North-Western European Journal of Mathematics

A representation for the derivative with respect to the initial data of the solution of an SDE with a non-regular drift

Olga Viktorovna Aryasova[1] Andrey Yurievich Pilipenko[2]

Received: June 10, 2016/Accepted: January 10, 2017/Online: January 25, 2017

Abstract

We consider a multidimensional SDE with a Gaussian noise and a drift vector being a vector function of bounded variation. We prove the existence of generalized derivative of the solution with respect to the initial conditions and represent the derivative as a solution of a linear SDE with coefficients depending on the initial process. The obtained representation is a natural generalization of the expression for the derivative in the smooth case. In the proof we use the results on continuous additive functionals.

Keywords: Stochastic flow; Continuous additive functional; Differentiability with respect to initial data.

MSC: 60J65, 60H10.

Introduction

Consider a d-dimensional nonhomogeneous stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}\varphi_t(x) = a(t,\varphi_t(x))\,\mathrm{d}t + \sum_{k=1}^{m} \sigma_k(t,\varphi_t(x))\,\mathrm{d}w_k(t), \\ \varphi_0(x) = x, \end{cases} \tag{1}$$

where $x \in \mathbb{R}^d$, $d \geq 1$, $m \geq 1$, $(w(t))_{t\geq 0} = (w_1(t),\ldots,w_m(t))_{t\geq 0}$ is a standard m-dimensional Wiener process, the drift coefficient $a : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ is Borel measurable and bounded, and the diffusion coefficient $\sigma : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d\times\mathbb{R}^m$ is bounded and continuous.

In what follows we assume that σ satisfies the following conditions:

[1] Institute of Geophysics, National Academy of Sciences of Ukraine, Palladin pr. 32, 03680, Kiev-142, Ukraine

[2] Institute of Mathematics, National Academy of Sciences of Ukraine, Tereshchenkivska str. 3, 01601, Kiev, Ukraine; National Technical University of Ukraine "KPI", Kiev, Ukraine

(C_1) $\sigma \in W^{0,1}_{2d+2,\mathrm{loc}}([0,\infty)\times\mathbb{R}^d)$.

(C_2) *Uniform ellipticity*. For each $T>0$, there exists an ellipticity constant $B>0$ such that for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $\theta\in\mathbb{R}^d$,

$$\theta^*\sigma(t,x)\sigma^*(t,x)\theta \geq B|\theta|^2,$$

where $|\cdot|$ is a norm in $\mathbb{R}^d$.

(C_3) *Hölder continuity*. For each $T>0$, there exist $L>0$, $0<\alpha\leq 1$ such that for all $t_1,t_2\in[0,T]$, $x_1,x_2\in\mathbb{R}^d$, $1\leq i\leq d$, $1\leq k\leq m$,

$$\left|\sigma^i_k(t_1,x_1)-\sigma^i_k(t_2,x_2)\right| \leq L\left(|t_1-t_2|^{\alpha/2}+|x_1-x_2|^{\alpha}\right).$$

Under these assumptions on the coefficients there exists a unique strong solution to Equation (1) on the previous page[3]).

It is well known[4] that if the coefficients of Equation (1) on the previous page are continuously differentiable in the spatial variable and the derivatives are bounded and Hölder continuous uniformly in t, then there exists a modification of $\varphi_t(x)$ (denoted by the same symbol) which is continuous in (t,x) and continuously differentiable in x almost surely. Moreover, the derivative $\nabla\varphi_t(x) := Y_t(x)$ is a solution of the equation

$$\mathrm{d}Y_t(x) = \nabla a(t,\varphi_t(x))Y_t(x)\,\mathrm{d}t + \sum_{k=1}^{m}\nabla\sigma_k(t,\varphi_t(x))Y_t(x)\,\mathrm{d}w_k(t), \tag{2}$$

where for a function $f:\mathbb{R}^d\to\mathbb{R}^d$, we set $\nabla f = \left(\frac{\partial f^i}{\partial x_j}\right)_{1\leq i,j\leq d}$.

Flandoli, Gubinelli, and Priola (2010) show that in the case of a smooth, bounded, uniformly non-degenerate noise and a bounded, uniformly in time Hölder continuous drift term the conditions on the coefficients can be essentially weakened, and the solution is continuously differentiable with respect to the spatial parameter.

The case of discontinuous drift is studied in Fedrizzi and Flandoli (2013a,b), Meyer-Brandis and Proske (2010), and Mohammed, Nilssen, and Proske (2015) and the weak differentiability of the solution to Equation (1) on the previous page is proved under rather weak assumptions on the drift. Fedrizzi and Flandoli (2013a) consider Equation (1) on the previous page with the identity diffusion matrix and a drift vector belonging to $L_q(0,T;L_p(\mathbb{R}^d))$ for some p, q such that

$$p\geq 2,\quad q>2,\quad \frac{d}{p}+\frac{2}{q}<1.$$

[3]See Veretennikov, 1981, "On strong solutions and explicit formulas for solutions of stochastic integral equations".

[4]Kunita, 1990, *Stochastic Flows and Stochastic Differential Equations*.

Applying a Zvonkin-type transformation they establish the existence of the Gâteaux derivative with respect to the initial data in $L_2(\Omega\times[0,T];\mathbb{R}^d)$. The authors of Meyer-Brandis and Proske (2010) and Mohammed, Nilssen, and Proske (2015) use the Malliavin calculus. In particular, Mohammed, Nilssen, and Proske (2015) prove that the solution of Equation (1) on p. 1 with a bounded measurable drift vector a and the identity diffusion matrix belongs to the space $L^2(\Omega;W^{1,p}(U))$ for each $t\in\mathbb{R}^d, p>1$, and any open and bounded $U\in\mathbb{R}^d$. Unfortunately, in these works no representations for the derivatives are given.

The one-dimensional case is considered in Aryasova and Pilipenko (2012) and Attanasio (2010) and explicit expressions for the Sobolev derivative are obtained. The formulas involve the local time of the initial process. There is no direct generalization of the expressions for the Sobolev derivative to the multidimensional case because the local time at a point does not exist in the multidimensional situation.

The aim of this paper is to get a natural representation for the derivative $\nabla_x\varphi_t(x)$ of the solution to Equation (1) on p. 1. We assume that σ satisfies conditions (C_1) to (C_3) on the preceding page, and for some $\rho>0$ and all $1\le k\le m$, $1\le i,j\le d$, the function $\left|\frac{\partial\sigma_k^i}{\partial y_j}(s,y)\right|^{2+\rho}$ belongs to the Kato-type class $\mathcal{K}$, i.e.,

$$\lim_{t\downarrow 0}\sup_{\substack{t_0\in[0,\infty)\\ x_0\in\mathbb{R}^d}}\int_{t_0}^{t_0+t}\mathrm{d}s\int_{\mathbb{R}^d}\frac{1}{(2\pi(s-t_0))^{d/2}}\exp\left\{-\frac{\left|y-x_0\right|^2}{2(s-t_0)}\right\}\times\left|\frac{\partial\sigma_k^i}{\partial y_j}(s,y)\right|^{2+\rho}\mathrm{d}y=0.$$

We show that the derivative $Y_t(x)=\nabla_x\varphi_t(x)$ is a solution to the SDE

$$Y_t(x)=E+\int_0^t\mathrm{d}A_s(\varphi(x))Y_s(x)+\sum_{k=1}^m\int_0^t\nabla\sigma_k(s,\varphi_s(x))Y_s(x)\,\mathrm{d}w_k(s),\tag{3}$$

where E is the d-dimensional identity matrix, $A_s(\varphi(x))$ is a continuous additive functional of the process $(t,\varphi_t(x))_{t\ge 0}$, which is equal to $\int_0^t\nabla a(s,\varphi_s(x))\,\mathrm{d}s$ if a is differentiable. This representation is a natural generalization of the expressions for the smooth case.

We prove the main result under the assumption that for each $t\ge 0$ and all $1\le i\le d$, $a^i(t,\cdot)$ is a function of bounded variation on $\mathbb{R}^d$, i.e., for each $1\le j\le d$, the generalized derivative $\mu^{ij}(t,\mathrm{d}y)=\frac{\partial a^i}{\partial x_j}(t,\mathrm{d}y)$ is a signed measure on $\mathbb{R}^d$. Besides, we suppose that for all $1\le i,j\le d$, $\mu^{ij}(t,\mathrm{d}y)\,\mathrm{d}t$ is of the class $\mathcal{K}$, i.e.,

$$\lim_{t\downarrow 0}\sup_{\substack{t_0\in[0,\infty)\\ x_0\in\mathbb{R}}}\int_{t_0}^{t_0+t}\mathrm{d}s\int_{\mathbb{R}^d}\frac{1}{(2\pi(s-t_0))^{d/2}}\exp\left\{-\frac{\left|y-x_0\right|^2}{2(s-t_0)}\right\}\left|\mu\right|^{ij}(s,\mathrm{d}y)=0,$$

where $\left|\mu\right|^{ij}=\mu^{ij,+}+\mu^{ij,-}$ is the variation of μ^{ij}; $\mu^{ij,+}$, $\mu^{ij,-}$ are measures from the Hahn-Jordan decomposition $\mu^{ij}=\mu^{ij,+}-\mu^{ij,-}$.

Similar results for a homogeneous SDE with the identity diffusion matrix and a drift being a vector function of bounded variation are obtained in Aryasova and Pilipenko (2014). In this case, there is no martingale member in the right-hand side of Equation (3) on the previous page. This essentially simplifies the proof. The argument is based on the theory of additive functionals of homogeneous Markov processes developed by Dynkin (1965). In Bogachev and Pilipenko (2015) the same method is applied to a homogeneous SDE with Lévi noise and a drift being a vector function of bounded variation. The authors prove the existence of a strong solution and the differentiability of the solution with respect to the initial data. Unfortunately, Dynkin's theory can not be directly applied to our problem because $(\varphi_t(x))_{t\geq 0}$ is not homogeneous.

The paper is organized as follows. In Section 1 we collect some facts from the theory of additive functionals of homogeneous Markov processes[5]. We consider a homogeneous process $(t,\varphi_t)_{t\geq 0}$ and adapt Dynkin's theory to the functionals of this process. The main result is formulated in Section 2 on p. 14 and proved in Section 3 on p. 15. The idea of the proof is to approximate the solution of Equation (1) on p. 1 by solutions of equations with smooth coefficients. The key point is the convergence of continuous homogeneous additive functionals of the approximating processes to a functional of the process being the solution to Equation (1) on p. 1 (Lemma 6 on p. 21). The proof of the corresponding statement uses essentially the result on the convergence of the transition probability densities of the approximating processes, which is obtained in Section 4 on p. 29.

The suggested method can be considered as a generalization of the local time approach used in the one-dimensional case.

1 Preliminaries: continuous additive functionals

Let $(\xi_t, \mathscr{F}_t, P_z)$ be a càdlàg homogeneous Markov process with a phase space $(E, \mathscr{B})$, where σ-algebra $\mathscr{B}$ contains all one-point sets[6]. Assume that $(\xi_t)_{t\geq 0}$ has the infinite life-time. Denote $\mathcal{N}_t = \sigma\{\xi_s : 0 \leq s \leq t\}$.

Definition 1 – A random function A_t, $t \geq 0$, adapted to the filtration $\{\mathcal{N}_t\}$ is called a non-negative continuous additive functional of the process $(\xi_t)_{t\geq 0}$ if it is

- non-negative;
- continuous in t;
- homogeneous additive, i.e., for all $t \geq 0$, $s > 0$, $z \in E$,

$$A_{t+s} = A_s + \theta_s A_t \quad P_z - \text{almost surely,} \tag{4}$$

  where θ is the shift operator.

[5] Dynkin, 1965, *Markov Processes*.
[6] See notations in ibid.

If in addition for each $t \geq 0$,

$$\sup_{z \in E} \mathbb{E}_z A_t < \infty,$$

then A_t, $t \geq 0$, is called a W-functional.

Remark 1 – It follows from Definition 1 on the preceding page that a W-functional is non-decreasing in t, and for all $z \in E$

$$P_z\{A_0 = 0\} = 1.$$

Definition 2 – The function

$$f_t(z) = \mathbb{E}_z A_t$$

is called the characteristic of a W-functional A_t.

Remark 2 (Dynkin[7]) – For all $s \geq 0$, $t \geq 0$,

$$\|f_{t+s}\|_E \leq \|f_t\|_E + \|f_s\|_E,$$

where $\|f_t\|_E = \sup_{z \in E} |f_t(z)|$.

The following theorem states the relation between the convergence of W-functionals and the convergence of their characteristics.

Theorem 1 (Dynkin[8]) – *Let $A_{n,t}$, $n \geq 1$, be W-functionals of the process $(\xi_t)_{t \geq 0}$ and $f_{n,t}(z) = \mathbb{E}_z A_{n,t}$ be their characteristics. Suppose that for each $t > 0$, a function $f_t(z)$ satisfies the condition*

$$\lim_{n \to \infty} \sup_{0 \leq u \leq t} \sup_{z \in E} \left| f_{n,u}(z) - f_u(z) \right| = 0. \tag{5}$$

Then $f_t(z)$ is the characteristic of a W-functional A_t. Moreover,

$$A_t = \underset{n \to \infty}{\text{l.i.m.}} A_{n,t},$$

where l.i.m. *denotes the convergence in the mean square sense (for any initial distribution ξ_0).*

Proposition 1 (Dynkin[9]) – *If for any $t \geq 0$ the sequence of non-negative additive functionals $\{A_{n,t} : n \geq 1\}$ of the Markov process $(\xi_t)_{t \geq 0}$ converges in probability to a continuous functional A_t, then the convergence in probability is uniform, i.e.*

$$\forall\ T > 0 \sup_{t \in [0,T]} \left| A_{n,t} - A_t \right| \to 0,\ n \to \infty, \textit{ in probability.}$$

[7]Dynkin, 1965, *Markov Processes*, Properties 6.15.
[8]Ibid., Theorem 6.3.

Example 1 – Let $E = \mathbb{R}^d$, h be a non-negative bounded measurable function on E, and suppose that the process $(\xi_t)_{t\geq 0}$ has a transition probability density $g_t(z_1, z_2)$. Then

$$A_t := \int_0^t h(\xi_s)\,\mathrm{d}s$$

is a W-functional of the process $(\xi_t)_{t\geq 0}$ and its characteristic is equal to

$$f_t(z) = \int_E \left(\int_0^t g_s(z,v)\,\mathrm{d}s \right) h(v)\,\mathrm{d}v = \int_E k_t(z,v) h(v)\,\mathrm{d}v,$$

where

$$k_t(z,v) = \int_0^t g_s(z,v)\,\mathrm{d}s.$$

Let a measure ν be such that $\int_E k_t(z,v)\nu(\mathrm{d}v)$ is well defined. If we can choose a sequence of non-negative bounded continuous functions $\{h_n : n \geq 1\}$ such that for each $T > 0$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \sup_{z\in E} \left| \int_E k_t(z,v) h_n(v)\,\mathrm{d}v - \int_E k_t(z,v)\nu(\mathrm{d}v) \right| = 0,$$

then by Theorem 1 on the previous page there exists a W-functional A_t^ν corresponding to the measure ν with its characteristic being equal to $\int_E k_t(z,v)\nu(\mathrm{d}v)$.

For a given measure ν, we have a sufficient condition for the existence of a corresponding W-functional.

Theorem 2 (Dynkin[10]) – *Suppose that*

$$\lim_{t\downarrow 0} \sup_{z\in E} f_t(z) = \lim_{t\downarrow 0} \sup_{z\in E} \int_E k_t(z,y)\nu(\mathrm{d}y) = 0. \tag{6}$$

Then $f_t(z)$ is the characteristic of a W-functional A_t^ν. Moreover,

$$A_t^\nu = \operatorname*{l.i.m.}_{\varepsilon\downarrow 0} \int_0^t \frac{f_\varepsilon(\xi_u)}{\varepsilon}\,\mathrm{d}u,$$

and the sequence of characteristics of integral functionals $\int_0^t \frac{f_\varepsilon(\xi_u)}{\varepsilon}\,\mathrm{d}u$ converges to $f_t(z)$ in sense of Equation (5) *on the previous page.*

[9] Dynkin, 1965, *Markov Processes*, Lemma 6.1'.

Consider a process $\eta_t = (\eta_t^1, \eta_t^2)$, $t \geq 0$, which is a (unique) solution to the system of SDEs:

$$\begin{cases} d\eta_t^1 = dt, \\ d\eta_t^2 = a(\eta_t^1, \eta_t^2)\,dt + \sum_{k=1}^{m} \sigma_k(\eta_t^1, \eta_t^2)\,dw_k(t). \end{cases} \tag{7}$$

For the initial condition $\eta_0^1 = t_0$, $\eta_0^2 = x_0$, denote the corresponding distribution of the process $(\eta_t)_{t\geq 0}$ by $\mathbf{P}_{t_0,x_0}$.

Since $(\eta_t)_{t\geq 0}$ is a homogeneous Markov process, we can apply for its investigation the theory of additive functionals.

Let h be a non-negative bounded measurable function on $E = [0,\infty) \times \mathbb{R}^d$. Then (cf. Example 1 on the preceding page)

$$A_t = \int_0^t h(\eta_s)\,ds$$

is a W-functional of the process $(\eta_t)_{t\geq 0}$. Its characteristic is equal to

$$\begin{aligned} f_t(t_0, x_0) &= \mathbb{E}_{t_0,x_0} \int_0^t h(\eta_s)\,ds = \int_0^t ds \int_{\mathbb{R}^d} G(t_0, x_0, t_0 + s, y) h(t_0 + s, y)\,dy \\ &= \int_{t_0}^{t_0+t} ds \int_{\mathbb{R}^d} G(t_0, x_0, s, y) h(s, y)\,dy, \end{aligned} \tag{8}$$

where $G(s,x,t,y)$, $0 \leq s \leq t$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, is the transition probability density of the process $(\eta_t^2)_{t\geq 0}$.

Let a measure ν on $[0,\infty) \times \mathbb{R}^d$ be such that $\int_{t_0}^{t_0+t} \int_{\mathbb{R}^d} G(t_0, x_0, s, y)\nu(ds, dy) < \infty$ for all $t \geq 0$, $t_0 \geq 0$, $x_0 \in \mathbb{R}^d$. If there exists a sequence of non-negative bounded continuous functions $\{h_n : n \geq 1\}$ such that for each $T > 0$,

$$\lim_{n\to\infty} \sup_{\substack{t\in[0,T] \\ t_0\in[0,\infty) \\ x_0\in\mathbb{R}^d}} \left| \int_{t_0}^{t_0+t} ds \int_{\mathbb{R}^d} G(t_0, x_0, s, y) h_n(s, y)\,dy - \int_{t_0}^{t_0+t} \int_{\mathbb{R}^d} G(t_0, x_0, s, y)\nu(ds, dy) \right| = 0,$$

then by Theorem 1 on p. 5 there exists a W-functional corresponding to the measure ν with its characteristic being equal to $\int_{t_0}^{t_0+t} \int_{\mathbb{R}^d} G(t_0, x_0, s, y)\nu(ds, dy)$.

Theorem 3 (Corollary of Theorem 2 on the preceding page) – *Suppose that*

$$\lim_{t\downarrow 0} \sup_{\substack{t_0\in[0,\infty) \\ x_0\in\mathbb{R}^d}} f_t(t_0, x_0) = \lim_{t\downarrow 0} \sup_{\substack{t_0\in[0,\infty) \\ x_0\in\mathbb{R}^d}} \int_{t_0}^{t_0+t} \int_{\mathbb{R}^d} G(t_0, x_0, s, y)\nu(ds, dy) = 0. \tag{9}$$

---

[10] Dynkin, 1965, *Markov Processes*, Theorem 6.6.

Then $f_t(z), z \in [0,\infty)\times\mathbb{R}^d$, *is the characteristic of a W-functional* A_t^ν. *Moreover,*

$$A_t^\nu = \underset{\varepsilon\downarrow 0}{\mathrm{l.i.m.}} \int_0^t \frac{f_\varepsilon(\eta_u)}{\varepsilon}\,\mathrm{d}u,$$

and the sequence of characteristics of integral functionals $\int_0^t \frac{f_\varepsilon(\eta_u)}{\varepsilon}\,\mathrm{d}u$ *converges to* $f_t(z)$ *in sense of Equation* (5) *on p. 5.*

Let $\eta_t(t_0,x_0) = (\eta_t^1(t_0,x_0), \eta_t^2(t_0,x_0))$ be a solution of Equation (7) on the previous page with the initial condition (t_0,x_0) and defined on a probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$. Let $\mathbb{P}_{t_0,x_0}$ be the distribution of the process $(\eta_t(t_0,x_0))_{t\ge 0}$, where $t_0\ge 0$, $x_0\in\mathbb{R}^d$. In Dynkin's notation[11] $(\eta_t(t_0,x_0))_{t\ge0},\mathcal{F}_t,\mathbb{P})$, $t_0\ge 0$, $x_0\in\mathbb{R}^d$, is called a Markov family of random functions.

Suppose that the measure ν satisfies the condition of Theorem 3 on the previous page. Then there exists a W-functional A_t^ν of the process $(\eta_t)_{t\ge0}$. According to the definition of W-functionals, the functional is measurable w.r.t. σ-algebra generated by the process $(\eta_t)_{t\ge0}$. Since the process $(\eta_t)_{t\ge0}$ is continuous and has the infinite life-time, we can consider $A_t^\nu = A_t^\nu(\cdot)$ as a measurable function on $[0,\infty)\times C([0,\infty),\mathbb{R}^d)$ that depends only on the behavior of the process on $[0,t]$. The composition $A_t^\nu(\eta_\cdot(t_0,x_0))$, $t\ge0$, is called a W-functional of $(\eta_t(t_0,x_0))_{t\ge0}$ corresponding to the measure ν. The function $A_t^\nu(\eta_\cdot(t_0,x_0))$ is defined for all $t_0\ge0, x_0\in\mathbb{R}^d$.

If $t_0=0$, $x_0=x$, the process $\eta_t^2(0,x)=\eta_t^2(x)$ is the solution of Equation (1) on p. 1 starting from x and therefore $\eta_t^2(x)=\varphi_t(x)$. Then $\eta_t(0,x)=(t,\varphi_t(x))$. Since the first coordinate $\eta_t^1(t_0,x_0)=t_0+t$ is non-random, we denote $A_t^\nu(\eta_\cdot(0,x))$ as $A_t^\nu(\varphi_\cdot(x))$.

Let us formulate the sufficient condition for Equation (9) on the previous page. If a and σ are bounded and measurable, and σ satisfies conditions conditions (C_2) and (C_3) on p. 2 then the transition probability density of the process $(\eta_2(t))_{t\ge0}$ satisfies the Gaussian estimate[12]:

$$G(s,x,t,y) \le \frac{C}{(t-s)^{d/2}}\exp\left\{-c\frac{|y-x|^2}{t-s}\right\} \tag{10}$$

valid in every domain of the form $0\le s<t\le T$, $x\in\mathbb{R}^d$, $y\in\mathbb{R}^d$, where $T>0$. Constants C, c are positive and depend only on d, T, $\|a\|_{T,\infty}$, $\|\sigma_k\|_{T,\infty}$, $1\le k\le m$, and the ellipticity constant B, where $\|a\|_{T,\infty}=\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\|a(t,x)\|$.

Denote by $p_0(s,x,t,y)$ the transition probability density of a Wiener process:

$$p_0(s,x,t,y) = \frac{1}{(2\pi(t-s))^{d/2}}\exp\left\{-\frac{|y-x|^2}{2(t-s)}\right\}. \tag{11}$$

[11] See Dynkin, 1965, *Markov Processes.*

[12] See N. I. Portenko, 1990, *Generalized Diffusion Processes*, Ch. II.

The following definition is analogous to the definition of the Kato class[13].

Definition 3 – A measure ν on $[0,\infty)\times\mathbb{R}^d$ is of the class $\mathcal{K}$, if

$$\lim_{t\downarrow 0}\sup_{\substack{t_0\in[0,\infty)\\ x_0\in\mathbb{R}^d}}\int_{t_0}^{t_0+t}\int_{\mathbb{R}^d} p_0(t_0,x_0,s,y)\nu(\mathrm{d}s,\mathrm{d}y)=0. \tag{12}$$

Taking into account Equation (10) on the preceding page it is easy to see that if ν is of the class $\mathcal{K}$ then it satisfies condition Equation (9) on p. 7.

Definition 4 – A signed measure ν is of the class $\mathcal{K}$ if the measure $|\nu|$ is of the class $\mathcal{K}$, where $|\nu|=\nu^+ + \nu^-$ is the variation of ν. Here ν^+, ν^- are the measures from the Hahn-Jordan decomposition $\nu=\nu^+ - \nu^-$.

Let $\nu=\nu^+ - \nu^-$ be a signed measure belonging to $\mathcal{K}$. Then by Theorem 2 on p. 6 there exist W-functionals $A_t^{\nu^\pm}$. Denote $A_t^{\nu}=A_t^{\nu^+}-A_t^{\nu^-}$.

Remark 3 – Suppose that the signed measure ν can be represented in the form $\nu=\widetilde{\nu}^+ - \widetilde{\nu}^-$, where $\widetilde{\nu}^+$, $\widetilde{\nu}^-$ are of the class $\mathcal{K}$ but are not necessarily orthogonal. Then one can see that $A_t^{\nu^+}-A_t^{\nu^-}=A_t^{\widetilde{\nu}^+}-A_t^{\widetilde{\nu}^-}$.

In what follows we often deal with measures which have densities with respect to the Lebesgue measure on $[0,\infty)\times\mathbb{R}^d$.

Definition 5 – A measurable function h on $[0,\infty)\times\mathbb{R}^d$ is called a function of the class $\mathcal{K}$ if the signed measure $\nu(\mathrm{d}s,\mathrm{d}y)=h(s,y)\,\mathrm{d}s\,\mathrm{d}y$ is of the class $\mathcal{K}$.

Remark 4 – Let $\nu(\mathrm{d}s,\mathrm{d}x)=\mu(\mathrm{d}x)\mathrm{d}s$, where μ is a measure on $\mathbb{R}^d$. Then the Equation (12) transforms into the following one

$$\lim_{t\downarrow 0}\sup_{x_0\in\mathbb{R}^d}\int_0^t \mathrm{d}s\int_{\mathbb{R}^d} p_0(0,x_0,s,y)\mu(\mathrm{d}y)=0. \tag{13}$$

It is proved[14] that μ satisfies the condition Equation (13) if and only if

$$\sup_{x\in\mathbb{R}}\int_{|x-y|\le 1}\mu(\mathrm{d}y)<\infty,\quad \text{when } d=1; \tag{14}$$

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{R}^2}\int_{|x-y|\le\varepsilon}\ln\frac{1}{|x-y|}\mu(\mathrm{d}y)=0,\quad \text{when } d=2; \tag{15}$$

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{R}^d}\int_{|x-y|\le\varepsilon}|x-y|^{2-d}\mu(\mathrm{d}y)=0,\quad \text{when } d\ge 3. \tag{16}$$

[13] Kuwae and Takahashi, 2007, “Kato class measures of symmetric Markov processes under heat kernel estimates”, Cf.

Consider now a measure ν of the form $\nu(\mathrm{d}s, \mathrm{d}x) = \mu(s, \mathrm{d}x)\,\mathrm{d}s$. Similarly to Equations (14) to (16) on the previous page one can show that if for each $T > 0$, μ satisfies the condition

$$\sup_{t\in[0,\infty)} \sup_{x\in\mathbb{R}} \int_{|x-y|\le 1} \mu(t, \mathrm{d}y) < \infty, \quad \text{when } d = 1, \tag{17}$$

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in[0,\infty)} \sup_{x\in\mathbb{R}^2} \int_{|x-y|\le\varepsilon} \ln \frac{1}{|x-y|} \mu(t, \mathrm{d}y) = 0, \quad \text{when } d = 2, \tag{18}$$

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in[0,\infty)} \sup_{x\in\mathbb{R}^d} \int_{|x-y|\le\varepsilon} |x-y|^{2-d} \mu(t, \mathrm{d}y) = 0, \quad \text{when } d \ge 3, \tag{19}$$

then $\nu \in \mathcal{K}$.

Remark 5 – Suppose that the measure $\nu(\mathrm{d}s, \mathrm{d}x) = \mu(s, \mathrm{d}x)\,\mathrm{d}s$ satisfies one of the conditions Equations (17) to (19). Then it can be verified[15] that for each $T > 0$, $r > 0$, there exists $K = K(r, T) > 0$ such that for all $x \in \mathbb{R}^d$, $t \in [0, T]$,

$$\mu(t, B(x, r)) < K,$$

where $B(x, r)$ is the ball centered at x and with radius r.

In the sequel we use the following modification of Khas'minskii's lemma[16].

Lemma 1 – *Let A_t be a W-functional with the characteristic f_t satisfying condition Equation (9) on p. 7. Then for all $p > 0$, $t \ge 0$, there exists a constant $C > 0$ depending on p, t, and the rate of convergence in Equation (9) on p. 7 such that*

$$\sup_{t_0\in[0,\infty), x_0\in\mathbb{R}^d} \mathbb{E}_{t_0,x_0} \exp\{pA_t\} < C.$$

Example 2 – Let $\nu(dt, \mathrm{d}x) = h(t, x)dt\,\mathrm{d}x$, where h is a non-negative bounded measurable function. Then the measure ν is of the class $\mathcal{K}$. The functional

$$A_t := \int_0^t h(\eta_s)\,\mathrm{d}s$$

is a W-functional of the process $(\eta_t)_{t\ge 0}$ with characteristic defined by Equation (8) on p. 7, and

$$A_t(\varphi(x)) = \int_0^t h(s, \varphi_s(x))\,\mathrm{d}s.$$

[14]E.g. in Chen, 2002, "Gaugeability and Conditional Gaugeability", Theorem 2.1.

[15]Cf. Dynkin, 1965, *Markov Processes*, Lemma 8.3.

[16]See Khasminskii, 1959, "On Positive Solutions of the Equation $\mathfrak{A}U + Vu = 0$"; or Sznitman, 1998, *Brownian Motion, Obstacles and Random Media.* Ch. 1, Lemma 2.1.

Example 3 – *Local time.* Let $d = 1$. It is well known that for any $x, y \in \mathbb{R}$ there exists a local time of the process $(\varphi_t(x))_{t\geq 0}$ at the point y, which is defined as

$$L_t^y(\varphi(x)) = \underset{\varepsilon\downarrow 0}{\text{l.i.m.}}\, \frac{1}{2\varepsilon}\int_0^t \mathbb{1}_{[y-\varepsilon,y+\varepsilon]}(\varphi_s(x))\,\mathrm{d}s.$$

It can be checked that $L_t^y(\varphi(x))$ is a W-functional of $(\varphi_t(x))_{t\geq 0}$ corresponding to the measure $\nu(\mathrm{d}s,\mathrm{d}x) = \mathrm{d}s\delta_y(\mathrm{d}x)$, where δ_y is the delta measure at the point y. Indeed, for fixed $y \in \mathbb{R}$ and each $\varepsilon > 0$, put

$$h^{\varepsilon,y}(t,x) = h^{\varepsilon,y}(x) = \frac{1}{2\varepsilon}\mathbb{1}_{[y-\varepsilon,y+\varepsilon]}(x), t \geq 0, x \in \mathbb{R},$$

and $\nu^{\varepsilon,y}(dt,\mathrm{d}x) = h^{\varepsilon,y}(t,x)dt\,\mathrm{d}x$. The function $h^{\varepsilon,y}$ is bounded and measurable. Then (see Example 2 on the preceding page) there exists a W-functional of the process $(\eta_t)_{t\geq 0}$ corresponding to the measure $\nu^{\varepsilon,y}$. This functional is defined by the formula

$$A_t^{\varepsilon,y} := A_t^{\nu^{\varepsilon,y}} = \int_0^t h^{\varepsilon,y}(\eta_s)\,\mathrm{d}s = \frac{1}{2\varepsilon}\int_0^t \mathbb{1}_{[y-\varepsilon,y+\varepsilon]}(\eta_s^2)\,\mathrm{d}s$$

and its characteristic is equal to

$$f_t^{\varepsilon,y}(t_0,x_0) = \mathbb{E}_{t_0,x_0}A_t^{\varepsilon,y}(\eta) = \int_{t_0}^{t_0+t}\mathrm{d}s\int_{\mathbb{R}^d} G(t_0,x_0,s,v)h^{\varepsilon,y}(s,v)\,\mathrm{d}v.$$

One can see that $f_t^{\varepsilon,y}(t_0,x_0)$ tends to

$$f_t^y(t_0,x_0) = \int_{t_0}^{t_0+t} G(t_0,x_0,s,y)\,\mathrm{d}s = \int_{t_0}^{t_0+t}\mathrm{d}s\int_{\mathbb{R}^d} G(t_0,x_0,s,v)\delta_y(\mathrm{d}v)$$

as $\varepsilon \to 0$ uniformly in $t \in [0,T]$, $t_0 \in [0,\infty]$, $x_0 \in \mathbb{R}$. Then by Theorem 1 on p. 5 there exists a functional

$$A_t^y = \underset{\varepsilon\downarrow 0}{\text{l.i.m.}}\, A_t^{\varepsilon,y} = \underset{\varepsilon\downarrow 0}{\text{l.i.m.}}\int_0^t h^{\varepsilon,y}(\eta_s)\,\mathrm{d}s.$$

In particular,

$$A_t^y(\varphi_t(x)) = L_t^y(\varphi(x)).$$

Note that if $d \geq 2$, the measure δ_y is not of the class $\mathcal{K}$. This agrees with the well-known fact that the local time for a multidimensional Wiener process does not exist.

The following lemma deals with the convergence of W-functionals of, generally speaking, different random functions.

Lemma 2 – *Let $\{(\xi_{n,t})_{t\ge 0} : n \ge 0\}$ be a sequence of homogeneous Markov random functions defined on a common probability space $(\Omega,\mathscr{F},P)$ with the common phase space $(E,\mathscr{B})$, where E is a metric space, $\mathscr{B}$ is the Borel σ-algebra. For $n \ge 0$, let $A_{n,t} = A_{n,t}(\xi_n)$ be a W-functional of the random function $(\xi_{n,t})_{t\ge 0}$ with the characteristic $f_{n,t}(z)$.*

Assume that

(A_1) *for each $t \ge 0$, $f_{0,t}(z)$ is continuous in $z \in E$;*

(A_2) *for each $t \ge 0$, $\xi_{n,t} \to \xi_{0,t}$, $n \to \infty$, in probability P;*

(A_3) *for all $n \ge 0$, $\lim_{\delta\downarrow 0}\|f_{n,\delta}\|_E = 0$, where*

$$\|f_{n,\delta}\|_E = \sup_{z\in E}|f_{n,\delta}(z)|;$$

(A_4) *for each $t > 0$, $\|f_{n,t} - f_{0,t}\|_E \to 0$, $n \to \infty$.*

Then for each $T > 0$,

$$\sup_{t\in[0,T]}\left|A_{n,t}(\xi_n) - A_{0,t}(\xi_0)\right| \to 0,\ n \to \infty,\ \textit{in probability } P.$$

Proof. Note that $A^\delta_{n,t} := \frac{1}{\delta}\int_0^t f_{n,\delta}(\xi_{n,s})\,\mathrm{d}s$ is a W-functional of the process $(\xi_{n,t})_{t\ge 0}$. Denote its characteristic by $f^\delta_{n,t}$. Then by Dynkin (1965, Lemma 6.5), for all $t \ge 0$, $z \in E$,

$$\mathbb{E}_z\left(A_{n,t} - \frac{1}{\delta}\int_0^t f_{n,\delta}(\xi_{n,s})\,\mathrm{d}s\right)^2 \le 2\left(f_{n,t}(z) + f^\delta_{n,t}(z)\right)\sup_{0\le u\le t}\left\|f_{n,u} - f^\delta_{n,u}\right\|_E.$$

Similarly to the proof of Dynkin (1965, Theorem 6.6), we get

$$\begin{aligned}\left|f^\delta_{n,t}(z) - f_{n,t}(z)\right| &\le \frac{1}{\delta}\int_t^{t+\delta}\left|f_{n,u}(z) - f_{n,t}(z)\right|\mathrm{d}u + \frac{1}{\delta}\int_0^\delta f_{n,u}(z)\,\mathrm{d}u\\ &\le \frac{1}{\delta}\int_t^{t+\delta}\left\|f_{n,u-t}\right\|_E\mathrm{d}u + \frac{1}{\delta}\int_0^\delta\left\|f_{n,u}\right\|_E\mathrm{d}u \le 2\left\|f_{n,\delta}\right\|_E.\end{aligned}$$

So for all $t \ge 0$,

$$\sup_{0\le u\le t}\left\|f^\delta_{n,u} - f_{n,u}\right\|_E \le 2\left\|f_{n,\delta}\right\|_E. \tag{20}$$

Using the calculations of the proof of Dynkin (1965, Theorem 6.6), once more we obtain

(Cont. next page)

$$f_{n,t}(z) + f^\delta_{n,t}(z) = f_{n,t}(z) + \frac{1}{\delta}\int_t^{t+\delta} f_{n,u}(z)\,\mathrm{d}u - \frac{1}{\delta}\int_0^\delta f_{n,u}(z)\,\mathrm{d}u$$

$$\begin{aligned}&\le \left\|f_{n,t}\right\|_E + \left\|f_{n,t+\delta}\right\|_E \\ &\le 2\left\|f_{n,t+\delta}\right\|_E. \end{aligned} \tag{21}$$

The inequalities Equation (20) on the preceding page and Equation (21) give us the estimate

$$\mathbb{E}_z\left(A_{n,t}(\xi_n) - \frac{1}{\delta}\int_0^t f_{n,\delta}(\xi_{n,s})\,\mathrm{d}s\right)^2 \le 8\left\|f_{n,\delta}\right\|_E\left\|f_{n,t+\delta}\right\|_E. \tag{22}$$

Further, we have

$$\begin{aligned}\mathbb{E}_z(A_{n,t}(\xi_n) - A_{0,t}(\xi_0))^2 \le 4\Bigg[&\mathbb{E}_z\left(A_{n,t}(\xi_n) - \frac{1}{\delta}\int_0^t f_{n,\delta}(\xi_{n,s})\,\mathrm{d}s\right)^2 \\ &+ \mathbb{E}_z\left(\frac{1}{\delta}\int_0^t f_{n,\delta}(\xi_{n,s})\,\mathrm{d}s - \frac{1}{\delta}\int_0^t f_{0,\delta}(\xi_{n,s})\,\mathrm{d}s\right)^2 \\ &+ \mathbb{E}_z\left(\frac{1}{\delta}\int_0^t f_{0,\delta}(\xi_{n,s})\,\mathrm{d}s - \frac{1}{\delta}\int_0^t f_{0,\delta}(\xi_{0,s})\,\mathrm{d}s\right)^2 \\ &+ \mathbb{E}_z\left(\frac{1}{\delta}\int_0^t f_{0,\delta}(\xi_{0,s})\,\mathrm{d}s - A_{0,t}(\xi_0)\right)^2\Bigg] = 4[I + II + III + IV]. \end{aligned} \tag{23}$$

By assertion (A_3) on the preceding page for any $\varepsilon > 0$ we can choose $\delta > 0$ such that $\left\|f_{0,\delta}\right\|_E < \varepsilon$. According to assertion (A_4) there exists $n_0 > 0$ such that for all $n > n_0$,

$$\left\|f_{n,\delta} - f_{0,\delta}\right\|_E < \varepsilon.$$

Then for all $n > n_0$,

$$\left\|f_{n,\delta}\right\|_E \le \left\|f_{n,\delta} - f_{0,\delta}\right\|_E + \left\|f_{0,\delta}\right\|_E < 2\varepsilon.$$

Note that for each $n \ge 0$, $k \ge 1$, we have $\left\|f_{n,k\delta}\right\|_E \le k\left\|f_{n,\delta}\right\|_E$. This implies that for any $t \ge 0$, $M_t := \sup_{n\ge 0}\left\|f_{n,t}\right\|_E < \infty$. Taking into account Equation (22), we obtain that for all $n > n_0$,

$$I \le 16 M_{t+\delta}\varepsilon,$$

and the same estimate holds for IV.

By the Hölder inequality,

$$II \le \frac{t}{\delta^2}\mathbb{E}_z\int_0^t (f_{n,\delta}(\xi_{n,s}) - f_{0,\delta}(\xi_{n,s}))^2\,\mathrm{d}s \le \frac{t^2}{\delta^2}\mathbb{E}_z \sup_{z\in E}(f_{n,\delta}(z) - f_{0,\delta}(z))^2$$

The assertion (A_4) on the preceding page yields the estimate $II \le \varepsilon$ valid for all $n \ge n_1 = n_1(\varepsilon, \delta)$.

Similarly,

$$III \le \frac{t}{\delta^2}\mathbb{E}_z \int_0^t (f_{0,\delta}(\xi_{n,s}) - f_{0,\delta}(\xi_{0,s}))^2 \, \mathrm{d}s.$$

The continuity of the function $f_{0,t}(\cdot)$ and assertion (A_2) provide the convergence $f_{0,\delta}(\xi_{n,s})$ to $f_{0,\delta}(\xi_{0,s})$ in probability as n tends to ∞. This convergence together with assertion (A_3) allow us to use the dominated convergence theorem and prove that $III \to 0$ as $n \to \infty$. Then the right-hand side of Equation (23) on the previous page tends to 0 as n tends to ∞. The uniform convergence follows from Proposition 1 on p. 5, which completes the proof. □

2 The main result

In the following theorem we formulate the main result of this paper.

Theorem 4 – *Suppose that $a = (a^1, \dots, a^d) : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded measurable function such that for each $t \ge 0$ and all $1 \le i \le d$, $a^i(t,\cdot)$ is a function of bounded variation on $\mathbb{R}^d$, i.e., for each $1 \le j \le d$, the generalized derivative $\mu^{ij}(t,\mathrm{d}y) = \frac{\partial a^i}{\partial y_j}(t,\mathrm{d}y)$ is a signed measure on $\mathbb{R}^d$. Assume that the signed measures $\nu^{ij}(dt,\mathrm{d}y) := \mu^{ij}(t,\mathrm{d}y)dt$, $1 \le i,j \le d$, are of the class $\mathcal{K}$.*

Let $\sigma : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$ be a bounded continuous function satisfying conditions (C_1) to (C_3) on p. 2 and the following condition

(C_4) *There exists $\rho > 0$ such that for all $1 \le k \le m$, $1 \le i,j \le d$, the function $\left|\frac{\partial \sigma_k^i}{\partial y_j}(s,y)\right|^{2+\rho}$ is of the class $\mathcal{K}$.*

Then there exists the derivative $Y_t(x) = \nabla\varphi_t(x)$ in L_p-sense: for all $p > 0$, $x \in \mathbb{R}^d$, $v \in \mathbb{R}^d$, $t \ge 0$,

$$\mathbb{E}\left|\frac{\varphi_t(x+\varepsilon v) - \varphi_t(x)}{\varepsilon} - Y_t(x)v\right|^p \to 0, \ \varepsilon \to 0. \tag{24}$$

This derivative is a unique solution of the integral equation

$$Y_t(x) = E + \int_0^t dA_s^\nu(\varphi(x))Y_s(x) + \sum_{k=1}^m \int_0^t \nabla\sigma_k(s,\varphi_s(x))Y_s(x)\,\mathrm{d}w_k(s), \tag{25}$$

where E is the $d \times d$-identity matrix, $\nabla\sigma_k(s,y) = \left(\frac{\partial \sigma_k^i}{\partial y_j}(s,y)\right)_{1\le i,j\le d}$; the first integral in the right-hand side of Equation (25) is the Lebesgue-Stieltjes integral with respect to the continuous function of bounded variation $t \to A_t^\nu(\varphi(x))$.

Moreover,

$$P\Big\{\forall t \geq 0 : \varphi_t(\cdot) \in W^1_{p,loc}(\mathbb{R}^d, \mathbb{R}^d), \nabla\varphi_t(x) = Y_t(x) \text{ for } \lambda\text{-a.a. } x\Big\} = 1, \tag{26}$$

where λ is the Lebesgue measure on $\mathbb{R}^d$.

Remark 6 – The W-functional $A^\nu_t = \left(A^{\nu^{ij}}_t\right)_{1\leq i,j\leq d}$ is well defined because the signed measure ν is of the class $\mathscr{K}$.

Remark 7 – Recall that for all $1 \leq i, j \leq d$, the mappings $A^{\nu^{ij,\pm}}_t$, which we will denote by $A^{ij,\pm}_t$, are continuous and monotonic in t. So for each $T > 0$, the function $t \to A^{ij}_t$ is a continuous function of bounded variation on $[0, T]$ almost surely.

3 The proof of Theorem 4 on the preceding page

The existence and uniqueness of the solution to Equation (25) on the preceding page follows from Protter (2004, Ch. V, Theorem 7). Indeed, condition (C_4) on the preceding page provides that for all $1 \leq k \leq m$, $\int_0^t |\nabla\sigma_k(s, \varphi_s(x))|^2 \, \mathrm{d}s < \infty$ almost surely and consequently

$$\int_0^t dA^\nu_s(\varphi(x)) + \sum_{k=1}^m \int_0^t \nabla\sigma_k(s, \varphi_s(x)) \, \mathrm{d}w_k(s), t \geq 0,$$

is a semimartingale.

It is well known that the statement of the theorem is true in the case of smooth coefficients, and the derivative satisfies Equation (2) on p. 2. To prove the theorem in the general case, we approximate the initial equation by equations with smooth coefficients.

The proof is divided into two steps.

3.1 First step

In the first step, we assume that there exists $R > 0$ such that for all $t \geq 0$, $x \in \mathbb{R}^d$, $|x| \geq R$, $a(t, x) = 0$, $\sigma(t, x) = \widetilde{\sigma} = \text{const}$, $\widetilde{\sigma}\widetilde{\sigma}^* > 0$.

For $n \geq 1$, let $\omega_n \in C^\infty_0(\mathbb{R}^d)$ be a non-negative function such that $\int_{\mathbb{R}^d} \omega_n(z) \, \mathrm{d}z = 1$, and $\omega_n(x) = 0$, $|x| \geq 1/n$. For all $t \geq 0$, $x \in \mathbb{R}^d$, $n \geq 1$, and $1 \leq k \leq m$, put

$$a_n(t, x) = (\omega_n * a)(t, x) = \int_{\mathbb{R}^d} \omega_n(x - y) a(t, y) \, \mathrm{d}y, \tag{27}$$

$$\sigma_{n,k}(t, x) = (\omega_n * \sigma_k)(t, x) = \int_{\mathbb{R}^d} \omega_n(x - y) \sigma_k(t, y) \, \mathrm{d}y. \tag{28}$$

Note that for each $T > 0$,

$$\sup_{n\geq 1}\|a_n\|_{T,\infty} \leq \|a\|_{T,\infty}, \tag{29}$$

$$\sup_{n\geq 1}\left\|\sigma_{n,k}\right\|_{T,\infty} \leq \|\sigma_k\|_{T,\infty},\ 1 \leq k \leq m, \tag{30}$$

where

$$\|a\|_{T,\infty} = \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}|a(t,x)|.$$

Besides, for all $n \geq 1$, σ_n satisfies condition (C_2) on p. 2, and the ellipticity constant can be chosen uniformly in n.

Remark 8 – For all $n \geq 1$ the transition probability density of the process $(\varphi_{n,t}(x))_{t\geq 0}$ satisfies the inequality Equation (10) on p. 8. It follows from Equation (29) and condition (C_2) on p. 2, which holds uniformly in n, that the constants in Equation (10) on p. 8 can be chosen uniformly in $n \geq 1$.

For each $T > 0$, we have $a_n \to a$, $n \to \infty$, in $L_1([0,T]\times\mathbb{R}^d)$. Choosing subsequences, without loss of generality we can assume that $a_n(t,x) \to a(t,x)$, $n \to \infty$, for almost all $t \geq 0$ and almost all x w.r.t. the Lebesgue measure. Then for all $n \geq 1,\ t \geq 0,\ x \in \mathbb{R}^d$ such that $|x| \geq R+1$,

$$a_n(t,x) = 0,\ \sigma_n(t,x) = \widetilde{\sigma}.$$

Without loss of generality we can assume that this is true for all x such that $|x| > R$. Moreover, from condition (C_3) on p. 2 we derive that for each $T > 0$, $\sigma_n \to \sigma$, $n \to \infty$, uniformly in $(t,x) \in [0,T]\times\mathbb{R}^d$.

Consider the SDE:

$$\begin{cases} d\varphi_{n,t}(x) = a_n(t,\varphi_{n,t}(x))\,\mathrm{d}t + \sum_{k=1}^{m}\sigma_{n,k}(t,\varphi_{n,t}(x))\,\mathrm{d}w_k(t), \\ \varphi_{n,0}(x) = x,\ x \in \mathbb{R}^d. \end{cases} \tag{31}$$

For each $n \geq 1$ there exists a unique strong solution of Equation (31).

Lemma 3 – *For each $p \geq 1$,*

1. *for all $t \geq 0$ and any compact set $U \in \mathbb{R}^d$,*

$$\sup_{x\in U,\ n\geq 1}\left(\mathbb{E}(\left|\varphi_{n,t}(x)\right|^p + |\varphi_t(x)|^p)\right) < \infty;$$

2. *for all $x \in \mathbb{R}^d,\ T \geq 0$,*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left|\varphi_{n,t}(x) - \varphi_t(x)\right|^p\right) \to 0 \text{ as } n \to \infty.$$

Proof. The first statements follows from the uniform boundedness of the coefficients, the second is the consequence of Luo (2011, Theorem 3.4). □

For $n \ge 1$, put

$$\nabla a_n = \left(\frac{\partial a_n^i}{\partial x_j}\right)_{1\le i,j\le d}, \ \nabla \sigma_{n,k} = \left(\frac{\partial \sigma_{n,k}^i}{\partial x_j}\right)_{1\le i,j\le d}.$$

Denote by $Y_{n,t}(x)$ the matrix of derivatives of $\varphi_{n,t}(x)$ in x, i.e., $Y_{n,t}^{ij}(x) = \frac{\partial \varphi_{n,t}^i(x)}{\partial x_j}$, $1 \le i,j \le d$. Then $Y_{n,t}(x)$ satisfies the equation

$$Y_{n,t}(x) = E + \int_0^t \nabla a_n(s,\varphi_{n,s}(x))Y_{n,s}(x)\,\mathrm{d}s + \sum_{k=1}^m \int_0^t \nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Y_{n,s}(x)\,\mathrm{d}w_k(s), \quad (32)$$

where E is the d-dimensional identity matrix.

By the properties of convolution of a generalized function[17],

$$\nabla a_n = \nabla a * \omega_n = a * \nabla\omega_n, \ n \ge 1. \quad (33)$$

Note that for all $n \ge 1$, $1 \le i,j \le d$, ∇a_n^{ij} is a bounded measurable function on $[0,\infty)\times\mathbb{R}^d$. Then (see Example 1 on p. 6) there exists a continuous homogeneous additive functional

$$A_{n,t}^{ij}(\varphi_n(x)) = \int_0^t \frac{\partial a_n^i}{\partial y_j}(s,\varphi_{n,s}(x))\,\mathrm{d}s$$

corresponding to the signed measure $\nabla a_n^{ij}(s,y)\,\mathrm{d}s\,\mathrm{d}y$.

Denote $\mu_n^{ij}(t,y)\,\mathrm{d}y = \frac{\partial a_n^i}{\partial y_j}(t,y)\,\mathrm{d}y$. For each $n \ge 1$, $1 \le i,j \le d$, put $\mu_n^{ij,\pm} = \mu^{ij,\pm} * \omega_n$ (recall that $\mu^{ij}(t,\mathrm{d}y) = \frac{\partial a^i}{\partial y_j}(t,\mathrm{d}y)$). Then $\mu_n^{ij} = \mu_n^{ij,+} - \mu_n^{ij,-}$. It can be easily seen that the measures $\nu_n^{ij,\pm}(dt,\mathrm{d}y) = \mu_n^{ij,\pm}(t,\mathrm{d}y)dt$, $n \ge 1$, are of the class $\mathcal{K}$. By Remark 8 on the preceding page, for each $x \in \mathbb{R}^d$ there exist W-functionals $A_t^{\nu_n^{ij,\pm}}(\varphi_{n,\cdot}(x))$, which we denote by $A_t^{ij,\pm}(\varphi_n(x))$. Generally speaking, $\mu_n^{ij,\pm} \ne (\mu^{ij} * \omega_n)^\pm$ but, by Remark 3 on p. 9,

$$A_{n,t}^{ij}(\varphi_n(x)) = A_t^{\nu_n^{ij}}(\varphi_{n,\cdot}(x)) = A_t^{ij,+}(\varphi_n(x)) - A_t^{ij,-}(\varphi_n(x)).$$

Denote $\varphi_{0,t}(x) = \varphi_t(x)$, $Y_{0,t}(x) = Y_t(x)$, $a_0 = a$, $\sigma_0 = \sigma$, $A_{0,t} = A_t$.

Lemma 4 – *For all $t \ge 0$, $p > 0$, $1 \le i,j \le d$, there exists a constant C such that*

$$\sup_{n\ge 0}\sup_{x\in\mathbb{R}^d} \mathbb{E}\exp\left\{pA_{n,t}^{ij,\pm}(\varphi_n(x))\right\} < C. \quad (34)$$

[17]See Vladimirov, 1967, *The Equation of Mathematical Physics*, Ch. 2, §7.

Proof. The statement of lemma follows from Lemma 1 on p. 10 and Remark 8 on p. 16. □

Lemma 5 – *For all $T \geq 0$, $x \in \mathbb{R}^d$, $p > 0$,*

$$\sup_{n\geq 0} \mathbb{E} \sup_{0\leq t\leq T} \left|Y_{n,t}(x)\right|^p < \infty.$$

Proof. For all $t > 0$, $n \geq 0$, define the variation of $A^{ij}_{n,\cdot}$ on $[0,t]$ by

$$\operatorname{Var} A^{ij}_{n,t}(\varphi(x)) := A^{ij,+}_{n,t}(\varphi(x)) + A^{ij,-}_{n,t}(\varphi(x)),$$

and denote

$$\operatorname{Var} A_{n,t}(\varphi(x)) := \Sigma_{1\leq i,j\leq d} \operatorname{Var} A^{ij}_{n,t}(\varphi(x)).$$

Set

$$\tau^N_n = \inf\left\{ t \geq 0 : \int_0^t \sum_{k=1}^m \left|\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))\right|^2 \mathrm{d}s + \operatorname{Var} A_{n,s}(\varphi(x)) + \left|Y_{n,s}(x)\right|^2 \geq N \right\}.$$

For the sake of brevity, denote

$$h_n(t,C,l) = -2l \operatorname{Var} A_{n,t}(\varphi(x)) - C \sum_{k=1}^m \int_0^t |\nabla\sigma_k(s,\varphi_s(x))|^2 \,\mathrm{d}s.$$

By Ito's formula, for all $n \geq 0$, $l \in \mathbb{N}$,

$$\mathrm{e}^{h_n(t\wedge\tau^N_n,C,l)} \left|Y_{n,t\wedge\tau^N_n}(x)\right|^{2l} = \left|Y_{n,0}(x)\right|^{2l} + M(t\wedge\tau^N_n) + I + II + III + IV,$$

where

$$M(t\wedge\tau^N_n) = 2l \int_0^{t\wedge\tau^N_n} \mathrm{e}^{h_n(s,C,l)} \left|Y_{n,s}(x)\right|^{2l-2}$$
$$\times \sum_{i,j=1}^d Y^{ij}_{n,s}(x) \sum_{k=1}^m \sum_{r=1}^d \nabla\sigma^{ir}_{n,k}(s,\varphi_{n,s}(x)) Y^{rj}_{n,s}(x) \,\mathrm{d}w_k(s),$$

$$I = -2l \int_0^{t\wedge\tau^N_n} \mathrm{e}^{h_n(s,C,l)} \left|Y_{n,s}(x)\right|^{2l} d \operatorname{Var} A_{n,s}(\varphi(x)),$$

$$II = -C \int_0^{t\wedge\tau^N_n} \mathrm{e}^{h_n(s,C,l)} \left|Y_{n,s}(x)\right|^{2l} \sum_{k=1}^m \left|\nabla\sigma_{n,k}(s,\varphi_s(x))\right|^2 \mathrm{d}s,$$

$$III = 2l \int_0^{t\wedge\tau^N_n} \mathrm{e}^{h_n(s,C,l)} \left|Y_{n,s}(x)\right|^{2l-2} \sum_{i,j=1}^d Y^{ij}_{n,s}(x) \sum_{r=1}^d dA^{ir}_{n,s}(\varphi(x)) Y^{rj}_{n,s}(x),$$

$$IV = 2l\int_0^{t\wedge\tau_n^N} \mathrm{e}^{h_n(s,C,l)}\left|Y_{n,s}(x)\right|^{2l-4}$$
$$\times\Big(\sum_{i,j=1}^{d}\sum_{v,q=1}^{d}\Big(2(l-1)Y_{n,s}^{ij}(x)Y_{n,s}^{vq}(x)+\left|Y_{n,s}(x)\right|^2\delta_{vi}\delta_{qj}\Big)$$
$$\times\sum_{k=1}^{m}\sum_{r=1}^{d}\nabla\sigma_{n,k}^{vr}(s,\varphi_{n,s}(x))Y_{n,s}^{rq}(x)\sum_{e=1}^{d}\nabla\sigma_{n,k}^{ie}(s,\varphi_{n,s}(x))Y_{n,s}^{ej}(x)\Big)\mathrm{d}s,$$

where $|\cdot|$ is the Hilbert-Schmidt norm.

Note that $I+III\le 0$. Besides, there exists a constant $\widetilde{C}=\widetilde{C}(d)>0$ such that

$$IV \le 2l\widetilde{C}\int_0^{t\wedge\tau_n^N} \mathrm{e}^{h_n(s,C,l)}\left|Y_{n,s}(x)\right|^{2l}\sum_{k=1}^{m}\left|\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))\right|^2\mathrm{d}s.$$

Then we can choose $C>0$ so large that $IV+II\le 0$. We obtain

$$\mathrm{e}^{h_n(t\wedge\tau_n^N,C,l)}\left|Y_{n,t\wedge\tau_n^N}(x)\right|^{2l}\le\left|Y_{n,0}(x)\right|^{2l}+M(t\wedge\tau_n^N), \tag{35}$$

where $M(t\wedge\tau_n^N), t\ge 0$, is a square integrable martingale. Then, for all $t\ge 0$,

$$\mathbb{E}\mathrm{e}^{h_n(t\wedge\tau_n^N,C,l)}\left|Y_{n,t\wedge\tau_n^N}(x)\right|^{2l}\le K,$$

where $K=\left|Y_{n,0}(x)\right|^{2l}=|E|^{2l}=d^l$. Passing to the limit as $N\to\infty$, we get that for all $T>0$ there exists $C=C(l,d)$ such that

$$\sup_{n\ge 0}\sup_{t\in[0,T]}\mathbb{E}\mathrm{e}^{h_n(t,C,l)}\left|Y_{n,t}(x)\right|^{2l}\le K. \tag{36}$$

By Equation (35), for all $T>0$,

$$\mathbb{E}\sup_{t\in[0,T]}\mathrm{e}^{2h_n(t,C,l)}\left|Y_{n,t}(x)\right|^{4l}\le 2\mathbb{E}\sup_{t\in[0,T]}\Big(\left|Y_{n,0}(x)\right|^{4l}+M^2(t)\Big)$$
$$\le K'\left(1+\sum_{k=1}^{m}\mathbb{E}\int_0^T \mathrm{e}^{2h_n(s,C,l)}\left|Y_{n,s}(x)\right|^{4l}\left|\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))\right|^2\mathrm{d}s\right).$$

Using the Hölder inequality with $p=1+\frac{\rho}{2}$, we get

(Cont. next page)

$$\mathbb{E}\sup_{t\in[0,T]}\mathrm{e}^{2h_n(t,C,l)}\left|Y_{n,t}(x)\right|^{4l}\le K'\left[1+\left(\mathbb{E}\int_0^T\Big(\mathrm{e}^{2h_n(s,C,l)}\left|Y_{n,s}(x)\right|^{4l}\Big)^{\frac{2+\rho}{\rho}}\mathrm{d}s\right)^{\frac{\rho}{2+\rho}}\right.$$

$$\times \sum_{k=1}^{m} \left(\mathbb{E} \int_0^T \left| \nabla \sigma_{n,k}(s, \varphi_{n,s}(x)) \right|^{2+\rho} \mathrm{d}s \right)^{\frac{2}{2+\rho}} \Bigg]. \tag{37}$$

Since for all $1 \le k \le m$, $1 \le i, j \le d$, the function $\left| \frac{\partial \sigma_k^i}{\partial y_j}(s,y) \right|^{2+\rho}$ is of the class $\mathcal{K}$, the functions $\left| \frac{\partial \sigma_{n,k}^i}{\partial y_j}(s,y) \right|^{2+\rho}$, $n \ge 1$, are of the class $\mathcal{K}$ too. It follows from Lemma 4 on p. 17 that for each $T > 0$,

$$\sup_{n\ge 1} \mathbb{E} \exp \left\{ \int_0^T \left| \nabla \sigma_{n,k}(s, \varphi_{n,s}(x)) \right|^{2+\rho} \right\} \mathrm{d}s < C(T), \tag{38}$$

where $C(T)$ is a constant which depends on T. Consequently,

$$\sup_{n\ge 0} \mathbb{E} \int_0^T \left| \nabla \sigma_{n,k}(s, \varphi_{n,s}(x)) \right|^{2+\rho} \mathrm{d}s < \infty. \tag{39}$$

By Equation (36) on the previous page we have

$$\sup_{n\ge 0} \mathbb{E} \int_0^T \left(\mathrm{e}^{2h_n(s,C,l)} \left| Y_{n,s}(x) \right|^{4l} \right)^{\frac{2+\rho}{\rho}} \mathrm{d}s < \infty. \tag{40}$$

From Equations (39) and (40) we get

$$\sup_{n\ge 0} \mathbb{E} \sup_{t\in[0,T]} \mathrm{e}^{2h_n(t,C,l)} \left| Y_{n,t}(x) \right|^{4l} < \infty. \tag{41}$$

Finally, for any $T > 0$, by the Hölder inequality,

$$\begin{aligned} \sup_{n\ge 0} \mathbb{E} \sup_{t\in[0,T]} \left| Y_{n,s}(x) \right|^{2l} &= \sup_{n\ge 0} \mathbb{E} \sup_{t\in[0,T]} \left[\left(\mathrm{e}^{h_n(t,C,l)} \left| Y_{n,s}(x) \right|^{2l} \right) \mathrm{e}^{-h_n(t,C,l)} \right] \\ &\le \sup_{n\ge 0} \left[\left(\mathbb{E} \sup_{t\in[0,T]} \mathrm{e}^{2h_n(t,C,l)} \left| Y_{n,t}(x) \right|^{4l} \right)^{1/2} \right. \\ &\quad \times \left(\mathbb{E} \exp \left\{ 4l \operatorname{Var} A_{n,T}(\varphi_n(x)) \right. \right. \\ &\quad \left. \left. \left. + \ 2C \sum_{k=1}^{m} \int_0^T \left| \nabla \sigma_{n,k}(s, \varphi_{n,s}(x)) \right|^2 \mathrm{d}s \right\} \right)^{1/2} \right]. \end{aligned}$$

Now the assertion of the lemma follows from Equation (39), Equation (41), and the fact that for each $T > 0$, $\sup_{n\ge 0} \operatorname{Var} A_{n,T}(\varphi_n(x)) < \infty$, which is a consequence of Lemma 4 on p. 17. □

Lemma 6 – *For each $T > 0$, $x \in \mathbb{R}^d$, $1 \le i, j \le d$,*

$$\sup_{0\le t\le T}\left|A^{ij,\pm}_{n,t}(\varphi_n(x)) - A^{ij,\pm}_t(\varphi(x))\right| \to 0,\ n \to \infty,\ \textit{in probability}\ \mathbb{P}.$$

Proof. To prove the lemma we use Lemma 2 on p. 12 in which we put $\xi_{n,t} = \eta_{n,t}$, $A_{n,t} = A_{n,t}(\eta_n)$, $\xi_{0,t} = \eta_t$, and $A_{0,t} = A_t(\eta)$, $n \ge 1$, $t \ge 0$. Here $(\eta_{n,t})_{t\ge0}$ is a solution to the system of the form Equation (7) on p. 7 with coefficients $a_n, \sigma_{n,k}$. Then

$$f_{0,t}(t_0, x_0) = \int_{t_0}^{t+t_0} \mathrm{d}s \int_{\mathbb{R}^d} G(t_0, x, s, y)\mu(\mathrm{d}y),$$

where $G(s, x, t, y)$, $0 \le s \le t$, $x, y \in \mathbb{R}^d$, is the transition probability density of the process $(\eta^2_t)_{t\ge0}$. For each $T > 0$, the function $G(s, x, t, y)$ is continuous on $0 \le s < t \le T$, $x, y \in \mathbb{R}^d$[18]. Taking into account the inequality Equation (10) on p. 8, which holds locally uniformly in x, we obtain assertion (A_1) of Lemma 2 on p. 12 from the dominated convergence theorem. Assertion (A_2) is a consequence of Lemma 3 on p. 16. Assertion (A_3) is obvious. Assertion (A_4) follows from Lemma 9 on p. 34, which is proved in Section 4 on p. 29. □

Lemma 7 – *For all $T \ge 0$, $x \in \mathbb{R}^d$,*

$$\sup_{0\le t\le T}\left|Y_{n,t}(x) - Y_t(x)\right| \to 0,\ n \to \infty,\ \textit{in probability}\ \mathbb{P}.$$

To prove the lemma we need three auxiliary propositions. The first one is a version of the Gronwall inequality and can be obtained by a standard argument.

Proposition 2 – *Let $x(t)$, $C(t)$ be non-negative continuous functions on $[0, +\infty)$, $K(t)$ be a non-negative, non-decreasing function, and $K(0) = 0$. If for all $0 \le t \le T$,*

$$x(t) \le C(t) + \int_0^t x(s)\,\mathrm{d}K(s),$$

then

$$x(T) \le \left(\sup_{0\le t\le T} C(t)\right)\exp\{K(T)\}.$$

The following simple proposition is technical.

Proposition 3 – *Let $\{h_n : n \ge 1\}$ be a sequence of continuous monotonic functions on $[0, T]$, and $f \in C([0, T])$. Suppose that $t \in [0, T]$, $h_n(t) \to h_0(t)$, as $n \to \infty$, $t \in [0, T]$. Then*

$$\sup_{t\in[0,T]}\left|\int_0^t f(s)\,\mathrm{d}h_n(s) - \int_0^t f(s)\,\mathrm{d}h_0(s)\right| \to 0,\ n \to \infty.$$

[18] See N. I. Portenko, 1990, *Generalized Diffusion Processes*, Ch. 2, §2.

Proof. By Helly's theorem $\int_0^t f(s)\,\mathrm{d}h_n(s) \to \int_0^t f(s)\,\mathrm{d}h_0(s), n \to \infty$, $t \in [0,T]$, pointwise. If f is non-negative, then the uniform convergence follows from Dini's theorem. In the general case we can consider a decomposition $f = f_+ - f_-$, where $f_\pm$ are continuous non-negative functions, and apply the above argument to $f_\pm$. □

Proposition 4 – *Let X, Y be complete separable metric spaces, $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let measurable mappings $\xi_n : \Omega \to X$, $h_n : X \to Y$, $n \geq 0$, be such that*

1. *$\xi_n \to \xi_0$, $n \to \infty$, in probability $\mathbb{P}$;*
2. *$h_n \to h_0$, $n \to \infty$, in measure ν, where ν is a probability measure on X;*
3. *for all $n \geq 1$ the distribution P_{ξ_n} of ξ_n is absolutely continuous w.r.t. the measure ν;*
4. *the family of densities $\{\frac{dP_{\xi_n}}{d\nu} :\ n \geq 1\}$ is uniformly integrable w.r.t. the measure ν.*

Then $h_n(\xi_n) \to h_0(\xi_0)$, $n \to \infty$, in probability.

The proof can be found, for example, in Bogachev (2007, Corollary 9.9.11), or Kulik and Pilipenko (2000, Lemma 2).

Proof (of Lemma 7 on the previous page). Let $Z_n(t)$, $n \geq 0$, be a solution to the equation

$$\begin{cases} \mathrm{d}Z_n(t) = -Z_n(t)\,\mathrm{d}A_{n,t}(\varphi_n(x)),\ t \in [0,T], \\ Z_n(0) = E. \end{cases}$$

where E is the d-dimensional identity matrix, $T > 0$. For each $t \in [0,T]$, $n \geq 0$ the matrix $Z_n(t)$ is invertible, and

$$\begin{cases} \mathrm{d}Z_n^{-1}(t) = \mathrm{d}A_{n,t}(\varphi_n(x))Z_n^{-1}(t),\ t \in [0,T], \\ Z_n^{-1}(0) = E, \end{cases}$$

We get

$$|Z_n(t)| \leq |E| + \int_0^t |Z_n(s)| d\,\mathrm{Var}\,A_{n,s}(\varphi_n(x)).$$

It follows from Proposition 2 on the previous page that

$$\sup_{t\in[0,T]} |Z_n(t)| \leq d^{1/2} \exp\{\mathrm{Var}\,A_{n,T}(\varphi_n(x))\}. \tag{42}$$

Here we use that $|E| = d^{1/2}$. Similarly,

$$\sup_{t\in[0,T]} \left|Z_n^{-1}(t)\right| \leq d^{1/2} \exp\{\mathrm{Var}\,A_{n,T}(\varphi_n(x))\}. \tag{43}$$

Let us prove that

$$\sup_{t\in[0,T]} |Z_n(t)-Z_0(t)| + \sup_{t\in[0,T]} \left|Z_n^{-1}(t)-Z_0^{-1}(t)\right| \to 0,\ n\to\infty,\ \text{in probability } \mathbb{P}. \quad (44)$$

We have

$$\begin{aligned}|Z_n(t)-Z_0(t)| &\le \left|\int_0^t (Z_0(s)-Z_n(s))\,\mathrm{d}A_{n,s}(\varphi_n(x))\right| \\ &\qquad + \left|\int_0^t Z_0(s)\,(\mathrm{d}A_{0,s}(\varphi_0(x)) - dA_{n,s}(\varphi_n(x)))\right| \\ &\le \int_0^t |Z_0(s)-Z_n(s)|\,\mathrm{d}\,\mathrm{Var}\,A_{n,s}(\varphi_n(x)) \\ &\qquad + \left|\int_0^t Z_0(s)\,(\mathrm{d}A_{0,s}(\varphi_0(x)) - \mathrm{d}A_{n,s}(\varphi_n(x)))\right|.\end{aligned}$$

By Proposition 2 on p. 21,

$$\begin{aligned}|Z_n(t)-Z_0(t)| &\le \sup_{0\le u\le t}\left|\int_0^u Z_0(s)\,(\mathrm{d}A_{0,s}(\varphi_0(x)) - \mathrm{d}A_{n,s}(\varphi_n(x)))\right| \\ &\qquad \times \exp\{\mathrm{Var}\,A_{n,t}(\varphi_n(x))\} \\ &\le \sup_{0\le u\le t}\left(\left|\int_0^u Z_0(s)\left(\mathrm{d}A^+_{0,s}(\varphi_0(x)) - \mathrm{d}A^+_{n,s}(\varphi_n(x))\right)\right|\right. \\ &\qquad \left. + \sup_{0\le u\le t}\left|\int_0^u Z_0(s)\left(\mathrm{d}A^-_{0,s}(\varphi_0(x)) - \mathrm{d}A^-_{n,s}(\varphi_n(x))\right)\right|\right) \\ &\qquad \times \exp\{\mathrm{Var}\,A_{n,t}(\varphi_n(x))\}. \end{aligned} \quad (45)$$

Let us apply Proposition 3 on p. 21. Put $h_n(s) = A^+_{n,s}(\varphi_n(x))$, $n\ge 0$, $f(s) = Z_0(s)$. Taking into account Lemma 4 on p. 17 we get that the first summand in the right-hand side of Equation (45) tends to 0 as $n\to\infty$ in probability $\mathbb{P}$ uniformly in $t\in[0,T]$. The second summand can be treated analogously. Thus we have proved:

$$\sup_{t\in[0,T]} |Z_n(t)-Z_0(t)| \to 0,\ n\to\infty,\ \text{in probability } \mathbb{P}.$$

The same convergence can be obtained for Z_n^{-1}.

Using Ito's formula we get

$$\begin{aligned}&Z_n(t)Y_{n,t}(x) - Z_0(t)Y_{0,t}(x) \\ &\quad = \sum_{k=1}^m \int_0^t \Big(Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Y_{n,s}(x) - Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Y_{0,s}(x)\Big)\mathrm{d}w_k(s).\end{aligned}$$

Applying Ito's formula again, we get for any $K > 0$,

$$\left|Z_n(t)Y_{n,t}(x) - Z_0(t)Y_{0,t}(x)\right|^2 \exp\left\{-K\int_0^t \sum_{k=1}^m \left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|^2 \mathrm{d}s\right\}$$
$$= \int_0^t \exp\left\{-K\sum_{k=1}^m \int_0^s \left|\nabla\sigma_{0,k}(u,\varphi_{0,u}(x))\right|^2 \mathrm{d}u\right\}\times$$
$$\sum_{k=1}^m \Big(\left|Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Y_{n,s}(x) - Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Y_{0,s}(x)\right|^2 -$$
$$K\left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|^2\left|Z_n(s)Y_{n,s}(x) - Z_0(s)Y_{0,s}(x)\right|^2\Big)\mathrm{d}s +$$
$$2\int_0^t \exp\left\{-K\int_0^t \sum_{k=1}^m \left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|^2 \mathrm{d}s\right\}\Big(Z_n(s)Y_{n,s}(x) - Z_0(s)Y_{0,s}(x)\Big)\times$$
$$\sum_{k=1}^m \Big(Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Y_{n,s}(x) - Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Y_{0,s}(x)\Big)\mathrm{d}w_k(s).$$

Taking into account the inequalities Equations (38) and (42) and Lemma 5 on p. 18, on p. 20 and on p. 22, one can see that the last summand in the right-hand side of the previous equation is a square integrable martingale. The same estimates allow us to write

$$\mathbb{E}\left|Z_n(t)Y_{n,t}(x) - Z_0(t)Y_{0,t}(x)\right|^2 \exp\left\{-K\int_0^t \sum_{k=1}^m \left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|^2 \mathrm{d}s\right\} \le I + II, \quad (46)$$

where

$$I = \mathbb{E}\int_0^t \exp\left\{-K\sum_{k=1}^m \int_0^s \left|\nabla\sigma_{0,k}(u,\varphi_{0,u}(x))\right|^2 \mathrm{d}u\right\}$$
$$\times \sum_{k=1}^m \left|Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Z_n^{-1}(s) - Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Z_0^{-1}(s)\right|^2\left|Z_n(s)Y_{n,s}(x)\right|^2 \mathrm{d}s,$$
$$II = \mathbb{E}\int_0^t \exp\left\{-K\sum_{k=1}^m \int_0^s \left|\nabla\sigma_{0,k}(u,\varphi_{0,u}(x))\right|^2 \mathrm{d}u\right\}$$
$$\times \sum_{k=1}^m \Big[\Big(\left|Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Z_0^{-1}(s)\right|^2 - K\left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|^2\Big)$$
$$\times \left|Z_n(s)Y_{n,s}(x) - Z_0(s)Y_{0,s}(x)\right|^2\Big]\mathrm{d}s. \quad (47)$$

It follows from the estimates Equations (42) and (43) on p. 22 that for large enough K, $II \le 0$.

Consider I. First using Proposition 4 on p. 22 let us show that for $1 \le k \le m$, $s \ge 0$, and $x \in \mathbb{R}^d$, $\nabla\sigma_{n,k}(s,\varphi_{n,s}(x)) \to \nabla\sigma_{0,k}(s,\varphi_{0,s}(x))$, $n \to \infty$, in probability. Fix $s \ge 0$, $x \in \mathbb{R}^d$, and $1 \le k \le m$. We apply Proposition 4 on p. 22 to $\xi_n = \varphi_{n,s}(x)$, $n \ge 0$. The convergence $\xi_n \to \xi_0$, $n \to \infty$, in probability, follows from Lemma 3 on p. 16. Put $X = \mathbb{R}^d$, $Y = \mathbb{R}^d \times \mathbb{R}^d$, $\nu(\mathrm{d}x) = C\frac{\mathrm{d}x}{1+|x|^{d+1}}$, where C is a constant such that ν is a probability measure on $\mathbb{R}^d$. For fixed s, x, and k put $h_n = \nabla\sigma_{n,k}(s,\cdot)$, $h_0 = \nabla\sigma_{0,k}(s,\cdot)$. Since for each $s \in [0,T]$, $\nabla\sigma_{n,k}(s,\cdot) \to \nabla\sigma_{0,k}(s,\cdot)$, $n \to \infty$, in $L_2(\mathbb{R}^d)$ we can assume without lost of generality that $\nabla\sigma_{n,k}(s,y) \to \nabla\sigma_{0,k}(s,y)$, $n \to \infty$, for each $1 \le k \le m$ and almost all $s \in [0,T]$, $y \in \mathbb{R}^d$, with respect to the Lebesgue measure. Then for almost all $s \in [0,T]$, $h_n \to h_0$, $n \to \infty$, in ν. Note that the processes $(\varphi_{n,t}(x))_{t\ge 0}$, $n \ge 0$, possess transition probability densities. Thus the distributions P_{ξ_n}, $n \ge 0$, are absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}^d$ and, consequently, w.r.t. the measure ν. Using Equation (10) on p. 8 it is easy to see that the sequence of densities $\left\{\frac{dP_{\xi_n}}{d\nu} : n \ge 1\right\}$ is uniformly integrable w.r.t. the measure ν. Therefore, all the assumptions of Proposition 4 on p. 22 are fulfilled, and for almost all $s \in [0,T]$, and all $x \in \mathbb{R}^d$,

$$\nabla\sigma_{n,k}(s,\varphi_{n,s}(x)) \to \nabla\sigma_{0,k}(s,\varphi_{0,s}(x)),\ n \to \infty, \text{ in probability } \mathbb{P}. \tag{48}$$

Let us return to I. We have

$$\begin{aligned}
&\left|Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))Z_n^{-1}(s) - Z_0(s)\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Z_0^{-1}(s)\right| \\
&\quad \le \left|Z_n(s)\nabla\sigma_{n,k}(s,\varphi_{n,s}(x))\right|\left|Z_n^{-1}(s) - Z_0^{-1}(s)\right| \\
&\quad + |Z_n(s)|\left|\nabla\sigma_{n,k}(s,\varphi_{n,s}(x)) - \nabla\sigma_{0,k}(s,\varphi_{0,s}(x))\right|\left|Z_0^{-1}(s)\right| \\
&\qquad + |Z_n(s) - Z_0(s)|\left|\nabla\sigma_{0,k}(s,\varphi_{0,s}(x))Z_0^{-1}(s)\right|.
\end{aligned}$$

Using the Hölder inequality as it was done in Equation (37) on p. 20 and taking into account the estimates Equation (42) on p. 22, Equation (43) on p. 22, and the relations Equation (44) on p. 23 and Equation (48), we get that the first expectation in the right-hand side of Equation (46) on the preceding page tends to 0 as $n \to \infty$. Thus we obtained that

$$\sup_{t\in[0,T]}\left|Z_n(t)Y_{n,t}(x) - Z_0(t)Y_{0,t}(x)\right| \to 0,\ n \to \infty, \text{ in probability } \mathbb{P}.$$

Now the assertion of the lemma can be deduce from the inequality

$$\left|Y_{n,t}(x) - Y_{0,t}(x)\right| \le \left|Z_n^{-1}(t)\right|\left|Z_n(t)Y_{n,t}(x) - Z_0(t)Y_0(t)\right| + \left|Z_n^{-1}(t) - Z_0^{-1}(t)\right||Z_0(t)Y_0(t)|$$

using standard arguments for the proof of uniform convergence, which completes the proof of Lemma 7 on p. 21. □

Making use of Lemma 3 on p. 16 and the dominated convergence theorem, for each $T > 0$, $p \geq 1$, we get the relation

$$\mathbb{E} \sup_{0\leq t\leq T} \int_U \left|\varphi_{n,t}(x) - \varphi_t(x)\right|^p \mathrm{d}x \to 0,\ n \to \infty,$$

valid for any bounded domain $U \subset \mathbb{R}^d$. Then there exists a subsequence $\{n_k :\ k \geq 1\}$ such that

$$\sup_{0\leq t\leq T} \int_U \left|\varphi_{n_k,t}(x) - \varphi_t(x)\right|^p \mathrm{d}x \to 0 \text{ a.s. as } k \to \infty.$$

Without loss of generality we can assume that

$$\sup_{0\leq t\leq T} \int_U \left|\varphi_{n,t}(x) - \varphi_t(x)\right|^p \mathrm{d}x \to 0 \text{ a.s. as } n \to \infty. \tag{49}$$

It follows from Lemma 7 on p. 21 in the similar way that for each $T > 0$, $p \geq 0$,

$$\sup_{0\leq t\leq T} \int_U \left|Y_{n,t}(x) - Y_t(x)\right|^p \mathrm{d}x \to 0,\ n \to \infty, \text{ a.s.} \tag{50}$$

Since the Sobolev space is a Banach space, the relations Equations (49) and (50) mean that $Y_t(x)$ is the matrix of the Sobolev derivatives of the solution to Equations (1) and (26) on p. 1 and on p. 15 holds.

3.2 Second step

Now we treat the general case using localization. Let the coefficients of Equation (1) on p. 1 satisfy the assumptions of Theorem 4 on p. 14. Let the functions $\beta, \gamma \in C^1(\mathbb{R}^d)$ be such that $\left|\beta(x)\right| \leq 1$; $\beta(x) = 1$, if $|x| \leq 2$; $\beta(x) = 0$, if $|x| > 3$; $\left|\gamma(x)\right| \leq 1$; $\gamma(x) = 0$, if $|x| \leq 1$; $\gamma(x) = 1$, if $|x| > 3/2$. For $R > 1$, put $\beta_R(x) = \beta(x/R)$, $\gamma_R(x) = \gamma(x/R)$. Consider the SDE

$$\begin{cases} \mathrm{d}\varphi_{R,t}(x) = a(t, \varphi_{R,t}(x))\beta_R(\varphi_{R,t}(x))\,\mathrm{d}t \\ \qquad\qquad + \sum_{k=1}^{m} \sigma_k(t, \varphi_{R,t}(x))\beta_R(\varphi_{R,t}(x))\,\mathrm{d}w_k(t) \\ \qquad\qquad + \sum_{j=1}^{m} \widetilde{\sigma}_j \gamma_R(\varphi_{R,t}(x))\,\mathrm{d}\widetilde{w}_j(t), \\ \varphi_{R,0}(x) = x, \end{cases} \tag{51}$$

where $\widetilde{\sigma}$ is a $d \times m$ constant matrix such that $\widetilde{\sigma}\widetilde{\sigma}^* > 0$; $(\widetilde{w}(t))_{t\geq 0} = (\widetilde{w}_1(t), \ldots, \widetilde{w}_m(t))_{t\geq 0}$ is an m-dimensional Wiener process independent of $(w(t))_{t\geq 0}$.

Similarly to Lemma 3 on p. 16, for each $x \in \mathbb{R}^d$, we get

$$\sup_{R>1} \mathbb{E}\Big(\big|\varphi_{R,t}(x)\big|^p + |\varphi_t(x)|^p \Big) < \infty.$$

Note that $\varphi_{R,t}(x)$ coincides with $\varphi_t(x)$ for $t \le \tau_R$, where $\tau_R = \inf\{s \ge 0 : \varphi_s(x) \ge R\}$. Then from the boundedness of the coefficients of Equation (1) on p. 1 we obtain that for all $x \in \mathbb{R}^d$,

$$\mathbb{P}\left\{ \sup_{0\le t\le T} \big|\varphi_{R,t}(x) - \varphi_t(x)\big| > \varepsilon \right\} \le \mathbb{P}\left\{ \sup_{0\le t\le T} |\varphi_t(x)| > R \right\} \to 0,\ R \to \infty.$$

It is not difficult, by analogy to Equation (49) on the preceding page, to arrive at the relation

$$\sup_{0\le t\le T} \int_U \big|\varphi_{R_k,t}(x) - \varphi_t(x)\big|^p \,\mathrm{d}x \to 0,\ k \to \infty,\ \text{a.s.,} \tag{52}$$

valid for all $x \in \mathbb{R}^d$, $p \ge 1$, and a sequence $\{R_k : k \ge 1\}$ such that $R_k \to \infty$, $k \to \infty$. It follows from Lemma 5 on p. 18 that for all $x \in \mathbb{R}^d$,

$$\sup_{R>1} \mathbb{E}\Big(\sup_{0\le t\le T} \big(\big|Y_{R,t}(x)\big|^p + |Y_t(x)|^p\big) \Big) < \infty. \tag{53}$$

According to Section 3.1 on p. 15, for each $k \ge 1$ there exists the derivative $\nabla\varphi_{R_k,t}(x)$ which, for almost all $x \in \mathbb{R}^d$, is equal to the solution of the equation

$$\begin{aligned} Y_{R_k,t}(x) = E &+ \int_0^t \beta_{R_k}(\varphi_{R_k,s}(x))\,\mathrm{d}A_{R_k,s}(\varphi_{R_k}(x))Y_{R_k,s}(x) \\ &+ \int_0^t \nabla\beta_{R_k}(\varphi_{R_k,s}(x))a(s,\varphi_{R_k,s}(x))Y_{R_k,s}(x)\,\mathrm{d}s \\ &+ \sum_{k=1}^m \int_0^t \nabla\sigma_k(s,\varphi_{R_k,s}(x))\beta_{R_k}(\varphi_{R_k,s}(x))Y_{R_k,s}(x)\,\mathrm{d}w_k(s) \\ &+ \sum_{k=1}^m \int_0^t \sigma_k(s,\varphi_{R_k,s}(x))\nabla\beta_{R_k}(\varphi_{R_k,s}(x))Y_{R_k,s}(x)\,\mathrm{d}w_k(s) \\ &+ \sum_{j=1}^m \int_0^t \widetilde{\sigma}_j \nabla\gamma_{R_k}(\varphi_{R_k,s}(x))\,\mathrm{d}\widetilde{w}_j(s). \end{aligned} \tag{54}$$

Note that $A_{R_k,t}(\varphi_{R_k}(x)) = A_t(\varphi(x))$, for $t \le \tau_{R_k}$, where $\tau_{R_k} = \inf\{t : \varphi_t(x) \ge R_k\}$. Therefore, for $t \le \tau_{R_k}$, $k \ge 1$, Equation (54) coincides with Equation (25) on p. 14. As $\tau_{R_k} \to \infty$, $k \to \infty$, we deduce that

$$\sup_{0\le t\le T} \int_U \big|Y_{R_k,t}(x) - Y_t(x)\big|^p \,\mathrm{d}x \to 0,\ k \to \infty,\ \text{almost surely,} \tag{55}$$

for $T > 0$, any bounded domain $U \subset \mathbb{R}^d$, and a sequence $\{R_k : k \geq 1\}$ such that $R_k \to \infty$ as $k \to \infty$. From Equations (52) and (55) on the previous page we get that $Y_t(x) = \nabla\varphi_t(x)$, $t \geq 0$, for λ-a.a. $x \in \mathbb{R}^d$, almost surely.

Let us verify Equation (24) on p. 14. Given $R > 1$, the coefficients of Equation (51) on p. 26 satisfy all the localizing conditions imposed on the coefficients of Equation (1) on p. 1 in Section 3.1 on p. 15. Denote by $\varphi^R_{n,t}$, $n \geq 1$, a solution to equation of the form Equation (51) on p. 26 with smooth coefficients such that for $p \geq 1$, $T > 0$, and $x \in \mathbb{R}^d$,

$$\mathbb{E}\left(\sup_{0\le t\le T}\left|\varphi^R_{n,t}(x) - \varphi_{R,t}(x)\right|^p\right) \to 0,\ n \to \infty, \tag{56}$$

$$\mathbb{E}\left(\sup_{0\le t\le T}\left|Y^R_{n,t}(x) - Y_{R,t}(x)\right|^p\right) \to 0,\ n \to \infty. \tag{57}$$

Then for all $x, h \in \mathbb{R}^d$, $v \in \mathbb{R}$,

$$\varphi^R_{n,t}(x + vh) = \varphi^R_{n,t}(x) + h\int_0^v Y^R_{n,t}(x + uh)\,\mathrm{d}u.$$

This equation, Equations (56) and (57), and Lemma 5 on p. 18 imply that for all $x, h \in \mathbb{R}^d$, $v \in \mathbb{R}$, and $R > 1$,

$$\varphi_{R,t}(x + vh) = \varphi_{R,t}(x) + h\int_0^v Y_{R,t}(x + uh)\,\mathrm{d}u.$$

By Equations (52), (53) and (55) on the previous page we get the equality

$$\varphi_t(x + vh) = \varphi_t(x) + h\int_0^v Y_t(x + uh)\,\mathrm{d}u \tag{58}$$

valid for all $x, h \in \mathbb{R}^d, v \in \mathbb{R}$, and $R > 1$, To obtain Equation (24) on p. 14 it remains to prove the L_p-continuity of $Y_t(x)$ w.r.t. x. Note that Lemma 6 on p. 21 implies the convergence

$$A_t(\varphi(x)) \to A_t(\varphi(x_0)),\ x \to x_0,\ \text{in probability.}$$

Then

$$Y_t(x) \to Y_t(x_0),\ x \to x_0,\ \text{in probability.} \tag{59}$$

This together with Lemma 5 on p. 18 entails convergence in L_p, $p > 0$. Now Equation (24) on p. 14 follows from Equations (58) and (59). This completes the proof of Theorem 4 on p. 14.

4 Appendix. Convergence of transition probability densities

In this section we prove the convergence of the transition probability densities of the processes $(\varphi_{n,t})_{t\geq 0}$, $n \geq 1$, to that of the process $(\varphi_t)_{t\geq 0}$ (Lemma 8 on the next page), which entails the convergence of characteristics of W-functionals (Lemma 9 on p. 34). The latter result is the basis of the proof of Lemma 6 on p. 21. We use the parametrix method, the transition probability densities of the processes with $a_n \equiv 0$, $n \geq 1$, being considered as the initial ones.

Suppose that σ satisfies the conditions of Theorem 4 on p. 14 and $\sigma(t,x) = \widetilde{\sigma} = const$ for $t \geq 0$, $x \in \mathbb{R}^d$ such that $|x| \geq R$, $\widetilde{\sigma}\widetilde{\sigma}^* > 0$. Let $\sigma_n, n \geq 1$, be defined by Equation (28) on p. 15. Then $\sigma_n \to \sigma, n \to \infty$, uniformly in $(t,x) \in [0,T] \times \mathbb{R}^d$. Recall that we can assume that $\sigma_n(t,x) = \widetilde{\sigma}$ for all $n \geq 1$, $t \geq 0$, and $x \in \mathbb{R}^d$ such that $|x| \geq R$.

Denote $\sigma_0 = \sigma$, $\varphi_0 = \varphi$, and for $n \geq 0$ put

$$b_n = \sigma_n \sigma_n^*.$$

Then $b_n \to b_0$, $n \to \infty$, uniformly in $(t,x) \in [0,T] \times \mathbb{R}^d$, $T > 0$.

Consider the parabolic equation

$$\frac{\partial u_n(s,x)}{\partial s} + \frac{1}{2} \sum_{i,j=1}^{d} b_n^{ij}(s,x) \frac{\partial^2 u_n(s,x)}{\partial x_i \partial x_j} = 0, \ n \geq 0.$$

It is well known that the Hölder continuity and uniform ellipticity of b_n provide the existence of a fundamental solution[19], which we denote by $g_n(s,x,t,y)$ (recall that now $a_n \equiv 0$). The function $g_n(s,x,t,y)$, $0 \leq s < t \leq T$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, is the transition probability density of the diffusion process which is a solution of the SDE

$$x_n(t) = x_n(s) + \sum_{k=1}^{m} \int_s^t \sigma_{n,k}(u, x_n(u))\, \mathrm{d}w_k(u).$$

By M. Portenko (1995, Ch. II, Lemma 3),

$$g_n(s,x,t,y) \to g_0(s,x,t,y), \ n \to \infty, \tag{60}$$

$$\frac{\partial g_n(s,x,t,y)}{\partial x_i} \to \frac{\partial g_0(s,x,t,y)}{\partial x_i}, \ 1 \leq i \leq d, \ n \to \infty, \tag{61}$$

uniformly in every domain

$$\mathcal{D}_\delta^T = \{(s,x,t,y) : 0 \leq s < t \leq T, x \in \mathbb{R}^d, y \in \mathbb{R}^d, t - s + \left|x - y\right| \geq \delta\},$$

[19] E.g. Ladyženskaja, Solonnikov, and Ural'ceva, 1967, *Linear and Quasi-Linear Equations of Parabolic Type*, Ch. IV, § 11.

for any fixed $\delta > 0$, $T > 0$.

Furthermore, for $0 \le s < t \le T$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, the estimates

$$\left|\nabla_x^l g_n(s,x,t,y)\right| \le C(t-s)^{-\frac{d+l}{2}} \exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\} \tag{62}$$

hold true. Here $n \ge 0$, $l = 0,1,2$, C, c are positive constants which depend only on d, T and $\|b_0\|_{T,\infty}$.

Now let a satisfy the condition of Theorem 4 on p. 14, and $a(t,x) = 0$ for $t \ge 0, |x| > R$. Put $a_0 = a$, and $\varphi_{0,t}(x) = \varphi_t(x)$, $t \ge 0, x \in \mathbb{R}^d$, where $\varphi_t(x)$ is the solution to Equation (1) on p. 1. Let for $n \ge 1$, a_n be defined by Equation (27) on p. 15, and $\varphi_{n,t}(x)$ be a solution of Equation (31) on p. 16. Denote by $G_n(s,x,t,y)$, $n \ge 0$, the transition probability density of the process $(\varphi_{n,t})_{t\ge 0}$. Then $G_n(s,x,t,y)$ can be constructed by the perturbation method[20] as a solution of the integral equation:

$$G_n(s,x,t,y) = g_n(s,x,t,y) + \int_s^t d\tau \int_{\mathbb{R}^d} g_n(s,x,\tau,z)\,(\nabla_z G_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z, \tag{63}$$

which satisfies the estimate

$$\left|\nabla_x^l G_n(s,x,t,y)\right| \le C'(t-s)^{-\frac{d+l}{2}} \exp\left\{-c'\frac{\left|y-x\right|^2}{t-s}\right\} \tag{64}$$

in any domain $0 \le s < t \le T$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^d$, for $n \ge 0$, $l = 0,1$. The constants C', c' can be chosen uniformly in n.

It follows from N. I. Portenko (1990, Theorem 2.1), that for $n \ge 1$, the function $G_n(s,x,t,y)$ is a fundamental solution of the parabolic equation

$$\frac{\partial u_n(s,x)}{\partial s} + \frac{1}{2}\sum_{i,j=1}^d b_n^{ij}(s,x)\frac{\partial^2 u_n(s,x)}{\partial x_i \partial x_j} + \sum_{i=1}^d a_n^i(s,x)\frac{\partial u_n(s,x)}{\partial x_i} = 0.$$

Remark 9 – Following the construction of $G_0(s,x,t,y)$ in N. I. Portenko (1990) one can observe that $G_0(s,x,t,y)$ is uniformly continuous in y uniformly on $|t-s| > \delta$, $x \in \mathbb{R}^d$, $\delta > 0$.

Lemma 8 – *$G_n(s,x,t,y) \to G_0(s,x,t,y)$, $n \to \infty$, uniformly on $\mathcal{D}_\delta^T$ for any fixed $\delta > 0, T > 0$.*

Proof. We use the idea of the proof from N. I. Portenko (1990, lemma 2.6). Denote

$$U_n(s,x,t,y) = \nabla_x G_n(s,x,t,y) - \nabla_x G_0(s,x,t,y). \tag{65}$$

[20] See N. I. Portenko, 1990, *Generalized Diffusion Processes*, Ch. 2.

From Equation (63) on the preceding page we get

$$
\begin{aligned}
U_n(s,x,t,y) &= \nabla_x g_n(s,x,t,y) - \nabla_x g_0(s,x,t,y) \\
&\quad + \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} \nabla_x g_n(s,x,\tau,z)\,(\nabla_z G_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z \\
&\quad - \int_s^t d\tau \int_{\mathbb{R}^d} \nabla_x g_0(s,x,\tau,z)\,(\nabla_z G_0(\tau,z,t,y), a_0(\tau,z))\,\mathrm{d}z \\
&= \nabla_x g_n(s,x,t,y) - \nabla_x g_0(s,x,t,y) \\
&\quad + \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} \nabla_x g_0(s,x,\tau,z)\,(U_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z \\
&\quad + \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} (\nabla_x g_n(s,x,\tau,z) - \nabla_x g_0(s,x,\tau,z)) \\
&\qquad\qquad \times (\nabla_z G_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z \\
&\quad + \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} \nabla_x g_0(s,x,\tau,z)\,(\nabla_z G_0(\tau,z,t,y), a_n(\tau,z) - a_0(\tau,z))\,\mathrm{d}z.
\end{aligned}
$$

Therefore,

$$
U_n(s,x,t,y) = \mathscr{A}_n U_n(s,x,t,y) + r_n(s,x,t,y), \tag{66}
$$

where

$$
\begin{aligned}
\mathscr{A}_n U_n(s,x,t,y) &= \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} \nabla_x g_0(s,x,\tau,z)\,(U_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z, \\
r_n(s,x,t,y) &= \sum_{k=1}^{3} I_n^k(s,x,t,y), \\
I_n^1(s,x,t,y) &= \nabla_x g_n(s,x,t,y) - \nabla_x g_0(s,x,t,y), \\
I_n^2(s,x,t,y) &= \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} (\nabla_x g_n(s,x,\tau,z) - \nabla_x g_0(s,x,\tau,z)) \\
&\qquad \times (\nabla_z G_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z, \\
I_n^3(s,x,t,y) &= \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} \nabla_x g_0(s,x,\tau,z)\,(\nabla_z G_0(\tau,z,t,y), a_n(\tau,z) - a_0(\tau,z))\,\mathrm{d}z.
\end{aligned}
$$

Recall that $a_n(t,x)$, $n \geq 0$, are bounded measurable and they have compact supports in x. So $a_n \in L_p([0,T]\times\mathbb{R}^d)$ for all $T>0, p>0, n\geq 0$. Fix $p>d+2$. Making use of the Hölder inequality and the estimate Equation (64) on the preceding page we have

(Cont. next page)

$$
I_n^2(s,x,t,y) \leq \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} |\nabla_x g_n(s,x,\tau,z) - \nabla_x g_0(s,x,\tau,z)| \left|\nabla_z G_n(\tau,z,t,y)\right| |a_n(\tau,z)|\,\mathrm{d}z
$$

$$\leq K\Bigg(\int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} |\nabla_x g_n(s,x,\tau,z) - \nabla_x g_0(s,x,\tau,z)|^q \times (t-\tau)^{-\frac{d+1}{2}q} \exp\left\{-cq\frac{\left|y-z\right|^2}{t-\tau}\right\} \mathrm{d}z\Bigg)^{1/q} \Bigg(\int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} |a_n(\tau,z)|^p \,\mathrm{d}z\Bigg)^{1/p}, \tag{67}$$

where K, c are positive constants, $1/p+1/q=1$. It follows from Equation (61) on p. 29 and M. Portenko (1995, Ch. II, Lemma 2), that $I_n^2(s,x,t,y) \to 0,\ n\to\infty$, uniformly on $\mathcal{D}_\delta^T$ for any $\delta>0,\ T>0$. The Equation (61) on p. 29 gives also that $I_n^1(s,x,t,y)\to 0$, $n\to\infty$, uniformly on $\mathcal{D}_\delta^T$. Consider $I_n^3(s,x,t,y)$. We have

$$\begin{aligned} I_n^3(s,x,t,y) &\leq \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} |\nabla_x g_0(s,x,\tau,z)|\left|\nabla_z G_0(\tau,z,t,y)\right||a_n(\tau,z)-a_0(\tau,z)|\,\mathrm{d}z \\ &\leq K\Bigg(\int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} |a_n(\tau,z)-a_0(\tau,z)|^p\,\mathrm{d}z\Bigg)^{1/p} \\ &\quad\times\Bigg(\int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} (\tau-s)^{-\frac{d+1}{2}q}\exp\left\{-cq\frac{|z-x|^2}{\tau-s}\right\} \\ &\quad\times (t-\tau)^{-\frac{d+1}{2}q}\exp\left\{-cq\frac{\left|y-z\right|^2}{t-\tau}\right\}\mathrm{d}z\Bigg)^{1/q} \\ &= K'\|a_n-a_0\|_{p,T}(t-s)^{-\frac{d+1}{2}+\gamma}\exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\}, \end{aligned} \tag{68}$$

where K' is a constant, $\gamma = \frac{p-d-2}{2p}$, $p>d+2$, $\|a\|_{p,T} = \|a\|_{L_p([0,T]\times\mathbb{R}^d)}$. For the proof of the last equality in Equation (68)[21]. Then $I_n^3(s,x,t,y)\to 0,\ n\to\infty$, uniformly on $\mathcal{D}_\delta^T$. Thus we conclude that

$$r_n(s,x,t,y)\to 0,\ n\to\infty,\ \text{uniformly on } \mathcal{D}_\delta^T \text{ for any } \delta>0,\ T>0.$$

Moreover, from Equations (62), (67) and (68) on p. 30 and on the current page we obtain the following estimate

$$\left|r_n(s,x,t,y)\right| \leq H(t-s)^{-\frac{d+1}{2}}\exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\} \tag{69}$$

valid in every domain of the form $0\leq s<t\leq T,\ x,y\in\mathbb{R}^d$. Here H is a positive constant. We obtain the above inequality for I_n^2 in the way similar to that for I_n^3.

By Equations (64) and (65) on p. 30 for all $0\leq s<t\leq T$, $x,y\in\mathbb{R}^d$,

$$\left|U_n(s,x,t,y)\right| \leq H'(t-s)^{-\frac{d+1}{2}}\exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\}, \tag{70}$$

where H' is a positive constant. Denote by $\mathscr{A}_n^k$ is the k-th power of the operator $\mathscr{A}_n$. Repeating the argument of Equation (68) on the preceding page, we get

$$\left|\mathscr{A}_n^k U_n(s,x,t,y)\right| \le C_k \|a\|_{p,T}^k (t-s)^{-\frac{d+1}{2}+k\gamma} \exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\},$$

where

$$C_k = H' C^k \left(\frac{\pi}{cq}\right)^{\frac{kd}{2q}} \left(\frac{\Gamma(\beta)}{\Gamma((k+1)\beta)}\right)^{1/q},\ q = \frac{p}{p-1},\ \gamma = \frac{p-d-2}{2p},\ \beta = q\gamma.$$

Here $k = 0,1,2,\ldots$, $0 \le s < t \le T$, $x,y \in \mathbb{R}^d$, C is the constant from the inequality of Equation (62) on p. 30. It follows from these estimates that

$$\lim_{k\to\infty} \sup_n \sup_{0\le s<t\le T,\ x,y\in\mathbb{R}^d} \left|\mathscr{A}_n^k U_n(s,x,t,y)\right| = 0. \tag{71}$$

Using estimate Equation (69) on the preceding page and arguing similarly we get for $k = 0,1,2,\ldots$, that

$$\left|\mathscr{A}_n^k r_n(s,x,t,y)\right| \le C_k' \|a\|_{p,T}^k (t-s)^{-\frac{d+1}{2}+k\gamma} \exp\left\{-c\frac{\left|y-x\right|^2}{t-s}\right\}, \tag{72}$$

where

$$C_k' = H' C^k \left(\frac{\pi}{cq}\right)^{\frac{kd}{2q}} \left(\frac{\Gamma(\beta)}{\Gamma((k+1)\beta)}\right)^{1/q}.$$

Iterating the relation Equation (66) on p. 31 and taking into account Equation (71) we deduce that

$$U_n(s,x,t,y) = \sum_{k=0}^{\infty} \mathscr{A}_n^k r_n(s,x,t,y). \tag{73}$$

The estimates Equation (72) provide the convergence of the series in the right-hand side of Equation (73) uniformly in n on $\mathscr{D}_\delta^T$. To prove that

$$\lim_{n\to\infty} U_n(s,x,t,y) = 0 \tag{74}$$

on $\mathscr{D}_\delta^T$ it is enough to show that $\mathscr{A}_n^k r_n(s,x,t,y) \to 0$, $n \to \infty$, for every fixed $k = 0,1,2,\ldots$. This can be easily obtained by induction.

For the difference $G_n - G_0$ from Equation (63) on p. 30 we have

$$G_n(s,x,t,y) - G_0(s,x,t,y) = \sum_{k=1}^{4} H_n^k(s,x,t,y),$$

where

$$H_n^1(s,x,t,y) = g_n(s,x,t,y) - g_0(s,x,t,y),$$
$$H_n^2(s,x,t,y) = \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} (g_n(s,x,\tau,z) - g_0(s,x,\tau,z))\,(\nabla G_n(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z,$$
$$H_n^3(s,x,t,y) = \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} g_0(s,x,\tau,z)\,(\nabla G_n(\tau,z,t,y) - \nabla G_0(\tau,z,t,y), a_n(\tau,z))\,\mathrm{d}z,$$
$$H_n^4(s,x,t,y) = \int_s^t \mathrm{d}\tau \int_{\mathbb{R}^d} g_0(s,x,\tau,z)\nabla G_0(\tau,z,t,y)(a_n(\tau,z) - a_0(\tau,z))\,\mathrm{d}z.$$

By Equation (60) on p. 29 we have $H_n^1(s,x,t,y) \to 0, n \to \infty$, uniformly on $\mathcal{D}_\delta^T$ for any $\delta > 0, T > 0$. It follows from Equations (60) and (74) on p. 29 and on the previous page, and the dominated convergence theorem that $H_n^2(s,x,t,y) \to 0, n \to \infty$, and $H_n^3(s,x,t,y) \to 0, n \to \infty$, uniformly on $\mathcal{D}_\delta^T$. Finally, H_n^4 satisfies the inequality

$$\left|H_n^4(s,x,t,y)\right| \le K\|a_n - a\|_{p,T}(t-s)^{-\frac{d}{2}+\gamma} \exp\left\{-c\frac{|y-x|^2}{t-s}\right\}.$$

This implies that $H_n^4(s,x,t,y) \to 0$ as $n \to \infty$ uniformly on $\mathcal{D}_\delta^T$. □

Lemma 9 – *Let* $\widetilde{\nu}(\mathrm{d}t,\mathrm{d}y) = \nu(t,\mathrm{d}y)dt$ *be a measure of the class* $\mathcal{K}$ *such that* $\mathrm{supp}(\widetilde{\nu}) \subset [0,T]\times U$ *for some* $T > 0$ *and compact set* $U \in \mathbb{R}^d$. *Then*

$$\int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} G_n(t_0,x,s,y)(\nu * \omega_n)(s,\mathrm{d}y) \to \int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} G_0(t_0,x,s,y)\nu(s,\mathrm{d}y),\ n \to \infty,$$

uniformly on $0 \le t_0 < t_0 + t \le T, x \in \mathbb{R}^d$.

Proof. We can write

$$\left|\int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} G_n(t_0,x,s,y)(\nu * \omega_n)(s,\mathrm{d}y) - \int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} G_0(t_0,x,s,y)\nu(s,\mathrm{d}y)\right|$$
$$\le I_n^1(t_0,t,x) + I_n^2(t_0,t,x),$$

where

$$I_n^1(t_0,t,x) = \left|\int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} \Big(\big(G_n(t_0,x,s,\cdot) * \omega_n\big)(y) - \big(G_0(t_0,x,s,\cdot) * \omega_n\big)(y)\Big)\nu(s,\mathrm{d}y)\right|,$$
$$I_n^2(t_0,t,x) = \left|\int_{t_0}^{t_0+t}\mathrm{d}s \int_{\mathbb{R}^d} \Big(\big(G_0(t_0,x,s,\cdot) * \omega_n\big)(y) - G_0(t_0,x,s,\cdot)(y)\Big)\nu(s,\mathrm{d}y)\right|.$$

[21] See, e.g., Friedman, 1964, *Partial differential equations of parabolic type*, Ch. 1, § 4, Lemma 3.

For any $\delta > 0$ we have

$$I_n^1(t_0,t,x) \le \int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \Big(G_n(t_0,x,s,\cdot) * \omega_n\Big)(y)\nu(s,\mathrm{d}y) + \int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \Big(G_0(t_0,x,s,\cdot) * \omega_n\Big)(y)\nu(s,\mathrm{d}y) + \int_{t_0+\delta}^{t_0+t} \mathrm{d}s \int_{\mathbb{R}^d} \Big|G_n(t_0,x,s,y) - G_0(t_0,x,s,y)\Big|\nu(s,\mathrm{d}y). \tag{75}$$

From Equation (10) on p. 8 we get

$$\begin{aligned}
&\int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \Big(G_n(t_0,x,s,\cdot) * \omega_n\Big)(y)\nu(s,\mathrm{d}y) \\
&\quad\le \int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \nu(s,\mathrm{d}y) \int_{\mathbb{R}^d} G_n(t_0,x,s,y-z)\omega_n(z)\,\mathrm{d}z \\
&\quad\le C \int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \nu(s,\mathrm{d}y) \int_{\mathbb{R}^d} \exp\left\{-c\frac{\big|y-(z+x)\big|^2}{s-t_0}\right\}\omega_n(z)\,\mathrm{d}z \\
&\quad\le C \sup_{\widetilde{x}\in\mathbb{R}^d} \int_{t_0}^{t_0+\delta} \mathrm{d}s \int_{\mathbb{R}^d} \exp\left\{-c\frac{\big|y-\widetilde{x}\big|^2}{s-t_0}\right\}\nu(s,\mathrm{d}y).
\end{aligned} \tag{76}$$

Because of the condition Equation (12) on p. 9, for each $\varepsilon > 0$, we can choose δ so small that for all $t_0 \in [0, T-\delta]$ the right-hand side of Equation (76) does not exceed $\varepsilon/2$. The same estimate for the second summand in the right-hand side of Equation (75) can be obtained similarly.

To prove the convergence of the last item in the right-hand side of Equation (75) to zero we note that for each $T > 0$ and compact set $U \subset \mathbb{R}^d$ there exists $C > 0$ such that

$$\sup_{t_0\in[0,\infty)} \int_{t_0+\delta}^{t_0+T} \mathrm{d}s \int_U \nu(s,\mathrm{d}y) < C. \tag{77}$$

Indeed, let $R > 0$ be such that $U \subset B(0,R)$. We have that for all $s \in [t_0+\delta, t_0+T]$, $x \in \mathbb{R}^d$, and $y \in \mathbb{R}^d$ such that $\big|y\big| \le R$,

$$p_0(t_0,x,s,y) \ge \frac{1}{(2\pi\delta)^{d/2}} \exp\left\{-\frac{(R+|x|)^2}{2\delta}\right\} := \frac{1}{K_\delta},$$

where $p_0(t,x,s,y)$, $0 \le t \le s$, $x \in \mathbb{R}^d, y \in \mathbb{R}^d$, is a transition probability density of a d-dimensional Wiener process. For each $x \in \mathbb{R}^d$,

(Cont. next page)

$$\sup_{t_0\in[0,\infty)} \int_{t_0+\delta}^{t_0+T} \mathrm{d}s \int_U \nu(s,\mathrm{d}y) \le K_\delta \sup_{t_0\in[0,\infty)} \int_{t_0+\delta}^{t_0+T} \mathrm{d}s \int_{|y|\le R} p_0(t_0,x,s,y)\nu(s,\mathrm{d}y)$$

$$\leq K_\delta \sup_{t_0\in[0,\infty)} \sup_{\widetilde{x}\in\mathbb{R}^d} \int_{t_0}^{t_0+T} \mathrm{d}s \int_{\mathbb{R}^d} p_0(t_0,\widetilde{x},s,y)\nu(s,\mathrm{d}y).$$

Fixed $C_0 > 0$, by the relation Equation (12) on p. 9 there exists $T_0 > 0$ such that

$$\sup_{t_0\in[0,\infty)} \sup_{\widetilde{x}\in\mathbb{R}^d} \int_{t_0}^{t_0+T_0} \mathrm{d}s \int_{\mathbb{R}^d} p_0(t_0,\widetilde{x},s,y)\nu(s,\mathrm{d}y) < C_0.$$

Then Remark 2 on p. 5 implies that there exists $C_1 > 0$ such that

$$\sup_{t_0\in[0,\infty)} \sup_{\widetilde{x}\in\mathbb{R}^d} \int_{t_0}^{t_0+T} \mathrm{d}s \int_{\mathbb{R}^d} p_0(t_0,\widetilde{x},s,y)\nu(s,\mathrm{d}y) < C_1,$$

which entails Equation (77) on the previous page. Now by Lemma 8 on p. 30 and Equation (77) the last summand in the right-hand side of Equation (75) on the previous page tends to zero uniformly on $0 \leq t_0 < t_0 + t \leq T$, $x \in \mathbb{R}^d$.

Thus we obtained that

$$\sup_{0\leq t_0<t_0+t\leq T} \sup_{x\in\mathbb{R}^d} I_n^1(t_0,t,x) \to 0, \quad n\to\infty.$$

Using the similar argument we get

$$\sup_{0\leq t_0<t_0+t\leq T} \sup_{x\in\mathbb{R}^d} I_n^2(t_0,t,x) \to 0, \quad n\to\infty.$$

This ends the proof. □

Acknowledgments

The authors appreciate Prof. Mykola Portenko for fruitful discussions.

References

Aryasova, O. V. and A. Y. Pilipenko (2012). “On properties of a flow generated by an SDE with discontinuous drift”. *Electronic Journal of Probability* **17**, no. 106, 1–20. ISSN: 1083-6489. DOI: 10.1214/EJP.v17-2138 (cit. on p. 3).

Aryasova, O. V. and A. Y. Pilipenko (2014). “On differentiability of stochastic flow for a multidimensional SDE with discontinuous drift”. *Electronic Communications in Probability* **19**, no. 45, 1–17. ISSN: 1083-589X. DOI: 10.1214/ECP.v19-2886 (cit. on p. 4).

Attanasio, S. (2010). “Stochastic flows of diffeomorphisms for one-dimensional SDE with discontinuous drift”. *Electronic Communications in Probability* **15**, no. 20, 213–226. ISSN: 1083-589X. DOI: 10.1214/ECP.v15-1545 (cit. on p. 3).

Bogachev, V. I. (2007). *Measure Theory*. 2. Berlin: Springer (cit. on p. 22).

Bogachev, V. I. and A. Y. Pilipenko (2015). "Strong solutions to stochastic equations with Lévy noise and a discontinuous drift coefficient". *Doklady Mathematics* **92** (1), pp. 471–475. ISSN: 1531-8362. DOI: 10.1134/S1064562415040213 (cit. on p. 4).

Chen, Z.-Q. (2002). "Gaugeability and Conditional Gaugeability". *Transactions of the American Mathematical Society* **354**, no. 11, 4639–4679 (cit. on p. 10).

Dynkin, E. B. (1965). *Markov Processes*. Trans. from the Russian by J. Fabius et al. 2 vols. Berlin: Springer (cit. on pp. 4–8, 10, 12).

Fedrizzi, E. and F. Flandoli (2013a). "Hölder Flow and Differentiability for SDEs with Nonregular Drift". *Stochastic Analysis and Applications* **31** (4), pp. 708–736. DOI: 10.1080/07362994.2012.628908 (cit. on p. 2).

Fedrizzi, E. and F. Flandoli (2013b). "Noise prevents singularities in linear transport equations". *Journal of Functional Analysis* **264** (6), pp. 1329–1354. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2013.01.003 (cit. on p. 2).

Flandoli, F., M. Gubinelli, and E. Priola (2010). "Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift". *Bulletin des Sciences Mathematiques* **134** (4), pp. 405–422. ISSN: 0007-4497. DOI: 10.1016/j.bulsci.2010.02.003 (cit. on p. 2).

Friedman, A. (1964). *Partial differential equations of parabolic type*. Englewood Cliffs, N.J: Prentice-Hall, Inc., 347 p. (cit. on p. 34).

Khasminskii, R. Z. (1959). "On Positive Solutions of the Equation $\mathfrak{A}U + Vu = 0$". *Theory of Probability and Its Applications* **4** (3), pp. 309–318. DOI: 10.1137/1104030 (cit. on p. 10).

Kulik, A. M. and A. Y. Pilipenko (2000). "Nonlinear Transformations of Smooth Measures on Infinite-Dimensional Spaces". *Ukrainian Mathematical Journal* **52** (9), pp. 1403–1431. ISSN: 0041-5995. DOI: 10.1023/A:1010380119199 (cit. on p. 22).

Kunita, H. (1990). *Stochastic Flows and Stochastic Differential Equations*. Cambridge Univ. Press (cit. on p. 2).

Kuwae, K. and M. Takahashi (2007). "Kato class measures of symmetric Markov processes under heat kernel estimates". *Journal of Functional Analysis* **250** (1), pp. 86–113. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2006.10.010 (cit. on p. 9).

Ladyženskaja, O. A., V. A. Solonnikov, and N. N. Ural'ceva (1967). *Linear and Quasi-Linear Equations of Parabolic Type*. Trans. from the Russian by S. Smith. **23**. Translations of Mathematical Monographs. Providence, Phode Island: American Mathematical Society. xi+648 (cit. on p. 29).

Luo, D. (2011). "Absolute continuity under flows generated by SDE with measurable drift coefficients". *Stochastic Processes and their Applications* **121** (10), pp. 2393–2415. ISSN: 0304-4149. DOI: 10.1016/j.spa.2011.05.012 (cit. on p. 17).

Meyer-Brandis, T. and F. Proske (2010). "Construction of strong solutions of SDE's via Malliavin calculus". *Journal of Functional Analysis* **258** (11), pp. 3922–3953. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2009.11.010 (cit. on pp. 2, 3).

Mohammed, S. E. A., T. Nilssen, and F. Proske (2015). "Sobolev Differentiable Stochastic Flows for SDE's with Singular Coefficients: Applications to the Transport Equation". *Annals of Probability* **43** (3), pp. 1535–1576 (cit. on pp. 2, 3).

Portenko, M. (1995). *Diffusion processes in media with membranes*. Ukrainian. **10**. Kyiv: Proc. of Institute of Mathem. of National Academy of Sci. of Ukraine (cit. on pp. 29, 32).

Portenko, N. I. (1990). *Generalized Diffusion Processes*. Translations of mathematical monographs. Providence, Phode Island: American Mathematical Society. ISBN: 9780821845387. DOI: 10.1090/mmono/083 (cit. on pp. 8, 21, 30).

Protter, P. E. (2004). *Stochastic Integration and Differential Equations*. Berlin: Springer-Verlag (cit. on p. 15).

Sznitman, A.-S. (1998). *Brownian Motion, Obstacles and Random Media*. Berlin: Springer: Springer Monographs in Mathematics (cit. on p. 10).

Veretennikov, A. Y. (1981). "On strong solutions and explicit formulas for solutions of stochastic integral equations". *Math. USSR Sborn* **39(3)**, pp. 387–403 (cit. on p. 2).

Vladimirov, V. S. (1967). *The Equation of Mathematical Physics*. Ed. by A. Jeffrey. Trans. from the Russian by A. Littlewood. **3**. Pure and Applied Mathematics. New York: Dekker, Marcel (cit. on p. 17).

Contents