Strong regularization by noise for kinetic SDEs

Giacomo Lucertini * Stefano Pagliarani † Andrea Pascucci ‡

This version: July 21, 2022

Abstract

In this paper we prove strong well-posedness for a system of stochastic differential equations driven by a degenerate diffusion satisfying a weak-type Hörmander condition, assuming Hölder regularity assumptions on the drift coefficient. This framework encompasses, as particular cases, stochastic Langevin systems of kinetic SDEs. The drift coefficient of the velocity component is allowed to be $\alpha$-Hölder continuous without any restriction on the index $\alpha$, which can be any positive number in $[0,1]$. As the deterministic counterparts of these differential systems are not well-posed, this result can be viewed as a phenomenon known as regularization by noise.

Keywords: regularization by noise, kinetic stochastic equations, Hörmander condition, Hölder estimates, anisotropic diffusion.

MSC: 60H10, 60J60, 35K65, 35K70.

1 Introduction

Let $W$ be a standard $d$-dimensional Brownian motion on some filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \geq 0})$. The $\mathbb{R}^d \times \mathbb{R}^d$-valued system of SDEs

$$
\begin{align*}
    dX^1_t &= F(t, X^1_t, X^2_t)dt + \sigma(t, X^1_t, X^2_t)dW_t, \\
    dX^2_t &= X^3_t dt,
\end{align*}
$$

(1.1)
is well-known in kinetic theory as it describes the motion of particles in the classical Langevin model. In a quite recent and interesting paper, Chaudru de Raynal [1] (see also [4]) proved the strong uniqueness for (1.1) assuming the Hölder regularity index of the drift coefficient $F$ strictly greater than $2/3$ and the diffusion coefficient $\sigma$ Lipschitz continuous in space, uniformly in time. Such a result, when the drift is less than Lipschitz continuous, is a remarkable instance of a well-known phenomenon called regularization by noise, see [7]. We note that, due to its degeneracy, the classical results for the martingale problem do not apply to (1.1): in fact, as in the present paper, the crucial structural assumption of [1] is the weak Hörmander condition. Uniqueness in law for (1.1) was proved by Menozzi [17] for $F$ Lipschitz continuous and by Chaudru de Raynal [2] assuming the regularity index $\alpha$ of $F$ to be strictly greater than $1/3$: the case $\alpha \in [0,\frac{1}{3}]$ was left as an open problem.

* Dipartimento di Matematica, Università di Bologna, Bologna, Italy. e-mail: giacomo.lucertini3@unibo.it.
† Dipartimento di Matematica, Università di Bologna, Bologna, Italy. e-mail: stefano.pagliarani9@unibo.it
‡ Dipartimento di Matematica, Università di Bologna, Bologna, Italy. e-mail: andrea.pascucci@unibo.it
In this paper we provide a direct proof of the \emph{strong uniqueness} for \eqref{eq:1.1} without restrictions on the Hölder exponent of the drift, which can be any $\alpha \in ]0,1[$, and under weak assumptions on the regularity on the diffusion coefficient that can be \emph{less than Lipschitz continuous}. We prove our results in a generalized framework, when the Lie algebra of the Hörmander condition can be of any step, including \eqref{eq:1.1} as the particular case of step two. More precisely, for fixed $N \geq d$ and $T > 0$, we consider the $N$-dimensional SDE
\begin{equation}
    dX_t = (BX_t + b(t, X_t)) dt + \Sigma(t, X_t) dW_t \tag{1.2}
\end{equation}
where $B$ is a $N \times N$ constant matrix and the coefficients are bounded measurable functions
\begin{align*}
    b : ]0, T[ \times \mathbb{R}^N &\rightarrow \mathbb{R}^N, \\
    \Sigma : ]0, T[ \times \mathbb{R}^N &\rightarrow \mathbb{R}^{N \times d}.
\end{align*}
Later we will specify the regularity of $b$ and $\Sigma$ in terms of some intrinsic Hölder and Lipschitz continuity in space, uniform in time.

\textbf{Remark 1.1.} System \eqref{eq:1.1} is a special case of \eqref{eq:1.2} with $N = 2d$. In particular, when $d = 1$ we have
\begin{equation*}
    B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} F \\ 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma \\ 0 \end{pmatrix}.
\end{equation*}
For constant $\sigma > 0$ and $F = 0$, Kolmogorov \cite{10} constructed the transition density of the Markov diffusion $X$ and characterized it as the Gaussian fundamental solution of the backward Kolmogorov operator
\begin{equation*}
    \mathcal{K} = \frac{\sigma^2}{2} \partial_{x_1 x_1} + x_1 \partial_{x_2} + \partial_t, \quad (x_1, x_2) \in \mathbb{R}^2.
\end{equation*}
Note that $\mathcal{K}$ is a hypoelliptic, non-uniformly parabolic operator on $\mathbb{R}^3$.

We now state our main structural assumptions.

\textbf{Assumption 1.2.} The matrix $\Sigma$ admits the block decomposition
\begin{equation*}
    \Sigma = \begin{pmatrix} \sigma \\ 0 \end{pmatrix}
\end{equation*}
where $\sigma = \sigma(t, x)$ is a $d \times d$ matrix such that $\sigma \sigma^* = (a_{ij})_{i, j = 1, \ldots, d}$ verifies the coercivity condition
\begin{equation*}
    \mu^{-1} |\xi|^2 \leq \sum_{i, j = 1}^d a_{ij}(t, x) \xi_i \xi_j \leq \mu |\xi|^2, \quad (t, x) \in ]0, T[ \times \mathbb{R}^N, \quad \xi \in \mathbb{R}^d,
\end{equation*}
for some positive constant $\mu$. We also have
\begin{equation*}
    b_{d+1} = \cdots = b_N = 0.
\end{equation*}

\textbf{Assumption 1.3 (Weak Hörmander condition).} The vector fields $\partial_{x_1}, \ldots, \partial_{x_d}$ and
\begin{equation*}
    Y := \langle Bx, \nabla \rangle + \partial_t
\end{equation*}
satisfy
\begin{equation*}
    \text{rank Lie}(\partial_{x_1}, \ldots, \partial_{x_d}, Y) = N + 1. \tag{1.3}
\end{equation*}
Remark 1.4. Formula (1.3) is usually referred to as a weak (or parabolic) Hörmander condition because of the crucial role of the drift $Y$ in generating the Lie algebra. When $d = N$ we recover the standard uniformly parabolic setting.

Remark 1.5. In [1] (see also [3, 4]) a more general dynamics for the $X_t^2$ component in (1.1) is allowed. In general, in their framework, the drift term $BX_t$ in (1.2) can be replaced with a non-linear function $G(X_t)$, so long as a weak-type Hörmander condition remains satisfied. It seems possible to extend our analysis to cover such a situation and this is left for future research. However, we should point out that some of the regularity thresholds on the function $G$, appearing in the aforementioned papers, are optimal and cannot be improved (see also [2, Section 4]).

Condition (1.3) is equivalent to the well-known Kalman rank condition for controllability of linear systems (cf., for instance, Section 9.5 in [20]). Also, it was shown in [12] that (1.3) is equivalent to $B$ having the block-form

$$B = \begin{pmatrix} * & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & 0 & \cdots & * & * \\ 0 & B_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_q & * \end{pmatrix}$$

(1.4)

where the $*$-blocks are arbitrary and $B_j$ is a $(d_j-1 \times d_j)$-matrix of rank $d_j$ with

$$d \equiv d_0 \geq d_1 \geq \cdots \geq d_q \geq 1, \quad \sum_{j=0}^q d_j = N.$$ 

The particular structure (1.4) of $B$ naturally induces an anisotropic metric on $\mathbb{R}^N$ and functional spaces that are suitable for the study of the regularity properties in this framework (cf. [15], [14], [18], [5] and [16]).

Definition 1.6 (Anisotropic norm and Hölder spaces). For any $x \in \mathbb{R}^N$ let

$$|x|_B := \sum_{j=0}^q \sum_{i=d_{j-1}+1}^{d_j} |x_i|^{\frac{1}{d_j}}, \quad \bar{d}_j := \sum_{k=0}^j d_k.$$ 

(1.5)

We denote by $C_B^\alpha$, for $\alpha \in ]0, 1]$, the set of continuous functions $g$ on $]0, T[\times\mathbb{R}^N$ such that

$$\|g\|_{C_B^\alpha} := \sup_{t \in ]0, T[} \sup_{x \in \mathbb{R}^N} |g(t, x)| + \sup_{t \in ]0, T[} \sup_{x, y \in \mathbb{R}^N, x \neq y} \frac{|g(t, x) - g(t, y)|}{|x - y|_B^{\alpha}} < \infty.$$ 

(1.6)

We write $g \in \text{Lip}_B$ if the norm (1.6) with $\alpha = 1$ is finite.

Remark 1.7. According to Definition 1.6 a function $g \in \text{Lip}_B$ is only $\frac{1}{d_j-1}$-Hölder continuous in the standard Euclidean sense, with respect to the variables $x_i$ for $\bar{d}_{j-1} < i \leq \bar{d}_j$ and $0 \leq j \leq q$. For instance, for the kinetic SDE (1.1) we have

$$|(x_1, x_2)|_B = |x_1| + |x_2|^{\frac{1}{d}}, \quad (x_1, x_2) \in \mathbb{R}^d \times \mathbb{R}^d.$$ 

In this case, $g \in \text{Lip}_B$ is actually only $\frac{1}{d}$-Hölder continuous as a function of $x_2$. Similarly, we have the strict inclusions

$$C_b^\alpha \subset C_B^\alpha \subset C_b^{\frac{1}{d}}.$$ 

(1.7)

where $C_b^\alpha$ denotes the space of bounded $\alpha$-Hölder continuous functions in the Euclidean sense.
The main result of this paper is the following

**Theorem 1.8.** Let Assumptions (1.2) and (1.3) be satisfied. If \( b \in C_B^\alpha \) for some \( \alpha \in [0,1] \) and \( \Sigma \in \text{Lip}_B \), then strong existence and uniqueness hold for SDE (1.2).

In light of (1.7), assumption \( b \in C_B^\alpha \) is equivalent to the standard Hölder continuity in the Euclidean sense. Only for sake of simplicity in Theorem 1.8 we assume coefficients that are continuous also in the time variable. Indeed, the results in [13], on which the proof Theorem 1.8 relies, are obtained for coefficients that are merely measurable in time. Therefore, also Theorem 1.8 could be proved in the latter setting, with some additional technicalities.

The proof of Theorem 1.8 makes use of the so-called Itô-Tanaka trick (see [7], [6], [9]). A crucial step in this technique is obtaining suitable intrinsic regularity estimates for the solution to the Cauchy problem for the Kolmogorov operator of (1.2). In particular, it is essential to prove intrinsic Lipschitz estimates on the solution and its gradient with respect to the first \( d \) space variables, which is done in Proposition 3.1. Note that these estimates are weaker than Lipschitz estimates with respect to the Euclidean distance, which do not hold under our regularity assumptions. Our proof, however, overcomes this lack of regularity by taking full advantage of the intrinsic geometry induced by the Hörmander condition (see Remark 3.3).

The rest of the paper is structured as follows. Section 2 contains preliminary definitions and results, while Theorem 1.8 is proved in Section 3. Section 3.1 contains the proof of Proposition 3.1.

## 2 Preliminary results

This section is devoted to recall some definitions and preliminary results that are necessary to construct a fundamental solution for the Kolmogorov operator of (1.2) and to establish its regularity. In the rest of the paper, to shorten notations we set

\[ S_T := [0, T] \times \mathbb{R}^N. \]

### 2.1 Intrinsic Hölder spaces and Itô formula

We introduce the intrinsic Hölder regularity along the vector fields appearing in the Hörmander condition (1.3). As it is standard in the framework of functional analysis on homogeneous groups (cf. [8]), the idea is to weight the Hölder exponent in terms of the formal degree of the vector fields, which is equal to 1 for \( \partial_{x_1}, \ldots, \partial_{x_d} \) and equal to 2 for \( Y \).

**Definition 2.1.** Let \( \alpha \in [0,1] \) and \( \beta \in ]0,2[ \). We denote respectively by \( C_B^\alpha \) and \( C_Y^\beta \) the set of the functions \( f : S_T \to \mathbb{R} \) such that the following semi-norms are finite

\[
\|f\|_{C_B^\alpha} := \sum_{i=1}^d \sup_{(t,x) \in S_T} \frac{|f(t,x + he_i) - f(t,x)|}{|he_i|^\alpha},
\]

\[
\|f\|_{C_Y^\beta} := \sup_{(t,x) \in S_T} \frac{|f(t,e^{(s-t)B}x) - f(t,x)|}{|t-s|^\beta}.
\]

Here \( e_i \) denotes the \( i \)-th element of the canonical basis of \( \mathbb{R}^N \).

Next we give the definition of intrinsic Hölder spaces of order 0, 1 and 2.
**Definition 2.2.** For \( \alpha \in [0, 1] \), \( C^{0,\alpha}_B \), \( C^{1,\alpha}_B \) and \( C^{2,\alpha}_B \) denote, respectively, the set of the functions \( f : S_T \to \mathbb{R} \) such that the following semi-norms are finite

\[
\| f \|_{C^{0,\alpha}_B} := \| f \|_{C^0_Y} + \| f \|_{C^2_Y},
\]

\[
\| f \|_{C^{1,\alpha}_B} := \| f \|_{C^{1+\alpha}_Y} + \sum_{i=1}^d \| \partial_i f \|_{C^{\alpha}_B} = \| f \|_{C^{1+\alpha}_Y} + \sum_{i=1}^d (\| \partial_i f \|_{C^\alpha_Y} + \| \partial_i f \|_{C^2_Y}),
\]

\[
\| f \|_{C^{2,\alpha}_B} = \sum_{i=1}^d \| \partial_i f \|_{C^{1+\alpha}_B} + \| Yf \|_{C^{\alpha}_B} = \sum_{i=1}^d (\| \partial_i f \|_{C^{1+\alpha}_B} + \sum_{i,j=1}^d (\| \partial_{ij}f \|_{C^{\alpha}_Y} + \| \partial_{ij}f \|_{C^2_Y}) + \| Yf \|_{C^{\alpha}_B}).
\]

For any \( n = 0, 1, 2 \), we say that a function \( f \in C^{n,\alpha}_B, \psi \in C^{n,\alpha}_B \) for any \( \psi \in C^{\infty}_0(S_T) \). Here \( Yf \) is intended as a Lie (or directional) derivative.

Intrinsic Taylor formulas for functions in \( C^{0,\alpha}_B \), \( C^{1,\alpha}_B \) were proved in [18] and [19]. We note explicitly that, to cope with the lack of regularity of the coefficients in the time-direction, the definition of \( C^{2,\alpha}_B \) differs from the one given in [18], specifically with regards to the regularity along \( Y \), namely, in [18] \( Yf \in C^2_Y \). We refer the reader to [13] Remark 1.15 for a comparison with other notions of intrinsic Hölder spaces proposed in the literature. Using intrinsic Taylor formulas it is possible to derive the following lemma (see also [11] Theorem 1.3).

**Lemma 2.3 (Intrinsic Itô formula).** Let \( X \) be a solution of equation (1.2) and \( f \in C^{2,\alpha}_B \). Then, we have

\[
df(t, X_t) = \mathcal{K}f(t, X_t)dt + (\nabla df \cdot \sigma)(t, X_t)dW_t,
\]

where \( \nabla_d := (\partial_{x_1}, \ldots, \partial_{x_d}) \).

### 2.2 Fundamental solution

In [18] the authors proved that under the assumptions of Theorem 1.8 the backward Kolmogorov operator of (1.2)

\[
\mathcal{K} := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, x) \partial_{x_i} \partial_{x_j} + \sum_{i=1}^d b_i(t, x) \partial_{x_i} + \langle Bx, \nabla \rangle + \partial_t, \quad (t, x) \in S_T,
\]

(2.1)

has a fundamental solution \( p = p(t, x; s, y) \). In [13] it is also proved that, for any fixed \( (s, y) \in S_T \) and \( \beta < \alpha \), the function \( p(\cdot, \cdot; s, y) \) belongs to \( C^{2,\beta}_B \) on the strip \( S_T \) for any \( \tau < T \) this is an optimal result since no further regularity can be expected under these assumptions on the coefficients. Hereafter, we say that a function \( u \) solves the PDE

\[
\mathcal{K}u = f \quad \text{in } S_T,
\]

(2.2)

if there exist the Euclidean derivatives \( \partial_{x_i} u, \partial_{x_i} x_j u \) for \( i, j = 1, \ldots, d \), the Lie derivative \( Y u \) and (2.2) is pointwise satisfied.

**Definition 2.4.** A fundamental solution to \( \mathcal{K} \) on \( S_T \) is a function \( p = p(t, x; s, y) \), defined for \( 0 < t < s < T \) and \( x, y \in \mathbb{R}^N \), such that, for any \( (s, y) \in S_T \):

1. \( \mathcal{K}p(\cdot, \cdot; s, y) = 0 \) in \( S_s \).
(ii) for any \( g \in C_b(\mathbb{R}^N) \) (i.e. \( g \) is bounded and continuous) we have
\[
\lim_{(t,x) \to (s,y)} \int_{\mathbb{R}^N} p(t,x; s, \eta) g(\eta) d\eta = g(y).
\]

The next result is a particular case of Theorem 1.8-1.9 in [13].

**Theorem 2.5 (Existence and regularity of fundamental solution).** Under the assumptions of Theorem 1.8, the operator \( K \) has a fundamental solution \( p \) on \( S_T \). Moreover, \( p(\cdot, \cdot; s, y) \in C^{2,\beta}_B \) on \( S_\tau \) for every \( 0 < \tau < s < T, y \in \mathbb{R}^N \) and \( \beta < \alpha \).

**3 Proof of Theorem 1.8**

Weak existence follows from a compactness argument (see [21]). Therefore, it is enough to prove strong uniqueness in order to have strong well-posedness. To get rid of the irregular drift term of SDE (1.2), we develop a modification of the so-called Itô-Tanaka trick (cf. [7]) suitably designed to exploit the intrinsic regularity of the fundamental solution of the generator \( K \) of (1.2).

We shall need the following result whose proof is postponed to Section 3.1.

**Proposition 3.1.** Under the assumptions of Theorem 1.8, for \( \lambda > 0 \) consider the function
\[
u_{\lambda}(t,x; T) := \int_t^T \int_{\mathbb{R}^N} e^{-\lambda (s-t)} p(t,x; s,y) b(s, y) dy ds, \quad (t,x) \in [0,T] \times \mathbb{R}^N,
\]
where \( p \) is the fundamental solution in Theorem 2.5. Then we have:

(i) \( \nu_{\lambda}(\cdot, \cdot; T) \in C^{2,\beta}_B \), for \( \beta < \alpha \), solves the vector-valued Cauchy problem
\[
\begin{cases}
Ku = -b + \lambda u, & \text{in } S_T, \\
u(T, \cdot) = 0, & \text{in } \mathbb{R}^N;
\end{cases}
\]

(ii) there exists \( c > 0 \), dependent on \( N, \mu, T, B \), the Lip_B-norm of \( \sigma \) and the \( C^\alpha_B \)-norm of \( b \), such that
\[
|\nu_{\lambda}(t,x; T) - \nu_{\lambda}(t,y; T)| \leq \frac{c}{\sqrt{\lambda}} |x - y|_B, \quad (3.1)
\]
\[
|\partial_x \nu_{\lambda}(t,x; T) - \partial_x \nu_{\lambda}(t,y; T)| \leq \frac{c}{\lambda^2} |x - y|_B, \quad (3.2)
\]

for any \( 0 < t < T, x, y \in \mathbb{R}^N \) and for any \( i = 1, \ldots, d \).

**Remark 3.2.** Since \( b_{d+1} = \cdots = b_N \equiv 0 \), the \( j \)-th component of \( \nu_{\lambda} \) is equal to zero, for any \( j = d+1, \ldots, N \).

Let \( X \) be a solution to (1.2) with respect to a given Brownian motion \( W \). By Proposition 3.1(i) and by the intrinsic Itô formula of Lemma 2.3 we have
\[
d\nu_{\lambda}(t, X_t) = Ku_{\lambda}(t, X_t) dt + (\nabla u_{\lambda} \cdot \Sigma)(t, X_t) dW_t
= (\lambda u_{\lambda}(t, X_t) - b(t, X_t)) dt + (\nabla u_{\lambda} \cdot \Sigma)(t, X_t) dW_t,
\]
which is, for \( t \in [0,T] \),
\[
\int_0^t b(s, X_s) ds = \nu_{\lambda}(0, X_0) - \nu_{\lambda}(t, X_t) + \lambda \int_0^t u_{\lambda}(s, X_s) ds + \int_0^t (\nabla u_{\lambda} \cdot \Sigma)(s, X_s) dW_s. \quad (3.3)
\]
Here, we a slight abuse of notation, \( \nabla u_\lambda \cdot \Sigma \) represents the \( N \)-dimensional vector with components

\[
(\nabla u_\lambda \cdot \Sigma)_i := \begin{cases} 
(\nabla_d u_\lambda \cdot \sigma)_i & \text{if } i = 1, \ldots, d, \\
0 & \text{if } i = d + 1, \ldots, N.
\end{cases}
\]

Plugging \((3.3)\) into \((1.2)\), we obtain

\[
X_t = X_0 + u_\lambda(0, X_0) - u_\lambda(t, X_t) + \int_0^t (\lambda u_\lambda(s, X_s) + B X_s)ds + \int_0^t (\nabla u_\lambda(s, X_s) + I) \cdot \Sigma(s, X_s)dW_s. \tag{3.4}
\]

Let now \(X'\) be a second solution to \((1.2)\) with respect to the same Brownian motion \(W\) and set \(Z := X - X'\). Writing also \(X'\) as in \((3.3)\) and subtracting the two equations we obtain

\[
Z_t = -u_\lambda(t, X_t) + u_\lambda(t, X'_t) + \lambda \int_0^t (u_\lambda(s, X_s) - u_\lambda(s, X'_s))ds + \int_0^t B Z_s ds \\
+ \int_0^t ((\nabla u_\lambda \cdot \Sigma)(s, X_s) - (\nabla u_\lambda \cdot \Sigma)(s, X'_s) + \Sigma(s, X_s) - \Sigma(s, X'_s))dW_s.
\]

Set now \(q > 2r + 1\) and note that the function \(x \mapsto |x|_B^q\) is convex on \(\mathbb{R}^N\). Since, by Remark \(3.2\) \(|u_\lambda| = |u_\lambda|_B\), we have

\[
\frac{E[|Z_t|_B^q]}{4q-1} \leq E[|u_\lambda(t, X_t) - u_\lambda(t, X'_t)|^q] + \lambda^q E\left[\int_0^t (u_\lambda(s, X_s) - u_\lambda(s, X'_s))ds\right]^q + E\left[\int_0^t B Z_s ds\right]^q \\
+ E\left[\int_0^t ((\nabla u_\lambda \cdot \Sigma)(s, X_s) - (\nabla u_\lambda \cdot \Sigma)(s, X'_s) + \Sigma(s, X_s) - \Sigma(s, X'_s))dW_s\right]^q
\]

(by the employing Jensen and Burkholder inequalities)

\[
\leq E[|u_\lambda(t, X_t) - u_\lambda(t, X'_t)|^q] + c \left( \lambda^q E\left[\int_0^t |(u_\lambda(s, X_s) - u_\lambda(s, X'_s))|^q ds\right] + E\left[\int_0^t |B Z_s|^q_B ds\right]\right) \\
+ c E\left[\int_0^t |(\nabla u_\lambda \cdot \Sigma)(s, X_s) - (\nabla u_\lambda \cdot \Sigma)(s, X'_s) + \Sigma(s, X_s) - \Sigma(s, X'_s)|^q ds\right]
\]

(by Proposition \(3.1(3)\) and the fact that \(\Sigma \in \text{Lip}_B\))

\[
\leq \frac{c}{\lambda^{q/2}} E[|Z_t|_B^q] + c(1 + \lambda^2) \int_0^t E[|Z_s|_B^q] ds,
\]

with \(c > 0\) dependent on \(N, \mu, T, B\) on the \(\text{Lip}_B\) norm of \(\sigma\) and on the \(C^0_B\) norm of \(b\). Now, taking \(\lambda\) suitably large, Gronwall inequality yields that \(X\) and \(X'\) are indistinguishable.

**Remark 3.3.** The proof in \([1]\) is based on the Zvonkin transform, which is similar to the Itô-Tanaka trick that we have exploited. In the aforementioned paper the author uses an argument analogous to ours: the main difference is that we bound \(X_t - X'_t\) in terms of the homogeneous norm \(|\cdot|_B\) as opposed to the Euclidean one. This enables us to complete the proof by using the estimates \((3.1)-(3.2)\), which are Lipschitz estimates with respect to the homogeneous norm. The latter, weaker than Euclidean Lipschitz estimates in \([1\text{ Proposition 1.3}]\), can be proved under the assumption of \(b\) being \(\alpha\)-Hölder continuous with no threshold for the Hölder exponent \(\alpha\).

The remainder of this section is devoted to the proof of Proposition \(3.1\)
3.1 Proof of Proposition 3.1

We start by recalling the setting in [13] for the construction of a fundamental solution of $K$. Let $G(C, \cdot)$ be the Gaussian kernel

$$G(C, x) := \frac{1}{\sqrt{(2\pi)^N \det C}} e^{-\frac{1}{2}(C^{-1}x, x)}, \quad x \in \mathbb{R}^N,$$

where $C$ is a symmetric and positive definite $N \times N$ matrix. For $\delta > 0$, we also set

$$\Gamma^\delta(t, x; s, y) := G(C(s-t), y - e^{(s-t)B}x), \quad C(\tau) := \int_0^\tau e^{(\tau - r)B} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(\tau - r)B^*} dr.$$

Throughout the remainder of this section, $C_\varepsilon$ will denote, indistinctly, a positive constant that only depends on a given $\varepsilon > 0$, on $N, T, \mu, B, \alpha$, and on the $C_B^a$ norms of the coefficients $b_i, a_{ij}$ of the Kolmogorov operator $K$ in (2.1).

**Notation 3.4.** Let $\nu = (\nu_1, \ldots, \nu_N) \in \mathbb{N}_0^N$ be a multi-index. We recall notation (1.5) and define the $B$-length of $\nu$ as

$$[\nu]_B := \sum_{j=0}^{q} (2^j + 1) \sum_{i=d_{j-1}+1}^{d_j} \nu_i.$$

**Remark 3.5.** Note that

$$|x|_B = \sum_{j=0}^{N} |x_j| \frac{e_j}{\varepsilon_j B},$$

where $e_j$ is the $j$-th vector of the canonical basis of $\mathbb{R}^N$.

Following the parametrix method, in [13] the authors constructed a fundamental solution $p$ for $K$ in the form

$$p(t, x; s, y) = P(t, x; s, y) + \Phi(t, x; s, y), \quad 0 \leq t < s \leq T, \quad x, y \in \mathbb{R}^N, \quad (3.5)$$

with

$$\Phi(t, x; s, y) := \int_t^s J(t, x; \tau; s, y) d\tau,$$

where:

- $P = P(t, x; s, y)$ is the so-called *parametrix function*, which is defined as
  $$P(t, x; s, y) := \Gamma(s, y)(t, x; s, y), \quad 0 \leq t < s \leq T, \quad x, y \in \mathbb{R}^N,$$

  where, for $(\tau, v) \in S_T$, we set
  $$\Gamma(\tau, v)(t, x; s, y) := G(C(\tau, v)(t, s), y - e^{(s-t)B}x), \quad 0 \leq t < s \leq T, \quad x, y \in \mathbb{R}^N,$$
  and
  $$C(\tau, v)(t, s) := \int_t^s e^{(s-r)B} A(\tau, v)(r) e^{(s-r)B^*} dr,$$
  $$A(\tau, v)(r) := \begin{pmatrix} A_0(r, e^{(r-\tau)B}v) & 0 \\ 0 & 0 \end{pmatrix}, \quad A_0 = (a_{ij})_{i,j=1,\ldots,d};$$
• \( J = J(t, x; s, y) \) is a measurable function such that, for every \( \varepsilon > 0 \), \( i = 1, \ldots, d \), and \( k = 1, \ldots, N \), we have

\[
|J(t, x; s, y)| \leq \frac{C_{\varepsilon}}{(s - \tau)^{1 - \frac{\alpha}{2}}} \Gamma^{\mu + \varepsilon}(t, x; T, y), \quad \text{(3.6)}
\]

\[
|\partial_{x_k} J(t, x; s, y)| \leq \frac{C_{\varepsilon}}{(s - \tau)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x; s, y)}, \quad \text{(3.7)}
\]

\[
|\partial_{x_i} J(t, x; \tau; s, y)| \leq \frac{C_{\varepsilon}}{(s - \tau)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x; T, y)}, \quad \text{(3.8)}
\]

for every \( 0 \leq t < \tau < s \leq T \), \( x, y \in \mathbb{R}^N \) (see \cite{13} Proposition B.1).

We first prove some preliminary results.

**Lemma 3.6.** For every \( \varepsilon > 0 \), \( k = 1, \ldots, N \) and \( i = 1, \ldots, d \), we have

\[
|\Phi(t, x + h e_k; s, y) - \Phi(t, x; s, y)| \leq \frac{C_{\varepsilon} |h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x + h e_k; s, y) + \Gamma^{\mu + \varepsilon}(t, x; s, y)), \quad \text{(3.9)}
\]

\[
|\partial_{x_i} \Phi(t, x + h e_k; s, y) - \partial_{x_i} \Phi(t, x; s, y)| \leq \frac{C_{\varepsilon} |h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x + h e_k; s, y) + \Gamma^{\mu + \varepsilon}(t, x; s, y)), \quad \text{(3.10)}
\]

for any \( 0 \leq t < s \leq T \), \( x, y \in \mathbb{R}^N \) and \( h \in \mathbb{R} \).

**Proof.** We prove \cite{13}. We set \( \beta := |e_k| \) and we have that

\[
\Phi(t, x + h e_k; s, y) - \Phi(t, x; s, y) = \int_t^s \underbrace{\partial_{x_i} J(t, x + h e_k; \tau; s, y) - J(t, x; \tau; s, y)}_{=: I(\tau)} d\tau.
\]

We split the domain of \((t, \tau)\) in two separate regions:

**Case** \( \tau - t \geq h^{2/\beta} \). By the mean-value theorem, there exists a real \( \bar{h} \) with \( |\bar{h}| \leq |h| \) such that

\[
I(\tau) = h \partial_{x_i} J(t, x + \bar{h} e_k; \tau; s, y).
\]

Therefore, estimate \cite{3.7} yields

\[
|I(\tau)| \leq C_{\varepsilon} \frac{|h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x + \bar{h} e_k; s, y)}
\]

(since \( \tau - t \geq h^{2/\beta} \))

\[
\leq C_{\varepsilon} \frac{|h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x + \bar{h} e_k; s, y)}
\]

(by standard estimates on \( \Gamma^{\mu + \varepsilon}(t, x + \bar{h} e_k; s, y) \) with \( \tau - t \geq h^{2/\beta} \))

\[
\leq C_{\varepsilon} \frac{|h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} \Gamma^{\mu + \varepsilon}(t, x + \bar{h} e_k; s, y) + \Gamma^{\mu + \varepsilon}(t, x; s, y))}.
\]

Therefore, by integrating in \( \tau \) we obtain

\[
\left| \int_t^s I(\tau) d\tau \right| \leq C_{\varepsilon} \frac{|h|^{1/\beta}}{(s - t)^{1 - \frac{\alpha}{2}}(\tau - t)^{\frac{\alpha - n}{2}} (\Gamma^{\mu + \varepsilon}(t, x + h e_k; s, y) + \Gamma^{\mu + \varepsilon}(t, x; s, y))}
\]
which proves \([3.9]\) when \(\tau - t \geq h^{2/\beta}\).

**Case** \(\tau - t < h^{2/\beta}\). Employing triangular inequality and estimate \([3.10]\) we get

\[
|I(\tau)| \leq \frac{C_\varepsilon}{(s - \tau)^{1 - \frac{\alpha}{2}}}(\Gamma^{\alpha+\varepsilon}(t, x + h\mathbf{e}_k; s, y) + \Gamma^{\alpha+\varepsilon}(t, x; s, y))
\]

(since \(\tau - t < h^{2/\beta}\))

\[
\leq C\frac{|h|^{1/\beta}}{(s - \tau)^{1 - \frac{\alpha}{2}(\tau - t)^{\frac{1}{\beta}}}}(\Gamma^{\alpha+\varepsilon}(t, x + h\mathbf{e}_k; s, y) + \Gamma^{\alpha+\varepsilon}(t, x; s, y)).
\]

Integrating in \(\tau\) yields \([3.9]\) when \(\tau - t < h^{2/\beta}\).

Equation \([3.10]\) can be proved following the previous arguments, in particular employing \([3.8]\) in the case \(\tau - t \geq h^{2/\beta}\). □

We also need the estimates in the following

**Lemma 3.7.** Set

\[
V_0(t, x; s) := \int_{\mathbb{R}^N} P(t, x; s, y)b(s, y)dy, \quad 0 \leq t < s \leq T, \ x \in \mathbb{R}^N.
\]

Then, for every \(k = 1, \ldots, N\) and \(i = 1, \ldots, d\) we have

\[
|V_0(t, x + h\mathbf{e}_k; s) - V_0(t, x; s)| \leq C_1|h|^{\frac{1}{\alpha+\varepsilon}}\frac{1}{(s - t)^{\frac{1}{2} - \frac{\alpha}{2}}},
\]

(3.11)

\[
|\partial_x V_0(t, x + h\mathbf{e}_k; s) - \partial_x V_0(t, x; s)| \leq C_1|h|^{\frac{1}{\alpha+\varepsilon}}\frac{1}{(s - t)^{\frac{1}{2} - \frac{\alpha}{2}}},
\]

(3.12)

for any \(0 \leq t < s \leq T, \ x \in \mathbb{R}^N\) and \(h \in \mathbb{R}\).

**Proof.** We set \(\beta := |b_k|_{B}\). Proceeding as in [13] Proposition B.1] we can prove that, for any \(k = 1, \ldots, N\) and \(i = 1, \ldots, d\), we have

\[
|V_0(t, x; s)| \leq C_1,
\]

(3.13)

\[
|\partial_x V_0(t, x; s)| \leq C_1\frac{1}{(s - t)^{\frac{1}{2} + \alpha}},
\]

(3.14)

\[
|\partial_{x,i} V_0(t, x; s)| \leq C_1\frac{1}{(s - t)^{\frac{1}{2} + \alpha}}.
\]

(3.15)

for any \(0 \leq t < s \leq T\) and \(x \in \mathbb{R}^N\).

We first prove \([3.11]\). As in the previous proof, we split the domain of \((t, s)\) into two separate regions:

**Case** \(s - t \geq h^{2/\beta}\). By the mean-value theorem, there exists a real \(\bar{h}\) with \(|\bar{h}| \leq |h|\) such that

\[
|V_0(t, x + h\mathbf{e}_k; s) - V_0(t, x; s)| = |h| |\partial_x V_0(t, x + \bar{h}\mathbf{e}_k; s)|
\]

(by the estimate \([3.14]\) and since \(s - t \geq h^{2/\beta}\))

\[
\leq C_1\frac{|h|^{1/\beta}}{(s - t)^{\frac{1}{2} - \frac{\alpha}{2}}} \leq C_1\frac{|h|^{1/\beta}}{(s - t)^{\frac{1}{2}}}.
\]

**Case** \(s - t < h^{2/\beta}\). Employing estimate \([3.10]\) we get
\[ |V_b(t, x + h \mathbf{e}_k; s) - V_b(t, x; s)| \leq C_1 \]

(since \( s - t < h^{2/\beta} \))

\[ \leq C_1 \frac{|h|^{1/\beta}}{(s - t)^{\frac{\beta}{2}}}. \]

With the same argument, we can prove (3.12) utilizing estimate (3.15) in the case \( s - t \geq h^{2/\beta} \) and (3.14) in the case \( s - t < h^{2/\beta} \).

We are now in the position to prove Proposition 3.1.

**Proof of Proposition 3.1.** Part (i). A direct computation shows that the function

\[ p_\lambda(t, x, s, y) := e^{-\lambda(s-t)} p(t, x; s, y) \]

is a fundamental solution of the operator \((K - \lambda)\). The result then stems from [13, Proposition 4.2].

Part (ii). We first note that the \( C^\alpha_B \) norms of the coefficients \( a_{ij} \) of the Kolmogorov operator \( K \) can be bounded by the square of the \( \text{Lip}_B \) norms of the coefficients \( \sigma_{ij} \).

By the representation (3.5) we have

\[ u_\lambda(t, x) - u_\lambda(t, x') = \int_t^T \int_{\mathbb{R}^N} (p(t, x; s, y) - p(t, x'; s, y)) b(s, y) dy e^{-\lambda(s-t)} ds \]

\[ = \int_t^T \int_{\mathbb{R}^N} (P(t, x; s, y) - P(t, x'; s, y)) b(s, y) dy e^{-\lambda(s-t)} ds \]

\[ + \int_t^T \int_{\mathbb{R}^N} (\Phi(t, x; s, y) - \Phi(t, x'; s, y)) b(s, y) dy e^{-\lambda(s-t)} ds, \]

and thus

\[ |u_\lambda(t, x) - u_\lambda(t, x')| \leq \int_t^T \int_{\mathbb{R}^N} |V_b(t, x; s) - V_b(t, x'; s)| e^{-\lambda(s-t)} ds \]

\[ + \int_t^T \int_{\mathbb{R}^N} |\Phi(t, x; s, y) - \Phi(t, x'; s, y)| \|b\|_\infty dy e^{-\lambda(s-t)} ds \]

(by (3.9) and (3.11) together with Remark 3.5 and integrating in \( y \))

\[ \leq c |x - x'| B \int_t^T \frac{e^{-\lambda(s-t)}}{(s-t)\frac{\beta}{2}} ds \leq c \frac{|x - x'| B}{\lambda^{\frac{\beta}{2}}}, \]

which is (3.11).

Finally, employing the same argument, with (3.10) and (3.12), yields (3.2) and concludes the proof.

**References**

[1] P. E. Chaudru de Raynal, *Strong existence and uniqueness for degenerate SDE with Hölder drift*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 259–286.

[2] P.-E. Chaudru de Raynal, *Weak regularization by stochastic drift: result and counter example*, Discrete Contin. Dyn. Syst., 38 (2018), pp. 1269–1291.

[3] P.-E. Chaudru de Raynal and S. Menozzi, *Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result*, Trans. Amer. Math. Soc., 375 (2022), pp. 1–45.
[4] P.-E. C. de Raynal, I. Honoré, and S. Menozzi, *Strong regularization by brownian noise propagating through a weak Hörmander structure*, preprint, arXiv:1810.12225, (2018).

[5] M. Di Francesco and S. Polidoro, *Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form*, Adv. Differential Equations, 11 (2006), pp. 1261–1320.

[6] F. Flandoli, *Random perturbation of PDEs and fluid dynamic models*, vol. 2015 of Lecture Notes in Mathematics, Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].

[7] F. Flandoli, M. Gubinelli, and E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math., 180 (2010), pp. 1–53.

[8] G. B. Folland and E. M. Stein, *Hardy spaces on homogeneous groups*, vol. 28 of Mathematical Notes, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

[9] R. Huang, *Scaling limit for a second-order particle system with local annihilation*, Preprint arxiv.org/abs/2205.07312, (2022).

[10] A. Kolmogorov, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2), 35 (1934), pp. 116–117.

[11] A. Lanconelli, S. Pagliarani, and A. Pascucci, *Local densities for a class of degenerate diffusions*, Ann. Inst. Henri Poincaré Probab. Stat., 56 (2020), pp. 1440–1464.

[12] E. Lanconelli and S. Polidoro, *On a class of hypoelliptic evolution operators*, vol. 52, 1994, pp. 29–63. Partial differential equations, II (Turin, 1993).

[13] G. Lucertini, S. Pagliarani, and A. Pascucci, *Optimal regularity for degenerate Kolmogorov equations with rough coefficients*, preprint, arXiv:2204.14158, (2021).

[14] A. Lunardi, *Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in \( \mathbb{R}^n \)*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 24 (1997), pp. 133–164.

[15] M. Manfredini, *The Dirichlet problem for a class of ultraparabolic equations*, Adv. Differential Equations, 2 (1997), pp. 831–866.

[16] S. Menozzi, *Parametrix techniques and martingale problems for some degenerate Kolmogorov equations*, Electron. Commun. Probab., 16 (2011), pp. 234–250.

[17] ———, *Martingale problems for some degenerate Kolmogorov equations*, Stochastic Process. Appl., 128 (2018), pp. 756–802.

[18] S. Pagliarani, A. Pascucci, and M. Pignotti, *Intrinsic Taylor formula for Kolmogorov-type homogeneous groups*, J. Math. Anal. Appl., 435 (2016), pp. 1054–1087.

[19] S. Pagliarani, A. Pascucci, and M. Pignotti, *Intrinsic expansions for averaged diffusion processes*, Stochastic Process. Appl., 127 (2017), pp. 2560–2585.
[20] A. Pascucci, *PDE and martingale methods in option pricing*, vol. 2 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2011.

[21] D. W. Stroock and S. S. Varadhan, *Multidimensional diffusion processes*, vol. 233, Springer Science & Business Media, 1997.