

# LARGE DEVIATIONS AND METASTABILITY ANALYSIS FOR HEAVY-TAILED DYNAMICAL SYSTEMS

BY XINGYU WANG<sup>1,a</sup> AND CHANG-HAN RHEE<sup>1,b</sup>

<sup>1</sup>Industrial Engineering and Management Sciences, Northwestern University

<sup>a</sup>xingyuwang2017@u.northwestern.edu; <sup>b</sup>chang-han.rhee@northwestern.edu

This paper introduces a novel framework that connects large deviations and metastability analysis in heavy-tailed stochastic dynamical systems. Employing this framework in the context of stochastic difference equations  $X_{j+1}^\eta(x) = X_j^\eta(x) + \eta a(X_j^\eta(x)) + \eta \sigma(X_j^\eta(x)) Z_{j+1}$  and its variations with truncated dynamics, we first establish locally uniform sample path large deviations and then translate such asymptotics into a precise 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. Our large deviations asymptotics are sharp enough to identify how rare events arise in heavy-tailed dynamical systems and characterize the *catastrophe principle*. Moreover, it also unveils the discrete hierarchy of phase transitions in the asymptotics of the first exit times and locations under truncated heavy-tailed noises. Our results in this paper open up the possibility of systematic analysis of the global dynamics of heavy-tailed stochastic processes. In the supplementary material, we also present the corresponding results for the Lévy driven SDEs.

## CONTENTS

|       |  |    |
|-------|--|----|
| 1     | Introduction . . . . .   | 2  |
| 2     | Main Results . . . . .   | 7  |
| 2.1   | Uniform $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ -Convergence . . . . . | 7  |
| 2.2   | Heavy-Tailed Large Deviations . . . . .                                      | 9  |
| 2.2.1 | Sample-Path Large Deviations . . . . .                                       | 9  |
| 2.2.2 | Catastrophe Principle . . . . .  | 12 |
| 2.3   | Metastability Analysis . . . . .   | 13 |
| 2.3.1 | First Exit Times and Locations . . . . .                                     | 13 |
| 2.3.2 | General Framework . . . . .  | 15 |
| 3     | Uniform $\mathbb{M}$ -Convergence and Sample Path Large Deviations . . . . . | 16 |
| 3.1   | Proof of Theorem 2.2 . . . . .   | 16 |
| 3.2   | Technical Lemmas for Theorems 2.3 and 2.4 . . . . .                          | 18 |
| 3.3   | Proofs of Theorems 2.3 and 2.4 . . . . .                                     | 29 |
| 3.3.1 | Proof of Proposition 3.10 . . . . .  | 35 |
| 4     | Metastability Analysis . . . . .   | 44 |
| 4.1   | Proof of Theorem 2.8 . . . . .   | 44 |
| 4.2   | Proof of Theorem 2.6 . . . . .   | 49 |
|       | Supplementary Material . . . . .   | 57 |
|       | References . . . . .   | 57 |

*MSC2020 subject classifications:* Primary 60F10, 60G17; secondary 60B10.

*Keywords and phrases:* Sample path large deviations, metastability, regular variation, stochastic difference and differential equations, first exit analysis,  $\mathbb{M}$ -convergence, catastrophe principle.

**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., [39, 32]), large deviations for finite dimensional processes under relaxed assumptions (see, e.g., [17, 20, 19, 1, 21]), Freidlin–Wentzell-type bounds for infinite dimensional processes (see, e.g., [9, 10, 31]), and large deviations for stochastic partial differential equations (see, e.g., [50, 12, 47, 38]), to name a few. On the other hand, the exponential scaling and the pre-exponents in the asymptotics of first exit times under Brownian perturbations were characterized in the Eyring–Kramers law (see [23, 34]). 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 [37]), results for discrete-time processes (see, e.g., [33, 11]), 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 [8].

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 [26, 28, 29, 27], 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 and the metastability 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. 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  $M(S \setminus \mathbb{C})$ -convergence [36, 46]. 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 equations. Moreover, this new formulation of the heavy-tailed large deviations paves the way to the analysis of local stability and global dynamics.
- **Metastability 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 distribution 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 [29] in the context of Weibull tails; see also [30].

In a companion paper [53], we show that the above framework is powerful enough to identify a scaling limit and characterize the global behavior of the heavy-tailed dynamical systems over a multi-well potential at the process level. In particular, the scaling limit is a Markov jump process whose state space consists of the local minima of the potential; under the truncated dynamics, the state space consists of only the widest minima. These findings systematically characterize a curious phenomena that the truncated heavy-tailed processes avoid narrow local minima altogether in the limit. As a result, it can be shown 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 understanding and further improving the curious effectiveness of the stochastic gradient descent (SGD) algorithms in training deep neural networks.

In this paper, we focus on the class of heavy-tailed phenomena captured by the notion of regular variation. To be specific, 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  $x \rightarrow \infty$  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  $(X_j^\eta(x))_{j \geq 0}$  be the solution of the following stochastic difference equation (under initial value  $X_0^\eta(x) = x$ )

$$(1.1) \quad X_{j+1}^\eta(x) = X_j^\eta(x) + \eta a(X_j^\eta(x)) + \eta \sigma(X_j^\eta(x)) Z_{j+1} \quad \forall j \geq 0.$$

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  $X_{j+1}^\eta(x)$  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  $(X_j^{\eta|b}(x))_{j \geq 0}$  with the following recursion

$$(1.2) \quad X_{j+1}^{\eta|b}(x) = X_j^{\eta|b}(x) + \varphi_b\left(\eta a(X_j^{\eta|b}(x)) + \eta \sigma(X_j^{\eta|b}(x)) Z_{j+1}\right) \quad \forall j \geq 0$$

under initial value  $X_0^{\eta|b}(x) = x$ . In other words,  $X_j^{\eta|b}(x)$  is a modulated version of  $X_j^\eta(x)$  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. Note that although (1.1) and (1.2) are probably the most natural scaling regime, more general scaling can be considered. In Section A of the Supplementary Material, we present the corresponding results under more general scaling regimes, i.e., (under initial value  $X_0^\eta(x) = x$  and  $X_0^{\eta|b}(x) = x$ )

$$(1.3) \quad \begin{aligned} X_j^\eta(x) &= X_{j-1}^\eta(x) + \eta a(X_{j-1}^\eta(x)) + \eta^\gamma \sigma(X_{j-1}^\eta(x)) Z_j \quad \forall j \geq 1; \\ X_j^{\eta|b}(x) &= X_{j-1}^{\eta|b}(x) + \varphi_b\left(\eta a(X_{j-1}^{\eta|b}(x)) + \eta^\gamma \sigma(X_{j-1}^{\eta|b}(x)) Z_j\right) \quad \forall j \geq 1. \end{aligned}$$

with some  $\gamma > 0$ . Similar results hold for Lévy-driven SDEs, which are summarized in Section B of the Supplementary Material.

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 [40, 41]. The summary of the subsequent developments in the context of processes with independent increments can be found in, for example, [6, 18, 22, 24]. 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 [25]. More recently, [46] 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  $\mathbf{S}^n \triangleq \{S_{[nt]}^n/n : t \in [0, 1]\}$  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 [46] takes the following form: for “general”  $B \in \mathbb{D}$ ,

$$(1.4) \quad \begin{aligned} 0 < \mathbf{C}_k(B^\circ) &\leq \liminf_{n \rightarrow \infty} \frac{\mathbf{P}(\mathbf{S}^n \in B)}{(n\mathbf{P}(|Z_1| > n))^k} \\ &\leq \limsup_{n \rightarrow \infty} \frac{\mathbf{P}(\mathbf{S}^n \in B)}{(n\mathbf{P}(|Z_1| > n))^k} \leq \mathbf{C}_k(B^-) < \infty, \end{aligned}$$

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. See also [4] where asymptotic bounds similar to (1.4) were obtained for random walks under more general scaling.

Note that in contrast to the standard log-asymptotics in the classical LDP framework, (1.4) provides exact asymptotics. This formulation provides a powerful framework in heavy-tailed contexts; for instance, this formulation has enabled the design and analysis of strongly efficient rare-event simulation algorithms for a wide variety of rare events associated with  $\mathbf{S}^n$ , as demonstrated in [15]. Moreover, [46, Section 4.4] proves that it is impossible to establish the classical large deviation principle w.r.t.  $J_1$  topology at the sample-path level for regularly varying Lévy processes. On a related note, it is worth mentioning that by relaxing the upper bound of the standard LDP, an alternative formulation known as “extended LDP” was proposed in [7], and such a formulation is also feasible for heavy-tailed processes; see, for example, [5, 2, 3]. However, the extended LDP only provides log-asymptotics. For regularly varying processes, it is often desirable and possible to obtain exact asymptotics; for example, the extended LDP wouldn’t suffice for analyzing the strong efficiency of the aforementioned rare-event simulation algorithm in [15]. We will also see that exact asymptotics are crucial in Section 2.3 and Section 4 for sharp exit time and exit location analysis. In fact, it demands an even stronger version of (1.4), which we will introduce shortly. Below, we describe the main contributions of this paper in greater 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 metastability analysis. To be specific, let  $\mathbf{X}^\eta(x) \triangleq \{X_{\lfloor t/\eta \rfloor}^\eta(x) : t \in [0, 1]\}$  be the time-scaled version of  $X_j^\eta(x)$  embedded in the continuous time, and note that  $\mathbf{X}^\eta(x)$  is a random element in  $\mathbb{D}$ . As  $\eta$  decreases,  $\mathbf{X}^\eta(x)$  converges to a deterministic limit  $\{\mathbf{y}_t(x) : t \in [0, 1]\}$ , where  $d\mathbf{y}_t(x)/dt = a(\mathbf{y}_t(x))$  with initial value  $\mathbf{y}_0(x) = x$ . 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  $k$ :

$$(1.5) \quad \begin{aligned} \inf_{x \in A} \mathbf{C}^{(k)}(B^\circ; x) &\leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}(\mathbf{X}^\eta(x) \in B)}{(\eta^{-1} \mathbf{P}(|Z_1| > \eta^{-1}))^k} \\ &\leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}(\mathbf{X}^\eta(x) \in B)}{(\eta^{-1} \mathbf{P}(|Z_1| > \eta^{-1}))^k} \leq \sup_{x \in A} \mathbf{C}^{(k)}(B^-; x). \end{aligned}$$

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 (1.5) is the minimum number of jumps that needs to be added to the path of  $\mathbf{y}_t(x)$  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 (1.5) become tighter, and hence, (1.5) is a *locally uniform* version of the large deviations formulation in (1.4). An important implication of (1.5) is a sharp characterization of the catastrophe principle. Specifically, Section 2.2.2 proves that the conditional distribution of  $\mathbf{X}^\eta(x)$  given the rare event of interest converges to the distribution of a piecewise deterministic random function  $\mathbf{X}_{|B}^*(x)$  with precisely  $k$  random jumps whose sizes are bounded from below:

$$\mathcal{L}(\mathbf{X}^\eta(x) | \mathbf{X}^\eta(x) \in B) \rightarrow \mathcal{L}(\mathbf{X}_{|B}^*(x)).$$

Note that the perturbation associated with  $Z_i$  is modulated by  $\eta \sigma(X_{i-1}^\eta(x))$ . 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  $\{\mathbf{X}^\eta(x) \in B\}$  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 [36] and further developed in [46], was a key technical tool behind (1.4). In this paper, we introduce a uniform version of the  $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ -convergence to establish the uniform asymptotics in (1.5) and prove an associated Portmanteau theorem (Theorem 2.2) in Section 2.1. These developments form the backbone that supports our proofs of the uniform sample-path large deviations in (1.5). Furthermore, we also establish the locally uniform asymptotics for  $\mathbf{X}^{\eta|b}(x)$  in Theorem 2.4. As Section 2.3 elaborates, such large deviations of  $\mathbf{X}^{\eta|b}(x)$  leads to exit times and locations with structurally different asymptotic limits compared to those associated with  $\mathbf{X}^\eta(x)$ .

**Metastability Analysis.** The second contribution of this paper is the first exit-time analysis for heavy-tailed systems. The first exit time problem finds applications in numerous contexts, including chemical reactions [34], physics [13, 14], extreme climate events [43], mathematical finance [48], and queueing systems [49]. A classical result in this literature is the Eyring-Kramers law [23, 34], which characterizes the exit time of Brownian particles; see also [37].

Unlike in the classical light-tailed context where 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 [26, 27]. These



results were extended to the multi-dimensional settings in [28] and, later, stochastic difference equations driven by  $\alpha$ -stable noises in [42] 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  $X^{\eta|b}(x)$ . 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  $d\mathbf{y}_t(x)/dt = -U'(\mathbf{y}_t(x))$  with initial condition  $\mathbf{y}_0(x) = x$ , it holds that  $\lim_{t \rightarrow \infty} \mathbf{y}_t(x) = 0$  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  $X_j^\eta(x)$  and  $X_j^{\eta|b}(x)$  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  $\eta \rightarrow \infty$ . Theorem 2.6 establishes that the joint law of the first exit time  $\tau^{\eta|b}(x) = \min\{j \geq 0 : X_j^{\eta|b}(x) \notin I\}$  and the exit location  $X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \triangleq X_{\tau^{\eta|b}(x)}^{\eta|b}(x)$  admits the following limit (for all  $x \in I$ ):

$$(1.6) \quad (\lambda_b^I(\eta) \cdot \tau^{\eta|b}(x), X_{\tau^{\eta|b}(x)}^{\eta|b}(x)) \Rightarrow (E, V_b) \quad \text{as } \eta \downarrow 0$$

with some (deterministic) time-scaling function  $\lambda_b^I(\eta)$ . 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  $\lambda_b^I(\eta)$  are provided in Section 2.3.1. Here, we note that  $\lambda_b^I(\eta)$  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 (2.27) 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  $X_j^\eta(x)$  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 [29] also investigated the exit events driven by multiple jumps. However, the mechanism through which multiple jumps arise in their context is due to 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 [26, 27, 28] for regularly varying SDEs, [29] paints interesting picture of the hierarchy in the asymptotics of the first exit times. See [30] 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 (1.6), where we find that the first exit time  $\tau^{\eta|b}(x)$  is (roughly) of order  $1/\eta^{1+\mathcal{J}_b^*(\alpha-1)}$  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  $\mathcal{J}_b^*$ .

Some of the results in Section 2.3 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 and Section 4 provide the proofs of Sections 2.1, 2.2, and 2.3. Results for stochastic difference equations under more general scaling regimes are presented in Section A of the Supplementary Material. Results for SDEs driven by Lévy processes with regularly varying increments are collected in Section B of the Supplementary Material.

**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, and Section 2.3 carries out the metastability analysis. All the proofs are deferred to the later sections.

Before presenting the main results, we set frequently used notations. Let  $[n] \triangleq \{1, 2, \dots, n\}$  for any positive integer  $n$ . Let  $\mathbb{N} = \{0, 1, 2, \dots\}$  be the set of non-negative integers. Let  $(\mathbb{S}, d)$  be a metric space with  $\mathcal{S}_{\mathbb{S}}$  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  $E^r \triangleq \{y \in \mathbb{S} : d(E, y) \leq r\}$  be the  $r$ -enlargement of a set  $E$ . Here for any set  $A \subseteq \mathbb{S}$  and any  $x \in \mathbb{S}$ , we define  $d(A, x) \triangleq \inf\{d(y, x) : y \in A\}$ . 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  $\inf_{x \in A, y \in B} d(x, y) > 0$ . For any Borel measure  $\mu$  on  $(\mathbb{S}, \mathcal{S}_{\mathbb{S}})$ , 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  $g : \mathbb{S} \rightarrow \mathbb{R}$ , let  $\text{supp}(g) \triangleq (\{x \in \mathbb{S} : g(x) \neq 0\})^-$ .

**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 [36, 46] 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  $(\mathbb{S}, d)$  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  $\mathcal{S}_{\mathbb{S} \setminus \mathbb{C}} \triangleq \{A \in \mathcal{S}_{\mathbb{S}} : A \subseteq \mathbb{S} \setminus \mathbb{C}\}$ . 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  $n \rightarrow \infty$  if  $\lim_{n \rightarrow \infty} |\mu_n(f) - \mu(f)| = 0$  for all  $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ . See [36] 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 [46], 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  $\mu_\theta^\eta, \mu_\theta \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$  for each  $\eta > 0$  and  $\theta \in \Theta$ . We say that  $\mu_\theta^\eta$  converges to  $\mu_\theta$  in  $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$  uniformly in  $\theta$  on  $\Theta$  as  $\eta \rightarrow 0$  if

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

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  $\mu_\theta^\eta, \mu_\theta \in \mathbb{M}(\mathbb{S} \setminus \mathbb{C})$  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*

$$(2.1) \quad \lim_{k \rightarrow \infty} \mu_{\theta_{n_k}}(f) = \mu_{\theta^*}(f) \quad \forall f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C}),$$

*then the next three statements are equivalent:*

- (i)  $\mu_\theta^\eta$  converges to  $\mu_\theta$  in  $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$  uniformly in  $\theta$  on  $\Theta$  as  $\eta \downarrow 0$ ;
- (ii)  $\lim_{\eta \downarrow 0} \sup_{\theta \in \Theta} |\mu_\theta^\eta(f) - \mu_\theta(f)| = 0$  for each  $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$  that is also uniformly continuous on  $\mathbb{S}$ ;
- (iii)  $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu_\theta^\eta(F) - \mu_\theta(F^\epsilon) \leq 0$  and  $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu_\theta^\eta(G) - \mu_\theta(G_\epsilon) \geq 0$  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)  $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu_\theta^\eta(F) \leq \sup_{\theta \in \Theta} \mu_\theta(F)$  and  $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu_\theta^\eta(G) \geq \inf_{\theta \in \Theta} \mu_\theta(G)$  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

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

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  $\{\mu_\theta^\eta : \theta \in \Theta\}$  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

$$\mu_\theta^\eta \triangleq \delta_{\theta-\eta}, \quad \mu_\theta \triangleq \delta_\theta,$$

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

$$\sup_{\theta \in \Theta} \mu_\theta^\eta(F) - \mu_\theta(F) \geq \delta_{-\eta/2}([-1, 0]) - \delta_{\eta/2}([-1, 0]) \quad \text{by picking } \theta = \eta/2$$

$$= \mathbb{I}\left\{\frac{-\eta}{2} \in [-1, 0]\right\} - \mathbb{I}\left\{\frac{\eta}{2} \in [-1, 0]\right\} = 1,$$

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

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  $\mu_\theta^\eta$  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  $\mathbb{C} = \emptyset$ ,  $\mathbb{S} = \mathbb{R}$ ,  $\Theta = [-1, 1]$ ,  $\mu_\theta^\eta = \delta_{-\theta}$ , and  $\mu_\theta = \delta_\theta$ .



**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  $Z_1, Z_2, \dots$  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  $Z_1, Z_2, \dots, Z_j$  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  $\{X_j^\eta(x) : j \geq 0\}$ , which is driven by the recursion

$$(2.3) \quad X_0^\eta(x) = x; \quad X_j^\eta(x) = X_{j-1}^\eta(x) + \eta a(X_{j-1}^\eta(x)) + \eta \sigma(X_{j-1}^\eta(x)) Z_j, \quad \forall j \geq 1$$

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  $\phi : (0, \infty) \rightarrow (0, \infty)$ , we say that  $\phi$  is regularly varying as  $x \rightarrow \infty$  with index  $\beta$  (denoted as  $\phi(x) \in \mathcal{RV}_\beta(x)$  as  $x \rightarrow \infty$ ) if  $\lim_{x \rightarrow \infty} \phi(tx)/\phi(x) = t^\beta$  for all  $t > 0$ . For details of the definition and properties of regularly varying functions, see, for example, Chapter 2 of [45]. 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

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

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

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

Next, we introduce the following assumptions on the drift coefficient  $a : \mathbb{R} \rightarrow \mathbb{R}$  and diffusion coefficient  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ . 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

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

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

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

Here,  $\Lambda_T$  is the set of all homeomorphism on  $[0, T]$ . Given any  $A \subseteq \mathbb{R}$ , let  $A^{k\uparrow} \triangleq \{(t_1, \dots, t_k) \in A^k : t_1 < t_2 < \dots < t_k\}$  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  $h_{[0, T]}^{(k)} : \mathbb{R} \times \mathbb{R}^k \times (0, T]^{k\uparrow} \rightarrow \mathbb{D}[0, T]$  as follows. Given any  $x_0 \in \mathbb{R}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ , and  $\mathbf{t} = (t_1, \dots, t_k) \in (0, T]^{k\uparrow}$ , let  $\xi = h_{[0, T]}^{(k)}(x_0, \mathbf{w}, \mathbf{t}) \in \mathbb{D}[0, T]$  be the solution to

$$(2.6) \quad \xi_0 = x_0$$

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

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

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  $h_{[0, T]}^{(k)}(x_0, \mathbf{w}, \mathbf{t})$  produces the ODE path perturbed by jumps  $w_1, \dots, w_k$  (modulated by the drift coefficient  $\sigma(\cdot)$ ) at times  $t_1, \dots, t_k$ . We adopt the convention that  $\xi = h_{[0, T]}^{(0)}(x_0)$  is the solution to the ODE  $d\xi_s/ds = a(\xi_s) \quad \forall s \in [0, T]$  under the initial condition  $\xi_0 = x_0$ . For any  $\alpha > 1$ , let  $\nu_\alpha$  be the (Borel) measure on  $\mathbb{R}$  with

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

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

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

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

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

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

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

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  $\mathbb{D}$ ,  $\mathbf{d}_{J_1}$ ,  $h^{(k)}$ ,  $\mathbf{C}^{(k)}$ ,  $\mathbb{D}_A^{(k)}$ , and  $\mathbf{X}^\eta(x)$  to denote  $\mathbb{D}[0, 1]$ ,  $\mathbf{d}_{J_1}^{[0, 1]}$ ,  $h_{[0, 1]}^{(k)}$ ,  $\mathbf{C}_{[0, 1]}^{(k)}$ ,  $\mathbb{D}_A^{(k)}[0, 1]$ , and  $\mathbf{X}_{[0, 1]}^\eta(x)$ , respectively.

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

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

uniformly in  $x$  on  $A$  as  $\eta \rightarrow 0$ . Furthermore, for any  $B \in \mathcal{S}_{\mathbb{D}[0,T]}$  that is bounded away from  $\mathbb{D}_A^{(k-1)}[0,T]$ ,

$$(2.13) \quad \inf_{x \in A} \mathbf{C}_{[0,T]}^{(k)}(B^\circ; x) \leq \liminf_{\eta \downarrow 0} \frac{\inf_{x \in A} \mathbf{P}(X_{[0,T]}^\eta(x) \in B)}{\lambda^k(\eta)} \\ \leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A} \mathbf{P}(X_{[0,T]}^\eta(x) \in B)}{\lambda^k(\eta)} \leq \sup_{x \in A} \mathbf{C}_{[0,T]}^{(k)}(B^-; x) < \infty.$$

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  $\inf_{x \in A} \mathbf{P}(X_{[0,T]}^\eta(x) \in B)$  and  $\sup_{x \in A} \mathbf{P}(X_{[0,T]}^\eta(x) \in B)$  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  $X_{[0,T]}^\eta(x)$  will degenerate to a Lévy process, and  $\mathcal{J}_{A;T}(\cdot)$  will reduce to  $\mathcal{J}(\cdot)$  defined in equation (3.3) of [46]. This confirms that Theorem 2.3 is a proper generalization of the heavy-tailed large deviations for Lévy processes and random walks in [46].

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

$$(2.14) \quad X_j^{\eta|b}(x) = X_{j-1}^{\eta|b}(x) + \varphi_b\left(\eta a(X_{j-1}^{\eta|b}(x)) + \eta \sigma(X_{j-1}^{\eta|b}(x)) Z_j\right) \quad \forall j \geq 1,$$

under initial value  $X_0^{\eta|b}(x) = x$ , where the truncation operator  $\varphi(\cdot)$  is defined as

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

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  $\mathbf{X}_{[0,T]}^{\eta|b}(x) \triangleq \{X_{[t/\eta]}^{\eta|b}(x) : t \in [0, T]\}$  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  $h_{[0,T]}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0, T]^{k\uparrow} \rightarrow \mathbb{D}[0, T]$  as follows. Given any  $x_0 \in \mathbb{R}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ , and  $\mathbf{t} = (t_1, \dots, t_k) \in (0, T]^{k\uparrow}$ , let  $\xi = h_{[0,T]}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t})$  be the solution to

$$(2.16) \quad \xi_0 = x_0;$$

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

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

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,  $h_{[0,T]}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t})$  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  $k = 0, 1, 2, \dots$ , let

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

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 C of the Supplementary Material.

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

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

Again, in case that  $T = 1$ , we set  $\mathbf{X}^{\eta|b}(x) \triangleq \mathbf{X}_{[0, 1]}^{\eta|b}(x)$ ,  $h^{(k)|b} \triangleq h_{[0, 1]}^{(k)|b}$ ,  $\mathbb{D}_A^{(k)|b} \triangleq \mathbb{D}_A^{(k)|b}[0, 1]$ , and  $\mathbf{C}^{(k)|b} \triangleq \mathbf{C}_{[0, 1]}^{(k)|b}$ . Now, we are ready to state the main result. It is worth noticing that Assumption 4 is not required in the truncated case. 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  $\lambda^{-k}(\eta)\mathbf{P}(\mathbf{X}_{[0, T]}^{\eta|b}(x) \in \cdot) \rightarrow \mathbf{C}_{[0, T]}^{(k)|b}(\cdot; x)$  in  $\mathbb{M}(\mathbb{D}[0, T] \setminus \mathbb{D}_A^{(k-1)|b}[0, T])$  uniformly in  $x$  on  $A$  as  $\eta \rightarrow 0$ . Furthermore, for any  $B \in \mathcal{S}_{\mathbb{D}[0, T]}$  that is bounded away from  $\mathbb{D}_A^{(k-1)|b}[0, T]$ ,*

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

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  $\mathbf{X}_{[0, T]}^{\eta|b}(x)$  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  $b \rightarrow \infty$  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  $\mathbf{X}^\eta(x)$  and  $\mathbf{X}^{\eta|b}(x)$ , 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  $\mathbb{D}_{\{x\}}^{(k-1)|b}$ ,  $B \cap \mathbb{D}_{\{x\}}^{(k)|b} \neq \emptyset$ , and  $\mathbf{C}^{(k)|b}(B^\circ) = \mathbf{C}^{(k)|b}(B^-) > 0$ . Then*

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

- (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  $\mathbb{D}_{\{x\}}^{(k-1)}$ ,  $B \cap \mathbb{D}_{\{x\}}^{(k)} \neq \emptyset$ , and  $\mathbf{C}^{(k)}(B^\circ) = \mathbf{C}^{(k)}(B^-) > 0$ . Then*

$$\mathbf{P}(X^\eta(x) \in \cdot \mid X^\eta(x) \in B) \Rightarrow \frac{\mathbf{C}^{(k)}(\cdot \cap B; x)}{\mathbf{C}^{(k)}(B; x)} \quad \text{as } \eta \downarrow 0.$$

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

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

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 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  $X^\eta(x)$  in part (ii).

**2.3. Metastability Analysis.** This section analyzes the metastability of  $X_j^\eta(x)$  and  $X_j^{\eta|b}(x)$ . 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 and Locations.** In this section, we analyze the first exit times and locations of  $X_j^\eta(x)$  and  $X_j^{\eta|b}(x)$  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  $\lim_{t \rightarrow \infty} \mathbf{y}_t(x) = 0$  holds for all  $x \in I$  where  $\mathbf{y}_t(x)$  is the solution of ODE

$$(2.22) \quad \mathbf{y}_0(x) = x, \quad \frac{d\mathbf{y}_t(x)}{dt} = a(\mathbf{y}_t(x)) \quad \forall t \geq 0.$$



It is worth noticing that Assumption 5 is more flexible than the assumptions in other related works in the literature. For instance, [27, 26] 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

$$(2.23) \quad \tau^\eta(x) \triangleq \min \{j \geq 0 : X_j^\eta(x) \notin I\}, \quad \tau^{\eta|b}(x) \triangleq \min \{j \geq 0 : X_j^{\eta|b}(x) \notin I\},$$

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

$$(2.24) \quad \check{g}^{(k)|b}(x, \mathbf{w}, \mathbf{t}) \triangleq h_{[0, t_{k+1}]}^{(k)|b}(x, \mathbf{w}, \mathbf{t})(t_k)$$

where  $\mathbf{t} = (t_1, \dots, t_k) \in (0, \infty)^{k\uparrow}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ , and  $h_{[0, T]}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0, T]^{k\uparrow} \rightarrow \mathbb{D}[0, T]$  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  $b > 0$ )

$$(2.25) \quad \check{\mathbf{C}}^{(k)|b}(\cdot; x) \triangleq \int \mathbb{I}\left\{\check{g}^{(k-1)|b}(x + \varphi_b(\sigma(x) \cdot w_0), \mathbf{w}, \mathbf{t}) \in \cdot\right\} \nu_\alpha(dw_0) \times \nu_\alpha^{k-1}(d\mathbf{w}) \times \mathcal{L}_\infty^{k-1\uparrow}(d\mathbf{t})$$

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

$$(2.26) \quad \check{\mathbf{C}}(\cdot; x) \triangleq \int \mathbb{I}\left\{x + \sigma(x) \cdot w \in \cdot\right\} \nu_\alpha(dw).$$

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

$$(2.27) \quad l \triangleq \inf_{x \in I^c} |x| = |s_{\text{left}}| \wedge |s_{\text{right}}|, \quad \mathcal{J}_b^* \triangleq \lceil l/b \rceil.$$

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  $\lambda^k(\eta) = (\lambda(\eta))^k$ . 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  $X_j^{\eta|b}(x)$  and  $X_j^\eta(x)$ . 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  $s_{\text{left}}/b \notin \mathbb{Z}$  and  $s_{\text{right}}/b \notin \mathbb{Z}$ . For any  $\epsilon > 0$ ,  $t \geq 0$ , and measurable set  $B \subseteq I^c$ ,*

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

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

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  $B \subseteq I^c$ ,

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

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

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  $(\mathbb{S}, d)$  and a family of  $\mathbb{S}$ -valued Markov chains  $\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\}$  parameterized by  $\eta$ , where  $x \in \mathbb{S}$  denotes the initial state and  $j$  denotes the time index. We use  $\mathbf{V}_{[0,T]}^\eta(x) \triangleq \{V_{\lfloor t/\eta \rfloor}^\eta(x) : t \in [0, T]\}$  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  $\mathbf{V}_{[0,T]}^\eta(x)$  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,  $V^\eta(x)$  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.**  $\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\}$  possesses an asymptotic atom  $\{A(\epsilon) \subseteq \mathbb{S} : \epsilon > 0\}$  associated with the domain  $I$ , location measure  $C(\cdot)$ , scale  $\gamma : (0, \infty) \rightarrow (0, \infty)$ , and covering  $\{I(\epsilon) \subseteq I : \epsilon > 0\}$  if the following holds: For each measurable set  $B \subseteq \mathbb{S}$ , there exist  $\delta_B : (0, \infty) \times (0, \infty) \rightarrow (0, \infty)$ ,  $\epsilon_B > 0$ , and  $T_B > 0$  such that

(2.28)

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

$$(2.29) \quad \leq \limsup_{\eta \downarrow 0} \frac{\sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \leq \frac{T}{\eta}; V_{\tau_\epsilon}^\eta(x) \in B)}{\gamma(\eta)T/\eta} \leq C(B^-) + \delta_B(\epsilon, T)$$

$$(2.30) \quad \limsup_{\eta \downarrow 0} \frac{\sup_{x \in I(\epsilon)} \mathbf{P}(\tau_{(I(\epsilon) \setminus A(\epsilon))^c}^\eta(x) > \frac{T}{\eta})}{\gamma(\eta)T/\eta} = 0$$

$$(2.31) \quad \liminf_{\eta \downarrow 0} \inf_{x \in I(\epsilon)} \mathbf{P}(\tau_{A(\epsilon)}^\eta(x) \leq T/\eta) = 1$$

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

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

To see how Definition 2.7 asymptotically characterize the atoms in  $V^\eta(x)$  for the first exit analysis from domain  $I$ , note that the condition (2.31) requires the process to efficiently return to the asymptotic atoms  $A(\epsilon)$ . The conditions (2.28) and (2.29) 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 (2.30) prevents the process  $V^\eta(x)$  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  $\{\{V_j^\eta(x) : j \geq 0\} : \eta > 0\}$  of Markov chains possesses an asymptotic atom  $\{A(\epsilon) \subseteq \mathbb{S} : \epsilon > 0\}$  associated with the domain  $I$ , location measure  $C(\cdot)$ , scale  $\gamma : (0, \infty) \rightarrow (0, \infty)$ , 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  $\tau_{I^c}^\eta(x)$  scales as  $1/\gamma(\eta)$ , and the distribution of the location  $V_\tau^\eta(x)$  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  $t \geq 0$ ,*

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

**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  $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu_\theta^\eta(F) - \mu_\theta(F^\epsilon) \leq 0$  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  $\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) > \delta \forall n \geq 1$ . Now, we make two observations. First, using Urysohn's lemma (see, e.g., lemma 2.3 of [36]), one can identify some  $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ , which is also uniformly continuous on  $\mathbb{S}$ , such that  $\mathbb{I}_F \leq f \leq \mathbb{I}_{F^\epsilon}$ . This leads to the bound  $\mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq \mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)$  for each  $n$ . Secondly, from statement (ii) we get  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)| = 0$ . In summary, we yield the contradiction

$$\limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq \limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f) \leq \lim_{n \rightarrow \infty} |\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)| = 0.$$

Analogously, if the claim  $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu_\theta^\eta(G) - \mu_\theta(G^\epsilon) \geq 0$ , 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  $\mathbb{I}_{G^\epsilon} \leq g \leq \mathbb{I}_G$ . 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  $\lim_{\eta \downarrow 0} \sup_{\theta \in \Theta} |\mu_\theta^\eta(g) - \mu_\theta(g)| = 0$  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

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

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  $\limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq 0$  and  $\liminf_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(G) - \mu_{\theta_n}(G^\epsilon) \geq 0$  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  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}(f) - \mu_{\theta^*}(f)| = 0$  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 [36]), we yield  $\limsup_{n \rightarrow \infty} \mu_{\theta_n}(F^\epsilon) \leq \mu_{\theta^*}(F^\epsilon)$  and  $\liminf_{n \rightarrow \infty} \mu_{\theta_n}(G^\epsilon) \geq \mu_{\theta^*}(G^\epsilon)$ . In summary, for any  $\epsilon > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(F) &\leq \limsup_{n \rightarrow \infty} \mu_{\theta_n}(F^\epsilon) + \limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(F) - \mu_{\theta_n}(F^\epsilon) \leq \mu_{\theta^*}(F^\epsilon), \\ \liminf_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(G) &\geq \liminf_{n \rightarrow \infty} \mu_{\theta_n}(G^\epsilon) + \liminf_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(G) - \mu_{\theta_n}(G^\epsilon) \geq \mu_{\theta^*}(G^\epsilon). \end{aligned}$$

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  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}^{\eta_n}(g) - \mu_{\theta^*}(g)| = 0$  for the  $g \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$  fixed in (3.1). However, recall that we have already obtained  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}(g) - \mu_{\theta^*}(g)| = 0$  using assumption (2.1). We now arrive at the contradiction

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

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  $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu_\theta^\eta(F) \leq \sup_{\theta \in \Theta} \mu_\theta(F)$  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  $\mu_{\theta_n}^{\eta_n}(F) > \sup_{\theta \in \Theta} \mu_\theta(F) + \delta \forall n \geq 1$ . 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  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}(f) - \mu_{\theta^*}(f)| = 0$  for all  $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ . Meanwhile,  $(i)$  implies that  $\lim_{n \rightarrow \infty} |\mu_{\theta_n}^{\eta_n}(f) - \mu_{\theta_n}(f)| = 0$  for all  $f \in \mathcal{C}(\mathbb{S} \setminus \mathbb{C})$ . Therefore,

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

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  $\limsup_{n \rightarrow \infty} \mu_{\theta_n}^{\eta_n}(F) \leq \mu_{\theta^*}(F) \leq \sup_{\theta \in \Theta} \mu_\theta(F)$ . In summary, we have established the claim  $\limsup_{\eta \downarrow 0} \sup_{\theta \in \Theta} \mu_\theta^\eta(F) \leq \sup_{\theta \in \Theta} \mu_\theta(F)$  for all closed  $F$  bounded away from  $\mathbb{C}$ . The same approach can also be applied to show  $\liminf_{\eta \downarrow 0} \inf_{\theta \in \Theta} \mu_\theta^\eta(G) \geq \inf_{\theta \in \Theta} \mu_\theta(G)$  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 [16].

DEFINITION 3.1. Let  $X_n$  and  $Y_n$  be random elements taking values in a complete separable metric space  $(\mathbb{S}, \mathbf{d})$ . 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  $B \in \mathcal{S}_{\mathbb{S}}$  bounded away from  $\mathbb{C}$ ,

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

In case that  $\mathbb{C} = \emptyset$ , Definition 3.1 simply degenerates to the standard notion of asymptotic equivalence; see definition 1 of [46]. 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 [16]). *Let  $X_n$  and  $Y_n$  be random elements taking values in a complete separable metric space  $(\mathbb{S}, \mathbf{d})$  and let  $\mathbb{C} \subseteq \mathbb{S}$  be Borel measurable. Suppose that  $\epsilon_n^{-1} \mathbf{P}(X_n \in \cdot) \rightarrow \mu(\cdot)$  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  $\epsilon_n^{-1} \mathbf{P}(Y_n \in \cdot) \rightarrow \mu(\cdot)$  in  $\mathbb{M}(\mathbb{S} \setminus \mathbb{C})$ .*

3.2. *Technical Lemmas for Theorems 2.3 and 2.4.* Our analysis hinges on the separation of *large noises* among  $(Z_j)_{j \geq 1}$  from the rest, and we pay special attention to  $Z_j$ 's that are large enough so that  $\eta|Z_j|$  exceed 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

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

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

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

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

Recall the definition of filtration  $\mathbb{F} = (\mathcal{F}_j)_{j \geq 0}$  where  $\mathcal{F}_j$  is the  $\sigma$ -algebra generated by  $Z_1, Z_2, \dots, Z_j$  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  $\Gamma_M$  denote the collection of families of random variables, over which we will prove the uniform asymptotics:

$$(3.5) \quad \Gamma_M \triangleq \left\{ (V_j)_{j \geq 0} \text{ is adapted to } \mathbb{F} : |V_j| \leq M \quad \forall j \geq 0 \text{ almost surely} \right\}.$$

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*

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



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

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

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

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

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

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

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

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

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

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,

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

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

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

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

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

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

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

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

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

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

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

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  $\eta \in (0, \eta_0)$ ,

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

from which we immediately get (3.8).

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

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

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

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

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 upper and lower bounds of Burkholder-Davis-Gundy inequality (Theorem 48, Chapter IV of [44]), 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  $s \in [0, t]$ ,  $y \geq 1$

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

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

$$\begin{aligned} \mathbf{E}[(\widehat{Z}_1^{(1)})^2] &= \mathbf{E}[(Z_1^{(1)} - \mathbf{E}Z_1^{(1)})^2] \\ &\leq 2\mathbf{E}[(Z_1^{(1)})^2] + 2[\mathbf{E}Z_1^{(1)}]^2 \leq 2\mathbf{E}[(Z_1^{(1)})^2] + 2[\mathbf{E}|Z_1^{(1)}|]^2. \end{aligned}$$

Also, it has been shown earlier that  $\mathbf{E}|Z_1^{(1)}| \in \mathcal{RV}_{(\alpha-1)\beta}(\eta)$ , and hence  $[\mathbf{E}|Z_1^{(1)}|]^2 \in \mathcal{RV}_{2(\alpha-1)\beta}(\eta)$ . 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  $\frac{t}{\eta} \cdot \eta^2 \cdot 2[\mathbf{E}|Z_1^{(1)}|]^2 < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$  for any  $\eta > 0$  sufficiently small. Next,  $\mathbf{E}[(Z_1^{(1)})^2] = \int_0^\infty 2x\mathbf{P}(|Z_1^{(1)}| > x)dx = \int_0^{1/\eta^\beta} 2x\mathbf{P}(|Z_1| > x)dx$ . 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  $\frac{t}{\eta} \cdot \eta^2 \cdot 2\mathbf{E}[(Z_1^{(1)})^2] < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$ . 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  $\frac{t}{\eta} \cdot \eta^2 \cdot 2\mathbf{E}[(Z_1^{(1)})^2] < \frac{1}{6}\eta^{1-\beta+\frac{2N}{p}}$ . In summary,

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

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

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

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

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

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  $I(\eta) \triangleq \#\{i \leq \lfloor t/\eta \rfloor : Z_i^{(2)} \neq 0\}$ . Besides, fix  $\delta_0 = \frac{\epsilon}{3MJ}$ . For any  $\delta \in (0, \delta_0)$  and  $(V_i)_{i \geq 0} \in \mathbf{\Gamma}_M$ , note that on event  $\{I(\eta) < J\}$ , we must have  $\max_{j \leq \lfloor t/\eta \rfloor} \eta \left|\sum_{i=1}^j V_{i-1} Z_i^{(2)}\right| < \eta \cdot M \cdot J \cdot \delta_0/\eta < MJ\delta_0 < \epsilon/3$ . On the other hand,

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

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

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

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  $\sigma(X_n^{\eta|b}(x)) \leq C$ . Therefore,  $\{\sigma(X_i^{\eta|b}(x))\}_{i \geq 0} \in \Gamma_C$ . From the strong Markov property at stopping times  $(\tau_i^{>\delta}(\eta))_{j \geq 1}$ ,

$$\begin{aligned} \sup_{x \in \mathbb{R}} \mathbf{P} \left( \left( \bigcap_{i=1}^k A_i(\eta, b, \epsilon, \delta, x) \right)^c \right) &\leq \sum_{i=1}^k \sup_{x \in \mathbb{R}} \mathbf{P} \left( (A_i(\eta, b, \epsilon, \delta, x))^c \right) \\ &\leq k \cdot \sup_{(V_i)_{i \geq 0} \in \Gamma_C} \mathbf{P} \left( \max_{j \leq \lfloor 1/\eta \rfloor \wedge (\tau_1^{>\delta}(\eta)-1)} \eta \left| \sum_{i=1}^j V_{i-1} Z_i \right| > \epsilon/2 \right) \end{aligned}$$

where  $C < \infty$  is the constant in Assumption 4 and the set  $\Gamma_C$  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

$$\sup_{(V_i)_{i \geq 0} \in \Gamma_C} \mathbf{P} \left( \max_{j \leq \lfloor 1/\eta \rfloor \wedge (\tau_1^{>\delta}(\eta)-1)} \eta \left| \sum_{i=1}^j V_{i-1} Z_i \right| > \epsilon/2 \right) = o(\eta^N)$$

(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  $(\tau_j^{>\delta}(\eta))_{j \geq 1}$  and  $(W_j^{>\delta}(\eta))_{j \geq 1}$  conditioned on event

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

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

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

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

LEMMA 3.4. *Let Assumption 1 hold. For any  $\delta > 0, c \geq \delta$  and  $k \in \mathbb{Z}^+$ ,*

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

and

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

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

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

For  $q_1(\eta)$ , note that

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

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

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

and hence,

$$(3.20) \quad \begin{aligned} \lim_{\eta \downarrow 0} \frac{q_3(\eta) \cdot (\mathbf{P}(\eta|W_1^{>\delta}(\eta)| > c))^k}{(H(1/\eta))^k} &= \lim_{\eta \downarrow 0} \frac{(H(\delta/\eta))^k \cdot (H(c/\eta)/H(\delta/\eta))^k}{(H(1/\eta))^k} \\ &= \lim_{\eta \downarrow 0} \frac{(H(c/\eta))^k}{(H(1/\eta))^k} = \frac{1}{c^{\alpha k}}. \end{aligned}$$

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

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

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

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

As a result, we must have

$$\mathcal{L}(\eta W_1^{>\delta}(\eta), \eta W_2^{>\delta}(\eta), \dots, \eta W_k^{>\delta}(\eta) \mid E_{c,k}^\delta(\eta)) \rightarrow \mathcal{L}(W_1^*(c), \dots, W_k^*(c)).$$

Moreover, notice that the sequences  $\eta W_1^{>\delta}(\eta), \dots, \eta W_k^{>\delta}(\eta)$  and  $\eta \tau_1^{>\delta}(\eta), \dots, \eta \tau_k^{>\delta}(\eta)$  are conditionally independent on event  $E_{c,k}^\delta(\eta)$ . Indeed, for any  $1 \leq i_1 < \dots < i_k \leq \lfloor 1/\eta \rfloor$  and  $c_1, \dots, c_k > c$ ,

$$\begin{aligned} &\frac{\mathbf{P}(\tau_j^{>\delta}(\eta) = i_j \text{ and } \eta|W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k])}{\mathbf{P}(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta); \ \eta|W_j^{>\delta}(\eta)| > c \ \forall j \in [k])} \\ &= \frac{\mathbf{P}(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1) \mathbf{P}(\eta|W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k])}{\mathbf{P}(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)) \mathbf{P}(\eta|W_j^{>\delta}(\eta)| > c \ \forall j \in [k])} \end{aligned}$$

due to the independence between  $(\tau_i^{>\delta}(\eta))_{i \geq 1}$  and  $(W_i^{>\delta}(\eta))_{i \geq 1}$



$$\begin{aligned}
&= \mathbf{P}\left(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1 \mid \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\right) \\
&\quad \cdot \mathbf{P}\left(\eta|W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k] \mid \eta|W_j^{>\delta}(\eta)| > c \ \forall j \in [k]\right) \\
&= \mathbf{P}\left(\tau_j^{>\delta}(\eta) = i_j \ \forall j \geq 1 \mid \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta); \eta|W_j^{>\delta}(\eta)| > c \ \forall j \in [k]\right) \\
&\quad \cdot \mathbf{P}\left(\eta|W_j^{>\delta}(\eta)| > c_j \ \forall j \in [k] \mid \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta); \eta|W_j^{>\delta}(\eta)| > c \ \forall j \in [k]\right).
\end{aligned}$$

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

$$\mathcal{L}(\eta \tau_1^{>\delta}(\eta), \eta \tau_2^{>\delta}(\eta), \dots, \eta \tau_k^{>\delta}(\eta) \mid E_{c,k}^\delta(\eta)) \rightarrow \mathcal{L}(U_{(1;k)}, \dots, U_{(k;k)}).$$

Note that since both  $\{\eta \tau_i^{>\delta}(\eta) : i = 1, \dots, k\}$  and  $\{U_{(i;k)} : i = 1, \dots, k\}$  are sorted in an ascending order, the joint CDFs are completely characterized by  $\{t_i : i = 1, \dots, k\}$ 's such that  $0 \leq t_1 \leq t_2 \leq \dots \leq t_k \leq 1$ . For any such  $(t_1, \dots, t_k) \in [0, 1]^k$ , note that

$$\begin{aligned}
&\mathbf{P}\left(\eta \tau_1^{>\delta}(\eta) > t_1, \eta \tau_2^{>\delta}(\eta) > t_2, \dots, \eta \tau_k^{>\delta}(\eta) > t_k \mid E_{c,k}^\delta(\eta)\right) \\
&= \mathbf{P}\left(\eta \tau_1^{>\delta}(\eta) > t_1, \eta \tau_2^{>\delta}(\eta) > t_2, \dots, \eta \tau_k^{>\delta}(\eta) > t_k \mid \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\right) \\
&= \frac{\mathbf{P}\left(\eta \tau_1^{>\delta}(\eta) > t_1, \eta \tau_2^{>\delta}(\eta) > t_2, \dots, \eta \tau_k^{>\delta}(\eta) > t_k; \tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\right)}{\mathbf{P}\left(\tau_k^{>\delta}(\eta) < \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\right)}
\end{aligned}$$

and observe that

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

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

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

Together with (3.19), we obtain

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

and conclude the proof.  $\square$

Next, we present several results about mappings  $h_{[0,T]}^{(k)}$  defined in (2.6)–(2.8) and  $h_{[0,T]}^{(k)|b}$  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 and 3.6 in Section C of the Supplementary Material.

Recall the definitions of the sets  $\mathbb{D}_A^{(k)}$  and  $\mathbb{D}_A^{(k)|b}$  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  $\mathbb{D}_A^{(k)}$  and  $\mathbb{D}_A^{(k)|b}$  when Assumptions 2 and 4 hold.

**LEMMA 3.5.** *Let Assumptions 2 and 4 hold. Let  $A \subseteq \mathbb{R}$  be compact and let  $B \in \mathcal{S}_{\mathbb{D}}$ . Let  $k = 0, 1, 2, \dots$ . If  $B$  is bounded away from  $\mathbb{D}_A^{(k-1)}$ , then there exist  $\bar{\epsilon} > 0$  and  $\bar{\delta} > 0$  such that the following claims hold:*

- (a) *Given any  $x \in A$ , the condition  $|w_j| > \bar{\delta} \forall j \in [k]$  must hold if  $h^{(k)}(x, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}}$ ;*
- (b)  $d_{J_1}(B^{3\bar{\epsilon}}, \mathbb{D}_A^{(k-1)}) > 0$ .

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 6 (Uniform Nondegeneracy).**  $\inf_{x \in \mathbb{R}} \sigma(x) \geq c$  for some  $c > 0$ .

Now, we state a result for  $\mathbb{D}_A^{(k)|b}$  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  $B \in \mathcal{S}_{\mathbb{D}}$ . Let  $k = 0, 1, 2, \dots$ . If  $B$  is bounded away from  $\mathbb{D}_A^{(k-1)|b}$ , then there exist  $\bar{\epsilon} > 0$  and  $\bar{\delta} > 0$  such that the following claims hold:*

- (a) *Given any  $x \in A$ , the condition  $|w_j| > \bar{\delta} \forall j \in [k]$  must hold if  $h^{(k)|b}(x, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}}$ ;*
- (b)  $d_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b}) > 0$ .

Furthermore, suppose that Assumption 6 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  $h^{(k)|b+\bar{\epsilon}}(x, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}}$ ,*
- (d)  $d_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}) > 0$ .

Lemma 3.7 reveals that the image of  $h^{(1)}$  (resp.  $h^{(1)|b}$ ) provides good approximations of the sample path of  $X_j^\eta$  (resp.  $X_j^{\eta|b}$ ) up until  $\tau_1^{>\delta}(\eta)$ , i.e. the arrival time of the first “large noise”; see (3.2), (3.3) for the definition of  $\tau_i^{>\delta}(\eta)$ ,  $W_i^{>\delta}(\eta)$ .

**LEMMA 3.7.** *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*

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

that

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

where

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

(b) Furthermore, suppose that Assumption 6 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

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

that

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

$$(3.23) \quad \sup_{t \in [0, 1]: t \leq \eta \tau_1^{>\delta}(\eta)} |\xi_t - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x)| \leq \rho \cdot \left(1 + \frac{bD}{c}\right) (\epsilon + |x - y| + 2\eta C)$$

where

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

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

Let  $A \triangleq \left\{ \max_{i \leq \lfloor 1/\eta \rfloor \wedge (\tau_1^{>\delta}(\eta) - 1)} \eta \left| \sum_{j=1}^i \sigma(X_{j-1}^{\eta}(x)) Z_j \right| \leq \epsilon \right\}$ . Let  $\mathbf{x}^\eta$  be the deterministic process defined in equation (C.11) in the Supplementary Material. Applying discrete version of Gronwall's inequality (see, for example, Lemma A.3 of [35]) we know that on event  $A$ ,

$$(3.24) \quad \left| \mathbf{x}_j^\eta(x) - X_j^\eta(x) \right| \leq \epsilon \cdot \exp(\eta D \cdot \lfloor 1/\eta \rfloor) \leq \rho \epsilon \quad \forall j \leq \lfloor 1/\eta \rfloor \wedge (\tau_1^{>\delta}(\eta) - 1).$$

On the other hand, recall that  $\mathbf{y}_t(y)$  defined in (2.22) is the solution to ODE  $d\mathbf{y}_t(y)/dt = a(\mathbf{y}_t(y))$  under initial condition  $\mathbf{y}_0(y) = y$ . Since  $\xi_t = \mathbf{y}_t(y)$  on  $t < \eta \tau_1^{>\delta}(\eta)$ , by applying Lemma C.5 in the Supplementary Material, we get

$$(3.25) \quad \sup_{t \in [0, 1]: t < \eta \tau_1^{>\delta}(\eta)} \left| \xi_t - \mathbf{x}_{\lfloor t/\eta \rfloor}^\eta(x) \right| \leq (\eta C + |x - y|) \cdot \rho.$$

Therefore,

$$(3.26) \quad \sup_{t \in [0, 1]: t < \eta \tau_1^{>\delta}(\eta)} \left| \xi_t - X_{\lfloor t/\eta \rfloor}^\eta(x) \right| \leq \rho \cdot (\epsilon + |x - y| + \eta C).$$

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

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

Also, for any  $w$  with  $|w| \leq \delta < \frac{b}{2C}$ , note that  $\varphi_b(\eta a(x) + \sigma(x)w) = \eta a(x) + \sigma(x)w \forall x \in \mathbb{R}$  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,  $X_j^\eta(x) = X_j^{\eta|b}(x)$  for all  $x \in \mathbb{R}$  and  $j < \tau_1^{>\delta}(\eta)$ . It then follows

directly from (3.26) that  $\sup_{t \in [0,1]: t < \eta\tau_1^{>\delta}(\eta)} |\xi_t - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x)| \leq \rho \cdot (\epsilon + |x - y| + \eta C)$ . A direct consequence is (we write  $\mathbf{y}(u; y) = \mathbf{y}_u(y)$ ,  $\mathbf{y}(s-; y) = \lim_{u \uparrow s} \mathbf{y}_u(y)$ , and  $\xi(t) = \xi_t$  in this proof)

$$(3.27) \quad \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) - X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right| \leq \rho \cdot (\epsilon + |x - y| + \eta C).$$

Therefore,

$$\begin{aligned} & \left| \xi(\eta\tau_1^{>\delta}(\eta)) - X_{\tau_1^{>\delta}(\eta)}^{\eta|b}(x) \right| \\ &= \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) + \varphi_b \left( \eta\sigma \left( \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) \right) W_1^{>\delta}(\eta) \right) \right. \\ & \quad \left. - \left[ X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) + \varphi_b \left( \eta a \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) + \eta\sigma \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) W_1^{>\delta}(\eta) \right) \right] \right| \\ &\leq \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) - X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right| \\ & \quad + \left| \varphi_b \left( \eta\sigma \left( \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) \right) W_1^{>\delta}(\eta) \right) - \varphi_b \left( \eta\sigma \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) W_1^{>\delta}(\eta) \right) \right| \\ & \quad + \left| \varphi_b \left( \eta\sigma \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) W_1^{>\delta}(\eta) \right) \right. \\ & \quad \quad \left. - \varphi_b \left( \eta a \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) + \eta\sigma \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) W_1^{>\delta}(\eta) \right) \right| \\ &\leq \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) - X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right| \\ & \quad + |\varphi_{b/c}(\eta W_1^{>\delta}(\eta))| \cdot \left| \sigma \left( \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) \right) - \sigma \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) \right| \\ & \quad + \left| \eta a \left( X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right) \right| \\ & \quad \text{by property (C.8) in Supplementary Material and } |\varphi_b(x) - \varphi_b(y)| \leq |x - y| \\ &\leq \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) - X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right| + \frac{b}{c} \cdot D \cdot \left| \mathbf{y}(\eta\tau_1^{>\delta}(\eta)-; y) - X_{\tau_1^{>\delta}(\eta)-1}^{\eta|b}(x) \right| + \eta C \\ &\leq \left( 1 + \frac{bD}{c} \right) \cdot \rho \cdot (\epsilon + |x - y| + \eta C) + \eta C \\ &\leq \left( 1 + \frac{bD}{c} \right) \cdot \rho \cdot (\epsilon + |x - y| + 2\eta C). \end{aligned}$$

This concludes the proof of part (b).  $\square$

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

**LEMMA 3.8.** *Let Assumptions 2, 4, and 6 hold. Let  $A_i(\eta, b, \epsilon, \delta, x)$  be defined as in (3.6). For any  $k \geq 0$ ,  $x \in \mathbb{R}$ ,  $\epsilon, b > 0$ ,  $\delta \in (0, \frac{b}{2C})$ , and  $\eta \in (0, \frac{b \wedge \epsilon}{2C})$ , it holds on event  $(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)) \cap \left\{ \tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta) \right\}$  that*

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

where  $\xi \triangleq h^{(k)|b}(x, \eta W_1^{\geq \delta}(\eta), \dots, \eta W_k^{\geq \delta}(\eta), \eta \tau_1^{\geq \delta}(\eta), \dots, \eta \tau_k^{\geq \delta}(\eta))$ ,  $\rho = \exp(D) \geq 1$ ,  $D \in [1, \infty)$  is the Lipschitz coefficient in Assumption 2,  $C \geq 1$  is the constant in Assumption 4, and  $c \in (0, 1)$  is the constant in Assumption 6.

PROOF. First of all, on  $A_1(\eta, b, \epsilon, \delta, x)$ , one can apply (3.22) of Lemma 3.7 and obtain

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

where we applied our choice of  $\eta C < \epsilon/2$ . In case that  $k = 0$ , we can already conclude the proof. Henceforth in the proof, we focus on the case where  $k \geq 1$ . Now we can instead apply (3.23) of Lemma 3.7 to get

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

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

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

To highlight the timestamp in the ODE  $\mathbf{y}_t(y)$  we write  $\mathbf{y}(t; y) = \mathbf{y}_t(y)$  in this proof. Note that for any  $t \in [\eta \tau_j^{\geq \delta}(\eta), \eta \tau_{j+1}^{\geq \delta}(\eta))$ , we have  $\xi_t = \mathbf{y}\left(t - \eta \tau_j^{\geq \delta}(\eta); \xi_{\eta \tau_j^{\geq \delta}(\eta)}\right)$ . Therefore, by applying (3.23) of Lemma 3.7 again, we obtain

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

Arguing inductively, we yield  $\sup_{t \in [0, \eta \tau_k^{\geq \delta}(\eta)]} \left| \xi_t - X_{[t/\eta]}^{\eta|b}(x) \right| \leq R_k = \left[ 3\rho \cdot \left(1 + \frac{bD}{c}\right) \right]^k \epsilon$ . Lastly, due to (3.21) of Lemma 3.7 and the fact that  $\eta \tau_{k+1}^{\geq \delta}(\eta) > 1$ ,

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

This concludes the proof.  $\square$

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  $\lim_{n \rightarrow \infty} x_n = x^*$ ,*

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



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

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

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 C.4 of the Supplementary Material), for any sequence  $y_m \in \mathbb{R}$  with  $\lim_{m \rightarrow \infty} y_m = y \in \mathbb{R}$ , we have

$$\lim_{m \rightarrow \infty} f(h^{(k)}(y_m, \mathbf{w}, \mathbf{t})) = f(h^{(k)}(y, \mathbf{w}, \mathbf{t})) \quad \forall \mathbf{w} \in \mathbb{R}^k, \mathbf{t} \in (0, 1)^{k\uparrow}.$$

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  $h^{(k)}(x, \mathbf{w}, \mathbf{t}) \in B \implies |w_j| > \bar{\delta} \forall j \in [k]$ . As a result,  $|f(h^{(k)}(x, \mathbf{w}, \mathbf{t}))| \leq \|f\| \cdot \mathbb{I}(|w_j| > \bar{\delta} \forall j \in [k])$ . Also, note that  $\int \mathbb{I}(|w_j| > \bar{\delta} \forall j \in [k]) \nu_\alpha^k(d\mathbf{w}) \times \mathcal{L}_1^{k\uparrow}(d\mathbf{t}) \leq 1/\bar{\delta}^{k\alpha} < \infty$ . This allows us to apply Dominated Convergence theorem and establish the continuity of  $\phi(x)$ . This implies

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

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 C.3 (resp. Lemma 3.6) instead of Lemma C.4 (resp. Lemma 3.5) so we omit the details.  $\square$

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 nearly 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 [36]. Nevertheless, under the conditions provided in Theorem 2.3 or 2.4 regarding the initial conditions of  $\mathbf{X}^\eta$  or  $\mathbf{X}^{\eta|b}$ , 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  $\mathbf{X}^{\eta|b}$ .

PROPOSITION 3.10. *Let  $\eta_n$  be a sequence of strictly positive real numbers with  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Let compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  be such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Under Assumptions 1, 2, and 3, it holds for any  $k = 0, 1, 2, \dots$  and  $b > 0$  that*

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

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

PROPOSITION 3.11. *Let  $\eta_n$  be a sequence of strictly positive real numbers with  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Let compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  be such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Under Assumptions 1, 2, 3, and 4, it holds for any  $k = 0, 1, 2, \dots$  that*

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

PROOF. Fix some  $k = 0, 1, 2, \dots$  and some  $g \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)})$ . By virtue of Portmanteau theorem for  $\mathbb{M}$ -convergence (see theorem 2.1 of [36]), it suffices to show that

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

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

$$\begin{aligned} \mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))] &= \mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))\mathbb{I}(\mathbf{X}^{\eta_n}(x_n) \in B)] \\ &= \mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))\mathbb{I}(\tau_{k+1}^{>\delta}(\eta_n) < \lfloor 1/\eta_n \rfloor; \mathbf{X}^{\eta_n}(x_n) \in B)] \\ &\quad + \mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))\mathbb{I}(\tau_k^{>\delta}(\eta_n) > \lfloor 1/\eta_n \rfloor; \mathbf{X}^{\eta_n}(x_n) \in B)] \\ &\quad + \mathbf{E}\left[g(\mathbf{X}^{\eta_n}(x_n))\mathbb{I}(\tau_k^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau_{k+1}^{>\delta}(\eta_n); \right. \\ &\quad \left. \eta_n |W_j^{>\delta}(\eta_n)| > \frac{b}{2C} \text{ for some } j \in [k]; \mathbf{X}^{\eta_n}(x_n) \in B)\right] \\ &\quad + \mathbf{E}\left[g(\mathbf{X}^{\eta_n}(x_n))\mathbb{I}(\tau_k^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau_{k+1}^{>\delta}(\eta_n); \right. \\ &\quad \left. \eta_n |W_j^{>\delta}(\eta_n)| \leq \frac{b}{2C} \forall j \in [k]; \mathbf{X}^{\eta_n}(x_n) \in B)\right] \end{aligned}$$

where  $C \geq 1$  is the constant in Assumption 4 such that  $|a(x)| \vee \sigma(x) \leq C$  for any  $x \in \mathbb{R}$ . Focusing on the last term in the display above, let

$$\tilde{A}(n, b, \delta) \triangleq \left\{ \tau_k^{>\delta}(\eta_n) \leq \lfloor 1/\eta_n \rfloor < \tau_{k+1}^{>\delta}(\eta_n); \eta_n |W_j^{>\delta}(\eta_n)| \leq \frac{b}{2C} \forall j \in [k]; \mathbf{X}^{\eta_n}(x_n) \in B \right\}.$$

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  $\tilde{A}(n, b, \delta)$  the step-size (before truncation)  $\eta a(X_{j-1}^{\eta|b}(x)) + \eta \sigma(X_{j-1}^{\eta|b}(x))Z_j$  of  $X_j^{\eta|b}$  is less than  $b$  for each  $j \leq \lfloor 1/\eta_n \rfloor$ , and hence  $\mathbf{X}^{\eta_n}(x_n) = \mathbf{X}^{\eta_n|b}(x_n)$ . 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

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

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

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

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

$$(3.28) \quad \lim_{n \rightarrow \infty} p_1(n, \delta) / \lambda^k(\eta_n) = 0,$$

$$(3.29) \quad \lim_{n \rightarrow \infty} p_2(n, \delta) / \lambda^k(\eta_n) = 0.$$

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

$$(3.30) \quad \limsup_{n \rightarrow \infty} p_3(n, b, \delta) / \lambda^k(\eta_n) \leq \psi_\delta(b) \triangleq \frac{k}{\delta^{\alpha k}} \cdot \left(\frac{\delta}{2C}\right)^\alpha \cdot \frac{1}{b^\alpha}.$$

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

$$(3.31) \quad \lim_{b \rightarrow \infty} \mathbf{C}^{(k)|b}(g; x^*) = \mathbf{C}^{(k)}(g; x^*).$$

Then by combining (3.28)–(3.30) with the upper and lower bounds for  $\mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))]$  established earlier, we see that for any  $b$  large enough (such that  $\frac{b}{2C} > \delta$ ),

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbf{E}[g(\mathbf{X}^{\eta_n|b}(x_n))]}{\lambda^k(\eta_n)} - \|g\| \psi_\delta(b) &\leq \lim_{n \rightarrow \infty} \frac{\mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))]}{\lambda^k(\eta_n)} \\ &\leq \lim_{n \rightarrow \infty} \frac{\mathbf{E}[g(\mathbf{X}^{\eta_n|b}(x_n))]}{\lambda^k(\eta_n)} + \|g\| \psi_\delta(b). \end{aligned}$$

As a result,

$$- \|g\| \psi_\delta(b) + \mathbf{C}^{(k)|b}(g; x^*) \leq \lim_{n \rightarrow \infty} \frac{\mathbf{E}[g(\mathbf{X}^{\eta_n}(x_n))]}{\lambda^k(\eta_n)} \leq \|g\| \psi_\delta(b) + \mathbf{C}^{(k)|b}(g; x^*).$$

In the last line of the display, we applied Proposition 3.10. Letting  $b$  tend to  $\infty$  and applying the limit (3.31), we conclude the proof. Now it only remains to prove (3.28) (3.29) (3.30) (3.31).

**Proof of Claim (3.28):**

Applying (3.4), we see that  $p_1(n, \delta) \leq (H(\frac{\delta}{\eta_n})/\eta_n)^{k+1}$  holds for any  $\delta > 0$  and any  $n \geq 1$ . Due to the regularly varying nature of  $H(\cdot)$ , we then yield  $\limsup_{n \rightarrow \infty} \frac{p_1(n, \delta)}{\lambda^{k+1}(\eta_n)} \leq 1/\delta^{\alpha(k+1)} < \infty$ . To show that claim (3.28) holds for any  $\delta > 0$  we only need to note that

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

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 (3.29):**

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

$$(3.32) \quad \left\{ \tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor; \mathbf{X}^\eta(x) \in B \right\} \cap \left( \bigcap_{i=1}^{k+1} A_i(\eta, \infty, \epsilon, \delta, x) \right) = \emptyset \quad \forall x \in A, \delta > 0, \eta \in (0, \frac{\epsilon}{C\rho})$$

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

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

Applying Lemma 3.3 (b) with some  $N > k(\alpha - 1)$ , we conclude that claim (3.29) holds for all  $\delta > 0$  small enough. Now it only remains to find  $\epsilon > 0$  that satisfies condition (3.32). 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  $d_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)}) > \bar{\epsilon}$ . Now we show that (3.32) 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  $\check{X}^{\eta, \delta}(x) \triangleq \{\check{X}_t^{\eta, \delta}(x) : t \in [0, 1]\}$  as the solution to (under initial condition  $\check{X}_0^{\eta, \delta}(x) = x$ )

$$\begin{aligned} \frac{d\check{X}_t^{\eta, \delta}(x)}{dt} &= a(\check{X}_t^{\eta, \delta}(x)) \quad \forall t \geq 0, t \notin \{\eta\tau_j^{\delta}(\eta) : j \geq 1\}, \\ \check{X}_{\eta\tau_i^{\delta}(\eta)}^{\eta, \delta}(x) &= X_{\tau_i^{\delta}(\eta)}^{\eta}(x) \quad \forall j \geq 1. \end{aligned}$$

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

$$\begin{aligned} &d_{J_1}(\check{X}^{\eta, \delta}(x), X^{\eta}(x)) \\ &\leq \sup_{t \in [0, \eta\tau_1^{\delta}(\eta)] \cup [\eta\tau_1^{\delta}(\eta), \eta\tau_2^{\delta}(\eta)] \cup \dots \cup [\eta\tau_k^{\delta}(\eta), \eta\tau_{k+1}^{\delta}(\eta)]} |\check{X}_t^{\eta, \delta}(x) - X_{\lfloor t/\eta \rfloor}^{\eta}(x)| \\ &\leq \rho \cdot (\epsilon + \eta C) \leq \rho\epsilon + \epsilon < \bar{\epsilon} \quad \text{because of (3.21) of Lemma 3.7.} \end{aligned}$$

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  $\check{X}^{\eta, \delta}(x) \in \mathbb{D}_A^{(k-1)}$ . As a result, on event  $(\cap_{i=1}^{k+1} A_i(\eta, \infty, \epsilon, \delta, x)) \cap \{\tau_k^{\delta}(\eta) > \lfloor 1/\eta \rfloor\}$  we must have  $d_{J_1}(\mathbb{D}_A^{(k-1)}, X^{\eta}(x)) < \bar{\epsilon}$ , and hence  $X^{\eta}(x) \notin B$  due to the fact that  $d_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)}) > \bar{\epsilon}$ . This establishes (3.32).

**Proof of Claim (3.30):**

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

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

Due to  $H(x) \in \mathcal{RV}_{-\alpha}(x)$  as  $x \rightarrow \infty$ , we conclude that  $\limsup_{n \rightarrow \infty} \frac{p_3(n, b, \delta)}{\lambda^k(\eta_n)} \leq \frac{k}{\delta^{\alpha k}} \cdot \left(\frac{\delta}{2C}\right)^{\alpha} \cdot \frac{1}{b^{\alpha}} = \psi_{\delta}(b)$ .

**Proof of Claim (3.31):**

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

$$(3.33) \quad \lim_{b \rightarrow \infty} \mathbf{C}^{(k)|b}(S; x^*) = \mathbf{C}^{(k)}(S; x^*).$$

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

$(c_g^{(i)})_{i=1}^N$ , some sequence  $(B_g^{(i)})_{i=1}^N$  of Borel measurable sets on  $\mathbb{D}$  that are bounded away from  $\mathbb{D}_A^{(k-1)}$  such that the following claims hold for  $g^\Delta(\cdot) \triangleq \sum_{i=1}^N c_g^{(i)} \mathbb{I}(\cdot \in B_g^{(i)})$ :

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

Now observe that

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

First, note that

$$\mathbf{C}^{(k)|b}(g^\Delta; x^*) = \sum_{i=1}^N c_g^{(i)} \mathbf{C}^{(k)|b}(B_g^{(i)}; x^*), \quad \mathbf{C}^{(k)}(g^\Delta; x^*) = \sum_{i=1}^N c_g^{(i)} \mathbf{C}^{(k)}(B_g^{(i)}; x^*).$$

Therefore, applying (3.33), we get  $\limsup_{b \rightarrow \infty} \left| \mathbf{C}^{(k)|b}(g^\Delta; x^*) - \mathbf{C}^{(k)}(g^\Delta; x^*) \right| = 0$ . Next,  $\left| \mathbf{C}^{(k)|b}(g^\Delta; x^*) - \mathbf{C}^{(k)|b}(g; x^*) \right| \leq \Delta \cdot \mathbf{C}^{(k)|b}(B; x^*)$  and  $\left| \mathbf{C}^{(k)}(g^\Delta; x^*) - \mathbf{C}^{(k)}(g; x^*) \right| \leq \Delta \cdot \mathbf{C}^{(k)}(B; x^*)$ . Thanks to (3.33) again, we get  $\limsup_{b \rightarrow \infty} \left| \mathbf{C}^{(k)|b}(g; x^*) - \mathbf{C}^{(k)}(g; x^*) \right| \leq 2\Delta \cdot \mathbf{C}^{(k)}(B; x^*)$ . The arbitrariness of  $\Delta > 0$  allows us to conclude the proof of (3.30).

We prove (3.33) by applying Dominated Convergence theorem. From the definition in (2.20),

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

where  $S \in \mathcal{S}_{\mathbb{D}}$  is bounded away from  $\mathbb{D}_A^{(k-1)}$ . First, for any  $\mathbf{w} \in \mathbb{R}^k$ ,  $\mathbf{t} \in (0, 1)^{k\uparrow}$  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  $h^{(k)|b}(x^*, \mathbf{w}, \mathbf{t}) = h^{(k)}(x^*, \mathbf{w}, \mathbf{t})$ . This implies  $\lim_{b \rightarrow \infty} \mathbb{I}\{h^{(k)|b}(x^*, \mathbf{w}, \mathbf{t}) \in S\} = \mathbb{I}\{h^{(k)}(x^*, \mathbf{w}, \mathbf{t}) \in S\}$  for all  $\mathbf{w} \in \mathbb{R}^k$  and  $\mathbf{t} \in (0, 1)^{k\uparrow}$ . In order to apply Dominated Convergence theorem and conclude the proof of (3.33), it suffices to find an integrable function that dominates  $\mathbb{I}\{h^{(k)|b}(x^*, \mathbf{w}, \mathbf{t}) \in S\}$ . Specifically, since  $S$  is bounded away from  $\mathbb{D}_A^{(k-1)}$ , we can find some  $\bar{\epsilon} > 0$  such that  $d_{J_1}(S, \mathbb{D}_A^{(k-1)}) > \bar{\epsilon}$ . 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

$$(3.34) \quad \mathbb{I}\{h^{(k)|b}(x^*, \mathbf{w}, \mathbf{t}) \in S\} \leq \mathbb{I}\{|w_j| > \bar{\delta} \quad \forall j \in [k]\} \quad \forall b > 0, \mathbf{w} \in \mathbb{R}^k, \mathbf{t} \in (0, 1)^{k\uparrow}.$$

From  $\int \mathbb{I}\{|w_j| > \bar{\delta} \quad \forall j \in [k]\} \nu_\alpha^k(d\mathbf{w}) \times \mathcal{L}_1^{k\uparrow}(d\mathbf{t}) \leq 1/\bar{\delta}^{k\alpha} < \infty$  we conclude the proof. Now it only remains to prove (3.34). Fix some  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ ,  $\mathbf{t} = (t_1, \dots, t_k) \in (0, 1)^{k\uparrow}$ , and  $b > 0$ . Let  $\xi_b = h^{(k)|b}(x^*, \mathbf{w}, \mathbf{t})$ . 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  $\mathbf{y}(\cdot)$  is the ODE defined in (2.22))

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

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

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

which implies  $d_{J_1}(\xi, \xi_b) < \bar{\epsilon}$ . However, due to  $\xi \in \mathbb{D}_A^{(k-1)}$  and  $d_{J_1}(S, \mathbb{D}_A^{(k-1)}) > \bar{\epsilon}$ , we must have  $\xi_b \notin S$ . This concludes the proof of (3.34).  $\square$

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  $k = 0, 1, \dots$  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  $\lim_{n \rightarrow \infty} \eta_n = 0$ , some sequence  $x_n \in A$ , and  $\epsilon > 0$  such that  $|\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f; x_n)| > \epsilon \forall n \geq 1$  where  $\mu_n^{(k)}(\cdot) \triangleq \mathbf{P}(\mathbf{X}^{\eta_n}(x_n) \in \cdot) / \lambda^k(\eta_n)$ . Since  $A$  is compact, by picking a proper subsequence we can assume w.l.o.g. that  $\lim_{n \rightarrow \infty} x_n = x^*$  for some  $x^* \in A$ . This allows us to apply Proposition 3.11 and yield  $\lim_{n \rightarrow \infty} |\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f; x^*)| = 0$ . On the other hand, using part (a) of Lemma 3.9, we get  $\lim_{n \rightarrow \infty} |\mathbf{C}^{(k)}(f; x_n) - \mathbf{C}^{(k)}(f; x^*)| = 0$ . Therefore, we arrive at the contradiction

$$\begin{aligned} &\lim_{n \rightarrow \infty} |\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f; x_n)| \\ &\leq \lim_{n \rightarrow \infty} |\mu_n^{(k)}(f) - \mathbf{C}^{(k)}(f; x^*)| + \lim_{n \rightarrow \infty} |\mathbf{C}^{(k)}(f; x^*) - \mathbf{C}^{(k)}(f; x_n)| = 0 \end{aligned}$$

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  $\sup_{x \in A} \mathbf{C}^{(k)}(B^-; x) < \infty$ .

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

Then by the definition of  $\mathbf{C}^{(k)|b}$  in (2.10),

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

This concludes the proof.  $\square$



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.  $\square$

**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  $M \geq 1$ ,

$$(3.35) \quad a_M(x) \triangleq \begin{cases} a(M) & \text{if } x > M, \\ a(-M) & \text{if } x < -M, \\ a(x) & \text{otherwise.} \end{cases} \quad \sigma_M(x) \triangleq \begin{cases} \sigma(M) & \text{if } x > M, \\ \sigma(-M) & \text{if } x < -M, \\ \sigma(x) & \text{otherwise.} \end{cases}$$

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 6. 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  $h_{M\downarrow}^{(k)|b} : \mathbb{R} \times \mathbb{R}^k \times (0, 1]^{k\uparrow} \rightarrow \mathbb{D}$  as follows. Given any  $x_0 \in \mathbb{R}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ ,  $\mathbf{t} = (t_1, \dots, t_k) \in (0, 1]^{k\uparrow}$ , let  $\xi = h_{M\downarrow}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t})$  be the solution to

$$(3.36) \quad \xi_0 = x_0;$$

$$(3.37) \quad \frac{d\xi_t}{dt} = a_M(\xi_t) \quad \forall t \in [0, 1], \quad t \neq t_1, t_2, \dots, t_k;$$

$$(3.38) \quad \xi_t = \xi_{t-} + \varphi_b(\sigma_M(\xi_{t-})w_j) \quad \text{if } t = t_j \text{ for some } j \in [k].$$

Also, we let

$$(3.39) \quad \mathbb{D}_{A;M\downarrow}^{(k)|b} \triangleq h_{M\downarrow}^{(k)|b}(\mathbb{R} \times \mathbb{R}^k \times (0, 1]^{k\uparrow}).$$

One can see that the key difference between  $h_{M\downarrow}^{(k)|b}$  and  $h^{(k)|b}$  is that, when constructing  $h_{M\downarrow}^{(k)|b}$ , 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 boundedness condition in Assumption 4 and the stronger uniform nondegeneracy condition in Assumption 6.

**PROPOSITION 3.12.** *Let  $\eta_n$  be a sequence of strictly positive real numbers with  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Let compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  be such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Under Assumptions 1, 2, 4, and 6, it holds for any  $k = 0, 1, 2, \dots$  and  $b > 0$  that*

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

**PROOF OF PROPOSITION 3.10.** Fix some  $b > 0, k \geq 0$ , as well as some  $g \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)|b})$  that is also uniformly continuous on  $\mathbb{D}$ . Thanks to the Portmanteau theorem for  $\mathbb{M}$ -convergence (see theorem 2.1 of [36]), it suffices to show that

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

Let  $B \triangleq \text{supp}(g)$ . Note that  $B$  is bounded away from  $\mathbb{D}_A^{(k-1)|b}$ . Applying Corollary C.2 of the Supplementary Material, we can fix some  $M_0$  such that the following claim holds for any  $M \geq M_0$ : for any  $\xi = h_{M\downarrow}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t})$  with  $\mathbf{t} = (t_1, \dots, t_k) \in (0, 1]^{k\uparrow}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$  and  $x_0 \in A$ ,

$$(3.40) \quad \xi = h_{M\downarrow}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t}) = h_{M\downarrow}^{(k)|b}(x_0, \mathbf{w}, \mathbf{t}); \quad \sup_{t \in [0,1]} |\xi(t)| \leq M_0.$$

Here the mapping  $h_{M\downarrow}^{(k)|b}$  is defined in (3.36)–(3.38). Now fix some  $M \geq M_0 + 1$  and recall the definitions of  $a_M, \sigma_M$  in (3.35). Also, define stochastic processes  $\widetilde{\mathbf{X}}^{\eta|b}(x) \triangleq \{\widetilde{X}_{\lfloor t/\eta \rfloor}^{\eta|b}(x) : t \in [0, 1]\}$  as

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

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

$$(3.41) \quad \sup_{t \in [0,1]} |\widetilde{X}_{\lfloor t/\eta \rfloor}^{\eta|b}(x)| > M \iff \sup_{t \in [0,1]} |X_{\lfloor t/\eta \rfloor}^{\eta|b}(x)| > M,$$

$$(3.42) \quad \sup_{t \in [0,1]} |X_{\lfloor t/\eta \rfloor}^{\eta|b}(x)| \leq M \implies \mathbf{X}^{\eta|b}(x) = \widetilde{\mathbf{X}}^{\eta|b}(x).$$

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

$$(3.43) \quad \begin{aligned} \mathbf{E}[g(\mathbf{X}^{\eta_n|b}(x_n))] &= \mathbf{E}\left[g(\mathbf{X}^{\eta_n|b}(x_n)) \mathbb{I}\left\{\mathbf{X}^{\eta_n|b}(x_n) \in B; \sup_{t \in [0,1]} |X_{\lfloor t/\eta \rfloor}^{\eta_n|b}(x_n)| \leq M\right\}\right] \\ &\quad + \mathbf{E}\left[g(\mathbf{X}^{\eta_n|b}(x_n)) \mathbb{I}\left\{\mathbf{X}^{\eta_n|b}(x_n) \in B; \sup_{t \in [0,1]} |X_{\lfloor t/\eta \rfloor}^{\eta_n|b}(x_n)| > M\right\}\right]. \end{aligned}$$

An upper bound then follows immediately from (3.41) and (3.42):

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

Similarly, by bounding the first term on the R.H.S. of (3.43) using (3.41) and (3.42), we obtain

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

To conclude the proof, it only remains to show that

$$(3.44) \quad \lim_{n \rightarrow \infty} \lambda^{-k}(\eta_n) \mathbf{E}[g(\widetilde{\mathbf{X}}^{\eta_n|b}(x_n))] = \mathbf{C}^{(k)|b}(g; x^*),$$

$$(3.45) \quad \lim_{n \rightarrow \infty} \lambda^{-k}(\eta_n) \mathbf{P}\left(\sup_{t \in [0,1]} |\widetilde{X}_{\lfloor t/\eta \rfloor}^{\eta_n|b}(x_n)| > M\right) = 0.$$

**Proof of Claim (3.44):**

Under Assumption 3, one can easily see that  $a_M, \sigma_M$  would satisfy Assumption 4 and 6. This allows us to apply Proposition 3.12 and obtain  $\lim_{n \rightarrow \infty} \lambda^{-k}(\eta_n) \mathbf{E}[g(\widetilde{\mathbf{X}}^{\eta_n|b}(x_n))] = \widetilde{\mathbf{C}}^{(k)|b}(g; x^*)$  where

$$\widetilde{\mathbf{C}}^{(k)|b}(\cdot; x) \triangleq \int \mathbb{I}\left\{h_{M\downarrow}^{(k)|b}(x, \mathbf{w}, t) \in \cdot\right\} \nu_\alpha^k(d\mathbf{w}) \times \mathcal{L}_1^{k\uparrow}(dt).$$

Given (3.40) 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 (3.44).

**Proof of Claim (3.45):**

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  $\mathbb{D}_A^{(k)|b}$ , then by applying Proposition 3.12 again we get

$$\limsup_{n \rightarrow \infty} \mathbf{P}\left(\widetilde{\mathbf{X}}^{\eta_n|b}(x_n) \in E\right) / \lambda^{k+1}(\eta_n) < \infty,$$

which then implies (3.45). To see why  $E$  is bounded away from  $\mathbb{D}_A^{(k)|b}$ , note that it follows directly from (3.40) that

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

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

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  $\hat{\mathbf{X}}^{\eta|b;(k)}$  that is not only asymptotically equivalent to  $\mathbf{X}^{\eta|b}$  (as  $\eta \downarrow 0$ ) but also (under the right scaling) approaches to  $\mathbf{C}_b^{(k)}$  in the sense of  $\mathbb{M}$ -convergence. To properly introduce the process  $\hat{\mathbf{X}}^{\eta|b;(k)}$ , a few new definitions are in order. For any  $j \geq 1$  and  $n \geq j$  let

$$(3.46) \quad \mathcal{J}_Z(c, n) \triangleq \#\{i \in [n] : |Z_i| \geq c\} \quad \forall c \geq 0; \quad \mathbf{Z}^{(j)}(\eta) \triangleq \max\left\{c \geq 0 : \mathcal{J}_Z(c, \lfloor 1/\eta \rfloor) \geq j\right\}.$$

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  $\mathbf{Z}^{(j)}(\eta)$  identifies the value of the  $j^{\text{th}}$  largest element in  $\{|Z_i| : i \leq \lfloor 1/\eta \rfloor\}$ . Moreover, let

$$(3.47) \quad \tau_i^{(j)}(\eta) \triangleq \min\{k > \tau_{i-1}^{(j)}(\eta) : |Z_k| \geq \mathbf{Z}^{(j)}(\eta)\}, \quad W_i^{(j)}(\eta) \triangleq Z_{\tau_i^{(j)}(\eta)} \quad \forall i = 1, 2, \dots, j$$

with the convention that  $\tau_0^{(j)}(\eta) = 0$ . Note that  $(\tau_i^{(j)}(\eta), W_i^{(j)}(\eta))_{i \in [j]}$  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  $\hat{\mathbf{X}}^{\eta|b;(j)}(x) \triangleq \{\hat{X}_t^{\eta|b;(j)}(x) : t \in [0, 1]\}$  as the solution to

$$(3.48) \quad \frac{d\hat{X}_t^{\eta|b;(j)}(x)}{dt} = a(\hat{X}_t^{\eta|b;(j)}(x)) \quad \forall t \in [0, 1], \quad t \notin \{\eta\tau_i^{(j)}(\eta) : i \in [j]\},$$

$$(3.49) \quad \hat{X}_t^{\eta|b;(j)}(x) = \hat{X}_{t-}^{\eta|b;(j)}(x) + \varphi_b\left(\eta\sigma(\hat{X}_{t-}^{\eta|b;(j)}(x))W_i^{(j)}(\eta)\right) \quad \text{if } t = \eta\tau_i^{(j)}(\eta) \text{ for some } i \in [j].$$

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

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

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

$$(3.50) \quad \tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta) \quad \implies \quad \hat{\mathbf{X}}^{\eta|b;(k)}(x) = h^{(k)|b}(x, \eta \mathbf{W}^{>\delta}(\eta), \eta \boldsymbol{\tau}^{>\delta}(\eta))$$

with  $\mathbf{W}^{>\delta}(\eta) = (W_1^{>\delta}(\eta), \dots, W_k^{>\delta}(\eta))$  and  $\boldsymbol{\tau}^{>\delta}(\eta) = (\tau_1^{>\delta}(\eta), \dots, \tau_k^{>\delta}(\eta))$ . The following two results allow us to apply Lemma 3.2, thus bridging the gap between  $\mathbf{X}^{\eta|b}$  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  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Let compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  be such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Under Assumptions 1, 2, 4, and 6, it holds for any  $k = 0, 1, 2, \dots$  and  $b > 0$  that  $\mathbf{X}^{\eta_n|b}(x_n)$  is asymptotically equivalent to  $\hat{\mathbf{X}}^{\eta_n|b;(k)}(x_n)$  (as  $n \rightarrow \infty$ ) 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  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Let compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  be such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Under Assumptions 1, 2, 4, and 6, it holds for any  $k = 0, 1, 2, \dots$  and  $b > 0$  that*

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

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.  $\square$

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  $\lim_{n \rightarrow \infty} \eta_n = 0$ . Also, fix a compact set  $A \subseteq \mathbb{R}$  and  $x_n, x^* \in A$  such that  $\lim_{n \rightarrow \infty} x_n = x^*$ . Meanwhile, arbitrarily pick some  $\Delta > 0$  and some  $B \in \mathcal{S}_{\mathbb{D}}$  that is bounded away from  $\mathbb{D}_A^{(k-1)|b}$ . It suffices to show that

$$(3.51) \quad \lim_{n \rightarrow \infty} \mathbf{P}\left(\mathbf{d}_{J_1}(\mathbf{X}^{\eta_n|b}(x_n), \hat{\mathbf{X}}^{\eta_n|b;(k)}(x_n)) \mathbb{I}(\mathbf{X}^{\eta_n|b}(x_n) \text{ or } \hat{\mathbf{X}}^{\eta_n|b;(k)}(x_n) \in B) > \Delta\right) / \lambda^k(\eta_n) = 0.$$

Applying Lemma 3.6, we can fix some  $\bar{\epsilon} > 0$  and  $\bar{\delta} \in (0, \frac{b}{3C})$  such that for any  $x \in A, \mathbf{t} = (t_1, \dots, t_k) \in (0, 1]^{k\uparrow}$ , and  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ ,

$$(3.52) \quad h^{(k)|b}(x, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}} \text{ or } h^{(k)|b+\bar{\epsilon}}(x, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}} \implies |w_i| > 3C\bar{\delta}/c \quad \forall i \in [k].$$

$$(3.53) \quad \mathbf{d}_{J_1}(B^{\bar{\epsilon}}, \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}) > \bar{\epsilon}$$

where  $C \geq 1$  and  $0 < c \leq 1$  are the constants in Assumptions 4 and 6, respectively. Meanwhile, let

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

Note that

(3.54)

$$B_0 = (B_0 \cap B_1^c) \cup (B_0 \cap B_1 \cap B_2^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3^c) \cup (B_0 \cap B_1 \cap B_2 \cap B_3).$$

To proceed, set  $\rho^{(k)} \triangleq \left[3\rho \cdot (1 + \frac{bD}{c})\right]^k \cdot 3\rho$  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

$$(3.55) \quad \limsup_{\eta \downarrow 0} \mathbf{P}\left(B_0 \cap B_1^c\right) / \lambda^k(\eta) = 0,$$

$$(3.56) \quad \limsup_{\eta \downarrow 0} \mathbf{P}\left(B_0 \cap B_1 \cap B_2^c\right) / \lambda^k(\eta) = 0,$$

$$(3.57) \quad \limsup_{\eta \downarrow 0} \mathbf{P}\left(B_0 \cap B_1 \cap B_2 \cap B_3^c\right) / \lambda^k(\eta) = 0,$$

$$(3.58) \quad \limsup_{\eta \downarrow 0} \mathbf{P}\left(B_0 \cap B_1 \cap B_2 \cap B_3\right) / \lambda^k(\eta) = 0$$

if we pick  $\delta > 0$  sufficiently small. Now fix such  $\delta$ . Combining these claims with the decomposition of event  $B_0$  in (3.54), we establish (3.51). Now we conclude the proof of this proposition with the proofs of claims (3.55)–(3.58).

**Proof of (3.55):**

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

**Proof of (3.56):**

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

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

In particular, we focus on  $\delta \in (0, \bar{\delta} \wedge \frac{b}{2})$  with  $\bar{\delta}$  characterized in (3.52). By definition,  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) = h^{(k)|b}(x, \eta \tau_1^{(k)}(\eta), \dots, \eta \tau_k^{(k)}(\eta), \eta W_1^{(k)}(\eta), \dots, \eta W_k^{(k)}(\eta))$ . Moreover, on  $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$  we must have  $\#\{i \in [\lfloor 1/\eta \rfloor] : \eta |Z_i| > \delta\} < k$ . From the definition of  $\mathbf{Z}^{(k)}(\eta)$  in (3.46), we then have that  $\min_{i \in [k]} \eta |W_i^{(k)}(\eta)| \leq \delta$ . In light of (3.52), we yield  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) \notin B^{\bar{\epsilon}}$  on  $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ , and hence

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

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

$$(3.59) \quad \{\mathbf{X}^{\eta|b}(x) \in B\} \cap \{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\} \cap \left(\bigcap_{i=1}^k A_i(\eta, b, \epsilon, \delta, x)\right) = \emptyset$$

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

$$\limsup_{\eta \downarrow 0} \mathbf{P}(\tilde{B}) / \lambda^k(\eta) \leq \limsup_{\eta \downarrow 0} \mathbf{P}\left(\left(\bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x)\right)^c\right) / \lambda^k(\eta).$$

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 (3.59). To proceed, let process  $\check{X}_t^{\eta|b;\delta}(x)$  be the solution to

$$(3.60) \quad \frac{d\check{X}_t^{\eta|b;\delta}(x)}{dt} = a(\check{X}_t^{\eta|b;\delta}(x)) \quad \forall t \in [0, \infty) \setminus \{\eta\tau_j^{>\delta}(\eta) : j \geq 1\},$$

$$(3.61) \quad \check{X}_{\eta\tau_j^{>\delta}(\eta)}^{\eta|b;\delta}(x) = X_{\tau_j^{>\delta}(\eta)}^{\eta|b}(x) \quad \forall j \geq 1$$

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

$$\begin{aligned} & d_{J_1}(\check{\mathbf{X}}^{\eta|b;\delta}(x), \mathbf{X}^{\eta|b}(x)) \\ & \leq \sup_{t \in [0, \eta\tau_1^{>\delta}(\eta)] \cup [\eta\tau_1^{>\delta}(\eta), \eta\tau_2^{>\delta}(\eta)] \cup \dots \cup [\eta\tau_{j-1}^{>\delta}(\eta), \eta\tau_j^{>\delta}(\eta)]} \left| \check{X}_t^{\eta|b;\delta}(x) - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x) \right| \\ (3.62) \quad & \leq \rho \cdot (\epsilon + \eta C) \leq 2\rho\epsilon < \bar{\epsilon} \quad \text{due to (3.22) of Lemma 3.7.} \end{aligned}$$

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

$$\begin{aligned} & \left| \Delta \check{X}_{\eta\tau_j^{>\delta}(\eta)}^{\eta|b;\delta}(x) \right| = \left| \check{X}_{\eta\tau_j^{>\delta}(\eta)-}^{\eta|b;\delta}(x) - X_{\tau_j^{>\delta}(\eta)}^{\eta|b}(x) \right| \quad \text{see (3.61)} \\ & \leq \left| \check{X}_{\eta\tau_j^{>\delta}(\eta)-}^{\eta|b;\delta}(x) - X_{\tau_j^{>\delta}(\eta)-1}^{\eta|b}(x) \right| + \left| X_{\tau_j^{>\delta}(\eta)-1}^{\eta|b}(x) - X_{\tau_j^{>\delta}(\eta)}^{\eta|b}(x) \right| \\ (3.63) \quad & < \bar{\epsilon} + b. \end{aligned}$$

As a result, on event  $\left(\bigcap_{i=1}^k A_i(\eta, b, \epsilon, \delta, x)\right) \cap \{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ , we have  $\check{\mathbf{X}}^{\eta|b;\delta}(x) \in \mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}$ . Considering the facts that  $\mathbb{D}_A^{(k-1)|b+\bar{\epsilon}}$  is bounded away from  $B^c$  (see (3.53)) as well as  $d_{J_1}(\check{\mathbf{X}}^{\eta|b;\delta}(x), \mathbf{X}^{\eta|b}(x)) < \bar{\epsilon}$  shown in (3.62), we have just established that  $\mathbf{X}^{\eta|b}(x) \notin B$ , thus establishing (3.59).

**Proof of (3.57):**

On event  $B_1 \cap B_2 = \{\tau_k^{>\delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{>\delta}(\eta)\}$ , it follows from (3.50) that  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) = h^{(k)|b}(x, \eta W_1^{>\delta}(\eta), \dots, \eta W_k^{>\delta}(\eta), \eta\tau_1^{>\delta}(\eta), \dots, \eta\tau_k^{>\delta}(\eta))$ . 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 (3.52), on event  $B_1 \cap B_2 \cap B_3^c$  we have  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) \notin B$ , and hence

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



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

(3.64)

$$\{\mathbf{X}^{\eta|b}(x) \in B\} \cap \{\tau_k^{\geq \delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{\geq \delta}(\eta); \eta|W_i^{\geq \delta}(\eta)| \leq \bar{\delta} \text{ for some } i \in [k]\} \\ \cap \left( \bigcap_{i=1}^{k+1} A_i(\eta, b, \epsilon, \delta, x) \right) = \emptyset.$$

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

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

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

We are left with proving the claim (3.64). First, note that on this event, there exists some  $J \in [k]$  such that  $\eta|W_J^{\geq \delta}(\eta)| \leq \bar{\delta}$ . Next, recall the definition of  $\check{X}_t^{\eta|b;\delta}(x)$  in (3.60)-(3.61), and note that it has been shown in (3.62) (with  $j = k + 1$ ) that

$$(3.65) \quad \sup_{t \in [0,1]} \left| \check{X}_t^{\eta|b;\delta}(x) - X_{\lfloor t/\eta \rfloor}^{\eta|b}(x) \right| < 2\rho\epsilon < \bar{\epsilon}.$$

If we can show that  $\check{X}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$ , then (3.65) immediately leads to  $\mathbf{X}^{\eta|b}(x) \notin B$ , thus proving claim (3.64). To proceed, first note that

$$\begin{aligned} & \left| \Delta \check{X}_{\eta\tau_J^{\geq \delta}(\eta)}^{\eta|b;\delta}(x) \right| \\ & \leq \left| \check{X}_{\eta\tau_J^{\geq \delta}(\eta)-}^{\eta|b;\delta}(x) - X_{\tau_J^{\geq \delta}(\eta)-1}^{\eta|b}(x) \right| + \left| X_{\tau_J^{\geq \delta}(\eta)-1}^{\eta|b}(x) - X_{\tau_J^{\geq \delta}(\eta)}^{\eta|b}(x) \right| \quad \text{see (3.61)} \\ & \leq 2\rho\epsilon + \eta \left| a(X_{\tau_J^{\geq \delta}(\eta)-1}^{\eta|b}(x)) + \sigma(X_{\tau_J^{\geq \delta}(\eta)-1}^{\eta|b}(x))W_J^{\geq \delta}(\eta) \right| \quad \text{using (3.65)} \\ & \leq 2\rho\epsilon + \eta C + C\bar{\delta} < 3C\bar{\delta} \quad \text{due to } 2\rho\epsilon < \bar{\delta}, \eta < \bar{\delta}, \text{ and } C \geq 1. \end{aligned}$$

Meanwhile, the calculations in (3.63) can be repeated to show that  $\check{X}^{\eta|b;\delta}(x) \in \mathbb{D}_A^{(k)|b+\bar{\epsilon}}$ , and hence  $\check{X}^{\eta|b;\delta}(x) = h^{(k)|b+\bar{\epsilon}}(x, \tilde{w}_1, \dots, \tilde{w}_k, \eta\tau_1^{\geq \delta}(\eta), \dots, \eta\tau_k^{\geq \delta}(\eta))$  for some  $(\tilde{w}_1, \dots, \tilde{w}_k) \in \mathbb{R}^k$ . Due to  $0 < c \leq \sigma(y) \leq C \forall y \in \mathbb{R}$  (see Assumptions 4 and 6),

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

which implies  $|\tilde{w}_J| < 3C\bar{\delta}/c$ . In light of our choice of  $\bar{\delta}$  in (3.52), we yield  $\check{X}^{\eta|b;\delta}(x) \notin B^{\bar{\epsilon}}$  and conclude the proof.

**Proof of (3.58):**

We focus on  $\delta \in (0, \bar{\delta} \wedge \frac{b}{2C})$ . On event  $B_1 \cap B_2 = \{\tau_k^{\geq \delta}(\eta) \leq \lfloor 1/\eta \rfloor < \tau_{k+1}^{\geq \delta}(\eta)\}$ ,  $\hat{\mathbf{X}}^{\eta|b;(k)}$  admits the expression in (3.50). 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

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

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

Therefore,

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

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

**LEMMA 3.15.** *Let Assumption 1 hold. Let  $A \subseteq \mathbb{R}$  be a compact set. Let bounded function  $\Phi : \mathbb{R} \times \mathbb{R}^k \times (0, 1]^{k\uparrow} \rightarrow \mathbb{R}$  be continuous on  $\mathbb{R} \times \mathbb{R}^k \times (0, 1)^{k\uparrow}$ . For any  $\delta > 0, c > \delta$  and  $k = 0, 1, 2, \dots$ ,*

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

$$\text{where } \phi_{c,k}(x) \triangleq \mathbf{E} \left[ \Phi(x, W_1^*(c), \dots, W_k^*(c), U_{(1;k)}, \dots, U_{(k;k)}) \right].$$

**PROOF.** Fix some  $\delta > 0, c > \delta$  and  $k = 0, 1, \dots$ . 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

$$(3.66) \quad \left| \lambda^{-k}(\eta_n) \mathbf{E} \left[ \Phi(x_n, \mathbf{W}^{\eta_n}, \boldsymbol{\tau}^{\eta_n}) \mathbb{I}_{E_{c,k}^\delta(\eta_n)} \right] - (1/k!) \cdot c^{-\alpha k} \cdot \phi_{c,k}(x_n) \right| > \epsilon \quad \forall n \geq 1$$

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

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

where  $\mathbf{W}^* \triangleq (W_j^*(c))_{j=1}^k$ ,  $\mathbf{U}^* \triangleq (U_{(j;k)})_{j=1}^k$ . However, by Bounded Convergence theorem, we see that  $\phi_{c,k}$  is also continuous, and hence  $\phi_{c,k}(x_n) \rightarrow \phi_{c,k}(x^*)$ . This leads to a contradiction with (3.66) and allows us to conclude the proof.  $\square$

We are now ready to prove Proposition 3.14.

**PROOF OF PROPOSITION 3.14.** Fix some  $b > 0$ ,  $k \in \mathbb{Z}_+$ , and  $g \in \mathcal{C}(\mathbb{D} \setminus \mathbb{D}_A^{(k-1)b})$  (i.e.,  $g : \mathbb{D} \rightarrow [0, \infty)$  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  $\mathbf{t} = (t_1, \dots, t_k) \in (0, 1)^{k\uparrow}$ ,  $\mathbf{w} = (w_1, \dots, w_k) \in \mathbb{R}^k$ :

$$(3.67) \quad h^{(k)b}(x_0, \mathbf{w}, \mathbf{t}) \in B^{\bar{\epsilon}} \implies |w_j| > \bar{\delta} \quad \forall j \in [k].$$

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

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

For term  $I_1(\eta, x)$ , it follows from (3.4) that  $\sup_{x \in \mathbb{R}} \mathbf{E}[I_1(\eta, x)] \leq \|g\| \cdot \left[ \frac{1}{\eta_n} \cdot H(\delta/\eta_n) \right]^{k+1}$ .

Therefore,  $\lim_{\eta \downarrow 0} \sup_{x \in A} \mathbf{E}[I_1(\eta, x)] / (\eta^{-1} H(\eta^{-1}))^k \leq \frac{\|g\|}{\delta^{\alpha(k+1)}} \cdot \lim_{n \rightarrow \infty} \frac{H(1/\eta)}{\eta} = 0$ .

By definition,  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) = h^{(k)|b}(x, \eta \tau_1^{(k)}(\eta), \dots, \eta \tau_k^{(k)}(\eta), \eta W_1^{(k)}(\eta), \dots, \eta W_k^{(k)}(\eta))$ . Moreover, on event  $\{\tau_k^{>\delta}(\eta) > \lfloor 1/\eta \rfloor\}$ , we must have  $\#\{i \in [\lfloor 1/\eta \rfloor] : \eta |Z_i| > \delta\} < k$ . From the definition of  $\mathbf{Z}^{(k)}(\eta)$  in (3.46), we then have that  $\min_{i \in [k]} \eta |W_i^{(k)}(\eta)| \leq \delta$ . In light of (3.67) 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  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) \notin B$  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  $\hat{\mathbf{X}}^{\eta|b;(k)}(x)$  admits the expression in (3.50), which implies  $\hat{\mathbf{X}}^{\eta|b;(k)}(x) \notin B$ . due to (3.67) and our choice of  $\delta < \bar{\delta}$ . In summary, we get  $I_3(\eta, x) = 0$ .

Lastly, on event  $E_{\bar{\delta}, k}^\delta(\eta)$ , the process  $\hat{\mathbf{X}}^{\eta|b;(k)}(x)$  would again admits the expression in (3.50). As a result, for any  $\eta > 0$  and  $x \in A$ , we have

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

where  $\mathbf{W}^{>\delta}(\eta) = (W_1^{>\delta}(\eta), \dots, W_k^{>\delta}(\eta))$ ,  $\boldsymbol{\tau}^{>\delta}(\eta) = (\tau_1^{>\delta}(\eta), \dots, \tau_k^{>\delta}(\eta))$ , and  $\Phi : \mathbb{R}^k \times (0, 1)^{k\uparrow} \rightarrow \mathbb{R}$  is defined as  $\Phi(x_0, \mathbf{w}, \mathbf{t}) \triangleq g(h^{(k)|b}(x_0, \mathbf{w}, \mathbf{t}))$ . Meanwhile, let  $\phi(x) \triangleq \mathbf{E}[\Phi(x, W_1^*(\bar{\delta}), \dots, W_k^*(\bar{\delta}), U_{(1;k)}, \dots, U_{(k;k)})]$ . First, the continuity of mapping  $\Phi$  on  $\mathbb{R}^k \times (0, 1)^{k\uparrow}$  follows directly from the continuity of  $g$  and  $h^{(k)|b}$  (see Lemma C.3 of the Supplementary Material). 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

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

due to Lemma 3.15. Meanwhile, due to continuity of  $\phi(\cdot)$ , for any  $x_n, x^* \in A$  with  $\lim_{n \rightarrow \infty} x_n = x^*$ , we have  $\lim_{n \rightarrow \infty} \phi(x_n) = \phi(x^*)$ . 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

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

First, using (3.67), we must have  $g(h^{(k)}(x^*, w_1, \dots, w_k, t)) = 0$  if there is some  $j \in [k]$  with  $|w_j| \leq \bar{\delta}$ . Next,  $\mathbf{P}\left(U_{(1;k)} \in dt_1, \dots, U_{(k;k)} \in dt_k\right) = k! \cdot \mathbb{I}\{0 < t_1 < t_2 < \dots < t_k < 1\} \mathcal{L}_1^{k\uparrow}(dt_1, \dots, dt_k)$  where  $\mathcal{L}_1^{k\uparrow}$  is the Lebesgue measure restricted on  $(0, 1)^{k\uparrow}$ . The conclusion of the proof then follows from

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

where we appealed to the definition in (2.20) in the last equality.  $\square$

**4. Metastability 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*

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

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

**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

$$\rho_i^\eta(x) \triangleq \inf \left\{ j \geq \rho_{i-1}^\eta(x) + \lfloor rT/\eta \rfloor : V_j^\eta(x) \in A(\epsilon) \right\}$$

where  $\rho_0^\eta(x) = 0$ . 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  $0 < r < R$ ,

$$\begin{aligned} \mathbf{P}\left(\tau_{I(\epsilon)^c}^\eta(y) \in (RT/\eta, \rho_1^\eta(y)]\right) &\leq \mathbf{P}\left(\tau_{I(\epsilon)^c}^\eta(y) \wedge \rho_1^\eta(y) > RT/\eta\right) \\ &\leq \mathbf{P}\left(V_j^\eta(y) \in I(\epsilon) \setminus A(\epsilon) \quad \forall j \in [\lfloor rT/\eta \rfloor, RT/\eta]\right) \\ &\leq \sup_{z \in I(\epsilon) \setminus A(\epsilon)} \mathbf{P}\left(\tau_{(I(\epsilon) \setminus A(\epsilon))^c}^\eta(z) > \frac{R-r}{2}T/\eta\right) \\ &= \gamma(\eta)T/\eta \cdot o(1) \end{aligned} \tag{4.1}$$

where the last equality is from (2.30) of Condition 1, and

$$\begin{aligned} &\sup_{y \in A(\epsilon)} \mathbf{P}\left(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)\right) \\ &\leq \sup_{y \in A(\epsilon)} \mathbf{P}\left(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq RT/\eta\right) + \sup_{y \in A(\epsilon)} \mathbf{P}\left(\tau_{I(\epsilon)^c}^\eta(y) \in (RT/\eta, \rho_1^\eta(y)]\right) \end{aligned} \tag{4.2}$$

$$\begin{aligned}
&\leq \sup_{y \in A(\epsilon)} \mathbf{P}\left(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq RT/\eta\right) + \gamma(\eta)T/\eta \cdot o(1) \\
&\leq (C(B^-) + \delta_B(\epsilon, RT) + o(1)) \cdot \gamma(\eta)RT/\eta
\end{aligned}$$

where the second inequality is from (4.1) and the last equality is from (2.29) of Condition 1.

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

$$\begin{aligned}
&\inf_{x \in A(\epsilon)} \mathbf{P}(\gamma(\eta)\tau_{I(\epsilon)^c}^\eta(x) > t; V_{\tau_\epsilon}^\eta(x) \in B) \\
&\geq \inf_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_K^\eta(x); V_{\tau_\epsilon}^\eta(x) \in B) \\
&= \inf_{x \in A(\epsilon)} \sum_{j=K}^{\infty} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \in (\rho_j^\eta(x), \rho_{j+1}^\eta(x)]; V_{\tau_\epsilon}^\eta(x) \in B) \\
&\geq \inf_{x \in A(\epsilon)} \sum_{j=K}^{\infty} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \in (\rho_j^\eta(x), \rho_j^\eta(x) + T/\eta]; V_{\tau_\epsilon}^\eta(x) \in B) \\
&\geq \inf_{x \in A(\epsilon)} \sum_{j=K}^{\infty} \inf_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \leq T/\eta; V_{\tau_\epsilon}^\eta(y) \in B) \cdot \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x)). \\
(4.3) \quad &\geq \inf_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \leq T/\eta; V_{\tau_\epsilon}^\eta(y) \in B) \cdot \sum_{j=K}^{\infty} \inf_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x)).
\end{aligned}$$

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

$$\begin{aligned}
(4.4) \quad &\sum_{j=K}^{\infty} \inf_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x)) \\
&\geq \sum_{j=0}^{\infty} \left( \inf_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) > \rho_1^\eta(y)) \right)^{K+j} = \sum_{j=0}^{\infty} \left( 1 - \sup_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)) \right)^{K+j} \\
&= \frac{1}{\sup_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y))} \cdot \left( 1 - \sup_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)) \right)^{\left\lceil \frac{t/\gamma(\eta)}{T/\eta} \right\rceil} \\
&\geq \frac{1}{(1 + \delta_{\mathbb{S}}(\epsilon, RT) + o(1)) \cdot \gamma(\eta)RT/\eta} \cdot \left( 1 - (1 + \delta_{\mathbb{S}}(\epsilon, RT) + o(1)) \cdot \gamma(\eta)RT/\eta \right)^{\frac{t/\gamma(\eta)}{T/\eta} + 1}.
\end{aligned}$$

where the last inequality is from (4.2) with  $B = \mathbb{S}$ . From (4.3), (4.4), and (2.28) of Condition 1, we have

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

$$\geq \frac{C(B^\circ) - \delta_B(\epsilon, T)}{1 + \delta_{\mathbb{S}}(\epsilon, RT)} \cdot \exp\left(- (1 + \delta_{\mathbb{S}}(\epsilon, RT)) \cdot R \cdot t\right).$$

By taking limit  $T \rightarrow \infty$  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  $k = \left\lfloor \frac{t/\gamma(\eta)}{T/\eta} \right\rfloor$  and note that

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

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

$$\{\tau_{I(\epsilon)^c}^\eta(x) > t/\gamma(\eta); \rho_k^\eta(x) > t/\gamma(\eta)\} \subseteq \bigcup_{j=1}^k \{\tau_{I(\epsilon)^c}^\eta(x) \wedge \rho_j^\eta(x) - \rho_{j-1}^\eta(x) \geq T/\eta\}$$

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

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

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

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

which proves the claim.

Now, with the claim in hand, we have that

$$\begin{aligned} \text{(II)} &\leq \sum_{j=1}^k \sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \wedge \rho_j^\eta(x) - \rho_{j-1}^\eta(x) \geq T/\eta) \\ &= \sum_{j=1}^k \sup_{x \in A(\epsilon)} \mathbf{E}\left[\mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \wedge \rho_j^\eta(x) - \rho_{j-1}^\eta(x) \geq T/\eta \mid \mathcal{F}_{\rho_{j-1}^\eta(x)}^\eta)\right] \\ &\leq \sum_{j=1}^k \sup_{y \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(y) \wedge \rho_1^\eta(y) \geq T/\eta) \\ &\leq \frac{t}{\gamma(\eta)T/\eta} \cdot \gamma(\eta)T/\eta \cdot o(1) = o(1) \end{aligned}$$



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.

$$\begin{aligned}
(I) &= \sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > t/\gamma(\eta) \geq \rho_K^\eta(x); V_{\tau_\epsilon}^\eta(x) \in B) \\
&\leq \sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_K^\eta(x); V_{\tau_\epsilon}^\eta(x) \in B) \\
&= \sum_{j=k}^{\infty} \sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) \in (\rho_j^\eta(x), \rho_{j+1}^\eta(x)]; V_{\tau_\epsilon}^\eta(x) \in B) \\
&= \sum_{j=k}^{\infty} \sup_{x \in A(\epsilon)} \mathbf{E} \left[ \mathbf{E} \left[ \mathbb{I}\{V_{\tau_\epsilon}^\eta(x) \in B\} \cdot \mathbb{I}\{\tau_{I(\epsilon)^c}^\eta(x) \leq \rho_{j+1}^\eta(x)\} \middle| \mathcal{F}_{\rho_j^\eta(x)} \right] \right. \\
&\quad \left. \cdot \mathbb{I}\{\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x)\} \right] \\
&\leq \sum_{j=k}^{\infty} \sup_{x \in A(\epsilon)} \mathbf{E} \left[ \sup_{y \in A(\epsilon)} \mathbf{P}(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)) \cdot \mathbb{I}\{\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x)\} \right] \\
&= \sup_{y \in A(\epsilon)} \mathbf{P}(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)) \cdot \sum_{j=k}^{\infty} \sup_{x \in A(\epsilon)} \mathbf{P}(\tau_{I(\epsilon)^c}^\eta(x) > \rho_j^\eta(x))
\end{aligned}$$

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

$$\begin{aligned}
&\sup_{y \in A(\epsilon)} \mathbf{P}(V_{\tau_\epsilon}^\eta(y) \in B; \tau_{I(\epsilon)^c}^\eta(y) \leq \rho_1^\eta(y)) \\
&\leq (C(B^-) + \delta_B(\epsilon, T) + o(1)) \cdot \gamma(\eta)T/\eta + \frac{1-r}{2} \cdot \gamma(\eta)T/\eta \cdot o(1)
\end{aligned}$$

whereas the second term is bounded via (2.28) of Condition 1 as follows:

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

Therefore,

$$\begin{aligned}
&\limsup_{\eta \downarrow 0} \sup_{x \in A(\epsilon)} \mathbf{P}(\gamma(\eta)\tau_{I(\epsilon)^c}^\eta(x) > t; V_{\tau_\epsilon}^\eta(x) \in B) \\
&\leq \frac{C(B^-) + \delta_B(\epsilon, T)}{r \cdot (1 - \delta_B(\epsilon, rT))} \cdot \exp \left( -r \cdot (1 - \delta_B(\epsilon, rT)) \cdot t \right).
\end{aligned}$$

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

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  $B \subseteq \mathbb{S}$ ,

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

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

$$\begin{aligned} \mathbf{P} \left( \gamma(\eta) \cdot \tau_{I^c}^\eta(x) > t, V_\tau^\eta(x) \in B \right) &= \underbrace{\mathbf{P} \left( \gamma(\eta) \cdot \tau_{I^c}^\eta(x) > t, V_\tau^\eta(x) \in B, V_{\tau_\epsilon}^\eta(x) \in I \right)}_{(I)} \\ &\quad + \underbrace{\mathbf{P} \left( \gamma(\eta) \cdot \tau_{I^c}^\eta(x) > t, V_\tau^\eta(x) \in B, V_{\tau_\epsilon}^\eta(x) \notin I \right)}_{(II)}. \end{aligned}$$

Since

$$(I) \leq \mathbf{P} \left( V_{\tau_\epsilon}^\eta(x) \in I \right) \quad \text{and} \quad (II) = \mathbf{P} \left( \gamma(\eta) \cdot \tau_\epsilon^\eta(x) > t, V_{\tau_\epsilon}^\eta(x) \in B \setminus I \right),$$

we have that

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

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

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

Taking  $\epsilon \rightarrow 0$ , 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  $x \in I$ ,

$$(4.6) \quad \begin{aligned} \mathbf{P} \left( \gamma(\eta) \cdot \tau_\epsilon^\eta(x) > t, V_{\tau_\epsilon}^\eta(x) \in B \right) &= \mathbf{E} \left[ \mathbf{P} \left( \gamma(\eta) \cdot \tau_\epsilon^\eta(x) > t, V_{\tau_\epsilon}^\eta(x) \in B \middle| \mathcal{F}_{\tau_{A(\epsilon)}^\eta(x)} \right) \right. \\ &\quad \left. \cdot \left( \mathbb{I} \{ \tau_{A(\epsilon)}^\eta(x) \leq T/\eta \} + \mathbb{I} \{ \tau_{A(\epsilon)}^\eta(x) > T/\eta \} \right) \right]. \end{aligned}$$

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

$$\begin{aligned}
& \mathbf{P}\left(\gamma(\eta) \cdot \tau_\epsilon^\eta(x) > t, V_{\tau_\epsilon^\eta}^\eta(x) \in B\right) \\
& \leq \mathbf{E}\left[\sup_{y \in A(\epsilon)} \mathbf{P}\left(\tau_\epsilon^\eta(y) > t/\gamma(\eta) - T/\eta, V_{\tau_\epsilon^\eta}^\eta(y) \in B\right) \cdot \mathbb{I}\{\tau_{A(\epsilon)}^\eta(x) \leq T/\eta\}\right] \\
& \quad + \mathbf{P}\left(\tau_{A(\epsilon)}^\eta(x) > T/\eta\right) \\
& \leq \sup_{y \in A(\epsilon)} \mathbf{P}\left(\gamma(\eta) \cdot \tau_\epsilon^\eta(y) > t - s, V_{\tau_\epsilon^\eta}^\eta(y) \in B\right) + \mathbf{P}\left(\tau_{A(\epsilon)}^\eta(x) > T/\eta\right)
\end{aligned}$$

for sufficiently small  $\eta$ 's; here, we applied  $\gamma(\eta)/\eta \rightarrow 0$  as  $\eta \downarrow 0$  in the last inequality. In light of (2.31) of Condition 1, by taking  $\eta \rightarrow 0$  uniformly over  $x \in I(\epsilon')$  and then  $T \rightarrow \infty$  we yield

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

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

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

which is the desired lower bound of the claim (4.5). This concludes the proof.  $\square$

**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

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

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

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

for the ODE  $\mathbf{y}_t(x)$  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  $t(\cdot)$  immediately implies

$$(4.9) \quad \mathbf{y}_t(y) \in [-\epsilon, \epsilon] \quad \forall y \in I_\epsilon^-, \quad t \geq t(\epsilon).$$

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

LEMMA 4.2. *The following claim holds for some  $\bar{t} > 0$ : 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  $B \subseteq (I_{\bar{\epsilon}/2})^c$ ,*

$$\begin{aligned} (T - \bar{t}) \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B_\Delta) &\leq \inf_{x \in [-\epsilon, \epsilon]} \mathbf{C}_{[0, T]}^{(\mathcal{J}_b^*)|b} \left( (\check{E}(\epsilon, B, T))^{\circ}; x \right) \\ &\leq \sup_{x \in [-\epsilon, \epsilon]} \mathbf{C}_{[0, T]}^{(\mathcal{J}_b^*)|b} \left( (\check{E}(\epsilon, B, T))^{-}; x \right) \leq T \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\Delta) \end{aligned}$$

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

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  $C_b^* = \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(I^c)$ , and note that it is established in Lemma D.3 in the Supplementary Material 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

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

Note that  $\partial I = \{s_{\text{left}}, s_{\text{right}}\}$  and recall our assumption  $s_{\text{left}}/b \notin \mathbb{Z}$  and  $s_{\text{right}}/b \notin \mathbb{Z}$ . Also, our choice of constant  $\bar{\epsilon}$  in (4.7) ensures that  $l = |s_{\text{left}}| \wedge s_{\text{right}} > (\mathcal{J}_b^* - 1) \cdot b + \bar{\epsilon}$ . Lemma D.2 in the Supplementary Material then verifies  $\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(\partial I) = 0$  and hence  $C(\partial I) = 0$ . Besides, note that  $\gamma(\eta)T/\eta = C_b^*T \cdot (\lambda(\eta))^{\mathcal{J}_b^*}$ .

We start by establishing conditions (2.28) and (2.29). First, given any  $B \subseteq \mathbb{R}$  we specify the choice of function  $\delta_B(\epsilon, T)$  in Condition 1. From the continuity of measures, we get  $\lim_{\Delta \downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((B^\Delta \cap I^c) \setminus (B^- \cap I^c)) = 0$  and  $\lim_{\Delta \downarrow 0} \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((B^\circ \cap I^c) \setminus (B_\Delta \cap I^c)) = 0$ . This allows us to fix a sequence  $(\Delta^{(n)})_{n \geq 1}$  such that  $\Delta^{(n+1)} \in (0, \Delta^{(n)}/2)$  and

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

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

$$(4.12) \quad \sup_{x \in [-\epsilon, \epsilon]} \mathbf{C}_{[0, T]}^{(\mathcal{J}_b^*)|b} \left( (\check{E}(\epsilon, \tilde{B}(\epsilon), T))^{-}; x \right) \leq T \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((B \setminus I_\epsilon)^{\Delta^{(n)}}),$$

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

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

$$\begin{aligned} \check{\delta}_B(\epsilon, T) &= T \cdot \max \left\{ \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((B^{\Delta^{(n)}} \cap I^c) \setminus (B^- \cap I^c)), \right. \\ &\quad \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((B^\circ \cap I^c) \setminus (B_{\Delta^{(n)}} \cap I^c)), \\ &\quad \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}((\partial I)^{\epsilon + \Delta^{(n)}}) \left. \right\} \\ &\quad + \bar{t} \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b}(B^\circ \setminus I) \end{aligned} \tag{4.14}$$

and  $\delta_B(\epsilon, T) = \check{\delta}_B(\epsilon, T)/(C_b^* \cdot T)$ . First, due to (4.11) and  $\check{C}^{(\mathcal{J}_b^*)|b}(B \setminus I) \leq \check{C}^{(\mathcal{J}_b^*)|b}(I^c) < \infty$ , we get

$$\lim_{T \rightarrow \infty} \delta_B(\epsilon, T) \leq \frac{1}{C_b^*} \cdot \left[ \frac{1}{2^{n_\epsilon}} \vee \check{C}^{(\mathcal{J}_b^*)|b}((\partial I)^{\epsilon + \Delta^{(n_\epsilon)}}) \right]$$

where  $n_\epsilon$  is the unique positive integer satisfying  $\epsilon \in (\epsilon^{(n+1)}, \epsilon^{(n)}]$ . Moreover, as  $\epsilon \downarrow 0$  we get  $n_\epsilon \rightarrow \infty$ . Since  $\partial I$  is closed, we get  $\cap_{r>0} (\partial I)^r = \partial I$ , which then implies  $\lim_{r \downarrow 0} \check{C}^{(\mathcal{J}_b^*)|b}((\partial I)^r) = \check{C}^{(\mathcal{J}_b^*)|b}(\partial I) = 0$  due to continuity of measures. In summary, we have verified that  $\lim_{\epsilon \downarrow 0} \lim_{T \rightarrow \infty} \delta_B(\epsilon, T) = 0$ .

Now, we are ready to verify conditions (2.28) and (2.29). Specifically, we introduce stopping times

$$(4.15) \quad \tau_\epsilon^{\eta|b}(x) \triangleq \min \{j \geq 0 : X_j^{\eta|b}(x) \notin I_\epsilon\}.$$

LEMMA 4.3 (Verifying conditions (2.28) and (2.29)). *Let  $\bar{t}$  be the constant in Lemma D.1 of Supplementary Material. Given measurable  $B \subseteq \mathbb{R}$ ,  $T > \bar{t}$ , and any  $\epsilon > 0$  small enough,*

$$\begin{aligned} \liminf_{\eta \downarrow 0} \inf_{x \in (-\epsilon, \epsilon)} \frac{\mathbf{P}\left(\tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B\right)}{\gamma(\eta)T/\eta} &\geq C(B^\circ) - \delta_B(\epsilon, T), \\ \limsup_{\eta \downarrow 0} \sup_{x \in (-\epsilon, \epsilon)} \frac{\mathbf{P}\left(\tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B\right)}{\gamma(\eta)T/\eta} &\leq C(B^-) + \delta_B(\epsilon, T). \end{aligned}$$

PROOF. Recall that  $\gamma(\eta)T/\eta = C_b^*T \cdot (\lambda(\eta))^{\mathcal{J}_b^*}$ ,  $C(\cdot) = \check{C}^{(\mathcal{J}_b^*)|b}(\cdot \setminus I)/C_b^*$ , and  $\delta_B(\epsilon, T) = \check{\delta}_B(\epsilon, T)/(C_b^* \cdot T)$ . By rearranging the terms, it suffices to show that

$$(4.16) \quad \limsup_{\eta \downarrow 0} \sup_{x \in (-\epsilon, \epsilon)} \frac{\mathbf{P}\left(\tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B\right)}{(\lambda(\eta))^{\mathcal{J}_b^*}} \leq T \cdot \check{C}^{(\mathcal{J}_b^*)|b}(B^- \setminus I) + \check{\delta}_B(\epsilon, T),$$

$$(4.17) \quad \liminf_{\eta \downarrow 0} \inf_{x \in (-\epsilon, \epsilon)} \frac{\mathbf{P}\left(\tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B\right)}{(\lambda(\eta))^{\mathcal{J}_b^*}} \geq T \cdot \check{C}^{(\mathcal{J}_b^*)|b}(B^\circ \setminus I) - \check{\delta}_B(\epsilon, T).$$

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

$$\begin{aligned} \left\{ \tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B \right\} &= \left\{ \tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in \tilde{B}(\epsilon) \right\} \\ &= \left\{ \mathbf{X}_{[0, T]}^{\eta|b}(x) \in \check{E}(\epsilon, \tilde{B}(\epsilon), T) \right\}. \end{aligned}$$

For any  $\epsilon \in (0, \bar{\epsilon})$  and  $\xi \in \check{E}(\epsilon, \tilde{B}(\epsilon), T)$ , 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 D.1, it holds for all  $\xi \in \mathbb{D}_{[-\epsilon, \epsilon]}^{(\mathcal{J}_b^* - 1)|b}[0, T]$  that  $\sup_{t \in [0, T]} |\xi(t)| < l - 2\bar{\epsilon}$ . In summary, we have established that

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

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

$$\begin{aligned}
 (4.18) \quad & \limsup_{\eta \downarrow 0} \sup_{x \in [-\epsilon, \epsilon]} \frac{\mathbf{P}\left(\tau_\epsilon^{\eta|b}(x) \leq T/\eta; X_{\tau_\epsilon^{\eta|b}(x)}^{\eta|b}(x) \in B\right)}{(\lambda(\eta))^{\mathcal{J}_b^*}} \\
 & \leq \sup_{x \in [-\epsilon, \epsilon]} \mathbf{C}_{[0, T]}^{(\mathcal{J}_b^*)|b} \left( (\check{E}(\epsilon, \tilde{B}(\epsilon), T))^-; x \right) \\
 & \leq T \cdot \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B \setminus I_\epsilon)^{\Delta^{(n)}} \right);
 \end{aligned}$$

here the last inequality we applied property (4.12). Furthermore,

$$\begin{aligned}
 & \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B \setminus I_\epsilon)^{\Delta^{(n)}} \right) \\
 & \leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}} \right) \quad \text{due to } (E \cup F)^\Delta \subseteq E^\Delta \cup F^\Delta \\
 & = \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}} \cap I^c \right) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^{\Delta^{(n)}} \cup (I_\epsilon^c)^{\Delta^{(n)}} \cap I \right) \\
 & \leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^{\Delta^{(n)}} \setminus I \right) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (I_\epsilon^c)^{\Delta^{(n)}} \cap I \right) \\
 & \leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^{\Delta^{(n)}} \setminus I \right) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (\partial I)^{\epsilon + \Delta^{(n)}} \right) \\
 & \leq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^- \setminus I \right) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B^{\Delta^{(n)}} \cap I^c) \setminus (B^- \cap I^c) \right) + \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (\partial I)^{\epsilon + \Delta^{(n)}} \right)
 \end{aligned}$$

Considering the definition of  $\check{\delta}_B$  in (4.14) and the choice of  $C(\cdot)$  in (4.10), 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

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

Furthermore, from the preliminary bound  $(E \cap F)_\Delta \supseteq E_\Delta \cap F_\Delta$  we get

$$\check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B \setminus I_\epsilon)_{\Delta^{(n)}} \right) \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B \setminus I)_{\Delta^{(n)}} \right) \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B_{\Delta^{(n)}} \cap I_{\Delta^{(n)}}^c \right).$$

Together with the fact that  $B_\Delta \setminus I = B_\Delta \cap I^c \subseteq (B_\Delta \cap (I^c)_\Delta) \cup (I^c \setminus (I^c)_\Delta)$ , we yield

$$\begin{aligned}
 & \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B \setminus I_\epsilon)_{\Delta^{(n)}} \right) \\
 & \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B_{\Delta^{(n)}} \setminus I \right) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( I^c \setminus I_{\Delta^{(n)}}^c \right) \\
 & \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B_{\Delta^{(n)}} \setminus I \right) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (\partial I)^{\Delta^{(n)}} \right) \\
 & \geq \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( B^\circ \setminus I \right) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (B^\circ \cap I^c) \setminus (B_{\Delta^{(n)}} \cap I^c) \right) - \check{\mathbf{C}}^{(\mathcal{J}_b^*)|b} \left( (\partial I)^{\Delta^{(n)}} \right).
 \end{aligned}$$

Plugging this bound back into (4.19), we establish the lower bound (4.17) and conclude the proof.  $\square$

The next two results verify conditions (2.30) and (2.31). Let

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

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  $\{\tau_{(I(\epsilon) \setminus A(\epsilon))^c}^\eta(x) > T/\eta\}$  in condition (2.30) means that  $X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon)$  for all  $j \leq T/\eta$ . Also, recall that  $\gamma(\eta)T/\eta = C_b^* T \cdot (\lambda(\eta))^{\mathcal{J}_b^*}$ . Therefore, to verify condition (2.30), it suffices to prove the following result.

LEMMA 4.4 (Verifying condition (2.30)). *Given any  $k \geq 1$  and  $\epsilon \in (0, \bar{\epsilon})$ , it holds for all  $T \geq k \cdot t(\epsilon/2)$  that*

$$\lim_{\eta \downarrow 0} \sup_{x \in I_\epsilon^-} \frac{1}{\lambda^{k-1}(\eta)} \mathbf{P} \left( X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon) \quad \forall j \leq T/\eta \right) = 0.$$

PROOF. First,  $\{X_j^{\eta|b}(x) \in I_\epsilon \setminus (-\epsilon, \epsilon) \quad \forall j \leq T/\eta\} = \{\mathbf{X}_{[0,T]}^{\eta|b}(x) \in E(\epsilon)\}$  where

$$E(\epsilon) \triangleq \{\xi \in \mathbb{D}[0, T] : \xi(t) \in I_\epsilon \setminus (-\epsilon, \epsilon) \quad \forall t \in [0, T]\}.$$

Recall the definition of  $\mathbb{D}_A^{(k)|b}[0, T]$  in (2.19). We claim that  $E(\epsilon)$  is bounded away from  $\mathbb{D}_{I_\epsilon^-}^{(k-1)|b}[0, T]$ . This allows us to apply Theorem 2.4 and conclude that

$$\sup_{x \in I_\epsilon^-} \mathbf{P} \left( \mathbf{X}_{[0,T]}^{\eta|b}(x) \in E(\epsilon) \right) = \mathcal{O}(\lambda^k(\eta)) = o(\lambda^{k-1}(\eta)) \quad \text{as } \eta \downarrow 0.$$

Now it only remains to verify that  $E(\epsilon)$  is bounded away from  $\mathbb{D}_{I_\epsilon^-}^{(k-1)|b}[0, T]$ , which can be established if we show that for any  $\xi \in \mathbb{D}_{I_\epsilon^-}^{(k-1)|b}[0, T]$  and  $\xi' \in E(\epsilon)$ ,

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

First, if  $\xi(t) \notin I_{\epsilon/2}$  for some  $t \leq T$ , then by definition of  $E(\epsilon)$  we get  $\mathbf{d}_{J_1}^{[0,T]}(\xi, \xi') \geq \frac{\epsilon}{2}$ . Now suppose that  $\xi(t) \in I_{\epsilon/2}$  for all  $t \leq T$ . Let  $x_0 \in I_\epsilon^-$ ,  $(w_1, \dots, w_{k-1}) \in \mathbb{R}^{k-1}$ , and  $(t_1, \dots, t_{k-1}) \in (0, T]^{k-1 \uparrow}$  be such that  $\xi = h_{[0,T]}^{(k-1)|b}(x_0, w_1, \dots, w_{k-1}, t_1, \dots, t_{k-1})$ . With the convention that  $t_0 = 0$  and  $t_k = T$ , we have

$$(4.22) \quad \xi(t) = \mathbf{y}_{t-t_{j-1}}(\xi(t_{j-1})) \quad \forall t \in [t_{j-1}, t_j].$$

for each  $j \in [k]$ . Here  $\mathbf{y}_\cdot(x)$  is the ODE defined in (2.22). Also, note that due to the assumption  $T \geq k \cdot t(\epsilon/2)$ , there exists some  $j \in [k]$  such that  $t_j - t_{j-1} \geq t(\epsilon/2)$ . 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  $\mathbf{d}_{J_1}^{[0,T]}(\xi, \xi') \geq \frac{\epsilon}{2}$ . This concludes the proof.  $\square$

Lastly, we establish condition (2.31). Note that the first visit time  $\tau_{A(\epsilon)}^\eta(x)$  therein coincides with  $R_\epsilon^{\eta|b}(x)$  defined in (4.20) due to our choice of  $A(\epsilon) = (-\epsilon, \epsilon)$ .

LEMMA 4.5 (Verifying condition (2.31)). *Let  $t(\cdot)$  be defined as in (4.8) and*

$$E(\eta, \epsilon, x) \triangleq \left\{ R_\epsilon^{\eta|b}(x) \leq \frac{t(\epsilon/2)}{\eta}; X_j^{\eta|b}(x) \in I_{\epsilon/2} \quad \forall j \leq R_\epsilon^{\eta|b}(x) \right\}.$$

*For each  $\epsilon \in (0, \bar{\epsilon})$  we have  $\lim_{\eta \downarrow 0} \sup_{x \in I_\epsilon^-} \mathbf{P} \left( (E(\eta, \epsilon, x))^c \right) = 0$ .*



PROOF. First, note that  $(E(\eta, \epsilon, x))^c \subseteq \{X_{[0, t(\epsilon/2)]}^{\eta|b}(x) \in E_1^*(\epsilon) \cup E_2^*(\epsilon) \cup E_3^*(\epsilon)\}$  where

$$E_1^*(\epsilon) \triangleq \{\xi \in \mathbb{D}[0, t(\epsilon/2)] : \xi(t) \notin (-\epsilon, \epsilon) \forall t \in [0, t(\epsilon/2)]\},$$

$$E_2^*(\epsilon) \triangleq \{\xi \in \mathbb{D}[0, t(\epsilon/2)] : \exists 0 \leq s \leq t \leq t(\epsilon/2) \text{ s.t. } \xi(t) \in (-\epsilon, \epsilon), \xi(s) \notin I_{\epsilon/2}\}.$$

Recall the definition of  $\mathbb{D}_A^{(k)|b}[0, T]$  in (2.19). We claim that both  $E_1^*(\epsilon)$  and  $E_2^*(\epsilon)$  are bounded away from

$$\mathbb{D}_{I_\epsilon^-}^{(0)|b}[0, t(\epsilon/2)] = \left\{ \{y_t(x) : t \in [0, t(\epsilon/2)]\} : x \in I_\epsilon^- \right\}.$$

To see why, note that from Assumption 5 and property (4.9), we get  $y_{t(\epsilon/2)}(x) \in [-\frac{\epsilon}{2}, \frac{\epsilon}{2}]$  and  $y_t(x) \in I_\epsilon$ ,  $|y_t(x)| \leq |x|$  for all  $t$  and  $x$  such that  $t \in [0, t(\epsilon/2)]$  and  $x \in I_\epsilon^-$ . Therefore,

$$(4.23) \quad d_{J_1}^{[0, t(\epsilon/2)]}(\mathbb{D}_{I_\epsilon^-}^{(0)|b}[0, t(\epsilon/2)], E_1^*(\epsilon)) \geq \frac{\epsilon}{2} > 0,$$

$$(4.24) \quad d_{J_1}^{[0, t(\epsilon/2)]}(\mathbb{D}_{I_\epsilon^-}^{(0)|b}[0, t(\epsilon/2)], E_2^*(\epsilon)) \geq \frac{\epsilon}{2} > 0.$$

This allows us to apply Theorem 2.4 and obtain

$$\sup_{x \in I_\epsilon^-} \mathbf{P}\left((E(\eta, \epsilon, x))^c\right) \leq \sup_{x \in I_\epsilon^-} \mathbf{P}(X_{[0, t(\epsilon/2)]}^{\eta|b}(x) \in E_1^*(\epsilon) \cup E_2^*(\epsilon)) = \mathcal{O}(\lambda(\eta)) \quad \text{as } \eta \downarrow 0.$$

To conclude the proof, note that  $\lambda(\eta) \in \mathcal{RV}_{\alpha-1}(\eta)$  (with  $\alpha > 1$ ) so  $\lim_{\eta \downarrow 0} \lambda(\eta) = 0$ .  $\square$

We provide some straightforward bounds for the law of geometric random variables.

LEMMA 4.6. *Let  $a : (0, \infty) \rightarrow (0, \infty)$ ,  $b : (0, \infty) \rightarrow (0, \infty)$  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  $k \in \mathbb{N}$ . For any  $c > 1$ , there exists  $\epsilon_0 > 0$  such that*

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

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

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

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

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 D.3 of the Supplementary Material 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)  $\eta a(X_{j-1}^{\eta|b}(x)) + \eta \sigma(X_{j-1}^{\eta|b}(x))Z_j$  of  $X_j^{\eta|b}$  is less than  $b$  for each  $j \leq t$ . Therefore,  $X_j^{\eta|b}(x)$  and  $X_j^\eta(x)$  coincide for such  $j$ 's. In other words, for any  $\eta \in (0, \frac{b}{2C})$ , on event  $\{\eta|Z_j| \leq \frac{b}{2C} \ \forall j \leq t\}$  we have

$$(4.25) \quad X_j^{\eta|b}(x) = X_j^\eta(x) \ \forall j \leq t.$$

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

$$(4.26) \quad \lim_{b \rightarrow \infty} \check{\mathbf{C}}^{(1)|b}(A) = \check{\mathbf{C}}(A).$$

This claim follows from a simple application of the dominated convergence theorem. Indeed, by definition of  $\check{\mathbf{C}}^{(1)|b}$ , we get  $\check{\mathbf{C}}^{(1)|b}(A) = \int \mathbb{I}\{\varphi_b(\sigma(0) \cdot w) \in A\} \nu_\alpha(dw)$ . For  $f_b(w) \triangleq \mathbb{I}\{\varphi_b(\sigma(0) \cdot w)\}$ , we first note that given  $w \in \mathbb{R}$ , we have  $f_b(w) = f(w) \triangleq \mathbb{I}\{\sigma(0) \cdot w\}$  for all  $b > |w| \cdot \sigma(0)$ . Therefore,  $\lim_{b \rightarrow \infty} f_b(w) = f(w)$  holds for all  $w \in \mathbb{R}$ . Next, due to  $r_A > 0$ , we have  $f_b(w) \leq \mathbb{I}\{|w| \geq r_A/\sigma(0)\}$  for all  $b > 0$  and  $w \in \mathbb{R}$ . Meanwhile, note that  $\int \mathbb{I}\{|w| \geq r_A/\sigma(0)\} \nu_\alpha(dw) = (\sigma(0)/r_A)^\alpha < \infty$ . This allows us to apply dominated convergence theorem and establish (4.26). Similarly, for all  $b > |s_{\text{left}}| \vee s_{\text{right}}$ , we have

$$(4.27) \quad C_b^* = \check{\mathbf{C}}^{(1)|b}(I^c) = \int \mathbb{I}\{\sigma(0) \cdot w \in I^c\} \nu_\alpha(dw) = \check{\mathbf{C}}(I^c) \triangleq C^*.$$

To see why, it suffices to notice that for such  $b$ ,

$$\varphi_b(\sigma(0) \cdot w) \notin I \quad \Longleftrightarrow \quad \sigma(0) \cdot w \notin I.$$

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  $A(\eta, x) = \{C^* H(\eta^{-1}) \tau^\eta(x) > t, X_{\tau^\eta(x)}^\eta(x) \in B\}$ , we arbitrarily pick some  $T > t$  and observe that

$$(4.28) \quad \begin{aligned} & A(\eta, x) \\ &= \underbrace{\left\{ C^* H(\eta^{-1}) \tau^\eta(x) \in (t, T], X_{\tau^\eta(x)}^\eta(x) \in B \right\}}_{\triangleq A_1(\eta, x, T)} \cup \underbrace{\left\{ C^* H(\eta^{-1}) \tau^\eta(x) > T, X_{\tau^\eta(x)}^\eta(x) \in B \right\}}_{\triangleq A_2(\eta, x, T)}. \end{aligned}$$

Let  $E_b(\eta, T) \triangleq \{\eta|Z_j| \leq \frac{b}{2C} \ \forall j \leq \frac{T}{C^* H(\eta^{-1})}\}$ . To analyze the probability of  $A_1(\eta, x, T)$ , we further decompose the event as  $A_1(\eta, x, T) = (A_1(\eta, x, T) \cap E_b(\eta, T)) \cup (A_1(\eta, x, T) \setminus E_b(\eta, T))$ . First, for all  $\eta \in (0, \frac{b}{2C})$ ,

$$\begin{aligned} & \mathbf{P}\left(A_1(\eta, x, T) \cap E_b(\eta, T)\right) \\ &= \mathbf{P}\left(\left\{ C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) \in (t, T], X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B \right\} \cap E_b(\eta, T)\right) \quad \text{due to (4.25) and (4.27)} \\ &\leq \mathbf{P}\left(C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) \in (t, T], X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right) \\ &= \mathbf{P}\left(C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) > t, X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right) \end{aligned}$$

$$- \mathbf{P} \left( C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) > T, X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B \right).$$

Using part (a) of Theorem 2.6 and observation (4.27), we get

(4.29)

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

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

$$\begin{aligned} & \limsup_{\eta \downarrow 0} \mathbf{P} \left( \eta|Z_j| > \frac{b}{2C} \text{ for some } j \leq \frac{T}{C^*H(\eta^{-1})} \right) \\ &= 1 - \liminf_{\eta \downarrow 0} \mathbf{P} \left( \text{Geom} \left( H \left( \frac{b}{\eta \cdot 2C} \right) \right) > \frac{T}{C^*H(\eta^{-1})} \right) \\ &\leq 1 - \lim_{\eta \downarrow 0} \exp \left( - \frac{T \cdot H(\eta^{-1} \cdot \frac{b}{2C})}{C^*H(\eta^{-1})} \right) \\ (4.30) \quad &= 1 - \exp \left( - \frac{T}{C^*} \cdot \left( \frac{2C}{b} \right)^\alpha \right). \end{aligned}$$

Similarly,

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

On  $\{C^*H(\eta^{-1})\tau^\eta(x) > T\} \cap E_b(\eta, T)$ , due to (4.25) we have  $\tau^\eta(x) = \tau^{\eta|b}(x)$ . Also, from (4.27) we get  $C^* = C_b^*$ . Using part (a) of Theorem 2.6 again, we get

$$\begin{aligned} & \limsup_{\eta \downarrow 0} \mathbf{P} \left( \left\{ C^*H(\eta^{-1})\tau^\eta(x) > T \right\} \cap E_b(\eta, T) \right) \\ (4.31) \quad & \leq \limsup_{\eta \downarrow 0} \mathbf{P} \left( C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) > T \right) \leq \exp(-T). \end{aligned}$$

Meanwhile, the limit of  $\sup_{x \in I_\epsilon} \mathbf{P} \left( \{C^*H(\eta^{-1})\tau^\eta(x) > T\} \cap E_b(\eta, T) \right)$  as  $\eta \downarrow 0$  is again bounded by (4.30). Collecting (4.29), (4.30), and (4.31), we have shown that for all  $b > 0$  large enough and all  $T > t$ ,

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

In light of claim (4.26), we can drive  $b \rightarrow \infty$  and obtain  $\limsup_{\eta \downarrow 0} \sup_{x \in I_\epsilon} \mathbf{P} \left( A(\eta, x) \right) \leq \frac{\check{\mathbf{C}}(B^-)}{C^*} \cdot \exp(-t) - \frac{\check{\mathbf{C}}(B^\circ)}{C^*} \cdot \exp(-T) + \exp(-T)$ . 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 (4.28), we get

$$\begin{aligned}
& \inf_{x \in I_\epsilon} \mathbf{P}(A(\eta, x)) \\
& \geq \inf_{x \in I_\epsilon} \mathbf{P}(A_1(\eta, x, T)) \geq \inf_{x \in I_\epsilon} \mathbf{P}(A_1(\eta, x, T) \cap E_b(\eta, T)) \\
& = \inf_{x \in I_\epsilon} \mathbf{P}\left(\left\{C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) \in (t, T], X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right\} \cap E_b(\eta, T)\right) \\
& \quad \text{due to (4.25), (4.27)} \\
& \geq \inf_{x \in I_\epsilon} \mathbf{P}\left(C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) \in (t, T], X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right) - \mathbf{P}\left((E_b(\eta, T))^c\right) \\
& \geq \inf_{x \in I_\epsilon} \mathbf{P}\left(C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) > t, X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right) \\
& \quad - \sup_{x \in I_\epsilon} \mathbf{P}\left(C_b^* \eta \cdot \lambda(\eta) \tau^{\eta|b}(x) > T, X_{\tau^{\eta|b}(x)}^{\eta|b}(x) \in B\right) - \mathbf{P}\left((E_b(\eta, T))^c\right).
\end{aligned}$$

Using part (a) of Theorem 2.6 and the limit in (4.30), we yield (for all  $b > 0$  large enough and all  $T > t$ )

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

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

## SUPPLEMENTARY MATERIAL

The supplementary material of the paper is collected in a separate file. Below, we provide a short description for each section of the supplementary material.

### A. Results under General Scaling

This section present results analogous to those in Section 2 under a general scaling, i.e., for the iterations  $X_j^\eta(x)$ 's and  $X_j^{\eta|b}(x)$ 's defined in (1.3).

### B. Results for Lévy-Driven Stochastic Differential Equations

This section collects results analogous to those in Section 2 for stochastic differential equations driven by Lévy processes with regularly varying increments.

### C. Properties of Mappings $h_{[0,T]}^{(k)}$ and $h_{[0,T]}^{(k)|b}$

This section collects useful results for mappings  $h_{[0,T]}^{(k)}$  defined in (2.6)–(2.8) and  $h_{[0,T]}^{(k)|b}$  defined in (2.16)–(2.18). In particular, we provide the proof of Lemmas 3.5 and 3.6.

### D. Properties of Measures $\check{\mathbf{C}}^{(k)|b}$

This section collects important properties of the measure  $\check{\mathbf{C}}^{(k)|b}(\cdot)$  defined in (2.25). In particular, the proof of Lemma 4.2 will be provided at the end of this section.

## REFERENCES

- [1] BALDI, P. and CARAMELLINO, L. (2011). General Freidlin–Wentzell Large Deviations and positive diffusions. *Statistics & Probability Letters* **81** 1218–1229. <https://doi.org/10.1016/j.spl.2011.03.020>

- [2] BAZHBA, M., BLANCHET, J., RHEE, C.-H. and ZWART, B. (2020). Sample path large deviations for Lévy processes and random walks with Weibull increments. *The Annals of Applied Probability* **30** 2695 – 2739. <https://doi.org/10.1214/20-AAP1570>
- [3] BAZHBA, M., RHEE, C.-H. and ZWART, B. (2022). Large deviations for stochastic fluid networks with Weibullian tails. *Queueing Systems* **102** 25–52.
- [4] BERNHARD, H., DAS, B. et al. (2020). Heavy-tailed random walks, buffered queues and hidden large deviations. *Bernoulli* **26** 61–92.
- [5] BOROVKOV, A. A. (2011). Large deviation principles for random walks with regularly varying distributions of jumps. *Siberian mathematical journal* **52** 402–410. <https://doi.org/10.1134/S0037446611030037>
- [6] BOROVKOV, A. A. and BOROVKOV, K. A. (2008). *Asymptotic Analysis of Random Walks: Heavy-Tailed Distributions. Encyclopedia of Mathematics and its Applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511721397>
- [7] BOROVKOV, A. A. and MOGUL'SKII, A. A. (2010). On large deviation principles in metric spaces. *Siberian Mathematical Journal* **51** 989–1003. <https://doi.org/10.1007/s11202-010-0098-0>
- [8] BOVIER, A. and DEN HOLLANDER, F. (2016). *Metastability: a potential-theoretic approach* **351**. Springer.
- [9] BUDHIRAJA, A. and DUPUIS, P. (2000). A variational representation for positive functionals of infinite dimensional Brownian motion. *Probability and mathematical statistics-Wroclaw University* **20** 39–61.
- [10] BUDHIRAJA, A., DUPUIS, P. and MAROULAS, V. (2008). Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability* **36** 1390 – 1420. <https://doi.org/10.1214/07-AOP362>
- [11] CATONI, O. and CERF, R. (1997). The exit path of a Markov chain with rare transitions. *ESAIM: Probability and Statistics* **1** 95–144. <https://doi.org/10.1051/ps:1997105>
- [12] CERRAI, S. and RÖCKNER, M. (2004). Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-Lipshitz reaction term. *The Annals of Probability* **32** 1100 – 1139. <https://doi.org/10.1214/aop/1079021473>
- [13] CHECHKIN, A. V., GONCHAR, V. Y., KLAFTER, J. and METZLER, R. (2005). Barrier crossing of a Lévy flight. *Europhysics Letters* **72** 348.
- [14] CHECHKIN, A. V., SLIUSARENKO, O. Y., METZLER, R. and KLAFTER, J. (2007). Barrier crossing driven by Lévy noise: Universality and the role of noise intensity. *Physical Review E* **75** 041101.
- [15] CHEN, B., BLANCHET, J., RHEE, C.-H. and ZWART, B. (2019). Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound Poisson processes. *Mathematics of Operations Research* **44** 919–942.
- [16] CHEN, B., RHEE, C.-H. and ZWART, B. (2023). Sample-path large deviations for a class of heavy-tailed Markov additive processes. *arXiv preprint arXiv:2010.10751*.
- [17] CHIGANSKY, P. and LIPTSER, R. (2007). *The Freidlin-Wentzell LDP with Rapidly Growing Coefficients In Stochastic Differential Equations: Theory And Applications: A Volume in Honor of Professor Boris L Rozovskii* 197–219. World Scientific. [https://doi.org/10.1142/9789812770639\\_0007](https://doi.org/10.1142/9789812770639_0007)
- [18] DENISOV, D., DIEKER, A. B. and SHNEER, V. (2008). Large deviations for random walks under subexponentiality: The big-jump domain. *The Annals of Probability* **36** 1946 – 1991. <https://doi.org/10.1214/07-AOP382>
- [19] DONATI-MARTIN, C. (2008). Large deviations for Wishart processes. *Probability and Mathematical Statistics* **28**.
- [20] DONATI-MARTIN, C., ROUAULT, A., YOR, M. and ZANI, M. (2004). Large deviations for squares of Bessel and Ornstein-Uhlenbeck processes. *Probability Theory and Related Fields* **129** 261–289. <https://doi.org/10.1007/s00440-004-0338-y>
- [21] DUPUIS, P. and ELLIS, R. S. (2011). *A weak convergence approach to the theory of large deviations*. John Wiley & Sons.
- [22] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (2013). *Modelling extremal events: for insurance and finance* **33**. Springer Science & Business Media. <https://doi.org/10.1007/978-3-642-33483-2>
- [23] EYRING, H. (1935). The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics* **3** 107–115. <https://doi.org/10.1063/1.1749604>
- [24] FOSS, S., KORSHUNOV, D., ZACHARY, S. et al. (2011). *An introduction to heavy-tailed and subexponential distributions* **6**. Springer. <https://doi.org/10.1007/978-1-4419-9473-8>
- [25] HULT, H., LINDSKOG, F., MIKOSCH, T. and SAMORODNITSKY, G. (2005). Functional large deviations for multivariate regularly varying random walks. *The Annals of Applied Probability* **15** 2651 – 2680. <https://doi.org/10.1214/105051605000000502>
- [26] IMKELLER, P. and PAVLYUKOVICH, I. (2006). First exit times of SDEs driven by stable Lévy processes. *Stochastic Processes and their Applications* **116** 611–642. <https://doi.org/10.1016/j.spa.2005.11.006>
- [27] IMKELLER, P. and PAVLYUKOVICH, I. (2008). Metastable behaviour of small noise Lévy-Driven diffusions. *ESAIM: PS* **12** 412–437. <https://doi.org/10.1051/ps:2007051>

- [28] IMKELLER, P., PAVLYUKEVICH, I. and STAUCH, M. (2010). First exit times of non-linear dynamical systems in  $\mathbb{R}^d$  perturbed by multifractal Lévy noise. *Journal of Statistical Physics* **141** 94–119.
- [29] IMKELLER, P., PAVLYUKEVICH, I. and WETZEL, T. (2009). First exit times for Lévy-driven diffusions with exponentially light jumps. *The Annals of Probability* **37** 530 – 564. <https://doi.org/10.1214/08-AOP412>
- [30] IMKELLER, P., PAVLYUKEVICH, I. and WETZEL, T. (2010). The hierarchy of exit times of Lévy-driven Langevin equations. *The European Physical Journal Special Topics* **191** 211–222.
- [31] KALLIANPUR, G. and XIONG, J. (1995). Stochastic Differential Equations in Infinite Dimensional Spaces. *Lecture Notes-Monograph Series* **26** iii–342.
- [32] KIFER, Y. (1990). A Discrete-Time Version of the Wentzell-Friedlin Theory. *The Annals of Probability* **18** 1676 – 1692. <https://doi.org/10.1214/aop/1176990641>
- [33] KLEBANER, F. and ZEITOUNI, O. (1994). The exit problem for a class of period doubling systems. *Ann. Appl. Probab* **4** 1188–1205.
- [34] KRAMERS, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7** 284–304.
- [35] KRUSE, R. (2014). *Strong and weak approximation of semilinear stochastic evolution equations*. Springer.
- [36] LINDSKOG, F., RESNICK, S. I., ROY, J. et al. (2014). Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probability Surveys* **11** 270–314.
- [37] M. I. FREIDLIN, A. D. W. (1998). *Random Perturbations of Dynamical Systems*. Springer New York, NY. <https://doi.org/10.1007/978-1-4612-0611-8>
- [38] MOHAN, M. T. (2021). Wentzell–Freidlin Large Deviation Principle for Stochastic Convective Brinkman–Forchheimer Equations. *Journal of Mathematical Fluid Mechanics* **23** 62. <https://doi.org/10.1007/s00021-021-00587-x>
- [39] MORROW, G. J. and SAWYER, S. (1989). Large Deviation Results for a Class of Markov Chains Arising from Population Genetics. *The Annals of Probability* **17** 1124–1146.
- [40] NAGAEV, A. (1969). Limit theorems for large deviations where Cramér’s conditions are violated. *Fiz-Mat. Nauk* **7** 17–22.
- [41] NAGAEV, A. V. (1978). On a property of sums of independent random variables. *Theory of Probability & Its Applications* **22** 326–338.
- [42] NGUYEN, T. H., SIMSEKLI, U., GURBUZBALABAN, M. and RICHARD, G. (2019). First Exit Time Analysis of Stochastic Gradient Descent Under Heavy-Tailed Gradient Noise. In *Advances in Neural Information Processing Systems* (H. WALLACH, H. LAROCHELLE, A. BEYGELZIMER, F. D’ALCHÉ-BUC, E. FOX and R. GARNETT, eds.) **32**. Curran Associates, Inc.
- [43] PENLAND, C. and EWALD, B. D. (2008). On modelling physical systems with stochastic models: diffusion versus Lévy processes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **366** 2455–2474.
- [44] PROTTER, P. E. (2005). *Stochastic integration and differential equations*. Springer.
- [45] RESNICK, S. I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media.
- [46] RHEE, C.-H., BLANCHET, J., ZWART, B. et al. (2019). Sample path large deviations for Lévy processes and random walks with regularly varying increments. *The Annals of Probability* **47** 3551–3605.
- [47] RÖCKNER, M., ZHANG, T. and ZHANG, X. (2010). Large deviations for stochastic tamed 3D Navier-Stokes equations. *Applied Mathematics and Optimization* **61** 267–285. <https://doi.org/10.1007/s00245-009-9089-6>
- [48] SCALAS, E., GORENFLO, R. and MAINARDI, F. (2000). Fractional calculus and continuous-time finance. *Physica A: Statistical Mechanics and its Applications* **284** 376–384.
- [49] SHWARTZ, A. and WEISS, A. (1995). *Large deviations for performance analysis: queues, communication and computing* **5**. CRC Press.
- [50] SOWERS, R. B. (1992). Large Deviations for a Reaction-Diffusion Equation with Non-Gaussian Perturbations. *The Annals of Probability* **20** 504–537.
- [51] VENTSEL’, A. D. and FREIDLIN, M. I. (1970). ON SMALL RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS. *Russian Mathematical Surveys* **25** 1. <https://doi.org/10.1070/RM1970v025n01ABEH001254>
- [52] WANG, X., OH, S. and RHEE, C.-H. (2022). Eliminating Sharp Minima from SGD with Truncated Heavy-tailed Noise. In *International Conference on Learning Representations*.
- [53] WANG, X. and RHEE, C.-H. (2024). Global Dynamics of Heavy-Tailed SGDs in Nonconvex Loss Landscape: Characterization and Control. Working Paper. <https://chrhee.github.io/papers/WangRhee24b.pdf>.