Rev. Mat. Iberoam. **29** (2013), no. 1, 25–52
DOI 10.4171/RMI/711

Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients

Xicheng Zhang

Abstract. In this article we prove existence and uniqueness for degenerate stochastic differential equations with Sobolev (possibly singular) drift and diffusion coefficients in a generalized sense. In particular, our result covers the classical DiPerna–Lions flows and we also obtain well-posedness for degenerate Fokker–Planck equations with irregular coefficients. Moreover, a large deviation principle of Freidlin–Wenzell type for this type of SDEs is established.

1. Introduction

The celebrated DiPerna–Lions theory [10] says that if a vector field $b \in W^{1,1}_{\rm loc}(\mathbb{R}^d)$ has bounded divergence and $\frac{b(x)}{1+|x|} \in L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$, then there exists a unique regular Lagrangian flow for the ordinary differential equation (ODE) in $\mathbb{R}^d$:

$$\mathrm{d}X_t(x) = b(X_t(x))\mathrm{d}t, \quad X_0(x) = x. \tag{1.1}$$

This theory was later extended to the case of BV vector field by Ambrosio [1]. Their methods were based on the connection between ODEs and transport or continuity equations. Recently, Crippa and De Lellis [9] developed a more direct argument to treat this problem by using the Hardy–Littlewood maximal functions for b assumed to be in $W^{1,p}_{\rm loc}(\mathbb{R}^d)$ for some $p > 1$. Moreover, Cipriano and Cruzeiro [8] studied the non-smooth flows associated to (1.1) when the exponential of the divergence of b satisfies some $L^p(\mathbb{R}^d, \mu)$-type hypothesis, where μ is the standard Gaussian measure on $\mathbb{R}^d$. Such a theory has also been extended to the classical Wiener space by Ambrosio and Figalli [2] (see also Fang and Luo [12]).

We now turn to the following Itô stochastic differential equation (SDE) in $\mathbb{R}^d$:

$$\mathrm{d}X_t(x) = b(X_t(x))\mathrm{d}t + \sigma(X_t(x))\mathrm{d}W_t, \quad X_0(x) = x. \tag{1.2}$$

Mathematics Subject Classification (2010): Primary 60H10; Secondary 35Q84.
Keywords: DiPerna–Lions theory, generalized stochastic flow, Hardy–Littlewood maximal function, large deviation.

Here $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$ are measurable functions, $(W_t)_{t\in[0,1]}$ is an m-dimensional standard Brownian motion on the classical Wiener space $(\Omega, \mathscr{F}, P)$, i.e., Ω is the space of all $\mathbb{R}^m$-valued continuous functions on $[0,1]$, $\mathscr{F}$ is the associated Borel σ-field, and P is the standard Wiener measure. For a generic point $\omega \in \Omega$, $W_t(\omega) = \omega_t$ is the coordinate process. Let $\mathscr{F}_t$ be the natural Brownian filtration generated by $\{W_s, s \leqslant t\}$.

In [14], Figalli has proved the well-posedness of martingale solutions for the SDE (1.2) with Sobolev coefficients by studying the associated Fokker–Planck equations. His strategy is similar to [1]. Recently, in [28] we gave a direct construction of the almost everywhere stochastic flow of (1.2) by using the same argument as in Crippa and De Lellis [9]. Furthermore, through linearizing Brownian motion, we also proved ([23]) a classical limit theorem that the solutions of ODE (1.1) converge, in a generalized sense, to the solutions of a Stratonovich SDE. In the papers [9], [28], and [23], the vector field b needs to be in $W^{1,q}_{\mathrm{loc}}(\mathbb{R}^d)$ for some $q > 1$. In the case of nondegenerate and regular diffusion coefficients, there have been numerous results about the existence and uniqueness of strong solutions to SDE (1.2) with singular drift b (cf. [30], [15], [18], [27], etc.).

The present work is a continuation of [28] and [23], and the main aims of this paper are twofold: First, we try to relax the assumptions on the diffusion and drift coefficients so that the diffusion coefficients can be discontinuous for Stratonovich SDEs, b can be in $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$, and the divergence of b can be polynomial growth. Secondly, we prove a Freidlin–Wentzell large deviation principle for SDEs with Sobolev coefficients.

In order to obtain a Freidlin–Wentzell large deviation estimate for the SDE (1.2) with discontinuous coefficients, we shall employ the weak convergence method of Dupuis and Ellis [11]. This method has proved to be very effective for various stochastic systems (cf. [4], [6], [22], etc.), where the key point is to use the variational representation of certain exponential Brownian functionals (cf. [3] and [29]) to prove an equivalent Laplace principle.

This paper is organized as follows: In Section 2, we state our main results. In Section 3, some preliminaries are given. In Section 4, the well-posedness theorems are proven. In Section 5, we shall prove a large deviation principle for the SDE (1.2).

2. Statement of main results

Let $\mathscr{M}(\mathbb{R}^d)$ be the total of all locally finite Borel measures on $\mathbb{R}^d$. For $p \geqslant 1$ and $\mu \in \mathscr{M}(\mathbb{R}^d)$, let $L^p_\mu = L^p_\mu(\mathbb{R}^d)$ be the usual L^p-space over $(\mathbb{R}^d, \mu)$ and $W^{p,k}_{\mathrm{loc}}(\mathbb{R}^d)$ the usual local Sobolev space. If $\mu = \mathscr{L}(\mathrm{d}x)$ is the Lebesgue measure, we simply write $L^p_\mu =: L^p$. For $R > 0$, by B_R we denote the ball in $\mathbb{R}^d$ with center zero and radius R.

First of all, we introduce the following general notion about μ-almost everywhere stochastic flow of SDE (1.2) (cf. [19], [28]):

Definition 2.1. Let $X_t(\omega, x)$ be a $\mathbb{R}^d$-valued measurable stochastic field on $[0,1]\times\Omega\times\mathbb{R}^d$. For $\mu\in\mathscr{M}(\mathbb{R}^d)$, we say X a *μ-almost everywhere stochastic flow* of the SDE (1.2) corresponding to (b,σ) if

(A) for some $p\geqslant 1$, there exists a constant $K_p>0$ such that for any nonnegative measurable function $\varphi\in L^p_\mu(\mathbb{R}^d)$,

$$\sup_{t\in[0,1]}\mathbb{E}\int_{\mathbb{R}^d}\varphi(X_t(x))\mu(\mathrm{d}x)\leqslant K_p\|\varphi\|_{L^p_\mu};\tag{2.1}$$

(B) for μ-almost all $x\in\mathbb{R}^d$, $t\mapsto X_t(x)$ is a continuous $(\mathscr{F}_t)$-adapted process satisfying that

$$\int_0^1|b(X_s(x))|\mathrm{d}s+\int_0^1|\sigma(X_s(x))|^2\mathrm{d}s<+\infty,\quad P-\text{a.s.},\quad\text{and}$$

$$X_t(x)=x+\int_0^t b(X_s(x))\mathrm{d}s+\int_0^t\sigma(X_s(x))\mathrm{d}W_s,\quad\forall t\in[0,1].$$

We first consider the Stratonovich SDE

$$\mathrm{d}X_t(x)=b(X_t(x))\mathrm{d}t+\sigma(X_t(x))\circ\mathrm{d}W_t,\quad X_0(x)=x,$$

or its equivalent Itô form:

$$\mathrm{d}X_t(x)=\left[b+\tfrac12\sigma^{jl}\partial_j\sigma^{\cdot l}\right](X_t(x))\,\mathrm{d}t+\sigma(X_t(x))\,\mathrm{d}W_t,\quad X_0(x)=x.$$

Here and below, we use the conventions that indices repeated in a product are summed automatically, and all derivatives and divergence are taken in the distributional sense. By definition, $\operatorname{div}\sigma^{\cdot l}:=\partial_i\sigma^{il}$, $l=1,\dots,m$.

The following result extends Theorem 2.6 in [28] to the Stratonovich SDE.

Theorem 2.2. *Assume that for some $r\in[0,+\infty)$,*

$$\frac{|b|+|\nabla\sigma|}{1+|x|},|\sigma|\in L^\infty(B_r^c),\quad b\in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d),\quad\sigma\in W^{2,2}_{\mathrm{loc}}(\mathbb{R}^d),\tag{2.2}$$

and for some $\varepsilon\in(0,1)$,

$$[\operatorname{div}b]^-,\quad|\operatorname{div}\sigma|,\quad\sup_{|z|\leqslant\varepsilon}|\sigma(\cdot-z)|\cdot|\nabla\operatorname{div}\sigma|\in L^\infty(\mathbb{R}^d).\tag{2.3}$$

Then there exists a unique $\mathscr{L}$-almost everywhere stochastic flow $X_t(x)$ (in the sense of Definition 2.1) corresponding to (b_σ,σ) with $p=1$ in (2.1), where $b_\sigma=b+\frac12\sigma^{jl}\partial_j\sigma^{\cdot l}$.

Remark 2.3. If $\operatorname{div}\sigma=\operatorname{div}b=0$, then from the proof below, one can see that

$$\int_{\mathbb{R}^d}\varphi(X_t(x))\,\mathrm{d}x=\int_{\mathbb{R}^d}\varphi(x)\,\mathrm{d}x\quad\text{a.s.},\ \forall t\in[0,1],$$

which means that the stochastic flow $x\mapsto X_t(x)$ is incompressible. In this case, b and σ in Theorem 2.2 only need to satisfy (2.2) and so are allowed to be singular in a finite ball. If σ vanishes, our result covers the classical DiPerna–Lions flow.

Our next aim is to relax the assumption $[\mathrm{div}\, b]^- \in L^\infty(\mathbb{R}^d)$ so that $[\mathrm{div}\, b]^-$ can have polynomial growth. We shall prove:

Theorem 2.4. *Assume that, for some $q > 1$,*

$$|\nabla b|, |\nabla \sigma|^2 \in L^q_{\mathrm{loc}}(\mathbb{R}^d), \quad \frac{|b| + |\sigma|}{1 + |x|} \in L^\infty(\mathbb{R}^d), \tag{2.4}$$

and there exist functions $\lambda \in C^2(\mathbb{R}^d)$ and $\gamma_1, \gamma_2, \gamma_3$ satisfying that for all small y in B_ε and all $x \in \mathbb{R}^d$,

$$\lambda(x) \leqslant \gamma_1(x - y), \quad |\nabla \lambda(x)| \leqslant \gamma_2(x - y), \quad |\nabla^2 \lambda(x)| \leqslant \gamma_3(x - y), \tag{2.5}$$

such that for all $p \geqslant 1$,

$$\int_{\mathbb{R}^d} \exp\Big\{p\Big([\mathrm{div}\, b]^- + |b|\gamma_2 + |\sigma|^2(\gamma_2^2 + \gamma_3) + |\nabla\sigma|^2\Big)(x) + \gamma_1(x)\Big\}\, \mathrm{d}x < +\infty. \tag{2.6}$$

Let $\mu(\mathrm{d}x) = \mathrm{e}^{\lambda(x)}\mathrm{d}x$. Then there exists a unique μ-almost everywhere stochastic flow $X_t(x)$ in the sense of Definition 2.1 corresponding to (b, σ) for any $p > 1$ in (2.1).

Remark 2.5. In this theorem, assumptions (2.5) and (2.6) are a little bit complicated. We now explain them by introducing two examples.

(1) Let $\lambda(x) = -\alpha \log(1 + |x|^2)$ for some $\alpha > \frac{d}{2}$. For all $|y| \leqslant \frac{1}{2}$ and $x \in \mathbb{R}^d$, we have

$$\begin{aligned}\lambda(x) &\leqslant -\alpha \log\big(1 + (|x - y| - |y|)^2\big) \leqslant -\alpha \log\big(1 + \tfrac{1}{2}|x - y|^2 - |y|^2\big) \\ &\leqslant -\alpha \log\big(\tfrac{3}{4} + \tfrac{1}{2}|x - y|^2\big) \leqslant -\alpha \log\big(1 + |x - y|^2\big) + \alpha \log 2 =: \gamma_1(x - y),\end{aligned}$$

and

$$\begin{aligned}|\nabla \lambda(x)| &\leqslant \frac{2\alpha|x|}{1 + |x|^2} \leqslant \frac{4\alpha}{1 + |x|} \leqslant \frac{8\alpha}{1 + |x - y|} =: \gamma_2(x - y), \\ |\nabla^2 \lambda(x)| &\leqslant \frac{6\alpha}{1 + |x|^2} \leqslant \frac{6\alpha}{1 + \frac{1}{2}|x - y|^2 - |y|^2} \leqslant \frac{12\alpha}{1 + |x - y|^2} =: \gamma_3(x - y).\end{aligned}$$

In this case, if b and σ have linear growth, then condition (2.6) reduces to

$$\int_{\mathbb{R}^d} \frac{\exp\{p([\mathrm{div}\, b]^- + |\nabla\sigma|^2)(x)\}}{(1 + |x|^2)^\alpha}\, \mathrm{d}x < +\infty, \quad \forall p \geqslant 1.$$

(2) Let $\lambda(x) = -|x|^{2\alpha}$ for some $\alpha \geqslant 1$. For all $|y| \leqslant \frac{1}{2}$ and $x \in \mathbb{R}^d$, we have

$$\lambda(x) \leqslant -(|x - y| - |y|)^{2\alpha} \leqslant -(|x - y| - \tfrac{1}{2})^{2\alpha} \leqslant C_\alpha - \tfrac{1}{2}|x - y|^{2\alpha} =: \gamma_1(x - y),$$

and

$$\begin{aligned}|\nabla \lambda(x)| &\leqslant 2\alpha|x|^{2\alpha - 1} \leqslant 2\alpha(|x - y| + \tfrac{1}{2})^{2\alpha - 1} =: \gamma_2(x - y), \\ |\nabla^2 \lambda(x)| &\leqslant 4\alpha^2|x|^{2\alpha - 2} \leqslant 4\alpha^2(|x - y| + \tfrac{1}{2})^{2\alpha - 2} =: \gamma_3(x - y).\end{aligned}$$

In this case, if for some $\beta \in [0,1)$,

$$\frac{|b(x)|}{1+|x|^\beta}, \ \frac{|\sigma(x)|}{(1+|x|)^{\beta-\alpha}} \in L^\infty(\mathbb{R}^d),$$

then by Young's inequality, condition (2.6) reduces to

$$\int_{\mathbb{R}^d} \exp\left\{p([\mathrm{div}\, b]^- + |\nabla\sigma|^2)(x) - \tfrac{1}{4}|x|^{2\alpha}\right\} \mathrm{d}x < +\infty, \quad \forall p \geqslant 1.$$

Remark 2.6. Recently, Fang–Luo–Thalmaier [13] also studied stochastic differential equations in the Gaussian space with Sobolev coefficients. However, our result is more general than Theorem 1.3 in [13]. In particular, from the previous example (1), one can see that condition 1.3 in Theorem 1.2 of [13] is not necessary.

As an easy consequence of Theorem 2.4 and Theorem 1.1 in [24], we have:

Corollary 2.7. *Assume that b and σ are bounded measurable functions and for some $q > 1$,*

$$|\nabla b|, |\nabla\sigma|^2 \in L^q_{\mathrm{loc}}(\mathbb{R}^d),$$

and (2.6) holds. Then for any probability density function ϕ with

$$\int_{\mathbb{R}^d} \phi(x)^r \mathrm{e}^{(1-r)\lambda(x)}\mathrm{d}x < +\infty,$$

where $r > \frac{q}{q-1} =: p$, and $\lambda(x)$ is from Theorem 2.4, there exists a unique distribution solution to the Fokker–Planck equation

$$\partial_t u_t = -\,\mathrm{div}(bu_t) + \tfrac{1}{2}\partial^2_{ij}\big([\sigma^{il}\sigma^{jl}]u_t\big), \quad u_0 = \phi, \tag{2.7}$$

in the class

$$\mathcal{M}_p := \Big\{u_t \in L^p_{\mathrm{loc}}(\mathbb{R}^d) : u_t(x) \geqslant 0, \int_{\mathbb{R}^d} u_t(x)\,\mathrm{d}x = 1,$$
$$\sup_{t\in[0,1]} \int_{\mathbb{R}^d} u_t(x)^p \mathrm{e}^{(1-p)\lambda(x)}\,\mathrm{d}x < +\infty\Big\}.$$

Proof. Let X_0 be an $\mathscr{F}_0$-measurable random variable with distribution $\phi(x)\mathrm{d}x$. It is easy to see that $Y_t := X_t(X_0)$ solves the SDE:

$$Y_t = X_0 + \int_0^t b(Y_s)\mathrm{d}s + \int_0^t \sigma(Y_s)\mathrm{d}W_s.$$

Let $\mu(\mathrm{d}x) = \mathrm{e}^{\lambda(x)}\mathrm{d}x$. Now for any $\varphi \in C^\infty_c(\mathbb{R}^d)$, by Hölder's inequality, we have

$$\begin{aligned}
\mathbb{E}\varphi(Y_t) &= \mathbb{E}(\mathbb{E}\varphi(X_t(x))|x = X_0) = \int_{\mathbb{R}^d} \mathbb{E}\varphi(X_t(x))\phi(x)\,\mathrm{d}x \\
&\leqslant \Big(\int_{\mathbb{R}^d} |\mathbb{E}\varphi(X_t(x))|^{\frac{r}{r-1}}\mu(\mathrm{d}x)\Big)^{1-\frac{1}{r}}\Big(\int_{\mathbb{R}^d} \phi(x)\mathrm{e}^{-r\lambda(x)}\mu(\mathrm{d}x)\Big)^{\frac{1}{r}} \\
&\leqslant \Big(\mathbb{E}\int_{\mathbb{R}^d} |\varphi(X_t(x))|^{\frac{r}{r-1}}\mu(\mathrm{d}x)\Big)^{1-\frac{1}{r}}\Big(\int_{\mathbb{R}^d} \phi(x)^r\mathrm{e}^{(1-r)\lambda(x)}\mathrm{d}x\Big)^{\frac{1}{r}} \leqslant C_\phi\|\varphi\|_{L^q_\mu}.
\end{aligned}$$

Hence, there exists a $u \in \mathcal{M}_p$ such that for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $t \in [0,1]$,

$$\int_{\mathbb{R}^d} \varphi(x) u_t(x)\,\mathrm{d}x = \mathbb{E}\varphi(Y_t) \leqslant C_\phi \|\varphi\|_{L^q_\mu}.$$

By Itô's formula, it is easy to check that u is a distribution solution of (2.7). The uniqueness follows from Theorem 1.1 in [24]. □

Remark 2.8. In Proposition 5 in [20] of Le Bris and Lions, the well-posedness of equation (2.7) was shown in the following space:

$$\{u \in L^\infty(0,1;(L^1 \cap L^\infty)(\mathbb{R}^d)),\ \sigma^{\mathrm{t}}\nabla u \in L^2(0,1;L^2(\mathbb{R}^d))\}.$$

Moreover, the conditions on b and σ are different.

Next, we consider Freidlin–Wentzell's large deviation estimate for the SDE (1.2) in the situation of Theorem 2.4. For $\varepsilon \in (0,1)$, let $X_{\varepsilon,t}(x)$ solve the following SDE in the sense of Definition 2.1:

$$\mathrm{d}X_{\varepsilon,t}(x) = b(X_{\varepsilon,t}(x))\,\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_{\varepsilon,t}(x))\,\mathrm{d}W_t, \quad X_{\varepsilon,0}(x) = x. \tag{2.8}$$

We need to fix another weighted measure $\nu(\mathrm{d}x) = \mathrm{e}^{\rho(x)}\mathrm{d}x$ such that

$$\int_{\mathbb{R}^d} |x|^{2p}\nu(\mathrm{d}x) < +\infty, \quad \forall p \geqslant 1.$$

Thus we can consider equation (2.8) as an infinite-dimensional stochastic equation in the Banach space $L^{2p}_\nu(\mathbb{R}^d)$, $p \geqslant 1$:

$$X_{\varepsilon,t} = \mathrm{Id} + \int_0^t b(X_{\varepsilon,s})\,\mathrm{d}s + \sqrt{\varepsilon}\int_0^t \sigma(X_{\varepsilon,s})\,\mathrm{d}W_s.$$

The large deviation result is stated as follows:

Theorem 2.9. *Assume that b and σ satisfy the same assumptions as in Theorem 2.4. Then the family of random variables $(X_\varepsilon)_{\varepsilon\in(0,1)}$ taking values in the space $\mathbb{S} := L^{2p}_\nu(\mathbb{R}^d; C([0,1];\mathbb{R}^d))$, $p \geqslant 1$, satisfies the large deviation principle. More precisely, for any $B \in \mathcal{B}(\mathbb{S})$, we have*

$$-\inf_{f\in B^o} I(f) \leqslant \varliminf_{\varepsilon\to 0} \varepsilon \log P(X_\varepsilon \in B) \leqslant \varlimsup_{\varepsilon\to 0} \varepsilon \log P(X_\varepsilon \in B) \leqslant -\inf_{f\in \bar{B}} I(f),$$

where $I(f) := \frac{1}{2}\inf_{\{h\in L^2(0,1):\ f=X^h\}} \|h\|^2_{L^2}$, and X^h solves the equation

$$X_t = \mathrm{Id} + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,h_s\,\mathrm{d}s. \tag{2.9}$$

Here the closure and interior are taken in $\mathbb{S}$.

Remark 2.10. Although Corollary 2.7 and Theorem 2.9 are given under the assumptions of Theorem 2.4, similar results also hold for Stratonovich SDEs in the setting of Theorem 2.2.

3. Preliminaries

3.1. Two estimates on regular stochastic flows

In this subsection, we assume that $b, \sigma \in C_b^\infty(\mathbb{R}^d)$ are bounded and have bounded derivatives of all orders. In this case, it is well known that the SDE (1.2) defines a C^∞-diffeomorphism flow $X_t(x), x \in \mathbb{R}^d, t \in [0,1]$ (cf. [16], [17], [21]). We first recall the following well known result about the Jacobian determinant (for example, see Lemma 3.1 in [28]).

Lemma 3.1. *For any $t \in [0,1]$ and $x \in \mathbb{R}^d$, we have*

$$\begin{aligned}(3.1)\quad &\det(\nabla X_t(x)) \\ &= \exp\left\{\int_0^t \operatorname{div}\sigma(X_s(x))\mathrm{d}W_s + \int_0^t \left[\operatorname{div} b - \tfrac{1}{2}\partial_i\sigma^{jl}\partial_j\sigma^{il}\right](X_s(x))\,\mathrm{d}s\right\},\end{aligned}$$

and for any $p \geqslant 1$,

$$\begin{aligned}(3.2)\quad &\mathbb{E}\,|\det(\nabla X_t^{-1}(x))|^p \\ &\leqslant \exp\left\{tp\Big(\|[-\operatorname{div} b + \tfrac{1}{2}\partial_i\sigma^{jl}\partial_j\sigma^{il} + \sigma^{il}\partial_{ij}^2\sigma^{jl} + \tfrac{p}{2}|\operatorname{div}\sigma|^2]^+\|_\infty\Big)\right\}.\end{aligned}$$

Below, let λ be a C^2-function on $\mathbb{R}^d$ and define

$$\mu(\mathrm{d}x) := \mathrm{e}^{\lambda(x)}\mathrm{d}x.$$

We write

$$\mathcal{J}_t(\omega, x) := \frac{(X_t(\omega,\cdot))_\sharp\mu(\mathrm{d}x)}{\mu(\mathrm{d}x)}, \quad \mathcal{J}_t^-(\omega, x) := \frac{(X_t^{-1}(\omega,\cdot))_\sharp\mu(\mathrm{d}x)}{\mu(\mathrm{d}x)},$$

which means that for any nonnegative measurable function φ on $\mathbb{R}^d$,

$$(3.3)\qquad \int_{\mathbb{R}^d}\varphi(X_t(\omega,x))\mu(\mathrm{d}x) = \int_{\mathbb{R}^d}\varphi(x)\mathcal{J}_t(\omega,x)\mu(\mathrm{d}x),$$

$$(3.4)\qquad \int_{\mathbb{R}^d}\varphi(X_t^{-1}(\omega,x))\mu(\mathrm{d}x) = \int_{\mathbb{R}^d}\varphi(x)\mathcal{J}_t^-(\omega,x)\mu(\mathrm{d}x).$$

It is easy to see that for almost all ω and all $(t,x) \in [0,1]\times\mathbb{R}^d$,

$$(3.5)\qquad \mathcal{J}_t(\omega,x) = [\mathcal{J}_t^-(\omega, X_t^{-1}(\omega,x))]^{-1},$$

and by Itô's formula and (3.1),

$$\begin{aligned}(3.6)\quad \mathcal{J}_t^-(x) &= \mathrm{e}^{\lambda(X_t(x))-\lambda(x)}\det(\nabla X_t(x)) \\ &= \exp\left\{\int_0^t \Lambda_1^\sigma(X_s(x))\,\mathrm{d}W_s + \int_0^t \Lambda_2^{b,\sigma}(X_s(x))\,\mathrm{d}s\right\},\end{aligned}$$

where $\Lambda_1^\sigma(x) := \left[\operatorname{div}\sigma + \sigma^{i\cdot}\partial_i\lambda\right](x)$ and

$$\Lambda_2^{b,\sigma}(x) := \left[\operatorname{div} b + b^i\partial_i\lambda + \tfrac{1}{2}(\sigma^{il}\sigma^{jl}\partial_{ij}^2\lambda - \partial_i\sigma^{jl}\partial_j\sigma^{il})\right](x).$$

We now give an L^p estimate for $\mathcal{J}_t(x)$, that is crucial for Theorem 2.4 and is inspired by [7] and [8].

Lemma 3.2. *Assume that $\mu(\mathbb{R}^d)<+\infty$. Then for any $t\in[0,1]$ and $p>1$, we have*

$$\begin{aligned}(3.7)\qquad &\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}_t(x)|^p\mu(\mathrm{d}x)\\ &\quad\leqslant \mu(\mathbb{R}^d)^{\frac{p}{p+1}}\left(\sup_{t\in[0,1]}\int_{\mathbb{R}^d}\exp\Big\{tp^3|\Lambda^\sigma_1(x)|^2-tp^2\Lambda^{b,\sigma}_2(x)\Big\}\mu(\mathrm{d}x)\right)^{\frac{1}{p+1}}.\end{aligned}$$

Proof. By (3.4) and (3.5), we have

$$\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}_t(x)|^p\mu(\mathrm{d}x)=\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}^-_t(x)|^{1-p}\mu(\mathrm{d}x). \tag{3.8}$$

Since for any $\alpha\in\mathbb{R}$,

$$t\mapsto\exp\Big\{\alpha\int_0^t\Lambda^\sigma_1(X_s(x))\mathrm{d}W_s-\frac{\alpha^2}{2}\int_0^t|\Lambda^\sigma_1(X_s(x))|^2\mathrm{d}s\Big\}$$

is a continuous exponential martingale, by (3.6) and Hölder's inequality, for any $\alpha\in\mathbb{R}$ and $q>1$, we have

$$\mathbb{E}\,|\mathcal{J}^-_t(x)|^\alpha\leqslant\left(\mathbb{E}\exp\Big\{\int_0^t\Big[\tfrac{q^2\alpha^2}{2(q-1)}|\Lambda^\sigma_1(X_s(x))|^2+\alpha q\Lambda^{b,\sigma}_2(X_s(x))\Big]\mathrm{d}s\Big\}\right)^{\frac1q}.$$

For notational simplicity, we write

$$\phi_{\alpha,q}(x):=\tfrac{q^2\alpha^2}{2(q-1)}|\Lambda^\sigma_1(x)|^2+\alpha q\Lambda^{b,\sigma}_2(x).$$

By Jensen's inequality, we have

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}^-_t(x)|^{1-p}\mu(\mathrm{d}x)&\leqslant\int_{\mathbb{R}^d}\Big(\mathbb{E}\,\mathrm{e}^{\int_0^t\phi_{1-p,q}(X_s(x))\mathrm{d}s}\Big)^{\frac1q}\mu(\mathrm{d}x)\\ &\leqslant\int_{\mathbb{R}^d}\Big(\frac1t\int_0^t\mathbb{E}\,\mathrm{e}^{t\phi_{1-p,q}(X_s(x))}\mathrm{d}s\Big)^{\frac1q}\mu(\mathrm{d}x)\\ &\leqslant\mu(\mathbb{R}^d)^{1-\frac1q}\Big(\frac1t\int_0^t\mathbb{E}\int_{\mathbb{R}^d}\mathrm{e}^{t\phi_{1-p,q}(X_s(x))}\mu(\mathrm{d}x)\mathrm{d}s\Big)^{\frac1q}\\ &\overset{(3.3)}{=}\mu(\mathbb{R}^d)^{1-\frac1q}\Big(\frac1t\int_0^t\mathbb{E}\int_{\mathbb{R}^d}\mathrm{e}^{t\phi_{1-p,q}(x)}\mathcal{J}_s(x)\mu(\mathrm{d}x)\mathrm{d}s\Big)^{\frac1q}\\ &\leqslant\mu(\mathbb{R}^d)^{1-\frac1q}\Big(\int_{\mathbb{R}^d}\mathrm{e}^{\frac{pt}{p-1}\phi_{1-p,q}(x)}\mu(\mathrm{d}x)\Big)^{\frac{p-1}{pq}}\Big[\sup_{s\in[0,1]}\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}_s(x)|^p\mu(\mathrm{d}x)\Big]^{\frac{1}{pq}},\end{aligned}$$

which together with (3.8) implies that

$$\sup_{s\in[0,1]}\mathbb{E}\int_{\mathbb{R}^d}|\mathcal{J}_s(x)|^p\mu(\mathrm{d}x)\leqslant\mu(\mathbb{R}^d)^{\frac{p(q-1)}{pq-1}}\Big(\sup_{t\in[0,1]}\int_{\mathbb{R}^d}\mathrm{e}^{\frac{pt}{p-1}\phi_{1-p,q}(x)}\mu(\mathrm{d}x)\Big)^{\frac{p-1}{pq-1}}.$$

The proof is completed by simplifying the above expression with $q=p$. □

Remark 3.3. From (3.7), one sees that by letting $p \downarrow 1$,

$$\mathbb{E}\int_{\mathbb{R}^d} |\mathcal{J}_t(x)|\mu(\mathrm{d}x) \leqslant \mu(\mathbb{R}^d)^{\frac{1}{2}}\Big(\int_{\mathbb{R}^d} \exp\Big\{|\Lambda_1^\sigma(x)|^2 + |\Lambda_2^{b,\sigma}(x)|\Big\}\mu(\mathrm{d}x)\Big)^{\frac{1}{2}}.$$

3.2. Two lemmas related to (2.1)

The following lemma will play a crucial role for taking limits below (cf. [28], [23]).

Lemma 3.4. *Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and let $(X_n)_{n\in\mathbb{N}}$ be a family of random fields on $\Omega\times\mathbb{R}^d$. Suppose that X_n converges to X for $P\otimes\mu$-almost all (ω, x), and that for some $p \geqslant 1$, there is a constant $K_p > 0$ such that for any nonnegative measurable function $\varphi \in L^p_\mu(\mathbb{R}^d)$,*

$$\sup_n \mathbb{E}\int_{\mathbb{R}^d} \varphi(X_n(x))\mu(\mathrm{d}x) \leqslant K_p\|\varphi\|_{L^p_\mu}. \tag{3.9}$$

Then we have:

(i) *For any nonnegative measurable function $\varphi \in L^p_\mu(\mathbb{R}^d)$,*

$$\mathbb{E}\int_{\mathbb{R}^d} \varphi(X(x))\mu(\mathrm{d}x) \leqslant K_p\|\varphi\|_{L^p_\mu}. \tag{3.10}$$

(ii) *If φ_n converges to φ in $L^p_\mu(\mathbb{R}^d)$, then for any $N > 0$,*

$$\lim_{n\to\infty} \mathbb{E}\int_{B_N} |\varphi_n(X_n(x)) - \varphi(X(x))|\mu(\mathrm{d}x) = 0. \tag{3.11}$$

Proof. (i) First, for any nonnegative continuous function $\varphi \in C_c(\mathbb{R}^d)$ with compact support, by Fatou's lemma and (3.9), we have

$$\mathbb{E}\Big(\int_{\mathbb{R}^d} \varphi(X(x))\mathrm{d}x\Big) \leqslant \varliminf_{n\to\infty} \mathbb{E}\Big(\int_{\mathbb{R}^d} \varphi(X_n(x))\mu(\mathrm{d}x)\Big) \leqslant K_p\|\varphi\|_{L^p_\mu}.$$

Let $O \subset \mathbb{R}^d$ be a bounded open set. Define

$$\varphi_n(x) := 1 - \Big(\frac{1}{1 + \mathrm{distance}(x, O^c)}\Big)^n.$$

Then $\varphi_n \in C_c(\mathbb{R}^d)$ and for every $x \in \mathbb{R}^d$,

$$\varphi_n(x) \uparrow 1_O(x) \text{ as } n \to \infty.$$

By the monotone convergence theorem, we find that (3.10) holds for $\varphi = 1_O$.

We now extend (3.10) to the indicator function of any bounded Borel set. Without loss of generality, we consider Borel sets in $(0, 1]^d$, and define

$$\mathscr{C} := \Big\{A \in \mathcal{B}((0,1]^d) : \mathbb{E}\Big(\int_{\mathbb{R}^d} 1_A(X(x))\mu(\mathrm{d}x)\Big) \leqslant K_p\mu(A)^{1/p}\Big\}$$

and

$$\mathscr{A} := \big\{A = \Pi_{i=1}^d(\alpha_i, \beta_i] : 0 < \alpha_i \leqslant \beta_i \leqslant 1\big\}.$$

It is easy to see that $\mathscr{C}$ is a monotone class and $\mathscr{A}$ is a semi-algebra on $(0,1]^d$. Let $\mathscr{A}_{\Sigma f}$ be the algebra generated by $\mathscr{A}$ through finite disjoint unions. Since all open subsets of $(0,1]^d$ belong to $\mathscr{C}$, by another approximation, one finds that $\mathscr{A}_{\Sigma f} \subset \mathscr{C}$. Hence, by the monotone class theorem,

$$\mathcal{B}((0,1]^d) \supset \mathscr{C} \supset \sigma(\mathscr{A}_{\Sigma f}) = \mathcal{B}((0,1]^d).$$

Let φ be a bounded nonnegative measurable function on some bounded open set O. By Lusin's theorem, there exists a sequence of bounded continuous functions φ_ε with supports in O such that

$$\|\varphi_\varepsilon\|_\infty \leqslant \|\varphi\|_\infty, \quad \lim_{\varepsilon\to 0} \mu(A_\varepsilon) = 0,$$

where $A_\varepsilon := \{x \in \mathbb{R}^d : \varphi(x) \neq \varphi_\varepsilon(x)\}$. Hence,

$$\begin{aligned}\mathbb{E}\Big(\int_{\mathbb{R}^d} |\varphi - \varphi_\varepsilon|(X(x))\mu(\mathrm{d}x)\Big) &\leqslant 2\|\varphi\|_\infty \mathbb{E}\Big(\int_{\mathbb{R}^d} 1_{A_\varepsilon}(X(x))\mu(\mathrm{d}x)\Big)\\ &\leqslant 2\|\varphi\|_\infty K_p \mu(A_\varepsilon)^{1/p} \overset{\varepsilon\to 0}{\longrightarrow} 0.\end{aligned}$$

For a general unbounded nonnegative measurable function φ on $\mathbb{R}^d$, we can approximate it by the monotone convergence theorem again.

(ii) Let $\varphi_m \in C_c(\mathbb{R}^d)$ converge to φ in $L^p_\mu(\mathbb{R}^d)$. By (3.9) and (3.10), we have

$$\begin{aligned}&\mathbb{E}\int_{B_N} |\varphi_n(X_n(x)) - \varphi(X(x))|\mu(\mathrm{d}x)\\ &\leqslant K_p\|\varphi_n - \varphi\|_{L^p_\mu} + \mathbb{E}\int_{B_N} |\varphi(X_n(x)) - \varphi(X(x))|\mu(\mathrm{d}x)\\ &\leqslant K_p\|\varphi_n - \varphi\|_{L^p_\mu} + 2K_p\|\varphi_m - \varphi\|_{L^p_\mu} + \mathbb{E}\int_{B_N} |\varphi_m(X_n(x)) - \varphi_m(X(x))|\mu(\mathrm{d}x),\end{aligned}$$

which converges to zero by first letting $n \to \infty$ and then $m \to \infty$. □

Let $\varrho \geqslant 0$ be a smooth function in $\mathbb{R}^d$ with $\mathrm{supp}\varrho \subset B_1$ and $\int_{\mathbb{R}^d} \varrho(x)\mathrm{d}x = 1$. For $\varepsilon > 0$, set

$$\varrho_\varepsilon(x) := \varepsilon^{-d}\varrho(\varepsilon^{-1}x). \tag{3.12}$$

For a function $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, define

$$b_\varepsilon(x) := b * \varrho_\varepsilon(x) = \int_{\mathbb{R}^d} b(y)\varrho_\varepsilon(x-y)\mathrm{d}y, \tag{3.13}$$

and for any $R > 0$ and $\varphi \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$,

$$M_R\varphi(x) := \sup_{0<s<R} \fint_{B_s} \varphi(x+y)\mathrm{d}y,$$

where

$$\fint_{B_s} \varphi(x+y)\mathrm{d}y := \frac{1}{|B_s|}\int_{B_s} \varphi(x+y)\mathrm{d}y.$$

We have the following elementary estimate:

Lemma 3.5. *Let $b \in W^{1,1}_{\rm loc}(\mathbb{R}^d)$. Then there exists an $\mathscr{L}$-null set $A \subset \mathbb{R}^d$ such that for all $x, y \notin A$,*

$$|b(x)-b(y)| \leqslant 2^d \int_0^{|x-y|} \fint_{B_s} |\nabla b|(x+z)\mathrm{d}z\mathrm{d}s + 2^d \int_0^{|x-y|} \fint_{B_s} |\nabla b|(y+z)\mathrm{d}z\mathrm{d}s.$$

In particular, for any $R>0$ and $x, y \notin A$ with $|x-y| \leqslant R$,

$$|b(x)-b(y)| \leqslant 2^d |x-y| (M_R|\nabla b|(x) + (M_R|\nabla b|(y)). \tag{3.14}$$

Proof. Let $b_\varepsilon(x)$ be defined by (3.13). For $r>0$, let $\Pi(\mathrm{d}z)$ denote the surface measure on the ball $\{z \in \mathbb{R}^d : |z| = r\}$. Noting that

$$|b_\varepsilon(x) - b_\varepsilon(x+z)| \leqslant |z| \int_0^1 |\nabla b_\varepsilon|(x+sz)\mathrm{d}s,$$

we have

$$\begin{aligned}\int_{|z|=r} |b_\varepsilon(x) - b_\varepsilon(x+z)|\Pi(\mathrm{d}z) &\leqslant r \int_0^1 \int_{|z|=r} |\nabla b_\varepsilon|(x+sz)\Pi(\mathrm{d}z)\mathrm{d}s \\ &= r \int_0^1 s^{1-d} \int_{|z|=sr} |\nabla b_\varepsilon|(x+z)\Pi(\mathrm{d}z)\mathrm{d}s.\end{aligned}$$

Hence, for any $\ell > 0$,

$$\begin{aligned}\int_{B_\ell} |b_\varepsilon(x) - b_\varepsilon(x+z)|\,\mathrm{d}z &= \int_0^\ell \int_{|z|=r} |b_\varepsilon(x) - b_\varepsilon(x+z)|\Pi(\mathrm{d}z)\,\mathrm{d}r \\ &\leqslant \int_0^\ell r \int_0^1 s^{1-d} \int_{|z|=sr} |\nabla b_\varepsilon|(x+z)\Pi(\mathrm{d}z)\,\mathrm{d}s\,\mathrm{d}r \\ &= \int_0^1 s^{-1-d} \int_0^{s\ell} r \int_{|z|=r} |\nabla b_\varepsilon|(x+z)\Pi(\mathrm{d}z)\,\mathrm{d}r\,\mathrm{d}s \\ &\leqslant \int_0^1 s^{-d}\ell \int_{B_{s\ell}} |\nabla b_\varepsilon|(x+z)\,\mathrm{d}z\mathrm{d}s = \ell^d \int_0^\ell s^{-d} \int_{B_s} |\nabla b_\varepsilon|(x+z)\,\mathrm{d}z\mathrm{d}s.\end{aligned}$$

For any $x, y \in \mathbb{R}^d$, set $\ell := |x-y|$, then

$$\begin{aligned}|b_\varepsilon(x) - b_\varepsilon(y)| &\leqslant \fint_{B_{\ell/2}} |b_\varepsilon(x) - b_\varepsilon(\tfrac{x+y}{2}+z)|\mathrm{d}z + \fint_{B_{\ell/2}} |b_\varepsilon(y) - b_\varepsilon(\tfrac{x+y}{2}+z)|\mathrm{d}z \\ &\leqslant 2^d \fint_{B_\ell} |b_\varepsilon(x) - b_\varepsilon(x+z)|\mathrm{d}z + 2^d \fint_{B_\ell} |b_\varepsilon(y) - b_\varepsilon(y+z)|\mathrm{d}z \\ &\leqslant 2^d \int_0^\ell \fint_{B_s} |\nabla b_\varepsilon|(x+z)\mathrm{d}z\mathrm{d}s + 2^d \int_0^\ell \fint_{B_s} |\nabla b_\varepsilon|(y+z)\mathrm{d}z\,\mathrm{d}s. \end{aligned} \tag{3.15}$$

Since for any $R, \ell > 0$,

$$\lim_{\varepsilon \to 0} \int_0^1 \int_{B_R} |b_\varepsilon - b|(x)\,\mathrm{d}x\,\mathrm{d}t = 0$$

and

$$\lim_{\varepsilon\to 0}\int_0^1\int_{B_R}\Big(\int_0^\ell \fint_{B_s}|\nabla(b_\varepsilon-b)|(x+z)\,\mathrm{d}z\,\mathrm{d}s\Big)\,\mathrm{d}x\,\mathrm{d}t=0,$$

we can take the limit $\varepsilon\to 0$ in (3.15) and obtain the desired estimate. □

Lemma 3.6. *Let $b\in W^{1,1}_{\rm loc}(\mathbb{R}^d)$. There exists an $\mathscr{L}$-null set $A\subset\mathbb{R}^d$ such that for any $\delta,\varepsilon\in(0,\frac14)$, and all $x,y\in\mathbb{R}^d\setminus A$ with $|x-y|\leqslant\sqrt{\delta}$,*

$$\frac{|b(x)-b(y)|}{\sqrt{|x-y|^2+\delta^2}}\leqslant 2^d(f_{\delta,\varepsilon}(x)+f_{\delta,\varepsilon}(y)), \tag{3.16}$$

where

$$f_{\delta,\varepsilon}(x):=\varepsilon^{-d}\|\varrho\|_\infty\int_{B_1}|\nabla b|(x+z)\,\mathrm{d}z+\frac{1}{\delta}\int_0^\delta\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s$$
$$+\int_\delta^{\sqrt{\delta}}\frac{1}{s}\Big(\fint_{B_s}|\nabla(b_\varepsilon-b)|(x+z)\mathrm{d}z\Big)\,\mathrm{d}s,$$

and $b_\varepsilon(x)=b*\varrho_\varepsilon(x)$ is the mollifying vector field. Moreover, for any $R>0$,

$$\int_{B_R}f_{\delta,\varepsilon}(x)\mathrm{d}x\leqslant C_{\varrho,d}\varepsilon^{-d}\|\nabla b\|_{L^1(B_{R+1})}+\frac{\log\delta^{-1}}{2}\|\nabla(b_\varepsilon-b)\|_{L^1(B_{R+1})}, \tag{3.17}$$

where $C_{\varrho,d}$ only depends on $\|\varrho\|_\infty$ and d.

Proof. Set $\ell:=|x-y|\leqslant\sqrt{\delta}$. By Lemma 3.5, we have

$$\frac{|b(x){-}b(y)|}{\sqrt{|x-y|^2+\delta^2}}\leqslant 2^d\Big(\frac{1}{\delta}\wedge\frac{1}{\ell}\Big)\Big(\int_0^\ell\fint_{B_s}|\nabla b|(x+z)\mathrm{d}z\mathrm{d}s+\int_0^\ell\fint_{B_s}|\nabla b|(y+z)\mathrm{d}z\mathrm{d}s\Big).$$

We make the following estimate:

$$\begin{aligned}
&\Big(\frac{1}{\delta}\wedge\frac{1}{\ell}\Big)\int_0^\ell\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s\\
&\leqslant\frac{1}{\delta}\int_0^\delta\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s+\frac{1_{\ell>\delta}}{\ell}\int_\delta^\ell\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s\\
&\leqslant\frac{1}{\delta}\int_0^\delta\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s+\frac{1}{\ell}\int_\delta^\ell\fint_{B_s}|\nabla b_\varepsilon|(x+z)\,\mathrm{d}z\,\mathrm{d}s\\
&\quad+\frac{1_{\ell>\delta}}{\ell}\int_\delta^\ell\fint_{B_s}|\nabla(b_\varepsilon-b)|(x+z)\,\mathrm{d}z\,\mathrm{d}s\\
&\leqslant\frac{1}{\delta}\int_0^\delta\fint_{B_s}|\nabla b|(x+z)\,\mathrm{d}z\,\mathrm{d}s+\sup_{z\in B_{\sqrt{\delta}}}|\nabla b_\varepsilon(x+z)|\\
&\quad+\int_\delta^{\sqrt{\delta}}\frac{1}{s}\Big(\fint_{B_s}|\nabla(b_\varepsilon-b)|(x+z)\mathrm{d}z\Big)\mathrm{d}s.
\end{aligned}$$

The estimate (3.16) now follows by noting that, provided that $\varepsilon, \delta < \frac{1}{4}$,

$$\sup_{z\in B_{\sqrt{\delta}}} |\nabla b_\varepsilon|(x+z) \leqslant \varepsilon^{-d}\|\varrho\|_\infty \int_{B_1} |\nabla b|(x+z)\mathrm{d}z$$

As for (3.17), by Fubini's theorem, we have

$$\begin{aligned}\int_0^1\int_{B_R} f_{\delta,\varepsilon}(x)\mathrm{d}x\mathrm{d}s &\leqslant \varepsilon^{-d}\|\varrho\|_\infty \int_{B_R}\int_{B_1} |\nabla b|(x+z)\mathrm{d}z\mathrm{d}x + \int_0^1\int_{B_{R+1}} |\nabla b|(z)\mathrm{d}z\mathrm{d}t \\ &\quad + \int_\delta^{\sqrt{\delta}} \frac{1}{s}\mathrm{d}s \int_0^1\int_{B_{R+1}} |\nabla(b_\varepsilon - b)|(z)\mathrm{d}z\mathrm{d}t \\ &\leqslant (\varepsilon^{-d}\|\varrho\|_\infty|B_1|+1)\int_0^1\int_{B_{R+1}} |\nabla b|(z)\mathrm{d}z\mathrm{d}t \\ &\quad + \log\left(\frac{1}{\sqrt{\delta}}\right)\int_0^1\int_{B_{R+1}} |\nabla(b_\varepsilon - b)|(z)\mathrm{d}z\mathrm{d}t.\end{aligned}$$

The proof is complete. □

We also recall the following well known result (cf. [26]):

Lemma 3.7. *For any $p>1$, there exists $C_{d,p}>0$ such that for any $N,R>0$ and $\varphi\in L^p_{\mathrm{loc}}(\mathbb{R}^d)$,*

$$\int_{B_N} (M_R\varphi(x))^p\mathrm{d}x \leqslant C_{d,p}\int_{B_{N+R}} |\varphi(x)|^p\mathrm{d}x. \tag{3.18}$$

3.3. An abstract criterion for the Laplace principle

Let $\mathbb{H}$ be the Cameron–Martin space over the classical Wiener space, the space of all absolutely continuous functions from $[0,1]$ to $\mathbb{R}^d$, which is isomorphic to $L^2(0,1;\mathbb{R}^d)$ through the mapping $h\mapsto \int_0^\cdot h_s\mathrm{d}s$. Below, we always regard $\mathbb{H}$ as $L^2(0,1;\mathbb{R}^d)$. For $M>0$, set

$$\mathcal{D}_M := \{h\in\mathbb{H} : \|h\|_{\mathbb{H}} \leqslant M\}$$

and

$$\mathcal{A}_M := \left\{\begin{array}{l} h:[0,1]\to\mathbb{H} \text{ is a simple and } (\mathscr{F}_t)\text{-adapted} \\ \text{process, and for almost all } \omega,\quad h(\cdot,\omega)\in\mathcal{D}_M \end{array}\right\}. \tag{3.19}$$

We equip $\mathcal{D}_M$ with the topology of weak convergence in $\mathbb{H}$ so that $\mathcal{D}_M$ becomes a compact Polish space. Let $\mathbb{S}$ be a Polish space. A function $I:\mathbb{S}\to[0,\infty]$ is given.

Definition 3.8. The function I is called a *rate function* if for every $a<\infty$, the set $\{f\in\mathbb{S} : I(f)\leqslant a\}$ is compact in $\mathbb{S}$.

Let $\{Z^\varepsilon : \Omega\to\mathbb{S}, \varepsilon\in(0,1)\}$ be a family of measurable mappings. Assume that there is a measurable map $Z_0:\mathbb{H}\to\mathbb{S}$ such that

(LD)$_1$ For any $M>0$, if a family $\{h_\varepsilon, \varepsilon\in(0,1)\}\subset\mathcal{A}_M$ (as random variables in $\mathcal{D}_M$) converges in distribution to $h\in\mathcal{A}_M$, then for some subsequence ε_k, $Z^{\varepsilon_k}\big(\cdot+\frac{1}{\sqrt{\varepsilon_k}}\int_0^\cdot h^{\varepsilon_k}_s(\cdot)\mathrm{d}s\big)$ converges in distribution to $Z_0(h)$ in $\mathbb{S}$.

(LD)$_2$ For any $M>0$, if $\{h_n, n\in\mathbb{N}\}\subset\mathcal{D}_M$ converges weakly to $h\in\mathbb{H}$, then for some subsequence h_{n_k}, $Z_0(h_{n_k})$ converges to $Z_0(h)$ in $\mathbb{S}$.

For each $f\in\mathbb{S}$, define

$$I(f):=\frac{1}{2}\inf_{\{h\in\mathbb{H}:\ f=Z_0(h)\}}\|h\|^2_{\mathbb{H}}, \tag{3.20}$$

where $\inf\emptyset=\infty$ by convention. Then under (LD)$_2$, $I(f)$ is a rate function.

We recall the following result due to [5] (see also Theorem 4.4 in [29]).

Theorem 3.9. *Under* (LD)$_1$ *and* (LD)$_2$, $\{Z^\varepsilon,\varepsilon\in(0,1)\}$ *satisfies the Laplace principle with the rate function* $I(f)$ *given by* (3.20). *More precisely, for each real bounded continuous function* g *on* $\mathbb{S}$*:*

$$\lim_{\varepsilon\to0}\varepsilon\log\mathbb{E}\Big(\exp\Big[-\frac{g(Z^\varepsilon)}{\varepsilon}\Big]\Big)=-\inf_{f\in\mathbb{S}}\{g(f)+I(f)\}. \tag{3.21}$$

In particular, the family $\{Z^\varepsilon,\varepsilon\in(0,1)\}$ *satisfies the large deviation principle in* $(\mathbb{S},\mathcal{B}(\mathbb{S}))$ *with the rate function* $I(f)$.

4. Proofs of Theorems 2.2 and 2.4

We first establish the following key stability estimate:

Lemma 4.1. *Assume that for some* $q\geqslant1$,

$$b,\hat b\in L^q_{\rm loc}(\mathbb{R}^d),\quad |\nabla b|\in L^q_{\rm loc}(\mathbb{R}^d)\quad \textit{and}\quad \sigma,\hat\sigma\in L^{2q}_{\rm loc}(\mathbb{R}^d),\quad |\nabla\sigma|\in L^{2q}_{\rm loc}(\mathbb{R}^d).$$

Let $\mu(\mathrm{d}x)=\mathrm{e}^{\lambda(x)}\mathrm{d}x$ *with* $\lambda\in C(\mathbb{R}^d)$. *Let* $X_t(x)$ *and* $\hat X_t(x)$ *be two* μ*-almost everywhere stochastic flows of* (1.2) *corresponding to* (b,σ) *and* $(\hat b,\hat\sigma)$ *in the sense of Definition* 2.1 *with* $p=q$ *in* (2.1). *Then for any* $N,R>1$, *there exist constants* $C_1,C_2,C_3>0$ *such that for all* $\eta,\delta,\varepsilon\in(0,1)$,

$$\begin{aligned}
&\mathbb{E}\int_{B_N}\Big(\sup_{t\in[0,1]}|X_t(x)-\hat X_t(x)|^2\wedge1\Big)\mu(\mathrm{d}x)\\
&\qquad\leqslant\eta+\frac{2\mu(B_N)}{R\eta}\mathbb{E}\int_{B_N}\Big(\sup_{t\in[0,1]}|X_t(x)|\vee|\hat X_t(x)|\Big)\mu(\mathrm{d}x)\\
&\qquad\quad+\frac{C_1(\varepsilon^{-d}1_{q=1}+1_{q>1})}{\eta\log\delta^{-1}}+\frac{C_2}{\eta}\|\nabla(b_\varepsilon-b)\|_{L^1(B_{R+1})}1_{q=1}\\
&\qquad\quad+\frac{C_3}{\eta\delta\log(4\delta)^{-1}}\Big(\|b-\hat b\|_{L^q(B_R)}+\|\sigma-\hat\sigma\|_{L^{2q}(B_R)}\Big),
\end{aligned}$$

where $b_\varepsilon(x)=b*\varrho_\varepsilon(x)$, $C_1=C(R,N,\|\nabla b\|_{L^q(B_{R+1})},\|\nabla\sigma\|_{L^{2q}(B_{R+1})},K_q,\lambda)$, *and* $C_2=C_3=C(R,N,K_q,\lambda)$. *Here,* K_q *is from* (2.1).

Proof. For $\delta > 0$, let $\xi_\delta : \mathbb{R}_+ \to \mathbb{R}_+$ be a smooth function with $0 \leqslant \xi'_\delta(s) \leqslant 1$, $0 \leqslant \xi''_\delta(s) \leqslant \frac{4}{\delta}$ and

$$\xi_\delta(s) = \begin{cases} s, & s \in [0, \delta/4], \\ \delta/2, & s \in [\delta, \infty). \end{cases}$$

By elementary calculations, we have

$$s \leqslant 2\xi_\delta(s), \quad s \in [0, \delta]. \tag{4.1}$$

Set

$$Z_t(\omega, x) := X_t(\omega, x) - \hat{X}_t(\omega, x)$$

and

$$\Phi_\delta(\omega, x) := \sup_{t\in[0,1]} \xi_\delta(|Z_t(\omega, x)|^2) = \xi_\delta\Big(\sup_{t\in[0,1]} |Z_t(\omega, x)|^2\Big).$$

We divide the proof into two steps.

Step 1. In this step we prove that for any $N, R > 1$, there exist constants $C_1, C_2, C_3 > 0$ as in the statement of the lemma such that for all $\delta, \varepsilon \in (0, 1)$,

$$\begin{aligned} \mathbb{E} \int_{B_N \cap G_R} \log\Big(\frac{\Phi_\delta(x)}{\delta^2} + 1\Big)\mu(\mathrm{d}x) &\leqslant C_1 \varepsilon^{-d} + C_2 \log \delta^{-1} \int_{B_{R+1}} |\nabla(b_\varepsilon - b)|(z)\mathrm{d}z \\ &\quad + \frac{C_3}{\delta}\Big(\|b - \hat{b}\|_{L^q(B_R)} + \|\sigma - \hat{\sigma}\|_{L^{2q}(B_R)}\Big), \end{aligned} \tag{4.2}$$

where

$$G_R(\omega) := \Big\{x \in \mathbb{R}^d : \sup_{t\in[0,1]} |X_t(\omega, x)| \vee |\hat{X}_t(\omega, x)| \leqslant R\Big\}.$$

Noticing that for μ-almost all $x \in \mathbb{R}^d$ and all $t \in [0, 1]$

$$Z_t(x) = \int_0^t (b(X_s(x)) - \hat{b}(\hat{X}_s(x)))\,\mathrm{d}s + \int_0^t (\sigma(X_s(x)) - \hat{\sigma}(\hat{X}_s(x)))\,\mathrm{d}W_s,$$

by Itô's formula, we have

$$\begin{aligned} &\log\Big(\frac{\xi_\delta(|Z_t(x)|^2)}{\delta^2} + 1\Big) \\ &\quad = 2\int_0^t \frac{\xi'_\delta(|Z_s(x)|^2)\langle Z_s(x), b(X_s(x)) - \hat{b}(\hat{X}_s(x))\rangle}{\xi_\delta(|Z_s(x)|^2) + \delta^2}\,\mathrm{d}s \\ &\qquad + 2\int_0^t \frac{\xi'_\delta(|Z_s(x)|^2)\langle Z_s(x), (\sigma(X_s(x)) - \hat{\sigma}(\hat{X}_s(x)))\mathrm{d}W_s\rangle}{\xi_\delta(|Z_s(x)|^2) + \delta^2} \\ &\qquad + \int_0^t \frac{\xi'_\delta(|Z_s(x)|^2)\|\sigma(X_s(x)) - \hat{\sigma}(\hat{X}_s(x))\|^2}{\xi_\delta(|Z_s(x)|^2) + \delta^2}\,\mathrm{d}s \\ &\qquad + 2\int_0^t \frac{\xi''_\delta(|Z_s(x)|^2)|(\sigma(X_s(x)) - \hat{\sigma}(\hat{X}_s(x)))^{\mathrm{t}} \cdot Z_s(x)|^2}{\xi_\delta(|Z_s(x)|^2) + \delta^2}\,\mathrm{d}s \\ &\qquad - 2\int_0^t \frac{(\xi'_\delta(|Z_s(x)|^2))^2|(\sigma(X_s(x)) - \hat{\sigma}(\hat{X}_s(x)))^{\mathrm{t}} \cdot Z_s(x)|^2}{(\xi_\delta(|Z_s(x)|^2) + \delta^2)^2}\,\mathrm{d}s \\ &\quad =: I_1(t, x) + I_2(t, x) + I_3(t, x) + I_4(t, x) + I_5(t, x). \end{aligned}$$

Since $I_5(t,x)$ is negative, we can drop it. For $I_1(t,x)$, by (4.1), we have

$$\begin{aligned}\sup_{t\in[0,1]}|I_1(t,x)| &\leqslant 4\int_0^1 \frac{|b(X_s(x))-b(\hat{X}_s(x))|\cdot 1_{|Z_s(x)|\leqslant\sqrt{\delta}}}{\sqrt{|Z_s(x)|^2+\delta^2}}\,\mathrm{d}s\\ &\quad+\frac{2}{\delta}\int_0^1|b(\hat{X}_s(x))-\hat{b}(\hat{X}_s(x))|\,\mathrm{d}s\\ &=: I_{11}(x)+I_{12}(x).\end{aligned}$$

Noting that

$$G_R(\omega)\subset\{x:|X_t(\omega,x)|\leqslant R\}\cap\{x:|\hat{X}_t(\omega,x)|\leqslant R\},\quad \forall t\in[0,1],$$

by (2.1), we have

$$\begin{aligned}\mathbb{E}\int_{G_R}|I_{12}(x)|\mu(\mathrm{d}x) &\leqslant \frac{2}{\delta}\,\mathbb{E}\int_0^1\!\!\int_{\mathbb{R}^d}|1_{B_R}(b-\hat{b})|(\hat{X}_s(x))\mu(\mathrm{d}x)\mathrm{d}s\\ &\leqslant \frac{2K_q}{\delta}\,\|1_{B_R}(b-\hat{b})\|_{L^q_\mu}\leqslant\frac{C_{q,R,\lambda}}{\delta}\,\|b-\hat{b}\|_{L^q(B_R)}.\end{aligned}\tag{4.3}$$

For $I_{11}(x)$, if $q=1$, by Lemma 3.6, we have

$$\begin{aligned}\mathbb{E}\int_{G_R}|I_{11}(x)|\mu(\mathrm{d}x) &\leqslant 2^{d+2}\,\mathbb{E}\int_0^1\!\!\int_{G_R}[f_{\delta,\varepsilon}(X_s(x))+f_{\delta,\varepsilon}(\hat{X}_s(x))]\mu(\mathrm{d}x)\mathrm{d}s\\ &\leqslant C_d\int_{B_R}f_{\delta,\varepsilon}(x)\mu(\mathrm{d}x)\leqslant C_{d,R,\lambda}\int_{B_R}f_{\delta,\varepsilon}(x)\mathrm{d}x\\ &\leqslant C_{d,R,\lambda,\varrho}\left(\varepsilon^{-d}\|\nabla b\|_{L^1(B_{R+1})}+\log\delta^{-1}\|\nabla(b_\varepsilon-b)\|_{L^1(B_{R+1})}\right);\end{aligned}\tag{4.4}$$

if $q>1$, by Lemma 3.7, we have

$$\begin{aligned}\mathbb{E}\int_{G_R}|I_{11}(x)|\mu(\mathrm{d}x) &\leqslant C\,\mathbb{E}\int_0^1\!\!\int_{G_R}(M_{\sqrt{\delta}}|\nabla b|(X_s(x))+M_{\sqrt{\delta}}|\nabla b|(\hat{X}_s(x)))\mu(\mathrm{d}x)\mathrm{d}s\\ &\leqslant C\left(\int_{B_R}(M_{\sqrt{\delta}}|\nabla b|(x))^q\mu(\mathrm{d}x)\right)^{1/q}\leqslant C\|\nabla b\|_{L^q(B_{R+1})}.\end{aligned}\tag{4.5}$$

For $I_2(t,x)$, set

$$\tau_R(\omega,x):=\inf\Big\{t\in[0,1]:|X_t(\omega,x)|\vee\hat{X}_t(\omega,x)>R\Big\},$$

then

$$G_R(\omega)=\{x:\tau_R(\omega,x)=1\}.$$

By Burkholder's inequality, Fubini's theorem and (4.1), we have

$$
\begin{aligned}
&\mathbb{E}\int_{B_N\cap G_R}\sup_{t\in[0,1]}|I_2(t,x)|\mu(\mathrm{d}x)\\
&\leqslant \int_{B_N}\mathbb{E}\left(\sup_{t\in[0,\tau_R(x)]}\left|\int_0^t\frac{\xi'_\delta(|Z_s(x)|^2)\langle Z_s(x),(\sigma(X_s(x))-\hat\sigma(\hat X_s(x)))\mathrm{d}W_s\rangle}{\xi_\delta(|Z_s(x)|^2)+\delta^2}\right|\right)\mu(\mathrm{d}x)\\
&\leqslant C\int_{B_N}\mathbb{E}\left[\int_0^{\tau_R(x)}\frac{(\xi'_\delta(|Z_s(x)|^2))^2|Z_s(x)|^2|\sigma(X_s(x))-\hat\sigma(\hat X_s(x))|^2}{(\xi_\delta(|Z_s(x)|^2)+\delta^2)^2}\mathrm{d}s\right]^{\frac12}\mu(\mathrm{d}x)\\
&\leqslant C\mu(B_N)^{\frac12}\left[\mathbb{E}\int_0^1\int_{B_N\cap G_R}\frac{|\sigma(X_s(x))-\hat\sigma(\hat X_s(x))|^2\cdot 1_{|Z_s(x)|\leqslant\sqrt\delta}}{|Z_s(x)|^2+\delta^2}\mu(\mathrm{d}x)\mathrm{d}s\right]^{\frac12}.
\end{aligned}
$$

As the treatment of $I_1(t,x)$, by Lemma 3.7, we can prove that

$$
\mathbb{E}\int_{B_N\cap G_R}\sup_{t\in[0,1]}|I_2(t,x)|\mu(\mathrm{d}x)\leqslant C\|\nabla\sigma\|_{L^{2q}(B_{R+1})}+\frac{C}{\delta}\|\sigma-\hat\sigma\|_{L^{2q}(B_R)}, \tag{4.6}
$$

and similarly,

$$
\mathbb{E}\int_{B_N\cap G_R}\sup_{t\in[0,1]}|I_3(t,x)|\mu(\mathrm{d}x)\leqslant C\|\nabla\sigma\|_{L^{2q}(B_{R+1})}+\frac{C}{\delta}\|\sigma-\hat\sigma\|_{L^{2q}(B_R)}, \tag{4.7}
$$

$$
\mathbb{E}\int_{B_N\cap G_R}\sup_{t\in[0,1]}|I_4(t,x)|\mu(\mathrm{d}x)\leqslant C\|\nabla\sigma\|_{L^{2q}(B_{R+1})}+\frac{C}{\delta}\|\sigma-\hat\sigma\|_{L^{2q}(B_R)}. \tag{4.8}
$$

Combining (4.3)–(4.8), we obtain (4.2).

Step 2. Notice that

$$
s\wedge 1\leqslant \xi_4(s)\leqslant 2,\quad s\geqslant 0.
$$

By definition of Φ_δ, it is enough to prove the estimate for $\mathbb{E}\int_{B_N}\Phi_4(x)\mu(\mathrm{d}x)$. For any $\eta>0$, we have

$$
\begin{aligned}
\mathbb{E}\int_{B_N}\Phi_4(x)\mu(\mathrm{d}x)&\leqslant \eta+\mu(B_N)\ P\Big\{\int_{B_N}\Phi_4(x)\mu(\mathrm{d}x)\geqslant\eta\Big\}\\
&\leqslant \eta+\mu(B_N)\ P\Big\{\int_{B_N\cap G_R^c}\Phi_4(x)\mu(\mathrm{d}x)\geqslant\frac{\eta}{2}\Big\}\\
&\quad+\mu(B_N)\ P\Big\{\int_{B_N\cap G_R}\Phi_4(x)\mu(\mathrm{d}x)\geqslant\frac{\eta}{2}\Big\}.
\end{aligned} \tag{4.9}
$$

In view of $\Phi_4(x)\leqslant 2$, by Chebyshev's inequality, we have

$$
\begin{aligned}
P\Big\{\int_{B_N\cap G_R^c}\Phi_4(x)\mu(\mathrm{d}x)\geqslant\frac{\eta}{2}\Big\}&\leqslant P\left\{\mu(B_N\cap G_R^c)\geqslant\frac{\eta}{4}\right\}\leqslant\frac{4}{\eta}\,\mathbb{E}\mu(B_N\cap G_R^c)\\
&\leqslant\frac{4}{R\eta}\,\mathbb{E}\int_{B_N}\Big(\sup_{t\in[0,1]}|X_t(x)|\vee|\hat X_t(x)|\Big)\mu(\mathrm{d}x).
\end{aligned} \tag{4.10}
$$

Set now

$$\Psi_\delta(x) := \log\Big(\frac{\Phi_\delta(x)}{\delta^2}+1\Big).$$

Notice that if $\Psi_\delta(x) \leqslant \log(4\delta)^{-1}$, then $\Phi_\delta(x) \leqslant \frac{\delta}{4}$, and so $\Phi_4(x) \leqslant \frac{\delta}{4}$ by definition. Hence, for any $\delta < \frac{\eta}{\mu(B_N)}$, we have

$$\begin{aligned}(4.11)\qquad P\Big\{\int_{B_N\cap G_R}\Phi_4(x)\mu(\mathrm{d}x) \geqslant \frac{\eta}{2}\Big\}&\\ &\leqslant P\Big\{\int_{B_N\cap G_R}\Phi_4(x)\cdot 1_{\{\Psi_\delta(x)>\log(4\delta)^{-1}\}}\mu(\mathrm{d}x)\geqslant\frac{\eta}{4}\Big\}\\ &\quad + P\Big\{\int_{B_N\cap G_R}\Phi_4(x)\cdot 1_{\{\Psi_\delta(x)\leqslant\log(4\delta)^{-1}\}}\mu(\mathrm{d}x)\geqslant\frac{\eta}{4}\Big\}\\ &\leqslant P\Big\{\int_{B_N\cap G_R}\Psi_\delta(x)\mu(\mathrm{d}x)\geqslant\frac{\eta\log(4\delta)^{-1}}{8}\Big\}+0\\ &\leqslant \frac{8}{\eta\log(4\delta)^{-1}}\mathbb{E}\int_{B_N\cap G_R}\Psi_\delta(x)\mu(\mathrm{d}x).\end{aligned}$$

The result now follows by combining (4.2), (4.9), (4.10) and (4.11). □

Let $\chi \in C^\infty(\mathbb{R}^d)$ be a nonnegative cutoff function with

$$(4.12)\qquad \|\chi\|_\infty \leqslant 1,\quad \chi(x) = \begin{cases}1, & |x|\leqslant 1,\\ 0, & |x|\geqslant 2.\end{cases}$$

Set $\chi_n(x) := \chi(x/n)$ and define

$$(4.13)\qquad b_n := b*\rho_n\cdot\chi_n,\quad \sigma_n := \sigma*\rho_n\cdot\chi_n,$$

where $\rho_n = \varrho_{1/n}$ is the mollifier given by (3.12).

We are now in a position to give the proofs of Theorems 2.2 and 2.4.

Proof of Theorem 2.2. Let b_n and σ_n be defined by (4.13). Let $X_t^n(x)$ be the solution of the Stratonovich SDE

$$\begin{aligned}X_t^n(x) &= x+\int_0^t b_n(X_s^n(x))\,\mathrm{d}s+\int_0^t\sigma_n(X_s^n(x))\circ\mathrm{d}W_s\\ &= x+\int_0^t \tilde b_n(X_s^n(x))\,\mathrm{d}s+\int_0^t\sigma_n(X_s^n(x))\,\mathrm{d}W_s,\end{aligned}$$

where

$$\tilde b_n := b_n+\frac{1}{2}\sigma_n^{jl}\partial_j\sigma_n^{\cdot l}.$$

We divide the proof into three steps.

Step 1. By Lemma 3.1 and the properties of the convolution operator, for all $x \in \mathbb{R}^d$ and $t \in [0,1]$, we have

$$\begin{aligned}
&\mathbb{E}\,|\det(\nabla[X_t^n(x)]^{-1})| \\
&\quad\leqslant \exp\Big\{\|[-\operatorname{div}\tilde{b}_n + \tfrac{1}{2}\partial_i\sigma_n^{jl}\partial_j\sigma_n^{il} + \sigma_n^{il}\partial_{ij}^2\sigma_n^{jl} + \tfrac{1}{2}|\operatorname{div}\sigma_n|^2]^+\|_\infty\Big\} \\
&\quad= \exp\Big\{\|[-\operatorname{div} b_n + \tfrac{1}{2}\sigma_n^{il}\partial_{ij}^2\sigma_n^{jl} + \tfrac{1}{2}|\operatorname{div}\sigma_n|^2]^+\|_\infty\Big\} \\
&\quad\leqslant \exp\Big\{\|[\operatorname{div} b_n]^-\|_\infty + \tfrac{1}{2}\||\sigma_n|\cdot|\nabla\operatorname{div}\sigma_n|\|_\infty + \tfrac{1}{2}\|\operatorname{div}\sigma_n\|_\infty^2\Big\}.
\end{aligned}$$

Noticing that

$$\begin{aligned}
\operatorname{div} b_n &= \partial_i\chi_n(b^i * \rho_n) + (\operatorname{div} b * \rho_n)\chi_n, \\
\sigma_n^{il}\partial_{ij}^2\sigma_n^{jl} &= (\sigma^{ij} * \rho_n)[(\partial_{ij}^2\sigma * \rho_n)\chi_n + 2(\partial_i\sigma * \rho_n)\partial_j\chi_n + (\sigma * \rho_n)\partial_{ij}^2\chi_n],
\end{aligned}$$

by (2.2), the definition of χ_n, and elementary calculus, for $n > 2(\frac{1}{\varepsilon} \vee r)$, where r is from (2.2), we find

$$\begin{aligned}
\|[\operatorname{div} b_n]^-\|_\infty &\leqslant C + \|[\operatorname{div} b]^-\|_\infty, \\
\||\sigma_n|\cdot|\nabla\operatorname{div}\sigma_n|\|_\infty &\leqslant C + \Big\|\sup_{|z|\leqslant\varepsilon}|\sigma(\cdot - z)|\cdot|\nabla\operatorname{div}\sigma|\Big\|_\infty, \\
\|\operatorname{div}\sigma_n\|_\infty^2 &\leqslant C + \|\operatorname{div}\sigma\|_\infty^2.
\end{aligned}$$

Here and below, C is independent of n. Thus,

$$\sup_{n\in\mathbb{N}}\sup_{(t,x)\in[0,1]\times\mathbb{R}^d}\mathbb{E}\,|\det(\nabla[X_t^n(x)]^{-1})| < +\infty.$$

Hence, for any nonnegative measurable function $\varphi \in L^1(\mathbb{R}^d)$,

$$\begin{aligned}
&\sup_{t\in[0,1]}\mathbb{E}\int_{\mathbb{R}^d}\varphi(X_t^n(x))\mathrm{d}x \\
&\qquad= \sup_{t\in[0,1]}\mathbb{E}\int_{\mathbb{R}^d}\varphi(x)\cdot|\det(\nabla[X_t^n(x)]^{-1})|\mathrm{d}x \leqslant K\|\varphi\|_{L^1}.
\end{aligned} \tag{4.14}$$

Step 2. In this step we prove that for any $N > 0$,

$$\sup_{n\in\mathbb{N}}\mathbb{E}\int_{B_N}\sup_{t\in[0,1]}|X_t^n(x)|^2\mathrm{d}x < +\infty. \tag{4.15}$$

Set

$$g_t(x) := \mathbb{E}\Big(\sup_{s\in[0,t]}|X_s^n(x)|^2\Big).$$

By Itô's formula, Burkholder's inequality, and Young's inequality, we have

$$
\begin{aligned}
g_t(x) \leqslant\ & |x|^2 + 2\mathbb{E}\int_0^t |X^n_s(x)|\cdot|\tilde b_n(X^n_s(x))|\mathrm{d}s + \mathbb{E}\int_0^t \|\sigma_n(X^n_s(x))\|^2\mathrm{d}s\\
&+ C\,\mathbb{E}\Big(\int_0^t |X^n_s(x)|^2\cdot\|\sigma_n(X^n_s(x))\|^2\,\mathrm{d}s\Big)^{1/2}\\
\leqslant\ & |x|^2 + 2\mathbb{E}\int_0^t |X^n_s(x)|\cdot|\tilde b_n(X^n_s(x))|\cdot(1_{|X^n_s(x)|\leqslant r} + 1_{|X^n_s(x)|>r})\,\mathrm{d}s\\
&+ \mathbb{E}\int_0^t \|\sigma_n(X^n_s(x))\|^2\mathrm{d}s + C\,\mathbb{E}\Big(\sup_{s\in[0,t]}|X^n_s(x)|\Big[\int_0^t \|\sigma_n(X^n_s(x))\|^2\mathrm{d}s\Big]^{1/2}\Big)\\
\leqslant\ & |x|^2 + 2r\,\mathbb{E}\int_0^t |\tilde b_n(X^n_s(x))|\cdot 1_{|X^n_s(x)|\leqslant r}\,\mathrm{d}s + C_r\,\mathbb{E}\int_0^t (1+|X^n_s(x)|^2)\,\mathrm{d}s\\
&+ \frac{1}{2}g_t(x) + C\,\mathbb{E}\int_0^t \|\sigma_n(X^n_s(x))\|^2\,\mathrm{d}s,
\end{aligned}
$$

where r is from (2.2) and we have used (2.2) in the last step. Hence,

$$
\begin{aligned}
g_t(x) \leqslant\ & 2|x|^2 + 4r\,\mathbb{E}\int_0^t |\tilde b_n(X^n_s(x))|\cdot 1_{|X^n_s(x)|\leqslant r}\,\mathrm{d}s\\
&+ 2\,C_r\int_0^t (1+g_s(x))\mathrm{d}s + C\,\mathbb{E}\int_0^t \|\sigma_n(X^n_s(x))\|^2\mathrm{d}s.
\end{aligned}
$$

By Gronwall's inequality, we obtain that

$$
g_1(x) \leqslant C_r\Big(|x|^2 + \mathbb{E}\int_0^1 |\tilde b_n(X^n_s(x))|\cdot 1_{|X^n_s(x)|\leqslant r}\mathrm{d}s + \mathbb{E}\int_0^1 \|\sigma_n(X^n_s(x))\|^2\mathrm{d}s\Big).
$$

Now, by (4.14) and (2.2), we have

$$
\begin{aligned}
&\mathbb{E}\int_{B_N} g_t(x)\mathrm{d}x \leqslant C_{N,r} + C_r\|\tilde b_n\|_{L^1(B_r)} + C_{N,r}(\|\sigma_n\|^2_{L^\infty(B_r^c)} + \|\sigma_n\|^2_{L^2(B_r)})\\
&\leqslant C_{N,r} + C_r\|b_n\|_{L^1(B_r)} + C_r\|\sigma_n\|_{L^2(B_r)}\|\nabla\sigma_n\|_{L^2(B_r)} + C_{N,r}(\|\sigma\|^2_{L^\infty(B_r^c)} + \|\sigma\|^2_{L^2(B_r)})\\
&\leqslant C_{N,r} + C_r\|b\|_{L^1(B_r)} + C_r\|\sigma\|_{L^2(B_r)}\|\nabla\sigma\|_{L^2(B_r)} + C_{N,r}(\|\sigma\|^2_{L^\infty(B_r^c)} + \|\sigma\|^2_{L^2(B_r)}),
\end{aligned}
$$

which gives (4.15).

Step 3. Noting that, for $n > R+1$,

$$
\|\nabla b_n\|_{L^1(B_{R+1})} \leqslant \|\nabla b\|_{L^1(B_{R+1})}, \quad \|\nabla\sigma_n\|_{L^2(B_{R+1})} \leqslant \|\nabla\sigma\|_{L^2(B_{R+1})},
$$

by (4.14), (4.15) and Lemma 4.1, we have that for any $\delta,\eta,\varepsilon\in(0,1)$,

$$
\begin{aligned}
&\mathbb{E}\int_{B_N}\Big(\sup_{t\in[0,1]}|X^n_t(x) - X^m_t(x)|^2\wedge 1\Big)\mathrm{d}x\\
&\qquad\leqslant \eta + \frac{C(N,r)}{R\eta} + \frac{C_2}{\eta}\|\nabla(b_n*\varrho_\varepsilon - b_n)\|_{L^1(B_{R+1})} + \frac{C_1\varepsilon^{-d}}{\eta\log\delta^{-1}}\\
&\qquad\quad + \frac{C_3}{\eta\delta\log\delta^{-1}}\left(\|b_n - b_m\|_{L^1(B_R)} + \|\sigma_n - \sigma_m\|_{L^2(B_R)}\right),
\end{aligned}
$$

where C_1, C_2 and C_3 are independent of n, ε and δ. We take limits in the following order: $n, m \to \infty$, $\delta \to 0$, $\varepsilon \to 0$, $R \to \infty$, $\eta \to 0$. We then find

$$\lim_{n,m\to\infty} \mathbb{E} \int_{B_N} \Big(\sup_{t\in[0,1]} |X_t^n(x) - X_t^m(x)|^2 \wedge 1 \Big) \mathrm{d}x = 0,$$

which together with (4.15) gives further that for any $p \in [1, 2)$,

$$\lim_{n,m\to\infty} \mathbb{E} \int_{B_N} \Big(\sup_{t\in[0,1]} |X_t^n(x) - X_t^m(x)|^p \Big) \mathrm{d}x = 0.$$

Therefore, there exists a continuous $\mathscr{F}_t$-adapted stochastic field $X_t(x)$ such that for any $N > 0$ and $p \in [1, 2)$,

$$\lim_{n\to\infty} \mathbb{E} \int_{B_N} \Big(\sup_{t\in[0,1]} |X_t^n(x) - X_t(x)|^p \Big) \mathrm{d}x = 0.$$

In particular, there exists a subsequence still denoted by n such that for $P \otimes \mu$-almost all (ω, x),

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |X_t^n(\omega, x) - X_t(\omega, x)| = 0.$$

Condition (A) in Definition 2.1 now follows by (4.14) and (i) of Lemma 3.4. For verifying (B) in Definition 2.1, it suffices to prove that for any $N > 0$ and $s \in [0, 1]$,

$$\lim_{n\to\infty} \mathbb{E} \int_{B_N} |b_n(X_s^n(x)) - b(X_s(x))| \, \mathrm{d}x = 0, \tag{4.16}$$

$$\lim_{n\to\infty} \mathbb{E} \int_{B_N} |(\sigma_n^{jl} \partial_j \sigma_n^{il})(X_s^n(x)) - (\sigma^{jl} \partial_j \sigma^{il})(X_s(x))| \, \mathrm{d}x = 0, \tag{4.17}$$

$$\lim_{n\to\infty} \mathbb{E} \int_{B_N} |\sigma_n(X_s^n(x)) - \sigma(X_s(x))|^2 \, \mathrm{d}x = 0. \tag{4.18}$$

We only prove (4.16). The others are analogous. We make the following decomposition:

$$\int_{B_N} |b_n(X_s^n(x)) - b(X_s(x))| \mathrm{d}x \leqslant \int_{B_N} |b_n\chi_m - b\chi_m|(X_s^n(x))| \mathrm{d}x$$
$$+ \int_{B_N} |b_n(1 - \chi_m)|(X_s^n(x)) \mathrm{d}x + \int_{B_N} |b(1 - \chi_m)|(X_s(x)) \mathrm{d}x =: I_1^{nm} + I_2^{nm} + I_3^m.$$

For fixed $m \in \mathbb{N}$, by (ii) of Lemma 3.4, we have

$$\lim_{n\to\infty} \mathbb{E}\, I_1^{nm} = 0. \tag{4.19}$$

On the other hand, for $m > r$, we have

$$I_2^{nm} \leqslant C \int_{B_N} (1 + |X_s^n(x)|) \cdot 1_{|X_s^n(x)| \geqslant m} \mathrm{d}x \leqslant \frac{C}{m} \int_{B_N} (1 + |X_s^n(x)|^2) \mathrm{d}x,$$

which together with (4.15) yields

$$\lim_{m\to\infty} \sup_n \mathbb{E}\, I_2^{nm} = 0. \tag{4.20}$$

Similarly,

$$\lim_{m\to\infty} \mathbb{E}\, I_3^m = 0. \tag{4.21}$$

Combining (4.19), (4.20) and (4.21), we get (4.16). The proof is thus complete. □

Proof of Theorem 2.4. Let b_n and σ_n be defined by (4.13). Since b and σ have linear growth, we have

$$|b_n(x)| + |\sigma_n(x)| \leqslant C(1+|x|),$$

where C is independent of n. It is then standard to prove that for any $p \geqslant 1$,

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left(\sup_{t\in[0,1]} |X_t^n(x)|^{2p}\right) < +\infty.$$

Note that

$$\partial_j \sigma_n^{il} = \partial_j \sigma^{il} * \rho_n \cdot \chi_n + \sigma_n^{il} \cdot \partial_j \chi_n,$$

and by the linear growth of σ

$$|\sigma_n \cdot \nabla \chi_n| \leqslant \frac{C\, 1_{n\leqslant |x| \leqslant 2n}}{n} \int_{\mathbb{R}^d} (1+|x-y|)\rho_n(y)\mathrm{d}y \leqslant C.$$

By Jensen's inequality and (2.5), for $n \geqslant \frac{1}{\varepsilon}$, we have

$$\begin{aligned}|\Lambda_1^{\sigma_n}|^2 &= |\operatorname{div} \sigma_n + \sigma_n^{i\cdot}\partial_i\lambda|^2 \\ &\leqslant C\left(|\operatorname{div}\sigma|^2 * \rho_n + |\sigma|^2 * \rho_n \cdot |\nabla\lambda|^2 + 1\right) \leqslant C\left(|\nabla\sigma|^2 + |\sigma|^2\gamma_2^2\right) * \rho_n + C\end{aligned}$$

and

$$\begin{aligned}-\Lambda_2^{b_n,\sigma_n} &= -\Big[\operatorname{div} b_n + b_n^i \partial_i \lambda + \frac{1}{2}(\sigma_n^{il}\sigma_n^{jl}\partial_{ij}^2\lambda - \partial_i\sigma_n^{jl}\partial_j\sigma_n^{il})\Big] \\ &\leqslant C\Big[[\operatorname{div} b]^- * \rho_n + |b| * \rho_n |\nabla\lambda| + (|\sigma| * \rho_n)^2 |\nabla^2\lambda| + (|\nabla\sigma| * \rho_n)^2) + 1\Big] \\ &\leqslant C\Big[[\operatorname{div} b]^- + |b|\gamma_2 + |\sigma|^2\gamma_3 + |\nabla\sigma|^2\Big] * \rho_n + C.\end{aligned}$$

Hence, for all $t \in [0,1]$ and $p > 1$, by Lemma 3.2 and Jensen's inequality again,

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^d} |\mathcal{J}_t^n(x)|^p \mu(\mathrm{d}x) &\leqslant C_N \sup_{t\in[0,1]} \int_{\mathbb{R}^d} \exp\left\{tp^3|\Lambda_3^{\sigma_n}(x)|^2 - tp^2\Lambda_2^{b_n,\sigma_n}(x)\right\}\mu(\mathrm{d}x) \\ &\leqslant C_N \int_{\mathbb{R}^d} \mathrm{e}^{C\left([\operatorname{div} b]^- + |b|\gamma_2 + |\sigma|^2(\gamma_2^2+\gamma_3) + |\nabla\sigma|^2\right) * \rho_n(x)} \cdot \mathrm{e}^{\lambda(x)}\mathrm{d}x \\ &\leqslant C_N \int_{\mathbb{R}^d} \mathrm{e}^{\left[C\left([\operatorname{div} b]^- + |b|\gamma_2 + |\sigma|^2(\gamma_2^2+\gamma_3) + |\nabla\sigma|^2\right) + \gamma_1\right] * \rho_n(x)}\mathrm{d}x \\ &\leqslant C_N \int_{\mathbb{R}^d} \mathrm{e}^{C\left([\operatorname{div} b]^- + |b|\gamma_2 + |\sigma|^2(\gamma_2^2+\gamma_3) + |\nabla\sigma|^2\right) + \gamma_1} * \rho_n(x)\mathrm{d}x \\ &= C_N \int_{\mathbb{R}^d} \mathrm{e}^{\left[C\left([\operatorname{div} b]^- + |b|\gamma_2 + |\sigma|^2(\gamma_2^2+\gamma_3) + |\nabla\sigma|^2\right) + \gamma_1\right](x)}\mathrm{d}x < +\infty.\end{aligned}$$

Thus, by (3.3) and Hölder's inequality, we obtain that for any $p > 1$,

$$\mathbb{E}\int_{\mathbb{R}^d} \varphi(X_t^n(x))\mu(\mathrm{d}x) = \mathbb{E}\int_{\mathbb{R}^d} \varphi(x)\mathcal{J}_t^n(x)\mu(\mathrm{d}x)$$
$$\leqslant \|\varphi\|_{L^p_\mu}\Big(\mathbb{E}\int_{\mathbb{R}^d} |\mathcal{J}_t^n(x)|^{\frac{p}{p-1}}\mu(\mathrm{d}x)\Big)^{1-\frac{1}{p}} \leqslant C.$$

The rest of the proof is the same as that of Step 3 in the proof of Theorem 2.2. □

5. Proof of Theorem 2.9

For proving Theorem 2.9, our task is to check $(\mathrm{LD})_1$ and $(\mathrm{LD})_2$. By the infinite-dimensional Yamada–Watanabe theorem (cf. [25]), there exists a measurable functional

$$\Phi_\varepsilon : \Omega \to \mathbb{S} = L^{2p}_\nu(\mathbb{R}^d; C([0,1];\mathbb{R}^d)), \quad p \geqslant 1,$$

such that

$$X_{\varepsilon,t}(\omega, x) = \Phi_\varepsilon(\omega)(t, x).$$

For $\varepsilon \in (0,1)$, let $h^\varepsilon \in \mathcal{A}_M$, where $\mathcal{A}_M$ is defined by (3.19). By Girsanov's theorem, one sees that

$$X_t^\varepsilon(\omega, x) = \Phi_\varepsilon\Big(W_\cdot(\omega) + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_s^\varepsilon(\omega)\mathrm{d}s\Big)(t,x)$$

solves the controlled equation:

$$\mathrm{d}X_t^\varepsilon(x) = b(X_t^\varepsilon(x))\mathrm{d}t + \sigma(X_t^\varepsilon(x))h_t^\varepsilon\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^\varepsilon(x))\mathrm{d}W_t, \quad X_0^\varepsilon(x) = x.$$

For $h \in \mathcal{A}_M$, let $X_t^h(x)$ solve equation (2.9). We have:

Lemma 5.1. (i) *For any $p \geqslant 1$ and $h \in \mathcal{A}_M$,*

$$\mathbb{E}\Big(\sup_{t\in[0,1]}|X_t^h(x)|^{2p}\Big) + \sup_{\varepsilon\in(0,1)}\mathbb{E}\Big(\sup_{t\in[0,1]}|X_t^\varepsilon(x)|^{2p}\Big) \leqslant C(1+|x|^{2p}).$$

(ii) *For any $p > 1$, $h^\varepsilon \in \mathcal{A}_M$ and nonnegative function $\varphi \in L^p_\mu(\mathbb{R}^d)$,*

$$\mathbb{E}\int_{B_N} \varphi(X_t^\varepsilon(x))\mu(\mathrm{d}x) \leqslant C_{N,M}\|\varphi\|_{L^p_\mu}.$$

Proof. (i) It follows in a standard way from the linear growth of b and σ.

(ii) Define b_n and σ_n by (4.13). Consider the following SDE:

$$\mathrm{d}X_t^{\varepsilon,n}(x) = b_n(X_t^{\varepsilon,n}(x))\mathrm{d}t + \sigma_n(X_t^{\varepsilon,n}(x))h_t^\varepsilon\mathrm{d}t + \sqrt{\varepsilon}\sigma_n(X_t^{\varepsilon,n}(x))\mathrm{d}W_t,$$
$$X_0^{\varepsilon,n}(x) = x.$$

From the proofs of Lemma 3.2 and Theorem 2.4, one can see that for any $p > 1$ and $\varphi \in L^p_\mu(\mathbb{R}^d)$,

$$\mathbb{E}\int_{B_N} \varphi(X^{\varepsilon,n}_t(x))\mu(\mathrm{d}x) \leqslant C_{N,M}\|\varphi\|_{L^p_\mu},$$

where $C_{N,M}$ is independent of ε. Now taking the limit $n \to \infty$ gives the result (see Lemma 3.4). □

Set

$$w^\varepsilon_t(x) := \int_0^t \sigma(X^h_s(x))(h^\varepsilon_s - h_s)\mathrm{d}s. \tag{5.1}$$

Lemma 5.2. *Suppose that h_ε converges weakly to h a.s. in $\mathcal{D}_M$. Then for any $p \geqslant 1$, we have*

$$\lim_{\varepsilon\to 0}\mathbb{E}\int_{B_N} \sup_{t\in[0,1]} |w^\varepsilon_t(x)|^{2p}\,\mathrm{d}x = 0.$$

Proof. For fixed (ω, x), let us first prove that

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,1]} |w^\varepsilon_t(\omega, x)| = 0. \tag{5.2}$$

By the weak convergence of $h^\varepsilon_\cdot(\omega)$ to $h_\cdot(\omega)$, one sees that, for fixed $t \in [0, 1]$,

$$\lim_{\varepsilon\to 0} w^\varepsilon_t(\omega, x) = \lim_{\varepsilon\to 0}\int_0^t \sigma(X^h_s(\omega, x))(h^\varepsilon_s(\omega) - h_s(\omega))\mathrm{d}s = 0.$$

Since for $t' < t$

$$\begin{aligned}|w^\varepsilon_t(\omega, x) - w^\varepsilon_{t'}(\omega, x)| &\leqslant \int_{t'}^t |\sigma(X^h_s(\omega, x))(h^\varepsilon_s(\omega) - h_s(\omega))|\,\mathrm{d}s\\ &\leqslant 2M\Big(\int_{t'}^t |\sigma(X^h_s(\omega, x))|^2\mathrm{d}s\Big)^{\frac{1}{2}} \to 0,\end{aligned}$$

uniformly in ε as $|t - t'| \to 0$, we immediately have (5.2). In view of

$$\sup_{t\in[0,1]} |w^\varepsilon_t(x)|^{2p} \leqslant C_{M,p}\int_0^1 |\sigma(X^h_s(x))|^{2p}\mathrm{d}s,$$

the desired limit now follows by the dominated convergence theorem and (5.2). □

Lemma 5.3. *Suppose that h^ε converges weakly to h a.s. in $\mathcal{D}_M$. Then for some subsequence ε_k, X^{ε_k} converges to X^h in probability in the space $\mathbb{S}$, where X^h solves equation (2.9).*

Proof. Set

$$Z^\varepsilon_t(x) := X^\varepsilon_t(x) - X^h_t(x).$$

By Itô's formula, for any $\delta > 0$, we have

$$\begin{aligned}\log\Big(\frac{|Z^\varepsilon_t(x)|^2}{\delta^2}+1\Big) &= 2\int_0^t \frac{\langle Z^\varepsilon_s(x), b(X^\varepsilon_s(x)) - b(X^h_s(x))\rangle}{|Z^\varepsilon_s(x)|^2+\delta^2}\,\mathrm{d}s\\ &\quad + 2\int_0^t \frac{\langle Z^\varepsilon_s(x), (\sigma(X^\varepsilon_s(x)) - \sigma(X^h_s(x)))h^\varepsilon_s\rangle}{|Z^\varepsilon_s(x)|^2+\delta^2}\,\mathrm{d}s\\ &\quad + 2\int_0^t \frac{\langle Z^\varepsilon_s(x), \sigma(X^h_s(x))(h^\varepsilon_s - h_s)\rangle}{|Z^\varepsilon_s(x)|^2+\delta^2}\,\mathrm{d}s\\ &\quad + 2\sqrt{\varepsilon}\int_0^t \frac{\langle Z^\varepsilon_s(x), \sigma(X^\varepsilon_s(x))\mathrm{d}W_s\rangle}{|Z^\varepsilon_s(x)|^2+\delta^2}\\ &\quad + \varepsilon\int_0^t \frac{\|\sigma(X^\varepsilon_s(x))\|^2}{|Z^\varepsilon_s(x)|^2+\delta^2}\,\mathrm{d}s - 2\varepsilon\int_0^t \frac{|(\sigma(X^\varepsilon_s(x)))^{\mathrm{t}}\cdot Z^\varepsilon_s(x)|^2}{(|Z^\varepsilon_s(x)|^2+\delta^2)^2}\,\mathrm{d}s\\ &=: I^\varepsilon_1(t,x) + I^\varepsilon_2(t,x) + I^\varepsilon_3(t,x) + I^\varepsilon_4(t,x) + I^\varepsilon_5(t,x) + I^\varepsilon_6(t,x).\end{aligned}$$

We want to prove that for any $N, R > 0$,

$$\mathbb{E}\int_{B_N\cap G^\varepsilon_R} \log\Big(\frac{\sup_{t\in[0,1]}|Z^\varepsilon_t(x)|^2}{\delta^2}+1\Big)\mu(\mathrm{d}x) \leqslant C_1 + \frac{C_2(\varepsilon)}{\delta}, \tag{5.3}$$

where C_1 is independent of ε and δ, $C_2(\varepsilon)\to 0$ as $\varepsilon\to 0$, and

$$G^\varepsilon_R(\omega) := \Big\{x\in\mathbb{R}^d : \sup_{t\in[0,1]} |X^\varepsilon_t(\omega,x)| \vee |X^h_t(\omega,x)| \leqslant R\Big\}.$$

First of all, $I^\varepsilon_6(t,x)$ is negative so can be dropped. By Lemmas 3.7 and 5.1, as in the proof of Lemma 4.1, it is easy to see that

$$\mathbb{E}\int_{B_N\cap G^\varepsilon_R} \sup_{t\in[0,1]} (|I^\varepsilon_1(t,x)| + I^\varepsilon_2(t,x)|)\mu(\mathrm{d}x) \leqslant C_1.$$

Moreover, by Burkholder's inequality, we also have

$$\mathbb{E}\int_{B_N\cap G^\varepsilon_R} \sup_{t\in[0,1]} (|I^\varepsilon_4(t,x)| + I^\varepsilon_5(t,x)|)\mu(\mathrm{d}x) \leqslant \frac{C\varepsilon}{\delta^2}.$$

We now deal with the hard term $I^\varepsilon_3(t,x)$. Set

$$\xi(x) := \frac{x}{|x|^2+\delta^2}.$$

Recalling (5.1), we have

$$I^\varepsilon_3(t,x) = 2\int_0^t \langle \xi(Z^\varepsilon_s(x)), \mathrm{d}w^\varepsilon_s(x)\rangle = 2\langle \xi(Z^\varepsilon_t(x)), w^\varepsilon_t(x)\rangle - 2\int_0^t \langle w^\varepsilon_s(x), \mathrm{d}\xi(Z^\varepsilon_s(x))\rangle.$$

By Itô's formula, we have

$$\begin{aligned}\mathrm{d}\xi(Z^\varepsilon_t(x)) &= \nabla\xi(Z^\varepsilon_t(x))(b(X^\varepsilon_t(x)) - b(X^h_t(x)))\mathrm{d}t + \nabla\xi(Z^\varepsilon_t(x))(\sigma(X^\varepsilon_t(x))h^\varepsilon_t\\ &\quad - \sigma(X^h_t(x))h_t)\mathrm{d}t + \frac{\varepsilon}{2}\partial^2_{ij}\xi(Z^\varepsilon_t(x))\sigma^{il}(X^\varepsilon_t(x))\sigma^{jl}(X^\varepsilon_t(x))\mathrm{d}t\\ &\quad + \sqrt{\varepsilon}\nabla\xi(Z^\varepsilon_t(x))\sigma(X^\varepsilon_t(x))\mathrm{d}W_t.\end{aligned}$$

Hence,

$$\begin{aligned}I_3^\varepsilon(t,x)&=2\langle\xi(Z_t^\varepsilon(x)),w_t^\varepsilon(x)\rangle-2\int_0^t\langle\nabla\xi(Z_s^\varepsilon(x))(b(X_s^\varepsilon(x))-b(X_s^h(x))),w_s^\varepsilon(x)\rangle\mathrm{d}s\\&\quad-2\int_0^t\langle\nabla\xi(Z_s^\varepsilon(x))(\sigma(X_s^\varepsilon(x))h_s^\varepsilon-\sigma(X_s^h(x))h_s),w_s^\varepsilon(x)\rangle\mathrm{d}s\\&\quad-\varepsilon\int_0^t\langle\partial_{ij}^2\xi(Z_s^\varepsilon(x))\sigma^{il}(X_s^\varepsilon(x))\sigma^{jl}(X_s^\varepsilon(x)),w_s^\varepsilon(x)\rangle\mathrm{d}s\\&\quad-2\sqrt{\varepsilon}\int_0^t\langle\nabla\xi(Z_s^\varepsilon(x))\sigma(X_s^\varepsilon(x))\mathrm{d}W_s,w_s^\varepsilon(x)\rangle\\&=:I_{31}^\varepsilon(t,x)+I_{32}^\varepsilon(t,x)+I_{33}^\varepsilon(t,x)+I_{34}^\varepsilon(t,x)+I_{35}^\varepsilon(t,x).\end{aligned}$$

Noticing that

$$\partial_i\xi^k(x)=\frac{1_{i=k}}{|x|^2+\delta^2}-\frac{2x^ix^k}{(|x|^2+\delta^2)^2}$$

and

$$\partial_{ij}^2\xi^k(x)=-\frac{2\cdot 1_{i=k}x^j}{(|x|^2+\delta^2)^2}+\frac{4x^ix^jx^k}{(|x|^2+\delta^2)^3},$$

we have

$$|\xi(x)|\leqslant\frac{1}{\delta},\quad|\nabla\xi(x)|\leqslant\frac{2}{\delta^2},\quad|\nabla^2\xi(x)|\leqslant\frac{6}{\delta^3}.$$

Using Lemma 5.2, as above, one finds that

$$\mathbb{E}\int_{B_N\cap G_R^\varepsilon}\sup_{t\in[0,1]}|I_3^\varepsilon(t,x)|\mu(\mathrm{d}x)\leqslant\frac{C(\varepsilon)}{\delta^3},$$

where $C(\varepsilon)\to 0$ as $\varepsilon\to 0$.

Combining the above estimates, we obtain (5.3). Thus, by (5.3) and Lemma 5.1, as in Step 2 in the proof of Lemma 4.1, there exists a subsequence ε_k such that for $P\otimes\mu$-almost all (ω,x)

$$\sup_{t\in[0,1]}|X_t^{\varepsilon_k}(\omega,x)-X_t^h(\omega,x)|\to 0,\quad\text{as } k\to\infty.$$

Using (i) of Lemma 5.1, there exists another subsequence ε'_k such that $X^{\varepsilon'_k}$ converges to X^h in probability in the space $\mathbb{S}$. □

Proof of Theorem 2.9. Let h^ε be a sequence in $\mathcal{A}_M$ converging to h in distribution. Since $\mathcal{D}_M$ is compact and the law of W is tight, $\{h^\varepsilon,W\}$ is tight in $\mathcal{D}_M\times\Omega$ by the definition of tightness. Without loss of generality, we assume that the law of $\{h^\varepsilon,W\}$ weakly converges to some $\mathbb{P}$ on $\mathcal{D}_M\times\Omega$. Then the law of h is just $\mathbb{P}(\cdot,\Omega)$. By Skorokhod's representation theorem, there are a probability space $(\tilde\Omega,\tilde{\mathscr{F}},\tilde P)$, and random elements $\{\tilde h^\varepsilon,\tilde W^\varepsilon\}$ and $\{\tilde h,\tilde W\}$ in $\mathcal{D}_M\times\Omega$ such that

(1) $(\tilde h^\varepsilon,\tilde W^\varepsilon)$ a.s. converges to $(\tilde h,\tilde W)$;

(2) $(\tilde h^\varepsilon,\tilde W^\varepsilon)$ has the same law as (h^ε,W);

(3) The law of $\{\tilde h,\tilde W\}$ is $\mathbb{P}$, and the law of h is the same as that of $\tilde h$.

Using Lemma 5.3, we get for some subsequence ε_k,

$$\Phi_{\varepsilon_k}\Big(\tilde{W}^{\varepsilon_k}_{\cdot} + \frac{1}{\sqrt{\varepsilon_k}}\int_0^{\cdot} \tilde{h}^{\varepsilon_k}_s \mathrm{d}s\Big) \to X^{\tilde{h}}, \quad \text{in probability.}$$

From this, we derive

$$\Phi_{\varepsilon_k}\Big(W_{\cdot} + \frac{1}{\sqrt{\varepsilon_k}}\int_0^{\cdot} h^{\varepsilon_k}_s \mathrm{d}s\Big) \to X^{h}, \quad \text{in distribution.}$$

Thus, $(\mathrm{LD})_1$ holds. $(\mathrm{LD})_2$ can be simply verified as in Lemma 5.3. □

Acknowledgements. The author is very grateful to Professor Nicolas Privault for providing him an opportunity to work in the City University of Hong Kong. This work was done during his very pleasant stay in Hong Kong. Deep thanks go also to the referee for his/her very kind report.

References

[1] Ambrosio, L.: Transport equation and Cauchy problem for BV vector fields. *Invent. Math.* **158** (2004), no. 2, 227–260.

[2] Ambrosio, L. and Figalli, A.: On flows associated to Sobolev vector fields in Wiener spaces: an approach á la DiPerna–Lions. *J. Funct. Anal.* **256** (2009), 179–214.

[3] Boué, M. and Dupuis, P.: A variational representation for certain functionals of Brownian motion. *Ann. Probab.* **26** (1998), no. 4, 1641–1659.

[4] Boué, M., Dupuis, P. and Ellis, R. S.: Large deviations for small noise diffusions with discontinuous statistics. *Probab. Theory Related Fields* **116** (2000), 125–149.

[5] Budhiraja, A. and Dupuis, P.: A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Statist.* **20** (2000), no. 1, *Acta Univ. Wratislav. No. 2246*, 39–61.

[6] Budhiraja, A., Dupuis, P. and Maroulas, V.: Large deviations for infinite dimensional stochastic dynamical systems. *Ann. Probab.* **36** (2008), no. 4, 1390–1420.

[7] Cruzeiro, A. B.: Équations différentielles ordinaires: non explosion et mesures quasi-invariantes. *J. Funct. Anal.* **54** (1983), no. 2, 193–205.

[8] Cipriano, F. and Cruzeiro, A. B.: Flows associated with irregular $\mathbb{R}^d$-vector fields. *J. Differential Equations* **219** (2005), no. 1, 183–201.

[9] Crippa G. and De Lellis C.: Estimates and regularity results for the DiPerna–Lions flow. *J. Reine Angew. Math.* **616** (2008), 15–46.

[10] DiPerna R. J. and Lions P.-L.: Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989), no. 3, 511–547.

[11] Dupuis, P. and Ellis, R. S.: *A weak convergence approach to the theory of large deviations.* Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, New York, 1997.

[12] Fang, S. and Luo, D.: Transport equations and quasi-invariant flows on the Wiener space. *Bull. Sci. Math.* **134** (2010), no. 3, 295–328.

[13] Fang, S., Luo, D. and Thalmaier, A.: Stochastic differential equations with coefficients in Sobolev spaces. *J. Funct. Anal.* **259** (2010), no. 5, 1129–1168.

[14] Figalli, A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.* **254** (2008), no. 1, 109–153.

[15] Gyöngy I. and Martínez, T.: On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.* **51** (2001), no. 4, 763–783.

[16] Ikeda, N. and Watanabe S.: *Stochastic differential equations and diffusion processes.* North-Holland/Kodansha, Amsterdam-New York/Tokyo, 1981.

[17] Kunita, H.: *Stochastic flows and stochastic differential equations.* Cambridge Studies in Advanced Mathematics 24, Cambridge University Press, Cambridge, 1990.

[18] Krylov, N. V. and Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2005), 154–196.

[19] Le Bris, C. and Lions, P.-L.: Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications. *Ann. Mat. Pura Appl. (4)* **183** (2004), no. 1, 97–130.

[20] Le Bris, C. and Lions, P.-L.: Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Comm. Partial Differential Equations* **33** (2008), no. 7-9, 1272–1317.

[21] Malliavin, P.: *Stochastic analysis.* Fundamental Principles of Mathematical Sciences 313, Springer-Verlag, Berlin, 1997.

[22] Ren, J., Xu, S. and Zhang, X.: Large deviations for multivalued stochastic differential equations. *J. Theoret. Probab.* **23** (2010), no. 4, 1142–1156.

[23] Ren, J. and Zhang, X.: Limit theorems for stochastic differential equations with discontinuous coefficients. *SIAM J. Math. Anal.* **43** (2011), no. 1, 302–321.

[24] Röckner, M. and Zhang, X.: Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris* **348** (2010), no. 7-8, 435–438.

[25] Röckner, M., Schmuland, B. and Zhang, X.: Yamada–Watanabe theorem for stochastic evolution equations in infinite dimensions. *Condensed Matter Physics* **11** (2008), no. 2(54), 247–259.

[26] Stein, E. M.: *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series 30, Princeton University Press, Princeton, NJ, 1970.

[27] Zhang, X.: Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.* **115** (2005), no. 11, 1805–1818.

[28] Zhang, X.: Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math.* **134** (2010), no. 4, 340–378.

[29] Zhang, X.: A variational representation for random functionals on abstract Wiener spaces. *J. Math. Kyoto Univ.* **49** (2009), no. 3, 475–490.

[30] Zvonkin, A. K.: A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)* **93** (1974), no. 1, 129–149.

Received December 18, 2010.

Xicheng Zhang: School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P. R. China.
E-mail: XichengZhang@gmail.com

The support of the NSF of China (grants no. 10971076 and 11271294) is acknowledged.