Weak averaging principle for multiscale stochastic dynamical systems driven by $\alpha$-stable processes

Yanjie Zhang$^{a,1}$, Qiao Huang$^{b,2}$, Xiao Wang$^c$, Zibo Wang$^b4$, and Jinqiao Duan$^d5$

$^a$ School of Mathematics
South China University of Technology, Guangzhou 510641, China
$^b$ Center for Mathematical Sciences
Huazhong University of Science and Technology, Wuhan 430074, China
$^c$ School of Mathematics and Statistics,
Henan University, Kaifeng 475001, China
$^d$ Department of Applied Mathematics,
Illinois Institute of Technology, Chicago, IL 60616, USA

Abstract

We study the averaging principle for a family of multiscale stochastic dynamical systems. The fast and slow components of the systems are driven by two independent stable Lévy noises, whose stable indexes may be different. Moreover, the slow components contain homogenization terms, whose homogenizing index $r_0$ has a relation with the stable index $\alpha_1$ of the noise of fast components given by $0 < r_0 < 1 - 1/\alpha_1$. By first studying a nonlocal Poisson equation and then constructing suitable correctors, we obtain the tightness of slow processes. It turns out that the slow components weakly converge to a Lévy process as the scale parameter goes to zero.

Keywords: Nonlocal Poisson equation, averaging principle, tightness, stable Lévy noises.

1. Introduction.

The multiscale stochastic dynamical systems arise widely in various areas. Finding a coarse-grained model that can effectively characterize the asymptotic behavior of such systems has always been an active research field. The multiscale stochastic dynamical systems driven by Gaussian noises were originally studied by Khasminski. Later on, Stroock et al. and Pardoux et al. verified the relatively weak compactness of slow components by different approaches, one is martingale method and the other is to construct a corrector by utilizing the Poisson equation. Then, E et al. showed the weak and strong convergence results by using the heterogeneous multiscale method. Recently, Cerrai et al. extended the averaging principles to infinite dimensional case, while Hairer et al. investigated a multiscale system driven by fractional Brownian motion. Röckner et al. figured out the sharp rates, normal derivations and the functional central limit theorem for multiscale stochastic dynamical systems.

However, random fluctuations are often non-Gaussian in nonlinear systems. The multiscale dynamical systems driven by stable Lévy noises have also been studied by some authors. To name a few, in Bao et al. illustrated that their limit process is a solution of either a stochastic partial differential equation (SPDE) or an SPDE with switching. In the first author and the collaborators combined the averaging principles with the stochastic filtering problems and showed the
convergence of filters. Sun et al. [18] studied the strong and weak convergence order respectively. It is noteworthy that all the above results are based on the fact that the ratio of the time between fast and slow components in their systems is $O(1/\varepsilon)$. A natural and important question is: for the multiscale stochastic dynamical systems with homogenization terms in the slow components, how to obtain the corresponding averaged systems when the fast-slow ratio of the time is $O(1/\varepsilon^2)$?

More precisely, let us consider the following inhomogeneous multiscale system in $\mathbb{R}^{n+m}$:

$$
\begin{aligned}
&dX^\varepsilon_t = \frac{1}{\varepsilon^2} b(X^\varepsilon_t, Y^\varepsilon_t)dt + \frac{1}{\varepsilon^\alpha_1} dL^\alpha_1, \quad X_0^\varepsilon = x \in \mathbb{R}^n, \\
&dY^\varepsilon_t = F(X^\varepsilon_t, Y^\varepsilon_t)dt + \frac{1}{\varepsilon^\alpha_2} G(X^\varepsilon_t, Y^\varepsilon_t)dt + dL_i^\alpha_2, \quad Y_0^\varepsilon = y \in \mathbb{R}^m,
\end{aligned}
$$

where $F,G$ are two Borel measurable functions, the noises $L^\alpha_i, i = 1,2$ are independent symmetric $\alpha_i$-stable ($1 < \alpha_i < 2$) Lévy processes with triplets $(0, 0, \nu_i)$, and $\nu_i$'s are symmetric $\alpha_i$-stable Lévy measures on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively (see, e.g. [23, Section 14]), $\varepsilon > 0$ is a scaling parameter, the homogenizing index $r_0$ in the slow component satisfies $0 < r_0 < 1 - 1/\alpha_1$. The solution process $(X^\varepsilon, Y^\varepsilon)$ is an $\mathbb{R}^n \times \mathbb{R}^m$-valued process, in which $X^\varepsilon$ is called the fast component and $Y^\varepsilon$ is the slow one. The detailed assumptions on the coefficients can be found in Section 2. The infinitesimal generator $\mathcal{L}^\varepsilon$ corresponding to the solution $(X^\varepsilon, Y^\varepsilon)$ of system (1.1) has the following form

$$
\mathcal{L}^\varepsilon := \frac{1}{\varepsilon^2} \left[ -(-\Delta_x) \frac{\alpha_1}{\alpha_2} + b \cdot \nabla_x \right] + \frac{1}{\varepsilon^\alpha_2} G \cdot \nabla_y + \left[ -(-\Delta_y) \frac{\alpha_2}{\alpha_1} + F \cdot \nabla_y \right].
$$

We shall study the weak convergence of the averaging principle for the system (1.1). Our study is divided into two parts. In the first part, we examine the well-posedness of a nonlocal Poisson equation corresponding to an ergodic jump process, and give the probability representation of the solution. In the second part, we study the weak averaging principle for the system (1.1). By constructing suitable correctors via the nonlocal Poisson equation in the first part, we show that the slow component $Y^\varepsilon$ of system (1.1) weakly converges to the averaged system as the scale parameter tends to zero. The main result is as follows.

**Theorem 1.** Under Hypotheses $(A_b)$, $(A_F)$, $(A_{G1})$, $(A_{G2})$, the slow component $Y^\varepsilon$ converges weakly to a process $Y$ as the scale parameter $\varepsilon$ goes to zero. Moreover, the limit process $Y$ is the unique solution of the martingale problem associated to the following operator,

$$
\mathcal{L}_2 = -(-\Delta_y) \frac{\alpha_2}{\alpha_1} + \bar{F}(y) \cdot \nabla_y,
$$

where $\bar{F}$ is the homogenized drift given by

$$
\bar{F}(y) = \int_{\mathbb{R}^n} F(x, y) \mu(\varepsilon)(dx).
$$

Moreover, for any bounded function $\Phi$ on $\mathbb{D}([0, T]; \mathbb{R}^m)$ that is measurable with respect to the sigma-field $\sigma(\omega, \omega \in \mathbb{D}([0, T]; \mathbb{R}^m), t \leq t')$ and any function $\phi \in C^\infty_0(\mathbb{R}^m)$,

$$
\mathbb{E} \left[ \left( \phi(Y_t) - \phi(Y_{t_0}) - \int_{t_0}^t \mathcal{L}_2 \phi(Y_s)ds \right) \Phi_{t_0}(Y) \right] = 0.
$$

In comparison with the existing work for the case of Brownian noise in [1], there are several difficulties to be solved. First, since the Lévy noise in our case is not square integrable, the solution $(X^\varepsilon_t, Y^\varepsilon_t)$ has finite $p$-th moment only for $p \in (0, \alpha_1 \land \alpha_2)$ (cf. [23, Theorem 25.3]). Second, we need to show the existence for a Poisson equation and give the
probability representation for its solution, while this Poisson equation associates to the generator of a jump process and is nonlocal. Third, we need to determine the homogenizing index \( r_0 \) in the slow component of (1.1) so that \( \{ Y^\varepsilon \} \varepsilon > 0 \) is tight.

This paper is organized as follows. In Section 2, we list all assumptions for the coefficients. In Section 3, we prove the existence of nonlocal Poisson equation in the whole space and give the probability representation. In Section 4, we apply the technique of nonlocal Poisson equation to examine the weak averaging principle for the multiscale stochastic dynamical system driven by stable noises. In Section 5, we present a specific example to illustrate our result. Some concluding remarks are made in Section 6.

To end this section, we introduce some notations. The capital letter \( C \) denotes a positive constant, whose value may changes from one line to another. The notation \( C_p \) is used to emphasize that the constant only depends on the parameter \( p \). When there are more than one parameters, we use \( C(\cdots) \) to emphasize the dependence. We will use \( \langle \cdot, \cdot \rangle \) and \( |\cdot| \) to denote the scalar product and norm in Euclidean space. For any positive integer \( k \), we define

\[
B_0(\mathbb{R}^n) := \{ f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ is Borel bounded measurable } \},
\]
\[
C_0(\mathbb{R}^n) := \{ f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ is continuous with compact support } \},
\]
\[
C_0^0(\mathbb{R}^n) := \{ f \in C_0(\mathbb{R}^n) \mid f \text{ is centered with respect to the invariant measure } \mu \},
\]
\[
C_k(\mathbb{R}^n) := \{ f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ and all its partial derivative up to order } k \text{ are continuous} \},
\]
\[
C_k^k(\mathbb{R}^n) := \{ f \in C_k(\mathbb{R}^n) \mid 1 \leq i \leq k, \text{ the } i\text{-th order partial derivative are bounded} \},
\]
\[
C_b^k(\mathbb{R}^n) := \{ f(x,y) \mid 1 \leq |\beta|_1 \leq k \text{ and } 1 \leq |\beta|_2 \leq l, \nabla_x^\beta \nabla_y^{\beta_2} f \text{ is uniformly bounded} \}.
\]

Set \( \|f\|_0 = \sup_{x \in \mathbb{R}^n} f(x) \), then the space \( C^k_b(\mathbb{R}^n) \) is a Banach space endowed with the norm \( \|f\|_k = \|f\|_0 + \sum_{j=1}^k \|\nabla^\otimes f\|_j \).

2. Assumptions

In this section, we introduce all assumptions for the coefficients that we will need to obtain our main result. First of all, we need some regularity assumptions for the drifts \( b, F \) and \( G \).

(A) Assume that \( b \in C^{2,2}_b \), and there exists a positive constant \( \gamma \) such that for all \( x_1, x_2 \in \mathbb{R}^n, y \in \mathbb{R}^m \),

\[
\sup_{y \in \mathbb{R}^m} |b(0, y)| < \infty, \quad \langle b(x_1, y) - b(x_2, y), x_1 - x_2 \rangle \leq -\gamma |x_1 - x_2|^2. \tag{2.1}
\]

Notably, the assumption (A) will be applied to the coupled multiscale stochastic dynamical systems (1.1).

(\( A_F \)): The function \( F \) satisfies the Lipschitz condition and linear growth condition, i.e., there exists a positive constant \( K_1 > 0 \) such that

\[
|F(x_1, y_1) - F(x_2, y_2)|^2 \leq K_1 |x_1 - x_2|^2 + |y_1 - y_2|^2,
\tag{2.2}
\]

and

\[
|F(x, y)| \leq K_1. \tag{2.3}
\]

(\( A_G \) 1) The function \( G \) satisfies the following condition: there exists a positive constant \( K_2 \) such that

\[
|G(x_1, y_1) - G(x_2, y_2)|^2 \leq K_2 |x_1 - x_2|^2 + |y_1 - y_2|^2,
\tag{2.4}
\]

\( Y^\varepsilon \),
\[ |G(x, y)| \leq K_2, \quad \sup_{x, y} |\nabla_x G(x, y)| \leq K_2, \]
\[ \sup_y |\nabla_y G(x, y)| \leq K_2(1 + |x|), \quad \sup_{x, y} |\nabla^2_x G(x, y)| \leq K_2, \]
\[ \sup_{x, y} |\nabla^2_x \nabla_y G(x, y)| \leq K_2, \quad \sup_{x, y} |\nabla^3_x G(x, y)| \leq K_2. \]
\[(2.5)\{??\}\]

\[ (A_{G2}): \text{For each } y \in \mathbb{R}^m, \text{ the function } G(\cdot, y) \text{ is } \text{“centered” with respect to } \mu^y, \text{ i.e.,} \]
\[ \int_{\mathbb{R}^m} G(x, y) \mu^y(dx) = 0, \quad (2.6)\{??\} \]

where \( \mu^y(dx) \) is the unique invariant measure of an ergodic Markov process \( X^{x,y} \)(see (4.1) below).

3. The Poisson equation in \( \mathbb{R}^n \).

In this section, we devote to study the Poisson equation in \( \mathbb{R}^n \) and give the probability representation. More precisely, consider the following Poisson equation in \( \mathbb{R}^n \),
\[ \mathcal{L} u(x) = -f(x), \quad (3.1) \text{poisson} \]
where
\[ \mathcal{L} = -(-\triangle_x)^{\alpha_1} + \sum_i b_i(x) \partial_{x_i}, \quad \int_{\mathbb{R}^n} f(x) \mu(dx) = 0, \quad (3.2)\{??\} \]
and \( \mu \) is the invariant probability measure of the following Markov process \( X \), which satisfies the following stochastic differential equation (SDE)
\[ dX_t = b(X_t)dt + dL_t^{\alpha_1}, \quad X_0 = x \in \mathbb{R}^n. \]
\[ (3.3) \text{gnde} \]

Here we need to give the regularity assumption for the drifts \( b \) in \( (3.3) \).

\( (A'_b) \): The function \( \hat{g} \) defined by
\[ \hat{g}(r) := \inf \left\{ -\frac{(b(x_1) - b(x_2), x_1 - x_2)}{|x_1 - x_2|^2} : x_1, x_2 \in \mathbb{R}^n, |x_1 - x_2| = r \right\}, \quad (3.4) \text{cond} \]
satisfies
\[ \lim_{r \to \infty} \inf \hat{g}(r) > 0. \quad (3.5) \{??\} \]

Remark 1. The condition \( (3.4) \) implies that the drift coefficient \( b \) is dissipative outside certain open ball. To be more precise, there exist positive constants \( M \) and \( R \) such that for all \( x_1, x_2 \in \mathbb{R}^n \) with \( |x_1 - x_2| \geq R \), we have
\[ (b(x_1) - b(x_2), x_1 - x_2) \leq -M|x_1 - x_2|^2. \quad (3.6) \{??\} \]

In the following, one expects that under Hypothesis \( (A'_b) \) the solution of \( (3.1) \) has the probability representation
\[ u(x) = \int_{0}^{\infty} \mathbb{E}f(X_s)ds, \quad (3.7) \text{rep_u} \]
where \( X_t \) is given in \( (3.3) \).
3.1. Exponential ergodicity property.

In the following, we will give the exponential ergodicity and moments estimate for the SDE (3.3).

Lemma 1. Let \((A'_b)\) hold, and assume further that the semigroup \(\{P_t\}_{t \geq 0}\) preserves finite first moment, then the process \(X\) possesses an invariant distribution \(\mu\) on \(\mathbb{R}^n\). Moreover, there exist positive constants \(C_x\) and \(\rho\), such that for any \(Q \in \mathcal{B}_b(\mathbb{R}^n)\) and \(t \geq 0\), we have

\[
|P_tQ(x) - \int_{\mathbb{R}^n} Q(y)\mu(dy)| \leq C_x \sup_{x \in \mathbb{R}^n} |Q(x)|e^{-\rho t}.
\]

(3.8)

Proof. The existence of invariant distribution \(\mu\) follows from [17, Corollary 1.8]. Also by the same corollary, for all \(r > 0\), there exist a concave function \(\phi : [0, \infty) \to [0, \infty)\) with \(\phi(0) = 0\) and \(\phi(r) > 0\), and two positive constants \(C\) and \(\rho\), such that for all nonnegative \(t \geq 0\) and any probability measure \(\eta\), we have

\[
\|\mu - P^*_t\eta\|_{TV} \leq Ce^{-\rho t}W_{\phi}(\mu, \eta),
\]

where \(\{P^*_t\}_{t \geq 0}\) is the dual semigroup of \(\{P_t\}_{t \geq 0}\), \(W_{\phi}\) is the \(p\)-Wasserstein distance associated with the cost function \(\phi\).

Now we fix \(x \in \mathbb{R}^n\) and take \(\eta = \delta_x\), then we have

\[
W_{\phi}(\mu, \eta) = W_{\phi}(\mu, \delta_x) = \left(\int_{\mathbb{R}^n} \phi(|x - y|)\mu(dy)\right)^{1/p}.
\]

Hence,

\[
|P_tQ(x) - \int_{\mathbb{R}^n} Q(y)\mu(dy)| \leq \sup_{x \in \mathbb{R}^n} |Q(x)| \cdot \|P_t(x, \cdot) - \mu\|_{TV} = \sup_{x \in \mathbb{R}^n} |Q(x)| \cdot \|P^*_t\delta_x - \mu\|_{TV}
\]

\[
\leq Ce^{-\rho t}W_{\phi}(\mu, \delta_x) = C_x e^{-\rho t}.
\]

Remark 2. The semigroup \(\{P_t\}_{t \geq 0}\) preserves finite first moment, i.e., if a measure \(\eta\) has a finite first moment, then the measure \(P^*_t\eta\) also has a finite first moment.

Remark 3. The function \(W_{\phi}(\mu, \eta)\) is defined by the following formula

\[
W_{\phi}(\mu, \eta) = \left(\inf_{\pi \in \Pi(\mu, \eta)} \int_{\mathbb{R}^n \times \mathbb{R}^n} \phi(|x - y|)\pi(dx, dy)\right)^{1/p},
\]

where \(\Pi(\mu, \eta)\) denotes the collection of all measures on \(\mathbb{R}^n \times \mathbb{R}^n\) with marginals \(\mu\) and \(\eta\) respectively.

Lemma 2. Let \((A'_b)\) hold. Then for any \(1 \leq p < \alpha_1\) and \(T > 1\), we have

\[
\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t|^p\right)^{\frac{1}{p}} \leq C_p T^{\frac{p}{\alpha_1}} + |x|^p.
\]

(3.9)
By the same lines of arguments as in [18, Lemma 3.2], we have

\[ E \left[ \sup_{t \in [0,T]} U_T(X_t) \right] \leq \frac{1}{2} E \left( \sup_{t \in [0,T]} |X_t|^p \right) + C_p T^\frac{p}{2} + |x|^p. \]  

(3.11) (?)

This implies

\[ E \left[ \sup_{t \in [0,T]} |X_t|^p \right] \leq C_p T^\frac{p}{2} + |x|^p. \]  

(3.12) (?)

3.2. Existence of the solution of nonlocal Poisson equation.

Now, we are in the position to show the existence of the solution of nonlocal Poisson equation (3.1).

Lemma 3. Under Hypothesis (A^λ_b), for all \( f \in C^0_0(\mathbb{R}^n) \), the function \( u \) defined in (3.7) is a solution to the equation (3.1) in \( C^0_0(\mathbb{R}^n) \).

Proof. We divide the proof into the following three steps.  

Step 1. We firstly show that the right hand side of (3.7) does make sense. This is directly verified as follows, using the estimate (3.8) and the fact that \( f \) is centered with respect to \( \mu \), we have

\[ \left| \int_0^\infty E f(X_s) ds \right| = \left| \int_0^\infty P_s f(x) ds \right| \leq C(x) \sup_{x \in \mathbb{R}^n} |f(x)| \int_0^\infty e^{-\rho s} ds < \infty. \]

Obviously the semigroup \( \{P_t\}_{t \geq 0} \) is a Feller semigroup on \((C_0(\mathbb{R}^n), \| \cdot \|_0) \) (see [16, Theorem 6.7.4]). The classical theory of semigroups of operators yields (see [19, Lemma II.1.3])

\[ \mathcal{L} \int_0^t P_s f ds = P_t f - f. \]

Step 2. We fix \( x \in \mathbb{R}^n \). Since \( f \in C^0_0(\mathbb{R}^n) \), the estimate (3.8) implies that \( P_t f(x) \) converges uniformly in \( t \) to 0, as \( t \to \infty \). Hence, a straightforward interchange of limits yields

\[ \mathcal{L} u(x) = \mathcal{L} \left( \lim_{t \to \infty} \int_0^t P_s f(x) ds \right) = \lim_{t \to \infty} \mathcal{L} \int_0^t P_s f(x) ds \]

\[ = \lim_{t \to \infty} P_t f(x) - f(x) = -f(x). \]

This shows that the function \( u \) defined in (3.7) is a solution of (3.1).

Step 3. We prove that \( u \) is also centered with respect to \( \mu \). By Fubini’s theorem, we have

\[ \int_{\mathbb{R}^n} u(x) \mu(dx) = \int_{\mathbb{R}^n} \int_0^\infty P_s f(x) ds \mu(dx) = \int_0^\infty \int_{\mathbb{R}^n} P_s f(x) \mu(dx) ds \]

\[ = \int_0^\infty \int_{\mathbb{R}^n} f(x) \mu(dx) ds = 0. \]

This implies that the function \( u \) is centered. 

Remark 4. One can also use the Fredholm alternative to obtain the existence of the equation (3.1) directly, as in [15, Proposition 4.12].
4. Weak convergence in the averaging principle.

Now we are going to apply the technique of nonlocal Poisson equation to exam the averaging principle for the multiscale stochastic dynamical systems driven by stable Lévy noises in (1.1), i.e.,

\[
\begin{cases}
   dX_t^\varepsilon = \frac{1}{\varepsilon^2} b(X_t^\varepsilon, Y_t^\varepsilon)dt + \frac{1}{\varepsilon^\alpha_1} dL_{t}^{\alpha_1}, \quad X_0^\varepsilon = x \in \mathbb{R}^n, \\
   dY_t^\varepsilon = F(X_t^\varepsilon, Y_t^\varepsilon)dt + \frac{1}{\varepsilon^\alpha_2} G(X_t^\varepsilon, Y_t^\varepsilon)dt + dL_{t}^{\alpha_2}, \quad Y_0^\varepsilon = y \in \mathbb{R}^m.
\end{cases}
\]

Introduce the following frozen equation associated to the fast component,

\[
dX_{t}^{x,y} = b(X_{t}^{x,y}, y)dt + dL_{t}^{\alpha_1}, \quad X_0^{x,y} = x \in \mathbb{R}^n.
\]

One can show that, for any frozen \( y \in \mathbb{R}^m \), the equation (4.1) has a unique strong solution \( X_{t}^{x,y} \). Moreover, Lemma 1 ensure that the solution process \( X_{t}^{x,y} \) possesses an invariant distribution \( \mu^y \) on \( \mathbb{R}^n \).

Motivated by Lemma 3, we consider the following Poisson equation

\[
L_1 \tilde{G}_j(x, y) = -G_j(x, y),
\]

where

\[
L_1 \tilde{G}(x, y) = -(-\Delta)^{\alpha_1} \tilde{G}(x, y) + \left\langle b(x, y), \nabla_x \tilde{G}(x, y) \right\rangle.
\]

Then Lemma 3 yields that the nonlocal Poisson equations (4.2) have a unique centered solution

\[
\tilde{G}_j(x, y) = \int_0^{\infty} E G_j(X_{t}^{x,y}, y)dt.
\]

4.1. Some a-priori estimates of \( X_{t}^{x,y} \).

In the following, we will give some prior estimates for the jump process \( X_{t}^{x,y} \).

Lemma 4. Under Hypotheses (A\( _b \)), (A\( _F \)), (A\( _G1 \)),(A\( _G2 \)), for all \( t \geq 0 \), and \( x_i \in \mathbb{R}^n, y_i \in \mathbb{R}^m, i = 1, 2 \), we have

\[
|X_{t}^{x_1,y_1} - X_{t}^{x_2,y_2}|^2 \leq e^{-\frac{\gamma}{2}t}|x_1 - x_2|^2 + C(||b||_1, \gamma)|y_1 - y_2|^2,
\]

where \( C(||b||_1, \gamma) \) is a constant independent of \( t \).

Proof. See Appendix A.1.

Lemma 5. Under Hypotheses (A\( _b \)), (A\( _F \)), (A\( _G1 \)),(A\( _G2 \)), for all \( t \geq 0 \), and \( x_i \in \mathbb{R}^n, y_i \in \mathbb{R}^m, i = 1, 2 \), we have

\[
|\nabla_y X_{t}^{x_1,y_1} - \nabla_y X_{t}^{x_2,y_2}|^2 \leq C(||b||_2, \gamma)te^{-\frac{\gamma}{2}t}|x_1 - x_2|^2 + C(||b||_2, \gamma)|y_1 - y_2|^2,
\]

where \( C(||b||_2, \gamma) \) is a constant independent of \( t \).

Proof. See Appendix A.2.
Lemma 6. Under Hypotheses \((A_b), (A_F), (A_{G1}), (A_{G2})\), for all \(t \geq 0\), \(x_i \in \mathbb{R}^n\) and \(y_i \in \mathbb{R}^m\), \(i = 1, 2\), we have

\[
|\nabla_x X_t^{x_1:y_1} - \nabla_x X_t^{x_2:y_2}|^2 \leq C(\|b\|_2, \gamma)te^{-\gamma t} (|y_1 - y_2|^2 + |x_1 - x_2|^2),
\]

where \(C(\|b\|_2, \gamma)\) is a constant independent of \(t\).

Proof. See Appendix A.3.

Now, we give the exponential ergodicity for the equation \((4.1)\).

Proposition 1. Under Hypotheses \((A_b), (A_F), (A_{G1}), (A_{G2})\), for each function \(\tilde{\varphi} \in C_b^1\), there exists a positive constant \(C\) such that for all \(t \geq 0\) and \(x \in \mathbb{R}^n\), we have

\[
\sup_{y \in \mathbb{R}^m} |P^y_t \tilde{\varphi}(x) - \mu^y(\tilde{\varphi})| \leq C\|\tilde{\varphi}\|_1 e^{-\gamma t} (1 + |x|^\frac{1}{2}),
\]

where

\[
P^y_t \tilde{\varphi}(x) = E^y_t(\tilde{\varphi}^{x,y}).
\]

Proof. By the definition of invariant measure, Lemma 5, Lemma 3 and Hölder inequality, we have

\[
|E^y_t(\tilde{\varphi}^{x,y}) - \mu^y(\tilde{\varphi})| = |E^y_t(\tilde{\varphi}^{x,y}) - \int_{\mathbb{R}^n} \tilde{\varphi}(z) \mu^y(dz)|
\leq \int_{\mathbb{R}^n} |E^y_t(\tilde{\varphi}^{x,y}) - E^y_t(\tilde{\varphi}^{x,y})| \mu^y(dz)
\leq 2(\|\tilde{\varphi}\|_0 + \|\nabla \tilde{\varphi}\|_0) \int_{\mathbb{R}^n} E|X_t^{x,y} - X_t^{x,y}|^\frac{1}{2} \mu^y(dz)
\leq C\|\tilde{\varphi}\|_1 e^{-\gamma t} \int_{\mathbb{R}^n} |x - z|^\frac{1}{2} \mu^y(dz)
\leq C\|\tilde{\varphi}\|_1 e^{-\gamma t} (1 + |x|^\frac{1}{2}).
\]

Lemma 7. Under Hypotheses \((A_b), (A_F), (A_{G1}), (A_{G2})\), there exists a positive constant \(C\) such that

(i)

\[
\sup_{y \in \mathbb{R}^m} |\tilde{G}(x, y)| \leq C(1 + |x|^\frac{1}{2}).
\]

(ii)

\[
\sup_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} |\nabla_x \tilde{G}(x, y)| \leq C.
\]

(iii)

\[
\sup_{y \in \mathbb{R}^m} |\nabla_y \tilde{G}(x, y)| \leq C \left(1 + |x|^\frac{1}{2}\right).
\]

(iv)

\[
\sup_{y \in \mathbb{R}^m} |\nabla^2_y \tilde{G}(x, y)| \leq C(1 + |x|).
\]

Proof. See Appendix A.4.

One can also deduce the following moment property for fast component \(X_t\).
Lemma 8. Let \((A_b)\) hold. Then for each \(1 \leq p < \alpha_1\), \(r_1 > \frac{2p}{\alpha_1}\) and \(T > 1\), we have

\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^\varepsilon|^p \right) \leq C_p T^\frac{p}{\alpha_1} \varepsilon^{-2p/\alpha_1}. \tag{4.15}\]

This implies

\[
\varepsilon^{r_1} \mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^\varepsilon|^p \right) \rightarrow 0, \text{ as } \varepsilon \to 0. \tag{4.16}\]

**Proof.** Note that for each \(\varepsilon > 0\), we have

\[
X_t^{\varepsilon, x} = x + \frac{1}{\varepsilon} \int_0^{t \varepsilon} b(X_s^\varepsilon, Y_s^\varepsilon) ds + \frac{1}{\varepsilon^{\alpha_1}} L_t^\alpha_1 \tag{4.17},
\]

where \(\{\tilde{L}_t^\alpha_1 = \frac{1}{\varepsilon^{\alpha_1}} L_t^{\alpha_1}, t \geq 0\}\) is also \(\alpha\)-stable process with the same law as \(L_t^{\alpha_1}\).

Define the auxiliary process \(\{\tilde{X}_t^\varepsilon\}\) by

\[
\tilde{X}_t^\varepsilon = y + \int_0^t b(\tilde{X}_s^\varepsilon, Y_s^\varepsilon) ds + \tilde{L}_t^\alpha_1. \tag{4.18}.
\]

Using the condition \(\sup_{y \in \mathbb{R}^m} |b(0, y)| < \infty\) and the same argument as Lemma 2, we yield that

\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\tilde{X}_t^\varepsilon|^p \right] \leq C_p T^{p/\alpha_1} + |x|^p. \tag{4.19}
\]

This implies

\[
\mathbb{E} \left( \sup_{t \in [0, T]} |X_t^\varepsilon|^p \right) = \mathbb{E} \left( \sup_{0 \leq t \leq T/\varepsilon^2} |\tilde{X}_t^\varepsilon|^p \right) \leq C_p T \varepsilon^{-2p/\alpha_1} + |x|^p. \tag{4.20}
\]

Therefore for each \(1 \leq p < \alpha_1\), \(r_1 > \frac{2p}{\alpha_1}\) and \(T > 1\), we have

\[
\varepsilon^{r_1} \mathbb{E} \left( \sup_{0 \leq t \leq T} |X_t^\varepsilon|^p \right) \rightarrow 0, \text{ as } \varepsilon \to 0. \tag{4.21}.
\]

\[
\square
\]

4.2. The tightness of slow component.

In this subsection, we will show that the effective low dimensional system weakly converges to the slow component of original system as the scale parameter tends to zero, by viewing the solution of nonlocal Poisson equation as a corrector.

Define

\[
f^\varepsilon(x, y) = f_1(y) + \varepsilon^{2-r_0} u(x, y), \tag{4.22}\]

where \(\varepsilon u\) is a corrector to \(f_1\), \(u\) is the solution of Poisson equation

\[
\mathcal{L}_1 u(x, y) = - \langle \nabla_y f_1(y), G(x, y) \rangle. \tag{4.23}
\]

By Lemma 3, we have

\[
u(x, y) = \langle \nabla_y f_1(y), \tilde{G}(x, y) \rangle. \tag{4.24}\]
Lemma 9. In the following, we will show the relative compactness of

By Itô’s formula, we get

\[
f^{\varepsilon}(X_t^\varepsilon, Y_t^\varepsilon) - f^\varepsilon(x, y) = f_1(Y_t^\varepsilon) - f_1(y) + \varepsilon^{2-\alpha} u(X_t^\varepsilon, Y_t^\varepsilon) - \varepsilon^{2-\alpha} u(x, y) \\
= \int_0^t \left( - (\Delta_y) \frac{\partial}{\partial y} f_1(Y_s^\varepsilon) \right) ds + \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} [f_1(Y_{s-}^\varepsilon) - f_1(Y_s^\varepsilon)] \tilde{N}_2(ds, dy) \\
+ \varepsilon^{2-\alpha} \int_0^t \left( - (\Delta_y) \frac{\partial}{\partial y} f_1(Y_s^\varepsilon) \right) ds + \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} [u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) - u(X_s^\varepsilon, Y_s^\varepsilon)] \tilde{N}_2(ds, dy) \tag{4.25}
\]

where

\[- (\Delta_y) \frac{\partial}{\partial y} f_1(y) = \int_{\mathbb{R}^m \setminus \{0\}} [f_1(y + z) - f_1(y) - I_{\{z \in \{0\}\}}(z, \partial_y f_1(y)) \nu_2(dz). \tag{4.26}\]

By the definition of \( u \), the summation of all terms of order \( \varepsilon^{-1} \) in (4.25) vanishes. Then we obtain

\[
f_1(Y_t^\varepsilon) = f_1(y) + \int_0^t \left( \nabla_y f_1(Y_s^\varepsilon), F(X_s^\varepsilon, Y_s^\varepsilon) + \varepsilon^{2-2\alpha} \sum_i G_i(X_s^\varepsilon, Y_s^\varepsilon) \partial_y \tilde{G}(X_s^\varepsilon, Y_s^\varepsilon) \right) ds \\
+ \varepsilon^{2-\alpha} \int_0^t \left( - (\Delta_y) \frac{\partial}{\partial y} f_1(Y_s^\varepsilon) \right) ds + \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} [u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) - u(X_s^\varepsilon, Y_s^\varepsilon)] \tilde{N}_2(ds, dy) + \varepsilon^{2-\alpha} R^\varepsilon_0(0, t), \tag{4.27}
\]

where

\[
R^\varepsilon_0(0, t) = u(x, y) - u(X_t^\varepsilon, Y_t^\varepsilon) + \int_0^t \left( \nabla_y u(X_s^\varepsilon, Y_s^\varepsilon), F(X_s^\varepsilon, Y_s^\varepsilon) \right) ds \\
+ \int_0^t \left( - (\Delta_y) \frac{\partial}{\partial y} u(X_s^\varepsilon, Y_s^\varepsilon) \right) ds + \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} [u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) - u(X_s^\varepsilon, Y_s^\varepsilon)] \tilde{N}_2(ds, dy) \tag{4.28}
\]

In the following, we will show the relative compactness of \( \{Y^\varepsilon\}_{\varepsilon > 0} \) in the metric space \( D([0, T], \mathbb{R}^m) \).

**Lemma 9.** The family \( \{Y^\varepsilon\}_{\varepsilon > 0} \) of second solution processes of (1.1) satisfies the following conditions:

(i) For all \( T > 0 \) and \( \delta > 0 \), there exists \( N > 0 \) such that

\[
\mathbb{P} \left( \sup_{0 \leq t \leq T} |Y_t^\varepsilon| > N \right) \leq \delta, \quad \forall 0 < \varepsilon < 1. \tag{4.29}
\]

(ii) For all \( T > 0 \), it holds that

\[
\lim_{\delta \to 0} \limsup_{\varepsilon \to 0} \sup_{0 \leq \tau < T-\delta} \mathbb{P}(|Y_{\tau+\delta}^\varepsilon - Y_\tau^\varepsilon| > \lambda) = 0, \quad \forall \lambda > 0, \tag{4.30}
\]

where the second supremum is taken over all stopping time \( \tau \) satisfying \( 0 \leq \tau < T - \delta \).
Therefore, by Theorem VI.4.5, yields that the family \( \{Y^x\}_{x > 0} \) is relative compact.

**Proof. Step 1.** Choose in (4.22) that \( f_1(y) = \log(1 + |y|^2) \). Then we get

\[
(1 + |y|) \cdot |\nabla_y f_1(y)| + (1 + |y|)^2 \cdot |\nabla^2_y f_1(y)| + (1 + |y|)^3 \cdot |\nabla^3_y f_1(y)| 
\]

For the term by Lemma 7 and (4.24), we have for every \( x \in \mathbb{R}^n, y \in \mathbb{R}^m \),

\[
|u(x, y)| \leq C(1 + |x|^\frac{1}{2}),
\]

and

\[
|\nabla_x u(x, y)| \leq C, \quad |\nabla_y u(x, y)| \leq C(1 + |x|^\frac{1}{2}), \quad |\nabla^2_y u(x, y)| \leq C(1 + |x|^\frac{1}{2}) + C(1 + |x|).
\]

By corollary we have

\[
E_{x,y} \sup_{0 \leq t \leq T} |u(X^x_t, Y^x_t)| \leq CE_{x,y} \sup_{0 \leq t \leq T} (1 + |X^x_t|^\frac{1}{2}),
\]

and

\[
E_{x,y} \sup_{0 \leq t \leq T} |\nabla_y u(X^x_t, Y^x_t)|^2 \leq CE_{x,y} \sup_{0 \leq t \leq T} (1 + |X^x_t|),
\]

\[
E_{x,y} \sup_{0 \leq t \leq T} |\nabla^2_y u(X^x_t, Y^x_t)| \leq CE_{x,y} \sup_{0 \leq t \leq T} (1 + |X^x_t|).
\]

We recall

\[
f_1(Y^x_t) = f_1(Y^x_0) + \int_0^t \left\langle \nabla_y f_1(Y^x_s), F(X^x_s, Y^x_s) + \varepsilon^{2-2\alpha} \sum_i G_i(X^x_s, Y^x_s) \partial_y G(X^x_s, Y^x_s) \right\rangle ds
\]

\[
+ \varepsilon^{2-2\alpha} \int_0^t \left\langle \nabla_y f_1(Y^x_s), \sum_i \partial_y \partial_{x_i} f_1(Y^x_s) \tilde{G}_i(X^x_s, Y^x_s) \right\rangle ds + \int_0^t \left\langle \varepsilon^{2-2\alpha} f_1(Y^x_s), \right\rangle ds
\]

\[
+ \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} \left[ f_1(Y^x_s - y) - f_1(Y^x_s - y) \varepsilon^{2-\alpha} \left\langle u(X^x_0, Y^x_0) - u(X^x_t, Y^x_t) \right\rangle ds, dy \right.
\]

\[
+ \varepsilon^{2-\alpha} \int_0^t \left\langle u(X^x_s, Y^x_s) \right\rangle ds \left[ \varepsilon^{2-\alpha} \int_0^t \left\langle \nabla_y u(X^x_s, Y^x_s) \right\rangle ds \right] + \varepsilon^{2-\alpha} \left[ \varepsilon^{2-\alpha} \int_0^t \left\langle \nabla_y u(X^x_s, Y^x_s) \right\rangle ds \right]
\]

\[
+ \varepsilon^{2-\alpha} \left[ \varepsilon^{2-\alpha} \int_0^t \left\langle u(X^x_s, Y^x_s) \right\rangle ds \right] + \varepsilon^{2-\alpha} \left[ \varepsilon^{2-\alpha} \int_0^t \left\langle \nabla_y u(X^x_s, Y^x_s) \right\rangle ds \right]
\]

\[
=: f_1(Y^x_0) + I_1 + I_2 + I_3 + I_4 + J_1 + J_2 + J_3 + J_4 + J_5.
\]

For the term \( I_1 \), by the inequality (4.31), Hypothesis (A_F), Hypothesis (A_G_1), \( 0 < r_0 < 1 - 1/\alpha \) and Lemma (iii), we have

\[
E \left[ \sup_{0 \leq t \leq T} |I_1| \right] \leq C_T + \varepsilon^{2-2\alpha} E \left[ \int_0^T \left( 1 + |X^x_t|^\frac{1}{2} \right) \right]
\]

\[
\leq C \int_0^T E \left[ \left( 1 + |X^x_t| \right) \right] ds.
\]

For the term \( I_2 \), by the same argument as \( I_1 \), we also have

\[
E \left[ \sup_{0 \leq t \leq T} |I_2| \right] \leq C \int_0^T E \left[ \left( 1 + |X^x_t| \right) \right] ds.
\]
For the term $I_3$, by the choice of $f_1$ and (4.31), we have
\[
E \left[ \sup_{0 \leq t \leq T} \int_0^t \left( -(-\triangle_y)^{\frac{s}{2}} f_1(Y_s^\varepsilon) \right) ds \right] 
= E \left[ \sup_{0 \leq t \leq T} \int_0^t \left\{ \int_{[|z| < 1]} \left[ f_1(Y_s^\varepsilon + z) - f_1(Y_s^\varepsilon) - I_{\{|z| \leq 1\}} \langle z, \nabla_y f_1(Y_s^\varepsilon) \rangle \right] \nu_2(dz) \right\} ds \right] 
\leq \int_0^T \|\nabla_y f_1\|_0 \int_{|z| > 1} z \nu_2(dz) + C_T \|\nabla_y f_1\|_0 \int_{|z| \leq 1} z^2 \nu_2(dz) \right\} ds 
\leq C_T. 
\] (4.39) ?

For the term $I_4$, by the Burkholder-Davis-Gundy inequality, Jensen’s inequality, (4.31) and Hölder inequality, we have
\[
E \left( \sup_{0 \leq t \leq T} |I_4| \right) \leq E \left[ \sup_{t \in [0, T]} \int_0^t \int_{|y| < 1} \left[ f_1(Y_s^\varepsilon + y) - f_1(Y_s^\varepsilon) \right] \bar{N}_2(ds, dy) \right] 
+ E \left[ \sup_{t \in [0, T]} \int_0^t \int_{|y| \geq 1} \left[ f_1(Y_s^\varepsilon + y) - f_1(Y_s^\varepsilon) \right] \bar{N}_2(ds, dy) \right] 
\leq \left\{ \int_0^T \int_{|y| < 1} E \left[ f_1(Y_s^\varepsilon + y) - f_1(Y_s^\varepsilon) \right]^2 \nu_2(dy) ds \right\} ^{\frac{1}{2}} 
\quad + \left\{ \int_0^T \int_{|y| \geq 1} E \left[ f_1(Y_s^\varepsilon + y) - f_1(Y_s^\varepsilon) \right]^2 \nu_2(dy) ds \right\} ^{\frac{1}{2}} 
\leq C \left[ \int_0^T \int_{|y| < 1} |y|^2 \nu_2(dy) ds \right] ^{\frac{1}{2}} + C \left[ \int_0^T \int_{|y| \geq 1} |y|^2 \nu_2(dy) ds \right] 
\leq C_T. 
\] (4.40) ?

For the term $J_1$, by the inequality (4.32), we have
\[
E \left[ \sup_{0 \leq t \leq T} |J_1| \right] \leq C \varepsilon^{2-\gamma_0} \left\{ 1 + E \left[ \sup_{0 \leq t \leq T} |X_t^\varepsilon| \right] \right\} ^{1/2}. 
\] (4.41) ?

For the term $J_2$, by the inequality (4.34), Hölder inequality and Hypotheses (A_F), we have
\[
E \left[ \sup_{0 \leq t \leq T} |J_2| \right] \leq \varepsilon^{2-\gamma_0} E \left[ \left( 1 + \sup_{0 \leq t \leq T} |X_t^\varepsilon| \right) \right]. 
\] (4.42) ?
For the term $J_3$, by (4.24), Lemma 7, Hypotheses (A_F) and (A_{G1}), (4.31) and (4.35), we obtain

$$
E \left[ \sup_{0 \leq t \leq T} |J_3| \right] = E \left[ \sup_{0 \leq t \leq T} \left\{ \int_0^t \left\{ \int_{|z| \leq 1} \left[ u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon + z) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) - I_{([z] \leq 1)}(z, \nabla_y u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon)) \right] v_2(dz) \right\} ds \right\} \right]
$$

Lower bounded by (4.32), inequality (4.33) and Lemma 7, we have

$$
\leq \varepsilon^{2-r_0} \int_0^T E \left\{ \|\nabla_y u(X_{s-}^\varepsilon, \cdot)\|_0 \int_{|z| > 1} z^2 \nu_2(dz) + \|\nabla_y^2 u(X_{s-}^\varepsilon, \cdot)\|_0 \int_{|z| \leq 1} z^2 \nu_2(dz) \right\} ds
$$

$$
\leq \varepsilon^{2-r_0} C_{\|u\|_2, \alpha_2} \int_0^T \left( 1 + E \left[ \sup_{0 \leq t \leq T} |X_{s-}^\varepsilon| \right] \right) ds.
$$

For the term $J_4$, by the Burkholder-Davis-Gundy inequality, Jensen’s inequality, inequality (4.32), inequality (4.33) and Lemma 7, we have

$$
E \left( \sup_{0 \leq t \leq T} |J_4| \right) \leq \varepsilon^{2-r_0} E \left[ \sup_{t \in [0, T] \cap \mathbb{R}} \left( \int_0^t \int_{|y| < 1} \left[ u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon + y) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right] \bar{N}_2(ds, dy) \right) \right]
$$

$$
+ \varepsilon^{2-r_0} E \left[ \sup_{t \in [0, T] \cap \mathbb{R}} \left( \int_0^t \int_{|y| \geq 1} \left[ u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon + y) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right] \bar{N}_2(ds, dy) \right) \right]
$$

$$
\leq \varepsilon^{2-r_0} \left( \int_0^T \int_{|y| < 1} E \left[ u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon + y) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right] ^2 \nu_2(dy) ds \right)^{\frac{1}{2}}
$$

$$
+ \varepsilon^{2-r_0} \left( \int_0^T \int_{|y| \geq 1} E \left[ u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon + y) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right] ^2 \nu_2(dy) ds \right)^{\frac{1}{2}}
$$

$$
\leq C \varepsilon^{2-r_0} \left[ E \int_0^T \int_{|y| < 1} |y|^2 \nu_2(dy)(1 + |X_{s-}^\varepsilon|) ds \right]^{\frac{1}{p}}
$$

$$
+ C \varepsilon^{2-r_0} \left[ E \int_0^T \int_{|y| \geq 1} |y| \nu_2(dy)(1 + |X_{s-}^\varepsilon|) ds \right]^{\frac{1}{p}}
$$

$$
\leq \varepsilon^{2-r_0} \left( 1 + \sup_{0 \leq t \leq T} |X_{s-}^\varepsilon| \right).
$$

For the term $J_5$, by the Burkholder-Davis-Gundy inequality, Jensen’s inequality, inequality (4.33) and Lemma 7, we
have
\[
E \left( \sup_{0 \leq t \leq T} |J_5| \right) \leq \varepsilon^{2-\rho_0} E \left[ \sup_{t \in [0,T]} \left| \int_0^t \int_{|x| < 1} \left[ u(X_{s-} + \varepsilon^{-\frac{2}{\alpha_1}} x, Y_{s-}^\varepsilon) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right] \tilde{N}_i(ds, dx) \right| \right]
+ \varepsilon^{2-\rho_0} E \left[ \sup_{t \in [0,T]} \left( \int_0^T \int_{|x| \geq 1} \left[ u(X_{s-} + \varepsilon^{-\frac{2}{\alpha_1}} x, Y_{s-}^\varepsilon) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right]^2 \tilde{N}_1(ds, dx) \right)^{\frac{1}{2}} \right]
\leq \varepsilon^{2-\rho_0} \left\{ \int_0^T \int_{|x| < 1} E \left[ u(X_{s-} + \varepsilon^{-\frac{2}{\alpha_1}} x, Y_{s-}^\varepsilon) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right]^2 \nu_1(dx)ds \right\}^{\frac{1}{2}}
+ \varepsilon^{2-\rho_0} \left\{ \int_0^T \int_{|x| \geq 1} E \left[ u(X_{s-} + \varepsilon^{-\frac{2}{\alpha_1}} x, Y_{s-}^\varepsilon) - u(X_{s-}^\varepsilon, Y_{s-}^\varepsilon) \right]^2 \nu_1(dx)ds \right\}^{\frac{1}{2}}
\leq C \varepsilon^{2-\rho_0-2/\alpha_1} \left[ \int_0^T \int_{|x| < 1} |x|^2 \nu_1(dx)ds \right]^{\frac{1}{2}} + C \varepsilon^{2-\rho_0-2/\alpha_1} \left[ \int_0^T \int_{|x| \geq 1} |x|^2 \nu_1(dx)ds \right]^{\frac{1}{2}}
\leq C \varepsilon^{2-\rho_0-2/\alpha_1}.
\]

Taking the above estimates (4.37)-(4.35) in (4.36) and (4.28), we obtain
\[
\sup_{0 < \varepsilon \leq 1} E_{\varepsilon,y} \left( \sup_{0 \leq t \leq T} \log \left( 1 + |Y_t^\varepsilon|^2 \right) \right) < \infty.
\]

In view of (4.46), we use Chebyshev’s inequality to get
\[
P \left( \sup_{0 \leq t \leq T} |Y_t^\varepsilon| > N \right) = P \left( \log \left( 1 + \sup_{0 \leq t \leq T} |Y_t^\varepsilon|^2 \right) > \log(1 + N^2) \right)
\leq \sup_{0 < \varepsilon \leq 1} E_{\varepsilon,y} \left( \sup_{0 \leq t \leq T} \log \left( 1 + |Y_t^\varepsilon|^2 \right) \right)
\log(1 + N^2)
\rightarrow 0, \quad N \rightarrow \infty.
\]

(ii) Let \( \tau \leq T - \delta_0 \) be a bounded stopping time. For any \( \delta \in (0, \delta_0) \), by the strong Markov property, we have
\[
P \left( |Y_{\tau+\delta}^\varepsilon - Y_{\tau}^\varepsilon| > \lambda \right) = E \left( P \left( |Y_{s+\delta}^\varepsilon - y| > \lambda \right) \right)_{(s,y) = (\tau, Y_{\tau}^\varepsilon)}.
\]

Define
\[
\overline{Y}_t^\varepsilon = Y_t^\varepsilon - y,
\]
then we have
\[
\left\{ \begin{array}{l}
\frac{d\overline{Y}_t^\varepsilon = F(X_t^\varepsilon, \overline{Y}_t^\varepsilon + y)dt + \frac{1}{\varepsilon} G(X_t^\varepsilon, \overline{Y}_t^\varepsilon + y)dt + dL_t^{\varepsilon^2}}
\overline{Y}_0^\varepsilon = 0 \in \mathbb{R}^m.
\end{array} \right.
\]
Let us write (4.30) in the particular case of the vector function $f_1(y) = y$. We obtain

$$
\bar{Y}_t^\varepsilon = \bar{Y}_0^\varepsilon + \int_0^t \left\langle I, F(X_s^\varepsilon, \bar{Y}_s^\varepsilon) + \varepsilon^{2-r_0} \sum_i G_i(X_s^\varepsilon, Y_s^\varepsilon) \right\rangle ds + \int_0^t \left( - (\Delta y) \overrightarrow{\bar{Y}_s^\varepsilon} \right) ds + \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} y \bar{N}_2(ds, dy) + \varepsilon^{2-r_0} \left[ u(X_0^\varepsilon, \bar{Y}_0^\varepsilon) - u(X_t^\varepsilon, \bar{Y}_t^\varepsilon) \right] dI_s ds,
$$

(4.51)

$$
\text{Combining with (4.48) and (4.52), we obtain such that}
$$
$$
+ \varepsilon^{2-r_0} \left[ \int_0^t \left( \nabla_y u(X_s^\varepsilon, \bar{Y}_s^\varepsilon), F(X_s^\varepsilon, \bar{Y}_s^\varepsilon) \right) ds \right] + \varepsilon^{2-r_0} \left[ \int_0^t \left( - (\Delta y) \overrightarrow{\bar{Y}_s^\varepsilon} \right) ds \right]
$$

(4.52)

Thus we have

$$
\bar{P}_{s,y} \left( |\bar{Y}_{s+\delta}^\varepsilon| > \lambda \right) \leq \frac{\bar{E}_{s,y} \left( |Y_{s+\delta}^\varepsilon| \right)}{\lambda} \leq C\delta^{1/\lambda}.
$$

(4.53)

Combining with (4.48) and (4.52), we obtain

$$
\bar{P} \left( |Y_{s+\delta}^\varepsilon - Y_s^\varepsilon| > \lambda \right) \leq \bar{P} (|Y_{s+\delta}^\varepsilon| > R) + \frac{C\delta^{1/\lambda}}{\lambda}
$$

(4.54)

Letting $\delta \to 0$ first and then $R \to \infty$, one sees that (ii) is satisfied.

\square

### 4.3. Proof of Theorem 2.

Consider the Skorokhod space $\mathbb{D}([0,T], \mathbb{R}^m)$ consisting of all $\mathbb{R}^m$-valued càdlàg functions on $[0,T]$, equipped with the Skorokhod topology. It is well-known that $\mathbb{D}([0,T], \mathbb{R}^m)$ is a Polish space (e.g., [24, Section VI.1] or [25, Section 14]).

Next, we will present the uniform approximation of càdlàg functions by step functions, which comes from [21, Lemma 9, Appendix A].

(stea)

#### Lemma 10

Let $h$ be a càdlàg function on $[0,T]$. If $(t^n_k)$ is a sequence of subdivisions $0 = t^n_0 < t^n_1 < \cdots < t^n_{k_n} = t$ of $[0,T]$ such that

$$
\sup_{0 \leq i \leq k-1} |t^n_{i+1} - t^n_i| \to 0, \quad \sup_{u \in [0,T] \setminus \{t^n_0, \cdots, t^n_{k_n}\}} |\Delta h(u)| \to 0, \quad \text{as} \quad n \to \infty.
$$

Then we have

$$
\sup_{u \in [0,T]} \left| h(u) - \sum_{i=0}^{k_n-1} h(t^n_i) I_{t^n_i, t^n_{i+1}}(u) + h(t^n_{k_n}) I_{t^n_{k_n}}(u) \right| \to 0.
$$

(4.55)

Due to the tightness of the family $\{Y^\varepsilon\}_{\varepsilon > 0}$, there exists a subsequence $\varepsilon_n \to 0$ and a stochastic process $Y$, such that $Y^\varepsilon_n$ converges weakly to $Y$, as $n \to \infty$. Based on Lemma 10, we have the following result.
Lemma 11. For sufficient small $\delta > 0$, there exist a $N \in \mathbb{N}$ and $\mathbb{R}^m$-valued step functions $y^1, y^2, \cdots, y^N$ s.t.

$$
\begin{align*}
\mathbb{P} \left( \bigcap_{k=1}^N \{ d_R(Y^{\varepsilon_n}, y^k) > \delta \} \right) &< \delta, \quad \forall n \in \mathbb{N}, \\
\mathbb{P} \left( \bigcap_{k=1}^N \{ d_R(Y, y^k) > \delta \} \right) &< \frac{\delta}{2}.
\end{align*}
$$

(4.56) (?)

Proof. Step 1. By the tightness of the set $\{Y; Y^{\varepsilon_n}, n \in \mathbb{N}\}$, there exists a compact set $K \subseteq D([0, T]; \mathbb{R}^m)$ such that for each $0 < \delta << 1$, we have

$$
\begin{align*}
\mathbb{P}(Y^{\varepsilon_n} \in K) &> 1 - \delta, \\
\mathbb{P}(Y \in K) &> 1 - \delta.
\end{align*}
$$

(4.57) (?)

Since $K$ is compact, it is totally bounded. Hence, $K$ admits a finite $\delta/2$-net, i.e., there exists a finite subset $\{\tilde{y}^1, \tilde{y}^2, \cdots, \tilde{y}^N\} \subseteq D([0, T]; \mathbb{R}^m)$ s.t.

$$
K \subseteq \bigcup_{k=1}^N \left\{ x \in D([0, T]; \mathbb{R}^m) : d_R(\tilde{y}^k, x) < \frac{\delta}{2} \right\}.
$$

(4.58) (?)

Set $A_k = \{ x \in D([0, T]; \mathbb{R}^m) : d_R(\tilde{y}^k, x) < \frac{\delta}{2} \}$, then we have

$$
\bigcap_{k=1}^N A_k^c \subseteq K^c.
$$

(4.59) (?)

This implies

$$
\begin{align*}
\mathbb{P} \left( \bigcap_{k=1}^N \left\{ d_R(Y^{\varepsilon_n}, \tilde{y}^k) > \frac{\delta}{2} \right\} \right) &< \frac{\delta}{2}, \quad \forall n \in \mathbb{N}, \\
\mathbb{P} \left( \bigcap_{k=1}^N \left\{ d_R(Y, \tilde{y}^k) > \frac{\delta}{2} \right\} \right) &< \frac{\delta}{2}.
\end{align*}
$$

(4.60) (?)

Step 2. By Lemma 10 for fixed $t \geq 0$, we can find the step function $y^k$, which is arbitrarily close to the càdlàg function in supremum norm, i.e.,

$$
\sup_{0 \leq t \leq T} |y^k - \tilde{y}^k| < \frac{\delta}{2}.
$$

(4.61) (?)

On the one hand, we have

$$
\begin{align*}
\{ d_R(Y^{\varepsilon_n}, y^k) > \delta \} &\subseteq \left\{ d_R(Y^{\varepsilon_n}, \tilde{y}^k) > \frac{\delta}{2} \right\} \cup \left\{ d_R(y^k, \tilde{y}^k) > \frac{\delta}{2} \right\} \\
&\subseteq \left\{ d_R(Y^{\varepsilon_n}, \tilde{y}^k) > \frac{\delta}{2} \right\} \cup \left\{ \sup |y^k - \tilde{y}^k| > \frac{\delta}{2} \right\},
\end{align*}
$$

(4.62) (?)

and

$$
\begin{align*}
\{ d_R(Y, y^k) > \delta \} &\subseteq \left\{ d_R(Y, \tilde{y}^k) > \frac{\delta}{2} \right\} \cup \left\{ d_R(y^k, \tilde{y}^k) > \frac{\delta}{2} \right\} \\
&\subseteq \left\{ d_R(Y, \tilde{y}^k) > \frac{\delta}{2} \right\} \cup \left\{ \sup |y^k - \tilde{y}^k| > \frac{\delta}{2} \right\}.
\end{align*}
$$

(4.63) (?)
Therefore we have
\[
\mathbb{P}
\left(\bigcap_{k=1}^{N} \left\{ d_R(Y^{\varepsilon_n}, y^k) > \delta \right\} \right) < \delta, \quad \forall n \in \mathbb{N},
\]

\[
\mathbb{P}
\left(\bigcap_{k=1}^{N} \left\{ d_R(Y, y^k) > \delta \right\} \right) < \delta.
\]

\[\text{(4.64)}\]

In what follows, we will fix a $\delta > 0$ and let $y^1, y^2, \cdots, y^N$ be the corresponding $N$ step functions in Lemma 11. For each $y \in \mathbb{D}([0,T]; \mathbb{R}^m)$ and $k = 1, 2, \cdots, N$, we define
\[
\beta_k(y) := d_R(y, y^k).
\]

Let $\psi, \varphi_1, \cdots, \varphi_N : \mathbb{D}([0,T]; \mathbb{R}^m) \to [0,1]$ be smooth mappings such that

(i) $\psi(y) + \sum_{k=1}^{N} \varphi_k(y) = 1, \quad \forall y \in \mathbb{D}([0,T]; \mathbb{R}^m)$;

(ii) $\text{supp} \psi \subset \bigcap_{k=1}^{N} \{ y; \beta_k(y) > \delta \};$

(iii) $\text{supp} \varphi_k \subset \bigcap_{k=1}^{N} \{ y; \beta_k(y) < 2\delta \}, \quad 1 \leq k \leq N$.

We shall introduce the following $[0,1]$-valued random variables:
\[
\xi_n := \psi(Y^{\varepsilon_n}), \quad \xi := \psi(Y),
\]

\[
\eta^k_n := \varphi_k(Y^{\varepsilon_n}), \quad \eta^k := \varphi_k(Y),
\]

and two measurable sets
\[
\tilde{A}_n := \bigcap_{k=1}^{N} \{ \omega; d_R(Y^{\varepsilon_n}(\omega), y^k) > \delta \},
\]
\[
\tilde{A} := \bigcap_{k=1}^{N} \{ \omega; d_R(Y(\omega), y^k) > \delta \}.
\]

In view of Lemma 11 we clearly have
\[
\text{supp} \xi_n \subseteq \tilde{A}_n, \quad \text{supp} \xi \subseteq \tilde{A},
\]

\[\text{(4.67)}\]

where for a real-valued random variable $\zeta$, its support $\text{supp} \zeta$ is defined by
\[
\text{supp} \zeta := \{ \omega \in \Omega; \zeta(\omega) \neq 0 \}.
\]

\[\text{(4.68)}\]

Set
\[
\Lambda_n(t) := \Gamma_n(t) - \varepsilon_n 2^{-r_0} R^\omega u(t_0,t) \Phi_{t_0} (Y^{\varepsilon_n}),
\]
\[
\Gamma_n(t) := \left[ \phi(Y^{\varepsilon_n}) - \phi(Y_{t_0}^{\varepsilon_n}) - \int_{0}^{t} (\nabla_y \phi(Y_s^{\varepsilon_n}), F(X_s^{\varepsilon_n}, Y_s^{\varepsilon_n})) \right] ds - \int_{0}^{t} \left( -(-\Delta)^{\frac{m}{2}} \phi(Y_s^{\varepsilon_n}) \right) ds \Phi_{t_0} (Y^{\varepsilon_n}),
\]

where $\Phi_{t_0} (\cdot)$ is a bounded function on $\mathbb{D}([0,T])$, which is measurable with respect to the sigma-field $\sigma(\omega, \omega \in \mathbb{D}([0,T]), 0 \leq t \leq t_0)$, $\phi$ is a $C_0^\infty$ function on $\mathbb{R}^m$.

For the convenience of expression, we assume that $X_{t_0}^{\varepsilon} = x, Y_{t_0}^{\varepsilon} = y$. It follows from the similar arguments used in the
proof of tightness in Lemma 9 we have
\[ \lim_{\varepsilon_n \to 0} \varepsilon_n^{2-r_0} E_{x,y} \left[ R_{\varepsilon_n}^c(t_0, t) \Phi_{t_0}(Y^\varepsilon) \right] = 0, \] (4.70)

and
\[ \lim_{\varepsilon_n \to 0} \varepsilon_n^{2-r_0} \sum_{i=1}^m E_{x,y} \left[ \int_{t_0}^t \left\langle \nabla_y \phi(Y_s^\varepsilon), G_t(X_s^\varepsilon, Y_s^\varepsilon) \partial_y G_t(X_s^\varepsilon, Y_s^\varepsilon) \right\rangle ds \Phi_{t_0}(Y^\varepsilon) \right] = 0. \] (4.71)

By the integral \( \int_0^t \int_{\mathbb{R}^m \setminus \{0\}} \vec{N}_2(ds, dy) \) in \( \mathbb{R}^m \) is a martingale with respect to the \( \sigma \)-algebras \( \mathcal{F}_t \) generated by \( \{ \vec{N}_2(s, dy); s \leq t \} \), we have
\[ E_{x,y} \left[ \int_{t_0}^t \int_{\mathbb{R}^m \setminus \{0\}} \left[ \phi(Y_s^\varepsilon + y) - \phi(Y_s^\varepsilon) \right] \vec{N}_2(ds, dy) \Phi_{t_0}(Y^\varepsilon) \right] = 0. \] (4.72)

In view of (4.71) and (4.70)–(4.72), we also have
\[ E_{x,y} [\Gamma_n(t)] \to 0, \quad n \to \infty. \] (4.73)

Define \( \Gamma_n^k \) as the random variable \( \Gamma_n \), where \( Y_s^\varepsilon \) is replaced by \( y^k \). Then we have the following result.

\textbf{Lemma 12.} For every \( \delta > 0 \), there exists a positive integer \( M_0 \), such that
\[ \sum_{k=1}^N \mathbb{E} \left[ (\Gamma_n(t) - \Gamma_n^k(t)) \eta_n^k(t) \right] \leq M_0 \delta. \] (4.74)

\textit{Proof.} By the definition of the function \( \psi \) and \( \varphi_k, k = 1, 2, \cdots, N \) in (4.65) and the equation (4.73), we have
\[ E_{x,y} (\Gamma_n \xi_n) + \sum_{k=1}^N E_{x,y} (\Gamma_n \eta_n^k) \to 0, \] (4.75)

as \( n \to \infty \).

For each positive integral \( k = 1, 2, \cdots, N \), we have
\[ \mathbb{E} \left[ (\Gamma_n(t) - \Gamma_n^k(t)) \eta_n^k(t) \right] = \mathbb{E} \left[ \phi(Y_s^\varepsilon) - (\phi(Y_s^\varepsilon) \right] \Phi_{t_0}(Y^\varepsilon) - \mathbb{E} \left[ \phi(y^k) - \phi(y^k) \right] \Phi_{t_0}(y^k) \right] \]
\[ - \left\{ \mathbb{E} \left[ \int_{t_0}^t \left\langle \nabla_y \phi(Y_s^\varepsilon), F(X_s^\varepsilon, Y_s^\varepsilon) \right\rangle ds \right] \Phi_{t_0}(Y^\varepsilon) - \mathbb{E} \left[ \int_{t_0}^t \left\langle \nabla_y \phi(y^k), F(X_s^\varepsilon, y^k) \right\rangle ds \right] \Phi_{t_0}(y^k) \right\} \]
\[ - \left\{ \mathbb{E} \left[ \int_{t_0}^t \left\langle -(-\Delta_y)^{\frac{a_2}{2}} \phi(Y_s^\varepsilon) \right\rangle ds \right] \Phi_{t_0}(Y^\varepsilon) - \mathbb{E} \left[ \int_{t_0}^t \left\langle -(-\Delta_y)^{\frac{a_2}{2}} \phi(y^k) \right\rangle ds \right] \Phi_{t_0}(y^k) \right\}. \] (4.76)

By \( \phi \in C_0^\infty \), Hypothesis (A_F) and the boundedness of \( \Phi_{t_0}(\cdot) \), we get the required result.

Set
\[ \Gamma(t) = \left[ \phi(Y_t) - \phi(Y_{t_0}) - \int_{t_0}^t \left\langle \nabla_y \phi(Y_s), \tilde{F}(Y_s) \right\rangle ds - \int_{t_0}^t \left( -(-\Delta_y)^{\frac{a_2}{2}} \phi(Y_s) \right) ds \right] \Phi_{t_0}(Y), \] (4.77)

where the function \( \tilde{F}(y) \) is defined by
\[ \tilde{F}(y) = \int_{\mathbb{R}^n} F(x, y) \mu^y(dx). \] (4.78)

Define \( \Gamma^k \) as the quantity obtained by replacing \( Y \) by \( y^k \) in the expression for \( \Gamma \). By the same technique as Lemma 12 we get the following corollary.
**Corollary 1.** For every \( \delta > 0 \), there exists a positive integer \( M \), such that

\[
\sum_{k=1}^{N} \mathbb{E} \left[ \left( \Gamma(t) - \Gamma^{k}(t) \right) \eta^{k}(t) \right] \leq M\delta. \tag{4.79}
\]

Next we will give the \( L^{1}(\Omega) \)-convergence of \( \Gamma^{k}_{n}(t) \).

**Lemma 13.** Let \( K \in \mathcal{C}^{1,0}_{\partial} \) and \( \bar{K}(y) := \int_{\mathbb{R}^{n}} K(x, y) \mu(dx) \), then for every \( 0 < t < T \), we have

\[
\mathbb{E} \left| \int_{0}^{t} (K(X_{s}^{x}, y_{s}^{k}) - \bar{K}(y_{s}^{k})) \, ds \right| \to 0, \quad n \to \infty. \tag{4.80}
\]

**Proof.** Let \( (a_{k}, b_{k}) \subseteq [0, T] \) be an interval on which \( y_{s}^{k} \) is a constant, denoted by \( z^{k} \). This can be done by Lemma 10. Then we will only show

\[
\mathbb{E} \left| \int_{a_{k}}^{b_{k}} [K(X_{s}^{x}, z^{k}) - \bar{K}(z^{k})] \, ds \right| \to 0, \quad n \to \infty. \tag{4.81}
\]

By the equation 1.1, we know that

\[
X_{t}^{x} = x + \int_{0}^{t} b(X_{u}^{x}, Y_{u}^{x}) \, du + \frac{1}{\varepsilon_{n}} L_{t}^{\alpha_{1}},
\]

\[
Y_{t}^{x} = y + \varepsilon_{n}^{2} \int_{0}^{t} F(X_{u}^{x}, Y_{u}^{x}) \, du + \varepsilon_{n}^{2} \int_{0}^{t} G(X_{u}^{x}, Y_{u}^{x}) \, du + L_{t}^{\alpha_{2}}.
\]

Then by Hypotheses (\( \mathcal{A}_{F} \)),(\( \mathcal{A}_{G} \)) and Lemma \( \mathcal{S} \) we have

\[
\mathbb{E} \left| Y_{t}^{x} - y \right| \to 0, \quad n \to \infty. \tag{4.83}
\]

On the one hand, by the self-similarity of \( \alpha \)-stable process, we have

\[
\mathbb{E} \left| \frac{1}{\varepsilon_{n}^{2}} L_{t}^{\alpha_{1}} \right| = \mathbb{E} |L_{t}^{\alpha_{1}}| < \infty. \tag{4.84}
\]

Therefore we have

\[
\mathbb{E} \left| X_{t}^{x} - X_{t}^{y} \right| = \mathbb{E} \left| \int_{0}^{t} \left[ b(X_{u}^{x}, Y_{u}^{x}) - b(X_{u}^{y}, y) \right] \, du \right| + \mathbb{E} \left| \frac{1}{\varepsilon_{n}^{2}} L_{t}^{\alpha_{1}} - L_{t}^{\alpha_{1}} \right| \leq \| \nabla b \|_{0} \int_{0}^{t} \mathbb{E} |X_{u}^{x} - X_{u}^{y}| \, ds + \| \nabla b \|_{0} \int_{0}^{t} \mathbb{E} |Y_{u}^{x} - y| \, du + \mathbb{E} \left| \frac{1}{\varepsilon_{n}^{2}} L_{t}^{\alpha_{1}} - L_{t}^{\alpha_{1}} \right|. \tag{4.85}
\]

By Grönwall’s inequality, (4.83) and (4.84), we have

\[
\mathbb{E} \left| X_{t}^{x} - X_{t}^{y} \right| \to 0, \quad n \to \infty. \tag{4.86}
\]
On the other hand, by $K \in C_b^{1,0}$ and Proposition 11, we know

\[ \mathbb{E} \left[ \int_{a_n}^{b_n} \left[ K(X_{s_n}^{\varepsilon_n}, z^k) - K(z^k) \right] ds \right] = \varepsilon_n^2 \mathbb{E} \left[ \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} \left[ K(X_{s_n}^{\varepsilon_n}, z^k) - K(z^k) \right] ds \right] \]
\[ \leq \varepsilon_n^2 \mathbb{E} \left[ \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} \left[ K(X_{s_n}^{\varepsilon_n}, z^k) - K(X_{s_n}^{z^k}, z^k) \right] ds + \varepsilon_n^2 \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} \mathbb{E} \left( K(X_{s_n}^{z^k}, z^k) \right) ds - \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} \int R^n K(x, z^k) \mu^k (dx) ds \right] \]
\[ \leq \varepsilon_n^2 \| \nabla x_k \|_0 \mathbb{E} \left[ \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} |X_{s_n}^{\varepsilon_n} - X_{s_n}^{z^k}| ds \right] + \varepsilon_n^2 \int_{a_k/\varepsilon_n^2}^{b_k/\varepsilon_n^2} C \| K \|_{1,0} e^{-\gamma s} \left( 1 + |x|^\frac{2}{q} \right) ds \]
\[ \to 0, \quad n \to \infty. \]

(4.87) \{?\}

4.3.1. Proof of Theorem 1.

Now, we are in the position to give:

Proof of Theorem 1. We divide the proof into the following two steps.

Step 1. By Lemma 9, there exists a subsequence $\{\varepsilon_n\} \to 0$, such that $Y^{\varepsilon_n}$ converges weakly to a limit point $Y$, as $n \to \infty$.

For any $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, by Hölder inequality, Lemma 11 and $\phi \in C_0^\infty$, we have

\[ |\mathbb{E} (\Gamma_n(t) \xi_n(t))| \leq (\mathbb{E}(\Gamma_n(t))^{p})^{1/p} \left( \mathbb{P}(\tilde{A}_n(t)) \right)^{1/q} \leq C \delta^{1/q}, \]

(4.88) \text{Step}

\[ |\mathbb{E} (\Gamma(t) \xi(t))| \leq (\mathbb{E}(\Gamma(t))^{p})^{1/p} \left( \mathbb{P}(\tilde{A}(t)) \right)^{1/q} \leq C \delta^{1/q}. \]

Since $\eta_n^k \Rightarrow \eta^k$ and $\eta_n^k \leq 1$. Thus for each $k = 1, 2, \cdots, N$, we have

\[ |\mathbb{E} (\Gamma_n^k \eta_n^k - \Gamma^k \eta^k)| \leq \mathbb{E} |(\Gamma_n^k - \Gamma^k) \eta_n^k| + |\eta_n^k| \mathbb{E} |(\eta_n^k - \eta^k)| \]
\[ \leq \mathbb{E} |\Gamma_n^k - \Gamma^k| + \mathbb{E} |\eta_n^k - \eta^k| \times \Gamma^k. \]

(4.89) \text{Real}

Step 2. By the definitions of $\psi$ and $\varphi$, we yield the following fact

\[ \mathbb{E} (\Gamma_n \xi_n(t)) + \sum_{k=1}^{N} \mathbb{E}_{x,y} (\Gamma_n \eta_n^k(t)) \to 0, \quad as \quad n \to \infty, \]

(4.90) \{?\}

\[ \mathbb{E}(\Gamma) = \mathbb{E}(\Gamma \xi_t) + \sum_{k=1}^{N} \mathbb{E}_{x,y} (\Gamma \eta_n^k), \]

and

\[ \sum_{k=1}^{N} \mathbb{E}(\Gamma_n \eta_n^k) = \sum_{k=1}^{N} \mathbb{E} \left[ (\Gamma_n - \Gamma^k) \eta_n^k \right] + \sum_{k=1}^{N} \mathbb{E}(\Gamma_n \eta^k), \]

(4.91) \text{Step}

\[ \sum_{k=1}^{N} \mathbb{E}(\Gamma \eta^k) = \sum_{k=1}^{N} \mathbb{E} \left[ (\Gamma - \Gamma^k) \eta_k \right] + \sum_{k=1}^{N} \mathbb{E}(\Gamma \eta^k). \]

By Lemma 12, 14.88, 14.89 and 14.91, we get

\[ \mathbb{E}[\Gamma_n] \to \mathbb{E}[\Gamma], \quad as \quad n \to \infty. \]

(4.92) \text{Conver}
Combining \((4.73)\) and \((4.92)\), we have
\[
E \left[ \left( \phi(Y_t) - \phi(Y_{t_0}) - \int_{t_0}^{t} \mathcal{L}_2 \phi(Y_s) ds \right) \Phi_{t_0}(Y_{t_0}) \right] = 0,
\]
(4.93) \(\square\)
where
\[
\mathcal{L}_2 \phi(y) = \langle \nabla_y \phi(y), \bar{F}(y) \rangle - (-\Delta_y)^{\frac{\alpha}{2}} \phi(y).
\]
(4.94) \(\square\)

5. Application to a specific example.

In this section, we will apply our main results to establish the weak averaging principle for a class of multiscale stochastic dynamical systems driven by \(\alpha\)-stable noises.

**Example.** Consider the following toy model
\[
\begin{cases}
    dX_t^\varepsilon = -\frac{X_t^\varepsilon}{\varepsilon^2} dt + \frac{1}{\varepsilon^{\alpha_1}} dL_t^{\alpha_1}, & X_0^\varepsilon = x \in \mathbb{R}, \\
    dY_t^\varepsilon = \sin \frac{X_t^\varepsilon}{\varepsilon^{r_0}} dt + dL_t^{\alpha_2}, & Y_0^\varepsilon = y \in \mathbb{R},
\end{cases}
\]
(5.1) \(\square\)
where \(b(x) = -x, \alpha_1 = 1.5, F(x,y) = 0, G(x,y) = \sin x\) and \(r_0 = 1 - \frac{1}{\alpha_1} = \frac{1}{3}\). It is easy to justify that \(b,F,G\) satisfy Hypotheses \((A_b), (A_F), (A_G1),(A_G2)\). Using a result in [22], we find the invariant measure \(\mu(dx) = \rho(x)dx\) with density
\[
\rho(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ix\xi} e^{-\frac{1}{\alpha_1} |\xi|^\alpha} d\xi = \frac{1}{\pi} \int_0^{\infty} \cos x\xi \cdot e^{-\frac{1}{\alpha_1} \xi^\alpha} d\xi.
\]
(5.2) \(\square\)

By Theorem \(\square\) the averaged equation for \(Y_t^\varepsilon\) is
\[
d\bar{Y}_t = dL_t^{\alpha_2}, \quad \bar{Y}_0 = y,
\]
(5.3) \(\square\)
and the slow component \(Y^\varepsilon\) converges weakly to a process \(Y\) as the scale parameter \(\varepsilon\) goes to zero.

6. Concluding remarks.

In this paper, we study the weak averaging principle for the multiscale systems driven by \(\alpha\)-stable noises. First we exam the existence of the solution of a Poisson equation for the nonlocal elliptic operator corresponding to an ergodic jump process. Then by constructing suitable correctors, we obtain the tightness of slow processes. It turns out that the slow components weakly converge to a Lévy process as the scale parameter goes to zero.

There are some limitations for this paper. Firstly, the condition \(1 < \alpha_i < 2, i = 1, 2\) plays an important role in deriving the effective dynamical system. How to obtain the effective low dimensional system and estimate the effect that the fast components have on slow ones is still an open problem with \(\alpha_i \in (0,1)\) ? Secondly, the slow components contain homogenization terms, whose homogenizing index \(r_0\) has a relation with the stable index \(\alpha_1\) of the noise of fast components given by \(0 < r_0 < 1 - 1/\alpha_1\). How to relax the restriction for the homogenizing index is also an active issue ? Thirdly, the multiscale stochastic dynamical systems are driven by additive stable Lévy noises. For a family of multiscale stochastic dynamical systems driven by multiplicative noises, how to obtain the effective low dimensional system and estimate the effect that the fast components have on slow ones is also an open problem ? Finally, it is worth emphasizing that the
estimate for (3.8) depends on the variable \(x\), so the classical semigroup method will be no longer suitable. Therefore, for such a case, it is necessary to find some new approaches to study.

In the future works, we will exam the uniqueness and regularity of nonlocal elliptic equation under the assumption of exponential ergodicity. Further, we will give the relationship between the convergence and regularity of the coefficients of the systems with respect to the slow variable. With that, we will exam the central limit theorem with homogenization and apply these results to a nonlinear filtering problem.

Acknowledgements. We would like to thank Wei Wei (Huazhong University of Science and Technology, China) and Xiaobin Sun (Jiangsu Normal University, China) for helpful discussions. The research of Y. Zhang was supported by the NSFC grant 11901202. The research of X. Wang was supported by the NSFC grant 11901159. The research of Z. Wang was supported by the NSFC grant 11531006 and 11771449.

7. Appendix A. Further Proofs

A.1. Proof of Lemma 4. By the equation (4.4), we have

\[
d(X_t^{x_1,y_1} - X_t^{x_2,y_2}) = [b(X_t^{x_1,y_1}, y_1) - b(X_t^{x_2,y_2}, y_2)] dt, \quad X_0^{x_1,y_1} - X_0^{x_2,y_2} = x_1 - x_2.
\] (7.1)

Multiplying both sides by \(2 (X_t^{x_1,y_1} - X_t^{x_2,y_2})\), by Assumption \((A_0)\) and Young’s inequality, we have

\[
\frac{d}{dt} |X_t^{x_1,y_1} - X_t^{x_2,y_2}|^2 = 2 \langle b(X_t^{x_1,y_1}, y_1) - b(X_t^{x_2,y_2}, y_2), X_t^{x_1,y_1} - X_t^{x_2,y_2}\rangle \\
\leq 2 \langle b(X_t^{x_1,y_1}, y_1) - b(X_t^{x_2,y_2}, y_2), X_t^{x_1,y_1} - X_t^{x_2,y_2}\rangle \\
+ 2 \langle b(X_t^{x_2,y_2}, y_1) - b(X_t^{x_2,y_2}, y_2), X_t^{x_1,y_1} - X_t^{x_2,y_2}\rangle \\
\leq -2\gamma |X_t^{x_1,y_1} - X_t^{x_2,y_2}|^2 + C(\|\nabla b\|_0, \gamma) |y_1 - y_2|^2 |X_t^{x_1,y_1} - X_t^{x_2,y_2}| \\
\leq -\gamma |X_t^{x_1,y_1} - X_t^{x_2,y_2}|^2 + C(\|\nabla b\|_0, \gamma) |y_1 - y_2|^2.
\] (7.2)

Hence, the comparison theorem yields that

\[
|X_t^{x_1,y_1} - X_t^{x_2,y_2}|^2 \leq e^{-2\gamma t} |x_1 - x_2|^2 + C(\|b\|_1, \gamma) |y_1 - y_2|^2.
\] (7.3)

A.2. Proof of Lemma 5. Note that

\[
d(\nabla_y X_t^{x,y}) = (\nabla_x b)(X_t^{x,y}, y) \nabla_y X_t^{x,y} dt + (\nabla_y b)(X_t^{x,y}, y) dt, \quad \nabla_y X_t^{x,y} = 0.
\] (7.4)

This implies that

\[
d(\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}) = ((\nabla_x b)(X_t^{x_1,y_1}, y_1) \nabla_y X_t^{x_1,y_1} - (\nabla_x b)(X_t^{x_2,y_2}, y_2) \nabla_y X_t^{x_2,y_2}) dt \\
+ ((\nabla_y b)(X_t^{x_1,y_1}, y_1) - (\nabla_y b)(X_t^{x_2,y_2}, y_2)) dt.
\] (7.5)
Multiplying both sides by \(2(\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2})\), we have

\[
\frac{d}{dt} |\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}|^2
= 2(\nabla_x b)(X_t^{x_1,y_1}, y_1) \nabla_y X_t^{x_1,y_1} - (\nabla_x b)(X_t^{x_2,y_2}, y_2) \nabla_y X_t^{x_2,y_2} \nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2})
+ 2(\nabla_y b)(X_t^{x_1,y_1}, y_1) - (\nabla_y b)(X_t^{x_2,y_2}, y_2), \nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2})
\]
(7.6) \(?\)

For the term \(\Sigma_1\), observe that by substituting \(x_2 = x_1 + \epsilon h\) and letting \(\epsilon \to 0\) in the dissipative condition 2.7, we have for all \(x, h \in \mathbb{R}^n, y \in \mathbb{R}^m,\)

\[ (\nabla_x b(x, y) h, h) \leq -\gamma |h|^2. \]  
(7.7) \(\text{diss}^y\)

Therefore we have

\[
\Sigma_1 \leq 2(\nabla_x b)(X_t^{x_1,y_1}, y_1) \nabla_y X_t^{x_1,y_1} - (\nabla_x b)(X_t^{x_1,y_1}, y_1) \nabla_y X_t^{x_2,y_2} \nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2})
+ 2(\nabla_y b)(X_t^{x_1,y_1}, y_1) - (\nabla_y b)(X_t^{x_2,y_2}, y_2), \nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2})
\leq -2\gamma |\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}|^2
\]
(7.8) \(?\)

For the term \(\Sigma_2\), we have

\[
\Sigma_2 \leq 2(\nabla_y b)(X_t^{x_1,y_1}, y_1) - (\nabla_y b)(X_t^{x_2,y_2}, y_2)||\nabla_y X_t^{x_2,y_2}|^2
+ 2(\nabla_y b)(X_t^{x_1,y_1}, y_1) - (\nabla_y b)(X_t^{x_2,y_2}, y_2)||\nabla_y X_t^{x_2,y_2}|^2
\leq 2||\nabla_y b||_0 |y_1 - y_2| ||\nabla_y X_t^{x_2,y_2}|^2
\]
(7.9) \(?\)

Obviously, Lemma 4 implies that

\[
\sup_{t \geq 0, x \in \mathbb{R}^n, y \in \mathbb{R}^m} |\nabla_y X_t^{x,y}| \leq C(\|b\|_1, \gamma). \]  
(7.10) \(\text{deryb}\)

Hence, by the assumption \(b \in C_b^{2,2}\) and Young’s inequality, we have

\[
\frac{d}{dt} |\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}|^2 \leq -\gamma |\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}|^2 + C(\|b\|_2, \gamma) \left( |y_1 - y_2|^2 + e^{-\frac{\gamma}{2}t} |x_1 - x_2|^2 \right). \]  
(7.11) \(?\)

By the comparison theorem, we yields that

\[
|\nabla_y X_t^{x_1,y_1} - \nabla_y X_t^{x_2,y_2}|^2 \leq C(\|b\|_2, \gamma) \left( t e^{-\frac{\gamma}{2}t} |x_1 - x_2|^2 + |y_1 - y_2|^2 \right). \]  
(7.12) \(?\)

A.3. Proof of Lemma 6. Note that

\[
d\nabla_x X_t^{x,y} = (\nabla_x b)(X_t^{x,y}, y) \cdot \nabla_x X_t^{x,y} dt, \quad \nabla_x X_0^{x,y} = I. \]  
(7.13) \(?\)
This implies

\[ d(\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}) = [(\nabla_x b)(X_t^{x_1,y_1}, y_1) \cdot \nabla_x X_{t}^{x_1,y_1} - (\nabla_x b)(X_t^{x_2,y_2}, y_2) \cdot \nabla_x X_{t}^{x_2,y_2}] dt. \] (7.14)

Multiplying both sides by \(2(\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2})\), and using the inequality (7.7), we have

\[
\frac{d}{dt}|\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|^2
\]

\[= 2((\nabla_x b)(X_t^{x_1,y_1}, y_1)\nabla_x X_{t}^{x_1,y_1} - (\nabla_x b)(X_t^{x_2,y_2}, y_2)\nabla_x X_{t}^{x_2,y_2}, \nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2})\]

\[\leq 2((\nabla_x b)(X_t^{x_1,y_1}, y_1)\nabla_x X_{t}^{x_1,y_1} - (\nabla_x b)(X_t^{x_2,y_2}, y_1)\nabla_x X_{t}^{x_2,y_2}, \nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2})\]

\[+ 2((\nabla_x b)(X_t^{x_1,y_1}, y_1)\nabla_x X_{t}^{x_2,y_2} - (\nabla_x b)(X_t^{x_1,y_1}, y_1)\nabla_x X_{t}^{x_2,y_2}, \nabla_x X_{t}^{x_2,y_2} - \nabla_x X_{t}^{x_2,y_2})\]

\[+ 2((\nabla_x b)(X_t^{x_1,y_1}, y_2)\nabla_x X_{t}^{x_2,y_2} - (\nabla_x b)(X_t^{x_2,y_2}, y_2)\nabla_x X_{t}^{x_2,y_2}, \nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2})\]

\[\leq -2\gamma|\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|^2\]

\[+ 2||\nabla_x b||_0|\nabla_x X_{t}^{x_2,y_2}||y_1 - y_2||\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|\]

\[+ 2||\nabla_x b||_0|\nabla_x X_{t}^{x_2,y_2}||X_{t}^{x_1,y_1} - X_{t}^{x_2,y_2}||\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|.\] (7.15)

By Lemma 4 we have

\[\sup_{t \geq 0, x \in \mathbb{R}^n, y \in \mathbb{R}^m}|\nabla_x X_{t}^{x,y}|^2 \leq e^{-\frac{2\gamma t}{2}}.\] (7.16)

Hence, by the assumption \(b \in C^2_b\) and Young’s inequality, we have

\[\frac{d}{dt}|\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|^2 \leq -\gamma|\nabla_y X_{t}^{x_1,y_1} - \nabla_y X_{t}^{x_2,y_2}|^2 + C(||b||_2, \gamma)e^{-\frac{2\gamma t}{2}} (|y_1 - y_2|^2 + |x_1 - x_2|^2),\] (7.17)

and then the comparison theorem yields,

\[|\nabla_x X_{t}^{x_1,y_1} - \nabla_x X_{t}^{x_2,y_2}|^2 \leq C(||b||_2, \gamma)t e^{-\frac{2\gamma t}{2}} (|y_1 - y_2|^2 + |x_1 - x_2|^2).\] (7.18)

**A.4. Proof of Lemma 7.** (i) Set

\[\bar{G}(y) = \int_{\mathbb{R}^n} G(x, y) \mu(dx).\] (7.19)

Then Hypothesis (A_G2) yields \(\bar{G} \equiv 0\). By Proposition 1 and Hypothesis (A_G1), we have

\[|\bar{G}(x, y)| \leq \int_0^\infty |\mathbb{E}[G(X_t^{x,y}, y)] - \bar{G}(y)| dt,
\]

\[\leq C(1 + |x|^\frac{1}{2}) \int_0^\infty e^{-\frac{\alpha t}{2}} dt\]

\[\leq C(1 + |x|^\frac{1}{2}).\] (7.20)

(ii) Note that

\[\nabla_x \bar{G}(x, y) = \int_0^\infty \mathbb{E}[\nabla_x G(X_t^{x,y}, y) \cdot \nabla_x X_t^{x,y}] dt,\] (7.21)

where \(\nabla_x X_t^{x,y}\) satisfies

\[\left\{ \begin{array}{l}
\frac{d}{dt}X_t^{x,y} = \nabla_x b(X_t^{x,y}, y) \cdot \nabla_x X_t^{x,y} dt, \\
\nabla_x X_t^{x,y}|_{t=0} = I.
\end{array} \right.\] (7.22)
By Lemma 4 we have

\[ \sup_{x,y} |\nabla_x X_t^{1,y}| \leq C e^{-\frac{t}{2}}. \tag{7.23} \]

Thus by Hypothesis (A1), we have

\[ \sup_{x,y} |\nabla_x \tilde{G}(x,y)| \leq C. \tag{7.24} \]

(iii) Set

\[ \tilde{G}_{t_0}(x,y,t) := EG(X_t^{x,y},y) - EG(X_t^{x,y},y) =: \tilde{G}(x,y,t) - \hat{G}(x,y,t + t_0). \tag{7.25} \]

Then Proposition 1 implies that

\[ \lim_{t_0 \to \infty} \tilde{G}_{t_0}(x,y,t) = EG(X_t^{x,y},y) - \hat{G}(y) = EG(X_t^{x,y},y). \tag{7.26} \]

On the one hand, by the Markov property, we have

\[ \tilde{G}_{t_0}(x,y,t) = \hat{G}(x,y,t) - \mathbb{E} \tilde{G}(X^{x,y}_{t_0},y,t). \tag{7.27} \]

Thus we have

\[ \nabla_y \tilde{G}_{t_0}(x,y,t) = \nabla_y \hat{G}(x,y,t) - \mathbb{E} \left[ \nabla_y \tilde{G}(X^{x,y}_{t_0},y,t) \right] - \mathbb{E} \left[ \nabla_x \tilde{G}(X^{x,y}_{t_0},y,t) \cdot \nabla_y X^{x,y}_{t_0} \right]. \tag{7.28} \]

Moreover, we also have

\[ \nabla_x \hat{G}(x,y,t) = \mathbb{E} \left[ \nabla_x G(X_t^{x,y},y) \cdot \nabla_x X_t^{x,y} \right]. \tag{7.29} \]

By Hypothesis (A1) and Lemma 4 we have

\[ \sup_{x,y} |\nabla_x \hat{G}(x,y,t)| \leq C e^{-\frac{t}{2}}. \tag{7.30} \]

On the other hand, we have

\[
\begin{align*}
|\nabla_y \hat{G}(x_1,y,t) - \nabla_y \hat{G}(x_2,y,t)| &= |\nabla_y (\mathbb{E} G(X_t^{x_1,y},y)) - \nabla_y (\mathbb{E} G(X_t^{x_2,y},y))| \\
&= \mathbb{E} |\nabla_x G(X_t^{x_1,y},y) \cdot \nabla_y X_t^{x_1,y} - \nabla_x G(X_t^{x_2,y},y) \cdot \nabla_y X_t^{x_2,y}| \\
&\quad + \mathbb{E} |\nabla_y G(X_t^{x_1,y},y) - \nabla_y G(X_t^{x_2,y},y)| \\
&\leq \mathbb{E} |\nabla_x G(X_t^{x_1,y},y) \cdot \nabla_y X_t^{x_1,y} - \nabla_x G(X_t^{x_2,y},y) \cdot \nabla_y X_t^{x_1,y}| \\
&\quad + \mathbb{E} |\nabla_x G(X_t^{x_2,y},y) \cdot \nabla_y X_t^{x_1,y} - \nabla_x G(X_t^{x_2,y},y) \cdot \nabla_y X_t^{x_2,y}| \\
&\quad + \mathbb{E} |\nabla_y G(X_t^{x_1,y},y) - \nabla_y G(X_t^{x_2,y},y)| \\
&:= S_1 + S_2 + S_3. \tag{7.31} \end{align*}
\]

For the term $S_1$, by the boundedness of $\nabla_x G$ and $\nabla_x \nabla_y G$ in Hypothesis (A1), (7.10) and Lemma 4 we have

\[ S_1 \leq C \mathbb{E} \left[ \frac{1}{2} \left( |X_t^{x_1,y} - X_t^{x_2,y}| \right)^{1/2} \right] \leq C e^{-\frac{t}{2}} |x_1 - x_2|. \tag{7.32} \]
For the term $S_2$, by the boundedness of $\nabla_x G$ and Lemma 5, we have

$$S_2 \leq C \mathbb{E} \left[ \left\| \nabla_y X^{x^{1:y}}_t - \nabla_y X^{x^{2:y}}_t \right\|^{1/2} \right] \leq C t^{\frac{4}{7}} e^{-\frac{2t}{7}} |x_1 - x_2|^{\frac{1}{7}}. \quad (7.33)$$

For the term $S_3$, by the boundedness of $\nabla_y G$ and $\nabla_x \nabla_y G$ and Lemma 4, we have

$$S_3 \leq C \mathbb{E} \left[ |X^{x^{1:y}}_t - X^{x^{2:y}}_t|^{1/2} \right] \leq C e^{-\frac{2t}{7}} |x_1 - x_2|^{\frac{1}{7}}. \quad (7.34)$$

Combining these together, we achieve from (7.31) that

$$|\nabla_y \tilde{G}(x_1, y, t) - \nabla_y \tilde{G}(x_2, y, t)| \leq C(1 + t^{\frac{4}{7}}) e^{-\frac{4}{7} \gamma_1} |x_1 - x_2|^{\frac{1}{7}}. \quad (7.35)$$

Therefore, by (7.30) - (7.35), Lemma 4 and Lemma 2, we have

$$\left| \nabla_y \tilde{G}_{t_0}(x, y, t) \right| = \left| \mathbb{E} \left[ \nabla_y \tilde{G}(x, y, t) - \nabla_y \tilde{G}(X^{x,y}_{t_0}, y, t) \right] \right| - \mathbb{E} \left[ \nabla_x \tilde{G} (X^{x,y}_{t_0}, x, y) \cdot \nabla_y X^{x,y}_{t_0} \right]$$

$$\leq C(1 + t^{\frac{4}{7}}) e^{-\frac{4}{7} \gamma_1} \mathbb{E} \left[ |X^{x,y}_{t_0} - x|^{\frac{1}{7}} \right] + C e^{-\frac{2t}{7}}$$

$$\leq C(1 + t^{\frac{4}{7}}) e^{-\frac{4}{7} \gamma_1} (1 + |x|^{\frac{1}{7}}). \quad (7.36)$$

This together with (7.26) implies that

$$\left| \nabla_y \tilde{G}(x, y) \right| = \left| \int_0^\infty \left( \lim_{t_0 \to \infty} \nabla_y \tilde{G}_{t_0}(x, y, t) \right) dt \right|$$

$$\leq \int_0^\infty C(1 + t^{\frac{4}{7}}) e^{-\frac{4}{7} \gamma_1} (1 + |x|^{\frac{1}{7}}) dt$$

$$\leq C(1 + |x|^{\frac{1}{7}}). \quad (7.37)$$

(iv) Note that by (7.26),

$$\left| \nabla^2_y \tilde{G}(x, y) \right| = \left| \int_0^\infty \left( \lim_{t_0 \to \infty} \nabla^2_y \tilde{G}_{t_0}(x, y, t) \right) dt \right|, \quad (7.38)$$

and

$$\nabla^2_y \tilde{G}_{t_0}(x, y, t) = \mathbb{E} \left[ \nabla^2_y \tilde{G}(x, y, t) - \nabla^2_y \tilde{G}(X^{x,y}_{t_0}, y, t) \right] - \mathbb{E} \left[ \nabla_x \nabla_y \tilde{G} (X^{x,y}_{t_0}, x, y) \cdot \nabla_y X^{x,y}_{t_0} \right]$$

$$- \mathbb{E} \left[ \nabla_x \tilde{G} (X^{x,y}_{t_0}, x, y) \cdot \nabla^2_y X^{x,y}_{t_0} \right] - \mathbb{E} \left[ \nabla_y \tilde{G} (X^{x,y}_{t_0}, y, t) \cdot \nabla^2_y X^{x,y}_{t_0} \right]$$

$$=: T_1 - T_2 - T_3 - T_4 - T_5. \quad (7.39)$$

where $\nabla_y X^{x,y}_t$ satisfies

$$\begin{cases} d\nabla_y X^{x,y}_t = \nabla_x b(X^{x,y}_t, y) \cdot \nabla_y X^{x,y}_t dt + \nabla_y b(X^{x,y}_t, y) dt, \\ \nabla_y X^{x,y}_t|_{t=0} = 0, \end{cases} \quad (7.40)$$

and $\nabla^2_y X^{x,y}_t$ satisfies

$$\begin{cases} d\nabla_y^2 X^{x,y}_t = \nabla^2_y b(X^{x,y}_t, y) \cdot (\nabla_y X^{x,y}_t)^2 dt + \nabla_y b(X^{x,y}_t, y) \cdot \nabla_y X^{x,y}_t dt \\ + (\nabla_x b(X^{x,y}_t, y) \cdot \nabla^2_y X^{x,y}_t dt + \nabla_x (\nabla_y b(X^{x,y}_t, y) \cdot \nabla_y X^{x,y}_t) dt + \nabla^2_y b(X^{x,y}_t, y) dt, \end{cases} \quad (7.41)$$

and $\nabla^2_y X^{x,y}_t|_{t=0} = 0$. \quad \square
To estimate the term $T_1$, we recall that $\hat{G}(x, y, t) = \mathbb{E}G(x_t^{x,y}, y)$. We derive

$$
\nabla_y^2 \hat{G}(x_t^{x,y}, y) - \nabla_y^2 G(x_t^{x,y}, y) = \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} - \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} \\
+ 2 \left[\nabla_x \nabla_y G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} - \nabla_x \nabla_y G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y}\right] \\
+ \left[\nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y}\right].
$$

Therefore we have

$$
= T_{11} + T_{12} + T_{13}.
$$

For the term $T_{11}$, we use Lemma 4 and Lemma 5 to get

$$
\mathbb{E} (|T_{11}|) \leq \mathbb{E} \left[ \left| \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} - \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} \right| \right] \\
+ \mathbb{E} \left[ \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} - \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} \right] \\
+ \mathbb{E} \left[ \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} - \left(\nabla_y x_t^{x,y}\right)^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} \right] \\
\leq \|\nabla_y^2 G\|_0 \mathbb{E} \left[ |x_t^{x,y} - x_t^{x,y}| \cdot \nabla_y x_t^{x,y} | \nabla_y x_t^{x,y} |^2 \right] + \|\nabla^2 G\|_0 \mathbb{E} \left[ |\nabla_y x_t^{x,y} + |\nabla_y x_t^{x,y}| \cdot |\nabla_y x_t^{x,y} - \nabla_y x_t^{x,y} | \right] \\
\leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}} |x_1 - x_2|.
$$

Similarly, for the term $T_{12}$ and $T_{13}$, we have

$$
\mathbb{E} (|T_{12}|) \leq 2 \|\nabla_y^2 G\|_0 \mathbb{E} \left[ |x_t^{x,y} - x_t^{x,y}| \cdot |\nabla_y x_t^{x,y} | \right] \\
+ 2 \|\nabla_y^2 G\|_0 \mathbb{E} \left[ |\nabla_y x_t^{x,y} - \nabla_y x_t^{x,y} | \right] \\
\leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}} |x_1 - x_2|,
$$

and

$$
\mathbb{E} (|T_{13}|) \leq \|\nabla x \nabla_y^2 G\|_0 |x_t^{x,y} - x_t^{x,y}| \leq C e^{-\frac{2t}{4}} |x_1 - x_2|.
$$

Combining (7.43) - (7.45), we obtain

$$
|T_1| \leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}} \mathbb{E}|X_t^{x,y} - x| \leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}} (1 + |x|).
$$

For the term $T_2$, again by the definition of $\hat{G}(x, y, t)$, we have

$$
\nabla_x \nabla_y \hat{G}(x, y, t) = \mathbb{E} \left[ (\nabla_x x_t^{x,y})^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_y x_t^{x,y} + \nabla_x G(x_t^{x,y}, y) \cdot \nabla_x \nabla_y G(x_t^{x,y}, y) \cdot \nabla_x x_t^{x,y} \right].
$$

By Lemma 4 and Lemma 5 we get

$$
|\nabla_x \nabla_y \hat{G}(x, y, t)| \leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}}.
$$

Therefore we have

$$
|T_2| \leq C(1 + t^{\frac{2}{4}}) e^{-\frac{2t}{4}}.
$$

For the term $T_3$, we have

$$
\nabla_x^2 \hat{G}(x, y, t) = \mathbb{E} \left[ (\nabla_x x_t^{x,y})^T \cdot \nabla_y^2 G(x_t^{x,y}, y) \cdot \nabla_x x_t^{x,y} + \nabla_x G(x_t^{x,y}, y) \cdot \nabla_x^2 x_t^{x,y} \right].
$$
By Lemma 4 and Lemma 6,

\[ |T_3| \leq C \left( e^{-\frac{t}{4}} + t^2 e^{-\frac{t}{4}} \right). \]  

(7.51) \{?\}

For the term \( T_4 \), we have

\[ \nabla_y \nabla_x \tilde{G}(x,y,t) = \mathbb{E} [(\nabla_y X_t^{x,y})^T \cdot \nabla^2_s G(X_t^{x,y},y) \cdot \nabla_x X_t^{x,y} + \nabla_y \nabla_x G(X_t^{x,y},y) \cdot \nabla_x X_t^{x,y} + \nabla_x G(X_t^{x,y},y) \cdot \nabla_y \nabla_x X_t^{x,y}]. \]  

(7.52) \{?\}

By Lemma 4 and 6, we have

\[ |T_4| \leq C(1 + t^\frac{1}{2})e^{-\frac{t}{4}}. \]  

(7.53) \{?\}

For the term \( T_5 \), we have

\[ \nabla_x \tilde{G}(x,y,t) = \mathbb{E} [\nabla_x G(X_t^{x,y},y) \cdot \nabla_x X_t^{x,y}]. \]  

(7.54) \{?\}

By Lemma 4 and Lemma 5 we get

\[ |T_5| \leq Ce^{-\frac{t}{4}}. \]  

(7.55) \{15\}

By (7.46)-(7.55), we get

\[ |\nabla_y^2 \tilde{G}(x,y)| \leq C \int_0^\infty \left[ (1 + t^\frac{1}{2})e^{-\frac{t}{4}}(1 + |x|) + e^{-\frac{t}{4}} \right] dt \leq C(1 + |x|). \]  

(7.56) \{?\}

8. Acknowledgments

We would like to thank Xiaobin Sun (Jiangsu Normal University) and Wei Wei (Huazhong University of Science and Technology) for helpful discussions. The research of Y. Zhang was supported by the NSFC grant 11901202. The research of X. Wang was supported by the NSFC grant 11901159. The research of Z. Wang was supported by the NSFC grants 11531006 and 11771449.

Reference

References

[EP][1] E. Pardoux, A. Y. Veretennilov, On the Poisson equation and diffusion approximation. I, Ann. Probab., 29 (2001) 1061-1085.

[GA][2] G. A. Pavliotis, A. M. Stuart, Multiscale methods: averaging and homogenization, Springer, New York, 2008.

[GA1][3] A. Papavasilion, Coarse-grained modeling of multiscale diffusion: The p-variation estimates, Spring, Heidelberg, 2011.

[JA][4] A. J. Majda, I. Timofeyev, E. V. Eijnden, A mathematical framework for stochastic climate models, Comm. Pure Appl. Math., 54 (2001) 891-974.

[MK][5] M. Katsoulakis, A. Majda, A. Sopasakis, Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part I, deterministic closures, Commun. Math. Sci., 2 (2004) 255-294.

[EA][6] R. Z. Khasminski, A limit theorem for solutions of differential equations with random right-hand side, Theory Probab. Appl., 11 (1966) 390-406.
[SD] G. C. Papanicolaou, D. W. Stroock, S. R. S. Varadhan, Martingale approach to some limit theorem, In Conference on Statistical Mechanics, Dynamical Systems and Turbulence, Duke Univ. Press, 1977.

[SM] S. Cerrai, M. Freidlin, Averaging principle for a class of stochastic reaction-diffusion equations, Probab. Theory Relat. Fields, 144 (2009) 137-177.

[VG] W. E, B. Engquist, Analysis of multiscale methods for stochastic differential equations, Commun. Pur. Appl. Math., 58 (2005) 1544-1585.

[MHL] M. Hairer, X.-M. Li, Averaging dynamics driven by fractional Brownian motion, Ann. Probab., 48 (2020) 1826-1860.

[Rock] M.R"ockner, L. J. Xie, Diffusion approximation for fully coupled stochastic differential equations, Ann. Probab., 49 (2021) 1205-1236.

[Rock11] M.R"ockner, L. J. Xie, Averaging principle and normal deviations for multiscale stochastic systems. Comm. Math. Phys., 383 (2021) 1889-1937.

[SM] J. Bao, G. Yin, C. Yuan, Two-time-scale stochastic partial differential equations driven by $\alpha$-stable noises: Averaging principles, Bernoulli, 23 (2018) 645-669.

[YZ] Y. Zhang, Z. Cheng, X. Zhang, X. Chen, J. Duan, X. Li, Data assimilation and parameter estimation for a multiscale stochastic system with $\alpha$-stable Lévy noise, J. Stat. Mech. Theory Exp., 11 (2017) 113401.

[Huang] Q. Huang, J. Duan, R. Song, Homogenization of periodic linear nonlocal partial differential equations, To be appear Bernoulli.

[Da] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge Univ. Press, Cambridge, 2004.

[Maj] M. B. Majka, Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes, Stoch. Process. Their Appl., 127 (2017) 4083-4125.

[XS1] X. Sun, L. Xie, Y. Xie, Strong and weak convergent rates for slow-fast stochastic differential equations driven by $\alpha$-stable process, To be appear Bernoulli.

[EN000] K. -J. Engel, R. Nagel, One-parameter semigroups for linear evolution equations, Springer-Verlag, New York, 2000.

[pl] P. Lachout, A skorohod space of discontinuous functions on a general set, Acta Univ. Carolinae. Math. Phys., 33 (1992) 91-97.

[FE] R. Cont, D. A. Fournié, Change of variable formulas for non-anticipative functionals on path space, J. Funct. Anal., 259 (2010) 91-97.

[SAT] S. Albeverio, B. Rüdiger, J. Wu, Invariant measures and symmetry property of Lévy type operators, Potential Anal., 13 (2000) 147-168.

[Sat99] K.-I. Sato, Lévy processes and infinitely divisible distributions, Cambridge university press, 1999.

[JAS13] J. Jacod and A. Shiryaev, Limit theorems for stochastic processes, second edition, vol. 288, Springer-Verlag Berlin Heidelberg, 2013.

[Bil68] Patrick Billingsley, Convergence of Probability Measures, John Wiley & Sons, Inc., 1968.