Homogenization of Dissipative Hamiltonian Systems under Lévy Fluctuations

Zibo Wang[a,1], **Li Lv**[a,2], **and Jinqiao Duan**[b,3]
[a] School of Mathematics and Statistics & Center for Mathematical Sciences,
Huazhong University of Science and Technology, Wuhan 430074, China
[b] Department of Applied Mathematics,
Illinois Institute of Technology, Chicago, IL 60616, USA

Abstract

This work is devoted to deriving small mass limiting equation for a class of Hamiltonian systems with multiplicative Lévy noise. Derivation of the limiting equation depends on the structure of the stochastic Hamiltonian systems, in which a noise-induced drift term arises. We prove convergence to the limiting equation in probability under appropriate assumptions on smoothness and boundedness. Furthermore, we demonstrate convergence in moment under stronger assumptions. A Lévy type Smoluchowski-Kramers approximation result is presented as an illustrative example.

Keywords: Homogenization; Hamiltonian systems; non-Gaussian Lévy noise; noise-induced drift; small mass limit; effective reduction

1. Introduction

The motion of a diffusing particle of mass m can be modeled by a stochastic differential equation (SDE)

$$dq_t = v_t dt, \qquad m dv_t = -\gamma v_t dt + \sigma dW_t,$$

where γ is the dissipation coefficient, σ is the diffusion coefficient and W is a Wiener process. The small mass limit problem was studied by Smoluchowski [1] and Kramers [2] when the mass $m \to 0$. Following their pioneering work, this subject has been investigated by a number of authors. For example, Nelson [3] derived the limiting equation when γ and σ are constants and a Fokker-Planck equation approach was provided by Doering [4]. Convergence in probability for γ constant and σ position-dependent was shown by Freidlin [5]. For the infinite dimensional case, the problem was studied by Cerrai-Freidlin [6]. These

[1]zibowang@hust.edu.cn
[2]Corresponding author: lilyu@hust.edu.cn
[3]duan@iit.edu

above problems can be illustrated in the framework of homogenization, for which a splendid relevant reference is given [7].

Recently, the phenomenon of presence of noise-induced drift term in the small mass limit problem attracted wide attentions. It arises when the dissipation and diffusion coefficients depend on the state variable. Then there will be an additional drift term which does not appear in the original system. This phenomenon was firstly discovered by Hanggi [8] for systems satisfying the fluctuation-dissipation relation. Then Volpe et al. [9] made an experimental observation for this phenomenon. Hottovy et al. [10] derived the limiting equation of SDEs with arbitrary state-dependent friction. Birrell et al. developed small mass limit theory on compact Riemannian manifolds [11] and for Hamiltonian systems [12]. A generalized homogenization theorem for Langevin systems was proved in [13]. Lim et al. [14] discussed generalized Langevin equation for non-Markovian anomalous diffusions. We point out that most existing works mentioned above are for Gaussian noise.

However, random fluctuations in nonlinear dynamical systems are often non-Gaussian [15]. The particle undergoing Lévy superdiffusion is performing motion with random jumps and step lengths following a power-law distribution [16]. As an important kind of non-Gaussian noise, Lévy noise have been found widely in atmospheric turbulence [17], epidemic spreading [18] and cell biological behaviour [19]. Lévy noise-driven non-equilibrium systems are known to manifest interesting physical properties. It is worth mentioning that Lévy noise-driven systems do not satisfy classical fluctuation dissipation relation. Therefore, linear response theory, which is viewed as a generalization of the fluctuation-dissipation theorem, has been studied for SDEs driven by Lévy noise [20, 21]. It is similar to the previous part that there are also some small mass limit results for SDEs driven by Lévy noise. For example, Talibi [22] developed Nelson theory for the α-stable Lévy process. Zhang [23] obtained Smoluchowski-Kramers approximation for SDEs driven by Lévy noise whose moment is finite.

Hamiltonian dynamics [24], as an equivalent description of Newton's second law in the framework of classical mechanics, form the framework of statistical mechanics. Dissipative Hamiltonian systems with noise have been investigated recently [25, 26].

In this present paper, we derive the small mass limiting equation of a class of dissipative Hamiltonian systems with Lévy noise

$$\begin{aligned} dq_t^\varepsilon &= \nabla_p H^\varepsilon(t, x_t^\varepsilon)dt, \\ dp_t^\varepsilon &= (-\gamma(t, x_t^\varepsilon)\nabla_p H^\varepsilon(t, x_t^\varepsilon) - \nabla_q H^\varepsilon(t, x_t^\varepsilon) + F(t, x_t^\varepsilon))dt + \sigma(t, x_{t-}^\varepsilon)dL_t, \end{aligned} \tag{1.1}$$

where $x_t^\varepsilon = (q_t^\varepsilon, p_t^\varepsilon)$ and H is a Hamiltonian function with small mass parameter ε. The functions γ, σ and F are dissipation coefficient, diffusion coefficient and external force dependent on $(q_t^\varepsilon, p_t^\varepsilon)$, respectively. Here the process $L = \{L_t\}_{t\geq 0}$ is a Lévy process. An inspiration for this paper goes back to the work by Birrell-Wehr [12]. The main idea of proof is the following: By means of the structure of Hamiltonian systems and a Lyapunov equation, we derive the limiting equation including a noise-induced drift term. Then, we prove that under appropriate assumptions, the original systems converge to the limiting equation in moment. Finally, utilizing non-explosion property of the solution of original systems, we show the convergence in probability for weaker assumptions.

This paper is organized as follows. In Section 2, we recall some basic notations and introduce a class of dissipative Hamiltonian systems with Lévy noise. In Section 3, we state and prove the homogenization result. More precisely, in Section 3.1, we obtain the moment estimation of kinetic function and get some relevant estimation results. In Section 3.2, we derive the limiting equation by using a Lyapunov equation. In Section 3.3, we finish the proof of the main results (Theorem 3.1 and Theorem 3.2). In Section 3.4, we extend the result to some more general systems. In Section 4, we present an illustrative example .

2. Preliminaries

2.1. *Lévy motion*

Let $(\Omega, \mathbb{P})$ be a probability space. An stochastic process $L_t = L(t)$ taking values in $\mathbb{R}^n$ with $L(0) = 0$ *a.s.* (almost surely) is called an n-dimensional Lévy process if it is stochastically continuous, with independent increments and stationary increments.

An n-dimensional Lévy process L_t can be expressed by Lévy-Itô decomposition, i.e., there exist a drift vector $b \in \mathbb{R}^n$, a covariance matrix Q such that

$$L_t = bt + B_Q(t) + \int_{||x||<1} x\widetilde{N}(t, dx) + \int_{||x||\geq 1} xN(t, dx),$$

where $N(dt, dx)$ is the Poisson random measure on $\mathbb{R} \times (\mathbb{R}^n \backslash \{0\})$, $\widetilde{N}(dt, dx) \triangleq N(dt, dx) - \nu(dx)dt$ is the compensated Poisson random measure, $\nu \triangleq \mathbb{E}N(1, \cdot)$ is the jump measure, and $B_Q(t)$ is an independent n-dimensional Brownian motion with covariance matrix Q. The triple (b, Q, ν) is called the generating triple for the Lévy process L_t. A Lévy process L_t has θ-th moment if and only if $\int_{||x||>1} ||x||^\theta \nu(dx) < \infty$.

2.2. *Dissipative Hamiltonian system with Lévy noise*

We consider the dissipative Hamiltonian system described in [12]. Given a time-dependent Hamiltonian function $H(t, x_t)$, where $x_t = (q_t, p_t) \in \mathbb{R}^n \times \mathbb{R}^n$. The following Hamiltonian system describe a system with dissipative force and an external force.

$$\begin{aligned} \dot{q}_t &= \nabla_p H(t, x_t), \\ \dot{p}_t &= -\gamma(t, x_t)\nabla_p H(t, x_t) - \nabla_q H(t, x_t) + F(t, x_t), \end{aligned} \tag{2.1}$$

with dissipation coefficient $\gamma : [0,\infty)\times\mathbb{R}^{2n} \to \mathbb{R}^{n\times n}$, and external forcing function $F : [0,\infty)\times\mathbb{R}^{2n} \to \mathbb{R}^n$. A natural example for Hamiltonian function is $H(q,p) = \frac{p^2}{2m} + V(q)$, where $\frac{p^2}{2m}$ represents kinetic energy of system and m represents mass. Hence we are interested in a family of Hamiltonians depending on some small parameter ε of the form

$$H^\varepsilon(t,q,p) \triangleq K^\varepsilon(t,q,p) + V(t,q) = K(\varepsilon,t,q,p/\sqrt{\varepsilon}) + V(t,q). \tag{2.2}$$

We remark that the notation K and V may not represent physical kinetic energy and potential energy. Actually, the splitting is more extensive as long as it satisfies the assumptions we will make below. However, we still call K kinetic energy and V potential energy function in the following sections.

In this paper, we study the following Hamiltonian system perturbed by Lévy fluctuation

$$\begin{aligned} dq_t^\varepsilon &= \nabla_p H^\varepsilon(t,x_t^\varepsilon)dt, \\ dp_t^\varepsilon &= (-\gamma(t,x_t^\varepsilon)\nabla_p H^\varepsilon(t,x_t^\varepsilon) - \nabla_q H^\varepsilon(t,x_t^\varepsilon) + F(t,x_t^\varepsilon))dt + \sigma(t,x_{t-}^\varepsilon)dL_t, \end{aligned} \tag{2.3}$$

with initial data $(q_0^\varepsilon,\ p_0^\varepsilon)$, where $\sigma : [0,\infty)\times\mathbb{R}^{2n} \to \mathbb{R}^{n\times d}$ is noise intensity function and $L = \{L_t\}_{t\geq 0}$ is a $\mathbb{R}^d$-valued pure jump Lévy process with triple $(0,0,\nu)$.

Remark 2.1. *We consider only pure jump Lévy process here, since by Lévy-Itô decomposition, Lévy process could be expressed as a sum of a Brownian motion and a pure jump Lévy process, in addition to a drift term which may be absorbed in the vector field in SDE. Homogenization of dissipative Hamiltonian systems with Brownian motion was studied in [12]. Thereby we use same notations as in [12] to make sure the influence of Brownian motion can be added to our results.*

We assume that the pure jump Lévy process has finite moment. More precisely, we make the following assumption for jump measure ν.

Assumption 1. There exists a constant θ such that the Lévy measure ν satisfies

$$\int_{|x|\geq 1} |x|^{2\vee\theta}\nu(dx) < \infty,$$

here $2\vee\theta = \max\{2,\theta\}$.

3. Homogenization of dissipative Hamiltonian systems under Lévy fluctuations

In this section we formulate the assumptions and state the main results Theorem 3.1 and Theroem 3.2.

3.1. ***Moment estimates***

In this subsection, we derive the moment estimation for kinetic energy K and some relevant estimation results. For the Hamiltonian function H we make the following assumptions.

Assumption 2. The Hamiltonian function H has form (2.2), where $K(\varepsilon, t, q, z)$ is non-negative and $\mathcal{C}^2$ in (t, q, z) for each ε. Moreover, there exists a constant $C_0 > 0$ such that $K^\varepsilon(0, x_0^\varepsilon) \leq C_0$. For every fixed constant $T > 0$ and $\varepsilon_0 > 0$, the following conditions hold on $(0, \varepsilon_0] \times [0, T] \times \mathbb{R}^{2n}$:

1. There exist positive constants C, M_1 such that

$$\max\{|\partial_t K(\varepsilon, t, q, z)|, ||\nabla_q K(\varepsilon, t, q, z)||, ||\nabla_z K(\varepsilon, t, q, z)||\} \leq M_1 + CK(\varepsilon, t, q, z).$$

2. There exist positive constants c, M_2 such that

$$||\nabla_z K(\varepsilon, t, q, z)||^2 + M_2 \geq cK(\varepsilon, t, q, z).$$

3. The kinetic energy $K(\varepsilon, t, q, z)$ is Lipschitz w.r.t (with respect to) z, i.e. there exists a constant L such that

$$|K(\varepsilon, t, q, z_1) - K(\varepsilon, t, q, z_2)| \leq L|z_1 - z_2|.$$

4. The potential energy $V(t, q)$ is $\mathcal{C}^1$ in (t, q) and $\nabla_q V$ is bounded.

For dissipative matrix function γ, external force F and noise intensity σ, we assume that

Assumption 3. For every $T > 0$, the following conditions hold on $[0, T] \times \mathbb{R}^{2n}$:

1. The function γ, F, σ are bounded and Lipschitz.
2. The matrix function γ is symmetric with eignevalues bounded below by some constant $\lambda > 0$.

Remark 3.1. *Under the Assumption 1-3 and additional Assumption 4 below, the solution x_t^ε to stochastic Hamiltonian system* (2.3) *exists and is unique. See Appendix for proof.*

At this point, we begin to prove the moment estimations of K. We firstly give an upper bound of kinetic energy K.

Lemma 3.1. *For every $\theta \geq 1$ and $T > 0$ there exist positive constants α_0, ε_0 such that for all constant $\alpha \in (0, \alpha_0], \epsilon \in (0, \varepsilon_0]$ and $t \in [0, T]$, we have*

$$K^\varepsilon(t, x_t^\varepsilon)^\theta \leq \frac{\kappa(\varepsilon)}{\alpha} + \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} e^{-\alpha(t-s)/\varepsilon} [K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon + \sigma(s, x_{s-}^\varepsilon)x)^\theta - K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon)^\theta] \widetilde{N}(ds, dx), \tag{3.1}$$

where $\kappa(\varepsilon) = \kappa_1 + \kappa_2 \varepsilon^{1-\theta/2}$ for positive constants κ_1 and κ_2.

Proof. Applying Itô formula to $e^{\alpha t/\varepsilon} K^\varepsilon(t, x_t^\varepsilon)^\theta$, we have

$$
\begin{aligned}
&e^{\alpha t/\varepsilon}K^\varepsilon(t,x_t^\varepsilon)^\theta \\
&= K^\varepsilon(0,x_0^\varepsilon)^\theta + \frac{\alpha}{\varepsilon}\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,x_s^\varepsilon)^\theta ds + \theta\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,x_s^\varepsilon)^{\theta-1}(\partial_s K)^\varepsilon(s,x_s^\varepsilon)ds \\
&+ \frac{\theta}{\varepsilon}\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,x_s^\varepsilon)^{\theta-1}(\nabla_z K)^\varepsilon(s,x_s^\varepsilon)(-\gamma(s,x_s^\varepsilon))(\nabla_z K)^\varepsilon(s,x_s^\varepsilon)ds \\
&+ \frac{\theta}{\sqrt{\varepsilon}}\int_0^t e^{\alpha s/\varepsilon}(\nabla_z K)^\varepsilon(s,x_s^\varepsilon)(-\nabla_q V(s,q_s^\varepsilon)+F(s,x_s^\varepsilon))ds \\
&+ \int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{\alpha s/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_{s-}^\varepsilon)x)^\theta - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)^\theta]\widetilde{N}(ds,dx) \\
&+ \int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{\alpha s/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_{s-}^\varepsilon)x)^\theta - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)^\theta]\nu(dx)ds \qquad (I_1) \\
&- \int_0^t\int_{|x|<1} e^{\alpha s/\varepsilon}\sigma^i(s,x_{s-}^\varepsilon)x\frac{\theta}{\sqrt{\varepsilon}}K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)^{\theta-1}(\nabla_{z_i}K)^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)]\nu(dx)ds, \qquad (I_2)
\end{aligned}
$$

where we denote the last two integrals by I_1, I_2 respectively. The notation $(\partial_s K)^\varepsilon(s,x)$ is equal to $\partial_s K(\varepsilon,s,q,p/\sqrt{\varepsilon})$ and similarly for $(\nabla_z K)^\varepsilon(s,x)$.

First we estimate terms I_1, I_2. Using mean value theorem and Lipschitz condition of K for the term I_1 we have

$$
\begin{aligned}
I_1 &= \int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{\alpha s/\varepsilon}[K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon+\sigma(s,x_s^\varepsilon)x)^\theta - K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^\theta]\nu(dx)ds \\
&\le 2^{\theta-2}\theta\int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{\alpha s/\varepsilon}\left[K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^{\theta-1}\,|K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon+\sigma(s,x_s^\varepsilon)x)-K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)|\right. \\
&\left.+|K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon+\sigma(s,x_s^\varepsilon)x)-K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)|^\theta\right]\nu(dx)ds \\
&\le \frac{2^{\theta-2}\theta L||\sigma||_\infty}{\sqrt{\varepsilon}}\int_{\mathbb{R}^d\setminus\{0\}}|x|\nu(dx)\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^{2\theta-1}ds + \frac{2^{\theta-2}\theta L^\theta||\sigma||_\infty}{\varepsilon^{\theta/2}}\int_{\mathbb{R}^d\setminus\{0\}}|x|^\theta\nu(dx)\int_0^t e^{\alpha s/\varepsilon}ds.
\end{aligned}
\tag{3.2}
$$

Under Assumption 2-3, for term I_2 we have

$$
\begin{aligned}
I_2 &= -\int_0^t\int_{|x|<1} e^{\alpha s/\varepsilon}\sigma^i(s,x_s^\varepsilon)x\frac{\theta}{\sqrt{\varepsilon}}K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^{\theta-1}(\nabla_{z_i}K)^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)]\nu(dx)ds \\
&\le \frac{\theta||\sigma||_\infty}{\varepsilon}\int_{|x|<1}|x|\nu(dx)\left(M_1\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^{\theta-1}ds + C\int_0^t e^{\alpha s/\varepsilon}K^\varepsilon(s,q_s^\varepsilon,p_s^\varepsilon)^\theta ds\right).
\end{aligned}
\tag{3.3}
$$

Then combining these two inequalities (3.2), (3.3) with Assumption 2-3, we obtain

$$
\begin{aligned}
& e^{\alpha t/\varepsilon} K^\varepsilon(t, x_t^\varepsilon)^\theta \\
& \le K^\varepsilon(0, x_0^\varepsilon)^\theta + \left(\frac{\alpha}{\varepsilon} + C\theta - \frac{\lambda c\theta}{\varepsilon} + \frac{C\theta}{\sqrt{\varepsilon}}||-\nabla_q V + F||_\infty\right)\int_0^t e^{\alpha s/\varepsilon} K^\varepsilon(s, x_s^\varepsilon)^\theta ds \\
& + \theta\left(M_1 + \frac{\lambda M_2}{\varepsilon} + \frac{M_1}{\sqrt{\varepsilon}}||-\nabla_q V + F||_\infty\right)\int_0^t e^{\alpha s/\varepsilon} K^\varepsilon(s, x_s^\varepsilon)^{\theta-1} ds \\
& + \left(\frac{2^{\theta-2}\theta L||\sigma||_\infty}{\sqrt{\varepsilon}}\int_{\mathbb{R}^d\setminus\{0\}}|x|\nu(dx) + \frac{\theta||\sigma||_\infty}{\varepsilon}\int_{|x|<1}|x|\nu(dx)\right)\int_0^t e^{\alpha s/\varepsilon} K^\varepsilon(s, x_s^\varepsilon)^{\theta-1} ds \\
& + \frac{C\theta||\sigma||_\infty}{\varepsilon}\int_0^t e^{\alpha s/\varepsilon} K^\varepsilon(s, q_s, p_s)^\theta ds + \frac{2^{\theta-2}\theta L^\theta||\sigma||_\infty^\theta}{\varepsilon^{\theta/2}}\int_{\mathbb{R}^d\setminus\{0\}}|x|^\theta\nu(dx)\int_0^t e^{\alpha s/\varepsilon} ds \\
& + \int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{\alpha s/\varepsilon}[K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon + \sigma(s, x_{s-}^\varepsilon)x)^\theta - K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon)^\theta]\widetilde{N}(ds, dx).
\end{aligned}
\tag{3.4}
$$

Note that Young inequality allows $K^{\theta-1} \le \frac{1}{\theta}\left(\frac{M}{\delta}\right)^{\theta-1} + \frac{\delta}{M}K^\theta$. Let $M = \max\{M_1, M_2\}$. We get

$$
\begin{aligned}
K^\varepsilon(t, x_t^\varepsilon)^\theta \le\ & e^{-\alpha t/\varepsilon} K^\varepsilon(0, x_0^\varepsilon) - \frac{D}{\varepsilon}\int_0^t e^{-\alpha(t-s)/\varepsilon} K^\varepsilon(s, x_s^\varepsilon)^\theta ds + \frac{d}{\alpha} \\
& + \int_0^t\int_{\mathbb{R}^d\setminus\{0\}} e^{-\alpha(t-s)/\varepsilon}[K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon + \sigma(s, x_{s-}^\varepsilon)x)^\theta - K^\varepsilon(s, q_{s-}^\varepsilon, p_{s-}^\varepsilon)^\theta]\widetilde{N}(ds, dx),
\end{aligned}
\tag{3.5}
$$

where

$$
\begin{aligned}
D =& \lambda c\theta - \alpha - C\theta\varepsilon - C\theta\sqrt{\varepsilon}||-\nabla_q V + F||_\infty - \theta\delta\varepsilon - \theta\delta\lambda - \theta\delta\sqrt{\varepsilon}||-\nabla_q V + F||_\infty \\
& - 2^{\theta-2}\theta L\delta||\sigma||_\infty M^{-1}\sqrt{\varepsilon}\int_{\mathbb{R}^d\setminus\{0\}}|x|\nu(dx) - \theta\delta||\sigma||_\infty\int_{|x|<1}|x|\nu(dx) - C\theta\delta||\sigma||_\infty,
\end{aligned}
\tag{3.6}
$$

and

$$
\begin{aligned}
d =& \left(\frac{M}{\delta}\right)^{\theta-1}\left(M\varepsilon + \lambda M + M\sqrt{\varepsilon}||-\nabla_q V + F||_\infty + 2^{\theta-2}L\sqrt{\varepsilon}||\sigma||_\infty\int_{\mathbb{R}^d\setminus\{0\}}|x|\nu(dx) + ||\sigma||_\infty\int_{|x|<1}|x|\nu(dx)\right) \\
& + \left(\frac{M}{\delta}\right)^{\theta-1} 2^{\theta-2}L^\theta||\sigma||_\infty^\theta\varepsilon^{1-\theta/2}\int_{\mathbb{R}^d\setminus\{0\}}|x|^\theta\nu(dx).
\end{aligned}
\tag{3.7}
$$

For all $\varepsilon, \delta, \alpha$ sufficiently small, D is non-negative. In addition, $K^\varepsilon(0, x_0^\epsilon)$ is bounded by Assumption 2. Thus we obtain the required inequality (3.1). □

Now we give the moment estimation of the kinetic energy $K^\varepsilon(t, x_t^\varepsilon)$ by means of above assumptions and lemma.

Lemma 3.2. (Supremum of expectation of the kinetic energy) *Under Assumption 1-3, for every* *positive T and θ, the kinetic energy K has the following uniform estimate*

$$\sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t,x_t^\varepsilon)^\theta\right] = O(\varepsilon^{1-\frac{2\vee\theta}{2}}), \ \ as\ \varepsilon\to 0. \tag{3.8}$$

Proof. We first consider $\theta \geq 1$. Note that

$$\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} e^{-\alpha(t-s)/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_s^\varepsilon)x)^\theta - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)^\theta]\widetilde{N}(ds,dx)$$

is a local martingale and it is in fact a martingale by using appropriate sequence of stopping times (see [16], page 266). Then we obtain the following equality

$$\mathbb{E}\left[\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} e^{-\alpha(t-s)/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_s^\varepsilon)x)^\theta - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)^\theta]\widetilde{N}(ds,dx)\right] = 0.$$

It follows that the equality (3.8) holds from Lemma 3.1 and preceding equation for $\theta \geq 1$. The results for $0<\theta<1$ follows by Hölder's inequality. $\square$

Lemma 3.3. (Expectation of supremum of the kinetic energy) *Under Assumption 1-3 and for every positive T and θ, the kinetic energy K has the following uniform estimate*

$$\mathbb{E}\left[\sup_{t\in[0,T]} K^\varepsilon(t,x_t^\varepsilon)^\theta\right] = O(\varepsilon^{-\frac{\theta}{2}}), \ \ as\ \varepsilon\to 0. \tag{3.9}$$

Proof. By Lemma 3.1 we have

$$K^\varepsilon(t,x_t^\varepsilon) \leq \frac{\kappa}{\alpha} + \int_0^t \int_{\mathbb{R}^d\setminus\{0\}} e^{-\alpha(t-s)/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_{s-}^\varepsilon)x) - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)]\widetilde{N}(ds,dx). \tag{3.10}$$

Itô's product formula implies that

$$\begin{aligned}
&\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} e^{-\alpha(t-s)/\varepsilon}[K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_{s-}^\varepsilon)x) - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)]\widetilde{N}(ds,dx)\\
=&\int_0^t \int_{\mathbb{R}^d\setminus\{0\}} [K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon+\sigma(s,x_{s-}^\varepsilon)x) - K^\varepsilon(s,q_{s-}^\varepsilon,p_{s-}^\varepsilon)]\widetilde{N}(ds,dx)\\
&+\int_0^t \frac{\alpha}{\varepsilon} e^{-\alpha(t-s)/\varepsilon} \int_0^s \int_{\mathbb{R}^d\setminus\{0\}} [K^\varepsilon(r,q_{r-}^\varepsilon,p_{r-}^\varepsilon+\sigma(r,x_{r-}^\varepsilon)x) - K^\varepsilon(r,q_{r-}^\varepsilon,p_{r-}^\varepsilon)]\widetilde{N}(dr,dx)ds.
\end{aligned} \tag{3.11}$$

We first show the proposition in the case when $\theta \geq 2$. Substituting (3.11) into (3.10) and taking supremum and expectation on both side, we have

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} K^{\varepsilon}(t,x_t^{\varepsilon})^{\theta}\right]\\
\leq& 2^{\theta-1}\left(\frac{\kappa}{\alpha}\right)^{\theta}+4^{\theta-1}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}[K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon}+\sigma(s,x_{s-}^{\varepsilon})x)-K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon})]\widetilde{N}(ds,dx)\right|^{\theta}\right]\\
&+4^{\theta-1}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{\alpha}{\varepsilon}e^{-\alpha(t-s)/\varepsilon}\int_0^s\int_{\mathbb{R}^d\setminus\{0\}}[K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon}+\sigma(r,x_{r-}^{\varepsilon})x)-K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon})]\widetilde{N}(dr,dx)ds\right|^{\theta}\right].
\end{aligned}
\tag{3.12}
$$

For the first Poisson stochastic integral term, Kunita first inequality ([16], Theorem 4.4.23) implies that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon}+\sigma(s,x_{s-}^{\varepsilon})x)-K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon})\widetilde{N}(ds,dx)\right|^{\theta}\right]\\
\leq& D(\theta)\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}^d\setminus\{0\}}|K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon}+\sigma(s,x_{s-}^{\varepsilon})x)-K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon})|^2\nu(dx)ds\right)^{\frac{\theta}{2}}\right]\\
&+\mathbb{E}\left[\int_0^T\int_{\mathbb{R}^d\setminus\{0\}}|K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon}+\sigma(s,x_{s-}^{\varepsilon})x)-K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon})|^{\theta}\nu(dx)ds\right]\\
\leq& D(\theta)\varepsilon^{-\frac{\theta}{2}}T^{\frac{\theta}{2}}L^{\theta}||\sigma||_{\infty}^{\theta}\left(\int_{\mathbb{R}^d\setminus\{0\}}|x|^2\nu(dx)\right)^{\frac{\theta}{2}}+\varepsilon^{-\frac{\theta}{2}}TL^{\theta}||\sigma||_{\infty}^{\theta}\int_{\mathbb{R}^d\setminus\{0\}}|x|^{\theta}\nu(dx)\\
=&O(\varepsilon^{-\frac{\theta}{2}}).
\end{aligned}
\tag{3.13}
$$

Next we deal with the second Poisson stochastic integral term

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{\alpha}{\varepsilon}e^{-\alpha(t-s)/\varepsilon}\int_0^s\int_{\mathbb{R}^d\setminus\{0\}}[K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon}+\sigma(r,x_{r-}^{\varepsilon})x)-K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon})]\widetilde{N}(dr,dx)ds\right|^{\theta}\right]\\
\leq&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\frac{\alpha}{\varepsilon}e^{-\alpha(t-s)/\varepsilon}\sup_{s\in[0,t]}\left|\int_0^s\int_{\mathbb{R}^d\setminus\{0\}}[K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon}+\sigma(r,x_{r-}^{\varepsilon})x)-K^{\varepsilon}(r,q_{r-}^{\varepsilon},p_{r-}^{\varepsilon})]\widetilde{N}(dr,dx)\right|ds\right|^{\theta}\right]\\
\leq&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon}+\sigma(s,x_{s-}^{\varepsilon})x)-K^{\varepsilon}(s,q_{s-}^{\varepsilon},p_{s-}^{\varepsilon})\widetilde{N}(ds,dx)\right|^{\theta}\right]\\
=&O(\varepsilon^{-\frac{\theta}{2}}),
\end{aligned}
\tag{3.14}
$$

where the last equality is obtained by utilizing (3.13). Therefore, equality (3.9) holds for $\theta\geq 2$ by (3.12), (3.13)and (3.14). It follows for all $\theta>0$ by Hölder's inequality.

□

We make an additional assumption for kinetic energy K as follows.

Assumption 4 For every $T > 0$, there exist $c > 0, \eta > 0$ such that

$$K(\varepsilon, t, q, z) \geq c||z||^{\eta}.$$

Now we can deduce an useful proposition under this assumption. Proposition 3.1 is a direct deduction from Lemma 3.2, Lemma 3.3 and Assumption 4.

Proposition 3.1. *Under Assumption 1-4, for every $T > 0$ we have*

$$\sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right] = \begin{cases} O(\varepsilon^{\frac{\theta}{2}}), & \text{if } \theta \leq 2\eta, \\ O(\varepsilon^{\frac{\theta}{2}+1-\frac{\theta}{2\eta}}), & \text{if } \theta > 2\eta, \end{cases} \quad \text{as } \varepsilon \to 0, \tag{3.15}$$

and

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}-\frac{\theta}{2\eta}}), \text{ as } \varepsilon \to 0. \tag{3.16}$$

Proof. From Assumption 4, we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right] \leq \varepsilon^{\frac{\theta}{2}} \sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^{\frac{\theta}{\eta}}\right].$$

Note that Lemma 3.2 implies $\sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^a\right] = O(1)$ for $a \leq 2$ and $\sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^a\right] = O(\varepsilon^{1-\frac{a}{2}})$ for $a > 2$. Hence we get (3.15). Equation (3.16) follows similar arguments and Lemma 3.3. $\square$

Remark 3.2. *If the parameter η in Assumption 4 was given, then proposition 1 told us the order of momentum p_t^ε convergence to zero. For example, assume that η in Assumption 4 equals to 2, we have $\sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}})$ when $\theta \leq 4$ and $\sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right] = O(\varepsilon^{1+\frac{\theta}{4}})$ when $\theta > 4$. Moreover, $\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{4}})$.*

3.2. ***Derivation of the limit equation***

In this subsection, we derive the limit equation of the system (2.3) as $\varepsilon \to 0$. To this end we make an additional assumption on γ.

Assumption 5 Every element γ_i^j in matrix function γ is C^1 and independent of p.

Note that stochastic Hamiltonian equation (2.3) can be simplified to

$$\begin{aligned} d(q_t^\varepsilon) &= \nabla_p H^\varepsilon(t, x_t^\varepsilon)dt \\ &= \gamma^{-1}(t, x_t^\varepsilon)(\nabla_q H^\varepsilon(t, x_t^\varepsilon) - F(t, x_t^\varepsilon))dt + \gamma^{-1}(t, x_t^\varepsilon)\sigma(t, x_{t-}^\varepsilon)dL_t - \gamma^{-1}(t, x_t^\varepsilon)d(p_t^\varepsilon). \end{aligned} \tag{3.17}$$

Since matrix function γ has bounded eigenvalues, γ is invertible. Taking stochastic integration by parts formula for the last term $\gamma^{-1}(t, x_t^\varepsilon)d(p_t^\varepsilon)$ on the right hand side of (3.17), we have

$$\begin{aligned}(\gamma^{-1})_i^j(t, q_t^\varepsilon)d(p_t^\varepsilon)_j =& - d((\gamma^{-1})_i^j(t, q_t^\varepsilon)(p_t^\varepsilon)_j) + (p_{t-}^\varepsilon)_j \partial_t(\gamma^{-1})_i^j(t, q_t^\varepsilon)dt \\ &+ (p_{t-}^\varepsilon)_j \partial_{q^l}(\gamma^{-1})_i^j(t, q_t^\varepsilon)\partial_{p_l} H^\varepsilon(t, x_t^\varepsilon)dt,\end{aligned}$$

where $\partial_{q^l}(\gamma^{-1})_i^j$ means the l-th component of $\nabla_q(\gamma^{-1})_i^j$, and $\partial_{p_l} H$ means the l-th component of $\nabla_q H$. Here we used Einstein summation notation. Therefore,

$$\begin{aligned}d(q_t^\varepsilon)_i =&(\gamma^{-1})_i^j(t, q_t^\varepsilon)(\partial_{q_j} H^\varepsilon(t, x_t^\varepsilon) - F_j(t, x_t^\varepsilon))dt + (\gamma^{-1})_i^j(t, q_t^\varepsilon)\sigma_j^\rho(t, x_{t-}^\varepsilon)d(L_t)_\rho \\ &- d((\gamma^{-1})_i^j(t, q_t^\varepsilon)(p_t^\varepsilon)_j) + (p_{t-}^\varepsilon)_j \partial_t(\gamma^{-1})_i^j(t, q_t^\varepsilon)dt + (p_{t-}^\varepsilon)_j \partial_{q^l}(\gamma^{-1})_i^j(t, q_t^\varepsilon)\partial_{p_l} H^\varepsilon(t, x_t^\varepsilon)dt.\end{aligned} \tag{3.18}$$

To simplify the last term $(p_t^\varepsilon)_j \partial_{p_l} H^\varepsilon(t, x_t^\varepsilon)dt$, we compute

$$\begin{aligned}&d((p_t^\varepsilon)_i(p_t^\varepsilon)_j) = (p_{t-}^\varepsilon)_i d(p_t^\varepsilon)_j + (p_{t-}^\varepsilon)_j d(p_t^\varepsilon)_i + d[p_i^\varepsilon, p_j^\varepsilon]_t \\ =&(p_{t-}^\varepsilon)_i \left[(-\gamma_j^k(t, p_t^\varepsilon)\partial_{p_k} H^\varepsilon(t, x_t^\varepsilon) - \partial_{q_j} H^\varepsilon(t, x_t^\varepsilon) + F_j(t, x_t^\varepsilon))dt + \sigma_j^\rho(t, x_{t-}^\varepsilon)d(L_t)_\rho\right] \\ &+(p_{t-}^\varepsilon)_j \left[(-\gamma_i^k(t, p_t^\varepsilon)\partial_{p_k} H^\varepsilon(t, x_t^\varepsilon) - \partial_{q_i} H^\varepsilon(t, x_t^\varepsilon) + F_i(t, x_t^\varepsilon))dt + \sigma_i^\rho(t, x_{t-}^\varepsilon)d(L_t)_\rho\right] \\ &+ \int_{\mathbb{R}^d\setminus\{0\}} \sigma_i^k(t, x_{t-}^\varepsilon)\sigma_j^l(t, x_{t-}^\varepsilon)x_k x_l N(dt, dx).\end{aligned} \tag{3.19}$$

Rewrite this equation in the form of the following Lyapunov equation [27]

$$\gamma_j^k(V_t)_{ki} + \gamma_i^k(V_t)_{kj} = (C_t)_{ij}, \tag{3.20}$$

where $(V_t)_{ij} = \partial_{p_i} H^\varepsilon(t, x_t^\varepsilon)(p_{t-}^\varepsilon)_j dt$, and

$$\begin{aligned}(C_t)_{ij} =& - d((p_t^\varepsilon)_i(p_t^\varepsilon)_j) + (p_{t-}^\varepsilon)_i \left[-\partial_{q_j} H^\varepsilon(t, x_t^\varepsilon) + F_j(t, x_t^\varepsilon)\right] dt + (p_{t-}^\varepsilon)_j \left[-\partial_{q_i} H^\varepsilon(t, x_t^\varepsilon) + F_i(t, x_t^\varepsilon)\right] dt \\ &+(p_{t-}^\varepsilon)_i\sigma_j^\rho(t, x_{t-}^\varepsilon)d(L_t)_\rho + (p_{t-}^\varepsilon)_j\sigma_i^\rho(t, x_{t-}^\varepsilon)d(L_t)_\rho + \int_{\mathbb{R}^d\setminus\{0\}} \sigma_i^k(t, x_{t-}^\varepsilon)\sigma_j^l(t, x_{t-}^\varepsilon)x_k x_l N(dt, dx).\end{aligned}$$

By solving Lyapunov equation (3.20), we have

$$(V_t)_{ij} = \int_0^\infty e^{-y\gamma_i^k}(C_t)_{kl}e^{-y\gamma_j^l}dy.$$

Hence, we have

$$
\begin{aligned}
&(p^\varepsilon_{t-})_j\partial_{p_l}H^\varepsilon(t,x^\varepsilon_t)dt = G^{ab}_{jl}(t,q^\varepsilon_t)(C_t)_{ab}\\
&= G^{ab}_{jl}(t,q^\varepsilon_t)\left[-d((p^\varepsilon_t)_a(p^\varepsilon_t)_b) + (p^\varepsilon_{t-})_a(-\partial_{p_b}H^\varepsilon(t,x^\varepsilon_t) + F_b(t,x^\varepsilon_t))dt\right.\\
&+(p^\varepsilon_{t-})_b(-\partial_{p_a}H^\varepsilon(t,x^\varepsilon_t) + F_a(t,x^\varepsilon_t))dt + (p^\varepsilon_{t-})_a\sigma^\rho_b(t,x^\varepsilon_{t-})d(L_t)_\rho + (p^\varepsilon_{t-})_b\sigma^\rho_a(t,x^\varepsilon_{t-})d(L_t)_\rho\\
&+\int_{\mathbb{R}^d\setminus\{0\}}\sigma^k_a(t,x^\varepsilon_{t-})\sigma^l_b(t,x^\varepsilon_{t-})x_kx_lN(dt,dx)],
\end{aligned}
\tag{3.21}
$$

where $G^{ab}_{jl}(t,q^\varepsilon_t) = \int_0^\infty e^{-y\gamma^a_j(t,q^\varepsilon_t)}e^{-y\gamma^b_l(t,q^\varepsilon_t)}dy$.

Combining Eq.(3.18) and Eq.(3.21) together, we see that q^ε_t satisfies the equation

$$
\begin{aligned}
d(q^\varepsilon_t)_i &= (\gamma^{-1})^j_i(t,q^\varepsilon_t)\left(\partial_{q_j}V(t,q^\varepsilon_t) + F_j(t,x^\varepsilon_t)\right)dt + (\gamma^{-1})^j_i(t,q^\varepsilon_t)\sigma^\rho_j(t,x^\varepsilon_{t-})d(L_t)_\rho\\
&\quad + (\gamma^{-1})^j_i(t,q^\varepsilon_t)\partial_{q_j}K^\varepsilon(t,x^\varepsilon_t)dt - \partial_{q^h}(\gamma^{-1})^j_i(t,q^\varepsilon_t)G^{ab}_{jh}(t,q^\varepsilon_t)\int_{\mathbb{R}^d\setminus\{0\}}\sigma^k_a(t,x^\varepsilon_{t-})\sigma^l_b(t,x^\varepsilon_{t-})x_kx_lN(dt,dx)\\
&\quad + d(R^\varepsilon_t)_i,
\end{aligned}
\tag{3.22}
$$

where

$$
\begin{aligned}
d(R^\varepsilon_t)_i &= d((\gamma^{-1})^j_i(t,q^\varepsilon_t)(p^\varepsilon_t)_j) - (p^\varepsilon_t)_j\partial_t(\gamma^{-1})^j_i(t,q^\varepsilon_t)dt\\
&\quad - \partial_{q^h}(\gamma^{-1})^j_i(t,q^\varepsilon_t)G^{ab}_{jh}(t,q^\varepsilon_t)\left[-d((p^\varepsilon_t)_a(p^\varepsilon_t)_b) + (p^\varepsilon_{t-})_a(-\partial_{p_b}H^\varepsilon(t,x^\varepsilon_t) + F_b(t,x^\varepsilon_t))dt\right.\\
&\quad \left. + (p^\varepsilon_{t-})_b(-\partial_{p_a}H^\varepsilon(t,x^\varepsilon_t) + F_a(t,x^\varepsilon_t))dt + (p^\varepsilon_{t-})_a\sigma^\rho_b(t,x^\varepsilon_{t-})d(L_t)_\rho + (p^\varepsilon_{t-})_b\sigma^\rho_a(t,x^\varepsilon_{t-})d(L_t)_\rho\right].
\end{aligned}
\tag{3.23}
$$

Note that term $(\gamma^{-1})^j_i(t,q^\varepsilon_t)\partial_{q_j}K^\varepsilon(t,x^\varepsilon_t)dt$ in (3.22) will survive in the limiting equation. Here we make another assumption.

Assumption 6 Every element $\partial_{q_j}K$ in ∇_qK is Lipschitz w.r.t q.

Remark 3.3. *This assumption seems a little strong. However, it is reasonable since we assume function K is $\mathcal{C}^2$, hence K is locally Lipschitz. Indeed we will extend our results to locally Lipshitz K in Section 3.4. If K is independent of q, then this term can be ignored. If K does not have additional assumption, we refer to [13] for estimations of this term.*

The proceeding calculations motivate the proposed lower dimensional limiting equation for the dynamics of position q:

$$
\begin{aligned}
d(q_t)_i &= (\gamma^{-1})^j_i(t,q_t)\left(\partial_{q_j}V(t,q_t) + F_j(t,x_t)\right)dt + (\gamma^{-1})^j_i(t,q_t)\sigma^\rho_j(t,x_{t-})d(L_t)_\rho\\
&\quad + (\gamma^{-1})^j_i(t,q_t)\partial_{q_j}K(t,x_t)dt - \partial_{q^h}(\gamma^{-1})^j_i(t,q_t)G^{ab}_{jh}(t,q_t)\int_{\mathbb{R}^d\setminus\{0\}}\sigma^k_a(t,x_{t-})\sigma^l_b(t,x_{t-})x_kx_lN(dt,dx),
\end{aligned}
\tag{3.24}
$$

where $x_t = (q_t, 0)$ since momentum p_t^ε converges to 0 from Proposition 3.1. Here we denote

$$S_i(t,x) = \int_0^t \partial_{q^h}(\gamma^{-1})_i^j(t,q) G_{jh}^{ab}(s,q) \int_{\mathbb{R}^d \backslash \{0\}} \sigma_a^k(s,x_{s-}) \sigma_b^l(s,x_{s-}) z_k z_l N(ds,dz). \tag{3.25}$$

Actually it is the noise-induced drift in limiting equation.

3.3. ***Proof of convergence to the limiting equation***

In this subsection, we show that the stochastic Hamiltonian system (2.3) converge to homogenized equation (3.24) in moment under an additional assumption:

Assumption 7. Assume that function γ is $\mathcal{C}^2$ and $\partial_t\gamma$, $\partial_{q^i}\gamma$, $\partial_t\partial q^i\gamma$ and $\partial_{q^i}\partial_{q^j}\gamma$ are bounded on $[0,T]\times\mathbb{R}^n$, for every T.

Now we demonstrate that the remainder term R_t^ε converges to zero. For convenience, we denote $\tilde{C}$ a finite positive constant whose value may vary from line to line and the notation $\tilde{C}(\cdot)$ to emphasize the dependence on the quantities appearing in the parentheses.

Lemma 3.4. *Under Assumption 1-7, for every $T>0, \eta>1$ and $\theta<\eta$, we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] = O(\varepsilon^\beta),\ \text{as } \varepsilon\to 0, \tag{3.26}$$

where R_t^ε was defined in Eq. (3.23) *and $\beta(\theta)$ is a piecewise function*

$$\beta(\theta) = \begin{cases} \frac{\theta}{2}\left(1-\frac{1}{\eta}\right), & 0<\theta\le\frac{2\eta}{\eta+1}, \\ 1-\frac{\theta}{\eta}, & \theta>\frac{2\eta}{\eta+1}. \end{cases}$$

Proof. Integrating Eq. (3.23) on $[0,T]$, then taking expectation and supremum on it, we have

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}||R_t^\varepsilon||^\theta\right] \le 8^{\theta-1}\left(\mathbb{E}\left[\sup_{t\in[0,T]}||(\gamma^{-1})_i^j(t,q_t^\varepsilon)(p_t^\varepsilon)_j||^\theta\right] + \mathbb{E}\left[\sup_{t\in[0,T]}||(\gamma^{-1})_i^j(0,q_0^\varepsilon)(p_0^\varepsilon)_j||^\theta\right]\right.\\
&+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t (p_s^\varepsilon)_j\partial_s(\gamma^{-1})_i^j(s,q_s^\varepsilon)ds\right|\right|^\theta\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t \partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)d((p_s^\varepsilon)_a(p_s^\varepsilon)_b)\right|\right|^\theta\right]\\
&+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t \partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_s^\varepsilon)_a\left(\partial_{q_b}K^\varepsilon(s,x_s^\varepsilon)+\partial_{q_b}V(s,q_s^\varepsilon)+F_b(s,x_s^\varepsilon)ds\right)\right|\right|^\theta\right]\\
&+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t \partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_s^\varepsilon)_b\left(\partial_{q_a}K^\varepsilon(s,x_s^\varepsilon)+\partial_{q_a}V(s,q_s^\varepsilon)+F_a(s,x_s^\varepsilon)ds\right)\right|\right|^\theta\right]\\
&+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t \partial_{q^h}(\gamma^{-1})_i^j(t,q_t^\varepsilon)G_{jh}^{ab}(t,q_t^\varepsilon)(p_{t-}^\varepsilon)_a\sigma_b^\rho(t,x_t^\varepsilon)d(L_t)_\rho\right|\right|^\theta\right]\\
&+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t \partial_{q^h}(\gamma^{-1})_i^j(t,q_t^\varepsilon)G_{jh}^{ab}(t,q_t^\varepsilon)(p_{t-}^\varepsilon)_b\sigma_a^\rho(t,x_t^\varepsilon)d(L_t)_\rho\right|\right|^\theta\right]\\
&:=\sum_{i=1}^8 J_i.
\end{aligned}
$$

We will now give upper bounds of terms $\{J_i\}_{i=1}^8$ for $\theta\ge 1$. For the first two terms,

$$
J_1+J_2\le 2||\gamma^{-1}||_\infty^\theta\mathbb{E}\left[\sup_{t\in[0,T]}||p_t^\varepsilon||^\theta\right].
$$

For the third term, we have

$$
J_3\le T^{\theta-1}||\partial_t\gamma^{-1}||_\infty^\theta\mathbb{E}\left[\int_0^T||p_s^\varepsilon||^\theta ds\right]\le T^\theta||\partial_t\gamma^{-1}||_\infty^\theta\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta\right].
$$

Note by Assumption 7 we can deduce that the function $\partial_q(\gamma^{-1})(t,q)G(t,q)$ is bounded and $\mathcal{C}^1$. Hence we have the following estimation (see Appendix)

$$
J_4\le\tilde{C}(\theta,T,M_1,C,\gamma)\left(\mathbb{E}[\sup_{t\in[0,T]}||p_t^\varepsilon||^{2\theta}]+\varepsilon^{-\frac{\theta}{2}}\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^{2\theta}\right]+\varepsilon^{-\frac{\theta}{2}}\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^{2\theta}K^\varepsilon(t,x_t^\varepsilon)^\theta\right]\right). \tag{3.27}
$$

Applying Hölder inequality and Assumption 2-3 we have

$$
\begin{aligned}
J_5&\le T^{\theta-1}\mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t||p_s^\varepsilon||^\theta\left(||\nabla_qK^\varepsilon(s,x_s^\varepsilon)||^\theta+||\nabla_qV+F||_\infty^\theta\right)ds\right]\\
&\le T^\theta\left(\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta||K^\varepsilon(t,x_t^\varepsilon)||^\theta\right]+(M_1^\theta+||\nabla_qV+F||_\infty^\theta)\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta\right]\right).
\end{aligned}
$$

The estimation of J_6 is similar to J_5. For the last two term (see Appendix), we have

$$J_7 \leq \tilde{C}(\theta, T, \nu) \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right]. \tag{3.28}$$

The estimation of J_8 is similar to J_7 as well. Substitute all these upper bound together, we obtain

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] \leq \tilde{C}\Bigg(&\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^\theta\right] + \mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^{2\theta}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right] + \varepsilon^{-\frac{\theta}{2}} \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^{2\theta}\right] \\
&+ \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta K^\varepsilon(t, x_t^\varepsilon)^\theta\right] + \varepsilon^{-\frac{\theta}{2}} \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^{2\theta} K^\varepsilon(t, x_t^\varepsilon)^\theta\right]\Bigg) \\
\leq \tilde{C}\Bigg(&\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^\theta\right] + \mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^\varepsilon||^{2\theta}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[||p_t^\varepsilon||^\theta\right]\Bigg) \\
+ \tilde{C}\varepsilon^{\frac{\theta}{2}}\Bigg(&\sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^{\theta+\frac{\theta}{\eta}}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^{\frac{2\theta}{\eta}}\right] + \sup_{t\in[0,T]} \mathbb{E}\left[K^\varepsilon(t, x_t^\varepsilon)^{\theta+\frac{2\theta}{\eta}}\right]\Bigg).
\end{aligned}$$

The last inequality follows from the similar arguments in proposition 3.1. Now we only need to compare order of ε in these terms. By means of Lemma 3.2 and Proposition 3.1, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}(1-\frac{1}{\eta})}) + O(\varepsilon^{1-\frac{\theta}{\eta}}). \tag{3.29}$$

Thus if $\theta > 2-\frac{2}{\eta+1}$, then $\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] = O(\varepsilon^{1-\frac{\theta}{\eta}})$. If $1 \leq \theta \leq 2-\frac{2}{\eta+1}$ then $\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}(1-\frac{1}{\eta})})$. As for the case $\theta < 1$, Hölder inequality implies that $\mathbb{E}\left[\sup_{t\in[0,T]} ||R_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}(1-\frac{1}{\eta})})$. □

Thus we can show that the stochastic Hamiltonian system (2.3) uniformly converges to the homogenized equation (3.24) in moment as follows.

Theorem 3.1. (Convergence to the limiting equation in moment) *Suppose Assumption 1-7 holds. Let x_t^ε be the solution of SDE* (2.3) *with initial condition* $(p_0^\varepsilon, q_0^\varepsilon)$ *and q_t be the solution of SDE* (3.24)*with initial condition q_0. Also suppose that for every $\varepsilon > 0, \eta > 1$, the initial condition satisfies integrable conditions $\mathbb{E}[||q_0^\varepsilon||^\theta] < \infty, \mathbb{E}[||q_0||^\theta] < \infty$ and $\mathbb{E}[||q_0^\varepsilon - q_0||^\theta] = O(\varepsilon^\beta)$. Then for every $T > 0$ and $\theta < \eta$, we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||q_t^\varepsilon - q_t||^\theta\right] = O(\varepsilon^\beta) \text{ as } \varepsilon \to 0. \tag{3.30}$$

Proof. First let $\theta \geq 2$. Define a vector $\widetilde{F}(t,x)$ and a matrix $\widetilde{\sigma}(t,x)$ as follows respectively

$$\widetilde{F}_i(t,x) = (\gamma^{-1})_i^j(t,q)(\partial_{p_j} K(t,x) + \partial_{p_j} V(t,q) + F_j(t,x)),$$

$$\widetilde{\sigma}_i^\rho(t,x) = (\gamma^{-1})_i^j(t,q)\sigma_j^\rho(t,x).$$

Hence we can rewrite Eq.(3.22) as

$$(q_t^\varepsilon)_i = (q_0^\varepsilon)_i + \int_0^t \widetilde{F}_i(s,x_s^\varepsilon)ds + \int_0^t \widetilde{\sigma}_i^\rho(s,x_s^\varepsilon)d(L_s)_\rho + S_i(t,x_t^\varepsilon) + (R_t^\varepsilon)_i, \tag{3.31}$$

and Eq.(3.24) as

$$(q_t)_i = (q_0)_i + \int_0^t \widetilde{F}_i(s,x_s)ds + \int_0^t \widetilde{\sigma}_i^\rho(s,x_s)d(L_s)_\rho + S_i(t,x_t). \tag{3.32}$$

Therefore, we obtain the following estimation

$$\begin{aligned} &\mathbb{E}\left[\sup_{s\in[0,t]}||q_s^\varepsilon - q_s||^\theta\right] \\ &\leq \tilde{C}\mathbb{E}\left[\sup_{s\in[0,t]}\left(||q_0^\varepsilon - q_0||^\theta + \left|\left|\int_0^s \widetilde{F}_i(r,x_r^\varepsilon) - \widetilde{F}_i(r,x_r)dr\right|\right|^\theta + \left|\left|\int_0^s \widetilde{\sigma}_i^\rho(r,x_r^\varepsilon) - \sigma_i^\rho(r,x_r)d(L_r)_\rho\right|\right|^\theta\right.\right. \\ &+ \left.\left. ||S_i(s,x_s^\varepsilon) - S_i(s,x_s)||^\theta + ||R_s^\varepsilon||^\theta\right)\right]. \end{aligned} \tag{3.33}$$

By the Lipschitz property of $\widetilde{F}$ and $\widetilde{\sigma}$ due to Assumptions, we have

$$\begin{aligned} \mathbb{E}\left[\sup_{s\in[0,t]}\left|\left|\int_0^s \widetilde{F}_i(r,x_r^\varepsilon) - \widetilde{F}_i(r,x_r)dr\right|\right|^\theta\right] &\leq \mathbb{E}\left[\sup_{s\in[0,t]} s^{\theta-1}\int_0^s ||\widetilde{F}_i(r,x_s^\varepsilon) - \widetilde{F}_i(r,x_s)||^\theta ds\right] \\ &\leq T^{\theta-1}\mathbb{E}\left[\int_0^t \left|\left|F_i(r,x_r^\varepsilon) - \widetilde{F}_i(r,x_r)\right|\right|^\theta dr\right] \\ &\leq \tilde{C}\left(\int_0^t \mathbb{E}[\sup_{r\in[0,s]}||q_r^\varepsilon - q_r||^\theta]ds + \sup_{s\in[0,t]}\mathbb{E}[||p_s^\varepsilon||^\theta]\right), \end{aligned} \tag{3.34}$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]}\left|\left|\int_0^s \widetilde{\sigma}_i^{\rho}(r,x_r^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(r,x_r)d(L_r)_{\rho}\right|\right|^{\theta}\right]\\
&\leq \tilde{C}\mathbb{E}\left[\sup_{s\in[0,t]}\left(\left|\left|\int_0^s\int_{\mathbb{R}^d\setminus\{0\}}(\widetilde{\sigma}_i^{\rho}(r,x_r^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(r,x_r))x\widetilde{N}(dr,dx)\right|\right|^{\theta}+\left|\left|\int_0^s\int_{|x|>1}(\widetilde{\sigma}_i^{\rho}(r,x_r^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(r,x_r))x\nu(dx)dr\right|\right|^{\theta}\right)\right]\\
&\leq \tilde{C}\left(\mathbb{E}\left[\left(\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}||\widetilde{\sigma}_i^{\rho}(s,x_s^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(s,x_s)||^2|x|^2\nu(dx)ds\right)^{\frac{\theta}{2}}\right]\right.\\
&\left.+\mathbb{E}\left[\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}||\widetilde{\sigma}_i^{\rho}(s,x_s^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(s,x_s)||^{\theta}|x|^{\theta}\nu(dx)ds\right]+\mathbb{E}\left[\int_0^t\left|\left|\int_{|x|>1}x\nu(dx)(\widetilde{\sigma}_i^{\rho}(s,x_s^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(s,x_s))\right|\right|^{\theta}ds\right]\right)\\
&\leq \tilde{C}\mathbb{E}\left(\int_0^t||\widetilde{\sigma}_i^{\rho}(s,x_s^{\varepsilon})-\widetilde{\sigma}_i^{\rho}(s,x_s)||^{\theta}ds\right)\\
&\leq \tilde{C}\left(\int_0^t\mathbb{E}[\sup_{r\in[0,s]}||q_r^{\varepsilon}-q_r||^{\theta}]dr+\sup_{s\in[0,t]}E[||p_s^{\varepsilon}||^{\theta}]\right).
\end{aligned}
\tag{3.35}
$$

We can also get a similar bound for the noise-induced term

$$
\mathbb{E}\left[\sup_{s\in[0,t]}||S_i(s,x_s^{\varepsilon})-S_i(s,x_s)||^{\theta}\right]\leq \tilde{C}\left(\int_0^t\mathbb{E}[\sup_{r\in[0,s]}||q_r^{\varepsilon}-q_r||^{\theta}]dr+\sup_{s\in[0,t]}\mathbb{E}[||p_s^{\varepsilon}||^{\theta}]\right). \tag{3.36}
$$

Consequently, estimations (3.34)-(3.36) together with Proposition 3.1 and Lemma 3.4 yield that

$$
\mathbb{E}\left[\sup_{s\in[0,t]}||q_s^{\varepsilon}-q_s||^{\theta}\right]\leq \tilde{C}\int_0^t\mathbb{E}\left[\sup_{r\in[0,s]}||q_r^{\varepsilon}-q_r||^{\theta}\right]ds+O(\varepsilon^{\beta}), \tag{3.37}
$$

for all $t\in[0,T]$. If $\mathbb{E}\left[\sup_{s\in[0,t]}||q_s^{\varepsilon}-q_s||^{\theta}\right]\in L^1[0,T]$. Then Gronwall's inequality implies

$$
\mathbb{E}\left[\sup_{s\in[0,t]}||q_s^{\varepsilon}-q_s||^{\theta}\right]\leq O(\varepsilon^{\beta})e^{\tilde{C}t}, \tag{3.38}
$$

which is precisely the result we want to prove. Indeed,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}||q_t^{\varepsilon}||^{\theta}\right]\leq C&\left(\mathbb{E}\left[\sup_{t\in[0,T]}||q_0^{\varepsilon}||^{\theta}\right]+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t\widetilde{F}(s,x_s^{\varepsilon})ds\right|\right|^{\theta}\right]\right.\\
&\left.+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t\widetilde{\sigma}^{\rho}(s,x_s^{\varepsilon})d(L_s)_{\rho}\right|\right|^{\theta}\right]+\mathbb{E}\left[\sup_{t\in[0,T]}||S(t,x_t^{\varepsilon})||^{\theta}\right]+\mathbb{E}\left[\sup_{t\in[0,T]}||(R_t^{\varepsilon})||^{\theta}\right]\right)\\
&<\infty,
\end{aligned}
$$

and similarly we can get $\mathbb{E}\left[\sup_{t\in[0,T]}||q_t||^\theta\right] < \infty$.

$\square$

3.4. *Extension*

In this section, we relax some assumptions that we make before. Actually we can extend all Lipschitz conditions to locally Lipschitz condition and remove all boundedness conditions. Organize and summarize the assumptions in the previous article, now we give a complete theorem.

Theorem 3.2. (Convergence to the limit equation in probability) *Suppose the family of Hamiltonians have the form*

$$H^\varepsilon(t, q, p) = K^\varepsilon(t, q, p) + V(t, q) = K(\varepsilon, t, q, p/\sqrt{\varepsilon}) + V(t, q),$$

and the following conditions hold:
1. The function $K^\varepsilon(t, q, p)$ is non-negative and $\mathcal{C}^2$.
2. There exist constant $C > 0, M_1 > 0$ such that

$$\max\{|\partial_t K(\varepsilon, t, q, z)|, ||\nabla_q K(\varepsilon, t, q, z)||, ||\nabla_z K(\varepsilon, t, q, z)||\} \leq M_1 + CK(\varepsilon, t, q, z).$$

3. There exist constant $c > 0, M_2 \geq 0$ such that

$$||\nabla_z K(\varepsilon, t, q, z)||^2 + M_2 \geq cK(\varepsilon, t, q, z).$$

4. For every $T > 0$, there exist constant $c > 0, \eta > 1$ such that

$$K(\varepsilon, t, q, z) \geq c||z||^\eta.$$

5. The potential energy function $V(t, q)$ is $\mathcal{C}^1$.
6. The dissipative coefficient γ is $\mathcal{C}^2$, independent of p and symmetric with eigenvalues bounded below by a constant $\lambda > 0$.
7. The external force F and noise intensity coefficient σ are continuous and locally Lipschitz.
Let x_t^ε be the solution of SDE (2.3) *with initial condition $(p_0^\varepsilon, q_0^\varepsilon)$ and q_t be the solution of SDE* (3.24)*with initial condition q_0. Also suppose that for every $\varepsilon > 0$ and $\theta \in (0, \eta)$, the initial condition satisfies integrable conditions $\mathbb{E}[||q_0^\varepsilon||^\theta] < \infty, \mathbb{E}[||q_0||^\theta] < \infty$ and $\mathbb{E}[||q_0^\varepsilon - q_0||^\theta] = O(\varepsilon^\beta)$. Then for every $T > 0, \delta > 0$ we have*

$$\lim_{\varepsilon\to 0} \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right) = 0. \tag{3.39}$$

Proof. Let $\chi : \mathbb{R}^n \to [0,1]$ be a C^∞ function. Define

$$V_r(t,q) = \chi_r(q)V(t,q), F_r(t,x) = \chi_r(q)\chi_r(p)F(t,x), \sigma_r(t,x) = \chi_r(q)\chi_r(p)\sigma(t,x),$$
$$K(\varepsilon,t,q,z) = \chi_r(z)K(\varepsilon,t,q,z), \gamma_r(t,q) = \chi_r(q)\gamma(t,q) + (1-\chi_r(q))\lambda I$$

Replacing the function V, F, K, γ, σ in (2.3) by $V_r, F_r, K_r, \gamma_r, \sigma_r$, we arrive at an SDE satisfying the condition in Theorem 3.1. Let $x_t^{r,\varepsilon}$ be solution to the corresponding SDE. Similarly, let q_t^r be the solution to the corresponding limiting SDE (3.24). Proposition 3.1 and Theorem 3.1 imply that, for every $T > 0$, $\eta > 1$ and $\theta \in (0, \eta)$

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^{r,\varepsilon}||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}-\frac{\theta}{2\eta}}) \text{ as } \varepsilon \to 0, \tag{3.40}$$

and

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r||^\theta\right] = O(\varepsilon^\beta) \text{ as } \varepsilon \to 0. \tag{3.41}$$

We will use this result to prove that q_t^ε converges to q_t in probability.

Denfine stopping times $\tau_r^\varepsilon = \inf\{t : ||q_t^\varepsilon|| \geq r\}$, $\eta_r^\varepsilon = \inf\{t : ||p_t^\varepsilon|| \geq \varepsilon r\}$ and $\tau_r = \inf\{t : ||q_t|| \geq r\}$. The drifts and diffusions of the modified and unmodified SDEs agree on the ball $\{||q|| < r, ||p|| < \varepsilon r\}$. Hence

$$q^\varepsilon_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge t} = q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge t},\ q_{\tau_r\wedge t} = q^r_{\tau_r\wedge t} \text{ for all } t \geq 0 \text{ a.s.}$$

For every $T > 0, \delta > 0$, we deduce that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right)\\
=&\mathbb{P}\left(\tau_r\wedge\tau_r^\varepsilon\wedge\eta_r^\varepsilon > T,\ \sup_{t\in[0,T]} ||q^\varepsilon_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge t} - q_{\tau_r\wedge t}|| > \delta\right) + \mathbb{P}\left(\tau_r\wedge\tau_r^\varepsilon\wedge\eta_r^\varepsilon \leq T,\ \sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right)\\
=&\mathbb{P}\left(\tau_r\wedge\tau_r^\varepsilon\wedge\eta_r^\varepsilon > T,\ \sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > \delta\right) + \mathbb{P}\left(\tau_r\wedge\tau_r^\varepsilon\wedge\eta_r^\varepsilon \leq T,\ \sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right)\\
\leq&\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > \delta\right) + \mathbb{P}\left(\tau_r\wedge\tau_r^\varepsilon\wedge\eta_r^\varepsilon \leq T\right),
\end{aligned} \tag{3.42}$$

where the first term on the right hand side converges to 0 as $\varepsilon \to 0$ by (3.41). Then we focus on the second term,

$$
\begin{aligned}
&\mathbb{P}\left(\tau_r \wedge \tau_r^\varepsilon \wedge \eta_r^\varepsilon \le T\right) \\
=&\mathbb{P}(\tau_r \le T) + \mathbb{P}\left(\tau_r > T, \tau_r^\varepsilon \wedge \eta_r^\varepsilon \le T\right) \\
\le&\mathbb{P}(\tau_r \le T) + \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > 1\right) + \mathbb{P}\left(\tau_r > T, \tau_r^\varepsilon \wedge \eta_r^\varepsilon \le T,\ \sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| \le 1\right) \\
\le&\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^r|| > r\right) + \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > 1\right) + \mathbb{P}\left(\tau_r^\varepsilon \wedge \eta_r^\varepsilon \le T, ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge T} - q^r_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge T}|| \le 1\right) \\
\le&\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^r|| > r\right) + \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > 1\right) + \mathbb{P}\left(\eta_r^\varepsilon > T, \tau_r^\varepsilon \le T, ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge T} - q^r_{\tau_r^\varepsilon\wedge T}|| \le 1\right) \\
&+ \mathbb{P}\left(\eta_r^\varepsilon \le T, ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge T} - q_{\tau_r^\varepsilon\wedge\eta_r^\varepsilon\wedge T}|| \le 1\right).
\end{aligned}
\tag{3.43}
$$

Note that when $\tau_r^\varepsilon \le T$, we have $||q_{\tau_r^\varepsilon\wedge T}|| \ge r$. Hence by $||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge T} - q^r_{\tau_r^\varepsilon\wedge T}|| \le 1$, we can deduce

$$||q^r_{\tau_r^\varepsilon\wedge T}|| \ge ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge T}|| - ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge T} - q^r_{\tau_r^\varepsilon\wedge T}|| > r - 1.$$

This implies that

$$\mathbb{P}\left(\tau_r^\varepsilon \le T, ||q^{r,\varepsilon}_{\tau_r^\varepsilon\wedge T} - q^r_{\tau_r^\varepsilon\wedge T}|| \le 1\right) \le \mathbb{P}\left(||q^r_{\tau_r^\varepsilon\wedge T}|| > r - 1\right) \le \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^r|| > r - 1\right). \tag{3.44}$$

Combining (3.42),(3.43) and (3.44) together, we have

$$
\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right) \\
&\le \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > \delta\right) + \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^r|| > r\right) + \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^{r,\varepsilon} - q_t^r|| > 1\right) \\
&+ \mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^r|| > r - 1\right) + \mathbb{P}\left(\eta_r^\varepsilon \le T\right).
\end{aligned}
\tag{3.45}
$$

On the other hand, by Chebyshev inequality and (3.40), we have

$$\mathbb{P}\left(\eta_r^\varepsilon \le T\right) \le \mathbb{P}\left(\sup_{t\in[0,T]} ||p_t^{r,\varepsilon}|| > \varepsilon r\right) \le (\varepsilon r)^{-2}\mathbb{E}\left[\sup_{t\in[0,T]} ||p_t^{r,\varepsilon}||^2\right] = O(\varepsilon^{-1-\frac{1}{\eta}})r^{-2}. \tag{3.46}$$

Then if we let $r^{-1} = o(\varepsilon^{\frac{1}{2}(1+\frac{1}{\eta})})$, i.e., the speed of r goes to infinity faster than $\varepsilon^{-\frac{1}{2}(1+\frac{1}{\eta})}$. We have

$$\mathbb{P}\left(\sup_{t\in[0,T]} ||q_t^\varepsilon - q_t|| > \delta\right) \to 0 \text{ as } r \to \infty,\ \varepsilon \to 0 \tag{3.47}$$

by the non-explosion property of q_t^r. $\square$

4. An Example

In this section, we present a prototypical example with Hamiltonian $H(m,t,q,p) = \frac{p^2}{2m} + V(t,q)$, where m is the mass of a particle. In this case, the small mass limit is also called Smoluchowski-Kramers limit. We consider the stochastic Hamiltonian system with external force $F(t,x)$ and Lévy noise L_t

$$\begin{aligned} dq_t^m &= \frac{1}{m} p_t^m dt, \\ dp_t^m &= \left(\frac{1}{m}\gamma(t,q_t^m)p_t^m - \nabla_q V(t,q_t^m) + F(t,x_t^m) \right) dt + \sigma(t,x_t^m) dL_t. \end{aligned} \tag{4.1}$$

By Proposition 3.1, p_t^m converges to zero. Then the homogenized equation in the small mass limit is

$$dq_t = \gamma^{-1}(t,q_t)(\nabla_q V(t,q_t) + F(t,q_t,0))dt + \gamma^{-1}(t,q_t)\sigma(t,q_t,0)dL_t + S(t,q_t), \tag{4.2}$$

where the noise induced drift is

$$S_i(t,q_t) = \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \partial_{q^h}(\gamma^{-1})_i^j(t,q_t) \int_0^\infty \left(e^{-y\gamma(s,q_s)}\right)_j^a \left(e^{-y\gamma(s,q_s)}\right)_l^b dy \sigma_a^k(s,q_t,0)\sigma_b^l(s,q_t,0) z_k z_l N(ds,dz). \tag{4.3}$$

Moreover, when dissipative coefficient γ is independent of q, the noise-induced drift (4.3) vanish, and the homogenized equation becomes

$$dq_t = \gamma^{-1}(t)(\nabla_q V(t,q_t) + F(t,q_t,0))dt + \gamma^{-1}(t)\sigma(t,q_t,0)dL_t. \tag{4.4}$$

This result coincide with that in [23].

5. Conclusion and Discussion

In this paper, we derive the small mass limiting equation for a class of Hamiltonian systems with multiplicative Lévy noise. Some interesting results appear. If the Hamiltonian function $H(\varepsilon,q,p)$ possesses appropriate properties, then momentum p will always converge to zero in finite time under uniform norm. The noise-induced drift term induced by pure jump Lévy noise is a Poisson process, which is rather different from that induced by Gaussian noise [12]. Our results could be applied to a class of stochastic Hamiltonian systems, such as a small mass particle in force field with state-dependent friction and a particle on a Riemannian manifold.

However, we have to mention that the pure jump Lévy noises in this paper have finite moment. In other words, it has bounded jumps. Large jumps could lead to some unpredictable dynamics although interlacing techniques allow us to deal with it. Hence an interesting problem is that how to accurately deal with Lévy noise without finite moments such as α-stable Lévy noise, which will be studied in the future.

Acknowledgments

The authors would like to thank Lingyu Feng, Jianyu Hu, Pingyuan Wei, Shenglan Yuan and Yanjie Zhang for helpful discussions. This work was partly supported by NSFC grants 11771449 and 11531006.

Appendix

A. Non-explosion of solution

In Appendix, we will prove that the solution of SDE (2.3) and limit equation are existence and unique under Assumption 1-4.

Lemma A.1. *Under Assumption 1-4, there exists a unique non-explosive solution to* (2.3) *in finite time interval* $[0, T]$.

Proof. First, we can verify that SDE with Assumption 1-3 satisfies Lipschitz condition and one side growth condition (refer to [16]) in every bounded cylinder $I \times U(R)$, where $U(R)$ is a ball with radius R. Then, we will prove that there is no explosion. Let τ_n be the first exit time of x_t^ε from the ball $B(0, n)$. From the right-continuity of the process x_t^ε we infer that

$$|x_{\tau_n}^\varepsilon| \geq n. \tag{A.1}$$

Define a function $U^\varepsilon(t, x_t^\varepsilon) = ||q_t^\varepsilon||^{2\eta} + K^\varepsilon(t, x_t^\varepsilon)$. By Assumption 4, we obtain that

$$\begin{aligned} U^\varepsilon(\tau_n, x_{\tau_n}^\varepsilon) &= ||q_{\tau_n}^\varepsilon||^{2\eta} + K^\varepsilon(\tau_n, x_{\tau_n}^\varepsilon) \\ &\geq ||q_{\tau_n}^\varepsilon||^{2\eta} + c\varepsilon^{-\eta}||p_{\tau_n}^\varepsilon||^{2\eta} \\ &\geq \min\{1, c\varepsilon^{-\eta}\}||x_{\tau_n}^\varepsilon||^{2\eta} \\ &\geq c|n|^{2\eta}. \end{aligned} \tag{A.2}$$

On the other hand, we have

$$
\begin{aligned}
&\mathbb{E}\left[U^\varepsilon(t\wedge\tau_n\wedge T, x^\varepsilon_{t\wedge\tau_n\wedge T})\right] \\
&=\mathbb{E}\left[U^\varepsilon(t\wedge\tau_n\wedge T, x^\varepsilon_{t\wedge\tau_n\wedge T})1_{\{\tau_n\wedge T\geq t\}}\right]+\mathbb{E}\left[U^\varepsilon(t\wedge\tau_n\wedge T, x^\varepsilon_{t\wedge\tau_n\wedge T})1_{\{\tau_n\wedge T< t\}}\right] \\
&=\mathbb{E}\left[U^\varepsilon(t, x^\varepsilon_{t})1_{\{\tau_n\wedge T\geq t\}}\right]+\mathbb{E}\left[U^\varepsilon(\tau_n\wedge T, x^\varepsilon_{\tau_n\wedge T})1_{\{\tau_n\wedge T< t\}}\right] \\
&=\mathbb{E}\left[U^\varepsilon(t, x^\varepsilon_{t})1_{\{\tau_n\wedge T\geq t\}}\right]+\mathbb{E}\left[U^\varepsilon(\tau_n, x^\varepsilon_{\tau_n})1_{\{\tau_n< T\}}1_{\{\tau_n< t\}}\right]+\mathbb{E}\left[U^\varepsilon(T, x^\varepsilon_{T})1_{\{\tau_n\geq T\}}1_{\{T< t\}}\right] \\
&\geq\mathbb{E}\left[U^\varepsilon(\tau_n, x^\varepsilon_{\tau_n})1_{\{\tau_n< t\}}\right].
\end{aligned}
\tag{A.3}
$$

Therefore, for all $n\in\mathbb{N}$

$$
\mathbb{P}(\tau_n<t)\leq c^{-1}n^{-2\eta}\mathbb{E}\left[U^\varepsilon(t\wedge\tau_n\wedge T, x^\varepsilon_{t\wedge\tau_n\wedge T})\right]. \tag{A.4}
$$

Notice that by Theorem 3.3 we have

$$
\mathbb{E}\left[U^\varepsilon(t\wedge\tau_n\wedge T, x^\varepsilon_{t\wedge\tau_n\wedge T})\right]\leq\mathbb{E}\left[\sup_{t\in[0,T]}||q^\varepsilon_t||^{2\eta}\right]+\mathbb{E}\left[\sup_{t\in[0,T]}K^\varepsilon(t,x^\varepsilon_t)\right]=O(1). \tag{A.5}
$$

Hence,

$$
\lim_{n\to\infty}\mathbb{P}(\tau_n<t)=0 \text{ for all } t. \tag{A.6}
$$

That is the desired assertion, as required. □

B. Proofs of (3.27) and (3.28)

We give calculations for estimations of (3.27) and (3.28) in remainder term.

Proof of (3.27)**.** By Assumption 7 we can deduce that the function $\partial_q(\gamma^{-1})(t,q)G(t,q)$ is bounded and $\mathcal{C}^1$. Let $f(t,q)=\partial_q(\gamma^{-1})(t,q)G(t,q)$. We have

$$
\begin{aligned}
J_4&=\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t\partial_{q^h}(\gamma^{-1})^j_i(s,q^\varepsilon_s)G^{ab}_{jh}(s,q^\varepsilon_s)d((p^\varepsilon_s)_a(p^\varepsilon_s)_b)\right|\right|^\theta\right] \\
&\leq\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t f(s,q^\varepsilon_s)d((p^\varepsilon_s)_i(p^\varepsilon_s)_j)\right|^\theta\right].
\end{aligned}
\tag{B.1}
$$

Since $f(s,q^\varepsilon_s)$ is a C^1-semimartingale, using integration by parts formula we obtain

$$
\begin{aligned}
\int_0^t f(s,q^\varepsilon_s)d((p^\varepsilon_s)_i(p^\varepsilon_s)_j)&=f(t,q^\varepsilon_t)(p^\varepsilon_t)_i(p^\varepsilon_t)_j-f(0,q^\varepsilon_0)(p^\varepsilon_0)_i(p^\varepsilon_0)_j \\
&\quad-\int_0^t(p^\varepsilon_s)_i(p^\varepsilon_s)_j\left(\partial_s f(s,q^\varepsilon_s)+\nabla_q f(s,q^\varepsilon_s)\nabla_p H^\varepsilon(s,x^\varepsilon_s)\right)ds.
\end{aligned}
\tag{B.2}
$$

Hence, for $\theta \geq 1$, we have

$$
\begin{aligned}
J_4 &\leq 3^{\theta-1}\left(2||f||_\infty^\theta \mathbb{E}[\sup_{t\in[0,T]}||p_t^\varepsilon||^{2\theta}] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t ||p_s^\varepsilon||^2\left(||\partial_s f||_\infty + ||\nabla_q f||_\infty |\nabla_p K^\varepsilon(s,x_s^\varepsilon)|\right)ds\right|^\theta\right]\right)\\
&\leq 3^{\theta-1}\left(2||f||_\infty^\theta \mathbb{E}[\sup_{t\in[0,T]}||p_t^\varepsilon||^{2\theta}] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t ||p_s^\varepsilon||^2\left(||\partial_s f||_\infty + ||\nabla_q f||_\infty \frac{1}{\sqrt{\varepsilon}}(M_1 + CK^\varepsilon(s,x_s^\varepsilon))\right)ds\right|^\theta\right]\right)\\
&\leq 3^{\theta-1}2||f||_\infty^\theta \mathbb{E}[\sup_{t\in[0,T]}||p_t^\varepsilon||^{2\theta}] + 6^{\theta-1}T^{\theta-1}\mathbb{E}\left[\int_0^T ||p_s^\varepsilon||^{2\theta}\left(||\partial_s f||_\infty^\theta + M_1^\theta||\nabla_q f||_\infty^\theta \varepsilon^{-\frac{\theta}{2}} + C^\theta K^\varepsilon(s,x_s^\varepsilon)^\theta\varepsilon^{-\frac{\theta}{2}}\right)ds\right].
\end{aligned}
\tag{B.3}
$$

□

Proof of (3.28)**.** Applying Kunita's first inequality [16] on J_7, we have

$$
\begin{aligned}
J_7 &= 2^{\theta-1}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\left|\int_0^t\int_{\mathbb{R}^d\setminus\{0\}}\partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_{s-}^\varepsilon)_a\sigma_b^\rho(s,x_s^\varepsilon)x\widetilde{N}(ds,dx)\right|\right|^\theta\right.\\
&\quad\left.+\sup_{t\in[0,T]}\left|\left|\int_0^t\int_{|x|>1}\partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_{s-}^\varepsilon)_a\sigma_b^\rho(s,x_s^\varepsilon)x\nu(dx)ds\right|\right|^\theta\right]\\
&\leq 2^{\theta-1}D(\theta)\mathbb{E}\left[\left(\int_0^T\int_{\mathbb{R}^d\setminus\{0\}}\left|\left|\partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_{s-}^\varepsilon)_a\sigma_b^\rho(s,x_s^\varepsilon)x\right|\right|^2\nu(dx)ds\right)^{\frac{\theta}{2}}\right]\\
&\quad+2^{\theta-1}\mathbb{E}\left[\int_0^T\int_{\mathbb{R}^d\setminus\{0\}}\left|\left|\partial_{q^h}(\gamma^{-1})_i^j(s,q_s^\varepsilon)G_{jh}^{ab}(s,q_s^\varepsilon)(p_{s-}^\varepsilon)_a\sigma_b^\rho(s,x_s^\varepsilon)x\right|\right|^\theta\nu(dx)ds\right]\\
&\quad+2^{\theta-1}T^\theta C\left(\int_{|x|>1}|x|\nu(dx)\right)^\theta\mathbb{E}\left[\sup_{t\in[0,T]}||p_t^\varepsilon||^\theta\right]\\
&\leq 2^{\theta-1}\left(D(\theta)T^{\frac{\theta}{2}}C\int_{\mathbb{R}\setminus\{0\}}|x|^2\nu(dx)^{\frac{\theta}{2}} + TC\int_{\mathbb{R}\setminus\{0\}}|x|^\theta\nu(dx) + T^\theta C\left(\int_{|x|>1}|x|\nu(dx)\right)^\theta\right)\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta\right].
\end{aligned}
\tag{B.4}
$$

We have to mention that Kunita's first inequality holds for $\theta \geq 2$. Actually $J_7 \leq \tilde{C}\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta\right]$ still holds for $\theta \in [1,2)$ since $\sup_{t\in[0,T]}\mathbb{E}\left[||p_t^\varepsilon||^\theta\right] = O(\varepsilon^{\frac{\theta}{2}})$ for $\theta \in (0,2\eta)$. □

References

References

[1] M. Smoluchowski, Drei vortrage uber diffusion, brownsche bewegung und koagulation von kolloidteilchen, Zeitschrift fur Physik 17 (1916) 557–585.

[2] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica 7 (4) (1940) 284–304.

[3] E. Nelson, Dynamical Theories of Brownian Motion, Vol. 106, Princeton university press, 2020.

[4] C. Doering, Modeling Complex Systems: Stochastic Processes, Stochastic Differential Equations, and Fokker-Planck Equations, in: 1990 Lectures In Complex Systems, Addison-Wesley, 1990, pp. 3–51.

[5] M. Freidlin, Some remarks on the Smoluchowski–Kramers approximation, Journal of Statistical Physics 117 (3) (2004) 617–634.

[6] S. Cerrai, M. Freidlin, On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom, Probability theory and related fields 135 (3) (2006) 363–394.

[7] G. Pavliotis, A. Stuart, Multiscale Methods: Averaging and Homogenization, Springer Science & Business Media, 2008.

[8] P. Hanggi, Nonlinear fluctuations: the problem of deterministic limit and reconstruction of stochastic dynamics, Physical Review A 25 (2) (1982) 1130.

[9] G. Volpe, L. Helden, T. Brettschneider, J. Wehr, C. Bechinger, Influence of noise on force measurements, Physical Review Letters 104 (17) (2010) 170602.

[10] S. Hottovy, A. McDaniel, G. Volpe, J. Wehr, The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction, Communications in Mathematical Physics 336 (3) (2015) 1259–1283.

[11] J. Birrell, S. Hottovy, G. Volpe, J. Wehr, Small mass limit of a langevin equation on a manifold, in: Annales Henri Poincaré, Vol. 18, Springer, 2017, pp. 707–755.

[12] J. Birrell, J. Wehr, Homogenization of dissipative, noisy, Hamiltonian dynamics, Stochastic Processes and their Applications 128 (7) (2018) 2367–2403.

[13] J. Birrell, J. Wehr, A homogenization theorem for langevin systems with an application to hamiltonian dynamics, in: Sojourns in Probability Theory and Statistical Physics-I, Springer, 2019, pp. 89–122.

[14] S. H. Lim, J. Wehr, M. Lewenstein, Homogenization for generalized langevin equations with applications to anomalous diffusion, in: Annales Henri Poincaré, Springer, 2020, pp. 1–59.

[15] J. Duan, An Introduction to Stochastic Dynamics, Vol. 51, Cambridge University Press, 2015.

[16] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge university press, 2009.

[17] R. Sánchez, D. E. Newman, J.-N. Leboeuf, V. Decyk, B. A. Carreras, Nature of transport across sheared zonal flows in electrostatic ion-temperature-gradient gyrokinetic plasma turbulence, Physical Review Letters 101 (20) (2008) 205002.

[18] B. Dybiec, A. Kleczkowski, C. A. Gilligan, Modelling control of epidemics spreading by long-range interactions, Journal of the Royal Society Interface 6 (39) (2009) 941–950.

[19] Y. Xu, Y. Li, H. Zhang, X. Li, J. Kurths, The switch in a genetic toggle system with Lévy noise, Scientific reports 6 (1) (2016) 1–11.

[20] B. Dybiec, J. M. Parrondo, E. Gudowska-Nowak, Fluctuation-dissipation relations under Lévy noises, EPL (Europhysics Letters) 98 (5) (2012) 50006.

[21] Q. Zhang, J. Duan, Linear response theory for nonlinear stochastic differential equations with α-stable Lévy noises, Journal of Statistical Physics 182 (2) (2021) 1–28.

[22] H. Al-Talibi, A. Hilbert, V. Kolokoltsov, Nelson-type limit for a particular class of Lévy processes, in: AIP Conference Proceedings, Vol. 1232, American Institute of Physics, 2010, pp. 189–193.

[23] S. Zhang, Smoluchowski-Kramers approximation for stochastic equations with Lévy-noise, Ph.D. thesis, Purdue University (2008).

[24] V. I. Arnol'd, Mathematical Methods of Classical Mechanics, Vol. 60, Springer Science & Business Media, 2013.

[25] P. Wei, Y. Chao, J. Duan, Hamiltonian systems with Lévy noise: Symplecticity, Hamilton's principle and averaging principle, Physica D: Nonlinear Phenomena 398 (2019) 69–83.

[26] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, Stochastic processes and their applications 91 (2) (2001) 205–238.

[27] J. M. Ortega, Matrix Theory: A Second Course, Springer Science & Business Media, 2013.