On sample-path moderate deviation principles for random walks

SUMITH REDDY ANUGU* AND GUODONG PANG*

Abstract. In this survey paper, we revisit the sample-path moderate deviation principles (MDPs) for random walks on the real line. The aim is to provide an overview of the existing techniques that can be used to prove these results. We focus on two approaches: (i) MDP for finite dimensional distributions and exponential tightness in a function space with an appropriate topology, and (ii) variational method via Varadhan's entropy formula. In particular, under weak tail conditions on the distribution of the i.i.d. variables, using both approaches, we prove the sample-path MDP in $\mathcal{C}$ under uniform topology and in $\mathcal{D}$ under the Skorohod J_1 topology.

1. Introduction

There has been tremendous development in large deviation theory as a field in probability theory and stochastic processes (see, for example, the monographs [14, 20, 22, 56, 59]). It has been applied broadly to many areas in science and engineering, as well as in economics and finance. The prominent application is closely related to estimating or simulating probabilities of rare events. On the other hand, closely related to large deviation principles (LDPs), moderate deviation principles (MDPs) have received relatively less attention, particularly, their applications, despite many developments in the theory of MDPs. It is quite often regarded as "in between LDP and central limit theorem", and also related to rare events, but "not so rare" events in comparison with those from applying the LDP theory. In this survey paper, we revisit the sample-path MDPs for random walks on the real line, and the goal is to provide an overview of the existing useful methods in the literature to establish sample-path MDPs for random walks. Since random walks are often building blocks for many stochastic systems, such as queueing and stochastic networks, the techniques will serve as fundamental tools for future work on sample-path MDPs for these systems.

For random walks, the sample-path LDPs have been well established for continuous-time interpolated processes in the spaces $(\mathcal{C}, \|\cdot\|_\infty)$ (see, e.g., [20, Theorem 5.1.2]) and $(\mathcal{D}, J_1)$ (see [36, Theorems 2.5 and 2.7]). The LDPs require a scaling for the processes like the functional law of large numbers for random walks (scaling up time and down space by the same order). Namely, letting Z^n be the continuous-time piecewise linear interpolations of $\{S_i \doteq \sum_{j=1}^i \vartheta_j\}_{1\le i\le n}$ such that $Z^n(\frac{i}{n}) = S_i$ where $\{\vartheta_i\}_{i\in\mathbb{N}}$ is a family of real-valued i.i.d. random variables, the classical sample-path LDP result of Mogulksii [20, Theorem 5.1.2] shows that for any set A of absolutely continuous functions on $[0,1]$ starting at zero and for large n,

$$\frac{1}{n}\log\mathbb{P}(n^{-1}Z^n\in A) \simeq -\inf_{\phi\in A}\frac{1}{2}\int_0^1 \Lambda(\dot\phi(t))dt, \tag{1.1}$$

where $\Lambda(\rho) = \sup_{\alpha\in\mathbb{R}}\left\{\alpha\rho - \log\mathbb{E}[\exp(\alpha\vartheta_1)]\right\}$ (assuming $\mathbb{E}[\exp(\alpha\vartheta_1)]<\infty$ for every α).

The sample-path MDPs for random walks have also been established for continuous-time interpolated processes in the spaces $(\mathcal{C}, \|\cdot\|_\infty)$ ([35, Theorem 2]) and $(\mathcal{D}, J_1)$ (see [36, Theorems 2.6 and 2.9]). However, the MDPs require a scaling for the processes that is in between the functional

* Department of Computational Applied Mathematics and Operations Research, George R. Brown School of Engineering, Rice University, Houston, TX 77005

E-mail addresses: anugu.reddy@rice.edu, gdpang@rice.edu.

Key words and phrases. Sample-path moderate deviation principle, random walk, exponential tightness, finite-dimensional MDP, variational method, Varadhan's entropy formula.

law of large numbers and the functional central limit theorem. In particular, the scaling of space is in between the scale of time and its square root together with a centering term, that is, the MDP-scaled process is given by

$$X^n \doteq \frac{Z^n - n\bar{\vartheta}}{a_n\sqrt{n}},$$

where $\bar{\vartheta} = \mathbb{E}[\vartheta_1]$ and $\{a_n\}_{n\in\mathbb{N}}$ is a positive sequence such that $a_n \to \infty$ and $\sqrt{n}/a_n \to \infty$ as $n \to \infty$ (such as $a_n \sim n^\alpha$ for $\alpha \in (0, 1/2)$). The sample-path MDP result ([35, Theorem 2]) shows that for any set A of absolutely continuous functions starting at zero and for large n,

$$\frac{1}{a_n^2}\log \mathbb{P}(X^n \in A) \simeq -\frac{1}{2\sigma^2}\inf_{\phi\in A}\int_0^1 |\dot{\phi}(t)|^2 dt\,, \tag{1.2}$$

where $\sigma^2 = \mathrm{Var}(\vartheta_1)$. It is easily seen that the rate function of the sample-path MDP only requires the first two moments of the i.i.d. variables $\{\vartheta_i\}_{i\in\mathbb{N}}$, while that for the sample-path LDP requires the entire distributional information of the variables $\{\vartheta_i\}_{i\in\mathbb{N}}$ via the moment generating function.

The proof techniques for the sample-path LDPs for stochastic processes can be adapted to prove the sample-path MDP by taking into account the differences in the scaling of the processes. In this survey paper, we give an overview of two techniques that are amenable for many stochastic processes. The first approach is to use the techniques that resemble the weak convergence of stochastic processes in the spaces $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$ (see [10, 26, 44, 60]). This theory has been developed for sample-path LDP and MDP in Puhalskii [45], Arcones [7, 9] and Feng and Kurtz [27]. Of many important results, two of them are (i) the exponential analog of Prokhorov's theorem which states that a family of random variables has a subsequence along which it satisfies an LDP (or MDP) if and only if it is exponentially tight (see Definition 1.3) and (ii) exponential analog of the statement that "tightness of random variables in an appropriate space together with weak convergence of their finite dimensional distribution implies weak convergence of the random variables." This approach requires to establish the MDP for finite-dimensional distributions of the properly scaled processes and exponential tightness in the functional space under an appropriate topology. We provide an overview for the necessary and sufficient criteria to prove exponential tightness and also for some sufficient criteria to prove the sample-path MDP in the spaces $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$ in Appendix A.

To prove the finite-dimensional MDPs (see Lemmas 3.4 and 3.7), we apply Gärtner-Ellis theorem [20, Theorem 2.3.6] by starting with calculations of log-moment generating function for certain transformed mutually independent variables, as done in the case of LDP. However, because of the centering term and the scaling of the MDP-scaled processes, the calculations differ significantly from the analogous calculations in the case of LDP (see, e.g., [20, Lemma 5.1.8]). This step is more or less the same for the interpolated processes both in $\mathcal{C}$ and $\mathcal{D}$. To prove the exponential tightness, we will need to employ exponential maximal inequalities (which is usually more challenging to establish than the usual maximal inequalities). The various criteria for exponential tightness in $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$ use the appropriate modulii of continuity under the different topologies, so the proofs to verify these various criteria differ greatly. To illustrate these criteria, we verify them for the MDP-scaled processes in both spaces (Section 3). Finally, to identify the rate function, one extracts from the rate function of the finite-dimensional MDP to a functional space, for which the cases in the spaces $\mathcal{C}$ and $\mathcal{D}$ differ. For random walks, this is relatively simple, since the rate function only takes finite values in the subspace of absolutely continuous functions. However, in general, the abstraction would require further substantial work with proper methods suitable for stochastic systems of interest.

The second approach is to use Varadhan's entropy formula (see Lemma B.1), which represents the logarithm of exponential functionals over a probability measure μ as a variational formula over the set of probability measures ν (that are absolutely continuous to μ). This variational approach has been used extensively to study LDPs for various stochastic models (see, e.g., [14, 22]). We are unable to identify a reference where this method is used to prove the sample-path MDP for random walks directly. The only relevant reference is [23], where the sample-path MDP for stochastic recursive equations as Markov chains is established, which includes random walks as a special case. We therefore have adapted this variational approach for the sample-path MDP for random walks in $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$. In writing this survey, we follow closely their approach with one key difference which will be highlighted below. In this approach, we establish an equivalent characterization of the sample-path MDP for processes $\{X^n\}$ in $(\mathcal{C}, \|\cdot\|_\infty)$: for an appropriate function F and large n,

$$\frac{1}{a_n^2}\log \mathbb{E}\big[\exp\big(-a_n^2 F(X^n)\big)\big] \simeq \inf_\phi \Big[F(\phi) + \frac{1}{2\sigma^2}\int_0^1 |\dot\phi(t)|^2 dt\Big].$$

We start with writing $X^n = H^n(\vartheta_1, \ldots, \vartheta_n)$ for some appropriate deterministic function H^n, for each n, and then apply the entropy formula under the n–product distribution $\mathbb{P}_\vartheta^n$, of the law $\mathbb{P}_\vartheta$ of ϑ_1. The main technical challenge lies in establishing tightness of the appropriate conditional mean of a certain family of probability measures (in particular, nearly optimal measures with respect to the aforementioned variational formula) from the uniform upper bound on the relative entropy of this family of probability measures (with respect to the law of the process X^n). This is addressed by showing that the upper bound of relative entropy implies that the conditional mean of the aforementioned family of probability measures is uniformly bounded in $\mathcal{L}_1^2$, the set of square integrable functions (with respect Lebesgue measure) on $[0,1]$ equipped with strong topology. This is the content of Lemma 4.1, whose proof differs from that for the analogous result in [23] (see further discussions in Remark 4.1). Then, we use weak* topology of $\mathcal{L}_1^2$ to infer the desired tightness property. This procedure can be also adapted to prove sample-path MDPs for processes in $(\mathcal{D}, J_1)$.

The variational representation in general involves solving a family of control problems, which can be challenging itself; see, for example, the control problem involved for the stochastic recursive equations in [23]. In the context of random walks, it only requires to show that the associated family of nearly optimal controls is compact in an appropriate sense and the limit points are nearly optimal to the associated limiting control problem. This leads to proving tightness of nearly optimal controls, which is a weaker notion than exponential tightness. To establish exponential tightness, it often requires to work with stronger quantitative estimates than those needed to establish tightness in the weak convergence sense. This may lead the reader into thinking that the second approach is easier to work with. However, it may not always be the case because the processes (for which we are interested in proving tightness) are defined under a new measure.

We have established the sample-path MDP results under a relatively weak condition on the tail behavior of the i.i.d. variables together with the scaling parameter (Assumption 2.1) in both approaches. The authors in [24] have proved that the condition is necessary and sufficient for the one-dimensional MDP of random walks to hold (see Theorem 2.2 therein). See further discussions in Remark 2.1, and also the historical development of the theory of MDP for random walks under various conditions on the variables and scaling parameters in the literature review below. In the second approach, when adapting [23] to random walks, the assumption of finite exponential moment generating function of the variables $\{\vartheta_i\}_{n\in\mathbb{N}}$ is also relaxed.

There are other approaches like 1.) generator/semigroup approach which is applicable, in general when the underlying probability measure is a solution to a martingale problem, and 2.) using the notion of idempotent probability. These approaches can be also used to establish the sample-path MDP for the random walks. We refer the reader to [27] and [47], respectively, for extensive exposition on these approaches.

In the following, we give a thorough survey of results on moderate deviations in the context of random walks. We start with results on MDP for one-dimensional marginals of $\{X^n\}_{n\in\mathbb{N}}$. The phrase "moderate deviations" was first introduced in [51, 52]. One of the first works on the moderate deviations of random walks was done by H. Cramér in [16] (see [17] for the english translation).

In this work, under the assumption that the exponential moment of ϑ_1 is finite around the origin (referred to simply as exponential moment condition from hereon), an asymptotic expansion (in x and large n) of the probability $\mathbb{P}(|X^n(1)| > x)$ was obtained for $a_n = \log n$ and $n^{\frac{1}{3}}$. This work was followed in different capacities by various authors in [1, 2, 4, 38, 39, 40, 41, 42, 49, 54]. The case of $a_n = n^\alpha$ for $\alpha \in (0, 1/2)$ is considered in [32, 37]. In [5, 24, 34, 46, 51, 52, 55], moderate deviations of $\{X^n(1)\}_{n\in\mathbb{N}}$ are considered where the condition on exponential moments is relaxed. In [3, 18, 29, 33, 50, 53], the explicit convergence rate of quantity $\mathbb{P}(|X^n(1)| > x)$ for $a_n = \log n$ is obtained. The case of ϑ_1 taking values in more general spaces like topological spaces is studied in [12, 15, 19, 21, 31].

We next discuss the relevant results on sample-path MDP for $\{X^n\}_{n\in\mathbb{N}}$. The earliest of works can be found in [11], where the author estimates the probabilities that the sample-path of X^n lies below a given deterministic path under the exponential moment condition and under the condition that $a_n(\log n)^{-\frac{1}{2}} \to \infty$. In [35], the sample-path MDP was established in its original form under slightly different assumptions on a_n and ϑ_1 given in Remark 2.1(ii), whereas in [30, 36, 43], exponential moment condition is assumed. In [13], the random variable ϑ_1 is assumed to take values in general topological spaces and under an appropriate exponential tightness condition, the corresponding sample-path MDP is shown. In [57], the case of Banach space valued ϑ_1 with finite sub-exponential moments and $a_n = n^\alpha$ with $\alpha \in (0, 1/2)$ is considered. In [7, 8, 9, 28], under much weaker condition than an exponentially decaying tail of ϑ_1, the sample-path MDP is shown.

Moderate deviations in other contexts such as triangular arrays, Markov processes, stationary sequences, empirical measures, and martingales have also been extensively studied. As we restrict ourselves only to the i.i.d. case, we do not get into the vast literature on these topics.

1.1. **Organization of the paper.** The rest of the paper is organized as follows: we introduce the necessary notation and definitions in the end of this section. In Section 2, we introduce the continuous-time interpolations of random walks and state the sample-path MDP results in the spaces $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$. Section 3 contains the proofs using Approach I, and Section 4 contains the proofs using Approach II. In Appendix A, we collect a few supplementary results on the criteria for exponential tightness of stochastic processes in $\mathcal{C}$ and $\mathcal{D}$, and for the sample-path MDP. In Appendix B, we recall the notion of relative entropy and give a few crucial relevant results. In Appendix C, we give the proof for a result (equation (3.11)) that is used to prove exponential equivalence.

1.2. **Notation and definitions.** We use $(\Omega, \mathcal{F}, \mathbb{P})$ to denote the underlying abstract probability space with $\mathbb{E}$ as the associated expectation. The set of nonnegative real numbers (integers) is denoted by $\mathbb{R}_+$ $(\mathbb{Z}_+)$, $\mathbb{N}$ stands for the set of natural numbers, and $\mathbb{1}_A(\cdot)$ denotes the indicator function corresponding to set A. The minimum (maximum) of two real numbers a and b is denoted by $a \wedge b$ $(a \vee b)$, respectively, and $a^\pm \doteq (\pm a) \vee 0$. The closure, interior, and complement of a set $A \subset \mathbb{R}^d$ are denoted by $\bar{A}$, A°, and A^c, respectively. For two probability measures μ_1 and μ_2, if μ_1 is absolutely continuous with respect to μ_2, we write $\mu_1 \ll \mu_2$.

For $T > 0$, the set of continuous functions on $[0, T]$ is denoted $\mathcal{C}_T$ and when it is equipped with uniform topology, we will denote it by $(\mathcal{C}_T, \|\cdot\|_\infty)$. $(\mathcal{D}_T, J_1)$ denotes the set of càdlàg processes equipped with the Skorohod topology and $d_{J_1}(\cdot,\cdot)$ denotes the corresponding metric. We set $x(t-) \doteq \lim_{s\uparrow t} x(t)$, for $x \in \mathcal{D}$. $\mathcal{L}^2_T$ denotes the Hilbert space consisting of square integrable (with respect to the Lebesgue measure) functions on $[0, T]$ equipped with the usual strong topology (with norm denoted by $\|\cdot\|_2$). When equipped with weak topology, we denote the same set by $\mathcal{L}^{2,*}_T$. Also, the set of all absolutely continuous functions ϕ on $[0, T]$ such that $\phi(0) = 0$ is denoted by $\mathcal{AC}_{T,0}$.

We collect some definitions of LDP below which is stated for general metric space. Let $\mathbb{S}$ be a Polish space with $\mathcal{S}$ being its Borel σ–algebra and $d(\cdot,\cdot)$ the corresponding metric.

Definition 1.1. A family of $\mathbb{S}$–valued random variables $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ is said to satisfy a large deviation principle (LDP) with rate $b_n \uparrow \infty$ (as $n \uparrow \infty$) and rate function $I : \mathbb{S} \to [0, \infty]$ if the following hold:

(1) For a Borel set $A \subset \mathbb{S}$,

$$-\inf_{x\in A^\circ} I(x) \le \liminf_{n\to\infty} \frac{1}{b_n} \log \mathbb{P}(\mathcal{X}^n \in A) \le \limsup_{n\to\infty} \frac{1}{b_n} \log \mathbb{P}(\mathcal{X}^n \in A) \le -\inf_{x\in \bar{A}} I(x)\,.$$

(2) For $l > 0$, $\{x \in \mathbb{S} : I(x) \le l\}$ is compact in $\mathbb{S}$.

Definition 1.2. Two families of $\mathbb{S}$–valued random variables $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ and $\{\mathcal{Y}^n\}_{n\in\mathbb{N}}$ are said to be exponentially equivalent with rate b_n, if for every $\delta > 0$,

$$\lim_{n\to\infty} \frac{1}{b_n} \log \mathbb{P}\left(d(\mathcal{X}^n, \mathcal{Y}^n) > \delta\right) = -\infty\,.$$

Definition 1.3. A family of $\mathbb{S}$–valued random variables $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ is said to be exponentially tight with rate b_n, if for every $l > 0$, there exists a compact set $K_l \subset \mathbb{S}$ such that

$$\limsup_{n\to\infty} \frac{1}{b_n} \log \mathbb{P}\left(\mathcal{X}^n \in K_l^c\right) \le -l\,.$$

There is an equivalent way to define LDP which is given by Varadhan's lemma below ([20, Theorem 4.3.1]).

Lemma 1.1. *A family of $\mathbb{S}$–valued random variables $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ satisfies an LDP with rate b_n and rate function I if and only if the following hold:*

(1) For every bounded continuous $F : \mathbb{S} \to \mathbb{R}$,

$$-\lim_{n\to\infty} \frac{1}{b_n} \log \mathbb{E}\Big[\exp\Big(-b_n F(\mathcal{X}^n)\Big)\Big] = \inf_{x\in\mathbb{S}} \Big[F(x) + I(x)\Big]\,.$$

(2) For every $l > 0$, $\{x \in \mathbb{S} : I(x) \le l\}$ is compact in $\mathbb{S}$.

Lemma 1.2. *Suppose $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ and $\{\mathcal{Y}^n\}_{n\in\mathbb{N}}$ are two exponentially equivalent families of $\mathbb{S}$–valued random variables with rate b_n and also suppose that $\{\mathcal{X}^n\}_{n\in\mathbb{N}}$ satisfies an LDP with rate b_n and rate function I. Then, the family $\{\mathcal{Y}^n\}_{n\in\mathbb{N}}$ also satisfies an LDP with rate b_n and rate function I.*

Let $\{a_n\}_{n\in\mathbb{N}}$ be a positive real valued sequence such that

$$a_n \uparrow \infty \quad \text{and} \quad \frac{\sqrt{n}}{a_n} \uparrow \infty \quad \text{as} \quad n \to \infty\,. \tag{1.3}$$

When we consider the LDP or MDP for a stochastic process $\mathcal{X}$ taking values in $\mathcal{C}_T$ or $\mathcal{D}_T$, it requires a proper scaling, that is, $\bar{\mathcal{X}}^n(t) = \frac{1}{n}\mathcal{X}(nt)$ and $\widetilde{\mathcal{X}}^n(t) = \frac{1}{a_n\sqrt{n}}(\mathcal{X}(nt) - n\bar{\mathcal{X}}(t))$ for some $\bar{\mathcal{X}}(t)$ (which is usually the limit of $\bar{\mathcal{X}}^n(t)$) and for a sequence of proper scaling parameters a_n satisfying (1.3). We say that $\bar{\mathcal{X}}^n$ in $\mathcal{C}_T$ or $\mathcal{D}_T$ satisfies an LDP with rate a_n^2 and rate function I as given in Definition 1.1 for $\mathbb{S} = \mathcal{C}_T$ or $\mathcal{D}_T$. Similarly, we say that $\widetilde{\mathcal{X}}^n$ in $\mathcal{C}_T$ or $\mathcal{D}_T$ satisfies an MDP with rate a_n^2 and rate function I as given in Definition 1.1 for $\mathbb{S} = \mathcal{C}_T$ or $\mathcal{D}_T$.

2. Sample-path moderate deviation principle for random walks

Let $\{\vartheta_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. $\mathbb{R}$-valued random variables with $\mathbb{P}_\vartheta$ being the distribution of ϑ_1 with mean $\bar{\vartheta}$ and variance σ^2. Let $\{a_n\}_{n\in\mathbb{N}}$ be a positive real valued sequence. The associated MDP-scaled random walk is given by

$$S_k^n \doteq \frac{1}{a_n\sqrt{n}} \sum_{i=1}^{k} (\vartheta_i - \bar{\vartheta})$$

for $k = 1, \ldots, n$ and $S_0^n \equiv 0$. We define a continuous-time interpolation and embed the above MDP-scaled random walk in space $\mathcal{C}_1$ and $\mathcal{D}_1$ (which we write as $\mathcal{C}$ and $\mathcal{D}$, respectively from hereon). Let

$$X^n(t) \doteq S^n_{\lfloor nt \rfloor} + \frac{1}{a_n\sqrt{n}}\left(nt - \lfloor nt \rfloor\right)\left(\vartheta_{\lfloor nt \rfloor + 1} - \bar{\vartheta}\right), \quad t \in [0,1],$$

be the continuous-time interpolation in $\mathcal{C}$ and

$$Y^n(t) \doteq S^n_{\lfloor nt \rfloor}, \quad t \in [0,1],$$

be the continuous-time interpolation in $\mathcal{D}$. (We restrict ourselves to $T = 1$, to keep the expressions simple. However, the general $T > 0$ case can be handled in the similar manner.)

We make the following assumption on the scaling parameter $\{a_n\}_{n\in\mathbb{N}}$ and variables $\{\vartheta_i\}_{i\in\mathbb{N}}$.

Assumption 2.1. The sequence $\{a_n\}_{n\in\mathbb{N}}$ in the definition of S_k^n satisfies (1.3) and the following tail condition on ϑ_1 holds

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log\left[n\mathbb{P}\left(|\vartheta_1| \geq a_n\sqrt{n}\right)\right] = -\infty\,. \tag{2.1}$$

Remark 2.1. The condition in (2.1) together with $\{a_n\}_{n\in\mathbb{N}}$ satisfying (1.3) is shown to be the necessary condition for the MDP of i.i.d. variables in [24, Theorem 2.2]. It will be seen in Lemma 3.1 that the condition in (2.1) implies the second moment of ϑ_1 is finite as well as a property on the behavior of its appropriately truncated third absolute moment under the scaling of $\sqrt{n}/a_n$.

In the literature, the MDP for i.i.d. variables has also been established under other conditions, for example, the following are given in [46, Example 7.2]:

(i) $\frac{\log(n)}{a_n^2} \to \infty$ and $\mathbb{E}\left[|\vartheta_1|^{2+\varepsilon}\right] < \infty$, for some $\varepsilon > 0$;

(ii) For some $\beta \in (0,1]$, $\frac{n^{\frac{\beta}{2}}}{a_n^{2-\beta}} \to \infty$ and $\mathbb{E}\left[\exp(b|\vartheta_1|^\beta)\right] < \infty$, for some $b > 0$.

It is clear that if $\{a_n\}_{n\in\mathbb{N}}$ satisfies (1.3), and if either the moment condition in (i) or the finite moment generating function condition in (ii) holds, then the condition (2.1) holds.

We now state the sample-path MDPs for $\{X^n\}_{n\in\mathbb{N}}$ and $\{Y^n\}_{n\in\mathbb{N}}$ in $\mathcal{C}$ and $\mathcal{D}$, respectively. In what follows, we write $\mathcal{AC}_{T,0}$ as $\mathcal{AC}_0$.

Theorem 2.1. *Under Assumption 2.1, the family of $\mathcal{C}$–valued random variables $\{X^n\}_{n\in\mathbb{N}}$ satisfies the sample-path MDP in $(\mathcal{C}, \|\cdot\|_\infty)$ with rate a_n^2 and rate function $I : \mathcal{C} \to [0,\infty]$ given by*

$$I(\phi) \doteq \begin{cases} \frac{1}{2\sigma^2}\int_0^1 |\dot{\phi}(t)|^2 dt\,, & \text{if } \phi \in \mathcal{AC}_0, \\ \infty, & \text{if } \phi \in \mathcal{C} \setminus \mathcal{AC}_0\,. \end{cases} \tag{2.2}$$

Theorem 2.2. *Under Assumption 2.1, the family of $\mathcal{D}$–valued random variables $\{Y^n\}_{n\in\mathbb{N}}$ satisfies the sample-path MDP in $(\mathcal{D}, J_1)$ with rate a_n^2 and rate function $I : \mathcal{D} \to [0,\infty]$ given by*

$$I(\phi) \doteq \begin{cases} \frac{1}{2\sigma^2}\int_0^1 |\dot{\phi}(t)|^2 dt\,, & \text{if } \phi \in \mathcal{AC}_0, \\ \infty, & \text{if } \phi \in \mathcal{D} \setminus \mathcal{AC}_0\,. \end{cases} \tag{2.3}$$

To begin with, we show that the level sets of I are compact subsets of $\mathcal{D}$ under uniform topology (consequently, also under the Skorohod J_1 topology).

Lemma 2.1. *For $l > 0$, $I_l \doteq \{\phi \in \mathcal{D} : I(\phi) \leq l\}$ for $I(\phi)$ in (2.3) is compact in $\mathcal{D}$ under uniform topology.*

Proof. Fix $l > 0$ and consider a sequence $\{\phi_k\}_{k\in\mathbb{N}} \subset I_l$. Clearly, for every k, ϕ_k is absolutely continuous with $\phi_k(0) = 0$ and satisfies the following:

$$\frac{1}{2\sigma^2}\int_0^1 |\dot{\phi}_k(t)|^2 dt \leq l\,. \tag{2.4}$$

Since

$$\phi_k(t) - \phi_k(s) = \int_s^t \dot{\phi}_k(u)du, \qquad \text{for } 0 \le s \le t \le 1,$$

an application of Cauchy-Schwartz inequality gives us

$$|\phi_k(t) - \phi_k(s)| \le 2\sigma^2 \sqrt{l}(t-s)^{\frac{1}{2}}\,.$$

With $\phi_k(0) = 0$, we can infer that $\sup_k \|\phi_k\|_\infty \le 2\sigma^2\sqrt{l}$. Then, Arzela-Ascoli's theorem implies that there exists a subsequence (still denoted by k) and $\phi_* \in \mathcal{C}$ such that $\phi_k \to \phi_*$ in $(\mathcal{C}, \|\cdot\|_\infty)$ as $k \to \infty$. It remains to show that $\phi_* \in I_l$. To do this, we restrict ourselves to the subsequence k obtained earlier and define $h_k(u) \doteq \dot{\phi}_k(u)$, for $u \in [0,1]$. From (2.4), we know that $\{h_k\}_{n\in\mathbb{N}}$ is compact in $\mathcal{L}_1^{2,*}$, i.e., there exists a subsequence (still denoted by k) and $h_* \in \mathcal{L}_1^2$ such that for every $g \in \mathcal{L}_1^2$,

$$\int_0^1 h_k(u)g(u)du \to \int_0^1 h_*(u)g(u)du, \text{ as } k \to \infty\,.$$

In particular, for $t \in [0,1]$ and $g(u) = \mathbb{1}_{[0,t]}(u)$, we have

$$\phi_k(t) = \int_0^t h_k(u)du \to \int_0^t h_*(u)du, \text{ as } k \to \infty\,. \tag{2.5}$$

Moreover, we have

$$\frac{1}{2\sigma^2}\int_0^1 |h_*(u)|^2 du \le l,$$

from the lower semicontinuity of the norm $\|\cdot\|_2^2$. To show that $\phi_* \in I_l$, it suffices to show that $h_* = \dot{\phi}_*$. This is easy to see from (2.5) and the fact that $\phi_k \to \phi_*$ in $(\mathcal{C}, \|\cdot\|_\infty)$, as $k \to \infty$. This completes the proof. □

Remark 2.2. For $n \in \mathbb{N}$, define

$$\bar{\mathcal{X}}^n(t) \doteq \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \vartheta_i + \frac{1}{n}\left(nt - \lfloor nt\rfloor\right)\vartheta_{\lfloor nt\rfloor+1}, \qquad t \in [0,1]\,,$$

and

$$\bar{\mathcal{Y}}^n(t) \doteq \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \vartheta_i\,, \qquad t \in [0,1]\,.$$

Under the assumption that $\mathbb{E}[e^{\rho\vartheta_1}] < \infty$ for any $\rho \in \mathbb{R}$, $\bar{\mathcal{X}}^n(t)$ satisfies a sample-path LDP in $(\mathcal{C}, \|\cdot\|_\infty)$ with rate n and rate function $\bar{I} : \mathcal{C} \to [0,\infty]$ given by (see, e.g.,[20, Theorem 5.1.2])

$$\bar{I}(\phi) \doteq \begin{cases} \frac{1}{2}\int_0^1 \Lambda\big(\dot{\phi}(t)\big)dt, & \text{if } \phi \in \mathcal{AC}_0, \\ \infty, & \text{if } \phi \in \mathcal{C}\setminus\mathcal{AC}_0, \end{cases} \tag{2.6}$$

where $\Lambda(\beta) \doteq \sup_{\rho\in\mathbb{R}}\left\{\beta\rho - \log\mathbb{E}\big[e^{\rho\vartheta_1}\big]\right\}$ is the Fenchel-Legendre transform of $\log\mathbb{E}\big[e^{\rho\vartheta_1}\big]$. Under the same assumption, $\bar{\mathcal{Y}}^n(t)$ satisfies an LDP in $(\mathcal{D}, J_1)$ with rate n and rate function being exactly $\bar{I}$ on $\mathcal{AC}_0$ and taking value ∞ on $\mathcal{D}\setminus\mathcal{AC}_0$ (see [36, Theorems 2.5 and 2.7]). By Lemma 1.1, for every bounded continuous function $F : \mathcal{C} \to \mathbb{R}$,

$$-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\Big[\exp\big(-nF(\bar{\mathcal{X}}^n)\big)\Big] = \inf_{\phi\in\mathcal{C}}\Big[F(\phi) + \bar{I}(\phi)\Big]\,, \tag{2.7}$$

and

$$-\lim_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{E}\Big[\exp\big(-a_n^2 F(X^n)\big)\Big] = \inf_{\phi\in\mathcal{C}}\Big[F(\phi) + I(\phi)\Big]\,. \tag{2.8}$$

An initial observation of (2.7) and (2.8) might suggest that one arrives at (2.8) by replacing ϕ in (2.7) by $n\bar{\vartheta}e + a_n\sqrt{n}\phi$ and evaluating the large n behavior. However, this suggestion is incorrect in general. We refer the reader to [61] for an extensive discussion relevant to this and sufficient conditions under which it is correct. In [58], the author discusses an example for which such a suggestion does not apply.

3. Proofs using finite-dimensional MDP and exponential tightness (Approach I)

3.1. **Proof of the MDP in the space $(\mathcal{C}, \|\cdot\|_\infty)$ (Theorem 2.1).** The proof uses a truncation argument (see [24, Pg. 212]). Let us define a truncation version of X^n which is exponentially equivalent to X^n. To that end, for a fixed $l > 0$, we first define the truncated version of ϑ_i as follows

$$\vartheta_i^n \doteq \vartheta_i \mathbb{1}_{[0, l\sqrt{n}a_n^{-1}]}(|\vartheta_i|)$$

and the associated MDP-scaled random walk is given as

$$\widehat{S}_k^n \doteq \frac{1}{a_n\sqrt{n}} \sum_{i=1}^{k} (\vartheta_i^n - \mathbb{E}[\vartheta_1^n]) \, . \tag{3.1}$$

Define the corresponding continuous time interpolation in $\mathcal{C}$ as

$$\widehat{X}^n(t) \doteq \widehat{S}_{\lfloor nt \rfloor}^n + \frac{1}{a_n\sqrt{n}} (nt - \lfloor nt \rfloor) \left(\vartheta_{\lfloor nt \rfloor + 1}^n - \mathbb{E}[\vartheta_1^n] \right), \quad t \in [0,1]. \tag{3.2}$$

The reason for defining $\widehat{X}^n$ is that the ϑ_i^n's are bounded real valued random variables, for every n. Before we proceed, we give the two important lemmas below.

Lemma 3.1. *Under Assumption 2.1,*

$$\mathbb{E}\left[|\vartheta_1|^2\right] < \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{a_n}{\sqrt{n}} \mathbb{E}\left[|\vartheta_1|^3 \mathbb{1}_{[0, l\sqrt{n}a_n^{-1}]}(|\vartheta_1|)\right] = 0, \quad \text{for every } l > 0 \, . \tag{3.3}$$

Proof. The result follows from [24]. See Lemma 2.5 and discussions on page 212 of that paper. □

Lemma 3.2. *For every $\delta > 0$,*

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\frac{1}{a_n\sqrt{n}} |\vartheta_i^n - \mathbb{E}\left[\vartheta_1^n\right]| > \delta \right) = -\infty \, .$$

Proof. Since ϑ_1^n is bounded, an application of Markov's inequality and Taylor's theorem for $\exp(x)$ gives us that for $\rho > 0$,

$$\frac{1}{a_n^2} \log \mathbb{P}\left(\frac{1}{a_n\sqrt{n}} |\vartheta_i^n - \mathbb{E}\left[\vartheta_1^n\right]| > \delta \right) \le -\rho\delta + \frac{1}{2n} \mathbb{E}\left[|\vartheta_i^n|^2\right] + \frac{a_n}{n\sqrt{n}} \mathbb{E}\left[|\vartheta_1^n|^3\right] .$$

Taking $n \to \infty$ and using Lemma 3.1, the second and third terms on the right hand side above go to zero. Now taking $\rho \uparrow \infty$ gives the result. □

The following lemma shows that the families $\{X^n\}_{n\in\mathbb{N}}$ and $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ are exponentially equivalent.

Lemma 3.3. *For every $\delta > 0$,*

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|X^n - \widehat{X}^n\|_\infty > \delta \right) = -\infty \, . \tag{3.4}$$

Proof. Fix $\delta > 0$ and denote

$$\tilde{\vartheta}_i^n \doteq (\vartheta_i - \bar{\vartheta}) \mathbb{1}_{\{|\vartheta_i - \bar{\vartheta}| \le \frac{\delta}{2} a_n\sqrt{n}\}} .$$

Define

$$\widetilde{S}_k^n \doteq \frac{1}{a_n\sqrt{n}} \sum_{i=1}^{k} \tilde{\vartheta}_i^n \, ,$$

and an intermediate process

$$\breve{X}^n(t) \doteq \widetilde{S}^n_{\lfloor nt \rfloor} + \frac{1}{a_n\sqrt{n}}\,(nt - \lfloor nt \rfloor)\,\tilde{\vartheta}^n_{\lfloor nt \rfloor + 1}\,, \quad t \in [0,1].$$

We show that X^n and $\breve{X}^n$ satisfy: for every $\delta > 0$,

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|X^n - \breve{X}^n\|_\infty > \delta\right) = -\infty\,. \tag{3.5}$$

First, observe from the definition of $\tilde{\vartheta}^n_i$ that

$$\left\{\|X^n - \breve{X}^n\|_\infty > \delta\right\} \subset \left\{\exists t \in [0,1],\, 1 \le i \le \lfloor nt \rfloor : |\vartheta_i - \bar{\vartheta}| > \frac{\delta}{2} a_n \sqrt{n}\right\}.$$

This implies

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|X^n - \breve{X}^n\|_\infty > \delta\right) \\
&\qquad \le \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\exists\, 1 \le i \le n :\, |\vartheta_i - \bar{\vartheta}| > \frac{\delta}{2} a_n \sqrt{n}\right) &(3.6)\\
&\qquad \le \limsup_{n\to\infty} \frac{1}{a_n^2} \log \left((n+1)\,\mathbb{P}\left(|\vartheta_1 - \bar{\vartheta}| > \frac{\delta}{2} a_n \sqrt{n}\right)\right). &(3.7)
\end{aligned}$$

In the above, to get (3.7), we use the union bound. Therefore, using (2.1) and (3.7), we have (3.5).

We next proceed to show that the processes $\breve{X}^n$ and $\widehat{X}^n$ satisfy the following:

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|\breve{X}^n - \widehat{X}^n\|_\infty > \delta\right) = -\infty, \text{ for every } \delta > 0\,. \tag{3.8}$$

Combining (3.5) and (3.8), we have (3.4).

To prove (3.8), define

$$w_i^n \doteq \tilde{\vartheta}^n_i - \vartheta^n_i + \mathbb{E}\big[\vartheta^n_1\big]\,.$$

For $\rho > 0$, we have

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|\breve{X}^n - \widehat{X}^n\|_\infty > \delta\right) \\
&\qquad \le \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\frac{1}{a_n\sqrt{n}} \left\| \sum_{i=1}^{\lfloor nt \rfloor} w_i^n + (nt - \lfloor nt \rfloor)\, w^n_{\lfloor nt \rfloor + 1} \right\|_\infty > \delta\right) \\
&\qquad \le \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\frac{1}{a_n\sqrt{n}} \sum_{i=1}^{n} |w_i^n| > \delta\right) \\
&\qquad \le -\rho\delta + \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{E}\left[\exp\left(\frac{\rho a_n}{\sqrt{n}}\left(\sum_{i=1}^n |w_i^n|\right)\right)\right] &(3.9)\\
&\qquad \le -\rho\delta + \limsup_{n\to\infty} \frac{n}{a_n^2} \log \mathbb{E}\left[\exp\left(\frac{\rho a_n}{\sqrt{n}} |w_1^n|\right)\right]. &(3.10)
\end{aligned}$$

In the above, to get (3.9), we use Markov's inequality; to get (3.10), we use the fact that $\{\vartheta_i\}_{i\in\mathbb{N}}$ is a family of i.i.d. random variables. We claim that for every $\rho > 0$,

$$\lim_{n\to\infty} \frac{n}{a_n^2} \log \mathbb{E}\left[\exp\left(\frac{\rho a_n}{\sqrt{n}} |w_1^n|\right)\right] = 0\,. \tag{3.11}$$

The proof of this claim is deferred to Appendix C. Now taking $\rho \uparrow \infty$ in (3.10) gives us (3.8) and also completes the proof. □

In the following, we establish the MDP for the finite dimensional distributions of $\widehat{X}^n$, that is, we show the MDP of $\{\widehat{X}^n_N\} \doteq \{\widehat{X}^n(t_1), \widehat{X}^n(t_2), \ldots, \widehat{X}^n(t_N)\}$, for every $N \in \mathbb{N}$ and $0 = t_0 < t_1 < t_2 < \ldots < t_N = 1$. We denote $\underline{v}_N \doteq (v_1, v_2, \ldots, v_N) \in \mathbb{R}^N$ and write $\underline{v}$ for brevity whenever there is no confusion.

Lemma 3.4. *Under Assumption 2.1, the family $\{\widehat{X}^n_N\}_{n\in\mathbb{N}}$ of $\mathbb{R}^N$–valued random variables satisfies an MDP with rate a_n^2 and rate function $I^N_{fdd} : \mathbb{R}^N \to [0, \infty]$ given by*

$$I^N_{fdd}(\underline{x}) = \frac{1}{2\sigma^2} \sum_{i=1}^{N} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}} \tag{3.12}$$

with $x_0 = 0$.

Proof. Define

$$Z^n_N = \{Z^{n,i}_N\}_{i=1}^N \doteq (\widehat{X}^n(t_1), \widehat{X}^n(t_2) - \widehat{X}^n(t_1), \ldots, \widehat{X}^n(t_N) - \widehat{X}^n(t_{N-1}))\,. \tag{3.13}$$

By definition, all of the components of Z^n_N are mutually independent. Observe that we can obtain $\widehat{X}^n_N$ from Z^n_N in the following way:

$$\widehat{X}^n_N = \left(Z^{n,1}_N, Z^{n,1}_N + Z^{n,2}_N, \ldots, \sum_{i=1}^{N} Z^{n,i}_N \right).$$

Additionally, the mapping from $\widehat{X}^n_N$ to Z^n_N is a continuous bijection. Since for every n, the process $\widehat{X}^n$ is uniformly bounded, we can infer that the $\mathbb{R}^N$–valued random variable Z^n_N is also uniformly bounded for every n. Therefore, using the fact that $(Z^{n,i}_N, Z^{n,j}_N)$ are mutually independent for $i \neq j$, we get

$$\frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(a_n^2 \sum_{i=1}^{N} \rho_i Z^{n,i}_N \Big) \Big] = \frac{1}{a_n^2} \sum_{i=1}^{N} \log \mathbb{E}\Big[\exp\Big(a_n^2 \rho_i Z^{n,i}_N \Big) \Big]\,. \tag{3.14}$$

Since

$$Z^{n,i}_N = \frac{1}{a_n\sqrt{n}} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i\rfloor} \left(\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \right),$$

using the fact that ϑ^n_i's are i.i.d, we have

$$\log \mathbb{E}\Big[\exp\Big(a_n^2 \rho_i Z^{n,i}_N \Big) \Big] = \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i\rfloor} \log \mathbb{E}\Big[\exp\Big(\frac{a_n}{\sqrt{n}} \rho_i \left(\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \right) \Big) \Big]\,. \tag{3.15}$$

Combining (3.14) and (3.15), we have

$$\begin{aligned}
&\frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(a_n^2 \sum_{i=1}^{N} \rho_i Z^{n,i}_N \Big) \Big] \\
&= \frac{1}{a_n^2} \sum_{i=1}^{N} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i\rfloor} \log \mathbb{E}\Big[\exp\Big(\frac{a_n}{\sqrt{n}} \rho_i \left(\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \right) \Big) \Big] \\
&= \frac{1}{a_n^2} \sum_{i=1}^{N} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i\rfloor} \Bigg(\frac{a_n}{\sqrt{n}} \rho_i \mathbb{E}\Big[\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \Big] + \frac{a_n^2 \rho_i^2}{2n} \mathbb{E}\Big[\left(\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \right)^2 \Big] \\
&\qquad + \mathcal{O}\Bigg(\frac{a_n^3}{n\sqrt{n}} \mathbb{E}\Big[\left(\vartheta^n_j - \mathbb{E}[\vartheta^n_1] \right)^3 \Big] \Bigg) \Bigg)
\end{aligned} \tag{3.16}$$

$$= \sum_{i=1}^{N} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i \rfloor} \left(\frac{\rho_i^2}{2n} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^2 \Big] + \mathcal{O}\left(\frac{a_n}{n\sqrt{n}} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^3 \Big] \right) \right). \tag{3.17}$$

In the above, to get (3.16), we apply Taylor's theorem to $\exp(x)$; to get (3.17), we use the fact that $\mathbb{E}\big[\vartheta_j^n - \mathbb{E}[\vartheta_1^n]\big] = 0$. Using Lemma 3.1 and dominated convergence theorem (notice that $\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \to \vartheta_i - \bar{\vartheta}$, $\mathbb{P}$–a.s. as $n \to \infty$), we can conclude that

$$\lim_{n\to\infty} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i \rfloor} \frac{\rho_i^2}{2n} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^2 \Big] = \frac{\rho_i^2 \sigma^2 (t_i - t_{i-1})}{2}.$$

Now observe that

$$\sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_{i-1} \rfloor} \frac{a_n}{n\sqrt{n}} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^3 \Big] \le \frac{a_n n}{n\sqrt{n}} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^3 \Big].$$

Using Lemma 3.1 again, we have

$$\lim_{n\to\infty} \sum_{j=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i \rfloor} \frac{a_n}{n\sqrt{n}} \mathbb{E}\Big[\big(\vartheta_j^n - \mathbb{E}[\vartheta_1^n] \big)^3 \Big] = 0.$$

Therefore, combining the above analysis, we have shown that

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(a_n^2 \sum_{i=1}^{N} \rho_i Z_N^{n,i} \Big) \Big] = \sum_{i=1}^{N} \frac{\rho_i^2 \sigma^2 (t_i - t_{i-1})}{2}. \tag{3.18}$$

Now using Gärtner-Ellis theorem [20, Theorem 2.3.6], we can conclude that $\{Z_N^n\}_{n\in\mathbb{N}}$ satisfies an MDP with rate a_n^2 and rate function $I_Z^N : \mathbb{R}^N \to [0, \infty]$ given by

$$\begin{aligned} I_Z^N(\underline{z}) &\doteq \sup_{\underline{\rho}\in\mathbb{R}^N} \Big(\sum_{i=1}^{N} \rho_i z_i - \sum_{i=1}^{N} \frac{\rho_i^2 \sigma^2 (t_i - t_{i-1})}{2} \Big) \\ &= \frac{1}{2\sigma^2} \sum_{i=1}^{N} (t_i - t_{i-1}) \Big(\frac{z_i}{t_i - t_{i-1}} \Big)^2. \end{aligned}$$

To get the last equation, we use the fact that

$$\frac{\partial}{\partial \rho_i} \Big(\sum_{i=1}^{N} \rho_i z_i - \sum_{i=1}^{N} \frac{\rho_i^2 \sigma^2 (t_i - t_{i-1})}{2} \Big) = 0 \implies \rho_i = \frac{z_i}{\sigma^2 (t_i - t_{i-1})}, \text{ for } 1 \le i \le N$$

and evaluate the expression inside the supremum at

$$\underline{\rho} \doteq \Big(\frac{z_1}{\sigma^2 (t_1 - t_0)}, \frac{z_2}{\sigma^2 (t_2 - t_1)}, \dots, \frac{z_N}{\sigma^2 (t_N - t_{N-1})} \Big).$$

From here, we obtain the MDP of $\{\widehat{X}_N^n\}_{n\in\mathbb{N}}$ in the following way. Recall that $G : \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$G(z) = \Big(z_1, z_2 + z_1, \dots, \sum_{i=1}^{N} z_i \Big)$$

is a continuous bijection and $G(Z_N^n) = \widehat{X}_N^n$. Applying the contraction principle [20, Theorem 4.2.1], we can conclude that $\{\widehat{X}_N^n\}_{n\in\mathbb{N}}$ satisfies an MDP with rate a_n^2 and rate function $I_{fdd}^N : \mathbb{R}^N \to [0, \infty]$

given by

$$I^N_{fdd}(\underline{x}) = \inf_{\{z\in\mathbb{R}^N : z=G(\underline{x})\}} I^N_Z(z) = \inf_{\{z\in\mathbb{R}^N : z=G(\underline{x})\}} \frac{1}{2\sigma^2}\sum_{i=1}^N \frac{z_i^2}{t_i - t_{i-1}}$$
$$= \frac{1}{2\sigma^2}\sum_{i=1}^N \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}$$

with $x_0 = 0$. This concludes the result. □

We now establish the exponential tightness of $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ in $(\mathcal{C}, \|\cdot\|_\infty)$.

Lemma 3.5. *The family $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ is exponentially tight with rate a_n^2 in $(\mathcal{C}, \|\cdot\|_\infty)$.*

Proof. Firstly, observe that $\widehat{X}^n(0) = 0$, using Theorem A.1, it suffices to show that for $\epsilon > 0$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \frac{1}{a_n^2}\log\mathbb{P}\left(w_\delta(\widehat{X}^n) \geq 3\epsilon\right) = -\infty\,. \tag{3.19}$$

From [10, Theorem 7.4], we know that

$$\mathbb{P}\left(w_\delta(\widehat{X}^n) \geq 3\epsilon\right) \leq \sum_{i=1}^k \mathbb{P}\left(\sup_{t_{i-1}\leq t\leq t_i} |\widehat{X}^n(t) - \widehat{X}^n(t_{i-1})| \geq \epsilon\right),$$

if $\min_{1\leq i\leq m}(t_i - t_{i-1}) \geq \delta$ and $t_m = 1$. In what follows, we take $t_i = i\lceil n\delta\rceil n^{-1}$ for $1 \leq i \leq k$, which clearly satisfy $\min_{1\leq i\leq m}(t_i - t_{i-1}) \geq \delta$. It is also clear that with $m_i \doteq nt_i = i\lceil n\delta\rceil$,

$$\mathbb{P}\left(\sup_{t_{i-1}\leq t\leq t_i} |\widehat{X}^n(t) - \widehat{X}^n(t_{i-1})| \geq \epsilon\right) \leq \mathbb{P}\left(\max_{m_{i-1}\leq k\leq m_i} |\widehat{S}^n_k - \widehat{S}^n_{m_{i-1}}| \geq \epsilon\right)$$
$$= \mathbb{P}\left(\max_{k\leq m_i - m_{i-1}} |\widehat{S}^n_k| \geq \epsilon\right)\,.$$

To obtain the last inequality, we use the fact that $\widehat{S}^n_k$ has stationary increments. Observe that $mn^{-1}\lceil n\delta\rceil = t_m = 1$. Consequently, we have $m = m(n) \to \delta^{-1}$ as $n \to \infty$ which in turn implies that $m(n) \leq 2\delta^{-1}$, for large n. Therefore, combining the above two displays, for large n, we have

$$\begin{aligned}\mathbb{P}\left(w_\delta(\widehat{X}^n) \geq 3\epsilon\right) &\leq \sum_{i=1}^m \mathbb{P}\left(\max_{1\leq k\leq m_i - m_{i-1}} |\widehat{S}^n_k| \geq \epsilon\right)\\ &\leq 2\delta^{-1}\mathbb{P}\left(\max_{1\leq k\leq\lceil n\delta\rceil} |\widehat{S}^n_k| \geq \epsilon\right)\\ &\leq 8\delta^{-1}\max_{1\leq k\leq\lceil n\delta\rceil}\mathbb{P}\left(|\widehat{S}^n_k| \geq \frac{\epsilon}{4}\right)\,.\end{aligned} \tag{3.20}$$

To obtain the last inequality, we use Etemadi's inequality [25, Theorem 1.1 (i)] which states that

$$\mathbb{P}\left(\max_{1\leq k\leq\lceil n\delta\rceil} |\widehat{S}^n_k| \geq \epsilon\right) \leq 4\max_{1\leq k\leq\lceil n\delta\rceil}\mathbb{P}\left(|\widehat{S}^n_k| \geq \frac{\epsilon}{4}\right)\,. \tag{3.21}$$

We then obtain

$$\begin{aligned}\frac{1}{a_n^2}\log\mathbb{P}\left(w_\delta(\widehat{X}^n) \geq 3\epsilon\right) &\leq \frac{1}{a_n^2}\log(8\delta^{-1}) + \frac{1}{a_n^2}\log\max_{1\leq k\leq\lceil n\delta\rceil}\mathbb{P}\left(|\widehat{S}^n_k| \geq \frac{\epsilon}{4}\right)\\ &\leq \frac{1}{a_n^2}\log(8\delta^{-1}) + \max_{1\leq k\leq\lceil n\delta\rceil}\frac{1}{a_n^2}\log\mathbb{P}\left(|\widehat{S}^n_k| \geq \frac{\epsilon}{4}\right)\,.\end{aligned} \tag{3.22}$$

In the following, we compute the probability component in the second term on the right hand side above for $1 \leq k \leq]n\delta[$:

$$\mathbb{P}\Big(|\widehat{S}_k^n| \geq \frac{\epsilon}{4}\Big) = \mathbb{P}\Big(\widehat{S}_k^n \geq \frac{\epsilon}{4}\Big) + \mathbb{P}\Big(\widehat{S}_k^n \leq -\frac{\epsilon}{4}\Big)$$
$$= \mathbb{P}\Big(\exp\Big(a_n^2\rho\widehat{S}_k^n\Big) \geq \exp\Big(\frac{a_n^2\rho\epsilon}{4}\Big)\Big) + \mathbb{P}\Big(\exp\Big(-a_n^2\rho\widehat{S}_k^n\Big) \geq \exp\Big(\frac{a_n^2\rho\epsilon}{4}\Big)\Big)$$
$$\leq \exp\big(-\frac{a_n^2\rho\epsilon}{4}\big)\Big(\mathbb{E}\Big[\exp\Big(a_n^2\rho\widehat{S}_k^n\Big)\Big] + \mathbb{E}\Big[\exp\Big(-a_n^2\rho\widehat{S}_k^n\Big)\Big]\Big), \tag{3.23}$$

where the second inequality follows from Markov inequality, and $\rho > 0$ is arbitrary. Next, using Taylor's series of $\exp(x)$ and the fact that the summands have zero mean and are pairwise independent random variables, we obtain

$$\log\mathbb{E}\Big[\exp\Big(a_n^2\rho\widehat{S}_k^n\Big)\Big] = \log\mathbb{E}\Big[\exp\Big(\frac{a_n^2\rho}{a_n\sqrt{n}}\sum_{i=1}^k(\vartheta_i^n - \mathbb{E}[\vartheta_1^n])\Big)\Big]$$
$$= \frac{a_n\rho}{\sqrt{n}}\sum_{i=1}^k\mathbb{E}\big[\vartheta_i^n - \mathbb{E}[\vartheta_1^n]\big] + \frac{a_n^2\rho^2}{2n}\sum_{i=1}^k\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big]$$
$$+ C^3|\rho|^3\mathcal{O}\Big(\Big(\frac{a_n}{\sqrt{n}}\Big)^3\sum_{i=1}^k\mathbb{E}\big[(\vartheta_i^n - \mathbb{E}[\vartheta_1^n])^3\big]\Big)$$
$$= \frac{a_n^2\rho^2}{2n}\sum_{i=1}^k\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big] + C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n - \mathbb{E}[\vartheta_1^n]|^3\big]\Big)$$
$$= \frac{a_n^2\rho^2}{2n}\sum_{i=1}^k\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big] + C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big).$$

In the above, the third inequality follows from the fact that $k \leq \lceil n\delta\rceil$ and that ϑ_i^n's are identically distributed. Similarly,

$$\log\mathbb{E}\Big[\exp\Big(-a_n^2\rho\widehat{S}_k^n\Big)\Big] = \frac{a_n^2\rho^2}{2n}\sum_{i=1}^k\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big] + C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big).$$

Using the above expressions, we have

$$\exp\big(-\frac{a_n^2\rho\epsilon}{4}\big)\Big(\mathbb{E}\Big[\exp\Big(a_n^2\rho\widehat{S}_k^n\Big)\Big] + \mathbb{E}\Big[\exp\Big(-a_n^2\rho\widehat{S}_k^n\Big)\Big]\Big)$$
$$\leq 2\exp\Big(\frac{a_n^2\rho^2}{2n}\sum_{i=1}^k\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big] - \frac{a_n^2\rho\epsilon}{4} + C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big)\Big). \tag{3.24}$$

Using Lemma 3.1, and combining (3.22), (3.23) and (3.24), we have

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(w_\delta(\widehat{X}^n) \geq 3\epsilon\Big) \leq \limsup_{n\to\infty}\frac{\rho^2}{2n}\sum_{i=1}^{\lceil n\delta\rceil}\mathbb{E}\big[|\vartheta_i^n - \mathbb{E}[\vartheta_1^n]|^2\big] - \frac{\rho\epsilon}{4}$$
$$\leq \frac{\rho^2\sigma^2\delta}{2} - \frac{\rho\epsilon}{4}.$$

Since $\rho > 0$ is arbitrary, choosing the optimal upper bound with $\rho = \frac{\epsilon}{4\sigma^2\delta}$, we get

$$\limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(w_\delta(\widehat{X}^n) \geq 3\epsilon\right) \leq -\frac{\epsilon^2}{32\sigma^2\delta}\,.$$

Now taking $\delta \downarrow 0$ gives the result. □

Completing the proof of Theorem 2.1. From Lemmas 1.2 and 3.3, it suffices to establish that the sample-path MDP holds for $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ with rate a_n^2 and rate function I as defined in (2.2). To achieve this, we use Corollary A.1 and Lemma 3.5 to conclude that $\{\widehat{X}^n\}_{n\in\mathbb{N}}$ satisfies an MDP in $(\mathcal{C}, \|\cdot\|_\infty)$ with rate a_n^2 and rate function

$$I^C(\phi) \doteq \sup_{0\leq t_1<t_2<\ldots<t_N=1} I^N_{fdd}(\phi(t_1), \phi(t_2), \ldots, \phi(t_N))\,.$$

It only remains to show that I^C is the same as the rate function I defined in (2.2). By (3.12) we have

$$I^C(\phi) = \frac{1}{2\sigma^2} \sup_{0<t_1<t_2<\ldots<t_N=1} \sum_{i=1}^N (t_i - t_{i-1})\Big(\frac{\phi(t_i)-\phi(t_{i-1})}{t_i - t_{i-1}}\Big)^2 \tag{3.25}$$

with $\phi(0) = 0$. Then, observe that

$$\sum_{i=1}^N (t_i - t_{i-1})\Big(\frac{\phi(t_i)-\phi(t_{i-1})}{t_i - t_{i-1}}\Big)^2$$

is the convex combination of

$$\left(\Big(\frac{\phi(t_1)-\phi(t_0)}{t_1-t_0}\Big)^2, \Big(\frac{\phi(t_2)-\phi(t_1)}{t_2-t_1}\Big)^2, \ldots, \Big(\frac{\phi(t_N)-\phi(t_{N-1})}{t_N-t_{N-1}}\Big)^2\right).$$

Therefore, using the convexity of the function $g(x) = |x|^2$, we can conclude that

$$\begin{aligned}
I^C(\phi) &= \sup_{0<t_1<t_2<\ldots<t_N=1} I^N_{fdd}(\phi(t_1), \phi(t_2), \ldots, \phi(t_N)) \\
&\leq \frac{1}{2\sigma^2} \sup_{0<t_1<t_2<\ldots<t_N=1} \Big(\phi(t_N) - \phi(t_0)\Big)^2 \\
&= \frac{1}{2\sigma^2} |\phi(1)|^2 && (3.26)\\
&\leq \frac{1}{2\sigma^2} \int_0^1 |\dot{\phi}(t)|^2 dt = I(\phi)\,. && (3.27)
\end{aligned}$$

To arrive at (3.26), we use the fact that $t_0 = 0$, $t_N = 1$ and $\phi(0) = 0$, and to arrive at (3.27), we apply Jensen's inequality to the function $g(x) = |x|^2$.

We now prove that $I^C(\phi) \geq I(\phi)$. To that end, choose $\phi \in \mathcal{AC}_0$ and let

$$f_k(t) \doteq k \int_{k^{-1}\lfloor kt\rfloor}^{k^{-1}(\lfloor kt\rfloor+1)} \dot{\phi}(s)ds \quad \text{and} \quad f_k(1) \doteq k\int_{1-k^{-1}}^1 \dot{\phi}(s)ds\,.$$

We obtain

$$\begin{aligned}
I^C(\phi) &\geq \frac{1}{2\sigma^2} \liminf_{k\to\infty} \sum_{i=1}^k \frac{1}{k}\Big(k\big(\phi(\frac{i}{k}) - \phi(\frac{i+1}{k})\big)\Big)^2 \\
&= \frac{1}{2\sigma^2} \liminf_{k\to\infty} \int_0^1 |f_k(t)|^2 dt && (3.28)\\
&\geq \frac{1}{2\sigma^2} \int_0^1 |\dot{\phi}(t)|^2 dt && (3.29)
\end{aligned}$$

$$= I(\phi)\,.$$

To arrive at (3.28), we use the definition of f_k and to arrive at (3.29), we use continuity of $|x|^2$. This completes the proof of Theorem 2.1. $\square$

3.2. **Proof of the MDP in the space** $(\mathcal{D}, J_1)$ **(Theorem 2.2).** Just like in the proof of Theorem 2.1, we adopt a truncation argument. Let

$$\widehat{Y}^n(t) \doteq \widehat{S}^n_{\lfloor nt\rfloor}\,, \quad t\in[0,1],$$

where $\widehat{S}^n_k$ is defined in (3.1) using the truncated variables.

In the following, we state the analogous results to Lemmas 3.3 and 3.4.

Lemma 3.6. *For every* $\delta>0$*,*

$$\lim_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(d_{J_1}\big(Y^n,\widehat{Y}^n\big)>\delta\right) = -\infty\,.$$

Proof. Let $e:[0,1]\to[0,1]$ be such that $e(t)=t$ and recall the definition of $d_{J_1}(\cdot,\cdot)$: for $x,y\in\mathcal{D}$,

$$d_{J_1}(x,y) = \inf\Big\{\|e-\lambda\|_\infty \vee \|x-y\circ\lambda\|_\infty\Big\}.$$

Here, λ is a strictly increasing continuous function of $[0,1]$ onto itself and the infimum is over all such λ. From the definition of $d_{J_1}(\cdot,\cdot)$, for sequences $\{x^n\}_{n\in\mathbb{N}}$ and $\{y^n\}_{n\in\mathbb{N}}$ of $\mathcal{D}$, $d_{J_1}(x^n,y^n)\to 0$ as $n\to\infty$, if and only if there exists a sequence $\{\lambda^n\}_{n\in\mathbb{N}}$ defined as above such that $\|e-\lambda^n\|_\infty\to 0$ and $\|x^n-y^n\circ\lambda^n\|_\infty\to 0$, as $n\to\infty$. So, it is clear that

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(d_{J_1}\big(Y^n,\widehat{Y}^n\big)>\delta\Big)\\
&=\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(\|Y^n-\widehat{Y}^n\circ\lambda^n\|_\infty>\delta, \text{for all sequences } \{\lambda^n\}_{n\in\mathbb{N}} \text{ such that } \|e-\lambda^n\|_\infty<\delta,\\
&\qquad\qquad \text{for every } n \text{ and } \|e-\lambda^n\|_\infty \xrightarrow{n\to\infty} 0\Big)\,.
\end{aligned}$$

We estimate the right hand side below. Fix a sequence $\{\lambda^n\}_{n\in\mathbb{N}}$ such that $\|e-\lambda^n\|_\infty<\delta$, for every n and $\|e-\lambda^n\|_\infty\xrightarrow{n\to\infty}0$. We have

$$\begin{aligned}
Y^n(t)-\widehat{Y}^n\circ\lambda^n(t) &= S^n_{\lfloor nt\rfloor}-\widehat{S}^n_{\lfloor n\lambda^n(t)\rfloor}\\
&=\frac{1}{a_n\sqrt{n}}\Bigg(\sum_{i=1}^{\lfloor nt\rfloor}(\vartheta_i-\bar{\vartheta})-\sum_{i=1}^{\lfloor n\lambda^n(t)\rfloor}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Bigg)\\
&=J_1^n(t)+J_2^n(t),
\end{aligned}$$

where

$$J_1^n(t)\doteq \mathbb{1}_{[\lambda^n(t),1]}(t)\frac{1}{a_n\sqrt{n}}\Bigg(\sum_{i=1}^{\lfloor n\lambda^n(t)\rfloor}\big[(\vartheta_i-\bar{\vartheta})-(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\big]+\sum_{i=\lfloor n\lambda^n(t)\rfloor+1}^{\lfloor nt\rfloor}(\vartheta_i-\bar{\vartheta})\Bigg)\,,$$

$$J_2^n(t)\doteq \mathbb{1}_{[0,\lambda^n(t))}(t)\frac{1}{a_n\sqrt{n}}\Bigg(\sum_{i=1}^{\lfloor nt\rfloor}\big[(\vartheta_i-\bar{\vartheta})-(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\big]-\sum_{i=\lfloor nt\rfloor+1}^{\lfloor n\lambda^n(t)\rfloor}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Bigg)\,.$$

To complete the proof, it suffices to show that

$$\lim_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(\|J_1^n\|_\infty>\delta\right)=-\infty$$
$$\lim_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(\|J_2^n\|_\infty>\delta\right)=-\infty\,.$$

The proofs of the above two equations follow closely the arguments in the proof of Lemma 3.3 and hence, are omitted. □

Remark 3.1. The above proof can also be done by employing the following bound on the metric d_{J_1} (see [10, Pg. 124]): for $x, y \in \mathcal{D}$, $d_{J_1}(x,y) \le \|x-y\|_\infty$.

Lemma 3.7. *Suppose* (2.1) *holds and let* $\{\widehat{Y}^n_N\} \doteq \{\widehat{Y}^n(t_1), \widehat{Y}^n(t_2), \ldots, \widehat{Y}^n(t_N)\}$*, for every* $N \in \mathbb{N}$ *and* $0 < t_1 < t_2 < \ldots < t_N = 1$*. Then the family* $\{\widehat{Y}^n_N\}_{n\in\mathbb{N}}$ *of* $\mathbb{R}^N$*–valued random variables satisfies an MDP with rate* a_n^2 *and rate function* I^N_{fdd} *(which is defined in* (3.12)*).*

The proof is omitted since it follows essentially identical arguments as those in the proof of Lemma 3.4.

In the following, we prove that $\{\widehat{Y}^n\}_{n\in\mathbb{N}}$ is exponentially tight in $(\mathcal{D}, J_1)$.

Lemma 3.8. *The family* $\{\widehat{Y}^n\}_{n\in\mathbb{N}}$ *is exponentially tight with rate* a_n^2 *in* $(\mathcal{D}, J_1)$*.*

Proof. We apply Theorem A.2. We first verify condition (A.5). For $R > 0$, we have

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,1]} |\widehat{Y}^n(t)| \ge R\right) &\le \mathbb{P}\left(\max_{1\le k\le n} |\widehat{S}^n_k| \ge R\right)\\ &\le 4\max_{1\le k\le n} \mathbb{P}\left(|\widehat{S}^n_k| \ge \frac{R}{4}\right).\end{aligned}$$

To obtain the last inequality, we use Etemadi's inequality; see (3.21). From here, we get

$$\begin{aligned}\frac{1}{a_n^2}\log\mathbb{P}\left(\sup_{t\in[0,1]} |\widehat{Y}^n(t)| \ge R\right) &\le \frac{\log(4)}{a_n^2} + \frac{1}{a_n^2}\log\max_{1\le k\le n}\mathbb{P}\left(|\widehat{S}^n_k| \ge \frac{R}{4}\right)\\ &\le \frac{\log(4)}{a_n^2} + \max_{1\le k\le n}\frac{1}{a_n^2}\log\mathbb{P}\left(|\widehat{S}^n_k| \ge \frac{R}{4}\right).\end{aligned}$$

Following exactly the computations leading up to (3.24) with ϵ and ρ replaced by R and 1, we get

$$\frac{1}{a_n^2}\log\mathbb{P}\Big(|\widehat{S}^n_k| \ge \frac{R}{4}\Big) \le \frac{\log(2)}{a_n^2} + \frac{1}{2n}\sum_{j=1}^{k}\mathbb{E}\big[|\vartheta^n_i - \mathbb{E}[\vartheta^n_1]|^2\big] - \frac{R}{4} + \mathcal{O}\Big(\frac{a_n}{\sqrt{n}}\mathbb{E}\big[|\vartheta^n_1|^3\big]\Big).$$

Combining the above two displays, taking $n \to \infty$ and using Lemma 3.1, we get

$$\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(\sup_{t\in[0,1]}|\widehat{Y}^n(t)| \ge R\right) \le -\frac{R}{4}.$$

Now taking $R \uparrow \infty$, we verify condition (A.5).

By Theorem A.2, it then suffices to verify either condition (A.6) or condition (A.7). For the sake of illustration, we verify both conditions.

Verifying condition (A.6)**:** Fix $\epsilon > 0$. It suffices to show that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(\inf_{\{t_i\}}\max_{1\le i\le m}\sup_{t_{i-1}\le s\le t<t_i}|\widehat{Y}^n(t)-\widehat{Y}^n(s)| \ge \epsilon\right) = -\infty, \tag{3.30}$$

where the infimum is over all possible sets $\{t_i\}_{i=1}^m$ such that $0 = t_0 < t_1 < t_2 < \ldots < t_m = 1$ and $\min_{1\le i\le m}(t_i - t_{i-1}) > \delta$. For now we choose such a set as follows. We take $t_i = i\lceil n\delta\rceil/n$ for $1 \le i \le m$, which satisfy $\min_{1\le i\le m}(t_i - t_{i-1}) \ge \delta$. It is clear that

$$\left\{\inf_{\{t_i\}}\max_{1\le i\le m}\sup_{t_{i-1}\le s\le t<t_i}|\widehat{Y}^n(t)-\widehat{Y}^n(s)| \ge \epsilon\right\} \subset \left\{\max_{1\le i\le m}\sup_{t_{i-1}\le s\le t<t_i}|\widehat{Y}^n(t)-\widehat{Y}^n(s)| \ge \epsilon\right\}.$$

It is also clear that with $m_i \doteq nt_i = i\lceil n\delta \rceil$,

$$\begin{aligned}
\mathbb{P}\left(\inf_{\{t_i\}} \max_{1\le i\le m} \sup_{t_{i-1}\le s\le t<t_i} |\widehat{Y}^n(t) - \widehat{Y}^n(s)| \ge \epsilon\right) &\le \mathbb{P}\left(\max_{1\le i\le m} \sup_{t_{i-1}\le s\le t<t_i} |\widehat{Y}^n(t) - \widehat{Y}^n(s)| \ge \epsilon\right) \\
&\le \mathbb{P}\left(\max_{1\le i\le m} \max_{m_{i-1}\le k\le l\le m_i} |\widehat{S}^n_k - \widehat{S}^n_l| \ge \epsilon\right) \\
&= \mathbb{P}\left(\max_{1\le i\le m} \max_{1\le k-l\le m_i-m_{i-1}} |\widehat{S}^n_{k-l}| \ge \epsilon\right) \\
&= \mathbb{P}\left(\max_{1\le i\le m} \max_{1\le k\le m_i-m_{i-1}} |\widehat{S}^n_k| \ge \epsilon\right).
\end{aligned}$$

To obtain the last two lines, we use the fact that $\widehat{S}^n_k$ has stationary increments and that we can replace $k-l$ by l. We then obtain

$$\begin{aligned}
\mathbb{P}\left(\max_{1\le i\le m} \max_{k\le m_i-m_{i-1}} |\widehat{S}^n_k| \ge \epsilon\right) &\le \sum_{i=1}^m \mathbb{P}\left(\max_{1\le k\le m_i-m_{i-1}} |\widehat{S}^n_k| \ge \epsilon\right) \\
&\le 2\delta^{-1}\mathbb{P}\left(\max_{1\le k\le \lceil n\delta\rceil} |\widehat{S}^n_k| \ge \epsilon\right).
\end{aligned}$$

We use union bound to get the first inequality above. To get the second inequality, observe that $mn^{-1}\lceil n\delta\rceil = t_m = 1$ which in turn implies that $m = m(n) \to \delta^{-1}$ as $n \to \infty$. Therefore, $m(n) \le 2\delta^{-1}$, for large enough n. From here, we conclude the verification of (A.6) using the same argument as that following (3.20) until the end of the proof of Lemma 3.5.

Verifying condition (A.7)**:** Fix $\delta > 0$. Since $\widehat{Y}^n(0) = 0$, we show that

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(|\widehat{Y}^n(\delta)| \ge \varepsilon\right) = -\infty .$$

For n large enough such that $\lfloor n\delta \rfloor < 1$, we clearly have $\widehat{Y}^n(\delta) = 0$. This immediately implies the above display. Next we prove

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(|\widehat{Y}^n(1-) - \widehat{Y}^n(1-\delta)| \ge \varepsilon\right) = -\infty . \tag{3.31}$$

Observe that for any n, $\widehat{Y}^n(1) = \widehat{Y}^n(1-)$ and

$$\widehat{Y}^n(1) - \widehat{Y}^n(1-\delta) = \frac{1}{a_n\sqrt{n}} \sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1]) .$$

We have

$$\begin{aligned}
&\mathbb{P}\Big(\frac{1}{a_n\sqrt{n}}\Big|\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1])\Big| \ge \epsilon\Big) \\
&= \mathbb{P}\Big(\frac{1}{a_n\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1]) \ge \epsilon\Big) + \mathbb{P}\Big(\frac{1}{a_n\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1]) \le -\epsilon\Big) \\
&= \mathbb{P}\Big(\exp\Big(\frac{a_n^2\rho}{a_n\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1])\Big) \ge \exp(a_n^2\rho\epsilon)\Big) \\
&\quad + \mathbb{P}\Big(\exp\Big(-\frac{a_n^2\rho}{a_n\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n} (\vartheta^n_i - \mathbb{E}[\vartheta^n_1])\Big) \ge \exp(a_n^2\rho\epsilon)\Big)
\end{aligned}$$

$$\leq \exp\big(-a_n^2\rho\epsilon\big)\Bigg(\mathbb{E}\Big[\exp\Big(\frac{a_n\rho}{\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Big)\Big]$$

$$+\,\mathbb{E}\Big[\exp\Big(-\frac{a_n\rho}{\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Big)\Big]\Bigg), \tag{3.32}$$

where the last line follows from Markov inequality, and $\rho>0$ is arbitrary. Next, by Taylor's series and the fact that the summands have zero mean and are pairwise independent random variables, we obtain

$$\begin{aligned}&\log\mathbb{E}\Big[\exp\Big(\frac{a_n\rho}{\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Big)\Big]\\&=\frac{a_n^2\rho^2(n-\lfloor n(1-\delta)\rfloor+1)}{2n}\mathbb{E}\big[|\vartheta_1^n-\mathbb{E}[\vartheta_1^n]|^2\big]+C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big)\\&\leq\frac{a_n^2\rho^2}{2}(\delta+\frac{1}{n})\mathbb{E}\big[|\vartheta_1^n-\mathbb{E}[\vartheta_1^n]|^2\big]+C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big)\,.\end{aligned}$$

Similarly,

$$\begin{aligned}\log\mathbb{E}\Big[\exp\Big(-\frac{a_n\rho}{\sqrt{n}}\sum_{i=\lfloor n(1-\delta)\rfloor+1}^{n}(\vartheta_i^n-\mathbb{E}[\vartheta_1^n])\Big)\Big]&\leq\frac{a_n^2\rho^2}{2}(\delta+\frac{1}{n})\mathbb{E}\big[|\vartheta_1^n-\mathbb{E}[\vartheta_1^n]|^2\big]\\&\quad+C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big)\,.\end{aligned}$$

Using the above expressions, we obtain that the right hand side of (3.32) is bounded by

$$2\exp\Bigg(\frac{a_n^2\rho^2}{2}(\delta+\frac{1}{n})\mathbb{E}\big[|\vartheta_1^n-\mathbb{E}[\vartheta_1^n]|^2\big]-a_n^2\rho\epsilon+C^3|\rho|^3\mathcal{O}\Big(\frac{a_n^3}{\sqrt{n}}\mathbb{E}\big[|\vartheta_1^n|^3\big]\Big)\Bigg)\,. \tag{3.33}$$

Combining (3.32) and (3.33), and using Lemma 3.1, we have

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(|\widehat{Y}^n(1)-\widehat{Y}^n(1-\delta)|\geq\epsilon\Big)&\leq\limsup_{n\to\infty}\frac{\rho^2}{2}(\delta+\frac{1}{n})\mathbb{E}\big[|\vartheta_1^n-\mathbb{E}[\vartheta_1^n]|^2\big]-\rho\epsilon\\&\leq\frac{\rho^2\sigma^2\delta}{2}-\rho\epsilon\,.\end{aligned}$$

Since $\rho>0$ is arbitrary, choosing the optimal upper bound with $\rho=\frac{2\epsilon}{\sigma^2\delta}$ and taking $\delta\downarrow 0$, we get (3.31).

Finally, we show that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(w_\delta''(\widehat{Y}^n)\geq\epsilon\Big)=-\infty\,.$$

From the definition of w_δ'', we have

$$w_\delta''(\widehat{Y}^n)=\sup_{\substack{t_1\leq t\leq t_2\\ t_2-t_1\leq\delta}}\Big\{|\widehat{Y}^n(t)-\widehat{Y}^n(t_1)|\wedge|\widehat{Y}^n(t)-\widehat{Y}^n(t_2)|\Big\}\,.$$

For any $0\leq s\leq t\leq 1$, we observe that

$$\widehat{Y}^n(t)-\widehat{Y}^n(s)=\frac{1}{a_n\sqrt{n}}\sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor}(\vartheta_i^n-\mathbb{E}[\vartheta_i^n])$$

and consider the event

$$\begin{aligned}\{w''_\delta(\widehat{Y}^n) \geq \epsilon\} &= \Big\{ \sup_{\substack{t_1 \leq t \leq t_2 \\ t_2 - t_1 \leq \delta}} \Big\{ |\widehat{Y}^n(t) - \widehat{Y}^n(t_1)| \wedge |\widehat{Y}^n(t) - \widehat{Y}^n(t_2)| \Big\} \geq \epsilon \Big\} \\ &= \Big\{ \exists t_1 \leq t \leq t_2 : t_2 - t_1 \leq \delta \,, |\widehat{Y}^n(t) - \widehat{Y}^n(t_1)| \geq \epsilon \Big\} \cap \\ &\qquad \Big\{ \exists t_1 \leq t \leq t_2 : t_2 - t_1 \leq \delta \,, |\widehat{Y}^n(t) - \widehat{Y}^n(t_2)| \geq \epsilon \Big\} \\ &\subset \Big\{ \sup_{\substack{t_1 \leq t \\ t - t_1 \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_1)| \geq \epsilon \Big\} \cap \Big\{ \sup_{\substack{t \leq t_2 \\ t_2 - t \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_2)| \geq \epsilon \Big\} \,.\end{aligned}$$

From the last expression in the above display and using the fact that $\widehat{Y}^n$ has independent increments, we have

$$\begin{aligned}\frac{1}{a_n^2} \log \mathbb{P}\Big(w''_\delta(\widehat{Y}^n) \geq \epsilon \Big) &\leq \frac{1}{a_n^2} \log \mathbb{P}\Big(\sup_{\substack{t_1 \leq t \\ t - t_1 \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_1)| \geq \epsilon \Big) \\ &\qquad + \frac{1}{a_n^2} \log \mathbb{P}\Big(\sup_{\substack{t \leq t_2 \\ t_2 - t \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_2)| \geq \epsilon \Big) \,.\end{aligned}$$

Both the terms on the right hand side of the above display can be handled using similar arguments as those used in proving (3.19). To see this, observe that $w_\delta(x)$ defined in (A.1) can also be represented as

$$w_\delta(x) = \sup_{\substack{s \leq t \\ t - s \leq \delta}} |x(t) - x(s)| \,.$$

We then obtain

$$\begin{aligned}\lim_{\delta \downarrow 0} \lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}\Big(\sup_{\substack{t_1 \leq t \\ t - t_1 \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_1)| \geq \epsilon \Big) = -\infty \,, \\ \lim_{\delta \downarrow 0} \lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}\Big(\sup_{\substack{t \leq t_2 \\ t_2 - t \leq \delta}} |\widehat{Y}^n(t) - \widehat{Y}^n(t_2)| \geq \epsilon \Big) = -\infty \,.\end{aligned}$$

Combining the above two displays verifies condition (A.7). This proves the lemma. □

Completing the proof of Theorem 2.2. We use Lemma 3.8 and apply Theorem A.3 to conclude that $\{\widehat{Y}^n\}_{n \in \mathbb{N}}$ satisfies an MDP in $(\mathcal{D}, J_1)$ with rate a_n^2 and rate function $I^D : \mathcal{D} \to [0, \infty]$ given by

$$I^D(\phi) \doteq \sup_{\substack{0 \leq t_1 < t_2 < \ldots < t_m \leq 1 \\ t_i \notin \Delta_\phi, \text{ for } 1 \leq i \leq m}} I^m_{fdd}(\phi(t_1), \phi(t_2), \ldots, \phi(t_m)), \tag{3.34}$$

where $\Delta_\phi \doteq \{t \in [0,1] : \phi(t-) \neq \phi(t)\}$ and I^m_{fdd} is the rate function from Lemma 3.7. From the computation in the proof of Theorem 2.1, we know that the right hand side of the above display is equal to

$$\frac{1}{2\sigma^2} \int_0^1 |\dot{\phi}(t)|^2 dt \,.$$

From here, we can conclude that $I^D = I$ with I as defined in (2.3). Finally, using Lemmas 1.2 and 3.6, we conclude the desired sample-path MDP of $\{Y^n\}_{n \in \mathbb{N}}$. This completes the proof of Theorem 2.2. □

Remark 3.2. One can also prove Theorem 2.2 by verifying the conditions of Theorem A.4. In particular, by verifying the condition (ii) in Theorem A.4, we immediately obtain the rate function

$$\mathcal{I}(\phi) \doteq \sup_{0\le t_1<t_2<\ldots<t_m\le 1} I^m_{fdd}(\phi(t_1), \phi(t_2), \ldots, \phi(t_m)),$$

without the factor $t_i \notin \Delta_\phi$ in (3.34). Moreover, $\mathcal{I}(\phi)$ only takes finite values for $\phi \in \mathcal{C}$.

However, in general, for $\mathcal{D}$–valued random variables $\{Z^n\}_{n\in\mathbb{N}}$, the condition (ii) of Theorem A.4 may not be satisfied, if the probability

$$\frac{1}{a_n^2}\log \mathbb{P}(Z^n \in \mathcal{D}\setminus\mathcal{C})$$

does not go to zero sufficiently fast as $n\to\infty$.

4. Proofs using the variational method (Approach II)

4.1. **Proof of Theorem 2.1.** Just like in Approach I, we know that the desired sample-path MDP of $\{X^n\}_{n\in\mathbb{N}}$ follows from the sample-path MDP of $\{\widehat{X}^n\}_{n\in\mathbb{N}}$. From Lemmas 1.1 and 2.1, it is then sufficient to prove the theorem below.

Theorem 4.1. *Under Assumption 2.1,*

$$\lim_{n\to\infty} -\frac{1}{a_n^2}\log \mathbb{E}\Big[\exp\Big(-a_n^2 F(\widehat{X}^n)\Big)\Big] = \inf_{x\in\mathcal{C}}\Big[F(x)+I(x)\Big], \tag{4.1}$$

where I is as defined in (2.2).

Proof. Recall that

$$\widehat{X}^n(t) \doteq \widehat{S}^n_{\lfloor nt\rfloor} + \frac{1}{a_n\sqrt{n}}\left(nt-\lfloor nt\rfloor\right)\left(\vartheta^n_{\lfloor nt\rfloor+1} - \mathbb{E}[\vartheta^n_{\lfloor nt\rfloor+1}]\right), \quad t\in[0,1].$$

From hereon, we set $\bar{\vartheta}^n \doteq \mathbb{E}[\vartheta^n_1]$.

The starting point of the proof is to apply Lemma B.1 to $\widehat{X}^n(t)$. For a fixed n, observe that for $H^n:\mathbb{R}^n\to\mathcal{C}$ given by

$$H^n(x_1,x_2,\ldots,x_n)(t) \doteq \frac{1}{a_n\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\left(x_i-\bar{\vartheta}^n\right) + \frac{1}{a_n\sqrt{n}}\left(nt-\lfloor nt\rfloor\right)\left(x_{\lfloor nt\rfloor+1}-\bar{\vartheta}^n\right),$$

we have $\widehat{X}^n(t) = H^n(\vartheta^n_1,\vartheta^n_2\ldots,\vartheta^n_n)(t)$, for $t\in[0,1]$. It is easy to see that the function H^n is continuous from $\mathbb{R}^n$ to $(\mathcal{C},\|\cdot\|_\infty)$ and in particular, it is a Borel measurable function. Let $\mathbb{P}^n$ be the n–product of $\mathbb{P}^n_\vartheta$ (where $\mathbb{P}^n_\vartheta$ is the distribution of ϑ^n_1) and also let $F:\mathcal{C}\to\mathbb{R}$ be a bounded continuous function. Applying Lemma B.1 to $(\Omega,\mathcal{F}) = \big(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n)\big)$, $\mu\equiv\mathbb{P}^n$ and $f(\cdot)\equiv a_n^2 F\circ H^n(\cdot)$, we immediately get

$$\begin{aligned}
-\frac{1}{a_n^2}\log\mathbb{E}\Big[\exp\Big(-a_n^2F(\widehat{X}^n)\Big)\Big] &= -\frac{1}{a_n^2}\log\mathbb{E}\Big[\exp\Big(-a_n^2F\circ H^n(\vartheta^n_1,\vartheta^n_2,\ldots,\vartheta^n_n)\Big)\Big]\\
&= -\frac{1}{a_n^2}\log\int_{\mathbb{R}^n} e^{-a_n^2F\circ H^n(x)}\mathbb{P}^n(dx)\\
&= \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)}\left[\frac{1}{a_n^2}R(\mathbb{Q}^n\|\mathbb{P}^n) + \int_{\mathbb{R}^n}F\circ H^n(x)\mathbb{Q}^n(dx)\right].
\end{aligned} \tag{4.2}$$

Now the rest of proof involves analyzing

$$\limsup_{n\to\infty}\inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)}\left[\frac{1}{a_n^2}R(\mathbb{Q}^n\|\mathbb{P}^n) + \int_{\mathbb{R}^n}F\circ H^n(x)\mathbb{Q}^n(dx)\right], \tag{4.3}$$

$$\liminf_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \left[\frac{1}{a_n^2} R(\mathbb{Q}^n \| \mathbb{P}^n) + \int_{\mathbb{R}^n} F\circ H^n(x)\mathbb{Q}^n(dx)\right]. \tag{4.4}$$

For the sake of ease of computations, we will deal with $\mathbb{R}^n$–valued random variables instead of measure $\mathbb{Q}^n \in \mathcal{P}(\mathbb{R}^n)$. To that end, for any n and measure $\mathbb{Q}^n \in \mathcal{P}(\mathbb{R}^n)$, let $\varpi^n \doteq (\varpi_1^n, \varpi_2^n, \ldots, \varpi_n^n)$ be a $\mathbb{R}^n$–valued random variable with distribution $\mathbb{Q}^n$. Let $\mathbb{Q}_i^n$ denote the conditional distribution of ϖ_i^n given $(\varpi_1^n, \varpi_2^n, \ldots, \varpi_{i-1}^n)$ with the MDP-scaled conditional mean

$$u_i^n \doteq \int_{\mathbb{R}} x\mathbb{Q}_i^n(dx) - \bar{\vartheta}^n. \tag{4.5}$$

From the construction of random variable with law $\mathbb{Q}^n$, it is a priori not clear if $\{\varpi_1^n, \varpi_2^n, \ldots, \varpi_n^n\}$ is an independent family of n real valued random variables under nearly optimal $\mathbb{Q}^n$ associated with (4.2). Hence, we have defined conditional distribution $\mathbb{Q}^n$ above to handle the possible dependence among $(\varpi_1^n, \varpi_2^n, \ldots, \varpi_n^n)$.

Now define

$$\widetilde{S}_k^n \doteq \frac{1}{a_n\sqrt{n}} \sum_{i=1}^{k} (\varpi_i^n - \bar{\vartheta}^n) \tag{4.6}$$

and its continuous time interpolation

$$\widetilde{X}^n(t) \doteq \widetilde{S}_{\lfloor nt\rfloor}^n + \frac{1}{a_n\sqrt{n}} (nt - \lfloor nt\rfloor) \left(\varpi_{\lfloor nt\rfloor+1}^n - \bar{\vartheta}^n\right), \quad t\in[0,1].$$

Note that $\widetilde{X}^n$ is defined under $\mathbb{Q}^n$, and that $\widetilde{X}^n$ is not a mean zero process as ϖ_i^n does not necessarily have mean $\bar{\vartheta}^n$.

Observe that

$$\widetilde{X}^n(t) = H^n(\varpi_1^n, \varpi_2^n, \ldots, \varpi_n^n)(t), \text{ for } t\in[0,1]. \tag{4.7}$$

This consequently implies that

$$\int_{\mathbb{R}^n} F\circ H^n(x)\mathbb{Q}^n(dx) = \mathbb{E}\big[F(\widetilde{X}^n)\big].$$

Next, applying Theorem B.1 repeatedly for n times and using the definition of $\mathbb{Q}_i^n$ gives us the following:

$$R(\mathbb{Q}^n\|\mathbb{P}^n) = \sum_{i=1}^{n} R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n).$$

Finally, (4.2) can be re-cast in terms of $\widetilde{X}^n$ as follows:

$$-\frac{1}{a_n^2}\log\mathbb{E}\Big[\exp\Big(-a_n^2 F(\widehat{X}^n)\Big)\Big] = \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^{n} R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big]. \tag{4.8}$$

To prove (4.1), it now suffices to prove

$$\limsup_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^{n} R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big] \le \inf_{x\in\mathcal{C}} \Big[F(x)+I(x)\Big], \tag{4.9}$$

and

$$\liminf_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^{n} R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big] \ge \inf_{x\in\mathcal{C}} \Big[F(x)+I(x)\Big]. \tag{4.10}$$

These two inequalities follow from Lemmas 4.3 and 4.4 below. This completes the proof of Theorem 4.1. □

In the following, we state and prove two crucial lemmas that will be used later in the proofs of (4.9) (see Lemma 4.3) and (4.10) (see Lemma 4.4). From the form of (4.9) and (4.10), it is clear that understanding the behavior of nearly optimal $\mathbb{Q}^n$ associated with the infima in (4.8) and (4.10) for large n is necessary. To that end, we have the following immediate upper bound on the relative entropy of nearly optimal $\mathbb{Q}^n$: fix $\delta > 0$ and let $\widetilde{\mathbb{Q}}^n$ (and the associated $\{\widetilde{\mathbb{Q}}^n_i\}_{1\leq i\leq n}$) be such that

$$-\frac{1}{a_n^2}\log \mathbb{E}\Big[\exp\Big(-a_n^2 F(\widehat{X}^n)\Big)\Big] = \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\mathbb{Q}^n_i\|\mathbb{P}^n_\vartheta) + F(\widetilde{X}^n)\Big]$$
$$\geq \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}^n_i\|\mathbb{P}^n_\vartheta) + F(\widetilde{X}^n)\Big] - \delta\,. \tag{4.11}$$

Since F is bounded with $\bar{F} \doteq \sup_{x\in\mathcal{C}}|F(x)| < \infty$, this immediately implies that

$$\sup_n \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}^n_i\|\mathbb{P}^n_\vartheta)\Big] \leq 2\bar{F} + \delta\,. \tag{4.12}$$

This bound is crucial in proving that $\{\widetilde{X}^n\}_{n\in\mathbb{N}}$ (associated with $\widetilde{Q}^n$ chosen as above) is tight.

Recall that the centered log-moment generating function $H^c_\mu : \mathbb{R}\to\mathbb{R}$ for a probability measure μ on $\mathbb{R}$ with mean m_μ is defined as

$$H^c_\mu(\alpha) = \sup_{\beta\in\mathbb{R}}\Big\{\alpha\beta - \log\int_{\mathbb{R}} e^{\beta(x-m_\mu)}\mu(dx)\Big\}\,. \tag{4.13}$$

Lemma 4.1. *Suppose $\widetilde{\mathbb{Q}}^n$ satisfies (4.11). Then, the following inequality holds:*

$$\liminf_{n\to\infty}\mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}^n_i\|\mathbb{P}^n_\vartheta)\Big] \geq \liminf_{n\to\infty}\frac{1}{2\sigma^2 n}\mathbb{E}\Big[\sum_{i=1}^n |\tilde{u}^n_i|^2\Big], \tag{4.14}$$

where $\tilde{u}^n_i = a_n^{-1}\sqrt{n}u^n_i$. Recall that $u^n_i = \int_{\mathbb{R}} x\widetilde{\mathbb{Q}}^n_i(dx) - \bar{\vartheta}^n$.

Proof. To begin with, we note that for every n, $\mathbb{P}^n_\vartheta$ has a compact support and hence, using Lemma B.2 gives us

$$\mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}^n_i\|\mathbb{P}^n_\vartheta)\Big] \geq \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n H^c_{\mathbb{P}^n_\vartheta}\Big(\int_{\mathbb{R}} y\widetilde{\mathbb{Q}}^n_i(dy) - \bar{\vartheta}^n\Big)\Big]\,. \tag{4.15}$$

In the following, we derive a desired lower bound on the right hand side. For every $\beta_i > 0$ with $i\geq 1$, define $\beta^n_i \doteq \frac{a_n}{\sqrt{n}}\beta_i$. Using the definitions of $H^c_{\mathbb{P}^n_\vartheta}$ and $\tilde{u}^n_i$, we have

$$H^c_{\mathbb{P}^n_\vartheta}\Big(\int_{\mathbb{R}} y\widetilde{\mathbb{Q}}^n_i(dy) - \bar{\vartheta}^n\Big) \geq \beta^n_i\frac{a_n}{\sqrt{n}}\tilde{u}^n_i - \log\int_{\mathbb{R}} e^{\beta^n_i(y-\bar{\vartheta}^n)}\mathbb{P}^n_\vartheta(dy)$$
$$\geq \beta_i\frac{a_n^2}{n}\tilde{u}^n_i - \log\int_{\mathbb{R}} e^{\frac{\beta_i a_n}{\sqrt{n}}(y-\bar{\vartheta}^n)}\mathbb{P}^n_\vartheta(dy)\,. \tag{4.16}$$

Using the Taylor's series for $\exp(x)$ and the definition of $\bar{\vartheta}^n$, we get

$$\log\int_{\mathbb{R}} e^{\frac{\beta_i a_n}{\sqrt{n}}(y-\bar{\vartheta}^n)}\mathbb{P}^n_\vartheta(dy) = \frac{\beta_i^2 a_n^2}{2n}\int_{\mathbb{R}}(|y|^2 - |\bar{\vartheta}^n|^2)\mathbb{P}^n_\vartheta(dy) + \mathcal{O}\Big(|\beta_i|^3\frac{a_n^3}{n\sqrt{n}}\int_{\mathbb{R}}|y|^3\mathbb{P}^n_\vartheta(dy)\Big)\,. \tag{4.17}$$

Plugging (4.16) and (4.17) in (4.15), we get

$$\mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}^n_i\|\mathbb{P}^n_\vartheta)\Big]$$

$$\geq \frac{1}{n}\sum_{i=1}^{n}\frac{n}{a_n^2}\left(\beta_i\frac{a_n^2}{n}\tilde{u}_i^n - \frac{\beta_i^2 a_n^2}{2n}\int_{\mathbb{R}}(|y|^2-|\bar{\vartheta}^n|^2)\mathbb{P}_{\vartheta}^n(dy) - \mathcal{O}\Big(|\beta_i|^3\frac{a_n^3}{n\sqrt{n}}\int_{\mathbb{R}}|y|^3\mathbb{P}_{\vartheta}^n(dy)\Big)\right)$$

$$\geq \frac{1}{n}\sum_{i=1}^{n}\left(\beta_i\tilde{u}_i^n - \frac{\beta_i^2}{2}\int_{\mathbb{R}}(|y|^2-|\bar{\vartheta}^n|^2)\mathbb{P}_{\vartheta}^n(dy) - \mathcal{O}\Big(|\beta_i|^3\frac{a_n}{\sqrt{n}}\int_{\mathbb{R}}|y|^3\mathbb{P}_{\vartheta}^n(dy)\Big)\right).$$

Let $K > 0$ and replace β_i by $\beta_i \wedge K$. Then the third sum above can be bounded by

$$K^3\mathcal{O}\Big(\frac{a_n}{\sqrt{n}}\int_{\mathbb{R}}|y|^3\mathbb{P}_{\vartheta}^n(dy)\Big) = K^3\mathcal{O}\Big(\frac{a_n}{\sqrt{n}}\mathbb{E}[|\vartheta_1^n|^3]\Big).$$

By Lemma 3.1, we conclude that the above expression goes to zero as $n \to \infty$. Therefore, we have proved that

$$\liminf_{n\to\infty}\ \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^{n}R(\widetilde{\mathbb{Q}}_i^n\|\mathbb{P}_{\vartheta}^n)\Big] \geq \liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Big((\beta_i\wedge K)\tilde{u}_i^n - \frac{(\beta_i\wedge K)^2}{2}\int_{\mathbb{R}}(|y|^2-|\bar{\vartheta}^n|^2)\mathbb{P}_{\vartheta}^n(dy)\Big)$$

$$\geq \liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Big((\beta_i\wedge K)\tilde{u}_i^n - \frac{(\beta_i\wedge K)^2\sigma^2}{2}\Big).$$

In the above, we have used dominated convergence theorem on each of the terms in the second sum. Choose $\beta_i = \tilde{u}_i^n/\sigma^2$. This gives us

$$\liminf_{n\to\infty}\ \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^{n}R(\widetilde{\mathbb{Q}}_i^n\|\mathbb{P}_{\vartheta}^n)\Big] \geq \liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Big((\beta_i\wedge K)\tilde{u}_i^n\wedge K - \frac{(\beta_i\wedge K)^2\sigma^2}{2}\Big)$$

$$\geq \liminf_{n\to\infty}\frac{1}{2\sigma^2 n}\sum_{i=1}^{n}\left|\tilde{u}_i^n\wedge(\sigma^2K)\right|^2.$$

Now taking $K \uparrow \infty$ gives us the result. ☐

Lemma 4.2. *Suppose $\widetilde{\mathbb{Q}}^n$ satisfies (4.11). Let U^n be an $\mathcal{L}_1^2$–valued random variable defined by*

$$U^n(t) = a_n^{-1}\sqrt{n}u_i^n, \quad \text{for} \quad t\in[in^{-1},(i+1)n^{-1}).$$

Here, u_i^n is as defined in (4.5) in association with $\widetilde{\mathbb{Q}}^n$. Also, let $\widetilde{U}^n(t) \doteq \int_0^t U^n(s)ds$. Then, we have

$$\widetilde{X}^n - \widetilde{U}^n \Rightarrow 0 \quad \text{in} \quad (\mathcal{C},\|\cdot\|_\infty) \quad \text{as} \quad n\to\infty. \tag{4.18}$$

Proof. Fix $\delta_0 > 0$. Recall (4.12). For $1 \leq i \leq n$, define

$$W_i^n \doteq \frac{1}{a_n\sqrt{n}}\sum_{j=1}^{i}(\varpi_j^n - \bar{\vartheta}^n - u_j^n)$$

and $R_i^n \doteq R(\widetilde{\mathbb{Q}}_i^n\|\mathbb{P}_{\vartheta}^n)$. Note that R_i^n is a non-negative random variable as $\widetilde{\mathbb{Q}}_i^n$ is a conditional distribution of ϖ_i^n given $(\varpi_1^n,\varpi_2^n,\ldots,\varpi_{i-1}^n)$ under $\widetilde{\mathbb{Q}}^n$.

Observe that $\widetilde{X}^n - \widetilde{U}^n$ is the continuous interpolation of W^n. Therefore, showing

$$\lim_{n\to\infty}\mathbb{P}\Big(\max_{1\leq i\leq n}|W_i^n| \geq \delta_0\Big) = 0 \quad \text{for every } \delta_0 > 0, \tag{4.19}$$

implies the desired convergence of $\widetilde{X}^n - \widetilde{U}^n$ to 0. From the definition of ϖ_i^n, $W_\cdot^n$ is a martingale with respect to filtration generated by $\{\varpi_i\}_{i\in\mathbb{N}}$. Therefore, using Doob's submartingale inequality, we obtain

$$\mathbb{P}\Big(\max_{1\leq i\leq n}|W_i^n| \geq \delta_0\Big) \leq \frac{1}{\delta_0^2}\mathbb{E}\Big[\Big|\frac{1}{a_n\sqrt{n}}\sum_{j=1}^{i}(\varpi_j^n - \bar{\vartheta}^n - u_j^n)\Big|^2\Big].$$

From the conditioning argument and the martingale property of W^n, we have

$$\mathbb{P}\Big(\max_{1\le i\le n}|W_i^n|\ge\delta_0\Big)\le\frac{1}{\delta_0^2a_n^2n}\sum_{i=1}^n\mathbb{E}\Big[\big|(\varpi_i^n-\bar{\vartheta}^n-u_i^n)\big|^2\Big]$$
$$\le\frac{1}{\delta_0^2a_n^2n}\sum_{i=1}^n\mathbb{E}\big[|\varpi_i^n|^2\big]\,.$$

To arrive at the last inequality, we use the fact that $\mathbb{E}\big[\big|Z-\mathbb{E}[Z]\big|^2\big]\le\mathbb{E}[Z^2]$ (for any integrable random variable Z) and the fact that $u_i^n=\int_{\mathbb{R}}y\widetilde{\mathbb{Q}}_i^n(dy)-\bar{\vartheta}^n$. Below we show that the right hand side of the above display goes to 0 as $n\to\infty$. To that end, we recall that from the definition of ϑ_1^n, $\mathbb{P}_\vartheta^n$ has a compact support $[-l\sqrt{n}a_n^{-1},l\sqrt{n}a_n^{-1}]$. We have

$$\frac{1}{a_n^2n}\sum_{i=1}^n\mathbb{E}\Big[\big|\varpi_i^n\big|^2\Big]\le\frac{1}{a_n^2n}\sum_{i=1}^n\int_{\mathbb{R}}|y|^2R_i^n(y)\mathbb{P}_\vartheta^n(dy)$$
$$\le\frac{1}{a_n^2n}(l\sqrt{n}a_n^{-1})^2\sum_{i=1}^n\int_{\mathbb{R}}R_i^n(y)\mathbb{P}_\vartheta^n(dy)$$
$$\le\frac{l^2}{a_n^4}\sum_{i=1}^n\mathbb{E}\big[R_i^n\big]$$
$$\le\frac{(2\bar{F}+\delta)l^2}{a_n^2}\,.$$

To get the second inequality, we use the compactness of the support of $\mathbb{P}_\vartheta^n$; the third inequality simply follows from the definition of R_i^n; the last inequality follows from (4.12). This proves (4.19) and also the desired result. □

We are finally in a position to prove (4.9) and (4.10).

Lemma 4.3. (4.9) *holds.*

Proof. Fix $\delta>0$ and let $x^*\in\mathcal{C}$ be such that

$$F(x^*)+I(x^*)\le\inf_{x\in\mathcal{C}}\Big[F(x)+I(x)\Big]+\delta\,.$$

From the definition of I, it is clear that $x^*\in\mathcal{AC}_0$. Let $U^*(t)\doteq\dot{x}^*(t)$. We now define a probability measure $\mathbb{Q}^{n,*}$ using $U^*(t)$ as follows. $\mathbb{Q}^{n,*}$ is a n-product of measures $\mathbb{Q}_1^{n,*}$ given by

$$\mathbb{Q}_1^{n,*}(dy)\doteq\exp\Big(\frac{yU^*(t)a_n}{\sqrt{n}}-H^c_{\mathbb{P}_\vartheta^n}\Big(\frac{a_nU^*(t)}{\sqrt{n}}\Big)\Big)\mathbb{P}_\vartheta^n(dy)\,.$$

It is clear that

$$\limsup_{n\to\infty}\mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^nR(\mathbb{Q}_i^{n,*}\|\mathbb{P}_\vartheta^n)\Big]=\limsup_{n\to\infty}\frac{1}{2\sigma^2n}\sum_{i=1}^n\Big|U^*\Big(\frac{i}{n}\Big)\Big|^2=\frac{1}{2\sigma^2}\|U^*\|_2^2=I(x^*)\,.$$

From Lemma 4.2, we know that

$$\widetilde{X}^n\Rightarrow\widetilde{U}^*(\cdot)\doteq\int_0^\cdot U^*(s)ds=x^*\quad\text{in}\quad(\mathcal{C},\|\cdot\|_\infty)\tag{4.20}$$

as $n\to\infty$. Therefore, we have

$$\limsup_{n\to\infty}\inf_{\mathbb{Q}^n\in\mathcal{P}(\mathcal{C})}\mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^nR(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n)+F(\widetilde{X}^n)\Big]$$

$$
\begin{aligned}
&\leq \limsup_{n\to\infty} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\mathbb{Q}_i^{n,*}\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big] \\
&= I(x^*) + F(x^*) \\
&\leq \inf_{x\in\mathcal{C}} \Big[I(x) + F(x)\Big] + \delta\,.
\end{aligned}
$$

This proves (4.9). □

Lemma 4.4. (4.10) *holds.*

Proof. For $\delta > 0$ and $n \in \mathbb{N}$, let $\widetilde{\mathbb{Q}}^n$ be such that it satisfies (4.11). Let U^n and $\widetilde{U}^n$ be as in Lemma 4.2. Then, from Lemma 4.1 and (4.12), we can conclude that

$$\sup_{n\in\mathbb{N}} \|U^n\|_2^2 \leq 4\bar{F} + 2\delta\,.$$

This means that $\{U^n\}_{n\in\mathbb{N}}$ is a tight family of $\mathcal{L}_1^{2,*}$–valued random variables and hence, along a subsequence (still denoted by n), U^n converges to some U^* in $\mathcal{L}_1^{2,*}$. Consequently, it is easy to see that

$$\widetilde{U}^n(\cdot) = \int_0^\cdot U^n(s)ds \Rightarrow \widetilde{U}^*(\cdot) \doteq \int_0^\cdot U^*(s)ds \quad \text{in} \quad (\mathcal{C}, \|\cdot\|_\infty)\,.$$

Combining the above display and Lemma 4.2, we can conclude that

$$\widetilde{X}^n \Rightarrow \widetilde{U}^* \quad \text{in} \quad (\mathcal{C}, \|\cdot\|_\infty) \tag{4.21}$$

as $n \to \infty$. Finally, we obtain

$$
\begin{aligned}
&\liminf_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathcal{C})} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big] \\
&\qquad \geq \liminf_{n\to\infty} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\widetilde{\mathbb{Q}}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{X}^n)\Big] - \delta && (4.22) \\
&\qquad \geq \liminf_{n\to\infty} \frac{1}{a_n^2}\mathbb{E}\Big[\sum_{i=1}^n R(\widetilde{\mathbb{Q}}_i^n\|\mathbb{P}_\vartheta^n)\Big] + \liminf_{n\to\infty} \mathbb{E}[F(\widetilde{X}^n)] - \delta \\
&\qquad \geq \liminf_{n\to\infty} \frac{1}{2\sigma^2}\mathbb{E}\Big[\|U^n\|_2^2\Big] + \mathbb{E}[F(\widetilde{U}^*)] - \delta && (4.23) \\
&\qquad \geq \frac{1}{2\sigma^2}\mathbb{E}\Big[\|U^*\|_2^2\Big] + \mathbb{E}[F(\widetilde{U}^*)] - \delta && (4.24) \\
&\qquad \geq \inf_{x\in\mathcal{C}}[I(x) + F(x)] - \delta\,.
\end{aligned}
$$

In the above, we get (4.22) from the choice of $\widetilde{\mathbb{Q}}^n$. To get the first term of (4.23), we use Lemma 4.1 and to get the second term, we apply the continuous mapping theorem along with Lemma 4.2. To get (4.24), we use the weak lower semi-continuity of $\|\cdot\|_2^2$. The last display follows trivially from the definition of I. Now the arbitrariness of $\delta > 0$ proves (4.10). □

Remark 4.1. In adapting the approach in [23], we have modified the method to prove the result in Lemma 4.1. The corresponding result is proved in Theorems 2.5 in [23], which provides the sufficient condition for all the necessary tightness results. More elaborately, define a measure $\widetilde{\mathbb{Q}}^{n,*} \in \mathcal{P}(\mathbb{R}\times[0,1])$ as $\widetilde{\mathbb{Q}}^{n,*}(dy \times dt) \doteq \delta_{U^n(t)}(dy)dt$ with δ_x being the Dirac delta measure at x and U^n is as defined in Lemma 4.2. It is then shown that under appropriate conditions $\{(\widetilde{\mathbb{Q}}^{n,*}, \widetilde{X}^n)\}_{n\in\mathbb{N}}$ is tight in $\mathcal{P}(\mathbb{R}\times[0,1])\times\mathcal{C}$, where $\mathcal{P}(\mathbb{R}\times[0,1])$ is equipped with the topology of weak convergence and $\mathcal{C}$ is equipped with uniform topology. It also characterizes the limit $(\widetilde{\mathbb{Q}}^*, \widetilde{X}^*) \in \mathcal{P}(\mathbb{R}\times[0,1])\times\mathcal{C}$, in particular, $\widetilde{X}^*(t) = \int_0^t \int_{\mathbb{R}} y \widetilde{\mathbb{Q}}^*_{1|2}(dy|t)$, where $\widetilde{\mathbb{Q}}^*_{1|2}(dy|s)$ is the conditional distribution of $\widetilde{\mathbb{Q}}^*$ on $\mathbb{R}$ given t. The function U^n is identified as a measure δ_{U^n} under the 'very weak' topology, and hence any possible weak limit of δ_{U^n} is a priori a non-trivial measure. In the setting of i.i.d. variables, the bounds in (4.12) and (4.14) suggest to make use of the Hilbert space $\mathcal{L}^2_1$ and its weak* topology, which facilitates to directly prove various bounds in order to establish the results in Lemmas 4.3 and 4.4. We think that this approach can be further extended to non-i.i.d. settings, for example, the space $\mathcal{L}^2_1$ could be replaced by a reproducing kernel Hilbert space associated with certain Gaussian processes.

4.2. **Proof of Theorem 2.2.** Recall that

$$\widehat{Y}^n(t) \doteq \widehat{S}^n_{\lfloor nt \rfloor} .$$

To begin with, using Lemmas 1.2 and 3.6, we know that the desired sample-path MDP of $\{Y^n\}_{n\in\mathbb{N}}$ follows from the sample-path MDP of $\{\widehat{Y}^n\}_{n\in\mathbb{N}}$. Again, an application of Lemmas 1.1 and 2.1 implies that it is sufficient to prove the theorem below.

Theorem 4.2. *Under Assumption 2.1,*

$$\lim_{n\to\infty} -\frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(-a_n^2 F(\widehat{Y}^n)\Big)\Big] = \inf_{x\in\mathcal{D}} \Big[F(x) + I(x)\Big] , \tag{4.25}$$

where I is as defined in (2.3).

Proof Sketch. The proof of this theorem follows very closely the proof of Theorem 4.1. Hence, we only sketch the proof, point out the significant differences and omit other details to avoid repetition. As in Section 4.1, the key ingredient here is Lemma B.1.

For a fixed n, define a function $G^n : \mathbb{R}^n \to \mathcal{D}$ as

$$G^n(x_1, x_2, \ldots, x_n)(t) \doteq \frac{1}{a_n\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \big(x_i - \bar{\vartheta}^n\big) .$$

It is easy to see that $\widehat{Y}^n(t) = G^n(\vartheta^n_1, \vartheta^n_2 \ldots, \vartheta^n_n)(t)$, for $t \in [0,1]$ and that G^n is a continuous function from $\mathbb{R}^n$ to $(\mathcal{D}, J_1)$. Now applying Lemma B.1 to $(\Omega, \mathcal{F}) = \big(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\big)$, $\mu \equiv \mathbb{P}^n$ with $\mathbb{P}^n$ defined in Section 4.1, and $f(\cdot) \equiv a_n^2 F \circ G^n(\cdot)$, we immediately get

$$\begin{aligned}
-\frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(-a_n^2 F(\widehat{Y}^n)\Big)\Big] &= -\frac{1}{a_n^2} \log \mathbb{E}\Big[\exp\Big(-a_n^2 F \circ G^n(\vartheta^n_1, \vartheta^n_2, \ldots, \vartheta^n_n)\Big)\Big] \\
&= -\frac{1}{a_n^2} \log \int_{\mathbb{R}^n} e^{-a_n^2 F\circ G^n(x)} \mathbb{P}^n(dx) \\
&= \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \left[\frac{1}{a_n^2} R(\mathbb{Q}^n \| \mathbb{P}^n) + \int_{\mathbb{R}^n} F \circ G^n(x) \mathbb{Q}^n(dx)\right] .
\end{aligned} \tag{4.26}$$

For $\varpi^n \doteq (\varpi^n_1, \varpi^n_2, \ldots, \varpi^n_n)$ with distribution $\mathbb{Q}^n \in \mathcal{P}(\mathbb{R}^n)$, let

$$\widetilde{Y}^n(t) \doteq \widetilde{S}^n_{\lfloor nt \rfloor} , \quad t \in [0,1].$$

Here, $\widetilde{S}^n_k$ is as defined in (4.6). Then,

$$\widetilde{Y}^n(t) = G^n(\varpi^n_1, \varpi^n_2, \ldots, \varpi^n_n)(t), \quad \text{for} \quad t \in [0,1] , \tag{4.27}$$

and consequently,

$$\int_{\mathbb{R}^n} F \circ G^n(x) \mathbb{Q}^n(dx) = \mathbb{E}\big[F(\widetilde{Y}^n)\big] .$$

As in Section 4.1, applying Theorem B.1 repeatedly for n times and using the definition of $\mathbb{Q}_i^n$ gives us the following:

$$R(\mathbb{Q}^n\|\mathbb{P}^n) = \sum_{i=1}^n R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n)\,.$$

Therefore, to prove Theorem 4.2, it suffices to prove

$$\limsup_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{Y}^n)\Big] \le \inf_{x\in\mathcal{D}} \Big[F(x)+I(x)\Big]\,, \tag{4.28}$$

and

$$\liminf_{n\to\infty} \inf_{\mathbb{Q}^n\in\mathcal{P}(\mathbb{R}^n)} \mathbb{E}\Big[\frac{1}{a_n^2}\sum_{i=1}^n R(\mathbb{Q}_i^n\|\mathbb{P}_\vartheta^n) + F(\widetilde{Y}^n)\Big] \ge \inf_{x\in\mathcal{D}} \Big[F(x)+I(x)\Big]\,. \tag{4.29}$$

The main difference in the key lemmas needed for proving (4.28) and (4.29) is the lemma analogous to Lemma 4.2, where (4.18) is replaced by

$$\widetilde{Y}^n - \widetilde{U}^n \Rightarrow 0 \quad \text{in} \quad (\mathcal{D}, J_1) \quad \text{as} \quad n\to\infty\,.$$

Consequently, the proofs of (4.28) and (4.29) follow closely the proofs of (4.9) and (4.10) with the only difference being the analogous replacement of (4.20) and (4.21). □

Appendix A. Criteria for sample-path MDP in $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$

In this section, we state some sufficient criteria to establish sample-path MDPs in $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$. Before proceeding, we state necessary and sufficient criteria for exponential tightness for processes in $(\mathcal{C}, \|\cdot\|_\infty)$ and $(\mathcal{D}, J_1)$. Recall the moduli of continuity (see, e.g., [10]): for every $0<\delta<1$ and $x\in\mathcal{D}$,

$$w_\delta(x) \doteq \sup_{t,s\in[0,1],|t-s|<\delta} |x(t)-x(s)|, \tag{A.1}$$

$$w'_\delta(x) \doteq \inf_{\{t_i\}} \max_{1\le i\le m} \sup_{t_{i-1}\le s\le t<t_i} |x(t)-x(s)|\,, \tag{A.2}$$

where the infimum is over all possible sets $\{t_i\}_{i=1}^m$ such that $0=t_0<t_1<t_2<\ldots<t_m=1$ and $\min_{1\le i\le m}(t_i-t_{i-1})>\delta$. Also recall another modulus of continuity:

$$w''_\delta(x) \doteq \sup_{\substack{t_1\le t\le t_2\\ t_2-t_1\le\delta}} \Big\{|x(t)-x(t_1)|\wedge|x(t)-x(t_2)|\Big\}\,. \tag{A.3}$$

The following result characterizes exponential tightness for processes in $(\mathcal{C}, \|\cdot\|_\infty)$.

Theorem A.1. *[45, Theorem 4.1] A family of $\mathcal{C}$–valued random variables $\{Z^n\}_{n\in\mathbb{N}}$ is exponentially tight with rate a_n^2, if and only if the following hold:*

$$\begin{aligned} &\lim_{R\to\infty}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(|Z^n(0)|\ge R\right) = -\infty\,,\\ &\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(w_\delta(Z^n)\ge\epsilon\right) = -\infty\,, \end{aligned} \tag{A.4}$$

for every $\epsilon>0$.

The following theorem gives two equivalent conditions for exponential tightness in $(\mathcal{D}, J_1)$.

Theorem A.2. *[45, Theorem 4.2] A family of $\mathcal{D}$–valued random variables $\{Z^n\}_{n\in\mathbb{N}}$ is exponentially tight with rate a_n^2, if and only if*

$$\lim_{R\to\infty}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\Big(\sup_{t\in[0,1]}|Z^n(t)|\geq R\Big)=-\infty\,, \tag{A.5}$$

and for every $\epsilon>0$, either

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(w'_\delta(Z^n)\geq\epsilon\right)=-\infty \tag{A.6}$$

holds or

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(|Z^n(0)-Z^n(\delta)|\geq\epsilon\right)=-\infty,\\
&\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(|Z^n(1-)-Z^n(1-\delta)|\geq\epsilon\right)=-\infty,\\
&\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{1}{a_n^2}\log\mathbb{P}\left(w''_\delta(Z^n)\geq\epsilon\right)=-\infty
\end{aligned} \tag{A.7}$$

hold.

Remark A.1. In [45, Theorem 4.2], only the proof that (A.5)–(A.6) together is equivalent to exponential tightness in $(\mathcal{D},J_1)$ is provided. Therefore, it suffices to show that Conditions (A.6) and (A.7) are equivalent. This follows from the relationships between w'_δ and w''_δ on Page 132 in [10]: for $j(x)\doteq\sup_{0\leq s\leq 1}|x(s)-x(s-)|$, $x\in\mathcal{D}$ and $0<\delta<\frac{1}{4}$,

$$\begin{aligned}
w'_\delta(x)&\geq w''_{\frac{\delta}{2}}(x)\vee|x(0)-x(\delta)|\vee|x(1-)-x(1-\delta)|\,,\\
w'_\delta(x)&\leq 24\Big\{w''_{4\delta}(x)\vee|x(0)-x(4\delta)|\vee|x(1-)-x(1-4\delta)|\Big\}\,.
\end{aligned}$$

We now state some useful sufficient criteria for sample-path MDPs in $(\mathcal{C},\|\cdot\|_\infty)$ and $(\mathcal{D},J_1)$. The rate function is represented in terms of the rate function associated with finite-dimensional MDP, which can often be further developed into certain functional forms for various processes. For example, the rate functions in (2.2) and (2.3) for random walks, and the rate function for shot noise processes in equation (2.18) of [6]. The following result is stated in [27] for the general theory of sample-path LDP, but can be used to establish sample-path MDP for the appropriately scaled processes in $\mathcal{D}$ as discussed in Section 1.2. Here for our purpose we only state it for MDP in $(\mathcal{D},J_1)$.

Theorem A.3. *[27, Theorem 4.28] Let $\{Z^n\}_{n\in\mathbb{N}}$ be a family of $\mathcal{D}$–valued random variables such that*

(i) for any $0\leq t_1<t_2<\ldots<t_m\leq 1$, the family of random variables

$$\{Z^n(t_1),Z^n(t_2),\ldots,Z^n(t_m)\}_{n\in\mathbb{N}}$$

satisfies an MDP in $\mathbb{R}^m$ with rate a_n^2 and rate function $I^Z_{t_1,t_2,\ldots,t_m}:\mathbb{R}^m\to[0,\infty]$, and

(ii) it is exponentially tight in $(\mathcal{D},J_1)$ with rate a_n^2.

Then, $\{Z^n\}_{n\in\mathbb{N}}$ satisfies an MDP in $(\mathcal{D},J_1)$ with rate a_n^2 and rate function $I^Z:\mathcal{D}\to[0,\infty]$ given by

$$I^Z(x)=\sup_{\substack{0\leq t_1<t_2<\ldots<t_m\leq 1\\ t_i\notin\Delta_x,\ \text{for } 1\leq i\leq m}} I^Z_{t_1,t_2,\ldots,t_m}(x(t_1),x(t_2),\ldots,x(t_m)),$$

where $\Delta_x\doteq\{t\in[0,1]:x(t-)\neq x(t)\}$.

We have the following immediate corollary for a sufficient criterion in $(\mathcal{C},\|\cdot\|_\infty)$.

Corollary A.1. *Let $\{Z^n\}_{n\in\mathbb{N}}$ be a family of $\mathcal{C}$–valued random variables such that Condition (i) of Theorem A.3 holds and is exponentially tight in $(\mathcal{C}, \|\cdot\|_\infty)$ with rate a_n^2. Then, $\{Z^n\}_{n\in\mathbb{N}}$ satisfies an MDP in $(\mathcal{C}, \|\cdot\|_\infty)$ with rate a_n^2 and rate function $I^Z : \mathcal{C} \to [0,\infty]$ given by*

$$I^Z(x) = \sup_{0\le t_1<t_2<\ldots<t_m\le 1} I^Z_{t_1,t_2,\ldots,t_m}(x(t_1), x(t_2), \ldots, x(t_m)) .$$

The following result from [48, Theorem A.3] gives another sufficient condition for sample-path MDP of $\mathcal{D}$–valued family of random variables using a sufficient condition for exponential tightness (see (A.8)). The criterion also guarantees that the rate function only takes finite values on $\mathcal{C}$.

Theorem A.4. *Suppose that $\{Z^n\}_{n\in\mathbb{N}}$ is a $\mathcal{D}$–valued family of random variables such that the Condition (i) of Theorem A.3 holds and for every $\epsilon > 0$,*

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{t\in[0,T]} \frac{1}{a_n^2} \log \mathbb{P}\left(\sup_{0\le s\le\delta} |Z^n(t+s) - Z^n(t)| \ge \epsilon \right) = -\infty . \tag{A.8}$$

Then, $\{Z^n\}_{n\in\mathbb{N}}$ satisfies an MDP in $(\mathcal{D}, J_1)$ with rate a_n^2 and rate function $I^Z : \mathcal{D} \to [0,\infty]$ given by

$$I^Z(x) = \begin{cases} \sup_{0\le t_1<t_2<\ldots<t_m\le 1} I^Z_{t_1,t_2,\ldots,t_m}(x(t_1), x(t_2), \ldots, x(t_m)), & \text{if } x \in \mathcal{C}, \\ \infty, & \text{if } x \in \mathcal{D} \setminus \mathcal{C} . \end{cases}$$

Appendix B. Relative entropy and the associated variational formula

This section contains the preliminary results necessary to prove Theorem 2.1 using the variational approach. We recall the notion of relative entropy and some relevant results. Suppose $(\Omega, \mathcal{F})$ is a measurable space. Let $R(\mu_1\|\mu_2)$ be the relative entropy of μ_1 with respect to μ_2, *i.e.,*

$$R(\mu_1\|\mu_2) \doteq \begin{cases} \int_\Omega \log\left(\frac{d\mu_1}{d\mu_2}(x)\right) d\mu_1(x), & \text{if } \mu_1 \ll \mu_2, \\ \infty, & \text{otherwise} . \end{cases}$$

Lemma B.1. *[14, Proposition 2.2] Let μ be a probability measure on $(\Omega, \mathcal{F})$ and $f : \Omega \to \mathbb{R}$ be a bounded measurable function. Then the following variational formula holds:*

$$-\log \int_\Omega e^{-f(x)} d\mu(x) = \inf_{\nu\in\mathcal{P}(\Omega)} \Big[R(\nu\|\mu) + \int_\Omega f(x) d\nu(x) \Big] . \tag{B.1}$$

The theorem below gives us the well-known chain rule for relative entropy.

Theorem B.1. *[22, Page 332] Let Ω_1 and Ω_2 be two Polish spaces with μ and ν being the probability measures on $\Omega_1 \times \Omega_2$. Let μ_1 and ν_1 denote the first marginals of μ and ν, respectively, and $\mu_{|1}$, $\nu_{|1}$ be the conditional distributions of the second component given the first of μ and ν, respectively,* i.e.,

$$\mu(dx \times dy) = \mu_1(dx)\mu_{|1}(dy|x) \text{ and } \nu(dx \times dy) = \nu_1(dx)\nu_{|1}(dy|x) .$$

Then, we have

$$R(\mu\|\nu) = R(\mu_1\|\nu_1) + \int_{\Omega_1} R\big(\mu_{|1}(\cdot|x)\|\nu_{|1}(\cdot|x)\big)\mu_1(dx) . \tag{B.2}$$

Lemma B.2. *Suppose $\mu \in \mathcal{P}(\mathbb{R})$ is a compactly supported measure with mean m_μ. Then the following bound holds for any $\nu \ll \mu$,*

$$R(\nu\|\mu) \ge H^c_\mu\Big(\int_\mathbb{R} x\nu(dx) - m_\mu\Big) .$$

Here, H^c_μ is as defined in (4.13).

Proof. Since $\mu \in \mathcal{P}(\mathbb{R})$ has a compact support, for any $\alpha \in \mathbb{R}$, we have $\int_{\mathbb{R}} e^{\alpha x}\mu(dx) < \infty$. We now apply Lemma B.1 to $f(x) = -\alpha(x - m_\mu)$, to get

$$-\log \int_{\mathbb{R}} e^{\alpha(x-m_\mu)}\mu(dx) = \inf_{\nu \in \mathcal{P}(\mathbb{R})} \Big[R(\nu\|\mu) - \alpha \int_{\mathbb{R}} x\nu(dx) + \alpha m_\mu\Big]$$
$$\leq R(\nu\|\mu) - \alpha \int_{\mathbb{R}} x\nu(dx) + \alpha m_\mu,$$

for any measure $\nu \ll \mu$. Re-arranging the last inequality, we have

$$R(\nu\|\mu) \geq \alpha\Big(\int_{\mathbb{R}} x\nu(dx) - m_\mu\Big) - \log \int_{\mathbb{R}} e^{\alpha(x-m_\mu)}\mu(dx)\,.$$

Since the above inequality holds for all $\alpha \in \mathbb{R}$, we immediately have

$$R(\nu\|\mu) \geq \sup_{\alpha\in\mathbb{R}} \Big\{\alpha\Big(\int_{\mathbb{R}} x\nu(dx) - m_\mu\Big) - \log \int_{\mathbb{R}} e^{\alpha(x-m_\mu)}\mu(dx)\Big\}$$
$$= H^c_\mu\Big(\int_{\mathbb{R}} x\nu(dx) - m_\mu\Big)\,.$$

The last line follows from the definition of $H^c_\mu(\cdot)$. This proves the lemma. □

Appendix C. Proof of (3.11)

The claim in (3.11) states that for every $\rho > 0$,

$$\lim_{n\to\infty} \frac{n}{a_n^2} \log \mathbb{E}\left[\exp\left(\frac{\rho a_n}{\sqrt{n}}|w_1^n|\right)\right] = 0\,.$$

Recall that

$$w_1^n = \tilde{\vartheta}_1^n - \vartheta_1^n + \mathbb{E}\Big[\vartheta_1^n\Big]$$
$$= (\vartheta_1 - \bar{\vartheta})\mathbb{1}_{\{|\vartheta_1-\bar{\vartheta}|\leq\frac{\delta}{2}a_n\sqrt{n}\}} - \vartheta_1\mathbb{1}_{[0,l\sqrt{n}a_n^{-1}]}(|\vartheta_1|) + \mathbb{E}\Big[\vartheta_1\mathbb{1}_{[0,l\sqrt{n}a_n^{-1}]}(|\vartheta_1|)\Big]\,.$$

Using Taylor's series of $\exp(x)$, we get

$$\frac{n}{a_n^2} \log \mathbb{E}\left[\exp\left(\frac{\rho a_n}{\sqrt{n}}|w_1^n|\right)\right] = \frac{n}{a_n^2}\Big(\frac{\rho a_n}{\sqrt{n}}\mathbb{E}[|w_1^n|] + \frac{\rho^2 a_n^2}{2n}\mathbb{E}[|w_1^n|^2] + \mathcal{O}\Big(\frac{a_n^3}{n\sqrt{n}}\mathbb{E}[|w_1^n|^3]\Big)\Big)$$
$$= \frac{\rho\sqrt{n}}{a_n}\mathbb{E}[|w_1^n|] + \frac{\rho^2}{2}\mathbb{E}[|w_1^n|^2] + \mathcal{O}\Big(\frac{a_n}{\sqrt{n}}\mathbb{E}[|w_1^n|^3]\Big)\,.$$

Since $\mathbb{E}[|\vartheta_1|^2] < \infty$, from the definition of w_1^n, the expectation in the second term above goes to zero as $n \to \infty$. It remains to show that $a_n^{-1}n^{\frac{1}{2}}\mathbb{E}[|w_1^n|]$ and $a_n n^{-\frac{1}{2}}\mathbb{E}[|w_1^n|^3]$ go to zero, as $n \to \infty$. In the following, we only consider the term $a_n^{-1}n^{\frac{1}{2}}\mathbb{E}[|w_1^n|]$ in detail, as the analysis for $a_n n^{-\frac{1}{2}}\mathbb{E}[|w_1^n|^3]$ follows analogously. We begin by recalling two well-known identities concerning a real-valued random variable Z:

$$\mathbb{E}[|Z|\mathbb{1}_{(a,b)}(|Z|)] = \int_a^b \mathbb{P}(|Z| > z)dz\,, \tag{C.1}$$

$$\mathbb{E}[|Z|^2\mathbb{1}_{(a,b)}(|Z|)] = \int_a^b z\mathbb{P}(|Z| > z)dz\,. \tag{C.2}$$

To that end, since $a_n \uparrow \infty$ as $n \to \infty$, we can choose n large enough such that

$$l\frac{\sqrt{n}}{a_n} < \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}\,. \tag{C.3}$$

Now, consider

$$\frac{\sqrt{n}}{a_n}\mathbb{E}[|w_1^n|] = \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{[0,la_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big] + \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{(la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] .$$

We now evaluate the two terms on the right hand side separately. We do this by considering three exclusive and exhaustive cases *viz.,*

(i) $|\vartheta_1| \le la_n^{-1}\sqrt{n}$,
(ii) $la_n^{-1}\sqrt{n} < |\vartheta_1| \le \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}$,
(iii) $|\vartheta_1| > \frac{\delta}{2}a_n\sqrt{n} + \bar{\vartheta}$.

It turns out that in each of these cases, w_1^n takes a simple form which we then bound by making use of (C.1) and (C.2).

From the definition of w_1^n and (C.3), whenever $|\vartheta_1| \le la_n^{-1}\sqrt{n}$, we have

$$w_1^n = \mathbb{E}\Big[\vartheta_1\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] .$$

This implies that

$$\begin{aligned}\frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{[0,la_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big] &= \frac{\sqrt{n}}{a_n}\Big|\mathbb{E}\Big[\vartheta_1\mathbb{1}_{[l\sqrt{n}a_n^{-1},\infty)}(|\vartheta_1|)\Big]\Big|\mathbb{E}\Big[\mathbb{1}_{[0,la_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big]\\ &\le \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1|\mathbb{1}_{[l\sqrt{n}a_n^{-1},\infty)}(|\vartheta_1|)\Big] .\end{aligned}$$

To arrive at the second inequality, we use the fact that $\mathbb{1}_{[0,la_n^{-1}\sqrt{n}]}(|\vartheta_1|) \le 1$ and that $|\mathbb{E}[Z]| \le \mathbb{E}[|Z|]$.

Next, we get

$$\begin{aligned}\frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1|\mathbb{1}_{[l\sqrt{n}a_n^{-1},\infty)}(|\vartheta_1|)\Big] &= \frac{\sqrt{n}}{a_n}\int_{la_n^{-1}\sqrt{n}}^{\infty}\mathbb{P}(|\vartheta_1| > x)dx\\ &\le l^{-1}\int_{l\sqrt{n}a_n^{-1}}^{\infty}x\mathbb{P}(|\vartheta_1| > x)dx\\ &\le l^{-1}\mathbb{E}\Big[|\vartheta_1|^2\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] . \end{aligned} \tag{C.4}$$

In the above, to get the first line, we use the identity (C.1); to get the second line, we use the fact the integral is on $[l\sqrt{n}a_n^{-1},\infty)$ and to get the last line, we use the identity (C.2).

Now consider

$$\begin{aligned}\frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] &= \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{(la_n^{-1}\sqrt{n},\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta}]}(|\vartheta_1|)\Big] + \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|w_1^n|\mathbb{1}_{(\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta},\infty)}(|\vartheta_1|)\Big]\\ &\doteq J_1 + J_2 .\end{aligned}$$

Whenever $la_n^{-1}\sqrt{n} < |\vartheta_1| \le \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}$, we have

$$w_1^n = \vartheta_1 + \mathbb{E}\Big[\vartheta_1\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] .$$

Then,

$$\begin{aligned}J_1 &= \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[\Big|\vartheta_1 + \mathbb{E}\Big[\vartheta_1\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big]\Big|\mathbb{1}_{(la_n^{-1}\sqrt{n},\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta}]}(|\vartheta_1|)\Big]\\ &\le \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1|\mathbb{1}_{[la_n^{-1}\sqrt{n},\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta}]}(|\vartheta_1|)\Big] + \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1|\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big]\mathbb{E}\Big[\mathbb{1}_{[la_n^{-1}\sqrt{n},\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta}]}(|\vartheta_1|)\Big]\\ &\le \frac{2\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1|\mathbb{1}_{[la_n^{-1}\sqrt{n},\infty)}(|\vartheta_1|)\Big] . \end{aligned} \tag{C.5}$$

In the above, to arrive at the last line, we use the fact that $[la_n^{-1}\sqrt{n}, \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}] \subset [la_n^{-1}\sqrt{n},\infty)$ and $\mathbb{1}_{[la_n^{-1}\sqrt{n},\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta}]}(|\vartheta_1|) \le 1$.

Whenever $|\vartheta_1| > \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}$, we have

$$w_1^n = \mathbb{E}\Big[\vartheta_1 \mathbb{1}_{[0, l a_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big] .$$

This implies that

$$\begin{aligned} J_2 &= \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[\Big|\mathbb{E}\Big[\vartheta_1 \mathbb{1}_{[0, l a_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big]\Big| \mathbb{1}_{[\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta},\infty)}(|\vartheta_1|)\Big] \\ &\le \frac{\sqrt{n}}{a_n}\mathbb{E}\Big[|\vartheta_1| \mathbb{1}_{[0, l a_n^{-1}\sqrt{n}]}(|\vartheta_1|)\Big]\mathbb{E}\Big[\mathbb{1}_{[\frac{\delta}{2}a_n\sqrt{n}-\bar{\vartheta},\infty)}(|\vartheta_1|)\Big] \\ &\le \frac{\sqrt{n}}{a_n}\mathbb{E}[|\vartheta_1|]\mathbb{P}\left(|\vartheta_1| > \frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta}\right) \\ &\le \frac{\sqrt{n}}{a_n}\mathbb{E}[|\vartheta_1|]\frac{\mathbb{E}[|\vartheta_1|^2]}{(\frac{\delta}{2}a_n\sqrt{n} - \bar{\vartheta})^2} . \end{aligned} \tag{C.6}$$

In the above, to get the second line, we use the fact that $|\mathbb{E}[Z]| \le \mathbb{E}[|Z|]$; to get the third line, we use the fact that $\mathbb{1}_{[0, l a_n^{-1}\sqrt{n}]}(|\vartheta_1|) \le 1$ and to get the last line, we apply Markov's inequality. From (C.4)–(C.6), we have

$$\lim_{n\to\infty} \frac{\sqrt{n}}{a_n}\mathbb{E}[|w_1^n|] = 0 .$$

As mentioned already, following similar arguments as above and applying Lemma 3.1, we obtain

$$\lim_{n\to\infty} \frac{a_n}{\sqrt{n}}\mathbb{E}[|w_1^n|^3] = 0 .$$

This completes the proof of (3.11).

Acknowledgement

This work is partly funded by the NSF Grant DMS 2216765.

References

[1] N. N. Amosova. Local limit theorems for probabilities of moderate deviations. *Lithuanian mathematical transactions of the Academy of Sciences of the Lithuanian SSR*, 14(3):401–407, 1974.

[2] N. N. Amosova. Local limit theorems for probabilities of moderate deviations. *Mathematical notes of the Academy of Sciences of the USSR*, 28(3):685–689, 1980.

[3] N. N. Amosova. Probabilities of moderate deviations. *Journal of Soviet Mathematics*, 20(3):2123–2130, 1982.

[4] N. N. Amosova. A supplement to W. Richter's theorem. *Mathematical Notes of the Academy of Sciences of the USSR*, 42(1):577–578, 1987.

[5] N. N. Amosova and W. D. Richter. Moderate deviations for densities in $\mathbb{R}^k$. *Journal of Soviet Mathematics*, 27:2997–3002, 1984.

[6] S. R. Anugu and G. Pang. Sample path moderate deviations for shot noise processes in the high intensity regime. *Stochastic Processes and Applications*, to appear, 2024.

[7] M. A. Arcones. The large deviation principle for stochastic processes. Part I. *Theory of Probability & Its Applications*, 47(4):567–583, 2003.

[8] M. A. Arcones. Moderate deviations of empirical processes. In *Stochastic inequalities and applications*, pages 189–212. Springer, 2003.

[9] M. A. Arcones. The large deviation principle for stochastic processes. II. *Theory of Probability & Its Applications*, 48(1):19–44, 2004.

[10] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 1999.

[11] A. A. Borovkov. Boundary-value problems for random walks and large deviations in function spaces. *Theory of Probability & Its Applications*, 12(4):575–595, 1967.

[12] A. A. Borovkov and A. A. Mogul'skii. Probabilities of large deviations in topological spaces. I. *Siberian Mathematical Journal*, 19(5):697–709, 1978.

[13] A. A. Borovkov and A. A. Mogul'skii. Probabilities of large deviations in topological spaces. II. *Siberian Mathematical Journal*, 21(5):653–664, 1980.

[14] A. Budhiraja and P. Dupuis. *Analysis and approximation of rare events*, volume 94. Springer, 2019.

[15] X. Chen. The moderate deviations of independent random vectors in a Banach space. *Chinese J. Appl. Probab. Statist*, 7(1):24–32, 1991.

[16] H. Cramér. Sur un nouveau théoreme-limite de la théorie des probabilités. In *Collected Works II*, pages 895–913. Springer, 1994.

[17] H. Cramér and H. Touchette. On a new limit theorem in probability theory (translation of 'sur un nouveau théoréme-limite de la théorie des probabilités'). *arXiv:1802.05988*, 2018.

[18] J. A. Davis. Convergence rates for probabilities of moderate deviations. *The Annals of Mathematical Statistics*, 39(6):2016–2028, 1968.

[19] A. De Acosta. Moderate deviations and associated Laplace approximations for sums of independent random vectors. *Transactions of the American Mathematical Society*, 329(1):357–375, 1992.

[20] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*. Springer Science & Business Media, 2009. Stochastic Modelling and Applied Probability (SMAP, Volume 38).

[21] C. M. Deo and G. J. Babu. Probabilities of moderate deviations in a Banach space. *Proceedings of the American Mathematical Society*, 83(2):392–397, 1981.

[22] P. Dupuis and R. S. Ellis. *A weak convergence approach to the theory of large deviations*. John Wiley & Sons, 1997.

[23] P. Dupuis and D. Johnson. Moderate deviations for recursive stochastic algorithms. *Stochastic Systems*, 5(1):87–119, 2015.

[24] P. Eichelsbacher and M. Löwe. Moderate deviations for I.I.D. random variables. *ESAIM: Probability and Statistics*, 7:209–218, 2003.

[25] N. Etemadi. On some classical results in probability theory. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 215–221, 1985.

[26] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 2009.

[27] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*. American Mathematical Society, 2006.

[28] F. Gao. Moderate deviations and functional limits for random processes with stationary and independent increments. *Science in China Series A: Mathematics*, 49:1753–1767, 2006.

[29] A. Gut. Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *The Annals of Probability*, 8(2):298–313, 1980.

[30] Y. Hu and T.-Y. Lee. Moderate deviation principles for trajectories of sums of independent Banach space valued random variables. *Transactions of the American Mathematical Society*, 355(8):3047–3064, 2003.

[31] M. Ledoux. Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. In *Annales de l'IHP Probabilités et statistiques*, volume 28, pages 267–280, 1992.

[32] Y. V. Linnik. On the probability of large deviations for the sums of independent variables. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob*, volume 2, pages 289–306, 1961.

[33] R. Michel. Results on probabilities of moderate deviations. *The Annals of Probability*, 2(2):349–353, 1974.

[34] R. Michel. Nonuniform central limit bounds with applications to probabilities of deviations. *The Annals of Probability*, 4(1):102–106, 1976.

[35] A. A. Mogul'skii. Large deviations for trajectories of multi-dimensional random walks. *Theory of Probability & its Applications*, 21(2):300–315, 1977.

[36] A. A. Mogul'skii. Large Deviations for Processes with Independent Increments. *The Annals of Probability*, 21(1):202 – 215, 1993.

[37] L. V. Osipov. On probabilities of large deviations for sums of independent random variables. *Theory of Probability & Its Applications*, 17(2):309–331, 1973.

[38] V. V. Petrov. On the probabilities of large deviations for sums of independent random variables. *Theory of Probability & Its Applications*, 10(2):287–298, 1965.

[39] V. V. Petrov. Sum of independent random variables. *Yu. V. Prokhorov. V. StatuleviCius (Eds.)*, 1972.

[40] V. V. Petrov. On probabilities of moderate deviations. *Journal of Mathematical Sciences*, 109:2189–2191, 2002.

[41] V. V. Petrov. The law of the iterated logarithm and probabilities of moderate deviations of sums of dependent random variables. *Journal of Mathematical Sciences*, 251(1):128–130, 2020.

[42] V. V. Petrov and J. Robinson. Large deviations for sums of independent non identically distributed random variables. *Communications in Statistics - Theory and Methods*, 37(18):2984–2990, 2008.

[43] D. Picard and J. Deshayes. Grandes et moyennes déviations pour les marches aléatoires. In *Grandes déviations et applications statistiques (Séminaire Orsay 1977-1978)*, number 68 in Astérisque, pages 53–71. Société mathématique de France, 1979.

[44] D. Pollard. *Convergence of stochastic processes*. Springer Science & Business Media, 2012.

[45] A. A. Puhalskii. *On Functional Principle of Large Deviations*, pages 198–218. De Gruyter, Berlin, Boston, 1991.

[46] A. A. Puhalskii. Large deviations of semimartingales via convergence of the predictable characteristics. *Stochastics: An International Journal of Probability and Stochastic Processes*, 49(1-2):27–85, 1994.

[47] A. A. Puhalskii. *Large deviations and idempotent probability*. Chapman and Hall/CRC, 2001.

[48] A. A. Puhalskii. Moderate deviations of many-server queues in the Halfin-Whitt regime and weak convergence methods. *arXiv:2305.01612*, 2023.

[49] W. Richter. Local limit theorems for large deviations. *Theory of Probability & Its Applications*, 2(2):206–220, 1957.

[50] V. Rohatgi. On the rate of convergence of probabilities of moderate deviations. *Journal of the Australian Mathematical Society*, 11:91 – 94, 02 1970.

[51] H. Rubin and J. Sethuraman. Bayes risk efficiency. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 347–356, 1965.

[52] H. Rubin and J. Sethuraman. Probabilities of moderate deviations. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 27(2/4):325–346, 1965.

[53] Z. Rychlik. Non-uniform central limit bounds with applications to probabilities of deviations. *Theory of Probability & its Applications*, 28(4):681–687, 1984.

[54] L. Saulis and V. Statulevicius. *Limit theorems for large deviations*, volume 73. Springer Science & Business Media, 2012.

[55] A. D. Slastnikov. On large deviations for sums of non-identically distributed random variables. *Theory of Probability & Its Applications*, 27(1):37–48, 1982.

[56] D. W. Stroock. *An introduction to the theory of large deviations*. Springer Science & Business Media, 2012.

[57] J. Tiefeng, M. B. Rao, W. Xiangchen, and L. Deli. Laws of large numbers and moderate deviations for stochastic processes with stationary and independent increments. *Stochastic processes and their applications*, 44(2):205–219, 1993.

[58] B. Tsirelson. Moderate deviations on different scales: no relations. *Israel Journal of Mathematics*, 203:161–71, 2014.

[59] S. R. S. Varadhan. *Large deviations and applications*. CBMS-NSF regional conference series in applied mathematics ; 46. Society for Industrial and Applied Mathematics, Philadelphia, Pa, 1984.

[60] W. Whitt. *Stochastic-process limits: An introduction to stochastic-process limits and their application to queues*. Springer, 2002.

[61] L. Wu. Moderate Deviations of Dependent Random Variables Related to CLT. *The Annals of Probability*, 23(1):420 – 445, 1995.