LARGE DEVIATIONS FOR TWO SCALED DIFFUSIONS

Robert Liptser

Dept. of Electrical Engineering-Systems (e-mail liptser@eng.tau.ac.il) Tel Aviv University, Tel Aviv 69978, ISRAEL

Summary. We formulate large deviations principle (LDP) for diffusion pair $(X^\varepsilon, \xi^\varepsilon) = (X_t^\varepsilon, \xi_t^\varepsilon)$, where first component has a small diffusion parameter while the second is ergodic Markovian process with fast time. More exactly, the LDP is established for $(X^\varepsilon, \nu^\varepsilon)$ with $\nu^\varepsilon(dt, dz)$ being an occupation type measure corresponding to ξ_t^ε. In some sense we obtain a combination of Freidlin-Wentzell's and Donsker-Varadhan's results. Our approach relies the concept of the exponential tightness and Puhalskii's theorem.

1. Introduction

Let ε be a small positive parameter, $(X^\varepsilon, \xi^\varepsilon) = (X_t^\varepsilon, \xi_t^\varepsilon)_{t \geq 0}$ be a diffusion pair defined on some stochastic basis $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ by Itô's equations w.r.t. independent Wiener processes W_t and V_t:

$$
\begin{aligned}
dX_t^\varepsilon &= A(X_t^\varepsilon, \xi_t^\varepsilon)dt + \sqrt{\varepsilon}B(X_t^\varepsilon, \xi_t^\varepsilon)dW_t \\
d\xi_t^\varepsilon &= \frac{1}{\varepsilon}b(\xi_t^\varepsilon)dt + \frac{1}{\sqrt{\varepsilon}}\sigma(\xi_t^\varepsilon)dV_t
\end{aligned}
\tag{1.1}
$$

subject to fixed initial point (x_0, z_0).

Assume $(X^\varepsilon, \xi^\varepsilon)$ is an ergodic process in the following sense. Let $p(z)$ be the unique invariant density of ξ^ε,

$$\nu^{(p)}(dt, dz) = p(z)dtdz,$$

and $\overline{X}_t$ is a solution of an ordinary differential equation $\dot{\overline{X}}_t = \overline{A}(\overline{X}_t)$ with $\overline{A}(x) = \int_R A(x, z)p(z)dz$ subject to the same initial point x_0 . Then for any bounded continuous function $h(t, z)$ and $T > 0$

$$
\begin{aligned}
\mathbf{P} - \lim_{\varepsilon \to 0} \int_0^T h(t, \xi_t^\varepsilon)dt &= \int_0^T \int_R h(t, z)\nu^{(p)}(dt, dz), \\
\mathbf{P} - \lim_{\varepsilon \to 0} r_T(X^\varepsilon, \overline{X}) &= 0,
\end{aligned}
\tag{1.2}
$$

where r_T is the uniform metric on $[0, T]$. The above-mentioned ergodic property is a motivation to examine LDP for pair $(X^\varepsilon, \xi^\varepsilon)$, or more exactly for pair $(X^\varepsilon, \nu^\varepsilon)$, where

$\nu^\varepsilon = \nu^\varepsilon(dt, dz)$ is an occupation measure on $(R_+ \times R, \mathcal{B}(R_+) \otimes \mathcal{B}(R))$ $(\mathcal{B}(R_+), \mathcal{B}(R)$ are the Borel σ-algebras on R_+ and R respectively) corresponding to ξ^ε:

$$\nu^\varepsilon(\Delta \times \Gamma) = \int_0^\infty I(t \in \Delta, \, \xi_t^\varepsilon \in \Gamma) dt, \quad \Delta \in \mathcal{B}(R_+), \, \Gamma \in \mathcal{B}(R). \tag{1.3}$$

A choice of ν^ε as the occupation measure is natural since the first ergodic property in (1.2) is nothing but

$$\mathbf{P} - \lim_{\varepsilon \to 0} \rho_T(\nu^\varepsilon, \nu^{(p)}) = 0,$$

where ρ_T is Levy-Prohorov's distance for restrictions of measures ν^ε and $\nu^{(p)}$ on $[0,T] \times R$. Also the first Itô's equation in (1.1) and the predictable quadratic variation $\langle M^\varepsilon \rangle_t$ of a martingale $M_t^\varepsilon = \int_0^t B(X_s^\varepsilon, \xi_s^\varepsilon) dW_s$ can be represented in the term of ν^ε:

$$X_t^\varepsilon = x_0 + \int_0^t \int_R A(X_s^\varepsilon, z) \nu^\varepsilon(ds, dz) + \sqrt{\varepsilon} M_t^\varepsilon,$$

$$\langle M^\varepsilon \rangle_t = \int_0^t \int_R B^2(X_s^\varepsilon, z) \nu^\varepsilon(ds, dz).$$

The random measure ν^ε obeys the disintegration $\nu^\varepsilon(dt, dz) = dt K_{\nu^\varepsilon}(t, dz)$ with the transition kernel $K_{\nu^\varepsilon}(t, dz)$ being probabilistic Dirac's measure that is ν^ε values in space $\mathbb{M} = \mathbb{M}_{[0,\infty)}$ of σ-finite (locally in t) measures $\nu = \nu(dt, dz)$ on $(R_+ \times R, \mathcal{B}(R_+) \otimes \mathcal{B}(R))$ obeying the disintegration $\nu(dt, dz) = K_\nu(t, dz) dt$ with the probabilistic transition kernel $K_\nu(t, dz)$ $(\int_R K_\nu(t, dz) \equiv 1)$. X^ε values in the space $\mathbb{C} = \mathbb{C}_{[0,\infty)}$ of continuous function. Define metrics r and ρ in $\mathbb{C}$ and $\mathbb{M}$ respectively, letting

$$r(X', X'') = \sum_{k \geq 1} \frac{r_k(X', X'') \wedge 1}{2^k} \quad \text{and} \quad \rho(\nu', \nu'') = \sum_{k \geq 1} \frac{\rho_k(\nu', \nu'') \wedge 1}{2^k}.$$

Evidently ergodic properties (1.2) are equivalent to

$$\mathbf{P} - \lim_{\varepsilon \to 0} \left[r(X^\varepsilon, \overline{X}) + \rho(\nu^\varepsilon, \nu^{(p)}) \right] = 0$$

and so for examination of the LDP for $(X^\varepsilon, \nu^\varepsilon)$ we choose the metric space $(\mathbb{C} \times \mathbb{M}, r \times \rho)$.

Recall the definition of LDP from Varadhan [1] adapted to our setting. The family $(X^\varepsilon, \nu^\varepsilon)$ obeys the LDP in the metric space $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ if

(**0.**) there exists a function $L(X, \nu)$, $X \in \mathbb{C}, \nu \in \mathbb{M}$, values in $[0, \infty]$, such that its level sets are compacts in $(\mathbb{C} \times \mathbb{M}, r \times \rho)$;

(**1.**) for any open set G from $(\mathbb{C} \times \mathbb{M}, r \times \rho)$

$$\underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big((X^\varepsilon, \nu^\varepsilon) \in G\big) \geq - \inf_{(X, \nu) \in G} L(X, \nu);$$

(**2.**) for any closed set F from $(\mathbb{C} \times \mathbb{M}, r \times \rho)$

$$\overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big((X^\varepsilon, \nu^\varepsilon) \in F\big) \leq - \inf_{(X,\nu) \in F} L(X, \nu).$$

The function $L(X, \nu)$, meeting in (**0.**), (**1.**), and (**2.**), is named rate function (action functional in the terminology of Freidlin and Wentzell [2] or good rate function in the terminology of Stroock [3]).

Below we recall well known particular results in LDP's related to pair $(X^\varepsilon, \xi^\varepsilon)$ and give corresponding forms of rate functions which will be inherited by a rate function for our setting. Note at first LDP for family $\mu^\varepsilon(dz) = \nu^\varepsilon([0,1], dz)$ (on the space of probability measures supplied by Levy-Prohorov's metric) proved by Donsker and Varadhan [4], [5], [6], [7] for a wide class of Markov processes $\xi_t^\varepsilon = \xi_{t/\varepsilon}$. Corresponding rate function obeys an invariant form: for any probabilistic measure μ on R

$$I(\mu) = - \inf \int_R \frac{\mathcal{L}u(z)}{u(z)} \mu(dz),$$

where $\mathcal{L}$ is backward Kolmogorov's operator, respecting to ξ, and where 'inf' is taken over all functions $u(z)$ from the domain of definition for the operator $\mathcal{L}$. For the diffusion case, Gärtner's type of $I(\mu)$ is well known ([8]):

$$I(\mu) = \begin{cases} \frac{1}{8} \int_R \sigma^2(z) \left[\frac{m'(z)}{m(z)} - \frac{p'(z)}{p(z)} \right]^2 m(z)dz, & d\mu = m(z)dz, \ dm(z) = m'(z)dz, \\ \infty, & \text{otherwise.} \end{cases} \qquad (1.4)$$

Freidlin-Wentzell's result, [2], is devoted to LDP for diffusion X^ε with drift $A(x)$ and diffusion $B^2(x)$ (independent of z) in the space of continuous functions on every finite time interval, supplied by the uniform metric. A rate function, say, for $[0, T]$ time interval is given by

$$S(X) = \begin{cases} \frac{1}{2} \int_0^T \frac{[\dot{X}_t - A(X_t)]^2}{B^2(X_t)} dt, & dX_t = \dot{X}_t dt, \ X_0 = x_0 \\ \infty, & \text{otherwise.} \end{cases} \qquad (1.5)$$

Other type of LDP for a degenerate diffusion X^ε defined by the first equation in (1.1) with $B(x, z) \equiv 0$ and $\xi_t^\varepsilon = \xi_{t/\varepsilon}$, where ξ_t is Markov process values in a finite state space, also is well known from Freidlin [9]. In this case rate function has a form similar to (1.5) ($H(y, x)$ is some non negative function):

$$S(X) = \begin{cases} \int_0^T H(\dot{X}_t, X_t)dt, & dX_t = \dot{X}_t dt, \ X_0 = x_0 \\ \infty, & \text{otherwise.} \end{cases} \qquad (1.6)$$

All above-mentioned LDP's are inspired the examination of the LDP for $(X^\varepsilon, \nu^\varepsilon)$. In some sense, the LDP for $(X^\varepsilon, \nu^\varepsilon)$ is a combination of Donsker-Varadhan's and Freidlin-Wenzell's results. Namely LDP for ν^ε is a generalization one for μ^ε while LDP for X^ε is implied by LDP for ν^ε and for a diffusion martingale scaled by $\sqrt{\varepsilon}$. Hence, a rate function for

$(X^\varepsilon, \nu^\varepsilon)$, is defined as a sum: $L(X, \nu) = L_1(X, \nu) + L_2(\nu)$, where $L_1(X, \nu)$ and $L_2(\nu)$ respect to X^ε and ν^ε and what is more $L_1(X, \nu)$ has the same form as $S(x)$ in (1.5) with $A(X_t)$ and $B^2(X_t)$ replaced on $A_\nu(t, X_t) = \int_R A(X_t, z) K_\nu(t, dz)$ and $B_\nu^2(t, X_t) = \int_R B^2(X_t, z) K_\nu(t, dz)$, where $K_\nu(t, dz)$ is the transition kernel of measure ν.

Note that $\xi_t^\varepsilon \in R$ and so the LDP for its occupation measure responds to a non compact diffusion case. Also note that diffusion parameter $B^2(x, z)$ is not assumed to be non singular and consequently $B^2(x, z) \equiv 0$ is admissible. The last allows to derive LDP for a singular diffusion parameter case from LDP for ν^ε using the contraction principle of Varadhan [1] (continuous mapping method of Freidlin [10]). This result extends above-mentioned [9] for non compact case.

In contrast with Freidlin and Wentzell [2], Donsker and Varadhan [4-7], Gärtner [8], and Veretennikov [11-12], and many others (see e.g. Acosta [13], Dupuis and Elis [14]) our method of proof is based on Puhalskii's theorem [15-16] and relies concepts of exponential tightness and LD relative compactness.

The paper is organized as follows. In Section 2, we formulate the general assumptions and the main result. Section 3 contains the method of proving LDP which also has been used in [17]. In Section 4, we check the exponential tightness while in Section 5 and 6 the upper and lower bounds in local LDP are verified. The main results are proved in Section 7. All technical results are gathered in Appendix.

2. Assumptions. Main result

1. We fix the following conditions which are assumed to be fulfilled hereafter.

(A.1) $A(x, z)$ and $B(x, z)$ are continuous in (x, z), Lipschitz continuous in x uniformly in z, and $\sup_z \left(|A(0, z)| + |B(0, z)| \right) < \infty$;

(A.2) $\sigma^2(z)$ is bounded and and uniformly positive function; it is continuously differentiable, having bounded and Lipschitz continuous derivative;

(A.3) $b(z)$ is Lipschitz continuous, satisfying

$$\lim_{|z| \to \infty} b(z)\mathrm{sign}\, z = -\infty.$$

It would be noted that (A.2) and (A.3) imply, so called, assumption (H*) from [6].

2. It is well known (see [18]) that under (A.2) and (A.3) ξ^ε is ergodic process obeying the unique invariant density

$$p(z) = \text{const.} \frac{\exp\left(2 \int_0^z \frac{b(y)}{\sigma^2(y)} dy \right)}{\sigma^2(z)}. \tag{2.1}$$

For any ν from $\mathbb{M}$ with the transition kernel $K_\nu(t, dz)$, define $K_\nu(t, dz)$-averaged drift $A_\nu(t, x) = \int_R A(x, z) K_\nu(t, dz)$ and diffusion parameter $B_\nu^2(t, x) = \int_R B^2(x, z) K_\nu(t, dz)$. If

ν is absolutely continuous w.r.t. $\Lambda(dt, dz) = dtdz$, put

$$n(t, z) = \frac{d\nu}{d\Lambda}(t, z). \tag{2.2}$$

If the density $n(t, z)$ is absolutely continuous w.r.t. dz: $d_z n(t, z) = n'_z(t, z)dz$, a function $n'_z(t, z)$ is chosen to be measurable in t, z.

Throughout the paper, we use conventions $0/0 = 0$ and $\min(\inf)(\emptyset) = \infty$.

For every $\nu \in \mathbb{M}$ and $X \in \mathbb{C}$ define two quantities (comp. (1.4) and (1.5)):

$$F(\nu) = \begin{cases} \int_0^\infty \int_R \sigma^2(z) \left[\frac{n'_z(t,z)}{n(t,z)} - \frac{p'(z)}{p(z)} \right]^2 n(t, z)dzdt, & d\nu = nd\Lambda, \ d_z n = n'_z dz \\ \infty, & \text{otherwise}; \end{cases}$$

$$S(X, \nu) = \begin{cases} \int_0^\infty \frac{[\dot{X}_t - A_\nu(t, X_t)]^2}{B_\nu^2(t, X_t)}dt, & dX = \dot{X}dt, \quad X_0 = x_0 \\ \infty, & \text{otherwise}. \end{cases} \tag{2.3}$$

3. Now we are in the position to formulate the main result.

Theorem 2.1. *Under (A.1), (A.2), and (A.3) the family $(X^\varepsilon, \nu^\varepsilon)$ obeys the LDP in $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ with rate function*

$$L(X, \nu) = \frac{1}{2}S(X, \nu) + \frac{1}{8}F(\nu).$$

4. LDP's for families (X^ε) and (ξ^ε) run out from Theorem 2.1.

Corollary 2.1. *(ν^ε) obeys the LDP in $(\mathbb{M}, \rho)$ with rate function $\frac{1}{8}F(\nu)$.*

Corollary 2.2. *(comp. [9]) (X^ε) obeys the LDP in $(\mathbb{C}, r)$ with rate function $S(X) = \inf_{\nu \in \mathbb{M}} L(X, \nu)$. In particular, if $B(x, z) \equiv 0$, it is sufficiently to take 'inf' over all ν from $\mathbb{M}$ with the transition kernel $K_\nu(t, dz) \equiv \mu(dz)$ with $d\mu = m(z)dz$ such that the density $m(z) = \frac{d\mu}{dz}(z)$ is absolutely continuous w.r.t. dz $(m'(z) = \frac{dm(z)}{dz})$. In this case, rate function*

$$S(X) = \begin{cases} \frac{1}{8} \int_0^\infty H(\dot{X}_t, X_t)dt, & dX = \dot{X}dt, \quad X_0 = x_0 \\ \infty, & \text{otherwise}, \end{cases} \tag{2.4}$$

where

$$H(y, x) = \inf \int_R \sigma^2(z) \left[\frac{m'(z)}{m(z)} - \frac{p'(z)}{p(z)} \right]^2 m(z)dz, \tag{2.5}$$

and where 'inf' is taken over all above-mentioned measures μ such that

$$y = \int_R A(x, z)m(z)dz.$$

As an example, also the LDP for the family of the Donsker and Varadhan occupation measures $\mu^\varepsilon(dz) = \nu^\varepsilon([0, 1] \times dz)$, corresponding to diffusion case, can be derived from

Theorem 2.1. In fact, due to the contraction principle, (μ^ε) obeys the LDP with Gärtner's type rate function (see (1.4)) $I(\mu) = \inf \frac{1}{8} F(\nu)$, where '$inf$' is taken over all $\nu \in \mathbb{M}$ such that

$$\nu(dt, dz) = I(1 \geq t) dt \mu(dz) + I(1 < t) \nu^{(p)}(dt, dz).$$

3. Preliminaries

For proving LDP for the family $(X^\varepsilon, \nu^\varepsilon)$ in the metric space $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ we apply Dawson-Gärtner's type theorem (see e.g. [19]. Following it the LDP in $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ is implied by LDP's in the metric spaces $(\mathbb{C}_{[0,n]} \times \mathbb{M}_{[0,n]}, r_n \times \rho_n)$, $n \geq 1$, where $\mathbb{C}_{[0,n]}$ is the space of continuous functions on the time interval $[0, n]$, $\mathbb{M}_{[0,n]}$ is the space of finite measures on $[0, n] \times R$, having probabilistic transition kernel w.r.t. dt, r_n is the uniform metric, and ρ_n is Levy-Prohorov's metric. The definition of the LDP in $(\mathbb{C}_{[0,n]} \times \mathbb{M}_{[0,n]}, r_n \times \rho_n)$ is given in terms of (**0.**), (**1.**), and (**2.**) with obvious modifications. Moreover, if $L_n(X, \nu), n \geq 1$ are rate functions, corresponding to LDP's in $(\mathbb{C}_{[0,n]} \times \mathbb{M}_{[0,n]}, r_n \times \rho_n), n \geq 1$, then rate function in $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ is defined as

$$L(X, \nu) = \sup_n L_n(X, \nu). \tag{3.1}$$

Hence only the LDP in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$ has to be checked for any $T > 0$. Our approach in proving the LDP in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T), T > 0$ relies the concept of the exponential tightness and notions of LD relative compactness and local LDP. Below we give necessary definitions.

Definition 1. The family $(X^\varepsilon, \nu^\varepsilon)$ is said to be exponentially tight in the metric space $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$, if there exists an increasing sequence of compacts $(K_j)_{j \geq 1}$ such that

$$\lim_j \overline{\lim_{\varepsilon \to 0}} \varepsilon \log \mathbf{P}\big((X^\varepsilon, \nu^\varepsilon) \in \{\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}\} \setminus K_j\big) = -\infty. \tag{3.2}$$

(Deuschel and Stroock [20], Lynch and Sethuraman [21].)

Definition 2. The family $(X^\varepsilon, \nu^\varepsilon)$ is said to be LD relatively compact in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$, if any decreasing to zero sequence (ε_k) contains further subsequence $(\bar{\varepsilon}_k)$ $((\bar{\varepsilon}_k) \subseteq (\varepsilon_k))$ such that the family $(X^{\bar{\varepsilon}_k}, \nu^{\bar{\varepsilon}_k})$ obeys the LDP in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$ (with rate function $\overline{L}_T(X, \nu)$).
(Puhalskii [15-16].)

Definition 3 The family $(X^\varepsilon, \nu^\varepsilon)$ is said to be obey the local LDP in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$ with local rate function $\widehat{L}_T(X, \nu)$, if for any (X, ν) from $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$

$$\lim_{\delta \to 0} \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big((r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \leq \delta\big)$$
$$= \lim_{\delta \to 0} \underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big((r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \leq \delta\big)$$
$$= -\widehat{L}_T(X, \nu). \tag{3.3}$$

(Freidlin and Wentzell [2].)

The connecting component of these notions used in the proof of the next result is Puhalskii's theorem [15-16]. Below we formulate only the first part of it.

Theorem P. *If $(X^\varepsilon, \nu^\varepsilon)$ is exponentially tight family in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$, then it is LD relatively compact.*

The following result is a reformulation of Theorem 1.3 from [17].

Proposition 3.1. *The exponential tightness and the local LDP for the family $(X^\varepsilon, \nu^\varepsilon)$ in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$ imply the LDP in $(\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}, r_T \times \rho_T)$ for this family with (good) rate function $L_T(X, \nu) \equiv \widehat{L}_T(X, \nu)$, where $\widehat{L}_T(X, \nu)$ is the local rate function.*

4. Exponential tightness in $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$

Theorem 4.1. *Under assumptions (A.1), (A.2), and (A.3) the family $(X^\varepsilon, \nu^\varepsilon)$ is exponentially tight in $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$.*

Proof. Following Definition 1, (3.2) has to be checked. It is clear it takes place if

$$\lim_j \varlimsup_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\left(X^\varepsilon \in \mathbb{C}_{[0,T]} \setminus K_j'\right) = -\infty$$

$$\lim_j \varlimsup_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\left(\nu^\varepsilon \in \mathbb{M}_{[0,T]} \setminus K_j''\right) = -\infty, \tag{4.1}$$

where K_j' and K_j'' are appropriate increasing sequences of compacts from $\mathbb{C}_{[0,T]}$ and $\mathbb{M}_{[0,T]}$ respectively. It is naturally to use as compacts K_j' increasing sets of uniformly bounded and equicontinuous functions from $\mathbb{C}_{[0,T]}$ parametrized by j. Since the process $(X_t^\varepsilon, \xi_t^\varepsilon)_{t \geq 0}$ is defined on a stochastic basis with the filtration $\mathbf{F}$ one can use Aldous-Puhalskii's type sufficient conditions (see [15], and also Theorem 3.1 in [17]) for C-exponential tightness:

$$\lim_j \varlimsup_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\left(\sup_{t \leq T} |X_t^\varepsilon| > j\right) = -\infty$$

$$\lim_{\delta \to 0} \varlimsup_{\varepsilon \to 0} \varepsilon \log \sup_{\tau \leq T-\delta} \mathbf{P}\left(\sup_{t \leq \delta} |X_{\tau+t}^\varepsilon - X_\tau^\varepsilon| > \eta\right) = -\infty, \ \forall \eta > 0, \tag{4.2}$$

where τ is a stopping time w.r.t. the filtration $\mathbf{F}$. Following Theorem 3.1 in [17], (4.2) implies the validity of the first part in (4.1) with above-mentioned compacts K_j' of uniformly bounded and equicontinuous functions. Now, choose relevant compacts $K_j'', j \geq 1$:

$$K_j'' = \bigcap_{m \geq j} \left\{\nu \in \mathbb{M}_{[0,T]} : \int_0^T \int_{|z|>m} \nu(dt, dz) \leq g(m)\right\}, \tag{4.3}$$

where $g(y), y > 0$ is positive continuous decreasing function with $\lim_{y \to \infty} g(y) = 0$. In fact, if $\nu_k \in K_j'', k \geq 1$ then we have for any $m \geq j$ $\sup_k \int_0^T \int_{|z| > m} \nu_k(dt, dz) \leq g(m)$ that is the set K_j'' is tight and by Prohorov's theorem (see [22]) is relatively compact. On the other hand, since the set $\{z : |z| > m\}$ is open a limit of any converging sequence from K_j'' also belongs to K_j'' that is K_j'' is compact in $(\mathbb{M}_{[0,T]}, \rho_T)$. Evidently $K_j'' \subseteq K_{j+1}''$. Below we choose a special function $g(y)$, suited to assumption (A.3), to satisfy the second part in (4.1).

We check the validity of (4.1) in the next two lemmas.

Lemma 4.1. *Under (A.1) the first relation in (4.1) holds.*

Lemma 4.2. *Under (A.2) and (A.3) the second relation in (4.1) holds.*

Proof of Lemma 4.1. Put
$$Z_t^* = \sup_{t' \leq t} |Z_{t'}|.$$

By virtue of (A.1) we have $|A(x, z)| \leq \ell(1 + |x|)$. Therefore, with $t \leq T$, we derive from (1.1)
$$X_t^{\varepsilon *} \leq |x_0| + \ell \int_0^t (1 + X_s^{\varepsilon *}) ds + \sqrt{\varepsilon} M_T^{\varepsilon *}, \tag{4.4}$$

where $M_t^\varepsilon = \int_0^t B(X_s^\varepsilon, \xi_s^\varepsilon) dW_s$. Due to Bellman-Gronwall's inequality, (4.4) implies $X_T^{\varepsilon *} \leq$ const.$(1 + \sqrt{\varepsilon} M_T^{\varepsilon *})$ with const., depending only on $|x_0|, \ell$, and T. Therefore, the first part of (4.2) holds if
$$\lim_j \overline{\lim_{\varepsilon \to 0}} \varepsilon \log \mathbf{P}(M_T^{\varepsilon *} > j) = -\infty. \tag{4.5}$$

On the other hand, by Chebyshev's inequality $\mathbf{P}(M_T^{\varepsilon *} > j) \leq j^{-1/\varepsilon} \mathbf{E}(M_T^{\varepsilon *})^{1/\varepsilon}$ and so, $\varepsilon \log \mathbf{P}(M_T^{\varepsilon *} > j) \leq -\log j + \varepsilon \log \mathbf{E}(M_T^{\varepsilon *})^{1/\varepsilon}$. Thereby (4.5) holds if
$$\overline{\lim_{\varepsilon \to 0}} \varepsilon \log \mathbf{E}(M_T^{\varepsilon *})^{1/\varepsilon} < \infty. \tag{4.6}$$

Below we check the validity (4.6). Assuming $1/\varepsilon > 2$ and applying Itô's formula to $|M_t^\varepsilon|^{1/\varepsilon}$, we get

$$|M_t^\varepsilon|^{1/\varepsilon} = \frac{1}{\sqrt{\varepsilon}} \int_0^t |M_s^\varepsilon|^{1/\varepsilon - 1} (\text{sign } M_s^\varepsilon) B(X_s^\varepsilon, \xi_s^\varepsilon) dW_s$$
$$+ \frac{1 - \varepsilon}{2\varepsilon} \int_0^t |M_s^\varepsilon|^{1/\varepsilon - 2} B^2(X_s^\varepsilon, \xi_s^\varepsilon) ds$$

that is $|M_t^\varepsilon|^{1/\varepsilon}$ is a submartingale obeying a decomposition: $|M_t^\varepsilon|^{1/\varepsilon} = N_t^\varepsilon + U_t^\varepsilon$ with local martingale N_t^ε and predictable increasing process
$$U_t^\varepsilon = \frac{1 - \varepsilon}{2\varepsilon} \int_0^t |M_s^\varepsilon|^{1/\varepsilon - 2} B^2(X_s^\varepsilon, \xi_s^\varepsilon) ds. \tag{4.7}$$

Then, due to a modification of Doob's inequality (see [23], Theorem 1.9.2)

$$\mathbf{E}\big(M_t^{\varepsilon*}\big)^{1/\varepsilon} \leq \Big(\frac{1}{1-\varepsilon}\Big)^{1/\varepsilon} \mathbf{E}U_t^{\varepsilon}. \qquad (4.8)$$

Now evaluate from above $|M_s^{\varepsilon}|^{1/\varepsilon-2}B^2(X_s^{\varepsilon},\xi_s^{\varepsilon})$. By virtue of (A.1) $|B(x,z)| \leq \ell(1+|x|)$. Thereby, due to above-mentioned upper bound $X_T^{\varepsilon*} \leq$ const.$(1+M_T^{\varepsilon*})$ which remains true with replacing T on s for any $s < T$, we arrive at

$$|M_s^{\varepsilon}|^{1/\varepsilon-2}B^2(X_s^{\varepsilon},\xi_s^{\varepsilon}) \leq \text{const.}(1+|M_s^{\varepsilon}|^{1/\varepsilon-2}+|M_s^{\varepsilon}|^{1/\varepsilon})$$
$$\leq \text{const.}(1+(M_s^{\varepsilon*})^{1/\varepsilon}).$$

Substituting the last upper bound in (4.7) and using (4.8) we obtain ($t \leq T$) $\mathbf{E}(M_t^{\varepsilon*})^{1/\varepsilon} \leq \frac{\text{const.}}{\varepsilon} \int_0^t \big[1 + \mathbf{E}(M_t^{\varepsilon*})^{1/\varepsilon}\big]ds$. Hence, by Bellman-Gronwall's inequality, an upper bound $\mathbf{E}(M_T^{\varepsilon*})^{1/\varepsilon} \leq \frac{\text{const.}T}{\varepsilon} \exp\{\frac{\text{const.}T}{\varepsilon}\}$ holds and implies (4.6). Consequently the first part in (4.2) is valid. To check the second part in (4.2), first use obvious estimates:

$$\mathbf{P}\big(\sup_{t\leq\delta}|X_{\tau+t}^{\varepsilon} - X_{\tau}^{\varepsilon}| > \eta\big) \leq \mathbf{P}\big(\sup_{t\leq\delta}|X_{\tau+t}^{\varepsilon} - X_{\tau}^{\varepsilon}| > \eta, X_T^{\varepsilon*} \leq j\big) + \mathbf{P}\big(X_T^{\varepsilon*} > j\big)$$

$$\leq 2\max\Big[\mathbf{P}\big(\sup_{t\leq\delta}|X_{\tau+t}^{\varepsilon} - X_{\tau}^{\varepsilon}| > \eta, X_T^{\varepsilon*} \leq j\big), \mathbf{P}\big(X_T^{\varepsilon*} > j\big)\Big].$$

Thence, due to proved above the first part of (4.2), the validity of the second part follows if

$$\lim_{\delta\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon\log\sup_{\tau\leq T-\delta}\mathbf{P}\big(\sup_{t\leq\delta}|X_{\tau+t}^{\varepsilon} - X_{\tau}^{\varepsilon}| > \eta, X_T^{\varepsilon*} \leq j\big) = -\infty, \; j \geq 1, \eta > 0. \qquad (4.9)$$

The simplest way for verifying of (4.9) consists in checking the validity of both

$$\lim_{\delta\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon\log\sup_{\tau\leq T-\delta}\mathbf{P}\Big(\sup_{t\leq\delta}\Big|\int_{\tau}^{\tau+t} A(X_s^{\varepsilon},\xi_s^{\varepsilon})ds\Big| > \eta, X_T^{\varepsilon*} \leq j\Big) = -\infty$$

$$\lim_{\delta\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon\log\sup_{\tau\leq T-\delta}\mathbf{P}\Big(\sup_{t\leq\delta}\Big|\sqrt{\varepsilon}\int_{\tau}^{\tau+t} B(X_s^{\varepsilon},\xi_s^{\varepsilon})dW_s\Big| > \eta, X_T^{\varepsilon*} \leq j\Big) = -\infty. \qquad (4.10)$$

Obviously, the first part in (4.10) holds. To verify the second, note that the process $Y_t^{\varepsilon} = \sqrt{\varepsilon}\int_{\tau}^{\tau+t} B(X_s^{\varepsilon},\xi_s^{\varepsilon})dW_s$ is continuous martingale w.r.t. the new filtration $\mathbf{F}^{\tau} = (\mathcal{F}_{\tau+t})_{t\geq 0}$ (see Ch.4, §7 in [23]). It has the predictable quadratic variation $\langle Y^{\varepsilon}\rangle_t = \varepsilon\int_{\tau}^{\tau+t} B^2(X_s^{\varepsilon},\xi_s^{\varepsilon})ds$. Also define a positive continuous local martingale (w.r.t. the same filtration $\mathbf{F}^{\tau}$)

$$Z_t^{\varepsilon} = \exp\big(\lambda Y_t^{\varepsilon} - \frac{1}{2}\lambda^2\langle Y^{\varepsilon}\rangle_t\big), \; \lambda \in R \qquad (4.11)$$

which is simultaneously a supermartingale (see [23], Problem 1.4.4) and so for any Markov time σ (w.r.t. $\mathbf{F}^{\tau}$) $\mathbf{E}Z_{\sigma}^{\varepsilon} \leq 1$. Take $\sigma = \inf\{t \leq \delta : |Y_t^{\varepsilon}| \geq \eta\}$. Evidently the second part of (4.10) holds if

$$\lim_{\delta\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon\log\sup_{\tau\leq T-\delta}\mathbf{P}\big(Y_{\sigma}^{\varepsilon} \geq \eta \;(\text{or}\; \leq -\eta), \sigma \leq \delta, X_T^{\varepsilon*} \leq j\big) = -\infty, \qquad (4.12)$$

By virtue of an obvious inequality $\mathbf{E}Z_\sigma^\varepsilon I\left(Y_\sigma^\varepsilon \geq \eta, X_T^{\varepsilon*} \leq j\right) \leq 1$ we find that

$$\varepsilon \log \mathbf{P}\left(Y_\sigma^\varepsilon \geq \eta, \sigma \leq \delta, X_T^{\varepsilon*} \leq j\right) \leq -\sup_{\lambda>0}\left[\lambda\eta - \text{const.}\frac{\lambda^2}{2}\delta\varepsilon\right] \tag{4.13}$$

and since $\sup_{\lambda>0}[\lambda\eta - \text{const.}\frac{\lambda^2}{2}\delta\varepsilon] = \frac{\eta^2}{2\text{const.}\delta\varepsilon}$ (4.12) with '$\geq \eta$' is implied by (4.13). The validity (4.12) with '$\leq -\eta$' is proved in the same way.

Proof of Lemma 4.2. It is clear that $\{\nu^\varepsilon \in \mathbb{M}_{[0,T]} \setminus K_j''\} = \{\ell(j,\nu^\varepsilon) < \infty\}$, where

$$\ell(j,\nu) = \min\{m \geq j : \int_0^T \int_{|z|>m} \nu(dt,dz) > g(m)\}. \tag{4.14}$$

Therefore, the second part of (4.1) is equivalent to

$$\lim_j \overline{\lim_{\varepsilon \to 0}} \varepsilon \log \mathbf{P}(\ell(j,\nu^\varepsilon) < \infty) = -\infty. \tag{4.15}$$

To verify (4.15), choose a special function $g(y)$ satisfying above-mentioned properties. To this end introduce non linear operator

$$\mathcal{D} = b(z)\frac{\partial}{\partial z} + \frac{\sigma^2(z)}{2}\left[\frac{\partial^2}{\partial z^2} + \left(\frac{\partial}{\partial z}\right)^2\right] \tag{4.16}$$

and choose a non negative twice continuously differentiable function $u(z)$ such that

$$-\sup_{v \in R} \mathcal{D}u(v) = -d > -\infty,$$
$$\lim_{j \to \infty} \inf_{|z|>j}\left[-\mathcal{D}u(z) + \sup_{v \in R} \mathcal{D}u(v)\right] = \infty. \tag{4.17}$$

Under assumptions (A.2) and (A.3) one can take any of function $u(z)$ with properties: $u(0) = 0$, $u'(z) = \text{sign } z$, $|z| > 1$, and $0 \leq u''(z) \leq 1$. With chosen $u(z)$ put

$$g(y) = \inf_{|z|>y}\left[-\mathcal{D}u(z) + \sup_v \mathcal{D}u(v)\right]^{-1/2}. \tag{4.18}$$

Introduce a positive continuous local martingale (the martingale property is checked by Itô's formula)

$$Z_t^\varepsilon = \exp\left(u(\xi_t^\varepsilon) - u(\xi_0) - \int_0^t \mathcal{D}u(\xi_s^\varepsilon)ds\right). \tag{4.19}$$

It is simultaneously a supermartingale (see Problem 1.4.4. in [23]) and so $\mathbf{E}Z_T^\varepsilon \leq 1$. The last implies

$$\mathbf{E}I(\ell(j,\nu^\varepsilon) < \infty)Z_T^\varepsilon \leq 1. \tag{4.20}$$

Inequality (4.20) can be sharpen by changing of Z_T^ε on its lower bound on the set $\{\ell(j,\nu^\varepsilon) < \infty\}$ which can be chosen non random. Taking into account that $\int_0^T \mathcal{D}u(\xi_s^\varepsilon)ds = \int_0^T \int_R \mathcal{D}u(z)\nu^\varepsilon(ds,dz)$ and $\ell(j,\nu^\varepsilon) \geq j$ we

$$\log Z_T^\varepsilon \geq -u(\xi_0) - \frac{dT}{\varepsilon} + \frac{1}{\varepsilon}\int_0^T \int_{|z|>\ell(j,\nu^\varepsilon)} \big[-\mathcal{D}u(z) + d\big]\nu^\varepsilon(ds,dz)$$

$$\geq -u(\xi_0) - \frac{dT}{\varepsilon} + \frac{1}{\varepsilon}\inf_{|z|>\ell(j,\nu^\varepsilon)}\big[-\mathcal{D}u(z)+d\big]\int_0^T\int_{|z|>\ell(j,\nu^\varepsilon)}\nu^\varepsilon(ds,dz)$$

$$\geq -u(\xi_0) - \frac{dT}{\varepsilon} + \frac{1}{\varepsilon}\inf_{|z|>\ell(j,\nu^\varepsilon)}\big[-\mathcal{D}u(z)+d\big]^{1/2} \ (= \log Z_*).$$

$$\geq -u(\xi_0) - \frac{dT}{\varepsilon} + \frac{1}{\varepsilon}\inf_{|z|>j}\big[-\mathcal{D}u(z)+d\big]^{1/2} \ (= \log Z_*).$$

Thereby, from (4.20), with Z_T^ε repalced on Z_*, we derive

$$\varepsilon\log\mathbf{P}\big(\ell(j,\nu^\varepsilon) < \infty\big) \leq \varepsilon u(\xi_0) + dT - \inf_{|z|>j}\big[-\mathcal{D}u(z)+d\big]^{1/2},$$

i.e. (4.15) is implied by (4.17).

5. Upper bound for local LDP in $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$

In this Section, we consider family $(X^\varepsilon, \nu^\varepsilon)$ from $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$. Parallel to $F(\nu)$ and $S(X,\nu)$, given in (2.3), let us define $F_T(\nu)$ and $S_T(X,\nu)$ by changing integrals '$\int_0^\infty$' in (2.3) on '$\int_0^T$'. Put

$$L_T(X,\nu) = \frac{1}{2}S_T(X,\nu) + \frac{1}{8}F_T(\nu). \tag{5.1}$$

Theorem 5.1. *Under (A.1), (A.2), and (A.3) for every (X,ν) from $\mathbb{C}_{[0,T]} \times \mathbb{M}_{[0,T]}$*

$$\varlimsup_{\delta\to 0}\varlimsup_{\varepsilon\to 0}\varepsilon\log\mathbf{P}\big(r_T(X^\varepsilon,X) + \rho_T(\nu^\varepsilon,\nu) \leq \delta\big) \leq -L_T(X,\nu).$$

Proof of this theorem is based on

Lemma 5.1. *Assume (A.1), (A.2), and (A.3). Then for every piece wise constant function $\lambda(t) = \sum_i \lambda(t_i)I(t_i \leq t < t_{i+1})$ (with not overlapping intervals $[t_i,t_{i+1})$), and for every compactly supported in z and continuously differentiable (once in t and twice in z) function $u(t,z)$, and $X \in \mathbb{C}_{[0,T]}$, $\nu \in \mathbb{M}_{[0,T]}$*

$$\varlimsup_{\delta\to 0}\varlimsup_{\varepsilon\to 0}\varepsilon\log\mathbf{P}\big(r_T(X^\varepsilon,X) + \rho_T(\nu^\varepsilon,\nu) \leq \delta\big)$$

$$\leq -\Big\{\sum_i \lambda(t_i)[X_{T\wedge t_{i+1}} - X_{T\wedge t_i}] - \int_0^T\int_R \lambda(t)A(X_t,z)\nu(dt,dz)$$

$$-\frac{1}{2}\int_0^T\int_R \lambda^2(t)B^2(X_t,z)\nu(dt,dz)\Big\} + \int_0^T\int_R \mathcal{D}u(t,z)\nu(dt,dz),$$

where $\mathcal{D}$ is the non linear operator defined in (4.16).

Proof. The following well known fact will be used hereafter. If N_t ($N_0 = 0$) is continuous local martingale and $\langle N \rangle_t$ is its predictable quadratic variation, then the exponential process $Z_t = \exp\left(N_t - (1/2)\langle N \rangle_t\right)$ is a continuous local martingale too, and what is more if N'_t, N''_t are continuous local martingales ($N'_0 = N''_0 = 0$) with the mutual predictable quadratic variation $\langle N', N'' \rangle_t \equiv 0$ and Z'_t, Z''_t are corresponding exponential processes, then the process $Z'_t Z''_t$ is also local martingale which, being positive, is a supermartingale too (Problem 1.4.4 in [23]) and so $\mathbf{E} Z'_t Z''_t \leq 1, t \geq 0$.

Let $\lambda(t)$ and $u(t, z)$ be functions involving in Lemma 5.1. Put

$$N'_t = \frac{1}{\sqrt{\varepsilon}} \int_0^t \lambda(s) B(X_s^\varepsilon, \xi_s^\varepsilon) dW_s$$

$$N''_t = \frac{1}{\sqrt{\varepsilon}} \int_0^t u'_z(s, \xi_s^\varepsilon) \sigma(\xi_s^\varepsilon) dV_s.$$

Evidently

$$\langle N' \rangle_t = \frac{1}{\varepsilon} \int_0^t \lambda^2(s) B^2(X_s^\varepsilon, \xi_s^\varepsilon) ds$$

$$\langle N'' \rangle_t = \frac{1}{\varepsilon} \int_0^t (u')_z^2(s, \xi_s^\varepsilon) \sigma^2(\xi_s^\varepsilon) ds. \tag{5.2}$$

Since Wiener processes W_t and V_t are independent and so $\langle N', N'' \rangle_t \equiv 0$ a process

$$Z_t = \exp\left(N'_t + N''_t - \frac{1}{2}\left[\langle N' \rangle_t + \langle N'' \rangle_t\right]\right) \tag{5.3}$$

is local martingale and also a supermartingale with

$$\mathbf{E} Z_t \leq 1, t \geq 0. \tag{5.4}$$

Note that

$$N'_t = \frac{1}{\varepsilon} \int_0^t \lambda(s)\left[dX_s^\varepsilon - A(X_s^\varepsilon, \xi_s^\varepsilon) ds\right] \tag{5.5}$$

and also find similar representation for N''_t. Due to Itô's formula we obtain

$$\frac{1}{\sqrt{\varepsilon}} \int_0^t u'_z(s, \xi_s^\varepsilon) \sigma(\xi_s^\varepsilon) dV_s = u(t, \xi_t^\varepsilon) - u(0, \xi_0) - \int_0^t u'_t(s, \xi_s^\varepsilon) ds - \frac{1}{\varepsilon} \int_0^t \mathcal{L}u(s, \xi_s^\varepsilon) ds,$$

where $\mathcal{L} = b(z)\frac{\partial d}{\partial z} + \frac{\sigma^2(z)}{2}\frac{\partial^2}{z^2}$ and consequently

$$N''_t = u(t, \xi_t^\varepsilon) - u(0, \xi_0) + \int_0^t u'_t(s, \xi_s^\varepsilon) ds - \frac{1}{\varepsilon} \int_0^t \mathcal{L}u(s, \xi_s^\varepsilon) ds. \tag{5.6}$$

(5.4) implies an obvious the inequqality

$$\mathbf{E}I(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \leq \delta)Z_T \leq 1. \tag{5.7}$$

which can be sharpen by changing of Z_T by its lower bound. To this end evaluate from below $\log Z_T$ on the set $\{r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \leq \delta\}$. For both $N_T' - \frac{1}{2}\langle N'\rangle_T$ and $N_T'' - \frac{1}{2}\langle N''\rangle_T$ we get

$$\begin{aligned}
N_T' - \frac{1}{2}\langle N'\rangle_T \geq{}& \frac{1}{\varepsilon}\sum_i \lambda(t_i)[X_{T\wedge t_{i+1}} - X_{T\wedge t_i}] \\
&- \frac{1}{\varepsilon}\int_0^T \int_R \lambda(t)A(X_t, z)\nu(dt, dz) - \frac{1}{2}\int_0^T \int_R \lambda^2(t)B^2(X_t, z)\nu(dt, dz) \\
&- \frac{1}{\varepsilon}\left\{\sum_i |\lambda(t_i)|\left[|X_{T\wedge t_{i+1}}^\varepsilon - X_{T\wedge t_{i+1}}| + |X_{T\wedge t_i}^\varepsilon - X_{T\wedge t_i}|\right]\right. \\
&+ \int_0^T |\lambda(t)||A(X_t^\varepsilon, \xi_t^\varepsilon) - A(X_t, \xi_t^\varepsilon)|ds \\
&+ \frac{1}{2}\int_0^T \lambda^2(t)|B^2(X_t^\varepsilon, \xi_t^\varepsilon) - B^2(X_t, \xi_t^\varepsilon)|ds \\
&+ \left.\left|\int_0^T \int_R [\lambda(t)A(X_t, z)| + \lambda^2(t)B^2(X_t, z)][\nu^\varepsilon - \nu](dz, dt)\right|\right\} \tag{5.8}
\end{aligned}$$

and

$$\begin{aligned}
N_T'' - \frac{1}{2}\langle N''\rangle_T ={}& -\frac{1}{\varepsilon}\int_0^T [\mathcal{L}u(s, \xi_s^\varepsilon) + \frac{1}{2}u_z^2(s, \xi_s^\varepsilon)]ds + u(T, \xi_T^\varepsilon) - u(0, \xi_0) - \int_0^T u_t(s, \xi_s^\varepsilon) \\
={}& -\frac{1}{\varepsilon}\int_0^T \int_R \mathcal{D}u(s, z)\nu^\varepsilon(dz, ds) + u(T, \xi_T^\varepsilon) - u(0, \xi_0) - \int_0^T u_t(s, \xi_s^\varepsilon)ds \\
\geq{}& -\frac{1}{\varepsilon}\int_0^T \int_R \mathcal{D}u(s, z)\nu(dz, ds) \\
&- \frac{1}{\varepsilon}\left\{\left|\int_0^T \int_R \mathcal{D}u(s, z)[\nu^\varepsilon - \nu](dz, ds)\right|\right. \\
&+ \left.\varepsilon|u(T, \xi_T^\varepsilon)| + \varepsilon|u(0, \xi_0)| + \varepsilon\int_0^T |u_t(s, \xi_s^\varepsilon)|ds\right\}. \tag{5.9}
\end{aligned}$$

The terms in the curly brackets in the right hand sides of (5.8) and (5.9) are random variables. Nevertheless, they can be evaluated from above on the set $\{r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon.\nu) \leq \delta\}$ by non random quantities. Evidently $\int_0^T \int_R |\lambda(t)||A(X_t^\varepsilon, \xi_t^\varepsilon) - A(X_t, \xi_t^\varepsilon)|ds \leq \text{const.}T\delta$ and $\frac{1}{2}\int_0^T \lambda^2(t)|B^2(X_t^\varepsilon, \xi_t^\varepsilon) - B^2(X_t, \xi_t^\varepsilon)|ds \leq \text{const.}\delta \int_0^T [1 + |X_t|]ds$. Denote by $H(s, z) = \lambda(s)A(X_s, z) + \frac{\lambda^2(s)B^2(X_s, z)}{2}]$. Since $\lambda(s)$ is piece wise constant function without loss of a generality one can assume that it is simply constant. Then function $H(s, z)$ is bounded

continuous function and so, by Lemma A.1 (see Appendix) for any $\gamma > 0$ and $k \geq 1$ there exist increasing continuous function $h_k^\gamma(y), y \geq 0$ with $h_k^\gamma(0) = 0$ and decreasing sequence $\varphi_k, k \geq 1$ with $\lim\limits_k \varphi_k = 0$ both dependent on $H(s, z)$ and ν only such that

$$\Big| \int_0^T \int_R H(s, z)[\nu^\varepsilon - \nu](ds, dz) \Big| \leq \gamma + h_k^\gamma(\delta) + \varphi_k.$$

Further, by the remark to Lemma A.1

$$\Big| \int_0^T \int_R \mathcal{D}u(t, z)[\nu^\varepsilon - \nu](dt, dz) \Big| \leq \gamma + h^\gamma(\delta),$$

where $h^\gamma(y)$ is an icreasing continuos function with $h^\gamma(0) = 0$ depending on $\mathcal{D}u(s, z)$ and ν only.

Hence, we the lower bounds (with positive const.'s):

$$N_T' - \frac{1}{2}\langle N' \rangle_T \geq \frac{1}{\varepsilon}\Big[\sum_i \lambda(t_i)[X_{T \wedge t_{i+1}} - X_{T \wedge t_i}] - \int_0^T \int_R \lambda(t)A(X_t, z)\nu(dt, dz)$$

$$- \int_0^T \int_R \lambda^2(t)B^2(X_t, z)\nu(dt, dz)\Big]$$

$$- \frac{\text{const.}}{\varepsilon}\Big(\gamma + h_k^\gamma(\delta) + \varphi_k\Big)\Big) \tag{5.10}$$

and

$$N_T'' - \frac{1}{2}\langle N'' \rangle_T \geq -\frac{1}{\varepsilon}\int_0^T \int_R \mathcal{D}u(t, z)\nu(dt, dz) - \frac{\text{const.}}{\varepsilon}\Big(\varepsilon + \gamma + h^\gamma(\delta) + \varphi_k\Big). \tag{5.11}$$

By virtue of (5.10) and (5.11) one can choose a non random lower bound:

$$\log Z_T \geq \frac{1}{\varepsilon}\Big[\sum_i \lambda(t_i)[X_{T \wedge t_{i+1}} - X_{T \wedge t_i}] - \int_0^T \int_R \lambda(t)A(X_t, z)\nu(dt, dz)$$

$$- \int_0^T \int_R \lambda^2(t)B^2(X_t, z)\nu(dt, dz) - \frac{1}{\varepsilon}\int_0^T \int_R \mathcal{D}u(t, z)\nu(dt, dz)\Big]$$

$$- \frac{\text{const.}}{\varepsilon}\Big(\varepsilon + \gamma + h^\gamma(\delta) + h_k^\gamma(\delta) + \varphi_k\Big).$$

$$= \log Z_*.$$

Hence and from (5.7), with replacing of Z_T on Z_*, it follows

$$\varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho(\nu^\varepsilon, \nu) \leq \delta \big)$$

$$\leq -\Big[\sum_i \lambda(t_i)[X_{T \wedge t_{i+1}} - X_{T \wedge t_i}] - \int_0^T \int_R \lambda(t)A(X_t, z)\nu(dt, dz)$$

$$- \frac{1}{2}\int_0^T \int_R \lambda^2(t)B^2(X_t, z)\nu(dt, dz)\Big] + \int_0^T \int_R \mathcal{D}u(t, z)\nu(dt, dz)$$

$$+ \text{const.}\Big\{\big(\varepsilon + \gamma + h^\gamma(\delta) + h_k^\gamma(\delta) + \varphi_k\big)\Big\}, \tag{5.12}$$

The desired result holds since the term in the curly brackets of the right hand side in (5.12) goes to zero if limit '$lim_{k\to\infty}lim_{\gamma\to 0}lim_{\delta\to 0}\lim_{\varepsilon\to 0}$' is taken.

Proof of Theorem 5.1. follows from Lemmas 5.1, A.2, and A.3 (see Appendix) since

$$
-\sup_{\lambda}\left\{\sum_{i}\lambda(t_i)[X_{T\wedge t_{i+1}}-X_{T\wedge t_i}]-\int_0^T\int_R\lambda(t)A(X_t,z)\nu(dt,dz)\right.
$$
$$
\left.-\frac{1}{2}\int_0^T\int_R\lambda^2(t)B^2(X_t,z)\nu(dt,dz)\right\}+\inf_u\int_0^T\int_R\mathcal{D}u(t,z)\nu(dt,dz)
$$
$$
=-\left[\frac{1}{2}S_T(X,\nu)+\frac{1}{8}F_T(\nu)\right]=-L_T(X,\nu).
$$

6. Lower bound for local LDP in $\mathbb{C}_{[0,T]}\times\mathbb{M}_{[0,T]}$

Theorem 6.1. *Under (A.1), (A.2), and (A.3), for every (X,ν) from $\mathbb{C}_{[0,T]}\times\in\mathbb{M}_{[0,T]}$*

$$
\underline{\lim}_{\delta\to 0}\underline{\lim}_{\varepsilon\to 0}\varepsilon\log\mathbf{P}\big(r_T(X^\varepsilon,X)+\rho_T(\nu^\varepsilon,\nu)\le\delta\big)\ge-L_T(X,\nu).
$$

Evidently for X,ν such that $L_T(X,\nu)=\infty$ it is nothing to prove. Therefore below we consider only the case $L_T(X,\nu)<\infty$ which distinguishes subsets from $\mathbb{C}_{[0,T]}\times\mathbb{M}_{[0,T]}$:

(i) $dX_t\ll dt$ and $\frac{1}{2}S_T(X,\nu)=\frac{1}{2}\int_0^T\frac{[\dot{X}_t-A_\nu(t,X_t)]^2}{B_\nu^2(t,X_t)}dt<\infty$;

(ii) $d\nu=nd\lambda$, $d_z n=n_z' dz$ and $\frac{1}{8}F_T(\nu)=\frac{1}{2}\int_0^T\int_R\frac{v_\nu^2(t,z)}{\sigma^2(z)}n(t,z)dzdt<\infty$, where

$$
v_\nu(t,z)=\frac{\sigma^2(z)}{2}\left[\frac{n_z'(t,z)}{n(t,z)}-\frac{p'(z)}{p(z)}\right]. \tag{6.1}
$$

It is convenient to consider further subset (ii') of (ii):

(ii') the function $v_\nu(t,z)$ is compactly supported in z and continuously differentiable in (t,z), having bounded partial derivatives.

The central role in proving Theorem 6.1 plays

Lemma 6.1. *Assume (i), (ii'), and $\inf_{x,z}B^2(x,z)>0$. Then for any $\delta>0$ and $\gamma>0$ there exists an increasing continuous function $h_\gamma(y)$ with $h_\gamma(0)=0$, depending on $\frac{v_\nu^2(s,z)}{\sigma^2(z)}$ and ν only, such that*

$$
\underline{\lim}_{\varepsilon\to 0}\varepsilon\log\mathbf{P}\big(r_T(X^\varepsilon,X)+\rho_T(\nu^\varepsilon,\nu)\le\delta\big)\ge-L_T(X,\nu)-\gamma-h_\gamma(\delta).
$$

Proof. Put

$$
b_\nu(t,z)=b(z)+v_\nu(t,z)
$$
$$
G_\nu(t,x,z)=\frac{\dot{X}_t-A_\nu(t,X_t)}{B_\nu(t,X_t)}B(x,z)+A(x,z) \tag{6.2}
$$

and parallel to $(X_t^\varepsilon, \xi_t^\varepsilon)$ introduce, on the same stochastic basis, new diffusion pair $(\widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)$:

$$d\widetilde{X}_t^\varepsilon = G_\nu(t, \widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)dt + \sqrt{\varepsilon}B(\widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)dW_t$$

$$d\widetilde{\xi}_t^\varepsilon = \frac{1}{\varepsilon}b_\nu(t, \widetilde{\xi}_t^\varepsilon)dt + \frac{1}{\sqrt{\varepsilon}}\sigma(\widetilde{\xi}_t^\varepsilon)dV_t \tag{6.3}$$

subject to the same initial point (x_0, ξ_0). Also denote by $\widetilde{\nu}^\varepsilon(dt, dz)$ the occupation measure corresponding to $\widetilde{\xi}^\varepsilon$: $\widetilde{\nu}^\varepsilon(\Delta \times \Gamma) = \int_0^\infty I(t \in \Delta, \widetilde{\xi}_t^\varepsilon \in \Gamma)dt$.

By virtue of the formula $b(z) = \frac{1}{2}\frac{p'(z)}{p(z)} + \sigma'(z)\sigma(z)$ (see (2.1)) we get $\frac{2b_\nu(t,z)}{\sigma^2(z)} = \frac{n'_z(t,z)}{n(t,z)} + 2\frac{\sigma'(z)}{\sigma(z)}$ and so $p_\nu(t, z) = c(t)\frac{\exp\left(2\int_0^z \frac{b_\nu(t,y)}{\sigma^2(y)}dy\right)}{\sigma^2(z)}$, with norming constant $c(t)$ such that $\int_R p_\nu(t, z) = 1$, coincides with $n(t, z)$. Then by Lemma A.5 (see Appendix)

$$\mathbf{P} - \lim_{\varepsilon \to 0} \rho_T(\widetilde{\nu}^\varepsilon, \nu) = 0 \quad \text{and} \quad \mathbf{P} - \lim_{\varepsilon \to 0} r_T(\widetilde{X}^\varepsilon, X) = 0. \tag{6.4}$$

Denote by Q^ε and $\widetilde{Q}^\varepsilon$ distributions of $(X_t^\varepsilon, \xi_t^\varepsilon)_{t \leq T}$, $(\widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)_{t \leq T}$ respectively. By Theorem 7.18 (Ch. 7 in [24] Q^ε is absolutely continuous w.r.t. $\widetilde{Q}^\varepsilon$ and

$$\frac{dQ^\varepsilon}{d\widetilde{Q}^\varepsilon}(\widetilde{X}^\varepsilon, \widetilde{\xi}^\varepsilon) = \exp\left(\frac{1}{\sqrt{\varepsilon}}M_T^\varepsilon - \frac{1}{2\varepsilon}\langle M^\varepsilon\rangle_T + \frac{1}{\sqrt{\varepsilon}}M_T - \frac{1}{2\varepsilon}\langle M\rangle_T\right), \tag{6.5}$$

where

$$M_t^\varepsilon = -\int_0^t \frac{v_\nu(s, \widetilde{\xi}_s^\varepsilon)}{\sigma(\widetilde{\xi}_s^\varepsilon)}dV_s \quad \text{and} \quad M_t = -\int_0^t \frac{\dot{X}_s - A_\nu(s, X_s)}{B_\nu(s, X_s)}dW_s$$

$$\langle M^\varepsilon\rangle_t = \int_0^t \frac{v_\nu^2(s, \widetilde{\xi}_s^\varepsilon)}{\sigma^2(\widetilde{\xi}_s^\varepsilon)}ds \quad \text{and} \quad \langle M\rangle_t = \int_0^t \frac{[\dot{X}_s - A_\nu(s, X_s)]^2}{B_\nu^2(s, X_s)}ds.$$

By virtue of (6.5) and the rule of changing for probability measure we obtain

$$\mathbf{P}\left(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \leq \delta\right) = \mathbf{E}\left[\frac{dQ^\varepsilon}{d\widetilde{Q}^\varepsilon}(\widetilde{X}^\varepsilon, \widetilde{\xi}^\varepsilon)I(r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \leq \delta)\right]. \tag{6.6}$$

The desired lover bound can be derived from (6.6) provided that a relevant lover bound for the right hand side of (6.6) can be found. Use an obvious inequality:

$$I\left(r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \leq \delta\right) \geq I(r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \leq \delta, \ |M_T^\varepsilon| \leq k, \ |M_T| \leq k)$$

and estimate from below $\log \frac{dQ^\varepsilon}{d\widetilde{Q}^\varepsilon}(\widetilde{X}^\varepsilon, \widetilde{\xi}^\varepsilon)$ on the set $\{r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \leq \delta\} \cap \{|M_T^\varepsilon| \leq k\} \cap \{|M_T| \leq k\}$. Noticing that $\frac{1}{2}\langle M\rangle_T = \frac{1}{2}S(X, \nu)_T$ and $\frac{1}{2}\int_0^T \int_R \frac{v_\nu^2(s,z)}{\sigma^2(z)}n(t, z)dzdt = \frac{1}{8}F_T(\nu)$ we obtain

$$\log \frac{dQ^\varepsilon}{d\widetilde{Q}^\varepsilon}(\widetilde{X}^\varepsilon, \widetilde{\xi}^\varepsilon) \geq -\frac{2k}{\sqrt{\varepsilon}} - \frac{1}{\varepsilon}L_T(X, \nu) - \frac{1}{2\varepsilon}\left|\int_0^T \int_R \frac{v_\nu^2(s, z)}{\sigma^2(z)}[\widetilde{\nu}^\varepsilon(dt, dz) - n(t, z)dzdt]\right|.$$

By the remark to Lemma A.1 (see Appendix), for any $\gamma > 0$ there exists increasing continuous function $h_\gamma(y)$ with $h\gamma(0) = 0$, depending on $\frac{v_\nu^2(s,z)}{\sigma^2(z)}$ and ν only, such that

$$\frac{1}{2}\left| \int_0^T \int_R \frac{v_\nu^2(s,z)}{\sigma^2(z)} [\widetilde{\nu}^\varepsilon(dt,dz) - n(t,z)dzdt \right| \le \gamma + h_\gamma(\delta).$$

Then the lower bound for the right hand side of (6.6) is the following: $-\frac{2k}{\sqrt{\varepsilon}} - \frac{1}{\varepsilon}[L_T(X,\nu) + \gamma + h_\gamma(\delta)]$. It implies

$$\varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta\big) \ge -L_T(X,\nu) - 2k\sqrt{\varepsilon} - \gamma - h_\gamma(\delta)$$
$$+ \varepsilon \log \mathbf{P}\big(r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \le \delta, \ |M_T^\varepsilon| \le k, \ |M_T| \le k\big).$$

Thus the statement of the lemma holds since

$$\lim_k \lim_{\varepsilon \to 0} \mathbf{P}\big(r_T(\widetilde{X}^\varepsilon, X) + \rho_T(\widetilde{\nu}^\varepsilon, \nu) \le \delta, \ |M_T^\varepsilon| \le k, \ |M_T| \le k\big) = 1$$

what follows by virtue of (6.4), obvious $\lim_k \mathbf{P}(|M_T| > k) = 0$ and

$$\mathbf{P}(|M_T^\varepsilon| > k) \le \frac{\mathbf{E}|M_T^\varepsilon|^2}{k^2} = \frac{\mathbf{E}\langle M^\varepsilon\rangle_T}{k^2} \le \frac{\text{const.}}{k^2} \to 0, k \to \infty.$$

Proof of Theorem 6.1. Assume (i), (ii), and $\inf_{x,z} B^2(x,z) > 0$. Due to Lemma A.4 (see Appendix), one can choose a sequence $\nu^{(k)}, k \ge 1$ of measures such that for every k the function $v_{\nu^{(k)}}(t,z)$ satisfies (ii') and what is more $\rho(\nu, \nu^{(k)}) \to 0$, $L_T(X, \nu^{(k)}) \to L_T(X,\nu)$. On the other hand, by Lemma 6.1 for any $\delta > 0$ and $\gamma > 0$ there exist increasing continuous function $h_{\gamma,k}(y)$ with $h_{\gamma,k}(0) = 0$, depending on $\nu^{(k)}$, such that

$$\underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu^{(k)}) \le \delta\big) \ge -L_T(X, \nu^{(k)}) - \gamma - h_{\gamma,k}(\delta).$$

Choose $k_\circ(\delta)$ such that for any $k \ge k_\circ(\delta)$ we have $0 < \delta - \rho_T(\nu, \nu^{(k)}) \le \delta/2$. Then, taking into account the triangular inequality: $\rho_T(\nu^\varepsilon, \nu) \le \rho_T(\nu^\varepsilon, \nu^{(k)}) + \rho_T(\nu, \nu^{(k)})$, we arrive at a lower bound:

$$\underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta\big)$$
$$\ge \underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu^{(k)}) \le \delta/2\big)$$
$$\ge -L_T(X, \nu^{(k)}) - \gamma - h_{\gamma,k}(\delta/2).$$

The right hand side of the last inequality converges to $-L_T(X,\nu)$ if limit '$lim_k lim_{\gamma \to 0} lim_{\delta \to 0}$' is taken.

Assume only (i) and (ii). Parallel to the process X_t^ε introduce new diffusion $X_t^{\varepsilon,\beta}, \beta \ne 0$:

$$dX_t^{\varepsilon,\beta} = A(X_t^{\varepsilon,\beta}, \xi_t^\varepsilon)dt + \sqrt{\varepsilon}\big[B(X_t^{\varepsilon,\beta}, \xi_t^\varepsilon)dW_t + \beta dW_t'\big]$$

subject to the same initial point x_0, where W_t' is a Wiener process independent of (W_t, ξ_t^ε). The diffusion parameter here is $B^2(x,z) + \beta^2$ and so, due to proved above,

$$\underline{\lim}_{\delta \to 0} \underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(r_T(X^{\varepsilon,\beta}, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta\big) \ge -L_T^\beta(X, \nu),$$

where $L_T^\beta(X, \nu) = \frac{1}{2} S_T^\beta(X, \nu) + \frac{1}{8} F_T(\nu)$, and where

$$S_T^\beta(X, \nu) = \int_0^\infty \frac{[\dot{X}_t - A_\nu(y, X_t)]^2}{B_\nu^2(t, X_t) + \beta^2} dt.$$

Evidently $\lim_{\beta \to 0} S_T^\beta(X, \nu) = S_T(X, \nu)$. On the other hand, by Lemma A.6 (see Appendix)

$$\lim_{\beta \to 0} \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(r_T(X^{\varepsilon,\beta}, X^\varepsilon) > \eta) = -\infty.$$

To get the desired result, we combine both these facts. Namely, using the triangular inequality: $r_T(X^\varepsilon, X) \le r_T(X^{\varepsilon,\beta}, X^\varepsilon) + r_T(X^{\varepsilon,\beta}, X)$ and taking $\eta = \delta/2$ we arrive at an upper bound

$$\begin{aligned} \mathbf{P}&\big(r_T(X^{\varepsilon,\beta}, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta\big) \\ &\le \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta/2\big) \\ &+ \mathbf{P}\big(r_T(X^{\varepsilon,\beta}, X^\varepsilon) > \delta/2\big) \\ &\le 2 \max\Big[\mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta/2\big), \mathbf{P}\big(r_T(X^{\varepsilon,\beta}, X^\varepsilon) > \delta/2\big)\Big] \end{aligned}$$

which implies

$$\underline{\lim}_{\delta \to 0} \underline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(r_T(X^\varepsilon, X) + \rho_T(\nu^\varepsilon, \nu) \le \delta\big) \ge -\lim_{\beta \to 0} L_T^\beta(X, \nu) = -L_T(X, \nu).$$

Other approach for establishing lower bound with singular diffusion parameter can be found in Puhalskii [25].

7. Proof of main result

Proof of Theorem 3.1. Due to Theorems 4.1, 5.1, and Proposition 3.1 the family $(X^\varepsilon, \nu^\varepsilon)$ obeys the LDP in $(\mathbb{C}_{[0,k]} \times \mathbb{M}_{[0,k]}, r_k \times \rho_k)$ with rate function $L_k(X, \nu)$. Then it obeys the LDP in the metric space $(\mathbb{C} \times \mathbb{M}, r \times \rho)$ with rate function $\sup_k L_k(X, \nu) = L(X, \nu)$.

Proof of Corollary 2.1. The result holds since $\inf_{X \in \mathbb{C}} S(X, \nu)$ is attained at X_t°, being a solution of a differential equation: $\dot{X}_t = A_\nu(t, X_t^\circ)$ subject to $X_0^\circ = x_0$, and so $S(X^\circ, \nu) = 0$.

Proof of Corollary 2.2. The first statement is obvious.

Assume $B^2(x,z) \equiv 0$. In this case $S(X,\nu) = 0$ for any X_t being a solution of a differential equation $\dot{X}_t = \int_R A(X_t,z)n(t,z)dz$ subject to $X_0 = x_0$; otherwise $S(X,\nu) = \infty$. Therefore

$$L(X,\nu) = \begin{cases} \frac{1}{8}\inf_{\nu:\dot{X}_t=\int_R A(X_t,z)n(t,z)dz, X_0=x_0} F(\nu) \\ \infty, \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise.} \end{cases}$$

On the other hand, since $F(\nu) < \infty$ implies $d\nu = nd\lambda$, $d_z n = n'_z dz$, assuming measurability in t of function

$$H(t,\dot{X}_t, X_t) = \inf_{\nu:\dot{X}_t=\int_R A(X_t,z)n(t,z)dz, X_0=x_0} \int_R \sigma^2(z)\Big[\frac{n'_z(t,z)}{n(t,z)} - \frac{p'(z)}{p(z)}\Big]^2 n(t,z)dz \quad (7.1)$$

we arrive at independent of t function $H(t,y,x) \equiv H(y,x)$, or by other words, 'inf' in (7.1) can be taken over all measures ν with densities $n(t,z) \equiv m(z)$. The last means the desired result holds if the function

$$H(y,x) = \inf_{m:\begin{cases} dm = m'dz \\ y = \int_R A(x,z)m(z)dz \end{cases}} \int_R \sigma^2(z)\Big[\frac{m'(z)}{m(z)} - \frac{p'(z)}{p(z)}\Big]^2 m(z)dz \quad (7.2)$$

is measurable. We check this by showing that level sets of $H(y,x)$ are closed.

Let $c \geq 0$ be fixed and $(y_n, x_n), n \geq 1$ be a sequence from $\{(y,x) : H(y,x) \leq c\}$ converging to a limit point (y_0, x_0). Show that $H(y_0, x_0) \leq c$. By virtue of assumption (A.1) the set $\mathcal{A}(y,x) = \{m : y = \int_R A(x,z)m(z)dz\}$ is closed in the Levy-Prohorov metric that is for every fixed (y,x) there exists a density $m^{(y,x)}$ from $\mathcal{A}(y,x)$ such that

$$H(y,x) = \begin{cases} \int_R \sigma^2(z)\Big[\frac{(m^{(y,x)}(z))'}{m^{(y,x)}(z)} - \frac{p'(z)}{p(z)}\Big]^2 m^{(y,x)}(z)dz, & dm^{(y,x)} = (m^{(y,x)})'dz \\ \infty, & \text{otherwise.} \end{cases} \quad (7.3)$$

Note that the function $H(y,x)$, defined in (7.3), obeys a following property: there exists a measure $\nu^{(y,x)}$ from $\mathbb{M}_{[0,1]}$, having density $m^{(y,x)}(z)$ w.r.t. $dtdz$, such that $H(y,x) = F_1(\nu^{(y,x)})$. Since $\frac{1}{8}F_1(\nu)$ is good rate function level sets $\{y,x : H(y,x) \leq c\}$ are compacts. Therefore $H(y_0, x_0) \leq c$.

Appendix

1. Evaluation via Levy-Prohorov's metric.

Lemma A.1. *Let $T > 0$, $\nu', \nu'' \in \mathbb{M}_{[0,T]}$, $\rho_T(\nu', \nu'') = q$, and $f(t,z)$ be bounded continuous function. Then for any $\gamma > 0$ and $k \geq 1$ one can choose increasing continuous function $h_k^\gamma(y), y \geq 0$ with $h_k^\gamma(0) = 0$ and decreasing sequence $\varphi_k, k \geq 1$ with $\lim_k \varphi_k = 0$ both depending on $f(t,z)$ and only from one of ν' or ν'' such that*

$$\Big|\int_0^T \int_R f(t,z)[\nu' - \nu''](dt,dz)\Big| \leq \gamma + h_k^\gamma(q) + \varphi_k.$$

Remark. *If $f(t,z)$ is bounded compactly supported continuos function, then the statement of the lemma remains true with $h_k^\gamma(y) \equiv h^\gamma(y)$ and $\varphi_k \equiv 0$.*

Proof. Assume $f(t,z)$ is continuously differentiable (one in z and twice in (t,z)) and compactly supported in z. Denote by $F'(t,z) = \nu'([0,t] \times (-\infty, z])$ that is $F'(t,z)$ is the distribution function correspoding to ν'. Integrating by parts we get

$$\int_0^T \int_R f(t,z)\nu'(dt,dz) = -\int_0^T \int_R \left[\frac{\partial f(t,z)}{\partial z} + \frac{\partial^2 f(t,z)}{\partial t \partial z} \right] F'(t,z)dzdt$$

and consequently (F'' is the distribution functions corresponding to ν'')

$$\left| \int_0^T \int_R f(t,z)[\nu' - \nu''](dt,dz) \right| \leq \int_0^T \int_R |F'(t,z) - F''(t,z)|m(t,z)dzdt,$$

where $m(t,z) = |\frac{\partial}{\partial z}f(t,z)| + |\frac{\partial^2}{\partial t \partial z}f(t,z)|$.

Assume $f(t,z)$ is compactly supported in z and continuous only. Then, approximating it by compactly supported and continuously differentiable in z function $f^\gamma(t,z)$ in a sense $\sup_{t,z} |f(t,z) - f^\gamma(t,z)| \leq \frac{\gamma}{2T}$, due the foregoing proof, we get

$$\left| \int_0^T \int_R f(t,z)[\nu' - \nu''](dt,dz) \right| \leq \gamma + \int_0^T \int_R |F'(t,z) - F''(t,z)|m^\gamma(t,z)dzdt$$

with $m^\gamma(t,z) = |\frac{\partial}{\partial z}f^\gamma(t,z)| + |\frac{\partial^2}{\partial t \partial z}f^\gamma(t,z)|$

In the general case, one can choose a decomposition $f(t,z) = f_k(t,z) + g_k(t,z)$, where $f_k(t,z)$ is continuous compactly supported in z on the interval $[-k,k]$ function while $g_k(t,z) \equiv 0$ on the interval $[-(k-1/2),(k-1/2)]$ and is bounded: $|g_k(t,z)| \leq L$. Then by foregoing result we get

$$\left| \int_0^T \int_R f(t,z)[\nu' - \nu''](dt,dz) \right| \leq \gamma + \int_0^T \int_R |F'(t,z) - F''(t,z)|m_k^\gamma(t,z)dzdt$$
$$+ L \int_0^T \int_{|z|>k-1/2} [\nu' + \nu''](dt,dz),$$

where $m_k^\gamma(t,z) = |\frac{\partial}{\partial z}f_k^\gamma(t,z)| + |\frac{\partial^2}{\partial t \partial z}f_k^\gamma(t,z)|$. Evaluate from above the last integral from the right hand side. To this end, choose an increasing sequences $z_k \nearrow \infty$, $k \to \infty$ such that $z_k \leq k-1/2$ and for every k z_k and $-z_k$ are points of continuity for the distribution function $F'(T,z)$. Then

$$\int_0^T \int_{|z|>k-1/2} [\nu' + \nu''](dt,dz) \leq 2\int_0^T \int_{|z|>z_k} \nu'(dt,dz)$$
$$+ \left| \int_0^T \int_{|z|>z_k} [\nu' - \nu''](dt,dz) \right|$$
$$\leq 2\int_0^T \int_{|z|>z_k} \nu'(dt,dz)$$
$$+ \left| F'(T,z_k) - F''(T,z_k) \right|$$
$$+ \left| F'(T,-z_k) - F''(T,-z_k) \right|.$$

Now evaluate from above $|F'(t,z) - F''(t,z)|$ via q and $F'(t,z)$. From the definition of the Levy-Prohorov metric (see e.g. [22], [26]) it follows: $q + F'(t - q, z - q) - F'(t,z) \leq F'(t,z) - F''(t,z) \leq q + F'(t + q, z + q) - F'(t,z)$ and so

$$|F'(t,z) - F''(t,z)| \leq q + [F'(t + q, z + q) - F'(t - q, z - q)].$$

Hence, combining all obtained upper estimates, we arrive at the desired result with

$$h_k^\gamma(y) = y\left(2L + \int_0^T \int_R m_k^\gamma(t,z)dtdz\right)$$

$$+ \int_0^T \int_R [F'(t + y, z + y) - F'(t - y, z - y)]m_k^\gamma(t,z)dzdt+$$

$$+ L\big|F'(T + y, z_k + y) - F'(T - y, z_k - y)\big|$$

$$+ L\big|F'(T + y, -z_k + y) - F'(T - y, -z_k - y)\big|$$

and

$$\varphi_k = 2L \int_0^T \int_{|z| > z_k} \nu'(dt, dz).$$

The same proof takes place with F'' instead of F'.

2. The Fenchel-Legendre transform.

Let $\lambda(t) = \sum_i \lambda(t_i)I(t_i \leq t < t_{i+1})$ with non overlapping intervals $[t_i, t_{i+1})$. For any $X \in \mathbb{C}_{[0,T]}$ and $\nu \in \mathbb{M}_{[0,T]}$ put $\int_0^T \lambda(t)dX_t = \sum_i \lambda(t_i)[X_{T \wedge t_{i+1}} - X_{T \wedge t_i}]$, $A_\nu(t, X_t) = \int_R A(X_t, z)K_\nu(t, dz)$, and $B_\nu^2(t, X_t) = \int_R B^2(X_t, z)K_\nu(t, dz)$. Let $\mathcal{D}$ be non linear operator defined in (4.16).

Lemma A.2. For any $X \in \mathbb{C}_{[0,T]}$ and $\nu \in \mathbb{M}_{[0,T]}$

$$\sup \int_0^T \left[\lambda(t)dX_t - (A_\nu(t, X_t) - \frac{1}{2}\lambda^2(t)B_\nu^2(t, X_t))dt\right]$$

$$= \begin{cases} \frac{1}{2} \int_0^T \frac{[\dot{X}_t - A_\nu(t, X_t)]^2}{B_\nu^2(t, X_t)} dt & dX_t = \dot{X}_t dt \\ \infty, & otherwise, \end{cases}$$

where 'sup' is taken over all piece wise constant functions $\lambda(t)$.

Lemma A.3. For any $\nu \in \mathbb{M}_{[0,T]}$

$$\inf \int_0^T \int_R \mathcal{D}u(t,z)\nu(dt, dz)$$

$$= \begin{cases} -\frac{1}{8} \int_0^T \int_R \sigma^2(z)\left[\frac{n'_z(t,z)}{n(t,z)} - \frac{p'(z)}{p(z)}\right]^2 n(t,z)dzdt, & d\nu = nd\lambda, \ d_z n = n'_z dz \\ -\infty, & otherwise, \end{cases}$$

where 'inf' is taken over all twice continuously differentiable (once in t and twice in z) compactly supported in z functions u(t, z).

Proof of Lemma A.2. For $dX_t \ll dt$ the result follows from Lemma 6.1 in [17] (see also Lemma 2.1 in [27]). For $dX_t = \dot{X}_t dt$ by lemma 6.1 [17] '$sup \int_0^T$' is equal $\int_0^T \sup_{\lambda \in R} \left\{ \lambda \left(\dot{X}_t - A_\nu(t, X_t) \right) - \frac{1}{2} \lambda^2 B_\nu^2(t, X_t) \right) \right\} dt = \frac{1}{2} \int_0^T \frac{[\dot{X}_t - A_\nu(t, X_t)]}{B_\nu^2(t, X_t)} dt.$

Proof of Lemma A.3. Assume $d\nu = nd\lambda$, $d_z n = n_z' dz$. Due to (2.1), $\frac{p'(z)}{p(z)} = \frac{2b(z) - 2\sigma(z)\sigma'(z)}{\sigma^2(z)}$ and so $b(z) = \frac{1}{2} \left[\sigma^2(z) \frac{p'(z)}{p(z)} + 2\sigma(z)\sigma'(z) \right]$. Putting $v(t, z) = u_z'(t, z)$ and taking into account the formula for $b(z)$ we arrive at

$$\int_0^T \int_R \mathcal{D}u(t, z)n(t, z)dzdt = \frac{1}{2} \int_0^T \int_R \left\{ \left[\sigma^2(z) \frac{p'(z)}{p(z)} + 2\sigma(z)\sigma'(z) \right] v(t, z) \right.$$
$$\left. + \sigma^2(z)(v_z'(t, z) + v^2(t, z)) \right\} n(t, z)dzdt. \qquad (8.1)$$

Then, integrating by parts:

$$\int_R \sigma^2(z)(v_z'(t, z)n(t, z)dz = -\int_R v(t, z)[2\sigma(z)\sigma'(z)n(t, z) + \sigma^2(z)n_z'(t, z)]dz,$$

we obtain

$$\int_0^T \int_R \mathcal{D}u(t, z)n(t, z)dzdt$$
$$= \frac{1}{2} \int_0^T \int_R \sigma^2(z) \left(v^2(t, z)n(t, z) + v(t, z) \left[\frac{p'(z)}{p(z)} n(t, z) - n_z(t, z) \right] \right) dzdt. \qquad (8.2)$$

(8.2) and the method of proving for lemma 6.1 in [17] imply

$$\inf \int_0^T \int_R \mathcal{D}u(t, z)n(t, z)dzdt = \frac{1}{2} \int_0^T \int_R \sigma^2(z) \inf_{v \in R} \left(v^2 n(t, z) \right.$$
$$\left. + v \left[\frac{p'(z)}{p(z)} n(t, z) - n_z(t, z) \right] \right) dzdt$$
$$= -\frac{1}{8} \int_0^T \int_R \sigma^2(z) \left[\frac{n_z(t, z)}{n(t, z)} - \frac{p'(z)}{p(z)} \right]^2 n(t, z)dzdt.$$

Thus for '$d\nu = nd\lambda, d_z n = n_z' dz$', the result holds.

Assume $d\nu = n\lambda, d_z n \ll dz$. Show that $\inf \int_0^T \int_R \mathcal{D}u(t, z)n(t, z)dzdt = -\infty$. To this end, take $u(t, z) \equiv u(z)$ and put $v(z) = u'(z)$. The function $v(z)$ is compactly supported and continuously differentiable and, in particular, has the finite total variation. Put $n(z) = \int_0^T n(t, z)dt$ and $w(z) = \frac{1}{2}\sigma^2(z)n(z)$. It is clear that there exists a positive constant, say,

ℓ such that $I(v) = \ell \int_R [v^2(z) + |v(z)|] n(z) dz + \int_R w(z) dv(z)$ is an upper bound for the right hand side of (8.1). Show that $I(v)$ can be chosen less than any negative quantity. Use the fact that $I(v)$ is well defined not only for compactly supported and continuously differentiable function $v(z)$ but also for any compactly supported function $v^\alpha(z)$ obeying finite total variation. Assume that there exists a family of $v^\alpha(z), \alpha \in (0, 1]$ such that

$$\lim_{\alpha \to 0} I(v^\alpha) = -\infty \qquad (8.3)$$

and every function $v^\alpha(z)$ obeys an approximation by $v_m^\alpha(z), m \geq 1$ of continuously differentiable compactly supported functions in a sense

$$\lim_m I(v_m^\alpha) = I(v^\alpha). \qquad (8.4)$$

We show that under (8.3) and (8.4) the desired result holds. In fact, for fixed α one can choose a number m_α such that $|I(v^\alpha) - I(v_{m_\alpha}^\alpha)| \leq 1$. Hence we obtain

$$\inf \int_0^T \int_R \mathcal{D}u(t, z) n(t, z) dz dt \leq I(v_{m_\alpha}^\alpha) \leq 1 + I(v^\alpha) \to -\infty, \ \alpha \to 0.$$

Therefore, only (8.3) and (8.4) have to be checked. Since $d_z n \not\ll dz$ the function $n(z)$ is not absolutely continuous and $w(z)$ is inherited the same property. Therefore by the definition of the negation for absolute continuity [28] a constant k can be chosen such that for any $\alpha > 0$ there exists a positive constant c and non overlapping intervals $(z_i', z_i'') \in [-c, c]$, such that $\sum_i |w(z_i'') - w(z_i')| \geq k$ and $\sum_i \int_{z_i'}^{z_i''} n(z) dz \leq \alpha$. Put

$$v^\alpha(z) = \begin{cases} -\frac{1}{\sqrt{\alpha}} \text{sign} [w(z_i'') - w(z_i')], & z_i' < z \leq z_i'' \\ 0, & \text{otherwise.} \end{cases}$$

Show that (8.3) holds. Evaluate from above $I(v^\alpha)$:

$$I(v^\alpha) = \ell \int_R \left[(v^\alpha(z))^2 + |v^\alpha(z)| \right] n(z) dz + \int_R w(z) dv^\alpha(z)$$

$$\leq \ell \left(\frac{1}{\alpha} + \frac{1}{\sqrt{\alpha}} \right) \sum_i \int_{z_i'}^{z_i''} n(z) dz + \sum_i w(z_i') [v^\alpha(z_i'') - v^\alpha(z_i')].$$

$$\leq \ell (1 + \sqrt{\alpha}) + \sum_i w(z_i') [v^\alpha(z_i'') - v^\alpha(z_i')].$$

Now, summizing by parts, we find $\sum_i w(z_i')[v^\alpha(z_i'') - v^\alpha(z_i')] = -\sum_i v^\alpha(z_i'')[w(z_i'') - w(z_i')]$. On the other hand, from the definition of $v^\alpha(z)$ it follows $\sum_i v^\alpha(z_i'')[w(z_i'') - w(z_i')] = \frac{1}{\sqrt{\alpha}} \sum_i |w(z_i'') - w(z_i')| \geq \frac{k}{\sqrt{\alpha}}$. Thereby

$$I(v^\alpha) \leq \ell(1 + \sqrt{\alpha}) - \frac{k}{\sqrt{\alpha}} \to \infty, \ \alpha \to 0.$$

Evidently to satisfy (8.4), it is sufficient to choose approximating functions $v_m^\alpha(z), m \geq 1$ which are compactly supported and continuously differentiable and such that $\lim_m v_m^\alpha(z) = v^\alpha(z)$ in every point of continuity of $v^\alpha(z)$.

Assume $\nu \not\ll \lambda$. Put $K^\nu(dz) = \int_0^T K_\nu(t, dz)dt$ and note that $K^\nu(dz) \not\ll dz$. Use Lebesgue's decomposition: $K^\nu(dz) = q(z)dz + K^\perp(dz)$, where $q(z)$ is a density of absolutely continuous part of $K^\nu(dz)$ and $K^\perp(dz)$ is its singular part. Taking $u(t, z) \equiv u(z)$ which is compactly supported, say, on $[-c, c]$ we find

$$\int_0^T \mathcal{D}u(z)\nu(dt, dz) = \int_{-c}^c \mathcal{D}u(z)q(z)dz + \int_{-c}^c \mathcal{D}u(z)K^\perp(dz).$$

Since $|u'(z)| \leq |u'(0)| + \int_{-c}^c |u''(y)|dy$ there exists constant, say, ℓ, such that $\int_{-c}^c \mathcal{D}u(z)q(z)dz \leq \ell(1 + \int_{-c}^c |u''(y)|dy)$ and so we arrive at an upper estimate

$$\int_0^T \mathcal{D}u(z)\nu(dt, dz) \leq \ell\Big(1 + \int_{-c}^c |u''(z)|dz\Big) + \frac{1}{2}\int_{-c}^c \sigma^2(z)u''(z)K^\perp(dz).$$

Then, using the singularity of $K^\perp(dz)$ and dz, one can choose $u''(z)$ such that the second integral is less any negative quantity while the first remains bounded.

3. <u>Approximation of rate function.</u>
For '$dX = \dot{X}_t dt$, $d\nu = nd\lambda$, $d_z n = n_z' dz$' denote by

$$S_T(X, \nu) = \int_0^T \frac{[\dot{X}_t - A_\nu(t, X_t)]^2}{B_\nu^2(t, X_t)}dt,$$

$$F_T(\nu) = \int_0^T \int_R \sigma^2(z)\Big[\frac{n_z'(t, z)}{n(t, z)} - \frac{p'(z)}{p(z)}\Big]^2 n(t, z)dzdt.$$

Also note one to one correspondence between density $n(t, z)$ and function $v_\nu(t, z)$ defined in (6.1):

$$n(t, z) = n(t, 0)\frac{p(z)}{p(0)}\exp\Big(2\int_0^z \frac{v_\nu(t, y)}{\sigma^2(y)}dy\Big). \tag{8.5}$$

Put

$$\phi(t) = \int_R |n_z'(t, y)|dy. \tag{8.6}$$

Lemma A.4. Let $B^2(x, z) \geq \beta^2 > 0$. If $S_T(X, \nu) < \infty$, $F_T(\nu) < \infty$, then ν can be approximated by a sequence of measures $\nu^{(k)}, k \geq 1$, satisfying the property: $d\nu^{(k)} = n^{(k)}d\lambda$, $d_z n^{(k)} = n_z^{(k)}dz$, such that the function $v_{\nu^{(k)}}(t, z)$, corresponding to $n^{(k)}(t, z)$, is compactly supported in z and continuously differentiable in (t, z) and what is more

$$\lim_k \rho_T(\nu, \nu^{(k)}) = 0$$

$$\lim_k S_T(X, \nu^{(k)}) = S_T(X, \nu)$$

$$\lim_k F_T(\nu^{(k)}) = F_T(\nu). \tag{8.7}$$

Proof. Introduce a chain of expanding subclasses of measures ν characterized in terms of $n(t, z)$ and $v_\nu(t, z)$:

  0) $v_\nu(t, z)$ is compactly supported in z and continuously differentiable in (t, z);
  1) $v_\nu(t, z)$ is compactly supported in z and bounded;
  2) $v_\nu(t, z)$ is compactly supported, $\inf_{t \leq T, z \in R} \frac{n(t,z)}{p(z)} > 0$ and $\sup_{t \leq T}[n(t, 0) + \phi(t)] < \infty$;
  3) $v_\nu(t, z)$ is compactly supported, $\inf_{t \leq T, z \in R} \frac{n(t,z)}{p(z)} > 0$;
  4) $v_\nu(t, z)$ is compactly supported;
  5) $v_\nu(t, z)$ satisfies the assumptions of the lemma.

The proof is based on the following fact. If measure ν from calss 'i' (i=1,...,5) can be approximated by $\nu^{(k)}, k \geq 1$ from class 'i-1' in a sense (8.7), then the statement of the lemma holds.

Assume $\nu^{(k)}, k \geq 1$ is such that

$$\Lambda_T - \lim_k n^{(k)}(t, z) = n(t, z), \ (\Lambda_T(dt, dz) = I_{[0,T]}dtdz),$$

$$\lim_k F_T(\nu^{(k)}) = F_T(\nu). \tag{8.8}$$

Then by Scheffe's theorem [29] [22] we have $\lim_k \int_0^T \int_R |n(t, z) - n^{(k)}(t, z)|dtdz = 0$ that is $\nu^{(k)}$ converges to ν in the total variation norm which implies convergence in Levy-Prohorov's metric too: $\rho_T(\nu, \nu^{(k)}) \to 0$. Since $A_{\nu^{(k)}}(t, X_t) = \int_R A(X_t, z)n^{(k)}(t, z)dz$ and $B^2_{\nu^{(k)}}(t, X_t) = \int_R B^2(X_t, z)n^{(k)}(t, z)dz$ by Lebesque dominated theorem $S_T(X, \nu^{(k)}) \to S_T(X, \nu)$. Therefore, for all steps of approximations only (8.8) has to be checked.

Assume ν is from class '1'. Approximate $v_\nu(t, z)$ by $v_\nu^{(k)}(t, z)$:

$$\lim_k \int_0^T \int_R [v_\nu(t, z) - v_\nu^{(k)}(t, z)]^2 (1 + n(t, z))dtdz = 0,$$

where for all k the functions $v_\nu^{(k)}(t, z), k \geq 1$ are compactly supported continuously differentiable in (t, z). Without loss of a generality one can assume that all function are bounded by the same constant. Similarly to (8.5) define a density of $\nu^{(k)}$:

$$n^{(k)}(t, z) = n^{(k)}(t, 0)\frac{p(z)}{p(0)} \exp \Big(\int_0^z \frac{v_\nu^{(k)}(t, y)}{\sigma^2(z)}dy \Big), \tag{8.9}$$

with $n^{(k)}(t, 0) = \Big(\int_R \frac{p(z)}{p(0)} \exp \Big(\int_0^z \frac{v_\nu^{(k)}(t,y)}{\sigma^2(z)}dy \Big)dz \Big)^{-1}$. Put

$$\nu^{(k)}(dt, dz) = n^{(k)}(t, z)dtdz.$$

Evidently $\nu^{(k)}$ belongs to class'0'. It is easy to check that

$$v_\nu^{(k)}(t, z) \equiv v_{\nu^{(k)}}(t, z) \tag{8.10}$$

and the validity of the first part in (8.8). To verify the second part in (8.8), note that $F_T(\nu^{(k)}) = 4\int_0^T \int_R \frac{(v_{\nu^{(k)}}(t,z))^2}{\sigma^2(z)} n^{(k)}(t,z)dzdt$ and consequently

$$\left|F_T(\nu) - F_T(\nu^{(k)})\right| \leq \text{const.} \int_0^T \int_R |n(t,z) - n^{(k)}(t,z)|dtdz$$

$$+ \text{const.} \int_0^T \int_R \left|v_\nu^2(t,z) - (v_{\nu^{(k)}}(t,z))^2\right|dtdz$$

$$\to 0, \ k \to \infty.$$

Assume ν is from class '2'. For the definiteness assume that there exists positive constant $z_\circ$ such that $v_\nu(t,z) \equiv 0$ out of $[-z_\circ, z_\circ]$. Put $v_\nu^{(k)}(t,z) = v_\nu(t,z)I(|n_z'(t,z)| \leq k)$, define $n^{(k)}(t,z)$ by (8.9) and take $\nu^{(k)}$ with this density. It belongs to class '1'. Herewith, $v_\nu^{(k)}(t,z)$ is defined by (8.10). It is clear that the first part in (8.8) holds and below we check the validity of the second part. We have

$$F_T(\nu^{(k)}) = 4\int_0^T \int_{|z|\leq z_\circ} \frac{v_\nu^2(t,z)}{\sigma^2(z)} I(|n_z(t,z)| \leq k)n^{(k)}(t,z)dzdt$$

$$F_T(\nu) = 4\int_0^T \int_{|z|\leq z_\circ} \frac{v_\nu^2(t,z)}{\sigma^2(z)} n(t,z)dzdt.$$

The required convergence $F_T(\nu^{(k)}) \to F_T(\nu)$ holds by Lebesgue dominated theorem since $n^{(k)}(t,z) \leq p(z)\exp\left(2\phi(t)\right) \leq \text{const.} n(t,z)$.

Assume ν is from class '3'. Putting $v_\nu^{(k)}(t,z) = v_\nu(t,z)I(n(t,0) + \phi(t) \leq k)$ we arrive at

$$n^{(k)}(t,z) = \begin{cases} n(t,z), & n(t,0) + \phi(t) \leq k \\ p(z), & n(t,0) + \phi(t) > k. \end{cases}$$

and since $n^{(k)}(t,0) \leq k + p(0)$ and $\phi^{(k)}(t) \leq k + \int_R |p'(z)|dz$ measure $\nu^{(k)}$ with density $n^{(k)}(t,z)$ belongs to class '2'. It is clear that the first part in (8.8) holds and

$$|F_T(\nu) - F_T(\nu^{(k)})| = 4\int_0^T \int_{|z|\leq z_\circ} \frac{v_\nu^2(t,z)}{\sigma^2(z)} I(n(t,0) + \phi(t) > k)p(z)dzdt$$

$$\leq \text{const.} \int_0^T \int_{|z|\leq z_\circ} \frac{v_\nu^2(t,z)}{\sigma^2(z)} I(n(t,0) + \phi(t) > k)n(t,z)dzdt$$

$$\to 0, \ k \to \infty.$$

Assume ν is from class '4'. Put $n^{(k)}(t,z) = c^{(k)}(t)\left(n(t,z) \vee p(z)\right)$, where $c^{(k)}(t) = \left(\int_R \left(n(t,z) \vee p(z)\right)dz\right)^{-1}$ is norming constant. $\nu^{(k)}$ with this density belongs to class

'3'. The first part in (8.8) holds and what is more $\lim_k c^{(k)}(t) = 1$. On the other hand, since $v_{\nu^{(k)}}(t,z) = v_\nu(t,z)I(n(t,z) \geq p(z)/k)$ we obtain

$$F_T(\nu^{(k)}) = 4 \int_0^T \int_R \frac{v_\nu^2(t,z)}{\sigma^2(z)} I(n(t,z) \geq p(z)/k)c^{(k)}(t)n(t,z)dzdt$$

$$\rightarrow 4 \int_0^T \int_R \frac{v_\nu^2(t,z)}{\sigma^2(z)} n(t,z)dzdt = F_T(\nu).$$

Assume ν is from class '5'. Put $v_\nu^{(k)}(t,z) = v_\nu(t,z)T(|z| \leq k)$ and define $n^{(k)}(t,z)$ by (8.9). Then

$$n^{(k)}(t,z) = n^{(k)}(t,0)\begin{cases} p(z)\frac{n(t,k)}{p(k)}, & z > k \\ n(t,z), & |z| \leq k \\ p(z)\frac{n(t,-k)}{p(-k)}, & z < -k. \end{cases}$$

Taking $\nu^{(k)}$ with this density and noticing that $\lim_k n^{(k)}(t,0) = 1$ we find

$$F_T(\nu^{(k)}) = 4 \int_0^T \int_{|z| \leq k} \frac{\left(v_\nu(t,z)\right)^2}{\sigma^2(z)} c^{(k)}(t)n(t,z)dzdt \rightarrow F_T(\nu),$$

i.e. both parts in (8.8) hold.

4. Ergodic property.

Consider diffusion pair $(\widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)$ defined by Itô's differential equations w.r.t. independent Wiener processes W_t and V_t:

$$d\widetilde{X}_t^\varepsilon = G(t, \widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)dt + \sqrt{\varepsilon}B(\widetilde{X}_t^\varepsilon, \widetilde{\xi}_t^\varepsilon)dW_t$$

$$d\widetilde{\xi}_t^\varepsilon = \frac{1}{\varepsilon}b(t, \widetilde{\xi}_t^\varepsilon)dt + \frac{1}{\sqrt{\varepsilon}}\sigma(\widetilde{\xi}_t^\varepsilon)dV_t \tag{8.11}$$

subject to (x_0, z_0), where $B(x,z)$ and $\sigma(z)$ are functions involving in (1.1). Assume $b(t,z)$ is continuous it (t,z), continuously differentiable in t, Lipschitz continuous in z uniformly in t, and $zb(t,z)$ is negative for large $|z|$ uniformly in t. Also assume that

$$G(t,x,z) = \frac{\dot{X}_t - A_p(t,x)}{B_p(t,x)}B(x,z) + A(x,z), \tag{8.12}$$

where $A(x,z)$ involves in (1.1), $\dot{X}_t$ is the Radon-Nykodim derivative of absolute continuous function X_t from $\mathbb{C}$ with $X_0 = x_0$, $A_p(t,x) = \int_R A(x,z)p(t,z)dz$, $B_p(t,x) = \sqrt{\int_R B^2(x,z)p(t,z)dz}$, and where (comp. (2.1))

$$p(t,z) = c(t)\frac{\exp\left(2\int_0^z \frac{b(t,y)}{\sigma^2(y)}dy\right)}{\sigma^2(z)}$$

with norming function $c(t)$ such that $\int_R p(t,z)dz = 1$. Introduce an occupation measure $\widetilde{\nu}^\varepsilon(dt, dz)$: $\widetilde{\nu}^\varepsilon(\Delta \times \Gamma) = \int_0^\infty I(t \in \Delta, \widetilde{\xi}_t^\varepsilon \in \Gamma)dt$ and put $\nu(dt, dz) = p(t,z)dzdt$.

Lemma A.5.

$$\mathbf{P} - \lim_{\varepsilon \to 0} \rho_T(\widetilde{\nu}^\varepsilon, \nu) = 0 \quad and \quad \mathbf{P} - \lim_{\varepsilon \to 0} r_T(\widetilde{X}^\varepsilon, X) = 0$$

Proof. It is clear, the first statement of the lemma is equivalent to: for any bounded and continuous function $h(t, z)$ $\int_0^T \int_R h(t, \widetilde{\xi}_t^\varepsilon) dt \to \int_0^T \int_R h(t, z) p(t, z) dz dt$ in probability or, for $h^\circ(t, z) = h(t, z) - \int_0^T \int_R h(t, y) p(t, y) dy dt$,

$$\mathbf{P} - \lim_{\varepsilon \to 0} \int_0^T \int_R h^\circ(t, \widetilde{\xi}_t^\varepsilon) dt = 0.$$

First we check it for continuously differentiable in t, z function $h(t, z)$, having bounded partial derivatives. Straightforward calculation brings Kolmogorov's type differential equation (t is fixed):

$$\frac{1}{2} \frac{\partial}{\partial z}(\sigma^2(x) p(t, x)) = b(t, z) p(t, z).$$

A conjugate differential equation

$$\frac{1}{2} \sigma^2(z) \frac{\partial w(t, z)}{\partial z} + b(t, z) w(t, z) = h^\circ(t, z). \tag{8.13}$$

obeys a solution

$$w(t, z) = \frac{2}{\sigma^2(z) p(t, z)} \int_{-\infty}^z h^\circ(t, y) p(t, y) dy.$$

It is clear that properties of $h(t, z)$ are inherited by $w(t, z)$ and so function $u(t, z) = \int_0^z w(t, y) dy$ is continuously differentiable once in t and twice in z and what is more, due to the boundness of $w(t, z)$, there exists a positive constant, say ℓ, such that $|u(t, z)| \leq \ell|z|$ and $|u_t(t, z)| \leq \ell|z|$. Applying Itô's formula to $u(t, \widetilde{\xi}_t^\varepsilon)$ and taking into account that $w(t, z)$ is solution of differential equation (8.13) we find $u(T, \widetilde{\xi}_T^\varepsilon) = u(0, \xi_0) + \int_0^T u_t'(t, \widetilde{\xi}_t^\varepsilon) dt + \frac{1}{\sqrt{\varepsilon}} \int_0^T w(t, \widetilde{\xi}_t^\varepsilon) \sigma(\widetilde{\xi}_t^\varepsilon) dV_t + \frac{1}{\varepsilon} \int_0^T h^\circ(t, \widetilde{\xi}_t^\varepsilon) dt$ that is

$$\int_0^T h^\circ(t, \widetilde{\xi}_t^\varepsilon) dt = \varepsilon u(T, \widetilde{\xi}_T^\varepsilon) - \varepsilon u(0, \xi_0)$$

$$- \varepsilon \int_0^T u_t'(t, \widetilde{\xi}_t^\varepsilon) dt - \sqrt{\varepsilon} \int_0^T w(t, \widetilde{\xi}_t^\varepsilon) \sigma(\widetilde{\xi}_t^\varepsilon) dV_t. \tag{8.14}$$

The second term in the right hand of (8.10) converges to zero; the last term converges to zero in probability since by Problem 1.9.2 in [23] the mentioned convergence follows from $\varepsilon \int_0^T w^2(t, \widetilde{\xi}_t^\varepsilon) \sigma^2(\widetilde{\xi}_t^\varepsilon) dt \to 0$; other two terms converge to zero in probability if $\lim_{\varepsilon \to 0} \varepsilon^2 \mathbf{E} \sup_{t \leq T} (\widetilde{\xi}_t^\varepsilon)^2 = 0$. To check the last, apply Itô's formula to $(\varepsilon \widetilde{\xi}_t^\varepsilon)^2$:

$$(\varepsilon \widetilde{\xi}_t^\varepsilon)^2 = (\varepsilon \xi_0)^2 + 2\varepsilon \int_0^t b(s, \widetilde{\xi}_s^\varepsilon) \widetilde{\xi}_s^\varepsilon ds + \varepsilon \int_0^t \sigma^2(\widetilde{\xi}_s^\varepsilon) ds$$

$$+ 2\varepsilon^{3/2} \int_0^t \widetilde{\xi}_s^\varepsilon \sigma(\widetilde{\xi}_s^\varepsilon) dV_s.$$

The function $b(s, z)$ is such that $zb(t, z)$ is negative for large $|z|$ what implies $(\varepsilon \widetilde{\xi_t^\varepsilon})^2 \le (\varepsilon \xi_0)^2 + T\varepsilon \text{const.} + 2\varepsilon^{3/2} \int_0^t \widetilde{\xi_s^\varepsilon} \sigma(\widetilde{\xi_s^\varepsilon}) dV_s$. Thereby $\mathbf{E}(\varepsilon \widetilde{\xi_t^\varepsilon})^2 \le (\varepsilon \xi_0)^2 + T\varepsilon \text{const.}$ In turn, using Doob's inequality (see e.g. Theorem 1.9.1 in [23]), we arrive at $\mathbf{E} \sup_{t \le T}(\varepsilon \widetilde{\xi_t^\varepsilon})^2 \le (\varepsilon \xi_0)^2 + T\varepsilon \text{const.} + \text{const.}\varepsilon^3 \int_0^T \mathbf{E}(\varepsilon \widetilde{\xi_t^\varepsilon})^2 dt$ and, due to the obtained above upper bound for $\mathbf{E}(\varepsilon \widetilde{\xi_t^\varepsilon})^2$, the result holds.

If $h(t, z)$ is bounded and continuous only, it can be approximated by smooth functions $h_m(t, z), m \ge 1$ in the following sense: for any $k \ge 1$ $\lim_m \sup_{t \le T, |z| \le k} |h(t, z) - h_m(t, z)| = 0$. Since for every $h_m(t, z)$ the statement of the lemma is proved, it holds for $h(t, z)$ if

$$\lim_m \int_0^T \int_R |h(t, z) - h_m(t, z)| p(t, z) dz dt = 0$$

$$\mathbf{P} - \lim_m \overline{\lim_{\varepsilon \to 0}} \int_0^T |h(t, \widetilde{\xi_t^\varepsilon}) - h_m(t, \widetilde{\xi_t^\varepsilon})| n(t, z) dt = 0.$$

The first takes place since $\lim_k \int_0^T \int_{|z|>k} n(t, z) dz dt = 0$ while the second from $\mathbf{P} - \lim_m \overline{\lim}_{\varepsilon \to 0} \int_0^T I(|\widetilde{\xi_t^\varepsilon})| > k) dt = 0$ and the fact that one can choose smooth bounded functions $g_k(z), k \ge 1$ such that $I(|z| > k) \le g_k(z), \lim_k g_k(z) = 0, z \in R$ and by proved above $\int_0^T g_k(\widetilde{\xi_t^\varepsilon}) dt \to \int_0^T \int_R g_k(z) p(t, z) dz dt \to 0, \ k \to \infty$.

To check the second statement, put $\Delta_t = \widetilde{X_t^\varepsilon} - X_t$. From the first equation in (8.11) we find

$$\Delta_t = \int_0^t \left[\frac{\dot{X}_t - A_p(s, X_s)}{B_p(s, X_s)} B(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) + A(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) - \dot{X}_s \right] ds$$

$$+ \sqrt{\varepsilon} \int_0^t B(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) dW_s$$

$$= \int_0^t \left[\frac{\dot{X}_t - A_p(s, X_s)}{B_p(s, X_s)} \left(B(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) - B(X_s, \widetilde{\xi_s^\varepsilon}) \right) \right] ds$$

$$+ \int_0^t \left[\frac{\dot{X}_s - A_p(s, X_s)}{B_p(s, X_s)} \left(B(X_s, \widetilde{\xi_s^\varepsilon}) - B_p(s, X_s) \right) \right] ds$$

$$+ \int_0^t \left(A(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) - A(X_s, \widetilde{\xi_s^\varepsilon}) \right) ds$$

$$+ \int_0^t \left(A(X_s, \widetilde{\xi_s^\varepsilon}) - A_p(s, X_s) \right) ds$$

$$+ \sqrt{\varepsilon} \int_0^t B(\widetilde{X_s^\varepsilon}, \widetilde{\xi_s^\varepsilon}) dW_s.$$

For brevity put $\varphi_s = \frac{\dot{X}_s - A_p(s, X_s)}{B_p(s, X_s)}$. Then by the Lipschitz continuity of $A(x, z), B(x, z)$ in

x uniformly in z, say, with constant ℓ, we obtain

$$|\Delta_t| \le \ell \int_0^t (1 + |\varphi_s|)\Delta_s ds$$
$$+ \sup_{t \le T} \left| \int_0^t \varphi_s \big(B(X_s, \widetilde{\xi}_s^\varepsilon) - B_p(s, X_s)\big)ds \right|$$
$$+ \sup_{t \le T} \left| \int_0^t \big(A(X_s, \widetilde{\xi}_s^\varepsilon) - A_p(s, X_s)\big)ds \right|$$
$$+ \sqrt{\varepsilon}\sup_{t \le T} \left| \int_0^t B(\widetilde{X}_s^\varepsilon, \widetilde{\xi}_s^\varepsilon)dW_s \right|.$$

Therefore, by Gronwall-Bellman's inequality

$$\sup_{t \le T}|\Delta_t| \le \exp\left(\ell \int_0^T (1 + |\varphi_s|)ds\right)$$
$$\times \left[\sup_{t \le T} \left| \int_0^t \varphi_s \big(B(X_s, \widetilde{\xi}_s^\varepsilon) - B_p(s, X_s)\big)ds \right| \right.$$
$$+ \sup_{t \le T} \left| \int_0^t \big(A(X_s, \widetilde{\xi}_s^\varepsilon) - A_p(s, X_s)\big)ds \right|$$
$$+ \left. \sqrt{\varepsilon}\sup_{t \le T} \left| \int_0^t B(\widetilde{X}_s^\varepsilon, \widetilde{\xi}_s^\varepsilon)dW_s \right| \right]$$

Hence, the second statement holds if

$$\mathbf{P} - \lim_{\varepsilon \to 0}\sup_{t \le T} \sqrt{\varepsilon}\sup_{t \le T} \left| \int_0^t B(\widetilde{X}_s^\varepsilon, \widetilde{\xi}_s^\varepsilon)dW_s \right| = 0 \tag{8.15}$$

and for any measurable function ψ_s such that $\int_0^T \psi_s^2 ds < \infty$ and any continuous function $C(x, z)$, being Lipschitz's continuous in x uniformly in z,

$$\mathbf{P} - \lim_{\varepsilon \to 0}\sup_{t \le T} \left| \int_0^t \psi_s \big(C(X_s, \widetilde{\xi}_s^\varepsilon) - C_p(s, X_s)\big)ds \right| = 0, \tag{8.16}$$

where $C_p(s, X_s) = \int_R C(X_s, z)p(s, z)dz$. It can be shown (see e.g. the method of proving the statement (2) of Theorem 4.6 Ch. 4 in [24]) that $\sup_{t \le T}\mathbf{E}(\widetilde{X}_t^\varepsilon)^2 \le$ const. and so $\mathbf{E}\int_0^T B^2(\widetilde{X}_s^\varepsilon, \widetilde{\xi}_s^\varepsilon)ds \le$ const. Consequently, by Doob's inequality (see e.g. Theorem 1.9.1 [23]) we get $\mathbf{E}\sup_{t \le T}\left| \sqrt{\varepsilon}\int_0^t B(\widetilde{X}_s^\varepsilon, \widetilde{\xi}_s^\varepsilon)dW_s \right|^2 \le \varepsilon$const. that is (8.15) holds. To check the validity of (8.16) with $\psi_s C(x, z) \ge 0$, note that, due to Problem 5.3.2 in[23], it is sufficient to prove

$$\mathbf{P} - \lim_{\varepsilon \to 0} \int_0^t \psi_s \big(C(X_s, \widetilde{\xi}_s^\varepsilon) - C_p(s, X_s)\big)ds = 0, \ \forall t \le T \tag{8.17}$$

and what is more, due to an arbitrariness of ψ_s and $C(x, z)$, (8.17) implies (8.16) in the general case since one can use separately (8.17) for positive $(\psi_s C(x, z))^+$ and negative $(\psi_s C(x, z))^-$ parts. Therefore, (8.17) remains to be verified. If ψ_s is bounded and continuous, (8.17) takes place by virtue of the first statement of the lemma. If only $\int_0^T \psi_s^2 ds < \infty$, approximate ψ_s by bounded and continuous functions $\psi_s^{(k)}$, $k \geq 1$ such that $\lim_k \int_0^T (\psi_s - \psi_s^{(k)})^2 ds = 0$ and, due to the boundness in z of $C(x, z)$ and Cauchy-Schwartz's inequality, we find that

$$\left| \int_0^t (\psi_s - \psi_s^{(k)})\big(C(X_s, \widetilde{\xi}_s^\varepsilon) - C_p(s, X_s)\big) ds \right| \leq \text{const.} \sqrt{\int_0^T (\psi_s - \psi_s^{(k)})^2 ds} \to 0, \ k \to \infty$$

that is (8.17) takes place since it holds for every $\psi_s^{(k)}$.

5. <u>LD-regularization.</u>

Parallel to X_t^ε, defined in (1.1), determine new diffusion $X_t^{\varepsilon,\beta}$ with uniformly non singular diffusion parameter $B^2(x, z) + \beta^2, \beta^2 > 0$, letting

$$dX_t^{\varepsilon,\beta} = A(X_t^{\varepsilon,\beta}, \xi_t^\varepsilon)dt + \sqrt{\varepsilon}\big[B(X_t^{\varepsilon,\beta}, \xi_t^\varepsilon)dW_t + \beta dW_t'\big] \qquad (8.18)$$

subject to the same initial point x_0, where W_t' is a Wiener process independent of (W_t, ξ_t^ε).

Lemma A.6. *Under assumption (A.1) for every $T > 0$ and $\eta > 0$*

$$\lim_{\beta \to 0} \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(r_T(X^{\varepsilon,\beta}, X^\varepsilon) > \eta) = -\infty.$$

Proof. Put $\Delta_t = X_t^{\varepsilon,\beta} - X_t^\varepsilon$, and

$$a_1(x', x'', z) = \frac{A(x'', z) - A(x', z)}{x'' - x'}, \quad a_2(x', x'', z) = \frac{B(x'', z) - B(x', z)}{x'' - x'},$$

where for $x' = x''$ $a_i(x', x', z), i = 1, 2$ are Radon-Nikodym's derivatives. Note that for $x' \neq x''$ $a_i(x', x'', z)$, are bounded, say, by constant ℓ, and so $a_i(x', x', z)$ can be taken bounded by the same constant. For brevity, denote by $\alpha_i(t) = a_i(X_t^{\varepsilon,\beta}, X_t^\varepsilon, \xi_t^\varepsilon), i = 1, 2$. (8.18) and (1.1) imply: $\Delta_t = \int_0^t \alpha_1(s)\Delta_s ds + \sqrt{\varepsilon}\int_0^t \alpha_2(s)\Delta_s dW_s + \sqrt{\varepsilon}\beta W_t'$. Letting $\mathcal{E}_t = \exp\left(\int_0^t [\alpha_1(s) - (\varepsilon/2)\alpha_2^2(s)]ds + \sqrt{\varepsilon}\int_0^t \alpha_2(s)dW_s\right)$ and using Itô's formula, we find that $\Delta_t = \sqrt{\varepsilon}\beta\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}dW_s'$ and thereby

$$\sup_{t \leq T} |\Delta_t| \leq \sqrt{\varepsilon}\beta \sup_{t \leq T} \mathcal{E}_t \sup_{t \leq T} \left| \int_0^t \mathcal{E}_s^{-1}dW_s' \right|.$$

Put $\Gamma_N = \{\frac{1}{N} \leq \inf_{t \leq T} \mathcal{E}_t \leq \sup_{t \leq T} \mathcal{E}_t \leq N\}$ and use an upper estimate

$$\mathbf{P}(\sup_{t \leq T} |\Delta_t| > \eta) \leq \mathbf{P}(\sup_{t \leq T} |\Delta_t| > \eta, \Gamma_N) + \mathbf{P}(\Omega \setminus \Gamma_N)$$

$$\leq 2\max\left[\mathbf{P}(\sup_{t \leq T} |\Delta_t| > \eta, \Gamma_N), \mathbf{P}(\Omega \setminus \Gamma_N)\right]$$

which implies, due to the boundness of $\alpha_i(s), i = 1, 2$, the desired statement if

$$\lim_N \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(\sqrt{\varepsilon} \sup_{t \leq T} \left| \int_0^t \alpha_2(s) dW_s \right| > N) = -\infty$$

$$\lim_{\beta \to 0} \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(\sqrt{\varepsilon} \beta \sup_{t \leq T} \left| \int_0^t \mathcal{E}_s^{-1} dW'_s \right| > \eta, \Gamma_N) = -\infty, \forall N \geq 1. \qquad (8.19)$$

Let $\tau = \{ t \leq T : \left| \int_0^t \alpha_2(s) dW_s \right| > (N/\sqrt{\varepsilon}) \}$ and $\sigma = \{ t \leq t : \left| \int_0^t \mathcal{E}_s^{-1} dW'_s \right| > (\eta/\sqrt{\varepsilon}\beta) \}$. It is clear that (8.19) is equivalent to:

$$\lim_N \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(\sqrt{\varepsilon} \int_0^\tau \alpha_2(s) dW_s \geq N \ (\text{or} \ \leq -N) \big) = -\infty$$

$$\lim_{\beta \to 0} \overline{\lim}_{\varepsilon \to 0} \varepsilon \log \mathbf{P}\big(\sqrt{\varepsilon} \beta \int_0^\sigma \mathcal{E}_s^{-1} dW'_s \geq \eta \ (\text{or} \ \leq -\eta), \Gamma_N \big) = -\infty, \forall N \geq 1. \quad (8.20)$$

Below we check (8.20). To this end with $\lambda \in R$, introduce continuous local martingales: $Z_t^1 = \exp\left(\lambda \int_0^t \alpha_2(s) dW_s - (\lambda^2/2) \int_0^t \alpha_2^2(s) ds \right)$ and $Z_t^2 = \exp\left(\lambda \int_0^t \mathcal{E}_s^{-1} dW'_s - (\lambda^2/2) \int_0^t \mathcal{E}_s^{-2} ds \right)$, where each of them is a supermartingale too (see Problem 1.4.4. in [23]) that is $\mathbf{E} Z_\tau^1 \leq 1$ and $\mathbf{E} Z_\sigma^2 \leq 1$. Then we use obvious inequalities: $\mathbf{E} I(\sqrt{\varepsilon} \int_0^\tau \alpha_2(s) dW_s \geq N) Z_\tau^1 \leq 1$ and $\mathbf{E} I(\sqrt{\varepsilon} \beta \int_0^\sigma \mathcal{E}_s dW'_s \geq \eta, \Gamma_N) Z_\sigma^2 \leq 1$. Since for $\lambda > 0$, $\log Z_\tau^1 \geq (\lambda N/\sqrt{\varepsilon}) - (\lambda^2 \ell^2 T/2)$ and $\log Z_\sigma^1 \geq (\lambda \eta/\sqrt{\varepsilon}\beta) - (\lambda^2 N^2 T/2)$ on sets $\{ \sqrt{\varepsilon} \int_0^\tau \alpha_2(s) dW_s \geq N \}$ and $\{ \sqrt{\varepsilon} \beta \int_0^\sigma \mathcal{E}_s dW'_s \geq \eta, \Gamma_N \}$ respectively, we arrive at (8.20) for '$\geq N$' and '$\geq \eta$', taking $\lambda^1 = (N/\sqrt{\varepsilon}\ell^2 T)$ and $\lambda^2 = (\eta/\sqrt{\varepsilon}\beta N^2 T)$. For '$\leq -N$' and '$\leq -\eta$' the validity of (8.20) is proved in the same way.

References

[1] Varadhan, S.R.S.: Large deviations and Applications. Philadelphia: SIAM 1984

[2] Freidlin, M.L., Wentzell, A.D.: Random Perturbations of Dynamical Systems. New York: Springer 1984

[3] Stroock, D.W.: An Introduction to the Theory of Large Deviations. Berlin: Springer 1984

[4] Donsker, M.D., Varadhan, S.R.S: Asymptotic evaluation of certain Markov process expectations for large time **I**. Communication of Pure and Applied Mathematics **28**, 1-47 (1975)

[5] Donsker, M.D., Varadhan, S.R.S: Asymptotic evaluation of certain Markov process expectations for large time **II**. Communication of Pure and Applied Mathematics **28**, 279-301 (1975)

[6] Donsker, M.D., Varadhan, S.R.S: Asymptotic evaluation of certain Markov process expectations for large time **III**. Communication of Pure and Applied Mathematics **29**, 389-462 (1976)

[7] Donsker, M.D., Varadhan, S.R.S: Asymptotic evaluation of certain Markov process expectations for large time **IV**. Communication of Pure and Applied Mathematics **36**, 183-212 (1983)

[8] Gärtner, J: On the large deviation for invariant measure. Theory of Probability and its Application **22**, 27-42 (1977)

[9] Freidlin, M.I.: Averaging principle and theorem on large deviations. Uspekhi Mat. **33**, 107-160 (1978)

[10] Freidlin, M.I.: The action functional for a class of stochastic processes. Theory Probab. Appl. **17**, 511-575 (1972)

[11] Veretennikov, A.Yu.: On large deviations for ergodic process empirical measures. Advances in Soviet Mathematics **12**, 125-133 (1992)

[12] Veretennikov, A.Yu.: On the large deviations for additive functionals of Markov processes **I**. Theory of Probability and its Application **38**, (1993)

[13] Acosta, A.: Large deviations for empirical measures of Markov chains. J. Theoret. Probab. **3**, 395-431 (1990)

[14] Dupuis, P., Ellis, R.S,: A stochastic optimal control approach to the theory of large deviations, Preprint (1992).

[15] Puhalskii, A.A.: On functional principle of large deviations. New trends in Probability and Statistics. V.Sazonov and Shervashidze (eds.), Vilnius, Lithuania, VSP/Mokslas, 198-218 (1991)

[16] Puhalskii, A.A.: On large deviation theory. Theory Probab., Appl. **38** (1993)

[17] Liptser, R.S., Pukhalskii,A.A.: Limit theorems on large deviations for semimartingales. Stochastic and Stochastics Reports **38**, 201-249 (1992)

[18] Skorokhod, A.V.: Asymptotic Methods in the Theory of Stochastic Differential Equations. AMS. Providence Rhode Island. Translation of Mathematical monographs, **78**, 1989

[19] Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Jones and Bartlet 1993

[20] Deushel, J.D., Stroock, D.W.: Large Deviations. New York: Academic Press 1989

[21] Lynch, J., Sethuraman, J.: Large deviations for the processes with independent increments. Ann. Probab. **15**, 610-627 (1987)

[22] Billingsley, P.: Convergence of Probability Measures. New York: Wiley and Sons 1968

[23] Liptser, R.S., Shiryayev, A.N.: Theory of Martingales. Dordrecht: Kluwer 1989

[24] Liptser, R.S., Shiryayev, A.N.: Statistic of Random Processes **I**: General Theory. New York: Springer 1977

[25] Puhalskii, A.A.: Large deviations of semimartingales via convergence of the predictable characteristics. Stochastics and Stochastics Reports **49**, 27-85 (1994)

[26] Hennequin, P.L., Tortrat, A.: Théorie des Probabilitś et quelques Applications. Paris: Masson 1965

[27] Puhalskii, A.A.: The method of stochastic exponential for large deviations. Stochastic processes and their applications, (Preprint)

[28] Titchmash, E.C.: The Theory of Functions. Oxford: University Press 1939