ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 259 (2010) 1129–1168

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Stochastic differential equations with coefficients in Sobolev spaces

Shizan Fang [a,*], Dejun Luo [b,c], Anton Thalmaier [b]

[a] I.M.B., BP 47870, Université de Bourgogne, Dijon, France

[b] UR Mathématiques, Université du Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg

[c] Key Laboratory of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

Received 3 February 2010; accepted 24 February 2010

Available online 16 March 2010

Communicated by Paul Malliavin

Abstract

We consider the Itô stochastic differential equation $\mathrm{d}X_t = \sum_{j=1}^m A_j(X_t)\,\mathrm{d}w_t^j + A_0(X_t)\,\mathrm{d}t$ on $\mathbb{R}^d$. The diffusion coefficients $A_1, \ldots, A_m$ are supposed to be in the Sobolev space $W_{\mathrm{loc}}^{1,p}(\mathbb{R}^d)$ with $p > d$, and to have linear growth. For the drift coefficient A_0, we distinguish two cases: (i) A_0 is a continuous vector field whose distributional divergence $\delta(A_0)$ with respect to the Gaussian measure γ_d exists, (ii) A_0 has Sobolev regularity $W_{\mathrm{loc}}^{1,p'}$ for some $p' > 1$. Assume $\int_{\mathbb{R}^d} \exp[\lambda_0(|\delta(A_0)| + \sum_{j=1}^m (|\delta(A_j)|^2 + |\nabla A_j|^2))]\,\mathrm{d}\gamma_d < +\infty$ for some $\lambda_0 > 0$. In case (i), if the pathwise uniqueness of solutions holds, then the push-forward $(X_t)_{\#}\gamma_d$ admits a density with respect to γ_d. In particular, if the coefficients are bounded Lipschitz continuous, then X_t leaves the Lebesgue measure Leb_d quasi-invariant. In case (ii), we develop a method used by G. Crippa and C. De Lellis for ODE and implemented by X. Zhang for SDE, to establish existence and uniqueness of stochastic flow of maps.

Keywords: Stochastic flows; Sobolev space coefficients; Density; Density estimate; Pathwise uniqueness; Gaussian measure; Ornstein–Uhlenbeck semigroup

* Corresponding author.
E-mail address: fang@u-bourgogne.fr (S. Fang).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2010.02.014

1. Introduction

Let $A_0, A_1, \dots, A_m : \mathbb{R}^d \to \mathbb{R}^d$ be continuous vector fields on $\mathbb{R}^d$. We consider the following Itô stochastic differential equation on $\mathbb{R}^d$ (abbreviated as SDE)

$$\mathrm{d}X_t = \sum_{j=1}^{m} A_j(X_t)\,\mathrm{d}w_t^j + A_0(X_t)\,\mathrm{d}t, \quad X_0 = x, \tag{1.1}$$

where $w_t = (w_t^1, \dots, w_t^m)$ is the standard Brownian motion on $\mathbb{R}^m$. It is a classical fact in the theory of SDE (see [16,17,21,30]) that, if the coefficients A_j are globally Lipschitz continuous, then SDE (1.1) has a unique strong solution which defines a stochastic flow of homeomorphisms on $\mathbb{R}^d$; however contrary to ordinary differential equations (abbreviated as ODE), the regularity of the homeomorphisms is only Hölder continuity of order $0 < \alpha < 1$. Thus it is not clear whether the Lebesgue measure Leb_d on $\mathbb{R}^d$ admits a density under the flow X_t. In the case where the vector fields A_j, $j = 0, 1, \dots, m$, are in $C_b^\infty(\mathbb{R}^d, \mathbb{R}^d)$, the SDE (1.1) defines a flow of diffeomorphisms, and Kunita [21] showed that the measures on $\mathbb{R}^d$ which have a strictly positive smooth density with respect to Leb_d are quasi-invariant under the flow. This result was recently generalized in [27] to the case where the drift A_0 is allowed to be only log-Lipschitz continuous. Studies on SDE beyond the Lipschitz setting attracted great interest during the last years, see for instance [10,13,12,19,20,23,24,29,34,35].

In the context of ODE, existence of a flow of quasi-invariant measurable maps associated to a vector field A_0 belonging to Sobolev spaces appeared first in [6]. In the seminal paper [7], Di Perna and Lions developed transport equations to solve ODE without involving exponential integrability of $|\nabla A_0|$. On the other hand, L. Ambrosio [1] took advantage of using continuity equations which allowed him to construct quasi-invariant flows associated to vector fields A_0 with only BV regularity. In the framework for Gaussian measures, the Di Perna–Lions method was developed in [4], also in [2,11] on the Wiener space.

The situation for SDE is quite different: even for vector fields $A_0, A_1, \dots, A_m$ in C^∞ with linear growth, if no conditions were imposed on the growth of the derivatives, the SDE (1.1) may not define a flow of diffeomorphisms (see [25,26]). More precisely, let τ_x be the life time of the solution to (1.1) starting from x. The SDE (1.1) is said to be *complete* if for each $x \in \mathbb{R}^d$, $\mathbb{P}(\tau_x = +\infty) = 1$; it is said to be *strongly complete* if $\mathbb{P}(\tau_x = +\infty,\ x \in \mathbb{R}^d) = 1$. The goal in [26] is to construct examples for which the coefficients are smooth, but such that the SDE (1.1) is not strongly complete (see [13,25] for positive examples). Now consider

$$\Sigma = \big\{(w, x) \in \Omega \times \mathbb{R}^d:\ \tau_x(w) = +\infty\big\}.$$

Suppose that SDE (1.1) is complete, then for any probability measure μ on $\mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \bigg(\int_{\Omega} \mathbf{1}_\Sigma(w, x)\,\mathrm{d}\mathbb{P}(w) \bigg)\,\mathrm{d}\mu(x) = 1.$$

Thus, by Fubini's theorem, $\int_\Omega (\int_{\mathbb{R}^d} \mathbf{1}_\Sigma(w, x)\,\mathrm{d}\mu(x))\,\mathrm{d}\mathbb{P}(w) = 1$. It follows that there exists a full measure subset $\Omega_0 \subset \Omega$ such that for all $w \in \Omega_0$, $\tau_x(w) = +\infty$ holds for μ-almost every $x \in \mathbb{R}^d$.

Now under the existence of a complete unique strong solution to SDE (1.1), we have a flow of measurable maps $x \to X_t(w, x)$.

Recently, inspired by previous work due to Ambrosio, Lecumberry and Maniglia [3], Crippa and De Lellis [5] obtained some new type of estimates of perturbation for ODE whose coefficients have Sobolev regularity. More precisely, the absence of Lipschitz condition was filled by the following inequality: for $f \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$,

$$\left|f(x) - f(y)\right| \leqslant C_d|x - y|\big(M_R|\nabla f|(x) + M_R|\nabla f|(y)\big)$$

holds for $x, y \in N^c$ and $|x - y| \leqslant R$, where N is a negligible set of $\mathbb{R}^d$ and $M_R g$ is the maximal function defined by

$$M_R g(x) = \sup_{0<r\leqslant R} \frac{1}{\mathrm{Leb}_d(B(x,r))} \int_{B(x,r)} \left|g(y)\right| \mathrm{d}y,$$

where $B(x,r) = \{y \in \mathbb{R}^d\colon\ |y - x| \leqslant r\}$; the classical moment estimate is replaced by estimating the quantity

$$\int_{B(0,r)} \log\left(\frac{|X_t(x) - \tilde{X}_t(x)|}{\sigma} + 1\right) \mathrm{d}x,$$

where $\sigma > 0$ is a small parameter. This method has recently been successfully implemented to SDE by X. Zhang in [36].

The aim in this paper is two-fold: first we shall study absolute continuity of the push-forward measure $(X_t)_\# \mathrm{Leb}_d$ with respect to Leb_d, once the SDE (1.1) has a unique strong solution; secondly we shall construct strong solutions (for almost all initial values) using the approach mentioned above for SDE with coefficients in Sobolev space. The key point is to obtain an *a priori* L^p estimate for the density. To this end, we shall work with the standard Gaussian measure γ_d; this will be done in Section 2. The main result in Section 3 is the following

Theorem 1.1. *Let $A_0, A_1, \ldots, A_m$ be continuous vector fields on $\mathbb{R}^d$ of linear growth. Assume that the diffusion coefficients $A_1, \ldots, A_m$ are in the Sobolev space $\bigcap_{q>1} \mathbb{D}^q_1(\gamma_d)$ and that $\delta(A_0)$ exists; furthermore there exists a constant $\lambda_0 > 0$ such that*

$$\int_{\mathbb{R}^d} \exp\left[\lambda_0\left(\left|\delta(A_0)\right| + \sum_{j=1}^m \left(\left|\delta(A_j)\right|^2 + |\nabla A_j|^2\right)\right)\right] \mathrm{d}\gamma_d < +\infty. \tag{1.2}$$

Suppose that pathwise uniqueness holds for SDE (1.1). *Then $(X_t)_\#\gamma_d$ is absolutely continuous with respect to γ_d and the density is in $L^1 \log L^1$.*

A consequence of this theorem concerns the following classical situation.

Theorem 1.2. *Let $A_0, A_1, \dots, A_m$ be globally Lipschitz continuous. Suppose that there exists a constant $C > 0$ such that*

$$\sum_{j=1}^{m} \langle x, A_j(x) \rangle^2 \leqslant C\left(1 + |x|^2\right) \quad \text{for all } x \in \mathbb{R}^d. \tag{1.3}$$

Then the stochastic flow of homeomorphisms X_t generated by SDE (1.1) *leaves the Lebesgue measure* Leb_d *quasi-invariant.*

Remark that condition (1.3) not only includes the case of bounded Lipschitz diffusion coefficients, but also, maybe more significant, indicates the role of dispersion: the vector fields $A_1, \dots, A_m$ should not go radially to infinity. The purpose of Section 4 is to find conditions that guarantee strict positivity of the density, in case where existence of the inverse flow is not known, see Theorem 4.4.

The main result of Section 5 is

Theorem 1.3. *Assume that the diffusion coefficients $A_1, \dots, A_m$ belong to the Sobolev space $\bigcap_{q>1} \mathbb{D}_1^q(\gamma_d)$ and the drift $A_0 \in \mathbb{D}_1^q(\gamma_d)$ for some $q > 1$. Assume condition* (1.2) *and that the coefficients $A_0, A_1, \dots, A_m$ are of linear growth, then there is a unique stochastic flow of measurable maps $X : [0, T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$, which solves* (1.1) *for almost all initial $x \in \mathbb{R}^d$ and the push-forward $(X_t(w, \cdot))_{\#}\gamma_d$ admits a density with respect to γ_d, which is in $L^1 \log L^1$.*

When the diffusion coefficients satisfy uniform ellipticity, a classical result due to Stroock and Varadhan [32] says that if the diffusion coefficients $A_1, \dots, A_m$ are bounded continuous and the drift A_0 is bounded Borel measurable, then weak uniqueness holds, that is uniqueness in law of the diffusion. This result was strengthened by Veretennikov [33], saying that in fact pathwise uniqueness holds. When A_0 is not bounded, some conditions on the diffusion coefficients were needed. In the case where the diffusion matrix $a = (a_{ij})$ is the identity, the drift A_0 in (1.1) can be quite singular: $A_0 \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$ with $p > d + 2$ implies that SDE (1.1) has the pathwise uniqueness (see Krylov and Röckner [20] for a more complete study); if the diffusion coefficients $A_1, \dots, A_m$ are bounded continuous, under a Sobolev condition, namely, $A_j \in W^{1,2(d+1)}_{\mathrm{loc}}$ for $j = 1, \dots, m$ and $A_0 \in L^{2(d+1)}_{\mathrm{loc}}(\mathbb{R}^d)$, X. Zhang proved in [34] that SDE (1.1) admits a unique strong solution. Note that even in this uniformly non-degenerated case, if the diffusion coefficients lose the continuity, there are counterexamples for which weak uniqueness does not hold, see [19,31].

Finally we would like to mention that under weaker Sobolev type conditions, the connection between weak solutions and Fokker–Planck equations has been investigated in [14,22]; some notions of "generalized solutions", as well as the phenomena of coalescence and splitting, have been explored in [23,24]. Stochastic transport equations are studied in [15,36].

2. L^p estimate of the density

The purpose of this section is to derive *a priori* estimates for the density of the push-forwards under the flow. We assume that the coefficients $A_0, A_1, \dots, A_m$ of SDE (1.1) are *smooth with compact support* in $\mathbb{R}^d$. Then the solution X_t, i.e., $x \mapsto X_t(x)$, is a stochastic flow of diffeomorphisms on $\mathbb{R}^d$. Moreover SDE (1.1) is equivalent to the following Stratonovich SDE

$$\mathrm{d}X_t = \sum_{j=1}^{m} A_j(X_t) \circ \mathrm{d}w_t^j + \tilde{A}_0(X_t)\,\mathrm{d}t, \quad X_0 = x, \tag{2.1}$$

where $\tilde{A}_0 = A_0 - \frac{1}{2}\sum_{j=1}^{m} \mathcal{L}_{A_j} A_j$ and $\mathcal{L}_A$ denotes the Lie derivative with respect to A.

Let γ_d be the standard Gaussian measure on $\mathbb{R}^d$, and $\gamma_t = (X_t)_\# \gamma_d$, $\tilde{\gamma}_t = (X_t^{-1})_\# \gamma_d$ the push-forwards of γ_d respectively by the flow X_t and its inverse flow X_t^{-1}. To fix ideas, we denote by $(\Omega, \mathscr{F}, \mathbb{P})$ the probability space on which the Brownian motion w_t is defined. Let $K_t = \frac{\mathrm{d}\gamma_t}{\mathrm{d}\gamma_d}$ and $\tilde{K}_t = \frac{\mathrm{d}\tilde{\gamma}_t}{\mathrm{d}\gamma_d}$ be the densities with respect to γ_d. By Lemma 4.3.1 in [21], the Radon–Nikodym derivative $\tilde{K}_t$ has the following explicit expression

$$\tilde{K}_t(x) = \exp\left(-\sum_{j=1}^{m} \int_0^t \delta(A_j)\big(X_s(x)\big) \circ \mathrm{d}w_s^j - \int_0^t \delta(\tilde{A}_0)\big(X_s(x)\big)\,\mathrm{d}s \right), \tag{2.2}$$

where $\delta(A_j)$ denotes the divergence of A_j with respect to the Gaussian measure γ_d:

$$\int_{\mathbb{R}^d} \langle \nabla\varphi, A_j\rangle\,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} \varphi\delta(A_j)\,\mathrm{d}\gamma_d, \quad \varphi \in C_c^1\big(\mathbb{R}^d\big).$$

It is easy to see that K_t and $\tilde{K}_t$ are related to each other by the equality below:

$$K_t(x) = \big[\tilde{K}_t\big(X_t^{-1}(x)\big)\big]^{-1}. \tag{2.3}$$

In fact, for any $\psi \in C_c^\infty(\mathbb{R}^d)$, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \psi(x)\,\mathrm{d}\gamma_d(x) &= \int_{\mathbb{R}^d} \psi\big[X_t\big(X_t^{-1}(x)\big)\big]\,\mathrm{d}\gamma_d(x)\\ &= \int_{\mathbb{R}^d} \psi\big[X_t(y)\big]\tilde{K}_t(y)\,\mathrm{d}\gamma_d(y)\\ &= \int_{\mathbb{R}^d} \psi(x)\tilde{K}_t\big(X_t^{-1}(x)\big)K_t(x)\,\mathrm{d}\gamma_d(x),\end{aligned}$$

which leads to (2.3) due to the arbitrariness of $\psi \in C_c^\infty(\mathbb{R}^d)$. In the following we shall estimate the $L^p(\mathbb{P} \times \gamma_d)$ norm of K_t.

We rewrite the density (2.2) with the Itô integral:

$$\tilde{K}_t(x) = \exp\left(-\sum_{j=1}^{m} \int_0^t \delta(A_j)\big(X_s(x)\big)\,\mathrm{d}w_s^j - \int_0^t \left[\frac{1}{2}\sum_{j=1}^{m} \mathcal{L}_{A_j}\delta(A_j) + \delta(\tilde{A}_0)\right]\big(X_s(x)\big)\,\mathrm{d}s \right). \tag{2.4}$$

Lemma 2.1. *We have*

$$\frac{1}{2}\sum_{j=1}^{m}\mathcal{L}_{A_j}\delta(A_j)+\delta(\tilde{A}_0)=\delta(A_0)+\frac{1}{2}\sum_{j=1}^{m}|A_j|^2+\frac{1}{2}\sum_{j=1}^{m}\langle\nabla A_j,(\nabla A_j)^*\rangle, \tag{2.5}$$

where $\langle\cdot,\cdot\rangle$ *denotes the inner product of* $\mathbb{R}^d\otimes\mathbb{R}^d$ *and* $(\nabla A_j)^*$ *the transpose of* ∇A_j.

Proof. Let A be a C^2 vector field on $\mathbb{R}^d$. From the expression

$$\delta(A)=\sum_{k=1}^{d}\left(x_kA^k-\frac{\partial A^k}{\partial x_k}\right),$$

we get

$$\mathcal{L}_A\delta(A)=\sum_{\ell,k=1}^{d}\left(A^\ell A^k\delta_{k\ell}+A^\ell x_k\frac{\partial A^k}{\partial x_\ell}-A^\ell\frac{\partial^2A^k}{\partial x_\ell\partial x_k}\right). \tag{2.6}$$

Note that

$$\frac{\partial}{\partial x_k}\left(A^\ell\frac{\partial A^k}{\partial x_\ell}\right)=\frac{\partial A^k}{\partial x_\ell}\frac{\partial A^\ell}{\partial x_k}+A^\ell\frac{\partial^2A^k}{\partial x_k\partial x_\ell}.$$

Thus, by means of (2.6), we obtain

$$\mathcal{L}_A\delta(A)=|A|^2+\delta(\mathcal{L}_AA)+\langle\nabla A,(\nabla A)^*\rangle. \tag{2.7}$$

Recall that $\delta(\tilde{A}_0)=\delta(A_0)-\frac{1}{2}\sum_{j=1}^m\delta(\mathcal{L}_{A_j}A_j)$. Hence, replacing A by A_j in (2.7) and summing over j, gives formula (2.5). □

We can now prove the following key estimate.

Theorem 2.2. *For* $p>1$,

$$\|K_t\|_{L^p(\mathbb{P}\times\gamma_d)} \leqslant\left[\int_{\mathbb{R}^d}\exp\left(pt\left[2\left|\delta(A_0)\right|+\sum_{j=1}^{m}\left(|A_j|^2+|\nabla A_j|^2+2(p-1)\left|\delta(A_j)\right|^2\right)\right]\right)\mathrm{d}\gamma_d\right]^{\frac{p-1}{p(2p-1)}}. \tag{2.8}$$

Proof. Using relation (2.3), we have

$$\int_{\mathbb{R}^d} \mathbb{E}\big[K_t^p(x)\big]\,\mathrm{d}\gamma_d(x) = \mathbb{E}\int_{\mathbb{R}^d} \big[\tilde{K}_t\big(X_t^{-1}(x)\big)\big]^{-p}\,\mathrm{d}\gamma_d(x)$$
$$= \mathbb{E}\int_{\mathbb{R}^d} \big[\tilde{K}_t(y)\big]^{-p}\tilde{K}_t(y)\,\mathrm{d}\gamma_d(y)$$
$$= \int_{\mathbb{R}^d} \mathbb{E}\big[\big(\tilde{K}_t(x)\big)^{-p+1}\big]\,\mathrm{d}\gamma_d(x). \tag{2.9}$$

To simplify the notation, denote the right-hand side of (2.5) by Φ. Then $\tilde{K}_t(x)$ rewrites as

$$\tilde{K}_t(x) = \exp\left(-\sum_{j=1}^m \int_0^t \delta(A_j)\big(X_s(x)\big)\,\mathrm{d}w_s^j - \int_0^t \Phi\big(X_s(x)\big)\,\mathrm{d}s\right).$$

Fixing an arbitrary $r > 0$, we get

$$\big(\tilde{K}_t(x)\big)^{-r} = \exp\left(r\sum_{j=1}^m \int_0^t \delta(A_j)\big(X_s(x)\big)\,\mathrm{d}w_s^j + r\int_0^t \Phi\big(X_s(x)\big)\,\mathrm{d}s\right)$$
$$= \exp\left(r\sum_{j=1}^m \int_0^t \delta(A_j)\big(X_s(x)\big)\,\mathrm{d}w_s^j - r^2\sum_{j=1}^m \int_0^t \big|\delta(A_j)\big(X_s(x)\big)\big|^2\,\mathrm{d}s\right)$$
$$\times \exp\left(\int_0^t \left(r^2\sum_{j=1}^m \big|\delta(A_j)\big|^2 + r\Phi\right)\big(X_s(x)\big)\,\mathrm{d}s\right).$$

By Cauchy–Schwarz's inequality,

$$\mathbb{E}\big[\big(\tilde{K}_t(x)\big)^{-r}\big] \leqslant \left[\mathbb{E}\exp\left(2r\sum_{j=1}^m \int_0^t \delta(A_j)\big(X_s(x)\big)\,\mathrm{d}w_s^j - 2r^2\sum_{j=1}^m \int_0^t \big|\delta(A_j)\big(X_s(x)\big)\big|^2\,\mathrm{d}s\right)\right]^{1/2}$$
$$\times \left[\mathbb{E}\exp\left(\int_0^t \left(2r^2\sum_{j=1}^m \big|\delta(A_j)\big|^2 + 2r\Phi\right)\big(X_s(x)\big)\,\mathrm{d}s\right)\right]^{1/2}$$
$$= \left[\mathbb{E}\exp\left(\int_0^t \left(2r^2\sum_{j=1}^m \big|\delta(A_j)\big|^2 + 2r\Phi\right)\big(X_s(x)\big)\,\mathrm{d}s\right)\right]^{1/2}, \tag{2.10}$$

since the first term on the right-hand side of the inequality in (2.10) is the expectation of a martingale. Let

$$\tilde{\Phi}_r = 2r\left|\delta(A_0)\right| + r\sum_{j=1}^{m}\left(|A_j|^2 + |\nabla A_j|^2 + 2r\left|\delta(A_j)\right|^2\right).$$

Then by (2.10), along with the definition of Φ and Cauchy–Schwarz's inequality, we obtain

$$\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\tilde{K}_t(x)\right)^{-r}\right] \mathrm{d}\gamma_d \leqslant \left[\int_{\mathbb{R}^d} \mathbb{E}\exp\left(\int_0^t \tilde{\Phi}_r\left(X_s(x)\right) \mathrm{d}s\right) \mathrm{d}\gamma_d\right]^{1/2}. \tag{2.11}$$

Following the idea of A.B. Cruzeiro ([6, Corollary 2.2], see also Theorem 7.3 in [8]) and by Jensen's inequality,

$$\exp\left(\int_0^t \tilde{\Phi}_r\left(X_s(x)\right) \mathrm{d}s\right) = \exp\left(\int_0^t t\tilde{\Phi}_r\left(X_s(x)\right) \frac{\mathrm{d}s}{t}\right) \leqslant \frac{1}{t}\int_0^t e^{t\tilde{\Phi}_r(X_s(x))}\,\mathrm{d}s.$$

Define $I(t) = \sup_{0\leqslant s\leqslant t}\int_{\mathbb{R}^d} \mathbb{E}[K_t^p(x)]\,\mathrm{d}\gamma_d$. Integrating on both sides of the above inequality and by Hölder's inequality,

$$\begin{aligned}
\int_{\mathbb{R}^d} \mathbb{E}\exp\left(\int_0^t \tilde{\Phi}_r\left(X_s(x)\right) \mathrm{d}s\right) \mathrm{d}\gamma_d(x) &\leqslant \frac{1}{t}\int_0^t \mathbb{E}\int_{\mathbb{R}^d} e^{t\tilde{\Phi}_r(X_s(x))}\,\mathrm{d}\gamma_d(x)\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t \mathbb{E}\int_{\mathbb{R}^d} e^{t\tilde{\Phi}_r(y)} K_s(y)\,\mathrm{d}\gamma_d(y)\,\mathrm{d}s \\
&\leqslant \frac{1}{t}\int_0^t \left\|e^{t\tilde{\Phi}_r}\right\|_{L^q(\gamma_d)} \|K_s\|_{L^p(\mathbb{P}\times\gamma_d)}\,\mathrm{d}s \\
&\leqslant \left\|e^{t\tilde{\Phi}_r}\right\|_{L^q(\gamma_d)} I(t)^{1/p},
\end{aligned}$$

where q is the conjugate number of p. Thus it follows from (2.11) that

$$\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\tilde{K}_t(x)\right)^{-r}\right] \mathrm{d}\gamma_d(x) \leqslant \left\|e^{t\tilde{\Phi}_r}\right\|_{L^q(\gamma_d)}^{1/2} I(t)^{1/2p}. \tag{2.12}$$

Taking $r = p-1$ in the above estimate and by (2.9), we obtain

$$\int_{\mathbb{R}^d} \mathbb{E}\left[K_t^p(x)\right] \mathrm{d}\gamma_d(x) \leqslant \left\|e^{t\tilde{\Phi}_{p-1}}\right\|_{L^q(\gamma_d)}^{1/2} I(t)^{1/2p}.$$

Thus we have $I(t) \leqslant \|e^{t\tilde{\Phi}_{p-1}}\|_{L^q(\gamma_d)}^{1/2} I(t)^{1/2p}$. Solving this inequality for $I(t)$ gives

$$\int_{\mathbb{R}^d} \mathbb{E}\big[K_t^p(x)\big]\,\mathrm{d}\gamma_d(x) \leqslant I(t) \leqslant \Bigg[\int_{\mathbb{R}^d} \exp\bigg(\frac{pt}{p-1}\tilde{\Phi}_{p-1}(x)\bigg)\,\mathrm{d}\gamma_d(x)\Bigg]^{\frac{p-1}{2p-1}}.$$

Now the desired estimate follows from the definition of $\tilde{\Phi}_{p-1}$. □

Corollary 2.3. *For any* $p > 1$,

$$\|\tilde{K}_t\|_{L^p(\mathbb{P}\times\gamma_d)} \leqslant \Bigg[\int_{\mathbb{R}^d} \exp\Bigg((p+1)t\Bigg[2\big|\delta(A_0)\big| + \sum_{j=1}^m \big(|A_j|^2 + |\nabla A_j|^2 + 2p\big|\delta(A_j)\big|^2\big)\Bigg]\Bigg)\,\mathrm{d}\gamma_d\Bigg]^{\frac{1}{2p+1}}. \tag{2.13}$$

Proof. Similar to (2.12), we have for $r > 0$,

$$\int_{\mathbb{R}^d} \mathbb{E}\big[\big(\tilde{K}_t(x)\big)^r\big]\,\mathrm{d}\gamma_d(x) \leqslant \big\|e^{t\tilde{\Phi}_r}\big\|_{L^q(\gamma_d)}^{1/2} I(t)^{1/2p}, \tag{2.14}$$

where $\tilde{\Phi}_r$ and $I(t)$ are defined as above. Since $I(t) \leqslant \|e^{t\tilde{\Phi}_{p-1}}\|_{L^q(\gamma_d)}^{p/(2p-1)}$, by taking $r = p-1$, we get

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathbb{E}\big[\big(\tilde{K}_t(x)\big)^{p-1}\big]\,\mathrm{d}\gamma_d(x) \\
&\quad\leqslant \big\|e^{t\tilde{\Phi}_{p-1}}\big\|_{L^q(\gamma_d)}^{p/(2p-1)} \\
&\quad= \Bigg[\int_{\mathbb{R}^d} \exp\Bigg(pt\Bigg[2\big|\delta(A_0)\big| + \sum_{j=1}^m \big(|A_j|^2 + |\nabla A_j|^2 + 2(p-1)\big|\delta(A_j)\big|^2\big)\Bigg]\Bigg)\,\mathrm{d}\gamma_d\Bigg]^{\frac{p-1}{2p-1}}.
\end{aligned}$$

Replacing p by $p+1$ in the last inequality gives the claimed estimate. □

3. Absolute continuity under flows generated by SDEs

Now assume that the coefficients A_j in SDE (1.1) are *continuous* and of linear growth. Then it is well known that SDE (1.1) has a weak solution of infinite life time. In order to apply the results of the preceding section, we shall regularize the vector fields using the Ornstein–Uhlenbeck semigroup $\{P_\varepsilon\}_{\varepsilon>0}$ on $\mathbb{R}^d$:

$$P_\varepsilon A(x) = \int_{\mathbb{R}^d} A\big(e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}\,y\big)\,\mathrm{d}\gamma_d(y).$$

We have the following simple properties.

Lemma 3.1. *Assume that A is continuous and $|A(x)| \leqslant C(1+|x|^q)$ for some $q \geqslant 0$. Then*

(i) *there is $C_q > 0$ independent of ε, such that*

$$\left|P_\varepsilon A(x)\right| \leqslant C_q\left(1+|x|^q\right), \quad \text{for all } x \in \mathbb{R}^d;$$

(ii) *$P_\varepsilon A$ converges uniformly to A on any compact subset as $\varepsilon \to 0$.*

Proof. (i) Note that $|e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y| \leqslant |x| + |y|$ and that there exists a constant $C > 0$ such that $(|x|+|y|)^q \leqslant C(|x|^q + |y|^q)$. Using the growth condition on A, we have for some constant $C > 0$ (depending on q),

$$\begin{aligned}\left|P_\varepsilon A(x)\right| &\leqslant \int_{\mathbb{R}^d} \left|A\left(e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y\right)\right| \mathrm{d}\gamma_d(y)\\ &\leqslant C\int_{\mathbb{R}^d}\left(1+|x|^q+|y|^q\right)\mathrm{d}\gamma_d(y) \leqslant C\left(1+|x|^q+M_q\right)\end{aligned}$$

where $M_q = \int_{\mathbb{R}^d} |y|^q \,\mathrm{d}\gamma_d(y)$. Changing the constant yields (i).

(ii) Fix $R > 0$ and x in the closed ball $B(R)$ of radius R, centered at 0. Let $R_1 > R$ be arbitrary. We have

$$\begin{aligned}\left|P_\varepsilon A(x) - A(x)\right| &\leqslant \int_{\mathbb{R}^d}\left|A\left(e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y\right) - A(x)\right|\mathrm{d}\gamma_d(y)\\ &= \left(\int_{B(R_1)} + \int_{B(R_1)^c}\right)\left|A\left(e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y\right) - A(x)\right|\mathrm{d}\gamma_d(y)\\ &=: I_1 + I_2.\end{aligned} \tag{3.1}$$

By the growth condition on A, for some constant $C_q > 0$, independent of ε, we have

$$\begin{aligned}I_2 &\leqslant \int_{B(R_1)^c}\left(\left|A\left(e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y\right)\right| + \left|A(x)\right|\right)\mathrm{d}\gamma_d(y)\\ &\leqslant C_q \int_{B(R_1)^c}\left(1+R^q+|y|^q\right)\mathrm{d}\gamma_d(y),\end{aligned}$$

where the last term tends to 0 as $R_1 \to +\infty$. For given $\eta > 0$, we may take R_1 large enough such that $I_2 < \eta$. Then there exists $\varepsilon_{R_1} > 0$ such that for $\varepsilon < \varepsilon_{R_1}$ and $|y| \leqslant R_1$,

$$\left|e^{-\varepsilon}x + \sqrt{1-e^{-2\varepsilon}}y\right| \leqslant e^{-\varepsilon}R + \sqrt{1-e^{-2\varepsilon}}R_1 \leqslant R_1.$$

Note that

$$\left|e^{-\varepsilon}x+\sqrt{1-e^{-2\varepsilon}}\,y-x\right|\leqslant \varepsilon R+\sqrt{2\varepsilon}R_1,\quad \text{for } |x|\leqslant R,\ |y|\leqslant R_1.$$

Since A is uniformly continuous on $B(R_1)$, there exits $\varepsilon_0\leqslant \varepsilon_{R_1}$ such that

$$\left|A\big(e^{-\varepsilon}x+\sqrt{1-e^{-2\varepsilon}}\,y\big)-A(x)\right|\leqslant \eta\quad \text{for all } y\in B(R_1),\ \varepsilon\leqslant\varepsilon_0.$$

As a result, the term $I_1\leqslant\eta$. Therefore by (3.1), for any $\varepsilon\leqslant\varepsilon_0$,

$$\sup_{|x|\leqslant R}\left|P_\varepsilon A(x)-A(x)\right|\leqslant 2\eta.$$

The result follows from the arbitrariness of $\eta>0$. □

The vector field $P_\varepsilon A$ is smooth on $\mathbb{R}^d$ but does not have compact support. We introduce cut-off functions $\varphi_\varepsilon\in C_c^\infty(\mathbb{R}^d,[0,1])$ satisfying

$$\varphi_\varepsilon(x)=1\quad \text{if } |x|\leqslant\frac{1}{\varepsilon},\qquad \varphi_\varepsilon(x)=0\quad \text{if } |x|\geqslant\frac{1}{\varepsilon}+2\quad\text{and}\quad \|\nabla\varphi_\varepsilon\|_\infty\leqslant 1.$$

Set

$$A_j^\varepsilon=\varphi_\varepsilon P_\varepsilon A_j,\quad j=0,1,\ldots,m.$$

Now consider the Itô SDE (1.1) with A_j being replaced by A_j^ε $(j=0,1,\ldots,m)$, and denote the corresponding terms by adding the superscript ε, e.g. X_t^ε, K_t^ε, etc.

In the sequel, we shall give a uniform estimate to K_t^ε. To this end, we need some preparations in the spirit of Malliavin calculus [28]. For a vector field A on $\mathbb{R}^d$ and $p>1$, we say that $A\in\mathbb{D}_1^p(\gamma_d)$ if $A\in L^p(\gamma_d)$ and if there exists $\nabla A:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ in $L^p(\gamma_d)$ such that for any $v\in\mathbb{R}^d$,

$$\nabla A(x)(v)=\partial_v A:=\lim_{\eta\to 0}\frac{A(x+\eta v)-A(x)}{\eta}\quad \text{holds in } L^{p'}(\gamma_d) \text{ for any } p'<p.$$

For such $A\in\mathbb{D}_1^p(\gamma_d)$, the divergence $\delta(A)\in L^p(\gamma_d)$ exists and the following relations hold:

$$\nabla P_\varepsilon A=e^{-\varepsilon}P_\varepsilon(\nabla A),\qquad \delta(P_\varepsilon A)=e^{\varepsilon}P_\varepsilon\big(\delta(A)\big).\tag{3.2}$$

Note that the second term in (3.2) holds once the divergence $\delta(A)\in L^p$ exists for some $p>1$. If $A\in L^p(\gamma_d)$, then $P_\varepsilon A\in\mathbb{D}_1^p(\gamma_d)$ and $\lim_{\varepsilon\to 0}\|P_\varepsilon A-A\|_{L^p}=0$.

Lemma 3.2. *Assume the vector field $A\in L^p(\gamma_d)$ admits the divergence $\delta(A)\in L^p(\gamma_d)$ for $p>1$, and denote by $A^\varepsilon=\varphi_\varepsilon P_\varepsilon A$. Then for $\varepsilon\in\,]0,1]$,*

$$\big|\delta(A^\varepsilon)\big| \leqslant P_\varepsilon\big(|A| + e\big|\delta(A)\big|\big),$$
$$\big|A^\varepsilon\big|^2 \leqslant P_\varepsilon\big(|A|^2\big),$$
$$\big|\delta(A^\varepsilon)\big|^2 \leqslant P_\varepsilon\big[2\big(|A|^2 + e^2\big|\delta(A)\big|^2\big)\big].$$

If furthermore $A \in \mathbb{D}_1^p(\gamma_d)$, *then*

$$\big|\nabla A^\varepsilon\big|^2 \leqslant P_\varepsilon\big[2\big(|A|^2 + |\nabla A|^2\big)\big].$$

Proof. Note that according to (3.2), $\delta(A^\varepsilon) = \delta(\varphi_\varepsilon P_\varepsilon A) = \varphi_\varepsilon e^\varepsilon P_\varepsilon \delta(A) - \langle \nabla\varphi_\varepsilon, P_\varepsilon A\rangle$, from where the first inequality follows. In the same way, the other results are obtained. □

Applying Theorem 2.2 to K_t^ε with $p = 2$, we have

$$\big\|K_t^\varepsilon\big\|_{L^2(\mathbb{P}\times\gamma_d)} \leqslant \left[\int_{\mathbb{R}^d} \exp\left(2t\left[2\big|\delta(A_0^\varepsilon)\big| + \sum_{j=1}^m\big(\big|A_j^\varepsilon\big|^2 + \big|\nabla A_j^\varepsilon\big|^2 + 2\big|\delta(A_j^\varepsilon)\big|^2\big)\right]\right)\mathrm{d}\gamma_d\right]^{1/6}. \tag{3.3}$$

By Lemma 3.2,

$$2\big|\delta(A_0^\varepsilon)\big| + \sum_{j=1}^m\big(\big|A_j^\varepsilon\big|^2 + \big|\nabla A_j^\varepsilon\big|^2 + 2\big|\delta(A_j^\varepsilon)\big|^2\big)$$
$$\leqslant P_\varepsilon\left[2|A_0| + 2e\big|\delta(A_0)\big| + \sum_{j=1}^m\big(7|A_j|^2 + 2|\nabla A_j|^2 + 4e^2\big|\delta(A_j)\big|^2\big)\right].$$

We deduce from Jensen's inequality and the invariance of γ_d under the action of the semigroup P_ε that

$$\big\|K_t^\varepsilon\big\|_{L^2(\mathbb{P}\times\gamma_d)} \leqslant \left[\int_{\mathbb{R}^d} \exp\left(4t\left[|A_0| + e\big|\delta(A_0)\big| + \sum_{j=1}^m\big(4|A_j|^2 + |\nabla A_j|^2 + 2e^2\big|\delta(A_j)\big|^2\big)\right]\right)\mathrm{d}\gamma_d\right]^{1/6} \tag{3.4}$$

for any $\varepsilon \leqslant 1$. According to (3.4), we consider the following conditions.

Assumptions (H).

(A1) For $j = 1, \dots, m$, $A_j \in \bigcap_{q\geqslant 1}\mathbb{D}_1^q(\gamma_d)$, A_0 is continuous and $\delta(A_0)$ exists.
(A2) The vector fields $A_0, A_1, \dots, A_m$ have linear growth.

(A3) There exists $\lambda_0 > 0$ such that

$$\int_{\mathbb{R}^d} \exp\left[\lambda_0\left(\left|\delta(A_0)\right| + \sum_{j=1}^m \left|\delta(A_j)\right|^2\right)\right] d\gamma_d < +\infty.$$

(A4) There exists $\lambda_0 > 0$ such that

$$\int_{\mathbb{R}^d} \exp\left(\lambda_0 \sum_{j=1}^m |\nabla A_j|^2\right) d\gamma_d < +\infty.$$

Note that by Sobolev's embedding theorem, the diffusion coefficients $A_1, \ldots, A_m$ admit Hölder continuous versions. In what follows, we consider these continuous versions. It is clear that under the conditions (A2)–(A4), there exists $T_0 > 0$ small enough, such that

$$\Lambda_{T_0} := \left[\int_{\mathbb{R}^d} \exp\left(4T_0\left[|A_0| + e\left|\delta(A_0)\right| + \sum_{j=1}^m \left(4|A_j|^2 + |\nabla A_j|^2 + 2e^2\left|\delta(A_j)\right|^2\right)\right]\right) d\gamma_d\right]^{1/6} < \infty. \tag{3.5}$$

In this case, for $t \in [0, T_0]$,

$$\sup_{0<\varepsilon\leqslant 1} \left\| K_t^\varepsilon \right\|_{L^2(\mathbb{P}\times\gamma_d)} \leqslant \Lambda_{T_0}. \tag{3.6}$$

Theorem 3.3. *Let $T > 0$ be given. Under* (A1)–(A4) *in Assumptions* (H)*, there are two positive constants C_1 and C_2, independent of ε, such that*

$$\sup_{0<\varepsilon\leqslant 1} \mathbb{E}\int_{\mathbb{R}^d} K_t^\varepsilon \left|\log K_t^\varepsilon\right| d\gamma_d \leqslant 2(C_1 T)^{1/2}\Lambda_{T_0} + C_2 T \Lambda_{T_0}^2, \quad \textit{for all } t \in [0, T].$$

Proof. We follow the arguments of Proposition 4.4 in [11]. By (2.3) and (2.4), we have

$$K_t^\varepsilon\left(X_t^\varepsilon(x)\right) = \left[\tilde{K}_t^\varepsilon(x)\right]^{-1} = \exp\left(\sum_{j=1}^m \int_0^t \delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon(x)\right) dw_s^j + \int_0^t \Phi_\varepsilon\left(X_s^\varepsilon(x)\right) ds\right),$$

where

$$\Phi_\varepsilon = \delta\left(A_0^\varepsilon\right) + \frac{1}{2}\sum_{j=1}^m \left|A_j^\varepsilon\right|^2 + \frac{1}{2}\sum_{j=1}^m \left\langle \nabla A_j^\varepsilon, \left(\nabla A_j^\varepsilon\right)^*\right\rangle.$$

Thus

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^d} K_t^\varepsilon\left|\log K_t^\varepsilon\right| \mathrm{d}\gamma_d &= \mathbb{E}\int_{\mathbb{R}^d}\left|\log K_t^\varepsilon\left(X_t^\varepsilon(x)\right)\right| \mathrm{d}\gamma_d(x)\\ &\leqslant \mathbb{E}\int_{\mathbb{R}^d}\left|\sum_{j=1}^m\int_0^t \delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon(x)\right)\mathrm{d}w_s^j\right| \mathrm{d}\gamma_d(x)\\ &\quad+\mathbb{E}\int_{\mathbb{R}^d}\left|\int_0^t \Phi_\varepsilon\left(X_s^\varepsilon(x)\right)\mathrm{d}s\right| \mathrm{d}\gamma_d(x)\\ &=: I_1+I_2. \end{aligned}\tag{3.7}$$

Using Burkholder's inequality, we get

$$\mathbb{E}\left|\sum_{j=1}^m\int_0^t \delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon(x)\right)\mathrm{d}w_s^j\right| \leqslant 2\mathbb{E}\left[\left(\sum_{j=1}^m\int_0^t \left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon(x)\right)\right|^2 \mathrm{d}s\right)^{1/2}\right].$$

For the sake of simplifying the notations, write $\Psi_\varepsilon=\sum_{j=1}^m|\delta(A_j^\varepsilon)|^2$. By Cauchy's inequality,

$$I_1\leqslant 2\left[\int_0^t \mathbb{E}\int_{\mathbb{R}^d}\left|\Psi_\varepsilon\left(X_s^\varepsilon(x)\right)\right| \mathrm{d}\gamma_d(x)\,\mathrm{d}s\right]^{1/2}.\tag{3.8}$$

Now we are going to estimate $\mathbb{E}\int_{\mathbb{R}^d}|\Psi_\varepsilon(X_s^\varepsilon(x))|^{2^\alpha}\,\mathrm{d}\gamma_d(x)$ for $\alpha\in\mathbb{Z}_+$ which will be done inductively. First if $s\in[0,T_0]$, then by (3.4) and (3.6), along with Cauchy's inequality,

$$\begin{aligned}\mathbb{E}\int_{\mathbb{R}^d}\left|\Psi_\varepsilon\left(X_s^\varepsilon(x)\right)\right|^{2^\alpha}\mathrm{d}\gamma_d(x) &= \mathbb{E}\int_{\mathbb{R}^d}\left|\Psi_\varepsilon(y)\right|^{2^\alpha} K_s^\varepsilon(y)\,\mathrm{d}\gamma_d(y)\\ &\leqslant \|\Psi_\varepsilon\|_{L^{2^{\alpha+1}}(\gamma_d)}^{2^\alpha}\left\|K_s^\varepsilon\right\|_{L^2(\mathbb{P}\times\gamma_d)}\\ &\leqslant \Lambda_{T_0}\|\Psi_\varepsilon\|_{L^{2^{\alpha+1}}(\gamma_d)}^{2^\alpha}. \end{aligned}\tag{3.9}$$

Now for $s\in\,]T_0,2T_0]$, we shall use the flow property of X_s^ε: let $(\theta_{T_0}w)_t:=w_{T_0+t}-w_{T_0}$ and X_t^{ε,T_0} be the solution of the Itô SDE driven by the new Brownian motion $(\theta_{T_0}w)_t$, then

$$X_{T_0+t}^\varepsilon(x,w)=X_t^{\varepsilon,T_0}\left(X_{T_0}^\varepsilon(x,w),\theta_{T_0}w\right),\quad \text{for all } t\geqslant 0,$$

and X_t^{ε,T_0} enjoys the same properties as X_t^ε. Therefore,

$$\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_s^\varepsilon(x)\big)\big|^{2^\alpha}\,\mathrm{d}\gamma_d(x)=\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_{s-T_0}^{\varepsilon,T_0}\big(X_{T_0}^\varepsilon(x)\big)\big)\big|^{2^\alpha}\,\mathrm{d}\gamma_d(x)$$
$$=\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_{s-T_0}^{\varepsilon,T_0}(y)\big)\big|^{2^\alpha}K_{T_0}^\varepsilon(y)\,\mathrm{d}\gamma_d(y)$$

which is dominated, using Cauchy–Schwarz inequality

$$\bigg(\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_{s-T_0}^{\varepsilon,T_0}(y)\big)\big|^{2^{\alpha+1}}\,\mathrm{d}\gamma_d(y)\bigg)^{1/2}\big\|K_{T_0}^\varepsilon\big\|_{L^2(\mathbb{P}\times\gamma_d)}$$
$$\leqslant\big(\Lambda_{T_0}\|\Psi_\varepsilon\|_{L^{2^{\alpha+2}}(\gamma_d)}^{2^{\alpha+1}}\big)^{1/2}\Lambda_{T_0}=\Lambda_{T_0}^{1+2^{-1}}\|\Psi_\varepsilon\|_{L^{2^{\alpha+2}}(\gamma_d)}^{2^\alpha}.$$

Repeating this procedure, we finally obtain, for all $s\in[0,T]$,

$$\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_s^\varepsilon(x)\big)\big|^{2^\alpha}\,\mathrm{d}\gamma_d(x)\leqslant\Lambda_{T_0}^{1+2^{-1}+\cdots+2^{-N+1}}\|\Psi_\varepsilon\|_{L^{2^{\alpha+N}}(\gamma_d)}^{2^\alpha},$$

where $N\in\mathbb{Z}_+$ is the unique integer such that $(N-1)T_0<T\leqslant NT_0$. In particular, taking $\alpha=0$ gives

$$\mathbb{E}\int_{\mathbb{R}^d}\big|\Psi_\varepsilon\big(X_s^\varepsilon(x)\big)\big|\,\mathrm{d}\gamma_d(x)\leqslant\Lambda_{T_0}^2\|\Psi_\varepsilon\|_{L^{2^N}(\gamma_d)}.\tag{3.10}$$

By Lemma 3.2,

$$\sup_{0<\varepsilon\leqslant1}\|\Psi_\varepsilon\|_{L^{2^N}(\gamma_d)}\leqslant\bigg\|2\sum_{j=1}^m\big(|A_j|^2+e^2\big|\delta(A_j)\big|^2\big)\bigg\|_{L^{2^N}(\gamma_d)}=:C_1$$

whose right-hand side is finite under the assumptions (A2)–(A4). This along with (3.8) and (3.10) leads to

$$I_1\leqslant2(C_1T)^{1/2}\Lambda_{T_0}.\tag{3.11}$$

The same manipulation works for the term I_2 and we get

$$I_2\leqslant C_2T\Lambda_{T_0}^2,\tag{3.12}$$

where

$$C_2=\bigg\||A_0|+e\big|\delta(A_0)\big|+\frac{3}{2}\sum_{j=1}^m|A_j|^2+\sum_{j=1}^m|\nabla A_j|^2\bigg\|_{L^{2^N}(\gamma_d)}<\infty.$$

Now we draw the conclusion from (3.7), (3.11) and (3.12). □

It follows from Theorem 3.3 that the family $\{K^\varepsilon_\cdot\}_{0<\varepsilon\leqslant 1}$ is weakly compact in $L^1([0,T]\times\Omega\times\mathbb{R}^d)$. Along a subsequence, $K^\varepsilon_\cdot$ converges weakly to some $K_\cdot\in L^1([0,T]\times\Omega\times\mathbb{R}^d)$ as $\varepsilon\to 0$. Let

$$\mathcal{C}=\left\{u\in L^1\big([0,T]\times\Omega\times\mathbb{R}^d\big):\ u_t\geqslant 0,\ \int_{\mathbb{R}^d}\mathbb{E}[u_t\log u_t]\,\mathrm{d}\gamma_d\leqslant 2(C_1T)^{1/2}\Lambda_{T_0}+C_2T\Lambda_{T_0}^2\right\}.$$

By convexity of the function $s\to s\log s$, it is clear that $\mathcal{C}$ is a convex subset of $L^1([0,T]\times\Omega\times\mathbb{R}^d)$. Since the weak closure of $\mathcal{C}$ coincides with the strong one, there exists a sequence of functions $u^{(n)}\in\mathcal{C}$ which converges to K in $L^1([0,T]\times\Omega\times\mathbb{R}^d)$. Along a subsequence, $u^{(n)}$ converges to K almost everywhere. Hence by Fatou's lemma, we get for almost all $t\in[0,T]$,

$$\int_{\mathbb{R}^d}\mathbb{E}(K_t\log K_t)\,\mathrm{d}\gamma_d\leqslant 2(C_1T)^{1/2}\Lambda_{T_0}+C_2T\Lambda_{T_0}^2. \tag{3.13}$$

Theorem 3.4. *Assume conditions* (A1)–(A4) *and that pathwise uniqueness holds for* SDE (1.1). *Then for each $t>0$, there is a full subset $\Omega_t\subset\Omega$ such that for all $w\in\Omega_t$, the density $\hat{K}_t$ of $(X_t)_\#\gamma_d$ with respect to γ_d exists and $\hat{K}_t\in L^1\log L^1$.*

Proof. Under these assumptions, we can use Theorem A in [18]. For the convenience of the reader, we include the statement:

Theorem 3.5. *(See [18].) Let $\sigma_n(x)$ and $b_n(x)$ be continuous, taking values respectively in the space of* $(d\times m)$*-matrices and* $\mathbb{R}^d$. *Suppose that*

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

and for any $R>0$,

$$\lim_{n\to+\infty}\sup_{|x|\leqslant R}\big(\|\sigma_n(x)-\sigma(x)\|+|b_n(x)-b(x)|\big)=0.$$

Suppose further that for the same Brownian motion $B(t)$, $X_n(x,t)$ solves the SDE

$$\mathrm{d}X_n(t)=\sigma_n\big(X_n(t)\big)\,\mathrm{d}B(t)+b_n\big(X_n(t)\big)\,\mathrm{d}t,\quad X_n(0)=x.$$

If pathwise uniqueness holds for

$$\mathrm{d}X(t)=\sigma\big(X(t)\big)\,\mathrm{d}B(t)+b\big(X(t)\big)\,\mathrm{d}t,\quad X(0)=x,$$

then for any $R>0$, $T>0$,

$$\lim_{n\to+\infty}\sup_{|x|\leqslant R}\mathbb{E}\Big(\sup_{0\leqslant t\leqslant T}\big|X_n(t,x)-X(t,x)\big|^2\Big)=0. \tag{3.14}$$

We continue the proof of Theorem 3.4. By means of Lemma 3.1 and Theorem 3.5, for any $T, R > 0$, we get

$$\lim_{\varepsilon\to 0}\sup_{|x|\leqslant R}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}\big|X_t^\varepsilon(x)-X_t(x)\big|^2\Big]=0. \tag{3.15}$$

Now fixing arbitrary $\xi\in L^\infty(\Omega)$ and $\psi\in C_c^\infty(\mathbb{R}^d)$, we have

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{R}^d}\big|\xi(\cdot)\big|\big|\psi\big(X_t^\varepsilon(x)\big)-\psi\big(X_t(x)\big)\big|\,\mathrm{d}\gamma_d(x)\\
&\quad\leqslant\|\xi\|_\infty\Big(\int_{B(R)}+\int_{B(R)^c}\Big)\mathbb{E}\big|\psi\big(X_t^\varepsilon(x)\big)-\psi\big(X_t(x)\big)\big|\,\mathrm{d}\gamma_d(x)\\
&\quad=:J_1+J_2.
\end{aligned}\tag{3.16}$$

By (3.15),

$$\begin{aligned}
J_1&\leqslant\|\xi\|_\infty\|\nabla\psi\|_\infty\int_{B(R)}\mathbb{E}\big|X_t^\varepsilon(x)-X_t(x)\big|\,\mathrm{d}\gamma_d(x)\\
&\leqslant\|\xi\|_\infty\|\nabla\psi\|_\infty\Big[\sup_{|x|\leqslant R}\mathbb{E}\Big(\sup_{0\leqslant t\leqslant T}\big|X_t^\varepsilon(x)-X_t(x)\big|^2\Big)\Big]^{1/2}\to 0,
\end{aligned}\tag{3.17}$$

as ε tends to 0. It is obvious that

$$J_2\leqslant 2\|\xi\|_\infty\|\psi\|_\infty\gamma_d\big(B(R)^c\big). \tag{3.18}$$

Combining (3.16), (3.17) and (3.18), we obtain

$$\limsup_{\varepsilon\to 0}\mathbb{E}\int_{\mathbb{R}^d}|\xi|\big|\psi\big(X_t^\varepsilon(x)\big)-\psi\big(X_t(x)\big)\big|\,\mathrm{d}\gamma_d(x)\leqslant 2\|\xi\|_\infty\|\psi\|_\infty\gamma_d\big(B(R)^c\big)\to 0$$

as $R\uparrow\infty$. Therefore

$$\lim_{\varepsilon\to 0}\mathbb{E}\int_{\mathbb{R}^d}\xi\psi\big(X_t^\varepsilon(x)\big)\,\mathrm{d}\gamma_d(x)=\mathbb{E}\int_{\mathbb{R}^d}\xi\psi\big(X_t(x)\big)\,\mathrm{d}\gamma_d. \tag{3.19}$$

On the other hand, by Theorem 3.3, for each fixed $t\in[0,T]$, up to a subsequence, K_t^ε converges weakly in $L^1(\Omega\times\mathbb{R}^d)$ to some $\hat{K}_t$, hence

$$\begin{aligned}
\mathbb{E}\int_{\mathbb{R}^d}\xi\psi\big(X_t^\varepsilon(x)\big)\,\mathrm{d}\gamma_d(x)&=\mathbb{E}\int_{\mathbb{R}^d}\xi\psi(y)K_t^\varepsilon(y)\,\mathrm{d}\gamma_d(y)\\
&\to\mathbb{E}\int_{\mathbb{R}^d}\xi\psi(y)\hat{K}_t(y)\,\mathrm{d}\gamma_d(y).
\end{aligned}\tag{3.20}$$

This together with (3.19) leads to

$$\mathbb{E}\int_{\mathbb{R}^d} \xi\psi\big(X_t(x)\big)\,\mathrm{d}\gamma_d(x) = \mathbb{E}\int_{\mathbb{R}^d} \xi\psi(y)\hat{K}_t(y)\,\mathrm{d}\gamma_d(y).$$

By the arbitrariness of $\xi \in L^\infty(\Omega)$, there exists a full measure subset Ω_ψ of Ω such that

$$\int_{\mathbb{R}^d} \psi\big(X_t(x)\big)\,\mathrm{d}\gamma_d(x) = \int_{\mathbb{R}^d} \psi(y)\hat{K}_t(y)\,\mathrm{d}\gamma_d(y), \quad \text{for any } \omega \in \Omega_\psi.$$

Now by the separability of $C_c^\infty(\mathbb{R}^d)$, there exists a full subset Ω_t such that the above equality holds for any $\psi \in C_c^\infty(\mathbb{R}^d)$. Hence $(X_t)_\#\gamma_d = \hat{K}_t\gamma_d$. □

Remark 3.6. The $K_t(w,x)$ appearing in (3.13) is defined almost everywhere. It is easy to see that $K_t(w,x)$ is a measurable modification of $\{\hat{K}_t(w,x);\ t \in [0,T]\}$.

Remark 3.7. Beyond the Lipschitz condition, several sufficient conditions guaranteeing pathwise uniqueness for SDE (1.1) can be found in the literature. For example in [12], the authors give the condition

$$\sum_{j=1}^m \big|A_j(x) - A_j(y)\big|^2 \leqslant C|x-y|^2 r\big(|x-y|^2\big),$$
$$\big|A_0(x) - A_0(y)\big| \leqslant C|x-y| r\big(|x-y|^2\big),$$

for $|x-y| \leqslant c_0$ small enough, where $r :]0,c_0] \to]0,+\infty[$ is C^1 satisfying

(i) $\lim_{s\to 0} r(s) = +\infty$,
(ii) $\lim_{s\to 0} \frac{sr'(s)}{r(s)} = 0$, and
(iii) $\int_0^{c_0} \frac{\mathrm{d}s}{sr(s)} = +\infty$.

Corollary 3.8. *Suppose that the vector fields $A_0, A_1, \ldots, A_m$ are globally Lipschitz continuous and there exists a constant $C > 0$, such that*

$$\sum_{j=1}^m \langle x, A_j(x)\rangle^2 \leqslant C\big(1+|x|^2\big) \quad \textit{for all } x \in \mathbb{R}^d. \tag{3.21}$$

Then $(X_t)_\#\,\mathrm{Leb}_d \ll \mathrm{Leb}_d$ for any $t \in [0,T]$.

Proof. It is obvious that hypotheses (A1), (A2) and (A4) are satisfied, and that for some constant $C > 0$,

$$\big|\delta(A_0)\big|(x) \leqslant C\big(1+|x|^2\big).$$

Hence there exists $\lambda_0 > 0$ such that $\int_{\mathbb{R}^d} \exp(\lambda_0|\delta(A_0)|)\,\mathrm{d}\gamma_d < +\infty$. Finally we have

$$\sum_{j=1}^m \big|\delta(A_j)\big|^2(x) \leqslant 2\sum_{j=1}^m \big\langle x, A_j(x)\big\rangle^2 + 2\sum_{j=1}^m \mathrm{Lip}(A_j)^2.$$

Therefore, under condition (3.21), there exists $\lambda_0 > 0$ such that

$$\int_{\mathbb{R}^d} \exp\Bigg(\lambda_0 \sum_{j=1}^m \big|\delta(A_j)\big|^2\Bigg)\,\mathrm{d}\gamma_d < +\infty.$$

Hence, hypothesis (A3) is satisfied as well. By Theorem 3.4, we have $(X_t)_\#\gamma_d = \hat{K}_t\gamma_d$. Let A be a Borel subset of $\mathbb{R}^d$ such that $\mathrm{Leb}_d(A) = 0$, then $\gamma_d(A) = 0$; therefore $\int_{\mathbb{R}^d} \mathbf{1}_{\{X_t(x)\in A\}}\,\mathrm{d}\gamma_d(x) = 0$. It follows that $\mathbf{1}_{\{X_t(x)\in A\}} = 0$ for Leb_d almost every x, which implies $\mathrm{Leb}_d(X_t \in A) = 0$; this means that $(X_t)_\#\,\mathrm{Leb}_d$ is absolutely continuous with respect to Leb_d. □

In the next section, we shall prove that under the conditions of Corollary 3.8, the density of $(X_t)_\#\,\mathrm{Leb}_d$ with respect to Leb_d is strictly positive, in other words, Leb_d is quasi-invariant under X_t.

Corollary 3.9. *Assume that conditions* (A1)–(A4) *hold. Let $\sigma = (A_j^i)$ and suppose that for some $C > 0$,*

$$\sigma(x)\sigma(x)^* \geqslant C\,\mathrm{Id}, \quad \textit{for all } x \in \mathbb{R}^d.$$

Then $(X_t)_\#\gamma_d$ is absolutely continuous with respect to γ_d.

Proof. The conditions (A1)–(A4) are stronger than those in Theorem 1.1 of [34] given by X. Zhang, so the pathwise uniqueness holds. Hence Theorem 3.4 applies to this case. □

4. Quasi-invariance under stochastic flow

In the sequel, by quasi-invariance we mean that the Radon–Nikodym derivative of the corresponding push-forward measure is strictly positive. First we prove that in the situation of Corollary 3.8, the Lebesgue measure is in fact quasi-invariant under the stochastic flow of homeomorphisms. To this end, we need some preparations. In what follows, $T_0 > 0$ is chosen small enough such that (3.5) holds.

Proposition 4.1. *Let $q \geqslant 2$. Then*

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d} \mathbb{E}\Bigg[\bigg|\sup_{0\leqslant t\leqslant T_0}\sum_{j=1}^m \int_0^t \big[\delta\big(A_j^\varepsilon\big)\big(X_s^\varepsilon\big) - \delta(A_j)(X_s)\big]\,\mathrm{d}w_s^j\bigg|^q\Bigg]\,\mathrm{d}\gamma_d = 0. \tag{4.1}$$

Proof. By Burkholder's inequality,

$$\mathbb{E}\left(\sup_{0\leqslant t\leqslant T_0}\left|\sum_{j=1}^m\int_0^t\left[\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)(X_s)\right]\mathrm{d}w_s^j\right|^q\right)$$

$$\leqslant C\mathbb{E}\left[\left(\int_0^{T_0}\sum_{j=1}^m\left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)(X_s)\right|^2\mathrm{d}s\right)^{q/2}\right]$$

$$\leqslant CT_0^{q/2-1}\sum_{j=1}^m\int_0^{T_0}\mathbb{E}\left(\left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)(X_s)\right|^q\right)\mathrm{d}s.$$

Again by the inequality $(a+b)^q\leqslant C_q(a^q+b^q)$, there exists a constant $C_{q,T_0}>0$ such that the above quantity is dominated by

$$C_{q,T_0}\sum_{j=1}^m\left[\int_0^{T_0}\mathbb{E}\left(\left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)\left(X_s^\varepsilon\right)\right|^q\right)\mathrm{d}s+\int_0^{T_0}\mathbb{E}\left(\left|\delta(A_j)\left(X_s^\varepsilon\right)-\delta(A_j)(X_s)\right|^q\right)\mathrm{d}s\right]. \tag{4.2}$$

Let I_1^ε and I_2^ε be the two terms in the squared bracket of (4.2). Note that

$$\begin{aligned}\int_{\mathbb{R}^d}\mathbb{E}\left(\left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)\left(X_s^\varepsilon\right)\right|^q\right)\mathrm{d}\gamma_d&=\mathbb{E}\int_{\mathbb{R}^d}\left|\delta\left(A_j^\varepsilon\right)-\delta(A_j)\right|^qK_s^\varepsilon\,\mathrm{d}\gamma_d\\&\leqslant\left\|\delta\left(A_j^\varepsilon\right)-\delta(A_j)\right\|_{L^{2q}(\gamma_d)}^q\left\|K_s^\varepsilon\right\|_{L^2(\mathbb{P}\times\gamma_d)}.\end{aligned}\tag{4.3}$$

According to (3.5), for $s\leqslant T_0$, we have $\|K_s^\varepsilon\|_{L^2(\mathbb{P}\times\gamma_d)}\leqslant\Lambda_{T_0}$. Remark that

$$\delta\left(A_j^\varepsilon\right)=\delta(\varphi_\varepsilon P_\varepsilon A_j)=\varphi_\varepsilon e^\varepsilon P_\varepsilon\delta(A_j)-\langle\nabla\varphi_\varepsilon,P_\varepsilon A_j\rangle,$$

which converges to $\delta(A_j)$ in $L^{2q}(\gamma_d)$. By (4.3),

$$\begin{aligned}\int_{\mathbb{R}^d}I_1^\varepsilon\,\mathrm{d}\gamma_d&=\int_0^{T_0}\left[\int_{\mathbb{R}^d}\mathbb{E}\left(\left|\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon\right)-\delta(A_j)\left(X_s^\varepsilon\right)\right|^q\right)\mathrm{d}\gamma_d\right]\mathrm{d}s\\&\leqslant T_0\Lambda_{T_0}\left\|\delta\left(A_j^\varepsilon\right)-\delta(A_j)\right\|_{L^{2q}(\gamma_d)}^q\end{aligned}$$

which tends to 0 as $\varepsilon\to0$.

For the estimate of I_2^ε, we remark that $\int_{\mathbb{R}^d} |\delta(A_j)|^{2q}\,\mathrm{d}\gamma_d < +\infty$. Let $\eta > 0$ be given. There exists $\psi \in C_c(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} \left|\delta(A_j) - \psi\right|^{2q} \mathrm{d}\gamma_d \leqslant \eta^2.$$

We have, for some constant $C_q > 0$,

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathbb{E}\left(\left|\delta(A_j)\left(X_s^\varepsilon\right) - \delta(A_j)(X_s)\right|^q\right) \mathrm{d}\gamma_d \\
&\quad \leqslant C_q \Bigg[\int_{\mathbb{R}^d} \mathbb{E}\left(\left|\delta(A_j)\left(X_s^\varepsilon\right) - \psi\left(X_s^\varepsilon\right)\right|^q\right) \mathrm{d}\gamma_d + \int_{\mathbb{R}^d} \mathbb{E}\left(\left|\psi\left(X_s^\varepsilon\right) - \psi(X_s)\right|^q\right) \mathrm{d}\gamma_d \\
&\qquad + \int_{\mathbb{R}^d} \mathbb{E}\left(\left|\psi(X_s) - \delta(A_j)(X_s)\right|^q\right) \mathrm{d}\gamma_d \Bigg]. \qquad (4.4)
\end{aligned}$$

Again by (3.6), we find

$$\begin{aligned}
\mathbb{E}\left[\int_{\mathbb{R}^d} \left|\delta(A_j)\left(X_s^\varepsilon\right) - \psi\left(X_s^\varepsilon\right)\right|^q \mathrm{d}\gamma_d\right] &= \mathbb{E}\left[\int_{\mathbb{R}^d} \left|\delta(A_j) - \psi\right|^q K_s^\varepsilon\,\mathrm{d}\gamma_d\right] \\
&\leqslant \left\|\delta(A_j) - \psi\right\|_{L^{2q}(\gamma_d)}^q \Lambda_{T_0} \leqslant \Lambda_{T_0}\eta.
\end{aligned}$$

In the same way,

$$\mathbb{E}\left[\int_{\mathbb{R}^d} \left|\delta(A_j)(X_s) - \psi(X_s)\right|^q \mathrm{d}\gamma_d\right] \leqslant \Lambda_{T_0}\eta.$$

To estimate the second term on the right-hand side of (4.4), we use Theorem 3.5: from (3.14), we see that up to a subsequence, $X_s^\varepsilon(w,x)$ converges to $X_s(w,x)$, for each $s \leqslant T_0$ and almost all $(w,x) \in \Omega \times \mathbb{R}^d$. By Lebesgue's dominated convergence theorem,

$$\lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} \mathbb{E}\left(\left|\psi\left(X_s^\varepsilon\right) - \psi(X_s)\right|^q\right) \mathrm{d}\gamma_d = 0.$$

In conclusion, $\lim_{\varepsilon\to 0} \int_{\mathbb{R}^d} I_2^\varepsilon\,\mathrm{d}\gamma_d = 0$. According to (4.2), the proof of (4.1) is complete. □

Proposition 4.2. *Let Φ be defined by*

$$\Phi = \delta(A_0) + \frac{1}{2}\sum_{j=1}^m |A_j|^2 + \frac{1}{2}\sum_{j=1}^m \left\langle \nabla A_j, (\nabla A_j)^* \right\rangle, \qquad (4.5)$$

and analogously Φ_ε where A_j is replaced by A_j^ε. Then

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\int_0^{T_0}\mathbb{E}\left(\left|\Phi_\varepsilon\left(X_s^\varepsilon\right)-\Phi(X_s)\right|^q\right)\mathrm{d}s\,\mathrm{d}\gamma_d=0. \tag{4.6}$$

Proof. Along the lines of the proof of Proposition 4.1, it is sufficient to remark that

$$\lim_{\varepsilon\to 0}\|\Phi_\varepsilon-\Phi\|_{L^{2q}(\gamma_d)}=0. \tag{4.7}$$

To see this, let us check convergence for the last term in the definition of Φ_ε. We have

$$\begin{aligned}&\left|\left\langle\nabla A_j^\varepsilon,\left(\nabla A_j^\varepsilon\right)^*\right\rangle-\left\langle\nabla A_j,(\nabla A_j)^*\right\rangle\right|\\ &\quad\leqslant\left\|\nabla A_j^\varepsilon-\nabla A_j\right\|\left\|\nabla A_j^\varepsilon\right\|+\|\nabla A_j\|\left\|\nabla A_j^\varepsilon-\nabla A_j\right\|.\end{aligned}$$

Note that $A_j^\varepsilon=\varphi_\varepsilon P_\varepsilon A_j$. Thus

$$\nabla A_j^\varepsilon=\nabla\varphi_\varepsilon\otimes P_\varepsilon A_j+e^{-\varepsilon}\varphi_\varepsilon P_\varepsilon(\nabla A_j),$$

which converges to ∇A_j in $L^{2q}(\gamma_d)$ as $\varepsilon\to 0$. □

Now we can prove

Proposition 4.3. *Under the conditions of Corollary* 3.8*, the Lebesgue measure* Leb_d *is quasi-invariant under the stochastic flow.*

Proof. Let k_t be the density of $(X_t)_\#\,\mathrm{Leb}_d$ with respect to Leb_d. We shall prove that k_t is strictly positive. Set

$$\tilde K_t^\varepsilon(x)=\exp\left(-\sum_{j=1}^m\int_0^t\delta\left(A_j^\varepsilon\right)\left(X_s^\varepsilon(x)\right)\mathrm{d}w_s^j-\int_0^t\Phi_\varepsilon\left(X_s^\varepsilon(x)\right)\mathrm{d}s\right), \tag{4.8}$$

where Φ_ε is defined in Proposition 4.2. By (2.3) we have

$$\int_{\mathbb{R}^d}\psi\left(X_t^\varepsilon\right)\tilde K_t^\varepsilon\,\mathrm{d}\gamma_d=\int_{\mathbb{R}^d}\psi\,\mathrm{d}\gamma_d,\qquad\psi\in C_c^1\left(\mathbb{R}^d\right). \tag{4.9}$$

Applying Propositions 4.1 and 4.2, up to a subsequence, for each $t\leqslant T_0$ and almost every (w,x), the term $\tilde K_t^\varepsilon(w,x)$ defined in (4.8) converges to

$$\tilde K_t(x)=\exp\left(-\sum_{j=1}^m\int_0^t\delta(A_j)\left(X_s(x)\right)\mathrm{d}w_s^j-\int_0^t\Phi\left(X_s(x)\right)\mathrm{d}s\right). \tag{4.10}$$

By Corollary 2.3 and Lemma 3.2, we may assume that T_0 is small enough so that for any $t \leqslant T_0$, the family $\{\tilde{K}_t^\varepsilon\colon\ \varepsilon \leqslant 1\}$ is also bounded in $L^2(\mathbb{P} \times \gamma_d)$. Therefore, by the uniform integrability, letting $\varepsilon \to 0$ in (4.9), we get $\mathbb{P}$-almost surely,

$$\int_{\mathbb{R}^d} \psi(X_t)\tilde{K}_t\,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} \psi\,\mathrm{d}\gamma_d, \quad \psi \in C_c^1\big(\mathbb{R}^d\big). \tag{4.11}$$

Now taking a Borel version of $x \to \tilde{K}_t(w,x)$. Under the assumptions, the solution X_t is a stochastic flow of homeomorphisms, hence the inverse flow X_t^{-1} exists. Consequently, if $t \leqslant T_0$, we deduce from (4.11) that the density $K_t(w,x)$ of $(X_t)_\#\gamma_d$ with respect to γ_d admits the expression $K_t(w,x) = [\tilde{K}_t(w, X_t^{-1}(w,x))]^{-1}$ which is strictly positive. For X_{t+T_0} with $t \leqslant T_0$, we use the flow property: $X_{t+T_0}(w,x) = X_t(\theta_{T_0}w, X_{T_0}(w,x))$. Thus, for any $\psi \in C_c^1(\mathbb{R}^d)$,

$$\begin{aligned}\int_{\mathbb{R}^d} \psi(X_{t+T_0})\,\mathrm{d}\gamma_d &= \int_{\mathbb{R}^d} \psi\big(X_t(X_{T_0})\big)\,\mathrm{d}\gamma_d \\ &= \int_{\mathbb{R}^d} \psi(X_t)K_{T_0}\,\mathrm{d}\gamma_d \\ &= \int_{\mathbb{R}^d} \psi K_{T_0}\big(X_t^{-1}\big)K_t\,\mathrm{d}\gamma_d.\end{aligned}$$

That is to say, the density $K_{t+T_0} = K_{T_0}(X_t^{-1})K_t$ is strictly positive. Continuing in this way, we obtain that K_t is strictly positive for any $t \geqslant 0$.

Now if $\rho(x)$ denotes the density of γ_d with respect to Leb_d, then

$$k_t(w,x) = \rho\big(X_t^{-1}(w,x)\big)^{-1}K_t(w,x)\rho(x) > 0$$

which concludes the proof. □

In what follows, we will give examples for which existence of the inverse flow is not known.

Theorem 4.4. *Let $A_1, \ldots, A_m$ be bounded C^1 vector fields on $\mathbb{R}^d$ such that their derivatives are of linear growth; furthermore let A_0 be continuous of linear growth such that $\delta(A_0)$ exists. Define*

$$\hat{A}_0 = A_0 - \sum_{j=1}^m \mathcal{L}_{A_j}A_j. \tag{4.12}$$

Suppose that $\delta(\hat{A}_0)$ exists and that

$$\int_{\mathbb{R}^d} \exp\big(\lambda_0\big(\big|\delta(A_0)\big| + \big|\delta(\hat{A}_0)\big|\big)\big)\,\mathrm{d}\gamma_d < +\infty, \quad \textit{for some } \lambda_0 > 0. \tag{4.13}$$

If pathwise uniqueness holds both for SDE (1.1) *and for*

$$dY_t = \sum_{j=1}^m A_j(Y_t)\,dw_t^j - \hat{A}_0(Y_t)\,dt, \tag{4.14}$$

then the solution X_t *to* SDE (1.1) *leaves the Gaussian measure* γ_d *quasi-invariant.*

Proof. Obviously the conditions in Theorem 3.4 are satisfied; hence $(X_t)_\# \gamma_d = K_t \gamma_d$. Let $t > 0$ be given, we consider the dual SDE to (1.1):

$$dY_s^t = \sum_{j=1}^m A_j\big(Y_s^t\big)\,dw_s^{t,j} - \hat{A}_0\big(Y_s^t\big)\,ds$$

for which pathwise uniqueness holds; here $w_s^t = w_{t-s} - w_t$ with $s \in [0, t]$. Let A_j^ε, $j = 0, 1, \ldots, m$, be the vector fields defined as above. Consider

$$dY_s^{t,\varepsilon} = \sum_{j=1}^m A_j^\varepsilon\big(Y_s^{t,\varepsilon}\big)\,dw_s^{t,j} - \hat{A}_0^\varepsilon\big(Y_s^{t,\varepsilon}\big)\,ds,$$

where $\hat{A}_0^\varepsilon = A_0^\varepsilon - \sum_{j=1}^m \mathcal{L}_{A_j^\varepsilon} A_j^\varepsilon$. Then it is known that $(X_t^\varepsilon)^{-1} = Y_t^{t,\varepsilon}$. It is easy to check that for some constant $C > 0$ independent of ε,

$$\big|\hat{A}_0^\varepsilon(x)\big| \leqslant C\big(1 + |x|\big). \tag{4.15}$$

Moreover,

$$\mathcal{L}_{A_j^\varepsilon} A_j^\varepsilon = \sum_{k=1}^d \big(A_j^\varepsilon\big)^k \left[\frac{\partial \varphi_\varepsilon}{\partial x_k} P_\varepsilon A_j + \varphi_\varepsilon e^{-\varepsilon} P_\varepsilon\left(\frac{\partial A_j}{\partial x_k}\right)\right]$$

which converges locally uniformly to $\mathcal{L}_{A_j} A_j$. Therefore $\hat{A}_0^\varepsilon$ converges uniformly over any compact subset to $\hat{A}_0$. By Theorem 3.5,

$$\lim_{\varepsilon \to 0} \sup_{|x| \leqslant R} \mathbb{E}\Big(\sup_{0 \leqslant s \leqslant t} \big|Y_s^{t,\varepsilon} - Y_s^t\big|^2\Big) = 0.$$

It follows that, along a sequence, $Y_t^{t,\varepsilon}$ converges to Y_t^t for almost every (w, x). Now let $\psi_1, \psi_2 \in C_b(\mathbb{R}^d)$, we have for $t \leqslant T_0$,

$$\int_{\mathbb{R}^d} \psi_1 \cdot \psi_2\big(X_t^\varepsilon\big) \tilde{K}_t^\varepsilon \, d\gamma_d = \int_{\mathbb{R}^d} \psi_1\big(Y_t^{t,\varepsilon}\big) \cdot \psi_2 \, d\gamma_d.$$

Letting $\varepsilon \to 0$ leads to

$$\int_{\mathbb{R}^d} \psi_1 \cdot \psi_2(X_t)\tilde{K}_t \,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} \psi_1\big(Y_t^t\big) \cdot \psi_2 \,\mathrm{d}\gamma_d. \tag{4.16}$$

Taking ψ_1 and ψ_2 positive in (4.16) and using a monotone class argument, we see that Eq. (4.16) holds for any positive Borel functions ψ_1 and ψ_2. Hence taking a Borel version of $\tilde{K}_t$ and setting $\psi_1 = 1/\tilde{K}_t$ in (4.16), we get

$$\int_{\mathbb{R}^d} \psi_2(X_t) \,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} \big[\tilde{K}_t\big(Y_t^t\big)\big]^{-1} \psi_2 \,\mathrm{d}\gamma_d. \tag{4.17}$$

It follows that $K_t = [\tilde{K}_t(Y_t^t)]^{-1} > 0$ for $t \leqslant T_0$. For X_{t+T_0} with $t \leqslant T_0$, we shall use repeatedly (4.16). By the flow property, $X_{t+T_0}(w,x) = X_t(\theta_{T_0}w, X_{T_0}(w,x))$ where $(\theta_{T_0}w)_t = w_{t+T_0} - w_{T_0}$. Letting $t = T_0$ and replacing ψ_2 by $\psi_2(X_t)$ we get

$$\int_{\mathbb{R}^d} \psi_1 \cdot \psi_2(X_{t+T_0})\tilde{K}_{T_0} \,\mathrm{d}\gamma_d = \int_{\mathbb{R}^d} \psi_1\big(Y_{T_0}^{T_0}\big)\psi_2(X_t) \,\mathrm{d}\gamma_d.$$

Taking $\psi_1 = 1/\tilde{K}_{T_0}$ in the above equality, we get

$$\begin{aligned}
\int_{\mathbb{R}^d} \psi_2(X_{t+T_0}) \,\mathrm{d}\gamma_d &= \int_{\mathbb{R}^d} \big[\tilde{K}_{T_0}\big(Y_{T_0}^{T_0}\big)\big]^{-1} \psi_2(X_t) \,\mathrm{d}\gamma_d \\
&= \int_{\mathbb{R}^d} \big[\tilde{K}_{T_0}\big(Y_{T_0}^{T_0}\big)\big]^{-1} \psi_2(X_t)\tilde{K}_t^{-1}\tilde{K}_t \,\mathrm{d}\gamma_d \\
&= \int_{\mathbb{R}^d} \big[\tilde{K}_{T_0}\big(Y_{T_0}^{T_0}\big(Y_t^t\big)\big)\big]^{-1} \big[\tilde{K}_t\big(Y_t^t\big)\big]^{-1} \psi_2 \,\mathrm{d}\gamma_d,
\end{aligned}$$

where in the last equality we have used (4.16) with $\psi_1 = [\tilde{K}_{T_0}(Y_{T_0}^{T_0})]^{-1}\tilde{K}_t^{-1}$. It follows that the density K_{t+T_0} of $(X_{t+T_0})_{\#}\gamma_d$ with respect to γ_d is strictly positive, and so on. □

Corollary 4.5. *Let $A_1, \ldots, A_m$ be bounded C^2 vector fields such that their derivatives up to order 2 grow at most linearly, and let A_0 be a continuous vector field of linear growth. Suppose that*

$$\big|A_0(x) - A_0(y)\big| \leqslant C_R|x-y| \log_k \frac{1}{|x-y|}$$
$$\text{for } |x| \leqslant R,\ |y| \leqslant R,\ |x-y| \leqslant c_0 \text{ small enough,} \tag{4.18}$$

where $\log_k s = (\log s)(\log\log s)\ldots(\log\ldots\log s)$. *Suppose further that*

$$\operatorname{div}(A_0) = \sum_{j=1}^{d} \frac{\partial A_0^j}{\partial x_j}$$

exists and is bounded. Then the stochastic flow X_t defined by SDE (1.1) *leaves the Lebesgue measure quasi-invariant.*

Proof. It is obvious that $\hat{A}_0$ defined in (4.12) satisfies condition (4.18); therefore by [12], pathwise uniqueness holds for SDE (1.1) and (4.14). Note that $\delta(A_0) = \langle x, A_0 \rangle - \operatorname{div}(A_0)$. Then condition (4.13) is satisfied; thus Theorem 4.4 yields the result. □

5. The case A_0 in a Sobolev space

From now on, A_0 is no longer supposed to be continuous, but to lie in some Sobolev space, that is, condition (A1) in (H) is replaced by

(A1′) For $i = 1, \dots, m$, $A_i \in \bigcap_{q \geqslant 1} \mathbb{D}_1^q(\gamma_d)$, $A_0 \in \mathbb{D}_1^q(\gamma_d)$ for some $q > 1$.

First we establish the following *a priori* estimate on perturbations, using the method developed in [36]. Let $\{A_0, A_1, \dots, A_m\}$ be a family of measurable vector fields on $\mathbb{R}^d$. We first give a precise meaning of solution to the following SDE

$$\mathrm{d}X_t = \sum_{i=1}^{m} A_i(X_t)\,\mathrm{d}w_t^i + A_0(X_t)\,\mathrm{d}t, \qquad X_0 = x. \tag{5.1}$$

Definition 5.1. We say that a measurable map $X : \Omega \times \mathbb{R}^d \to C([0,T], \mathbb{R}^d)$ is a solution to the Itô SDE (5.1) if

(i) for each $t \in [0,T]$ and almost all $x \in \mathbb{R}^d$, $w \to X_t(w,x)$ is measurable with respect to $\mathscr{F}_t$, i.e., the natural filtration generated by the Brownian motion $\{w_s\colon\ s \leqslant t\}$;
(ii) for each $t \in [0,T]$, there exists $K_t \in L^1(\mathbb{P} \times \mathbb{R}^d)$ such that $(X_t(w,\cdot))_{\#}\gamma_d$ admits K_t as the density with respect to γ_d;
(iii) almost surely

$$\sum_{i=1}^{m} \int_0^T \big|A_i\big(X_s(w,x)\big)\big|^2\,\mathrm{d}s + \int_0^T \big|A_0\big(X_s(w,x)\big)\big|\,\mathrm{d}s < +\infty;$$

(iv) for almost all $x \in \mathbb{R}^d$,

$$X_t(w,x) = x + \sum_{i=1}^{m} \int_0^t A_i\big(X_s(w,x)\big)\,\mathrm{d}w_s^i + \int_0^t A_0\big(X_s(w,x)\big)\,\mathrm{d}s;$$

(v) the flow property holds

$$X_{t+s}(w,x) = X_t\big(\theta_s w, X_s(w,x)\big).$$

Now let $\{\hat{A}_0, \hat{A}_1, \ldots, \hat{A}_m\}$ be another family of measurable vector fields on $\mathbb{R}^d$, and denote by $\hat{X}_t$ the solution to the SDE

$$\mathrm{d}\hat{X}_t = \sum_{i=1}^m \hat{A}_i(\hat{X}_t)\,\mathrm{d}w_t^i + \hat{A}_0(\hat{X}_t)\,\mathrm{d}t, \quad \hat{X}_0 = x. \tag{5.2}$$

Let $\hat{K}_t$ be the density of $(\hat{X}_t)_{\#}\gamma_d$ with respect to γ_d and define

$$\Lambda_{p,T} = \sup_{0\leqslant t\leqslant T}\big(\|K_t\|_{L^p(\mathbb{P}\times\gamma_d)} \vee \|\hat{K}_t\|_{L^p(\mathbb{P}\times\gamma_d)}\big). \tag{5.3}$$

Theorem 5.2. *Let $q > 1$. Suppose that $A_1, \ldots, A_m$ as well as $\hat{A}_1, \ldots, \hat{A}_m$ are in $\mathbb{D}_1^{2q}(\gamma_d)$ and $A_0, \hat{A}_0 \in \mathbb{D}_1^q(\gamma_d)$. Then, for any $T > 0$ and $R > 0$, there exist constants $C_{d,q,R} > 0$ and $C_T > 0$ such that for any $\sigma > 0$,*

$$\begin{aligned}
&\mathbb{E}\left[\int_{G_R} \log\left(\frac{\sup_{0\leqslant t\leqslant T}|X_t - \hat{X}_t|^2}{\sigma^2} + 1\right)\mathrm{d}\gamma_d\right] \\
&\quad\leqslant C_T\Lambda_{p,T}\left\{ C_{d,q,R}\left[\|\nabla A_0\|_{L^q} + \left(\sum_{i=1}^m \|\nabla A_i\|_{L^{2q}}^2\right)^{1/2} + \sum_{i=1}^m \|\nabla A_i\|_{L^{2q}}^2\right]\right. \\
&\qquad\left. + \frac{1}{\sigma^2}\sum_{i=1}^m \|A_i - \hat{A}_i\|_{L^{2q}}^2 + \frac{1}{\sigma}\left[\|A_0 - \hat{A}_0\|_{L^q} + \left(\sum_{i=1}^m \|A_i - \hat{A}_i\|_{L^{2q}}^2\right)^{1/2}\right]\right\},
\end{aligned}$$

where p is the conjugate number of q: $1/p + 1/q = 1$, and

$$G_R(w) = \Big\{x \in \mathbb{R}^d\colon \sup_{0\leqslant t\leqslant T}\big|X_t(w,x)\big| \vee \big|\hat{X}_t(w,x)\big| \leqslant R\Big\}. \tag{5.4}$$

Proof. Denote $\xi_t = X_t - \hat{X}_t$, then $\xi_0 = 0$. By Itô's formula,

$$\begin{aligned}
\mathrm{d}|\xi_t|^2 &= 2\sum_{i=1}^m \big\langle \xi_t, A_i(X_t) - \hat{A}_i(\hat{X}_t)\big\rangle\,\mathrm{d}w_t^i + 2\big\langle \xi_t, A_0(X_t) - \hat{A}_0(\hat{X}_t)\big\rangle\,\mathrm{d}t \\
&\quad + \sum_{i=1}^m \big|A_i(X_t) - \hat{A}_i(\hat{X}_t)\big|^2\,\mathrm{d}t.
\end{aligned} \tag{5.5}$$

For $\sigma > 0$, $\log(|\xi_t|^2/\sigma^2 + 1) = \log(|\xi_t|^2 + \sigma^2) - \log\sigma^2$. Again by Itô's formula,

$$\mathrm{d}\log\big(|\xi_t|^2 + \sigma^2\big) = \frac{\mathrm{d}|\xi_t|^2}{|\xi_t|^2 + \sigma^2} - \frac{1}{2}\frac{4\sum_{i=1}^m \langle \xi_t, A_i(X_t) - \hat{A}_i(\hat{X}_t)\rangle^2}{(|\xi_t|^2 + \sigma^2)^2}\,\mathrm{d}t.$$

Using (5.5), we obtain

$$
\begin{aligned}
\mathrm{d}\log\big(|\xi_t|^2+\sigma^2\big) &= 2\sum_{i=1}^m \frac{\langle \xi_t, A_i(X_t)-\hat{A}_i(\hat{X}_t)\rangle}{|\xi_t|^2+\sigma^2}\,\mathrm{d}w_t^i + 2\frac{\langle \xi_t, A_0(X_t)-\hat{A}_0(\hat{X}_t)\rangle}{|\xi_t|^2+\sigma^2}\,\mathrm{d}t \\
&\quad + \sum_{i=1}^m \frac{|A_i(X_t)-\hat{A}_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}t - 2\sum_{i=1}^m \frac{\langle \xi_t, A_i(X_t)-\hat{A}_i(\hat{X}_t)\rangle^2}{(|\xi_t|^2+\sigma^2)^2}\,\mathrm{d}t \\
&=: \mathrm{d}I_1(t)+\mathrm{d}I_2(t)+\mathrm{d}I_3(t)+\mathrm{d}I_4(t).
\end{aligned} \tag{5.6}
$$

Let $\tau_R(x) = \inf\{t \geqslant 0\colon\ |X_t(x)| \vee |\hat{X}_t(x)| > R\}$. Remark that almost surely, $G_R \subset \{x\colon \tau_R(x) > T\}$ and for any $t \geqslant 0$, $\{\tau_R > t\} \subset B(R)$. Therefore

$$
\mathbb{E}\bigg[\int_{G_R} \sup_{0\leqslant t\leqslant T} \big|I_1(t)\big|\,\mathrm{d}\gamma_d\bigg] \leqslant \mathbb{E}\bigg[\int_{B(R)} \sup_{0\leqslant t\leqslant T\wedge\tau_R} \big|I_1(t)\big|\,\mathrm{d}\gamma_d\bigg].
$$

By Burkholder's inequality,

$$
\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T\wedge\tau_R} \big|I_1(t)\big|^2\Big] \leqslant 4\mathbb{E}\bigg[\int_0^{T\wedge\tau_R} \sum_{i=1}^m \frac{\langle \xi_t, A_i(X_t)-\hat{A}_i(\hat{X}_t)\rangle^2}{(|\xi_t|^2+\sigma^2)^2}\,\mathrm{d}t\bigg],
$$

which is obviously less than

$$
4\mathbb{E}\bigg[\int_0^{T\wedge\tau_R} \sum_{i=1}^m \frac{|A_i(X_t)-\hat{A}_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}t\bigg].
$$

Hence

$$
\begin{aligned}
&\mathbb{E}\bigg[\int_{B(R)} \sup_{0\leqslant t\leqslant T\wedge\tau_R} \big|I_1(t)\big|\,\mathrm{d}\gamma_d\bigg] \\
&\quad \leqslant 4\bigg[\int_0^T \bigg(\mathbb{E}\int_{\{\tau_R>t\}} \sum_{i=1}^m \frac{|A_i(X_t)-\hat{A}_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}\gamma_d\bigg)\,\mathrm{d}t\bigg]^{1/2}.
\end{aligned} \tag{5.7}
$$

We have $A_i(X_t) - \hat{A}_i(\hat{X}_t) = A_i(X_t) - A_i(\hat{X}_t) + A_i(\hat{X}_t) - \hat{A}_i(\hat{X}_t)$. Using the density $\hat{K}_t$, it is clear that

$$
\begin{aligned}
\mathbb{E}\int_{\{\tau_R>t\}} \frac{|A_i(\hat{X}_t)-\hat{A}_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}\gamma_d &\leqslant \frac{1}{\sigma^2}\mathbb{E}\int_{\mathbb{R}^d} \big|A_i(\hat{X}_t)-\hat{A}_i(\hat{X}_t)\big|^2\,\mathrm{d}\gamma_d \\
&= \frac{1}{\sigma^2}\mathbb{E}\int_{\mathbb{R}^d} |A_i-\hat{A}_i|^2\hat{K}_t\,\mathrm{d}\gamma_d.
\end{aligned}
$$

Thus by Hölder's inequality and according to (5.3), we have

$$\mathbb{E}\int_{\{\tau_R>t\}} \frac{|A_i(\hat{X}_t)-\hat{A}_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}\gamma_d \leqslant \frac{\Lambda_{p,T}}{\sigma^2}\|A_i-\hat{A}_i\|_{L^{2q}}^2. \tag{5.8}$$

Now we shall use Theorem A.1 in Appendix A to estimate the other term. Note that on the set $\{\tau_R>t\}$, $X_t, \hat{X}_t \in B(R)$, thus $|X_t-\hat{X}_t|\leqslant 2R$. Since $(X_t)_{\#}\gamma_d \ll \gamma_d$ and $(\hat{X}_t)_{\#}\gamma_d \ll \gamma_d$, we can apply (A.2) so that

$$\left|A_i(X_t)-A_i(\hat{X}_t)\right| \leqslant C_d|X_t-\hat{X}_t|\big(M_{2R}|\nabla A_i|(X_t)+M_{2R}|\nabla A_i|(\hat{X}_t)\big).$$

Then

$$\mathbb{E}\left[\int_{\{\tau_R>t\}} \frac{|A_i(X_t)-A_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}\gamma_d\right] \leqslant C_d^2\mathbb{E}\int_{\{\tau_R>t\}} \big(M_{2R}|\nabla A_i|(X_t)+M_{2R}|\nabla A_i|(\hat{X}_t)\big)^2\,\mathrm{d}\gamma_d.$$

Notice again that on $\{\tau_R(x)>t\}$, $X_t(x)$ and $\hat{X}_t(x)$ are in $B(R)$, therefore

$$\begin{aligned}\mathbb{E}\left[\int_{\{\tau_R>t\}} \frac{|A_i(X_t)-A_i(\hat{X}_t)|^2}{|\xi_t|^2+\sigma^2}\,\mathrm{d}\gamma_d\right] &\leqslant 2C_d^2\mathbb{E}\int_{B(R)} \big(M_{2R}|\nabla A_i|\big)^2(K_t+\hat{K}_t)\,\mathrm{d}\gamma_d \\ &\leqslant 4C_d^2\Lambda_{p,T}\left(\int_{B(R)} \big(M_{2R}|\nabla A_i|\big)^{2q}\,\mathrm{d}\gamma_d\right)^{1/q}. \end{aligned} \tag{5.9}$$

Remark that the maximal function inequality does not hold for the Gaussian measure γ_d on the whole space $\mathbb{R}^d$. However, on each ball $B(R)$,

$$\gamma_d|_{B(R)} \leqslant \frac{1}{(2\pi)^{d/2}}\mathrm{Leb}_d|_{B(R)} \leqslant e^{R^2/2}\gamma_d|_{B(R)}.$$

Thus, according to (A.3),

$$\begin{aligned}\int_{B(R)} \big(M_{2R}|\nabla A_i|\big)^{2q}\,\mathrm{d}\gamma_d &\leqslant \frac{1}{(2\pi)^{d/2}}\int_{B(R)} \big(M_{2R}|\nabla A_i|\big)^{2q}\,\mathrm{d}x \\ &\leqslant \frac{C_{d,q}}{(2\pi)^{d/2}}\int_{B(3R)} |\nabla A_i|^{2q}\,\mathrm{d}x \\ &\leqslant C_{d,q}e^{9R^2/2}\int_{B(3R)} |\nabla A_i|^{2q}\,\mathrm{d}\gamma_d \\ &\leqslant C_{d,q}e^{9R^2/2}\|\nabla A_i\|_{L^{2q}}^{2q}.\end{aligned}$$

Therefore by (5.9), there exists a constant $C_{d,q,R} > 0$ such that

$$\mathbb{E}\Big[\int_{\{\tau_R > t\}} \frac{|A_i(X_t) - A_i(\hat{X}_t)|^2}{|\xi_t|^2 + \sigma^2}\,\mathrm{d}\gamma_d\Big] \leqslant C_{d,q,R}\,\Lambda_{p,T}\,\|\nabla A_i\|^2_{L^{2q}}.$$

Combining this estimate with (5.7) and (5.8), we get

$$\begin{aligned}&\mathbb{E}\Big[\int_{G_R} \sup_{0\leqslant t\leqslant T}\big|I_1(t)\big|\,\mathrm{d}\gamma_d\Big]\\ &\quad\leqslant C T^{1/2}\Lambda_{p,T}^{1/2}\Bigg(C_{d,q,R}\sum_{i=1}^m \|\nabla A_i\|^2_{L^{2q}} + \frac{1}{\sigma^2}\sum_{i=1}^m \|A_i - \hat{A}_i\|^2_{L^{2q}}\Bigg)^{1/2}.\end{aligned} \tag{5.10}$$

Now we turn to deal with $I_2(t)$ in (5.6). We have

$$\mathbb{E}\Big[\int_{G_R} \sup_{0\leqslant t\leqslant T}\big|I_2(t)\big|\,\mathrm{d}\gamma_d\Big] \leqslant 2\int_0^T\Big[\mathbb{E}\int_{G_R} \frac{|A_0(X_t) - \hat{A}_0(\hat{X}_t)|}{(|\xi_t|^2 + \sigma^2)^{1/2}}\,\mathrm{d}\gamma_d\Big]\,\mathrm{d}t.$$

Note that for $x \in G_R$, $\hat{X}_t(x) \in B(R)$ for each $t \in [0, T]$, thus

$$\mathbb{E}\Big[\int_{G_R} \frac{|A_0(\hat{X}_t) - \hat{A}_0(\hat{X}_t)|}{(|\xi_t|^2 + \sigma^2)^{1/2}}\,\mathrm{d}\gamma_d\Big] \leqslant \frac{1}{\sigma}\mathbb{E}\int_{B(R)} |A_0 - \hat{A}_0|\hat{K}_t\,\mathrm{d}\gamma_d \leqslant \frac{\Lambda_{p,T}}{\sigma}\|A_0 - \hat{A}_0\|_{L^q}.$$

Again using (A.2),

$$\mathbb{E}\Big[\int_{G_R} \frac{|A_0(X_t) - A_0(\hat{X}_t)|}{(|\xi_t|^2 + \sigma^2)^{1/2}}\,\mathrm{d}\gamma_d\Big] \leqslant C_d\mathbb{E}\int_{G_R} \big(M_{2R}|\nabla A_0|(X_t) + M_{2R}|\nabla A_0|(\hat{X}_t)\big)\,\mathrm{d}\gamma_d,$$

which is dominated by

$$C_d\mathbb{E}\Big[\int_{B(R)} \big(M_{2R}|\nabla A_0|\big)(K_t + \hat{K}_t)\,\mathrm{d}\gamma_d\Big] \leqslant C_{d,q,R}\|\nabla A_0\|_{L^q}\Lambda_{p,T}.$$

Therefore we arrive at the following estimate for I_2:

$$\mathbb{E}\Big[\int_{G_R} \sup_{0\leqslant t\leqslant T}\big|I_2(t)\big|\,\mathrm{d}\gamma_d\Big] \leqslant 2T\Lambda_{p,T}\Big(C_{d,q,R}\|\nabla A_0\|_{L^q} + \frac{1}{\sigma}\|A_0 - \hat{A}_0\|_{L^q}\Big). \tag{5.11}$$

In the same way we get

$$\mathbb{E}\bigg[\int_{G_R}\sup_{0\leqslant t\leqslant T}\big|I_3(t)\big|\,\mathrm{d}\gamma_d\bigg] \leqslant CT\Lambda_{p,T}\Bigg(C_{d,q,R}\sum_{i=1}^m\|\nabla A_i\|_{L^{2q}}^2+\frac{1}{\sigma^2}\sum_{i=1}^m\|A_i-\hat{A}_i\|_{L^{2q}}^2\Bigg). \tag{5.12}$$

The term $I_4(t)$ is negative and hence omitted. Combining (5.6) and (5.10)–(5.12), the proof is completed. □

Now we shall construct a solution to SDE (5.1). To this end, we take $\varepsilon = 1/n$ and write A_j^n instead of $A_j^{1/n}$ introduced in Section 3. Then by assumption (A2) and Lemma 3.1, there is a constant $C > 0$ independent of n and i, such that

$$\big|A_i^n(x)\big| \leqslant C\big(1+|x|\big). \tag{5.13}$$

Let X_t^n be the solution to Itô SDE (5.1) with coefficients A_i^n $(i = 0, 1, \ldots, m)$. Then for any $\alpha \geqslant 1$ and $T > 0$, there exists $C_{\alpha,T} > 0$ independent of n such that

$$\mathbb{E}\Big(\sup_{0\leqslant t\leqslant T}\big|X_t^n\big|^\alpha\Big) \leqslant C_{\alpha,T}\big(1+|x|^\alpha\big), \quad \text{for all } x \in \mathbb{R}^d. \tag{5.14}$$

Let K_t^n be the density of $(X_t^n)_\#\gamma_d$ with respect to γ_d. Under the hypotheses (A2)–(A4), there exists $T_0 > 0$ such that (recall that p is the conjugate number of $q > 1$):

$$\Lambda_{p,T_0} := \Bigg[\int_{\mathbb{R}^d}\exp\Bigg(2pT_0\Bigg[|A_0|+e\big|\delta(A_0)\big| + \sum_{j=1}^m\big(2p|A_j|^2+|\nabla A_j|^2+2(p-1)e^2\big|\delta(A_j)\big|^2\big)\Bigg]\Bigg)\,\mathrm{d}\gamma_d\Bigg]^{\frac{p-1}{p(2p-1)}} < \infty. \tag{5.15}$$

Similar to (3.6), we have

$$\sup_{t\in[0,T_0]}\sup_{n\geqslant 1}\big\|K_t^n\big\|_{L^p(\gamma_d\times\mathbb{P})} \leqslant \Lambda_{p,T_0} < +\infty. \tag{5.16}$$

Next we shall prove that the family $\{X_\cdot^n\colon\, n \geqslant 1\}$ converges to some stochastic field.

Theorem 5.3. *Let T_0 be given in* (5.15). *Then, under the assumptions* (A1′) *and* (A2)–(A4)*, there exists $X : \Omega \times \mathbb{R}^d \to C([0, T_0], \mathbb{R}^d)$ such that for any $\alpha \geqslant 1$,*

$$\lim_{n\to\infty} \mathbb{E}\left[\int_{\mathbb{R}^d} \Big(\sup_{0\leqslant t\leqslant T_0} \big|X_t^n - X_t\big|^\alpha\Big)\,\mathrm{d}\gamma_d\right] = 0. \tag{5.17}$$

Proof. We shall prove that $\{X^n:\ n \geqslant 1\}$ is a Cauchy sequence in $L^\alpha(\Omega \times \mathbb{R}^d; C([0, T_0], \mathbb{R}^d))$. Denote by $\|\cdot\|_{\infty,T_0}$ the uniform norm on $C([0, T_0], \mathbb{R}^d)$. We have to prove that

$$\lim_{n,k\to+\infty} \mathbb{E}\left[\int_{\mathbb{R}^d} \big\|X^n - X^k\big\|_{\infty,T_0}^\alpha\,\mathrm{d}\gamma_d\right] = 0. \tag{5.18}$$

First by (5.14), the quantity

$$J_{\alpha,T_0} := \sup_{n\geqslant 1} \mathbb{E}\left[\int_{\mathbb{R}^d} \big\|X^n\big\|_{\infty,T_0}^{2\alpha}\,\mathrm{d}\gamma_d\right] \leqslant C_{\alpha,T_0} \int_{\mathbb{R}^d} \big(1 + |x|^{2\alpha}\big)\,\mathrm{d}\gamma_d \tag{5.19}$$

is obviously finite. Let $R > 0$ and set

$$G_{n,R}(w) = \big\{x \in \mathbb{R}^d:\ \big\|X^n(w,x)\big\|_{\infty,T_0} \leqslant R\big\}.$$

Using (5.19), for any $\alpha \geqslant 1$ and $R > 0$, we have

$$\sup_{n\geqslant 1} \mathbb{E}\big[\gamma_d\big(G_{n,R}^c\big)\big] \leqslant \frac{J_{\alpha,T_0}}{R^{2\alpha}}.$$

Now by Cauchy–Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left[\int_{G_{n,R}^c \cup G_{k,R}^c} \big\|X^n - X^k\big\|_{\infty,T_0}^\alpha\,\mathrm{d}\gamma_d\right] \\
&\quad\leqslant \big(\mathbb{E}\big[\gamma_d\big(G_{n,R}^c \cup G_{k,R}^c\big)\big]\big)^{1/2} \cdot \left(\mathbb{E}\int_{\mathbb{R}^d} \big\|X^n - X^k\big\|_{\infty,T_0}^{2\alpha}\,\mathrm{d}\gamma_d\right)^{1/2} \\
&\quad\leqslant \left(\frac{2J_{\alpha,T_0}}{R^{2\alpha}}\right)^{1/2} \cdot \big(2^{2\alpha} J_{\alpha,T_0}\big)^{1/2}.
\end{aligned} \tag{5.20}$$

Given $\varepsilon > 0$, we may choose $R > 1$ sufficiently large such that the last quantity in inequality (5.20) is less than ε. Then, for any $n, k \geqslant 1$,

$$\mathbb{E}\left(\int_{G_{n,R}^c \cup G_{k,R}^c} \big\|X^n - X^k\big\|_{\infty,T_0}^\alpha\,\mathrm{d}\gamma_d\right) \leqslant \varepsilon. \tag{5.21}$$

Let

$$\sigma_{n,k} = \|A_0^n - A_0^k\|_{L^q} + \left(\sum_{i=1}^m \|A_i^n - A_i^k\|_{L^{2q}}^2\right)^{1/2},$$

which tends to 0 as $n, k \to +\infty$ since A_0^n converges to A_0 in $L^q(\gamma_d)$ and A_i^n converges to A_i in $L^{2q}(\gamma_d)$ for $i = 1, \ldots, m$. Now applying Theorem 5.2 with A_i and $\hat{A}_i$ being replaced respectively by A_i^n and A_i^k, we get

$$\begin{aligned} I_{n,k} &:= \mathbb{E}\left[\int_{G_{n,R}\cap G_{k,R}} \log\left(\frac{\|X^n - X^k\|_{\infty,T_0}^2}{\sigma_{n,k}^2} + 1\right) \mathrm{d}\gamma_d\right] \\ &\leqslant C_{T_0}\Lambda_{p,T_0}\left\{C_{d,q,R}\left[\|\nabla A_0^n\|_{L^q} + \left(\sum_{i=1}^m \|\nabla A_i^n\|_{L^{2q}}^2\right)^{1/2} + \sum_{i=1}^n \|\nabla A_i^n\|_{L^{2q}}^2\right] + 2\right\}. \end{aligned}$$

Recall that $A_i^n = \varphi_{1/n} P_{1/n} A_i$. Thus $\nabla A_i^n = \nabla\varphi_{1/n} \otimes P_{1/n} A_i + \varphi_{1/n} e^{-1/n} P_{1/n} \nabla A_i$ and

$$|\nabla A_i^n| \leqslant P_{1/n}(|A_i| + |\nabla A_i|).$$

We obtain the following uniform estimates

$$\|\nabla A_0^n\|_{L^q} \leqslant \|A_0\|_{\mathbb{D}_1^q}, \qquad \|\nabla A_i^n\|_{L^{2q}} \leqslant \|A_i\|_{\mathbb{D}_1^{2q}}.$$

Hence the quantity $I_{n,k}$ is uniformly bounded with respect to n, k. Let $\hat{\Pi}$ be the measure on $\Omega \times \mathbb{R}^d$ defined by

$$\int_{\Omega\times\mathbb{R}^d} \psi(w, x)\, \mathrm{d}\hat{\Pi}(w, x) = \mathbb{E}\left[\int_{G_{n,R}\cap G_{k,R}} \psi(w, x)\, \mathrm{d}\gamma_d(x)\right].$$

Obviously we have $\hat{\Pi}(\Omega \times \mathbb{R}^d) \leqslant 1$. Let $\eta > 0$ and consider

$$\Sigma_{n,k} = \left\{(w, x) \in \Omega \times \mathbb{R}^d\colon \|X^n(w, x) - X^k(w, x)\|_{\infty,T_0} \geqslant \eta\right\}$$

which equals

$$\left\{(w, x) \in \Omega \times \mathbb{R}^d\colon \log\left(\frac{\|X^n - X^k\|_{\infty,T_0}^2}{\sigma_{n,k}^2} + 1\right) \geqslant \log\left(\frac{\eta^2}{\sigma_{n,k}^2} + 1\right)\right\}.$$

It follows that as $n, k \to +\infty$,

$$\hat{\Pi}(\Sigma_{n,k}) \leqslant \frac{I_{n,k}}{\log(\eta^2/\sigma_{n,k}^2 + 1)} \to 0, \tag{5.22}$$

since $\sigma_{n,k} \to 0$ and the family $\{I_{n,k}\colon n,k \geqslant 1\}$ is bounded. Now

$$\mathbb{E}\Bigg[\int\limits_{G_{n,R}\cap G_{k,R}} \|X^n - X^k\|^\alpha_{\infty,T_0}\,\mathrm{d}\gamma_d\Bigg] = \int\limits_{\Omega\times\mathbb{R}^d} \|X^n - X^k\|^\alpha_{\infty,T_0}\,\mathrm{d}\hat{\Pi}$$
$$= \int\limits_{\Sigma^c_{n,k}} \|X^n - X^k\|^\alpha_{\infty,T_0}\,\mathrm{d}\hat{\Pi} + \int\limits_{\Sigma_{n,k}} \|X^n - X^k\|^\alpha_{\infty,T_0}\,\mathrm{d}\hat{\Pi}. \tag{5.23}$$

The first term on the right side of (5.23) is bounded by η^α, while the second one, due to (5.19) and (5.22), is dominated by

$$\sqrt{\hat{\Pi}(\Sigma_{n,k})}\cdot\sqrt{\mathbb{E}\int\limits_{\mathbb{R}^d}\|X^n - X^k\|^{2\alpha}_{\infty,T_0}\,\mathrm{d}\gamma_d} \leqslant 2^\alpha\sqrt{J_{\alpha,T_0}\hat{\Pi}(\Sigma_{n,k})} \to 0 \quad \text{as } n,k\to+\infty.$$

Now taking $\eta = \varepsilon^{1/\alpha}$ and combining (5.21) and (5.23), we see that

$$\limsup_{n,k\to+\infty}\mathbb{E}\Bigg[\int\limits_{\mathbb{R}^d}\|X^n - X^k\|^\alpha_{\infty,T_0}\,\mathrm{d}\gamma_d\Bigg] \leqslant 2\varepsilon,$$

which implies (5.18).

Let $X \in L^\alpha(\Omega\times\mathbb{R}^d; C([0,T_0],\mathbb{R}^d))$ be the limit of X^n in this space. We see that for each $t\in[0,T]$ and almost all $x\in\mathbb{R}^d$, $w\to X_t(w,x)$ is $\mathscr{F}_t$-measurable. □

Proposition 5.4. *There exists a family* $\{\hat{K}_t\colon t\in[0,T_0]\}$ *of density functions on* $\mathbb{R}^d$ *such that* $(X_t)_\#\gamma_d = \hat{K}_t\gamma_d$ *for each* $t\in[0,T_0]$. *Moreover,*

$$\sup_{0\leqslant t\leqslant T_0}\|\hat{K}_t\|_{L^p(\mathbb{P}\times\gamma_d)} \leqslant \Lambda_{p,T_0}$$

where Λ_{p,T_0} *is given by* (5.16).

Proof. It is the same as the proof of Theorem 3.4. □

The same arguments as in the proof of Propositions 4.1 and 4.2 yield the following

Proposition 5.5. *For any* $\alpha\geqslant 2$, *up to a subsequence,*

$$\lim_{n\to\infty}\int\limits_{\mathbb{R}^d}\mathbb{E}\Bigg[\sup_{0\leqslant t\leqslant T_0}\Bigg|\sum_{i=1}^m\int_0^t\big[A_i^n\big(X_s^n\big) - A_i(X_s)\big]\,\mathrm{d}w_s^i\Bigg|^\alpha\Bigg]\,\mathrm{d}\gamma_d = 0,$$

and

$$\lim_{n\to\infty}\int_{\mathbb{R}^d}\left[\mathbb{E}\int_0^{T_0}\left|A_0^n\left(X_s^n\right)-A_0(X_s)\right|^\alpha\,\mathrm{d}s\right]\mathrm{d}\gamma_d=0.$$

Now for regularized vector fields A_i^n, $i=0,1,\ldots,m$, we have

$$X_t^n(x)=x+\sum_{i=1}^m\int_0^t A_i^n\left(X_s^n\right)\mathrm{d}w_s^i+\int_0^t A_0^n\left(X_s^n\right)\mathrm{d}s. \tag{5.24}$$

When $n\to+\infty$, by Theorem 5.3 and Proposition 5.5, the two sides of (5.24) converge respectively to X and

$$x+\sum_{i=1}^m\int_0^{\cdot} A_i(X_s)\,\mathrm{d}w_s^i+\int_0^{\cdot} A_0(X_s)\,\mathrm{d}s$$

in the space $L^\alpha(\Omega\times\mathbb{R}^d;C([0,T_0],\mathbb{R}^d))$. Therefore, for almost all $x\in\mathbb{R}^d$, the following equality holds $\mathbb{P}$-almost surely:

$$X_t(x)=x+\sum_{i=1}^m\int_0^t A_i(X_s)\,\mathrm{d}w_s^i+\int_0^t A_0(X_s)\,\mathrm{d}s,\quad \text{for all } t\in[0,T_0].$$

That is to say, X_t solves SDE (5.1) over the interval $[0,T_0]$.

The following result proves pathwise uniqueness of SDE (5.1) for a.e. initial value $x\in\mathbb{R}^d$.

Proposition 5.6. *Under the conditions* (A1′) *and* (A2)–(A4)*, the* SDE (5.1) *has a unique solution on the interval* $[0,T_0]$.

Proof. Let $(Y_t)_{t\in[0,T_0]}$ be another solution. Set, for $R>0$,

$$G_R=\Big\{(w,x)\in\Omega\times\mathbb{R}^d\colon \sup_{0\leqslant t\leqslant T_0}\left|X_t(w,x)-Y_t(w,x)\right|\leqslant R\Big\}.$$

Remark that in Theorem 5.2, the terms involving $1/\sigma$ and $1/\sigma^2$ vanish. Therefore the term

$$\begin{aligned}I&:=\mathbb{E}\int_{G_R}\log\left(\frac{\sup_{0\leqslant t\leqslant T_0}|X_t-Y_t|^2}{\sigma^2}+1\right)\mathrm{d}\gamma_d\\&\leqslant C_{T_0}\Lambda_{p,T_0}C_{d,q,R}\left[\|A_0\|_{\mathbb{D}_1^q}+\left(\sum_{i=1}^m\|A_i\|_{\mathbb{D}_1^{2q}}^2\right)^{1/2}+\sum_{i=1}^m\|A_i\|_{\mathbb{D}_1^{2q}}^2\right]\end{aligned}$$

is bounded for any $\sigma > 0$. For $\eta > 0$ consider

$$\Sigma_\eta = \Big\{(w, x) \in \Omega \times \mathbb{R}^d\colon \sup_{0 \leqslant t \leqslant T_0} \big|X_t(w, x) - Y_t(w, x)\big| \geqslant \eta\Big\}.$$

Similar to (5.22), we have

$$\mathbb{E}\bigg[\int_{G_R} \mathbf{1}_{\Sigma_\eta}\, \mathrm{d}\gamma_d\bigg] \leqslant \frac{I}{\log(\eta^2/\sigma^2 + 1)} \to 0, \quad \text{as } \sigma \to 0.$$

Hence we obtain

$$\mathbf{1}_{G_R} \sup_{0 \leqslant t \leqslant T_0} |X_t - Y_t| = 0, \quad (\mathbb{P} \times \gamma_d)\text{-a.s.}$$

Letting $R \to \infty$, we obtain that $(\mathbb{P} \times \gamma_d)$-almost surely, $X_t = Y_t$ for all $t \in [0, T_0]$. □

Now we extend the solution to any time interval $[0, T]$. Let $\theta_{T_0} w$ be the time-shift of the Brownian motion w by T_0 and denote by $X_t^{T_0}$ the corresponding solution to SDE driven by $\theta_{T_0} w$. By Proposition 5.6, $\{X_t^{T_0}(\theta_{T_0} w, x)\colon 0 \leqslant t \leqslant T_0\}$ is the unique solution to the following SDE over $[0, T_0]$:

$$X_t^{T_0}(x) = x + \sum_{i=1}^m \int_0^t A_i\big(X_s^{T_0}(x)\big)\, \mathrm{d}(\theta_{T_0} w)_s^i + \int_0^t A_0\big(X_s^{T_0}(x)\big)\, \mathrm{d}s.$$

For $t \in [0, T_0]$, define $X_{t+T_0}(w, x) = X_t^{T_0}(\theta_{T_0} w, X_{T_0}(w, x))$. Note that X_t is well defined on the interval $[0, 2T_0]$ up to a $(\mathbb{P} \times \gamma_d)$-negligible subset of $\Omega \times \mathbb{R}_d$. Replacing x by $X_{T_0}(x)$ in the above equation, we get easily

$$X_{t+T_0}(x) = x + \sum_{i=1}^m \int_0^{t+T_0} A_i\big(X_s(x)\big)\, \mathrm{d}w_s^i + \int_0^{t+T_0} A_0\big(X_s(x)\big)\, \mathrm{d}s.$$

Therefore X_t defined as above is a solution to SDE on the interval $[0, 2T_0]$. Continuing in this way, the solution of SDE (5.1) on the interval $[0, T]$ is obtained.

Theorem 5.7. *The family $\{X_t\colon t \in [0, T]\}$ constructed above is the unique solution to* SDE (5.1) *in the sense of Definition* 5.1*. Moreover, for each $t \in [0, T]$, the density K_t of $(X_t)_\# \gamma_d$ with respect to γ_d is in $L^1 \log L^1$.*

Proof. Let Y_t, $t \in [0, T]$ be another solution in the sense of Definition 5.1. First by Proposition 5.6, we have $(\mathbb{P} \times \gamma_d)$-almost surely, $Y_t = X_t$ for all $t \in [0, T_0]$. In particular, $Y_{T_0} = X_{T_0}$. Next by the flow property, Y_{t+T_0} satisfies the following equation:

$$Y_{t+T_0}(x) = Y_{T_0}(x) + \sum_{i=1}^{m} \int_0^t A_i\big(Y_{s+T_0}(x)\big)\, \mathrm{d}(\theta_{T_0} w)_s^i + \int_0^t A_0\big(Y_{s+T_0}(x)\big)\, \mathrm{d}s,$$

that is, Y_{t+T_0} is a solution with initial value Y_{T_0}. But by the above discussion, X_{t+T_0} is also a solution with the same initial value $X_{T_0} = Y_{T_0}$. Again by Proposition 5.6, we have $(\mathbb{P} \times \gamma_d)$-almost surely, $X_{t+T_0} = Y_{t+T_0}$ for all $t \leqslant T_0$. Hence we proved that $X|[0, 2T_0] = Y|[0, 2T_0]$. Repeating this procedure, we obtain the uniqueness over $[0, T]$. Existence of the density K_t of $(X_t)_\#\gamma_d$ with respect to γ_d beyond T_0 is deduced from the flow property. However, to ensure that $K_t \in L^1 \log L^1$, we have to use Theorem 3.3 and the fact that

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \mathbb{E}\Big[\sup_{0\leqslant t\leqslant T} \big|X_t^n - X_t\big|^\alpha \Big]\, \mathrm{d}\gamma_d = 0,$$

which can be checked using the same arguments as in the proof of Propositions 4.1 and 4.2. □

Appendix A

For any locally integrable function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $R > 0$, the local maximal function $M_R f$ is defined by

$$M_R f(x) = \sup_{0<r\leqslant R} \frac{1}{\mathrm{Leb}_d(B(x,r))} \int_{B(x,r)} \big|f(y)\big|\, \mathrm{d}y, \tag{A.1}$$

where $B(x,r) = \{y \in \mathbb{R}^d\colon |y - x| \leqslant r\}$. The following result is the starting point for the approach concerning Sobolev coefficients, used in [5,36].

Theorem A.1. *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ be such that $\nabla f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Then there is a constant $C_d > 0$ (independent of f) and a negligible subset N, such that for $x, y \in N^c$ with $|x - y| \leqslant R$,*

$$\big|f(x) - f(y)\big| \leqslant C_d |x - y| \big(\big(M_R|\nabla f|\big)(x) + \big(M_R|\nabla f|\big)(y)\big). \tag{A.2}$$

Moreover for $p > 1$ and $f \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$, there is a constant $C_{d,p} > 0$ such that

$$\int_{B(r)} (M_R f)^p\, \mathrm{d}x \leqslant C_{d,p} \int_{B(r+R)} |f|^p\, \mathrm{d}x. \tag{A.3}$$

Since inequality (A.2) plays a key role in the proof of Theorem 5.2, we include its proof for the sake of the reader's convenience.

Proof of estimate (A.2). We follow the idea of the proof of Claim #2 on p. 253 in [9]. For any bounded measurable subset U in $\mathbb{R}^d$ of Lebesgue measure $\mathrm{Leb}_d(U) > 0$, define the average of

$f \in L^1_{loc}(\mathbb{R}^d)$ on U by

$$(f)_U = \fint_U f(y)\,\mathrm{d}y := \frac{1}{\mathrm{Leb}_d(U)} \int_U f(y)\,\mathrm{d}y.$$

Write $(f)_{x,r}$ instead of $(f)_{B(x,r)}$ for simplicity. Then $M_R f(x) = \sup_{0<r\leqslant R}(|f|)_{x,r}$. We use the following simple inequality: for any $C \in \mathbb{R}$,

$$\bigl|(f)_U - C\bigr| \leqslant \fint_U \bigl|f(y) - C\bigr|\,\mathrm{d}y. \tag{A.4}$$

First, for any $x \in \mathbb{R}^d$ and $r \in]0, R]$, by Poincaré's inequality with $p = 1$ and $p^* = d/(d-1)$ (see [9, p. 141]), there exists $C_d > 0$ such that

$$\begin{aligned}\fint_{B(x,r)} \bigl|f - (f)_{x,r}\bigr|\,\mathrm{d}y &\leqslant \biggl(\fint_{B(x,r)} \bigl|f - (f)_{x,r}\bigr|^{d/(d-1)}\,\mathrm{d}y\biggr)^{(d-1)/d} \\ &\leqslant C_d r \fint_{B(x,r)} |\nabla f|\,\mathrm{d}y \leqslant C_d M_R|\nabla f|(x) r. \end{aligned} \tag{A.5}$$

In particular, for any $k \geqslant 0$, by (A.4) and (A.5),

$$\begin{aligned}\bigl|(f)_{x,r/2^{k+1}} - (f)_{x,r/2^k}\bigr| &\leqslant \fint_{B(x,r/2^{k+1})} \bigl|f - (f)_{x,r/2^k}\bigr|\,\mathrm{d}y \\ &\leqslant 2^d \fint_{B(x,r/2^k)} \bigl|f - (f)_{x,r/2^k}\bigr|\,\mathrm{d}y \\ &\leqslant 2^d C_d M_R|\nabla f|(x) r/2^k. \end{aligned}$$

Since $f \in L^1_{loc}(\mathbb{R}^d)$, there is a negligible subset $N \subset \mathbb{R}^d$ such that for all $x \in N^c$, $f(x) = \lim_{r\to 0}(f)_{x,r}$. Thus for any $x \in N^c$, by summing up the above inequality, we get

$$\bigl|f(x) - (f)_{x,r}\bigr| \leqslant \sum_{k=0}^{\infty} \bigl|(f)_{x,r/2^{k+1}} - (f)_{x,r/2^k}\bigr| \leqslant 2^{1+d} C_d M_R|\nabla f|(x) r. \tag{A.6}$$

Next for $x, y \in N^c$, $x \neq y$ and $|x - y| \leqslant R$, let $r = |x - y|$. Then by the triangular inequality, (A.4) and (A.5),

$$\begin{aligned}\bigl|(f)_{x,r} - (f)_{y,r}\bigr| &\leqslant \fint_{B(x,r)\cap B(y,r)} \bigl(\bigl|(f)_{x,r} - f(z)\bigr| + \bigl|f(z) - (f)_{y,r}\bigr|\bigr)\,\mathrm{d}z \\ &\leqslant \tilde{C}_d \biggl[\fint_{B(x,r)} \bigl|(f)_{x,r} - f(z)\bigr|\,\mathrm{d}z + \fint_{B(y,r)} \bigl|f(z) - (f)_{y,r}\bigr|\,\mathrm{d}z\biggr] \\ &\leqslant \tilde{C}_d C_d \bigl(M_R|\nabla f|(x) + M_R|\nabla f|(y)\bigr) r. \end{aligned} \tag{A.7}$$

Now (A.2) follows from the triangular inequality and (A.6), (A.7):

$$\begin{aligned}\left|f(x)-f(y)\right| &\leqslant \left|f(x)-(f)_{x,r}\right|+\left|(f)_{x,r}-(f)_{y,r}\right|+\left|(f)_{y,r}-f(y)\right| \\ &\leqslant 2^{1+d}C_dM_R|\nabla f|(x)r+\tilde{C}_dC_d\big(M_R|\nabla f|(x)+M_R|\nabla f|(y)\big)r \\ &\quad +2^{1+d}C_dM_R|\nabla f|(y)r \\ &= C_d\big(2^{1+d}+\tilde{C}_d\big)|x-y|\big(M_R|\nabla f|(x)+M_R|\nabla f|(y)\big).\end{aligned}$$

We obtain (A.2). □

References

[1] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields, Invent. Math. 158 (2004) 227–260.
[2] L. Ambrosio, A. Figalli, On flows associated to Sobolev vector fields in Wiener space: an approach à la Di Perna–Lions, J. Funct. Anal. 256 (1) (2009) 179–214.
[3] L. Ambrosio, M. Lecumberry, S. Maniglia, Lipschitz regularity and approximate differentiability of the Di Perna–Lions flow, Rend. Sem. Mat. Univ. Padova 114 (2005) 29–50.
[4] F. Cipriano, A.B. Cruzeiro, Flows associated with irregular $\mathbb{R}^d$-vector fields, J. Differential Equations 210 (2005) 183–201.
[5] G. Crippa, C. De Lellis, Estimates and regularity results for the Di Perna–Lions flows, J. Reine Angew. Math. 616 (2008) 15–46.
[6] A.B. Cruzeiro, Équations différentielles ordinaires: non explosion et mesures quasi-invariantes, J. Funct. Anal. 54 (1983) 193–205.
[7] R.J. Di Perna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989) 511–547.
[8] B. Driver, Integration by parts and quasi-invariance for heat kernal measures on loop groups, J. Funct. Anal. 149 (1997) 470–547.
[9] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, Stud. Adv. Math., CRC Press, London, 1992.
[10] S. Fang, Canonical Brownian motion on the diffeomorphism group of the circle, J. Funct. Anal. 196 (2002) 162–179.
[11] Shizan Fang, Dejun Luo, Transport equations and quasi-invariant flows on the Wiener space, Bull. Sci. Math. (2009), doi:10.1016/j.bulsci.2009.01.001, in press.
[12] S. Fang, T. Zhang, A study of a class of stochastic differential equations with non-Lipschitzian coefficients, Probab. Theory Related Fields 132 (2005) 356–390.
[13] S. Fang, P. Imkeller, T. Zhang, Global flows for stochastic differential equations without global Lipschitz conditions, Ann. Probab. 35 (2007) 180–205.
[14] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, J. Funct. Anal. 254 (2008) 109–153.
[15] F. Flandoli, M. Gubinelli, E. Priola, Well-posedness of the transport equation by stochastic perturbation, Invent. Math. 180 (1) (2009) 1–53.
[16] Zhiyuan Huang, Foundations of Stochastic Analysis, second ed., Science Press of China, 2001 (in Chinese).
[17] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, second ed., North-Holland, Amsterdam, 1989.
[18] H. Kaneko, S. Nakao, A note on approximation for stochastic differential equations, in: Séminaire de Probabilités, XXII, in: Lecture Notes in Math., vol. 1321, Springer, Berlin, 1988, pp. 155–162.
[19] N.V. Krylov, On weak uniqueness for some diffusion with discontinuous coefficients, Stochastic Process. Appl. 113 (2004) 37–64.
[20] N.V. Krylov, M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, Probab. Theory Related Fields 131 (2005) 154–196.
[21] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, 1990.
[22] C. LeBris, P.L. Lions, Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients, Comm. Partial Differential Equations 33 (2008) 1272–1317.
[23] Y. Le Jan, O. Raimond, Integration of Brownian vector fields, Ann. Probab. 30 (2002) 826–873.
[24] Y. Le Jan, O. Raimond, Flows, coalescence and noise, Ann. Probab. 32 (2004) 1247–1315.
[25] X.-M. Li, Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds, Probab. Theory Related Fields 100 (1994) 485–511.

[26] X.-M. Li, M. Scheutzow, Lack of strong completeness for stochastic flows, http://arxiv.org/abs/0908.1839.
[27] Dejun Luo, Quasi-invariance of Lebesgue measure under the homeomorphic flow generated by SDE with non-Lipschitz coefficient, Bull. Sci. Math. 133 (2009) 205–228.
[28] P. Malliavin, Stochastic Analysis, Grundlehren Math. Wiss., vol. 313, Springer, 1997.
[29] P. Malliavin, The canonical diffusion above the diffeomorphism group of the circle, C. R. Acad. Sci. 329 (1999) 325–329.
[30] D. Revuz, M. Yor, Continuous Martingale and Brownian Motion, Grundlehren Math. Wiss., vol. 293, Springer-Verlag, 1991.
[31] M.V. Safonov, Non-uniqueness for second order elliptic equations with measurable coefficients, SIAM J. Math. Anal. 30 (1999) 879–895.
[32] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, Springer, New York, 1979.
[33] A.J. Veretennikov, On the strong solutions of stochastic differential equations, Theory Probab. Appl. 24 (1979) 354–366.
[34] Xicheng Zhang, Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients, Stochastic Process. Appl. 115 (2005) 1805–1818.
[35] Xicheng Zhang, Homeomorphic flows for multi-dimensional SDEs with non-Lipschitz coefficients, Stochastic Process. Appl. 115 (3) (2005) 435–448, Erratum to “Homeomorphic flows for multi-dimensional SDEs with non-Lipschitz coefficients”.
[36] Xicheng Zhang, Stochastic flows of SDEs with irregular coefficients and stochastic transport equations, Bull. Sci. Math. (2009), doi:10.1016/j.bulsci.2009.12.004, in press.