Markov Processes Relat. Fields **8**, 611–636 (2002)

Markov Processes and Related Fields

Freidlin – Wentzell Type Large Deviations for Smooth Processes*

R. Liptser[1] **, V. Spokoiny**[2] **and A.Yu. Veretennikov**[3]

[1] Department of Electrical Engineering-Systems, Tel Aviv University, 69978 Tel Aviv, Israel. E-mail: liptser@eng.tau.ac.il
[2] Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany. E-mail: spokoiny@wias-berlin.de
[3] Department of Mathematics, University of Kansas, Lawrence, KS, 66045, USA. School of Mathematics, University of Leeds, Woodhouse Lane, LS2 9JT, Leeds, UK. Institute for Information Transmission Problems, Russian Academy of Sciences, 19, B. Karetny, 101447 Moscow, Russia. E-mail: veretennikov@math.ukans.edu

Received December 14, 2001, revised June 7, 2002

Abstract. We establish large deviation principle for the family of vector-valued random processes $X^\varepsilon, \varepsilon \to 0$ defined by ordinary differential equations (under $0 < \kappa < 1/2$)

$$\dot{X}_t^\varepsilon = F(X_t^\varepsilon) + \varepsilon^{1/2-\kappa} G(X_t^\varepsilon) \dot{W}_t^\varepsilon,$$

where $\dot{W}_t^\varepsilon = \varepsilon^{-1/2} g(\xi_{t/\varepsilon})$, ξ_t is a vector-valued ergodic diffusion satisfying, so called, "recurrence condition" and g is a vector-function with zero barycenter with respect to the invariant measure of (ξ_t). A choice of $\kappa < 1/2$ provides the rate function of Freidlin – Wentzell type.

KEYWORDS: moderate deviations, Poisson decomposition, Puhalskii theorem

AMS SUBJECT CLASSIFICATION: 60J27, 60F10

1. Introduction

Large deviation principle (LDP) is a powerful tool of asymptotic analysis for various stochastic systems. A lot of important results in LDP are known for systems governed by Wiener process. In many applications Wiener process is only an approximation, in *central limit theorem scale*, for some intricate phenomena and a priori is unclear that such type of robustness remains valid in the LDP scale as well.

*The work was supported by the Weierstrass Institute for Applied Analysis, Berlin; for the third author also by grants RFBR 00-01-22000, INTAS 99-0590, CNRS PICS-99 "Homogénéisation, moyénnisation et milieux aléatoires. Méthodes probabilistes", and EPSRC GR/R40746/01.

The LDP in $\mathbf{C}_{[0,\infty)}(\mathbf{R}^\ell)$ is known from Freidlin and Wentzell, [14], for $\mathbf{R}^\ell$ valued diffusion with Wiener process W_t and "diffusion parameter" Υ (positive definite $(\ell \times \ell)$-matrix):

$$\begin{aligned} dx^\varepsilon_t &= F(x^\varepsilon_t)\,dt + \sqrt{\varepsilon}G(x^\varepsilon_t)\Upsilon^{1/2}\,dW_t \\ x^\varepsilon_0 &= x_0. \end{aligned}$$

It is characterized by the speed ε and rate function

$$J(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^\infty \| \dot{X}_t - F(X_t)) \|^2_{Q^{-1}(X_t)}\,dt, & X \in \mathfrak{F} \\ \infty, & \text{otherwise}, \end{cases} \tag{1.1}$$

where $\mathfrak{F} = \big\{X \in \mathbf{C}_{[0,\infty)}(\mathbf{R}^\ell) : dX_t = \dot{X}_t\,dt,\ X_0 = x_0\big\}$, * is transposition symbol, $Q(x) = G(x)\Upsilon G^*(x)$ is nonsingular matrix, and $\|\cdot\|_{Q^{-1}}$ is $\mathbf{L}^2$-norm: $\|x\|_{Q^{-1}} = \sqrt{\langle x, Q^{-1}x\rangle}$.

If $(x^\varepsilon_t)_{t\ge 0}$ is an approximation for a smooth process $(X^\varepsilon_t)_{t\ge 0}$ with

$$dX^\varepsilon_t = \big(F(X^\varepsilon_t) + \sqrt{\varepsilon}G(X^\varepsilon_t)\dot{W}^\varepsilon_t\big)\,dt, \tag{1.2}$$

where $\dot{W}^\varepsilon_t$ is "wide-band noise" so that $W^\varepsilon_t = \int_0^t \dot{W}^\varepsilon_s\,ds$ converges in distribution sense, as $\varepsilon \to 0$, to Wiener process with the diffusion matrix Υ, a natural question is important: may families $(x^\varepsilon_t)_{t\ge 0}$ and $(X^\varepsilon_t)_{t\ge 0}$ share the same LDP? In general, the answer is negative. However, under certain natural restrictions a result might be positive and this is the problem under consideration.

We assume that

$$\dot{W}^\varepsilon_t = \frac{1}{\sqrt{\varepsilon}}g(\xi_{t/\varepsilon}), \tag{1.3}$$

where $\xi = (\xi_t)_{t\ge 0}$, $\xi_0 = z_0$, is an ergodic $\mathbf{R}^d$-valued diffusion with respect to vector-valued Brownian motion B_t with d independent components:

$$d\xi_t = b(\xi_t)\,dt + \sigma(\xi_t)\,dB_t, \tag{1.4}$$

and $g(z)$ is a vector-valued function such that

$$\int_{\mathrm{R}^d} g(z)\mu(dz) = 0, \tag{1.5}$$

where μ is the invariant measure of ξ. Due to the above-mentioned remark on a closeness of distributions for (W^ε_t) and (W_t), our approach to LDP examination exploits ideas from proofs of functional central limit theorems (FCLT) based on a method of corrector (see, e.g., the papers by Papanicolaou, Stroock and Varadhan [29], Ethier and Kurtz [12], Bhattacharya [6], Pardoux and Veretennikov [30]). This method uses the Poisson equations

$$\mathcal{L}u_i = -g_i, \quad i = 1, \dots, d, \tag{1.6}$$

where g_i's are entries of g and

$$\mathcal{L} = \sum_i b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^*)_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}$$

is the diffusion operator of ξ. Under "recurrence conditions": (A_b) in Section 2 (see [18] and also [38]) (1.6) possess unique solutions (up to additive constants) in the class of functions with Sobolev's partial second derivatives locally integrable in any power and a polynomial growth. Henceforth

$$\psi = \begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}, \quad \Psi = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} \cdots \frac{\partial u_1}{\partial x_d} \\ \dots\dots\dots\dots \\ \frac{\partial u_d}{\partial x_1} \cdots \frac{\partial u_d}{\partial x_d} \end{pmatrix} := \begin{pmatrix} \nabla u_1 \\ \vdots \\ \nabla u_d \end{pmatrix}. \tag{1.7}$$

By embedding theorems [20], all entries of the matrix Ψ are continuous functions. So, by Krylov's generalization of Itô formula (see [19]), applied to $\psi(\xi_{t/\varepsilon})$, we find the Poisson decomposition for $\varepsilon^{-1/2}\int_0^t g(\xi_{s/\varepsilon})\,ds$ $(= W_t^\varepsilon)$:

$$W_t^\varepsilon = \varepsilon^{1/2}\big(\psi(z_0) - \psi(\xi_{t/\varepsilon})\big) + M_t^\varepsilon, \tag{1.8}$$

where

$$M_t^\varepsilon = \int_0^t \Psi(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big) \tag{1.9}$$

is a continuous martingale with

$$\big\langle M^\varepsilon\big\rangle_t = \int_0^t \Psi(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\sigma^*(\xi_{s/\varepsilon})\Psi^*(\xi_{s/\varepsilon})\,ds \tag{1.10}$$

the predictable quadratic covariation matrix-valued process. At the same time, by the Bogolubov homogenization principle (see e.g. [14]), for any $t > 0$ we have $\big\langle M^\varepsilon\big\rangle_t \xrightarrow[\varepsilon\to 0]{\text{prob.}} \Upsilon t$, with

$$\Upsilon = \int_{\mathrm{R}^d} \Psi(z)\sigma(z)\sigma^*(z)\Psi^*(z)\mu(dz). \tag{1.11}$$

This convergence provides the FCLT: $(M_t^\varepsilon)_{t\ge 0} \xrightarrow[\varepsilon\to 0]{\text{law}} (\Upsilon^{1/2}W_t)_{t\ge 0}$ for vector-valued Wiener process $(W_t)_{t\ge 0}$ with independent components (see e.g., Chapter 8 in [27]) and also $(W_t^\varepsilon)_{t\ge 0} \xrightarrow[\varepsilon\to 0]{\text{law}} (\Upsilon^{1/2}W_t)_{t\ge 0}$, since $\varepsilon^{1/2}\big(\psi(z_0) - \psi(\xi_{t/\varepsilon})\big)$ is asymptotically negligible in $\varepsilon \to 0$.

It may seem that families $(\varepsilon^{1/2}W_t^\varepsilon)_{t\ge 0}$ and $(\varepsilon^{1/2}\Upsilon^{1/2}W_t)_{t\ge 0}$ share the same LDP. Unfortunately, arguments valid for FCLT fail for LDP setting as far as $\sqrt{\varepsilon}W_t^\varepsilon = \varepsilon\big(\psi(\xi_0) - \psi(\xi_{t/\varepsilon})\big) + \sqrt{\varepsilon}M_t^\varepsilon$ and

$$\varepsilon\big(\psi(\xi_0) - \psi(\xi_{t/\varepsilon})\big) \quad \text{and} \quad \varepsilon\langle M^\varepsilon\rangle_t$$

have the same smallness in ε and, even if the LDP holds true for $(W_t^\varepsilon)_{t\geq 0}$, due to an influence of $\varepsilon\big(\psi(z_0)-\psi(\xi_{t/\varepsilon})\big)$, it is far from LDP for $(\varepsilon^{1/2}\Upsilon^{1/2}W_t)_{t\geq 0}$ (see examples in [13] and [22]).

To keep the same LDP as for $(\varepsilon^{1/2}\Upsilon^{1/2}W_t)_{t\geq 0}$ does, following Bayer and Freidlin [3], we replace $\varepsilon^{1/2}$ by $\varepsilon^{1/2-\kappa}$, $0<\kappa<1/2$. Then we have

$$\varepsilon^{1/2-\kappa}W_t^\varepsilon=\varepsilon^{1-\kappa}\big(\psi(z_0)-\psi(\xi_{t/\varepsilon})\big)+\varepsilon^{1/2-\kappa}M_t^\varepsilon$$

and

$$\varepsilon^{1-\kappa}\big(\psi(z_0)-\psi(\xi_{t/\varepsilon})\big)\quad\text{is smaller than } \varepsilon^{1-2\kappa}\big\langle M^\varepsilon\big\rangle_t \text{ in } \varepsilon\to 0.$$

Now, it might be expected that $(\varepsilon^{1/2-\kappa}W_t^\varepsilon)$, $(\varepsilon^{1/2-\kappa}M_t^\varepsilon)$, $(\varepsilon^{1/2}\Upsilon^{1/2-\kappa}W_t)$ share the same LDP.

Thus, we deal with LDP of a moderate deviation type (shortly, MDP). The MDP evaluation results are well known for many settings: see the papers by Borovkov, Mogulski [4, 5], Chen [8], Ledoux [21] (processes with independent increments); Guillin [16,17] (averaging principle and MDP); Dembo [9] (martingales with bounded jumps); Dembo and Zajic [10] (functional of empirical processes); Dembo and Zeitouni [11] (iterates of expanding maps); Liptser [23] (stationary process and MDP); Liptser and Spokoiny [28] (MDP for integral functionals of diffusion processes); Puhalskii [35] (queues in critical loading); Chang, Yao, Zajic [7] (queues with long-range dependent input); Wu [39] (Markov processes) and [40] (Hamiltonian systems).

In accordance to given above comparative analysis for

$$(\varepsilon^{1/2}W_t^\varepsilon)\quad\text{and}\quad(\varepsilon^{1/2-\kappa}W_t^\varepsilon)$$

we restrict ourselves by the LDP examination for $X^\varepsilon=\big((X_t^\varepsilon)_{t\geq 0},\varepsilon\to 0\big)$ when X_t^ε solves an ordinary differential equations (compare (1.2))

$$\dot X_t^\varepsilon=F(X_t^\varepsilon)+\varepsilon^{1/2-\kappa}G(X_t^\varepsilon)\dot W_t^\varepsilon, \tag{1.12}$$

where $\dot W_t^\varepsilon$ is defined in (1.3) and $0<\kappa<1/2$. We show that X^ε shares the LDP with a family $\widehat X^\varepsilon=\big((\widehat X_t^\varepsilon)_{t\geq 0},\varepsilon\to 0\big)$ of diffusion type non-Markovian processes defined by the Itô equation with respect to the continuous martingale M_t^ε defined in (1.9):

$$d\widehat X_t^\varepsilon=F(\widehat X_t^\varepsilon)\,dt+\varepsilon^{1/2-\kappa}G(\widehat X_t^\varepsilon)\,dM_t^\varepsilon, \tag{1.13}$$

subject to $\widehat X_0^\varepsilon=x_0$. So, the original problem is reduced to LDP examination for diffusion type processes with "stochastic homogenization" of diffusion parameter (see (1.10)). If G is a constant matrix, related LDP results can be found in [2,36], and also in Example 7.3 in [32]. In the case considered $G=G(\widehat X_t^\varepsilon)$, so that the direct use of the above-mentioned results is not applicable. We apply Theorem 2.3 from [33], which allows to establish in this case that $\widehat X^\varepsilon$ shares the LDP with a family $\widetilde X^\varepsilon$ of Markov diffusions

$$\begin{aligned}d\widetilde X_t^\varepsilon&=F(\widetilde X_t^\varepsilon)\,dt+\varepsilon^{1/2-\kappa}G(\widetilde X_t^\varepsilon)\Upsilon^{1/2}\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big)\\ \widetilde X_0^\varepsilon&=x_0\end{aligned}$$

if the matrix $Q(x) = G(x)\Upsilon G^*(x)$ is nonsingular for all $x \in \mathbf{R}^\ell$.

We include the case of singular matrix $Q(X_t)$ into consideration as well and show that the LDP remains valid, under additional condition, and the inverse matrix Q^{-1} is replaced by pseudoinverse one Q^+. The proof of this result uses some regularization procedure having an independent interest.

The paper is organized as follows. In Section 2, the assumptions are given and main results are formulated. The Poisson decomposition for an integral functional of Markov processes is given in Section 3. Sections 4–6 hold all proofs. Some auxiliary results are given in Appendix.

2. Assumptions. Main results

2.1. Assumptions

Equations (1.2) and (1.4) are subject to fixed x_0 and z_0 respectively. Henceforth, $\|\cdot\|$ is Euclidean norm. If x is vector with entries $x(1), \dots, x(\ell)$ write $|x| = \sum_{i=1}^{\ell} |x(i)|$.

(A$_F$) Entries of F are continuously differentiable functions; their partial derivatives are bounded; particularly for some constant $K > 0$

$$\sum_{i=1}^{\ell} |F_i(x)| \le K\Big(1 + \sum_{i=1}^{\ell} |x_i|\Big).$$

(A$_G$) Entries of G are bounded and twice continuously differentiable functions; their partial derivatives are bounded.

(A$_\sigma$) Entries of σ are bounded and Lipschitz continuous functions;

$$\inf_x \inf_{z:\|z\|=1} \big(\sigma\sigma^*(x)z, z\big) > 0.$$

(A$_b$) Entries of b are locally bounded functions; there exist positive constants C, r and α such that for $\|x\| > C$

$$\Big(b(x), \frac{x}{\|x\|}\Big) \le -r\|x\|^\alpha.$$

(A$_g$) Entries of g are measurable functions satisfying (1.5); with α from (A$_b$) and some $\beta < 0$ there is a constant $C > 0$ such that

$$\|g(z)\| \le C(1 + \|z\|)^{\beta+\alpha-1}.$$

2.2. The LDP

Denote by $Q^+(x)$ the Moore–Penrose pseudoinverse matrix (see, e.g. [1]) for $Q(x)$. Parallel to $\mathfrak{F} = \big\{X \in \mathbf{C}_{[0,\infty)}(\mathbf{R}^\ell) : dX_t = \dot{X}_t dt\big\}$ introduce (henceforth $\mathcal{I}$ is a unite matrix)

$$\widetilde{\mathfrak{F}} = \mathfrak{F} \bigcap \Big\{ \int\limits_{\{(QQ^+(X_t)-\mathcal{I})(\dot{X}_t - F(X_t))\neq 0\}} dt = 0 \Big\}. \tag{2.1}$$

Set (compare (1.1))

$$J(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int\limits_0^\infty \| \dot{X}_t - F(X_t) \|^2_{Q^+(X_t)}\, dt, & X \in \widetilde{\mathfrak{F}}, \\ \infty, & \text{otherwise.} \end{cases} \tag{2.2}$$

We examine the LDP in $\mathbf{C}_{[0,\infty)}(\mathbf{R}^\ell)$ supplied by the local uniform metric

$$\rho_\infty(X', X'') = \sum_{n=1}^\infty 2^{-n}\big(1 \wedge \rho_n(X', X'')\big),$$

where

$$\rho_T(X', X'') = \sup_{t \le T} \sum_{i=1}^\ell \big|X'_t(i) - X''_t(i)\big|, \quad T > 0 \tag{2.3}$$

and $X'_t(i), X''_t(i)$'s are components of X'_t, X''_t respectively. We follow here the standard Varadhan's definition of the LDP, [37].

Theorem 2.1. *Under* (A_F), (A_G), (A_b), (A_σ), (A_g) *and* (1.5) *the family* (X^ε), $\varepsilon \to 0$, *defined by* (1.12) *with* $\dot{W}^\varepsilon_t$ *from* (1.3), *obeys the LDP with the speed* $\varepsilon^{1-2\kappa}$ *and rate function* (2.2).

Remark 2.1. Theorem 2.1 serves the MDP in the context of Wu [39] and Guillin [16]

$$X^\varepsilon_t = \varepsilon^{1/2-\kappa} \int\limits_0^{t/\varepsilon} g(\xi_s)\, ds.$$

Under (A_g) the MDP holds with the speed $\varepsilon^{1-2\kappa}$ and rate function

$$J(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int\limits_0^\infty \|\dot{X}_t\|^2_{\Upsilon^+} dt, & X \in \widetilde{\mathfrak{F}}, \\ \infty, & \text{otherwise.} \end{cases}$$

Notice that non-singularity of Υ is not required. Moreover under $\alpha > 1$ and $\beta = 1 - \alpha$ the function g might be taken unbounded while in [39] and [16] g is bounded.

Theorem 2.2. *Let* $(\widehat{X}_t^\varepsilon)_{t\ge 0}$ *be defined in* (1.13). *Under the assumptions of Theorem* 2.1 *the family* $\widehat{X}^\varepsilon = \big((\widehat{X}^\varepsilon)_{t\ge 0}, \varepsilon \to 0\big)$, *obeys the LDP with the speed* $\varepsilon^{1-2\kappa}$ *and rate function* (2.2).

2.3. Preliminaries

In this section, we briefly describe methods for proving Theorem 2.1 and 2.2.

It is obvious that the statement of Theorem 2.1 holds true provided that Theorem 2.2 is valid and for any $T > 0$ and $\eta > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\left(\rho_T(X^\varepsilon, \widehat{X}^\varepsilon) > \eta\right) = -\infty. \tag{2.4}$$

The proof of Theorem 2.2 uses Puhalskii's result (Theorem 2.3 in [33], see also [34]). In accordance to it, Theorem 2.2 holds if

1) $Q(x) = G(x)\Upsilon G^*(x)$ is uniformly nonsingular in x matrix;

2) fast homogenization of diffusion parameter holds: for any $T > 0$, $\eta > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\left(\sup_{t\le T} \Big\| \int_0^t \big[G(\widehat{X}_s^\varepsilon)\{d\langle M^\varepsilon\rangle_s - \Upsilon\, ds\}G^*(\widehat{X}_s^\varepsilon)\big] \Big\| > \eta\right) = -\infty. \tag{2.5}$$

For singular $Q(x)$, the statement of Theorem 2.2 is proved with a help of mentioned in Introduction regularization procedure.

3. Poisson decomposition

Recall that $\mathcal{L}$ and μ are the diffusion operator and invariant measure of ξ respectively. Consider the Poisson equation

$$\mathcal{L}u = -f, \tag{3.1}$$

where f is any entry of g. Pardoux an Veretennikov (Theorem 2 in [30]) proved that, under (A_σ), (A_b) and (A_g), (3.1) possesses solution with properties (∇u is the gradient of u): if for some $\beta < 0$ there is $C > 0$ such that $\|f(x)\| \le C(1 + \|x\|)^{\beta+\alpha-1}$, then

$$|u(z)| \le C \sup_z \big\{|f(z)|(1 + \|z\|)^{-\beta-\alpha+1}\big\},$$
$$\|\nabla u(z)\| \le C(1 + \|z\|^{(\beta+\alpha-1)^+}).$$

In the sequel, we shall consider only centered solutions, i.e. $\int_{\mathrm{R}^d} u(z)\mu(dz) = 0$, which is unique in the above described class. It is proved in [30] that Sobolev's partial second derivatives of u exist and are locally integrable in any power. So, by Krylov's generalization of the Itô formula, [19], we find

$$u(\xi_t) = u(z_0) + \int_0^t \mathcal{L}u(\xi_s)\, ds + \int_0^t \nabla^* u(\xi_s)\sigma(\xi_s)\, dB_s$$

and, due to (3.1), we get the Poisson decomposition

$$\int_0^t f(\xi_s)\,ds = u(z_0) - u(\xi_t) + \int_0^t \nabla^* u(\xi_s)\sigma(\xi_s)\,dB_s. \tag{3.2}$$

4. Proof of Theorem 2.2. Nonsingular Q

4.1. Fast homogenization of diffusion parameter

It suffices to prove (2.5) in a coordinate form. By (1.10)

$$d\langle M^\varepsilon\rangle_s = \Psi(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\sigma^*(\xi_{s/\varepsilon})\Psi^*(\xi_{s/\varepsilon})\,ds.$$

Since entries of $G(\widehat{X}^\varepsilon_s)\big\{d\langle M^\varepsilon\rangle_s - \Upsilon\,ds\big\}G^*(\widehat{X}^\varepsilon_s)$ have the following structure: $h(\widehat{X}^\varepsilon_s)q(\xi_{s/\varepsilon})ds$, we shall show that

$$\lim_{\varepsilon\to 0}\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\sup_{t\le T}\Big|\int_0^t h(\widehat{X}^\varepsilon_s)q(\xi_{s/\varepsilon})ds\Big| > \eta\Big) = -\infty. \tag{4.1}$$

By (A_G), h is bounded, twice continuously differentiable and its partial derivatives are bounded. By (A_σ) and the boundedness of Ψ, the function q is bounded. Notice also that (1.11) provides

$$\int_{\mathrm{R}^d} q(z)\mu(dz) = 0.$$

To establish (4.1), we use a decomposition $q(z) = q'(z) + q''(z)$ with

$$q'(z) = \begin{cases} q(z), & \|z\|\le m,\\ 0, & \|z\|\ge m+1\end{cases} \quad \text{and} \quad \int_{\mathrm{R}^d} q'(z)\mu(dz) = 0.$$

For m large enough, q' with desired properties exist. In addition, $q'' = q - q'$ possesses the property

$$|q''(z)| \le C_m I(\|z\|\ge m),$$

where C_m is some constant. If suffices to show that

$$\lim_{\varepsilon\to 0}\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\sup_{t\le T}\Big|\int_0^t h(\widehat{X}^\varepsilon_s)q'(\xi_{s/\varepsilon})\,ds\Big| \ge \eta\Big) = -\infty \tag{4.2}$$

$$\lim_{\varepsilon\to 0}\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\int_0^T I(\|\xi_{t/\varepsilon})\| \ge m)\,dt \ge \eta\Big) = -\infty. \tag{4.3}$$

4.1.1. Proof of (4.2)

For verification of (4.2) we apply the Poisson decomposition. Since q' is bounded and compactly supported, the Poisson equation $\mathcal{L}u = -q'$ possesses a bounded solution u with bounded gradient ∇u. Then

$$\int_0^t q'(\xi_s)\,ds = u(z_0) - u(\xi_t) + \int_0^t \nabla^* u(\xi_s)\sigma(\xi_s)\,dB_s$$

and so

$$\varepsilon u(\xi_{t/\varepsilon}) = \varepsilon u(z_0) - \int_0^t q'(\xi_{s/\varepsilon})\,ds + \sqrt{\varepsilon}\int_0^t \nabla^* u(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big).$$

By the Itô formula

$$h\big(\widehat{X}^\varepsilon_t\big) = h\big(\widehat{X}^\varepsilon_0\big) + \int_0^t \mathfrak{h}_\varepsilon\big(\widehat{X}^\varepsilon_s, \xi_{s/\varepsilon}\big)\,ds + \frac{1}{2}\varepsilon^{1/2-\kappa}\int_0^t \mathfrak{h}\big(\widehat{X}^\varepsilon_s, \xi_{s/\varepsilon}\big)\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big),$$

where $\mathfrak{h}_\varepsilon(x,z)$ is a function of arguments $x \in \mathbf{R}^\ell$, $z \in \mathbf{R}^d$, depending on parameter ε, and $\mathfrak{h}$ is vector row of the same arguments. By (A_F), (A_G) and (A_σ), the function $\mathfrak{h}_\varepsilon(x,z)$ satisfies the linear growth condition

$$\|\mathfrak{h}_\varepsilon(x,z)\| \le r(1 + \|x\|) \tag{4.4}$$

with a constant r independent of z and entries of $\mathfrak{h}$ are bounded. Now, applying the Itô formula to $\varepsilon h(\widehat{X}^\varepsilon_t)u(\xi_{t/\varepsilon})$, we find

$$\begin{aligned}
\varepsilon h\big(\widehat{X}^\varepsilon_t\big)u(\xi_{t/\varepsilon}) &= \varepsilon h(x_0)u(z_0) - \int_0^t h\big(\widehat{X}^\varepsilon_s\big)q'(\xi_{s/\varepsilon})\,ds \\
&\quad + \sqrt{\varepsilon}\int_0^t h\big(\widehat{X}^\varepsilon_s\big)\nabla u(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big) \\
&\quad + \varepsilon\int_0^t u(\xi_{s/\varepsilon})\mathfrak{h}_\varepsilon\big(\widehat{X}^\varepsilon_s, \xi_{s/\varepsilon}\big)\,ds \\
&\quad + \frac{1}{2}\varepsilon^{3/2-\kappa}\int_0^t u(\xi_{s/\varepsilon})\mathfrak{h}\big(\widehat{X}^\varepsilon_s, \xi_{s/\varepsilon}\big)\,dM^\varepsilon_s \\
&\quad + \frac{1}{2}\varepsilon^{1-\kappa}\int_0^t \mathfrak{h}\big(\widehat{X}^\varepsilon_s, \xi_{s/\varepsilon}\big)\sigma(\xi_{s/\varepsilon})\nabla(\xi_{s/\varepsilon})\,ds.
\end{aligned}$$

Set

$$N_t^\varepsilon = \varepsilon^k \int_0^t h\big(\widehat{X}_s^\varepsilon\big) \nabla u(\xi_{s/\varepsilon}) \sigma(\xi_{s/\varepsilon})\, d\big(\sqrt{\varepsilon} B_{s/\varepsilon}\big)$$
$$+ \frac{\varepsilon}{2} \int_0^t u(\xi_{s/\varepsilon}) \mathfrak{h}\big(\widehat{X}_s^\varepsilon, \xi_{s/\varepsilon}\big)\, dM_s^\varepsilon.$$

Since h, u, ∇u and $\mathfrak{h}$ are bounded and $\mathfrak{h}_\varepsilon$ satisfies (4.4) with a generic constant r, we have

$$\sup_{t \le T} \Big| \int_0^t h\big(\widehat{X}_s^\varepsilon\big) q'(\xi_{s/\varepsilon})\, ds \Big| \le r\varepsilon^{1-\kappa} \big(1 + \sup_{t \le T} \big|\widehat{X}_t^\varepsilon\big|\big) + \varepsilon^{1/2-\kappa} \sup_{t \le T} |N_t^\varepsilon|. \tag{4.5}$$

With a help of Lemma A.1 (see Appendix), it is readily to derive that

$$\lim_{C \to \infty} \overline{\lim_{\varepsilon \to 0}}\, \varepsilon^{1-2\kappa} \log \mathsf{P} \left(\sup_{t \le T} \big|\widehat{X}_t^\varepsilon\big| > C \right) = -\infty.$$

Hence, the proof of (4.2) is reduced to

$$\lim_{\varepsilon \to 0} \varepsilon^{1-2\kappa} \log \mathsf{P} \left(\varepsilon^{1/2-\kappa} \sup_{t \le T} \big|N_t^\varepsilon\big| > C \right) = -\infty.$$

The process N_t^ε is a continuous martingale with $d\langle N^\varepsilon \rangle_t \le r\varepsilon^k dt$, so that the desired result is implied by Corollary A.1 to Lemma A.1. □

4.1.2. Proof of (4.3)

Introduce a nonlinear operator $\mathcal{D}$ acting on twice continuously differentiable function v as

$$\mathcal{D}v(z) = \mathcal{L}v(z) + \frac{1}{2} \|\nabla v(z) \sigma(z)\|^2. \tag{4.6}$$

We apply $\mathcal{D}$ to

$$v(z) = \frac{\|z\|^2}{1 + \|z\|}.$$

The gradient

$$\nabla v(z) = \frac{\|z\|(2 + \|z\|)}{(1 + \|z\|)^2} \frac{z}{\|z\|}$$

is bounded ($\|\nabla v(z)\| \le \text{const}$, so $\|\nabla v(z)\sigma(z)\|^2 \le \text{const}$) and the boundedness of the second partial derivatives of v is readily verified.

Set $U_t^\varepsilon = v(\xi_{t/\varepsilon}) - v(z_0) - \varepsilon^{-1}\int_0^t \mathcal{D}v(\xi_{s/\varepsilon})\,ds$. By the Itô formula

$$\begin{aligned}
U_t^\varepsilon &= \frac{1}{\sqrt{\varepsilon}}\int_0^t \nabla^* v(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big) + \frac{1}{\varepsilon}\int_0^t (\mathcal{L}-\mathcal{D})v(\xi_{s/\varepsilon})\,ds \\
&= \frac{1}{\sqrt{\varepsilon}}\int_0^t \nabla^* v(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}\big) - \frac{1}{2\varepsilon}\int_0^t \|\nabla v(\xi_{s/\varepsilon})\sigma(\xi_{s/\varepsilon})\|^2\,ds \\
&= \text{“continuous martingale}-\frac{1}{2}\text{ of quadratic variation process”}.
\end{aligned}$$

Hence, $Z_t^\varepsilon = \exp\big(U_t^\varepsilon\big)$ is a positive continuous local martingale with $Z_0^\varepsilon = 1$ and supermartingale as well (see Problem 1.4.4 in [27]). Then $\mathsf{E}\,Z_T^\varepsilon \le 1$. So, with

$$\mathfrak{A} = \Big\{\int_0^T I\big(\|\xi_{t/\varepsilon})\| \ge m\big)\,dt \ge \eta\Big\}$$

we have $1 \ge \mathsf{E}\,I_{\mathfrak{A}} Z_T^\varepsilon$. In the latter inequality, we replace Z_T^ε by its lower bound on $\mathfrak{A}$. With $\mathcal{A}v := \sum\limits_{ij}(\sigma\sigma^*)_{ij}v''_{ij}$, write $\mathcal{L}v(z) = \nabla v(z)b(z) + \mathcal{A}v(z)/2$. Then (see (4.6))

$$\mathcal{D}v(z) = \frac{(2+\|z\|)}{(1+\|z\|)^2}z^*b(z) + \frac{1}{2}\mathcal{A}v(z) + \frac{1}{2}\|\nabla v(z)\sigma(z)\|^2.$$

Recall $\|\mathcal{A}v(z)\|$ and $\|\nabla v(z)\sigma(z)\|$ are bounded and by (A_b) for $\|z\| > C$ we have $z^*b(z) \le -r\|z\|^{1+\alpha}$ with $r>0$ and $\alpha>0$. Hence

$$\sup_{z\in\mathrm{R}^d}\mathcal{D}v(z) = \widetilde{v} < \infty, \quad \lim_{y\to\infty}\inf_{\|z\|>y}(-\mathcal{D}v(z)+\widetilde{v}) = \infty.$$

We give the lower bound for Z_T^ε expressed in terms of $\widetilde{v}$ and $(-\mathcal{D}v(z)+\widetilde{v})$. With m such that $\widetilde{v}T < \inf\limits_{\|z\|>m}\big(-\mathcal{D}v(z)+\widetilde{v}\big)\eta$, write

$$\begin{aligned}
\log Z_T^\varepsilon &= v(\xi_{T/\varepsilon}) - v(z_0) - \frac{1}{\varepsilon}\int_0^T \mathcal{D}v(\xi_{s/\varepsilon})\,ds \\
&\ge -v(z_0) - \frac{\widetilde{v}}{\varepsilon}T + \frac{1}{\varepsilon}\int_0^T I\big(\|\xi_{s/\varepsilon}\| > m\big)\big(-\mathcal{D}v(\xi_{s/\varepsilon})+\widetilde{v}\big)\,ds \\
&\ge -v(z_0) - \frac{\widetilde{v}}{\varepsilon}T + \frac{1}{\varepsilon}\inf_{\|z\|>m}\big(-\mathcal{D}v(z)+\widetilde{v}\big)\eta \quad := \quad \log Z_*.
\end{aligned}$$

Since obviously $1 \ge \mathsf{E}\,I_{\mathfrak{A}}Z_*$ and Z_* is nonrandom, we find

$$\varepsilon^{1-2\kappa}\log\mathsf{P}\,(\mathfrak{A}) \le \varepsilon^{1-2\kappa}v(z_0) + \frac{1}{\varepsilon^{2\kappa}}\Big(\widetilde{v}T - \inf_{\|z\|>m}\big(-\mathcal{D}v(z)+\widetilde{v}\big)\eta\Big) \xrightarrow[\varepsilon\to 0]{} -\infty.$$

□

Thus, for nonsingular Q the statement of Theorem 2.2 is valid. □

5. Proof of Theorem 2.2 for singular Q

5.1. Preliminaries

In this section, $Q(X_t)$ is not assumed to be nonsingular for any X_t. Despite (2.5) remains to hold, the statement of Puhalskii's Theorem 2.3 from [33] is not longer valid. So, we apply another way for proving Theorem 2.2 which based on obtained LDP result for nonsingular Q.

The Dawson – Gärtner theorem (see e.g. [11] or [31]), adapted to the case considered, states that announced LDP for family $\widehat{X}^\varepsilon$ holds, if for any $T > 0$ the LDP for the family $(\widehat{X}^\varepsilon)_T = \big((\widehat{X}^\varepsilon_t)_{0\le t\le T}, \varepsilon \to 0\big)$ holds in the metric space $(\mathbf{C}_{[0,T]}(R^\ell), \rho_T)$ with the speed $\varepsilon^{1-2\kappa}$ and rate function (compare (2.2))

$$J_T(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \| \dot{X}_t - F(X_t) \|^2_{Q^+(X_t)}\, dt, & X \in \widetilde{\mathfrak{F}}_T, \\ \infty, & \text{otherwise}, \end{cases} \tag{5.1}$$

where $\widetilde{\mathfrak{F}}_T$ is the restriction of $\widetilde{\mathfrak{F}}$ on $[0,T]$. The proof of this statement requires the verification of exponential tightness with speed $\varepsilon^{1-2\kappa}$ and local LDP with the same speed and rate function J_T: for any $X \in \mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$

$$\varliminf_{\delta\to 0}\varliminf_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X) \le \delta\right) \ge -J_T(X)z \tag{5.2}$$

and

$$\varlimsup_{\delta\to 0}\varlimsup_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X) \le \delta\right) \le -J_T(X). \tag{5.3}$$

5.2. Main lemma

With some $\gamma > 0$ we introduce a process

$$\widehat{X}^{\varepsilon,\gamma}_t = x_0 + \int_0^t F\big(\widehat{X}^{\varepsilon,\gamma}_s\big)\, ds + \varepsilon^{1/2-\kappa} M^{\varepsilon,\gamma}_t \tag{5.4}$$

with

$$M^{\varepsilon,\gamma}_t = \int_0^t G\big(\widehat{X}^{\varepsilon,\gamma}_s\big)\, dM^\varepsilon_s + \gamma^{1/2}\widehat{B}_t, \tag{5.5}$$

and $(\widehat{B}_t)_{t\ge 0}$ the standard vector-valued Wiener process (of suitable size) independent of $(\xi_t)_{t\ge 0}$.

Lemma 5.1. *Under the assumptions from Section 2, for any $T > 0$, $\eta > 0$*

$$\lim_{\gamma\to 0}\varlimsup_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\left(\rho_T(\widehat{X}^{\varepsilon,\gamma}, \widehat{X}^\varepsilon) > \eta\right) = -\infty.$$

Proof. Set $\triangle_t^{\varepsilon,\gamma} = \widehat{X}_t^{\varepsilon,\gamma} - \widehat{X}_t^{\varepsilon}$. Due to (1.13), (5.4) and (5.5) we find

$$\triangle_t^{\varepsilon,\gamma} = \int_0^t \big(F\big(\widehat{X}_s^{\varepsilon,\gamma}\big) - F\big(\widehat{X}_s^{\varepsilon}\big)\big)\,ds + \varepsilon^{1/2-\kappa}\int_0^t \big(G\big(\widehat{X}_s^{\varepsilon,\gamma}\big) - G\big(\widehat{X}_s^{\varepsilon}\big)\big)\,dM_s^{\varepsilon} + \varepsilon^{1/2-\kappa}\gamma^{1/2}\widehat{B}_t. \tag{5.6}$$

Denote $|\triangle_t^{\varepsilon,\gamma}| = \sum_{i=1}^{\ell}|\triangle_t^{\varepsilon,\gamma}(i)|$. Since entries of F and G are continuously differentiable and their derivatives are bounded, entries of

$$\mathfrak{f}(s) = \frac{\big(F\big(\widehat{X}_s^{\varepsilon,\gamma}\big) - F\big(\widehat{X}_s^{\varepsilon}\big)\big)}{\big|\triangle_s^{\varepsilon,\gamma}\big|} \quad \text{and} \quad \mathfrak{g}(s) = \frac{\big(G\big(\widehat{X}_s^{\varepsilon,\gamma}\big) - G\big(\widehat{X}_s^{\varepsilon}\big)\big)}{\big|\triangle_s^{\varepsilon,\gamma}\big|}$$

are well defined and bounded. Now, we may rewrite (5.6) to

$$\triangle_t^{\varepsilon,\gamma} = \int_0^t \big|\triangle_s^{\varepsilon,\gamma}\big|\mathfrak{f}_s\,ds + \varepsilon^{1/2-\kappa}\int_0^t \big|\triangle_s^{\varepsilon,\gamma}\big|\,dN_s^{\varepsilon} + \varepsilon^{1/2-\kappa}\gamma^{1/2}\widehat{B}_t, \tag{5.7}$$

where $N_t^{\varepsilon} = \int_0^t \mathfrak{g}_s\,dM_s^{\varepsilon}$ and entries $N_t^{\varepsilon}(i)$ of N_t^{ε} are continuous martingales with the predictable quadratic variation processes $\langle N^{\varepsilon}(i)\rangle_t$ absolutely continuous with respect to dt with bounded densities, i.e.

$$d\langle N^{\varepsilon}(i)\rangle_t \le r\,dt \tag{5.8}$$

(henceforth r is positive generic constant). Owing to $\|\triangle_t^{\varepsilon,\gamma}\|^2 = (\triangle_t^{\varepsilon,\gamma})^*\triangle_t^{\varepsilon,\gamma}$, by the Itô formula we find

$$\begin{aligned}\|\triangle_t^{\varepsilon,\gamma}\|^2 &= \int_0^t 2|\triangle_s^{\varepsilon,\gamma}|(\triangle_s^{\varepsilon,\gamma})^*\mathfrak{f}(s)\,ds + \varepsilon^{1/2-\kappa}\int_0^t 2|\triangle_s^{\varepsilon,\gamma}|(\triangle_s^{\varepsilon,\gamma})^*\,dN_s^{\varepsilon}\\ &\quad + \varepsilon^{1/2-\kappa}\gamma^{1/2}\int_0^t 2(\triangle_s^{\varepsilon,\gamma})^*\,d\widehat{B}_t + \varepsilon^{1-2\kappa}\int_0^t |\triangle_s^{\varepsilon,\gamma}|^2\,d(\operatorname{trace}\langle N^{\varepsilon}\rangle_s)\\ &\quad + \varepsilon^{1-2\kappa}\gamma\ell t.\end{aligned} \tag{5.9}$$

Letting $0/0 = 0$, introduce $\mathfrak{i}^*(s) = 2|\triangle_s^{\varepsilon,\gamma}|(\triangle_s^{\varepsilon,\gamma})^* \,/\, \|\triangle_s^{\varepsilon,\gamma}\|^2$, $\mathfrak{j}(s) = 2\mathfrak{i}(s)\mathfrak{f}(s)$ and $\mathfrak{r}(s) = |\triangle_s^{\varepsilon,\gamma}|^2 \,/\, \|\triangle_s^{\varepsilon,\gamma}\|^2$. Obviously, $\mathfrak{r}(s)$, $\mathfrak{j}(s)$ and entries of $\mathfrak{i}^*(s)$ are bounded. Set

$$U(t) = \varepsilon^{1/2-\kappa}\gamma^{1/2}\int_0^t 2(\triangle_s^{\varepsilon,\gamma})^*d\widehat{B}_s + \varepsilon^{1-2\kappa}\gamma\ell t.$$

With $U(t)$, we rewrite (5.9) into a linear Itô equation with respect to $\|\triangle_t^{\varepsilon,\gamma}\|^2$:

$$\|\triangle_t^{\varepsilon,\gamma}\|^2 = \int_0^t \|\triangle_s^{\varepsilon,\gamma}\|^2 \big\{\mathfrak{j}(s)\,ds + \varepsilon^{1/2-\kappa}\mathfrak{i}^*(s)\,dN_s^\varepsilon + \varepsilon^{1-2\kappa}\mathfrak{r}(s)\,d(\mathrm{trace}\langle N^\varepsilon\rangle_s)\big\} + U(t). \tag{5.10}$$

Since $\langle N^\varepsilon, \widehat{B}\rangle_t \equiv 0$, applying the Itô formula to $\mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}\,dU(s)$, with

$$\mathcal{E}_t = \exp\Big(\int_0^t \Big\{\mathfrak{j}(s)\,ds + \varepsilon^{1/2-\kappa}\mathfrak{i}^*(s)\,dN_s^\varepsilon + \varepsilon^{1-2\kappa}\mathfrak{r}(s)\,d(\mathrm{trace}\langle N^\varepsilon\rangle_s) - \frac{1}{2}\varepsilon^{1-2\kappa}\mathfrak{i}^*(s)\,d\langle N^\varepsilon\rangle_s\mathfrak{i}(s)\Big\}\Big),$$

we find $\|\triangle_t^{\varepsilon,\gamma}\|^2 = \mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}\,dU(s)$. The statement of lemma is valid, if

$$\lim_{\gamma\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log P\Big(\sup_{t\le T}\|\triangle_t^{\varepsilon,\gamma}\|^2 > \eta\Big) = -\infty. \tag{5.11}$$

For (5.11) to hold, it suffices

$$\begin{aligned}
&\lim_{C\to\infty}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\sup_{t\le T}\mathcal{E}_t > C\Big) = -\infty,\\
&\lim_{C\to\infty}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\sup_{t\le T}\mathcal{E}_t^{-1} > C\Big) = -\infty,\\
&\lim_{C\to\infty}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\sup_{t\le T}\|\triangle_t^{\varepsilon,\gamma}\|^2 > C\Big) = -\infty, \quad \gamma > 0.
\end{aligned} \tag{5.12}$$

In fact, if (5.12) is valid, (5.11) is reduced to: for any $C > 0$

$$\lim_{\gamma\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\varepsilon^{1/2-\kappa}\gamma^{1/2}\sup_{t\le T}\Big|\int_0^t \mathcal{E}_s^{-1}2(\triangle_s^{\varepsilon,\gamma})^*\,d\widehat{B}_s\Big| > \eta, \tag{5.13}$$
$$\sup_{t\le T}\mathcal{E}_t \le C,\quad \sup_{t\le T}\mathcal{E}_t^{-1} \le C,\quad \sup_{t\le T}\|\triangle^{\varepsilon,\gamma}\|_t^2 \le C\Big) = -\infty.$$

Further, on the set $\big\{\sup_{t\le T}\mathcal{E}_t^{-1} \le C,\ \ \sup_{t\le T}\|\triangle^{\varepsilon,\gamma}\|_t^2 \le C\big\}$, the integral $\int_0^t \mathcal{E}_s^{-1}2(\triangle_s^{\varepsilon,\gamma})^*\,d\widehat{B}_s$ coincides with

$$I_{A_t}\int_0^t I_{A_s}\mathcal{E}_s^{-1}2(\triangle_s^{\varepsilon,\gamma})^*\,d\widehat{B}_s,$$

where $A_s = \big\{\sup_{s'\le s}\mathcal{E}_{s'}^{-1} \le C,\ \sup_{s'\le s}\|\triangle^{\varepsilon,\gamma}\|_t^2 s' \le C\big\}$. Hence, (5.13) is reduced to

$$\lim_{\gamma\to 0}\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\Big(\varepsilon^{1/2-\kappa}\sup_{t\le T}\big|M_t\big| > \frac{\eta}{\gamma^{1/2}}\Big) = -\infty, \tag{5.14}$$

where $M_t = \int_0^t I_{A_s}\mathcal{E}_s^{-1} 2(\triangle_s^{\varepsilon,\gamma})^* d\widehat{B}_s$ is continuous martingale, $\langle M\rangle_T \le \text{const}$ So, the validity of (5.14) is established with a help of Lemma A.1.

The first and second conditions from (5.12) are implied by Lemma A.1. The proof of the third condition from (5.12) is valid by remark to Lemma A.1 and by $\triangle_t^{\varepsilon,\gamma} = \widehat{X}_t^{\varepsilon,\gamma} - \widehat{X}_t^{\varepsilon}$. □

5.3. The LDP for $\widehat{X}^{\varepsilon,\gamma}$

The family $\widehat{X}^{\varepsilon,\gamma} = \big((\widehat{X}^{\varepsilon,\gamma})_{t\ge 0}, \varepsilon \to 0\big)$, defined in (5.4), (5.5), obeys the LDP, since the matrix

$$Q^\gamma(x) = Q(x) + \gamma\mathcal{I} \tag{5.15}$$

is uniformly in x nonsingular. This LDP is characterized by the speed $\varepsilon^{1-2\kappa}$ and rate function

$$J^\gamma(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^\infty \|\dot{X}_t - F(X_t))\|^2_{(Q^\gamma(X_t))^{-1}}\, dt, & X \in \mathfrak{F}, \\ \infty, & \text{otherwise.} \end{cases}$$

Parallel to $J_T(X)$ introduce

$$J_T^\gamma(X) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \|\, \dot{X} - F(X_t)\,\|^2_{(Q^\gamma)^{-1}(X_t)}\; dt, & X \in \mathfrak{F}_T, \\ \infty, & \text{otherwise} \end{cases}$$

(here $\mathfrak{F}_T$ is the restriction of $\mathfrak{F}$ on $[0,T]$).

Lemma 5.2.

1) $\lim\limits_{\gamma\to 0} J_T^\gamma(X) = J_T(X)$, $X \in \mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$, $T > 0$.

2) *The function $J_T(X)$ is semicontinuous from below.*

Proof. 1) If $X \notin \mathfrak{F}_T$, then $J_T^\gamma(X) \equiv \infty$ as well as $J_T(X) \equiv \infty$. Let $X \in \mathfrak{F}_T \setminus \widetilde{\mathfrak{F}}_T$. Then $J_T(X) = \infty$. On the other hand, since

$$\|\dot{X}_t - F(X_t))\|^2_{(Q^\gamma(X_t))^{-1}}$$

increases in $\gamma \downarrow 0$, by the monotone convergence theorem

$$\lim_{\gamma\to 0} J_T^\gamma(X) = \frac{1}{2}\int_0^T \lim_{\gamma\to 0} \|\dot{X} - F(X_t)\|^2_{(Q^\gamma)^{-1}(X_t)}\, dt = \infty.$$

Let $X \in \widetilde{\mathfrak{F}}_T$. It suffices to show

$$\lim_{\gamma\to 0} \|\dot{X} - F(X_t)\|^2_{(Q^\gamma)^{-1}(X_t)} = \|\dot{X} - F(X_t)\|^2_{Q^+(X_t)}.$$

Notice that $QQ^+(X_t)(\dot{X}_t - F(X_t)) = (\dot{X}_t - F(X_t))$ provides

$$\|\dot{X} - F(X_t)\|^2_{(Q^\gamma)^{-1}(X_t)} = \|\dot{X}_t - F(X_t))\|^2_{QQ^+(Q^\gamma)^{-1}QQ^+(X_t)}$$

and it remains to check that $\lim_{\gamma\to 0} QQ^+(Q^\gamma)^{-1}QQ^+ = Q^+$. With S the orthogonal matrix ($S^* = S^{-1}$) transforming Q to a diagonal form: $S^*QS = \text{diag}Q$ we have $S^*Q^+S = \text{diag}Q^+$ (see e.g. [1]). So, it suffices to prove

$$\lim_{\gamma\to 0} SQQ^+(Q^\gamma)^{-1}QQ^+S^* = SQ^+S^*. \tag{5.16}$$

The right-hand side of (5.16) is the scalar matrix with nonnegative entries q^+_{ii}, where q_{ii} are entries of diag$Q$ and

$$q^+_{ii} = \begin{cases} q^{-1}_{ii}, & q^{-1}_{ii} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

At the same time the left-hand side of (5.16) is a scalar matrix as well

$$SQQ^+(Q^\gamma)^{-1}QQ^+S^* = \text{diag}Q\,\text{diag}Q^+(\gamma\mathcal{I} + \text{diag}Q)^{-1}\text{diag}Q\,\text{diag}Q^+$$

with entries $(q_{ii}q^+_{ii})^2/(\gamma + q_{ii}) \to q^+_{ii}$, $\gamma \to 0$.

2) As was mentioned above, $J^\gamma_T(X)$ is increasing in $\gamma \downarrow 0$. This remark and 1) provide $J^\gamma_T(X) \le J_T(X)$. Let X^n, $n \ge 1$, converge to X in the metric ρ_T, i.e. $\lim_n \rho_T(X^n, X) = 0$. Since J^γ_T is semicontinuous from below, we have

$$\varliminf_n J_T(X^n) \ge \varliminf_n J^\gamma_T(X^n) \ge J^\gamma_T(X) \to J_T(X), \quad \gamma \to 0.$$

□

5.4. The LDP for the family $(\widehat{X}^\varepsilon)_T$

Obviously, $\widehat{X}^{\varepsilon,\gamma}$ obeys the LDP in the metric space $\big(\mathbf{C}_{[0,T]}(\mathbf{R}^\ell), \rho_T\big)$ with the speed $\varepsilon^{1-2\kappa}$ and rate function J^γ_T. Hence, this family is exponentially tight in $\big(\mathbf{C}_{[0,T]}(\mathbf{R}^\ell), \rho_T\big)$ with the speed $\varepsilon^{1-2\kappa}$. The latter remark and Lemma 5.1 provide the exponential tightness in the same metric space and speed for family $(\widehat{X}^\varepsilon)_T$.

We apply the LDP result for $(\widehat{X}^{\varepsilon,\gamma})_T = ((\widehat{X}^{\varepsilon,\gamma})_{0\le t\le T}, \varepsilon \to 0)$ to establish the lower bound in local LDP for $(\widehat{X}^\varepsilon)_T$.

Lemma 5.3. *Under the assumptions from Section 2, for any $T > 0$ and $X \in \mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$*

$$\varliminf_{\delta\to 0} \varliminf_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\big(\rho_T(\widehat{X}^\varepsilon, X) \le \delta\big) \ge -J_T(X).$$

Proof. By the triangular inequality

$$\rho_T(\widehat{X}^\varepsilon, X) \le \rho_T(\widehat{X}^{\varepsilon,\gamma}, X) + \rho_T(\widehat{X}^{\varepsilon,\gamma}, \widehat{X}^\varepsilon)$$

we have

$$\begin{aligned}
\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},X)<\frac{\delta}{2}\Big) &\le \mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon},X)-\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})<\frac{\delta}{2}\Big)\\
&\le \mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})>\frac{\delta}{2}\Big)+\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)\le\delta\big)\\
&\le 2\Big\{\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})>\frac{\delta}{2}\Big)\bigvee\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)\le\delta\big)\Big\}
\end{aligned}$$

and so

$$\begin{aligned}
&\Big\{\varlimsup_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})>\frac{\delta}{2}\Big)\Big\}\\
&\bigvee\Big\{\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)\le\delta\big)\Big\}\\
&\qquad\ge\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},X)<\frac{\delta}{2}\Big).
\end{aligned}\tag{5.17}$$

The LDP for the family $(\widehat{X}^{\varepsilon,\gamma})_T$ provides

$$\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},X)<\frac{\delta}{2}\Big)\ge-\inf_{\{Y:\rho_T(Y,X)<\delta/2\}}J_T^{\gamma}(Y)\big(\ge-J_T^{\gamma}(X)\big)$$

while by Lemma 5.2 $\lim\limits_{\gamma\to0}J_T^{\gamma}(X)=J_T(X)$. Hence

$$\varliminf_{\gamma\to0}\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},X)<\frac{\delta}{2}\Big)\ge-J_T(X).$$

Now, owing to (5.17)

$$\begin{aligned}
&\Big\{\varlimsup_{\gamma\to0}\varlimsup_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})>\frac{\delta}{2}\Big)\Big\}\\
&\bigvee\Big\{\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)\le\delta\big)\Big\}\\
&\qquad\ge\varliminf_{\gamma\to0}\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\rho_T(\widehat{X}^{\varepsilon,\gamma},X)<\frac{\delta}{2}\Big)\ge-J_T(X).
\end{aligned}$$

By Lemma 5.1

$$\varlimsup_{\gamma\to0}\varlimsup_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^{\varepsilon})>\delta/2\big)=-\infty.$$

Consequently,

$$\varliminf_{\delta\to0}\varliminf_{\varepsilon\to0}\varepsilon^{1-2\kappa}\log\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)<\delta\big)\ge-J_T(X)$$

and it remains to notice that $\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)\ge\delta\big)\ge\mathsf{P}\big(\rho_T(\widehat{X}^{\varepsilon},X)<\delta\big)$. $\square$

5.4.1. Upper bound in the local LDP for $(\widehat{X}^\varepsilon)_T$

Obviously, the case $X_0 = x_0$ only has to be analyzed.

Denote (see (2.5))

$$\mathfrak{A}^{\varepsilon,\gamma} = \Big\{ \sup_{t\le T} \Big\| \int_0^t \big[G(\widehat{X}^{\varepsilon,\gamma}_s)\{d\langle M^\varepsilon\rangle_s - (\Upsilon + \gamma\mathcal{I})\,ds\}G^*(\widehat{X}^{\varepsilon,\gamma}_s)\big]\Big\| > \eta\Big\},$$

$\overline{\mathfrak{A}}^{\varepsilon,\gamma} = \Omega \setminus \mathfrak{A}^{\varepsilon,\gamma}$ and introduce the piece-wise constant vector-valued function $\lambda(s)$, $\lambda(s) = \lambda(s_k), s_{k-1} \le s < s_k$.

Proposition 5.1. *Under the assumptions from Section 2, for any X from $\mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$ with $X_0 = x_0$*

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}}\, \varepsilon^{1-2\kappa} \log \mathsf{P}\big(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X) \le \delta\big) \\
&\quad\le\; K_1(\lambda)\delta + K_2(\lambda)\eta \\
&\qquad - \int_0^T \Big(\lambda^*(s)\big(dX_s - F(X_s)ds\big) - \frac{1}{2}\lambda^*(s)Q^\gamma(X_s)\lambda(s)\Big)ds, \qquad (5.18)
\end{aligned}$$

where $\int_0^t \lambda(s)dX_s = \sum_k \lambda(s_k)(X_{s_k} - X_{s_{k-1}})$ and $K_i(\lambda)$, $i = 1, 2$, are positive constants depending on λ and independent of δ and γ.

Proof. With $M^{\varepsilon,\gamma}_t$ defined in (5.5) let us introduce a positive continuous local martingale

$$Z^{\varepsilon\gamma}_t = \exp\Big(\frac{1}{\varepsilon^{1/2-\kappa}} \int_0^t \lambda^*(s)\,dM^{\varepsilon,\gamma}_s - \frac{1}{2\varepsilon^{1-2\kappa}} \int_0^t \lambda^*(s)\,d\langle M^{\varepsilon,\gamma}\rangle_s \lambda(s)\Big).$$

By Problem 1.4.4 in [27], $Z^{\varepsilon\gamma}_t$ is also supermartingale, $\mathsf{E}\,Z^{\varepsilon\gamma}_T \le 1$. Taking into account

$$\begin{aligned}
dM^{\varepsilon,\gamma}_s &= \frac{d\widehat{X}^{\varepsilon,\gamma}_s - F(\widehat{X}^{\varepsilon,\gamma}_s)\,ds}{\varepsilon^{1/2-\kappa}} \\
d\langle M^{\varepsilon,\gamma}\rangle_s &= G^*(\widehat{X}^{\varepsilon,\gamma}_s)Q^\gamma(\widehat{X}^{\varepsilon,\gamma}_s)G(\widehat{X}^{\varepsilon,\gamma}_s)ds,
\end{aligned}$$

it is readily to derive that on the set $\{\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X) \le \delta\}$ the random variable $Z^{\varepsilon\gamma}_T$ is bounded from below by a positive nonrandom parameter

$$\begin{aligned}
Z_* = \exp\Big(\frac{1}{\varepsilon^{1-2\kappa}}\Big[&K_1(\lambda)\delta + K_2(\lambda)\eta \\
&- \int_0^T \Big(\lambda^*(s)\big(dX_s - F(X_s)\,ds\big) - \frac{1}{2}\lambda^*(s)Q^\gamma(X_s)\lambda(s)\Big)\,ds\Big]\Big).
\end{aligned}$$

Obviously, we have $1 \ge Z_* P\big(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X) \le \delta\big)$ and the result is done. $\square$

Lemma 5.4. *Under the assumptions from Section 2, for any $T > 0$ and $X \in \mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$*

$$\overline{\lim_{\delta\to 0}}\,\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\le -J_T(X).$$

Proof. The use of $\mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\le \mathsf{P}\left(\mathfrak{A}^{\varepsilon,\gamma}\right)+\mathsf{P}\left(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)$ and triangular inequality $\rho_T(\widehat{X}^{\varepsilon,\gamma}, X)\le\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)+\rho_T(\widehat{X}^\varepsilon, X)$ provides

$$\begin{aligned}
&\mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\\
&\quad\le\ \mathsf{P}\left(\mathfrak{A}^{\varepsilon,\gamma}\right)+\mathsf{P}\left(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X)-\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)\le\delta\right)\\
&\quad\le\ \mathsf{P}\left(\mathfrak{A}^{\varepsilon,\gamma}\right)+\mathsf{P}\left(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X)\le\frac{\delta}{2}\right)+\mathsf{P}\left(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)>\frac{\delta}{2}\right)\\
&\quad\le\ 3\Big\{\mathsf{P}\left(\mathfrak{A}^{\varepsilon,\gamma}\right)\bigvee \mathsf{P}\left(\overline{\mathfrak{A}}^{\varepsilon,\gamma}, \rho_T(\widehat{X}^{\varepsilon,\gamma}, X)\le\frac{\delta}{2}\right)\bigvee \mathsf{P}\left(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)>\frac{\delta}{2}\right)\Big\}.
\end{aligned}$$

By (4.1) $\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log\mathsf{P}\left(\mathfrak{A}^{\varepsilon,\gamma}\right)=0$. Taking also into account Proposition 5.1 we find

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\\
&\qquad\le\ \Big\{\Big[K_1(\lambda)\frac{\delta}{2}+K_2(\lambda)\eta\\
&\qquad\qquad-\int_0^T\Big(\lambda^*(s)\big(dX_s-F(X_s)\,ds\big)-\frac{1}{2}\lambda^*(s)Q^\gamma(X_s)\lambda(s)\Big)ds\Big]\\
&\qquad\qquad\bigvee\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log\mathsf{P}\left(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)>\frac{\delta}{2}\right)\Big\}.
\end{aligned}$$

By Lemma 5.1 $\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log\mathsf{P}\left(\rho_T(\widehat{X}^{\varepsilon,\gamma},\widehat{X}^\varepsilon)>\delta/2\right)\to-\infty,\ \gamma\to 0$. Hence, whereas $Q^\gamma(X_s)\to Q(X_s)$, $\gamma\to 0$, it holds

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\\
&\qquad\le\ \Big[K_1(\lambda)\frac{\delta}{2}+K_2(\lambda)\eta\\
&\qquad\qquad-\int_0^T\Big(\lambda^*(s)\big(dX_s-F(X_s)\,ds\big)-\frac{1}{2}\lambda^*(s)Q(X_s)\lambda(s)\Big)ds\Big]
\end{aligned}$$

and in turn

$$\begin{aligned}
&\overline{\lim_{\delta\to 0}}\,\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right)\\
&\qquad\le\Big[K_2(\lambda)\eta-\int_0^T\Big(\lambda^*(s)\big(dX_s-F(X_s)\,ds\big)-\frac{1}{2}\lambda^*(s)Q(X_s)\lambda(s)\Big)\,ds\Big].
\end{aligned}$$

Moreover, since the left-hand side of this inequality is independent of η, it also holds

$$\overline{\lim_{\delta\to 0}}\,\overline{\lim_{\varepsilon\to 0}}\,\varepsilon^{1-2\kappa}\log \mathsf{P}\left(\rho_T(\widehat{X}^\varepsilon, X)\le\delta\right) \le -\int_0^T \Big(\lambda^*(s)\big(dX_s - F(X_s)\big) - \frac{1}{2}\lambda^*(s)Q(X_s)\lambda(s)\Big)ds. \quad (5.19)$$

Finally, minimization of the right-hand side of (5.19) in λ gives a lower bound $-J'_T$ with $J'_T(X)=\infty$ for $X\in \mathbf{C}_{[0,t]}(\mathbf{R}^\ell)\setminus\widetilde{\mathfrak{F}}_T$ (see, e.g. Theorem 6.1 in [25]) and for $X\in\widetilde{\mathfrak{F}}_T$

$$J'_T=\int_0^T \sup_{\nu\in\mathrm{R}^\ell}\Big(\nu^*\big(\dot{X}_s-F(X_s)\big)-\frac{1}{2}\nu^*Q(X_s)\nu\Big)ds=\frac{1}{2}\int_0^T\big\|\dot{X}_s-F(X_s)\big\|^2_{Q^+(X_s)}.$$

Thus $J'_T=J_T(X)$. □

6. Proof of Theorem 2.1

In this section, we verify (2.4). By (1.8) we have

$$\varepsilon^{1/2-\kappa}W^\varepsilon_t := \int_0^t g(\xi_{s/\varepsilon})\,ds=\varepsilon^{1-\kappa}[\psi(z_0)-\psi(\xi_{t/\varepsilon})]+\varepsilon^{1/2-\kappa}M^\varepsilon_t.$$

The random process $\psi(\xi_{t/\varepsilon})$ is a continuous semimartingale, so that the Itô integral $V^\varepsilon_t=\varepsilon^{1-\kappa}\int_0^t G(X^\varepsilon_s)\,d\psi(\xi_{s/\varepsilon})$ is semimartingale as well. With V^ε_t (1.12) is transformed to

$$X^\varepsilon_t = x_0+\int_0^t F(X^\varepsilon_s)\,ds+\varepsilon^{1/2-\kappa}\int_0^t G(X^\varepsilon_s)\,dM^\varepsilon_s - V^\varepsilon_t. \quad (6.1)$$

Lemma 6.1. *For any $T>0$ there is $K>0$, so that $\rho_T(V^\varepsilon,0)\le KT\varepsilon^{1-2\kappa}$.*

Proof. It suffices to give the proof for any entries $V^\varepsilon_t(i)$ of V^ε_t. Write

$$V^\varepsilon_t(i)=\varepsilon^{1-\kappa}\int_0^t\sum_j G_{ij}(X^\varepsilon_s)\,du_j(\xi_{s/\varepsilon}),\quad i=1,\dots,d.$$

By the Itô formula, applied to $\varepsilon^{1-\kappa}\sum_j G_{ij}(X^\varepsilon_t)u_j(\xi_{t/\varepsilon})$, we get

$$\begin{aligned} V^\varepsilon_t(i) &= -\varepsilon^{1-\kappa}\sum_j G_{ij}(X_0)u_j(z_0)+\varepsilon^{1-\kappa}\sum_j G_{ij}(X^\varepsilon_t)u_j(\xi_{t/\varepsilon}) \\ &\quad -\varepsilon^{1-\kappa}\int_0^t\sum_{jp}\frac{\partial G_{i,j}(X^\varepsilon_s)}{\partial x_k}F_p(X^\varepsilon_s)ds \\ &\quad -\varepsilon^{1-2\kappa}\int_0^t\sum_{ijpq}\frac{\partial G_{i,j}(X^\varepsilon_s)}{\partial x_k}G_{pq}(X^\varepsilon_s)g_q(\xi_{s/\varepsilon})ds \end{aligned}$$

and, owing to the boundedness of $G_{ij}, u_i, F_p, \partial G_{pq}/\partial x_k$, the result is done. $\square$

Notice that $\rho_T(X^\varepsilon,\widehat{X}^\varepsilon)=\sup_{t\le T}|\Delta^\varepsilon_t|$, where $\Delta^\varepsilon_t=X^\varepsilon_t-\widehat{X}^\varepsilon_t$ and $|\Delta^\varepsilon_t|=\sum_{i=1}^\ell\Delta^\varepsilon_t(i)$. By (6.1) and (1.13), it follows

$$\Delta^\varepsilon_t=\int_0^t\big(F(X^\varepsilon_s)-F(\widehat{X}^\varepsilon_s)\big)\,ds+\varepsilon^{1/2-\kappa}\int_0^t\big(G(X^\varepsilon_t)-G(\widehat{X}^\varepsilon_t)\big)\,dM^\varepsilon_s-V^\varepsilon_t. \tag{6.2}$$

Set $\mathfrak{f}(s)=(F(X^\varepsilon_s)-F(\widehat{X}^\varepsilon_s))/|\Delta^\varepsilon_s|$, $\mathfrak{g}(s)=(G(X^\varepsilon_s)-G(\widehat{X}^\varepsilon_s))/|\Delta^\varepsilon_s|$. By ($\mathrm{A}_F$) and ($\mathrm{A}_G$), the vector $\mathfrak{f}(s)$ and matrix $\mathfrak{g}(s)$ are well defined and have bounded entries. Let us rewrite (6.2) in a coordinate form

$$\Delta^\varepsilon_t(i)=\int_0^t|\Delta^\varepsilon_s|\mathfrak{f}_i(s)\,ds+\varepsilon^{1/2-\kappa}\int_0^t|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)-V^\varepsilon_t(i),$$

where $\mathfrak{f}_i(s)$ and $V^\varepsilon_t(i)$ are coordinates of $\mathfrak{f}(s)$ and V^ε_t respectively and

$$m^\varepsilon_t(i)=\int_0^t\sum_{jpk}\mathfrak{g}_{ij}(s)\Psi_{jp}(\xi_{s/\varepsilon})\sigma_{pq}(\xi_{s/\varepsilon})\,d\big(\sqrt{\varepsilon}B_{s/\varepsilon}(k)\big)$$

(here $\mathfrak{g}_{ij}(s)$ are entries of $\mathfrak{g}(s)$). Henceforth r is a positive generic constant. Notice that $m^\varepsilon_t(i)$ is a continuous martingale with $d\big\langle m^\varepsilon(i)\big\rangle_t\le r\,dt$. By ($\mathrm{A}_F$)) $|\mathfrak{f}_i(s)|\le r$ and by Lemma 6.1 $|V^\varepsilon_t(i)|\le r\varepsilon^{1-2\kappa}$. Hence, for $t'\le t$

$$|\Delta^\varepsilon_{t'}|\le r\Big[\int_0^{t'}|\Delta^\varepsilon_s|\,ds+\sum_{i=1}^\ell\varepsilon^{1/2-\kappa}\sup_{t'\le t}\Big|\int_0^{t'}|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)\Big|+\varepsilon^{1-2\kappa}\Big]$$

and by the Bellman – Gronwall inequality we have

$$\sup_{t'\le t}|\Delta^\varepsilon_{t'}|\le r\Big[\sum_{i=1}^\ell\varepsilon^{1/2-\kappa}\sup_{t'\le t}\Big|\int_0^{t'}|\Delta^\varepsilon_s|dm^\varepsilon_t(i)\Big|+\varepsilon^{1-2\kappa}\Big].$$

Since $\sup_{t'\le t}|\Delta^\varepsilon_{t'}|$ is continuous in t with $|\Delta^\varepsilon_0| = 0$, we may assume that $\sup_{t'\le T}|\Delta^\varepsilon_{t'}|$ is bounded (otherwise a localization procedure is applied). Assuming that ε is small enough so that $p=\varepsilon^{-(1/2-\kappa)} > 1$, by the Hölder inequality it holds

$$\mathsf{E}\sup_{t'\le t}|\Delta^\varepsilon_{t'}|^p \le r^p\Big[p^{-p}\sum_{i=1}^{\ell}\mathsf{E}\sup_{t'\le t}\Big|\int_0^{t'}|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)\Big|^p + p^{-2p}\Big]. \tag{6.3}$$

Further, by the Doob inequality

$$\mathsf{E}\sup_{t'\le t}\Big|\int_0^{t'}|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)\Big|^p \le \Big(\frac{p}{p-1}\Big)^p\,\mathsf{E}\Big|\int_0^{t}|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)\Big|^p.$$

Moreover, similarly to the proof of Lemma 4.12 (Chapter 4, Section 4.3 in [26]) it is possible to establish

$$\mathsf{E}\Big|\int_0^{t}|\Delta^\varepsilon_s|\,dm^\varepsilon_s(i)\Big|^p \le r^{p/2}p^2\int_0^t \mathsf{E}\,|\Delta^\varepsilon_s|^p\,ds \le r^{p/2}p^2\int_0^t \mathsf{E}\sup_{s'\le s}|\Delta^\varepsilon_{s'}|^p\,ds.$$

Thus, for $U_t = \mathsf{E}\sup_{t'\le t}|\Delta^\varepsilon_{t'}|^p$ we get the integral inequality

$$U_t \le r^p p^{-2p} + r^p p^{2-p}\int_0^t U_s\,ds.$$

Now, by the Bellman – Gronwall inequality we have $U_T \le r^p p^{-2p}\exp\big(r^p p^{2-p}T\big)$ and, by the Chebyshev inequality

$$\mathsf{P}\big(\sup_{t\le T}|\Delta^\varepsilon_t| \ge \eta\big) \le \frac{r^p p^{-2p}\exp\big(r^p p^{2-p}T\big)}{\eta^p}.$$

Taking now into consideration that $\varepsilon^{1-2\kappa} = 1/p$ and $\lim_{p\to\infty} r^p p^{1-p} = 0$, we obtain

$$\varepsilon^{1-2\kappa}\log\mathsf{P}\Big(\sup_{t\le T}|\Delta^\varepsilon_t|\ge\eta\Big) \le \log r - 2\log p + \frac{r^p p}{p^p} - \eta \to -\infty, \quad \varepsilon\to\infty.$$

□

A. Auxiliary results for exponential tightness

Let S^ε_t and N^ε_t be continuous semimartingale and martingale respectively with paths in $\mathbf{C}_{[0,T]}(\mathbf{R}^\ell)$. Set

$$Y^\varepsilon_t = S^\varepsilon_t + \varepsilon^{1/2-\kappa}N^\varepsilon_t.$$

Denote by $S^\varepsilon_t(i)$, $N^\varepsilon_t(i)$ the entries of S^ε_t, N^ε_t and write $Y^\varepsilon_t(i) = S^\varepsilon_t(i) + \varepsilon^{1/2-\kappa}N^\varepsilon_t(i)$. Define $|S^\varepsilon_t| = \sum_{i=1}^{\ell}|S^\varepsilon_t(i)|$ and similarly $|N^\varepsilon_t|$, $|Y^\varepsilon_t|$.

Lemma A.1. *Assume for some nonnegative* c_1, c_2, c_3

$$|S^\varepsilon_t| \le c_1 + c_2 \int_0^t \big(1 + |Y^\varepsilon_s|\big)\, ds,$$

$$d\langle N^\varepsilon(i)\rangle_t * \le c_3\, dt, \quad i = 1\ldots,\ell.$$

Then for any $T > 0$

$$\lim_{C\to\infty} \overline{\lim_{\varepsilon\to 0}}\, \varepsilon^{1-2\kappa} \log \mathsf{P}\Big(\sup_{t\le T} |Y^\varepsilon_t| > C\Big) = -\infty.$$

Remark A.1. The assumptions of Lemma A.1 are satisfied for Y^ε_t being X^ε_t (6.1), $\widehat{X}^\varepsilon_t$ (1.13) and $\widehat{X}^{\varepsilon,\gamma}_t$ (5.4).

Proof. Due to the first assumption

$$|Y^\varepsilon_t| \le c_1 + c_2 \int_0^t \big(1 + |Y^\varepsilon_s|\big) ds + \varepsilon^{1/2-\kappa} \sup_{t\le T} |N^\varepsilon_t|, \quad t \le T.$$

Hence, by the Bellman – Gronwall inequality, for any $t \le T$

$$|Y^\varepsilon_t| \le \exp\{c_2 T\}\Big(c_1 + c_2 T + \varepsilon^{1/2-\kappa} \sup_{t\le T} |N^\varepsilon_t|\Big),$$

that is the same upper bound is valid for $\sup_{t\le T} |Y^\varepsilon_t|$ as well. The latter proves the statement of the lemma, if

$$\lim_{C\to\infty} \overline{\lim_{\varepsilon\to 0}}\, \varepsilon^{1-2\kappa} \log \mathsf{P}\Big(\sup_{t\le T} |N^\varepsilon_t(i)| > C\Big) = -\infty \tag{A.1}$$

for any $i = 1, \ldots, \ell$.

Introduce Markov times

$$\tau^\varepsilon_\pm = \inf\Big\{t : N^\varepsilon_t \ge +\frac{C}{\varepsilon^{1/2-\kappa}} \ \Big(\le -\frac{C}{\varepsilon^{1/2-\kappa}}\Big)\Big\},$$

where $\inf\{\varnothing\} = \infty$. Owing to $\{\varepsilon^{1/2-\kappa} \sup_{t\le\Theta} |M_t| > C\} \subseteq \{\tau^\varepsilon_+ \le T\} \cup \{\tau^\varepsilon_- \le T\}$, (A.1) is provided by

$$\lim_{C\to\infty} \overline{\lim_{\varepsilon\to 0}}\, \varepsilon^{1-2\kappa} \log \mathsf{P}\big(\tau^\varepsilon_\pm \le T\big) = -\infty. \tag{A.2}$$

So, we shall verify (A.2). With $\phi \in \mathbf{R}$, let us define

$$Z_t(\phi) = \exp\Big(\phi N^\varepsilon_t(i) - \frac{\phi^2}{2}\big\langle N^\varepsilon(i)\big\rangle_t\Big).$$

By Problem 1.4.4 in [27] $Z_t(\phi)$ is a supermartingale, so that $\mathsf{E}\, Z_{\tau^\varepsilon_\pm \wedge T}(\phi) \le 1$. Since by the second assumption $\big\langle N^\varepsilon(i)\big\rangle_T \le c_3 T$, for a positive ϕ we have

$$1 \ge \mathsf{E}\, I(\tau^\varepsilon_+ \le T) Z_{\tau^\varepsilon_+ \wedge T}(\phi) \ge \mathsf{P}\big(\tau^\varepsilon_+ \le T\big) \exp\Big(\frac{\phi C}{\varepsilon^{1/2-\kappa}} - \frac{\phi^2 c_3 T}{2}\Big)$$

while the choice $\phi = C/c_3 T\varepsilon^{1/2-\kappa}$ provides

$$\mathsf{P}\left(\tau_+^\varepsilon \le T\right) \le \exp\Big(- \frac{C^2}{2c_3 T\varepsilon^{1-2\kappa}}\Big).$$

Then $\varepsilon^{1-2\kappa} \log \mathsf{P}\left(\tau_+^\varepsilon \le T\right) \le -C^2/2c_2TK$ and (A.2) holds for "+".

For "−", the proof is similar. □

Corollary A.1. *Assume* $S_t^\varepsilon \equiv 0$ *and* $d\langle N^\varepsilon(i)\rangle_t \le c_3\varepsilon^{c_4}\,dt,\ i = 1\dots,\ell,\ c_4 > 0.$ *Then for any* $T > 0$, $\eta > 0$

$$\lim_{\varepsilon\to 0} \varepsilon^{1-2\kappa} \log \mathsf{P}\Big(\sup_{t\le T} |Y_t^\varepsilon| > \eta\Big) = -\infty.$$

Acknowledgement

We would like to thank anonymous referee for helpful comments essentially improving the paper.

References

[1] A. Albert (1972) *Regression and the Moore – Penrose Pseudoinverse.* Academic Press, New York and London.

[2] P. Baldi (1991) Large deviations for diffusion processes with homogenization and applications. *Ann. Prob.* **19**, 509–524.

[3] U. Bayer and M.I. Freidlin (1977) Theorems on large deviations and stability under random perturbations. *Doclady AN SSSR* **235** (2), 253–256.

[4] A.A. Borovkov and A.A. Mogulski (1978) Probabilities of large deviations in topological vector space. I. *Sib. Math. J.* **19**, 697–709.

[5] A.A. Borovkov and A.A. Mogulski (1980) Probabilities of large deviations in topological vector space. II. *Sib. Math. J.* **19**, 12–26.

[6] R.N. Bhattacharya (1992) On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wharsch. verw. Geb.* **60**, 185–201.

[7] C.-S. Chang, D.D. Yao and T. Zajic Moderate deviations for queues with long-range dependent input. *Stochastic Networks*, New York, 275–298 (1995), *Lect. Notes Statist.* **117** Springer, New York (1996).

[8] X. Chen (1991) The moderate deviations of independent random vectors in a Banach space. *Chinese J. Appl. Probab. Statist.* **7**, 24–33.

[9] A. Dembo (1996) Moderate deviations for martingales with bounded jumps. *Elect. Comm. in Probab.* **1**, 11–17.

[10] A. Dembo and T. Zajic (1997) Uniform large and moderate deviations for functional empirical processes. *Stoch. Process. Appl.* **67**, 195–211.

[11] A. Dembo and O. Zeitouni (1998) Moderate deviations for iterates of expanding maps. In: *Liptser Festschrift*, London, 1–11.

[12] S.N. Ethier and T.G. Kurtz. (1986) *Markov Processes. Characterization and Convergence.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York.

[13] M.I. Freidlin (1978) Averaging principle and theorem on large deviations. *Uspekhi Mat. Nauk* **33**, 107–160.

[14] M.I. Freidlin and A.D. Wentzell (1984) *Random Perturbations of Dynamical Systems.* New York, Springer.

[15] J. Jacod and A.N. Shiryaev (1987) *Limit Theorems for Stochastic Processes.* Springer-Verlag, New York.

[16] A. Guillin (2001) Moderate deviations of inhomogeneous functionals of Markov processes and application to averaging. *Stoch. Process. Appl.* **92**, 287–313.

[17] A. Guillin (2000) Averaging principle of SDE with small diffusion: moderate deviation. To appear in *Ann. Prob.*

[18] R.Z. Khasminskii (1980) *Stochastic Stability of Differential Equations.* Sijthoff & Noordhoff.

[19] N.V. Krylov *Controlled Diffusion Processes.* Springer (1980), Moscow (1977).

[20] O. Ladyzenskaja, V. Solonnikov and N. Ural'ceva (1968) Linear and quasilinear equations of parabolic type. *Translation Monographs* **23**, AMS, Providence.

[21] M. Ledoux (1992) Sur les deviations modérées des sommes de variables aléatoires vectorielles independantes de même loi. *Ann. Inst. H. Poincaré, Probabilitès et Statististiques* **28**, 267–280.

[22] R. Liptser (1996) Large deviations for two scaled diffusions. *Theory Probability and Related Fields* **106**, 71–104.

[23] R. Liptser (1997) Moderate deviations for dynamic model governed by stationary process. In: *Festschrift A. Skorokhod.*

[24] R. Liptser (2000) Moderate deviation for dynamic model governed by stationary process. In: *Skorokhod's Ideas in Probability Theory*, National Academy of Science of Ukraine Institute of Mathematics, Kyiv, 256–270.

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

[26] R.S. Liptser and A.N. Shiryayev (2000) *Statistics of Random Processes.* I. Springer Verlag.

[27] R.S. Liptser and A.N. Shiryayev (1989) *Theory of Martingales.* Kluwer Acad. Publ.

[28] R.S. Liptser and V. Spokoiny (1999) Moderate Deviations type evaluation for integral functionals of diffusion process. *Electronic Journal of Probability* **4**, 1–25. `http://www.math.washington.edu/~ejpecp/EjpVol4/paper17.abs. html.`

[29] C.C. Papanicolaou, D.W. Stroock and S.R.S. Varadhan (1977) Martingale approach to some limit theorems. In: *Conference on Statistical Mechanics, Dynamical Systems and Turbulence*, M. Reed (ed.), Duke Univ. Math. Series **3**.

[30] E. Pardoux and A.Yu Veretennikov (2001) On Poisson equation and diffusion approximation. 1. To appear in *Ann. Prob.*

[31] A.A. Pukhalskii (1991) On functional principle of large deviations. In: *New Trends in Probability and Statistics*, V. Sazonov, T. Shervashidze (eds.), Vilnius, Lithuania, VSP/Mokslas, 198–218.

[32] A. Puhalskii (1994) Large deviations of semimartingales via convergence of the predictable characteristics. *Stoch. Stoch. Rep.* **49**, 27–85.

[33] A. Puhalskii (1999) Large deviations of semimartingales: a maxingale problem approach. II. Uniqueness for the maxingale problem. Applications. *Stoch. Stoch. Rep.* **68**, 65–143.

[34] A. Puhalskii (2001) *Large Deviations and Idempotent Probability.* Chapman & Hall/CRC Press.

[35] A.A. Pukhalskii (1999) Moderate deviations for queues in critical loading. *Queueing Systems* **31**, 359–392.

[36] B. Remillard and D.A. Dawson (1989) Laws of the iterated logarithm and large deviations for a class of diffusion processes. *Canad. J. Statist.* **17**, 349–376.

[37] S.R.S. Varadhan (1984) *Large Deviations and Applications.* Philadelphia, SIAM.

[38] A.Yu. Veretennikov (1999) On polynomial mixing and convergence rate for stochastic difference and differential equations. *Teor. Veroyatn. Primen.* **44** (2), 312–327. English version: Preprint N 393 (1998), WIAS, Berlin.

[39] L. Wu (1995) Moderate deviations of dependent random variables related to CLT. *Ann. Prob.* **23**, 420–445.

[40] L. Wu (2001) Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stoch. Process. Appl.* **91**, 205–238.