On existence of strong solutions to stochastic equations with Lévy noise and differentiability with respect to initial condition[1]

V.I. Bogachev, A.Yu. Pilipenko

Abstract

We prove existence of strong solutions to stochastic equations with Lévy noise and discontinuous drifts and differentiability of solutions with respect to initial conditions.

Keywords: stochastic equation, Lévy noise, strong solution, differentiability with respect to initial conditions

AMS Subject Classification: 60H20, 60J57, 60J75, 60G52

1 Introduction

Let $N(du, dt)$ be a Poisson random measure on $(\mathbb{R}^d \setminus \{0\}) \times [0, \infty)$ with a Lévy measure $\nu(du) \times dt$, where $\widetilde{N}$ is a compensated Lévy measure (about Lévy measures, Lévy noise and related notions, see, e.g., [1] and [27]), $W(t)$ is an m-dimensional Wiener process, $a\colon \mathbb{R}^d \to \mathbb{R}^d$, $b_1, b_2\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma\colon \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^d$ are measurable mappings. Let us consider the stochastic equation

$$\varphi_t = \varphi_0 + \int_0^t a(\varphi_{s-})ds + \int_0^t \int_{|u|\le 1} b_1(\varphi_{s-}, u)\widetilde{N}(du, ds) + \int_0^t \int_{|u|>1} b_2(\varphi_{s-}, u)N(du, ds) + \int_0^t \sigma(\varphi_{s-})dW(s), t \ge 0, \quad (1.1)$$

with a non-anticipating initial condition φ_0. In case where $\varphi_0 = x$ we denote the solution by $\varphi_t(x)$, omitting indication of the variable ω of the probability space. We assume that

$$N([0, T] \times \{|u| > 1\}) < \infty \quad \text{for each } T > 0.$$

Since the measure N is atomic, the integral of the form

$$\int_0^T \int_{|u|>1} f(t, u)N(dt, du)$$

is a finite sum and always exists. So there will be no other restrictions on b_2 (moreover, the general case reduces to that of $b_2 = 0$), but the conditions on other coefficients are important and will be given below.

If the coefficients in equation (1.1) are sufficiently regular, then it is known that there exists a unique strong solution to this equation. The goal of this paper is to prove an existence and uniqueness theorem in case where the drift coefficient a can be discontinuous. We shall also prove the differentiability in L_p of the solution $\varphi_t(x)$ with respect to the initial condition x.

The problem of existence and uniqueness of a strong solutions is well studied in the case where the diffusion coefficient σ is non-degenerate. For example, if

[1]This work has been supported by the joint project 14-01-90406 of the RFBR and the project 09-01-14 of the Presidium of the NAS of Ukraine.

$b_1 = b_2 = 0$, σ satisfies the Lipschitz condition, $\sigma^*(x)\sigma(x) > 0, x \in \mathbb{R}^d$, and a is a measurable function of at most linear growth, the equation

$$d\varphi_t = a(\varphi_t)dt + \sigma(\varphi_t)dW(t), t \geq 0,$$

has a unique strong solution (see [29]).

If $\sigma = 0, b_i \neq 0$ and a does not satisfy the Lipschitz condition, then the question about existence and uniqueness of a strong solution is much more difficult for investigation. In the multi-dimensional case, only recently the existence of a strong solution has been proved in [25, 15, 9] for the equation

$$d\varphi_t = a(\varphi_t)dt + d\xi_\alpha(t), t \geq 0 \tag{1.2}$$

with a Hölder continuous drift coefficient, where $\{\xi_\alpha(t), t \geq 0\}$ is an α-stable process with $\alpha \in (1, 2)$.

The case of a discontinuous drift coefficient was considered in the one-dimensional case, see, e.g., [28, 22], where the proof of the corresponding result essentially employed the linear ordering of the real line. The multi-dimensional case with a discontinuous drift has been studied only recently, see [6].

Unlike strong solutions, the existence and uniqueness of weak solutions to (1.1) have been studied in-depth, see, e.g., [7], [12], [16], [17], [4], [19], [18], [21] and references there. In many situations, estimates on transition densities of the corresponding Markov process have been obtained. For example, the existence of a weak solution to (1.2) for $\alpha \in (1, 2)$ is proved in the case $a \in L_p(\mathbb{R}^d, \mathbb{R}^d), p \in \left(\frac{d}{\alpha-1}; \infty\right]$, and even in a more general situation, where a belongs to the Kato class $K_d^{\alpha-1}$ (see Definition 2.5 in Section 2). The transition density in this case is estimated from above by the transition density of the corresponding stable process.

We suggest a new method of proving the existence of a strong solution to (1.1) that develops the approach employed in our previous paper [6] for an additive α-stable symmetric Lévy noise with $\alpha \in (1, 2)$. Concerning the drift coefficient a we shall assume that the derivative ∇a in the sense of distributions is a measure (not necessarily absolutely continuous) for which the additive functional

$$A_t(\varphi) = \int_0^t \nabla a(\varphi_s)ds$$

of the weak solution (1.1) is well-defined. We shall also prove the differentiability of $\varphi_t(x)$ in x and derive the natural equation for the derivative

$$\begin{aligned}\nabla\varphi_t(x) = I + \int_0^t A_{ds}(\varphi(x))\nabla\varphi_{s-}(x) + \int_0^t \int_{|u|\leq 1} \nabla b_1(\varphi_{s-}(x), u)\nabla\varphi_{s-}(x)\widetilde{N}(du, ds) \\ + \int_0^t \int_{|u|>1} \nabla b_2(\varphi_{s-}(x), u)\nabla\varphi_{s-}(x)N(du, ds) \\ + \int_0^t \nabla\sigma(\varphi_{s-}(x))\nabla\varphi_{s-}(x)dW(s), t \geq 0.\end{aligned} \tag{1.3}$$

Let us also mention the recent paper [3], where the equation of the form

$$d\varphi_t = a(\varphi_t)dt + dW_H(t), t \geq 0,$$

is considered, where W_H is a fractional Brownian motion with a sufficiently small Hurst parameter H and a is a possibly discontinuous function. The idea of the proof of existence of a strong solution in [3] is related to a study of local times of the process φ_t. In general, a solution is not a Markov process. However, local times are analogs of additive functionals for Markov processes and there is some connection between our approach and that of paper [3].

The organization of our paper is as follows.

In Section 2 we explain the idea of our proof of existence and uniqueness of a strong solution to (2.9) in the case of additive noise. As an illuminating example we obtain a sufficient condition for existence and uniqueness of solutions to stochastic differential equations in case where the noise is a symmetric α-stable process with $\alpha \in (1, 2)$.

In Section 3 we prove a general theorem on existence and uniqueness of strong solutions to equation (1.1). In case of a non-additive noise some additional difficulties arise in obtaining moment estimates for solutions to stochastic equations.

In Section 4 we justify the linear equation (1.3) for the derivative with respect to the deterministic initial condition for solutions to equation (1.1). Note that in case $b_1 = b_2 = 0, \sigma =$ const equations of this type for derivatives have been considered in [2].

Finally, in Section 5 we prove some auxiliary assertions on existence of solutions and properties and convergence of additive functionals of Markov processes that are employed in this paper.

2 The idea of proof: additive noise

In this section we explain the idea of our proof of existence of strong solutions to equations with additive noise in the case where there are some a priori estimates on the transition density. In particular, we obtain a sufficient condition for existence of strong solutions to equation (1.2).

Let us consider the equation

$$d\varphi_t = a(\varphi_t)dt + d\xi(t), \quad t \geq 0, \tag{2.1}$$

where $\xi(t)$ is some Lévy process.

Let $a_0, a_1 \in C_0^1(\mathbb{R}^d, \mathbb{R}^d)$. Denote by $\varphi_t^\lambda(x)$ the solution to (2.1) with the drift coefficient $a_0 + \lambda(a_1 - a_0)$. Then

$$\varphi_t^1(x) - \varphi_t^0(x) = \int_0^1 \frac{\partial \varphi_t^\lambda(x)}{\partial \lambda} d\lambda. \tag{2.2}$$

We observe that

$$\frac{\partial \varphi_t^\lambda(x)}{\partial \lambda} = \int_0^t \nabla(a_0 + \lambda(a_1 - a_0)) \circ \varphi_s^\lambda(x) \frac{\partial \varphi_s^\lambda(x)}{\partial \lambda} ds + \int_0^t (a_1 - a_0) \circ \varphi_s^\lambda(x) ds.$$

The Gronwall lemma yields the following estimate:

$$\sup_{s\in[0,t]}\Big|\frac{\partial\varphi_s^\lambda(x)}{\partial\lambda}\Big|$$
$$\le \exp\Big\{\int_0^t |\nabla(a_0+\lambda(a_1-a_0))\circ\varphi_s^\lambda(x)|ds\Big\}\int_0^t |(a_1-a_0)\circ\varphi_s^\lambda(x)|ds.$$

By Hölder's inequality with $\theta^{-1}+\theta'^{-1}=1$ we have

$$E\sup_{s\in[0,t]}\Big|\frac{\partial\varphi_s^\lambda(x)}{\partial\lambda}\Big| \le \Big(E\exp\Big\{\theta\int_0^t |\nabla(a_0+\lambda(a_1-a_0))\circ(\varphi_s^\lambda(x))|ds\Big\}\Big)^{1/\theta}\cdot$$
$$\cdot\Big(t^{\theta'-1}E\int_0^t |(a_1-a_0)\circ(\varphi_s^\lambda(x))|^{\theta'}ds\Big)^{1/\theta'}$$
$$\le \Big(E\exp\Big\{2\theta\int_0^t |\nabla a_0\circ(\varphi_s^\lambda(x))|ds\Big\}+E\exp\{2\theta\int_0^t |\nabla a_1\circ(\varphi_s^\lambda(x))|ds\}\Big)^{1/\theta}\cdot$$
$$\cdot\Big(t^{\theta'-1}E\int_0^t |(a_1-a_0)\circ(\varphi_s^\lambda(x))|^{\theta'}ds\Big)^{1/\theta'}. \quad (2.3)$$

Suppose that for every $\lambda\in[0,1]$ there exists the distribution density $p_t^\lambda(x,y)$ of $\varphi_t^\lambda(x)$ and one has the following upper bound uniformly in λ :

$$p_t^\lambda(x,y)\le p_t(x,y),\ t>0,\ x,y\in\mathbb{R}^d,\ \lambda\in[0,1],$$

where $p_t(x,y)$ is a function measurable in all variables, $p_t(x,\cdot)$ is integrable in y for all fixed $t>0, x\in\mathbb{R}^d$.

Set

$$k_t(x,y):=\int_0^t p_s(x,y)ds.$$

On account of (2.2) and (2.3) we obtain

$$E\sup_{s\in[0,t]}|\varphi_s^1(x)-\varphi_s^0(x)|\le C(t,\nabla a_1,\nabla a_0,\theta)\Big(\int_{\mathbb{R}^d}|(a_1-a_0)(y)|^{\theta'}k_t(x,y)dy\Big)^{1/\theta'}, \quad (2.4)$$

where $C(t,\nabla a_1,\nabla a_0,\theta)$ depends only on $t>0,\theta>1$ and the exponential moments of the random variable

$$\int_0^t |\nabla a_i\circ(\varphi_s^\lambda(x))|ds.$$

We observe that that this random variable is a continuous nonnegative additive functional of a Markov process. Hence the exponential moments exist under rather weak assumptions (see Section 5 and the reasoning below).

Let now $a\colon\mathbb{R}^d\to\mathbb{R}^d$ be a locally integrable mapping. Let $\varphi_{n,t}$ denote the solution to (1.1) with the drift coefficient a^n, where $a^n=a*g_n$, $g_n(x)=n^dg(nx)$, and g is a nonnegative infinitely differentiable symmetric probability density function with compact support.

Suppose that for every $n\ge1$ there exists the transition density of the process $\varphi_{n,t}$ satisfying the estimate

$$p_t^n(x,y)\le p_t(x,y),\ t>0,\ x,y\in\mathbb{R}^d, \quad (2.5)$$

where $p_t(x,y) = p_t(x-y)$ is a function measurable in all variables, $p_t(\cdot)$ is integrable for all $t > 0$. Then (2.4) yields the inequality

$$E \sup_{s\in[0,t]} |\varphi_{n,s}(x) - \varphi_{m,s}(x)| \\ \leq C(t, \nabla a_n, \nabla a_m, \theta) \left(\int_{\mathbb{R}^d} |(a_n - a_m)(y)|^{\theta'} k_t(x-y) dy \right)^{1/\theta'}. \quad (2.6)$$

For justification of convergence of $\{\varphi_{n,s}(x), s \in [0,t]\}_{n\geq 1}$ it suffices to establish the uniform boundedness of $C(t, \nabla a_n, \nabla a_m, \theta)$ in n, m and the fact that the sequence $\{a_n\}$ is a Cauchy sequence in the space $L_{\theta'}(\mathbb{R}^d, k_t(x-y)dy)$ (later we show that the limit satisfies the original equation). We need the following definition.

Definition 2.1. *A function f and the measure ν belong to the Kato class $\mathcal{K}(p)$ with respect to a nonnegative function $p_t(x)$ $(t > 0, x \in \mathbb{R}^d)$ if*

$$\int_{\mathbb{R}^d} |f(y)| k_t(x-y) dy < \infty, \ t > 0, \ x \in \mathbb{R}^d,$$

and

$$\lim_{t\downarrow 0} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} |f(y)| k_t(x-y) dy = 0,$$

where

$$k_t(x) = \int_0^t p_s(x)\, ds$$

and, respectively,

$$\int_{\mathbb{R}^d} k_t(x-y)\nu(dy) < \infty, \ t > 0, \ x \in \mathbb{R}^d,$$

$$\lim_{t\downarrow 0} \sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} k_t(x-y)\nu(dy) = 0.$$

If ν is a signed measure or a vector measure on $\mathbb{R}^d$, then we say that $\nu \in \mathcal{K}(p)$ if $|\nu| \in \mathcal{K}(p)$, where $|\nu|$ is the total variation of ν.

Theorem 2.2. *Suppose that a locally integrable function a satisfies the following conditions:*

1) *for every n there exists the transition density $p_t^n(x,y)$ satisfying the uniform estimate* (2.5) *with some common function $p_t(x)$ such that*

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} p_t(x) dx < \infty \quad \forall T > 0;$$

2) *there exists $\theta' > 1$ such that the function $|a|^{\theta'}$ belongs to the Kato class $\mathcal{K}(p)$;*

3) *the derivative ∇a in the sense of distributions is a measure from the Kato class $\mathcal{K}(p)$.*

Then there exists a strong solution to equation (2.1).

If, in addition, equation (2.1) *has a unique weak solution, then its strong solution is unique.*

Proof. We observe that

$$\sup_n \int_{\mathbb{R}^d} |a_n(y)|^{\theta'} k_t(x-y)dy = \sup_n \int_{\mathbb{R}^d} |a * g_n(y)|^{\theta'} k_t(x-y)dy$$
$$\leq \sup_n \int_{\mathbb{R}^d} |a|^{\theta'} * g_n(y) k_t(x-y)dy = \sup_n |a|^{\theta'} * g_n * k_t(x)$$
$$= \sup_n |a|^{\theta'} * k_t * g_n(x) \leq \sup_z |a|^{\theta'} * k_t(z) < \infty. \quad (2.7)$$

Let $\theta_1 > \theta'$. Then (2.6) yields that

$$E \sup_{s\in[0,t]} |\varphi_{n,s}(x) - \varphi_{m,s}(x)|$$
$$\leq C(t, \nabla a_n, \nabla a_m, \theta_1) \left(\int_{\mathbb{R}^d} |(a_n - a_m)(y)|^{\theta_1'} k_t(x-y)dy \right)^{1/\theta'}.$$

The assumptions of the theorem and Lemma 5.3 yield the boundedness of

$$C(t, \nabla a_n, \nabla a_m, \theta_1)$$

uniformly in n, m. Next, (2.7) and the fact that $\lim_{n\to\infty} a_n(y) = a(y)$ for a.e. y with respect to Lebesgue measure imply convergence of the integrals to zero as $n, m \to \infty$.

Hence there exists a process $\{\varphi_t(x), t \geq 0\}$ such that

$$E \sup_{s\in[0,t]} |\varphi_{n,s}(x) - \varphi_s(x)| \to 0, \ n \to \infty. \quad (2.8)$$

The fact that the limiting process $\{\varphi_t(x), t \geq 0\}$ is a strong solution follows from (2.8), Corollary 9.9.11 in [5] and the uniform estimates on the densities (2.5).

Remark 2.3. We have proved the theorem in the case where the initial condition is x and is not random. Similarly, we can prove our assertion for every non-anticipating initial condition. It can be shown that one can choose a version of the process $\varphi_t(x)$ jointly measurable in t, x.

Slightly modifying the proof in [11], we obtain that $\varphi_t(x)$ is a unique strong solution if the weak solution is unique. □

Example 2.4. Let us consider equation (1.2) in the case, where $\{\xi_\alpha(t), t \geq 0\}$ is a symmetric α-stable process with $\alpha \in (1, 2)$. Existence of weak solutions to (1.2) is known in the case where $a \in L_p(\mathbb{R}^d, \mathbb{R}^d), p \in \left(\frac{d}{\alpha-1}; \infty\right]$, see [23], and even in a more general case where a belongs to the Kato class $K_d^{\alpha-1}$. The transition density is estimated from above by the transition density of the corresponding stable process. Let us give the corresponding definition of the Kato class in this situation and formulate a result on properties of weak solutions.

Definition 2.5. *A function f and a measure ν belong to the Kato class K_d^β if*

$$\lim_{\varepsilon\downarrow 0} \sup_{x\in\mathbb{R}^d} \int_{|x-y|\leq\varepsilon} |x-y|^{\beta-d} |f(y)| dy = 0$$

and, respectively,

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathbb{R}^d}\int_{|x-y|\le\varepsilon}|x-y|^{\beta-d}\nu(dy)=0.$$

If ν is a signed measure or a vector measure on $\mathbb{R}^d$, then we say that $\nu\in K_d^\beta$ if $|\nu|\in K_d^\beta$, where $|\nu|$ is the total variation of ν.

Remark 2.6. One can show that the Kato class associated with the function

$$p_t(x)=\frac{t}{(t^{1/\alpha}+|x|)^{d+\alpha}}$$

coincides with the class K_d^α, see [30], [8].

The following theorem holds.

Theorem 2.7. *If $a\in K_d^{\alpha-1},\alpha\in(1,2)$, then there exists a unique weak solution to* (1.2) *and this solution is a homogeneous Markov process the transition density of which satisfies the following estimate: for each $T>0$ there is L_T such that*

$$\frac{L_T^{-1}t}{(t^{1/\alpha}+|y-x|)^{d+\alpha}}\le p_t(x,y)\le\frac{L_Tt}{(t^{1/\alpha}+|y-x|)^{d+\alpha}}\ \forall\, t\in[0,T],\ x,y\in\mathbb{R}^d, \quad (2.9)$$

where L_T depends only on $T,\alpha,c,d,$ and the rate of decreasing to zero of the function $\sup_{x\in\mathbb{R}^d}\int_{|x-y|\le\varepsilon}|x-y|^{\alpha-d}|a(y)|dy$ as $\varepsilon\to 0+$.

In particular, (2.9) *holds if $a\in L_p(\mathbb{R}^d,\mathbb{R}^d),p\in\left(\frac{d}{\alpha-1},\infty\right]$.*

Existence is proved in [7], for uniqueness, see [10].

Thus, the hypotheses of Theorem 2.2 are fulfilled if $a\in K_d^{\alpha-1}$, $|a|^{1+\delta}\in K_d^\alpha$ for some $\delta>0$, and the distributional derivatives $\mu_{ij}:=\frac{\partial a_i}{\partial x_j}$ are measures from the class K_d^α. In this case there exists a unique strong solution to equation (1.2). These conditions are satisfied, for example, by functions of the form $a=\sum_{k=1}^n f_k 1\!\!1_{A_k}$, where $f_k\in C_b^1$, and A_k is a bounded set with smooth boundary. This completes our discussion of Example 2.4.

3 The general case

Suppose that $\sigma\in C_b^1(\mathbb{R}^d,\mathbb{R}^m\otimes\mathbb{R}^d)$, the function b_1 is continuously differentiable in x and satisfies the conditions

$$\sup_x\int_{|u|\le 1}|b_1(x,u)|^2\nu(du)<\infty, \quad (3.1)$$

$$\sup_x\int_{|u|\le 1}|\nabla_x b_1(x,u)|^2\nu(du)<\infty \quad (3.2)$$

$$\sup_{x,u}|\nabla_x b_1(x,u)|<\infty. \quad (3.3)$$

Let $a_0,a_1\in C_0^1(\mathbb{R}^d,\mathbb{R}^d)$. Set $a_\lambda(x):=a_0(x)+\lambda(a_1(x)-a_0(x))$.

Suppose first that $b_2 \equiv 0$. Under the stated assumptions about coefficients, there exists a unique solution to the equation

$$\varphi_t^\lambda(x) = x + \int_0^t a_\lambda(\varphi_{s-}^\lambda(x))ds + \int_0^t \int_{|u|\le 1} b_1(\varphi_{s-}^\lambda(x), u)\widetilde{N}(du, ds)$$
$$+ \int_0^t \sigma(\varphi_{s-}^\lambda(x))dW(s), t \ge 0. \quad (3.4)$$

It is also readily verified (similarly to the reasoning in [20, § 3.3]) that there exists a derivative $\frac{\partial \varphi_t^\lambda(x)}{\partial \lambda}$ in quadratic mean. Moreover, $Y_t := Y_t^\lambda(x) := \frac{\partial \varphi_t^\lambda(x)}{\partial \lambda}$ satisfies the linear stochastic equation

$$Y_t = \int_0^t (\nabla a_\lambda)(\varphi_{s-}^\lambda(x))Y_{s-}ds + \int_0^t (a_1(\varphi_{s-}^\lambda(x)) - a_0(\varphi_{s-}^\lambda(x))ds$$
$$+ \int_0^t \int_{|u|\le 1} \nabla b_1(\varphi_{s-}^\lambda(x), u)Y_{s-}\widetilde{N}(du, ds) + \int_0^t \nabla\sigma(\varphi_{s-}^\lambda(x))Y_{s-}dW(s), t \ge 0 \quad (3.5)$$

and

$$\varphi_t^1(x) - \varphi_t^0(x) = \int_0^1 \frac{\partial \varphi_t^\lambda(x)}{\partial \lambda} d\lambda. \quad (3.6)$$

Our aim is to obtain estimates analogous to (2.6).

Set

$$\mathcal{E}_c(t) := \exp\left\{-c\int_0^t |\nabla a_\lambda(\varphi_{s-}^\lambda)|ds - ct\right\}.$$

Let us apply Itô's formula to $|Y_t|^2\mathcal{E}_c(t)$, see [1, § 4.4], [20, § 2.3]. For shortening notation let us consider only the case $d = 1, m = 1$. We have

$$|Y_t|^2\mathcal{E}_c(t) = \int_0^t |Y_{s-}|^2\mathcal{E}_c(s-)\Big(2\nabla a_\lambda(\varphi_{s-}^\lambda) - c|\nabla a_\lambda(\varphi_{s-}^\lambda)| - c + |\nabla\sigma(\varphi_{s-}^\lambda)|^2$$
$$+ \int_{|u|<1} \Big[(1 + \nabla b_1(\varphi_{s-}^\lambda, u))^2 - 1 - 2\nabla b_1(\varphi_{s-}^\lambda, u)\Big]\nu(du)ds\Big)$$
$$+ \int_0^t 2Y_{s-}\mathcal{E}_c(s-)(a_1 - a_0) \circ \varphi_{s-}^\lambda ds$$
$$+ \int_0^t \nabla\sigma(\varphi_{s-}^\lambda(x))Y_{s-}^2\mathcal{E}_c(s-)dW(s)$$
$$+ \int_0^t \int_{|u|<1} 2|Y_{s-}|^2\mathcal{E}_c(s-)((1 + \nabla b_1(\varphi_{s-}^\lambda, u))^2 - 1)\widetilde{N}(du, ds). \quad (3.7)$$

Remark 3.1. Under the stated assumptions about coefficients, one can show (see [20, § 3.1]) that

$$\sup_{x\in\mathbb{R}^d} \sup_{\lambda\in[0,1]} \sup_{t\in[0,T]} E|Y_t^\lambda(x)|^p < \infty, \ \forall\, p \ge 1.$$

Hence the expectation of the stochastic integrals in the right side of (3.7) is zero.

If we choose c sufficiently large, then the first integral in the right side of (3.7) is non-positive. Hence

$$\begin{aligned} E|Y_t|^2\mathcal{E}_c(t) &\le E\int_0^t 2Y_{s-}\mathcal{E}_c(s-)(a_1-a_0)\circ\varphi_{s-}^{\lambda}\,ds \\ &\le E\int_0^t Y_{s-}^2\mathcal{E}_c(s-)ds + E\int_0^t |(a_1-a_0)\circ\varphi_{s-}^{\lambda}|^2 ds \\ &= E\int_0^t Y_s^2\mathcal{E}_c(s)ds + E\int_0^t |(a_1-a_0)\circ\varphi_s^{\lambda}|^2 ds. \end{aligned}$$

The Gronwall lemma yields that

$$E|Y_t|^2\mathcal{E}_c(t) \le e^t E\int_0^t |(a_1-a_0)\circ\varphi_s^{\lambda}|^2 ds.$$

Suppose that there exists a measurable nonnegative function

$$p\colon\ (0,\infty)\times[0,\infty)\to[0,\infty),\quad (x,t)\mapsto p_t(r)$$

such that the transition densities $p_t^{\lambda}(x,y)$ of the processes φ_t^{λ} satisfy the uniform estimate

$$p_t^{\lambda}(x,y)\le p_t(|x-y|)\quad \forall\,\lambda\in[0,1],\forall\,t>0,\forall\,x,y\in\mathbb{R}^d.$$

Then we have

$$\begin{aligned} E|Y_t|^2\mathcal{E}_c(t) &\le e^t\int_0^t\int_{\mathbb{R}^d}|(a_1-a_0)(y)|^2 p_s(|x-y|)dyds \\ &= e^t\int_{\mathbb{R}^d}|(a_1-a_0)(y)|^2 k_t(|x-y|)dy, \end{aligned}$$

where

$$k_t(r)=\int_0^t p_s(r)ds.$$

By the Cauchy inequality

$$\begin{aligned} E|Y_t| = E|Y_t|(\mathcal{E}_c(t))^{1/2}(\mathcal{E}_c(t))^{-1/2} &\le \sqrt{E|Y_t|^2\mathcal{E}_c(t)}\sqrt{E(\mathcal{E}_c(t))^{-1}} \\ &\le e^t\int_{\mathbb{R}^d}|(a_1-a_0)(y)|^2k_t(|x-y|)dy\times \\ &\times\sqrt{E\exp\Big\{c\int_0^t(|\nabla a_0(\varphi_{s-}^{\lambda})|+|\nabla a_1(\varphi_{s-}^{\lambda})|)ds+ct\Big\}}. \end{aligned}\tag{3.8}$$

Theorem 3.2. *Suppose that* $\sigma\in C_b^1(\mathbb{R}^d,\mathbb{R}^m\otimes\mathbb{R}^d)$ *and that conditions* (3.1), (3.2) *and* (3.3) *are fulfilled. Let* b_2 *be Borel measurable. Let* $g\in C_0^1(\mathbb{R}^d)$ *be a nonnegative probability density. Let* $\varphi_{n,t}(x)$ *denote the solution to* (1.1) *with the drift coefficient* $a_n=a*g_n$*, where* $g_n(x):=n^dg(nx)$*, and initial condition* $\varphi_{n,0}(x)=x$.

Suppose that the transition densities $p_t^n(x,y)$ *of the processes* $\varphi_{n,t}$ *satisfy the uniform estimate*

$$p_t^n(x,y)\le p_t(|x-y|)\quad \forall\,n,\forall\,t>0,\forall\,x,y\in\mathbb{R}^d,$$

where

$$\sup_{t\in[0,T]}\int_{\mathbb{R}^d} p_t(|x|)dx<\infty \quad \forall T>0.$$

If $|a|^{2+\varepsilon}$ *belongs to the Kato class* $\mathcal{K}(p)$ *for some* $\varepsilon>0$ *and the distributional derivative of* a *is a vector measure* $\nabla a\in\mathcal{K}(p)$*, then there exists a strong solution to equation* (1.1)*.*

If, in addition, a weak solution is unique, then the indicated solution is the unique strong solution.

Proof. Note that the σ-algebra

$$\sigma\Big\{N(A\cap\{|u|\le 1\},[0,t]),\ t\ge 0, A\in\mathcal{B}(\mathbb{R}^d)\Big\}$$

and the σ-algebra

$$\sigma\Big\{N(A\cap\{|u|> 1\},[0,t]),\ t\ge 0, A\in\mathcal{B}(\mathbb{R}^d)\Big\}$$

are independent, the Poisson random measure N has only finitely many atoms on the set $\{|u|>1\}\times[0,t]$, and

$$P\Big(N(\{|u|>1\}\times[0,t])=0\Big)=e^{-\nu(\{|u|>1\})t}.$$

Our equation for an arbitrary b_2 coincides with the equation for $b_2=0$ on the time intervals where the restriction of the measure N to $\{|u|>1\}\times[0,\infty)$ has no atoms. In addition, the solution for an arbitrary b_2 on $[0,T]$ can be obtained from solutions with $b_2=0$ and finitely many jumps, see [20, § 3.5].

The following assertion is readily proved.

Lemma 3.3. *Suppose that for every initial condition there exists a unique* (*strong or weak*) *solution to equation* (1.1) *with some* b_2 *and this solution is a strong Markov process. Then, for every initial condition, there exists a unique* (*respectively, strong or weak*) *solution to equation* (1.1) *with an arbitrary* b_2 *and this solution is a strong Markov process. In addition, one has*

$$p_t^0(x,y)\le e^{\nu(\{|u|>1\})t}p_t(x,y),$$

where $p_t^0(x,y)$ *is the transition density of the solution with* $b_2=0$ *and* $p_t(x,y)$ *is the transition density of a solution with an arbitrary* b_2*.*

We can further assume without loss of generality that $b_2=0$. Convergence of the sequence $\varphi_t^n(x)$ in L_1 will follow from (3.6), (3.8) and Lemma 5.3 provided that we show that

$$\begin{aligned}&\lim_{t\to 0+}\sup_{n\ge 1}\sup_x\int_{\mathbb{R}^d}|\nabla a_n(y)|k_t(|x-y|)dy=0,\\&\lim_{n\to\infty}\sup_x\int_{\mathbb{R}^d}|(a_n-a)(y)|^2k_t(|x-y|)dy=0.\end{aligned}\tag{3.9}$$

Indeed,

$$\int_{\mathbb{R}^d} |\nabla a_n(y)| k_t(|x-y|)dy = \int_{\mathbb{R}^d} |\nabla a * g_n(y)| k_t(|x-y|)dy$$
$$\leq \int_{\mathbb{R}^d} (|\nabla a| * g_n(y)) k_t(|x-y|)dy$$
$$= |\nabla a| * g_n * k_t(x) = |\nabla a| * k_t * g_n(x) \leq \sup_z |\nabla a| * k_t(z) \to 0,\ t \to 0+,$$

because $\nabla a \in \mathcal{K}(p)$.

The proof of relationship (3.9) is similar to the justification of the corresponding step in Theorem 2.2.

Now [5, Corollary 9.9.11] and the uniform estimate for densities imply convergence in probability

$$a_n(\varphi_t^n(x)) \xrightarrow{P} a_n(\varphi_t(x)),\ \sigma(\varphi_t^n(x)) \xrightarrow{P} \sigma(\varphi_t(x)),\ b_1(\varphi_t^n(x), u) \xrightarrow{P} b_1(\varphi_t(x), u) \quad (3.10)$$

as $n \to \infty$ for every fixed $t > 0$. The possibility to pass to the limit under the integral sign in the equations for $\varphi_{n,t}(x)$ follows from (3.10) and the uniform estimates for the second moments of the expressions under the integral sign, which follow from our assumptions about the coefficients of the equation.

Uniqueness of a strong solution in the case where we have existence of a strong solution and uniqueness of a weak solution is proved similarly to Theorem 2.2. □

4 The equation for the derivative in initial condition

For investigation of differentiability of $\varphi_{n,t}(x)$ in x it is convenient to assume that equation (1.1) is written in the form

$$\varphi_t(x) = x + \int_0^t a(\varphi_{s-}(x))ds + \int_0^t \int_{u\in\mathbb{R}^d} b_1(\varphi_{s-}(x), u)\widetilde{N}(du, ds)$$
$$+ \int_0^t \sigma(\varphi_{s-}(x))dW(s), t \geq 0. \quad (4.1)$$

Suppose that

$$\sup_x \int_{\mathbb{R}^d} (|b_1(x,u)|^2 + |\nabla b_1(x,u)|^2)\nu(du) < \infty \quad (4.2)$$

and that for (4.1) the hypotheses of Theorem 3.2 are fulfilled. Then we have well-defined additive functionals

$$A_t^{\partial_{x_j} a^i}(\varphi(x)) = \int_0^t \partial_{x_j} a^i(\varphi_s(x))ds,$$

see Theorem 5.9, where $a = (a^1, \ldots, a^d)$. Let

$$A_t(\varphi(x)) = A_t^{\nabla a}(\varphi(x)) = \| A_t^{\partial_{x_j} a^i}(\varphi(x)) \|_{i,j=1}^d.$$

As a function of t the mapping $A_t(\varphi(x))$ has bounded variation.

Theorem 4.1. *Suppose that in addition to the stated assumption we have the locally uniform convergence of the transition densities of the processes $\varphi_{n,t}(x)$ from Theorem 3.2 to the transition density of $\varphi_t(x)$:*

$$\sup_{\delta\le t\le\delta^{-1}}\ \sup_{|x|,|y|\le\delta^{-1}} |p_t^{\varphi_n}(x,y)-p_t^{\varphi}(x,y)|=0 \quad \forall\,\delta>0. \tag{4.3}$$

Then $\varphi_{n,t}(x)$ is continuously differentiable in $L^p(\Omega)$ whenever $p\ge 1$, i.e., for all $x,h\in\mathbb{R}^d$ we have

$$\frac{\varphi_t(x+uh)-\varphi_t(x)}{u}\xrightarrow{L^p(\Omega)}\nabla\varphi_t(x)h \text{ as } u\to 0.$$

The derivative $Y_t(x)=\nabla\varphi_t(x)$ satisfies the linear stochastic differential equation

$$Y_t(x)=I+\int_0^t Y_{s-}(x)dA_s(\varphi(x))+\int_0^t\int_{u\in\mathbb{R}^d}\nabla b_1(\varphi_{s-}(x),u)Y_{s-}(x)\widetilde{N}(du,ds) \\ +\int_0^t\nabla\sigma(\varphi_{s-}(x))Y_{s-}(x)dW(s),t\ge 0.$$

Proof. The processes $\varphi_t^n(x)$ are continuously differentiable in $L^p(\Omega)$ with respect to the initial condition and

$$\nabla\varphi_{n,t}(x)=I+\int_0^t\nabla\varphi^n_{s-}(x)dA^n_s(\varphi(x)) \\ +\int_0^t\int_{u\in\mathbb{R}^d}\nabla b_1(\varphi_{n,s-}(x),u)\nabla\varphi_{n,s-}(x)\widetilde{N}(du,ds) \\ +\int_0^t\nabla\sigma(\varphi_{n,s-}(x))\nabla\varphi_{n,s-}(x)dW(s),t\ge 0, \tag{4.4}$$

where

$$A^n_t(\varphi_{n,t}(x))=\int_0^t\nabla a_n(\varphi_{n,s}(x))ds.$$

For all $n\ge 1,x,h\in\mathbb{R}^d,u\in\mathbb{R},t\ge 0$ one has

$$\varphi_{n,t}(x+uh)=\varphi_{n,t}(x)+\int_0^u\nabla\varphi_{n,t}(x+zh)hdz. \tag{4.5}$$

Let us justify passage to the limit in (4.5) as $n\to\infty$.

It follows from Lemma 5.3 that

$$\sup_n\sup_x E|A^n_t(\varphi_{n,t}(x))|^p+\sup_x E|A_t(\varphi(x))|^p<\infty \quad \forall\,p\ge 1.$$

It is readily verified that

$$\begin{aligned}&\sup_{t\in[0,T]}\ \sup_n\sup_x E|\nabla\varphi_{n,t}(x))|^p<\infty,\\ &\sup_{t\in[0,T]}\ \sup_x E|Y_t(x)|^p<\infty, \quad \forall\,p\ge 1,T>0.\end{aligned} \tag{4.6}$$

Set

$$Z_t^n := A_t^n(\varphi_{n,t}(x)) + \int_0^t \int_{u\in\mathbb{R}^d} \nabla b_1(\varphi_{n,s-}(x),u)\widetilde{N}(du,ds)$$
$$+ \int_0^t \nabla\sigma(\varphi_{n,s-}(x))dW(s) =: A_t^n(\varphi_{n,t}(x)) + M_t^n.$$

Then equation (4.4) can be written in the form

$$\nabla\varphi_t^n(x) = I + \int_0^t (dZ_s^n)\nabla\varphi_{n,s-}(x). \tag{4.7}$$

We need the following fact: for the martingales M_t^n and each $T>0$, there holds the uniform convergence in probability

$$\sup_{t\in[0,T]} |M_t^n - M_t| \xrightarrow{P} 0,\ n\to\infty. \tag{4.8}$$

where

$$M_t = \int_0^t \int_{u\in\mathbb{R}^d} \nabla b_1(\varphi_{s-}(x),u)\widetilde{N}(du,ds) + \int_0^t \nabla\sigma(\varphi_{s-}(x))dW(s),$$

and $\varphi_t(x)$ is the solution constructed in Theorem 3.2. The proof of this fact follows from the Kolmogorov inequality for the supremum of martingales combined with convergence (3.10), the boundedness of $\nabla\sigma$ and (3.1).

From Corollary 5.14 we obtain the uniform convergence in probability for the corresponding additive functionals:

$$\sup_{t\in[0,T]}\left| A_t^{(\partial_{x_j}a^i)^{\pm}}(\varphi(x)) - A_t^{(\partial_{x_j}a^i)^{\pm}*g_n}(\varphi_{n,t}(x))\right| \xrightarrow{P} 0,\ n\to\infty, \forall T>0.$$

We recall that $a_n = a * g_n$. Hence for each $T>0$ we have

$$\sup_{t\in[0,T]}\left|\int_0^t (\partial_{x_j}a_n^i(\varphi_{n,s}(x)))^{\pm}ds - A_t^{(\partial_{x_j}a^i)^{\pm}}(\varphi(x))\right| \xrightarrow{P} 0,\ n\to\infty. \tag{4.9}$$

Generally speaking,

$$\int_0^t (\partial_{x_j}a_n^i(\varphi_{n,s}))^{\pm}ds = \int_0^t (\partial_{x_j}(a^i * g_n(\varphi_{n,s}))^{\pm}ds \neq \int_0^t (\partial_{x_j}a^i)^{\pm} * g_n(\varphi_{n,s})ds,$$

but

$$\int_0^t \partial_{x_j}a_n^i(\varphi_{n,s})ds = \int_0^t (\partial_{x_j}a_n^i(\varphi_{n,s}))^{+}ds - \int_0^t (\partial_{x_j}a_n^i(\varphi_{n,s}))^{-}ds$$
$$= \int_0^t (\partial_{x_j}a^i)^{+} * g_n(\varphi_{n,s}) - \int_0^t (\partial_{x_j}a^i)^{-} * g_n(\varphi_{n,s})ds.$$

Now, Lemma 4.8, (4.6), (4.9) and [26, Theorem 14 in § 5.4] yield convergence $\nabla\varphi_{n,t}(x)\to Y_t(x)$, $n\to\infty$ in all L_p. Hence (4.5) implies that

$$\varphi_t(x+uh) = \varphi_t(x) + \int_0^u Y_t(x+zh)hdz.$$

Similarly to the proof of Theorem 3.2 one can show that $Y_t(x)$ is continuously in x in all L_p. Hence $Y_t(x)$ is the derivative of $\varphi_t(x)$ in x. □

Remark 4.2. It readily follows from the proof of the main result in [7] that for equation (1.2), where $\{\xi_\alpha(t), t \geq 0\}$ is a symmetric α-stable process with $\alpha \in (1,2)$, condition (4.3) is fulfilled.

5 Additive functionals of Markov processes

Let us consider a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\{\mathcal{F}_t\colon\ t \geq 0\}$. Let $(X_t)_{t\geq 0}$ be a homogeneous Markov process with a phase space E adapted to the filtration $\mathcal{F}_t$. Set $\mathcal{N}_t = \sigma\{\xi(s)\colon\ 0 \leq s \leq t\}$.

Definition 5.1. *A random function function $\{A_t, t \geq 0\}$ adapted to the filtration $\{\mathcal{N}_t\}$ is called a continuous additive functional of the process $\{X_t, t \geq 0\}$ if*

1) *it is nonnegative,*
2) *continuous in t,*
3) *homogeneously additive, i.e., for all $t \geq 0, s \geq 0, x \in E$ one has*

$$A_{t+s} = A_s + \theta_s A_t \quad P_x\text{-a.s.},$$

where θ is the shift operator.

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

$$\sup_x E_x A_t < \infty,$$

then $\{A_t, t \geq 0\}$ is called a W-functional and the function

$$f_t(x) = E_x A_t$$

is called its characteristic.

We need the following result (see [13, Theorem 6.3]).

Lemma 5.2. *A W-functional is determined by its characteristic uniquely up to equivalence.*

Lemma 5.3. *Suppose that $\{A_t\}$ is a W-functional with characteristic $f_t(x)$ and*

$$\lim_{h\to 0+} \|f_h\|_\infty = 0, \tag{5.1}$$

where $\|g\|_\infty := \sup_x |g(x)|$. Then

$$E_x \exp\{\gamma A_t\} \leq C < \infty \quad \forall\, \gamma > 0, t > 0,$$

where the constant C depends on t, γ, and the rate of decreasing to zero of the norm $\|f_h\|_\infty$ as $h \to 0+$.

The proof of this lemma follows from [14, Chapter II, § 6, Lemma 3] and [24, Lemma 1.1].

Example 5.4. Let $E = \mathbb{R}^d$ and let the process $\{X_t\}$ have a transition density $p_t(x,y)$. For a nonnegative measurable function g set

$$A_t = \int_0^t g(X_s)ds.$$

We observe that

$$f_t(x) := E_x A_t = \int_0^t \int_{\mathbb{R}^d} g(y)p_s(x,y)dyds = \int_{\mathbb{R}^d} k_t(x,y)g(y)dy, \tag{5.2}$$

where

$$k_t(x,y) := \int_0^t p_s(x,y)ds. \tag{5.3}$$

The process A_t is a W-functional if

$$\sup_x \int_{\mathbb{R}^d} k_t(x,y)g(y)dy < \infty,\ t > 0. \tag{5.4}$$

Condition (5.1) is fulfilled if

$$\lim_{h\to 0+} \sup_x \int_{\mathbb{R}^d} k_h(x,y)g(y)dy = 0, \tag{5.5}$$

i.e., $g \in \mathcal{K}(p)$, see Definition 2.1.

Let us consider the process $\{\varphi_t\}$ from Theorem 2.7. Then the transition density satisfies the estimate (2.9). It follows from our discussion in Example 2.4 that in order to have (5.5) it is necessary and sufficient that the function g be in the Kato class K_d^α, see Remark 2.6 and Theorem 2.7.

Remark 5.5. (i) It is readily verified that if (5.5) holds, then (5.4) is also true.

(ii) If $\{\varphi_t(x), t \geq 0\}$ is a solution to (1.1) defined on some probability space $(\Omega, \mathcal{F}, P)$ with $\varphi_0(x) = x$, then

$$E_x A_t = E\int_0^t g(\varphi_s(x))ds.$$

Our next results are concerned with convergence of additive functionals.

Theorem 5.6. 1) *Let $A_i(t), i = 1,2$ be W-functionals of a Markov process X, and let $f_t^i(x)$ be the characteristic of $A_i(t)$. Then for all $t \geq 0$ we have*

$$E(A_1(t) - A_2(t))^2 \leq 2 \sup_{s\in[0,t]} \|f_s^1 - f_s^2\|_\infty (\|f_t^1\|_\infty + \|f_t^2\|_\infty).$$

2) *Let $\{A_n(t),\ t \geq 0\}_{n\geq 1}$ be a sequence of W-functionals of a Markov process X. Suppose that one has the uniform convergence of characteristics*

$$\lim_{n\to\infty} \sup_{0\leq u\leq t} \|f_u^n - f_u\|_\infty = 0, t > 0, \tag{5.6}$$

where $f_t(x)$ is some function.

Then $f_t(x)$ is the characteristic of some W-functional $A(t)$. Moreover, for all $t \geq 0$ we have

$$\lim_{n\to\infty} E(A_n(t) - A(t))^2 = 0.$$

The proof follows from [13, Theorem 6.4, Lemma 6.5].

Remark 5.7. In case where our formulations contain no assumptions about the distribution of $X(0)$, this means that the corresponding result is valid for every initial distribution of the process X.

Theorem 5.8. *Suppose that* $\{A(t),\ t \geq 0\}$ *is a* W*-functional of a Markov process* $\{X(t), t \geq 0\}$ *with the characteristic* $f_t(x)$ *satisfying* (5.1)*. Set*

$$A_h(t) := \int_0^t \frac{f_h(X(s))}{h} ds. \tag{5.7}$$

Then

$$E(A(t) - A_h(t))^2 \leq 8\|f_h\|_\infty \|f_{t+h}\|_\infty \leq 8\|f_h\|_\infty \|f_{t+1}\|_\infty, \quad h \in (0, 1].$$

In particular, every W*-functional satisfying* (5.1) *is a limit in quadratic mean of integral functionals of the form* (5.7)*.*

The proof follows from the reasoning in [13, Lemma 6.5, Theorem 6.6].

Theorem 5.9. *Suppose that a Markov process* X *has a transition density* $p_t(x, y)$*. Suppose that* ν *is a measure such that* $\nu \in \mathcal{K}(p)$*, i.e.,*

$$\sup_x \int_{\mathbb{R}^d} k_t(x, y)\nu(dy) \to 0,\ t \to 0+, \tag{5.8}$$

where $k_t(x, y)$ *is defined in* (5.3)*. Then*

$$f_t(x) := \int_{\mathbb{R}^d} k_t(x, y)\nu(dy)$$

is the characteristic of some W*-functional* $A_t = A_t^\nu(X)$ *of the process* X *and one has convergence in quadratic mean*

$$A_t = l.i.m._{h\to 0+} \int_0^t \frac{f_h(X_s)}{h} ds.$$

We observe that the characteristic of the integral functional (5.7) equals

$$f_t^h(x) := \int_{\mathbb{R}^d} k_t(x, y) \frac{f_h(y)}{h} dy.$$

One can show that

$$\lim_{h\to 0+} \sup_{0\leq u\leq t} \|f_u^h - f_u\|_\infty = 0, \tag{5.9}$$

(see the reasoning in the proof of [13, Theorem 6.6]).

The proof of the theorem follows from Theorem 5.8 and Theorem 5.6.

Remark 5.10. Similarly to integral functionals, where the measure $\nu(dy)$ has the form $g(y)dy$, the functional $A_t^\nu = A_t^\nu(X)$ is sometimes denoted by $\int_0^t \frac{d\nu}{dy}(X_s)ds$ in spite of the fact that the density $\frac{d\nu}{dy}$ need not exist. If $X_0 = x$, then the corresponding W-functional will be denoted by $A_t^\nu(X(x))$.

Example 5.11. Suppose that the conditions of Theorem 2.7 are fulfilled and φ_t is the solution to equation (1.2). It follows from the reasoning in Example 5.4 that

$$f_t(x) = \int_{\mathbb{R}^d} k_t(x, y)\nu(dy)$$

is the characteristic of a W-functional satisfying (5.1) precisely when ν belongs to the Kato class K_d^α.

The next theorem gives sufficient conditions for convergence of W-functionals of different Markov processes defined on a common probability space.

Theorem 5.12. *Suppose that Markov random functions $\{X_n(t), t \geq 0\}_{n\geq 0}$ with values in $\mathbb{R}^d$ are defined on a common probability space and that for every $t > 0$ one has convergence in probability*

$$X_n(t) \xrightarrow{P} X_0(t),\ n \to \infty. \tag{5.10}$$

Let the sequence of characteristics of W-functionals $\{A^n(X_n)\}$ satisfy the conditions

$$\lim_{h\to 0+} \sup_{n\geq 0} \|f_h^n\|_\infty = 0; \tag{5.11}$$

$$\lim_{n\to\infty} \sup_{|x|\leq R} |f_h^n(x) - f_h^0(x)| = 0 \quad \forall\, h > 0, R > 0, \tag{5.12}$$

and let the function $f_h^0(\cdot)$ be continuous for every $h > 0$.

Then for all $p \geq 1$ and $T > 0$ we have

$$E \sup_{t\in[0,T]} |A_t^n(X_n) - A_t^0(X_0)|^p \to 0, n \to \infty.$$

Proof. We have

$$\begin{aligned}
&|A_t^n(X_n) - A_t^0(X_0)| \\
&\quad \leq \left|A_t^n(X_n) - \int_0^t \frac{f_h^n(X_n(s))}{h} ds\right| + \left|\int_0^t \frac{f_h^n(X_n(s))}{h} ds - \int_0^t \frac{f_h^0(X_n(s))}{h} ds\right| \\
&\qquad + \left|\int_0^t \frac{f_h^0(X_n(s))}{h} ds - \int_0^t \frac{f_h^0(X_0(s))}{h} ds\right| + \left|\int_0^t \frac{f_h^0(X_0(s))}{h} ds - A_t^0(X_0)\right|.
\end{aligned}$$

For estimating the second moment of the first and last terms we apply Theorem 5.8 for sufficiently small $h > 0$ (uniformly in n).

For fixed $h > 0$, the moments of the second term converge to zero due to conditions (5.10) and (5.12), and the same is true for the third term by the continuity and boundedness of f_0^h and convergence (5.10). Hence for each $t > 0$ we have

$$E|A_t^n(X_n) - A_t^0(X_0)|^2 \to 0, n \to \infty.$$

Since the functions $A_t^n(X_n)$ are continuous and monotone in t, this yields the uniform convergence in probability

$$\sup_{t\in[0,T]} |A_t^n(X_n) - A_t^0(X_0)| \xrightarrow{P} 0, n \to \infty,\ \forall\, T > 0.$$

For completing the proof it remains to observe that (see Lemma 5.3)

$$\sup_n \sup_{t\in[0,T]} |A_t^n(X_n)|^p < \infty$$

for all $p \geq 1$ and $T > 0$. □

Remark 5.13. The assumption of continuity of f_0^h can be replaced by the following one: for every $n \geq 1$, there exists a transition density of the process X_n satisfying the estimate

$$p_t^n(x,y) \leq p_t(x,y),\ t > 0,\ x,y \in \mathbb{R}^d,$$

where $p_t(x,y) = p_t(x-y)$ is a function jointly measurable in all variables and $p_t(\cdot)$ is integrable for all $t > 0$.

Corollary 5.14. *Suppose that Markov random functions $\{X_n(t), t \geq 0\}_{n\geq 0}$ with values in $\mathbb{R}^d$ are defined on a common probability space and for every $t > 0$ one has convergence in probability* (5.10). *Suppose that*

1) *one has the locally uniform convergence of the transition densities*

$$\sup_{\delta\leq t\leq \delta^{-1}} \sup_{|x|,|y|\leq \delta^{-1}} |p_t^n(x,y) - p_t^0(x,y)| \to 0,\ n \to \infty, \quad \forall\, \delta > 0,$$

2) *there exists a function $p_t\colon\ x \mapsto p_t(|x|)$ in $L_1(\mathbb{R}^d)$ such that*

$$p_t^n(x,y) \leq p_t(|x-y|) \quad \forall\, n \geq 0,\, t > 0,\, x,y \in \mathbb{R}^d,$$

3) $\nu \in \mathcal{K}(p)$,

4) *for each $\delta > 0$ one has*

$$\lim_{R\to\infty} \sup_{\delta\leq t\leq \delta^{-1}} \int_{|x|>R} p_t(x)\nu(dx) = 0.$$

Then one has convergence of additive functionals (*see notation in Remark* 5.10)

$$E \sup_{t\in[0,T]} |A_t^\nu(X_n) - A_t^\nu(X_0)|^p \to 0,\ n\to\infty, \quad \forall\, p \geq 1, T > 0.$$

Corollary 5.15. *Suppose that the conditions of the previous corollary are fulfilled. Let $g_n(x) = n^d g(nx)$, where g is a continuous symmetric probability density. Let ν_n denote the measure with density $g_n * \nu$. Then for all $p \geq 1$ and $T > 0$ we have*

$$E \sup_{t\in[0,T]} |A_t^{\nu_n}(X_n) - A_t^\nu(X_0)|^p \to 0, n \to \infty.$$

References

[1] Applebaum D. Lévy processes and stochastic calculus. Cambridge University Press, Cambridge, 2009.

[2] Aryasova O.V., Pilipenko A.Yu. On differentiability of stochastic flow for a multidimensional SDE with discontinuous drift. Electron. Commun. Probab. 2014. V. 19, N 45. P. 1–17.

[3] Banos D., Nilssen T., Proske F. Strong existence and higher order Fréchet differentiability of stochastic flows of fractional Brownian motion driven SDE's with singular drift. arXiv:1511.02717

[4] Bass R.F. Uniqueness in law for pure jump Markov processes. Probab. Theory Related Fields. 1988. V. 79. P. 271–287.

[5] Bogachev V.I. Measure theory. V. 2. Springer, Berlin, 2007.

[6] Bogachev V.I., Pilipenko A.Yu. Strong solutions to stochastic equations with Lévy noise and a discontinuous drift coefficient. Dokl. Akad. Nauk. 2015. V. 463, N 5. P. 509–513 (in Russian); English transl.: Doklady Math. 2015. V. 92, N 1. P. 471–475.

[7] Bogdan K., Jakubowski T. Estimates of heat kernel of fractional Laplacian perturbed by gradient operators. Comm. Math. Phys. 2007. V. 271, N 1. P. 179–198.

[8] Chen Z.-Q. Gaugeability and conditional gaugeability. Trans. Amer. Math. Soc. 2002. V. 354, N 11. P. 4639–4679.

[9] Chen Z.-Q., Song R., Zhang X. Stochastic flows for Lévy process with Hölder drifts. Arxiv 1501.04758.

[10] Chen Z.-Q., Wang L. Uniqueness of stable processes with drift. ArXiv:1309.6414.

[11] Chernyi A.S. On strong and weak uniqueness for stochastic differential equations. Teor. Verojatn. i Primen. 2001. V. 46, N 3. P. 483–497 (in Russian); English transl.: Theory Probab. Appl. 2002. V. 46, N 3. P. 406–419.

[12] Debussche A., Fournier N. Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. J. Funct. Anal. 2013. V. 264, N 8. P. 1757–1778.

[13] Dynkin E.B. Markov processes. Springer, Berlin – Heidelberg, 1965.

[14] Gikhman I.I., Skorokhod A.V. The theory of stochastic processes. II. Springer, Berlin, 2004.

[15] Haadem S., Proske F. On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients. J. Funct. Anal. 2014. V. 266, N 8. P. 5321–5359.

[16] Kim P., Song R. Stable process with singular drift. Stoch. Proc. Appl. 2014. V. 124. P. 2479–2516.

[17] Knopova V., Kulik A. Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. Preprint 2014. http://arxiv.org/abs/1412.8732.

[18] Kochubei A.N. Parabolic pseudo-differential equations, hypersingular integrals and Markov processes. Math. USSR Izestija. 1989. V. 33. P. 233–259.

[19] Komatsu T. On the martingale problem for generators of stable processes with perturbations. Markov processes associated with certain integro-differential operators. Osaka J. Math. 1984. V. 21. P. 113–132.

[20] Kunita H. Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: Real and stochastic analysis, pp. 305–373, Birkhäuser Boston, 2004.

[21] Mikulevicius R., Pragarauskas H. On the martingale problem associated with non-degenerate Lévy operators. Lith. Math. J. 1992. V. 32, N 3. P. 297–311.

[22] Pilipenko A. On existence and properties of strong solutions of one-dimensional stochastic equations with an additive Lévy noise. Theory Stoch. Proc. 2012. V. 18, N 2. P. 77–82.

[23] Podolynny S.I., Portenko N.I. On multidimensional stable processes with locally unbounded drift. Random Oper. Stoch. Equ. 1995. V. 3, N 2. P. 113–124.

[24] Portenko N.I. Generalized diffusion processes. Amer. Math. Soc., Providence, Rhode Island, 1990.

[25] Priola E. Pathwise uniqueness for singular SDE's driven by stable processes. Osaka J. Math. 2012. V. 49, N 2. P. 421–447.

[26] Protter P. Stochastic integration and differential equations. 2nd ed. Springer, Berlin – New York, 2005.

[27] Sato K.I. Lv́y processes and infinitely divisible distributions. Cambridge Univ. Press, Cambridge, 1999.

[28] Tanaka H., Tsuchiya M., Watanabe S. Perturbation of drift-type for Lévy processes. J. Math. Kyoto Univ. 1974. V. 14, N 1. P. 73–92.

[29] Veretennikov A.Yu. On the strong solutions of stochastic differential equations. Teor. Verojatn. i Primen. 1979. V. 24, N 2. P. 348–360 (in Russian); English transl.: Theory Probab. Appl. 1979. V. 24, N 2. P. 354–366.

[30] Zhao Z. A probabilistic principle and generalized Schrödinger perturbation. J. Funct. Anal. 1991. V. 101. P. 162–176.

V.I. Bogachev: Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia

A.Yu. Pilipenko: Institute of Mathematics, Tereschenkivska st., 3, Kiev, 01601, Ukraine