Sample Path Moderate Deviations for Shot Noise Processes in the High Intensity Regime

SUMITH REDDY ANUGU† AND GUODONG PANG‡

Abstract. We study the sample-path moderate deviation principle (MDP) for shot noise processes in the high intensity regime. The shot noise processes have a renewal arrival process, non-stationary noises (with arrival-time dependent distributions) and a general measurable shot response function of the noises. The rate function in the MDP exhibits a memory phenomenon in this asymptotic regime, which is in contrast with that in the conventional time-space scaling regime.

To prove the sample-path MDP, we first establish an exponential equivalence with another process that is easier to study. We will then establish the sample-path MDP of this process combining the Gärtner-Ellis (to prove the finite dimensional MDP) and Dawson-Gärtner Theorem (to prove the sample-path MDP under the topology of pointwise convergence). Finally, we prove exponential tightness and strengthen the MDP to the Skorohod topology. In the proofs, because of the inherent non-stationarity of shot noise process, we exploit a certain maximal inequality for the sum of independent random variables (which are not necessarily identically distributed) as well as Lusin’s theorem for exponential tightness. The rate function is derived using the tools of reproducing kernel Hilbert space.

1. Introduction

Shot noise process can be viewed as a natural model for a system which experiences shocks that occur according to an arrival process and have an enduring effect on its dynamics. In particular, they have been found very useful in the areas of physics ([7, 39, 53]), queueing theory ([10, 27, 38]) and teletraffic theory ([37, 48]), insurance and risk theory ([35, 36, 41, 42, 44, 55]), storage processes ([11, 40]) and so on.

Various asymptotic properties and scaling limits have been established for shot noise processes. There are two asymptotic scaling regimes that have been studied in the literature. The first one is the conventional time-space scaling regime (speeding up time and scaling down space; see (2.15) and (2.22)). This is studied in different settings, including functional central limit theorems (FCLTs) in [29, 30, 31, 32, 34, 35, 36] and sample-path large deviation principles (LDPs) in [15, 19, 22]. The second one is referred to as the high intensity regime (the arrival rate is scaled up while time is not scaled in the shot response function, and space is scaled down; see (2.5) and (2.7)). FCLTs and relevant asymptotic properties in this regime have been studied in [4, 26, 28, 46, 47]. Infinite-server queues can be regarded as a shot noise process with a particular indicator response function, and heavy-traffic limits (that is, in the high intensity regime with no scaling on service times) have been established in the literature (see, e.g., [45] and references therein). Large deviation principles are also established for infinite-server queues in heavy traffic/high intensity regime, see [24] for results at a fixed time, and [6] for the sample-path LDP of a two-parameter process tracking elapsed/residual service times. However, to our best knowledge, no MDPs have been established for shot noise processes in either asymptotic regime in the literature.

† 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, shot noise process, renewal arrival process, non-stationary noise distribution (arrival-time dependent distributions).
The main aim of this paper is to establish the sample-path MDP for shot noise processes in the high intensity regime. We consider the arrival process to be a renewal process satisfying a sample-path MDP, and assume that the noises are non-stationary, in particular, the noises are conditionally independent given the arrival times while the distribution of each noise depends on its own arrival time (see Assumptions A.1–A.3). The shot response function also satisfies rather general conditions (Assumptions A.4–A.5). The MDP-scaled process in this regime is given in (2.7), which has a centering term as in FCLTs, but has a scaling of arrival rate and space satisfying a certain condition (2.8). See further discussions on the scaling at the beginning of Section 2.2. The main result of the sample-path MDP is given in Theorem 2.1, in which the rate function consists of two components, one involving a “covariance” operator (corresponding to the covariance function in the Gaussian limit with memory of the FCLT in [46]) and the other being standard corresponding to that of a Brownian motion. We refer to Remark 2.7 on discussions about how the MDP rate functions are related to the Gaussian limit processes in the FCLT. As a corollary, we also state the sample-path MDP result for the $GI/G_t/\infty$ queueing model (Corollary 2.1).

To prove the sample-path MDP, we first prove that the scaled process of interest $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ (see (2.7)) is exponentially equivalent to another process $\{\tilde{\tilde{X}}^n\}_{n \in \mathbb{N}}$ that is easier to study (Lemma 3.1) and this helps us to conclude the sample-path MDP of $\{X^n\}_{n \in \mathbb{N}}$ from that of $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ with the same rate and the same rate function. In this construction, the component $\tilde{X}^n_1$ in the decomposition of $\tilde{X}^n$ (see (3.1), the two components are dependent although asymptotically independent) is replaced by $\tilde{\tilde{X}}^n_1$ in the decomposition of $\tilde{\tilde{X}}^n$ (see (3.6)) such that the two components in $\tilde{\tilde{X}}^n$ are independent.

We then use Gärtner-Ellis theorem to conclude the MDP of the finite dimensional distributions of $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ (Proposition 3.1 and Lemma 3.2). Then we use Dawson-Gärtner theorem to conclude sample-path MDP of $\{\tilde{X}^n_1\}_{n \in \mathbb{N}}$ in the topology of pointwise convergence (Theorem 3.1). Due to the inherent non-stationarity of shot noise process and the additional challenge of arrival-time dependence of the noises, going from finite dimensional rate function to the sample-path rate function (the content of Dawson-Gärtner theorem) is difficult. Therefore, we use the following observation: the rate function can be viewed as an appropriate inverse of covariance operator (corresponding to the limiting Gaussian process of the associated FCLT). This indicates us to express the rate function in the language of reproducing kernel Hilbert space.

Next, to strengthen the MDP to the Skorohod topology, we establish exponential tightness of $\{\tilde{X}^n_1\}_{n \in \mathbb{N}}$ in the Skorohod topology (Lemma 3.3). The proof of exponential tightness use a certain maximal inequality of a sum of independent random variables that are not necessarily identically distributed [20, Theorem 1.1]. This relates the tail probability of the running maximum of the sum to the running maximum of the tail probability of the sum (see Lemma A.2 and Corollary A.1 in the Appendix). Due to the lack of sufficient continuity on shot shape function, the proof is more involved and uses Lusin’s theorem to arrive at the relevant estimates. Then, the sample-path MDP for the second component $\tilde{X}^n_2$ in (3.1) can be established by using the known results for the MDP of renewal processes ([52]) and applying the contraction principle (see Proposition 3.2). Finally, combining all these results together we obtain the sample-path MDP for $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ in the space $\mathcal{D}_T$ endowed with the Skorohod topology.

As a comparison we also state the the sample-path MDP result in the conventional time-space scaling regime (Theorem 2.2 in Section 2.3). Although both the scalings give rise to MDP with the same rates, the rate functions are dramatically different. In the case of the conventional time-space scaling, the rate function looks like the inverse of covariance of a Brownian motion (with a time-varying covariance function), whereas in the high intensity regime, the rate function looks like the inverse of a certain non-stationary Gaussian process with memory. See further discussions in Remarks 2.8 and 2.9.
As another comparison, we also state the sample-path LDPs for the shot noise processes in the two scaling regimes (Section 2.4). It is expected that the rate functions in the LDPs are very different in the two regimes (see (2.19) and (2.23)), since in the conventional time-space scaling regime, the ‘lingering’ effect of noises vanishes (equivalent to that the response function \(H(t,x)\) is replaced by \(H(\infty,x)\)), while in the high intensity scaling regime, the effect of noises is indicated as ‘memory’. On the other hand, the rate functions in the LDPs involve the LDP rate function of the renewal arrival process, which further involves the log moment generating function of the interarrival times (see (2.20)). However, the rate functions in the MDPs in both regimes only involves the mean and variance of the interarrival times of the renewal arrival process (Theorems 2.1 and 2.2). We highlight that some of the LDP results in both regimes are also new to the literature since we consider non-stationary noises (see further discussions in Section 2.4). The sample-path large deviation principles (LDPs) in [15, 19, 22] all assume Poisson arrival processes and stationary noises in the conventional time-space scaling regime. The results we present are more general involving both a renewal arrival process and non-stationary noises. The methodology we develop also goes beyond those in [15, 22] since the well developed LDP methods for Poisson random measures (see the recent monograph [13]) could be used, which is impossible in our setting.

To put our paper in the context of the vast literature of MDPs of stochastic systems, we give a partial overview of the different methods used to prove MDPs. The general approach to prove sample-path LDP and MDP is to use Dawson-Gärtner theorem [18, Theorem 4.6.1] in conjunction with Gärtner-Ellis theorem [18, Theorem 2.3.6] and exponential tightness in appropriate functional spaces. This has been used in study various Markov and non-Markov systems (see, e.g., [6, 19, 21, 22, 49, 50, 51]). Specific properties of processes of interest are often exploited to establish the required properties, such as convergence of the non-linear semi-groups of Markov processes [21], and semimartingale representations [50, 51]. For certain Markov processes driven by Brownian motion or Poisson random measure, a weak convergence approach using the variational representation of certain functionals of Brownian motion and/or Poisson random measure and the associated control problem formulation (see [13, Section 3.2, 3.3, 8.1 and 8.2]) has been used extensively (see, e.g., [3, 12, 14, 16, 17, 23, 43] for Markov models and [1, 9, 33] for non-Markov models). However, shot noise processes in our paper bring in new challenges with the non-Markovian and non-stationary characteristics. This leads us to great difficulties in proving the exponential tightness (Lemma 3.3) and exponential equivalence (Lemma 3.1), as well as identifying the rate function due to the memory property. We have developed new methods to tackle these challenges, which may turn to be useful to study LDP and MDP for other non-Markovian stochastic systems in future work.

1.1. Organization of the paper. In the rest of this section, we introduce notation that we will use throughout the paper. In Section 2.1, we introduce the model and assumptions, and in Section 2.2, we state the main results on sample-path MDP in the high intensity regime. In Sections 2.3 and 2.4, we compare with the MDP in the conventional time-space scaling regime, and the LDP results in the two scaling regimes. In Section 3, we provide the proofs of the main result. In the Appendix, we state the auxiliary results that are used in the proofs.

1.2. Notation. \((\Omega, \mathcal{F}, \mathbb{P})\) denotes the abstract probability space. For a Polish space \(S\), \(\mathcal{B}(S)\) denotes the corresponding Borel \(\sigma\)-algebra. \(\mathcal{P}(S)\) denotes the space of probability measures on \(S\) equipped with the topology of weak convergence. For a fixed \(T > 0\), \((\mathcal{D}_T, J_1)\) denotes the space of real valued functions on \([0,T]\) that are right continuous with left limits equipped with the Skorohod topology \((d_{J_1}, \langle \cdot, \cdot \rangle)\) is the corresponding metric). Let \(\|x\|_T := \sup_{0 \leq t \leq T} |x(t)|\) for \(x \in \mathcal{D}_T\), and \(\langle \cdot, \cdot \rangle\) denote the Euclidean inner product on \(\mathbb{R}^n\), for any \(n \in \mathbb{N}\). We denote the set of absolutely continuous functions \(x : [0,T] \to \mathbb{R}\) such that \(x(0) = 0\) by \(\mathcal{AC}_0\). For \(x, y \in \mathbb{R}\), we write \(a \wedge b\) for \(\min(a, b)\). Finally, for any two real valued functions \(f\) and \(g\), we write \(f = O(g)\) whenever \(\limsup_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| < \infty\). Lebesgue measure on \(\mathbb{R}\) is denoted by \(\nu\).
2. Model and Results

2.1. The model and assumptions. For a Borel measurable function $H : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, consider the following $D_T$-valued shot noise process:

$$X(t) = \sum_{i=1}^{A(t)} H(t - \tau_i, \xi_i),$$

where $A(\cdot)$ is a renewal process with arrival times $\{\tau_i\}_{i \geq 1}$ and interarrival times $\{\eta_i\}_{i \geq 1}$, i.e.,

$$A(t) = \max\{i \geq 1 : \tau_i \leq t\} \quad \text{with} \quad A(0) = 0,$$

and

$$\xi_i = g(\tau_i, \vartheta_i),$$

where $g : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a measurable and deterministic function. In what follows, we use

$$H(t, \tau_i, \vartheta_i) \equiv H(t - \tau_i, \xi_i) = H(t - \tau_i, g(\tau_i, \vartheta_i))$$

(2.3)

to be more concise, and let

$$G_1(t, s) \equiv \int_{\mathbb{R}^d} H(t, s, x) F(dx).$$

(2.4)

Assumption 2.1. The following conditions hold:

- **A.1** $\{\eta_i\}_{i \in \mathbb{N}}$ is a family of $\mathbb{R}_+$-valued i.i.d. random variables, satisfying $\mathbb{E}[e^{\rho \eta_1}] < \infty$ for some $\rho > 0$. Let $\lambda \equiv (\mathbb{E}[\eta_1])^{-1}$ and $\sigma^2 \equiv (\mathbb{E}[\eta_1^2] - (\mathbb{E}[\eta_1])^2$.

- **A.2** $\{\vartheta_i\}_{i \in \mathbb{N}}$ is a family of $\mathbb{R}^d$-valued i.i.d. random variables with c.d.f. $F$.

- **A.3** The two sequences $\{\eta_i\}_{i \in \mathbb{N}}$ and $\{\vartheta_i\}_{i \in \mathbb{N}}$ are independent.

- **A.4** $H(\cdot, x)$ is assumed to be right continuous with left limits for each $x \in \mathbb{R}^d$. $H$ is uniformly bounded, that is, $C_H \equiv \sup\{|H(t, x)| : (t, x) \in [0, T] \times \mathbb{R}^d\} < \infty$.

- **A.5** The total variation of $G_1(t, \cdot)$ is uniformly finite over $t \in [0, T]$; $\sup_{t \in [0, T]} V_T(G_1(t, \cdot)) < \infty$. Here, $V_T(h(\cdot))$ is the total variation of any function $h(\cdot)$ on $[0, T]$.

Remark 2.1. Assumptions **A.1–A.3** are very mild assumptions and are often satisfied by models in practice. The conditions in Assumption **A.1** are standard for renewal processes under which a CLT is satisfied. See further discussions on the LDP and MDP for renewal processes in Section A.1.

Remark 2.2. The distribution of the noises $\xi_i = g(\tau_i, \vartheta_i)$ depends on the arrival time $\tau_i$, and given that $\{\vartheta_i\}$ is i.i.d., the sequence $\{\xi_i\}_{i \in \mathbb{N}}$ is independent given the arrival times $\{\tau_i\}_{i \in \mathbb{N}}$, and given that $\tau_i = s$, the distribution of $\xi_i$ is

$$F_s(x) = \mathbb{P}(\xi_i \leq x|\tau_i = s) = \mathbb{P}(g(s, \vartheta_i) \leq x) = \int_{\mathbb{R}^d} 1_{g(s, y) \leq x} F(dy).$$

This is one approach to model the non-stationarity of noises. For instance, $F_s(x) = 1 - e^{-\mu(s)x}$ for a positive function $\mu(s)$ can be regarded as a non-stationary exponential distribution with rate $\mu(s)$ for $s \geq 0$ and $x \geq 0$. In this case, one can take $g(s, \vartheta_i) = (\mu(s))^{-1}\vartheta_i$ with $\vartheta_i$ being an exponential random variable with mean 1. For another instance, if $\xi_i = g(s, \vartheta_i) = \int_s^{s+\vartheta_i} \phi(u) du$ for some $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $\vartheta_i \in \mathbb{R}_+$, that is, in the context of queueing systems, the realization of the service requirement $\vartheta_i$ is through a time-varying rate $\phi$ and $\xi_i$ is the realized service time, and in this case, $F_s(x) = \int_{\mathbb{R}_+} 1\left(\int_s^{s+y} \phi(u) du \leq x\right) F(dy)$. A third example is that the functions are piecewise: $g(s, x) = g_k(x)$ for $s \in [t_k, t_{k+1}]$ given $0 = t_1 < t_2 < \cdots < t_k < t_{k+1} < \cdots < t_K = T$ and $g_k : \mathbb{R}^d \rightarrow \mathbb{R}^d$ for each $k$, so that $\int_{\mathbb{R}^d} 1_{g_k(y) \leq x} F(dy)$ for $s \in [t_k, t_{k+1}]$. 
Remark 2.3. Assumption A.4 will ensure that the process $X(t)$ is $\mathcal{D}_T$-valued. A common form of the function $H(t, x)$ is multiplicative, taking the form $H(t, x) = \tilde{H}(t)\varphi(x)$, where $\tilde{H}: \mathbb{R}_+ \to \mathbb{R}$ and $\varphi: \mathbb{R}^d \to \mathbb{R}$. The condition in Assumption A.4 requires that $\tilde{H}$ is right continuous with left limits. The function $\tilde{H}(t - \tau_i)$ in the expression of $X(t)$ represents the decay effect of the noises as time progresses. A typical example is the exponential decay, that is, $\tilde{H}(t - \tau_i) = e^{-\beta(t-\tau_i)}$ for some constant $\beta > 0$. Another one is a power decay, that is, $\tilde{H}(t - \tau_i) = (t - \tau_i)^\beta$ for some $\beta > 0$. However, the function $H(t, x)$ can be also non-multiplicative, for example, in the $G/G/\infty$ queueing model, the queue length (number of customers/jobs) process has $H(t, x) = 1_{t<\infty}$ and the workload-input process has $H(t, x) = x1_{t<\infty}$ for $x \geq 0$.

Remark 2.4. Assumption A.5 is only used in the proof of Proposition 3.2. This is a technical assumption in order to apply the contraction principle.

2.2. Sample-path MDP in the high intensity regime. In this section, we consider a scaled version of $X(t)$ in the high intensity regime and establish a sample-path MDP. Define the following scaled version of the shot noise process $X(\cdot)$ in (2.1):

$$\bar{X}^n(t) = \frac{1}{n} \sum_{i=1}^{A^n(t)} H(t, \tau^n_i, \vartheta_i) ,$$

where

$$A^n(t) \triangleq A(nt), \quad \tau^n_i \triangleq \frac{\tau_i}{n} = \frac{1}{n} \sum_{k=1}^{i} \eta_k ,$$

We remark that the scaling in the process $X^n(t)$ should be regarded as a high intensity scaling regime. The scaling of the renewal process $A^n(t) = A(nt)$ has the arrival times $\tau^n_i$ (and the interarrival times $\eta_k$) being scaled down by $n$, which is equivalent to the arrival rate $\lambda$ being scaled up by $n$, that is, the arrival rate of $A^n(t)$ is $\lambda^n = n\lambda$. It can be thus regarded either as the usual scaling of (arrival) time, or as the scaling of the intensity (arrival rate). However, since the function $H(t, \tau^n_i, \vartheta_i)$ has no scaling in $t$, one should regard the scaling of $X^n(t)$ in (2.5) as the high intensity scaling regime. See further discussions and comparison of the MDP results in the time-space scaling regime in Section 2.3 and the LDP results in the two different scaling regimes in Section 2.4.

We now define the following MDP-scaled process in the high intensity regime:

$$\bar{X}^n(t) \triangleq \frac{\sqrt{n}}{a_n} \left( X^n(t) - \lambda \int_0^t G_1(t, s) ds \right)$$

$$= \frac{\sqrt{n}}{a_n} \left( \frac{1}{n} \sum_{i=1}^{A^n(t)} H(t, \tau^n_i, \vartheta_i) - \lambda \int_0^t G_1(t, s) ds \right) ,$$

with $\{a_n\}_{n \in \mathbb{N}}$ being such that

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

and $G_1(t, s)$ is defined in (2.4).

We also introduce the following quantities: for $0 \leq s, t \leq T$,

$$G_2(t, s, u) \triangleq \int_{\mathbb{R}^d} H(t, u, x)H(s, u, x)F(dx) ,$$

$$\Lambda(s, t) \triangleq \lambda \int_0^{s\wedge t} (G_2(t, s, u) - G_1(t, u)G_1(s, u)) du .$$

We now state the main result of the paper.
Theorem 2.1. Under Assumption 2.1 and the conditions on $a_n$ in (2.8), the family of $\mathcal{D}_T$-valued random variables $\{X^n\}_{n \in \mathbb{N}}$ satisfies the following

(i) For every Borel measurable set $A$ in $(\mathcal{D}_T, J_1)$,
\[
- \inf_{\phi \in A^o} I^{MDP}(\phi) \leq \liminf_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}(X^n \in A) \leq \limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}(X^n \in A) \leq - \inf_{\phi \in \tilde{A}} I^{MDP}(\phi),
\]
where $A^o$ and $\tilde{A}$ denote the interior and closure of the measurable set $A$, respectively;

(ii) For $l \geq 0$, $\{\phi : I^{MDP}(\phi) \leq l\}$ is a compact set in $(\mathcal{D}_T, J_1)$.

Here,
\[
I^{MDP}(\phi) = \inf_{(\phi_1, \phi_2) \in \mathcal{D}_T \times \mathcal{D}_T : \phi = \phi_1 + \phi_2} \left\{I_1^{MDP}(\phi_1) + I_2^{MDP}(\phi_2)\right\},
\]
with
\[
I_1^{MDP}(\phi_1) \doteq \frac{1}{2} \int_0^T \int_0^T z(s)\Lambda(s,t)z(t)dsdt,
\]
\[
I_2^{MDP}(\phi_2) \doteq \frac{1}{2\lambda^2 \sigma^2} \inf \left\{\int_0^T |\dot{x}(t)|^2 dt\right\},
\]
where $\Lambda(s,t)$ is defined in (2.9) and $z$ is a Lebesgue measurable function on $[0,T]$ such that $\phi_1(\cdot) = \int_0^T z(s)\Lambda(\cdot, s)ds$. If no such $z$ exists, then we take $I_1^{MDP}(\phi_1) = \infty$. The infimum in (2.13) is over $x \in \mathcal{A}C_0$ such that
\[
\phi_2(t) = \int_0^t G_1(t,u)dx(u) = x(t)G_1(t,t) - \int_0^t x(u-)duG_1(t,u), \text{ for } t \in [0,T].
\]

Remark 2.5. From hereon, if the family of $\mathcal{D}_T$-valued random variable $\{Z^n\}_{n \in \mathbb{N}}$ satisfies (i) and (ii) with a function $I$ in place of $I^{MDP}$, then we say that $\{Z^n\}_{n \in \mathbb{N}}$ satisfies a moderate deviation principle (MDP) with rate function $I$ and rate $a_n^2$.

In the special case of a multiplicative shot response function taking the form $H(t,x) = \bar{H}(t)\varphi(x)$, where $\bar{H} : \mathbb{R} \to \mathbb{R}$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$, we have the functions
\[
G_1(t,s) = \bar{H}(t-s)\int_{\mathbb{R}^d} \varphi(g(s,x))F(dx),
\]
\[
G_2(t,s,u) = \bar{H}(t-u)\bar{H}(s-u)\int_{\mathbb{R}^d} \varphi(g(u,x))^2F(dx),
\]
\[
\Lambda(s,t) = \lambda \bar{H}(t-s)\bar{H}(s-u)\int_{\mathbb{R}^d} \left(\varphi(g(u,x))^2 - \varphi(g(u,x))\varphi(g(u,x))\right)F(dx).
\]

We now apply Theorem 2.1 to a GI/G1/∞ queue which has a renewal arrival process and a time-varying service time. In this model, $H : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ with $H(t,x) = \mathbb{1}_{t<x}$ (non-multiplicative) and $X(t)$ is the number of customers/jobs in the system (in service) at time $t$. As a special case of Theorem 2.1, we obtain the following sample-path MDP for the queueing process $X(t)$.

Corollary 2.1. Under Assumption 2.1, in the GI/G1/∞ queueing model, the MDP-scaled queueing process $\tilde{X}^n(t)$ satisfies a MDP in $(\mathcal{D}_T, J_1)$ with rate $a_n^2$ and rate function in (2.11), in which
\[
G_1(t,s) = 1 - F_s(t-s) \doteq F_s^c(t-s), \quad G_2(t,s,u) = 1 - F_u(t\wedge s-u) \doteq F_u^c(t\wedge s-u),
\]
and
\[
\Lambda(t,s) = \int_0^{t\wedge s} (F_u^c(t\wedge s-u) - F_u^c(t-u)F_u^c(s-u))du.
\]
Remark 2.6. Puhalskii [51] recently established sample-path MDP for many-server queues with renewal arrival processes and i.i.d. service times in the so-called Halfin-Whitt regime (the arrival rate/intensity and number of servers are scaled up with fixed service rate in such a way that the system becomes critically loaded). For that model, the MDP for the sequential empirical process associated with the service times is proved and subsequently, the MDP for a process of the form in (2.1) with the arrival process being the entering service process of customers/jobs ($\tau_i$’s being the entering service times and hence the arrival process is no longer a renewal process due to the waiting times) and the functions $H(t, x) = 1_{t<x}$ and $g(t, x) = x$. The approach in that paper is different from ours. The approach in [51] makes use of a semi-martingale decomposition of the sequential empirical process along with the techniques of exponential martingales. That approach however cannot be adapted to general shot noise processes.

Remark 2.7. We remark that the function $\Lambda(t, s)$ is in fact the covariance function of the limit process for the CLT-scaled processes

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{A_n(t)} (H(t, \tau^n_i, \vartheta_i) - G_1(t, \tau^n_i)).$$

See Theorem 2.2 in [46] for shot noise processes and Theorem 3.2 in [45] for the GI/Gt/∞ queue.

The constant $\lambda^3\sigma^2$ is the variance coefficient of the Brownian limit process of the CLT-scaled renewal process

$${\hat{A}_n(t)} = \frac{1}{\sqrt{n}}(A_n(t) - \lambda nt). \quad (2.14)$$

See Theorem 17.3 in [5]. These indicate how the rate functions in the MDP are related to the FCLT results.

2.3. Comparison with the MDP for shot noise processes in a conventional time-space scaling regime. In this section, we discuss the MDP results in the conventional time-space scaling regime. To that end, we assume that $\lim_{t \to \infty} H(t, x)$ exists, for every $x \in \mathbb{R}^d$ and we denote it by $H(\infty, x)$.

Consider the following MDP-scaled process with both time and space scalings:

$$X_n^t(\tau, \vartheta) = \sqrt{\frac{n}{\sigma_\infty}} \left( \frac{1}{n} \sum_{i=1}^{A(nt)} H(nt - \tau_i, g(\tau_i, \vartheta_i)) - \lambda \int_0^t \mathbb{E}[H(\infty, g(s, \vartheta_i))] ds \right). \quad (2.15)$$

For $u \in [0, T]$, define

$$G_1^\infty(u) = \int_{\mathbb{R}^d} H(\infty, g(u, x)) F(dx),$$

$$G_2^\infty(u) = \int_{\mathbb{R}^d} H(\infty, g(u, x))^2 F(dx).$$

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

$$I_{\infty}^{MDP}(\phi) = \begin{cases} \frac{1}{2} \int_0^T \frac{|\phi(t)|^2}{\sigma_\infty^2(t)} dt, & \text{whenever } \phi \in \mathcal{AC}_0, \\ \infty, & \text{otherwise}, \end{cases} \quad (2.16)$$

where

$$\sigma_\infty^2(u) = \lambda(G_2^\infty(u) - (G_1^\infty(u))^2) + \lambda^3\sigma^2(G_1^\infty(u))^2. \quad (2.17)$$
Remark 2.8. The proof of this theorem can be carried out by taking a similar approach as the proof of Theorem 2.1 (details are omitted for brevity). To begin with, one can show that \( \{\tilde{X}^n\}_{n \in \mathbb{N}} \) is exponentially equivalent to \( \{X^n\}_{n \in \mathbb{N}} \) defined as

\[
\tilde{X}^n(t) = \frac{\sqrt{n}}{a_n} \left( \frac{1}{n} \sum_{i=1}^{A(nt)} H(nt - \tau_i, g(\tau_i, \vartheta_i)) - \lambda \int_0^t \mathbb{E}[H(\infty, g(s, \vartheta_1))] ds \right).
\]

This process can be decomposed into two processes (asymptotically independent):

\[
\tilde{X}^n(t) = \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{A(nt)} \left( H(nt - \tau_i, g(\tau_i, \vartheta_i)) - \mathbb{E}[H(\infty, g(\tau_i, \vartheta_1))] \right)
+ \int_0^t \mathbb{E}[H(\infty, g(s, \vartheta_1))] d \left( \frac{1}{a_n \sqrt{n}} (A(ns) - \lambda n s) \right).
\]

One can then prove the MDP for both components and obtain the MDP for \( \{\tilde{X}^n\}_{n \in \mathbb{N}} \) with the rate function \( I_\infty^{MDP} : \mathcal{D}_T \to [0, \infty] \) given by

\[
I_\infty^{MDP}(\phi) = \inf_{(\phi_1, \phi_2) \in \mathcal{D}_T \times \mathcal{D}_T: \phi = \phi_1 + \phi_2} \left\{ I_{1, \infty}^{MDP}(\phi_1) + I_{2, \infty}^{MDP}(\phi_2) \right\},
\]

where

\[
I_{1, \infty}^{MDP}(\phi_1) = \frac{1}{2\lambda} \int_0^T \frac{|\hat{\phi}_1(t)|^2}{G^\infty_2(u) - (G^\infty_1(u))^2} dt,
\]

\[
I_{2, \infty}^{MDP}(\phi_2) = \frac{1}{2\lambda^3 \sigma^2} \int_0^T \frac{|\hat{\phi}_2(t)|^2}{(G^\infty_1(u))^2} dt,
\]

and the two components in the rate function correspond to the two processes in the decomposition above. Note that the infimum in (2.18) is over a convex functional of \( \phi_1 \) and \( \phi_2 \) subject to a convex constraint (viz., \( \phi = \phi_1 + \phi_2 \)). This can then be explicitly solved to obtained the desired result in the theorem.

Remark 2.9. We remark on how the rate function is related to the FCLT for the diffusion-scaled process

\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{A(nt)} H(nt - \tau_i, \xi_i) - \lambda \int_0^t \mathbb{E}[H(\infty, g(s, \vartheta_1))] ds \right).
\]

Observe that it can be decomposed into two components:

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{A(nt)} \left( H(nt - \tau_i, \xi_i) - \mathbb{E}[H(\infty, g(\tau_i, \vartheta_1))] \right) + \int_0^t \mathbb{E}[H(\infty, g(s, \vartheta_1))] d\hat{A}^n(s)
\]

where \( \hat{A}^n(t) \) is defined in (2.14). It can be shown (for example, modifying the proofs in [29, 35]) that the diffusion-scaled processes converge in \( (\mathcal{D}_T, J_1) \) to

\[
B_1 \left( \int_0^t \lambda \text{Var}(H(\infty, g(s, \vartheta_1))) ds \right) + B_2 \left( \int_0^t \lambda^3 \sigma^2 \left( \mathbb{E}[H(\infty, g(s, \vartheta_1))] \right)^2 ds \right)
\]

where \( B_1 \) and \( B_2 \) are two independent Brownian motions. Also, observe that the sum of the two independent Brownian limits is equivalent in distribution to a Brownian motion with the variance function \( \int_0^t \sigma^2_\infty(u) du \), for \( \sigma^2_\infty(u) \) given in (2.17). The function \( \sigma^2_\infty(\cdot) \) is also exactly what appears in the rate function in (2.16).
In the special case of i.i.d. noises with \( g(t, x) = x \), the variance coefficient of the Brownian limit reduces to
\[
\lambda \text{Var}(H(\infty, \vartheta_1)) + \lambda^3 \sigma^2(E[H(\infty, \vartheta_1)])^2.
\]
This is the asymptotic variance of the compound renewal process \( \sum_{i=1}^{A(t)} H(\infty, \vartheta_i) \) as \( t \to \infty \).

2.4. Comparison with the LDPs for shot noise processes in the two scaling regimes. In this section, we discuss the differences of the MDP results above with the LDP results in the two scaling regimes. We assume here that \( \mathbb{P}(\eta_1 = 0) = 0 \). This assumption enables us to use the LDP result for renewal processes in Theorem A.2.

2.4.1. LDP in the high intensity regime. Recall that \( X^n \) in (2.5). The family of \( \mathcal{D}_T \)-valued random variables \( \{X^n\}_{n \in \mathbb{N}} \) can be shown to satisfy the following

(i) For every Borel measurable set \( A \) in \( (\mathcal{D}_T, J_1) \),
\[
- \inf_{\phi \in A^o} I^{\text{LDP}}[X^n](\phi) \leq \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(X^n \in A)
\]
\[
\leq \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}(X^n \in A) \leq - \inf_{\phi \in \bar{A}} I^{\text{LDP}}[X^n](\phi),
\]

where \( A^o \) and \( \bar{A} \) denote the interior and closure of the measurable set \( A \), respectively;
(ii) For \( l \geq 0 \), \( \{\phi : I^{\text{LDP}}[X^n](\phi) \leq l \} \) is a compact set in \( (\mathcal{D}_T, J_1) \).

Here,
\[
I^{\text{LDP}}[X^n](\phi) = \left\{ \begin{array}{ll}
\sup_{\rho \in \mathcal{C}_T} \int_0^T \left( \dot{\phi}(t) \int_t^T \rho(u)du - \Psi_A \left( \log \mathbb{E} \left[ \exp \left( \int_t^T \rho(s)H(s, t, \vartheta_1)ds \right) \right] \right) \right) dt, \\
0, \text{ otherwise.}
\end{array} \right.
\]

(2.19)

Here \( \mathcal{C}_T \) denotes the set of continuous functions on \( [0, T] \), and \( \Psi_A \) is defined as
\[
\Psi_A(\rho) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\rho A^n(1)}].
\]

From [25, Theorem 1], we have
\[
\Psi_A(\rho) = -\psi^{-1}_\eta(-\rho) \quad \text{with} \quad \psi_\eta(\rho) = \log \mathbb{E}[e^{\rho \eta^n}].
\]

(2.20)

Remark 2.10. In what follows, whenever a family of \( \mathcal{D}_T \)-valued random variables \( \{Z^n\}_{n \in \mathbb{N}} \) satisfy above conditions (i) and (ii) with function \( I \) in place of \( I^{\text{LDP}}[X^n] \), we say that \( \{Z^n\}_{n \in \mathbb{N}} \) satisfies a large deviation principle with rate function \( I \) and rate \( n \).

In the special case of \( GI/GI/\infty \) queueing model, with \( \xi_i = g(\tau_i, \vartheta_i) \) and \( H(t, x) = \mathbb{1}_{t < x} \), the rate function in the LDP becomes
\[
I^{\text{LDP}}[X^n](\phi) = \left\{ \begin{array}{ll}
\sup_{\rho \in \mathcal{C}_T} \int_0^T \left( \dot{\phi}(t) \int_t^T \rho(u)du - \Psi_A \left( \log \mathbb{E} \left[ \exp \left( \int_t^T \rho(s)H(s, t, \vartheta_1)ds \right) \right] \right) \right) dt, \\
0, \text{ otherwise.}
\end{array} \right.
\]

(2.21)

Note that in [24], the LDP for the \( GI/GI/\infty \) queue is established for the queueing process \( X(t) \) at each fixed time \( t = \hat{t} \geq 0 \). (That result can be obtained by setting \( \rho(\cdot) = \hat{\rho} \delta_{\hat{t}}(\cdot) \) for some constant \( \hat{\rho} \).
Our result extends that to a model with non-stationary service times and the LDP result is in the sample-path sense (see also [6] for the sample-path LDP is established for a two-parameter process tracking the elapsed/residual service times in \( GI/GI/\infty \) queues).
We also remark that the proof of the sample-path LDP for the scaled shot noise process $\bar{X}^n(t)$ differs from that of the MDP for the processes $\bar{X}^n(t)$ in (2.7). The main difference between LDP and MDP is similar to that of LLN and CLT viz., centering is necessary for MDP and CLT. Although one can follow a similar approach in Section 3, the proof of finite-dimensional LDP would require appropriate limits of log moment generating of $A^n(\cdot)$ and $\vartheta_1$, and the proof of exponential tightness could be modified appropriately without the centering terms. The details are omitted for brevity.

2.4.2. LDP in the conventional time-space scaling regime. Next, for the shot noise process $X(t)$ in (2.1), define the following scaled process

$$\bar{X}^n(t) = \frac{1}{n} \sum_{i=1}^{A(nt)} H(nt - \tau_i, \xi_i).$$

(2.22)

Here, we again assume that $\xi_i$ is of the form $g(\tau_i, \vartheta_i)$, for a family of i.i.d random variables $\{\vartheta_i\}_{i \in \mathbb{N}}$ and that $H(t,x) \to H(\infty, x)$ as $t \to \infty$, for every $x$. The LDP result gives the rate function:

$$I_{LDP}[\bar{X}^n](\phi) = \begin{cases} 
\sup_{\rho \in C_T} \int_0^T (\dot{\phi}(t)\rho(t) - \Psi_A \left( \log \mathbb{E} \left[ \exp \left( \rho(t)H(\infty, t, \vartheta_1) \right) \right] \right)) \, dt, & \text{if } \phi \in AC_0, \\
0, & \text{otherwise.}
\end{cases}$$

(2.23)

This can be established by using a similar argument discussed above (omitted for brevity).

In the case of i.i.d. noises with $g(t, x) = x$, the LDP is like that for the compound renewal process as studied in Borovkov [8]. However, with the general function $g(t, x)$, the LDP is an extension to compound renewal processes with non-stationary compound variables (arrival time dependent distributions).

In addition, in the special case of Poisson arrivals and i.i.d. noises, the LDP result coincides with that in [22] and also in [15] when their model does not have state-dependent component in $H$. In this case, the rate function is

$$I_{LDP}[\bar{X}^n](\phi) = \begin{cases} 
\int_0^T \Lambda^* (\dot{\phi}(t)) \, dt, & \text{if } \phi \in AC_0, \\
\infty, & \text{otherwise,}
\end{cases}$$

(2.24)

where

$$\Lambda_{H(\infty, \xi)}(\theta) = \log \mathbb{E}[\exp(\theta H(\infty, \xi_1))], \quad \Lambda(\theta) = \lambda \left( \exp \left( \Lambda_{H(\infty, \xi)}(\theta) \right) - 1 \right)$$

and $\Lambda^*(x)$ is the Legendre transform of $\Lambda$, that is,

$$\Lambda^*(x) = \sup_{\theta \in \mathbb{R}} \{ \theta x - \Lambda(\theta) \}.$$ 

One can easily see how the rate function in (2.23) reduces to that in (2.24) by noting $\Psi_A(\rho) = \lambda(e^\rho - 1)$ for Poisson arrival process $A(t)$. However, even with Poisson arrivals, for non-stationary noises, one cannot simplify the rate function in (2.23) except using $\Psi_A(\rho) = \lambda(e^\rho - 1)$.

3. Proof of Theorem 2.1

In this section we prove the sample-path MDP for $\bar{X}^n(t)$ in (2.7) as stated in Theorem 2.1. First, we observe that the process $\bar{X}^n$ can be decomposed into two processes:

$$\bar{X}^n(t) = \bar{X}^n_1(t) + \bar{X}^n_2(t),$$

(3.1)

where

$$\bar{X}^n_1(t) = \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{A_n(t)} \bar{H}(t, \tau^n_i, \vartheta_i),$$

(3.2)
\[ \widetilde{X}_2^n(t) = \int_0^t G_1(t, s) d\widetilde{A}^n(s), \]  
with
\[ \widetilde{A}^n(t) = \frac{1}{a_n \sqrt{n}} (A^n(t) - \lambda nt), \]
and
\[ \widetilde{H}(t, s, x) = H(t, s, x) - G_1(t, s). \]

Before proceeding to the proof, we provide an overview of the proof strategy. From the statement of the rate function in Theorem 2.1, as a sum of the two rate functions associated with the rate functions of \( \widetilde{X}_1^n(t) \) and \( \widetilde{X}_2^n(t) \), it may appear that the rate function comes from two independent processes. However, the two processes \( \widetilde{X}_1^n(t) \) and \( \widetilde{X}_2^n(t) \) are not independent since both depend on the arrival process \( A^n(t) \). On the other hand, the two processes \( \widetilde{X}_1^n(t) \) and \( \widetilde{X}_2^n(t) \) are asymptotically independent as \( n \to \infty \), in the sense that the variability of \( \widetilde{X}_1^n(t) \) (as given in \( \Lambda(t, s) \) in (2.9)) only depends on the arrival rate \( \lambda \) of the renewal process \( A^n(t) \), and comes rather from the variability of the noises, while the variability of \( \widetilde{X}_2^n(t) \) depends on the variability of the interarrival times of renewal process \( A^n(t) \) (as in the associated FCLT). This is also the case in the FCLT result for the shot noise processes in the high intensity regime [46]. Therefore, we construct another process \( \{\widetilde{X}_1^n\} \) (see (3.6) below), which is exponentially equivalent (see Lemma 3.1) to \( \{\widetilde{X}_2^n\} \), and more importantly, independent of \( \{\widetilde{X}_2^n\} \). Thus, by [18, Theorem 4.2.13], instead of establishing the MDP of \( \{\{\widetilde{X}_1^n, \widetilde{X}_2^n\}\}_{n \in \mathbb{N}} \), it then suffices to establish the MDPs of \( \{\widetilde{X}_1^n\}_{n \in \mathbb{N}} \) and \( \{\widetilde{X}_2^n\} \) separately.

The proof of the sample-path MDP of \( \{\widetilde{X}_1^n\} \) in \( D_T \) uses Gärtner–Ellis theorem [18, Theorem 2.3.6] and Dawson–Gärtner theorem [18, Theorem 4.6.1]. We first establish the MDP in the topology of pointwise convergence and then strengthen it to be in the Skorohod \( J_1 \) topology. To be more elaborate, we first prove the MDP for the finite dimensional distributions (see Lemma 3.2) by considering appropriate limits of the log-moment generating function (see Proposition 3.1). Using this, we arrive at the sample-path MDP of \( \widetilde{X}_1^n \) in topology of pointwise convergence by invoking Dawson–Gärtner theorem. We next establish exponential tightness to arrive at the desired MDP in \( (D_T, J_1) \). The MDP of \( \{\widetilde{X}_2^n\}_{n \in \mathbb{N}} \) is derived using the contraction principle and the existing MDP for renewal processes.

### 3.1. Exponential equivalence of \( \widetilde{X}_1^n \) and \( \widetilde{X}_2^n \). Let

\[ \widetilde{X}_1^n(t) = \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} \widetilde{H}(t, i \lambda n, \vartheta_i), \quad \text{for } t \in [0, T]. \]  

where \( \widetilde{H}(t, s, x) \) is defined in (3.5). To be concise, we write \( s_i^n = \frac{i \lambda n}{a_n} \) in the following. Note that comparing with \( \widetilde{X}_1^n \), in the definition of \( \widetilde{X}_1^n \), we have replaced the arrival times \( \tau_i^n = \frac{i}{a_n} \) by \( s_i^n = \frac{i \lambda n}{a_n} \) and \( A^n(t) \) by \( \lfloor \lambda nt \rfloor \), and thus removed the randomness from the arrival process \( A^n \). We prove that \( \widetilde{X}_1^n \) is exponentially equivalent to \( \widetilde{X}_1^n \) in \( (D_T, J_1) \).

**Lemma 3.1.** Under Assumptions A.1–A.4, for any \( \delta > 0 \),

\[ \lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \| \widetilde{X}_1^n - \widetilde{X}_2^n \|_{L_2} > \delta \right) = -\infty. \]

**Remark 3.1.** Using Lemma 3.1 (which is referred to exponential equivalence in [18, Pg. 130]), we can conclude the LDP of \( \{\widetilde{X}_1^n\}_{n \in \mathbb{N}} \) in \( (D_T, J_1) \) from the LDP of \( \{\widetilde{X}_2^n\}_{n \in \mathbb{N}} \) in \( (D_T, J_1) \).

**Sketch of the proof.** Firstly from Theorem A.1, it is evident that for large \( n \), \( A^n(t) \) satisfies the following with large probability: for any \( k \geq 0 \), there is \( n(k) \) such that for \( n \geq n(k) \),

\[ |\lambda nt| - |ka_n \sqrt{n}| \leq A^n(t)(\omega) \leq |\lambda nt| + |ka_n \sqrt{n}|, \quad \forall t \in [0, T]. \]  

(P)
This suggests that we split the probability
\[ P \left( \| \tilde{X}_1^n - \tilde{X}_1^n \|_{\mathcal{A}} > \delta \right) \]
into probabilities over two sets viz., over a set where the above property (P) holds (say \( P_1^n(k) \)) and over a set where the above property (P) does not hold (say \( P_2^n(k) \)).

We will show that \( P_2^n(k) \) can be ignored for large \( n \) and hence, the only relevant term for large \( n \) is \( P_1^n(k) \). We then go on to show that \( \{ \tau_i^n \}_{i \in \mathbb{N}} \) which are the corresponding arrival times for \( A^n \) has a following property: whenever \( A^n \) satisfies (P), we have
\[
\frac{i}{\lambda_n} - 2 \left\lfloor \frac{ka_n \sqrt{n}}{\lambda n} \right\rfloor \leq \tau_i^n \leq \frac{i}{\lambda_n} + 2 \left\lfloor \frac{ka_n \sqrt{n}}{\lambda n} \right\rfloor. \tag{3.7}
\]

Define the following sets:
\[
\mathcal{W}(k) = \left\{ \omega : \Lambda^{-}(t, k) < A^n(t)(\omega) < \Lambda^{+}(t, k), \text{ for } n \geq n(k) \text{ and } t \in [0, T] \right\} \tag{3.8}
\]
where \( \Lambda^{\pm}(s, k) = \lfloor \lambda sn \pm ka_n \sqrt{n} \rfloor \). As mentioned already, the reason behind defining and considering the above sets is as follows: from Theorem A.1, it is clear that for any \( k \geq 0 \), the event
\[
\left\{ \omega : \frac{\sqrt{n}}{a_n} \left| \frac{1}{n} A^n(t) - \lambda t \right| \geq k \right\}
\]
occurs with probability of the order of \( e^{-k a_n^2} \), for large \( n \). Therefore, defining the set \( \mathcal{W}(k) \) as above allows us to study events that are very probable. Note that on \( \mathcal{W}(k) \), (3.7) holds.

**Proof.** Fix \( \delta > 0 \) and consider the event \( \mathcal{W}(k) \) in (3.8). On \( \mathcal{W}(k) \),
\[
\tilde{X}_1^n(t) - \tilde{X}_1^n(t) = \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{A^n(t)} \tilde{H}(t, \tau_i^n, \vartheta_i) - \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} \tilde{H}(t, s_i^n, \vartheta_i) \\
= \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} \left( \tilde{H}(t, \tau_i^n, \vartheta_i) - \tilde{H}(t, s_i^n, \vartheta_i) \right) + \frac{1}{a_n \sqrt{n}} \sum_{j=\lfloor \lambda nt \rfloor + 1}^{A^n(t)} \tilde{H}(t, \tau_j^n, \vartheta_j). 
\]

It is clear that on \( \mathcal{W}(k) \),
\[
\frac{1}{a_n \sqrt{n}} \sum_{j=\lfloor \lambda nt \rfloor + 1}^{A^n(t)} \tilde{H}(t, \tau_j^n, \vartheta_j) \leq 2C_H k \frac{a_n \sqrt{n}}{a_n \sqrt{n}} \leq 2C_H k.
\]

Let \( J_{i,t} = \tilde{H}(t, \tau_i^n, \vartheta_i) - \tilde{H}(t, s_i^n, \vartheta_i) \). We next estimate the probability of event
\[
\left\{ \sup_{0 \leq t \leq T} \left| \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} J_{i,t} \right| > \frac{\delta}{2} \right\},
\]
and will apply a modified version of Corollary A.1 and prove that
\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log P \left( \left\{ \left\| \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} J_{i,t} \right\|_T > \frac{\delta}{2} \right\} \cap \mathcal{W}(k) \right) \leq \frac{-\rho \delta}{8}. \tag{3.9}
\]

From now on, we write \( P_k(A) \) for \( P(\mathcal{A} \cap \mathcal{W}(k)) \) and \( E_k[\cdot] \) for \( E[\cdot | \mathcal{W}(k)] \), for short. We use a modified version of Corollary A.1 for \( P_k \) and \( E_k \) to get
\[
\frac{1}{a_n^2} \log P_k \left( \sup_{0 \leq t \leq T} \left| \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nt \rfloor} J_{i,t} \right| > \frac{\delta}{2} \right)
\]
\[
\frac{\rho^2}{2n} \sum_{i=1}^{\lfloor \lambda nT \rfloor} \mathbb{E}_k [J_i, t^2] \leq \frac{\rho \delta}{8} + C_H^2 |\rho| |\beta (\frac{a_n}{\sqrt{n}})|.
\] (3.10)

To arrive at the third term in the final inequality, we used the fact that the expectation of cube of the sum of pairwise independent random variables with zero mean is equal to the sum of expectation of the cubes of individual random variables. (Since we use this fact multiple times, we omit mentioning it from hereon.)

We now compute
\[
\sum_{i=1}^{\lfloor \lambda nT \rfloor} \mathbb{E}_k [J_i, t^2] = \sum_{i=1}^{\lfloor \lambda nT \rfloor} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, \vartheta_i) - \tilde{H}(t, \frac{i}{\lambda n}, \vartheta_i) \right)^2 \right]
\]
\[
= \sum_{i=1}^{\lfloor \lambda nT \rfloor} \int_{\mathbb{R}^d} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, x) - \tilde{H}(t, \frac{i}{\lambda n}, x) \right)^2 \right] F(dx)
\]
\[
= \sum_{i=1}^{\lfloor \lambda nT \rfloor} \int_{\mathbb{R}^d} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, x) - \tilde{H}(t, \frac{i}{\lambda n}, x) \right)^2 \right] F(dx)
\]
\[
\leq \lfloor \lambda nT \rfloor \sup_{1 \leq i \leq \lfloor \lambda nT \rfloor} \int_{\mathbb{R}^d} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, x) - \tilde{H}(t, \frac{i}{\lambda n}, x) \right)^2 \right] F(dx).
\] (3.11)

This implies that
\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}_k \left( \| \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda nT \rfloor} J_i, t \|_T > \delta \right) \leq \frac{\lambda T \rho^2}{2} \limsup_{n \to \infty} \sup_{1 \leq i \leq \lfloor \lambda nT \rfloor} \int_{\mathbb{R}^d} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, x) - \tilde{H}(t, \frac{i}{\lambda n}, x) \right)^2 \right] F(dx) - \frac{\rho \delta}{8}.
\]

Observe that
\[
\limsup_{n \to \infty} \int_{\mathbb{R}^d} \mathbb{E}_k \left[ \left( \tilde{H}(t, \tau^n_i, x) - \tilde{H}(t, \frac{i}{\lambda n}, x) \right)^2 \right] F(dx) = 0
\]
from dominated convergence theorem and the fact that \( |\tau^n_i - s^n_i| = |\tau^n_i - \frac{i}{\lambda n}| \to 0 \) as \( n \to \infty \) on the event \( \mathcal{W}(k) \). Therefore, we have shown (3.9).

Now consider the event \( \mathcal{W}(k)^c \) which is a closed set. From Theorem A.1 and the definition of MDP (as \( \{A^n\}_{n \in \mathbb{N}} \) satisfies an MDP from Theorem A.1), we have
\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \| \bar{X}_1^n - \bar{X}_1^n \|_T > \delta \right) \cap \mathcal{W}(k)^c \leq \limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \mathcal{W}(k)^c \right)
\]
\[
\leq - \inf_{x \in \mathcal{W}(k)^c} I_A^{\text{MDP}}(x).
\]

Finally, we have
\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \| \bar{X}_1^n - \bar{X}_1^n \|_T > \delta \right) \leq \max \left\{ \limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \| \bar{X}_1^n - \bar{X}_1^n \|_T > \delta \right) \cap \mathcal{W}(k)^c, \right. \]
\[
\left. \limsup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \| \bar{X}_1^n - \bar{X}_1^n \|_T > \delta \right) \cap \mathcal{W}(k) \right\}
\]
\[
\leq \max \left\{ - \frac{\rho \delta}{8}, - \inf_{x \in \mathcal{W}(k)^c} I_A^{\text{MDP}}(x) \right\}.
\]

Taking \( \rho \uparrow \infty \) and then taking \( k \uparrow \infty \), we have the result. \( \square \)
3.2. MDP for the finite dimensional distributions of $\hat{X}^n_t$. We first prove the following property on the exponential moments of the finite dimensional distributions of $\hat{X}^n_t$.

**Proposition 3.1.** Under Assumptions A.1–A.4, for $N \geq 1$ and $0 < t_1 < t_2 < t_3 < \ldots < t_N \leq T$, we have

$$
\lim_{n \to \infty} \frac{1}{a^n} \log \mathbb{E} \left[ \exp \left( a^n \sum_{n}^{N} \rho \hat{X}^n_t(t_m) \right) \right] = \lambda \left( \frac{1}{2} \sum_{i,j=1}^{N} \rho_i \rho_j \Lambda(t_i, t_j) \right) \tag{3.12}
$$

for every $\{\rho \}_{n=1}^{N} \subseteq \mathbb{R}$, where $\Lambda(\cdot, \cdot)$ is defined in (2.9). In particular, for $t \in [0, T]$ and $\rho \in \mathbb{R}$,

$$
\lim_{n \to \infty} \frac{1}{a^n} \log \mathbb{E} \left[ \exp \left( a^n \rho \hat{X}^n_t(t) \right) \right] = \frac{1}{2} \rho^2 \lambda \Lambda(t, t).
$$

**Proof.** Recall the definition of $\hat{X}^n_t$:

$$
\hat{X}^n(t) = \frac{1}{a^n} \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i), \text{ for } t \in [0, T],
$$

where $\{\vartheta_i\}_{i \in \mathbb{N}}$ is a family of $\mathbb{R}^d$–valued i.i.d. random variables distributed according to $F$. So from Assumptions A.1 and A.4, we infer that $\mathbb{E} \left[ \exp \left( a^n \rho \hat{X}^n_t(t) \right) \right]$ is finite for every $\rho \in \mathbb{R}$, $t \in [0, T]$ and $n \in \mathbb{N}$.

Therefore, the following series expansion holds:

$$
\mathbb{E} \left[ \exp \left( a^n \rho \hat{X}^n_t(t) \right) \right] = \mathbb{E} \left[ \exp \left( \frac{a^n}{\sqrt{n}} \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i) \right) \right] \leq \sum_{k=0}^{\infty} \frac{\rho^k}{k!} \left( \frac{a^n}{\sqrt{n}} \right)^k \mathbb{E} \left[ \left( \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i) \right)^k \right].
$$

Taking the logarithm on both sides of the above and then using the Taylor’s series of the function $\log(1 + x)$, we get

$$
\frac{1}{a^n} \log \mathbb{E} \left[ \exp \left( a^n \rho \hat{X}^n_t(t) \right) \right] = \frac{1}{a^n} \mathbb{E} \left[ \left( \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i) \right)^2 \right] + C_H^3 |\rho|^3 O \left( \frac{a^n}{\sqrt{n}} \right). \tag{3.13}
$$

Note that the linear term in the above equation is zero as $\mathbb{E}[\tilde{H}(t, s^n_i, \vartheta_i)] = 0$. As $\frac{a^n}{\sqrt{n}} \to 0$ as $n \to \infty$, we can neglect the error term. We have

$$
\mathbb{E} \left[ \left( \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i) \right)^2 \right] = \sum_{i=1}^{\lfloor \lambda n t \rfloor} \sum_{j=1}^{\lfloor \lambda n t \rfloor} \mathbb{E} \left[ \tilde{H}(t, s^n_i, \vartheta_i) \tilde{H}(t, s^n_j, \vartheta_j) \right] = \sum_{i=1}^{\lfloor \lambda n t \rfloor} \sum_{j=1}^{\lfloor \lambda n t \rfloor} \mathbb{E} \left[ \tilde{H}(t, s^n_i, \vartheta_i) \tilde{H}(t, s^n_j, \vartheta_j) \right] = \sum_{i=1}^{\lfloor \lambda n t \rfloor} \mathbb{E} \left[ \tilde{H}(t, s^n_i, \vartheta_i)^2 \right], \tag{3.14}
$$

where we used the fact that $\vartheta_i$’s are mutually independent for any $i = 1, \ldots, \lfloor \lambda n t \rfloor$, and that $\mathbb{E}[\tilde{H}(t, s^n_i, \vartheta_i)] = 0$ for each $i$. Now, we have

$$
\mathbb{E} \left[ \left( \sum_{i=1}^{\lfloor \lambda n t \rfloor} \tilde{H}(t, s^n_i, \vartheta_i) \right)^2 \right] = \sum_{i=1}^{\lfloor \lambda n t \rfloor} \int_{\mathbb{R}^d} \tilde{H}(t, s^n_i, x)^2 F(dx).
$$
Now from (3.13), taking \( n \to \infty \), we obtain
\[
\lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{E} \left[ \exp \left( \frac{a_n^2 \rho \bar{X}_n^i(t)}{} \right) \right] = \lim_{n \to \infty} \frac{1}{n} \frac{\rho^2}{2} \int_{\mathbb{R}^d} \sum_{i=1}^{\lfloor \lambda n t \rfloor} \bar{H}(t, s^n_i, x)^2 F(dx)
\]
\[
= \frac{\rho^2 \lambda}{2} \int_{\mathbb{R}^d} \int_0^t \bar{H}(t, s, x)^2 ds F(dx)
\]
\[
= \frac{\rho^2 \lambda}{2} \Lambda(t, t).
\]

Here the second equality follows from the definition of Lebesgue integral and measurability of \( \bar{H}(t, s, x) \) in \( s \), and the last equation follows from the definition of \( \Lambda(\cdot, \cdot) \).

To prove (3.12) in the finite-dimensional case, we use a similar argument. From Assumptions A.1 and A.4, and Taylor’s series of \( \exp(x) \), we have
\[
\mathbb{E} \left[ \exp \left( \frac{a_n}{\sqrt{n}} \sum_{m=1}^{N} \rho_m \sum_{i=1}^{\lfloor \lambda n t_m \rfloor} \bar{H}(t_m, s^n_i, \vartheta_i) \right) \right] = 1 + \mathbb{E} \left[ \left( \frac{a_n}{\sqrt{n}} \sum_{m=1}^{N} \rho_m \sum_{i=1}^{\lfloor \lambda n t_m \rfloor} \bar{H}(t_m, s^n_i, \vartheta_i) \right)^2 \right] + C_H^2 \| \rho \|^2 \mathbb{O}(\frac{a_n}{\sqrt{n}}),
\]
\[
= 1 + \frac{a_n^2}{n} \sum_{k,m=1}^{N} \rho_m \rho_k \mathbb{E} \left[ \sum_{i=1}^{\lfloor \lambda n t_k \rfloor} \sum_{j=1}^{\lfloor \lambda n t_k \rfloor} \bar{H}(t_m, s^n_i, \vartheta_i) \bar{H}(t_k, s^n_j, \vartheta_j) \right] + C_H^2 \| \rho \|^2 \mathbb{O}(\frac{a_n}{\sqrt{n}}),
\]
\[
= 1 + \frac{a_n^2}{n} \sum_{k,m=1}^{N} \rho_m \rho_k \sum_{i=1}^{\lfloor \lambda n t_k \rfloor} (G_2(t_m, t_k, s^n_i) - G_1(t_m, s^n_i) G_1(t_k, s^n_i)) + C_H^2 \| \rho \|^2 \mathbb{O}(\frac{a_n}{\sqrt{n}}).
\]

Note that we have truncated the Taylor’s series at the third term. Using the fact that \( \mathbb{E}[\bar{H}(t_m, s^n_i, \vartheta_i)] = 0 \) and \( \{ \vartheta_i \} \) is an i.i.d. sequence, we get \( \frac{a_n}{\sqrt{n}} \). Noting that the above sum over \( i \) is an approximation of a Lebesgue integral and taking the limit as \( n \to \infty \) (in other words, as the sum over \( i \) approaches the corresponding integral), we get
\[
\lim_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{E} \left[ \exp \left( \frac{a_n}{\sqrt{n}} \sum_{m=1}^{N} \rho_m \sum_{i=1}^{\lfloor \lambda n t_m \rfloor} \bar{H}(t_m, s^n_i, \vartheta_i) \right) \right] = \lim_{n \to \infty} \frac{1}{n} \sum_{k,m=1}^{N} \rho_m \rho_k \sum_{i=1}^{\lfloor \lambda n t_k \rfloor} (G_2(t_m, t_k, s^n_i) - G_1(t_m, s^n_i) G_1(t_k, s^n_i))
\]
\[
= \frac{\lambda}{2} \sum_{k,m=1}^{N} \rho_m \rho_k \Lambda(t_m, t_k), \text{ from the definition of } \Lambda(\cdot, \cdot).
\]

This completes the proof of the lemma. \( \square \)

The following lemma gives us the MDP of \( \{ \tilde{X}_{1,N}^n \} = \{ \tilde{X}_{1}(t_1), \tilde{X}_{1}(t_2), \tilde{X}_{1}(t_3), \ldots, \tilde{X}_{1}(T) \} \), for every \( N \in \mathbb{N} \) and \( 0 < t_1 < t_2 < t_3 < \ldots < t_N \leq T \). In the following, we write \((v_1, v_2, \ldots, v_N) \in \mathbb{R}^N \) as \( \mathbf{v} \).

**Lemma 3.2.** Suppose Assumptions A.1–A.4 hold. Then the family of \( \mathbb{R}^N \)-valued random variables \( \{ \tilde{X}_{1,N}^n \} \) satisfies the following:
(i) For every Borel measurable set $A$ in $\mathbb{R}^N$,
\[
- \inf_{x \in A^o} I_f^N(x) \leq \liminf_{n \to \infty} \frac{1}{a_n} \log \mathbb{P}(\hat{X}_1^n \in A) \leq \limsup_{n \to \infty} \frac{1}{a_n} \log \mathbb{P}(\hat{X}_1^n \in A) \leq - \inf_{x \in \bar{A}} I_f^N(x),
\] (3.15)
where $A^o$ and $\bar{A}$ denote the interior and closure of the measurable set $A$, respectively;
(ii) For $l \geq 0$, $\{x : I_f^N(x) \leq l\}$ is a compact set in $\mathbb{R}^N$.

Here, $I_f^N : \mathbb{R}^N \to [0, \infty]$ given by
\[
I_f^N(x) = \frac{1}{2} \langle x, \hat{\Lambda}^{-1} x \rangle,
\] (3.16)
where $\{\hat{\Lambda}_{ij}\} = \{\Lambda(t_i, t_j)\}$ for $1 \leq i, j \leq N$ and $\Lambda(\cdot, \cdot)$ in (2.9).

Proof. From Proposition 3.1, it is clear that
\[
\chi(\rho) \doteq \lim_{n \to \infty} \frac{1}{a_n} \log \mathbb{E} \left[ \exp \left( a_n^2 \sum_{m=1}^{N} \rho \hat{X}_1^n(t_m) \right) \right] = \langle \rho, \hat{\Lambda} \rho \rangle
\]
for every $\rho \in \mathbb{R}^N$. The Fenchel-Legendre transform of $\chi(\cdot)$ is given by
\[
\chi^*(\bar{x}) \doteq \sup_{\rho \in \mathbb{R}^N} \left( \langle \bar{x}, \rho \rangle - \chi(\rho) \right) = \frac{1}{2} \langle \bar{x}, \hat{\Lambda}^{-1} \bar{x} \rangle.
\] (3.17)
The second equality follows from a simple calculation. Now applying [18, Theorem 2.3.6], we have the result with $I_f^N = \chi^*$ after noting that statement in (ii) is clearly true. \hfill \Box

We next extend the finite-dimensional MDP result to the sample-path MDP of $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ in $\mathcal{D}_T$. But the form of $I_f^N$ makes it difficult in taking $N \to \infty$. To overcome this difficulty, we simplify (3.16) in the following way: Instead of solving the optimization problem
\[
\sup_{\rho \in \mathbb{R}^N} (\langle \bar{x}, \rho \rangle - \chi(\rho))
\]
extplicitly in $x$ using calculus (say) which can be difficult to perform in the case of infinite dimensions, we solve it using the property of Euclidean inner product $\langle \cdot, \cdot \rangle$. The advantage in doing so is that this method works even when $\langle \cdot, \cdot \rangle$ denotes a general inner product. With this observation in mind, we first note that $\hat{\Lambda}$ is positive definite. Therefore, for every $\bar{x}$, there is a unique $\nu \in \mathbb{R}^N$ such that $\bar{x} = \hat{\Lambda}^{-1} \nu$. We now solve the aforementioned optimization problem using just the properties of Euclidean inner product $\langle \cdot, \cdot \rangle$.
\[
\langle \bar{x}, \rho \rangle - \frac{1}{2} \langle \rho, \hat{\Lambda} \rho \rangle = \langle \bar{x}, \nu \rangle - \frac{1}{2} \langle \rho, \hat{\Lambda} \nu \rangle = \langle \bar{x}, \hat{\Lambda} \nu \rangle - \frac{1}{2} \langle \rho - \nu, \hat{\Lambda} (\rho - \nu) \rangle,
\]
From above, it is clear that the supremum in (3.17) occurs when $\rho = \nu$ and the value is
\[
\frac{1}{2} \langle \bar{x}, \hat{\Lambda} \nu \rangle.
\]
Clearly, as mentioned above, this method is robust enough to be applied for an infinite dimensional Hilbert space. This is the motivation for the construction of a reproducing kernel Hilbert space in the proof of the Theorem 3.1 below.

We are now in a position to state the sample-path MDP of $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ in $\mathcal{D}_T$ with the topology of pointwise convergence. Recall that $\{f_n\}_{n \in \mathbb{N}} \subset \mathcal{D}_T$ converges to $f \in \mathcal{D}_T$ as $n \to \infty$ in this topology if $f_n(t) \to f(t)$, as $n \to \infty$, for $t \in [0, T]$. 
**Theorem 3.1.** Suppose Assumptions A.1–A.4 hold. Then the family of $D_T$–valued random variables $\{X^n_1\}_{n \in \mathbb{N}}$ satisfies an MDP in the topology of pointwise convergence with rate $a_n^2$ and rate function $I_{MDP}^1 : D_T \to [0, \infty]$ given by (2.12).

**Proof.** Using Lemma 3.2 and [18, Theorem 4.6.1], we have the MDP of $\{X^n_1\}_{n \in \mathbb{N}}$ with rate $a_n^2$ and rate function $I_{MDP}^1 : D_T \to [0, \infty]$ given by

$$I_{MDP}^1(x) = \sup_{0 < t_1 < t_2 < \ldots < t_N \leq T} I_N^1(x(t_1), x(t_2), \ldots, x(t_N)) = \frac{1}{2} \sup_{0 < t_1 < t_2 < \ldots < t_N \leq T} \sum_{i,j=1}^N v_i^N \Lambda(t_i, t_j)v_j^N,$$

(3.18)

where $v_j^N$'s are such that $x(t_i) = \sum_{j=1}^N \Lambda(t_i, t_j)v_j^N$, for $1 \leq i \leq N$.

The rest of the proof is to show that (2.12) and (3.18) are equal, when they are finite. To that end, we interpolate $(x(t_1), x(t_2), \ldots, x(t_N))$ in the following way:

$$\tilde{x}^N(t) = \sum_{j=1}^N \Lambda(t, t_j)v_j^N,$$

(3.19)

The reason for doing this is that it lies in the Hilbert space defined below (in particular, a reproducing kernel Hilbert space with norm denoted by $\| \cdot \|_\Lambda$),

$$\| \tilde{x}^N \|_\Lambda = \sum_{i,j=1}^N v_i^N \Lambda(t_i, t_j)v_j^N.$$

We now construct the aforementioned Hilbert space. This is a standard construction [2, Section 1.2 and 1.3], but we give it nonetheless for completeness. From the definition of $\Lambda(\cdot, \cdot)$, it is clear that it is positive definite, i.e., for any $v \in \mathbb{R}^k$, $k \in \mathbb{N}$,

$$\sum_{i,j=1}^k v_i v_j \Lambda(t_i, t_j) \geq 0,$$

for any $(t_1, t_2, \ldots, t_k) \in [0, T]^k$. Now consider the span (say $H_0$) of all functions of the form $f : [0, T] \to \mathbb{R}$,

$$f(t) = \sum_{i=1}^K a_i \Lambda(t_j, t),$$

where $K \in \mathbb{N}$, $(a_1, a_2, a_3, \ldots, a_K) \in \mathbb{R}^K$ and $(t_1, t_2, \ldots, t_K) \in [0, T]^K$. Now we define an inner product on $H_0$ in the following way: For $f, g \in H_0$ and $(s_1, s_2, \ldots, s_J) \in [0, T]^J$ with

$$f = \sum_{i=1}^K a_i \Lambda(t_i, t) \quad \text{ and } \quad g = \sum_{j=1}^J b_j \Lambda(s_j, t),$$

$$\langle f, g \rangle_\Lambda = \sum_{i=1}^K \sum_{j=1}^J a_i b_j \Lambda(t_i, s_j).$$

Finally, we get the reproducing kernel Hilbert space (say $H$) by completing $H_0$ under $\langle \cdot, \cdot \rangle_\Lambda$ [2, Section 1.2 and 1.3].
Suppose \( x \in \mathcal{H}_0 \). Then from the definition of \( \mathcal{H}_0 \), we know

\[
x(t) = \sum_{i=1}^{N} v^N_i \Lambda(t, t_i), \quad \text{for some } N \text{ and } (v^N_1, v^N_2, \ldots, v^N_N) \in \mathbb{R}^N.
\]

Therefore,

\[
\|x\|_\Lambda = \sum_{i,j=1}^{N} v^N_i \Lambda(t_i, t_j)v^N_j = \sum_{i,j=1}^{N} \tilde{v}^N_i \Lambda(t_i, t_j)\tilde{v}^N_j,
\]

for any other representation of \( x(\cdot) \) such that

\[
x(t) = \sum_{i=1}^{N} \tilde{v}^N_i \Lambda(t, t_i).
\]

This in turn means that for \( x \in \mathcal{H}_0 \),

\[
I^{\text{MDP}}_1(x) = \frac{1}{2} \sup_{0 < t_1 < t_2 < \ldots < t_N \leq T} \sum_{i,j=1}^{N} v^N_i \Lambda(t_i, t_j)v^N_j
\]

\[
= \frac{1}{2} \sup_{0 < t_1 < t_2 < \ldots < t_N \leq T} \|x\|_\Lambda^2 = \frac{1}{2} \|x\|_\Lambda^2.
\]

Since any \( x \in \mathcal{H} \) is a limit point of some sequence \( \{x_n\}_{n \in \mathbb{N}} \subset \mathcal{H}_0 \), we can conclude that

\[
I^{\text{MDP}}_1(x) = \frac{1}{2} \|x\|_\Lambda^2, \quad \text{for } \xi \in \mathcal{H}.
\]

In the following, we explicitly give the expression for \( \|x\|_\Lambda^2 \). To that end, we rewrite

\[
x_n(t) = \sum_{i=1}^{N_n} (w^n_j - w^n_{j-1}) \Lambda(t, t_j), \quad \text{for some } (w^n_1, w^n_2, \ldots, w^n_{N_n}) \in \mathbb{R}^{N_n}.
\]

Then, we have

\[
\|x_n\|_\Lambda^2 = \sum_{i,j=1}^{N_n} \frac{w^n_i - w^n_{i-1}}{t_i - t_{i-1}} \Lambda(t_i, t_j) \frac{w^n_j - w^n_{j-1}}{t_j - t_{j-1}} (t_i - t_{i-1})(t_j - t_{j-1}).
\]

Since \( \Lambda(\cdot, \cdot) \) is Lebesgue measurable on \([0, T] \times [0, T]\), as \( n \to \infty \), we have

\[
\|x\|_\Lambda^2 = \lim_{n \to \infty} \|x_n\|_\Lambda^2 = \int_0^T \int_0^T \dot{\omega}(s) \Lambda(s,t) \dot{\omega}(t) ds dt, \quad \text{where } x(\cdot) = \int_0^T \Lambda(\cdot, s) \dot{\omega}(s) ds.
\]

Since, only \( \dot{\omega} \) is involved in the above equation (and it is sufficient for \( \omega \) to lie in \( \mathcal{A}_0 \) to have the above integrals well-defined), we replace \( \dot{\omega}(\cdot) = z(\cdot) \), where \( z \) is a Lebesgue measurable function on \([0, T]\). This gives us

\[
\|x\|_\Lambda^2 = \lim_{n \to \infty} \|x_n\|_\Lambda^2 = \int_0^T \int_0^T z(s) \Lambda(s,t) z(t) ds dt, \quad \text{where } x(\cdot) = \int_0^T \Lambda(\cdot, s) z(s) ds.
\]

This proves the result. \(\square\)
3.3. Exponential tightness of $\tilde{X}_1^n$ in $(D_T, J_1)$. We now proceed to study $\{\tilde{X}_1^n\}_{n \in \mathbb{N}}$ first by establishing the exponential tightness (see [18, Pg. 8] for definition) of $\{\tilde{X}_1^n\}_{n \in \mathbb{N}}$ in $(D_T, J_1)$.

**Lemma 3.3.** Suppose Assumptions A.1–A.4 hold. Then the following holds

$$
\lim \limsup_{\delta \to 0} \sup_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P} \left( \left\{ x \in D_T : \sup_{s \in [0, \delta]} |\tilde{X}_1^n(t + s) - \tilde{X}_1^n(t)| > \epsilon \right\} \right) = -\infty,
$$

(3.20)

for every $\epsilon > 0$.

**Proof.** Fix $0 < s < t \leq T$ and consider

$$
\tilde{X}_1^n(t) - \tilde{X}_1^n(s) = \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda t \rfloor} \widetilde{H}(t, \frac{i}{\lambda n}, \vartheta_{i}) - \frac{1}{a_n \sqrt{n}} \sum_{j=1}^{\lfloor \lambda s \rfloor} \widetilde{H}(s, \frac{j}{\lambda n}, \vartheta_{j})
$$

$$
= \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda s \rfloor} \widetilde{H}(t, \frac{i}{\lambda n}, \vartheta_{i}) - \widetilde{H}(s, \frac{i}{\lambda n}, \vartheta_{i}) + \frac{1}{a_n \sqrt{n}} \sum_{j=1}^{\lfloor \lambda t \rfloor} \widetilde{H}(t, \frac{j}{\lambda n}, \vartheta_{j}) - \widetilde{H}(s, \frac{j}{\lambda n}, \vartheta_{j})
$$

$$
\doteq S(n, s, t) + R(n, s, t).
$$

It is clear that the two terms $S(n, s, t)$ and $R(n, s, t)$ are independent.

Observe that for any $\epsilon > 0$,

$$
\left\{ \sup_{0 \leq t - s \leq \delta} |\tilde{X}_1^n(t) - \tilde{X}_1^n(s)| > \epsilon \right\} \subset \left\{ \sup_{0 \leq t - s \leq \delta} |S(n, s, t)| > \frac{\epsilon}{2} \right\} \cup \left\{ \sup_{0 \leq t - s \leq \delta} |R(n, s, t)| > \frac{\epsilon}{2} \right\}. \tag{3.21}
$$

In the following, we will estimate the probabilities of the above events on the right hand side using Corollary A.1.

We begin with estimating

$$
\left\{ \sup_{0 \leq t - s \leq \delta} |R(n, s, t)| > \frac{\epsilon}{2} \right\}.
$$

We can rewrite $R(n, s, t)$ as

$$
\frac{1}{a_n \sqrt{n}} \sum_{i=1}^{\lfloor \lambda n(t - s) \rfloor} \frac{\lambda i}{\lambda n} \vartheta_{i} \doteq \frac{1}{a_n \sqrt{n}} \sum_{i=1}^{N_{t-s}} H_{i,t}.
$$

Here, $N_{t-s} = \lfloor \lambda n(t - s) \rfloor$. It is easy to see that using a slightly modified version of Corollary A.1, we get

$$
\frac{1}{a_n^2} \log \mathbb{P} \left( \sup_{0 \leq t - s \leq \delta} |R(n, s, t)| > \frac{\epsilon}{2} \right) \leq \frac{\rho^2}{2n} \mathbb{E} \left[ \sum_{i=1}^{\lfloor \lambda n \delta \rfloor} H_{i,t}^2 \right] - \frac{\rho \epsilon}{8} + C_H^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
$$

$$
\leq \frac{\rho^2 4C_H^2 |\lambda n \delta|}{2n} - \frac{\rho \epsilon}{8} + C_H^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
$$

Therefore,

$$
\limsup_{n \to \infty} \sup_{t \in [0, T]} \frac{1}{a_n^2} \log \mathbb{P} \left( \sup_{0 \leq t - s \leq \delta} |R(n, s, t)| > \frac{\epsilon}{2} \right) \leq \frac{\rho^2 4C_H^2 |\lambda n \delta|}{2} - \frac{\rho \epsilon}{8}.
$$

Until now, $\rho$ is arbitrary positive number. We now choose the most optimum value of $\rho$ so that the above bound is minimum. This is achieved for $\rho = \frac{\epsilon}{32C_H^2 \lambda \delta}$ and the corresponding bound is
Again, from a slightly modified version of Corollary A.1, we have \( \gamma \) (and hence, uniformly continuous with \( \nu \)).

We next estimate the probability of

\[
\left\{ \sup_{0 \leq t-s \leq \delta} |S(n, s, t)| < \frac{\epsilon}{2} \right\}
\]

following exactly along the same lines we did for \( R(n, s, t) \) with

\[
S_{i, t} = \tilde{H}(t, \frac{i}{\lambda n}, \nu_i) - \tilde{H}(s, \frac{i}{\lambda n}, \nu_i).
\]

Again, from a slightly modified version of Corollary A.1, we have

\[
\limsup_{n \to \infty} \sup_{t \in [0, T]} \frac{1}{a_n^2} \log P \left( \sup_{0 \leq t-s \leq \delta} \left| S(n, s, t) \right| > \frac{\epsilon}{2} \right) = \limsup_{n \to \infty} \sup_{t \in [0, T]} \frac{1}{a_n^2} \log P \left( \sup_{0 \leq t-s \leq \delta} \frac{1}{a_n} \sum_{i=1}^{N_n} S_{i, t} > \frac{\epsilon}{2} \right)
\]

\[
\leq \rho^2 \limsup_{n \to \infty} \frac{1}{2n} \sum_{i=1}^{[\lambda n T]} \mathbb{E} \left[ S_{i, t}^2 \right] - \frac{\rho \epsilon^3}{8} + C_{\tilde{H}} |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
\]

We now compute

\[
\sum_{i=1}^{[\lambda n T]} \mathbb{E} \left[ S_{i, t}^2 \right] = \sum_{i=1}^{[\lambda n T]} \mathbb{E} \left[ \left( \tilde{H}(t, \frac{i}{\lambda n}, \nu_i) - \tilde{H}(s, \frac{i}{\lambda n}, \nu_i) \right)^2 \right]
\]

\[
= \sum_{i=1}^{[\lambda n T]} \int_{\mathbb{R}^d} \left( \tilde{H}(t, \frac{i}{\lambda n}, x) - \tilde{H}(s, \frac{i}{\lambda n}, x) \right)^2 F(dx)
\]

\[
= \sum_{i=1}^{[\lambda n T]} \int_{\mathbb{R}^d} \left( \tilde{H}(t, \frac{i}{\lambda n}, x) - \tilde{H}(s, \frac{i}{\lambda n}, x) \right)^2 F(dx).
\]

From Lemma A.1, we know that there is a closed set \( K_\delta \subset [0, T] \) such that \( \tilde{H}(\cdot, u, x) \) is continuous (and hence, uniformly continuous with \( \gamma_{u,x}(\cdot) \) being the modulus of continuity) on \( K_\delta \) and \( \nu([0, T] \setminus K_\delta) < \delta \). So we write \( \tilde{H}(s, u, x) = \tilde{H}(s, u, x) \mathbb{1}_{s \in K_\delta} + \tilde{H}(s, u, x) \mathbb{1}_{s \notin K_\delta} \). Using this, we have

\[
\int_{\mathbb{R}^d} \left( \tilde{H}(t, \frac{i}{\lambda n}, x) - \tilde{H}(s, \frac{i}{\lambda n}, x) \right)^2 F(dx)
\]

\[
= \int_{\mathbb{R}^d} \left( \left( \tilde{H}(t, \frac{i}{\lambda n}, x) \mathbb{1}_{t \in K_\delta} - \tilde{H}(s, \frac{i}{\lambda n}, x) \mathbb{1}_{s \in K_\delta} \right)^2 + \left( \tilde{H}(t, \frac{i}{\lambda n}, x) \mathbb{1}_{t \notin K_\delta} - \tilde{H}(s, \frac{i}{\lambda n}, x) \mathbb{1}_{s \notin K_\delta} \right)^2 \right) F(dx)
\]

\[
\leq \int_{\mathbb{R}^d} \left( \frac{\gamma_{\frac{i}{\lambda n}, x}(t-s)}{\lambda n} \right)^2 F(dx)
\]

\[
+ \int_{\mathbb{R}^d} \left( \tilde{H}(t, \frac{i}{\lambda n}, x) \mathbb{1}_{t \notin K_\delta} - \tilde{H}(s, \frac{i}{\lambda n}, x) \mathbb{1}_{s \notin K_\delta} \right)^2 F(dx)
\]

\[
= W_1(t-s, \frac{i}{\lambda n}) + W_2(t, s, \frac{i}{\lambda n})
\]
Using (3.28), we have

\[ W_1(\delta, \frac{i}{\lambda_n}) + W_2(t, s, \frac{i}{\lambda_n}). \]

This implies that

\[
\limsup_{n \to \infty} \sup_{t \in [0,T]} \frac{1}{a_n^2} \log P \left( \sup_{0 \leq t-s \leq \delta} |S(n, s, t)| > \frac{\epsilon}{2} \right) \\
\leq \frac{\rho^2}{2} \int_{K_{\delta}} W_1(\delta, u) du + \frac{\rho^2}{2} \int_{[0,T]\setminus K_{\delta}} \sup_{t \in [0,T]} W_2(t, s, u)^2 du - \frac{\rho \epsilon}{8}.
\]

As we did earlier, we choose \( \rho = \epsilon \left( \int_{c_{K_{\delta}}} W_1(\delta, u) du + 4C_H \delta \right)^{-1} \) such that the above bound is minimum and then take \( \delta \to 0 \). This completes the proof. \( \square \)

**Remark 3.2.** From Lemma 3.3 which implies exponential tightness and [49, Theorem P], for every subsequence \( n_k \), there exists a further subsequence (still denoted by \( n_k \)) such that \( \{\lambda_n^{n_k}\}_{k \in \mathbb{N}} \) satisfies MDP in \( \mathcal{D}_T \) under \( J_1 \) topology with with some rate function \( I_{(n)} \) (subscript \( (n) \) denotes the possible dependence of the rate function on the subsequence.) We now remark that the conclusion of [51, Theorem A.3] is stronger than what we used in the proof of Lemma 3.3. To be more precise, [51, Theorem A.3] says that any rate function \( I_{(n)} \) is such that \( I_{(n)}(x) = \infty \), whenever \( x \in \mathcal{D}_T \setminus \mathcal{C}_T \). Recall that \( \mathcal{C}_T \) is the space of continuous functions on \([0, T]\).

**Alternate proof of Lemma 3.3.** This alternate proof exploits the already established FCLT result. We are required to show that for every \( L > 0 \), there is a compact set (in \( J_1 \) topology) \( K_L \subset \mathcal{D}_T \) such that

\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log P(\lambda_n^{n_k} \in K_L^c) \leq -L.
\]

Here \( K_L^c \) is the complement of \( K_L \) in \( \mathcal{D}_T \).

From [46, Theorem 2.2], the family of \( \mathcal{D}_T \)-valued processes

\[
Y^n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \tilde{H}(t, \frac{i}{\lambda_n}, \tilde{d}_i) = a_n \lambda_n^{n_k}(t)
\]

is tight in \( (\mathcal{D}_T, J_1) \), i.e., for every \( 0 < \delta < 1 \), there is a compact set \( \hat{K}_\delta \subset \mathcal{D}_T \) such that

\[
P(Y^n \in \hat{K}_\delta) \leq \delta
\]

and the weak limit is some Gaussian process \( Y \) with continuous sample paths. Suppose

\[
M = \limsup_{n \to \infty} E \left[ e^{\alpha \|Y^n\|^2_T} \right] < \infty,
\]

for some \( \alpha > 0 \). Then for any compact set (in \( J_1 \) topology) \( K_L \) with diameter \( \sqrt{L/\alpha} > 0 \), we have

\[
P(\lambda_n^{n_k} \in K_L^c) = P \left( \frac{1}{a_n} Y^n \in K_L^c \right) \leq P(Y^n \in a_n K_L^c).
\]

Using (3.28), we have

\[
\mathbb{E} \left[ e^{\alpha \|Y^n\|^2_T} 1_{\{Y^n \notin a_n K_L^c\}} \right] + \mathbb{E} \left[ e^{\alpha \|Y^n\|^2_T} 1_{\{Y^n \in a_n K_L^c\}} \right] = \mathbb{E} \left[ e^{\alpha \|Y^n\|^2_T} \right] \leq M,
\]

\[
e^{\alpha^2 L} P(Y^n \in a_n K_L^c) \leq M,
\]

\[
\limsup_{n \to \infty} \frac{1}{a_n^2} \log P(\lambda_n^{n_k} \in K_L^c) = \limsup_{n \to \infty} \frac{1}{a_n^2} \log P(Y^n \in a_n K_L^c) \leq -L.
\]

Sample Path MDP for shot noise processes 21
Now it only remains to prove (3.28). To that end, we first observe from Portmanteau theorem [5, Theorem 2.1] that for an open ball $B_R$ (in $J_1$ topology) of radius $R$ in $D_T$ around the origin,

$$\lim_{n \to \infty} \sup_{n} \mathbb{P}(Y^n \in B_R) \leq \mathbb{P}(Y \in B_R) \leq e^{-CR^2},$$

for some $C > 0$. For $\alpha < C$, we have

$$\lim_{n \to \infty} \sup_{n} \mathbb{E} \left[ e^{\alpha \|Y^n\|_T^2} \right] = \lim_{n \to \infty} \sum_{R=1}^{\infty} \mathbb{E} \left[ e^{\alpha \|Y^n\|_T^2} \mathbb{1}_{\{R \leq \|Y^n\|_T < R+1\}} \right] \leq \lim_{n \to \infty} \sum_{R=1}^{\infty} e^{\alpha(R+1)^2} \mathbb{E} \left[ \mathbb{1}_{\{R \leq \|Y^n\|_T < R+1\}} \right] \leq \lim_{n \to \infty} \sum_{R=1}^{\infty} e^{\alpha(R+1)^2} \mathbb{P}(\|Y^n\|_T \geq R) \leq \sum_{R=1}^{\infty} e^{\alpha(R+1)^2} \lim_{n \to \infty} \mathbb{P}(\|Y^n\|_T \geq R) \leq \sum_{R=1}^{\infty} e^{\alpha(R+1)^2} e^{-CR^2} < \infty, \text{ since } \alpha < C.$$

This proves the desired result. \hfill \Box

Remark 3.3. We note that from the tightness of $\{Y^n\}_{n \in \mathbb{N}}$ and the fact that all the weak limit points of $\{Y^n\}_{n \in \mathbb{N}}$ are Gaussian process with continuous sample paths, we can conclude that the compact set (in $J_1$ topology) $K_L$ is also compact in uniform topology.

Remark 3.4. We have provided two proofs of Lemma 3.3 and, each of them implies stronger statements than just the statement of Lemma 3.3 viz., Remark 3.2 and Remark 3.3.

3.4. Completing the proof of Theorem 2.1. First, by the MDP of $\hat{X}_1^n$ in the topology of pointwise convergence in Theorem 3.1 and the exponential tightness of $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ in $(D_T, J_1)$, we obtain the sample-path MDP for $\hat{X}_1^n$ in $(D_T, J_1)$ below.

**Theorem 3.2.** Suppose Assumptions A.1–A.4 hold. Then the family of $D_T$–valued random variables $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ satisfies an MDP in $(D_T, J_1)$ with rate $a_n^2$ and rate function $I_1^{\text{MDP}}$ given by (2.12).

Remark 3.5. Using Remark 3.3 and the fact that the uniform topology on $[0, T]$ is finer than topology of pointwise convergence along with [18, Corollary 4.2.6], we can conclude that $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ in $(D_T, \|\cdot\|)$.

Since uniform topology of $[0, T]$ is finer than $J_1$ topology, we also have the statement of the theorem. Therefore, below we only provide the proof under the conclusion of Lemma 3.3 and the associated Remark 3.2.

**Proof.** From Remark 3.2, we know that for any sequence $n_k$, there is a further subsequence (still denoted by) $n_k$ such that $\{\hat{X}_1^n\}_{n \in \mathbb{N}}$ satisfies MDP with rate $a_n^2$ and some rate function $I_{(n)}$ such that $I_{(n)}(x) = \infty$ whenever $x \in D_T \setminus C_T$. This means that for any $L > 0$, the set $\tilde{K}_L \equiv \{I_{(n)} \leq L\}$ is a compact set of $D_T$ in $J_1$ topology and contains only continuous functions. This means that $\tilde{K}_L$ is also compact in the uniform topology (i.e., the topology that is induced by $\|\cdot\|_T$). This follows from the fact that $J_1$ topology restricted to $C_T$ is the same as the uniform topology.

Now consider a closed set $C \subset D_T$ in $J_1$ topology. We have

$$\lim_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\hat{X}_{I_{(n_k)}^k} \in C) \leq \lim_{k \to \infty} \frac{1}{a_{n_k}^2} \log \left( \mathbb{P}(\hat{X}_{I_{(n_k)}^k} \in C \cap \tilde{K}_L) + \mathbb{P}(\hat{X}_{I_{(n_k)}^k} \in C \cap \tilde{K}_L^c) \right)$$
\begin{align*}
\leq \max \left\{ \limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \mathcal{C} \cap \widehat{K}_L), \right. \\
\left. \limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \mathcal{C} \cap \widehat{K}_L^c) \right\}
\leq \max \left\{ \limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \mathcal{C} \cap \widehat{K}_L), \right. \\
\left. \limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \widehat{K}_L^c) \right\}
\leq \max \left\{ \limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \mathcal{C} \cap \widehat{K}_L), -L \right\}.
\end{align*}

We have obtained the last inequality after using the fact that \( \{\widehat{X}_{1}^{n_k}\}_{n \in \mathbb{N}} \) satisfies MDP in \((\mathcal{D}_T, J_1)\) with rate function \( I_{(n)} \) and the upper bound in the definition of MDP. Indeed,
\[
\limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \widehat{K}_L^c) \leq - \inf_{x \in \widehat{K}_L} I_{(n)}(x) \leq -L.
\]

Now observe that \( \mathcal{C} \cap \widehat{K}_L \) is closed in \( \mathcal{D}_T \) under uniform topology, and then it is also closed under the topology of pointwise convergence. Therefore, using Theorem \(3.1\) and the MDP upper bound, we get
\[
\limsup_{k \to \infty} \frac{1}{a_{n_k}^2} \log \mathbb{P}(\widehat{X}_{1}^{n_k} \in \mathcal{C}) \leq \max \left\{ - \inf_{x \in \mathcal{C} \cap \widehat{K}_L} I^{\text{MDP}}_1(x), -L \right\} \leq \max \left\{ - \inf_{x \in \mathcal{C}} I^{\text{MDP}}_1(x), -L \right\}
\leq - \inf_{x \in \mathcal{C}} I^{\text{MDP}}_1(x), \text{ after taking } L \uparrow \infty.
\]

Since the right hand above is independent of the subsequence \( n_k \), we have
\[
\limsup_{n \to \infty} \frac{1}{a_k^2} \log \mathbb{P}(\widehat{X}_n \in \mathcal{C}) \leq - \inf_{x \in \mathcal{C}} I^{\text{MDP}}_1(x).
\]

We now move on to proving the lower bound in the definition of MDP. To do that, we remark that it only suffices to prove the following: for any \( x \in \mathcal{D}_T \) such that \( I^{\text{MDP}}_1(x) < \infty \), for any subsequence \( n_k \) and any open ball (in \( J_1 \) topology) of radius \( \delta > 0 \) around \( x \) (denoted by \( O_x(\delta) \)), we have
\[
\liminf_{n \to \infty} \frac{1}{a_n^2} \log \mathbb{P}(\widehat{X}_n \in O_x(\delta)) \geq -I^{\text{MDP}}_1(x).
\]

Without loss of generality, we can take \( x \in \mathcal{C}_T \) (see Remark \(3.2\)). Since \( x \) is continuous on \([0, T]\), it is also uniformly continuous with modulus of continuity \( w_x(\cdot) \) (say). Then from the property of \( J_1 \) topology, there is an open \( r \)-ball \( B_x(\| \cdot \|_T, r) \) around \( x \) in uniform topology such that \( B_x(\| \cdot \|_T, r) \subset O_x(\delta) \). Let \( e(\cdot) \) denote the identity map on \([0, T]\). Using the definition of the \( J_1 \) topology,
\[
\max \{\|y - x \circ a\|_T, \|a - I\|_T\} < \max \{\|y - x\|_T + w_x(\|e(\cdot) - a(\cdot)\|_T), \|e(\cdot) - a(\cdot)\|_T\}
\]
with \( f \circ g \) denoting the composition of functions \( f \) and \( g \). Therefore, choosing \( r = \delta \) and \( a \equiv e \), we can ensure that
\[
B_x(\| \cdot \|_T, r) \subset O_x(\delta).
\]
Again fix a subsequence $n_k$, along which $\{\tilde{X}^{n_k}_1\}_{k \in \mathbb{N}}$ satisfies MDP in $(D_T, J_1)$. For $L > 0$, choose $\hat{K}_L$ corresponding to this subsequence as earlier. From the above discussion, we have

$$\liminf_{k \to \infty} \frac{1}{a^2_{n_k}} \log \mathbb{P}(\tilde{X}^{n_k}_1 \in B_x(\| \cdot \|_T, \delta)) \leq \liminf_{k \to \infty} \frac{1}{a^2_{n_k}} \log \mathbb{P}(\tilde{X}^{n_k}_1 \in O_x(\delta)).$$

Following the proof of the lower bound in [18, Theorem 4.2.4] (whose proof involves choosing a continuous function $g$ mapping $(D_T, \| \cdot \|_T)$ to $D_T$ with the topology of pointwise convergence and a compact set in $(D_T, \| \cdot \|_T)$; to that end, we choose $\hat{K}_L$ as the compact set and the function $g$ as the continuous injection of $(C_T, \| \cdot \|_T)$ into $C_T$ with topology of pointwise convergence), we get

$$\liminf_{k \to \infty} \frac{1}{a^2_{n_k}} \log \mathbb{P}(\tilde{X}^{n_k}_1 \in O_x(\delta)) \geq \liminf_{k \to \infty} \frac{1}{a^2_{n_k}} \log \mathbb{P}(\tilde{X}^{n_k}_1 \in B_x(\| \cdot \|_T, \delta)) \geq -I^1_{MDP}(x).$$

From the arbitrariness of sequence $n_k$, we have

$$\liminf_{n \to \infty} \frac{1}{a^2_n} \log \mathbb{P}(\tilde{X}^n_1 \in O_x(\delta)) \geq -I^1_{MDP}(x).$$

To show that $\{x \in D_T : I^1_{MDP}(x) \leq l\}$ is a compact set (in $J_1$ topology) of $D_T$ for every $l \geq 0$, we do the following: Using the same argument as earlier, over any subsequence, $\{\tilde{X}^n_1\}_{n \in \mathbb{N}}$ satisfies MDP in $(D_T, J_1)$ with some rate function $I^{(n)}_G$ which is such that $\{x \in D_T : I^{(n)}_G(x) \leq l\}$ is a compact set (in $J_1$ topology) of $D_T$. From [18, Lemma 4.1.4], the rate function $I^{(n)}_G$ and $I^1_{MDP}$ are identical. Hence, $I^1_{MDP}$ also satisfies the desired property. This completes the proof.

Next we prove the MDP of $\{\tilde{X}^n_2\}_{n \in \mathbb{N}}$ by using the MDP for renewal processes in Theorem A.1.

**Proposition 3.2.** Suppose Assumptions A.1–A.5 hold. Then the family of $D_T$–valued random variables $\{\tilde{X}^n_2\}_{n \in \mathbb{N}}$ satisfies an MDP in $(D_T, J_1)$ with rate $a^2_n$ and rate function $I^2_{MDP}$ given by (2.13).

**Proof.** From Theorem A.1, we know that $\{\tilde{A}^n\}_{n \in \mathbb{N}}$ satisfies MDP with rate $a^2_n$ and rate function $I^A_{MDP}$. By integration by parts, we can write

$$\tilde{X}^n_2(t) = \int_0^t G_1(t, s) d\tilde{A}^n(s)
= \tilde{A}^n(t)G_1(t, t) - \int_0^t \tilde{A}^n(u-)dG_1(t, u).$$

From [46, Lemma 6.1], Assumption A.4 and A.5, we know that the mapping $\phi : (D_T, J_1) \to (D_T, J_1)$ defined by

$$\phi(f) = f(t)G_1(t, t) - \int_0^t f(u-)dG_1(t, u)$$

is continuous. Therefore, using the contraction principle [18, Theorem 4.2.1], we can conclude that $\{\tilde{X}^n_2\}_{n \in \mathbb{N}}$ satisfies MDP in $(D_T, J_1)$ with rate $a^2_n$ and rate function given by (2.13).

**Proof of Theorem 2.1.** Let

$$\tilde{X}^n(t) = \tilde{X}^n_1(t) + \tilde{X}^n_2(t), \text{ for } t \in [0, T].$$

From Lemma 3.1 and [18, Theorem 4.2.13], the MDP of $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ in $(D_T, J_1)$ is implied by the MDP of $\{\tilde{X}^n\}_{n \in \mathbb{N}}$ in $(D_T, J_1)$ with the same rate and rate function. To get the MDP of $\{\tilde{X}^n\}_{n \in \mathbb{N}}$, we use Proposition 3.2 and Theorem 3.2. This concludes the proof of the theorem. 

**Appendix A. Some Auxiliary Results for the Proofs**

In this section, we give a few auxiliary results that are used in the paper.
A.1. The MDP and LDP of \( A^n \). We first present a version of the sample-path MDP for renewal processes.

**Theorem A.1.** Under Assumption A.1 and suppose that \( a_n \) satisfies (2.8), the family of \( \mathcal{D}_T \)-valued random variables \( \{ \tilde{A}^n \}_{n \in \mathbb{N}} \) defined by

\[
\tilde{A}^n(t) = \frac{1}{a_n \sqrt{n}} (A^n(t) - n \lambda t), \quad \text{for } t \in [0, T]
\]

where \( A^n(t) \) is defined in (2.6), obeys an MDP in \( (\mathcal{D}_T, J_1) \) with rate \( a_n^2 \) and the following rate function \( I_{\tilde{A}}^{MDP} \):

\[
I_{\tilde{A}}^{MDP}(x) = \begin{cases} \frac{1}{2 \lambda \sigma^2} \int_0^T |\dot{x}(t)|^2 dt, & \text{whenever } x \in \mathcal{A}_0, \\ \infty, & \text{otherwise}. \end{cases} \quad (A.1)
\]

**Remark A.1.** We remark that in [52, Theorem 6.2], the sample-path MDP for renewal processes is proved under either of the following two sets of conditions:

(i) \( \frac{\log(n)}{a_n^3} \to \infty \) and \( \mathbb{E}(\eta_1)^{2+\varepsilon} < \infty \), for some \( \varepsilon > 0 \);

(ii) For some \( \beta \in (0, 1] \), \( \frac{n^2}{a_n^3} \to \infty \) and \( \mathbb{E}\exp(b(\eta_1)^{\beta}) < \infty \), for some \( b > 0 \).

The proof is given in [50]. As noted right after Theorem 6.2 in [52], the case \( \beta = 1 \) which is not included there is dealt with by the same argument. In our paper, the MDP for \( \tilde{X}^n_t(t) \) is proved under the conditions of \( a_n \) in (2.8), while the MDP for \( \tilde{X}^n_t(t) \) is proved using contraction mapping theorem together with the known MDP result for renewal processes. Thus, to unify the conditions on \( a_n \), we have imposed the conditions on the interarrival times \( \{\eta_i\}_{i \in \mathbb{N}} \) in Assumption A.1.

**Theorem A.2.** [52, Theorem 6.1(b)] Assume that \( \mathbb{E}[e^{\gamma \eta_1}] < \infty \), for some \( \gamma > 0 \). Let \( \gamma^* = \sup \{\gamma : \mathbb{E}[e^{\gamma \eta_1}] < \infty\} \). Also assume that \( \mathbb{P}(\eta_1 = 0) = 0 \). Then \( \{n^{-1}A^n\}_{n \in \mathbb{N}} \) satisfies LDP in \( (\mathcal{D}_T, J_1) \) with rate \( n \) and rate function

\[
I_{\tilde{A}}^{LDP}(x) = \begin{cases} \int_0^T \varphi(\dot{x}(t)) dt, & \text{whenever } x \in \mathcal{A}_0, \\ \infty, & \text{otherwise}. \end{cases}
\]

where

\[
\varphi(x) = \sup \{\gamma x - \log \mathbb{E}[e^{\gamma \eta_1}] : \gamma < \gamma^*\}.
\]

A.2. A result related to Lusin’s theorem.

**Lemma A.1.** Suppose \( f : [0, T] \times [0, T] \times \mathbb{R}^d \to \mathbb{R}_+ \) be a bounded measurable function. Then, for small enough \( \delta > 0 \), there exists a closed set \( C_\delta \subset [0, T] \times [0, T] \) such that \( \nu([0, T] \times [0, T] \setminus C_\delta) < 2\delta \) and \( f(\cdot, \cdot, x) \) is continuous on \( C_\delta \), for every \( x \in \mathbb{R}^d \).

**Remark A.2.** The main content of this lemma is that we can choose the aforementioned closed set uniformly with respect to \( x \in \mathbb{R}^d \).

**Proof.** Fix a \( \delta > 0 \). From Lusin’s theorem [54, Pg. 66], there exists a closed set \( \tilde{C}_\delta(x) \subset [0, T] \times [0, T] \) such that \( \nu([0, T] \times [0, T] \setminus \tilde{C}_\delta(x)) < \delta \) and \( f(\cdot, \cdot, x) \) is continuous on \( \tilde{C}_\delta(x) \). Now, choose \( \delta < \frac{T^2}{2} \). Then, for any \( x, y \in \mathbb{R}^d \), \( \nu(\tilde{C}_\delta(x) \cap \tilde{C}_\delta(y)) > T^2 - 2\delta \). Now define \( C_\delta = \cap_{x \in \mathbb{R}^d} \tilde{C}_\delta(x) \) and this satisfies \( \nu([0, T] \times [0, T] \setminus C_\delta) < 2\delta \). \( \square \)
A.3. A maximal inequality.

**Lemma A.2.** Let \( \{Z_i\}_{i \in \mathbb{N}} \) be a family of independent bounded real valued random variables. Also assume that \( E[Z_i] = 0 \), for every \( i \). Then for the random variable \( S_i \doteq \sum_{j=1}^{i} Z_j \), we have the following

\[
\frac{1}{a_n^2} \log P \left( \sup_{1 \leq i \leq n} \frac{1}{a_n \sqrt{n}} |S_i| > 4 \epsilon \right) \leq \frac{\rho^2}{2n} E \left[ \sum_{j=1}^{n} Z_j^2 \right] - \rho \epsilon + C_Z^2 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right),
\]

for every \( \epsilon > 0 \). Here, \( C_Z \) is the upper bound of \( |Z_1| \).

**Proof.** To begin with, we use the following result from [20, Theorem 1.1]:

\[
P \left( \sup_{1 \leq i \leq n} \frac{1}{a_n \sqrt{n}} |S_i| > 4 \epsilon \right) \leq 4 \max_{1 \leq i \leq n} P \left( \frac{1}{a_n \sqrt{n}} |S_i| > \epsilon \right). \tag{A.2}
\]

We now estimate

\[
P \left( \frac{1}{a_n \sqrt{n}} |S_i| > \epsilon \right) = P \left( \frac{1}{a_n \sqrt{n}} S_i > \epsilon \right) + P \left( \frac{1}{a_n \sqrt{n}} S_i < -\epsilon \right)
= P \left( \exp \left( \frac{a_n \rho}{a_n \sqrt{n}} S_i \right) > \exp(\frac{a_n \rho}{a_n \sqrt{n}} \epsilon) \right) + P \left( \exp \left( - \frac{a_n \rho}{a_n \sqrt{n}} S_i \right) > \exp(\frac{a_n \rho}{a_n \sqrt{n}} \epsilon) \right)
\leq \exp \left( - \frac{a_n \rho^2}{2n} \right) \left[ \exp \left( \frac{a_n \rho^2}{a_n \sqrt{n}} |S_i| \right) + E \left[ \exp \left( - \frac{a_n \rho^2}{a_n \sqrt{n}} S_i \right) \right] \right), \tag{A.3}
\]

where the second inequality follows from Markov inequality, and \( \rho > 0 \) is arbitrary. Next, by Taylor’s series and the fact that \( E[Z_i] = 0 \) and that \( \{Z_i\}_{i \in \mathbb{N}} \) is family of pairwise independent random variables, we obtain

\[
\log E \left[ \exp \left( \frac{a_n \rho}{a_n \sqrt{n}} S_i \right) \right] = \frac{a_n \rho^2}{2n} E \left[ S_i \right] + \frac{a_n \rho^2}{2n} E \left[ |S_i|^2 \right] + C_Z^2 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right)
= \frac{a_n \rho^2}{2n} E \left[ \sum_{j=1}^{i} Z_j^2 \right] + C_Z^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
\]

Similarly,

\[
\log E \left[ \exp \left( - \frac{a_n \rho}{a_n \sqrt{n}} S_i \right) \right] = \frac{a_n \rho^2}{2n} E \left[ \sum_{j=1}^{i} Z_j^2 \right] + C_Z^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
\]

Using the above expressions, for \( 1 \leq j \leq n \), we have

\[
\exp \left( - \frac{a_n \rho^2}{2n} \right) \left[ \exp \left( \frac{a_n \rho^2}{a_n \sqrt{n}} S_i \right) \right] + \exp \left( - \frac{a_n \rho^2}{a_n \sqrt{n}} S_i \right)
\leq 2 \exp \left( \frac{a_n \rho^2}{2n} E \left[ \sum_{j=1}^{n} Z_j^2 \right] - a_n \rho \epsilon + C_Z^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right) \right). \tag{A.4}
\]

Combining (A.2), (A.3), (A.4) and taking logarithm on both sides, we have the result. \( \square \)

The following is a simple consequence of Lemma A.2.

**Corollary A.1.** Suppose the hypothesis of Lemma A.2 holds. Then we have

\[
\frac{1}{a_n^2} \log P \left( \sup_{0 \leq t \leq a_n \sqrt{n}} |S_{\lfloor \lambda t \rfloor}(t)| > 4 \epsilon \right) \leq \frac{\rho^2}{2n} E \left[ \sum_{j=1}^{\lfloor \lambda t \rfloor} Z_j^2 \right] - \rho \epsilon + C_Z^3 |\rho|^3 O \left( \frac{a_n}{\sqrt{n}} \right).
\]
for every $\epsilon > 0$. Here,

$$S_t(t) = \sum_{j=1}^{i} Z_{j,t}$$

with $\{Z_{j,t}\}_{j \in \mathbb{N}}$ being a family of independent (but parametrized by $t \in [0, t_0]$) real valued random variables.

References


