EXPONENTIAL MIXING FOR SDES UNDER THE TOTAL VARIATION*

XUHUI PENG

KEY LABORATORY OF HIGH PERFORMANCE COMPUTING AND STOCHASTIC INFORMATION PROCESSING(HPCSIP)(MINISTRY OF EDUCATION OF CHINA), COLLEGE OF MATHEMATICS AND COMPUTER SCIENCE, HUNAN NORMAL UNIVERSITY, CHANGSHA, HUNAN 410081, P. R. CHINA

Abstract. First, we establish an abstract ergodic result on \( \mathbb{R}^d \). Classical ergodic results on \( \mathbb{R}^d \) require that the process is irreducible, we weaken it to some weak form of irreducibility in this article. The main method used in this article is coupling. Then, we apply our abstract ergodic result to stochastic differential equations driven by a Lévy noise and obtain a new result.

Keywords: Exponential Mixing, Coupling Method, Ergodic.

MSC 2000: 60H15; 60H07

1. Motivations and Main Results

1.1. Motivations. We consider a time homogeneous Markov process \( X = \{X_t, t \in \mathbb{R}^+\} \) on \( \mathbb{R}^d \). When this process is starting from \( x \in \mathbb{R}^d \) at time \( t = 0 \), we also denote this process by \( X = \{X_t(x), t \in \mathbb{R}^+\} \). The process \( X \) is supposed to be strong Markov, be adapted to a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) and to have càdlàg trajectories. The transition function for the process \( X \) is denoted by \( P_t(x, dy), t \in \mathbb{R}^+, x \in \mathbb{R}^d \).

A family of linear operators \( \{P_t\}_{t \geq 0} \) on \( \mathcal{B}_b(\mathbb{R}^d) \) (the space of bounded and Borel measurable functions), defined by

\[
P_tf(x) := \mathbb{E}[f(X_t(x))], \quad x \in \mathbb{R}^d, \; t \geq 0, \; f \in \mathcal{B}_b(\mathbb{R}^d),
\]

is associated with the process \( X \).

This article is mainly motivated by [18]. In [18], under the following two hypotheses \( H_a \), \( H_b \), and some type of Lyapunov structure, the authors proved

\[
\|P_t(x, \cdot) - \mu\|_W \leq K(x)e^{-\gamma t}, \quad \forall t \geq 0,
\]

Email addresses: xhpeng@hunnu.edu.cn & pengxuhui@amss.ac.cn

* The author was supported by NSFC (No.11501195), a Scientific Research Fund of Hunan Provincial Education Department (No.17C0953), the Youth Scientific Research Fund of Hunan Normal University (No.Math140650) and the Construct Program of the Key Discipline in Hunan Province.
here $\gamma > 0$, $K : \mathbb{R}^d \to \mathbb{R}^+$ is a function depending on the Lyapunov structure, $\|v\|_W$ denotes a weighted total variation norm for the signed measure $v$, $\mu$ is the invariant measure for $P_t$, i.e,

$$\mu(A) = P_t^* \mu(A) := \int_{\mathbb{R}^d} P_t(x, A) \mu(dx), \quad \forall t > 0, \forall A \in \mathcal{B}(\mathbb{R}^d),$$

here $\mathcal{B}(\mathbb{R}^d)$ denotes the collection of all Borel measurable sets on $\mathbb{R}^d$.

- $H_a$. The Markov process $X_t$ is irreducible aperiodic, i.e. there exists a $t_0 > 0$ such that

$$P_{t_0}(x, A) > 0, \quad \forall x \in \mathbb{R}^d, \forall \text{ open set } A;$$

- $H_b$. $P_t$ is strong Feller.

Other classical abstract ergodic results were given in [4][13][14][15] etc. Assume that

$\Phi_t$ is a right process, and that all compact sets are petite for some skeleton chain, for some $c, d > 0$ and some norm-like function $V$

$$AV(x) \leq -cV(x) + d,$$

then there exist $\beta < 1$ and $B < \infty$ such that

$$\sup_{|g| \leq V + 1} \left| \int_X g(y) P_t(x, dy) - \pi(g) \right| \leq B(V(x) + 1)^\beta, \quad x \in X, \quad t \in \mathbb{R}^+. $$

It isn’t easy to verify $H_a$ in many situations. Compared to the results of [18], in subsection 1.2 below, we will give an abstract ergodic result in Theorem 1.1. Our Hypotheses $H_2$ and $H_3$ in Theorem 1.1 are weaker than Hypotheses $H_a, H_b$ in some situations. Then in subsection 1.3, we apply our abstract result to stochastic differential equations driven by a Lévy noise and obtain a new result.

1.2. Abstract Ergodic Result. Assume $F(u) \geq 1$ is a continuous function on $\mathbb{R}^d$ tending to $+\infty$ as $|u| \to \infty$. Our Hypotheses in this article are

- $H_1$ (Lyapunov function condition). There are positive constants $t_*, R_*, C_*$ and $a < 1$ such that

$$\mathbb{E}\left[F(X_{t_*}(x))\right] \leq aF(x), \quad \text{for } |x| \geq R_*, $$

$$\mathbb{E}\left[F(X_{t}(x))\right] \leq C_*, \quad \text{for } |x| \leq R_*, \forall t \geq 0.$$

- $H_2$. The following equality holds,

$$\lim_{y \to x} \sup_{|f| \leq 1} \left[ P_t f(x) - P_t f(y) \right] = 0, \quad \forall x \in \mathbb{R}^d,$$
or the following inequality holds,
\[ |\nabla P_t f(x)| \leq C_t(|x|)|f|_{\infty}, \quad \forall f \in \mathcal{B}(\mathbb{R}^d), \]
here \( C_t(\cdot) \) is a locally bounded function from \( \mathbb{R}^+ \) to \( \mathbb{R}^+ \) for fixed \( t \) and
\[ \|f\|_{\infty} := \text{ess. sup}_{x \in \mathbb{R}^d} |f(x)|. \]

- **H3 (Weak Form of Irreducibility).** For any \( R, \delta > 0 \), there exist positive constants \( R_0 := R_0(R) > 0 \) and \( T_0 := T_0(R, \delta) \) such that for any \( t \geq T_0 \) and any \( x, y \in B_R := \{ u \in \mathbb{R}^d, |u| \leq R \} \),
\[
\sup_{x \in \Gamma(P_t \delta, P_t \delta)} \pi \{ (u, v) \in \mathbb{R}^d \times \mathbb{R}^d, |u - v| \leq \delta, u, v \in B_{R_0} \} > 0,
\]
here \( P_t \delta \) is a measure on \( \mathbb{R}^d \) such that \( P_t \delta(A) = P_t(x, A), \forall A \in \mathcal{B}(\mathbb{R}^d) \) and \( \Gamma(\mu, \nu) \) denotes the set of probability measures \( \pi \) on \( \mathbb{R}^d \times \mathbb{R}^d \) such that \( \pi(A \times \mathbb{R}^d) = \mu(A), \pi(\mathbb{R}^d \times A) = \nu(A), \forall A \in \mathcal{B}(\mathbb{R}^d) \).

Hairer and Mattingly \cite{5, 6} gave the concept of weak form of irreducibility first. Our abstract ergodic result in this article is the following theorem, its proof will be given in Section 2.

**Theorem 1.1.** Assume Hypotheses \( H_1, H_2 \) and \( H_3 \) hold, then the process \( X_t \) is exponential ergodic under the total variation, i.e., there exist a unique invariant probability measure \( \mu \) for \( P_t \) and some positive constants \( \theta, C \) such that for any \( x \in \mathbb{R}^d \),
\[
\|P_t(x, \cdot) - \mu\|_{\text{var}} \leq Ce^{-\theta t}(1 + F(x)), \quad \forall t \geq 0, \tag{1.1}
\]
i.e.,
\[
\sup_{\|f\|_{\infty} \leq 1} |P_t f(x) - \mu(f)| \leq Ce^{-\theta t}(1 + F(x)), \quad \forall t \geq 0.
\]

Although the Hypothesis \( H_2 \) is stronger than \( H_b \), there is no essential difference between the verification of Hypothesis \( H_2 \) and Hypothesis \( H_b \) for stochastic differential equations. Usually, the sufficient condition for Hypothesis \( H_b \) is Hörmander condition in Brownian case or some similar condition in Lévy case, for example, \cite{3, 7, 20} etc. Under these conditions, \( H_2 \) also holds. For this, one can see \cite{2, Theorem 4.2} for the Brownian case, \cite{3, Remark 1.2} and Theorem 1.3\] and \cite{20, Theorem 1.1} for the Lévy case.

Under Hörmander condition or some similar condition, the Malliavin matrix of the solution to SDEs is almost invertible (cf. \cite{16}). Once the Malliavin matrix is almost invertible, the continuity of \( X_t(x) \) with respect to \( x \) under the total variation follows. For this, one can see \cite{1, Corollary 9.6.12} \cite{3, Theorem 1.1} for more details.

The advantage of Hypothesis \( H_3 \) is that, we only need to give some moment estimates such as
\[
\mathbb{E}|X_t(x) - X_t(y)|^p \quad \text{and} \quad \mathbb{E}|X_t(x)|^p, \quad p > 0
\]
in some situations, and don’t need to verify Hypothesis $H_a$, which is always complicated and
needs control theory, one can see Section 3 and [18, Theorem 6.1] for more details.

Our Hypotheses $H_2$ and $H_3$ in Theorem [1.1] are weaker than Hypotheses $H_a, H_b$ in some
situations. In order to show this, we consider the following simple example.

**Example 1.1.** Fix $k > 0, B \neq 0$. Consider the following SDEs,

\[
\begin{align*}
\text{d}X_t &= -Y_t^2 \text{d}t - kX_t \text{d}t \\
\text{d}Y_t &= -kY_t \text{d}t + BdW_t \\
(X_0, Y_0)|_{t=0} &= (x, y) \in \mathbb{R} \times \mathbb{R},
\end{align*}
\]  

(1.2)

then the Hypotheses in Theorem [1.1] hold, but the Hypothesis $H_a$ doesn’t hold.

**Proof.** It is easy to see that

\[
\begin{align*}
Y_t &= Y_t(x, y) = e^{-kt}y + Be^{-kt} \int_0^t e^{ks} \text{d}W_s, \\
X_t &= X_t(x, y) = e^{-kt}x - \int_0^t e^{-k(t-s)}Y_s^2 \text{d}s.
\end{align*}
\]  

(1.3)

(1) The verification of $H_1$. Let $F(x, y) = |x| + |y|^2 + 1$. By (1.3), we obtain

\[
\mathbb{E}[F(X_t, Y_t)] = \mathbb{E}|Y_t|^2 + \mathbb{E}|X_t| + 1
\]

\[
\leq \mathbb{E}|Y_t|^2 + e^{-kt}|x| + \int_0^t e^{-k(t-s)}Y_s^2 \text{d}s + 1
\]

\[
\leq e^{-2kt}|y|^2 + \frac{B^2}{2k} + e^{-kt}|x| + \int_0^t e^{-k(t-s)} \left| e^{-2ks}y^2 + \frac{B^2}{2k} \right| \text{d}s + 1
\]

\[
\leq e^{-2kt}|y|^2 + \frac{B^2}{2k} + e^{-kt}|x| + |y|^2 \frac{1}{k}(1 - e^{-kt}) + \frac{B^2}{2k^2}(1 - e^{-kt}) + 1,
\]

which gives the desired result.

(II) The verification of $H_2$. By [2, Theorem 4.2], $H_2$ holds.

(III) The verification of $H_3$. Let $B_t$ be a Brownian motion which is independent of $W_t$. For
any $(x', y') \in \mathbb{R}^2$. Let $(\hat{X}_t, \hat{Y}_t)$ be the solution to Eq.(1.2) with initial value $(x', y')$ and noise $B_t$, that is

\[
\begin{align*}
\hat{Y}_t &= \hat{Y}_t(x', y') = e^{-kt}y' + Be^{-kt} \int_0^t e^{ks} \text{d}B_s, \\
\hat{X}_t &= \hat{X}_t(x', y') = e^{-kt}x' - \int_0^t e^{-k(t-s)}\hat{Y}_s^2 \text{d}s.
\end{align*}
\]  

(1.4)

For any $R, \delta > 0$ and $(x, y), (x', y') \in B_R := \{u \in \mathbb{R}^2, |u| \leq R\}$, it is easy to see that there exists
a $T_0 := T_0(R, \delta)$ such that for any $t \geq T_0$,

\[
e^{-kt}|x - x'| < \delta.
\]  

(1.5)
For any $R_0, \delta, \varepsilon > 0$ and $t > T_0$, let $h, \tilde{h} \in C([0, t], \mathbb{R})$ be two functions such that
\[
h_0 = y, \quad \tilde{h}_0 = y', \quad |h_t - \tilde{h}_t| < \frac{\delta}{2}, \quad |h_t| \leq \frac{R_0}{2}, \quad |\tilde{h}_t| \leq \frac{R_0}{2},
\]
\[
\int_0^t e^{-ks}|h_t^2 - \tilde{h}_t^2|ds < \frac{\delta}{8},
\]
and denote
\[
B^h_{\varepsilon} = \{u(\cdot) \in C([0, t], \mathbb{R}) : u(0) = y, \sup_{s \in [0, t]} |u(s) - h_s| \leq \varepsilon \land \frac{R_0}{2}\},
\]
\[
B^\tilde{h}_{\varepsilon} = \{u(\cdot) \in C([0, t], \mathbb{R}) : u(0) = y', \sup_{s \in [0, t]} |u(s) - \tilde{h}_s| \leq \varepsilon \land \frac{R_0}{2}\}.
\]
Then there exists a positive constant $\varepsilon = \varepsilon(t, \delta, k)$, such that for any $u(\cdot) \in B^h_{\varepsilon}$, $\tilde{u} \in B^\tilde{h}_{\varepsilon}$
\[
\int_0^t e^{-ks}|u_s^2 - \tilde{u}_s^2|ds < \delta,
\]
\[
|u_t - \tilde{u}_t| < \delta,
\]
\[
|u_t| < R_0, \quad |\tilde{u}_t| < R_0.
\]
By the independence of $\{B_s\}_{s \in [0, t]}$ and $\{W_s\}_{s \in [0, t]}$, the definitions of $B^h_{\varepsilon}$, $B^\tilde{h}_{\varepsilon}$, and (1.3)(1.4) (1.5)(1.6), we derive that
\[
\mathbb{P}(|X_t - \tilde{X}_t| < 2\delta, |Y_t - \tilde{Y}_t| < 2\delta, |Y_t| \leq R_0, |\tilde{Y}_t| \leq R_0)
\]
\[
\geq \mathbb{P}((Y_s)_{s \in [0, t]} \in B^h_{\varepsilon}, (\tilde{Y}_s)_{s \in [0, t]} \in B^\tilde{h}_{\varepsilon})
\]
\[
\geq \mathbb{P}((Y_s)_{s \in [0, t]} \in B^h_{\varepsilon}) \cdot \mathbb{P}((\tilde{Y}_s)_{s \in [0, t]} \in B^\tilde{h}_{\varepsilon})
\]
\[
> 0,
\]
where in the last inequality we have used [17] Theorem 3.2. This completes the verification of $H_3$.

(IV) By (1.3), we obtain
\[
\mathbb{P}(X_t \leq e^{-kt}x) = 1.
\]
Therefore, $H_d$ doesn’t hold. □

Now, in the end of this subsection, we introduce the ideas in the proof of Theorem 1.1 in Section 2.

The main tool to prove Theorem 1.1 is the coupling method. One can see [22] for an introduction of the coupling method. In the proof of Theorem 1.1 we also borrow some ideas from [23]. For any $x, y \in \mathbb{R}^d$, Hypothesis $H_1$ is used to ensure that the processes $(X_t(x), X_t(y))$ enter a ball $B_{R'_1} = \{u \in \mathbb{R}^d \times \mathbb{R}^d, |u| \leq R'_1\}$ very quickly, see lemmas 2.2 2.4 below for more details. Denote $\tilde{\tau}$ the time of the two processes $(X_t(x), X_t(y))$ enter this ball $B_{R'_1}$.
Hypothesis $H_3$ is used to ensure that
\[
\mathbb{P}\left(\left|X_{\tilde{\tau}+T}(x) - X_{\tilde{\tau}+T}(y)\right| \text{ is very small}
\right.
\text{ and } X_{\tilde{\tau}+T}(x), X_{\tilde{\tau}+T}(y) \text{ stay at some ball}\bigg) > 0
\]
holds for some big but finite $T$. Denote $u_x = X_{\tilde{\tau}+T}(x), u_y = X_{\tilde{\tau}+T}(y)$.

By Hypothesis $H_2$, one finds a $t_0 > 0$ such that (see Lemma 2.3 below for more details),
\[
\|P_{t_0}(u_x, \cdot) - P_{t_0}(u_y, \cdot)\|_{\text{var}} = \min \mathbb{P}(Z_1 \neq Z_2) < 1,
\]
the minimum is taken over all couplings $(Z_1, Z_2)$ of $(P_{t_0}(u_x, \cdot), P_{t_0}(u_y, \cdot))$. Then, there is a coupling $(Z_1, Z_2)$ of $(X_0(u_x), X_0(u_y))$ such that
\[
\mathbb{P}(Z_1 = Z_2) > 0,
\]
which means with a positive probability, the coupling time is $\tilde{\tau} + t_0$.

By the arguments of the above and strong Markov property, the coupling of $X_t(x), X_t(y)$
will be successful with fast speed. This completes the proof of Theorem 1.1.

### 1.3. Application: The SDEs Driven by Lévy Processes

Consider the following stochastic differential equation (abbreviated as SDE) in $\mathbb{R}^d$
\[
dX_t = b(X_t)dt + A_1dW_t + A_2dL_t, \quad X_0 = x, \quad (1.7)
\]
where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth vector field, $A_1$ and $A_2$ are two constant $d \times d$-matrices, $W_t$
is a $d$-dimensional standard Brownian motion and $L_t$ is a purely jump d-dimensional Lévy
process with Lévy measure $\nu(dz)$. Let $\Gamma_0 := \{z \in \mathbb{R}^d : 0 < |z| < 1\}$. Throughout this work, we assume that $\frac{\nu(dz)}{dz}|_{z=0} = \kappa(z)$ satisfies the following conditions: for some $\alpha \in (0, 2)$ and $m \in \mathbb{N}$,$(H_m^\alpha) : \kappa \in C^m(\Gamma_0, (0, \infty))$ is symmetric (i.e. $\kappa(-z) = \kappa(z)$) and satisfies the following Oreys
order condition (cf. [19] Proposition 28.3))
\[
\lim_{\epsilon \to 0} \epsilon^{-\alpha/2} \int_{|z| \leq \epsilon} |z|^2 \kappa(z) dz := c_1 > 0,
\]
and bounded condition: for $j = 1, \cdots, m$ and some $C_j > 0$,
\[
|\nabla^j \log \kappa(z)| \leq C_j |z|^{-j}, \quad z \in \Gamma_0.
\]
Let $A^*$ be the transpose of $A$, and
\[
\nabla^2_{A_1 A_1^*} f := \sum_{i,j=1}^d (A_1 A_1^*)_{ij} \partial^2_{ij} f,
\]
Let $B_0 = I_{d \times d}$ be the identity matrix and define for $n \in \mathbb{N}$,
\[
B_n(x) := b(x) \cdot \nabla b_{n-1}(x) - \nabla b(x) \cdot B_{n-1}(x) + \frac{1}{2} \nabla^2_{A_1 A_1^*} B_{n-1}(x).
\]
Here and below $(\nabla b)_{ij} := \partial_j b^i(x)$. 


Let $P_t(x, \cdot)$ be the transition probability associated with equation (1.7), that is for any $A \in \mathcal{B}(\mathbb{R}^d)$, $P_t(x, A) = \mathbb{P}(X_t \in A)$.

Apply Theorem 1.1 to Eq.(1.7), the following theorem holds.

**Theorem 1.2.** Assume ($\mathbf{H}_1^n$) holds, $\int_{\mathbb{R}^d} |z|^2 \nu(dz) < \infty$, and

1. for some $k > 0$,
   \[ \langle x - y, b(x) - b(y) \rangle \leq -k|x - y|^2, \tag{1.8} \]
   here $\langle \cdot, \cdot \rangle$ denotes the usual inner product on $\mathbb{R}^d$,
2. for any $x \in \mathbb{R}^d$, there exists some $n = n(x)$ such that
   \[ \text{Rank}\{A_1, B_1(x)A_1, \cdots, B_n(x)A_1, A_2, B_2(x)A_2, \cdots, B_n(x)A_2\} = d, \tag{1.9} \]
   then there exist a unique invariant probability measure $\mu$ for $P_t$ and a positive constant $\theta$ such that for any $x \in \mathbb{R}^d$,
   \[ \|P_t(x, \cdot) - \mu\|_{\text{var}} \leq Ce^{-\theta t}(1 + |x|^2), \quad \forall t \geq 0. \]

The proof of Theorem 1.2 will be given in Section 3.

**Remark 1.1.** When ($\mathbf{H}_1^n$) and (1.9) hold, Hypothesis $\mathbf{H}_2$ also holds by the proof of [20, Theorem 1.1]. (1.8) is used to verify Hypotheses $\mathbf{H}_1, \mathbf{H}_3$.

**Remark 1.2.** When $\|A^{-1}_1\| \leq \lambda$ for some $\lambda > 0$, here $A^{-1}_1$ denotes the inverse of $A_1$, Lan and Wu [8, Theorem 1.3] proved that $X_t$ is irreducible aperiodic. But generally speaking, (1.9) doesn’t imply $\mathbf{H}_2$. For this, one can see Example 1.1.

2. **The proof of Theorem 1.1**

This section is organized as follows. In subsection 2.1, we will give a construction of the coupling Markov chain and list some lemmas. In these lemmas listed in this section, we have already assumed that Hypotheses $\mathbf{H}_1, \mathbf{H}_2$ and $\mathbf{H}_3$ hold. In subsection 2.2, we will give a proof of Theorem 1.1. For any $x = (x_1, \cdots, x_d), y = (y_1, \cdots, y_d) \in \mathbb{R}^d$, $|x| := \sqrt{\sum_{i=1}^d x_i^2}$, $|(x, y)| := \sqrt{|x|^2 + |y|^2}$.

2.1. **Construction of the coupling Markov chain and some lemmas.** Denote by $\mathcal{D}(X)$ the law of random variable $X$. Let $(\Lambda_1, \Lambda_2)$ be two probability measures on a metric space $(E, \mathcal{E})$, here $\mathcal{E}$ denotes the Borelian subsets of $E$. Let $(Z_1, Z_2)$ be two random variables $(\Omega, \mathcal{F}) \to (E, \mathcal{E})$, we say $(Z_1, Z_2)$ is a coupling of $(\Lambda_1, \Lambda_2)$ if $\Lambda_1 = \mathcal{D}(Z_1), \Lambda_2 = \mathcal{D}(Z_2)$. Let $\Lambda$ be a sign measure on $(E, \mathcal{E})$, the total variation is given by

\[ ||\Lambda||_{\text{var}} := \sup\{||\Lambda(\Gamma)|| \mid \Gamma \in \mathcal{E}\}. \tag{2.1} \]

In this subsection, we first recall a fundamental result in the coupling methods, and then we list some lemmas which will be used in the proof of Theorem 1.1.
Lemma 2.1. (III etc.) Let \((\Lambda_1, \Lambda_2)\) be two probability measures on a metric space \((E, \mathcal{E})\). Then

\[ \|\Lambda_1 - \Lambda_2\|_{\text{var}} = \min \mathbb{P}(Z_1 \neq Z_2). \]

The minimum is taken over all couplings \((Z_1, Z_2)\) of \((\Lambda_1, \Lambda_2)\). There exists a coupling which reaches the minimum value. It is called a maximal coupling.

Lemma 2.2. (i) For any \(x \in \mathbb{R}^d\) and \(k \in \mathbb{N}\), we have

\[ \mathbb{E}[F(X_{kt}(x))] \leq a^k F(x) + C_* \frac{1}{1 - a}. \]

(ii) There exist positive constants \(R'_*, C'_*\) and \(a' \in (0, 1)\) such that for any \(k \in \mathbb{N}\),

\[ \mathbb{E}[F(X_{kt}(x)) + F(X_{kt}(y))] \leq a'[F(x) + F(y)], \text{ for } |(x, y)| \geq R'_*, \]

\[ \mathbb{E}[F(X_{kt}(x)) + F(X_{kt}(y))] \leq C'_*, \quad \text{for } |(x, y)| \leq R'_*. \]

Proof. By Hypothesis \(H_1\), we obtain

\[ \mathbb{E}[F(X_{kt}(x))] = \mathbb{E}[\mathbb{E}[F(X_{kt}(x))|\mathcal{F}_{(k-1)t}]] \]

\[ \leq a \mathbb{E}[F(X_{(k-1)t}(x))] + C_* \]

\[ \leq a [a \mathbb{E}[F(X_{(k-2)t}(x))] + C_*] + C_* \]

\[ = a^2 \mathbb{E}[F(X_{(k-2)t}(x))] + C_*(a + 1) \]

\[ \leq \cdots \]

\[ \leq a^k \mathbb{E}[F(X_{(k-k)t}(x))] + C_*(a^{k-1} + a^{k-2} + \cdots + 1) \]

\[ = a^k F(x) + C_*(a^{k-1} + a^{k-2} + \cdots + 1) \]

\[ \leq a^k F(x) + C_* \frac{1}{1 - a}. \quad (2.2) \]

(2.2) implies that,

\[ \mathbb{E}[F(X_{kt}(x))] + \mathbb{E}[F(X_{kt}(y))] \leq a^k(F(x) + F(y)) + C_* \frac{2}{1 - a}. \]

Noticing that \(a \in (0, 1)\) and

\[ \lim_{|u| \to \infty} F(u) = +\infty, \]

we get the desired the result (ii).

Recall that \(R'_* > 0\) is a constant given by Lemma 2.2.

Lemma 2.3. There exists \(T_0 = T_0(R'_*) > 1\) such that for any \(x, y \in B_{R'_*}\) and any \(t > T_0\),

\[ \|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{var}} = \sup_{0 \leq f \leq 1} \left[ P_t f(x) - P_t f(y) \right] < 1, \quad (2.3) \]
and furthermore
\[ p := \sup_{|u| \leq R', |v| \leq R'} \| P_t(x, \cdot) - P_t(y, \cdot) \|_{\text{var}} < 1. \]  
(2.4)

**Proof.** The proof of (2.3). By Hypothesis H3, for any \( \delta > 0 \), there exist \( R_0 = R_0(R'), T_0 = T_0(R', \delta) > 1 \) such that for any \( x, y \in B_{R'} \) and \( t > T_0 \)

\[ \sup_{\pi \in \Gamma(P_{t-\delta}(\delta), P_{t-\delta}(\delta))} \pi \left\{ (u, v) \in \mathbb{R}^d \times \mathbb{R}^d : |u - v| \leq \delta, u, v \in B_{R_0} \right\} > 0. \]  
(2.5)

We emphasis that \( R_0 \) is independent of \( \delta \), and the constant \( \delta \) will be given in the next paragraph.

By H2, there exists a constant \( \delta > 0 \), such that

\[ \sup_{0 < f < 1} \left[ P_1 f(u) - P_1 f(v) \right] \leq \frac{1}{4} \]  
(2.6)

holds for any \( u, v \in B_{R_0} \) with \( |u - v| < \delta \).

For any \( \pi \in \Gamma(P_{t-1}(\delta), P_{t-1}(\delta)) \) and \( f \) with \( 0 \leq f \leq 1 \), we have

\[
\begin{align*}
& P_t f(x) - P_t f(y) \\
& = \int_{\mathbb{R}^d \times \mathbb{R}^d} \pi(du, dv) \left[ P_1 f(u) - P_1 f(v) \right] \\
& = \int_{\Theta} \pi(du, dv) \left[ P_1 f(u) - P_1 f(v) \right] + \int_{\Theta^c} \pi(du, dv) \left[ P_1 f(u) - P_1 f(v) \right] \\
& \leq \frac{1}{2} \pi(\Theta) + \pi(\Theta^c)
\end{align*}
\]

here \( \Theta = \{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d : |u - v| \leq \delta, u, v \in B_{R_0}\}, \Theta^c = (\mathbb{R}^d \times \mathbb{R}^d) \setminus \Theta \). Combing the above inequality, (2.6) and (2.5), we obtain (2.3).

**The proof of (2.4).** By Hypothesis H2, (2.3) and the compactness of the set \( \{(u, v) \in \mathbb{R}^d \times \mathbb{R}^d : |u| \leq R', |v| \leq R'\} \), one arrives at (2.4).

Let \( T = \left[ \frac{T_0 + 1}{t_*} + 1 \right] \cdot t_* := k_0 \cdot t_* \), here \( [b] \) denotes an integer with \( b - 1 < [b] \leq b, t_* > 0 \) appears in Hypothesis H1 and \( T_0 > 0 \) is given by lemma 2.3 then \( T > T_0 + 1 \). Let \( P_{x,y}(\cdot) \) be the law of the maximal coupling of \( P_T(x, \cdot) \) and \( P_T(y, \cdot) \).

**Proposition 2.1.** There exist a probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\) and an \( \mathbb{R}^d \times \mathbb{R}^d \) valued Markov chain \( \{S(k)\}_{k \geq 0} \) on \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\) with transition probability family \( P_{x,y}(\cdot)(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \). Moreover, for every \( (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \), the marginal chain \( \{S^x(k)\}_{k \geq 0} \) has the same distribution as \( \{X_{kT}(x)\}_{k \geq 0} \) and the marginal chain \( \{S^y(k)\}_{k \geq 0} \) has the same distribution as \( \{X_{kT}(y)\}_{k \geq 0} \).

The sequence \( \{S(k)\}_{k \geq 0} \) constructed above is a Markov chain on the probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\). \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\) is not necessarily the same as \((\Omega, \mathcal{F}, P)\). Without loss of generality, we assume that \( (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}) = (\Omega, \mathcal{F}, P) \).
Otherwise, we can consider the product space $(\tilde{\Omega} \times \Omega, \tilde{\mathcal{F}} \times \mathcal{F}, \tilde{\mathbb{P}} \times \mathbb{P})$.

For any $k \in \mathbb{N}$, define stopping times

$$
\tau_1 = \inf \{m > 0, |(S^x(m), S^y(m))| \leq R'_1\},
$$

$$
\tau_{k+1} = \inf \{m > \tau_k, |(S^x(m), S^y(m))| \leq R'_1\},
$$

$$
\tau_{x,y} = \inf \{m \geq 0, S^x(m) = S^y(m)\}.
$$

In this article, we use the letter $C$ to denote an unimportant constant whose value may change in different places.

**Lemma 2.4.** For some positive constant $\theta$, we have

$$
\mathbb{E}[e^{\theta \tau_1}] \leq C(1 + F(x) + F(y)),
$$

$$
\mathbb{E}[e^{\theta \tau_k}] \leq C^k [1 + F(x) + F(y)].
$$

**Proof.** By [21, Proposition 3.1], Lemma 2.3 and the fact $T = k_0 t_*$, there exists $\theta > 0$ such that

$$
\mathbb{E}\exp(\theta \tau_1) \leq C(1 + F(x) + F(y)).
$$

Let $u_x = S^x(\tau_{k-1}), u_y = S^y(\tau_{k-1})$. By strong Markov property and the above inequality, we get

$$
\mathbb{E}\exp(\theta \tau_k) = \mathbb{E}\left[\mathbb{E}\exp(\theta \tau_k) | \mathcal{F}_{\tau_{k-1}}\right]
$$

$$
= \mathbb{E}\left[\mathbb{E}\exp(\theta \tau_k - \theta \tau_{k-1} + \theta \tau_{k-1}) | \mathcal{F}_{\tau_{k-1}}\right]
$$

$$
= \mathbb{E}\left[\exp(\theta \tau_{k-1}) \cdot \mathbb{E}\left[\exp(\theta \tau_k - \theta \tau_{k-1}) | \mathcal{F}_{\tau_{k-1}}\right]\right]
$$

$$
\leq \mathbb{E}\left[\exp(\theta \tau_{k-1}) \cdot \left(C + CF(u_x) + CF(u_y)\right)\right]
$$

$$
\leq C \mathbb{E}\left[\exp(\theta \tau_{k-1})\right],
$$

which gives the desired result. \[\square\]

**Lemma 2.5.** For any $k \in \mathbb{N}$, we have

$$
\mathbb{P}(\tau_{x,y} > \tau_k + 1) \leq p^k,
$$

here $p$ appears in Lemma 2.3.
Proof. By Lemmas 2.1, 2.3, we have
\[
\begin{align*}
\mathbb{P}(\tau_{xy} > \tau_k + 1) & \leq \mathbb{P}(S^y(k) \neq S^y(k), k = 0, \ldots, \tau_k + 1) \\
& = \mathbb{E}\left[\mathbb{P}(S^y(k) \neq S^y(k), k = 0, \ldots, \tau_k + 1 | S^y(k), \tau_k)\right] \\
& = \mathbb{E}\left[I_{(S^y(k) \neq S^y(k), k = 0, \ldots, \tau_k)} \mathbb{P}(S^y(\tau_k + 1) \neq S^y(\tau_k + 1) | S^y(\tau_k))\right] \\
& \leq p \mathbb{P}(S^y(k) \neq S^y(k), k = 0, \ldots, \tau_k) \\
& \leq p \mathbb{P}(S^y(k) \neq S^y(k), k = 0, \ldots, \tau_{k-1} + 1) \\
& \leq \ldots \\
& \leq p^k.
\end{align*}
\]
\[\Box\]

Lemma 2.6. There exists a constant \(\varepsilon > 0\), such that for any \(x \neq y\),
\[
\mathbb{E}\left[e^{\varepsilon \tau_{xy}}\right] \leq C(1 + F(x) + F(y)).
\]

Proof. Denote \(\tau_{-1} \equiv -1, \tau_0 \equiv 0\). For any \(\varepsilon > 0\) and \(p' > 1, q' > 1\) with \(\frac{1}{p'} + \frac{1}{q'} = 1\), we have
\[
\begin{align*}
\mathbb{E}\left[e^{\varepsilon \tau_{xy}}\right] & = \sum_{k=0}^{\infty} \mathbb{E}\left[e^{\varepsilon \tau_{xy}} I_{(\tau_k + 1 < \tau_{xy} < \tau_{k+1})}\right] \\
& \leq \sum_{k=0}^{\infty} \mathbb{E}\left[e^{\varepsilon (\tau_k + 1)}\right]^{1/p'} \mathbb{P}(\tau_{xy} > \tau_{k-1} + 1)^{1/q'} \\
& \leq \sum_{k=0}^{\infty} C \mathbb{E}\left[e^{\varepsilon (\tau_k + 1)}\right]^{1/p'} p^{k/q'}.
\end{align*}
\]

Setting \(\varepsilon \leq \theta\) and \(p'\) big enough, one arrives at that
\[
\begin{align*}
\mathbb{E}\left[e^{\varepsilon \tau_{xy}}\right] & \leq \sum_{k=1}^{\infty} C \cdot C_1 \frac{k^{p'}}{p'} \left[1 + CF(x) + CF(y)\right]^{1/p'} \cdot p^{k/q'} \\
& \leq C(1 + F(x) + F(y)),
\end{align*}
\]
here in the first inequality, we have used Lemma 2.4 \[\Box\]

2.2. The proof of Theorem 1.1

Proof. By Lemma 2.6 for some \(\theta > 0\) and any \(f \in \mathcal{F}_b(\mathbb{R}^d)\) with \(\|f\|_{\infty} \leq 1\), we obtain
\[
\mathbb{E}\left[f(S^y(k)) - f(S^y(k))\right] \leq Ce^{-\theta k}(1 + F(x) + F(y)).
\]

Since the marginal chain \(\{S^y(k)\}_{k \geq 0}\) has the same distribution as \(\{X_{kt}(x)\}_{k \geq 0}\) and the marginal chain \(\{S^y(k)\}_{k \geq 0}\) has the same distribution as \(\{X_{kt}(y)\}_{k \geq 0}\), then for any \(\|f\|_{\infty} \leq 1\), we have
\[
\mathbb{E}\left[f(X_{kt}(x)) - f(X_{kt}(y))\right] \leq Ce^{-\theta k}(1 + F(x) + F(y)).
\]
(2.7)
For any $t > 0$ and $f$ with $\|f\|_\infty \leq 1$, let $k_t = [\frac{t}{\theta}]$, $\tilde{f} = P_{t-k_tf}$. By (2.7), we get

$$P_tf(x) - P_tf(y) = \mathbb{E}\left[P_{t-k_tf}(X_{kT}(x)) - P_{t-k_tf}(X_{kT}(y))\right]$$

$$= \mathbb{E}\left[\tilde{f}(X_{kT}(x)) - \tilde{f}(X_{kT}(y))\right]$$

$$\leq Ce^{-\theta k}(1 + F(x) + F(y))$$

$$\leq Ce^{-\theta k}(1 + F(x) + F(y))$$

$$\leq Ce^{-\theta t}(1 + F(x) + F(y)).$$

Hence, we have proven that the following equality holds for some $\theta > 0$ and any $t > 0$, $f \in \mathcal{B}_b(\mathbb{R}^d)$ with $\|f\|_\infty \leq 1$,

$$P_tf(x) - P_tf(y) \leq C\left[1 + F(x) + F(y)\right]e^{-\theta t}. \quad (2.8)$$

According to [21, Section 2.2], (2.8) implies (1.1) which finishes the proof of Theorem 1.1. For the convenience of reading, we still give its details here.

Denote by $\langle f, P_t(x, \cdot) \rangle = P_t f(x)$. For any $s > t$ and $f \in C_b(\mathbb{R}^d)$ with $\|f\|_\infty \leq 1$, it is not difficult to see that

$$|\langle f, P_t(x, \cdot) - P_s(y, \cdot) \rangle| = \left|\int_{\mathbb{R}^d} P_{s-t}(y, dz) \int_{\mathbb{R}^d} (P_t(x, dv) - P_s(z, dv)) f(v)\right|$$

$$\leq Ce^{-\theta t} \int_{\mathbb{R}^d} P_{s-t}(y, dz) (1 + F(x) + F(z))$$

$$= Ce^{-\theta t} \left(1 + F(x) + \mathbb{E}_y[F(X_{s-t})]\right). \quad (2.9)$$

By the Prokhorov theorem, $\mathcal{P}(\mathbb{R}^d)$ is a complete metric space under the weak topology (here $\mathcal{P}(\mathbb{R}^d)$ denotes the set of all probability measures on $\mathbb{R}^d$). Let $y = 0$ and $s \to +\infty$ in (2.9), for any $f \in C_b(\mathbb{R}^d)$ with $\|f\|_\infty \leq 1$, one arrives at that

$$|\langle f, P_t(x, \cdot) - \mu \rangle| \leq Ce^{-\theta t}(1 + F(x)),$$

which implies (1.1). \hfill \Box

3. The proof of Theorem 1.2

By Theorem 1.1, we only need to verify Hypotheses $H_1, H_2$ and $H_3$. In the next three subsections, we will give a verification of Hypotheses $H_1, H_2$ and $H_3$ respectively.

3.1. The verification of Hypothesis $H_1$. Let $N$ be a Poisson random measure with density $d\nu(dz)$, where $\nu(dz) = \kappa(z)dz$ and let

$$L_t = \int_0^t \int_{|z|<1} z\tilde{N}(dsdz) + \int_0^t \int_{|z|>1} zN(dsdz),$$

$$\tilde{N}(drdz) = N(drdz) - d\nu(dz).$$
By Itô’s formula, we have

\[ |X_t(x)|^2 = |x|^2 + 2 \int_0^t \langle X_{s-}^x, b(X_s) \rangle ds + 2 \int_0^t \langle X_{s-}^x, A_1 dW_s \rangle + \int_0^t Tr(A_1 \cdot A_i^i) ds \]

\[ + \int_0^t \int_{|z| < 1} \left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 \right] N(ds dz) + \int_0^t \int_{|z| \geq 1} \left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 \right] N(ds dz) \]

\[ + \int_0^t \int_{|z| < 1} \left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 - \sum_{i=1}^d (A_2 z, X^i_i(s_i)) \right] ds v(dz). \]

For any \( \varepsilon > 0 \), one easily sees that

\[ \langle X_{s-}^x, A_2 z \rangle \leq \varepsilon |X_{s-}^x|^2 + C(\varepsilon) |z|^2, \forall \varepsilon > 0, \]

\[ \langle b(x), x \rangle = \langle b(x) - b(0), x - 0 \rangle + \langle b(0), x \rangle \leq -k|x|^2 + |b(0)| \cdot |x|, \]

and there exists \( C = C(\varepsilon) \) such that the following inequality holds,

\[ |X_t(x)|^2 \leq |x|^2 + 2 \int_0^t \left[ -k|X_{s-}^x|^2 + |b(0)| \cdot |X_{s-}^x| \right] ds + 2 \int_0^t \langle X_{s-}^x, A_1 dW_s \rangle + \int_0^t C ds \]

\[ + \int_0^t \int_{|z| < 1} \left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 \right] N(ds dz) + \int_0^t \int_{|z| \geq 1} \left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 \right] N(ds dz) \]

\[ + \int_0^t \int_{|z| < 1} \left[ \varepsilon |X_{s-}^x|^2 + C|z|^2 \right] ds v(dz). \]

Setting \( \varepsilon \) small enough and combing the above inequality with (3.1), one arrives at that for some \( C = C(k) \),

\[ \mathbb{E}[|X_t(x)|^2] \]

\[ \leq |x|^2 + \int_0^t \left[ -\frac{3k}{2} \mathbb{E}|X_{s-}^x|^2 + C \right] ds + \int_0^t \int_{|z| < 1} \mathbb{E}\left[ |X_{s-}^x + A_2 z|^2 - |X_{s-}^x|^2 \right] ds v(dz) \]

\[ \leq |x|^2 + \int_0^t \left[ -\frac{3k}{2} \mathbb{E}|X_{s-}^x|^2 + C \right] ds + \int_0^t \int_{|z| > 1} \mathbb{E}\left[ |X_{s-}^x + A_2 z|^2 + C|z|^2 \right] ds v(dz) \]

\[ \leq |x|^2 + \int_0^t \left[ -\frac{3k}{2} |X_{s-}^x|^2 + C \right] ds + \int_0^t \int_{|z| > 1} \left[ \frac{k}{2} \mathbb{E}|X_{s-}^x|^2 + C|z|^2 \right] ds v(dz) \]

\[ \leq |x|^2 - k \int_0^t \mathbb{E}|X_{s-}^x|^2 ds + \int_0^t C ds. \]

By the Gronwall’s inequality, we obtain that

\[ \mathbb{E}[|X_t(x)|^2] \leq |x|^2 e^{-kt} + C, \]

which gives the desired result.
3.2. The verification of Hypothesis $H_2$. The main ideas in this subsection are borrowed from the proof of [20, Theorem 1.1] and the location method in [2, Theorem 4.2]. For any $\ell \in \mathbb{N}$, let $h_{\ell}(x) : \mathbb{R}^d \to \mathbb{R}$ be a smooth function with compact support and $h_{\ell}(x) = 1$ when $x \in B_\ell^\circ := \{u \in \mathbb{R}^d : |u| < \ell\}$. Let $X_t^\ell(x)$ be the solution to the following SDEs

$$
\frac{dX_t^\ell}{X_t^\ell} = h_{\ell}(X_t^\ell) b(X_t^\ell) dt + A_1 dW_t + A_2 dL_t, \quad X_0^\ell = x.
$$

(3.3)

Following the proof of [20, Theorem 1.1], one obtains that for any $x \in B_\ell^\circ$,

$$
\lim_{y \to x} \sup_{\|f\|_{\infty} \leq 1} \mathbb{E} \left[ f(X_t^\ell(y)) - f(X_t^\ell(x)) \right] = 0.
$$

Define a sequence of stopping time

$$
\tau_{\ell}(x) = \inf\{s > 0, \; X_s^\ell(x) \notin B_\ell^\circ\}, \quad \ell \in \mathbb{N}, \; x \in \mathbb{R}^d.
$$

By Appendix A, for any $x \in \mathbb{R}^d$, $l \geq 2$, $t > 0$, we have

$$
\lim_{y \to x} \sup_{\|f\|_{\infty} \leq 1} I_{[t,\tau_{\ell}(x)]} \leq I_{[t,\tau_{\ell-1}(x)]}, \; a.s.
$$

(3.4)

For any $x,y \in B_\ell^\circ$, one sees that

$$
\begin{align*}
\mathbb{E}[f(X_t(x)) - f(X_t(y))] & \leq \mathbb{E}[f(X_t(x))1_{t < \tau_{\ell}(x)} - f(X_t(y))1_{t < \tau_{\ell}(y)}] + \|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(x)) + \|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(y)) \\
& \leq \mathbb{E}[f(X_t^\ell(x))1_{t < \tau_{\ell}(x)} - f(X_t^\ell(y))1_{t < \tau_{\ell}(y)}] + \|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(x)) + \|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(y)) \\
& \leq \mathbb{E}[f(X_t^\ell(x)) - f(X_t^\ell(y))] + 2\|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(x)) + 2\|f\|_{\infty} \mathbb{P}(t \geq \tau_{\ell}(y)).
\end{align*}
$$

Therefore,

$$
\begin{align*}
\lim_{y \to x} \sup_{\|f\|_{\infty} \leq 1} \mathbb{E}[f(X_t(x)) - f(X_t(y))] & \leq 2\mathbb{P}(t \geq \tau_{\ell}(x)) + 2\lim_{y \to x} \sup_{\|f\|_{\infty} \leq 1} \mathbb{P}(t \geq \tau_{\ell}(y)) \\
& \leq 2\mathbb{P}(t \geq \tau_{\ell}(x)) + 2\mathbb{P}(t \geq \tau_{\ell-1}(x)),
\end{align*}
$$

let $\ell \to \infty$ in the above inequality and by (3.4), we obtain that

$$
\lim_{y \to x} \sup_{\|f\|_{\infty} \leq 1} \mathbb{E}[f(X_t(x)) - f(X_t(y))] = 0.
$$

3.3. The verification of Hypothesis $H_3$. By calculating, we get

$$
\begin{align*}
\frac{d|X_t(x) - X_t(y)|^2}{X_t(x) - X_t(y)} & = \langle f(X_t(x)) - f(X_t(y)), X_t(x) - X_t(y) \rangle dt \\
& \leq -2k|X_t(x) - X_t(y)|^2 dt,
\end{align*}
$$

which implies that

$$
|X_t(x) - X_t(y)|^2 \leq |x - y|^2 e^{-k\tau}.
$$

By the above inequality and (3.2), Hypothesis $H_3$ holds.

Appendix A: The Proof of (3.4)
Proof. Let \( \Gamma \) be a measurable set with \( \mathbb{P}(\Gamma^c) = 0 \) and such that \( X_t^0(x, \omega) \) is càdlàg with respect to \( s \) for \( \omega \in \Gamma \). For \( \omega \in \Gamma \), the conclusion is apparent if
\[
\limsup_{y \to x} I_{[t > \tau_{\ell}(y)]}(\omega) = 0 \text{ or } I_{[\tau_{\ell} - 1(x)]}(\omega) = 1.
\]
Assume that \( \limsup_{y \to x} I_{[t > \tau_{\ell}(y)]}(\omega) = 1 \) and \( I_{[\tau_{\ell} - 1(x)]}(\omega) = 0 \), then
\[
\sup_{s \in [0, t]} |X_{s}^{\ell-1}(x, \omega)| \leq \ell - 1.
\]
Therefore,
\[
\sup_{s \in [0, t]} |X_{s}^{\ell}(x, \omega)| \leq \ell - 1. \tag{3.5}
\]
Since \( \limsup_{y \to x} I_{[t > \tau_{\ell}(y)]}(\omega) = 1 \), then there exists a sequence \( \{x_n\} \subset \mathbb{R}^d \) with \( x_n \to x \) as \( n \to \infty \), such that for \( n \) large enough
\[
\sup_{s \in [0, t]} |X_{s}^{\ell}(x_n, \omega)| \geq l. \tag{3.6}
\]
On the other hand, by (3.5), we have
\[
d|X_{t}^{\ell}(x_n) - X_{t}^{\ell}(x)| \leq C|X_{t}^{\ell}(x_n) - X_{t}^{\ell}(x)|dt,
\]
and furthermore
\[
\sup_{s \in [0, t]} |X_{s}^{\ell}(x_n) - X_{s}^{\ell}(x)| \leq e^{Ct}|x_n - x|,
\]
which contradicts (3.5) and (3.6).

\[\square\]

References

[1] Bogachev, V. I. (2010). Differentiable Measures and the Malliavin Calculus[M]. American Mathematical Society, Providence, Rhode Island.

[2] Dong, Z., & Peng, X. (2014). Malliavin matrix of degenerate sde and gradient estimate. *Electronic Journal of Probability*, 19(73), 1-34.

[3] Dong, Z., Peng, X., Song, Y., & Zhang, X. (2016). Strong feller properties for degenerate sdes with jumps. *Annales de l’Institut Henri Poincaré - Probabilites et Statistiques*, 52(2), 888-897.

[4] Down, D., Meyn, S. P., & Tweedie, R. L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.*, 23 (4), 1671-1691.

[5] Hairer, M., & Mattingly, J. C. (2006). Ergodicity of the 2d navier-stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, 164(3), 993-1032.

[6] Hairer, M., & Mattingly, J. C. (2008). Spectral gaps in wasserstein distances and the 2d stochastic navier-stokes equations. *Annals of Probability*, 36(6), 2050-2091.

[7] Ichihara, K., & Kunita, H. (1974). A classification of the second order degenerate elliptic operators and its probabilistic characterization *Probability Theory and Related Fields*, 30(3), 235-254.

[8] Lan, G., Wu, J.L. (2014). Stochastic continuity, irreducibility and non confluence for SDEs with jumps[J]. *arXiv:1407.1658*
[9] Kusuoka, S., & Stroock, D. (1984). Applications of the malliavin calculus, part I. *North-Holland Mathematical Library*, 32(5), 271-306.

[10] Kusuoka, S., & Stroock, D. (1987). Applications of the Malliavin calculus, III, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34, 391-442.

[11] Lindvall, T. (1992). Lectures on the coupling method. Wiley, New York.

[12] Malliavin, P. (1978). Stochastic calculus of variation and hypoelliptic operators. *Proc.intern.symp.sde Kyoto*, 1976(1), 195-263.

[13] Meyn, S.P., & Tweedie, R.L. (1993). Stability of markovian processes iii: foster-lyapunov criteria for continuous time processes. *Advances in Applied Probability*, 25(3), 518-548.

[14] Meyn, S. P., & Tweedie, R. L. (1993). Stability of markovian processes ii: continuous-time processes and sampled chains. *Advances in Applied Probability*, 25(3), 487-517.

[15] Meyn, S.P., & Tweedie, R. L. (1993). Markov chains and stochastic stability. Springer-Verlag.

[16] Nualart, D. (2006). The Malliavin Calculus and Related Topics, Springer.

[17] Pakkanen, M. S. (2010). Stochastic integrals and conditional full support. *Journal of Applied Probability*, 47(3), 650-667.

[18] Rey-Bellet, L. (2006). Ergodic Properties of Markov Processes. Open Quantum Systems II. Springer Berlin Heidelberg.

[19] Sato, K.(1999). Lévy processes and infinite divisible distributions. Cambridge Univ. Press, Cambridge.

[20] Song, Y., & Zhang, X. (2014). Regularity of density for sdes driven by degenerate lévy noises. *Electronic Journal of Probability*, 20.

[21] Shirikyan, A. (2008). Exponential mixing for randomly forced partial differential equations: method of coupling. *Instability in Models Connected with Fluid Flows II*, 7, 155-188.

[22] Wang, F.Y. (2010). Coupling and their applications, [arXiv:1012.5687](https://arxiv.org/abs/1012.5687)

[23] Xu, L.(2014). Exponential mixing of 2D SDEs forced by degenerate Lévy noises. *J.Evol.Equ*, 14, 249-272.