y^*_{\cdot} \in G_{2\epsilon})$. Driving $\epsilon \downarrow 0$, we conclude the proof for the $L_p$ convergence. The proof for the f.d.d. convergence is almost identical and hence we omit the details. □

In light of Lemma 5.3, a natural approach to Theorem 2.10 is to identify some $\hat{Y}^{\eta,\epsilon}_t$ that converges to $Y^{*|b}_t$ while staying close enough to $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)$. To this end, we introduce the next key component of our framework, i.e., a technical tool for establishing the weak convergence of jump processes. Inspired by the approach in [24], Lemma 5.5 shows that the convergence of jump processes can be established by verifying the convergence of the inter-arrival times and destinations of jumps. Specifically, we introduce the following mapping $\Phi$ for constructing jump processes.

**Definition 5.4.** *Let random variables* $((U_j)_{j\geq 1}, (V_j)_{j\geq 1})$ *be such that* $V_j \in \mathbb{R}\ \forall j \geq 1$ *and*

$$U_j \in [0,\infty) \quad \forall j \geq 1, \qquad \lim_{i\to\infty}\mathbf{P}\Big(\sum_{j=1}^{i} U_j > t\Big) = 1 \quad \forall t > 0. \tag{5.6}$$

*Let mapping $\Phi(\cdot)$ be defined as follows: the image* $Y_{\cdot} = \Phi\big((U_j)_{j\geq 1}, (V_j)_{j\geq 1}\big)$ *is a stochastic process taking values in $\mathbb{R}$ such that (under the convention $V_0 \equiv 0$)*

$$Y_t = V_{\mathcal{J}(t)}\ \forall t \geq 0 \qquad \textit{where} \qquad \mathcal{J}(t) \triangleq \max\{J \geq 0 :\ \sum_{j=1}^{J} U_j \leq t\}. \tag{5.7}$$

**Remark 4.** *We add two remarks regarding Definition 5.4. First, $(U_j)_{j\geq 1}$ and $(V_j)_{j\geq 1}$ 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 $\lim_{i\to\infty}\mathbf{P}(\sum_{j=1}^{i} U_j > t) = 1\ \forall t > 0$ 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 $U_j > 0\ \forall j \geq 1$, the process $Y_t$ admits a more standard expression and satisfies $Y_t = V_i$ for all $t \in [\sum_{j=1}^{i} U_j, \sum_{j=1}^{i+1} U_j)$. 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 $w \in \mathbb{R}$ the process $Y_t$ is absorbed at $w$, then a representation compatible with Definition 5.4 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 mentioned above, Lemma 5.5 states that the convergence of jump processes in f.d.d. follows from the convergence in distributions of the inter-arrival times and destinations of jumps.

**Lemma 5.5.** *Let mapping $\Phi$ be specified as in Definition 5.4. Let $Y_{\cdot} = \Phi\big((U_j)_{j\geq 1}, (V_j)_{j\geq 1}\big)$ and, for each $n \geq 1$, $Y^n_{\cdot} = \Phi\big((U^n_j)_{j\geq 1}, (V^n_j)_{j\geq 1}\big)$. 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 $\{Y^n_t : \ t > 0\} \overset{f.d.d.}{\to} \{Y^*_t : \ t > 0\}$ as $n \to \infty$.*

*Proof.* 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 find some $J(\epsilon) > 0$ such that $\mathbf{P}(\sum_{j=1}^{J(\epsilon)} U_j \leq t) < \epsilon$, and hence $\mathbf{P}(\sum_{j=1}^{J(\epsilon)} U^n_j \leq t) < \epsilon$ for all $n$ large enough. Also, we can fix $\Delta(\epsilon) > 0$ such that $\mathbf{P}\Big( \sum_{i=1}^j U_i \in \bigcup_{l\in[k]}[t_l - \Delta(\epsilon), t_l + \Delta(\epsilon)]$ for some $j \leq J(\epsilon)\Big) < \epsilon$. Throughout the proof, we may abuse the notation slightly and write $J = J(\epsilon)$ and $\Delta = \Delta(\epsilon)$ when there is no ambiguity.

For any probability measure $\mu$, let $\mathscr{L}_\mu(X)$ be the law of the random element $X$ under $\mu$. Applying Skorokhod's representation theorem, we can 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 (i) $\mathscr{L}_{\mathbf{P}}(U^n_1, V^n_1, U^n_2, V^n_2, \cdots) = \mathscr{L}_{\mathbf{Q}}(\widetilde{U}^n_1, \widetilde{V}^n_1, \widetilde{U}^n_2, \widetilde{V}^n_2 \cdots)$ for all $n \geq 1$, (ii) $\mathscr{L}_{\mathbf{P}}(U_1, V_1, U_2, V_2, \cdots) = \mathscr{L}_{\mathbf{Q}}(\widetilde{U}_1, \widetilde{V}_1, \widetilde{U}_2, \widetilde{V}_2, \cdots)$, and (iii) $\widetilde{U}^n_j \xrightarrow{\mathbf{Q}-a.s.} \widetilde{U}_j$ and $\widetilde{V}^n_j \xrightarrow{\mathbf{Q}-a.s.} \widetilde{V}_j$ as $n \to \infty$ for all $j \geq 1$. 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 $Y = \Phi\Big((\widetilde{U}_j)_{j\geq 1}, (\widetilde{V}_j)_{j\geq 1}\Big)$ and (for each $n \geq 1$) $Y^n = \Phi\Big((\widetilde{U}^n_j)_{j\geq 1}, (\widetilde{V}^n_j)_{j\geq 1}\Big)$. Next, for each $i \in [k]$, we define

$$\mathcal{I}^{\leftarrow}_i(\Delta) = \max\{j \geq 0 : \ \widetilde{U}_1 + \cdots \widetilde{U}_j \leq t_i - \Delta\}, \qquad \mathcal{I}^{\rightarrow}_i(\Delta) = \min\{j \geq 0 : \ \widetilde{U}_1 + \cdots \widetilde{U}_j \geq t_i + \Delta\}.$$

That is, $\mathcal{I}^{\leftarrow}_i(\Delta)$ is the index of the last jump in $Y_s$ before time $t_i - \Delta$, and $\mathcal{I}^{\rightarrow}_i(\Delta)$ is the index of the first jump after time $t_i + \Delta$. Recall that we have fixed $0 < t_1 < \cdots < t_k = t < \infty$. On event

$$A_n = \Big\{ \sum_{i=1}^j \widetilde{U}_i \notin \bigcup_{l\in[k]} [t_l - \Delta, t_l + \Delta] \ \forall j \leq J \Big\} \cap \Big\{ \sum_{j=1}^J \widetilde{U}_j > t, \ \sum_{j=1}^J \widetilde{U}^n_j > t \Big\},$$

we have $\mathcal{I}^{\rightarrow}_i(\Delta) = \mathcal{I}^{\leftarrow}_i(\Delta) + 1 \leq J$ for all $i \in [k]$. Then, on $A_n$ it holds $\mathbf{Q}$-a.s. that (for all $i \in [k]$)

$$\lim_{n\to\infty} \sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)} \widetilde{U}^n_j = \sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)} \widetilde{U}_j \leq t_i - \Delta, \qquad \lim_{n\to\infty} \sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)+1} \widetilde{U}^n_j = \sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)+1} \widetilde{U}_j \geq t_i + \Delta,$$

As a result, on $A_n$ it holds for all $n$ large enough that $\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)} \widetilde{U}^n_j < t_i$ and $\sum_{j=1}^{\mathcal{I}^{\leftarrow}_i(\Delta)+1} \widetilde{U}^n_j > t_i$ for all $i \in [k]$, implying that $Y^n_{t_i} = \widetilde{V}_{\mathcal{I}^{\leftarrow}_i(\Delta)} \ \forall i \in [k]$. Furthermore, due to $\widetilde{V}^n_j \to \widetilde{V}_j$ $\mathbf{Q}$-a.s. for all $j \leq J$, it holds $\mathbf{Q}$-a.s. that $\lim_{n\to\infty} |\widetilde{V}^n_{I^{\leftarrow}_i(\Delta)} - \widetilde{V}_{I^{\leftarrow}_i(\Delta)}| \leq \lim_{n\to\infty} \max_{j\leq J} |\widetilde{V}^n_j - \widetilde{V}_j| = 0$. Therefore, on $A_n$ 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 $i \in [k]$. Then, for any $g : \mathbb{R}^k \to \mathbb{R}$ 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 $\|g\| = \sup_{\boldsymbol{y}\in\mathbb{R}^k} |g(\boldsymbol{y})|$)

$$\limsup_{n\to\infty} \Big| \mathbf{E}g(\boldsymbol{Y}^n) - \mathbf{E}g(\boldsymbol{Y}) \Big| \leq \limsup_{n\to\infty} \mathbf{E}_{\mathbf{Q}} \Big| g(\boldsymbol{Y}^n) - g(\boldsymbol{Y}) \Big|$$

$$
\begin{aligned}
&= \limsup_{n\to\infty} \mathbf{E}_{\mathbf{Q}} \Big| g(\boldsymbol{Y}^n) - g(\boldsymbol{Y}) \Big| \mathbb{I}_{A_n} + \limsup_{n\to\infty} \mathbf{E}_{\mathbf{Q}} \Big| g(\boldsymbol{Y}^n) - g(\boldsymbol{Y}) \Big| \mathbb{I}_{(A_n)^c} \\
&\leq 0 + 2\,\|g\| \limsup_{n\to\infty} \mathbf{Q}\big((A_n)^c\big) \qquad \text{due to } \boldsymbol{Y}^n \xrightarrow{\mathbf{Q}-a.s.} \boldsymbol{Y} \text{ on } A_n \\
&\leq 2\,\|g\| \cdot \bigg( \limsup_{n\to\infty} \mathbf{Q}\big(\sum_{i=1}^{J} \widetilde{U}_j \leq t\big) + \limsup_{n\to\infty} \mathbf{Q}\big(\sum_{i=1}^{J} \widetilde{U}_j^n \leq t\big) \\
&\qquad\qquad + \limsup_{n\to\infty} \mathbf{Q}\Big( \sum_{i=1}^{j} \widetilde{U}_i \in \bigcup_{l\in[k]} [t_l - \Delta, t_l + \Delta] \text{ for some } j \leq J \Big) \bigg) \\
&\leq 6\,\|g\| \cdot \epsilon.
\end{aligned}
$$

The last inequality follows from our choice of $J = J(\epsilon)$ and $\Delta = \Delta(\epsilon)$ at the beginning. From the arbitrariness of the mapping $g$ and $\epsilon > 0$, we conclude the proof using Portmanteau theorem. $\square$

## 5.2 Proof of Theorem 2.10

Now, we explain how to apply the framework developed in Section 5.1 and establish Theorem 2.10. In particular, the verification of the part $(i)$ of Condition 2 hinges on the chioce of the approximator $\hat{Y}^{\eta,\epsilon}_t$. Here, we construct a process $\hat{X}^{\eta,\epsilon|b}_t(x)$ as follows. Let $\hat{\tau}^{\eta,\epsilon|b}_0(x) \triangleq 0$,

$$
\hat{\tau}^{\eta,\epsilon|b}_1(x) \triangleq \min\Big\{ j \geq 0 :\ X^{\eta|b}_j(x) \in \bigcup_{i\in[n_{\min}]} (m_i - \epsilon, m_i + \epsilon) \Big\}, \tag{5.8}
$$

and

$$
\hat{\mathcal{I}}^{\eta,\epsilon|b}_1(x) \triangleq i \iff X^{\eta|b}_{\hat{\tau}^{\eta,\epsilon|b}_1(x)}(x) \in I_i. \tag{5.9}
$$

For $k \geq 2$,

$$
\hat{\tau}^{\eta,\epsilon|b}_k(x) \triangleq \min\Big\{ j \geq \hat{\tau}^{\eta,\epsilon|b}_{k-1}(x) :\ X^{\eta|b}_j(x) \in \bigcup_{i\neq\hat{\mathcal{I}}^{\eta,\epsilon|b}_{k-1}(x)} (m_i - \epsilon, m_i + \epsilon) \Big\} \qquad \forall k \geq 2. \tag{5.10}
$$

and

$$
\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x) \triangleq i \iff X^{\eta|b}_{\hat{\tau}^{\eta,\epsilon|b}_k(x)}(x) \in I_i. \tag{5.11}
$$

Essentially, $\hat{\tau}^{\eta,\epsilon|b}_k(x)$ records the $k$-th time $X^{\eta|b}_j(x)$ visits (the $\epsilon$-neighborhood of) a local minimum and $\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)$ denotes the index of the visited local minimum. Let

$$
\hat{X}^{\eta,\epsilon|b}_{\cdot}(x) \triangleq \Phi\bigg( \Big( \big(\hat{\tau}^{\eta,\epsilon|b}_k(x) - \hat{\tau}^{\eta,\epsilon|b}_{k-1}(x)\big) \cdot \lambda^*_b(\eta) \Big)_{k\geq 1}, \big(m_{\hat{\mathcal{I}}^{\eta,\epsilon|b}_k(x)}\big)_{k\geq 1} \bigg).
$$

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 a time scaling of $\lambda^*_b(\eta)$.

Using Lemma 5.5, the convergence of $\hat{X}^{\eta,\epsilon|b}_{\cdot}(x)$ 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 local minima over the potential $U$. 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 5.6.

**Proposition 5.6.** *Let Assumptions 1, 2, 3, 6, and 7 hold. Let $i \in [n_{min}]$ and $x \in I_i$. For any $\epsilon > 0$ small enough, the following claims hold.*

*(i)* $\{\hat{X}^{\eta,\epsilon|b}_t(x):\ t>0\} \stackrel{f.d.d.}{\to} \{Y^{*|b}_t:\ t>0\}$ *as* $\eta \downarrow 0$;

*(ii) Given any* $T \in (0,\infty)$, $p \in [1,\infty)$, *and any sequence of strictly positive reals* $\eta_n$*'s such that* $\lim_{n\to\infty}\eta_n = 0$, *the laws of* $\hat{X}^{\eta_n,\epsilon|b}_{\cdot}$ *are tight in* $(\mathbb{D}[0,T], \boldsymbol{d}^{[0,T]}_{L_p})$.

Proposition 5.7 then verifies part (*ii*) of Condition 2 in Lemma 5.3, under the choice of $Y^\eta_t = X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)$ and $\hat{Y}^{\eta,\epsilon}_t = \hat{X}^{\eta,\epsilon|b}_t(x)$. We give the proof in Section 5.2.

**Proposition 5.7.** *Let Assumptions 1, 2, 3, 6, and 7 hold. Let* $x \in \bigcup_{i\in[n_{min}]} I_i$. *Given any* $T>0$ *and* $p \in [1,\infty)$, *it holds for all* $\epsilon>0$ *small enough that*

$$\lim_{\eta\downarrow 0}\mathbf{P}\bigg(\boldsymbol{d}^{[0,T]}_{L_p}\Big(X^{\eta|b}_{\lfloor\cdot/\lambda^*_b(\eta)\rfloor}(x), \hat{X}^{\eta,\epsilon|b}_{\cdot}(x)\Big) \geq 2\epsilon\bigg) = 0, \qquad \lim_{\eta\downarrow 0}\mathbf{P}\Big(\Big|X^{\eta|b}_T(x) - \hat{X}^{\eta,\epsilon|b}_T(x)\Big| \geq \epsilon\Big) = 0.$$

Now, we are ready to prove Theorem 2.10.

*Proof of Theorem 2.10.* Fix some $i \in [n_{\min}]$ and $x \in I_i$. From Lemma 5.2 and Proposition 5.6, we verify part (*i*) of Condition 2, i.e., given any $T>0$, the claim

$$\{\hat{X}^{\eta,\epsilon|b}_t(x):\ t>0\} \stackrel{f.d.d.}{\to} \{Y^{*|b}_t:\ t>0\} \quad \text{and} \quad \hat{X}^{\eta,\epsilon|b}_{\cdot}(x) \Rightarrow Y^{*|b}_{\cdot} \text{ in } (\mathbb{D}[0,T], \boldsymbol{d}^{[0,T]}_{L_p}) \text{ as } \eta\downarrow 0$$

holds for all $\epsilon>0$ small enough. Meanwhile, given any $T\in(0,\infty)$ and $p\in[1,\infty)$, Proposition 5.7 verifies part (*ii*) of Condition 2 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 $Y^*_t = Y^{*|b}_t$. Applying Lemma 5.3, we obtain that (for any $T\in(0,\infty)$ and $p\in[1,\infty)$)

$$\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x):\ t>0\} \stackrel{f.d.d.}{\to} \{Y^{*|b}_t:\ t>0\} \quad \text{and} \quad X^{\eta|b}_{\lfloor\cdot/\lambda^*_b(\eta)\rfloor}(x) \Rightarrow Y^{*|b}_{\cdot} \text{ in } (\mathbb{D}[0,T], \boldsymbol{d}^{[0,T]}_{L_p})$$

as $\eta\downarrow 0$. This allows us to conclude the proof using Lemma 5.1. □

The remainder of Section 5.2 is devoted to proving Propositions 5.6 and 5.7. Henceforth in Section 5.2, we fix some $b\in(0,\infty)$ be such that Assumption 7 holds. In particular,

$$|s_j - m_i|/b \notin \mathbb{Z} \qquad \forall i\in[n_{\min}],\ j\in[n_{\min}-1]. \tag{5.12}$$

This allows us to fix some $\bar{\epsilon}\in(0,1\wedge b)$ such that

$$l_i > \big(\mathcal{J}^*_b(i)-1\big)b + 3\bar{\epsilon} \text{ and } [m_i-\bar{\epsilon}, m_i+\bar{\epsilon}] \subseteq [s_{i-1}+\bar{\epsilon}, s_i-\bar{\epsilon}] \qquad \forall i\in[n_{\min}] \tag{5.13}$$

with $l_i$ and $\mathcal{J}^*_b(i)$ defined in (2.31) and (2.32), respectively. In other words, we fix some $\bar{\epsilon}$ small enough such that, even with $\bar{\epsilon}$-shrinkage, the number of jumps required to exit from ($\epsilon$-shrunken) $I_i$ remains $\mathcal{J}^*_b(i)$.

We start by highlighting a few properties of the limiting Markov jump process $Y^{*|b}$ in Theorem 2.10, using results for the measure $\check{\mathbf{C}}^{(k)|b}$ collected in Section C of the Appendix. Recall the definitions of $q_b(i)$ and $q_b(i,j)$ in (2.40). First, by definition,

$$q_b(i) = \sum_{j\in[n_{\min}]:\ j\neq i} q_b(i,j) + \sum_{j\in[n_{\min}-1]} \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(\{s_j\}; m_i).$$

From (5.13), we have $|s_j - m_i| > (\mathcal{J}^*_b(i)-1)\cdot b + \bar{\epsilon}$. Then, by applying Lemma C.2, we get $\sum_{j\in[n_{\min}-1]} \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(\{s_j\}; m_i) = 0$. Together with Lemma C.3, we yield that

$$\sum_{j\in[n_{\min}]:\ j\neq i} q_b(i,j) = q_b(i) \in (0,\infty). \tag{5.14}$$

Furthermore, Lemma C.4 verifies that

$$q_b(i,j) > 0 \qquad \Longleftrightarrow \qquad \mathcal{J}_b^*(i,j) = \mathcal{J}_b^*(i). \tag{5.15}$$

As a result, in Definition 2.9 we know that the typical transition graph associated with threshold $b$ contains an edge $(m_i \to m_j)$ if and only if $q_b(i,j) > 0$.

Next, we stress that the law of the Markov jump process $Y^{*|b}$ can be expressed using the mapping $\Phi$ introduced in Definition 5.4. Given any $m_{\text{init}} \in \{m_1, m_2, \ldots, m_{n_{\min}}\}$, we set $V_1 = m_{\text{init}}$, $U_1 = 0$, and (for any $t > 0$, $l \geq 1$, and $i, j \in [n_{\min}]$ with $i \neq j$)

$$\begin{aligned} &\mathbf{P}\Big(U_{l+1} < t,\ V_{l+1} = m_j \ \Big|\ V_l = m_i, (V_j)_{j=1}^{l-1},\ (U_j)_{j=1}^{l}\Big) = \mathbf{P}\Big(U_{l+1} < t,\ V_{l+1} = m_j \ \Big|\ V_l = m_i\Big) \\ &= \begin{cases} \frac{q_b(i,j)}{q_b(i)} & \text{if } m_i \notin V_b^*, \\ \frac{q_b(i,j)}{q_b(i)} \cdot \Big(1 - \exp\big(-q_b(i)t\big)\Big) & \text{if } m_i \in V_b^*. \end{cases} \end{aligned} \tag{5.16}$$

In other words, conditioning on $V_l = m_i$, 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 claim that

$$Y_{\cdot}^{*|b} \stackrel{d}{=} \Phi\Big((U_j)_{j\geq 1}, (V_j)_{j\geq 1}\Big). \tag{5.17}$$

In fact, under the conditions in Theorem 2.10, it is straightforward to show that

- $(i)$ For any $t > 0$, $\lim_{i\to\infty} \mathbf{P}(\sum_{j\leq i} U_j > t) = 1$;
- $(ii)$ For any $u > 0$ and $i \geq 1$, $\mathbf{P}(U_1 + \cdots + U_i = u) = 0$;
- $(iii)$ $Y_{\cdot}^{*|b} \stackrel{d}{=} \Phi\Big((U_j)_{j\geq 1}, (V_j)_{j\geq 1}\Big)$; that is, it is a continuous-time Markov chain with state space $V_b^*$, generator

$$\mathbf{P}(Y_{t+h}^{*|b} = m_j \mid Y_t^{*|b} = m_i) = h \cdot \sum_{j' \in [n_{\min}]:\ j' \neq i} q_b(i,j')\theta_b(m_j|m_{j'}) + \boldsymbol{o}(h) \qquad \text{as } h \downarrow 0,$$

  and initial distribution $\mathbf{P}(Y_0^{*|b} = m_j) = \theta_b(m_j|m_{\text{init}})$; see (2.40) and (2.41) for the definitions of $q_b(i,j)$ and $\theta_b(m_j|m_i)$, respectively.

For the sake of completeness, we collect the proof in Section D. The representation (5.17) and the properties stated above will greatly facilitate the proofs below.

The proofs of Propositions 5.6 and 5.7 hinge on the first exit analysis in Theorem 2.6 and Section 4. Note that Theorem 2.6 focuses on some bounded interval $I$. In contrast, regarding the potential $U$ characterized in Assumption 6, while for all $i = 2, \cdots, n_{\min}$ the attraction field $I_i$ is indeed bounded, for $i = 1$ or $n_{\min}$ (that is, the leftmost or the rightmost attraction field) we have $I_1 = (-\infty, s_1)$ and $I_{n_{\min}} = (s_{n_{\min}-1}, \infty)$, both of which are unbounded. Besides, our analysis below involves $S(\delta) \triangleq \bigcup_{i\in[n_{\min}-1]}[s_i - \delta, s_i + \delta]$ (i.e., the union of the $\delta$-neighborhood of any boundary point $s_i$). As a result, we will frequently consider sets of form

$$I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M, M) = (I_i)_\delta \cap (-M, M) \tag{5.18}$$

for some $\delta, M > 0$. For any $M > 0$ large enough such that $-M < m_1 < s_1 < \cdots < s_{n_{\min}-1} < m_{n_{\min}} < M$, 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 $i = 2, 3, \cdots, n_{\min} - 1$, and 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 $I_{n_{\min};\delta,M} = (s_{n_{\min}-1} + \delta, s_{n_{\min}} - \delta) \cap (-M, M) = (s_{n_{\min}-1} + \delta, M)$ (due to $s_{n_{\min}} = \infty$). We also set

$$\sigma_{i;\epsilon}^{\eta|b}(x) \triangleq \min\Big\{j \geq 0 :\ X_j^{\eta|b}(x) \in \bigcup_{l\neq i}(m_l - \epsilon, m_l + \epsilon)\Big\}, \tag{5.19}$$

$$\tau^{\eta|b}_{i;\delta,M}(x) \triangleq \min\Big\{j \geq 0 :\ X^{\eta|b}_j(x) \notin I_{i;\delta,M}\Big\}. \tag{5.20}$$

In other words, $\tau^{\eta|b}_{i;\delta,M}(x)$ is the first exit time from $I_{i;\delta,M}$, and $\sigma^{\eta|b}_{i;\epsilon}(x)$ is the first time visiting the $\epsilon$-neighborhood of a local minimum different from $m_i$.

We first state Lemmas 5.8 and 5.9. To give an overview, we establish these two lemmas by adapting and applying the first exit analysis in Section 4 to the slightly more general settings in Propositions 5.6 and 5.7. We collect their proofs in Section E to omit repetitions.

First, Lemma 5.8 states that it is unlikely to get close to any of the boundary points $s_i$'s or exit a wide enough compact set.

**Lemma 5.8.** *Let Assumptions 1, 2, 3, 4, and 6 hold. Let $b \in (0,\infty)$ be such that (5.12) holds. There exists $M > 0$ such that*

$$\max_{i\in[n_{min}]} \breve{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big((-M,M)^c; m_i\big) = 0, \tag{5.21}$$

*Furthermore, given any $\Delta > 0$ and any $\epsilon \in (0,\bar{\epsilon})$ (with $\bar{\epsilon}$ specified in (5.13)), it holds for all $\delta > 0$ small enough that*

$$\limsup_{\eta\downarrow 0}\max_{i\in[n_{min}]}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\Big(\exists j < \sigma^{\eta|b}_{i;\epsilon}(x)\ s.t.\ X^{\eta|b}_j(x) \in S(\delta)\ or\ \big|X^{\eta|b}_j(x)\big| \geq M+1\Big) < \Delta. \tag{5.22}$$

Recall the scale function $\lambda^*_b$ defined in (2.38). Lemma 5.9 then provides an analogue of Theorem 2.6 for the current setup.

**Lemma 5.9.** *Let Assumptions 1, 2, 3, 4 and 6 hold. Let $b \in (0,\infty)$ be such that (5.12) holds. Let $\bar{\epsilon} > 0$ be specified as in (5.13).*

*(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 $t > 0$,*

$$\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\bigg) \geq \exp\big(-q_b(i)\cdot t\big)\cdot\frac{q_b(i,j)}{q_b(i)},$$

$$\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\bigg) \leq \exp\big(-q_b(i)\cdot t\big)\cdot\frac{q_b(i,j)}{q_b(i)}.$$

*If $m_i \notin V^*_b$, then for any $\epsilon \in (0,\bar{\epsilon})$ and any $t > 0$,*

$$\begin{aligned}\frac{q_b(i,j)}{q_b(i)} &\leq \liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\bigg)\\ &\leq \limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}\bigg(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) \leq t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\bigg) \leq \frac{q_b(i,j)}{q_b(i)}.\end{aligned}$$

Now, we are ready to prove Proposition 5.6.

*Proof of Proposition 5.6.* We first show that claims $(i)$ and $(ii)$ follow directly from the next claim: for any $\epsilon > 0$ small enough,

$$(U^{\eta,\epsilon}_1, V^{\eta,\epsilon}_1, U^{\eta,\epsilon}_2, V^{\eta,\epsilon}_2, \cdots) \Rightarrow (U_1, V_2, U_2, V_2, \cdots) \qquad \text{as } \eta\downarrow 0, \tag{5.23}$$

where the laws of $U_j$'s and $V_j$'s are defined in (5.16). Specifically, we only consider $\epsilon > 0$ small enough such that claim (5.23) holds. In light of Lemma 5.5 and Proposition D.1, (5.23) immediately leads to the claims in $(i)$. Regarding claim $(ii)$, note that $\hat{X}^{\eta,\epsilon,|b}_t$ is a step function (i.e., piece-wise constant) that only takes values in $\mathcal{M} \triangleq \{m_j : \ j = 1, 2, \cdots, n_{\min}\}$, which is a finite set. Let

$$A_N \triangleq \{\xi \in \mathbb{D}[0, T] : \ \xi \text{ is a step function with at most } N \text{ jumps and only takes values in } \mathcal{M}\}.$$

First, the finite-dimensional nature of $A_N$ (i.e., at most $N$ jumps over $[0, T]$, only $n_{\min}$ possible values) implies that $A_N$ is a compact set in $(\mathbb{D}[0, T], \boldsymbol{d}^{[0,T]}_{L_p})$. Besides,

$$\limsup_{n\to\infty} \mathbf{P}(\hat{X}^{\eta_n,\epsilon,|b}_{\cdot} \notin A_N) = \limsup_{n\to\infty} \mathbf{P}(\sum_{j=1}^{N+1} U^{\eta_n,\epsilon}_j \leq T) \leq \mathbf{P}(\sum_{j=1}^{N+1} U_j \leq T),$$

where the last inequality follows from $(U^{\eta_n,\epsilon}_1, \cdots, U^{\eta_n,\epsilon}_N) \Rightarrow (U_1, \cdots, U_N)$. Using part $(i)$ of Proposition D.1, we have $\lim_{N\to\infty} \limsup_{n\to\infty} \mathbf{P}(\hat{X}^{\eta_n,\epsilon,|b}_{\cdot} \notin A_N) = 0$, which verifies the tightness of $\hat{X}^{\eta_n,\epsilon,|b}_{\cdot}$.

Now, it only remains to prove (5.23). This is equivalent to proving that, for each $N \geq 1$, $(U^{\eta,\epsilon}_1, V^{\eta,\epsilon}_1, \cdots, U^{\eta,\epsilon}_N, V^{\eta,\epsilon}_N)$ converges in distribution to $(U_1, V_1, \cdots, U_N, V_N)$ as $\eta \downarrow 0$. Fix some $N = 1, 2, \cdots$. First, note that $U_1 = 0$ and $V_1 = m_i$. From part $(i)$ of Lemma 5.9, we get $(U^{\eta,\epsilon}_1, V^{\eta,\epsilon}_1) \Rightarrow (0, m_i) = (U_1, V_1)$ 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 any $t > 0$, $i, j \in [n_{\min}]$ with $i \neq j$, it follows directly from part $(ii)$ of Lemma 5.9 that

$$\begin{aligned}
&\lim_{\eta\downarrow 0} \mathbf{P}\Big(U^{\eta,\epsilon}_{n+1} \leq t, \ V^{\eta,\epsilon}_{n+1} = m_j \ \Big| \ V^{\eta,\epsilon}_n = m_i, \ V^{\eta,\epsilon}_l = v_l \ \forall l \in [n-1], \ U^{\eta,\epsilon}_l \leq t_l \ \forall l \in [n]\Big) \\
&= \begin{cases} \frac{q_b(i,j)}{q_b(i)} & \text{if } m_i \notin V^*_b, \\ \frac{q_b(i,j)}{q_b(i)} \cdot \Big(1 - \exp\big(-q_b(i)t\big)\Big) & \text{if } m_i \in V^*_b. \end{cases}
\end{aligned}$$

This coincides with the conditional law of $\mathbf{P}\Big(U_{n+1} < t, \ V_{n+1} = m_j \ \Big| \ V_n = m_i, \ (V_j)^{n-1}_{j=1}, \ (U_j)^n_{j=1}\Big)$ specified in (5.16). By arguing inductively, we conclude the proof. □

Moving onto the proof of Proposition 5.7, we first prepare a lemma that establishes the weak convergence from $X^{\eta|b}_{\lfloor \cdot/\lambda^*_b(\eta)\rfloor}(x)$ to $\hat{X}^{\eta,\epsilon|b}_{\cdot}(x)$ in terms of finite dimensional distributions.

**Lemma 5.10.** *Let Assumptions 1, 2, 3, 4, 6, and 7 hold. Given any $t > 0$ and $x \in \bigcup_{i\in[n_{min}]} I_i$,*

*(i)* $\lim_{\eta\downarrow 0} \mathbf{P}\Big(X^{\eta|b}_j(x) \notin (-M, M) \text{ for some } j \leq t/\lambda^*_b(\eta)\Big) = 0$ *for the constant $M > 0$ specified in Lemma 5.8;*

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

*Proof.* Throughout this proof, let $\bar{\epsilon}$ be specified as in (5.13).

$(i)$ We prove a stronger result. Let $I_{M,\delta} = (-M, M)\backslash S(\delta)$ where $S(\delta) = \bigcup_{i\in[n_{\min}-1]}[s_i - \delta, s_i + \delta]$. Recall the definition of $\hat{\tau}^{\eta,\epsilon|b}_j(x)$ in (5.8) and (5.10). For any $N \in \mathbb{Z}_+$, on event

$$\Big(\bigcap_{k=1}^{N-1} \underbrace{\Big\{X^{\eta|b}_j(x) \in I_{M,\delta} \ \forall j \in \big[\hat{\tau}^{\eta,\epsilon|b}_k(x), \hat{\tau}^{\eta,\epsilon|b}_{k+1}(x)\big]\Big\}}_{A_k(\eta,\delta)}\Big) \cap \underbrace{\Big\{\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq t/\lambda^*_b(\eta)\Big\}}_{B_1(\eta)} \cap \underbrace{\Big\{\hat{\tau}^{\eta,\epsilon|b}_N > t/\lambda^*_b(\eta)\Big\}}_{B_2(\eta)}$$

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 $\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq t/\lambda^*_b(\eta) < \hat{\tau}^{\eta,\epsilon|b}_N(x)$. Therefore, it suffices to show that for any $\Delta > 0$, there are some positive integer $N$ and $\delta > 0$ such that

$$\limsup_{\eta\downarrow 0} \big[\mathbf{P}\big(B^c_1(\eta)\big) + \mathbf{P}\big(B^c_2(\eta)\big) + \sum_{k=1}^{N-1} \mathbf{P}\big(A^c_k(\eta, \delta)\big)\big] < \Delta. \tag{5.24}$$

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 $\bigcup_{l \in [n_{\min}]}(m_l - \epsilon, m_l + \epsilon)$, we have $\hat{\tau}^{\eta,\epsilon|b}_1(x) \leq R^{\eta|b}_{i;\epsilon}(x)$ and hence

$$\begin{aligned}\limsup_{\eta \downarrow 0} \mathbf{P}\Big(B^c_1(\eta)\Big) &\leq \limsup_{\eta \downarrow 0} \mathbf{P}\Big(\hat{\tau}^{\eta,\epsilon|b}_1(x) > t/\lambda^*_b(\eta)\Big) \leq \limsup_{\eta \downarrow 0} \mathbf{P}\Big(\lambda^*_b(\eta) \cdot R^{\eta|b}_{i;\epsilon}(x) > t\Big) \\ &= 0 \qquad \text{using Lemma 5.9 } (i).\end{aligned} \tag{5.25}$$

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 important properties summarized in Section D. In particular, we can fix some $N$ large enough such that $\mathbf{P}(U_1 + \cdots + U_N \leq t) < \Delta/2$. From part $(i)$ of Proposition 5.6, we now get

$$\begin{aligned}\limsup_{\eta \downarrow 0} \mathbf{P}\Big(B^c_2(\eta)\Big) &\leq \limsup_{\eta \downarrow 0} \mathbf{P}\bigg(\sum_{n=1}^{N}\big(\tau^{\eta,\epsilon|b}_n(x) - \tau^{\eta,\epsilon|b}_{n-1}(x)\big) \cdot \lambda^*_b(\eta) \leq t\bigg) \\ &\leq \mathbf{P}(U_1 + \cdots + U_N \leq t) < \Delta/2.\end{aligned} \tag{5.26}$$

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 $\epsilon$-neighborhood of some $m_l$ that is different from $m_k$); also, for all $j \geq 2$, $\hat{\tau}^{\eta,\epsilon|b}_j(x)$ is the first time since $\hat{\tau}^{\eta,\epsilon|b}_{j-1}(x)$ that $X^{\eta|b}_{\cdot}(x)$ visits the $\epsilon$-neighborhood of some $m_l$ that is different from the one visited at $\hat{\tau}^{\eta,\epsilon|b}_{j-1}(x)$. From the strong Markov property at $\hat{\tau}^{\eta,\epsilon|b}_k(x)$,

$$\sup_{k \geq 1} \mathbf{P}\Big(A^c_k(\eta)\Big) \leq \max_{l \in [n_{\min}]} \sup_{y \in [m_l - \epsilon, m_l + \epsilon]} \mathbf{P}\Big(\exists j < \sigma^{\eta|b}_{l;\epsilon}(y)\ s.t.\ X^{\eta|b}_j(y) \in S(\delta) \text{ or } \big|X^{\eta|b}_j(y)\big| \geq M\Big).$$

Applying Lemma 5.8, 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 (5.25) and (5.26), we finish the proof of (5.24).

$(ii)$ 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 (see (5.8)–(5.11) 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 $\epsilon \in (0, \bar{\epsilon})$

$$\lim_{\eta \downarrow 0} \mathbf{P}\Big(X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x) \in \bigcup_{l \in [n_{\min}]}(m_l - \epsilon, m_l + \epsilon)\Big) = 1.$$

Pick some $\delta_t \in (0, \frac{t}{3})$, $\delta > 0$. Recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$, and define event

$$(\mathrm{I}) = \Big\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor - \lfloor 2\delta_t/H(\eta^{-1})\rfloor}(x) \in I_{M,\delta}\Big\}.$$

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, define event

$$(\mathrm{II}) = \Big\{R^\eta - t_1(\eta) \leq \delta_t/H(\eta^{-1})\Big\}.$$

On event (I) $\cap$ (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

$$(\mathrm{III}) = \Big\{\tau^\eta - R^\eta > 2\delta_t/H(\eta^{-1})\Big\}.$$

On event (I) $\cap$ (II) $\cap$ (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 $\delta_t \in (0, \frac{t}{3})$ and $\delta > 0$ such that

$$\liminf_{\eta \downarrow 0} \mathbf{P}\Big( \text{(I)} \Big) \geq 1 - \Delta, \tag{5.27}$$

$$\liminf_{\eta \downarrow 0} \mathbf{P}\Big( \text{(II)} \ \Big| \ \text{(I)} \Big) \geq 1, \tag{5.28}$$

$$\liminf_{\eta \downarrow 0} \mathbf{P}\Big( \text{(III)} \ \Big| \ \text{(I)} \cap \text{(II)} \Big) \geq 1 - \Delta. \tag{5.29}$$

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 (5.27) (5.28) (5.29). Throughout the remainer of this proof, we fix some $\epsilon \in (0, \bar{\epsilon})$ and $\Delta > 0$.

**Proof of** (5.27). This has been established in the proof for part $(i)$.

**Proof of** (5.28). 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 $\eta$ 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

$$\begin{aligned}
\mathbf{P}\Big( \text{(II)}^c \ \Big| \ \text{(I)} \Big) &\leq \max_{k \in [n_{\min}]} \sup_{y \in I_{k;\delta,M}} \mathbf{P}\bigg( X^{\eta|b}_j(y) \notin \bigcup_{l \in [n_{\min}]} (m_l - \frac{\epsilon}{2}, m_l + \frac{\epsilon}{2}) \ \forall j \leq \delta_t/H(\eta^{-1}) \bigg) \\
&\leq \max_{k \in [n_{\min}]} \sup_{y \in I_{k;\delta,M}} \mathbf{P}\Big( R^{\eta|b}_{k;\epsilon/2}(y) > \delta_t/H(\eta^{-1}) \Big) \\
&\leq \max_{k \in [n_{\min}]} \sup_{y \in I_{k;\delta,M}} \mathbf{P}\Big( R^{\eta|b}_{k;\epsilon/2}(y) > T/\eta \Big)
\end{aligned}$$

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 6, $\boldsymbol{t}_k(x, \frac{\epsilon}{4}) < \infty$ 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 4.5 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 (5.28).

**Proof of** (5.29). We prove the claim for all $\delta_t$ small enough. By strong Markov property at $R^\eta$,

$$\mathbf{P}\Big( \text{(III)}^c \ \Big| \ \text{(I)} \cap \text{(II)} \Big) \leq \max_{k \in [n_{\min}]} \sup_{y \in [m_k - \epsilon/2, m_k + \epsilon/2]} \mathbf{P}\Big( \exists j \leq \frac{2\delta_t}{H(\eta^{-1})} \ s.t. \ X^{\eta|b}_j(y) \notin (m_k - \epsilon, m_k + \epsilon) \Big).$$

Also, note that $\epsilon < \bar{\epsilon} < b$; see (5.13). 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 $\delta_t$ 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 (5.29). $\square$

The next result provides a bound over the proportion of time that $X^{\eta|b}_j(x)$ is not close enough to a local minimum.

**Lemma 5.11.** *Let Assumptions 1, 2, 3, 4, 6, and 7 hold. Given any $\epsilon \in (0,\bar{\epsilon})$, it holds for all $t \in (0,1)$ small enough that*

$$\limsup_{\eta \downarrow 0} \max_{i:\ m_i \in V^*_b} \sup_{x \in (m_i - \frac{\epsilon}{2}, m_i + \frac{\epsilon}{2})} \mathbf{P}\bigg( \int_0^t \mathbb{I}\Big\{ X^{\eta|b}_{\lfloor s/\lambda^*_b(\eta) \rfloor}(x) \notin (m_i - \epsilon, m_i + \epsilon) \Big\} ds > t^2 \bigg) < q^* t,$$

*where $q^* \in (0,\infty)$ is a constant that does not vary with $t$.*

*Proof.* There are only finitely many elements in $V^*_b$. Therefore, it suffices to fix some $m_i \in V^*_b$ (recall that $I_i = (s_{i-1}, s_i)$ is the attraction field associated with $m_i$, and w.l.o.g. we assume $m_i = 0$) and prove that

$$\limsup_{\eta \downarrow 0} \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\bigg( \int_0^t \mathbb{I}\Big\{ X^{\eta|b}_{\lfloor s/\lambda^*_b(\eta) \rfloor}(x) \notin (-\epsilon, \epsilon) \Big\} ds > t^2 \bigg) < q^* t \tag{5.30}$$

holds for all $t > 0$ small enough, where $q^* \in (0,\infty)$ is a constant that does not vary with $\epsilon, \Delta$, or $t$.

Let $T^\eta_0 = 0$, and (for all $i \geq 1$)

$$S^\eta_i \triangleq \min\{ j > T^\eta_{i-1} :\ X^{\eta|b}_j(x) \notin (-\epsilon, \epsilon) \}, \qquad T^\eta_i \triangleq \min\{ j > S^\eta_i :\ X^{\eta|b}_j(x) \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2}) \}.$$

Note that if $X^{\eta|b}_j(x) \notin (-\epsilon,\epsilon)$, then there is some $i \geq 1$ such that $j \in [S^\eta_i, T^\eta_i - 1]$. Next, let $N^\eta \triangleq \max\{ i \geq 0 :\ S^\eta_i \leq t/\lambda^*_b(\eta)\}$, and note that $\#\{ j \leq \lfloor t/\lambda^*_b(\eta) \rfloor :\ X^{\eta|b}_j(x) \notin (-\epsilon,\epsilon)\} \leq \sum_{i=1}^{N^\eta} T^\eta_i - S^\eta_i$.

Now, recall that $\alpha > 1$ is the heavy-tailed index in Assumption 1, and the scale function $\lambda^*_b(\eta)$ is defined in (2.38) with $\lambda^*_b(\eta) \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)+1}(\eta)$. Let $\beta \in (0, \alpha - 1)$ and $k(\eta) = 1/\eta^{(J^*_b(V)-1)(\alpha-1)+\beta}$. For any $x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})$, define events

$$A^\eta_{t,M}(x) \triangleq \big\{ X^{\eta|b}_j(x) \in (s_{i-1} + \frac{\epsilon}{2}, s_i - \frac{\epsilon}{2}) \cap (-M, M) \text{ for all } j \leq \lfloor t/\lambda^*_b(\eta) \rfloor \big\},$$
$$B^\eta_{t,\delta}(x) \triangleq \big\{ \text{for each } i \leq k(\eta), \exists j \in [T^\eta_{i-1} + 1, S^\eta_i] \text{ s.t. } \eta|Z_j| > \delta \big\}.$$

On $B^\eta_{t,\delta}(x)$, we must have $N^\eta \leq \#\{ j \leq \lfloor t/\lambda^*_b(\eta) \rfloor :\ \eta|Z_j| > \delta\}$. Furthermore, given some constant $T \in (0,\infty)$, let $E^\eta_{t,T}(x) \triangleq \{ T^\eta_i \wedge \lfloor t/\lambda^*_b(\eta) \rfloor - S^\eta_i \leq T/\eta\ \forall i \geq 1\}$. On event $B^\eta_{t,\delta}(x) \cap E^\eta_{t,T}(x)$ we get

$$\begin{aligned} \#\{ j \leq \lfloor t/\lambda^*_b(\eta) \rfloor :\ X^{\eta|b}_j(x) \notin (-\epsilon,\epsilon)\} &\leq \sum_{i=1}^{N^\eta} T^\eta_i \wedge \lfloor t/\lambda^*_b(\eta) \rfloor - S^\eta_i \\ &\leq k(\eta) \cdot T/\eta = T/\eta^{1+\beta+(J^*_b(V)-1)(\alpha-1)}, \end{aligned}$$

and hence

$$\int_0^t \mathbb{I}\Big\{ X^{\eta|b}_{\lfloor s/\lambda^*_b(\eta) \rfloor}(x) \notin (-\epsilon,\epsilon) \Big\} ds \leq \frac{T/\eta^{1+\beta+(J^*_b(V)-1)(\alpha-1)} + 1}{\lfloor t/\lambda^*_b(\eta) \rfloor}.$$

However, due to $\lambda^*_b(\eta) \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)+1}(\eta)$ and $J^*_b(V)\cdot(\alpha-1)+1 = (J^*_b(V)-1)\cdot(\alpha-1)+\alpha > (J^*_b(V)-1)\cdot(\alpha-1)+1+\beta$, we have (for any $t, T > 0$ and $\beta \in (0,\alpha-1)$)

$$\lim_{\eta \downarrow 0} \frac{T/\eta^{1+\beta+(J^*_b(V)-1)(\alpha-1)} + 1}{\lfloor t/\lambda^*_b(\eta) \rfloor} = 0.$$

The discussion above implies the following: to prove (5.30), it suffices to find some $t, T, M, \delta \in (0,\infty)$ such that

$$\limsup_{\eta \downarrow 0} \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\Big( (A^\eta_{t,M}(x))^c \Big) < q^* t, \tag{5.31}$$

$$\lim_{\eta \downarrow 0} \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\Big( \big(B^\eta_{t,\delta}(x)\big)^c \Big) = 0, \tag{5.32}$$

$$\lim_{\eta \downarrow 0} \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\Big( A^\eta_{t,M}(x) \cap B^\eta_{t,\delta}(x) \cap \big(E^\eta_{t,T}(x)\big)^c \Big) = 0. \tag{5.33}$$

In particular, $q^* \in (0,\infty)$ is a constant that does not vary with $\epsilon, \Delta, M, \delta$, or $t$.

**Proof of** (5.31). This follows immediately from the first exit time analysis. Specifically, recall that we have assumed w.l.o.g. that the local minimum $m_i \in V^*_b$ at hand is located at the origin, i.e., $m_i = 0$. This implies $\mathcal{J}^*_b(V) = \lceil \min\{|s_{i-1}|, s_i\}/b \rceil$; that is, starting from the local minimum, it requires at least $\mathcal{J}^*_b(V)$ jumps (each bounded by $b$) to escape from the attraction field $(s_{i-1}, s_i)$. Furthermore, by our choice of $\bar{\epsilon}$ in (5.13) (which is essentially due to the assumption that $|s_j - m_i|/b \notin \mathbb{Z}$ for all $i \in [n_{\min}]$ and $j \in [n_{\min} - 1]$), it holds for all $\epsilon \in (0, \bar{\epsilon})$ that $\mathcal{J}^*_b(V) = \lceil \min\{|s_{i-1} + \frac{\epsilon}{2}|, s_i - \frac{\epsilon}{2}\}/b \rceil$. For any $M \in (0,\infty)$ large enough, we then have $\mathcal{J}^*_b(V) = \lceil \min\{|s_{i-1} + \frac{\epsilon}{2}|, s_i - \frac{\epsilon}{2}, M\}/b \rceil$, thus implying that, starting from the origin, it also requires at least $\mathcal{J}^*_b(V)$ jumps to escape from $(s_{i-1} + \frac{\epsilon}{2}, s_i - \frac{\epsilon}{2}) \cap (-M, M)$. By applying part $(a)$ of Theorem 2.6 onto $(s_{i-1} + \frac{\epsilon}{2}, s_i - \frac{\epsilon}{2}) \cap (-M, M)$, we can find $q \in (0,\infty)$ such that

$$\limsup_{\eta \downarrow 0} \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\Big( \big(A^\eta_{t,M}(x)\big)^c \Big) \leq 1 - \exp(-qt) \qquad \forall t > 0.$$

For all $t > 0$ small enough, we have $1 - \exp(-qt) \leq 2qt$. By picking $q^* = 2q$, we conclude the proof.

**Proof of** (5.32). By strong Markov property at $T^\eta_i$,

$$\sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\Big( \big(B^\eta_{t,\delta}(x)\big)^c \Big) \leq k(\eta) \cdot \sup_{y \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\big( X^{\eta|b}_j(y) \notin (-\epsilon, \epsilon) \text{ for some } j < \tau^{>\delta}_1(\eta) \big)$$

Here, $\tau^{>\delta}_1(\eta)$ is the stopping time defined in (3.2) as the first arrival time of some $\eta|Z_j| > \delta$. Applying Lemma 4.7, it holds for all $\delta > 0$ small enough that $\sup_{y \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\big( X^{\eta|b}_j(y) \notin (-\epsilon, \epsilon)$ for some $j < \tau^{>\delta}_1(\eta) \big) = \boldsymbol{o}\big(1/k(\eta)\big)$. This concludes the proof of claim (5.32).

**Proof of** (5.33). On $A^\eta_{t,M}(x) \cap B^\eta_{t,\delta}(x)$, we have $T^\eta_i \wedge \lfloor t/\lambda^*_b(\eta) \rfloor = \tilde{T}^\eta_i \wedge \lfloor t/\lambda^*_b(\eta) \rfloor$ for each $i \geq 1$, where

$$\tilde{T}^\eta_i \triangleq \min\big\{ j > S^\eta_i :\ X^{\eta|b}_j(x) \notin \big( (s_{i-1} + \frac{\epsilon}{2}, s_i - \frac{\epsilon}{2}) \cap (-M, M) \big) \setminus [-\frac{\epsilon}{2}, \frac{\epsilon}{2}] \big\}.$$

Furthermore, it has been shown above that, on $B^\eta_{t,\delta}(x)$ we have $N^\eta \leq k(\eta)$. Therefore,

$$\begin{aligned}
&\sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\big( A^\eta_{t,M}(x) \cap B^\eta_{t,\delta}(x) \cap \big(E^\eta_{t,T}(x)\big)^c \big) \\
&\leq \sup_{x \in (-\frac{\epsilon}{2}, \frac{\epsilon}{2})} \mathbf{P}\big( \tilde{T}^\eta_i - S^\eta_i > T/\eta \text{ for some } i \leq k(\eta) \big) \\
&\leq k(\eta) \cdot \underbrace{\sup_{y \in (-\epsilon, \epsilon)} \mathbf{P}\big( X^{\eta|b}_j(x) \in \big( (s_{i-1} + \frac{\epsilon}{2}, s_i - \frac{\epsilon}{2}) \cap (-M, M) \big) \setminus [-\frac{\epsilon}{2}, \frac{\epsilon}{2}] \text{ for all } j \leq \lfloor T/\eta \rfloor \big)}_{\triangleq p^*_T(\eta)}.
\end{aligned}$$

The last step is due to the strong Markov property at $\{S^\eta_i : i \in [k(\eta)]\}$. Applying Lemma 4.4, we can find $T$ large enough such that $p^*_T(\eta) = \boldsymbol{o}\big(1/k(\eta)\big)$ as $\eta \downarrow 0$ and complete the proof. □

Now, we are ready to prove Proposition 5.7.

*Proof of Proposition 5.7.* The claim $\lim_{\eta \downarrow 0} \mathbf{P}(|X^{\eta|b}_T(x) - \hat{X}^{\eta,\epsilon|b}_T(x)| \geq \epsilon) = 0$ has already been proved in part $(ii)$ of Lemma 5.10. In the remainder of this proof, we focus on establishing the claim

$\lim_{\eta\downarrow 0}\mathbf{P}\big(\boldsymbol{d}^{[0,T]}_{L_p}(X^{\eta|b}_{\lfloor\cdot/\lambda^*_b(\eta)\rfloor}(x),\hat{X}^{\eta,\epsilon|b}_{\cdot}(x))\geq 2\epsilon\big)=0$. For simplicity of notations, we focus on the case where $T=1$. Nevertheless, the proof below can be easily generalized for arbitrary $T>0$.

By definition of $\hat{X}^{\eta,\epsilon|b}_t(x)$, we have $|X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)-\hat{X}^{\eta,\epsilon|b}_t(x)|<\epsilon$ whenever $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)\in\bigcup_{i\in[n_{\min}]}(m_i-\epsilon,m_i+\epsilon)$. Now, we make a few observations. For any $\eta>0$ and any positive integer $N$, let $\mathcal{I}^{(\eta)}_N(n)\triangleq\mathbb{I}\big\{\boldsymbol{i}^{(\eta)}_N(n)>1/N^2\big\}$ where

$$\boldsymbol{i}^{(\eta)}_N(n)\triangleq\int_{n/N}^{(n+1)/N}\mathbb{I}\Big\{X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)\notin\bigcup_{i\in[n_{\min}]}(m_i-\epsilon,m_i+\epsilon)\Big\}dt\qquad\forall n=0,1,\cdots,N-1.$$

That is, $\boldsymbol{i}^{(\eta)}_N(n)$ is the amount of time over $[\frac{n}{N},\frac{n+1}{N})$ that $X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)$ is not close enough to any local minima, and $\mathcal{I}^{(\eta)}_N(n)$ is the indicator that $\boldsymbol{i}^{(\eta)}_N(n)>1/N^2$.

Let $K^{(\eta)}_N\triangleq\sum_{n=1}^{N-1}\mathcal{I}^{(\eta)}_N(n)$. The proof hinges on the following claims: there exist some $C\in(0,\infty)$, a family of events $A^\eta_N$, and some constant $M>0$ such that

- $(i)$ on $A^\eta_N$, we have $X^{\eta|b}_j(x)\in[-M,M]$ for all $j\leq\lfloor 1/\lambda^*_b(\eta)\rfloor$;
- $(ii)$ for all positive integer $N$ large enough, $\lim_{\eta\downarrow 0}\mathbf{P}(A^\eta_N)=1$;
- $(iii)$ for all positive integer $N$ large enough, there exists $\bar{\eta}=\bar{\eta}(N)>0$ such that under any $\eta\in(0,\bar{\eta})$,

$$\mathbf{P}(K^{(\eta)}_N\geq j\mid A^\eta_N)\leq\mathbf{P}\bigg(\text{Binom}(N,\frac{2C}{N})\geq j\bigg)\qquad\forall j=1,2,\cdots,N.$$

Here, $\text{Binom}(n,p)$ is the RV denoting the number of successful trials among $n$ independent Bernoulli trials, each with success rate $p$. W.l.o.g., in claim $(i)$ we can assume that $M$ is sufficiently large such that $x\in[-M,M]$ and $m_j\in[-M,M]$ for all $j\in[n_{\min}]$. As a result, we must have $\hat{X}^{\eta,\epsilon|b}_t(x)\in[-M,M]$ as well for all $t\geq 0$. To see how we apply these claims, let

$$\boldsymbol{d}^{(\eta)}_p(n)\triangleq\int_{n/N}^{(n+1)/N}\Big|X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)-\hat{X}^{\eta,\epsilon|b}_t(x)\Big|^p dt.$$

On event $A^\eta_N$, for any $n=0,1,\cdots,N-1$, if $\boldsymbol{i}^{(\eta)}_N(n)\leq 1/N^2$, we have $\boldsymbol{d}^{(\eta)}_p(n)\leq\epsilon^p\cdot\frac{1}{N}+(2M)^p\cdot\frac{1}{N^2}$; Otherwise, we have the trivial bound $\boldsymbol{d}^{(\eta)}_p(n)\leq(2M)^p\cdot\frac{1}{N}$. Therefore, on $A^\eta_N$,

$$\begin{aligned}\boldsymbol{\Delta}(\eta)&\triangleq\bigg(\boldsymbol{d}_{L_p}\Big(X^{\eta|b}_{\lfloor\cdot/\lambda^*_b(\eta)\rfloor}(x),\hat{X}^{\eta,\epsilon|b}_{\cdot}(x)\Big)\bigg)^p\\&=\sum_{n=0}^{N-1}\int_{n/N}^{(n+1)/N}\Big|X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)-\hat{X}^{\eta,\epsilon|b}_t(x)\Big|^p dt\\&\leq(2M)^p\cdot\frac{1}{N}+\sum_{n=1}^{N-1}\int_{n/N}^{(n+1)/N}\Big|X^{\eta|b}_{\lfloor t/\lambda^*_b(\eta)\rfloor}(x)-\hat{X}^{\eta,\epsilon|b}_t(x)\Big|^p dt\\&\leq(2M)^p\cdot\frac{1}{N}+K^{(\eta)}_N\cdot\frac{(2M)^p}{N}+(N-1-K^{(\eta)}_N)\cdot\Big(\frac{\epsilon^p}{N}+\frac{(2M)^p}{N^2}\Big)\leq(2M)^p\cdot\frac{1+K^{(\eta)}_N+\frac{1}{N}}{N}+\epsilon^p.\end{aligned}$$

Then, given any $N$ large enough, $\eta\in(0,\bar{\eta}(N))$ and any $\beta\in(0,1)$,

$$\boldsymbol{r}(\eta)\triangleq\mathbf{P}\bigg(\boldsymbol{\Delta}(\eta)\geq\underbrace{\frac{1+\frac{1}{N}+2C+\sqrt{N^\beta}}{N}}_{\triangleq\delta(N,\beta)}\cdot(2M)^p+\epsilon^p\bigg)$$

$$
\begin{aligned}
&\le \mathbf{P}(K_N^{(\eta)} \ge 2C + \sqrt{N^\beta}) = \mathbf{P}\big(\{K_N^{(\eta)} \ge 2C + \sqrt{N^\beta}\} \cap A_N^\eta\big) + \mathbf{P}\big(\{K_N^{(\eta)} \ge 2C + \sqrt{N^\beta}\} \setminus A_N^\eta\big) \\
&\le \mathbf{P}\bigg(\text{Binom}(N, \frac{2C}{N}) \ge 2C + \sqrt{N^\beta}\bigg) + \mathbf{P}\big((A_N^\eta)^c\big) \qquad \text{by claim } (iii) \\
&\le \frac{\text{var}\Big[\text{Binom}(N, \frac{2C}{N})\Big]}{N^\beta} + \mathbf{P}\big((A_N^\eta)^c\big) \le \frac{2C}{N^\beta} + \mathbf{P}\big((A_N^\eta)^c\big).
\end{aligned}
$$

Driving $\eta \downarrow 0$, it follows from claim $(ii)$ that $\limsup_{\eta\downarrow 0} \boldsymbol{r}(\eta) \le 2C/N^\beta$ for all $N$ large enough. Lastly, note that $C/N^\beta \to 0$ as $N \to \infty$; also, due to $\beta \in (0,1)$ we have $\lim_{N\to\infty} \delta(N,\beta) = 0$, and hence $\delta(N,\beta)\cdot(2M)^p + \epsilon^p < 2^p\epsilon^p$ eventually for all $N$ large enough. In summary, we get $\lim_{\eta\downarrow 0}\mathbf{P}(\boldsymbol{\Delta}(\eta) > 2^p\epsilon^p) = 0$ and conclude the proof. Now, it only remains to verify claims $(i)$, $(ii)$, and $(iii)$.

**Proof of Claims** $(i)$ **and** $(ii)$. We start by specifying events $A_N^\eta$. Let $t_N(n) = n/N$ and

$$
\begin{aligned}
&A_N^\eta(n) \\
&\triangleq \underbrace{\Big\{X^{\eta|b}_{\lfloor t_N(k)/\lambda_b^*(\eta)\rfloor}(x) \in \bigcup_{i:\ m_i \in V_b^*}\Big(m_i - \frac{\epsilon}{2}, m_i + \frac{\epsilon}{2}\Big)\ \forall k \in [n]\Big\}}_{\triangleq A^\eta_{N,1}(n)} \cap \underbrace{\Big\{X^{\eta|b}_{\lfloor t/\lambda_b^*(\eta)\rfloor}(x) \in [-M,M]\ \forall t \le t_N(k)\Big\}}_{\triangleq A^\eta_{N,2}(n)}
\end{aligned}
$$

and let $A_N^\eta = A_N^\eta(N)$. Note that $A_N^\eta(1) \supseteq A_N^\eta(2) \supseteq \cdots \supseteq A_N^\eta(N) = A_N^\eta$. Furthermore, $\{X^{\eta|b}_{\lfloor t/\lambda_b^*(\eta)\rfloor} :\ t > 0\} \stackrel{f.d.d.}{\to} \{Y_t^{*|b} :\ t > 0\}$ due to $\{\hat{X}_t^{\eta,\epsilon|b} :\ t > 0\} \stackrel{f.d.d.}{\to} \{Y_t^{*|b} :\ t > 0\}$ and $\lim_{\eta\downarrow 0}\mathbf{P}(|X_T^{\eta|b}(x) - \hat{X}_T^{\eta,\epsilon|b}(x)| \ge \epsilon) = 0$ for any $T > 0$; this is the content of Lemma 5.3. Next, by definition, $Y_t^{*|b}$ only visits states in $V_b^*$. Combining this fact with the weak convergence in f.d.d. we get $\lim_{\eta\downarrow 0}\mathbf{P}\big(A^\eta_{N,1}\big) = 1$ for any $N \ge 1$. On the other hand, part $(i)$ of Lemma 5.10 gives $\lim_{\eta\downarrow 0}\mathbf{P}\big(A^\eta_{N,2}\big) = 1\ \forall N \ge 1$ for any $M$ large enough. This verifies claims $(i)$ and $(ii)$.

**Proof of Claim** $(iii)$. Consider a random vector $\big(\widetilde{\mathcal{I}}_N^\eta(n)\big)_{n\in[N-1]}$ with law $\mathcal{L}\Big(\big(\mathcal{I}_N^\eta(n)\big)_{n\in[N-1]}\ \Big|\ A_N^\eta\Big)$. It suffices to find 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 any sequence $i_j \in \{0,1\}\ \forall j \in [n-1]$,

$$
\mathbf{P}\Big(\widetilde{\mathcal{I}}_N^\eta(n) = 1\ \Big|\ \widetilde{\mathcal{I}}_N^\eta(j) = i_j\ \forall j \in [n-1]\Big) < 2C/N \qquad \forall \eta \in (0,\bar\eta). \tag{5.34}
$$

To see why, under condition (5.34) and for any $\eta \in (0,\bar\eta(N))$, there exists a coupling between iid Bernoulli RVs $(\mathcal{Z}_N(n))_{n\in[N-1]}$ with success rate $2C/N$ and $(\widetilde{\mathcal{I}}_N^\eta(n))_{n\in[N-1]}$ such that $\widetilde{\mathcal{I}}_N^\eta(n) \le \mathcal{Z}_N(n)\ \forall n \in [N-1]$ almost surely. This stochastic dominance between $(\mathcal{Z}_N(n))_{n\in[N-1]}$ and $(\widetilde{\mathcal{I}}_N^\eta(n))_{n\in[N-1]}$ immediately verifies claim $(iii)$.

To prove condition (5.34) note that given any $N$, any $n \in [N-1]$, and any sequence $i_j \in \{0,1\}\ \forall j \in [n-1]$,

$$
\begin{aligned}
&\mathbf{P}\Big(\widetilde{\mathcal{I}}_N^\eta(n) = 1\ \Big|\ \widetilde{\mathcal{I}}_N^\eta(j) = i_j\ \forall j \in [n-1]\Big) \\
&= \frac{\mathbf{P}\Big(\widetilde{\mathcal{I}}_N^\eta(n) = 1;\ \widetilde{\mathcal{I}}_N^\eta(j) = i_j\ \forall j \in [n-1]\Big)}{\mathbf{P}\Big(\widetilde{\mathcal{I}}_N^\eta(j) = i_j\ \forall j \in [n-1]\Big)} \\
&= \frac{\mathbf{P}\Big(\big\{\mathcal{I}_N^\eta(n) = 1;\ \mathcal{I}_N^\eta(j) = i_j\ \forall j \in [n-1]\big\} \cap A_N^\eta\Big)\Big/\mathbf{P}(A_N^\eta)}{\mathbf{P}\Big(\big\{\mathcal{I}_N^\eta(j) = i_j\ \forall j \in [n-1]\big\} \cap A_N^\eta\Big)\Big/\mathbf{P}(A_N^\eta)} \qquad \text{by definition of } \big(\widetilde{\mathcal{I}}_N^\eta(n)\big)_{n\in[N-1]}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(n)=1;\ \mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)} \qquad \text{due to } A^\eta_N(n)\supseteq A^\eta_N\\
&= \frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(n)=1;\ \mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}\cdot\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)}\\
&= \underbrace{\mathbf{P}\Big(\mathcal{I}^\eta_N(n)=1\ \Big|\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}_{\triangleq p^\eta_1(N)}\cdot\underbrace{\frac{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N(n)\Big)}{\mathbf{P}\Big(\big\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\big\}\cap A^\eta_N\Big)}}_{\triangleq p^\eta_2(N)}.
\end{aligned}
$$

For term $p^\eta_1(N)$, note that on $A^\eta_N(n)$ we have $X^{\eta|b}_j(x)\in\bigcup_{i:\ m_i\in V^*_b}(m_i-\frac{\epsilon}{2},m_i+\frac{\epsilon}{2})$ at $j=\lfloor t_N(n)/\lambda^*_b(\eta)\rfloor$, and hence (using Markov property)

$$
p^\eta_1(N)\leq \max_{i:\ m_i\in V^*_b}\ \sup_{y\in(m_i-\frac{\epsilon}{2},m_i+\frac{\epsilon}{2})}\mathbf{P}\bigg(\int_0^{1/N}\mathbb{I}\Big\{X^{\eta|b}_{\lfloor s/\lambda^*_b(\eta)\rfloor}(y)\notin(m_i-\epsilon,m_i+\epsilon)\Big\}ds>1/N^2\bigg).
$$

Applying Lemma 5.11, for all $N$ large enough there exist $\bar\eta=\bar\eta(N)>0$, such that $p^\eta_1\leq C/N\ \forall\eta\in(0,\bar\eta)$, where $C\in(0,\infty)$ is independent of the value of $N$ and $\eta$. As for term $p^\eta_2$, note that for any event $B$ with $\mathbf{P}(B)>0$, we have

$$
\frac{\mathbf{P}(B\cap A^\eta_N(n))}{\mathbf{P}(B\cap A^\eta_N)}\leq\frac{\mathbf{P}(B)}{\mathbf{P}(B)-\mathbf{P}\big((A^\eta_N)^c\big)}\to 1 \qquad \text{as } \eta\downarrow 1 \text{ due to } \lim_{\eta\downarrow 0}\mathbf{P}(A^\eta_N)=1. \tag{5.35}
$$

In the definition of $p^\eta_2$, 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 $B=\{\mathcal{I}^\eta_N(j)=i_j\ \forall j\in[n-1]\}$ in (5.35), we can find some $\bar\eta=\bar\eta(N)$ such that $p^\eta_2<2\ \forall\eta\in(0,\bar\eta)$ uniformly for all those choices. Combining the bounds $p^\eta_1<C/N$ and $p^\eta_2<2$, we verify condition (5.34) and conclude the proof. □

## 5.3 Proof of Theorem 2.11

*Proof of Theorem 2.11.* For any $b>\max_{i\in[n_{\min}],\ j\in[n_{\min}-1]}|m_i-s_j|$, by definitions in (2.32) 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 $b$.

Pick some closed set $A\subseteq\mathbb{D}[0,T]$ (w.r.t. $L_p$ topology), and observe that

$$
\begin{aligned}
\mathbf{P}\Big(X^\eta_{\lfloor\cdot/H(\eta^{-1})\rfloor}(x)\in A\Big)&=\mathbf{P}\Big(X^\eta_{\lfloor\cdot/H(\eta^{-1})\rfloor}(x)\in A;\ X^{\eta|b}_j(x)=X^\eta_j(x)\ \forall j\leq\lfloor T/H(\eta^{-1})\rfloor\Big) \qquad (5.36)\\
&\quad+\mathbf{P}\Big(X^\eta_{\lfloor\cdot/H(\eta^{-1})\rfloor}(x)\in A;\ X^{\eta|b}_j(x)\neq X^\eta_j(x)\text{ for some } j\leq\lfloor T/H(\eta^{-1})\rfloor\Big)\\
&\leq\underbrace{\mathbf{P}\Big(X^{\eta|b}_{\lfloor\cdot/H(\eta^{-1})\rfloor}(x)\in A\Big)}_{\text{(I)}}+\underbrace{\mathbf{P}\Big(X^{\eta|b}_j(x)\neq X^\eta_j(x)\text{ for some } j\leq\lfloor T/H(\eta^{-1})\rfloor\Big)}_{\text{(II)}}.
\end{aligned}
$$

For term (I), it follows from Theorem 2.10 that $\limsup_{\eta\downarrow 0}\text{(I)}\leq\mathbf{P}\big(Y^{*|b}_\cdot(m_i)\in A\big)$. 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 $\{\eta|Z_j|\leq\frac{b}{2C}\ \forall j\leq\lfloor T/H(\eta^{-1})\rfloor\}$ 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 $j\leq\lfloor T/H(\eta^{-1})\rfloor$. Therefore, $X^{\eta|b}_j(x)$ and $X^\eta_j(x)$ coincide for such $j$'s. In other words, for any $\eta\in(0,\frac{b}{2C})$, we have $\{\eta|Z_j| \leq \frac{b}{2C}\ \forall j \leq \lfloor T/H(\eta^{-1})\rfloor\} \subseteq \{X_j^{\eta|b}(x) = X_j^{\eta}(x)\ \forall j \leq \lfloor T/H(\eta^{-1})\rfloor\}$. which leads to (recall that $H(\cdot) = \mathbf{P}(|Z_1| > \cdot)$)

$$\limsup_{\eta\downarrow 0}\,(\mathrm{II}) \leq \limsup_{\eta\downarrow 0} \mathbf{P}\Big(\exists j \leq \lfloor T/H(\eta^{-1})\rfloor\ s.t.\ \eta|Z_j| > \frac{b}{2C}\Big)$$
$$\leq \limsup_{\eta\downarrow 0} \frac{T}{H(\eta^{-1})}\cdot H(\eta^{-1}\cdot\frac{b}{2C}) = T\cdot\Big(\frac{2C}{b}\Big)^{\alpha} \qquad \text{due to } H(x) \in \mathcal{RV}_{-\alpha}(x).$$

In summary, $\limsup_{\eta\downarrow 0}\mathbf{P}\big(X^{\eta}_{\lfloor \cdot/H(\eta^{-1})\rfloor}(x) \in A\big) \leq \mathbf{P}\big(Y^{*|b}_{\cdot}(m_i) \in A\big) + T\cdot\big(\frac{2C}{b}\big)^{\alpha}$. 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 (2.40) and (2.42) 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 infinitesimal generator. Therefore, $\lim_{b\to\infty}\mathbf{P}\big(Y^{*|b}_{\cdot}(m_i) \in A\big) = \mathbf{P}\big(Y^*_{\cdot}(m_i) \in A\big)$. Together with the fact that $\lim_{b\to\infty}\big(\frac{2C}{b}\big)^{\alpha} = 0$, in (5.36) we obtain $\limsup_{\eta\downarrow 0}\mathbf{P}\big(X^{\eta}_{\lfloor\cdot/H(\eta^{-1})\rfloor}(x) \in A\big) \leq \mathbf{P}\big(Y^*_{\cdot}(m_i) \in A\big)$. From the arbitrariness of the closed set $A$, we conclude the proof with Portmanteau theorem. $\square$

# 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] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons, 2nd ed edition, 1999.

[4] 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.

[5] A. Bovier and F. Den Hollander. *Metastability: a potential-theoretic approach*, volume 351. Springer, 2016.

[6] 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.

[7] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability*, 36(4):1390 – 1420, 2008.

[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] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 1935.

[21] S. Foss, D. Korshunov, S. Zachary, et al. *An introduction to heavy-tailed and subexponential distributions*, volume 6. Springer, 2011.

[22] 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.

[23] 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.

[24] P. Imkeller and I. Pavlyukevich. Metastable behaviour of small noise lévy-driven diffusions. *ESAIM: PS*, 12:412–437, 2008.

[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] G. Kallianpur and J. Xiong. Stochastic differential equations in infinite dimensional spaces. *Lecture Notes-Monograph Series*, 26:iii–342, 1995.

[29] J. Kemeny and J. Snell. Finite markov chains. with a new appendix. In *Generalization of a fundamental matrix*. Springer, 1983.

[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*, 7:284–304, 1940.

[33] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[34] R. Kruse. *Strong and weak approximation of semilinear stochastic evolution equations*. Springer, 2014.

[35] 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.

[36] A. D. W. M. I. Freidlin. *Random Perturbations of Dynamical Systems*. Springer New York, NY, 1998.

[37] M. T. Mohan. Wentzell–freidlin large deviation principle for stochastic convective brinkman–forchheimer equations. *Journal of Mathematical Fluid Mechanics*, 23(3):62, 2021.

[38] 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.

[39] A. Nagaev. Limit theorems for large deviations where cramér's conditions are violated. *Fiz-Mat. Nauk*, 7:17–22, 1969.

[40] A. V. Nagaev. On a property of sums of independent random variables. *Theory of Probability & Its Applications*, 22(2):326–338, 1978.

[41] T. H. Nguyen, U. Simsekli, M. Gurbuzbalaban, and G. RICHARD. First exit time analysis of stochastic gradient descent under heavy-tailed gradient noise. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[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.

# A Results for Stochastic Differential Equations

In this section, we collect the results for stochastic differential equations driven by Lévy processes with regularly varying increments. Specifically, any one-dimensional Lévy process $\boldsymbol{L} = \{L_t : \ t \geq 0\}$ can be characterized by its generating triplet $(c_L, \sigma_L, \nu)$ 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 $\nu$ is the Lévy measure of the Lévy process $L_t$ characterizing the intensity of jumps in $L_t$. More precisely, we have the following Lévy–Itô decomposition

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

where $B$ is a standard Brownian motion, the measure $\nu$ 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 9.** $\mathbf{E}L_1 = 0$. *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 the 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 $\sigma$-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 $Y^\eta_t$ be the solution to SDE

$$Y^\eta_0(x) = x, \qquad dY^\eta_t(x) = a\big(Y^\eta_{t-}(x)\big)dt + \sigma\big(Y^\eta_{t-}(x)\big)d\bar{L}^\eta_t. \tag{A.2}$$

Below, we state the results regarding the sample-path large deviations, first exit times, and global dynamics of $Y^\eta_t$.

## A.1 Sample-Path Large Deviations

Recall the definitions of the mapping $h^{(k)}_{[0,T]}$ in (2.6)-(2.8) as well as the measure $\mathbf{C}^{(k)}_{[0,T]}(\ \cdot\ ; x)$ in (2.10). Also, recall the notion of uniform $\mathbb{M}$-convergence introduced in Definition 2.1. Define $\boldsymbol{Y}^\eta_{[0,T]}(x) = \{Y^\eta_t(x) : \ t \in [0,T]\}$ 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 $\boldsymbol{Y}^\eta_{[0,T]}(x)$ 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 A.1.** *Under Assumptions 2, 3, 4, and 9, 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 $X_j^{\eta|b}(x)$, we introduce processes $Y_t^{\eta|b}(x)$ that can be seen as a modulated version of $Y_t^\eta(x)$ where all jumps are truncated under the threshold value $b$. More generally, we consider the construction of a sequence of stochastic processes $\big(\boldsymbol{Y}_t^{\eta|b;(k)}(x;f,g)\big)_{k\geq 0}$ 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

$$dY_t^{\eta|b;(0)}(x;f,g) \triangleq f\big(Y_{t-}^{\eta|b;(0)}(x;f,g)\big)dt + g\big(Y_{t-}^{\eta|b;(0)}(x;f,g)\big)d\bar{L}_t^\eta \tag{A.3}$$

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

$$\tau_Y^{\eta|b;(1)}(x;f,g) \triangleq \min\Big\{t>0:\ \Big|g\big(Y_{t-}^{\eta|b;(0)}(x;f,g)\big)\cdot\Delta\bar{L}_t^\eta\Big| = \Big|\Delta Y_t^{\eta|b;(0)}(x;f,g)\Big| > b\Big\}, \tag{A.4}$$

$$W_Y^{\eta|b;(1)}(x;f,g) \triangleq \Delta Y^{\eta|b;(0)}_{\tau_Y^{\eta|b;(1)}(x;f,g)}(x;f,g) \tag{A.5}$$

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

$$Y^{\eta|b;(k)}_{\tau_Y^{\eta|b;(k)}(x;f,g)}(x;f,g) \triangleq Y^{\eta|b;(k)}_{\tau_Y^{\eta|b;(k)}(x;f,g)-}(x;f,g) + \varphi_b\Big(W_Y^{\eta|b;(k)}(x;f,g)\Big), \tag{A.6}$$

$$dY_t^{\eta|b;(k)}(x;f,g) \triangleq f\big(Y_{t-}^{\eta|b;(k)}(x;f,g)\big)dt + g\big(Y_{t-}^{\eta|b;(k)}(x;f,g)\big)d\bar{L}_t^\eta \quad \forall t>\tau_Y^{\eta|b;(k)}(x;f,g), \tag{A.7}$$

$$\tau_Y^{\eta|b;(k+1)}(x;f,g) \triangleq \min\Big\{t>\tau_Y^{\eta|b;(k)}(x;f,g):\ \Big|g\big(Y_{t-}^{\eta|b;(k)}(x;f,g)\big)\cdot\Delta\bar{L}_t^\eta\Big| > b\Big\}, \tag{A.8}$$

$$W_Y^{\eta|b;(k+1)}(x;f,g) \triangleq \Delta Y^{\eta|b;(k)}_{\tau_Y^{\eta|b;(k+1)}(x;f,g)}(x;f,g) \tag{A.9}$$

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

$$Y_t^{\eta|b}(x) \triangleq \sum_{k\geq 0} Y_t^{\eta|b;(k)}(x;a,\sigma)\cdot\mathbb{I}\Big\{t\in\Big[\tau_Y^{\eta|b;(k)}(x;a,\sigma),\tau_Y^{\eta|b;(k+1)}(x;a,\sigma)\Big)\Big\} \tag{A.10}$$

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

$$Y_t^{\eta|b}(x) = Y_t^{\eta|b;(k)}(x;a,\sigma) \qquad\Longleftrightarrow\qquad t\in\Big[\tau_Y^{\eta|b;(k)}(x;a,\sigma),\tau_Y^{\eta|b;(k+1)}(x;a,\sigma)\Big). \tag{A.11}$$

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 $Y_t^{\eta|b}(x)$. Once again, the proof is omitted as it closely resembles that of $X_j^{\eta|b}(x)$.

**Theorem A.2.** *Under Assumptions 2, 3, and 9, 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 $\mathbb{M}\big(\mathbb{D}[0,T]\setminus\mathbb{D}_A^{(k-1)|b}[0,T]\big)$ 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 $\mathbb{D}_A^{(k-1)|b}[0,T]$,*

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

To conclude this subsection, we present the conditional limit results for $\boldsymbol{Y}^\eta$ and $\boldsymbol{Y}^{\eta|b}$.

**Corollary A.3.** *Let Assumptions 2, 3, and 9 hold.*

*(i) For some $b > 0$, $k = 0, 1, 2, \cdots$, $x \in \mathbb{R}$, and measurable $B \subseteq \mathbb{D}$, suppose that $B$ is bounded away from $\mathbb{D}^{(k-1)|b}_{\{x\}}$, $B \cap \mathbb{D}^{(k)|b}_{\{x\}} \neq \emptyset$, and $\mathbf{C}^{(k)|b}(B^\circ) = \mathbf{C}^{(k)|b}(B^-) > 0$. Then*

$$\mathbf{P}\big(\boldsymbol{Y}^{\eta|b}(x) \in \cdot \,|\, \boldsymbol{Y}^{\eta|b}(x) \in B\big) \Rightarrow \frac{\mathbf{C}^{(k)|b}(\,\cdot \cap B; x)}{\mathbf{C}^{(k)|b}(B; x)}$$

*as $\eta \downarrow 0$.*

*(ii) Furthermore, suppose that Assumption 4 holds. For some $k = 0, 1, 2, \cdots$, $x \in \mathbb{R}$, and measurable $B \subseteq \mathbb{D}$, suppose that $B$ is bounded away from $\mathbb{D}^{(k-1)}_{\{x\}}$, $B \cap \mathbb{D}^{(k)}_{\{x\}} \neq \emptyset$, and $\mathbf{C}^{(k)}(B^\circ) = \mathbf{C}^{(k)}(B^-) > 0$. Then*

$$\mathbf{P}\big(\boldsymbol{Y}^{\eta}(x) \in \cdot \,|\, \boldsymbol{Y}^{\eta}(x) \in B\big) \Rightarrow \frac{\mathbf{C}^{(k)}(\,\cdot \cap B; x)}{\mathbf{C}^{(k)}(B; x)}$$

*as $\eta \downarrow 0$.*

## A.2 First Exit Time Analysis

Consider some open interval $I = (s_{\text{left}}, s_{\text{right}})$ where $s_{\text{left}} < 0 < s_{\text{right}}$. 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 $I = (s_{\text{left}}, s_{\text{right}})$. 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 (2.24) and $\check{\mathbf{C}}(\cdot)$ defined in (2.25). We omit the proof due to its similarity to that of Theorem 2.6.

**Theorem A.4.** *Let Assumptions 2, 3, 5, and 9 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 $B \subseteq I^c$,*

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

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

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

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

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

## A.3 Sample-Path Convergence of Global Dynamics

To conclude, we collect the sample-path convergence results for SDEs $Y_t^{\eta|b}(x)$ and $Y_t^{\eta}(x)$. We skip the proof as they are almost identical to those of $X_j^{\eta|b}(x)$ and $X_j^{\eta}(x)$.

Analogous to the treatment in Section 2.4, we set $a(\cdot) = -U'(\cdot)$ for some potential function $U : \mathbb{R} \to \mathbb{R}$ satisfying Assumption 6. Recall the definition of $V_b^*$ in (2.34) 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 $\nu$ is the Lévy measure of the Lévy process $L_t$. Define scale function

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

**Theorem A.5.** *Let Assumptions 2, 3, 6, 7, and 9 hold. Given any $i \in [n_{min}]$, $x \in I_i$, and $p \geq 1$,*

$$\{Y^{\eta|b}_{\lfloor t/\lambda^*_{b;L}(\eta)\rfloor}(x) :\ t > 0\} \stackrel{f.d.d.}{\to} \{Y_t^{*|b} :\ t > 0\} \quad \text{and} \quad Y^{\eta|b}_{\lfloor \cdot/\lambda^*_{b;L}(\eta)\rfloor}(x) \Rightarrow Y_\cdot^{*|b} \text{ in } (\mathbb{D}[0,\infty), \boldsymbol{d}_{L_p}^{[0,\infty)})$$

*as $\eta \downarrow 0$, where the Markov jump process $Y_t^{*|b}$ is characterized in Theorem 2.10.*

**Theorem A.6.** *Let Assumptions 2, 3, 4, 6, and 9 hold. Given any $i \in [n_{min}]$, $x \in I_i$, and $p \geq 1$,*

$$\{Y^{\eta}_{\lfloor t/\lambda_L(\eta)\rfloor}(x) :\ t > 0\} \stackrel{f.d.d.}{\to} \{Y_t^{*} :\ t > 0\} \quad \text{and} \quad Y^{\eta}_{\lfloor \cdot/\lambda_L(\eta)\rfloor}(x) \Rightarrow Y_\cdot^{*} \text{ in } (\mathbb{D}[0,\infty), \boldsymbol{d}_{L_p}^{[0,\infty)})$$

*as $\eta \downarrow 0$, where the Markov jump process $Y_t^{*}$ is characterized in Theorem 2.11.*

# B Properties of Mappings $h^{(k)}_{[0,T]}$ and $h^{(k)|b}_{[0,T]}$

In this section, we collect a few useful results about the mapping $h^{(k)}_{[0,T]}$ defined in (2.6)–(2.8) and $h^{(k)|b}_{[0,T]}$ defined in (2.16)–(2.18). In particular, we provide the proof of Lemmas 3.5–3.8.

For any $\xi \in \mathbb{D}$, let $\|\xi\| \triangleq \sup_{t\in[0,1]}|\xi(t)|$. Also, recall the definition of $\mathbb{D}_A^{(k)|b}$ in (2.19). Lemma B.1 shows that $\|\xi\|$ is uniformly bounded for all $\xi \in \mathbb{D}_A^{(k)|b}$.

**Lemma B.1.** *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 $\|\xi\| \leq M\ \forall \xi \in \mathbb{D}^{(k)|b}_{[u,v]}$.*

*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 $\xi^*$ and the upper bound $b$ on the jump size of $\xi$ 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 $M = A_{k+1} + \|\xi^*\| = (2\rho)^{k+1}N + \|\xi^*\|$. □

Recall the definitions of functions $a_M, \sigma_M$ in (3.31), mapping in $h^{(k)|b}_{M\downarrow}$ in (3.32)–(3.34), and sets $\mathbb{D}^{(k)|b}_{A;M\downarrow}$ in (3.35). Next, we present a corollary that follows directly from the boundedness of $\mathbb{D}^{(k)|b}_{A}$ shown in Lemma B.1.

**Corollary B.2.** *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 $M \geq M_0$*

- $\sup_{t \leq 1} |\xi(t)| \leq M_0 \ \forall \xi \in \mathbb{D}^{(k)|b}_{A} \cup \mathbb{D}^{(k)|b}_{A;M\downarrow}$;
- *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 $\xi \triangleq h^{(k)|b}(x_0, \boldsymbol{w}, \boldsymbol{t}) \in \mathbb{D}^{(k)|b}_{A} \subseteq \mathbb{D}^{(k)|b}_{[u,v]}$. Let $M_0 < \infty$ be the uniform upper bound associated with $\mathbb{D}^{(k)|b}_{[u,v]}$ in Lemma B.1: 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 $\mathbb{D}^{(k)|b}_{A;M\downarrow} = \mathbb{D}^{(k)|b}_{A}$ This concludes the proof. □

Now, we are ready to study the continuity of mappings $h^{(k)}$ and $h^{(k)|b}$.

**Lemma B.3.** *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

$$|x^* - x'| < \delta, \qquad |w'_j - w^*_j| \vee |t'_j - t_j| < \delta \ \forall j \in [k]. \tag{B.1}$$

In particular, by applying Corollary B.2 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 (B.1). 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 (3.31)). Let $C^* = \sup_{x \in [-M,M]} |a(x)| \vee \sigma(x) \vee 1$. Let $h^* = h^{(k)|b}_{M\downarrow}$, see (3.32)-(3.34). 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

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

Also, by picking $\delta > 0$ small enough, it is guaranteed that (under convention $t^*_0 = t'_0 = 0$, $t^*_{k+1} = t'_{k+1} = 1$)

$$\delta < \widetilde{\delta} \vee 1; \qquad \max_{j \in [k]} \left| \frac{t^*_{j+1} - t^*_j}{t'_{j+1} - t'_j} - 1 \right| < \widetilde{\delta} \ \forall \boldsymbol{t}' = (t'_1, \cdots, t'_k) \in (0,1)^{k\uparrow}, \ \max_{j \in [k]} |t'_j - t^*_j| < \delta. \tag{B.3}$$

Now it only remains to show that, under the current the choice of $\delta$, the bound $\boldsymbol{d}_{J_1}(\xi, \xi') < \epsilon$ follows from condition (B.1). To proceed, fix some $\xi'$ satisfying condition (B.1). 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

$$
\begin{aligned}
|\lambda(u) - u| &= \left| t^*_j + \frac{t^*_{j+1} - t^*_j}{t'_{j+1} - t'_j} \cdot (u - t'_j) - u \right| = \left| t^*_j + \frac{t^*_{j+1} - t^*_j}{t'_{j+1} - t'_j} \cdot v - (v + t'_j) \right| \quad \text{with } v \triangleq u - t'_j \\
&\le |t^*_j - t'_j| + \left| \frac{t^*_{j+1} - t^*_j}{t'_{j+1} - t'_j} - 1 \right| \cdot v \\
&\le \widetilde{\delta} + \widetilde{\delta} \cdot 1 < \epsilon. \qquad \text{(B.4)}
\end{aligned}
$$

In summary, $\sup_{u \in [0,1]} |\lambda(u) - u| < \epsilon$. Moving on, we show $\sup_{u \in [0,1]} \big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| < \epsilon$. with an inductive argument. First, Assumption 2 allows us to apply Gronwall's inequality and get $\sup_{u \in (0, t^*_1 \wedge t'_1)} \big|\xi^*(u) - \xi'(u)\big| \le \exp\big(D \cdot (t^*_1 \wedge t'_1)\big)|x^* - x'| \le \rho\delta$. As a result, for any $u \in (0, t^*_1 \wedge t'_1)$,

$$
\begin{aligned}
\big|\xi^*\big(\lambda(u)\big) - \xi'(u)\big| &= \left|\xi^*\left(\frac{t^*_1}{t'_1} \cdot u\right) - \xi'(u)\right| \le \left|\xi^*\left(\frac{t^*_1}{t'_1} \cdot u\right) - \xi^*(u)\right| + \big|\xi'(u) - \xi^*(u)\big| \\
&\le \left|\xi^*\left(\frac{t^*_1}{t'_1} \cdot u\right) - \xi^*(u)\right| + \rho\delta \\
&\le \sup_{x \in \mathbb{R}} |a^*(x)| \cdot \left|\frac{t^*_1}{t'_1} - 1\right| \cdot u + \rho\delta \qquad \text{due to } \xi^* = h^*(x^*, \boldsymbol{w}^*, \boldsymbol{t}^*) \\
&\le C^*\widetilde{\delta} + \rho\widetilde{\delta} = (C^* + \rho)\widetilde{\delta} \qquad \text{due to (B.3).}
\end{aligned}
$$

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 (B.3)(B.4),

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

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 $|\varphi_b(x) - \varphi_b(y)| \le |x - y|$,

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

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

$$
\begin{aligned}
\big|\xi^*\big(\lambda(t'_j+v)\big) - \xi'(t'_j+v)\big| &\leq \Big|\xi^*\big(\lambda(t'_j+v)\big) - \xi^*(t'_j+v)\Big| + \big|\xi^*(t'_j+v) - \xi'(t'_j+v)\big| \\
&\leq \Big|\xi^*\big(\lambda(t'_j+v)\big) - \xi^*(t'_j+v)\Big| + \rho R_j\widetilde{\delta} \qquad \text{Using Gronwall's inequality} \\
&\leq \sup_{x\in\mathbb{R}} |a^*(x)| \cdot |\lambda(t'_j+v) - (t'_j+v)| + \rho R_j \widetilde{\delta} \\
&\leq 2C^*\widetilde{\delta} + \rho R_j\widetilde{\delta} \qquad \text{due to (B.4).}
\end{aligned}
$$

Again, in case that $t'_{j+1} \leq t^*_{j+1}$, 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 (B.3)(B.4) to yield

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

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 $\Big|\xi^*(1) - \xi'(1)\Big| \leq \limsup_{t\uparrow 1}\Big|\xi^*\big(\lambda(t)\big) - \xi'(t)\Big| \leq \big(5C^* + \rho R_j\big)\widetilde{\delta} \leq R_{j+1}\widetilde{\delta}$. In case that $j \leq k$, using $|\varphi_b(x) - \varphi_b(y)| \leq |x-y|$,

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

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 (B.2). Combining this bound with (B.4), we get $\boldsymbol{d}_{J_1}(\xi^*, \xi') < \epsilon$ and conclude the proof. $\square$

**Lemma B.4.** *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

$$|x^* - x'| < \delta; \qquad |w_j' - w_j^*| \vee |t_j' - t_j^*| < \delta \ \forall j \in [k], \tag{B.5}$$

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 B.3. 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 (B.5), 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. □

Now, we move onto the proofs of Lemmas 3.5–3.8.

*Proof of Lemma 3.5.* 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

$$\bar{\delta} \triangleq \frac{\bar{\epsilon}}{\rho C + 1}. \tag{B.6}$$

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 (2.8), and construct $\xi'$ 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, $\xi'$ is driven by the same ODE as $\xi$ on $[t_J, t_{J+1})$, except that at the beginning of the intervals, $\xi'$ starts from $\xi(t_J-)$ instead of $\xi(t_J)$. On the other hand, $\xi'$ coincides with $\xi$ outside of $[t_J, t_{J+1})$. To see how close $\xi$ and $\xi'$ 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

$$\begin{aligned} |\xi(s) - \xi'(s)| &\leq \exp\big((t_{J+1} - t_J)D\big)\big|\xi(t_J) - \xi'(t_J-)\big| \\ &\leq \rho\big|\xi(t_J) - \xi(t_J-)\big| \\ &\leq \rho C\bar{\delta} < \bar{\epsilon}, \end{aligned} \tag{B.7}$$

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 $\xi'$ 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 $s \in [t_k, 1]$, □

Before establishing Lemma 3.6, we make one observation related to Assumption 8 and the truncation operator $\varphi_b$ defined in (2.15). 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 $z_1, z_2 \geq c$)

$$|\varphi_b(z_1 \cdot w_1) - \varphi_b(z_2 \cdot w_2)| \leq |z_1 \widetilde{w}_1 - z_2 \widetilde{w}_2| \qquad \text{where } \widetilde{w}_1 = \varphi_{b/c}(w_1),\ \widetilde{w}_2 = \varphi_{b/c}(w_2). \tag{B.8}$$

*Proof of Lemma 3.6.* The same arguments in Lemma 3.5 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

$$\xi = h^{(k)|b}(x, \boldsymbol{w}, \boldsymbol{t}),\ \xi' = h^{(k)|b+\epsilon}(x, \boldsymbol{w}, \boldsymbol{t}) \quad \Longrightarrow \quad \boldsymbol{d}_{J_1}(\xi, \xi') \leq \Big[2\rho\Big(1 + \frac{bD}{c}\Big)\Big]^k \epsilon \tag{B.9}$$

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 (B.9) 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 (B.9), 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 (B.9). 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

$$\sup_{t \in [t_j, t_{j+1})} \big|\xi(t) - \xi'(t)\big| \leq \exp\big(D(t_{j+1} - t_j)\big)\big|\xi(t_j) - \xi'(t_j)\big| \leq \rho R_j. \tag{B.10}$$

Lastly, at $t = t_{j+1}$, if $j = k$ (so $t_{j+1} = 1$), the continuity of $\xi, \xi'$ 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 $z_* \triangleq \xi(t_{j+1}-), z'_* \triangleq \xi'(t_{j+1}-)$)

$$\begin{aligned}
&\big|\xi(t_{j+1}) - \xi'(t_{j+1})\big| \\
&= \big|z_* + \varphi_b\big(\sigma(z_*)w_{j+1}\big) - \big[z'_* + \varphi_{b+\epsilon}\big(\sigma(z'_*)w_{j+1}\big)\big]\big| \\
&\leq \big|z_* - z'_*\big| + \big|\varphi_b\big(\sigma(z_*)w_{j+1}\big) - \varphi_b\big(\sigma(z'_*)w_{j+1}\big)\big| + \big|\varphi_b\big(\sigma(z'_*)w_{j+1}\big) - \varphi_{b+\epsilon}\big(\sigma(z'_*)w_{j+1}\big)\big| \\
&\leq \big|z_* - z'_*\big| + \big|\varphi_b\big(\sigma(z_*)w_{j+1}\big) - \varphi_b\big(\sigma(z'_*)w_{j+1}\big)\big| + \epsilon \\
&\leq \big|z_* - z'_*\big| + |\sigma(z_*) - \sigma(z'_*)| \cdot |\varphi_{b/c}(w_{j+1})| + \epsilon \qquad \text{using (B.8)}
\end{aligned}$$

$$
\begin{aligned}
&\leq \left|z_* - z_*'\right| + D \cdot \left|z_* - z_*'\right| \cdot (b/c) + \epsilon \qquad \text{due to Lipschitz continuity of } \sigma\text{; see Assumption 2} \\
&= (1 + \frac{bD}{c})|z_* - z_*'| + \epsilon \leq (1 + \frac{bD}{c})\rho R_j + \epsilon \qquad \text{due to (B.10)} \\
&\leq \rho\big(1 + \frac{bD}{c}\big) \cdot \big[2\rho\big(1 + \frac{bD}{c}\big)\big]^{j-1}\epsilon + \epsilon \\
&\leq \big[2\rho\big(1 + \frac{bD}{c}\big)\big]^{j}\epsilon.
\end{aligned}
$$

The proof to (B.9) can be completed by arguing inductively for $j = 1, 2, \cdots, k$. □

Before proving Lemma 3.7, we develop the following tool. Let $\boldsymbol{x}^\eta_j(x)$ be the solution to

$$
\boldsymbol{x}^\eta_0(x) = x, \qquad \boldsymbol{x}^\eta_j(x) = \boldsymbol{x}^\eta_{j-1}(x) + \eta a\big(\boldsymbol{x}^\eta_{j-1}(x)\big) \quad \forall j \geq 1. \tag{B.11}
$$

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

**Lemma B.5.** *Let Assumptions 2 and 4 hold. For any $\eta > 0, t > 0$ and $x, y \in \mathbb{R}$,*

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

*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 $\boldsymbol{x}^\eta_s(x) \triangleq \boldsymbol{x}^\eta_{\lfloor s \rfloor}(x)$. Also, we set $\boldsymbol{y}^\eta_s(y) \triangleq \boldsymbol{y}_{s\eta}(y)$ for any $s \geq 0$. Now observe that (for any $s \geq 0$)

$$
\begin{aligned}
\boldsymbol{y}^\eta_s(y) &= \boldsymbol{y}^\eta_{\lfloor s \rfloor}(y) + \eta\int_{\lfloor s \rfloor}^s a(\boldsymbol{y}^\eta_u(y))du \\
\boldsymbol{y}^\eta_{\lfloor s \rfloor}(y) &= y + \eta\int_0^{\lfloor s \rfloor} a(\boldsymbol{y}^\eta_u(y))du \\
\boldsymbol{x}^\eta_{\lfloor s \rfloor}(y) &= x + \eta\int_0^{\lfloor s \rfloor} a(\boldsymbol{x}^\eta_u(y))du.
\end{aligned}
$$

Let $b(u) \triangleq \boldsymbol{y}^\eta_u(y) - \boldsymbol{x}^\eta_u(x)$. 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 $s > 0$)

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

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

*Proof of Lemma 3.7.* (a) By definition of $\xi$, we have $\xi_t = \boldsymbol{y}_t(y) = h^{(0)}(y)(t)$ for any $t \in [0,1]$ with $t < \eta\tau^{>\delta}_1(\eta)$. Also, since $\tau^{>\delta}_1(\eta)$ only takes values in $\{1, 2, \cdots\}$, we know that $\eta\tau^{>\delta}_1(\eta) \leq 1 \iff \tau^{>\delta}_1(\eta) \leq \lfloor 1/\eta \rfloor$ and $\eta\tau^{>\delta}_1(\eta) > 1 \iff \tau^{>\delta}_1(\eta) > \lfloor 1/\eta \rfloor$.

Let $A \triangleq \Big\{\max_{i \leq \lfloor 1/\eta \rfloor \wedge \left(\tau^{>\delta}_1(\eta) - 1\right)} \eta\Big|\sum_{j=1}^i \sigma\big(X^\eta_{j-1}(x)\big)Z_j\Big| \leq \epsilon\Big\}$. Recall the definition of the deterministic process $\boldsymbol{x}^\eta$ defined in (B.11). Applying discrete version of Gronwall's inequality (see, for example, Lemma A.3 of [34]) we know that on event $A$,

$$
\Big|\boldsymbol{x}^\eta_j(x) - X^\eta_j(x)\Big| \leq \epsilon \cdot \exp(\eta D \cdot \lfloor 1/\eta \rfloor) \leq \rho\epsilon \qquad \forall j \leq \lfloor 1/\eta \rfloor \wedge \big(\tau^{>\delta}_1(\eta) - 1\big). \tag{B.12}
$$

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 B.5 we get

$$\sup_{t\in[0,1]:\ t<\eta\tau_1^{>\delta}(\eta)} \left|\xi_t - \boldsymbol{x}^{\eta}_{\lfloor t/\eta\rfloor}(x)\right| \leq \big(\eta C + |x-y|\big)\cdot\rho. \tag{B.13}$$

Therefore,

$$\sup_{t\in[0,1]:t<\eta\tau_1^{>\delta}(\eta)} \left|\xi_t - X^{\eta}_{\lfloor t/\eta\rfloor}(x)\right| \leq \rho\cdot\Big(\epsilon + |x-y| + \eta C\Big). \tag{B.14}$$

(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 (B.14) that $\sup_{t\in[0,1]:t<\eta\tau_1^{>\delta}(\eta)}\left|\xi_t - X^{\eta|b}_{\lfloor t/\eta\rfloor}(x)\right| \leq \rho\cdot\big(\epsilon + |x-y| + \eta C\big)$. 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)

$$\left|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\right| \leq \rho\cdot\big(\epsilon + |x-y| + \eta C\big). \tag{B.15}$$

Therefore,

$$\begin{aligned}
&\left|\xi\big(\eta\tau_1^{>\delta}(\eta)\big) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)}(x)\right| \\
&= \Bigg|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) + \varphi_b\bigg(\eta\sigma\Big(\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)\Big)W_1^{>\delta}(\eta)\bigg) \\
&\qquad - \bigg[X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x) + \varphi_b\bigg(\eta a\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big) + \eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\bigg)\bigg]\Bigg| \\
&\leq \left|\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y) - X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\right| \\
&\qquad + \underbrace{\left|\varphi_b\bigg(\eta\sigma\Big(\boldsymbol{y}(\eta\tau_1^{>\delta}(\eta)-;y)\Big)W_1^{>\delta}(\eta)\bigg) - \varphi_b\bigg(\eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\bigg)\right|}_{\triangleq I_1} \\
&\qquad + \underbrace{\left|\varphi_b\bigg(\eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\bigg) - \varphi_b\bigg(\eta a\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big) + \eta\sigma\Big(X^{\eta|b}_{\tau_1^{>\delta}(\eta)-1}(x)\Big)W_1^{>\delta}(\eta)\bigg)\right|}_{\triangleq I_2}.
\end{aligned}$$

Based on observation (B.8), 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 (B.15). 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,

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

$$\leq \Big(1 + \frac{bD}{c}\Big) \cdot \rho \cdot \big(\epsilon + |x-y| + 2\eta C\big).$$

This concludes the proof of part (b). $\square$

By applying Lemma 3.7 inductively, we can now establish Lemma 3.8.

*Proof of Lemma 3.8.* First of all, on $A_1(\eta, b, \epsilon, \delta, x)$, one can apply (3.22) of Lemma 3.7 and obtain

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

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 (3.23) of Lemma 3.7 to get

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

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

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

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 (3.23) of Lemma 3.7 again, we obtain

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

Arguing inductively, we yield $\sup_{t \in [0, \eta\tau_k^{>\delta}(\eta)]} \big|\xi_t - X^{\eta|b}_{\lfloor t/\eta \rfloor}(x)\big| \leq R_k = \big[3\rho \cdot (1 + \frac{bD}{c})\big]^k \epsilon$. Lastly, due to (3.21) of Lemma 3.7 and the fact that $\eta\tau_{k+1}^{>\delta}(\eta) > 1$,

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

This concludes the proof. $\square$

# C Properties of Measures $\check{\mathbf{C}}^{(k)|b}$

This section collects several important properties of measure $\check{\mathbf{C}}^{(k)|b}(\cdot)$ defined in (2.24). In particular, the proof of Lemma 4.2 will be provided at the end of this section.

Throughout this section, we impose Assumptions 1, 2, 3, and 5 on some $I = (s_{\text{left}}, s_{\text{right}})$ where $s_{\text{left}} < 0 < s_{\text{right}}$. Besides, we adopt the choices of $\bar{\epsilon} > 0$ and $\boldsymbol{t}(\epsilon)$ in (4.7) and (4.8) throughout this section.

Recall that $I^- = [s_{\text{left}}, s_{\text{right}}]$. Also, recall that $l = |s_{\text{left}}| \wedge s_{\text{right}}$ and $\mathcal{J}_b^* = \lceil l/b \rceil$. By studying the mapping $h^{(k)|b}_{[0,T]}$ defined in (2.16)-(2.18) in the current context, we establish useful properties for the mapping $\breve{g}^{(k)|b}$ defined in (2.23).

**Lemma C.1.** *Let Assumption 2 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*

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

*(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*

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

*(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*

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

*then*

- $\sup_{t \in [0, t_{\mathcal{J}_b^*-1})} |\xi(t)| \leq (\mathcal{J}_b^* - 1)b + \bar{\epsilon} < l - 2\bar{\epsilon}$;
- $|\xi(t_{\mathcal{J}_b^*-1})| \geq l - \bar{\epsilon}$;
- $\inf_{t \in [0, t_{\mathcal{J}_b^*-1}]} |\xi(t)| \geq \bar{\epsilon}$;
- $|w_j| > \bar{\delta}$ *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 $T > 0$, $\theta > l - \bar{\epsilon}$, $x \in [-\epsilon_0, \epsilon_0]$, $\boldsymbol{w} = (w_1, \cdots, w_{\mathcal{J}_b^*}) \in \mathbb{R}^{\mathcal{J}_b^*}$, 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 values 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 $\sup_{t\in[0,T]}|\xi(t)|<l=|s_{\text{left}}|\wedge s_{\text{right}}$ 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

$$\frac{d|\xi(t)|}{dt}=-\big|a\big(\xi(t)\big)\big|\quad\forall t\in[0,T]\setminus\{t_1,\cdots,t_j\}\tag{C.1}$$

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

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

This concludes the proof of part $(a)$.

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

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

$(c)$ We start from the claim that $\sup_{t\in[0,t_{\mathcal{J}^*_b-1})}|\xi(t)|<l-2\bar\epsilon$. 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 $\sup_{t\in[0,t_{\mathcal{J}^*_b-1})}|\xi(t)|=\sup_{t\in[0,t_{\mathcal{J}^*_b-1})}|\hat\xi(t)|\leq(\mathcal{J}^*_b-1)b+\bar\epsilon<l-2\bar\epsilon$.

Next, to see why the claim $|\xi(t_{\mathcal{J}^*_b-1})|\geq l-\bar\epsilon$ is true, note that $\sup_{t\in[0,T]}|\xi(t)|\geq l-\bar\epsilon$ and we have just shown that $\sup_{t\in[0,t_{\mathcal{J}^*_b-1})}|\xi(t)|<l-2\bar\epsilon$. Now consider the following proof by contradiction. Suppose that $|\xi(t_{\mathcal{J}^*_b-1})|<l-\bar\epsilon$. 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 (C.1), we yield the contradiction that $\sup_{t\in[t_{\mathcal{J}^*_b-1},T]}|\xi(t)|\leq|\xi(t_{\mathcal{J}^*_b-1})|\wedge\big(\sup_{t\in[0,t_{\mathcal{J}^*_b-1})}|\xi(t)|\big)<l-\bar\epsilon$. 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 (C.1) implies that

$$|\xi(t_{\mathcal{J}^*_b-1})|\leq|\xi(t)|+\sum_{s\in(t,t_{\mathcal{J}^*_b-1}]}|\Delta\xi(s)|$$

$$
\begin{aligned}
&\le \bar{\epsilon} + (\mathcal{J}_b^* - 1)b \qquad \text{due to truncation operators } \varphi_b \text{ in } h_{[0,T]}^{(\mathcal{J}_b^*-1)|b} \\
&< l - 2\bar{\epsilon} \qquad \text{due to (4.7).}
\end{aligned}
$$

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

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

$$
(\mathcal{J}_b^* - 1)b + 3\bar{\epsilon} + C\bar{\delta} < l,
$$

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 $|w_{j^*}| \le \bar{\delta}$. From observation (C.1), we get

$$
\begin{aligned}
|\xi(t_{\mathcal{J}_b^*-1})| &\le |\xi(0)| + \sum_{t\in[0,t_{\mathcal{J}_b^*-1}]} |\Delta\xi(t)| \\
&\le |x_0| + \varphi_b\Big(\Big|\sigma(x_0)\cdot w_0\Big|\Big) + \sum_{j=1}^{\mathcal{J}_b^*-1} \varphi_b\Big(\Big|\sigma\big(\xi(t_j-)\big)\cdot w_j\Big|\Big) \\
&\le \bar{\epsilon} + (\mathcal{J}_b^* - 1)b + C\bar{\delta} \qquad \text{due to } |x_0| \le \bar{\epsilon}, |w_{j^*}| \le \bar{\delta} \text{ and } |\sigma(y)| \le C \text{ for all } y \in \mathbb{R} \\
&< l - 2\bar{\epsilon} \qquad \text{due to our choice of } \bar{\delta}.
\end{aligned}
$$

This contradiction with the fact $|\xi(t_{\mathcal{J}_b^*-1})| \ge l - \bar{\epsilon}$ 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 $\mathcal{J}_b^* \ge 2$ and start by specifying the constant $\bar{t}$. From the continuity of $a(\cdot)$ (see Assumption 2) and the fact that $a(y) \ne 0\ \forall y \in (-l, 0) \cup (0, l)$, we can find some $c_{\bar{\epsilon}} > 0$ such that $|a(y)| \ge c_{\bar{\epsilon}}$ for all $y \in [-l + \bar{\epsilon}, -\bar{\epsilon}] \cup [\bar{\epsilon}, l - \bar{\epsilon}]$. Now we pick some

$$
t_{\bar{\epsilon}} \triangleq l/c_{\bar{\epsilon}}, \qquad \bar{t} = (\mathcal{J}_b^* - 1)\cdot t_{\bar{\epsilon}}.
$$

We proceed with a proof by contradiction. Suppose that $t_{\mathcal{J}_b^*-1} \ge \bar{t} = (\mathcal{J}_b^* - 1)\cdot t_{\bar{\epsilon}}$, then we can find some $j^* = 1, 2, \cdots, \mathcal{J}_b^* - 1$ such that $t_{j^*} - t_{j^*-1} \ge t_{\bar{\epsilon}}$. First, recall that we have shown that $|\xi(t_{j^*-1})| < l - \bar{\epsilon}$. Next, note that we must have $|\xi(t)| < \bar{\epsilon}$ for some $t \in [t_{j^*-1}, t_{j^*})$. Indeed, suppose that $|\xi(t)| \ge \bar{\epsilon}$ for all $t \in [t_{j^*-1}, t_{j^*})$. Then from observation (C.1) and the fact that $|a(y)| \ge c_{\bar{\epsilon}}$ for all $y \in [-\gamma, -\bar{\epsilon}] \cup [\bar{\epsilon}, \gamma]$, we yield

$$
|\xi(t_{j^*}-)| \le |\xi(t_{j^*-1})| - c_{\bar{\epsilon}}\cdot t_{\bar{\epsilon}} \le l - c_{\bar{\epsilon}}\cdot \frac{l}{c_{\bar{\epsilon}}} = 0.
$$

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 $\inf_{t\in[0,t_{\mathcal{J}_b^*-1}]} |\xi(t)| \ge \bar{\epsilon}$. 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 (C.1), we get $|\xi(t_1-)| \le |x_0| \le \epsilon$. As a result,

$$
\begin{aligned}
R_0 &\le \epsilon_0 + \Big|\varphi_b\Big(\sigma\big(\xi(t_1-)\big)\cdot w_1\Big) - \varphi_b\big(\sigma(0)\cdot w_1\big)\Big| \\
&\le \epsilon + \big|\sigma\big(\xi(t_1-)\big) - \sigma(0)\big|\cdot\big|\varphi_{b/c}(w_1)\big| \qquad \text{using (B.8)} \\
&\le \epsilon + D\epsilon\cdot\frac{b}{c} = (1 + \frac{bD}{c})\cdot\epsilon \qquad \text{because of Assumption 2.}
\end{aligned}
$$

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 $\hat{t}_{j+1} \triangleq t_{j+1} - t_1$)

$$\begin{aligned}
R_{j+1} &= |\hat{\xi}(\hat{t}_{j+1}) - \xi(t_{j+1})| \\
&= \Big|\hat{\xi}(\hat{t}_{j+1}-) + \varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big) \cdot w_{j+1}\Big) - \Big[\xi(t_{j+1}-) + \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big) \cdot w_{j+1}\Big)\Big]\Big| \\
&\leq \Big|\hat{\xi}(\hat{t}_{j+1}-) - \xi(t_{j+1}-)\Big| + \Big|\varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big) \cdot w_{j+1}\Big) - \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big) \cdot w_{j+1}\Big)\Big| \\
&\leq \exp(DT)R_j + \Big|\varphi_b\Big(\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big) \cdot w_{j+1}\Big) - \varphi_b\Big(\sigma\big(\xi(t_{j+1}-)\big) \cdot w_{j+1}\Big)\Big| \\
&\leq \exp(DT)R_j + \Big|\sigma\big(\hat{\xi}(\hat{t}_{j+1}-)\big) - \sigma\big(\xi(t_{j+1}-)\big)\Big| \cdot |\varphi_{b/c}(w_{j+1})| \qquad \text{using (B.8)} \\
&\leq \exp(DT)R_j + D\big|\hat{\xi}(t_{j+1}-) - \xi(t_{j+1}-)\big| \cdot \frac{b}{c} \qquad \text{due to Assumption 2} \\
&\leq \exp(DT)R_j + \frac{bD}{c} \cdot \exp(DT)R_j = \Big(1 + \frac{bD}{c}\Big)\exp(DT)R_j \leq \rho^{j+2} \cdot \epsilon.
\end{aligned}$$

By arguing inductively we conclude the proof.

($e$) Note that the statement is not affected by the values of $\xi$ 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

$$|\xi(t_{\mathcal{J}_b^*}) - \hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| < \underbrace{\Big[2\exp\big(D(\bar{t}+1)\big) \cdot \Big(1 + \frac{bD}{c}\Big)\Big]^{\mathcal{J}_b^*+1}}_{\triangleq \rho^*} \cdot \epsilon_0 \qquad \forall \epsilon_0 \in (0, \bar{\epsilon}]. \tag{C.2}$$

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 (C.2). From observation (C.1), we get $|\xi(t_1-)| \leq |\xi(0)| = |x| \leq \epsilon_0$. 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 $|\hat{\xi}(t_{\mathcal{J}_b^*} - t_1)| > \theta > l - \bar{\epsilon}$, then $t_{\mathcal{J}_b^*} - t_1 < \bar{t}$ from part ($c$). Likewise, from part ($c$), $|\xi(t_{\mathcal{J}_b^*})| > \theta > l - \bar{\epsilon}$, then $t_{\mathcal{J}_b^*} < \bar{t}$. Therefore, in either case, $t_{\mathcal{J}_b^*} - t_1 + 1 \leq \bar{t} + 1$. This concludes the proof. $\square$

The next few lemmas study the continuity of measure $\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}$ as well as the mass it charges on different sets.

**Lemma C.2.** *Let $\bar{\epsilon} \in (0, b)$ be defined as in (4.7). For any $|\gamma| > (\mathcal{J}_b^* - 1)b + \bar{\epsilon}$ such that $\gamma/b \notin \mathbb{Z}$,*

$$\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\{\gamma\}) = 0.$$

*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 $\gamma \neq b$, we know that $\big\{w : \ \varphi_b\big(\sigma(0)\cdot w\big) = \gamma\big\} \subseteq \{\frac{\gamma}{\sigma(0)}\}$. The absolute continuity of $\nu_\alpha$ (w.r.t the Lebesgue measure) then implies that $\breve{\mathbf{C}}^{(1)|b}(\{\gamma\}) = 0$.

Now we focus on the case where $\mathcal{J}_b^* \geq 2$. Observe that

$$\begin{aligned}
&\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\{\gamma\}) \\
&= \int \mathbb{I}\Bigg(\mathbb{I}\Big\{g^{(\mathcal{J}_b^*-1)|b}\big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*-2}, w^*, t_1, \cdots, t_{\mathcal{J}_b^*-2}, t_{\mathcal{J}_b^*-2}+t^*\big) = \gamma\Big\} \\
&\qquad\qquad \times \nu_\alpha(dw^*)\mathcal{L}(dt^*)\Bigg)\nu_\alpha^{\mathcal{J}_b^*-1}(dw_1, \cdots, dw_{\mathcal{J}_b^*-1}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-2\uparrow}(dt_1, \cdots, dt_{\mathcal{J}_b^*-2}) \\
&= \int \Bigg(\int_{(t^*,w^*)\in E(\boldsymbol{w},\boldsymbol{t})} \nu_\alpha(dw^*)\mathcal{L}(dt^*)\Bigg)\nu_\alpha^{\mathcal{J}_b^*-1}(d\boldsymbol{w}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-2\uparrow}(d\boldsymbol{t})
\end{aligned}$$

where

$$\begin{aligned}
E(\boldsymbol{w},\boldsymbol{t}) &= \Bigg\{(w,t) \in \mathbb{R}\times(0,\infty) : \ \varphi_b\Bigg(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) + \sigma\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\Big)w\Bigg) = \gamma\Bigg\}, \\
\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t}) &= \breve{g}^{(\mathcal{J}_b^*-2)|b}\big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*-2}, t_1, \cdots, t_{\mathcal{J}_b^*-2}\big).
\end{aligned}$$

Here $\boldsymbol{y}_t(x)$ is the ODE defined in (2.22). 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

$$E(\boldsymbol{w},\boldsymbol{t}) \subseteq \big\{(w,t) \in \mathbb{R}\times(0,\infty) : w = w^*(t) \text{ or } t = t^*\big\}. \tag{C.3}$$

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 (C.3). 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 C.1 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 5 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 $\varphi_b\Bigg(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big) + \sigma\Big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\Big)w\Bigg) = \gamma$, 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 $w = \frac{\gamma - \boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)}{\sigma\big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\big)}$. (Note that this quantity is well-defined due to $\sigma(x) > 0 \ \forall x \in \mathbb{R}$; see Assumption 3.) By setting $w^*(t) \triangleq \frac{\gamma - \boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)}{\sigma\big(\boldsymbol{y}_t\big(\widetilde{\boldsymbol{x}}(\boldsymbol{w},\boldsymbol{t})\big)\big)}$ we conclude the proof. $\square$

**Lemma C.3.** $\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(I^c) \in (0,\infty)$.

*Proof.* Let $\bar{t}, \bar{\delta}$ be the constants characterized in Lemma C.1. We start with the proof of finiteness. Recall that $l = |s_{\text{left}}| \wedge s_{\text{right}}$, and observe

$$\begin{aligned}
&\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big((-\infty, s_{\text{left}}] \cup [s_{\text{right}}, \infty)\Big) \\
&\leq \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(\mathbb{R} \setminus \big[-(l-\bar{\epsilon}), l-\bar{\epsilon}\big]\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \int \mathbb{I}\bigg\{\Big|\breve{g}^{(\mathcal{J}_b^*-1)|b}\big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), (w_1,\cdots,w_{\mathcal{J}_b^*-1}), (t_1,\cdots,t_{\mathcal{J}_b^*-1})\big)\Big| > l-\bar{\epsilon}\bigg\} \\
&\qquad\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1,\cdots,dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(dt_1,\cdots,dt_{\mathcal{J}_b^*-1}) \\
&= \int \mathbb{I}\bigg\{\Big|h^{(\mathcal{J}_b^*-1)|b}_{[0,1+t_{\mathcal{J}_b^*-1}]}\big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), (w_1,\cdots,w_{\mathcal{J}_b^*-1}), (t_1,\cdots,t_{\mathcal{J}_b^*-1})\big)(t_{\mathcal{J}_b^*-1})\Big| > l-\bar{\epsilon}\bigg\} \\
&\qquad\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1,\cdots,dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(dt_1,\cdots,dt_{\mathcal{J}_b^*-1}) \\
&\leq \int \mathbb{I}\Big(|w_j| > \bar{\delta}\ \forall j\in[\mathcal{J}_b^*];\ t_{\mathcal{J}_b^*-1} < \bar{t}\Big)\nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(d\boldsymbol{t}) \quad \text{using part } (c) \text{ of Lemma C.1} \\
&\leq \bar{t}^{\mathcal{J}_b^*-1}/\bar{\delta}^{\alpha\mathcal{J}_b^*} < \infty.
\end{aligned}
$$

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 $l/b \notin \mathbb{Z}$, 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,

$$
\breve{\mathbf{C}}^{(1)|b}\Big(\big[s_{\text{right}},\infty\big)\Big) = \int \mathbb{I}\big\{\varphi_b\big(\sigma(0)\cdot w\big) \geq s_{\text{right}}\big\}\nu_\alpha(dw) \geq \int_{w\in[\frac{b}{\sigma(0)},\infty)} \nu_\alpha(dw) = \left(\frac{\sigma(0)}{b}\right)^\alpha > 0.
$$

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

$$
\begin{aligned}
&\breve{g}^{(\mathcal{J}_b^*)|b}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1,\cdots,w_{\mathcal{J}_b^*-1},\boldsymbol{t}\Big) \\
&= h^{(\mathcal{J}_b^*-1)|b}_{[0,t_{\mathcal{J}_b^*-1}+1]}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1,\cdots,w_{\mathcal{J}_b^*-1},\boldsymbol{t}\Big)(t_{\mathcal{J}_b^*-1}) > s_{\text{right}}.
\end{aligned}
\tag{C.4}
$$

Then from the continuity of mapping $h^{(\mathcal{J}_b^*-1)|b}_{[0,t_{\mathcal{J}_b^*-1}+1]}$ (see Lemma B.3), 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

$$
\begin{aligned}
&\breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}\Big(\big[s_{\text{right}},\infty\big)\Big) \\
&\geq \int \mathbb{I}\Big\{|w_j' - w_j| < \Delta\ \forall j\in[\mathcal{J}_b^*];\ |t_j' - t_j| < \Delta\ \forall j \in [\mathcal{J}_b^*-1]\Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1',\cdots,dw'_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1}(dt_1',\cdots,dt'_{\mathcal{J}_b^*-1}) \\
&> 0.
\end{aligned}
$$

It only remains to show (C.4). 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 C.1 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 $j = 1,2,\cdots,\mathcal{J}_b^*-1$)

$$
\xi(t_j) = \xi(t_{j-1}) + \int_{s\in[t_{j-1},t_j)} a\big(\xi(s)\big)ds + \varphi_b\big(\sigma(\xi(t_j-)\cdot w_j\big)
$$

$$
\begin{aligned}
&= \xi(t_{j-1}) + \int_{s\in[t_{j-1},t_j)} a\big(\xi(s)\big)ds + b \quad \text{ due to } \sigma\big(\xi(t_j-)\big)\cdot w_j \ge c\cdot\frac{b}{c} = b\\
&\ge \xi(t_{j-1}) - C_0\cdot(t_j - t_{j-1}) + b \quad \text{because of } a(x)x\le 0 \text{ (see Assumption 5) and } \big|a\big(\xi(t)\big)\big|\le C_0\\
&= \xi(t_{j-1}) - C_0\Delta + b.
\end{aligned}
$$

By arguing inductively, we get

$$
\breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b(\sigma(0)\cdot w_{\mathcal{J}_b^*}), w_1,\cdots,w_{\mathcal{J}_b^*-1},\boldsymbol{t}\Big) = \xi(t_{\mathcal{J}_b^*-1}) \ge \mathcal{J}_b^* b - (\mathcal{J}_b^*-1)C_0\Delta.
$$

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 C.4.** *Let $\bar\epsilon \in (0,b)$ be defined as in (4.7). Given any open interval $S\subseteq\mathbb{R}$, let*

$$
r_S \triangleq \inf\{|x| :\ x\in S\}, \qquad d_S \triangleq \lceil r_S/b\rceil.
$$

*If $d_S \ge k$ and $r_S - (d_S-1)\cdot b > \bar\epsilon$ for some positive integer $k$, 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 (2.24), 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 $T = t_{k-1}+1$)

$$
h^{(k-1)|b}_{[0,T]}\big(\varphi_b(\sigma(0)\cdot w_0),\boldsymbol{w},\boldsymbol{t}\big)(t_{k-1}) \in S. \tag{C.5}
$$

However, part (a) of Lemma C.1 implies that $\big|h^{(k-1)|b}_{[0,T]}\big(\varphi_b(\sigma(0)\cdot w_0),\boldsymbol{w},\boldsymbol{t}\big)(t)\big| < (k-1)\cdot b + \bar\epsilon$ for all $t\in[0,t_{k-1})$. Therefore,

$$
\begin{aligned}
r_S \le \big|h^{(k-1)|b}_{[0,T]}\big(\varphi_b(\sigma(0)\cdot w_0),\boldsymbol{w},\boldsymbol{t}\big)(t_{k-1})\big| &\le \big|h^{(k-1)|b}_{[0,T]}\big(\varphi_b(\sigma(0)\cdot w_0),\boldsymbol{w},\boldsymbol{t}\big)(t_{k-1}-)\big| + b\\
&\le k\cdot b + \bar\epsilon.
\end{aligned}
$$

This leads to $r_S/b < k+1$, and hence $d_S = k$ or $k+1$. Furthermore, suppose that $d_S = k+1$. Then $r_S \le k\cdot b+\bar\epsilon$ 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 (C.5) holds under the choice of $T = t_{k-1}+1$. Then from the continuity of mapping $h^{(k)|b}_{[0,T]}$ (see Lemma B.3), 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, $\breve{\mathbf{C}}^{(k)|b}(S) \ge \Big(\prod_{i\in[k-1]}\int_{(t_i-\Delta,t_i+\Delta)}\mathcal{L}(dt)\Big)\cdot \Big(\prod_{i=0,1,\cdots,k-1}\int_{(w_i-\Delta,w_i+\Delta)}\nu_\alpha(dw)\Big) > 0$.

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 (C.5) 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. $\square$

To conclude, we provide the proof of Lemma 4.2.

*Proof of Lemma 4.2.* Let $\bar{\delta}, \bar{t}$ be characterized as in Lemma C.1. Using part $(e)$ of Lemma C.1, for the fixed $\Delta > 0$ we can fix some $\epsilon_0 \in (0, \Delta/2)$ such that the following claim holds (recall that $l = |s_{\text{left}}| \wedge s_{\text{right}}$): 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 $\boldsymbol{t} = (t_1, \cdots, t_{\mathcal{J}^*_b}) \in (0,T]^{\mathcal{J}^*_b \uparrow}$,

$$|\xi(t_{\mathcal{J}^*_b})| \vee |\hat{\xi}(t_{\mathcal{J}^*_b} - t_1)| > l - \bar{\epsilon} \qquad \Longrightarrow \qquad |\hat{\xi}(t_{\mathcal{J}^*_b} - t_1) - \xi(t_{\mathcal{J}^*_b})| < \Delta/2 \tag{C.6}$$

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 $\xi'$ such that $\boldsymbol{d}^{[0,T]}_{J_1}(\xi, \xi') < \epsilon$, 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 (2.10),

$$\begin{aligned}
&\mathbf{C}^{(\mathcal{J}^*_b)|b}_{[0,T]}\Big(\big(\breve{E}(\epsilon, B, T)\big)^-;\ x\Big) \\
&\leq \int \mathbb{I}\Big\{\exists t \in [0,T] \text{ s.t. } h^{(\mathcal{J}^*_b)|b}_{[0,T]}(x, \boldsymbol{w}, \boldsymbol{t})(t) \in B^{\Delta/2}\Big\}\ \nu^{\mathcal{J}^*_b}_\alpha(d\boldsymbol{w}) \times \mathcal{L}^{\mathcal{J}^*_b \uparrow}_T(d\boldsymbol{t}) \\
&\text{(by setting } u_j \triangleq t_j - t_1 \text{ for all } j = 2, 3, \cdots, \mathcal{J}^*_b) \\
&= \int \bigg(\int \mathbb{I}\Big\{\exists t \in [0,T] \text{ s.t. } h^{(\mathcal{J}^*_b)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}^*_b})(t) \in B^{\Delta/2}\Big\} \\
&\qquad\qquad \times \nu^{\mathcal{J}^*_b}_\alpha(d\boldsymbol{w}) \times \mathcal{L}^{\mathcal{J}^*_b - 1\uparrow}_{T - t_1}(du_2, \cdots, du_{\mathcal{J}^*_b})\bigg)\mathcal{L}(dt_1) \\
&= \int \phi_B(t_1, x)\mathcal{L}_T(dt_1)
\end{aligned} \tag{C.7}$$

where

$$\begin{aligned}
\phi_B(t_1, x) = \int \mathbb{I}\Big\{&\exists t \in [0,T] \text{ s.t. } h^{(\mathcal{J}^*_b)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}^*_b})(t) \in B^{\Delta/2}\Big\} \\
&\times \nu^{\mathcal{J}^*_b}_\alpha(d\boldsymbol{w}) \times \mathcal{L}^{\mathcal{J}^*_b - 1\uparrow}_{T - t_1}(du_2, \cdots, du_{\mathcal{J}^*_b}).
\end{aligned}$$

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 $\inf_{w \in B}|w| \geq l - \bar{\epsilon}/2$. Because of our choice of $\Delta \in (0, \bar{\epsilon}/2)$, we then have $\inf_{w \in B^{\Delta/2}}|w| > l - \bar{\epsilon}$. Using property (C.6), for all $t_1 \in (0,T]$ and $x \in [-\epsilon_0, \epsilon_0]$ we have the upper bound

$$\begin{aligned}
\phi_B(t_1, x) \leq \int \mathbb{I}\Big\{&\breve{g}^{(\mathcal{J}^*_b - 1)|b}\Big(\varphi_b\big(\sigma(0) \cdot w_1\big), w_2, \cdots, w_{\mathcal{J}^*_b}, u_2, \cdots, u_{\mathcal{J}^*_b}\Big) \in B^\Delta\Big\} \\
&\times \nu^{\mathcal{J}^*_b}_\alpha(dw_1, \cdots, dw_{\mathcal{J}^*_b}) \times \mathcal{L}^{\mathcal{J}^*_b - 1\uparrow}_{T - t_1}(du_2, \cdots, du_{\mathcal{J}^*_b})
\end{aligned} \tag{C.8}$$

due to part ($c$) of Lemma C.1. 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 (C.8) that

$$\begin{aligned}
\phi_B(t_1, x) &\leq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B^\Delta \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_\infty^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to } T - t_1 > \bar{t} \text{ (from } t_1 \in (0, T-\bar{t})) \text{ and } u_{J^*} < \bar{t} \text{ (see part (c) of Lemma C.1)} \\
&= \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\Delta).
\end{aligned}$$

On the other hand, for all $t_1 \in [T - \bar{t}, T)$ and $x \in [-\epsilon_0, \epsilon_0]$, from (C.8) we get

$$\begin{aligned}
\phi_B(t_1, x) &\leq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B^\Delta \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{T-t_1}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\leq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B^\Delta \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to } T - t_1 \leq \bar{t} \\
&\leq \int \mathbb{I}\Big\{ \Big| \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \Big| > l - \bar{\epsilon} \Big\} \\
&\qquad\qquad\qquad\qquad \times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to } \Delta < \bar{\epsilon} \text{ and recall } l = |s_{\text{left}}| \wedge s_{\text{right}} \\
&\leq \int \mathbb{I}\Big\{ |w_j| > \bar{\delta}\ \forall j \in [\mathcal{J}_b^*] \Big\} \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}_{\bar{t}}^{\mathcal{J}_b^*-1\uparrow}(du_2, \cdots, du_{\mathcal{J}_b^*}) \\
&\qquad \text{due to part } (c) \text{ of Lemma C.1} \\
&\leq (1/\bar{\delta})^{\alpha \mathcal{J}_b^*} \cdot \bar{t}^{\mathcal{J}_b^*-1}. \qquad \text{(C.9)}
\end{aligned}$$

Therefore, in (C.7) 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 $\xi'$ 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,

$$\begin{aligned}
&\mathbf{C}_{[0,T]}^{(\mathcal{J}_b^*)|b}\Big(\big(\breve{E}(\epsilon, B, T)\big)^\circ;\ x\Big) \\
&\geq \int \mathbb{I}\Big\{ \exists t \in [0,T] \text{ s.t. } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, \boldsymbol{t})(t) \in B_{\Delta/2} \text{ and } h_{[0,T]}^{(\mathcal{J}_b^*)|b}(x, \boldsymbol{w}, \boldsymbol{t})(s) \in I_{2\epsilon}\ \forall s \in [0,t) \Big\}\ \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}_T^{\mathcal{J}_b^*\uparrow}(d\boldsymbol{t})
\end{aligned}$$

$$= \int \widetilde{\phi}_B(t_1, x) \mathcal{L}_T(dt_1)$$

where

$$\widetilde{\phi}_B(t_1, x) = \int \mathbb{I}\Big\{ \exists t \in [0,T] \text{ s.t. } h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(t) \in B_{\Delta/2}$$
$$\text{and } h^{(\mathcal{J}_b^*)|b}_{[0,T]}(x, \boldsymbol{w}, t_1, t_1 + u_2, t_1 + u_3, \cdots, t_1 + u_{\mathcal{J}_b^*})(s) \in I_{2\epsilon}\ \forall s \in [0,t) \Big\}$$
$$\times \nu_\alpha^{\mathcal{J}_b^*}(d\boldsymbol{w}) \times \mathcal{L}^{\mathcal{J}_b^*-1\uparrow}_{T-t_1}(du_2, \cdots, du_{\mathcal{J}_b^*}).$$

Analogous to the argument for (C.8), using property (C.6) we yield that for all $t \in (0, T - \bar{t})$:

$$\phi_B(t_1, x) \geq \int \mathbb{I}\Big\{ \breve{g}^{(\mathcal{J}_b^*-1)|b}\Big(\varphi_b\big(\sigma(0)\cdot w_1\big), w_2, \cdots, w_{\mathcal{J}_b^*}, u_2, \cdots, u_{\mathcal{J}_b^*}\Big) \in B_\Delta \Big\}$$
$$\times \nu_\alpha^{\mathcal{J}_b^*}(dw_1, \cdots, dw_{\mathcal{J}_b^*}) \times \mathcal{L}^{\mathcal{J}_b^*-1\uparrow}_{T-t_1}(du_2, \cdots, du_{\mathcal{J}_b^*})$$
$$= \breve{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B_\Delta).$$

due to part $(c)$ of Lemma C.1 again. To avoid repetitions, we omit the details here. □

# D Properties of the Markov Jump Process $Y^{*|b}$

**Proposition D.1.** *Let Assumptions 6 and 7 hold. Given any $m_{init} \in \{m_1, \ldots, m_{n_{min}}\}$, the following claims hold for $((U_j)_{j\geq 1}, (V_j)_{j\geq 1})$ defined in (5.16):*

*(i) For any $t > 0$, $\lim_{i\to\infty} \mathbf{P}(\sum_{j\leq i} U_j > t) = 1$;*

*(ii) For any $u > 0$ and $i \geq 1$, $\mathbf{P}(U_1 + \cdots + U_i = u) = 0$;*

*(iii) $Y^{*|b}_{\cdot} \stackrel{d}{=} \Phi((U_j)_{j\geq 1}, (V_j)_{j\geq 1})$ holds for the mapping $\Phi$ defined in (5.4), ; that is, it is a continuous-time Markov chain with initial distribution (2.39) and generator*

$$\mathbf{P}(Y^{*|b}_{t+h} = m_j \mid Y^{*|b}_t = m_i) = h \cdot \sum_{j' \in [n_{min}]:\ j' \neq i} q_b(i, j')\theta_b(m_j|m_{j'}) + \boldsymbol{o}(h) \qquad \text{as } h \downarrow 0; \tag{D.1}$$

*see (2.40) and (2.41) for the definitions of $q_b(i,j)$ and $\theta_b(m_j|m_i)$, respectively.*

*Proof.* $(i)$ Recall the definitions of $q_b(i)$ and $q_b(i,j)$ in (2.40). Let $(S_n)_{n\geq 0}$ be a discrete-time Markov chain with state space $\{m_1, \ldots, m_{n_{\min}}\}$ and one-step transition kernel $\mathbf{P}(S_{n+1} = m_j | S_n = m_i) = q_b(i,j)/q_b(i)$. Note that the chain is well-defined due to (5.14). We also introduce notations $S_n(v)$ for the chain initialized under $S_0(v) = v$. For each $n \geq 0$, set $I^S_n(v) = i$ if and only if $S_n(v) = m_i$; that is, the sequence of indices $(I^S_n(v))_{n\geq 0}$ indicates the state of the chain at time $n$.

Let $(E_i)_{i\geq 0}$ be a sequence of iid Exponential RVs with rate 1, which is also independent of $(S_n(m_{\text{init}}))_{n\geq 0}$. For any $i \geq 2$, the law of $(U_j)_{j\geq 1}, (V_j)_{j\geq 1}$ defined in (5.16) then indicates that (recall that $U_1 = 0$ and $V_1 = m_{\text{init}}$)

$$\sum_{j\in[i]} U_j \stackrel{d}{=} \sum_{j=0,1,\cdots,i-2} \frac{E_j}{q_b\big(I^S_j(m_{\text{init}})\big)} \cdot \mathbb{I}\big\{S_j(m_{\text{init}}) \in V_b^*\big\} \tag{D.2}$$
$$\geq \frac{1}{q^*} \cdot \sum_{j=0,1,\cdots,i-2} E_j \cdot \mathbb{I}\big\{S_j(m_{\text{init}}) \in V_b^*\big\} \qquad \text{where } q^* \triangleq \max_{i\in[n_{\min}]:\ m_i \in V_b^*} q_b(i) \in (0,\infty)$$
$$\stackrel{d}{=} \sum_{j=0}^{N_{i-2}} \frac{E_j}{q^*} \qquad \text{where } N_i \triangleq \sum_{j=0}^{i} \mathbb{I}\big\{S_j(m_{\text{init}}) \in V_b^*\big\}.$$

To proceed with the proof of part $(i)$, we fix some $t > 0$. Note that for any positive integer $n$, it holds on event $\{\sum_{j=0}^{n} E_j/q^* > t\} \cap \{N_{i-2} > n\}$ that $\sum_{j=0}^{N_{i-2}} E_j/q^* > t$. This implies $\mathbf{P}(\sum_{j\leq i} U_j > t) \geq \mathbf{P}(\sum_{j=0}^{n} E_j/q^* > t) \cdot \mathbf{P}(N_{i-2} > n)$, due to the independence between $(E_j)_{j\geq 1}$ and $(S_j)_{j\geq 0}$ (and hence $N_i$). Therefore, it suffices to show that, for any $\epsilon > 0$, there exists some positive integer $n = n(\epsilon)$ such that

$$\mathbf{P}(\sum_{j=0}^{n} E_j/q^* > t) > 1-\epsilon, \qquad \lim_{i\to\infty} \mathbf{P}(N_i > n) = 1. \tag{D.3}$$

Furthermore, the inequality $\mathbf{P}(\sum_{j=0}^{n} E_j/q^* > t) > 1-\epsilon$ holds for any $n$ large enough due to $q^* \in (0,\infty)$; see (5.14). Meanwhile, since $S_n$'s is irreducible, the chain will visit $V_b^*$ (more generally, any subset of its state space) infinitely often. In other words, for any fixed $n$, we have $\lim_{i\to\infty} \mathbf{P}(N_i > n) = \lim_{i\to\infty} \mathbf{P}(\#\{j = 0,1,\ldots,i:\ S_j(m_{\text{init}}) \in V_b^*\} > n) = 1$. This concludes the proof of (D.3).

$(ii)$ Fix some $u > 0$ and positive integer $i$. Representation (D.2) implies $U_1+\cdots+U_i \stackrel{d}{=} \sum_{j=1}^{i} C_j \cdot E_j$ for an iid sequence $(E_i)_{i\geq 1}$ of Exponential RVs with rate 1 and another sequence of RVs $(C_i)_{i\geq 1}$ that is independent of $(E_i)_{i\geq 1}$. In particular, $C_i$'s only take values in the set $\mathscr{C} = \{0\} \cup \{1/q_b(i):\ m_i \in V_b^*\}$, which has finitely many elements. Therefore,

$$\mathbf{P}(U_1 + \cdots + U_i = u) = \sum_{(c_1,\cdots,c_i)\in\mathscr{C}^i} \mathbf{P}(c_1E_1 + \cdots + c_iE_i = u)\mathbf{P}(C_j = c_j\ \forall j \in [i]) = 0$$

due to the absolutely continuity of Exponential RVs.

$(iii)$ Recall that $U_1 \equiv 0$. We start by stating a useful property of the mapping $\Phi$. Set $\hat{T}_0 = 1$. For any $k \geq 1$, define (under the convention $U_0 \equiv 0$)

$$\hat{T}_k \triangleq \min\{j > \hat{T}_{k-1}:\ U_j \neq 0\}, \qquad \hat{V}_k \triangleq V_{-1+\hat{T}_k}, \qquad \hat{U}_k \triangleq \sum_{j=\hat{T}_{k-1}}^{-1+\hat{T}_k} U_j = U_{\hat{T}_{k-1}}. \tag{D.4}$$

Note that due to $U_1 \equiv 0$, we have $\hat{T}_1 \geq 2$ and hence $-1+\hat{T}_1 \geq 1$. This confirms that $\hat{V}_1$ is well-defined. In simple terms, $((\hat{U}_k)_{k\geq 1}, (\hat{V}_k)_{k\geq 1})$ can be interpreted as a transformation of $((U_j)_{j\geq 1}, (V_j)_{j\geq 1})$ with consecutive instantaneous jumps grouped together. As a result,

$$\Phi\Big((U_j)_{j\geq 1}, (V_j)_{j\geq 1}\Big) = \Phi\Big((\hat{U}_k)_{k\geq 1}, (\hat{V}_k)_{k\geq 1}\Big). \tag{D.5}$$

To proceed, we consider another representation of the Markov jump process $Y^{*|b}$ based on the following straightforward observation: the law of the process would remain the same if we allow the process to jump from any state $m_i$ to itself at any exponential rate (i.e., by including Markovian "dummy" jumps where the process does not move at all). Specifically, $Y^{*|b} \stackrel{d}{=} \Phi((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1})$ with $\tilde{U}_k$'s and $\tilde{V}_k$'s defined as follows. Let $\tilde{V}_1$ be sampled from the distribution $\theta_b(\cdot|m_{\text{init}})$ defined in (2.41), and set $\tilde{U}_1 \equiv 0$. Note that so far, we have $(\tilde{U}_1, \tilde{V}_1) \stackrel{d}{=} (\hat{U}_1, \hat{V}_1)$. Furthermore, for any $t > 0$, $l \geq 1$, and $m_i,\ m_j \in V_b^*$ (with possibly $m_i = m_j$),

$$\begin{aligned}&\mathbf{P}(\tilde{U}_{l+1} < t,\ \tilde{V}_{l+1} = m_j \mid \tilde{V}_l = m_i, (\tilde{V}_j)_{j=1}^{l-1},\ (\tilde{U}_j)_{j=1}^{l}) = \mathbf{P}(\tilde{U}_{l+1} < t,\ \tilde{V}_{l+1} = m_j \mid \tilde{V}_l = m_i)\\ &= r^{*|b}(i,j) \cdot \big(1 - \exp(-q_b(i)t)\big),\end{aligned} \tag{D.6}$$

where

$$r^{*|b}(i,j) \triangleq \sum_{j'\in[n_{\min}]:\ j'\neq i} \frac{q_b(i,j')}{q_b(i)} \cdot \theta_b(m_j|m_{j'}) \tag{D.7}$$

with $q_b(i)$ and $q_b(i,j)$ defined in (2.40). To see why $Y^{*|b}_{\cdot} \stackrel{d}{=} \Phi((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1})$, note that the process $\Phi((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1})$ is initialized under $\tilde{V}_1 \sim \theta_b(\cdot|m_{\text{init}})$, which is the same initial distribution of $Y^{*|b}_{\cdot}$. Moreover, any jump in $\Phi((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1})$ from $m_i \in V^*_b$ to $m_j \in V^*_b$ (with possibly $m_i = m_j$) is Markovian and occurs with exponential rate $\sum_{j'\neq i} q_b(i,j')\theta_b(m_j|m_{j'})$. In other words, $\Phi((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1})$ is simply a reformulation of $Y^{*|b}$ where we include "dummy" jumps from $m_i \in V^*_b$ to itself with exponential rate $\sum_{j'\neq i} q_b(i,j')\theta_b(m_i|m_{j'})$.

In light of (D.5), it only remains to show that

$$\Big((\hat{U}_k)_{k\geq 1}, (\hat{V}_k)_{k\geq 1}\Big) \stackrel{d}{=} \Big((\tilde{U}_k)_{k\geq 1}, (\tilde{V}_k)_{k\geq 1}\Big). \tag{D.8}$$

Specifically, fix some $k \geq 1$, $m_i, m_j \in V^*_b$, and some $t > 0$. Observe that

$$\begin{aligned}
&\mathbf{P}(\hat{U}_{k+1} < t,\ \hat{V}_{k+1} = m_j,\ \hat{V}_k = m_i) \\
&= \sum_{N\geq 1}\sum_{n\geq 1} \mathbf{P}(\hat{U}_{k+1} < t,\ V_{N+n} = m_j,\ \hat{T}_{k+1} - 1 = N+n,\ V_N = m_i,\ \hat{T}_k - 1 = N) \qquad \text{by (D.4)} \\
&= \sum_{N\geq 1}\sum_{n\geq 1} \mathbf{P}(U_{N+1} < t,\ V_p \notin V^*_b\ \forall N+1 \leq p \leq N+n-1; \\
&\qquad\qquad V_{N+n} = m_j,\ \hat{T}_{k+1} - 1 = N+n,\ V_N = m_i,\ \hat{T}_k - 1 = N) \qquad \text{by (D.4) and (5.16)} \\
&= \sum_{N\geq 1}\sum_{n\geq 1}\sum_{(i_1,\cdots,i_{n-1})\in\mathscr{I}(i,n-1)} \mathbf{P}(U_{N+1} < t,\ V_{N+p} = m_{i_p}\ \forall p \in [n-1]; \\
&\qquad\qquad V_{N+n} = m_j,\ \hat{T}_{k+1} - 1 = N+n,\ V_N = m_i,\ \hat{T}_k - 1 = N) \\
&\text{where } \mathscr{I}(i,n-1) \triangleq \big\{(i_1,\ldots,i_{n-1}) :\ i_p \neq i_{p-1} \text{ and } m_{i_p} \notin V^*_b\ \forall p \in [n-1]\big\} \text{ with convention } i_0 = i \\
&= \sum_{N\geq 1} \mathbf{P}(V_N = m_i,\ \hat{T}_k - 1 = N) \\
&\quad \cdot \sum_{n\geq 1}\sum_{(i_1,\cdots,i_{n-1})\in\mathscr{I}(i,n-1)} \frac{q_b(i,i_1)}{q_b(i)}\Big(1 - \exp\big(-q_b(i)t\big)\Big)\frac{q_b(i_1,i_2)}{q_b(i_1)}\cdots\frac{q_b(i_{n-2},i_{n-1})}{q_b(i_{n-2})}\frac{q_b(i_{n-1},j)}{q_b(i_{n-1})} \\
&\qquad\qquad \text{using (5.16)} \\
&= \sum_{N\geq 1} \mathbf{P}(V_N = m_i,\ \hat{T}_k - 1 = N) \\
&\quad \cdot \sum_{i_1\neq i} \frac{q_b(i,i_1)}{q_b(i)}\Big(1 - \exp\big(-q_b(i)t\big)\Big) \cdot \sum_{n\geq 1} \mathbf{P}(\tau^S(m_1) = n-1,\ S_{\tau^S(m_1)}(m_1) = m_j).
\end{aligned}$$

In the last line of the display above, we adopt the notations in part ($i$) and let $S_n(v)$ be a Markov chain with initial value $S_0(v) = v$ and transition kernel $\mathbf{P}(S_{n+1} = m_j|S_n = m_i) = q_b(i,j)/q_b(i)$. Furthermore, let $\tau^S(v) = \min\{n \geq 0 :\ S_n(v) \in V^*_b\}$ be the hitting time of any state in $V^*_b$. Now, observe that

$$\begin{aligned}
&\mathbf{P}(\hat{U}_{k+1} < t,\ \hat{V}_{k+1} = m_j,\ \hat{V}_k = m_i) \\
&= \sum_{N\geq 1} \mathbf{P}(V_N = m_i,\ \hat{T}_k - 1 = N) \cdot \sum_{i_1\in[n_{\min}]:\ i_1\neq i} \frac{q_b(i,i_1)}{q_b(i)}\Big(1 - \exp\big(-q_b(i)t\big)\Big)\theta_b(m_j|m_{i_1}) \\
&= \sum_{N\geq 1} \mathbf{P}(V_N = m_i,\ \hat{T}_k - 1 = N) \cdot r^{*|b}(i,j) \cdot \Big(1 - \exp\big(-q_b(i)t\big)\Big) \qquad \text{with } r^{*|b}(\cdot,\cdot) \text{ defined in (D.7)} \\
&= r^{*|b}(i,j) \cdot \Big(1 - \exp\big(-q_b(i)t\big)\Big) \cdot \mathbf{P}(\hat{V}_k = m_i).
\end{aligned}$$

This verifies $\mathbf{P}(\hat{U}_{k+1} < t,\ \hat{V}_{k+1} = m_j \mid \hat{V}_k = m_i) = r^{*|b}(i,j) \cdot \big(1 - \exp(-q_b(i)t)\big)$. Through (D.6) we conclude the proof of (D.8). $\square$

# E Technical Lemmas for Propositions 5.6 and 5.7

This section collects the proofs of Lemma 5.8 and Lemma 5.9, which adapts the first exit analysis in Section 4 to a more general setting.

*Proof of Lemma 5.8.* In light of Lemma C.4, it holds for all $M > 0$ large enough such that

$$\check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}((-M,M)^c; m_i) = 0 \qquad \forall i \in [n_{\min}].$$

This concludes the proof of (5.21).

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 (5.13). Also, we fix some $\epsilon \in (0, \bar{\epsilon})$ and show that (5.22) holds for such $\epsilon$. Recall the definition of $\tau^{\eta|b}_{i;\delta,M}(x)$ in (5.20) and $I_{i;\delta,M} = (s_{i-1} + \delta, s_i - \delta) \cap (-M, M)$. We make a few observations regarding the stopping time $\tau^{\eta|b}_{i;2\delta,M}(x) = \min\{j \geq 0 :\ X^{\eta|b}_j(x) \notin I_{i;2\delta,M}\}$. First, due to $I_{i;2\delta,M} \subseteq I_{i;\delta,M}$, we must have $\tau^{\eta|b}_{i;2\delta,M}(x) \leq \tau^{\eta|b}_{i;\delta,M}(x) \leq \sigma^{\eta|b}_{i;\epsilon}(x)$. Second, by definition of $\tau^{\eta|b}_{i;2\delta,M}(x)$, we have $X^{\eta|b}_j(x) \notin S(\delta)$, $|X^{\eta|b}_j(x)| < M$ for all $j < \tau^{\eta|b}_{i;2\delta,M}(x)$. On event

$$A_0(\eta,\delta,x) \triangleq \{X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in (-M,M);\ X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \notin S(2\delta)\},$$

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

$$A_1(\eta,\delta,x) \triangleq \{\exists j < \sigma^{\eta|b}_{i;\epsilon}(x)\ s.t.\ X^{\eta|b}_j(x) \in S(\delta)\},\ A_2(\eta,x) \triangleq \{\exists j < \sigma^{\eta|b}_{i;\epsilon}(x)\ s.t.\ |X^{\eta|b}_j(x)| \geq M+1\}.$$

Let $R^{\eta|b}_{j;\epsilon}(x) \triangleq \min\{k \geq 0 :\ X^{\eta|b}_k(x) \in (m_j - \epsilon, m_j + \epsilon)\}$ be the first time entering $(m_j - \epsilon, m_j + \epsilon)$. From the strong Markov property at $\tau^{\eta|b}_{i;2\delta,M}(x)$,

$$\begin{aligned}
&\max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(\big(A_1(\eta,\delta,x) \cup A_2(\eta,x)\big) \cap A_0(\eta,\delta,x)\Big) \\
&\leq \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\Big(A_1(\eta,\delta,x) \cup A_2(\eta,x)\ \Big|\ A_0(\eta,\delta,x)\Big) \\
&\leq \max_{j\in[n_{\min}]} \underbrace{\sup_{y\in[s_{j-1}+2\delta, s_j-2\delta]\cap[-M,M]} \mathbf{P}\bigg(\Big\{X^{\eta|b}_k(x) \in [s_{j-1}+\delta, s_j-\delta] \cap (-M-1, M+1)\ \forall \exists k < R^{\eta|b}_{j;\epsilon}(x)\Big\}^c\bigg)}_{p_j(\eta)}.
\end{aligned}$$

For any $j \in [n_{\min}]$ and any $\delta > 0$ small enough, by applying Lemma 4.5 onto $I_j \cap (-M-1, M+1)$ (with parameter $\epsilon$ therein set as $2\delta$) we get $\lim_{\eta\downarrow 0} p_j(\eta) = 0$. In summary, we have shown that $\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big((A_1(\eta,\delta,x) \cup A_2(\eta,x)) \cap A_0(\eta,\delta,x)\big) = 0$. Therefore, to establish (5.22), it only remains to show that $\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big((A_0(\eta,\delta,x))^c\big) < \Delta$. Now, it only remains to prove that for all $\delta > 0$ small enough,

$$\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in S(2\delta)\big) < \Delta, \tag{E.1}$$

$$\limsup_{\eta\downarrow 0} \max_{i\in[n_{\min}]} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \notin (-M,M)\big) = 0. \tag{E.2}$$

Note that $\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(X^{\eta|b}_{\tau^{\eta|b}_{i;2\delta,M}(x)}(x) \in S(2\delta)\big) \leq \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(S(2\delta); m_i\big) / \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(I^c_{i;2\delta,M}; m_i\big)$ can be established using part $(a)$ of Theorem 2.6. From Lemma C.2, we get $\lim_{\delta\downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(S(2\delta); m_i\big) = \check{\mathbf{C}}^{(\mathcal{J}_b^*(i))|b}\big(\{s_1, \cdots, s_{n_{\min}}\}; m_i\big) = 0$ and verifies claim (E.1). Similarly, claim (E.2) follows directly from part $(a)$ of Theorem 2.6 when applied onto $I_{i;\delta,M}$, combined with (5.21). $\square$

*Proof of Lemma 5.9.* (*i*) Fix some $t > 0$ and $\epsilon \in (0, \bar{\epsilon})$. Recall that $\lambda^*_b(\eta) \in \mathcal{RV}_{\mathcal{J}^*_b(V)\cdot(\alpha-1)+1}(\eta)$ as $\eta \downarrow 0$. Due to $\mathcal{J}^*_b(V) \geq 1$ and $\alpha > 1$, we have $\mathcal{J}^*_b(V)\cdot(\alpha-1)+1 \geq \alpha > 1$. This implies that $\lim_{\eta\downarrow 0} \frac{T/\eta}{t/\lambda^*_b(\eta)} = 0\ \forall T > 0$, and hence (given any $T > 0$)

$$\mathbf{P}\big(R^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_b(\eta) \leq t,\ X^{\eta|b}_j(x) \in I_i\ \forall j \leq R^{\eta|b}_{i;\epsilon}(x)\big) \geq \mathbf{P}\big(R^{\eta|b}_{i;\epsilon}(x) \leq T/\eta,\ X^{\eta|b}_j(x) \in I_i\ \forall j \leq R^{\eta|b}_{i;\epsilon}(x)\big)$$

for all $\eta$ small enough. Now, pick $M > 0$ large enough such that $|M| > \min\{|s_{i-1}+\epsilon|, |s_i - \epsilon|\}$. By picking $T > 0$ large enough, one can apply Lemma 4.5 onto $(-M, M) \cap I_j$ to conclude the proof of part (*i*).

(*ii*) Let $\lambda^*_{i;b}(\eta) \triangleq \eta\cdot\lambda^{\mathcal{J}^*_b(i)}(\eta)$. 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 $t > 0$,

$$\liminf_{\eta\downarrow 0} \inf_{x\in[m_i-\epsilon, m_i+\epsilon]} \mathbf{P}\big(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big) \geq \exp\big(-q_b(i)\cdot t\big)\cdot\frac{q_b(i,j)}{q_b(i)}, \tag{E.3}$$

$$\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon, m_i+\epsilon]} \mathbf{P}\big(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big) \leq \exp\big(-q_b(i)\cdot t\big)\cdot\frac{q_b(i,j)}{q_b(i)}. \tag{E.4}$$

Indeed, in case that $m_i \in V^*_b$, claims in part (*ii*) are equivalent to (E.3) and (E.4) due to $\mathcal{J}^*_b(i) = \mathcal{J}^*_b(V)$ (see (2.34)) and hence $\lambda^*_{i;b}(\eta) = \lambda^*_b(\eta) = \eta\cdot\lambda^{\mathcal{J}^*_b(V)}(\eta)$. In case that $m_i \notin V^*_b$ (i.e., $\mathcal{J}^*_b(i) < \mathcal{J}^*_b(V)$), we have $\lim_{\eta\downarrow 0}\frac{t/\lambda^*_{i;b}(\eta)}{T/\lambda^*_b(\eta)} = 0$ for all $t, T \in (0,\infty)$. We then recover the upper and lower bounds in part (*ii*) by letting $t \downarrow 0$ in (E.3) and (E.4).

The rest of this proof is devoted to establishing (E.3) and (E.4). Here, we collect a few useful facts about the measure $\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}$. By assumption (5.12), one can apply Lemma C.2 and obtain $\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(\{s_1, \cdots, s_{n_{\min}-1}\}; m_i\big) = 0$. Recall the definition of $q_b(i,j) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j; m_i)$ in (2.40). Due to $I_j = (s_{j-1}, s_j)$,

$$q_b(i,j) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j; m_i) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j^-; m_i) \qquad \forall i, j \in [n_{\min}] \text{ with } i \neq j. \tag{E.5}$$

Combining (E.5) with the continuity of measures, we have $\lim_{\delta\downarrow 0}\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big((s_{i-1}-\delta, s_i+\delta)^c; m_i\big) = q_b(i) = \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_i^c; m_i)$. Next, throughout the remainder of this proof, we only consider $M \in (0,\infty)$ large enough such that the claim (5.21) of Lemma 5.8 holds. Given any $\Delta > 0$, regarding the set $I_{i;\delta,M} = (-M, M) \cap (s_{i-1}+\delta, s_i - \delta)$ it holds for all $\delta > 0$ small enough that

$$\begin{aligned}\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_{i;\delta,M}; m_i) &\leq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}((-M,M)^c; m_i) + \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}((s_{i-1}+\delta, s_i-\delta)^c; m_i) \\ &= 0 + \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}((s_{i-1}+\delta, s_i-\delta)^c; m_i) < (1+\Delta)\cdot q_b(i).\end{aligned} \tag{E.6}$$

Lastly, due to $I_{i;\delta,M} \subseteq I_i$,

$$\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I^c_{i;\delta,M}; m_i\big) \geq \check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}\big(I^c_i; m_i\big) = q_b(i). \tag{E.7}$$

**Proof of Lower Bound** (E.3).

We fix some $i \neq j$ and $t > 0$ when proving (E.3). Recall the definitions of $I_{i;\delta,M}$ and $\tau^{\eta|b}_{i;\delta,M}(x)$ in (5.18) and (5.20), respectively. Observe that

$$\begin{aligned}&\big\{\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big\} \\ &\supseteq \underbrace{\big\{\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{j;\delta,M+1}\big\}}_{\text{(I)}} \cap \underbrace{\big\{X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big\}}_{\text{(II)}}.\end{aligned}$$

We first analyze $\mathbf{P}((\text{II})|(\text{I}))$. By strong Markov property at $\tau^{\eta|b}_{i;\delta,M}(x)$, $\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{II})\mid(\text{I}))\geq \inf_{y\in I_{j;\delta,M+1}}\mathbf{P}\big(X^{\eta|b}_k(y)\in I_j\ \forall k\leq R^{\eta|b}_{j;\epsilon}(y)\big)$. Here, recall that $R^{\eta|b}_{j;\epsilon}(x)=\min\{j\geq 0:\ \ X^{\eta|b}_k(x)\in(m_j-\epsilon,m_j+\epsilon)\}$. Applying Lemma 4.5, we yield

$$\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{II})\mid(\text{I}))=1. \tag{E.8}$$

Next, we move onto the analysis of $\mathbf{P}((\text{I}))$. Due to $I_{j;\delta,M+1}\subseteq I_j$,

$$(\text{I})=\underbrace{\{\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta)>t;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x))}(x)\in I_j\}}_{(\text{III})}\cap\underbrace{\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_{j;\delta,M+1}\}}_{(\text{IV})}.$$

Given any $\Delta>0$, by applying part $(a)$ of Theorem 2.6 onto $I_{i;\delta,M}$, we yield (for any $\delta$ small enough)

$$\begin{aligned}\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{III}))&\geq\exp\big(-\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_{i;\delta,M};m_i)\cdot t\big)\cdot\frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I_j;m_i)}{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_{i;\delta,M};m_i)}\\&>\frac{\exp(-(1+\Delta)q_b(i)\cdot t)}{1+\Delta}\cdot\frac{q_b(i,j)}{q_b(i)}\qquad\text{due to (E.5) and (E.6).}\end{aligned}$$

Meanwhile, note that $(\text{IV})^c=\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in S(\delta)\}\cup\{|X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)|\geq M+1\}$. Due to $\tau^{\eta|b}_{i;\delta,M}(x)\leq\sigma^{\eta|b}_{i;\epsilon}(x)$, the claim $\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{IV})^c)<\Delta$ follows directly from (5.22) of Lemma 5.8. In summary, for all $\delta>0$ small enough,

$$\liminf_{\eta\downarrow 0}\inf_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{I}))\geq\frac{\exp(-(1+\Delta)q_b(i)\cdot t)}{1+\Delta}\cdot\frac{q_b(i,j)}{q_b(i)}-\Delta. \tag{E.9}$$

Combining (E.8) and (E.9) and then driving $\Delta\downarrow 0$, we conclude the proof of the lower bound (E.3).

**Proof of Upper Bound** (E.4).

Let $(\text{I})=\{\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta)>t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x)\in I_j\}$. Arbitrarily pick some $\Delta>0$. Given $\delta>0$, define event $(\text{II})=\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in(-M-1,M+1)\setminus S(\delta)\}$. We start from the decomposition $(\text{I})=\big((\text{I})\setminus(\text{II})\big)\cup\big((\text{I})\cap(\text{II})\big)$. Applying (5.22) of Lemma 5.8, it holds for all $\delta>0$ small enough that

$$\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{I})\setminus(\text{II}))\leq\limsup_{\eta\downarrow 0}\sup_{x\in[m_i-\epsilon,m_i+\epsilon]}\mathbf{P}((\text{II})^c)<\Delta. \tag{E.10}$$

Next, recall that $\tau^{\eta|b}_{i;\delta,M}(x)$ is the first exit time from $I_{i;\delta,M}$, and $\sigma^{\eta|b}_{i;\epsilon}(x)$ is the first time visiting the $\epsilon$-neighborhood of a local minimum different from $m_i$; see (5.19) and (5.20). By definition of $\tau^{\eta|b}_{i;\delta,M}(x)$, on event $(\text{I})\cap(\text{II})$ there must be 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

$$(k)=(\text{I})\cap(\text{II})\cap\{X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x)\in I_{k;\delta,M+1}\}$$

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, as has been shown at the beginning of the proof of part $(ii)$, we have $\lim_{\eta\downarrow 0} \frac{T/\eta}{\Delta/\lambda^*_{i,b}(\eta)} = 0$ for all $T \in (0,\infty)$. Then,

$$
\begin{aligned}
&\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((k,1)) \\
&\leq \limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big((k) \cap \{\sigma^{\eta|b}_{i;\epsilon}(x) - \tau^{\eta|b}_{i;\delta,M}(x) > T/\eta\}\big) \\
&\leq \limsup_{\eta\downarrow 0} \sup_{y\in I_{k;\delta,M+1}} \mathbf{P}(\sigma^{\eta|b}_{i;\epsilon}(y) > T/\eta) \qquad \text{by strong Markov property at } \tau^{\eta|b}_{i;\delta,M}(x) \\
&\leq \limsup_{\eta\downarrow 0} \sup_{y\in I_{k;\delta,M+1}} \mathbf{P}(X^{\eta|b}_j(y) \notin (m_k-\epsilon, m_k+\epsilon)\ \forall j \leq T/\eta) \\
&= 0 \qquad \text{for all } T>0 \text{ large enough due to Lemma 4.5.}
\end{aligned}
\tag{E.11}
$$

Meanwhile,

$$
\begin{aligned}
&\sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((k,2)) \\
&\leq \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{k;\delta,M+1};\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big) \\
&\leq \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{k;\delta,M+1}\big) \\
&\qquad \cdot \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j \ \big|\ \tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_{k;\delta,M+1}\big) \\
&\leq \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \underbrace{\mathbf{P}\big(\tau^{\eta|b}_{i;\delta,M}(x)\cdot\lambda^*_{i;b}(\eta) > t-\Delta;\ X^{\eta|b}_{\tau^{\eta|b}_{i;\delta,M}(x)}(x) \in I_k\big)}_{(k,\mathrm{I})} \cdot \sup_{y\in I_{k;\delta,M+1}} \underbrace{\mathbf{P}\big(X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(y)}(y) \in I_j\big)}_{(k,\mathrm{II})}.
\end{aligned}
$$

Applying part $(a)$ of Theorem 2.6 onto $I_{i;\delta,M}$ and the bound (E.7), we yield (for any $\delta$ small enough)

$$
\begin{aligned}
\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((k,\mathrm{I})) &\leq \exp\big(-\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_{i;\delta,M};m_i)\cdot(t-\Delta)\big)\cdot\frac{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^-_k;m_i)}{\check{\mathbf{C}}^{(\mathcal{J}^*_b(i))|b}(I^c_{i;\delta,M};m_i)} \\
&\leq \exp\big(-q_b(i)\cdot(t-\Delta)\big)\cdot\frac{q_b(i,k)}{q_b(i)} \qquad \text{using (E.7) and (E.5).}
\end{aligned}
\tag{E.12}
$$

Moving on, we analyze the probability of event $(k,\mathrm{II})$. If $k=j$, we apply the trivial upper bound $\mathbf{P}((k,\mathrm{II})) \leq 1$. If $k \neq j$, on event $(k,\mathrm{II})$, $\big(X^{\eta|b}_n(y)\big)_{n\geq 0}$ visited $(m_j-\epsilon, m_j+\epsilon)$ before visiting any other local minima's $\epsilon$-neighborhood, despite the fact that the initial value $X^{\eta|b}_0(y) = y$ belongs to $I_{k;\delta,M+1} \subset I_k$. This implies that $\big(X^{\eta|b}_n(y)\big)_{n\geq 0}$ must have left $I_k$ before visiting its local minimum $m_k$. Applying Lemma 4.5, we obtain $\limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M+1}} \mathbf{P}((k,\mathrm{II})) = 0\ \forall k\neq j$ for all $\delta>0$ small enough In summary, for all $\delta$ small enough,

$$
\begin{aligned}
&\limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}\big(\sigma^{\eta|b}_{i;\epsilon}(x)\cdot\lambda^*_{i;b}(\eta) > t,\ X^{\eta|b}_{\sigma^{\eta|b}_{i;\epsilon}(x)}(x) \in I_j\big) \\
&\leq \limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((\mathrm{II})^c) \\
&\qquad + \sum_{k\in[n_{\min}]:\ k\neq i} \limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((k,\mathrm{I}))\cdot\limsup_{\eta\downarrow 0}\sup_{y\in I_{k;\delta,M}} \mathbf{P}((k,\mathrm{II})) \qquad \text{due to (E.11)} \\
&\leq \limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((\mathrm{II})^c) + \limsup_{\eta\downarrow 0} \sup_{x\in[m_i-\epsilon,m_i+\epsilon]} \mathbf{P}((j,\mathrm{I}))
\end{aligned}
$$

$$\leq \Delta + \exp\big(-q_b(i)\cdot(t-\Delta)\big)\cdot\frac{q_b(i,j)}{q_b(i)} \qquad \text{due to (E.10) and (E.12).}$$

Let $\Delta \downarrow 0$ and we conclude the proof of the upper bound. $\square$