WEAK EULER SCHEME FOR LÉVY-DRIVEN STOCHASTIC DIFFERENTIAL EQUATIONS*

R. MIKULEVIČIUS† AND CH. ZHANG‡

Abstract. This paper studies the rate of convergence of the weak Euler approximation for solutions to Lévy-driven stochastic differential equations with nondegenerate main part driven by a spherically symmetric stable process, under the assumption of Hölder continuity. The rate of convergence is derived for a full regularity scale based on solving the associated backward Kolmogorov equation and investigating the dependence of the rate on the regularity of the coefficients and driving processes.

Key words. stochastic differential equations, Lévy processes, weak Euler approximation, rate of convergence, Hölder conditions

1. Introduction. Stochastic differential equations arise from a broad variety of fields [21]. Let $T \in (0, \infty)$ and $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space with a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \in [0,T]}$ of σ-algebras satisfying usual conditions. Suppose $X = \{X_t\}_{t \in [0,T]}$ is an $\mathbb{F}$-adapted stochastic process solving a given stochastic differential equation. When the equation satisfies certain conditions, there may exist a closed-form expression for X. In general, this is unrealistic, and numerical approximations are used [10]. Among others, weak Euler is one of the most commonly applied approximations, provided that it converges.

The Euler approximation $Y = \{Y_t\}_{t \in [0,T]}$ is said to converge to X with order $\kappa > 0$ if for each smooth function g with bounded derivatives, there exists a constant C, depending only on g, such that

$$|\mathbf{E}[g(Y_T)] - \mathbf{E}[g(X_T)]| \leqslant C\kappa(\delta) = C\delta^\kappa,$$

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

For diffusion processes, the problem of estimating the rate of convergence has been well addressed, for instance, in the case of smooth conditions [19], [20], [26], [27] as well as in the case of Hölder continuity [13].

While diffusion processes are probably the most intensively investigated, they exhibit almost surely continuous sample paths, which is not necessarily the case in many applications. On the other hand, stochastic processes with jumps provide more flexibility in modeling dynamic phenomena with continuous and discontinuous uncertainty. In particular, Lévy processes [1], [4], [24] are the simplest generic class of processes having a.s. continuous paths interspersed with random jumps of arbitrary sizes occurring at random times. Stochastic differential equations driven by Lévy processes are therefore widely employed in modeling systems arising from, for instance, mathematics, science, engineering, and finance [2], [25], [29].

†Department of Mathematics, University of Southern California, Los Angeles, CA 90007 (mikulvcs@math.usc.edu).

‡Department of Finance and Banking, Curtin University, Miri, Sarawak 98009, Malaysia (changyong.zhang@curtin.edu.my).

For Lévy-driven stochastic differential equations under smooth conditions, numerical approximations have been consistently studied in, for example, [8], [9], [11], [22], [23].

The rate of convergence of the weak Euler scheme has also been estimated under the assumption that the coefficients are Hölder-continuous, with a minor restriction on the scale of regularity [18]. In the present paper, the restriction is removed, and for the first time the rate is derived for the whole Hölder–Zygmund scale, by looking into the dependence of the rate of convergence on the Hölder regularity of the coefficients and driving processes. Here, for a driving process a part of its regularity is the variation of the process. From this perspective, the Wiener process is the most "chaotic" among α-stable processes.

The paper is organized as follows. In section 2, the stochastic differential equations under consideration are introduced and the rate of convergence is stated, followed by the proof of the main theorem, along with the essential technical results, presented in section 3.

2. Model and result.

2.1. Lévy-driven stochastic differential equations. Let $\alpha \in (0, 2]$ and $\mathbf{R}_0^d = \mathbf{R}^d \setminus \{0\}$. Consider in $\mathbf{R}^d$ the model

$$(1) \qquad X_t = X_0 + \int_0^t a(X_{s-})\, ds + \int_0^t b(X_{s-})\, dS_s + \int_0^t c(X_{s-})\, dL_s, \qquad t \in [0, T],$$

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

$$a(x) = \big(a^i(x)\big)_{1 \leqslant i \leqslant d}, \ b(x) = \big(b^{ij}(x)\big)_{1 \leqslant i,j \leqslant d}, \ c(x) = \big(c^{ij}(x)\big)_{1 \leqslant i \leqslant d,\, 1 \leqslant j \leqslant m}$$

are measurable and bounded, with $a = 0$ if $\alpha \in (0, 1)$ and nondegenerate b.

The main part of the equation is driven by $S = \{S_t\}_{t \in [0,T]}$, a d-dimensional $\mathbb{F}$-adapted standard spherically symmetric α-stable process defined by

$$S_t = \int_0^t \int y \big[\big(1 - \chi^\alpha(y)\big)\, p_0(dy, ds) + \chi^\alpha(y)\, q_0(dy, ds)\big], \qquad \alpha \in (0, 2),$$

where $\chi^\alpha(y) = \mathbf{1}_{\{\alpha=1\}} \mathbf{1}_{\{|y| \leqslant 1\}} + \mathbf{1}_{\{\alpha \in (1,2)\}}$, $p_0(dy, dt)$ is a Poisson point measure on $\mathbf{R}_0^d \times [0, \infty)$ with

$$\mathbf{E}[p_0(dy, dt)] = \frac{dy}{|y|^{d+\alpha}}\, dt, \quad q_0(dy, dt) = p_0(dy, dt) - \frac{dy}{|y|^{d+\alpha}}\, dt.$$

The process S is a standard Wiener process in $\mathbf{R}^d$ if $\alpha = 2$.

The last term is driven by $L = \{L_t\}_{t \in [0,T]}$, an m-dimensional $\mathbb{F}$-adapted Lévy process whose characteristic function is $\exp\{t\eta(\xi)\}$ with

$$\eta(\xi) = \int_{\mathbf{R}_0^m} \big[e^{i(\xi, y)} - 1 - i(\xi, y)\mathbf{1}_{\{\alpha \in (1,2]\}} \mathbf{1}_{\{|y| \leqslant 1\}}\big]\, \pi(dy),$$

that is,

$$L_t = \int_0^t \int y \big[\big(1 - \widetilde{\chi}^\alpha(y)\big)\, p(dy, ds) + \widetilde{\chi}^\alpha(y)\, q(dy, ds)\big],$$

where $\widetilde{\chi}^\alpha(y) = \mathbf{1}_{\{\alpha \in (1,2]\}} \mathbf{1}_{\{|y| \leqslant 1\}}$, $p(dy, dt)$ is a Poisson point measure on $\mathbf{R}_0^m \times [0, \infty)$ with $\mathbf{E}[p(dy, dt)] = \pi(dy)\, dt$, and $q(dy, dt) = p(dy, dt) - \pi(dy)\, dt$ is the centered Poisson measure. It is assumed that

$$\int (|y|^\alpha \wedge 1)\, \pi(dy) < \infty.$$

Mathematical models defined by (1) are used to describe random dynamic phenomena arising from various fields. A typical example is the following intensively applied jump-diffusion process, the results of which are stated in Corollary 1.

Example 1. Consider an $\mathbb{F}$-adapted process $X = \{X_t\}_{t \in [0,T]}$ solving

$$(2) \quad X_t = X_0 + \int_0^t a(X_{s-})\, ds + \int_0^t b(X_{s-})\, dW_s + \int_0^t c(X_{s-})\, dL_s, \qquad t \in [0, T],$$

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

$$a(x) = \left(a^i(x)\right)_{1 \leqslant i \leqslant d}, \ b(x) = \left(b^{ij}(x)\right)_{1 \leqslant i,j \leqslant d}, \ c(x) = \left(c^{ij}(x)\right)_{1 \leqslant i \leqslant d,\, 1 \leqslant j \leqslant m}$$

are measurable and bounded with $\inf_x |\det b(x)| > 0$, $W = \{W_t\}_{t \in [0,T]}$ is a d-dimensional standard Wiener process, and $L = \{L_t\}_{t \in [0,T]}$ is an m-dimensional Lévy process such that there exists a number $\mu \in (0, 3)$ satisfying

$$\int_{|y| \leqslant 1} |y|^2\, \pi(dy) + \int_{|y| > 1} |y|^\mu\, \pi(dy) < \infty.$$

2.2. Weak Euler approximation. The weak Euler approximation of X defined in (1) is an $\mathbb{F}$-adapted stochastic process $Y = \{Y_t\}_{t \in (0,T]}$ defined by the stochastic equation

$$(3) \quad Y_t = X_0 + \int_0^t a(Y_{\tau_{i_s}})\, ds + \int_0^t b(Y_{\tau_{i_s}})\, dS_s + \int_0^t c(Y_{\tau_{i_s}})\, dL_s, \qquad t \in [0, T],$$

where $\{\tau\}_\delta = \{\tau_i\}_{i=0,\ldots,n_T}$ is a time discretization of the interval $[0, T]$ satisfying $0 = \tau_0 < \tau_1 < \cdots < \tau_{n_T} = T$ and $\max_i(\tau_i - \tau_{i-1}) \leqslant \delta \in (0, 1)$. Here $\tau_{i_s} = \tau_i$ if $s \in [\tau_i, \tau_{i+1})$, $i = 0, \ldots, n_T - 1$. Contrary to the coefficients in (1), those in (3) are piecewise constant in each time interval $[\tau_i, \tau_{i+1})$.

The rate of convergence is stated in Theorem 1. Without being explicitly specified, $C = C(\cdot, \ldots, \cdot)$ denotes possibly different constants depending only on the corresponding arguments, and the following notation are used.

Denote $\mathbf{H} = [0, T] \times \mathbf{R}^d$ and $\mathbf{N} = \{0, 1, \ldots\}$. For $x, y \in \mathbf{R}^d$, write $(x, y) = \sum_{i=1}^d x_i y_i$, $|x| = \sqrt{(x, x)}$, and for $B \in \mathbf{R}^{d \times d}$, $|B| = \sum_{i=1}^d |B^{ii}|$.

For $(t, x) \in \mathbf{H}$, multi-index $\gamma \in \mathbf{N}^d$, and $i, j = 1, \ldots, d$, denote

$$\partial_t u(t, x) = \frac{\partial}{\partial t} u(t, x), \quad \partial_x^\gamma u(t, x) = \frac{\partial^{|\gamma|}}{\partial^{\gamma_1} x_1 \cdots \partial^{\gamma_d} x_d} u(t, x),$$

$$\partial_i u(t, x) = \frac{\partial}{\partial x_i} u(t, x), \quad \partial_{ij}^2 u(t, x) = \frac{\partial^2}{\partial x_i x_j} u(t, x),$$

$$\partial_x u(t, x) = \nabla_x u(t, x) = \left(\partial_1 u(t, x), \ldots, \partial_d u(t, x)\right),$$

$$\partial^2 u(t, x) = \Delta u(t, x) = \sum_{i=1}^d \partial_{ii}^2 u(t, x).$$

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

$$|u|_\beta = \sum_{|\gamma| \leqslant [\beta]^-} \sup_{(t,x) \in \mathbf{H}} |\partial_x^\gamma u(t,x)|$$

$$+ \mathbf{1}_{\{\{\beta\}^+ < 1\}} \sup_{\substack{|\gamma| = [\beta]^-, \\ t,x,h \neq 0}} \frac{|\partial_x^\gamma u(t, x+h) - \partial_x^\gamma u(t,x)|}{|h|^{\{\beta\}^+}}$$

$$+ \mathbf{1}_{\{\{\beta\}^+ = 1\}} \sup_{\substack{|\gamma| = [\beta]^-, \\ t,x,h \neq 0}} \frac{|\partial_x^\gamma u(t, x+h) + \partial_x^\gamma u(t, x-h) - 2\partial_x^\gamma u(t,x)|}{|h|^{\{\beta\}^+}}$$

is finite. Accordingly, $C^\beta(\mathbf{R}^d)$ denotes the space on $\mathbf{R}^d$. The classes C^β are Hölder–Zygmund spaces and coincide with Hölder spaces if $\beta \notin \mathbf{N}$ [28].

For $u \in C^\beta(\mathbf{R}^d)$, $\beta \in (0, 1]$, denote $|u|_0 = \sup_x |u(x)|$ and

$$[u]_\beta = \mathbf{1}_{\{\beta \in (0,1)\}} \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\beta}$$

$$+ \mathbf{1}_{\{\beta = 1\}} \sup_{x, h \neq 0} \frac{|u(x+h) - 2u(x) + u(x-h)|}{|h|^\beta}.$$

Define $\widetilde{C}^\beta(\mathbf{R}^d)$, $\beta > 0$, as a space such that for $\beta \notin \mathbf{N}$, $\widetilde{C}^\beta(\mathbf{R}^d) = C^\beta(\mathbf{R}^d)$ and for $\beta \in \mathbf{N}$, $\widetilde{C}^\beta(\mathbf{R}^d)$ is a space of functions u on $\mathbf{R}^d$ having $\beta - 1$ continuous bounded derivatives and $\partial_x^\gamma u(x)$, $|\gamma| = \beta - 1$, are Lipschitz.

For $u \in \widetilde{C}^\beta(\mathbf{R}^d)$, $\beta > 0$, denote

$$\|u\|_\beta = \mathbf{1}_{\{\beta \notin \mathbf{N}\}} |u|_\beta + \mathbf{1}_{\{\beta \in \mathbf{N}\}} \left[\sum_{|\gamma| \leqslant \beta - 1} |\partial_x^\gamma u|_0 + \sup_{\substack{x \neq y, \\ |\gamma| = \beta - 1}} \frac{|\partial_x^\gamma u(x) - \partial_x^\gamma u(y)|}{|x - y|} \right].$$

THEOREM 1. *Let $Y = \{Y_t\}_{t \in (0,T]}$ be the weak Euler approximation with step size $\delta \in (0,1)$ of the stochastic process $X = \{X_t\}_{t \in (0,T]}$ defined by (1). For $\alpha \in (0,2]$, $\beta \in (0,3)$, and $\mu \in [\beta, \alpha + \beta)$, assume $a^i, b^{ij} \in \widetilde{C}^\beta(\mathbf{R}^d)$, $1 \leqslant i, j \leqslant d$, $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant m$, $\inf_x |\det b(x)| > 0$, and*

$$\int_{|y| \leqslant 1} |y|^\alpha \, \pi(dy) + \int_{|y| > 1} |y|^\mu \, \pi(dy) < \infty.$$

Then there exists a constant C such that

$$|\mathbf{E}[g(Y_T)] - \mathbf{E}[g(X_T)]| \leqslant C|g|_{\alpha + \beta} \kappa(\delta, \alpha, \beta) \qquad \forall g \in C^{\alpha + \beta}(\mathbf{R}^d),$$

$$\left| \mathbf{E}\left[\int_0^T f(Y_{\tau_{i_s}}) \, ds \right] - \mathbf{E}\left[\int_0^T f(X_s) \, ds \right] \right| \leqslant C|f|_\beta \kappa(\delta, \alpha, \beta) \qquad \forall f \in C^\beta(\mathbf{R}^d),$$

where

$$\kappa(\delta, \alpha, \beta) = \begin{cases} \delta^{\beta/\alpha}, & \beta < \alpha, \\ \delta(1 + |\ln \delta|), & \beta = \alpha, \\ \delta, & \beta > \alpha. \end{cases}$$

Remark 1. The assumption $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d, 1 \leqslant j \leqslant m$, indicates that if $\mu < 1$, a heavier tail of π can be balanced by a higher regularity of c^{ij}.

Applying Theorem 1 to Example 1 results in Corollary 1.

COROLLARY 1. *Let* $Y = \{Y_t\}_{t \in (0,T]}$ *be the weak Euler approximation of the stochastic process* $X = \{X_t\}_{t \in (0,T]}$ *defined by* (2). *For* $\beta \in (0,3)$ *and* $\mu \in [\beta, 2+\beta)$, *assume* $a^i, b^{ij} \in \widetilde{C}^\beta(\mathbf{R}^d)$, $1 \leqslant i, j \leqslant d$, $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant m$, $\inf_x |\det b(x)| > 0$, *and*

$$\int_{|y| \leqslant 1} |y|^2 \, \pi(dy) + \int_{|y| > 1} |y|^\mu \, \pi(dy) < \infty.$$

Then there exists a constant C such that

$$|\mathbf{E}[g(Y_T)] - \mathbf{E}[g(X_T)]| \leqslant C|g|_{2+\beta} \kappa(\delta, \beta) \qquad \forall \, g \in C^{2+\beta}(\mathbf{R}^d),$$

$$\left| \mathbf{E}\left[\int_0^T f(Y_{\tau_{i_s}}) \, ds \right] - \mathbf{E}\left[\int_0^T f(X_s) \, ds \right] \right| \leqslant C|f|_\beta \kappa(\delta, \beta) \qquad \forall \, f \in C^\beta(\mathbf{R}^d),$$

where

$$\kappa(\delta, \beta) = \begin{cases} \delta^{\beta/2}, & \beta < 2, \\ \delta(1 + |\ln \delta|), & \beta = 2, \\ \delta, & \beta > 2. \end{cases}$$

Similarly, Theorem 1 can also be applied directly to other representative cases, for instance, $\mu = \alpha$ and that of heavy tails.

Remark 2. The same rate of convergence is obtained in Theorem 1, while the assumptions are much weaker compared with a previous study [22]. For the special case of diffusion processes, Corollary 1 improves the rate of convergence obtained previously [13].

For the case when $\beta \notin \mathbf{N}$ and $\beta \neq \alpha$, the result was proved in [18]. This paper focuses on what is left.

3. Proof. To prove Theorem 1, standard techniques such as stochastic flows cannot be applied due to the lack of regularity. Instead, the solution to the corresponding backward Kolmogorov equation is used. The operators of the equation associated with X_t are defined as follows.

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

$$\mathcal{A}_z u(t,x) = \mathbf{1}_{\{\alpha=1\}} \big(a(z), \nabla_x u(t,x) \big) + \mathbf{1}_{\{\alpha=2\}} \frac{1}{2} \sum_{i,j=1}^d B^{ij}(z) \partial_{ij}^2 u(t,x)$$

$$+ \mathbf{1}_{\{\alpha \in (0,2)\}} \int \big[u\big(t, x + b(z)y\big) - u(t,x)$$

$$- \chi^\alpha(y) \big(\nabla u(t,x), b(z)y \big) \big] \frac{dy}{|y|^{d+\alpha}},$$

$$\mathcal{A}u(t,x) = \mathcal{A}_x u(t,x) = \mathcal{A}_z u(t,x) \big|_{z=x},$$

with $\chi^\alpha(y) = \mathbf{1}_{\{\alpha=1\}}\mathbf{1}_{\{|y|\leqslant 1\}} + \mathbf{1}_{\{\alpha\in(1,2)\}}$, $B = b^*b$, and

$$\mathcal{B}_z u(t,x) = \mathbf{1}_{\{\alpha\in(1,2]\}}\big(a(z), \nabla_x u(t,x)\big)$$
$$+ \int_{\mathbf{R}_0^m} \big[u\big(t, x + c(z)y\big) - u(t,x)$$
$$- \mathbf{1}_{\{\alpha\in(1,2]\}}\mathbf{1}_{\{|y|\leqslant 1\}}\big(\nabla_x u(t,x), c(z)y\big)\big]\,\pi(dy),$$
$$\mathcal{B}u(t,x) = \mathcal{B}_x u(t,x) = \mathcal{B}_z u(t,x)\big|_{z=x}.$$

The operator $\mathcal{L} = \mathcal{A} + \mathcal{B}$ is the generator of X_t defined in (1); $\mathcal{A}$ is the principal part and $\mathcal{B}$ is the lower-order or subordinated part. The corresponding operator for $u \in C^{\alpha+\beta}(\mathbf{R}^d)$ is defined similarly.

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

$$u(X_t) - \int_0^t (\mathcal{A} + \mathcal{B})u(X_s)\,ds \qquad \forall\,u \in C^{\alpha+\beta}(\mathbf{R}^d)$$

is a martingale [14].

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

$$(\partial_t + \mathcal{A} + \mathcal{B})v(t,x) = 0, \qquad 0 \leqslant t \leqslant T,$$
$$v(T,x) = g(x),$$

then by Itô's formula,

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

The regularity of v determines a one-step estimate and the rate of convergence of the approximation. For $\beta \in (0,1)$, the result for the Kolmogorov equation in Hölder classes is available (see [15], [16]). In a standard way it can be extended to the case $\beta \geqslant 1$. The main difficulty is to derive the one-step estimate stated in Lemma 3.

3.1. Backward Kolmogorov equation. For the one-step estimate, consider in Hölder–Zygmund spaces the backward Kolmogorov equation associated with X_t,

(4)
$$\big(\partial_t + \mathcal{A} + \mathcal{B}\big)v(t,x) = f(t,x),$$
$$v(T,x) = 0.$$

Given a measurable and bounded function f on $\mathbf{H}$, $u \in C^{\alpha+\beta}(\mathbf{H})$ is said to be a solution to (4) if

$$u(t,x) = \int_t^T [\mathcal{L}u(s,x) - f(s,x)]\,ds \qquad \forall\,(t,x) \in \mathbf{H}.$$

THEOREM 2. *Let $\beta \in (0,3)$ and $\mu \in [\beta, \alpha+\beta)$. Assume $a^i, b^{ij} \in \widetilde{C}^\beta(\mathbf{R}^d)$, $1 \leqslant i,j \leqslant d$, $c^{ij} \in \widetilde{C}^{\beta/(\mu\wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant m$, $\inf_x |\det b(x)| > 0$, and*

$$\int_{|y|\leqslant 1} |y|^\alpha\,\pi(dy) + \int_{|y|>1} |y|^\mu\,\pi(dy) < \infty.$$

Then, for each $f \in C^\beta(\mathbf{H})$, there exist a unique solution $v \in C^{\alpha+\beta}(\mathbf{H})$ to (4) and a constant C independent of f such that $|u|_{\alpha+\beta} \leqslant C|f|_\beta$.

Corollary 2 follows directly from Theorem 2.

COROLLARY 2. *Let $\beta \in (0,3)$ and $\mu \in [\beta, \alpha + \beta)$. Assume $a^i, b^{ij} \in \widetilde{C}^\beta(\mathbf{R}^d)$, $1 \leqslant i, j \leqslant d$, $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant m$, $\inf_x |\det b(x)| > 0$, and*

$$\int_{|y| \leqslant 1} |y|^\alpha \, \pi(dy) + \int_{|y| > 1} |y|^\mu \, \pi(dy) < \infty.$$

Then for each $f \in C^\beta(\mathbf{R}^d)$ and $g \in C^{\alpha+\beta}(\mathbf{R}^d)$, there exist a unique solution $v \in C^{\alpha+\beta}(\mathbf{H})$ to the Cauchy problem

$$(5) \qquad \begin{aligned} \big(\partial_t + \mathcal{A} + \mathcal{B}\big)v(t,x) &= f(x), \\ v(T,x) &= g(x) \end{aligned}$$

and a constant C independent of f and g such that $|v|_{\alpha+\beta} \leqslant C(|f|_\beta + |g|_{\alpha+\beta})$.

For the proof of Theorem 2 and Corollary 2, the equation with constant coefficients is solved first. The result is then extended to (4) by handling variable coefficients using partition of unity and deriving a priori Schauder estimates in Hölder–Zygmund spaces.

3.1.1. Kolmogorov equation with constant coefficients. By changing the variable of integration, the principal operator $\mathcal{A}$ can be rewritten as

$$\begin{aligned} \mathcal{A}_z u(t,x) = \mathbf{1}_{\{\alpha=1\}}\big(a(z), \nabla_x u(t,x)\big) + \mathbf{1}_{\{\alpha=2\}} \sum_{i,j=1}^d B^{ij}(z)\partial_{ij}^2 u(t,x) \\ + \mathbf{1}_{\{\alpha \in (0,2)\}} \int \big[u(t, x+y) - u(t,x) \\ - \chi^\alpha(y)\big(\nabla u(t,x), y\big)\big] m(z,y) \frac{dy}{|y|^{d+\alpha}}, \end{aligned}$$

where $B = b^* b$ and

$$(6) \qquad m(z,y) = \frac{1}{|\det b(z)|} \frac{1}{|b(z)^{-1}y/|y| \, |^{d+\alpha}}, \qquad \alpha \in (0,2).$$

It holds that

$$(7) \qquad \int_{S^{d-1}} y m(\,\cdot\,, y)\, \mu_{d-1}(dy) = 0.$$

Here S^{d-1} is the unit sphere in $\mathbf{R}^d$ and μ_{d-1} is the Lebesgue measure.

For various estimates, the representation of difference stated in Lemma 1, which is Lemma 2.1 in [12], and the result given in Lemma 2 need to be employed.

For $u \in C^{\alpha+\beta}(\mathbf{R}^d)$, $\alpha \in (0,2)$, define the fractional Laplacian

$$\partial^\alpha u(x) = \int \big[u(x+y) - u(x) - \chi^\alpha(y)\big(\nabla u(x), y\big)\big] \frac{dy}{|y|^{d+\alpha}}, \qquad x \in \mathbf{R}^d.$$

LEMMA 1. *Let $\delta \in (0,1)$ and $u \in C_0^\infty(\mathbf{R}^d)$. Denote*

$$k^\delta(y,z) = |z+y|^{-d+\delta} - |z|^{-d+\delta}.$$

Then there exist constants C and $K = K(\delta, d)$ such that

$$\int |k^\delta(y,z)|\, dz \leqslant C|y|^\delta \qquad \forall\, y \in \mathbf{R}^d$$

and

$$u(x+y) - u(x) = K \int k^\delta(y,z)\, \partial^\delta u(x-z)\, dz.$$

By taking pointwise limit and applying the dominated convergence theorem, the result of Lemma 1 can be extended to $u \in C^\delta(\mathbf{R}^d)$.

Given a measurable and bounded function $m(y)$ on $\mathbf{R}^d$, define

$$L^m u(x) = \int_{\mathbf{R}^d} \big[u(x+y) - u(x) - \chi^\alpha(y)\big(\nabla u(x), y\big)\big] m(y)\, \frac{dy}{|y|^{d+\alpha}}, \qquad u \in C^{\alpha+\beta}.$$

The following result holds.

LEMMA 2. *Let $\alpha \in (0,2)$, $\beta > 0$, $u \in C^{\alpha+\beta}(\mathbf{R}^d)$, and $|m| \leqslant K$. Assume, for $\alpha = 1$,*

$$\fint_{r < |y| \leqslant 1} y m(y)\, \frac{dy}{|y|^{d+1}} = 0 \qquad \forall\, r \in (0,1).$$

Then there exists a constant C independent of u such that

$$|\mathcal{L}^m u|_\beta \leqslant CK |u|_{\alpha+\beta}.$$

The result was proved for $\beta \in (0,1]$ (see [17]). If $\beta > 1$, for any multi-index $|\gamma| = [\beta]$, $\partial^\gamma u \in C^{\alpha+\beta-[\beta]}$ and

$$|\partial^\gamma(\mathcal{L}^m u)|_{\beta-[\beta]^-} = |\mathcal{L}^m(\partial^\gamma u)|_{\beta-[\beta]^-} \leqslant CK |\partial^\gamma u|_{\alpha+\beta-[\beta]^-}.$$

The statement then follows.

Denote $\mathcal{A}^0 u(x) = \mathcal{A}_{z_0} u(x)$ for $z_0 \in \mathbf{R}^d$ and consider a backward Kolmogorov equation with constant coefficients

$$(8) \qquad \begin{aligned} \big(\partial_t + \mathcal{A}^0 - \lambda\big) v(t,x) &= f(x), \\ v(T,x) &= 0, \end{aligned}$$

where $\lambda \geqslant 0$.

PROPOSITION 1. *Let $\beta > 0$, $f \in C^\beta(\mathbf{R}^d)$, and let K_1 and K_2 be two constants with $|\det b(z)| \geqslant K_1$ and $\mathbf{1}_{\{\alpha=1\}}|a(z)| + |b(z)| \leqslant K_2$ for any $z \in \mathbf{R}^d$. Then there exist a unique solution $u \in C^{\alpha+\beta}(\mathbf{H})$ to (8) and constants C (which may be different) such that*

$$|u|_{\alpha+\beta} \leqslant C(\alpha, \beta, T, d, K_1, K_2)|f|_\beta,$$
$$|u|_\beta \leqslant C(\alpha, d)(\lambda^{-1} \wedge T)|f|_\beta,$$
$$|u(t, \cdot) - u(s, \cdot)|_{\alpha/2+\beta} \leqslant C(t-s)^{1/2}|f|_\beta \qquad \forall\, s \leqslant t \leqslant T.$$

Proposition 1 was proved for $\beta \notin \mathbf{N}$ [18]. For $\beta \in \mathbf{N}$, the result follows by exactly the same steps.

3.1.2. Kolmogorov equation with variable coefficients. For Schauder estimates in the case of variable coefficients, it is essential to derive an estimate for the lower-order operator, as stated in Proposition 2.

PROPOSITION 2. *Let $\beta \in (0,3)$ and $\mu \in [\beta, \alpha + \beta)$. Assume $a \in \widetilde{C}^{\beta}(\mathbf{R}^d)$, $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, and*

$$\int_{|y| \leqslant 1} |y|^{\alpha} \, \pi(dy) + \int_{|y|>1} |y|^{\mu} \, \pi(dy) < \infty.$$

Then for each $\varepsilon > 0$, there exists a constant $C(\varepsilon)$ such that

$$|\mathcal{B}f|_{\beta} \leqslant \varepsilon |f|_{\alpha+\beta} + C(\varepsilon)|f|_0, \qquad f \in C^{\alpha+\beta}(\mathbf{R}^d).$$

For $\beta \notin \mathbf{N}$, Proposition 2 was already proved [18]. For $\beta \in \mathbf{N}$, it can be verified by following similar steps.

With the estimate of $\mathcal{B}f$, Theorem 2 is proved in a standard way by using partition of unity and the estimates for constant coefficients, which allow us to obtain a priori estimates. The continuation by parameter method is then applied to transfer from constant to variable coefficients. Corollary 2 follows naturally from Theorem 2. They are proved in a way similar to that in [18]. The details are thus omitted here.

Remark 4. If the assumptions of Corollary 2 hold and $v \in C^{\alpha+\beta}(\mathbf{H})$ is the solution to (5), then $\partial_t v = f - \mathcal{A}_x v - \mathcal{B}_x v$, and by Lemma 2 and Proposition 2, $|\partial_t v|_{\beta} \leqslant C(|g|_{\alpha+\beta} + |f|_{\beta})$.

3.2. One-step estimate. To determine the rate of convergence, a key step is to estimate the conditional expectation of each increment of the Euler approximation, which is provided in Lemma 3.

LEMMA 3. *Let $Y = \{Y_t\}_{t \in (0,T]}$ be the weak Euler approximation with step size $\delta \in (0,1)$ of the stochastic process $X = \{X_t\}_{t \in (0,T]}$ defined by (1). For $\alpha \in (0,2]$, $\beta \in (0,3)$, and $\mu \in [\beta, \alpha + \beta)$, assume $a^i, b^{ij} \in \widetilde{C}^{\beta}(\mathbf{R}^d)$, $1 \leqslant i,j \leqslant d$, $c^{ij} \in \widetilde{C}^{\beta/(\mu \wedge 1)}(\mathbf{R}^d)$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant m$, $\inf_x |\det b(x)| > 0$, and*

$$\int_{|y| \leqslant 1} |y|^{\alpha} \, \pi(dy) + \int_{|y|>1} |y|^{\mu} \, \pi(dy) < \infty.$$

Then there exists a constant C such that

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

where $i_s = i$ if $\tau_i \leqslant s < \tau_{i+1}$ and $\kappa(\delta, \alpha, \beta)$ is defined as in Theorem 1.

The proof of Lemma 3 is based on applying Itô's formula to $f(Y_s) - f(Y_{\tau_{i_s}})$, $f \in C^{\beta}(\mathbf{R}^d)$.

If $\beta \leqslant \alpha$, f is first smoothed by using $w \in C_0^{\infty}(\mathbf{R}^d)$, a nonnegative smooth function with support on $\{|x| \leqslant 1\}$ such that $w(x) = w(|x|)$, $x \in \mathbf{R}^d$, and $\int w(x) \, dx = 1$. Thanks to the symmetry,

$$(9) \qquad \int_{\mathbf{R}^d} x^i w(x) \, dx = 0, \qquad i = 1, \dots, d.$$

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

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

If $\beta > \alpha$, the inequality is clear from Remark 3 and Itô's formula.

In the following proof of Lemma 3, we will call Lemma 6 and Corollary 3, which provide, respectively, the estimates for $\mathcal{A}_z f^\varepsilon$ and $\mathcal{B}_z f^\varepsilon$.

Proof. If $\beta \leqslant \alpha$, by Remark 3, applying Itô's formula to f^ε yields

$$\mathbf{E}[f^\varepsilon(Y_s) - f^\varepsilon(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]$$
$$= \mathbf{E}\left[\int_{\tau_{i_s}}^s \left(\mathcal{A}_{Y_{\tau_{i_s}}} f^\varepsilon(Y_r) + \mathcal{B}_{Y_{\tau_{i_s}}} f^\varepsilon(Y_r)\right) dr \ \middle| \ \mathcal{F}_{\tau_{i_s}}\right], \qquad s \in [0, T].$$

By Lemma 6 and Corollary 3, for $\varepsilon \in (0, 1)$,

$$|\mathbf{E}[f(Y_s) - f(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]| \leqslant |\mathbf{E}[(f - f^\varepsilon)(Y_s) - (f - f^\varepsilon)(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]|$$
$$+ |\mathbf{E}[f^\varepsilon(Y_s) - f^\varepsilon(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]| \leqslant CK(\varepsilon, \delta, \alpha, \beta)|f|_\beta,$$

where C is a constant independent of ε, f and

$$K(\varepsilon, \delta, \alpha, \beta) = \begin{cases} \varepsilon^\beta + \delta\varepsilon^{-\alpha+\beta}, & \beta < \alpha, \\ \varepsilon^\beta + \delta(1 - \ln\varepsilon), & \beta = \alpha < 2, \\ (\varepsilon^2 + \delta)(1 - \ln\varepsilon), & \beta = \alpha = 2. \end{cases}$$

Minimizing $K(\varepsilon, \delta, \alpha, \beta)$ with respect to $\varepsilon \in (0, 1)$ gives

$$|\mathbf{E}[f(Y_s) - f(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]| \leqslant C\kappa(\delta, \alpha, \beta)|f|_\beta,$$

where $\kappa(\delta, \alpha, \beta)$ is defined as in Theorem 1.

If $\beta > \alpha$, by Remark 3, a direct application of Itô's formula yields

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

Lemmas 2 and 6 then imply

$$|\mathbf{E}[f(Y_s) - f(Y_{\tau_{i_s}}) \mid \mathcal{F}_{\tau_{i_s}}]| \leqslant C\delta|f|_\beta.$$

The conclusion of Lemma 3 follows.

The integral estimates provided in Lemma 4 will be useful for the estimate of $\mathcal{A}_z f^\varepsilon$. Recall that $m(z, y)$ defined by (6) in operator $\mathcal{A}_z$ is bounded, smooth, and 0-homogeneous and symmetric with respect to y.

LEMMA 4. *Let $v \in C_0^\infty(\mathbf{R}^d)$. Then*
(i) *for $\alpha \in (0, 2)$,*

$$\int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} |v(y + \widetilde{y}) - v(y) - \chi^\alpha(\widetilde{y})(\nabla v(y), \widetilde{y})| \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}} \, dy < \infty,$$

where $\chi^\alpha(y) = \mathbf{1}_{\{\alpha=1\}} \mathbf{1}_{\{|y|\leqslant 1\}} + \mathbf{1}_{\{\alpha\in(1,2)\}}$;
(ii) *for $\beta \in (0, 1]$, if $\beta < \alpha$,*

$$\sup_z \int_{\mathbf{R}^d} |(\mathcal{A}_z v)(y)| |y|^\beta \, dy < \infty,$$

and if $\beta = \alpha$,

$$\sup_z \int_{\mathbf{R}^d} |(\mathcal{A}_z v)(y)|(|y|^\alpha \wedge k) \, dy \leqslant C(1 + \ln k) \qquad \forall \, k > 1;$$

(iii) *for $\beta \in (1, 2)$, if $\beta < \alpha < 2$,*

$$\int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} \int_0^1 |v(y + s\widetilde{y}) - v(y)||y|^{\beta - 1} \, ds \, \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha - 1}} \, dy < \infty,$$

and if $\beta = \alpha < 2$,

$$\int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} \int_0^1 |v(y + s\widetilde{y}) - v(y)|(|y|^{\beta - 1} \wedge k) \, ds \, \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha - 1}} \, dy \leqslant C(1 + \ln k) \quad \forall \, k > 1.$$

Proof. (i) Clearly,

$$|v(y + \widetilde{y}) - v(y) - \chi^\alpha(\widetilde{y})(\nabla v(y), \widetilde{y})|$$

$$\leqslant \mathbf{1}_{\{|\widetilde{y}| \leqslant 1\}} \left\{ \int_0^1 \left[\max_{i,j} |\partial_{ij}^2 v(y + s\widetilde{y})||\widetilde{y}|^2 + \mathbf{1}_{\{\alpha \in (0,1)\}} |\nabla v(y + s\widetilde{y})||\widetilde{y}| \right] ds \right\}$$

$$+ \mathbf{1}_{\{|\widetilde{y}| > 1\}} \left\{ |v(y + \widetilde{y})| + |v(y)| + \mathbf{1}_{\{\alpha \in (1,2)\}} |\nabla v(y)||\widetilde{y}| \right\}, \qquad y, \widetilde{y} \in \mathbf{R}^d.$$

The claim then follows.

(ii) For $\beta \in (0, 1)$, $\beta < \alpha$, $z \in \mathbf{R}^d$,

$$\int_{\mathbf{R}^d} |(\mathcal{A}_z v)(y)||y|^\beta \, dy \leqslant \int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})||y|^\beta \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$+ \int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y)||y|^\beta \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$+ \max_{i,j} \int_{\mathbf{R}^d} \int_{|\widetilde{y}| \leqslant 1} \int_0^1 |\partial_{ij}^2 v(y + s\widetilde{y})||\widetilde{y}|^2 |y|^\beta \, ds \, \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

and

$$\int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})||y|^\beta \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$\leqslant C \left[\int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})||y + \widetilde{y}|^\beta \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy \right.$$

$$\left. + \int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})||\widetilde{y}|^\beta \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy \right].$$

For $\beta \in (0, 1]$, $\beta = \alpha$, assume $v(x) = 0$ if $|x| > R$. Then for $k > 1$ with $A = (R + 1)^{1/\alpha}$,

$$\int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})|(|y|^\alpha \wedge k) \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$\leqslant \int_{\mathbf{R}^d} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})|(|y|^\alpha \wedge Ak) \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$= \int_{|y| \leqslant (R+1)k^{1/\alpha}} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})||y|^\alpha \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy$$

$$+ k \int_{|y| > (R+1)k^{1/\alpha}} \int_{|\widetilde{y}| > 1} |v(y + \widetilde{y})| \frac{d\widetilde{y}}{|\widetilde{y}|^{d + \alpha}} \, dy = A_1 + A_2,$$

where

$$|A_1| \leqslant \int \int_{1 \leqslant |\widetilde{y}| \leqslant (R+1)(1+k^{1/\alpha})} |v(y+\widetilde{y})| \frac{d\widetilde{y}}{|\widetilde{y}|^d} \, dy \leqslant C(1+\ln k)$$

and

$$|A_2| \leqslant k \int \int_{|\widetilde{y}| \geqslant k^{1/\alpha}} |v(y+\widetilde{y})| \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}} \, dy \leqslant Ckk^{-1} = C,$$

since for $|y+\widetilde{y}| \leqslant R$ and $|y| > (R+1)k^{1/\alpha}$, $|\widetilde{y}| \geqslant (R+1)k^{1/\alpha} - R \geqslant k^{1/\alpha}$.

Then

$$\int \int_{|\widetilde{y}| \leqslant 1} |v(y+\widetilde{y}) - v(y) - \mathbf{1}_{\{\alpha=1\}}(\nabla v(y), \widetilde{y})| (|y|^\alpha \wedge k) \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}} \, dy$$

$$\leqslant \int \int_{|\widetilde{y}| \leqslant 1} \int_0^1 |\nabla v(y+s\widetilde{y}) - \mathbf{1}_{\{\alpha=1\}} \nabla v(y)| |y|^\alpha \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy < \infty.$$

Assertion (ii) follows.

(iii) For $1 < \beta < \alpha < 2$,

$$\int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} \int_0^1 |v(y+s\widetilde{y}) - v(y)| |y|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy$$

$$\leqslant \int_{\mathbf{R}^d} \int_{|\widetilde{y}|>1} \int_0^1 |v(y+s\widetilde{y})| |y|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy$$

$$+ \int_{\mathbf{R}^d} \int_{|\widetilde{y}|>1} \int_0^1 |v(y)| |y|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy$$

$$+ \int_{\mathbf{R}^d} \int_{|\widetilde{y}| \leqslant 1} \int_0^1 \int_0^1 |\nabla v(y+s\tau\widetilde{y})| |y|^{\beta-1} \, d\tau \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-2}} \, dy.$$

Also,

$$\int_{\mathbf{R}^d} \int_{|\widetilde{y}|>1} \int_0^1 |v(y+s\widetilde{y})| |y|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy$$

$$\leqslant C \left[\int_{\mathbf{R}^d} \int_{|\widetilde{y}|>1} \int_0^1 |v(y+s\widetilde{y})| |y+s\widetilde{y}|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy \right.$$

$$\left. + \int_{\mathbf{R}^d} \int_{|\widetilde{y}|>1} \int_0^1 |v(y+s\widetilde{y})| |\widetilde{y}|^{\beta-1} \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}} \, dy \right]$$

and

$$\int_{\mathbf{R}^d} \int_{|\widetilde{y}| \leqslant 1} \int_0^1 \int_0^1 |\nabla v(y+s\tau\widetilde{y})| |y|^{\beta-1} d\tau \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-2}} \, dy$$

$$\leqslant C \left[\int_{\mathbf{R}^d} \int_{|\widetilde{y}| \leqslant 1} \int_0^1 \int_0^1 |\nabla v(y+s\tau\widetilde{y})| |y+s\tau\widetilde{y}|^{\beta-1} \, d\tau \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-2}} \, dy \right.$$

$$\left. + \int_{\mathbf{R}^d} \int_{|\widetilde{y}| \leqslant 1} \int_0^1 \int_0^1 |\nabla v(y+s\tau\widetilde{y})| |\widetilde{y}|^{\beta-1} \, d\tau \, ds \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-2}} \, dy \right].$$

If $1 < \beta = \alpha < 2$, the same argument as in assertion (ii) applies.

Assertion (iii) follows. Lemma 4 is proved.

Remark 5. The estimate in assertion (iii) implies that (ii) can be extended to all $\beta \in (0, \alpha]$ with $\alpha < 2$.

Lemma 5 states the modulus of continuity estimate of a function $f \in C^1(\mathbf{R}^d)$ [5], [6]. The proof follows that of Lemma 5.6 in [5].

LEMMA 5. *Let $f \in C^1(\mathbf{R}^d)$ and $[f]_1 \leqslant K$. Then there exists a constant C such that for all $x, h \in \mathbf{R}^d$, $h \neq 0$,*

$$|f(x + h) - f(x)| \leqslant C|h|(1 + |\ln |h||)|f|_1.$$

Proof. Fix $x, h \in \mathbf{R}^d$, $h \neq 0$ with $0 < |h| < 1/2$, let k be a positive integer such that

$$2^{-k-1} \leqslant |h| < 2^{-k},$$

and set $\tau_0 = 2^k h$. Then $2^{-k} \leqslant 2|h|$, $\ln |h| < -k \ln 2$ or $k < -\ln |h|/\ln 2$, and $1/2 \leqslant \tau_0 < 1$. Define for $\tau \in \mathbf{R}^d$

$$v(\tau) = f(x + \tau) - f(x).$$

Then

$$\left| v(\tau) - 2v\left(\frac{\tau}{2}\right) \right| = \left| f(x + \tau) - 2f\left(x + \frac{\tau}{2}\right) + f(x) \right| \leqslant [f]_1 \frac{|\tau|}{2}.$$

Hence

$$\left| 2^{j-1} v\left(\frac{\tau_0}{2^{j-1}}\right) - 2^j v\left(\frac{\tau_0}{2^j}\right) \right| \leqslant [f]_1 2^{j-1} \frac{|\tau_0|}{2^j} = 2^{-1}[f]_1|\tau_0|$$

and

$$|v(\tau_0) - 2^k v(h)| = \left| \sum_{j=1}^{k} \left(2^{j-1} v\left(\frac{\tau_0}{2^{j-1}}\right) - 2^j v\left(\frac{\tau_0}{2^j}\right) \right) \right|$$

$$\leqslant \sum_{j=1}^{k} \left| 2^{j-1} v\left(\frac{\tau_0}{2^{j-1}}\right) - 2^j v\left(\frac{\tau_0}{2^j}\right) \right| \leqslant k 2^{-1}|\tau_0|[f]_1.$$

Since $|v(\tau_0)| \leqslant 2|f|_0$ or $|v(\tau_0)| \leqslant [f]_{1/2}$, then

$$|v(h)| \leqslant 2^{-k}|2^k v(h) - v(\tau_0)| + 2^{-k}|v(\tau_0)| \leqslant [f]_1 k 2^{-1-k}|\tau_0| + 2 \cdot 2^{-k}|f|_0$$
$$\leqslant [f]_1 k|h| + 4|h||f|_0 \leqslant C|f|_1|h|(1 - \ln |h|).$$

The statement follows.

In what follows, some estimates for Af^ε and $\mathcal{B}f^\varepsilon$ are proved.

LEMMA 6. *Let $f \in C^\beta(\mathbf{R}^d)$ and $\varepsilon \in (0, 1)$. The following statements hold.*
(i) *For $\beta \in (0, 2]$, there exists a constant C such that*

$$|f^\varepsilon(x) - f(x)| \leqslant C|f|_\beta K(\varepsilon, \beta) \qquad \forall x \in \mathbf{R}^d,$$

where $K(\varepsilon, \beta) = \varepsilon^\beta$ if $\beta < 2$ and $K(\varepsilon, 2) = \varepsilon^2(1 - \ln \varepsilon)$.

(ii) *For $\alpha \in (0, 2)$, there exists a constant C such that*

$$(10) \qquad |\mathcal{A}_z f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta) \qquad \forall\, z, x \in \mathbf{R}^d,$$

where $K(\varepsilon, \alpha, \beta) = \varepsilon^{-\alpha+\beta}$ if $\beta < \alpha$ and $K(\varepsilon, \alpha, \beta) = 1 - \ln \varepsilon$ if $\beta = \alpha$. In particular,

$$(11) \qquad |\partial^\alpha f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta) \qquad \forall\, x \in \mathbf{R}^d.$$

(iii) *For $\beta \in (0, 2]$, there exists a constant C such that*

$$(12) \qquad \begin{aligned} |\partial_k f^\varepsilon(x)| &\leqslant C|f|_\beta K(\varepsilon, 1, \beta) && \forall\, x \in \mathbf{R}^d, \quad k = 1, \ldots, d, \ \text{if } \beta \leqslant 1, \\ |\partial^2_{kl} f^\varepsilon(x)| &\leqslant C|f|_\beta K(\varepsilon, 2, \beta) && \forall\, x \in \mathbf{R}^d, \quad k, l = 1, \ldots, d, \\ |f^\varepsilon|_1 &\leqslant C|f|_1; \end{aligned}$$

for $\alpha \in [1, 2)$, $\beta \in (0, 1]$,

$$(13) \qquad |f^\varepsilon|_\alpha \leqslant C|f|_\beta \varepsilon^{-\alpha+\beta};$$

and for $\alpha \in (1, 2)$, $\beta \in (0, \alpha]$, $\beta \neq \alpha - 1$,

$$(14) \qquad |\partial^{\alpha-1} \nabla f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta) \qquad \forall\, x \in \mathbf{R}^d.$$

Proof. (i) For $\beta \in (0, 1]$,

$$f^\varepsilon(x) - f(x) = \int [f(x - y) - f(x)] w^\varepsilon(y)\, dy = \int [f(x + y) - f(x)] w^\varepsilon(y)\, dy$$

and

$$f^\varepsilon(x) - f(x) = \frac{1}{2} \int [f(x + y) + f(x - y) - 2f(x)] w^\varepsilon(y)\, dy.$$

Hence, $|f^\varepsilon(x) - f(x)| \leqslant C|f|_\beta \varepsilon^\beta$.

For $\beta \in (1, 2)$, by (9),

$$\begin{aligned} f^\varepsilon(x) - f(x) &= \int [f(x - y) - f(x)] w^\varepsilon(y)\, dy \\ &= \int \big[f(x + y) - f(x) - \big(\nabla f(x), y\big) \big] w^\varepsilon(y)\, dy \\ &= \int \int_0^1 \big(\nabla f(x + sy) - \nabla f(x), y\big)\, ds\, w^\varepsilon(y)\, dy \end{aligned}$$

and

$$|f^\varepsilon(x) - f(x)| \leqslant C|\nabla f|_{\beta-1} \int |y|^{1+(\beta-1)} w^\varepsilon(y)\, dy \leqslant C|f|_\beta \varepsilon^\beta.$$

For $\beta = 2$, by Lemma 5,

$$|f^\varepsilon(x) - f(x)| \leqslant C|\nabla f|_1 \int |y|^2 (1 + |\ln |y||) w^\varepsilon(y)\, dy \leqslant C|\nabla f|_1 \varepsilon^2 (1 - \ln \varepsilon).$$

(ii) Changing the variable of integration with $\overline{y} = y/\varepsilon$ and using (7) for $\alpha = 1$ gives

$$
\begin{aligned}
\mathcal{A}_z w^\varepsilon(x) &= \mathbf{1}_{\{\alpha=1\}}\big(a(z), \nabla w^\varepsilon(x)\big) \\
&\quad + \int \big[w^\varepsilon(x+y) - w^\varepsilon(x) - \chi^\alpha(y)\big(\nabla w^\varepsilon(x), y\big)\big] m(z,y) \frac{dy}{|y|^{d+\alpha}} \\
&= \varepsilon^{-\alpha}\varepsilon^{-d}(\mathcal{A}_z w)\left(\frac{x}{\varepsilon}\right) \qquad \forall\, z, x \in \mathbf{R}^d,
\end{aligned}
$$

where $\chi^\alpha(y) = \mathbf{1}_{\{\alpha=1\}}\mathbf{1}_{\{|y|\leqslant 1\}} + \mathbf{1}_{\{\alpha\in(1,2)\}}$. It then follows from Lemma 4(i), Fubini's theorem, and changing the variable of integration with $\overline{y} = y/\varepsilon$ that

$$
\begin{aligned}
\mathcal{A}_z f^\varepsilon(x) &= \int_{\mathbf{R}^d} \varepsilon^{-\alpha}\varepsilon^{-d}(\mathcal{A}_z w)\left(\frac{x-y}{\varepsilon}\right) f(y)\,dy \\
&= \int \varepsilon^{-\alpha}\varepsilon^{-d}(\mathcal{A}_z w)\left(\frac{y}{\varepsilon}\right) f(x-y)\,dy \\
&= \int \varepsilon^{-\alpha}(\mathcal{A}_z w)(y) f(x-\varepsilon y)\,dy \qquad \forall\, x, z \in \mathbf{R}^d.
\end{aligned}
$$

By Lemma 4(i) and Fubini's theorem,

$$
\int_{\mathbf{R}^d} \mathcal{A}_z w(y)\,dy = 0.
$$

Hence, if $\beta \in (0,1]$, $\beta \leqslant \alpha$, then

$$
\begin{aligned}
\mathcal{A}_z f^\varepsilon(x) &= \int \varepsilon^{-\alpha}(\mathcal{A}_z w)(y) f(x-\varepsilon y)\,dy \\
&= \frac{1}{2}\int \varepsilon^{-\alpha}(\mathcal{A}_z w)(y)[f(x-\varepsilon y) + f(x+\varepsilon y) - 2f(x)]\,dy,
\end{aligned}
$$

and by Lemma 4(ii)

$$
|\mathcal{A}_z f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta) \int_{\mathbf{R}^d} |(\mathcal{A}_z w)(y)|(|y|^\beta \wedge \varepsilon^{-\beta})\,dy \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta).
$$

If $\beta \in (1,2)$, $\beta \leqslant \alpha$, then by Theorem 2.27 in [7], switching differentiation and integration gives

$$
\begin{aligned}
\mathcal{A}_z w(y) &= \int \big[w(y+\widetilde{y}) - w(y) - \big(\nabla w(y), \widetilde{y}\big)\big] m(z, \widetilde{y}) \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}} \\
&= \int \int_0^1 \big(\nabla_y w(y+s\widetilde{y}) - \nabla_y w(y), \widetilde{y}\big)\,ds\, m(z, \widetilde{y}) \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}} \\
&= \sum_{i=1}^d \frac{\partial}{\partial y_i} \int \int_0^1 [w(y+s\widetilde{y}) - w(y)]\widetilde{y}_i\, ds\, m(z, \widetilde{y}) \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}}.
\end{aligned}
$$

Clearly,

$$
\int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} \int_0^1 |w(y+s\widetilde{y}) - w(y)|\,ds\, \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}}\,dy < \infty
$$

and

$$\int [w(y + s\widetilde{y}) - w(y)]\, dy = 0.$$

It then follows from integrating by parts and Fubini's theorem that, for $x, z \in \mathbf{R}^d$,

$$\mathcal{A}_z f^\varepsilon(x) = \int \varepsilon^{-\alpha} \mathcal{A}_z w(y) f(x - \varepsilon y)\, dy$$

$$= \varepsilon^{-\alpha+1} \int \int \int_0^1 [w(y + s\widetilde{y}) - w(y)]\big(\nabla f(x - \varepsilon y), \widetilde{y}\big) m(z, \widetilde{y})\, ds\, \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}}\, dy$$

$$= \varepsilon^{-\alpha+1} \int_{\mathbf{R}^d} \int_{\mathbf{R}_0^d} \int_0^1 [w(y + s\widetilde{y}) - w(y)]$$

$$\times \big(\nabla f(x - \varepsilon y) - \nabla f(x), \widetilde{y}\big) m(z, \widetilde{y})\, ds\, \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha}}\, dy.$$

Thus by Lemma 4(iii)

$$|\mathcal{A}_z f^\varepsilon(x)| \leqslant C\varepsilon^{-\alpha+1}\varepsilon^{\beta-1}|\nabla f|_{\beta-1} \int \int \int_0^1 |w(y + s\widetilde{y}) - w(y)|$$

$$\times (|y|^{\beta-1} \wedge \varepsilon^{-(\beta-1)})\, ds\, \frac{d\widetilde{y}}{|\widetilde{y}|^{d+\alpha-1}}\, dy \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta).$$

Relation (11) is obtained by taking $m = 1$.

(iii) If $\beta < 1$, by changing the variable of integration,

$$\partial_k f^\varepsilon(x) = \varepsilon^{-1} \int_{\mathbf{R}^d} \varepsilon^{-d} \partial_k w\left(\frac{x - y}{\varepsilon}\right) f(y)\, dy$$

$$= \varepsilon^{-1} \int_{\mathbf{R}^d} \varepsilon^{-d} \partial_k w\left(\frac{y}{\varepsilon}\right) f(x - y)\, dy$$

$$= \varepsilon^{-1} \int_{\mathbf{R}^d} \partial_k w(y)[f(x - \varepsilon y) - f(x)]\, dy.$$

If $\beta = 1$,

$$f^\varepsilon(x + h) + f^\varepsilon(x - h) - 2f^\varepsilon(x)$$

$$= \frac{1}{2} \int w_\varepsilon(y)[f(x - y + h) + f(x - y - h) - 2f(x - y)]\, dy$$

and $|f^\varepsilon|_1 \leqslant |f|_1$.

Since $\partial_{kl}^2 w(y) = \partial_{kl}^2 w(-y)$, $k, l = 1, \ldots, d, y \in \mathbf{R}^d$,

$$\partial_{kl}^2 f^\varepsilon(x) = \varepsilon^{-2} \int_{\mathbf{R}^d} \varepsilon^{-d} \partial_{kl}^2 w\left(\frac{x - y}{\varepsilon}\right) f(y)\, dy$$

$$= \varepsilon^{-2} \int_{\mathbf{R}^d} \varepsilon^{-d} \partial_{kl}^2 w\left(\frac{y}{\varepsilon}\right) f(x - y)\, dy$$

$$= \varepsilon^{-2} \int_{\mathbf{R}^d} \partial_{kl}^2 w(y)[f(x - \varepsilon y) - f(x)]\, dy$$

$$= \frac{1}{2}\varepsilon^{-2} \int_{\mathbf{R}^d} \partial_{kl}^2 w(y)[f(x + \varepsilon y) + f(x - \varepsilon y) - 2f(x)]\, dy,$$

then if $\beta \in (0,1]$, $|\partial_k f^\varepsilon(x)| \leqslant C\varepsilon^{-1+\beta}|f|_\beta$, and $|\partial_{kl}^2 f^\varepsilon(x)| \leqslant C\varepsilon^{-2+\beta}|f|_\beta$ for any $x \in \mathbf{R}^d$. Similarly, if $\beta \in (1,2]$,

$$\partial_k f^\varepsilon(x) = \int \varepsilon^{-d} w\left(\frac{y}{\varepsilon}\right) \partial_k f(x-y)\, dy = \int \varepsilon^{-d} w\left(\frac{x-y}{\varepsilon}\right) \partial_k f(y)\, dy$$

and

$$\partial_{kl}^2 f^\varepsilon(x) = \varepsilon^{-1} \int \varepsilon^{-d} \partial_l w\left(\frac{y}{\varepsilon}\right) \partial_k f(x-y)\, dy$$
$$= \varepsilon^{-1} \int \partial_l w(y)[\partial_k f(x-\varepsilon y) - \partial_k f(x)]\, dy.$$

Hence, by Lemma 5, $|\partial_{kl}^2 f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, 2, \beta)$.

Formula (13) is obtained by applying (12) and the interpolation theorem. For $\beta \in (0,1]$, consider an operator on C^β defined by $T^\varepsilon(f) = f^\varepsilon$. By (12),

$$|T^\varepsilon(f)|_k \leqslant C\varepsilon^{-k+\beta}|f|_\beta, \qquad k = 1,2, \quad f \in C^\beta(\mathbf{R}^d).$$

Hence, $T^\varepsilon \colon C^\beta(\mathbf{R}^d) \to C^k(\mathbf{R}^d)$, $k = 1,2$, is bounded. By Theorem 6.4.5 in [3], $T^\varepsilon \colon C^\beta(\mathbf{R}^d) \to C^\alpha(\mathbf{R}^d)$ is bounded and

$$|T^\varepsilon(f)|_\alpha \leqslant C\varepsilon^{(-1+\beta)(2-\alpha)} \varepsilon^{(-2+\beta)(\alpha-1)}|f|_\beta$$
$$= C|f|_\beta K(\varepsilon, \alpha, \beta), \qquad f \in C^\beta(\mathbf{R}^d).$$

Depending on the value of β, (14) is proved separately for three cases.

(a) If $\beta \in (0,1]$, $\alpha \in (1,2)$, $\beta < \alpha - 1$, then

$$\nabla f^\varepsilon = \varepsilon^{-1}\varepsilon^{-d} \int \nabla w\left(\frac{y}{\varepsilon}\right) f(x-y)\, dy = \varepsilon^{-1}\varepsilon^{-d} \int \nabla w\left(\frac{y}{\varepsilon}\right) f(x-y)\, dy$$

and

$$\partial^{\alpha-1}\nabla f^\varepsilon = \varepsilon^{-1-(\alpha-1)} \int \partial^{\alpha-1}(\nabla w)(y) f(x-\varepsilon y)\, dy.$$

As in the proof of assertion (i) of Lemma 6, by Lemma 4,

$$|\partial^{\alpha-1}\nabla f^\varepsilon(x)| \leqslant C|f|_\beta \varepsilon^{-\alpha+\beta}.$$

(b) If $\beta \in (0,1], \alpha \in (1,2), \beta > \alpha - 1$, then

$$\partial^{\alpha-1}\nabla f^\varepsilon = \varepsilon^{-1} \int \nabla w(y) \partial^{\alpha-1} f(x-\varepsilon y)\, dy$$
$$= \varepsilon^{-1} \int \nabla w(y)[\partial^{\alpha-1} f(x-\varepsilon y) - \partial^{\alpha-1} f(x)]\, dy$$

and

$$|\partial^{\alpha-1}\nabla f^\varepsilon| \leqslant C\varepsilon^{-1+\beta-\alpha+1}|\partial^{\alpha-1} f|_{\beta-\alpha+1} \leqslant C|f|_\beta \varepsilon^{-\alpha+\beta}.$$

(c) If $\beta \in (1,\alpha]$, then $\partial^{\alpha-1}\nabla f^\varepsilon = \partial^{\alpha-1}(\nabla f)^\varepsilon$ and by (11),

$$|\partial^{\alpha-1}\nabla f^\varepsilon(x)| = |\partial^{\alpha-1}(\nabla f)^\varepsilon(x)| \leqslant CK(\varepsilon, \alpha-1, \beta-1)|\nabla f|_{\beta-1}.$$

The statement thus follows.

COROLLARY 3. *Let* $\alpha \in (0, 2]$, $\beta \leqslant \alpha$, *and* $\varepsilon \in (0, 1)$. *Assume* $a(x)$ *is bounded and*

$$\int (|y|^\alpha \wedge 1)\, \pi(dy) < \infty.$$

Then there exists a constant C *such that*

$$|\mathcal{B}_z f^\varepsilon(x)| \leqslant C|f|_\beta K(\varepsilon, \alpha, \beta) \quad \forall\, z, x \in \mathbf{R}^d, \quad \forall\, f \in C^\beta(\mathbf{R}^d),$$

where $K(\varepsilon, \alpha, \beta) = \varepsilon^{-\alpha+\beta}$ *if* $\beta < \alpha$ *and* $K(\varepsilon, \alpha, \beta) = 1 - \ln \varepsilon$ *if* $\beta = \alpha$.

Proof. If $\alpha \in (0, 1)$, then by Lemma 1,

$$f^\varepsilon(x + y) - f^\varepsilon(x) = \int k^\alpha(y, \widetilde{y}) \partial^\alpha f^\varepsilon(x - \widetilde{y})\, d\widetilde{y}.$$

Then by Lemma 6,

$$|f^\varepsilon(x + y) - f^\varepsilon(x)| \leqslant C|f|_\beta (|y|^\alpha \wedge 1) K(\varepsilon, \alpha, \beta), \qquad x, y \in \mathbf{R}^d,$$

and

$$|f^\varepsilon\big(x + c(x)y\big) - f^\varepsilon(x)| \leqslant C|f|_\beta (|c(x)y|^\alpha \wedge 1) K(\varepsilon, \alpha, \beta)$$
$$\leqslant C|f|_\beta \big[\mathbf{1}_{\{|y| \leqslant 1\}} |c(x)y|^\alpha + \mathbf{1}_{\{|y| > 1\}} (|c(x)y|^\alpha \wedge 1)\big] K(\varepsilon, \alpha, \beta).$$

If $\alpha = 1$, then, by Lemma 6(ii) and (12),

$$|f^\varepsilon(x + y) - f^\varepsilon(x)| \leqslant C \sup_x [|f(x)| + |\nabla f^\varepsilon(x)|](|y| \wedge 1)$$
$$\leqslant C|f|_\beta (|y| \wedge 1) K(\varepsilon, 1, \beta), \qquad x, y \in \mathbf{R}^d,$$

and

$$|f^\varepsilon\big(x + c(x)y\big) - f^\varepsilon(x)| \leqslant C|f|_\beta (|c(x)y| \wedge 1) K(\varepsilon, 1, \beta)$$
$$\leqslant C|f|_\beta [\mathbf{1}_{|y| \leqslant 1} |c(x)y| + \mathbf{1}_{|y| > 1} (|c(x)y| \wedge 1)] K(\varepsilon, 1, \beta).$$

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

$$f^\varepsilon(x + y) - f^\varepsilon(x) - (\nabla f^\varepsilon(x), y)$$
$$= \int_0^1 \big(\nabla f^\varepsilon(x + sy) - \nabla f^\varepsilon(x), y\big)\, ds \qquad \forall\, x, y \in \mathbf{R}^d.$$

For $\alpha \in (1, 2)$ and $\beta \neq \alpha - 1$, by Lemmas 1 and 6,

$$|f^\varepsilon(x + y) - f^\varepsilon(x) - (\nabla f^\varepsilon(x), y)| \leqslant C|f|_\beta |y|^\alpha K(\varepsilon, \alpha, \beta).$$

For $\alpha \in (1, 2)$ and $\beta = \alpha - 1$,

$$\nabla f^\varepsilon(x + \widetilde{y}) - \nabla f^\varepsilon(x) = \varepsilon^{-1} \varepsilon^{-d} \int \nabla w\left(\frac{y}{\varepsilon}\right) [f(x + \widetilde{y} - y) - f(x - y)]\, dy$$

and

$$|\nabla f^\varepsilon(x + \widetilde{y}) - \nabla f^\varepsilon(x)| \leqslant C\varepsilon^{-1} |\widetilde{y}|^\beta |f|_\beta = C|f|_\beta |\widetilde{y}|^{\alpha-1} \varepsilon^{-\alpha+\beta}.$$

For $\alpha = 2$, by Lemma 6,

$$|\nabla f^\varepsilon(x + \widetilde{y}) - \nabla f^\varepsilon(x)| \leqslant \sup_x |\partial^2 f^\varepsilon(x)||\widetilde{y}| \leqslant C|f|_\beta|\widetilde{y}|K(\varepsilon, \alpha, \beta).$$

In all cases, it holds that

$$|f^\varepsilon(x + y) - f^\varepsilon(x) - (\nabla f^\varepsilon(x), y)| \leqslant C|f|_\beta|y|^\alpha K(\varepsilon, \alpha, \beta).$$

Hence, for $|y| \leqslant 1$,

$$|f^\varepsilon\big(x + c(x)y\big) - f^\varepsilon(x) - \big(\nabla f^\varepsilon(x), c(x)y\big)| \leqslant C|f|_\beta|c(x)y|^\alpha K(\varepsilon, \alpha, \beta).$$

Also, for $|y| > 1$,

$$|f^\varepsilon\big(x + c(x)y\big) - f^\varepsilon(x)| \leqslant 2|f|_\beta.$$

Therefore, the statement follows by the assumptions and Lemma 6.

3.3. Rate of convergence. With results on the backward Kolmogorov equations and one-step estimates, the rate of convergence stated in Theorem 1 is proved by applying Itô's formula.

Proof. Let $v \in C^{\alpha+\beta}(\mathbf{H})$ be a unique solution to (5). By Itô's formula and Remark 3,

$$\mathbf{E}[v(0, X_0)] = \mathbf{E}[v(T, X_T)]$$
$$- \mathbf{E}\left[\int_0^T \big(\partial_t v(s, X_s) + \mathcal{A}_{X_s} v(s, X_s) + \mathcal{B}_{X_s} v(s, X_s)\big) ds\right]$$
$$= \mathbf{E}\left[g(X_T) - \int_0^T f(X_s) ds\right]$$

and

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

By Proposition 2, Corollary 2, Remark 4, and Lemma 2, for $s \in [0, T]$,

$$|\mathcal{A}_z v(s, \cdot)|_\beta + |\mathcal{B}_z v(s, \cdot)|_\beta \leqslant C|v|_{\alpha+\beta} \leqslant C|g|_{\alpha+\beta},$$
$$|\partial_t v(s, \cdot)|_\beta \leqslant C|g|_{\alpha+\beta}.$$

Then, by Itô's formula and Corollary 2,

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

Hence, by Lemma 3, there exists a constant C independent of g and f such that

$$|\mathbf{E}[g(Y_T)] - \mathbf{E}[g(X_T)]| \leqslant C|g|_{\alpha+\beta}\kappa(\delta, \alpha, \beta)$$

and

$$\left|\mathbf{E}\left[\int_0^T f(Y_{\tau_{i_s}})\,ds\right] - \mathbf{E}\left[\int_0^T f(X_s)\,ds\right]\right| \leqslant C|f|_\beta\kappa(\delta, \alpha, \beta).$$

The statement of Theorem 1 follows.

4. Conclusions. This paper studies weak Euler approximation for stochastic differential equations driven by Lévy processes. The dependence of the rate of convergence on the regularity of the coefficients and driving processes is investigated under the assumption that the coefficients are β-Hölder continuous. The main part of the stochastic differential equation is driven by a spherically symmetric α-stable process, and the tail of the Lévy measure of the lower-order part has a μ-order finite moment with $\mu \in [\beta, \alpha + \beta)$. Depending on the values of α and β, three scenarios are looked into to derive the rate of convergence for a full regularity scale. To estimate the rate, the existence of a unique solution to the corresponding backward Kolmogorov equation in Hölder space is first proved. The assumptions on the regularity of coefficients and test functions are weaker than those in the existing literature.

One possible improvement could be to consider the asymptotic of the tails at infinity instead of the tail moment μ. Another direction could be to consider stochastic differential equations driven by point and martingale measures. That is, for $\alpha \in (0, 2]$ and a measurable space $(U, \mathcal{U})$ associated with a nonnegative σ-finite measure π such that there exists a decreasing sequence $U_n \in \mathcal{U}$, $U = \bigcup_n U_n^c$, and $\pi(U_n^c) < \infty$ for each n, consider an $\mathbb{F}$-adapted stochastic process $X = \{X_t\}_{t\in[0,T]}$ solving

$$
\begin{aligned}
X_t = X_0 &+ \int_0^t a(X_{s-})\,ds + \int_0^t b(X_{s-})\,dW_s \\
&+ \int_0^t \int_{\mathbf{R}_0^d} c(X_{s-}, y)\big[\big(\mathbf{1}_{\{\alpha\in(0,2)\}} - \chi^\alpha(y)\big)\,p^\alpha(ds, dy) + \chi^\alpha(y)\,q^\alpha(ds, dy)\big] \\
&+ \int_0^t \int_U l(X_{s-}, v)\big[\big(1 - \widetilde{\chi}^\alpha(v)\big)\,p(ds, dv) + \widetilde{\chi}^\alpha(v)\,q(ds, dv)\big],
\end{aligned}
$$

where a and b are measurable and bounded, c and l are measurable,

$$\chi^\alpha(y) = \mathbf{1}_{\{\alpha=1\}}\mathbf{1}_{\{|y|\leqslant 1\}} + \mathbf{1}_{\{\alpha\in(1,2)\}}, \quad \widetilde{\chi}^\alpha(v) = \mathbf{1}_{\{\alpha\in(1,2]\}}\mathbf{1}_{\{v\in U_1\}},$$

$W = \{W_t\}_{t\in[0,T]}$ is a standard Wiener process, and $p^\alpha(dt, dy)$ and $p(dt, dv)$ are independent Poisson point measures on $[0, T] \times \mathbf{R}_0^d$ and $[0, T] \times U$, respectively, with $q^\alpha(dt, dy) = p^\alpha(dt, dy) - dy\,dt/|y|^{d+\alpha}$ and $q(dt, dv) = p(dt, dv) - \pi(dv)\,dt$ being the corresponding martingale measures.

REFERENCES

[1] D. APPLEBAUM, *Lévy Processes and Stochastic Calculus*, Cambridge Stud. Adv. Math. 93, Cambridge Univ. Press, Cambridge, 2004.

[2] O. E. BARNDORFF-NIELSEN, TH. MIKOSCH, AND S. I. RESNICK, EDS., *Lévy Processes. Theory and Applications*, Birkhäuser Boston, Inc., Boston, MA, 2001.

[3] J. BERGH AND J. LÖFSTRÖM, *Interpolation Spaces. An Introduction*, Grundlehren Math. Wiss. 223, Springer-Verlag, Berlin, 1976.

[4] J. BERTOIN, *Lévy Processes*, Cambridge Tracts in Math. 121, Cambridge Univ. Press, Cambridge, 1996.

[5] L. A. CAFFARELLI AND X. CABRÉ, *Fully Nonlinear Elliptic Equations*, Amer. Math. Soc. Colloq. Publ. 43, Amer. Math. Soc., Providence, RI, 1995.

[6] H. DONG AND D. KIM, *Schauder estimates for a class of non-local elliptic equations*, Discrete Contin. Dyn. Syst., 33 (2013), pp. 2319–2347.

[7] G. B. FOLLAND, *Real Analysis. Modern Techniques and Their Applications*, 2nd ed., Pure Appl. Math. (N.Y.), John Wiley & Sons, Inc., New York, 1999.

[8] J. JACOD, *The Euler scheme for Lévy driven stochastic differential equations*: *Limit theorems*, Ann. Probab., 32 (2004), pp. 1830–1872.

[9] J. JACOD, TH. G. KURTZ, S. MÉLÉARD, AND PH. PROTTER, *The approximate Euler method for Lévy driven stochastic differential equations*, Ann. Inst. H. Poincaré Probab. Statist., 41 (2005), pp. 523–558.

[10] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. (N.Y.) 23, Springer-Verlag, Berlin, 1992.

[11] A. KOHATSU-HIGA AND P. TANKOV, *Jump-adapted discretization schemes for Lévy-driven SDEs*, Stochastic Process. Appl., 120 (2010), pp. 2258–2285.

[12] T. KOMATSU, *On the martingale problem for generators of stable processes with perturbations*, Osaka J. Math., 21 (1984), pp. 113–132.

[13] R. MIKULEVIČIUS AND E. PLATEN, *Rate of convergence of the Euler approximation for diffusion processes*, Math. Nachr., 151 (1991), pp. 233–239.

[14] R. MIKULEVIČIUS AND H. PRAGARAUSKAS, *On the martingale problem associated with nondegenerate Lévy operators*, Lithuanian Math. J., 32 (1992), pp. 297–311.

[15] R. MIKULEVIČIUS AND H. PRAGARAUSKAS, *On the Cauchy problem for certain integro-differential operators in Sobolev and Hölder spaces*, Lithuanian Math. J., 32 (1992), pp. 238–264.

[16] R. MIKULEVIČIUS AND H. PRAGARAUSKAS, *On Hölder solutions of the integro-differential Zakai equation*, Stochastic Process. Appl., 119 (2009), pp. 3319–3355.

[17] R. MIKULEVIČIUS AND H. PRAGARAUSKAS, *On the Cauchy problem for integro-differential operators in Hölder classes and the uniqueness of the martingale problem*, Potential Anal., 40 (2014), pp. 539–563.

[18] R. MIKULEVIČIUS AND C. ZHANG, *On the rate of convergence of weak Euler approximation for nondegenerate SDEs driven by Lévy processes*, Stochastic Process. Appl., 121 (2011), pp. 1720–1748.

[19] G. N. MIL'SHTEIN, *A method of second-order accuracy integration of stochastic differential equations*, Theory Probab. Appl., 23 (1979), pp. 396–401.

[20] G. N. MIL'SHTEIN, *Weak approximation of solutions of systems of stochastic differential equations*, Theory Probab. Appl., 30 (1986), pp. 750–766.

[21] B. ØKSENDAL, *Stochastic Differential Equations. An Introduction with Applications*, 6th ed., Universitext, Springer-Verlag, Berlin, 1998.

[22] PH. PROTTER AND D. TALAY, *The Euler scheme for Lévy driven stochastic differential equations*, Ann. Probab., 25 (1997), pp. 393–423.

[23] S. RUBENTHALER, *Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process*, Stochastic Process. Appl., 103 (2003), pp. 311–349.

[24] K. SATO, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge Stud. Adv. Math. 68, Cambridge Univ. Press, Cambridge, 1999.

[25] W. SCHOUTENS, *Levy Processes in Finance. Pricing Financial Derivatives*, Wiley Ser. Probab. Statist., John Wiley & Sons, Ltd., Hoboken, NJ, 2003.

[26] D. TALAY, *Efficient numerical schemes for the approximation of expectations of functionals of the solution of a S.D.E., and applications*, in Filtering and Control of Random Processes (Paris, 1983), Lect. Notes Control Inf. Sci. 61, Springer, Berlin, 1984, pp. 294–313.

[27] D. TALAY, *Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution*, RAIRO Modél. Math. Anal. Numér., 20 (1986), pp. 141–179.

[28] H. TRIEBEL, *Theory of Function Spaces* II, Monogr. Math. 84, Birkhäuser Verlag, Basel, 1992.

[29] R. L. WOLPERT AND M. S. TAQQU, *Fractional Ornstein–Uhlenbeck Lévy processes and the Telecom process*: *Upstairs and downstairs*, Signal Process., 85 (2005), pp. 1523–1545.