Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 122 (2012) 2730–2757

stochastic processes and their applications

www.elsevier.com/locate/spa

On the rate of convergence of simple and jump-adapted weak Euler schemes for Lévy driven SDEs

R. Mikulevicius

University of Southern California, Los Angeles, CA, United States

Received 13 October 2011; received in revised form 24 April 2012; accepted 24 April 2012
Available online 5 May 2012

Abstract

The paper studies the rate of convergence of a weak Euler approximation for solutions to possibly completely degenerate SDEs driven by Lévy processes, with Hölder-continuous coefficients. It investigates the dependence of the rate on the regularity of coefficients and driving processes and its robustness to the approximation of the increments of the driving process. A convergence rate is derived for some approximate jump-adapted Euler scheme as well.

MSC: primary 60J75; 60J60; 60H30; 45K05; 35S10

Keywords: Parabolic integro-differential equations; Weak Euler scheme; Approximate and jump-adapted Euler schemes

1. Introduction

The paper studies the weak Euler approximation for solutions to possibly completely degenerate SDEs driven by Lévy processes. As in [12], the main goal is to investigate the dependence of the convergence rate on the regularity of coefficients and driving processes. In addition, we consider the robustness of the results to the approximation of the law of the increments of the driving noise in the whole scale of time discretization errors.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space with a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ of σ-algebras satisfying the usual conditions and $\alpha \in (0, 2]$ be fixed. Consider the following model in $\mathbf{R}^d$:

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s + \int_0^t G(X_{s-})dZ_s, \quad t \in [0, T], \tag{1.1}$$

E-mail address: mikulvcs@math.usc.edu.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2012.04.013

where $a(x) = (a^i(x))_{1\le i\le d}$, $b(x) = (b^{ij}(x))_{1\le i\le d,1\le j\le n}$, $G(x) = (G^{ij}(x))_{1\le i\le d,1\le j\le m}$, $x \in \mathbf{R}^d$ are measurable and bounded, with $a = 0$ if $\alpha \in (0,1)$ and $b = 0$ if $\alpha \in (0,2)$. The process W_s is a standard Wiener in $\mathbf{R}^n$. The last term is driven by $Z = \{Z_t\}_{t\in[0,T]}$, an m-dimensional Lévy process whose characteristic function is $\exp\{t\eta(\xi)\}$ with

$$\eta(\xi) = \int_{\mathbf{R}_0^m} \left[e^{i(\xi,\upsilon)} - 1 - i\chi_\alpha(\upsilon)(\xi,\upsilon)\right]\pi(d\upsilon),$$

where $\chi_\alpha(\upsilon) = \chi_{\{|\upsilon|\le 1\}}\mathbf{1}_{\{\alpha\in(1,2]\}}$. Hence,

$$Z_t = \int_0^t \int (1-\chi_\alpha(\upsilon))\upsilon p(ds,d\upsilon) + \int_0^t \int \chi_\alpha(\upsilon)\upsilon q(ds,d\upsilon),$$

where $p(dt,d\upsilon)$ is a Poisson point measure on $[0,\infty) \times \mathbf{R}_0^m$ ($\mathbf{R}_0^m = \mathbf{R}^m \setminus \{0\}$) with $\mathbf{E}[p(dt,d\upsilon)] = \pi(d\upsilon)dt$, and $q(dt,d\upsilon) = p(dt,d\upsilon) - \pi(d\upsilon)dt$ is the centered Poisson measure. It is assumed that Z_t is a Lévy process of order α:

$$\int (|\upsilon|^\alpha \wedge 1)\pi(d\upsilon) < \infty.$$

Let the time discretization $\{\tau_i, i = 0,\ldots,n_T\}$ of the interval $[0,T]$ with maximum step size $\delta > 0$ be a partition of $[0,T]$ such that $0 = \tau_0 < \tau_1 < \cdots < \tau_{n_T} = T$ and $\max_i(\tau_i - \tau_{i-1}) \le \delta$. The Euler approximation of X is an $\mathbb{F}$-adapted stochastic process $Y = \{Y_t\}_{t\in[0,T]}$ defined by the stochastic equation

$$Y_t = X_0 + \int_0^t a(Y_{\tau_{i_s}})ds + \int_0^t b(Y_{\tau_{i_s}})dW_s + \int_0^t G(Y_{\tau_{i_s}})dZ_s, \quad t \in [0,T], \tag{1.2}$$

where $\tau_{i_s} = \tau_i$ if $s \in [\tau_i, \tau_{i+1})$, $i = 0,\ldots,n_T - 1$. Contrary to those in (1.1), the coefficients in (1.2) are piecewise constants in each time interval of $[\tau_i, \tau_{i+1})$.

The weak Euler approximation Y is said to converge with order $\kappa > 0$ if for each bounded smooth function g with bounded derivatives, there exists a constant C, depending only on g, such that

$$|\mathbf{E}g(Y_T) - \mathbf{E}g(X_T)| \le C\delta^\kappa,$$

where $\delta > 0$ is the maximum step size of the time discretization.

The weak Euler approximation of stochastic differential equations with smooth coefficients and $G = 0$ has been consistently studied. For diffusion processes, Milstein was one of the first to investigate the order of weak convergence and derived $\kappa = 1$ [13,14]. Talay considered a class of the second order approximations for diffusion processes [18,19]. For Itô processes with jump components (a finite number of jumps in a finite interval), it was shown in [9] the first-order convergence in the case in which the coefficient functions possess fourth-order continuous derivatives. Platen and Kloeden & Platen studied not only Euler but also higher order approximations; see [5,15] and the references therein.

Protter & Talay [17] analyzed the weak Euler approximation for (1.1) with $\alpha = 2$. They proved that the order of convergence is $\kappa = 1$, provided that G, b, a and g have four bounded derivatives and the Lévy measure of Z has finite moments of the order $\mu = 8$. In this paper, we show that $\kappa = 1$ can be achieved when $\mu = 4$ and there still is some order of convergence for $\mu \in (2,4]$. Moreover, we assume β-Lipschitz continuity of the coefficients and g and derive that for $\alpha < \beta \le \mu \le 2\alpha$ the order of convergence $\kappa = \frac{\beta}{\alpha} - 1$. In particular, when $\beta = \mu = 2\alpha$ with $\alpha \in (0,2)$ (the diffusion part is absent), the convergence order is still $\kappa = 1$.

As in [10,12], this paper employs the idea of Talay (see [18]) and uses the solution to the backward Kolmogorov equation associated with X_t, Itô's formula, and one-step estimates. Since one step estimates were derived in [12], the main difficulty is to solve the degenerate backward Kolmogorov equation in Lipschitz classes (see Theorem 4 below). We obtain the solution of the degenerate equation as a limit of solutions to regularized (nondegenerate) equations. Although the solution to (1.1) is strong and probabilistic arguments are applied for the uniform Lipschitz estimates of the approximating sequence, contrary to [17], we do not use derivatives of the stochastic flows.

If (1.1) has a nondegenerate main part, some assumptions imposed can be relaxed (see [12,10], Kubilius & Platen [8] and Platen & Bruti-Liberati [16]). More complex and higher order schemes were studied and discussed, for example, by Cont and Tankov, Jourdain and Kohatsu-Higa (see [1,4] and the references therein).

Motivated by the difficulty to approximate the increments of the driving processes, Jacod et al. in [3], studied the approximated Euler scheme where the increments of Z are substituted by i.i.d. random variables that are easier to simulate. There are two sources of errors in this case. One comes from time discretization and the other one from substitution. We extend some of the results in [3] to the whole rate scale and show that the errors add up. In particular, the driving process Z can be replaced with a Levy process $\tilde{Z}$ having finite number of jumps in $[0, T]$ by possibly cutting small jumps of Z and sometimes replacing them with a Wiener process or drift. In addition, we consider a simple jump-adapted Euler scheme and show that presence of $\tilde{Z}$-jump moments in the partition $\{\tau_i\}$ influences the convergence rate. The approximation itself is simpler and assumptions imposed are different than those introduced by Kohatsu-Higa and Tankov in [6] (see the references therein as well) for a more sophisticated (higher order) jump-adapted scheme.

The paper is organized as follows. In Section 2, some notation is introduced, the main results stated and the proof of the main theorem is outlined. In Section 3, we present the essential technical results about backward degenerate Kolmogorov equation, followed by the proof of the main theorem in Section 4. The robustness of the approximation and jump-adapted Euler scheme is considered as well. In the last section, we discuss the optimality of the imposed assumptions.

2. Notation and main result

Denote $H = [0, T] \times \mathbf{R}^d$, $\mathbf{N} = \{0, 1, 2, \ldots\}$, $\mathbf{R}^d_0 = \mathbf{R}^d \setminus \{0\}$. For $x, y \in \mathbf{R}^d$, write $(x, y) = \sum_{i=1}^d x_i y_i$. For $(t, x) \in H$, multi-index $\gamma \in \mathbf{N}^d$ with $D^\gamma = \frac{\partial^{|\gamma|}}{\partial x_1^{\gamma_1} \cdots \partial x_d^{\gamma_d}}$, and $i, j = 1, \ldots, d$, denote

$$
\begin{aligned}
&\partial_t u(t, x) = \frac{\partial}{\partial t} u(t, x), \qquad D^k u(t, x) = \left(D^\gamma u(t, x)\right)_{|\gamma|=k}, \quad k \in \mathbf{N},\\
&\partial_i u(t, x) = u_{x_i}(t, x) = \frac{\partial}{\partial x_i} u(t, x), \qquad \partial^2_{ij} u(t, x) = u_{x_i x_j}(t, x) = \frac{\partial^2}{\partial x_i x_j} u(t, x),\\
&\partial_x u(t, x) = \nabla u(t, x) = \nabla_x u(t, x) = (\partial_1 u(t, x), \ldots, \partial_d u(t, x)),\\
&\Delta u(t, x) = \sum_{i=1}^d u_{x_i x_i}(t, x).
\end{aligned}
$$

For a smooth function v on $\mathbf{R}^d$ and $k \in \mathbf{N}$, denote

$$v^{(k)}(x;\xi^1,\ldots,\xi^k)=\sum_{i_1,\ldots,i_k=1}^{d} v_{x_{i_1}\cdots x_{i_k}}(x)\xi^1_{i_1}\cdots\xi^k_{i_k},\quad x,\xi^i\in\mathbf{R}^d, i=1,\ldots,k.$$

In particular, $v^{(1)}(x;\xi)=(\nabla v(x),\xi), x,\xi\in\mathbf{R}^d$.

For $\beta=[\beta]^- + \{\beta\}^+ > 0$, where $[\beta]^-\in\mathbf{N}$ and $\{\beta\}^+\in(0,1]$, let $\tilde{C}^\beta(H)$ denote the Lipschitz space of measurable functions u on H such that the norm

$$|u|_\beta=\sum_{|\gamma|\le[\beta]^-}|D^\gamma_x u(t,x)|_0+\sup_{\substack{|\gamma|=[\beta]^-,\\ t,x\ne\tilde{x}}}\frac{|D^\gamma_x u(t,x)-D^\gamma_x u(t,\tilde{x})|}{|x-\tilde{x}|^{\{\beta\}^+}}$$

is finite, where $|v|_0=\sup_{(t,x)\in H}|v(t,x)|$. We denote $\tilde{C}^\beta(\mathbf{R}^d)$ the corresponding function space on $\mathbf{R}^d$.

$C=C(\cdot,\ldots,\cdot)$ denotes constants depending only on quantities appearing in parentheses. In a given context, the same letter is (generally) used to denote different constants depending on the same set of arguments.

The main result of this paper is the following statement.

Theorem 1. *Let $\alpha<\beta\le\mu\le 2\alpha$, and assume $a^i, b^{ij}\in\tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij}\in\tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$, and*

$$\int_{|\upsilon|\le 1}|\upsilon|^\alpha d\pi+\int_{|\upsilon|\ge 1}|\upsilon|^\mu\pi(d\upsilon)<\infty,$$

where π is the Lévy measure of the driving process Z. Then there is a constant C such that for all $g\in\tilde{C}^\beta(\mathbf{R}^d)$

$$|\mathbf{E}g(Y_T)-\mathbf{E}g(X_T)|\le C|g|_\beta\delta^{\frac{\beta}{\alpha}-1}.$$

Applying Theorem 1 to the case $\alpha=2$ we have an obvious consequence in the jump–diffusion case.

Corollary 1. *Consider the jump–diffusion case ($\alpha=2$)*

$$X_t=X_0+\int_0^t a(X_s)ds+\int_0^t b(X_s)dW_s+\int_0^t G(X_{s-})dZ_s,\quad t\in[0,T].$$

Let $2<\beta\le\mu\le 4$. Assume $a, b^{ij}, G^{ij}\in\tilde{C}^\beta(\mathbf{R}^d)$ and

$$\int_{|\upsilon|\le 1}|\upsilon|^2\pi(d\upsilon)+\int_{|\upsilon|>1}|\upsilon|^\mu\pi(d\upsilon)<\infty.$$

Then there is a constant C such that for all $g\in\tilde{C}^\beta(\mathbf{R}^d)$

$$|\mathbf{E}g(Y_T)-\mathbf{E}g(X_T)|\le C|g|_\beta\delta^{\frac{\beta}{2}-1}.$$

An immediate extension of Theorem 1 (for the test function $g\in\tilde{C}^\nu(\mathbf{R}^d)$ with $\nu\in(0,\beta]$) is the following statement.

Corollary 2. *Let $\alpha<\beta\le\mu\le 2\alpha$, and assume $a^i, b^{ij}\in\tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij}\in\tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$, and*

$$\int_{|\upsilon|\le 1}|\upsilon|^\alpha d\pi+\int_{|\upsilon|\ge 1}|\upsilon|^\mu\pi(d\upsilon)<\infty,$$

where π is the Lévy measure of the driving process Z. Let $\nu \in (0, \beta]$. Then there is a constant C such that for all $g \in \tilde{C}^{\nu}(\mathbf{R}^d)$

$$|\mathbf{E}g(Y_T) - \mathbf{E}g(X_T)| \leq C|g|_{\nu}\delta^{\nu\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)}.$$

Remark 1. In particular, if $\alpha \in [1, 2]$, $\mu = \beta = 2\alpha$ and g is Lipschitz ($\nu = 1$), then the convergence rate $\kappa = \frac{1}{2\alpha}$.

2.1. Approximate simple Euler scheme

Following [3], for $\sigma \in (0, 1), \delta > 0$, we choose a time discretization $\{\tau_i\}$ and replace the increments of the driving process $Z_{\tau_{i+1}} - Z_{\tau_i}$ in (1.2) by $\mathcal{F}_{\tau_i}$-conditionally independent random variables $\zeta_i, i = 0, \ldots, n_T - 1$. We assume that there is a function $\phi(\sigma)$ such that $\lim_{\sigma\to 0}\phi(\sigma) = 0$ and for $i = 0, \ldots, n_T - 1$,

$$\left|\mathbf{E}[h(Z_{\tau_{i+1}} - Z_{\tau_i}) - h(\zeta_{i+1})|\mathcal{F}_{\tau_i}]\right| \leq C|h|_{\beta}\phi(\sigma)(\tau_{i+1} - \tau_i), \quad h \in \tilde{C}^{\beta}(\mathbf{R}^d) \tag{2.1}$$

with some constant C, independent of σ, δ and h. Let $\xi_t = 0$ if $0 \leq t < \tau_1, \xi_t = \zeta_i$ if $t_i \leq t < t_{i+1}, i = 1, \ldots, n_T - 1$. We still assume that $\max_i(\tau_{i+1} - \tau_i) \leq \delta$ and approximate X_t by

$$\tilde{Y}_t = X_0 + \int_0^t a(\tilde{Y}_{\tau_{i_s}})ds + \int_0^t b(\tilde{Y}_{\tau_{i_s}})dW_s + \int_0^t G(\tilde{Y}_{\tau_{i_s}})d\xi_s, \quad t \in [0, T]. \tag{2.2}$$

In this case $\tilde{Y}_t$ depends on δ and σ.

In the following example, we approximate the increments of Z_t by the increments of a Lévy process with finite number of jumps in $[0, T]$. This approximation is constructed by cutting small jumps of Z_t. We replace the small jump part by appropriately chosen drift if $\alpha < \beta \in (1, 2], \alpha \in (0, 1]$. If $\alpha < \beta \in (2, 3], \alpha \in (1, 2]$, the small jump part is replaced by a Wiener process. Given $\sigma \in (0, 1)$, we denote B^{σ} the square root of the positive definite $m \times m$-matrix $\left(\int_{|\upsilon|\leq\sigma} \upsilon_i \upsilon_j d\pi\right)_{1\leq i,j\leq m}$. Let $\tilde{W}_t$ be a standard independent Wiener process in $\mathbf{R}^m$.

Example 1. For $\sigma \in (0, 1)$ we approximate

$$Z_t = \int_0^t \int (1 - \chi_{\alpha}(\upsilon))\upsilon p(ds, d\upsilon) + \int_0^t \int \chi_{\alpha}(\upsilon)\upsilon q(ds, d\upsilon), \quad t \in [0, T],$$

by

$$\tilde{Z}_t = Z_t^{\sigma} + R_t^{\sigma},$$

with

$$Z_t^{\sigma} = \int_0^t \int_{|\upsilon|>\sigma} (1 - \chi_{\alpha}(\upsilon))\upsilon p(ds, d\upsilon) + \int_0^t \int_{|\upsilon|>\sigma} \chi_{\alpha}(\upsilon)\upsilon q(ds, d\upsilon)$$

and

$$R_t^{\sigma} = \begin{cases} t\int_{|\upsilon|\leq\sigma} \upsilon\pi(d\upsilon) & \text{if } \alpha < \beta \in (1, 2], \alpha \in (0, 1], \\ B^{\sigma}\tilde{W}_t & \text{if } \alpha < \beta \in (2, 4], \alpha \in (1, 2], \\ 0 & \text{otherwise.} \end{cases}$$

In this case (see Lemma 6 below) (2.1) holds with

$$\phi(\sigma) = \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta\wedge 3} d\pi$$

and

$$\zeta_{i+1} = \tilde{Z}_{\tau_{i+1}} - \tilde{Z}_{\tau_i}, \quad i = 0, \ldots, n_T - 1. \tag{2.3}$$

We show that time discretization and substitution errors add up.

Theorem 2. *Let $\alpha < \beta \leq \mu \leq 2\alpha$, and let $a^i, b^{ij} \in \tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$, and*

$$\int_{|\upsilon|\leq 1} |\upsilon|^\alpha d\pi + \int_{|\upsilon|\geq 1} |\upsilon|^\mu \pi(d\upsilon) < \infty,$$

where π is the Lévy measure of the driving process Z. Assume that there is a function $\phi(\sigma)$ such that $\lim_{\sigma\to 0}\phi(\sigma) = 0$ and for $i = 0, \ldots, n_T - 1$,

$$\left|\mathbf{E}h(Z_{\tau_{i+1}} - Z_{\tau_i}) - \mathbf{E}h(\zeta_{i+1})\right| \leq C|h|_\beta \phi(\sigma)(\tau_{i+1} - \tau_i), \quad h \in \tilde{C}^\beta(\mathbf{R}^d), \tag{2.4}$$

for some constant C.

Then there is a constant C (independent of σ, δ) such that for all $g \in \tilde{C}^\beta(\mathbf{R}^d)$

$$|\mathbf{E}g(\tilde{Y}_T) - \mathbf{E}g(X_T)| \leq C|g|_\beta[\delta^{\frac{\beta}{\alpha}-1} + \phi(\sigma)].$$

The same way as Corollary 2 (see the proof below) we have the following statement.

Corollary 3. *Let assumptions of Theorem 2 hold and $\nu \in (0, \beta]$. Then there is a constant C such that for all $g \in \tilde{C}^\nu(\mathbf{R}^d)$*

$$|\mathbf{E}g(\tilde{Y}_T) - \mathbf{E}g(X_T)| \leq C|g|_\nu[\delta^{\nu(\frac{1}{\alpha}-\frac{1}{\beta})} + \phi(\sigma)^{\frac{\nu}{\beta}}].$$

Remark 2. (i) Assume the assumptions of Theorem 2 hold. Since $\lim_{\sigma\to 0}\phi(\sigma) = 0$, for each $\delta > 0$ there is $\sigma = \sigma(\delta)$ such that $\phi(\sigma(\delta)) \leq \delta^{\frac{\beta}{\alpha}-1}$ and therefore

$$|\mathbf{E}g(\tilde{Y}_T) - \mathbf{E}g(X_T)| \leq C|g|_\beta \delta^{\frac{\beta}{\alpha}-1}.$$

In particular, if $\phi(\sigma) \leq C\sigma^\mu$ with $\mu > 0$ (it is the case in Example 1 for a small jump α'-stable-like driving process Z with $\alpha' < \alpha$), then we can choose $\sigma^\mu = \delta^{\frac{\beta}{\alpha}-1}$ or $\sigma = \delta^{(\frac{\beta}{\alpha}-1)\mu^{-1}}$.

(ii) In order to study precisely the case of unbounded test functions (like one in [3]), one would need to solve first the backward Kolmogorov equation in Hölder spaces with weights that are defined by the powers of $w(x) = \left(1 + |x|^2\right)^{1/2}, x \in \mathbf{R}^d$.

Applying Theorem 2 to the model of Example 1 we have

Proposition 1. *Let $\alpha < \beta \leq \mu \leq 2\alpha$, and let $a^i, b^{ij} \in \tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$, and*

$$\int_{|\upsilon|\leq 1} |\upsilon|^\alpha d\pi + \int_{|\upsilon|\geq 1} |\upsilon|^\mu \pi(d\upsilon) < \infty,$$

where π is the Lévy measure of the driving process Z. For the approximate Euler scheme in Example 1, there is a constant C (independent of σ,δ) such that for all $g\in\tilde{C}^{\beta}(\mathbf{R}^d)$

$$|\mathbf{E}g(\tilde{Y}_T)-\mathbf{E}g(X_T)|\leq C|g|_{\beta}\left[\delta^{\frac{\beta}{\alpha}-1}+\int_{|\upsilon|\leq\sigma}|\upsilon|^{\beta\wedge 3}d\pi\right].$$

2.2. Approximate jump-adapted Euler scheme

As in Example 1, for $\sigma\in(0,1)$ we approximate the increments of the driving process

$$Z_t=\int_0^t\int(1-\chi_\alpha(\upsilon))\upsilon p(ds,d\upsilon)+\int_0^t\int\chi_\alpha(\upsilon)\upsilon q(ds,d\upsilon),\quad t\in[0,T],$$

by the increments of

$$\tilde{Z}_t=Z_t^{\sigma}+R_t^{\sigma},$$

with

$$Z_t^{\sigma}=\int_0^t\int_{|\upsilon|>\sigma}(1-\chi_\alpha(\upsilon))\upsilon p(ds,d\upsilon)+\int_0^t\int_{|\upsilon|>\sigma}\chi_\alpha(\upsilon)\upsilon q(ds,d\upsilon)$$

and

$$R_t^{\sigma}=\begin{cases}t\int_{|\upsilon|\leq\sigma}\upsilon\pi(d\upsilon) & \text{if } \alpha<\beta\in(1,2],\alpha\in(0,1],\\ B^{\sigma}\tilde{W}_t & \text{if } \alpha<\beta\in(2,4],\alpha\in(1,2],\\ 0 & \text{otherwise,}\end{cases}$$

where B^{σ} is the square root of the positive definite $m\times m$-matrix $\left(\int_{|\upsilon|\leq\sigma}\upsilon_i\upsilon_j d\pi\right)_{1\leq i,j\leq m}$ and $\tilde{W}_t$ is a standard independent Wiener process in $\mathbf{R}^m$.

Given $\sigma\in(0,1),\delta>0$, consider the following Z^{σ}-jump-adapted time discretization (see [9]): $\tau_0=0$,

$$\tau_{i+1}=\inf\left(t>\tau_i:\Delta Z_t^{\sigma}\neq 0\right)\wedge(\tau_i+\delta)\wedge T. \tag{2.5}$$

In this case, the time discretization $\{\tau_i,i=0,\ldots,n_T\}$ of the interval $[0,T]$ is random, τ_i are stopping times. We approximate X_t by

$$\hat{Y}_t=X_0+\int_0^t a(\hat{Y}_{\tau_{i_s}})ds+\int_0^t b(\hat{Y}_{\tau_{i_s}})dW_s+\int_0^t G(\hat{Y}_{\tau_{i_s}})d\tilde{Z}_s,\quad t\in[0,T]. \tag{2.6}$$

The following error estimate holds.

Theorem 3. *Let $\alpha<\beta\leq\mu\leq 2\alpha$, and let $a^i,b^{ij}\in\tilde{C}^{\beta}(\mathbf{R}^d)$, $G^{ij}\in\tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$, and*

$$\int_{|\upsilon|\leq 1}|\upsilon|^{\alpha}d\pi+\int_{|\upsilon|\geq 1}|\upsilon|^{\mu}\pi(d\upsilon)<\infty,$$

where π is the Lévy measure of the driving process Z.

Then there is a constant C (independent of σ,δ) such that for all $g\in\tilde{C}^{\beta}(\mathbf{R}^d)$

$$|\mathbf{E}g(\hat{Y}_T)-\mathbf{E}g(X_T)|\leq C|g|_{\beta}\left[\{(\delta\wedge\lambda_\sigma^{-1})\tilde{\lambda}_\sigma\}^{\frac{\beta}{\alpha}-1}+\int_{|\upsilon|\leq\sigma}|\upsilon|^{\beta\wedge 3}d\pi\right],$$

where $\lambda_\sigma = \pi(\{|\upsilon| > \sigma\})$ *and*

$$\tilde{\lambda}_\sigma = 1 + 1_{\alpha\in(1,2)} \left| \int_{\sigma<|\upsilon|\le 1} \upsilon d\pi \right|.$$

In particular, the following statement holds.

Corollary 4. *Suppose the assumptions of Theorem 3 hold.*

(i) *If* $\delta = T$ *(only jump moments are chosen for the time discretization), then*

$$|\mathbf{E}g(\hat{Y}_T) - \mathbf{E}g(X_T)| \le C|g|_\beta \left[\left(\frac{\tilde{\lambda}_\sigma}{\lambda_\sigma} \right)^{\frac{\beta}{\alpha}-1} + \int_{|\upsilon|\le\sigma} |\upsilon|^{\beta\wedge 3} d\pi \right].$$

(ii) *If* $\sup_{\sigma\in(0,1)} |\int_{\sigma<|\upsilon|\le 1} \upsilon d\pi| < \infty$ *for* $\alpha \in (1,2)$*, then*

$$|\mathbf{E}g(\hat{Y}_T) - \mathbf{E}g(X_T)| \le C|g|_\beta \left[(\delta \wedge \lambda_\sigma^{-1})^{\frac{\beta}{\alpha}-1} + \int_{|\upsilon|\le\sigma} |\upsilon|^{\beta\wedge 3} d\pi \right],$$

where $\lambda_\sigma = \pi(\{|\upsilon| > \sigma\})$.

2.3. *Outline of the proof of Theorem 1*

To prove Theorem 1, as in [10,12], the solution to the backward Kolmogorov equation associated with X_t is used. First we introduce the operator of the Kolmogorov equation associated with X_t.

For $u \in \tilde{C}^\beta(H)$, $\beta > \alpha$, denote

$$\begin{aligned} L_z u(t,x) &= (a(z), \nabla_x u(t,x)) + \frac{1}{2} \sum_{i,j=1}^d (b^i(z), b^j(z)) \partial^2_{ij} u(x) \\ &\quad + \int_{\mathbf{R}^m_0} [u(t, x + G(z)\upsilon) - u(t,x) - \chi_\alpha(\upsilon)(\nabla_x u(t,x), G(z)\upsilon)]\, \pi(d\upsilon), \\ Lu(t,x) &= L_x u(t,x) = L_z u(t,x)|_{z=x}, \end{aligned}$$

where $b^i(z) = (b^{ij}(z))_{1\le j\le m}$, $i = 1,\dots,d$.

Remark 3. Under assumptions of Theorem 1, there exists a unique strong solution to Eq. (1.1) and the stochastic process

$$u(X_t) - \int_0^t Lu(X_s) ds, \quad \forall u \in \tilde{C}^\beta(\mathbf{R}^d)$$

with $\beta > \alpha$ is a martingale. The operator L is the generator of X_t defined in (1.1).

If $v(t,x)$, $(t,x) \in H$ satisfies the backward Kolmogorov equation

$$\begin{aligned} &(\partial_t + L)\, v(t,x) = 0, \quad 0 \le t \le T, \\ &v(T,x) = g(x), \end{aligned}$$

then by Itô's formula

$$\mathbf{E}[g(Y_T)] - \mathbf{E}[g(X_T)] = \mathbf{E}[v(T,Y_T) - v(0,Y_0)] = \mathbf{E}\left[\int_0^T (\partial_t + L_{Y_{\tau_{i_s}}}) v(s, Y_s) ds \right].$$

The regularity of v determines the one-step estimate and the rate of convergence of the approximation.

3. Backward Kolmogorov equation

In Lipschitz spaces $\tilde{C}^{\beta}(H)$, consider the backward Kolmogorov equation associated with X_t:

$$(\partial_t + L)\, u(t,x) = f(t,x), \qquad (3.1)$$
$$u(T,x) = g(x).$$

Definition 1. Let f, g be measurable and bounded functions. We say that $u \in \tilde{C}^{\beta}(H)$ with $\beta > \alpha$ is a solution to (3.1) if

$$u(t,x) = g(x) + \int_t^T [Lu(s,x) - f(s,x)]\, ds, \quad \forall (t,x) \in H. \qquad (3.2)$$

First we show that $L : \tilde{C}^{\beta}(H) \to \tilde{C}^{\beta-\alpha}(H)$ is continuous.

Lemma 1. *Let $\alpha < \beta \le \mu \le 2\alpha$,*

$$\int_{|\upsilon|\le 1} |\upsilon|^{\alpha} d\pi + \int_{|\upsilon|>1} |\upsilon|^{\mu} d\pi < \infty$$

and $a^i, b^{ij} \in \tilde{C}^{\beta}(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$. Then for any $v \in \tilde{C}^{\beta}(\mathbf{R}^d)$ we have $Lv \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$ and there is a constant independent of v such that

$$|Lv|_{\beta-\alpha} \le C|v|_{\beta}.$$

Proof. Let

$$Bv(x) = \int [v(x + G(x)\upsilon) - v(x) - \chi_{\alpha}(\upsilon)(\nabla v(x), G(x)\upsilon)]\, d\pi.$$

Then

$$Lv = Bv + (a(x), \nabla v(x)) + \frac{1}{2}(b^i(x), b^j(x))\partial^2_{ij} v(x).$$

By Proposition 13 in [12], $Bv \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$ if $\beta - \alpha \notin \mathbf{N}$ and $|Bv|_{\beta-\alpha} \le C|v|_{\beta}$. In this case, obviously, $Lv \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$ as well.

If $\alpha > 1$, $\beta = 1 + \alpha$, then

$$Bv(x) = \int_{|\upsilon|\le 1}\int_0^1 [\nabla v(x + sG(x)\upsilon) - \nabla v(x)]G(x)\upsilon\, ds\, d\pi + \int_{|\upsilon|>1} [v(x + G(x)\upsilon) - v(x)]\, d\pi.$$

Since

$$\nabla(Bv(x)) = \int_{|\upsilon|\le 1}\int_0^1 [\partial^2 v(x + sG(x)\upsilon) - \partial^2 v(x)]G(x)\upsilon\, ds\, d\pi$$

$$+ \int_{|\upsilon|\leq 1}\int_0^1 \partial^2 v(x+sG(x)\upsilon)\nabla G(x)\upsilon G(x)\upsilon ds d\pi$$
$$+ \int_{|\upsilon|>1} [\nabla v(x+G(x)\upsilon)-\nabla v(x)]\, d\pi$$
$$+ \int_{|\upsilon|>1} \nabla v(x+G(x)\upsilon)\nabla G(x)\upsilon d\pi,$$

it follows that $\sup_x |\nabla(Bv(x))| \leq C|v|_\beta$. Therefore $|Lv|_{\beta-\alpha} \leq C|v|_\beta$ as well. If $\alpha = 1$ and $\beta = 2$, then

$$|\nabla Bv(x)| = \int [\nabla v(x+G(x)\upsilon)-\nabla v(x)]d\pi + \int \nabla v(x+G(x)\upsilon)G(x)\upsilon d\pi,$$
$$\sup_x |\nabla Bv(x)| \leq C|v|_\beta$$

and $|Lv|_{\beta-\alpha} \leq C|v|_\beta$. The case $\beta = 4, \alpha = 2$ is considered in a similar way. □

The main result of this section is the following statement.

Theorem 4. *Let $\alpha < \beta \leq \mu \leq 2\alpha$, and*

$$\int_{|\upsilon|\leq 1} |\upsilon|^\alpha \pi(d\upsilon) + \int_{|\upsilon|>1} |\upsilon|^\mu \pi(d\upsilon) < \infty.$$

Assume $a^i, b^{ij} \in \tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$. Then for each $f \in \tilde{C}^\beta(\mathbf{R}^d)$, $g \in \tilde{C}^\beta(\mathbf{R}^d)$, there exists a unique solution $u \in \tilde{C}^\beta(H)$ to (3.1) and a constant C independent of f, g such that $|u|_\beta \leq C(|f|_\beta + |g|_\beta)$.

To prove Theorem 4, for $\varepsilon \in (0, 1)$ we consider a nondegenerate equation

$$\begin{aligned}&\left(\partial_t + L^\varepsilon\right) u(t,x) = f(t,x),\\ &u(T,x) = g_\varepsilon(x),\end{aligned} \tag{3.3}$$

where $L^\varepsilon u = -\varepsilon^\alpha (-\Delta)^{\alpha/2} u + Lu$ and

$$g_\varepsilon(x) = \int g(y)w^\varepsilon(x-y)dy = \int g(x-y)w^\varepsilon(y)dy, \quad x \in \mathbf{R}^d$$

with $w^\varepsilon(x) = \varepsilon^{-d} w\left(\frac{x}{\varepsilon}\right)$, $x \in \mathbf{R}^d$, $w \in C_0^\infty(\mathbf{R}^d)$, $\int w dx = 1$.

An obvious consequence of Corollary 9 in [12] is the following statement.

Lemma 2 (*See Corollary 9 in [12]*)**.** *Let $\alpha < \beta \leq \mu \leq 2\alpha$,*

$$\int_{|\upsilon|\leq 1} |\upsilon|^\alpha \pi(d\upsilon) + \int_{|\upsilon|>1} |\upsilon|^\mu \pi(d\upsilon) < \infty,$$

and $a^i, b^{ij}, g, f \in \tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$. Then for each $\varepsilon \in (0,1)$ there is $\bar{\beta} > 2\alpha$ and a unique $u = u_\varepsilon \in \tilde{C}^{\bar{\beta}}(H)$ solving (3.3).

We separate in the operator L^ε its "bounded jump" part $\bar{L}^\varepsilon v(x) = \bar{L}_z^\varepsilon v(x)|_{z=x}$ with

$$\bar{L}_z^\varepsilon v(x) = -\varepsilon^\alpha (-\Delta)^{\alpha/2} u + (a(z), \nabla_x v(x)) + \frac{1}{2}\sum_{i,j=1}^d (b^i(z), b^j(z))\partial_{ij}^2 v(x)$$

$$+\int_{|\upsilon|\leq 1}[v(x+G\,(z)\,\upsilon)-v(x)-\chi_{\alpha}(\upsilon)(\nabla v(x),G(z)\upsilon)]\,d\pi,$$

$z,x\in R^{d}$, $v\in C_{0}^{\infty}(\mathbf{R}^{d})$, so that

$$L_{z}^{\varepsilon}v(x)=\bar{L}_{z}^{\varepsilon}v(x)+\int_{|\upsilon|>1}[v(x+G(z)y)-v(x)]d\pi,\quad x,z\in\mathbf{R}^{d}.$$

Remark 4. If the assumptions of Lemma 2 hold and $u_{\varepsilon}\in\tilde{C}^{\bar{\beta}}(H)$ solves (3.3) with $\bar{\beta}>2\alpha$, then u_{ε} satisfies the following equation as well:

$$\begin{aligned}&\left(\partial_{t}+\bar{L}^{\varepsilon}\right)u(t,x)=F(u,t,x),\\&u(T,x)=g_{\varepsilon}(x),\end{aligned}\tag{3.4}$$

where $F(u,t,x)=F_{z}(u,t,x)|_{z=x}$ with

$$F_{z}(u,t,x)=f(t,x)-\int_{|\upsilon|>1}[u(t,x+G(z)\upsilon)-u(t,x)]\,d\pi.$$

Using a probabilistic form of a maximum principle we will derive uniform (independent of ε) $\tilde{C}^{\beta}$-norm estimates of u_{ε} and passing to the limit as $\varepsilon\to 0$ we will obtain $u\in\tilde{C}^{\beta}(H)$ solving (3.1). First we prove some auxiliary statements.

Let

$$\begin{aligned}\tilde{Z}_{t}&=\int_{0}^{t}\int_{|\upsilon|\leq 1}[(1-\chi_{\alpha}(\upsilon))\upsilon p(dt,d\upsilon)+\chi_{\alpha}(\upsilon)\upsilon q(dt,d\upsilon)]\\&=\int_{0}^{t}\int_{|\upsilon|\leq 1}\upsilon q(dt,d\upsilon)+t\int_{|\upsilon|\leq 1}(1-\chi_{\alpha}(\upsilon))\upsilon d\pi.\end{aligned}\tag{3.5}$$

For $(s,x)\in H$, $h\in\mathbf{R}^{d}$, $\xi\in\mathbf{R}^{d}$, the following stochastic processes in $[s,T]$ are used to derive the uniform estimates:

$$\begin{aligned}&dU_{t}=\varepsilon dZ_{t}^{\alpha}+a(U_{t})dt+b(U_{t})dW_{t}+G(U_{t-})d\tilde{Z}_{t},\\&dH_{t}=[a(U_{t}+H_{t})-a(U_{t})]dt+[b(U_{t}+H_{t})-b(U_{t})]dW_{t}\\&\qquad+\left[G(U_{t-}+H_{t-})-G(U_{t-})\right]d\tilde{Z}_{t},\\&d\bar{V}_{t}=a^{(1)}(U_{t}+H_{t};\bar{V}_{t})dt+b^{(1)}(U_{t}+H_{t};\bar{V}_{t})dW_{t}\\&\qquad+\int_{|\upsilon|\leq 1}G^{(1)}(U_{t-}+H_{t-};\bar{V}_{t-})d\tilde{Z}_{t},\\&dV_{t}=a^{(1)}(U_{t};V_{t})dt+b^{(1)}(U_{t};V_{t})dW_{t}+G^{(1)}(U_{t-};V_{t-})d\tilde{Z}_{t},\\&U_{s}=x,\qquad H_{s}=h,\qquad V_{s}=\xi,\qquad\bar{V}_{s}=\xi,\end{aligned}\tag{3.6}$$

where Z^{α} is $\mathbf{R}^{d}$-valued spherically symmetric α-stable process corresponding to $(-\Delta)^{\alpha/2}$ and independent of Z. Recall for a function v on $\mathbf{R}^{d}$ we denote $v^{(1)}(x;\xi)=(\nabla v(x),\xi)$, $x,\xi\in\mathbf{R}^{d}$ and, for example, componentwise,

$$dV_{t}^{j}=(\nabla a^{j}(U_{t}),V_{t})dt+\sum_{i=1}^{n}(\nabla b^{ji}(U_{t}),V_{t})dW_{t}^{i}+\sum_{i=1}^{m}(\nabla G^{ji}(U_{t-}),V_{t-})d\tilde{Z}_{t}^{i},$$

$j=1,\ldots,d$.

Lemma 3. (a) *If $a^i, b^i_j, G^{ij} \in \tilde{C}^1(\mathbf{R}^d)$, then for each $l \geq 2$ there is a constant C such that*

$$\mathbf{E}\left[\sup_{s\leq t\leq T}|H_t|^l\right] \leq C|h|^l.$$

(b) *If $a^i, b^i_j, G^{ij} \in \tilde{C}^{1+\kappa}(\mathbf{R}^d)$ with $\kappa \in (0,1]$, then for each $l \geq 2$ there is a constant C such that*

$$\mathbf{E}\left[\sup_{s\leq t\leq T}|V_s|^l + \sup_{s\leq t\leq T}|\bar{V}_s|^l\right] \leq C|\xi|^l,$$

$$\mathbf{E}\left[\sup_{s\leq t\leq T}|V_t - \bar{V}_t|^l\right] \leq C|\xi|^l|h|^{l\kappa}.$$

Proof. (a) Since (3.5) holds, we have by the Hölder inequality and martingale moment estimates (see [11,17])

$$\begin{aligned}\mathbf{E}\sup_{s\leq r\leq t}|H_r|^l &\leq C\left\{|h|^l + \mathbf{E}\left[\left(\int_s^t|H_r|^2dr\right)^{l/2}\right] + \mathbf{E}\int_s^t|H_r|^l dr\right\}\\ &\leq C\left[|h|^l + \mathbf{E}\int_s^t\sup_{s\leq r'\leq r}|H_{r'}|^l dr\right], \quad s\leq t\leq T\end{aligned}$$

and inequality follows by the Gronwall lemma.

(b) Similarly, for each $l \geq 2$, there is a constant C so that

$$\mathbf{E}\left[\sup_{s\leq t\leq T}|V_s|^l + \sup_{s\leq t\leq T}|\bar{V}_s|^l\right] \leq C|\xi|^l.$$

Then

$$\begin{aligned}\mathbf{E}\sup_{s\leq r\leq t}|V_r-\bar{V}_r|^l &\leq C\left\{\mathbf{E}\left[\left(\int_s^t|H_r|^{2\kappa}|\bar{V}_r|^2dr\right)^{l/2}\right] + \mathbf{E}\int_s^t|H_r|^{\kappa l}|\bar{V}_r|^l dr\right.\\ &\quad\left. + \mathbf{E}\left[\left(\int_s^t|\bar{V}_r - V_r|^2dr\right)^{l/2}\right] + \mathbf{E}\int_s^t|\bar{V}_r - V_r|^l dr\right\}\\ &\leq C\left[\mathbf{E}\int_s^t|H_r|^{\kappa l}|\bar{V}_r|^l dr + \mathbf{E}\int_s^t|\bar{V}_r - V_r|^l dr\right], \quad s\leq t\leq T.\end{aligned}$$

By the Gronwall lemma,

$$\begin{aligned}\mathbf{E}\sup_{s\leq r\leq T}|V_r-\bar{V}_r|^l &\leq C\mathbf{E}\int_s^T|H_r|^{\kappa l}|\bar{V}_r|^l dr\\ &\leq C\int_s^T[\mathbf{E}(|H_r|^{2\kappa l})]^{1/2}[\mathbf{E}(|\bar{V}_r|^{2l})^{1/2}]dr\\ &\leq C|\xi|^l|h|^{\kappa l}. \quad \square\end{aligned}$$

3.1. Proof of Theorem 4

1. *Existence.* By Lemma 2, for each $\varepsilon \in (0,1)$ there is a unique solution $u_\varepsilon \in \tilde{C}^{\bar\beta}(H)$ to (3.3) for some $\bar\beta > 2\alpha$. By Remark 4, (3.4) holds as well. Let $(s,x) \in H$ and U_t solves (3.6). By Itô's formula,

$$\mathbf{E}g_\varepsilon(U_T) - u_\varepsilon(s,x) = \mathbf{E}\int_s^T F(u_\varepsilon, r, U_r)dr$$

and

$$|u_\varepsilon(s,\cdot)|_0 \le |g|_0 + \int_s^T |f(r,\cdot)|_0 + C|u_\varepsilon(r,\cdot)|_0 dr.$$

By the Gronwall lemma, there is a constant not depending on u_ε and ε such that

$$\sup_{0\le t\le T} |u_\varepsilon(t,\cdot)|_0 \le C\left[|g|_0 + \int_0^T |f(r,\cdot)|_0 dr\right].$$

As suggested in [7], we estimate multilinear forms associated to the derivatives of u. Let $k = [\beta]^-$, $(t,x) \in H$, $\xi^1,\ldots,\xi^k \in \mathbf{R}^d$ and

$$u_\varepsilon^{(k)}(t,x;\xi^1,\ldots,\xi^k) = \sum_{i_1,\ldots,i_k=1}^d \frac{\partial^k u(t,x)}{\partial x_{i_k}\cdots x_{i_1}}\xi^1_{i_1}\cdots\xi^k_{i_k} \quad \text{if } k\ge 1,$$

$$u_\varepsilon^{(0)}(t,x) = u_\varepsilon(t,x).$$

For $z \in \mathbf{R}^d$, $(t,x) \in H$, $\xi^1 \in \mathbf{R}^d,\ldots,\xi^k \in \mathbf{R}^d$, let

$$\begin{aligned}
&\mathcal{P}_z u_\varepsilon^{(k)}(t,x;\xi^1,\ldots,\xi^k)\\
&\quad = -\varepsilon^\alpha(-\Delta_x)^{\alpha/2}u_\varepsilon^{(k)}(t,x,\xi^1,\ldots,\xi^k)\\
&\qquad + \int_{|\upsilon|\le 1}\Big\{u_\varepsilon^{(k)}(x+G(z)\upsilon;\xi^1+G^{(1)}(z;\xi^1)\upsilon,\ldots,\xi^k\\
&\qquad + G_\varepsilon^{(1)}(z;\xi^k)\upsilon) - u_\varepsilon^{(k)}(x;\xi^1,\ldots,\xi^k) - \chi_\alpha(\upsilon)\Big[(\nabla_x u_\varepsilon^{(k)}(x;\xi_1,\ldots,\xi^k), G(z)\upsilon)\\
&\qquad - \sum_{l=1}^k(\nabla_{\xi^l}u_\varepsilon^{(k)}(x;\xi_1,\ldots,\xi^k), G^{(1)}(z;\xi^l)\upsilon)\Big]\Big\}d\pi\\
&\qquad + (a(z),\nabla_x u_\varepsilon^{(k)}(x;\xi^1,\ldots,\xi^k)) + \sum_{l=1}^k(\nabla_{\xi^l}u^{(k)}(x;\xi^1,\ldots,\xi^k), a^{(1)}(z;\xi^l))\\
&\qquad + \frac{1}{2}\sum_{i,j}\Big\{(b^i(z),b^j(z))\partial^2_{ij}u_\varepsilon^{(k)}(x;\xi^1,\ldots,\xi^k)\\
&\qquad + \sum_{l=1}^k[(b^{i,(1)}(z;\xi^l),b^j(z))\partial_{\xi^l_i x_j}u_\varepsilon^{(k)}(x;,\xi^1,\ldots,\xi^k)\\
&\qquad + (b^i(z),b^{j,(1)}(z;\xi^l))\partial_{x_i\xi^l_j}u^{(k)}(x;,\xi^1,\ldots,\xi^k)]\Big\}.
\end{aligned}$$

Differentiating both sides of (3.4) and multiplying by $\xi^1_{i_1}\cdots\xi^k_{i_k}$ we see that $u^{(k)}_\varepsilon(t,x;\xi^1,\dots,\xi^k)$ satisfies the equation

$$\partial_t u^{(k)}_\varepsilon(t,x,\xi^1,\dots,\xi^k)+\mathcal{P}^\varepsilon u^{(k)}_\varepsilon(t,x,\xi^1,\dots,\xi^k)=A(u_\varepsilon,t,x,\xi^1,\dots,\xi^k), \tag{3.7}$$

where

$$A(u_\varepsilon,t,x,\xi^1,\dots,\xi^k)=B(u_\varepsilon,t,x,\xi^1,\dots,\xi^k)+F^{(k)}(u_\varepsilon,t,x;\xi^1,\dots\xi^k)$$

and $B(u_\varepsilon,t,x,\xi^1,\dots,\xi^k)$ is a finite sum of the terms of the form

$$\begin{aligned}&[\nabla_x u^{(l)}_\varepsilon(t,x+G(x)\upsilon;\xi^{i_1},\dots,\xi^{i_l})-\nabla_x u^{(l)}(t,x+G(x)\upsilon;\xi^{i_1},\dots,\xi^{i_l})]\\&\quad\times G^{(k-l)}(x;\xi^{i_{l+1}},\dots,\xi^{i_k})\upsilon\\&\quad=\int_0^1\partial^2 u^{(l)}_x(t,x+sG(x)\upsilon;\xi^{i_1},\dots,\xi^{i_l})G(x)\upsilon ds G^{(k-l)}(x;\xi^{i_{l+1}},\dots,\xi^{i_k})\upsilon\end{aligned}$$

with $l\le k-2$ and

$$u^{(l)}(t,x+G(x)\upsilon;\xi^{i_1},\dots,\xi^{i_l})G^{(l_1)}(x;\xi^{i^1_1},\dots,\xi^{i^1_{k_1}})\cdots G^{(l_m)}(x;\xi^{i^m_1},\dots,\xi^{j^m_{k_m}})$$

with $m\ge 2, l\le k, l+l_1+\cdots+l_m=k$ and $(\xi^{i_1},\dots,\xi^{i_l},\dots,\xi^{i^m_{k_m}})$ being a permutation of $\xi^1,\dots,\xi^k$. In any case, there is a constant C independent of ε and u_ε so that for all $(t,x)\in[0,T]\times\mathbf{R}^d$, $\xi^i\in\mathbf{R}^d$,

$$\begin{aligned}&|A(u_\varepsilon,t,x,\xi^1,\dots,\xi^k)|\le C(|u_\varepsilon(t,\cdot)|_k+|f(t,\cdot)|_k)|\xi^1|\cdots|\xi^k|,\\&|A(u_\varepsilon,t,\cdot,\xi^1,\dots,\xi^k)|_{\beta-k}\le C(|u_\varepsilon(t,\cdot)|_\beta+|f(t,\cdot)|_\beta)|\xi^1|\cdots|\xi^k|,\end{aligned} \tag{3.8}$$

and

$$\begin{aligned}&|A(u_\varepsilon,t,x,\bar\xi^1,\dots,\bar\xi^k)-A(u_\varepsilon,t,x,\xi^1,\dots,\xi^k)|\\&\quad\le C(|f(t,\cdot)|_k+|u_\varepsilon(t,\cdot)|_k)\sum_{l=1}^k|\xi^1|\cdots|\xi^{l-1}\|\bar\xi^l-\xi^l\|\bar\xi^{l+1}|\cdots|\bar\xi^k|.\end{aligned} \tag{3.9}$$

On the other hand, for any $(s,x)\in H$ with the processes defined in (3.6), it follows by Itô's formula,

$$\begin{aligned}&\mathbf{E}[u^{(k)}_\varepsilon(T,U_T,V^1_T,\dots,V^k_T)-u^{(k)}_\varepsilon(s,x,\xi^1,\dots,\xi^k)]\\&\quad=\mathbf{E}[g^{(k)}_\varepsilon(U_T,V^1_T,\dots,V^k_T)-u^{(k)}_\varepsilon(s,x,\xi^1,\dots,\xi^k)]\\&\quad=\mathbf{E}\int_s^T[\partial_t u^{(k)}_\varepsilon(t,U_t,V^1_t,\dots,V^k_t)+\mathcal{P}^\varepsilon_{U_t}u^{(k)}_\varepsilon(t,U_t,V^1_t,\dots,V^k_t)]dt\\&\quad=\mathbf{E}\int_s^T[A(u_\varepsilon,t,U_t,V^1_t,\dots,V^k_t)]dt\end{aligned}$$

and

$$\begin{aligned}&\mathbf{E}[u^{(k)}_\varepsilon(T,U_T+H_T,\bar V^1_T,\dots,\bar V^k_T)-u^{(k)}_\varepsilon(T,U_T,V^1_T,\dots,V^k_T)]\\&\qquad-[u^{(k)}_\varepsilon(s,x+h,\xi^1,\dots,\xi^k)-u^{(k)}_\varepsilon(s,x,\xi^1,\dots,\xi^k)]\\&\quad=\mathbf{E}[g^{(k)}(U_T+H_T,\bar V^1_T,\dots,\bar V^k_T)-g^{(k)}(U_T,V^1_T,\dots,V^k_T)]\\&\qquad-[u^{(k)}_\varepsilon(s,x+h,\xi^1,\dots,\xi^k)-u^{(k)}_\varepsilon(s,x,\xi^1,\dots,\xi^k)]\end{aligned}$$

$$= \mathbf{E}\int_s^T \Big\{[\partial_t u_\varepsilon^{(k)}(t, U_t + H_t, \bar V_t^1, \dots, \bar V_t^k) + \mathcal{P}^\varepsilon_{U_t+H_t} u_\varepsilon^{(k)}(t, U_t + H_t, \bar V_t^1, \dots, \bar V_t^k)]$$
$$- [\partial_t u_\varepsilon^{(k)}(t, U_t, V_t^1, \dots, V_t^k) + \mathcal{P}^\varepsilon_{U_t} u_\varepsilon^{(k)}(t, U_t, V_t^1, \dots, V_t^k)]\Big\} dt$$
$$= \mathbf{E}\int_s^T [A(u_\varepsilon, t, U_t + H_t, \bar V_t^1, \dots, \bar V_t^k) - A(u_\varepsilon, t, U_t, V_t^1, \dots, V_t^k)]dt.$$

Since by (3.8)

$$|A(u_\varepsilon, t, U_t, V_t^1, \dots, V_t^k)| \le C(|u_\varepsilon(t,\cdot)|_k + |f(t,\cdot)|_k)|V_t^1|\cdots|V_t^k|,$$

it follows by Lemma 3 and the Hölder inequality,

$$\mathbf{E}|A(u_\varepsilon, t, U_t, V_t^1, \dots, V_t^k)| \le C(|u_\varepsilon(t,\cdot)|_k + |f(t,\cdot)|_k)|\xi^1|\cdots|\xi^k|. \tag{3.10}$$

Since

$$|A(u_\varepsilon, t, U_t + H_t, \bar V_t^1, \dots, \bar V_t^k) - A(u_\varepsilon, t, U_t, V_t^1, \dots, V_t^k)|$$
$$\le |A(u_\varepsilon, t, U_t + H_t, \bar V_t^1, \dots, \bar V_t^k) - A(u_\varepsilon, t, U_t, \bar V_t^1, \dots, \bar V_t^k)|$$
$$+ |A(u_\varepsilon, t, U_t, \bar V_t^1, \dots, \bar V_t^k) - A(u_\varepsilon, t, U_t, V_t^1, \dots, V_t^k)|$$
$$= A_1 + A_2,$$

it follows by the estimates (3.8), (3.9) and Lemma 3 that

$$\mathbf{E}A_1 \le C(\mathbf{E}|H_t|^{2(\beta-k)})^{1/2}\left(|f(t,\cdot)|_\beta + |u(t,\cdot)|_\beta\right)$$
$$\le C|h|^{\beta-k}\left(|f(t,\cdot)|_\beta + |u(t,\cdot)|_\beta\right)$$

and for $|h| \le 1$

$$\mathbf{E}A_2 \le C(|f(t,\cdot)|_k + |u(t,\cdot)|_k)\sum_{l=1}^k \mathbf{E}|V_t^1|\cdots|V_t^{l-1}||\bar V_t^l - V_t^l||\bar V_t^{l+1}|\cdots|\bar V_t^k|$$
$$\le C(|f(t,\cdot)|_k + |u(t,\cdot)|_k)\sum_l (\mathbf{E}[|\bar V_t^l - V_t^l|^2])^{1/2}|\xi^1|\cdots|\xi^{l-1}||\xi^{l+1}|\cdots|\xi^k|$$
$$\le C(|f(t,\cdot)|_k + |u(t,\cdot)|_k)|\xi^1|\cdots|\xi^k||h|^{\beta-k}.$$

Similarly, we estimate

$$\mathbf{E}|g_\varepsilon^{(k)}(U_T, V_T^1, \dots, V_T^k)| \le C|g|_k|\xi^1|\cdots|\xi^k|$$

and for $|h| \le 1$

$$\mathbf{E}|g_\varepsilon^{(k)}(U_T + H_T, \bar V_T^1, \dots, \bar V_T^k) - g_\varepsilon^{(k)}(U_T, V_T^1, \dots, V_T^k)| \le C|g|_\beta|h|^{\beta-k}|\xi^1|\cdots|\xi^k|.$$

So,

$$|u_\varepsilon^{(k)}(s, x; \xi^1, \dots, \xi^k)|_0 \le C|\xi^1|\cdots|\xi^k|\left[|g|_k + \int_s^T (|u_\varepsilon(t,\cdot)|_k + |f(t,\cdot)|_k)dt\right],$$
$$0 \le s \le T,$$

and by the Gronwall lemma,

$$\sup_{0\le s\le T}|u_\varepsilon^{(k)}(s, x; \xi^1, \dots, \xi^k)|_0 \le C|\xi^1|\cdots|\xi^k|\left[|g|_k + \int_0^T |f(t,\cdot)|_k dt\right].$$

Also, for $|h| \leq 1, x \in \mathbf{R}^d, 0 \leq s \leq T$,

$$
\begin{aligned}
&|u_\varepsilon^{(k)}(s, x+h, \xi^1, \ldots, \xi^k) - u_\varepsilon^{(k)}(s, x, \xi^1, \ldots, \xi^k)| \\
&\quad \leq C|h|^{\beta-k}|\xi^1|\cdots|\xi^k|\left[|g|_\beta + \int_s^T (|f(t,\cdot)|_\beta + |u(t,\cdot)|_\beta)dt\right],
\end{aligned}
$$

and by the Gronwall lemma,

$$
\sup_{0\leq s\leq T} |u^{(k)}(s, \cdot, \xi^1, \ldots, \xi^k)|_{\beta-k} \leq C|\xi^1|\cdots|\xi^k|\left[|g|_\beta + \int_0^T |f(t,\cdot)|_\beta dt\right].
$$

Therefore for each $\beta \in (\alpha, 2\alpha]$,

$$
\sup_{\varepsilon\in(0,1)} |u_\varepsilon|_\beta \leq C\left[|g|_\beta + \int_0^T |f(t,\cdot)|_\beta dt\right]. \tag{3.11}
$$

Since for each $(s, x) \in H$,

$$
u_\varepsilon(s, x) = g_\varepsilon(x) + \int_s^T [L^\varepsilon u_\varepsilon(t, x) - f(t, x)]dt, \tag{3.12}
$$

and there is a constant $C > 0$ so that for all $(t, x) \in H, h \in \mathbf{R}^d$,

$$
\begin{aligned}
|\partial_t u_\varepsilon(t, x+h) - \partial_t u_\varepsilon(t, x)| &\leq |L_{x+h}^\varepsilon u(t, x+h) - L^\varepsilon u_\varepsilon u(t, x)| \\
&\quad + |f(t, x+h) - f(t, x)| \\
&\leq C|h|^{\tilde{\beta}-\alpha}(|u_\varepsilon|_{\tilde{\beta}} + |f|_\beta)
\end{aligned} \tag{3.13}
$$

for some $\tilde{\beta} \in (\alpha, \alpha + \alpha \wedge 1)$. It follows from (3.11) and (3.13) that there is a sequence $\varepsilon_n \to 0$ and $u \in \tilde{C}^\beta(H)$ such that such $u_{\varepsilon_n} \to u$ uniformly on compact sets of H. By (3.11), $L^\varepsilon u_\varepsilon(t, x) \to Lu(t, x)$ pointwise and passing to the limit in (3.12), we see that $u \in \tilde{C}^\beta(H)$ is a solution to (3.1).

2. *Uniqueness*. Let $u^1, u^2 \in \tilde{C}^\beta(H)$ be two solutions to (3.1). Then $v = u^1 - u^2$ satisfies (3.1) with $g = 0, f = 0$. Let $X_t^{s,x}$ be the solution to (1.1) starting from $x \in \mathbf{R}^d$ at time moment s. Then by Itô's formula,

$$
\begin{aligned}
-v(s, x) &= \mathbf{E}v(T, X_T^{s,x}) - v(s, x) \\
&= \mathbf{E}\int_s^T \left[\partial_t v(r, X_r^{s,x}) + Lv(r, X_r^{s,x})\right]dr = 0
\end{aligned}
$$

and uniqueness follows.

4. One-step estimate and proof of main results

First, we modify the mollified function estimates for the Lipschitz spaces. Let $w \in C_0^\infty(\mathbf{R}^d)$, be a nonnegative smooth function with support in $\{|x| \leq 1\}$ such that $w(x) = w(|x|), x \in \mathbf{R}^d$, and $\int w(x)dx = 1$. Due to the symmetry,

$$
\int_{\mathbf{R}^d} x^i w(x)dx = 0, \quad i = 1, \ldots, d. \tag{4.1}
$$

For $x \in \mathbf{R}^d$ and $\varepsilon \in (0,1)$, define $w^\varepsilon(x) = \varepsilon^{-d} w\left(\frac{x}{\varepsilon}\right)$ and the convolution

$$f^\varepsilon(x) = \int f(y) w^\varepsilon(x-y)dy = \int f(x-y)w^\varepsilon(y)dy, \quad x \in \mathbf{R}^d. \tag{4.2}$$

Lemma 4. *Let $\alpha < \beta \le 2\alpha$, $f \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$. Then*

$$|f^\varepsilon(x) - f(x)| \le C\varepsilon^{\beta-\alpha}|f|_{\beta-\alpha}, \quad x \in \mathbf{R}^d, \tag{4.3}$$

and there is a constant C such that

$$|Lf^\varepsilon| \le C\varepsilon^{\beta-2\alpha}|f|_{\beta-\alpha}. \tag{4.4}$$

Proof. Indeed, if $\beta - \alpha \le 1$, then

$$\begin{aligned} |f^\varepsilon(x) - f(x)| &\le \int |f(x-y) - f(x)| w^\varepsilon(y)dy \\ &\le C|f|_{\beta-\alpha}\varepsilon^{\beta-\alpha}. \end{aligned}$$

If $\beta - \alpha \in (1,2]$, then

$$\begin{aligned} |f^\varepsilon(x) - f(x)| &= \left| \int f(x+y) - f(x) - (\nabla f(x), y) w^\varepsilon(y)dy \right| \\ &\le \int \int_0^1 |(\nabla f(x+sy) - \nabla f(x), y)| ds w^\varepsilon(y)dydy \\ &\le C\varepsilon^{\beta-\alpha}|f|_{\beta-\alpha}. \end{aligned}$$

According to Lemma 17(iii) and Corollary 18 in [12], for each β, so that $\beta - \alpha < \alpha$,

$$|Lf^\varepsilon| \le C\varepsilon^{(\beta-\alpha)-\alpha}|f|_{\beta-\alpha} = C\varepsilon^{\beta-2\alpha}|f|_{\beta-\alpha}.$$

Inequality (4.4) still holds for $\beta - \alpha = \alpha$ or $\beta = 2\alpha$ by a straightforward estimate. □

We modify one-step estimate in [12] for Lipschitz spaces as well.

Lemma 5. *Let $\alpha < \beta \le \mu \le 2\alpha$,*

$$\int_{|\upsilon|\le 1} |\upsilon|^\alpha d\pi + \int_{|\upsilon|>1} |\upsilon|^\mu d\pi < \infty,$$

and $a^i, b^{ij} \in \tilde{C}^\beta(\mathbf{R}^d)$, $G^{ij} \in \tilde{C}^{\beta\vee 1}(\mathbf{R}^d)$. Then there exists a constant C such that for all $f \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$,

$$\left|\mathbf{E}\left[f(Y_s) - f(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}\right]\right| \le C|f|_{\beta-\alpha}\delta^{\frac{\beta}{\alpha}-1}, \quad \forall s \in [0,T],$$

where $i_s = i$ if $\tau_i \le s < \tau_{i+1}$.

Proof. Applying Itô's formula, for $s \in [0,T]$,

$$\mathbf{E}[f^\varepsilon(Y_s) - f^\varepsilon(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}] = \mathbf{E}\left[\int_{\tau_{i_s}}^s \left(L_{Y_{\tau_{i_s}}} f^\varepsilon(Y_r)\right) dr \middle| \mathcal{F}_{\tau_{i_s}}\right].$$

Hence, for $\varepsilon \in (0,1)$, by (4.3) and (4.4),

$$\begin{aligned}|\mathbf{E}[f(Y_s)-f(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}]| &\le |\mathbf{E}[(f-f^{\varepsilon})(Y_s)-(f-f^{\varepsilon})(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}]| \\ &\quad +|\mathbf{E}[f^{\varepsilon}(Y_s)-f^{\varepsilon}(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}]| \\ &\le CF(\varepsilon,\delta)|f|_{\beta-\alpha},\end{aligned}$$

with a constant C independent of ε, f and $F(\varepsilon,\delta)=\varepsilon^{\beta-\alpha}+\varepsilon^{\beta-2\alpha}\delta$. Minimizing $F(\varepsilon,\delta)$ in $\varepsilon \in (0,1)$, we obtain

$$|\mathbf{E}[f(Y_s)-f(Y_{\tau_{i_s}})|\mathcal{F}_{\tau_{i_s}}]| \le C\delta^{\frac{\beta}{\alpha}-1}|f|_{\beta}. \quad \square$$

4.1. *Proof of Theorem 1*

Let $u \in \tilde{C}^{\beta}(H)$ be the unique solution to (3.1) with $f=0$. By Itô's formula,

$$\begin{aligned}\mathbf{E}[u(0,X_0)] &= \mathbf{E}[u(T,X_T)]-\mathbf{E}\left[\int_0^T \big(\partial_t u(s,X_s)+L_{X_s}u(s,X_s)\big)\,ds\right] \\ &= \mathbf{E}\left[g(X_T)\right]\end{aligned}$$

and

$$\mathbf{E}[u(0,X_0)] = \mathbf{E}[u(0,Y_0)]. \tag{4.5}$$

By Lemma 1,

$$|L_z u(s,\cdot)|_{\beta-\alpha} \le C|g|_{\beta}, \qquad |\partial_t u(s,\cdot)|_{\beta-\alpha} \le C|g|_{\beta}, \quad s\in[0,T], z\in\mathbf{R}^d. \tag{4.6}$$

Then, by Itô's formula and (4.6), it follows that

$$\begin{aligned}\mathbf{E}[g(Y_T)]-\mathbf{E}[g(X_T)] &= \mathbf{E}[u(T,Y_T)]-\mathbf{E}[u(0,Y_0)] \\ &= \mathbf{E}\left[\int_0^T \left\{\left[\partial_t u(s,Y_s)-\partial_t u(s,Y_{\tau_{i_s}})\right]\right.\right. \\ &\quad \left.\left. + \left[L_{Y_{\tau_{i_s}}}u(s,Y_s)-L_{Y_{\tau_{i_s}}}u(s,Y_{\tau_{i_s}})\right]\right\} ds\right].\end{aligned}$$

Hence, by (4.6) and Lemma 5, there exists a constant C independent of g such that

$$|\mathbf{E}g(Y_T)-\mathbf{E}g(X_T)| \le C\delta^{\frac{\beta}{\alpha}-1}|g|_{\beta}.$$

The statement of Theorem 1 follows.

4.1.1. *Proof of Corollary 2*

According to [2], there is a rapidly decreasing smooth function $w \in \mathcal{S}(\mathbf{R}^d)$, the Schwartz space, such that $\int w(x)dx = 1$ and all moments are zero:

$$\int w(x)x^{\gamma}dx = 0, \quad \gamma\in\mathbf{N}^d, \gamma \ne \mathbf{0},$$

where $x^{\gamma}=x_1^{\gamma_1}\cdots x_d^{\gamma_d}$, $x=(x_1,\ldots,x_d)\in\mathbf{R}^d$. Let $\varepsilon\in(0,1)$, $w_{\varepsilon}(x)=\varepsilon^{-d}w(x/\varepsilon)$, $x\in\mathbf{R}^d$,

$$g_{\varepsilon}(x)=\int g(x-y)w_{\varepsilon}(y)dy, \quad x\in\mathbf{R}^d.$$

We will show that for $\beta \in (0,4]$, $\nu \le \beta$,

$$\sup_x |g_\varepsilon(x) - g(x)| \le C|g|_\nu \varepsilon^\nu, \qquad (4.7)$$
$$|g_\varepsilon|_\beta \le C\varepsilon^{\nu-\beta}|g|_\nu.$$

(A standard mollifier could be taken if $\nu \le 2$; see Lemma 4). Since for $x \in \mathbf{R}^d$,

$$g_\varepsilon(x) - g(x) = \int \left[g(x-y) - g(x) - \sum_{1\le|\gamma|\le[\nu]^-} \frac{D^\gamma g(x)}{\gamma!} y^\gamma \right] w_\varepsilon(y)dy,$$

it follows that

$$\sup_x |g_\varepsilon(x) - g(x)| \le C|g|_\nu \varepsilon^\nu.$$

If β is an integer, $\gamma \in \mathbf{N}^d$, $|\gamma| = \beta$ and $\gamma = \mu + \mu'$ with $|\mu| = [\nu]$, $\mu' \ne \mathbf{0}$, then

$$\begin{aligned} D^\gamma g_\varepsilon(x) &= \varepsilon^{-[\gamma]} \int g(y)(D^\gamma w)_\varepsilon(x-y)dy = \varepsilon^{[\nu]-\beta} \int D^\mu g(y)(D^{\mu'} w)_\varepsilon(x-y)dy \\ &= \varepsilon^{[\nu]-\beta} \int [D^\mu g(y) - D^\mu g(x)](D^{\mu'} w)_\varepsilon(x-y)dydy \end{aligned}$$

and

$$|D^\gamma g_\varepsilon(x)| \le C\varepsilon^{\nu-\beta}|g|_\nu, \quad x \in \mathbf{R}^d.$$

If β is not an integer, the second inequality in (4.7) follows by interpolation.

According to Theorem 1 and (4.7),

$$\begin{aligned} |\mathbf{E}g(Y_T) - \mathbf{E}g(X_T)| &\le 2\sup_x |g_\varepsilon(x) - g(x)| + |\mathbf{E}g_\varepsilon(Y_T) - \mathbf{E}g_\varepsilon(X_T)| \\ &\le C|g|_\nu F(\varepsilon,\delta), \end{aligned}$$

where $F(\varepsilon,\delta) = \varepsilon^\nu + \varepsilon^{\nu-\beta}\delta^{\frac{\beta}{\alpha}-1}$. Minimizing F in $\varepsilon \in (0,1)$, the statement of Corollary 2 follows.

4.2. Approximate simple Euler scheme

Consider the approximation of X_t defined by the increments of $\tilde{Z}_t = Z_t^\sigma + R_t^\sigma$, $0 \le t \le T$, in Example 1. Obviously, $\tilde{Z}_t$ depends on α, β and σ. Its generator is

$$\tilde{L}v(x) = \int_0^t \int_{|\upsilon|>\varepsilon} [v(s,x+\upsilon) - v(s,x) - \chi_\alpha(\upsilon)(\nabla v(s,x),\upsilon)]\pi(d\upsilon) + R^{\alpha,\beta}v(x),$$

where

$$R^{\alpha,\beta}v(x) = \begin{cases} \int_{|\upsilon|\le\sigma} (\nabla v(x),\upsilon)\, d\pi & \text{if } \alpha < \beta \in (1,2], \alpha \in (0,1], \\ \frac{1}{2}\sum_{i,j}(B^{\sigma*}B^\sigma)_{ij}\partial^2_{ij}v(x) & \text{if } \alpha < \beta \in (2,4], \alpha \in (1,2], \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 6. *Let $\alpha < \beta \leq 2\alpha$ and $h \in \tilde{C}^{\beta}(\mathbf{R}^d)$. Then there is a constant C such that for every $\mathbb{F}^{Z^{\sigma}}$-stopping times $0 \leq \tau \leq \tau' \leq T$ we have*

$$|\mathbf{E}[h(Z_{\tau'} - Z_{\tau}) - h(\tilde{Z}_{\tau'} - \tilde{Z}_{\tau})|\mathcal{F}_{\tau}]| \leq C\phi(\sigma)|h|_{\beta}\mathbf{E}[\tau' - \tau|\mathcal{F}_{\tau}],$$

with

$$\phi(\sigma) = \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta\wedge 3} d\pi$$

(here $\mathbb{F}^{Z^{\sigma}}$ is the natural filtration of σ-algebras generated by Z^{σ}).

Proof. Let $\bar{Z}^{\sigma} = Z - Z^{\sigma}$. We show first that there is a constant C such that for any $s < t$, $g \in \tilde{C}^{\beta}(\mathbf{R}^d)$,

$$|\mathbf{E}g(\bar{Z}^{\sigma}_t - \bar{Z}^{\sigma}_s) - \mathbf{E}g(R^{\sigma}_t - R^{\sigma}_s)| \leq C\phi(\sigma)|g|_{\beta}|t - s|. \tag{4.8}$$

By Itô's formula

$$v(r, x) = \mathbf{E}[g(\bar{Z}^{\sigma}_t - \bar{Z}^{\sigma}_r + x)], \quad 0 \leq r \leq t, \tag{4.9}$$

is the solution of the backward Kolmogorov equation

$$\partial_t v(r, x) + \int_{|\upsilon|\leq\sigma} [v(r, x + \upsilon) - v(r, x) - \chi_{\alpha}(\upsilon)\,(\nabla v(r, x), \upsilon)]\pi(d\upsilon) = 0,$$
$$v(t, x) = g(x), \quad 0 \leq s \leq t. \tag{4.10}$$

Obviously, $v \in \tilde{C}^{\beta}([0, t] \times \mathbf{R}^d)$ and (see (4.9)) $|v|_{\beta} \leq |g|_{\beta}$. By Itô's formula and (4.10),

$$\begin{aligned}\mathbf{E}g(R^{\sigma}_t - R^{\sigma}_s) - \mathbf{E}g(\bar{Z}^{\sigma}_t - \bar{Z}^{\sigma}_s) &= \mathbf{E}v(t, R^{\sigma}_t - R^{\sigma}_s) - v(s, 0)\\ &= \mathbf{E}\int_s^t [R^{\alpha,\beta}v(r, R^{\sigma}_r - R^{\sigma}_s)\\ &\quad - \bar{L}v(r, R^{\sigma}_r - R^{\sigma}_s)]dr,\end{aligned} \tag{4.11}$$

where

$$\bar{L}v(r, x) = \int_{|\upsilon|\leq\sigma} [v(r, x + \upsilon) - v(r, x) - \chi_{\alpha}(\upsilon)\,(\nabla v(r, x), \upsilon)]\pi(d\upsilon), \quad (r, x) \in H.$$

If $\alpha < \beta \in (1, 2]$, $\alpha \in (0, 1]$, then for all $(r, x) \in H$,

$$\begin{aligned}&\left|R^{\alpha,\beta}v(r, x) - \int_{|\upsilon|\leq\sigma} [v(r, x + \upsilon) - v(r, x)]\pi(d\upsilon)\right|\\ &\leq \int_0^1 \int_{|\upsilon|\leq\sigma} |\nabla v(r, x + s\upsilon) - \nabla v(r, x)|\,|\upsilon| d\pi ds\\ &\leq C|v|_{\beta} \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta} d\pi \leq C|h|_{\beta} \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta} d\pi.\end{aligned}$$

If $\alpha < \beta \in (2, 4]$, $\alpha \in (1, 2]$, then for all $(r, x) \in H$,

$$\begin{aligned}&\left|R^{\alpha,\beta}v(r, x) - \int_{|\upsilon|\leq\sigma} [v(r, x + \upsilon) - v(r, x) - (\nabla v(r, x), \upsilon)]d\pi\right|\\ &\leq \int_0^1 \int_{|\upsilon|\leq\sigma} |D^2 v(r, x + s\upsilon) - D^2 v(r, x)|\,|\upsilon|^2 d\pi ds\end{aligned}$$

$$\leq C|\upsilon|_\beta \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta\wedge 3} d\pi \leq C|h|_\beta \int_{|\upsilon|\leq\sigma} |\upsilon|^{\beta\wedge 3} d\pi.$$

The estimate of the difference $R^{\alpha,\beta}v - \bar{L}v$ in other cases is straightforward and (4.8) follows by (4.11).

Since Z^σ, $\bar{Z}^\sigma$ and R^σ are independent and τ, τ' are $\mathbb{F}^{Z^\sigma}$ stopping times, we have by (4.8) that

$$|\mathbf{E}[h(Z^\sigma_{\tau'} - Z^\sigma_\tau + \bar{Z}^\sigma_{\tau'} - \bar{Z}^\sigma_\tau) - h(Z^\sigma_{\tau'} - Z^\sigma_\tau + R^\sigma_{\tau'} - R^\sigma_\tau)|\mathcal{F}_\tau]|$$
$$\leq C\phi(\sigma)|h|_\beta \mathbf{E}[\tau' - \tau|\mathcal{F}_\tau].$$

The statement follows. □

For the proof or Theorem 2 we will need the following estimate.

Lemma 7. *Let*

$$V_t = at + bW_t + GZ_t,$$

where $a \in \mathbf{R}^d$, b is a $d \times d$-matrix and G is an $m \times m$-matrix. We assume $b = 0$ if $\alpha \in (0,2)$ and $a = 0$ if $\alpha \in (0,1)$ and

$$|a| + |b| + |G| \leq K.$$

Let $\alpha < \beta \leq \mu \leq 2\alpha$ and $h \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$.

Then there is a constant $C = C(\alpha, \beta, K)$ such that

$$|\mathbf{E}h(V_t) - h(0)| \leq Ct^{\frac{\beta}{\alpha}-1}|h|_{\beta-\alpha}.$$

Proof. For $f \in \tilde{C}^\beta(\mathbf{R}^d)$, applying Itô's formula,

$$\mathbf{E}f(V_t) - f(0) = \mathbf{E}\int_0^t \mathcal{K}f(V_r)dr,$$

where for $x \in \mathbf{R}^d$,

$$\mathcal{K}f(x) = (a, \nabla f(x)) + \frac{1}{2}\sum_{i,j} b^* b \partial^2_{ij} f(x)$$
$$+ \int [f(x+\upsilon) - f(x) - \chi_\alpha(\upsilon)(\nabla f(x), \upsilon)]\pi(d\upsilon).$$

For $h \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$ we take $w \in C_0^\infty(\mathbf{R}^d)$ to be a nonnegative smooth function with support in $\{|x| \leq 1\}$ such that $w(x) = w(|x|)$, $x \in \mathbf{R}^d$, and $\int w(x)dx = 1$. For $x \in \mathbf{R}^d$ and $\varepsilon \in (0,1)$, define $w^\varepsilon(x) = \varepsilon^{-d} w\left(\frac{x}{\varepsilon}\right)$ and the convolution

$$h^\varepsilon(x) = \int f(y) w^\varepsilon(x-y)dy, \quad x \in \mathbf{R}^d.$$

Then by Lemma 4

$$|\mathbf{E}h(V_t) - h(0)| \leq 2\varepsilon^{\beta-\alpha}|h|_{\beta-\alpha} + \left|\mathbf{E}\int_0^t \mathcal{K}h^\varepsilon(V_r)dr\right|$$
$$\leq C|h|_{\beta-\alpha}(\varepsilon^{\beta-\alpha} + \varepsilon^{\beta-2\alpha}t)$$

for each $\varepsilon \in (0,1)$. The statement follows by minimizing the inequality in ε. □

4.2.1. Proof of Theorem 2

Let $u \in \tilde{C}^{\beta}(H)$ be the unique solution to the backward Kolmogorov equation

$$(\partial_t + L)\, u(t,x) = 0, \qquad (4.12)$$
$$u(T,x) = g(x).$$

Let for $\tau_i \leq t \leq \tau_{i+1}$

$$H_t^i = a(\tilde{Y}_{\tau_i})(t-\tau_i) + b(\tilde{Y}_{\tau_i})(W_t - W_{\tau_i}) + G(\tilde{Y}_{\tau_i})\left(Z_t - Z_{\tau_i}\right)$$

and denote $\Delta\tilde{Y}_{\tau_i} = \tilde{Y}_{\tau_{i+1}} - \tilde{Y}_{\tau_i}$. We approximate

$$\begin{aligned} u(T,\tilde{Y}_T) - u(0,Y_0) &= \sum_i u(\tau_{i+1},\tilde{Y}_{\tau_{i+1}}) - u(\tau_i,\tilde{Y}_{\tau_i}) \\ &= \sum_i [u(\tau_{i+1},\tilde{Y}_{\tau_i} + \Delta\tilde{Y}_{\tau_i}) - u(\tau_{i+1},\tilde{Y}_{\tau_i} + H^i_{\tau_{i+1}})] \\ &\quad + \sum_i [u(\tau_{i+1},\tilde{Y}_{\tau_i} + H^i_{\tau_{i+1}}) - u(\tau_i,\tilde{Y}_{\tau_i})] \\ &= D_1 + \sum_i D_{2i}. \end{aligned}$$

According to (2.4) (Lemma 6),

$$\mathbf{E}|D_1| \leq C\phi(\sigma)|u|_\beta \leq C\phi(\sigma)|g|_\beta.$$

Now, we estimate the second term. By Itô's formula for each i,

$$\begin{aligned} \mathbf{E}[D_{2i}|\mathcal{F}_{\tau_i}] &= \mathbf{E}[u(\tau_{i+1},\tilde{Y}_{\tau_i} + H^i_{\tau_{i+1}}) - u(\tau_{i+1},\tilde{Y}_{\tau_i})|\mathcal{F}_{\tau_i}] \\ &= \mathbf{E}\left\{\int_{\tau_i}^{\tau_{i+1}} [\partial_t u(r,\tilde{Y}_{\tau_i} + H^i_r) + L_{\tilde{Y}_{\tau_i}} u(r,\tilde{Y}_{\tau_i} + H^i_r)]dr|\mathcal{F}_{\tau_i}\right\} \\ &= \mathbf{E}\int_{\tau_i}^{\tau_{i+1}} [(\partial_t u(r,\tilde{Y}_{\tau_i} + H^i_r) - \partial_t u(r,\tilde{Y}_{\tau_i})) \\ &\quad + (L_{\tilde{Y}_{\tau_i}} u(r,\tilde{Y}_{\tau_i} + H^i_r) - L_{\tilde{Y}_{\tau_i}} u(r,\tilde{Y}_{\tau_i}))]dr \end{aligned}$$

and by Theorem 4 and Lemmas 1 and 7,

$$\begin{aligned} \left|\sum_i \mathbf{E}D_{2i}\right| &\leq \sum_i |\mathbf{E}D_{2i}| \leq C\delta^{\frac{\beta}{\alpha}-1}|Lu|_{\beta-\alpha} \\ &\leq C\delta^{\frac{\beta}{\alpha}-1}|u|_\beta \leq C\delta^{\frac{\beta}{\alpha}-1}|g|_\beta \end{aligned}$$

and the statement of Theorem 2 follows.

4.3. Approximate jump-adapted scheme

Consider the approximation of X_t defined by the increments of $\tilde{Z}_t = Z_t^\sigma + R_t^\sigma$, $0 \leq t \leq T$, in Example 1. For $\sigma \in (0,1)$, $\delta > 0$, consider the following Z^σ-jump-adapted time discretization: $\tau_0 = 0$,

$$\tau_{i+1} = \inf\left(t > \tau_i : \Delta Z_t^\sigma \neq 0\right) \wedge (\tau_i + \delta) \wedge T.$$

In this case, the time discretization $\{\tau_i, i = 0, \dots, n_T\}$ of the interval $[0, T]$ is random, τ_i are stopping times. We approximate X_t by

$$\hat{Y}_t = X_0 + \int_0^t a(\hat{Y}_{\tau_{i_s}})ds + \int_0^t b(\hat{Y}_{\tau_{i_s}})dW_s + \int_0^t G(\hat{Y}_{\tau_{i_s}})d\tilde{Z}_s, \quad t \in [0, T].$$

In this case,

$$\tau_{i+1} - \tau_i = \eta_{i+1} \wedge \delta \wedge (T - \tau_i)$$

with

$$\eta_{i+1} = \inf(t > 0 : p((\tau_i, \tau_i + t], \{|\upsilon| > \sigma\}) \geq 1)$$

and η_{i+1} is $\mathcal{F}_{\tau_i}$-conditionally exponential with parameter $\lambda_\sigma = \pi(\{|\upsilon| > \sigma\})$.

Lemma 8. *Let* $\delta_i' = \delta \wedge (T - \tau_i)$, $i \geq 0$, *and* $\lambda_\sigma = \pi(\{|\upsilon| > \sigma\})$.

(i) *There is a constant* $c > 0$ *such that for any* $i \geq 0$

$$c\left(\delta_i' \wedge \lambda_\sigma^{-1}\right) \leq \mathbf{E}[\tau_{i+1} - \tau_i|\mathcal{F}_{\tau_i}] \leq \delta_i' \wedge \lambda_\sigma^{-1}.$$

(ii) *There is a constant* C *such that for any* $i \geq 0$,

$$\begin{aligned} \mathbf{E}[(\tau_{i+1} - \tau_i)^2|\mathcal{F}_{\tau_i}] &\leq C\mathbf{E}[\delta_i'^2 \wedge \lambda_\sigma^{-2}|\mathcal{F}_{\tau_i}] \\ &\leq C(\delta \wedge \lambda_\sigma^{-1})\mathbf{E}[\tau_{i+1} - \tau_i|\mathcal{F}_{\tau_i}]. \end{aligned}$$

Proof. Since $\tau_{i+1} - \tau_i = \eta_{i+1} \wedge \delta \wedge (T - \tau_i)$ and

$$\eta_{i+1} = \inf(t > 0 : p((\tau_i, \tau_i + t], \{|\upsilon| > \sigma\}) \geq 1)$$

is $\mathcal{F}_{\tau_i}$-conditionally exponential with parameter λ_σ, we find

$$\begin{aligned} \mathbf{E}[\tau_{i+1} - \tau_i|\mathcal{F}_{\tau_i}] &= \mathbf{E}\left[\eta_{i+1} \wedge \delta_i'|\mathcal{F}_{\tau_i}\right] = \lambda_\sigma \int_0^{\delta_i'} te^{-\lambda_\sigma t}dt + \delta_i' e^{-\lambda_\sigma \delta_i'} \\ &= \frac{1 - e^{-\lambda_\sigma \delta_i'}}{\lambda_\sigma} \end{aligned}$$

and (i) follows. Similarly,

$$\begin{aligned} \mathbf{E}[(\tau_{i+1} - \tau_i)^2|\mathcal{F}_{\tau_i}] &= \lambda_\sigma \mathbf{E}\left[\int_0^{\delta_i'} t^2 e^{-\lambda_\sigma t}dt + \delta_i'^2 e^{-\lambda_\sigma \delta_i'}|\mathcal{F}_{\tau_i}\right]dt \\ &= \frac{2}{\lambda_\sigma^2}[-\lambda_\sigma \delta_i' e^{-\lambda_\sigma \delta_i'} + 1 - e^{-\lambda_\sigma \delta_i'}] \end{aligned}$$

and (ii) follows using (i). □

An immediate consequence of Lemma 8 is the following statement.

Corollary 5. (i) *There are constants* $c, C > 0$ *such that*

$$\begin{aligned} c\mathbf{E}\sum_i(\tau_{i+1} - \tau_i) &\leq \sum_i \mathbf{E}[(\delta \wedge \lambda_\sigma^{-1}) \wedge (T - \tau_i)] \\ &\leq C\mathbf{E}\sum_i(\tau_{i+1} - \tau_i) = CT. \end{aligned}$$

(ii) *There is* $C > 0$ *such that*

$$\sum_i \mathbf{E}[(\tau_{i+1} - \tau_i)^2] \leq CT(\delta \wedge \lambda_\sigma^{-1}).$$

Proof. We derive (i) by summing inequalities in Lemma 8(i). According to Lemma 8(ii) and (i),

$$\begin{aligned}\sum_i \mathbf{E}[(\tau_{i+1} - \tau_i)^2] &\leq C \sum_i \mathbf{E}[(T - \tau_i)^2 \wedge \delta^2 \wedge \lambda_\sigma^{-2}] \\ &\leq C(T \wedge \delta \wedge \lambda_\sigma^{-1}) \sum_i \mathbf{E}[(T - \tau_i) \wedge \delta \wedge \lambda_\sigma^{-1}] \\ &\leq CT(\delta \wedge \lambda_\sigma^{-1}).\end{aligned}$$

The statement follows. □

For the proof of Theorem 3 we will need the following estimate as well.

Lemma 9. *Let*

$$V_t = at + bW_t + GZ_t,$$

where $a \in \mathbf{R}^d$, b *is a* $d \times d$*-matrix and* G *is an* $m \times m$*-matrix. We assume* $b = 0$ *if* $\alpha \in (0, 2)$ *and* $a = 0$ *if* $\alpha \in (0, 1)$ *and*

$$|a| + |b| + |G| \leq K.$$

Let $\alpha < \beta \leq \mu \leq 2\alpha$ *and* $h \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$.

Then there is a constant $C = C(\alpha, \beta, K)$ *such that for any* $i \geq 0$

$$\left|\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} h(V_r) - h(V_{\tau_i}) | \mathcal{F}_{\tau_i}\right]\right| \leq C|h|_{\beta-\alpha} \tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1} \left(\delta \wedge \lambda_\sigma^{-1}\right)^{\frac{\beta}{\alpha}-1} \mathbf{E}[(\tau_{i+1} - \tau_i)|\mathcal{F}_{\tau_i}],$$

where $\lambda_\sigma = \pi(\{|\upsilon| > \sigma\})$,

$$\tilde{\lambda}_\sigma = 1 + 1_{\alpha\in(1,2)} \left|\int_{1\geq|\upsilon|>\sigma} \upsilon d\pi\right|.$$

Proof. For $f \in \tilde{C}^\beta(\mathbf{R}^d)$, $i \geq 0$, applying Itô's formula,

$$\begin{aligned}&\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} f(V_r) - f(V_{\tau_i}) | \mathcal{F}_{\tau_i}\right] dr \\ &\quad = \mathbf{E} \int_{\tau_i}^{\tau_{i+1}} \left[\int_{\tau_i}^{s} [\mathcal{K} f(V_r) dr + M_s - M_{\tau_i}] ds | \mathcal{F}_{\tau_i}\right] dr,\end{aligned}$$

where for $x \in \mathbf{R}^d$,

$$\begin{aligned}\mathcal{K} f(x) = (a, \nabla f(x)) &+ \frac{1}{2} \sum_{i,j} b^* b \partial_{ij}^2 f(x) \\ &+ \int [f(x + \upsilon) - f(x) - \chi_\alpha(\upsilon)(\nabla f(x), \upsilon)]\pi(d\upsilon)\end{aligned}$$

and

$$M_t = \int_0^t \int [f(V_{r-} + G\upsilon) - f(V_{r-})] q(dr, d\upsilon), \quad t \in [0, T].$$

Note that

$$\int_{\tau_i}^{\tau_{i+1}} (M_s - M_{\tau_i}) d(s - \tau_i) = (M_{\tau_{i+1}} - M_{\tau_i})(\tau_{i+1} - \tau_i) - \int_{\tau_i}^{\tau_{i+1}} (s - \tau_i) dM_s.$$

Since Z^σ and $\bar{Z}^\sigma = Z - Z^\sigma$ are independent and τ_i are $\mathbb{F}^{Z^\sigma}$-stopping times, it follows by the definition of τ_i that

$$\begin{aligned}
&\mathbf{E}\left[(M_{\tau_{i+1}} - M_{\tau_i})(\tau_{i+1} - \tau_i) - \int_{\tau_i}^{\tau_{i+1}} (s - \tau_i) dM_s | \mathcal{F}_{\tau_i}\right] \\
&\quad = \mathbf{E}\left[-(\tau_{i+1} - \tau_i)(U^\sigma_{\tau_{i+1}} - U^\sigma_{\tau_i}) + \int_{\tau_i}^{\tau_{i+1}} (s - \tau_i) dU^\sigma_s | \mathcal{F}_{\tau_i}\right] \\
&\quad = -\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} (U^\sigma_s - U^\sigma_{\tau_i}) ds | \mathcal{F}_{\tau_i}\right],
\end{aligned}$$

where

$$\begin{aligned}
U^\sigma_t &= \int_0^t \int_{|\upsilon| > \sigma} [f(V_{r-} + G\upsilon) - f(V_{r-})] d\pi dr \\
&= \int_0^t \int_{|\upsilon| > 1} [f(V_{r-} + G\upsilon) - f(V_{r-})] d\pi dr \\
&\quad + \int_0^t \int_{1 \geq |\upsilon| > \sigma} \chi_\alpha(\upsilon)(\nabla f(V_r), \upsilon) d\pi dr \\
&\quad + \int_0^t \int_{1 \geq |\upsilon| > \sigma} [f(V_{r-} + G\upsilon) - f(V_{r-}) - \chi_\alpha(\upsilon)(\nabla f(V_r), \upsilon)] d\pi dr.
\end{aligned}$$

Hence

$$\begin{aligned}
&\left|\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} f(V_r) - f(V_{\tau_i}) | \mathcal{F}_{\tau_i}\right] dr\right| \\
&\quad \leq C\left(1 + 1_{\alpha \in (1,2)} \left|\int_{1 \geq |\upsilon| > \varepsilon} \upsilon d\pi\right|\right) |f|_\beta \mathbf{E}[(\tau_{i+1} - \tau_i)^2 | \mathcal{F}_{\tau_i}]. \qquad (4.13)
\end{aligned}$$

For $h \in \tilde{C}^{\beta-\alpha}(\mathbf{R}^d)$ we take $w \in C_0^\infty(\mathbf{R}^d)$ to be a nonnegative smooth function with support in $\{|x| \leq 1\}$ such that $w(x) = w(|x|)$, $x \in \mathbf{R}^d$, and $\int w(x) dx = 1$. For $x \in \mathbf{R}^d$ and $\varepsilon \in (0, 1)$, define $w^\varepsilon(x) = \varepsilon^{-d} w\left(\frac{x}{\varepsilon}\right)$ and the convolution

$$h^\varepsilon(x) = \int f(y) w^\varepsilon(x - y) dy, \quad x \in \mathbf{R}^d.$$

Then by Lemma 4 and (4.13),

$$\begin{aligned}
&\left|\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} h(V_r) - h(V_\tau) | \mathcal{F}_{\tau_i}\right] dr\right| \\
&\quad \leq 2\varepsilon^{\beta-\alpha} |h|_{\beta-\alpha} \mathbf{E}[(\tau_{i+1} - \tau_i) \,| \mathcal{F}_\tau] + \left|\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} (h^\varepsilon(V_r) - h^\varepsilon(V_\tau)) dr | \mathcal{F}_\tau\right]\right| \\
&\quad \leq 2\varepsilon^{\beta-\alpha} |h|_{\beta-\alpha} \mathbf{E}[\tau_{i+1} - \tau_i | \mathcal{F}_{\tau_i}] + C\varepsilon^{\beta-2\alpha} \\
&\qquad \times \left[1 + 1_{\alpha \in (1,2)} \int_{1 \geq |\upsilon| > \varepsilon} \upsilon d\pi\right] |h|_{\beta-\alpha} \mathbf{E}[(\tau_{i+1} - \tau_i)^2 | \mathcal{F}_{\tau_i}].
\end{aligned}$$

Minimizing the inequality in ε we find by Lemma 8(ii) that

$$\left|\mathbf{E}\left[\int_{\tau_i}^{\tau_{i+1}} h(V_r) - h(V_\tau)|\mathcal{F}_{\tau_i}\right]dr\right|$$
$$\leq C|h|_{\beta-\alpha}\tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1}\mathbf{E}[\tau_{i+1}-\tau_i|\mathcal{F}_{\tau_i}]^{2-\frac{\beta}{\alpha}}\mathbf{E}[(\tau_{i+1}-\tau_i)^2|\mathcal{F}_{\tau_i}]^{\frac{\beta}{\alpha}-1}$$
$$\leq C|h|_{\beta-\alpha}\tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1}(\delta\wedge\lambda_\sigma^{-1})^{\frac{\beta}{\alpha}-1}\mathbf{E}[\tau_{i+1}-\tau_i|\mathcal{F}_{\tau_i}]. \quad \square$$

4.3.1. Proof of Theorem 3

Let $u \in \tilde{C}^\beta(H)$ be the unique solution to the backward Kolmogorov equation (see Theorem 4)

$$(\partial_t + L)\,u(t,x) = 0, \qquad (4.14)$$
$$u(T,x) = g(x).$$

Let for $\tau_i \leq t \leq \tau_{i+1}$

$$H_t^i = a(\hat{Y}_{\tau_i})(t-\tau_i) + b(\hat{Y}_{\tau_i})(W_t - W_{\tau_i}) + G(\hat{Y}_{\tau_i})\left(Z_t - Z_{\tau_i}\right)$$

and denote $\Delta\hat{Y}_{\tau_i} = \hat{Y}_{\tau_{i+1}} - \hat{Y}_{\tau_i}$. We approximate

$$\begin{aligned} u(T,\hat{Y}_T) - u(0,X_0) &= \sum_i u(\tau_{i+1},\hat{Y}_{\tau_{i+1}}) - u(\tau_i,\hat{Y}_{\tau_i}) \\ &= \sum_i [u(\tau_{i+1},\hat{Y}_{\tau_i}+\Delta\hat{Y}_{\tau_i}) - u(\tau_{i+1},\hat{Y}_{\tau_i}+H^i_{\tau_{i+1}})] \\ &\quad + \sum_i [u(\tau_{i+1},\hat{Y}_{\tau_i}+H^i_{\tau_{i+1}}) - u(\tau_i,\hat{Y}_{\tau_i})] \\ &= D_1 + \sum_i D_{2i}. \end{aligned}$$

According to Lemma 6,

$$\mathbf{E}|D_1| \leq C\phi(\sigma)|u|_\beta \leq C\phi(\sigma)|g|_\beta.$$

Now, we estimate the second term. By Itô's formula for each i,

$$\begin{aligned} \mathbf{E}[D_{2i}|\mathcal{F}_{\tau_i}] &= \mathbf{E}[u(\tau_{i+1},\hat{Y}_{\tau_i}+H^i_{\tau_{i+1}}) - u(\tau_{i+1},\hat{Y}_{\tau_i})|\mathcal{F}_{\tau_i}] \\ &= \mathbf{E}\left\{\int_{\tau_i}^{\tau_{i+1}}[\partial_t u(r,\hat{Y}_{\tau_i}+H_r^i) + L_{\hat{Y}_{\tau_i}}u(r,\hat{Y}_{\tau_i}+H_r^i)]dr|\mathcal{F}_{\tau_i}\right\} \\ &= \mathbf{E}\int_{\tau_i}^{\tau_{i+1}}[(\partial_t u(r,\hat{Y}_{\tau_i}+H_r^i) - \partial_t u(r,\hat{Y}_{\tau_i})) \\ &\quad + (L_{\hat{Y}_{\tau_i}}u(r,\hat{Y}_{\tau_i}+H_r^i) - L_{\hat{Y}_{\tau_i}}u(r,\hat{Y}_{\tau_i}))]dr \end{aligned}$$

and by Theorem 4 and Lemmas 1, 9 and Corollary 5,

$$\begin{aligned} \left|\sum_i \mathbf{E}D_{2i}\right| &\leq \sum_i |\mathbf{E}D_{2i}| \leq C\tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1}\left(\delta\wedge\lambda_\sigma^{-1}\right)^{\frac{\beta}{\alpha}-1}(|\partial_t u|_{\beta-\alpha} + |Lu|_{\beta-\alpha}) \\ &\leq C\tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1}\left(\delta\wedge\lambda_\sigma^{-1}\right)^{\frac{\beta}{\alpha}-1}|u|_\beta \leq C\tilde{\lambda}_\sigma^{\frac{\beta}{\alpha}-1}\left(\delta\wedge\lambda_\sigma^{-1}\right)^{\frac{\beta}{\alpha}-1}|g|_\beta \end{aligned}$$

and the statement of Theorem 3 follows.

5. Conclusion

The paper studies a simple weak Euler approximation of solutions to possibly completely degenerate stochastic differential equations driven by Lévy processes. The dependence of the rate of convergence on the regularity of coefficients and driving processes is investigated under the assumption of β-Lipschitz continuity of the coefficients. It is assumed that the SDE is driven by Levy processes of order $\alpha \in (0, 2]$ and that the tail of the Lévy measure of the driving process has a μ-order finite moment ($\mu \in (\alpha, 2\alpha]$). The resulting rate depends on β, α and μ. Following [3], the robustness of the results to the approximation of the law of the increments of the driving noise is studied as well. It is shown that time discretization and substitution errors add up. In addition, a jump-adapted approximate Euler scheme is considered as well. The derived error estimate shows that sometimes the inclusion of jump moments into time discretization $\{\tau_i\}$ could improve the convergence rate. In order to estimate the rate of convergence, the existence of a unique solution to the corresponding backward degenerate Kolmogorov equation in Lipschitz space is first proved.

On the other hand, there is a discrepancy in the model (1.1) between $\alpha = 2$ and $\alpha \in (0, 2)$. One would like to consider the equation

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s^\alpha + \int_0^t G(X_{s-})dZ_s, \quad t \in [0, T],$$

with a possibly degenerate b and a spherically symmetric α-stable W^α (in (1.1), $b = 0$ for $\alpha \in (0, 2)$).

Since (1.1) could be degenerate, a solution corresponding to a given $\alpha \in (0, 2]$ can be looked at as a solution corresponding to $\bar{\alpha} \in (\alpha, 2]$ as well. Therefore the rate for a fixed α cannot be "universally optimal": there is always a large subclass for which the rate claimed for α could be better and achieved under weaker assumptions. For example, if $\beta = \mu = 2\alpha$ with $\alpha \in (0, 2)$ (the diffusion part is absent), the convergence order is $\kappa = 1$ ($\mu = 4$ and $G \in \tilde{C}^4$ is not needed). Even "strictly at α", the assumption about the tail moment $\mu \in (\alpha, 2\alpha]$ is not optimal. It could be weakened for a subclass with the driving processes Z such that the compensator of the jump measure of X_t has a nice density with respect to a reference measure. For example, let us consider the following one dimensional model

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s + \int_0^t G(X_{s-})dZ_s, \quad t \in [0, T], \tag{5.1}$$

where Z is a symmetric λ-stable with $\lambda \in (0, 1)$ and $G \geq 0$. Assume $a, b, G^\lambda, g \in \tilde{C}^4(\mathbf{R})$. Although $\mu < 1$ in this case and the equation is possibly degenerate, a plausible convergence rate is still $\kappa = 1$ (or $\kappa = \nu/4$ if $g \in \tilde{C}^\nu(\mathbf{R})$, $\nu \in (0, 4]$), because the integral part of the generator of (5.1),

$$Iv(x) = \int [v(x + G(x)y) - v(x)]\frac{dy}{|y|^{1+\lambda}} = G(x)^\lambda \int [v(x + y) - v(x)]\frac{dy}{|y|^{1+\lambda}},$$

is differentiable without assuming much about the tail moments of the Lévy measure.

Acknowledgments

The author is very grateful to the referees for their valuable comments and suggestions.

References

[1] R. Cont, P. Tankov, Financial Modelling with Jump Processes, Chapman and Hall/CRC Press, 2004.
[2] R. Estrada, Vector moment problem for rapidly decreasing smooth functions of several variables, Proceedings of the American Mathematical Society 126 (1998) 761–768.
[3] J. Jacod, T. Kurtz, S. Méléard, P. Protter, The approximate Euler scheme for Lévy driven stochastic differential equations, Annales Henri Poincaré 41 (2005) 523–558.
[4] B. Jourdain, A. Kohatsu-Higa, A review of recent results on approximation of solutions of stochastic differential equations, Progress in Probability 65 (2011) 121–144.
[5] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer Verlag, 2000.
[6] A. Kohatsu-Higa, P. Tankov, Jump-adapted discretization schemes for Lévy-driven SDEs, Stochastic Processes and their Applications 120 (2010) 2258–2285.
[7] N.V. Krylov, Controlled Diffusion Processes, Springer Verlag, 1980.
[8] K. Kubilius, E. Platen, Rate of weak convergence of the Euler approximation for diffusion processes with jumps, Quantitative Finance Research Centre, University of Technology, Sydney, Research Paper Series 54, 2001.
[9] R. Mikulevičius, E. Platen, Time discrete Taylor approximations for Itô processes with jump component, Mathematische Nachrichten 138 (1988) 93–104.
[10] R. Mikulevičius, E. Platen, Rate of convergence of the Euler approximation for diffusion processes, Mathematische Nachrichten 151 (1991) 233–239.
[11] R. Mikulevičius, H. Pragarauskas, On L_p-estimates of some singular integrals related to jump processes, arXiv:1008.3044v3 [math.PR], 2010.
[12] R. Mikulevičius, C. Zhang, On the rate of convergence of weak Euler approximation for nondegenerate SDEs driven by Lévy processes, Stochastic Processes and their Applications 121 (2011) 1720–1748.
[13] G.N. Milstein, A method of second-order accuracy integration of stochastic differential equations, Theory of Probability and its Applications 23 (1979) 396–401.
[14] G.N. Milstein, Weak approximation of solutions of systems of stochastic differential equations, Theory of Probability and its Applications 30 (1986) 750–766.
[15] E. Platen, An introduction to numerical methods for stochastic differential equations, Acta Numerica 8 (1999) 197–246.
[16] E. Platen, N. Bruti-Liberati, Numerical Solutions of Stochastic Differential Equations with Jumps in Finance, Springer Verlag, 2010.
[17] P.E. Protter, D. Talay, The Euler scheme for Lévy driven stochastic differential equations, The Annals of Probability 25 (1997) 393–423.
[18] D. Talay, Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE. and applications, in: Filtering and Control of Random Processes, in: Lecture Notes in Control and Information Sciences, vol.61, 1984, pp. 294–313.
[19] D. Talay, Discretization of a stochastic differential equation and rough estimate of the expectations of functionals of the solution, ESAIM: Mathematical Modelling and Numerical Analysis—Modélisation Mathématique et Analyse Numérique 20 (1986) 141–179.