Homogenization of Dissipative Hamiltonian Systems Under Lévy Fluctuations

Zibo Wang · Li Lv · Jinqiao Duan

Received: 9 May 2021 / Accepted: 10 November 2022 / Published online: 23 November 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
We study the small mass limit for a class of Hamiltonian systems with multiplicative non-Gaussian Lévy noise. Derivation of the limiting equation depends on the structure of the stochastic Hamiltonian systems, in which a discontinuous noise-induced drift term arises. Firstly, we show that the momentum in the stochastic Hamiltonian system converges to zero when the kinetic energy has polynomial growth. Then, we prove that the stochastic Hamiltonian system with classical kinetic energy converges to the limiting equation in probability, with respect to Skorokhod topology as the mass tends to zero.

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)

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

Communicated by Eliot Fried.

---

Li Lv
lilyu@hust.edu.cn
Zibo Wang
zibowang@hust.edu.cn
Jinqiao Duan
duan@iit.edu

1 School of Mathematics and Statistics and Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan 430074, China
2 Department of Applied Mathematics and Department of Physics, Illinois Institute of Technology, Chicago, IL 60616, USA
where $\gamma$ is the dissipation coefficient, $\sigma$ is the diffusion coefficient and $W$ is a Wiener process. The small mass limit problem was studied by Smoluchowski (1916) and Kramers (1940) when the mass $m \to 0$. Following their pioneering work, this subject has been investigated by a number of authors. For example, Nelson (2020) derived the limiting equation when $\gamma$ and $\sigma$ are constants and a Fokker–Planck equation approach was provided by Doering (1990). Convergence in probability for $\gamma$ constant and $\sigma$ position-dependent was shown by Freidlin (2004). For the infinite dimensional case, the problem was studied by Cerrai and Freidlin (2006). These above problems can be illustrated in the framework of homogenization, for which a splendid relevant reference is given (Pavliotis and Stuart 2008).

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 (1982) for systems satisfying the fluctuation-dissipation relation. Then, Volpe et al. (2010) made an experimental observation for this phenomenon. Hottovy et al. (2015) derived the limiting equation of SDEs with arbitrary state-dependent friction. Birrell et al. developed small mass limit theory on compact Riemannian manifolds (Birrell et al. 2017) and for Hamiltonian systems (Birrell and Wehr 2018). A generalized homogenization theorem for Langevin systems was proved in Birrell and Wehr (2019). Lim et al. (2020) 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 (Duan 2015). The particle undergoing Lévy superdiffusion is performing motion with random jumps and step lengths following a power-law distribution (Applebaum 2009). As an important kind of non-Gaussian noise, Lévy noise has been found widely in atmospheric turbulence (Sánchez et al. 2008), epidemic spreading (Dybiec et al. 2009) and cell biological behavior (Xu et al. 2016). 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 (Dybiec et al. 2012; Zhang and Duan 2021).

It is similar to the previous part that there are also some small mass limit results for SDEs driven by Lévy noise. For example, Al-Talibi et al. (2010) developed Nelson theory for the $\alpha$-stable Lévy process. Zhang (2008) obtained Smoluchowski–Kramers approximation for SDEs driven by Lévy noise whose moment is finite. However, the dissipation drift coefficient in their model is a constant independent of the velocity and position, which accounts to a limiting equation where the noise-induced drift term does not exist.

Hamiltonian dynamics (Arnol’d 2013), 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 (Wei et al. 2019; Wu 2001; Yuan and Blömker 2022).
In this present paper, we consider the following stochastic differential equation with multiplicative Lévy noise:

\[
\begin{align*}
\frac{dq}{\varepsilon} &= \frac{1}{\varepsilon} p \, dt, \\
\frac{dp}{\varepsilon} &= \left( -1 \frac{\gamma(t, q)}{\varepsilon} p + F(t, x) \right) \, dt + \sigma(t, x) \, dL_t. 
\end{align*}
\]  

(1.1)

We derive the small mass limiting equation for (1.1) in the framework of a class of dissipative Hamiltonian systems

\[
\begin{align*}
\frac{dq}{\varepsilon} &= \nabla_p H(t, x) \, dt, \\
\frac{dp}{\varepsilon} &= ( -\gamma(t, x) \nabla_p H(t, x) - \nabla_q H(t, x) ) \, dt + \sigma(t, x) \, dL_t.
\end{align*}
\]  

(1.2)

where \( x = (q, p) \) with position \( q \) and momentum \( p \). The function \( H \) is a Hamiltonian function with small mass parameter \( \varepsilon \). The functions \( \gamma, \sigma \) and \( F \) are dissipation coefficient, diffusion coefficient and external force, which depend on \( (q, p) \), 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 and Wehr (2018), and the main novelty of our work is that we combine the features of Hamiltonian structure and non-Gaussian Lévy noise. We reduce the original Hamiltonian system to a stochastic differential equation which only depends on the position variable \( q \) by means of a Lyapunov equation and the Lévy–Itô formula. We formally derive the limiting equation with a noise-induced drift term (see Sect. 4 for details), which is discontinuous in time \( t \) compared with the continuous noise-induced drift term in Birrell and Wehr (2018). Nevertheless, the non-Gaussian Lévy noise brings some difficulties in the derivation of limiting equation. The methods in Birrell and Wehr (2018) cannot be directly extended to the case of Lévy noise, since Lévy noise has random jump sizes at random times and some moments of Lévy noise may not exist.

Based on these observations, we obtain the main results in the present paper in two folds. First, under the assumption that the kinetic energy \( K \) has a polynomial growth, we give some moment estimates of the kinetic energy \( K \) and the momentum \( p \), in the order of the small mass parameters \( \varepsilon \). We prove that the momentum \( p \) converges to zero, while the kinetic energy \( K \) diverges as the small parameter \( \varepsilon \) goes to zero (see Propositions 3.1, 3.2 and 3.3 in Sect. 3 for details).

Second, we show that (1.1) converges to the following limiting equation

\[
\frac{dq}{t} = \gamma^{-1}(t, q) F(t, q, 0) \, dt + \gamma^{-1}(t, q) \sigma(t, q, 0) \, dL_t + S(t, q),
\]  

(1.3)

in probability with respect to Skorokhod topology (i.e., the kinetic energy \( K \) is in form of the classical kinetic energy \( K = \frac{|p|^2}{2\varepsilon} \)), where \( S \) is the noise-induced drift with form

\[
S_t \left( d_\varepsilon, q \right) = d_q^j \left( \gamma^{-1} \right)_j (t, q) J_{\varepsilon k} (d_\varepsilon, q),
\]  

(1.4)
and \( J \) is the matrix solving the Lyapunov equation \( \gamma J + J \gamma^* = \int_{\mathbb{R}^d \setminus \{0\}} (\sigma x)(\sigma x)^* N(dt, dx) \). We show the convergence under two kinds of different conditions, respectively. First, we show the convergence under bounded assumptions. Then, we extend the convergence result to unbounded assumptions by means of the non-explosion property of the solution (see Theorems 5.1 and 5.2 in Sect. 5 for details). Actually, our conclusion holds for the kinetic energy \( K \) of general quadratic polynomial growth. In addition, we also clarify that the convergence to the limiting equation is not valid for the kinetic energy \( K \) of the growth order more than quadratic, although we have shown that the momentum \( p \) converges to zero (see Remark 5.1 for details).

This paper is organized as follows. In Sect. 2, we recall some basic notations and introduce a class of dissipative Hamiltonian systems with multiplicative Lévy noise. In Sect. 3, we derive the moment estimates of the kinetic energy \( K \) and the momentum \( p \). In Sect. 4, we derive the limiting equation by using a Lyapunov equation and Lévy–Itô formula. In Sect. 5, we prove the Hamiltonian system with classical kinetic energy (1.1) converges to the limiting Eq. (1.3) in probability with respect to Skorokhod topology. More precisely, in Sect. 5.1, we obtain the convergence to the limiting equation under bounded assumptions. In Sect. 5.2, we extend the convergence result under some unbounded assumptions. Finally, we present an illustrative example in Sect. 6.

2 Preliminaries

2.1 Lévy Motion

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

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

\[
L_t = bt + B_Q(t) + \int_{|x|<1} x \tilde{N}(t, dx) + \int_{|x|\geq 1} x N(t, dx),
\]

where \( N(dt, dx) \) is the Poisson random measure on \( \mathbb{R} \times (\mathbb{R}^d \setminus \{0\}) \), \( \tilde{N}(dt, dx) \triangleq N(dt, dx) - \nu(dx)dt \) is the compensated Poisson random measure, \( |\cdot| \) denotes the \( l^2 \)-norm on \( \mathbb{R}^d \), \( \nu \triangleq \mathbb{E}N(1, \cdot) \) is the jump measure, and \( B_Q(t) \) is an independent \( n \)-dimensional Brownian motion with covariance matrix \( Q \). The triple \((b, Q, \nu)\) is called the generating triple for the Lévy process \( L_t \). A Lévy process \( L_t \) has \( \theta \)-th moment if and only if \( \int_{|x|\geq 1} |x|^\theta \nu(dx) < \infty \).

We recall some results regarding the Skorokhod space. Let \( \mathcal{D}([0, T], \mathbb{R}^d) \) be the space of càdlàg \( \mathbb{R}^d \)-valued functions on \([0, T]\). The following metric was defined by Skorokhod

\[
d(\varphi, \psi) = \inf_{\lambda \in \Lambda} \left\{ ||\varphi||^0 + \sup_{t \in [0, T]} |\varphi(t) - \psi(\lambda(t))| \right\}, \quad \varphi, \psi \in \mathcal{D}([0, T], \mathbb{R}^d),
\]
where Λ denotes the class of strictly increasing continuous functions of [0, T] onto itself and $|\lambda|_c := \sup_{s, t \in [0, T], s \neq t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|$. The space $\mathcal{D}([0, T], \mathbb{R}^d)$ is separable and complete under Skorokhod metric. By definition, Skorokhod metric $d(\varphi, \psi)$ is bounded above by uniform metric $\sup_t |\varphi - \psi|$.

2.2 Dissipative Hamiltonian System with Lévy Noise

We consider the dissipative Hamiltonian system described in Birrell and Wehr (2018). Given a time-dependent Hamiltonian function $H(t, x_t)$, where $x_t = (q_t, p_t) \in \mathbb{R}^d \times \mathbb{R}^d$. The following Hamiltonian system describes a system with dissipative force and an external force.

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

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

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

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

$$
\begin{align*}
\text{d}q^\varepsilon_t &= \nabla_p H^\varepsilon(t, x_t^\varepsilon)\text{d}t, \\
\text{d}p^\varepsilon_t &= (-\gamma(t, x_t^\varepsilon)\nabla_p H^\varepsilon(t, x_t^\varepsilon) - \nabla_q H^\varepsilon(t, x_t^\varepsilon) + \tilde{F}(t, x_t^\varepsilon))\text{d}t + \sigma(t, x_{t-}^\varepsilon)\text{d}L_t,
\end{align*}
$$

with initial data $(q_0^\varepsilon, p_0^\varepsilon)$, where $\sigma : [0, \infty) \times \mathbb{R}^{2d} \to \mathbb{R}^{d \times d}$ is noise intensity function and $L = \{L_t\}_{t \geq 0}$ is a $\mathbb{R}^d$-valued pure jump Lévy process with triple $(0, 0, \nu)$. Since the Hamiltonian function we consider is in the form of (2.2), we can rewrite the stochastic Hamiltonian system (2.3) as following

$$
\begin{align*}
\text{d}q^\varepsilon_t &= \nabla_p K^\varepsilon(t, x_t^\varepsilon)\text{d}t, \\
\text{d}p^\varepsilon_t &= (-\gamma(t, x_t^\varepsilon)\nabla_p K^\varepsilon(t, x_t^\varepsilon) - \nabla_q K^\varepsilon(t, x_t^\varepsilon) + F(t, x_t^\varepsilon))\text{d}t + \sigma(t, x_{t-}^\varepsilon)\text{d}L_t,
\end{align*}
$$

where we let a function $F$ denote $\tilde{F} - \nabla_q V$ for readability reasons.

Remark 2.1 We consider only a 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 Birrell and Wehr (2018). Thereby, we use same notations as in Birrell and Wehr (2018) to make sure the influence of Brownian motion can be added to our results.

3 Moment Estimates of the Kinetic Energy $K$ and the Momentum $p$

3.1 Assumptions on the Kinetic Energy

In this section, we derive the moment estimation for kinetic energy $K$ and some relevant estimation results. For the kinetic energy $K$, we make the following assumption.

**Assumption 1** The kinetic energy $K(\varepsilon, t, q, z)$ is non-negative and $C^2$ in $(t, q, z)$ for each $\varepsilon$, where $z = p/\sqrt{\varepsilon}$. Moreover, there exists a constant $C_0 > 0$ such that $K(0, x_0^\varepsilon) \leq C_0$. For every fixed constant $T > 0$ and $\varepsilon_0 > 0$, the following conditions hold on $(0, \varepsilon_0] \times [0, T] \times \mathbb{R}^d$:

(i) There exist positive constants $C_1, M_1$ such that

$$\max \{|\partial_t K(\varepsilon, t, q, z)|, |\nabla_q K(\varepsilon, t, q, z)|\} \leq M_1 + C_1 K(\varepsilon, t, q, z).$$

(ii) There exist positive constants $c, M_2$ such that

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

(iii) The kinetic energy $K(\varepsilon, t, q, z)$ has polynomial growth w.r.t (with respect to) $z$, i.e., there exist positive constants $c_2, C_2$ and $\eta \geq 2$ such that

$$c_2 |z|^\eta \leq K(\varepsilon, t, q, z) \leq C_2 (|z|^\eta + 1).$$

(iv) There exist positive constants $C_3$ and $C_4$ such that

$$|\nabla_K K(\varepsilon, t, q, z)| \leq C_3 (|z|^{\eta-1} + 1), \quad |\nabla_z\nabla_z K(\varepsilon, t, q, z)| \leq C_4 (|z|^{\eta-2} + 1).$$

The two inequalities in (iv) come from a natural discovery that derivatives of $K$ with polynomial growth have lower growth order than $K$. We give some examples in the following remark.

**Remark 3.1** We can check that the classical kinetic energy $K^\varepsilon(p) = \frac{|p|^2}{2\varepsilon}$ satisfies the above assumptions. In addition, some other general ‘kinetic energies’, such as the kinetic energy function $K^\varepsilon(p) = \frac{|p|^2}{2\varepsilon} + \ln(1 + \frac{|p|^2}{\varepsilon})$ and kinetic energy in polynomial form $K^\varepsilon(p) = \sum_{l=k_1}^{k_2} (|p|^2/2\varepsilon)^l$, in Birrell and Wehr (2018), satisfy these assumptions as well. In particular, the kinetic energy of the form $K^\varepsilon(p, q) = \frac{1}{2\varepsilon} g^{ij}(t, q) p_i p_j$ with bounded Riemannian metric tensor $g^{ij}(t, q)$, which describes a particle on Riemannian manifold, also satisfies Assumption 1.
Lemma 3.1 Under Assumption 1, for every $\delta > 0$, there exists a positive constant $C$ such that,

\[ K(\varepsilon, t, q, z + z_1) \theta^{-1} |\nabla_z K(\varepsilon, t, q, z + z_1)| \]
\[ \leq C(K(\varepsilon, t, q, z)^\theta + |z_1|^{\eta\theta - 1} + 1), \]  
(3.1)

\[ K(\varepsilon, t, q, z + z_1) \theta^{-2} |\nabla_z^2 K(\varepsilon, t, q, z + z_1)|^2 \]
\[ \leq C(\delta K(\varepsilon, t, q, z)^\theta + |z_1|^{\eta\theta - 2} + 1 + \delta^{1-\eta\theta}), \] 
(3.2)

\[ K(\varepsilon, t, q, z + z_1) \theta^{-1} ||\nabla_z \nabla_z K(\varepsilon, t, q, z + z_1)|| \]
\[ \leq C(\delta K(\varepsilon, t, q, z)^\theta + |z_1|^{\eta\theta - 2} + 1 + \delta^{1-\eta\theta/2}). \]  
(3.3)

Proof By Assumption 1 and a direct calculation, we obtain that

\[ K(\varepsilon, t, q, z + z_1) \theta^{-1} |\nabla_z K(\varepsilon, t, q, z + z_1)| \]
\[ \leq C_3^{\theta-1} C_4 (1 + |z + z_1|^{\eta})^{\theta-1} (1 + |z + z_1|^{\eta - 1}) \]
\[ \leq C_3^{\theta-1} C_4 2^{\theta-2} (1 + |z + z_1|^{\eta(\theta-1)}) (1 + |z + z_1|^{\eta - 1}) \]
\[ \leq C_3^{\theta-1} C_4 2^{\theta-2} (3 + |z + z_1|^{\eta\theta - 1}) \]
\[ \leq C_3^{\theta-1} C_4 2^{\theta-2} (3 + 2^{\eta\theta - 2}(|z|^{\eta\theta - 1} + |z_1|^{\eta\theta - 1})) \]
\[ \leq C_3^{\theta-1} C_4 2^{\theta-2} (3 + 2^{\eta\theta - 2} (1 + |z|^{\eta\theta} + |z_1|^{\eta\theta - 1})) \]
\[ \leq C(K(\varepsilon, t, q, z)^\theta + |z_1|^{\eta\theta - 1} + 1), \]

where we use Young inequality $|z|^{\eta\theta - 1} \leq |z|^{\eta\theta} + 1$ in the fifth inequality. Applying similar calculations, for every $\delta > 0$, we obtain

\[ K(\varepsilon, t, q, z + z_1) \theta^{-2} |\nabla_z^2 K(\varepsilon, t, q, z + z_1)|^2 \]
\[ \leq C_3^{\theta-1} C_4^2 2^{\theta-2} (3 + |z + z_1|^{\eta\theta - 2}) \]
\[ \leq C_3^{\theta-1} C_4^2 2^{\theta-2} (3 + 2^{\eta\theta - 3}(|z|^{\eta\theta - 2} + |z_1|^{\eta\theta - 2})) \]
\[ \leq C_3^{\theta-1} C_4^2 2^{\theta-2} (3 + 2^{\eta\theta - 3} (\delta^{1-\eta\theta} + \delta |z|^{\eta\theta} + |z_1|^{\eta\theta - 2})) \]
\[ \leq C(\delta K(\varepsilon, t, q, z)^\theta + |z_1|^{\eta\theta - 2} + 1 + \delta^{1-\eta\theta}), \]

where we use a generalized Young inequality $|z|^{\eta\theta - 1} \leq \delta^{1-\eta\theta} + \delta |z|^{\eta\theta}$. The same calculations hold for the inequality (3.3). \qed

For the dissipative matrix function $\gamma$, the external force $F$ and the noise intensity $\sigma$, we assume that

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

(i) The function $\gamma$, $F$, $\sigma$ are bounded and locally Lipschitz.

(ii) The matrix function $\gamma$ is symmetric with eigenvalues bounded below by some constant $\lambda > 0$. 

 Springer
We assume that the pure jump Lévy process has finite moment. More precisely, we make the following assumption for jump measure \( \nu \).

**Assumption 3** There exists a constant \( \theta_0 \geq 1 \) such that the Lévy measure \( \nu \) satisfies

\[
\int_{|x| \geq 1} |x|^{2\theta_0 \eta} \nu(dx) < \infty,
\]

here, the constant \( \eta \) is the same as in Assumption 1.

**Remark 3.2** Under Assumptions 1–3, the solution \( x_t^\varepsilon \) to stochastic Hamiltonian system (2.4) exists and is unique. See Appendix A for proof.

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

**Lemma 3.2** Under Assumptions 1–3, for every \( 1 \leq \theta \leq 2\theta_0 \) and \( T > 0 \), there exist positive constants \( \alpha_0, \varepsilon_0 \) such that for all constant \( \alpha \in (0, \alpha_0) \) and \( \varepsilon \in (0, \varepsilon_0) \) and \( t \in [0, T] \), we have

\[
E[K^\varepsilon(t, x_t^\varepsilon)^\theta] \leq \frac{\kappa_1(\varepsilon)}{\alpha} + E[e^{-\alpha t} \int_0^t \int_{|x| < 1} e^{\alpha s/\varepsilon} [K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon + \sigma(s, x_s^\varepsilon)x)^\theta - K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon)^\theta] N(ds, dx)],
\]

and

\[
K^\varepsilon(t, x_t^\varepsilon)^\theta \leq \kappa_2(\varepsilon) + \int_0^t \int_{|x| < 1} [K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon + \sigma(s, x_s^\varepsilon)x)^\theta - K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon)^\theta] N(ds, dx) + \int_0^t \int_{|x| \geq 1} [K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon + \sigma(s, x_s^\varepsilon)x)^\theta - K^\varepsilon(s, q_s^\varepsilon, p_s^\varepsilon)^\theta] N(ds, dx),
\]

(3.4)

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

**Proof** Note that the function \( e^{\alpha t/\varepsilon} K^\varepsilon(t, x)^\theta \) is \( C^1 \) on time \( t \), \( C^2 \) on \( x \), together with \( \sup_{0 \leq s \leq \tau} \sup_{0 < |x| < 1} |\sigma(s, x_s^\varepsilon)x| < \infty \). Hence, we can apply Itô theorem (Applebaum 2009, Page 251) to \( e^{\alpha t/\varepsilon} K^\varepsilon(t, x_t^\varepsilon)^\theta \), and with probability 1, we have
\[ e^{\alpha t/\epsilon} K^\epsilon(t, x^\epsilon_1) \theta = K^\epsilon(0, x^\epsilon_1) \theta + \frac{\alpha}{\epsilon} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x^\epsilon_s) \theta \, ds \]

\[ + \frac{\theta}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x^\epsilon_s) \theta - (\partial_t K^\epsilon)(s, x^\epsilon_s) \, ds \]

\[ + \frac{\theta}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x^\epsilon_s) \theta - \nabla_x K^\epsilon(s, x^\epsilon_s) \nabla_x K^\epsilon(s, x^\epsilon_s) \, ds \]

\[ + \frac{\theta}{\sqrt{\epsilon}} \int_0^t e^{\alpha s/\epsilon} K^\epsilon(s, x^\epsilon_s) \theta - (\nabla_x K^\epsilon)(s, x^\epsilon_s)(-\nabla_y K^\epsilon)(s, x^\epsilon_s) + F(s, x^\epsilon_s) \, ds \]

\[ + \int_0^t \int_{|x| < 1} e^{\alpha s/\epsilon} [K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) + \sigma(s, x^\epsilon_{s-}) x] \theta - K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) \theta ] \tilde{N}(ds, dx) \]

\( (I_1) \)

\[ + \int_0^t \int_{|x| \geq 1} e^{\alpha s/\epsilon} [K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) + \sigma(s, x^\epsilon_{s-}) x] \theta - K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) \theta ] N(ds, dx) \]

\( (I_2) \)

\[ + \int_0^t \int_{|x| < 1} e^{\alpha s/\epsilon} [K^\epsilon(s, q^\epsilon_{s}, p^\epsilon_{s}) + \sigma(s, x^\epsilon_{s}) x] \theta - K^\epsilon(s, q^\epsilon_{s}, p^\epsilon_{s}) \theta ] \]

\[ - \frac{\theta}{\sqrt{\epsilon}} K^\epsilon(s, q^\epsilon_{s}, p^\epsilon_{s}) \theta - (\nabla_x K^\epsilon)(s, x^\epsilon_s)(\sigma(s, x^\epsilon_{s}) x)] v(dx) \, ds, \]

\( (I_3) \)

where we denote the last three integrals by \( I_1, I_2, I_3 \) in the equality, respectively. The notation \((\partial_t K^\epsilon)(s, x)\) is equal to \(\partial_t K(\epsilon, s, q, p/\sqrt{\epsilon})\) and similarly for \((\nabla_x K^\epsilon)(s, x)\).

First, we estimate terms \( I_2, I_3 \). Using mean value theorem and the growth condition of \( K \), for the term \( I_2 \), there exists some \( \xi \in [0, 1] \) such that

\[ \mathbb{E}[I_2] = \mathbb{E} \left[ \int_0^t \int_{|x| \geq 1} e^{\alpha s/\epsilon} [K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) + \sigma(s, x^\epsilon_{s-}) x] \theta - K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) \theta ] v(dx) \, ds \right] \]

\[ = \mathbb{E} \left[ \frac{\theta}{\sqrt{\epsilon}} \int_0^t \int_{|x| \geq 1} e^{\alpha s/\epsilon} K^\epsilon(s, q^\epsilon_{s}, p^\epsilon_{s}) + \xi \sigma(s, x^\epsilon_{s}) x \theta - K^\epsilon(s, q^\epsilon_{s}, p^\epsilon_{s}) \theta ] v(dx) \, ds \right] \]

\[ \leq \mathbb{E} \left[ C \theta \int_0^t \int_{|x| \geq 1} e^{\alpha s/\epsilon} \left( K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) + \xi \theta \sigma(s, x^\epsilon_{s}) x \theta \right) \theta \, ds \right] \]

\[ \leq \mathbb{E} \left[ C \theta \int_0^t \int_{|x| \geq 1} e^{\alpha s/\epsilon} \left( K^\epsilon(s, q^\epsilon_{s-}, p^\epsilon_{s-}) + \xi \theta \sigma(s, x^\epsilon_{s}) x \theta \right) \theta \, ds \right] \]

\[ + C \theta \int_0^t \int_{|x| \geq 1} |x| v(dx) \, ds \]. \hspace{1cm} (3.6)
For the term $I_3$, using mean value theorem and (3.2)–(3.3), for every $\delta > 0$, there exists some $\xi \in [0, 1]$ such that

\[
I_3 \leq \frac{\theta(\theta - 1)}{2\varepsilon} \int_0^t \int_{|x| < 1} e^{\alpha s / \varepsilon} K^e(s, q^e_x, p^e_x) + \xi \sigma(s, x^e_x) x^{\theta - 2} |\nabla_z K^e(s, q^e_x, p^e_x) + \xi \sigma(s, x^e_x) x| \sigma(s, x^e_x) x^2 v(dx) ds + \frac{\theta}{2\varepsilon} \int_0^t \int_{|x| < 1} e^{\alpha s / \varepsilon} K^e(s, q^e_x, p^e_x) + \xi \sigma(s, x^e_x) x^{\theta - 1} |\nabla_z \nabla_z K^e(s, q^e_x, p^e_x) + \xi \sigma(s, x^e_x) x| \sigma(s, x^e_x) x^2 v(dx) ds
\]

\[
\leq \frac{C \theta^2}{2\varepsilon} \int_0^t \int_{|x| < 1} e^{\alpha s / \varepsilon} \delta K^e(s, q^e_x, p^e_x)^{\theta} + |\xi \sigma(s, x^e_x) x^{\theta - 2} + 1 + \delta^{1 - \eta \theta} + \delta^{1 - \eta \theta / 2})| \sigma(s, x^e_x) x^2 v(dx) ds
\]

\[
\leq \frac{C \theta^2 |\sigma|_\infty^2}{2\varepsilon} \int_{|x| < 1} |x|^2 v(dx) \int_0^t e^{\alpha s / \varepsilon} \delta K^e(s, q^e_x, p^e_x)^{\theta} ds + C \theta^2 \left( \frac{|\sigma|_\infty^{\eta \theta^2}}{2\varepsilon^{\eta \theta^2 / 2}} \int_{|x| < 1} |x|^2 v(dx) \right) \int_0^t e^{\alpha s / \varepsilon} ds.
\]

By Assumption 1 and Assumption 2, we conclude

\[
e^{\alpha_t / \varepsilon} K^e(t, x^e_t)^\theta
\]

\[
\leq K^e(0, x^0_t)^\theta + \left( \frac{\alpha}{\varepsilon} + C_1 \theta - \frac{\lambda c \theta}{\varepsilon} + \frac{C \theta}{\sqrt{\varepsilon}} |F|_\infty \right) \int_0^t e^{\alpha s / \varepsilon} K^e(s, x^e_s)^\theta ds + \theta \left( M_1 + \frac{\lambda M_2}{\varepsilon} + \frac{C}{\sqrt{\varepsilon}} |F|_\infty \right) \int_0^t e^{\alpha s / \varepsilon} K^e(s, x^e_s)^{\theta - 1} ds + I_1 + I_2 + I_3
\]

\[
\leq K^e(0, x^0_t)^\theta - \frac{D}{\varepsilon} \int_0^t e^{\alpha s / \varepsilon} K^e(s, x^e_s)^\theta ds + \frac{d}{\alpha} + I_1 + I_2 + I_3,
\]

where

\[
D = \lambda c \theta - \alpha - \theta C_1 \varepsilon - \theta C |F|_\infty (\sqrt{\varepsilon} + \delta) - M_1 \theta \delta \varepsilon - \lambda M_2 \theta \delta,
\]

and

\[
d = \left( \frac{1}{\delta} \right)^{\theta - 1} (M_1 \varepsilon + \lambda M_2 + C |F|_\infty \sqrt{\varepsilon}).
\]

Taking expectation on both side of (3.8) and combining inequality (3.6) with (3.7), we get
According to Assumptions 1–3, all the terms in $\delta, \alpha$ is non-negative; thus, this term can be omitted on the right hand side of above inequality. In addition, $K^\varepsilon(0, x_0^\varepsilon)$ is bounded by Assumption 1. Hence, we obtain (3.4).

Similarly, taking $\alpha = 0$ in (3.8) and combining (3.7), we have

$$K^\varepsilon(t, x_i^\varepsilon)^\theta \leq K^\varepsilon(0, x_0^\varepsilon)^\theta + \theta \left( C_1 - \frac{\lambda c}{\varepsilon} + \frac{C}{\sqrt{\varepsilon}} \|F\|_\infty \right) \int_0^t K^\varepsilon(s, x_s^\varepsilon)^\theta ds + \theta \left( M_1 + \frac{\lambda M_2}{\varepsilon} + \frac{C}{\sqrt{\varepsilon}} \|F\|_\infty \right) \int_0^t K^\varepsilon(s, x_s^\varepsilon)^{\theta-1} ds + I'_1 + I'_2$$

$$+ \frac{C\theta^2}{2\varepsilon} \|\sigma\|_\infty^2 \int_{|x| < 1} |x|^{2\theta} v(dx) \int_0^t K^\varepsilon(s, x_s^\varepsilon)^\theta ds$$

$$+ C\theta^2 \left[ \|\sigma\|_\infty^{\theta} \int_{|x| < 1} |x|^\theta v(dx) \right] \int_0^t K^\varepsilon(s, x_s^\varepsilon)^\theta ds$$

$$+ \frac{1 + \delta^{1-\eta/2} + \delta^{1-\eta/2}}{2\varepsilon} \|\sigma\|_\infty^2 \int_{|x| < 1} |x|^\theta v(dx) \int_0^t ds$$

$$\leq K^\varepsilon(0, x_0^\varepsilon)^\theta - \frac{D_2}{\varepsilon} \int_0^t K^\varepsilon(s, x_s^\varepsilon)^\theta ds + d_2 + I'_1 + I'_2,$$

where $I'_i$ equals $I_i$ with $\alpha = 0$, for $i = 1, 2$, 

$$D_2 = \lambda \varepsilon c - \theta C_1 \varepsilon - \theta C \|F\|_\infty (\sqrt{\varepsilon} + \delta) - M_1 \theta \varepsilon - \lambda M_2 \theta \delta - \frac{C\theta^2}{2} \|\sigma\|_\infty^2 \delta \int_{|x| < 1} |x|^2 v(dx).$$
and
\[ d_2 = d + \frac{C\theta^2}{2} T ||\sigma||_\infty^{-1} (1 + \delta^{1-n\theta} + \delta^{1-n\theta/2}) \]
\[ \int_{|x|<1} |x|^2 v(dx) + \frac{C\theta^2}{2} T ||\sigma||_\infty^{\eta\theta/2} \int_{|x|<1} |x|^{\eta\theta} v(dx). \]

By choosing \( \delta, \alpha \) sufficiently small, we conclude that \( D_2 \) is non-negative. Then, we have (3.5).

Then, we give the moment estimation of the kinetic energy \( K^\varepsilon(t, x^\varepsilon) \) by means of above assumptions and lemma. Now, we first show a property of the compensated Poisson integral in (3.4).

**Lemma 3.3** Under Assumptions 1–3, the compensated Poisson integral in (3.4)
\[ \int_0^t \int_{|x|<1} e^{ax/\varepsilon} \left[ K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s + \sigma(s, x^\varepsilon s) x)^\theta - K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s)^\theta \right] \tilde{N}(ds, dx) \]
is a martingale for \( 1 \leq \theta \leq \theta_0 \).

**Proof** Actually, this property is equivalent to following inequality (see Applebaum (2009), page 266).
\[ \int_0^t \int_{|x|<1} e^{2ax/\varepsilon} \mathbb{E}\left[ K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s + \sigma(s, x^\varepsilon s) x)^\theta - K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s)^\theta \right]^2 v(dx)ds < \infty. \]

To demonstrate this inequality, we define a stopping time \( \tau_R := \inf\{t \geq 0; K^\varepsilon(t, x^\varepsilon t) \geq R\} \); here, \( R \) is a positive number. It follows from the inequality (3.5) that
\[ K^\varepsilon(t \land \tau_R, x^\varepsilon t \land \tau_R)^\theta \]
\[ \leq 3 \left\{ \int_0^{t \land \tau_R} \int_{|x|<1} \left[ K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s + \sigma(s, x^\varepsilon s) x)^\theta - K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s)^\theta \right] \tilde{N}(ds, dx) \right\}^2 
+ 3 \left\{ \int_0^{t \land \tau_R} \int_{|x|<1} \left[ K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s + \sigma(s, x^\varepsilon s) x)^\theta - K^\varepsilon(s, q^\varepsilon s, p^\varepsilon s)^\theta \right] v(dx)ds \right\}^2 
+ 3 \left| K^\varepsilon - \delta/4 \right|^2. \]

Note that by the mean value theorem and (3.1), there exists some \( \xi \in [0, 1] \) such that
\[ |K^\varepsilon(t, q, p + \sigma x)^\theta - K^\varepsilon(t, q, p)^\theta| \]
\[ = \frac{1}{\sqrt{\varepsilon}} K^\varepsilon(t, q, p + \xi \sigma x)^{\theta-1} |\nabla_z K^\varepsilon(t, q, p + \xi \sigma x)||\sigma x| \]
\[ \leq \frac{C|\sigma x|}{\sqrt{\varepsilon}} \left( K^\varepsilon(t, q, p)^\theta + \left| \frac{\sigma x}{\sqrt{\varepsilon}} \right|^{\frac{\theta-1}{\theta}} + 1 \right). \]
Applying Kunita inequality to the compensated Poisson integral in (3.12), we have

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| \int_0^{t \wedge \tau_R} \int_{|x| \geq 1} \left[ K^\delta(s, q_s^\delta, p_s^\delta + \sigma(s, x_s^\delta)x)^{\frac{\theta}{2}} - K^\delta(s, q_s^\delta, p_s^\delta)^{\frac{\theta}{2}} \right] \nu(dx) ds \right|^2 \right] \\
\leq \tilde{C} \mathbb{E} \left[ \int_0^{T \wedge \tau_R} \int_{|x| \geq 1} \left| K^\delta(s, q_s^\delta, p_s^\delta + \sigma(s, x_s^\delta)x)^{\frac{\theta}{2}} - K^\delta(s, q_s^\delta, p_s^\delta)^{\frac{\theta}{2}} \right|^2 \nu(dx) ds \right] \\
\leq \tilde{C} \mathbb{E} \left[ \int_0^{T \wedge \tau_R} \int_{|x| \geq 1} \left( \epsilon^{-1/2} |\sigma x| + \epsilon^{-\eta/4} |\sigma x|^{\eta/2} \right)^2 \nu(dx) ds \right] \\
\leq \tilde{C} \mathbb{E} \left[ \int_{|x| \geq 1} |x|^2 \nu(dx) \int_0^{T \wedge \tau_R} K^\delta(s, x_s^\delta)^{\theta} ds + \epsilon^{-\eta/2} \int_{|x| \geq 1} |x|^\eta \nu(dx) \\
+ \epsilon^{-1/2} |\sigma x| K^\delta(s, q_s^\delta, p_s^\delta)^{\frac{\theta}{2}} \nu(dx) ds \right] ,
\]

(3.13)

For the second integral in (3.12), we have

\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| \int_0^{t \wedge \tau_R} \int_{|x| \geq 1} \left[ K^\delta(s, q_s^\delta, p_s^\delta + \sigma(s, x_s^\delta)x)^{\frac{\theta}{2}} - K^\delta(s, q_s^\delta, p_s^\delta)^{\frac{\theta}{2}} \right] \nu(dx) ds \right|^2 \right] \\
\leq \mathbb{E} \left[ \left( \int_{|x| \geq 1} |x| \nu(dx) \right) \int_0^{T \wedge \tau_R} K^\delta(s, x_s^\delta)^{\theta} ds \right]^2 \\
+ \epsilon^{-\eta/2} \left( \int_{|x| \geq 1} |x|^\eta \nu(dx) \right)^2 + \epsilon^{-1} \left( \int_{|x| \geq 1} |x| \nu(dx) \right)^2
\]

(3.14)

where \( \tilde{C} \) is finite constants dependent on \( \theta, T, \nu, ||\sigma||_\infty, \eta \) vary from line to line. To show the order of \( \epsilon \) in the estimation, here we write it explicitly in the estimation and denote constants associated with \( \epsilon \) by \( C_\epsilon \) below. Combining (3.12), (3.13) and (3.14) and applying Hölder inequality, we have
\[
K^\varepsilon(t \land \tau_R, x^\varepsilon_{t \land \tau_R})^\theta \leq \bar{C} e^{-1} E \left[ \int_0^{T \land \tau_R} K^\varepsilon(s, x^\varepsilon_s)^\theta \, ds \right] + \bar{C} e^{-\eta\theta/2} \leq \bar{C} \int_0^t E \left[ \sup_{t \in [0, T]} K^\varepsilon(s \land \tau_R, x^\varepsilon_{s \land \tau_R})^\theta \right] \, ds + \bar{C} e. \tag{3.15}
\]

Since \( E \left[ \sup_{s \in [0, t]} K^\varepsilon(s \land \tau_R, x^\varepsilon_{s \land \tau_R})^\theta \right] \leq R^\theta \) for all \( t \). Thus, the integral on above inequality is finite for all \( t \leq T \), and this combining with Grönwall inequality gives
\[
E \left[ \sup_{t \in [0, T]} K^\varepsilon(t \land \tau_R, x^\varepsilon_{t \land \tau_R})^\theta \right] \leq \bar{C} e^\bar{C} e T,
\]
then we let \( R \to +\infty \) (thus \( \tau_R \to +\infty \)); hence, the monotone convergence theorem yields
\[
E \left[ \sup_{t \in [0, T]} K^\varepsilon(t, x^\varepsilon_t)^\theta \right] \leq \bar{C} e^\bar{C} e T < +\infty. \tag{3.16}
\]
This gives the result for \( 2 \leq \theta \leq 2\theta_0 \), which holds for all \( 1 \leq \theta < 2 \) by applying Hölder inequality. Then, for \( 1 \leq \theta \leq \theta_0 \), we can prove the inequality (3.11) as follows. There exists some \( \xi \in [0, 1] \) such that
\[
\int_0^t \int_{|x| < 1} e^{2ax^2/\varepsilon} E[K^\varepsilon(s, q^\varepsilon_s + \sigma(s, x^\varepsilon_s)x)^\theta - K^\varepsilon(s, q^\varepsilon_s + \xi \sigma(s, x^\varepsilon_s)x)^\theta] v(dx) \, ds
\leq e^{2ax^2/\varepsilon} \int_0^t \int_{|x| < 1} \left[ \frac{\theta}{\sqrt{\varepsilon}} K^\varepsilon(s, q^\varepsilon_s + \xi \sigma(s, x^\varepsilon_s)x)^{\theta-1} \nabla_x K^\varepsilon(s, q^\varepsilon_s + \xi \sigma(s, x^\varepsilon_s)x) \right]^2 v(dx) \, ds
\leq e^{2ax^2/\varepsilon} \int_0^t \int_{|x| < 1} \left[ \frac{\theta^2 C^2}{\varepsilon} \left( K^\varepsilon(s, q^\varepsilon_s + \xi \sigma(s, x^\varepsilon_s)x)^\theta + \left| \frac{\xi \sigma(s, x^\varepsilon_s)x}{\sqrt{\varepsilon}} \right|^{\theta-1} + 1 \right)^2 \right] \left| \sigma(s, x^\varepsilon_s)x \right|^2 v(dx) \, ds
\leq 3\theta^2 C^2 e^{2ax^2/\varepsilon} \left( e^{-1} |\sigma|^2_\infty \int_{|x| < 1} |x|^2 v(dx) \int_0^t E \left[ K^\varepsilon(s, x^\varepsilon_s)^{2\theta} \right] \, ds
\right.
\left. + e^{-\theta} |\sigma|^2_\infty \int_{|x| < 1} |x|^{2\theta} v(dx) t
\right.
\left. + e^{-1} |\sigma|^2_\infty \int_{|x| < 1} |x|^2 v(dx) t \right)
\]
\[=: J_1 + J_2 + J_3, \tag{3.17}\]
where we utilized Lemma 3.1 in the second inequality. We notice that the terms \( J_2, J_3 \) are finite due to Assumption 2–3. As for the term \( J_1 \), it follows from the inequality (3.16) that
\[
J_1 \leq \bar{C} e \left[ \sup_{t \in [0, T]} K^\varepsilon(t, x^\varepsilon_t)^{2\theta} \right] < +\infty.
\]
Therefore, the inequality (3.11) holds, that is to say, we complete the proof of this lemma. \( \square \)
Next, we are going to present the moment estimates of the kinetic energy $K$ combining the Lemmas 3.2 and 3.3, and we obtain the following two results.

**Proposition 3.1** (Supremum of expectation of the kinetic energy) Under Assumptions 1–3, for every positive $T$ and $0 < \theta \leq \theta_0$, the kinetic energy $K$ has the following uniform estimate

$$
\sup_{t \in [0,T]} \mathbb{E}[K^\theta(t,x^\varepsilon_t)] = \begin{cases} 
O(\varepsilon^{1-\frac{\eta}{2}}), & \text{if } 1 \leq \theta \leq \theta_0, \\
O(\varepsilon^{(1-\frac{\theta}{2})}), & \text{if } 0 < \theta < 1,
\end{cases}
$$

(3.18)

**Proof** Since the compensated Poisson integral is a martingale by Lemma 3.3, thus we obtain the following equality

$$
e^{-at/\varepsilon} \mathbb{E}\left[\int_0^T \int_{|x|<1} e^{as/\varepsilon} [K^\theta(s,q^\varepsilon_{s-},p^\varepsilon_{s-} + \sigma(s,x^\varepsilon_{s-})x) \theta - K^\theta(s,q^\varepsilon_{s-},p^\varepsilon_{s-})^\theta] \tilde{N}(ds, dx)\right] = 0.
$$

It follows that the equality (3.18) holds from Lemma 3.2 and above equation for $1 \leq \theta \leq \theta_0$. The result for $0 < \theta < 1$ follows by Hölder’s inequality.

**Proposition 3.2** (Expectation of supremum of the kinetic energy) Under Assumptions 1–3, for every positive $T$ and $\theta$ with $0 < \theta \leq \theta_0$, the kinetic energy $K$ has the following uniform estimate

$$
\mathbb{E}\left[\sup_{t \in [0,T]} K^\theta(t,x^\varepsilon_t)\right] = O(\varepsilon^{-\eta/2}).
$$

(3.19)

**Proof** Using the same argument as (3.15) for $K^\theta(t,x^\varepsilon_t)$, we have the following estimate

$$
K^\theta(t,x^\varepsilon_t) \leq \tilde{C}\varepsilon^{-1} \mathbb{E}\left[\int_0^T K^\theta(s,x^\varepsilon_s) ds\right] + \tilde{C}\varepsilon^{-\eta/2}.
$$

Taking supremum and expectation on both side, we obtain that

$$
\mathbb{E}\left[\sup_{t \in [0,T]} K^\theta(t,x^\varepsilon_t)\right] \leq \tilde{C}\varepsilon^{-1} T \sup_{t \in [0,T]} \mathbb{E}\left[K^\theta(t,x^\varepsilon_t)\right] + \tilde{C}\varepsilon^{-\eta/2}.
$$

Thus, we obtain the desired result by (3.18).

With these moment estimates for the kinetic energy, we obtain following Proposition 3.3 which is a direct deduction from Propositions 3.1 and 3.2.

**Proposition 3.3** Under Assumptions 1–3, for every $T > 0$ we have

$$
\sup_{t \in [0,T]} \mathbb{E}[|p^\varepsilon_t|^\theta] = \begin{cases} 
O(\varepsilon^\theta), & \text{if } 0 < \theta < \eta, \\
O(\varepsilon), & \text{if } \eta \leq \theta \leq \eta\theta_0,
\end{cases}
$$

(3.20)
and for $0 < \theta \leq \eta \theta_0$,
\[
\mathbb{E} \left[ \sup_{t \in [0,T]} \left| p_t^\varepsilon \right|^\theta \right] = O(1). \tag{3.21}
\]

**Proof** From Assumption 1, we have
\[
\sup_{t \in [0,T]} \mathbb{E} \left[ \left| p_t^\varepsilon \right| \right] \leq c_2^{-\theta} \varepsilon^\theta \sup_{t \in [0,T]} \mathbb{E} \left[ K^\varepsilon (t, x_t^\varepsilon)^\theta \right].
\]

Hence, we get (3.20) combining above inequality with Proposition 3.1. By the same calculation and Proposition 3.2, we further obtain (3.21).

**Remark 3.3** As we state above, the term $\mathbb{E} \left[ \sup_{t \in [0,T]} \left| p_t^\varepsilon \right|^\theta \right]$ will not converge to zero. But in the case of Brownian noise, each moment of the momentum $p_t^\varepsilon$ will converge to zero.

### 4 Derivation of the Limiting Equation

In this section, we first derive the limiting equation of the system (2.4) as $\varepsilon \to 0$, then we give the convergence in expectation. Actually, we can only prove the convergence in a special case, which we will illustrate later. To this end, we make an additional assumption on $\gamma$.

**Assumption 4** Every element $\gamma_{ij}^\varepsilon$ in the matrix function $\gamma$ is $C^1$ in $(t, q)$ and independent of $p$.

Note that the stochastic Hamiltonian equation (2.4) can be simplified to
\[
d(q_t^\varepsilon) = \nabla_p K^\varepsilon (t, x_t^\varepsilon) dt
= \gamma^{-1}(t, x_t^\varepsilon) F(t, x_t^\varepsilon) dt + \gamma^{-1}(t, x_t^\varepsilon) \sigma(t, x_t^\varepsilon) dL_t - \gamma^{-1}(t, x_t^\varepsilon) d(p_t^\varepsilon). \tag{4.1}
\]

Since the matrix function $\gamma$ has bounded eigenvalues, $\gamma$ 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 (4.1), we have
\[
-(\gamma^{-1})_i^j (t, q_t^\varepsilon) d(p_t^\varepsilon)_j = - d((\gamma^{-1})_i^j (t, q_t^\varepsilon) (p_t^\varepsilon)_j) + (p_t^\varepsilon)_j \partial_t (\gamma^{-1})_i^j (t, q_t^\varepsilon) dt + (p_t^\varepsilon)_j \partial_q (\gamma^{-1})_i^j (t, q_t^\varepsilon) \partial_p K^\varepsilon (t, x_t^\varepsilon) dt,
\]
where $\partial_q (\gamma^{-1})_i^j$ means the $l$-th component of $\nabla_q (\gamma^{-1})_i^j$, and $\partial_p K$ means the $l$-th component of $\nabla_p K$. Here, we used Einstein summation notation. Therefore,
\[
d(q_t^\varepsilon)_i = (\gamma^{-1})_i^j (t, q_t^\varepsilon) F_j(t, x_t^\varepsilon) dt + (\gamma^{-1})_i^j (t, q_t^\varepsilon) \sigma_j^\theta (t, x_t^\varepsilon) d(L_t)_\rho
- d((\gamma^{-1})_i^j (t, q_t^\varepsilon) (p_t^\varepsilon)_j)
+ (p_t^\varepsilon)_j \partial_t (\gamma^{-1})_i^j (t, q_t^\varepsilon) dt + (p_t^\varepsilon)_j \partial_q (\gamma^{-1})_i^j (t, q_t^\varepsilon) \partial_p K^\varepsilon (t, x_t^\varepsilon) dt. \tag{4.2}
\]
Assumption 5 The function $K(\varepsilon, t, q, z)$ is independent of $q$ and has symmetry form as $K(\varepsilon, t, q, z) = \tilde{K}(\varepsilon, t, |p|^2/\varepsilon)$.

Under the Assumption 6, we note that $\partial_p K^p(t, p) = \frac{2}{\varepsilon} \tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)p_i$, here $\tilde{K}'$ denotes the partial derivative with respect to the last variable. Hence, (4.2) becomes

$$d(q^*_i) = (\gamma^{-1})_i^j(t, q^*_i)F_j(t, x_i^t)dt + (\gamma^{-1})_i^j(t, |q^*_i|_t)\sigma_i^p(t, x_i^t)d(L_t)_p$$

$$- d((\gamma^{-1})_i^j(t, q^*_i)(p^*_i)_j) + (p^*_i)_j \partial_t (\gamma^{-1})_i^j(t, q^*_i)dt$$

$$+ \frac{2}{\varepsilon} \partial_q^i (\gamma^{-1})_i^j(t, q^*_i)\tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)(p^*_i)_j dt. \quad (4.3)$$

To simplify the last term $\tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)(p^*_i)_j dt$, utilizing Itô’s product formula for $(p^*_i)_i, (p^*_i)_j$ (see Applebaum (2009), Theorem 4.4.13), we obtain

$$d((p^*_i)_i(p^*_i)_j) = (p^*_i)_i d(p^*_i)_j + (p^*_i)_j d(p^*_i)_i + d[p^*_i, p^*_j]_t$$

$$= -\gamma^j_i(t, q^*_i)^2 \tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)(p^*_i)_j dt$$

$$- \gamma^j_i(t, q^*_i)^2 \tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)(p^*_i)_i dt$$

$$+ (p^*_i)_j F_j(t, x_i^t)dt + (p^*_i)_i \sigma_i^p(t, x_i^t)d(L_t)_p$$

$$+ (p^*_i)_j F_i(t, x_i^t)dt + (p^*_i)_j \sigma_i^p(t, x_i^t)d(L_t)_p$$

$$+ \int_{\mathbb{R}^d \setminus \{0\}} \sigma_i^k(t, x_i^t)\sigma_j^l(t, x_i^t) x_k x_l N(dt, dx). \quad (4.4)$$

Rewrite this equation in the form of the following Lyapunov equation (Ortega 2013)

$$\gamma^j_i(B_t)_{ii} + \gamma^j_i(B_t)_{ij} = (Q_t)_{ij}, \quad (4.5)$$

where $(B_t)_{ii} = \frac{2}{\varepsilon} \tilde{K}'(\varepsilon, t, |p|^2/\varepsilon)(p^*_i)_i dt$, and

$$(Q_t)_{ij} = -d((p^*_i)_i(p^*_i)_j) + (p^*_i)_i F_j(t, x_i^t)dt + (p^*_i)_j F_i(t, x_i^t)dt$$

$$+ (p^*_i)_i \sigma_i^p(t, x_i^t)d(L_t)_p$$

$$+ (p^*_i)_j \sigma_i^p(t, x_i^t)d(L_t)_p + \int_{\mathbb{R}^d \setminus \{0\}} \sigma_i^k(t, x_i^t)\sigma_j^l(t, x_i^t) x_k x_l N(dt, dx).$$

By solving Lyapunov equation (4.5), we have

$$(B_t)_{ij} = \int_0^\infty (e^{-\gamma^j_i}) (Q_t)_{ab} (e^{-\gamma^j_i})_j dy.$$
Hence, we have

\[ \frac{2}{\varepsilon} \tilde{K}'(\varepsilon, t, |p|^{2}/\varepsilon)(p_{t}^{\varepsilon})_{k}(p_{t}^{\varepsilon})_{l} dt = G_{kl}^{ab}(t, q_{t}^{\varepsilon})(Q_{t})_{ab} \]

\[ = G_{kl}^{ab}(t, q_{t}^{\varepsilon})[-d((p_{t}^{\varepsilon})_{a}(p_{t}^{\varepsilon})_{b}) + (p_{t}^{\varepsilon})_{a} F_{b}(t, x_{t}^{\varepsilon}) dt + (p_{t}^{\varepsilon})_{b} F_{a}(t, x_{t}^{\varepsilon}) dt + (p_{t}^{\varepsilon})_{a} \sigma_{b}^{p}(t, x_{t}^{\varepsilon}) d(L_{t})_{\rho} + (p_{t}^{\varepsilon})_{b} \sigma_{a}^{p}(t, x_{t}^{\varepsilon}) d(L_{t})_{\rho} + \int_{\mathbb{R}^{d}\setminus\{0\}} \sigma_{a}^{k}(t, x_{t}^{\varepsilon}) \sigma_{b}^{l}(t, x_{t}^{\varepsilon}) x_{k} x_{l} N(dt, dx)]. \]

(4.6)

where \( G_{kl}^{ab}(t, q_{t}^{\varepsilon}) = \int_{0}^{\infty} (e^{-\gamma(t, q_{t}^{\varepsilon})})_{k}^{a} (e^{-\gamma(t, q_{t}^{\varepsilon})})_{l}^{b} dy. \)

Combining Eqs. (4.2) and (4.6) together, we see that \( q_{t}^{\varepsilon} \) satisfies the equation

\[ d(q_{t}^{\varepsilon})_{i} = (\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) F_{j}(t, x_{t}^{\varepsilon}) dt + (\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) \sigma_{j}^{p}(t, x_{t}^{\varepsilon}) d(L_{t})_{\rho} - \tilde{G}_{i}^{ab}(t, q_{t}^{\varepsilon}) d((p_{t}^{\varepsilon})_{a}(p_{t}^{\varepsilon})_{b}) + G_{i}^{ab}(t, q_{t}^{\varepsilon}) \int_{\mathbb{R}^{d}\setminus\{0\}} \sigma_{a}^{k}(t, x_{t}^{\varepsilon}) \sigma_{b}^{l}(t, x_{t}^{\varepsilon}) x_{k} x_{l} N(dt, dx) - d(R_{t}^{\varepsilon})_{i}, \]

(4.7)

where

\[ \tilde{G}_{i}^{ab}(t, q_{t}^{\varepsilon}) = \partial_{q}^{k}(\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) G_{kl}^{ab}(t, q_{t}^{\varepsilon}) \]

\[ = \partial_{q}^{k}(\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) \int_{0}^{\infty} (e^{-\gamma(t, q_{t}^{\varepsilon})})_{k}^{a} (e^{-\gamma(t, q_{t}^{\varepsilon})})_{l}^{b} dy, \]

(4.8)

and

\[ d(R_{t}^{\varepsilon})_{i} = d((\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) (p_{t}^{\varepsilon})_{j}) - (p_{t}^{\varepsilon})_{j} \partial_{t}(\gamma^{-1})_{i}^{j}(t, q_{t}^{\varepsilon}) dt \]

\[ - \tilde{G}_{i}^{ab}(t, q_{t}^{\varepsilon}) [(p_{t}^{\varepsilon})_{a} F_{b}(t, x_{t}^{\varepsilon}) dt + (p_{t}^{\varepsilon})_{b} F_{a}(t, x_{t}^{\varepsilon}) dt + (p_{t}^{\varepsilon})_{a} \sigma_{b}^{p}(t, x_{t}^{\varepsilon}) d(L_{t})_{\rho} + (p_{t}^{\varepsilon})_{b} \sigma_{a}^{p}(t, x_{t}^{\varepsilon}) d(L_{t})_{\rho}]. \]

(4.9)

The terms \( \int \tilde{G}_{i}^{ab}(t, q_{t}^{\varepsilon}) d((p_{t}^{\varepsilon})_{a}(p_{t}^{\varepsilon})_{b}) \) and \( R_{t}^{\varepsilon} \) are expected to go to zero; thus, the proceeding calculations motivate the proposed lower dimensional limiting equation for the dynamics of position \( q_{t} \) :

\[ d(q_{t})_{i} = (\gamma^{-1})_{i}^{j}(t, q_{t}) F_{j}(t, x_{t}) dt + (\gamma^{-1})_{i}^{j}(t, q_{t}) \sigma_{j}^{p}(t, x_{t}) d(L_{t})_{\rho} \]

\[ + \tilde{G}_{i}^{ab}(t, q_{t}^{\varepsilon}) \int_{\mathbb{R}^{d}\setminus\{0\}} \sigma_{a}^{k}(t, x_{t}^{\varepsilon}) \sigma_{b}^{l}(t, x_{t}^{\varepsilon}) x_{k} x_{l} N(dt, dx), \]

(4.10)

where \( x_{t} = (q_{t}, 0) \) since momentum \( p_{t}^{\varepsilon} \) converges to 0 from Proposition 3.3. Here, we denote

\[ S_{i}(t, x_{t}) = \int_{0}^{\gamma} \tilde{G}_{i}^{ab}(s, q_{t}^{\varepsilon}) \int_{\mathbb{R}^{d}\setminus\{0\}} \sigma_{a}^{k}(s, x_{s}^{\varepsilon}) \sigma_{b}^{l}(s, x_{s}^{\varepsilon}) x_{k} x_{l} N(ds, dz). \]

(4.11)
Actually, it is the noise-induced drift in limiting equation, which is discontinuous.

5 Convergence to the Limiting Equation with Classical Kinetic Energy

\[ K(\varepsilon, t, q, p) = \frac{|p|^2}{2\varepsilon} \]

In this section, we prove the convergence to the limiting equation in the case of classical kinetic energy. In fact, we demonstrate convergence in probability under two kinds of different conditions, which we will illustrate in the following subsections, respectively.

We consider classical kinetic energy

\[ K(\varepsilon, t, q, p) = \frac{|p|^2}{2\varepsilon}, \]

where \( \varepsilon \) is the mass of a particle. In this case, the small mass limit is also called Smoluchowski–Kramers limit.

Then, the stochastic Hamiltonian system with external force \( F(t, x) \) and Lévy noise \( L_t \) is as follows

\[
\begin{align*}
\frac{dq^\varepsilon_t}{\varepsilon} &= \frac{1}{\varepsilon} p^\varepsilon_t dt, \\
\frac{dp^\varepsilon_t}{\varepsilon} &= \left( -\frac{1}{\varepsilon} \gamma(t, q^\varepsilon_t) p^\varepsilon_t + F(t, x^\varepsilon_t) \right) dt + \sigma(t, x^\varepsilon_t) dL_t. 
\end{align*}
\]

Then, the limiting equation in the small mass limit is

\[
\begin{align*}
dq_t &= \gamma^{-1}(t, q_t) F(t, q_t, 0) dt + \gamma^{-1}(t, q_{t-}) \sigma(t, q_{t-}) dL_t + S(dt, q_t), 
\end{align*}
\]

according to the Eq. (4.10), where the noise induced drift is

\[
S_t(t, q_t) = \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \partial_q h(\gamma^{-1})_i (s, q_s) \\
\int_0^\infty \left( e^{-\gamma y(s, q_s)} \right)^a \left( e^{-\gamma y(s, q_s)} \right)^b dy \sigma^k_a(s, q_s, 0) \sigma^j_b(s, q_s, 0) z_k z_l N(ds, dz).
\]

5.1 Convergence Under Bounded Assumptions

Our proof is based on a weak limit theorem for stochastic differential equations as follows

**Lemma 5.1** (Kurtz and Protter 1991, Theorem 5.4) Let \( U^\varepsilon \) and \( Y^\varepsilon \) be processes with sample paths in \( D([0, \infty), \mathbb{R}^k) \) and \( D([0, T], \mathbb{R}^m) \), where \( Y^\varepsilon_t \) is a semimartingale with \( Y^\varepsilon_t = M^\varepsilon_t + A^\varepsilon_t \) be its Doob–Meyer decomposition. Suppose that \( (X^\varepsilon, Y^\varepsilon, U^\varepsilon) \) and \( (X, Y, U) \) satisfy the following equation

\[
\begin{align*}
X^\varepsilon_t &= X_0 + U^\varepsilon_t + \int_0^t f(X^\varepsilon_s) dY^\varepsilon_s, \\
X_t &= X_0 + \int_0^t f(X_{s-}) dY_s,
\end{align*}
\]
where \( f : \mathbb{R}^k \to \mathbb{R}^{k \times m} \) is a matrix-valued function and \( X_0 \) is the same initial condition for \( X_t^\varepsilon, X_t \). Assume that the following conditions are satisfied:

(C1). The processes \((U^\varepsilon, Y^\varepsilon)\) converge to \((0, Y)\) in probability with respect to Skorokhod topology, i.e., for all \( \delta > 0 \),

\[
\lim_{\varepsilon \to 0} \mathbb{P}(d(U^\varepsilon, 0) + d(Y^\varepsilon, Y) > \delta) = 0.
\]

(C2). The total variations of \( A^\varepsilon \), denoted by \( V_t(A^\varepsilon) \), are stochastically bounded for each \( t \), i.e.,

\[
\lim_{L \to \infty} \mathbb{P}(V_t(A^\varepsilon) > L) = 0, \text{ uniformly in } \varepsilon.
\]

Moreover, if there exists a unique global solution \( X \) to (5.5). Then, \( X^\varepsilon \) converges to \( X \) in probability with respect to Skorokhod topology.

In our case, let \( X_t^\varepsilon := (t, q_t^\varepsilon, p_t^\varepsilon) \), \( X_t^0 := (0, q_t^0, p_t^0) \) and \( U_t^\varepsilon := (0, R_t^\varepsilon, p_t^\varepsilon) \). Define a vector function \( \tilde{F}_i(t, x) := \left( y_i^{-1} \right)_j(t, q) F_j(t, x) \), a matrix function \( \tilde{\sigma}_i^0(t, x) := \left( y_i^{-1} \right)_j(t, q) \sigma_j^0(t, x) \) and a tensor function \( \tilde{\sigma}_i^{kl}(t, x) := \tilde{G}_{iab}^b(t, q) \sigma_a^k(t, x) \sigma_b^l(t, x) \).

Then, we define a vector function \( f : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d \times (1+1+d^2+d^2+d+1)} \) as

\[
f_i(t, x) = \left( \tilde{f}_i(t, x), \tilde{F}_i(t, x), \tilde{\sigma}_i^1(t, x), ..., \tilde{\sigma}_i^d(t, x), \right. \\
\left. \tilde{\sigma}_i^{11}(t, x), ..., \tilde{\sigma}_i^{dd}(t, x), \tilde{G}_i^{11}(t, q), ..., \tilde{G}_i^{dd}(t, q), 0 \right),
\]

where \( \tilde{f}_i(t, x) = 1 \), and \( \tilde{f}_i(t, x) = 0 \) for \( i = 2, ..., d \). Also, define \( Y_t^\varepsilon \) with path in \( \mathcal{D}([0, T], \mathbb{R}^{1+1+d^2+d^2+d+1}) \) as

\[
Y_t^\varepsilon = \begin{pmatrix}
t \\
t \\
L_t \\
\int_0^L \int_{\mathbb{R}^d \setminus \{0\}} x_1 x N(ds, dx) \\
\vdots \\
\int_0^L \int_{\mathbb{R}^d \setminus \{0\}} x_d x N(ds, dx) \\
(p_t^1)_1(p_t^e) - (p_0^1)_1(p_0^e) \\
\vdots \\
(p_t^d)_d(p_t^e) - (p_0^d)_d(p_0^e) \\
0
\end{pmatrix}.
\]

Thus, we can rewrite (5.1) in the form

\[
X_t^\varepsilon = X_0 + U_t^\varepsilon + \int_0^t f(X_{s^-}^\varepsilon) dY_s^\varepsilon.
\]
Similarly, let $X_t := (t, q_t, 0)$. Then, (5.2) can be rewritten as

$$X_t = X_0 + \int_0^t f(X_s) dY_s,$$

with $f$ defined by (5.6), and $Y_t$ defined by

$$Y_t = \begin{pmatrix} t \\ t \\ L_t \\ \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} x_1 x N(ds, dx) \\ \vdots \\ \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} x_d x N(ds, dx) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

(5.10)

Now, we are ready to show that the stochastic Hamiltonian system (5.1) converges to the limiting Eq. (5.2). According to Lemma 5.1, it suffices to check the two conditions (C1)(C2). To complete the verification, we state four lemmas below which will be used later.

We divide the remainder term $R^\varepsilon_t$ into the sum of the term $\gamma^{-1}(t, q^\varepsilon_t) p^\varepsilon_t$ and the other term $\tilde{R}^\varepsilon_t$, where $\tilde{R}^\varepsilon_t$ is defined as

$$d(\tilde{R}^\varepsilon_t)_i = -(p^\varepsilon_t)_j \partial_t (\gamma^{-1})^j(t, q^\varepsilon_t) dt - \tilde{G}^{ab}_i (t, q^\varepsilon_t) [(p^\varepsilon_t)_a F_b(t, x^\varepsilon_t) dt + (p^\varepsilon_t)_b F_a(t, x^\varepsilon_t) dt + (p^\varepsilon_t)_a \sigma^\varepsilon_a (t, x^\varepsilon_t) d(L_1)_\rho + (p^\varepsilon_t)_b \sigma^\varepsilon_a (t, x^\varepsilon_t) d(L_1)_\rho].$$

(5.11)

Now, we will prove that $\tilde{R}^\varepsilon_t$ converge to zero in uniform metric on $\mathcal{D}[0, T]$, under an additional assumption below.

**Assumption 6** Assume that function $\gamma$ satisfies the following condition, $\partial_t \gamma$ and $\partial_q \gamma$ are bounded on $[0, T] \times \mathbb{R}^n$, for every $T > 0$.

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 5.2** Under Assumptions 1–6 with $K(\varepsilon, t, p) = |p|^2 / 2\varepsilon$, for every $T > 0$, we have

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

(5.12)
where \( \beta(\theta) \) is a piecewise function

\[
\beta(\theta) = \begin{cases} 
\frac{\theta}{2}, & \text{if } 0 < \theta < 2, \\
1, & \text{if } 2 \leq \theta \leq 2\theta_0.
\end{cases}
\] (5.13)

**Proof** Integrating (4.9) on \([0, T]\), then taking expectation and supremum on it, for \(2 \leq \theta \leq 2\theta_0\) we have

\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |(\tilde{R}^\varepsilon_t)|^\theta \right] \\
\leq 3^{\theta-1} \left( \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \int_0^T (p^\varepsilon_s)^j \partial_s (\gamma^{-1})_i(s, q^\varepsilon_s) ds \right|^\theta \right] \right) \\
+ \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \int_0^T \tilde{G}^{ab}_i(s, q^\varepsilon_s)(p^\varepsilon_s)_a F_b(s, x^\varepsilon_s) + \tilde{G}^{ab}_i(s, q^\varepsilon_s)(p^\varepsilon_s)_b F_a(s, x^\varepsilon_s) ds \right|^\theta \right] \\
+ \mathbb{E} \left[ \sup_{t \in [0, T]} \left| \int_0^T \left[ \tilde{G}^{ab}_i(s, q^\varepsilon_s)(p^\varepsilon_s)_a \sigma^\rho_b (t, x^\varepsilon_s) + \tilde{G}^{ab}_i(s, q^\varepsilon_s)(p^\varepsilon_s)_b \sigma^\rho_a (t, x^\varepsilon_s) \right] d(L_t)_\rho \right|^\theta \right] \\
=: \sum_{i=1}^3 J_i.
\] (5.14)

We will now give upper bounds of terms \(\{J_i\}_{i=1}^3\). For the first term, applying Hölder inequality and the boundedness of \(\partial_t \gamma^{-1}\) we have

\[
J_1 \leq \mathbb{E} \left[ \sup_{t \in [0, T]} t^{\theta-1} \int_0^T |(p^\varepsilon_s)^j \partial_s (\gamma^{-1})_i(s, q^\varepsilon_s)|^\theta ds \right] \\
\leq ||\partial_t \gamma^{-1}||^\theta_\infty \mathbb{E} \left[ T^{\theta-1} \int_0^T |p^\varepsilon_s|^\theta ds \right] \leq T^{\theta-1} ||\partial_t \gamma^{-1}||^\theta_\infty \sup_{t \in [0, T]} \mathbb{E} \left[ |p^\varepsilon_t|^\theta \right].
\] (5.15)

Note that by Assumption 6, we can deduce that the function \(\tilde{G}(t, q)\) is bounded. For the second term, we have

\[
J_2 \leq 2^\theta T^{\theta-1} \mathbb{E} \left[ \int_0^T |p^\varepsilon_s|^\theta ||F||^\theta_\infty ||\tilde{G}||^\theta_\infty ds \right] \leq (2T)^{\theta-1} ||F||^\theta_\infty ||\tilde{G}||^\theta_\infty \sup_{t \in [0, T]} \mathbb{E} \left[ |p^\varepsilon_t|^\theta \right].
\] (5.16)
For the last term \( J_3 \), applying Kunita’s first inequality (Applebaum 2009), we have

\[
J_3 \leq 2 \times 2^{\theta-1} \mathbb{E} \left[ \sup_{t \in [0,T]} \left| \int_0^t \int_{|x| > 1} \tilde{G}_i^{ab} (s, q_{s^-}, (p_{s^-})_a \sigma_b^\rho (s, x_{s^-}) x \tilde{N}(ds, dx) \right|^\theta \right]
\]

\[
+ 2 \times 2^{\theta-1} \mathbb{E} \left[ \sup_{t \in [0,T]} \left| \int_0^t \int_{\mathbb{R}^d \setminus [0]} \tilde{G}_i^{ab} (s, q_{s^-}, (p_{s^-})_a \sigma_b^\rho (s, x_{s^-}) x \nu(dx)ds \right|^\theta \right]
\]

\[
\leq 2^\theta \tilde{C} \mathbb{E} \left[ \left( \int_0^T \int_{\mathbb{R}^d \setminus [0]} \left| \tilde{G}_i^{ab} (s, q_{s^-}, (p_{s^-})_a \sigma_b^\rho (s, x_{s^-}) x \right|^2 \nu(dx)ds \right)^\theta \right]
\]

\[
+ 2^\theta \tilde{C} \mathbb{E} \left[ \left( \int_0^T \int_{\mathbb{R}^d \setminus [0]} \left| \tilde{G}_i^{ab} (s, q_{s^-}, (p_{s^-})_a \sigma_b^\rho (s, x_{s^-}) x \right| \nu(dx)ds \right)^\theta \right]
\]

\[
+ 2^\theta T^\theta \tilde{C} \left( \int_{|x| > 1} |x| \nu(dx) \right)^\theta \sup_{t \in [0, T]} \mathbb{E} \left[ |p_t^\varepsilon|^\theta \right]
\]

\[
\leq \tilde{C} (\theta, T, \nu, \tilde{G}, \sigma) \sup_{t \in [0, T]} \mathbb{E} \left[ |p_t^\varepsilon|^\theta \right], \tag{5.17}
\]

where we used the boundedness of \( \sigma \) and \( \tilde{G} \). Substitute all these upper bound together, we obtain that

\[
\mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{R}_t^\varepsilon|^\theta \right] \leq \tilde{C} (\theta, T, \nu, F, \gamma, \sigma) \sup_{t \in [0, T]} \mathbb{E} \left[ |p_t^\varepsilon|^\theta \right].
\]

This together with Proposition 3.3 and Hölder inequality yield desired result. \( \square \)

As we discussed in Remark 3.3, the process \( p_t^\varepsilon \) does not converge to zero in uniform metric. Therefore, for other terms in process \( (U_t^\varepsilon, Y_t^\varepsilon) \), we cannot show the convergence in uniform metric like \( \tilde{R}^\varepsilon \). However, we can prove that they converge in probability with respect to Skorokhod topology.

**Lemma 5.3** Under Assumptions 1–6 with \( K (\varepsilon, t, p) = \frac{|p|^2}{2\varepsilon} \). The family of stochastic processes \( p_t^\varepsilon \) converges to zero in probability with respect to Skorokhod topology, i.e., for all \( \delta > 0 \)

\[
\lim_{\varepsilon \to 0} \mathbb{P} \left( d(p_t^\varepsilon, 0) > \delta \right) = 0. \tag{5.18}
\]

**Proof** According to Proposition 3.3, we know that for each \( t \in [0, T] \), the processes \( p_t^\varepsilon \) converge to zero in distribution. Hence, the finite dimensional distributions of the processes \( p_t^\varepsilon \) converge to the finite dimensional distributions of zero process (Billingsley 2013, Theorem 3.9). Moreover, if the family of processes \( (p_t^\varepsilon)_\varepsilon \) is tight, then the processes \( p_t^\varepsilon \) converge in distribution to the zero process, which yields that the processes \( p_t^\varepsilon \) converge in probability to the zero process. To prove the tightness of \( (p_t^\varepsilon)_\varepsilon \), by Aldous’ criterion (Jacod and Shiryaev 2013, page 356), it suffices to check the following two conditions:
(i) For all $T > 0$ and $\delta > 0$, there exists $N > 0$ such that, for all $\varepsilon$

$$\mathbb{P}\left( \sup_{t \in [0, T]} |p^\varepsilon_t| > N \right) \leq \delta; \quad (5.19)$$

(ii) For every stopping time $\tau \in [0, T]$ and every $\lambda > 0$, it holds that

$$\lim_{\delta \to 0} \lim_{\varepsilon \to 0} \sup_{\tau \in [0, T-\delta]} \mathbb{P}\left( |p^\varepsilon_{\tau+\delta} - p^\varepsilon_{\tau}| > \lambda \right) = 0. \quad (5.20)$$

For the assertion (i), inequality (5.19) follows by (3.21) and Markov inequality. To verify (ii), let $\tau \leq T - \delta_0$ be a bounded stopping time. For every $\delta \in (0, \delta_0)$, by the Markov property, we have

$$\mathbb{P}(|p^\varepsilon_{\tau+\delta} - p^\varepsilon_{\tau}| > \lambda) = \mathbb{E}(\mathbb{P}(|p^\varepsilon_{s+\delta} - y| > \lambda)|_{(s,y)=(\tau, p^\varepsilon_{\tau})}). \quad (5.21)$$

Define a new process $\tilde{p}^\varepsilon_t := p^\varepsilon_t - y$. We get

$$d\tilde{p}^\varepsilon_t = \left( -\frac{1}{\varepsilon} \gamma(t, \tilde{q}^\varepsilon_t)(\tilde{p}^\varepsilon_t + y) + F(t, q^\varepsilon_t, \tilde{p}^\varepsilon_t + y) \right) dt + \sigma(t, q^\varepsilon_t, \tilde{p}^\varepsilon_t + y) dL_t, \quad \tilde{p}^\varepsilon_0 = 0. \quad (5.22)$$

By Itô formula, we have

$$(\tilde{p}^\varepsilon_t)^2 = \int_0^t 2\langle \tilde{p}^\varepsilon_s, \left( -\frac{1}{\varepsilon} \gamma(s, q^\varepsilon_s)(\tilde{p}^\varepsilon_s + y) + F(s, q^\varepsilon_s, \tilde{p}^\varepsilon_s + y) \right) \rangle ds $$

$$+ \int_0^t \int_{|x|<1} \left[ (\tilde{p}^\varepsilon_s + \sigma(s, q^\varepsilon_s, \tilde{p}^\varepsilon_s + y)x)^2 - (\tilde{p}^\varepsilon_s)^2 \right] \tilde{N}(ds, dx) $$

$$+ \int_0^t \int_{|x|\geq1} \left[ (\tilde{p}^\varepsilon_s + \sigma(s, q^\varepsilon_s, \tilde{p}^\varepsilon_s + y)x)^2 - (\tilde{p}^\varepsilon_s)^2 \right] N(ds, dx) $$

$$- 2(\sigma(s, q^\varepsilon_s, \tilde{p}^\varepsilon_s + y)x, \tilde{p}^\varepsilon_s)\nu(dx)ds. \quad (5.23)$$

For the first integral, utilizing the boundness of $\gamma$ and $F$, we have

$$\mathbb{E}_{s,y}\left[ \int_0^t \langle \tilde{p}^\varepsilon_s, -\frac{1}{\varepsilon} \gamma(s, q^\varepsilon_s)(\tilde{p}^\varepsilon_s + y) \rangle ds \right]$$

$$\leq \mathbb{E}_{s,y}\left[ \int_0^t \langle \tilde{p}^\varepsilon_s, -\frac{1}{\varepsilon} \gamma(s, q^\varepsilon_s)y \rangle ds \right] \leq \frac{||\gamma||_\infty |y|}{\varepsilon} \int_0^t \mathbb{E}_{s,y}[|\tilde{p}^\varepsilon_s|] ds, \quad (5.24)$$

and

$$\mathbb{E}_{s,y}\left[ \int_0^t 2\langle \tilde{p}^\varepsilon_s, F(s, q^\varepsilon_s, \tilde{p}^\varepsilon_s + y) \rangle ds \right] \leq 2||F||_\infty \int_0^t \mathbb{E}_{s,y}[|\tilde{p}^\varepsilon_s|] ds, \quad (5.25)$$
where $\mathbb{E}_{s,y}$ denotes the conditional expectation under the condition $p^ε_s = y$. The second integral is a martingale, since

\[
\int_0^t \int_{|x| < 1} \mathbb{E}_{s,x,y} \left[ \left( (\tilde{p}^ε_{s-} + \sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x)^2 - (\tilde{p}^ε_{s-})^2 \right)^2 \right] \nu(dx)ds \\
\leq \int_0^t \int_{|x| < 1} \mathbb{E}_{s,x,y} \left[ \left( 2||\sigma||_\infty |x| ||\sigma||^2_\infty |x|^2 \right)^2 \right] \nu(dx)ds \\
\leq \tilde{C} \int_0^t \left( \int_{|x| < 1} |x|^2 \nu(dx) \mathbb{E}_{s,x,y} \left[ |\tilde{p}^ε_s|^2 \right] + \int_{|x| < 1} |x|^4 \nu(dx) \right) ds < +\infty.
\]

For the third integral, we have

\[
\mathbb{E}_{s,x,y} \left[ \int_0^t \int_{|x| \geq 1} ( (\tilde{p}^ε_{s-} + \sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x)^2 - (\tilde{p}^ε_{s-})^2 ) N(ds, dx) \right] \\
= \int_0^t \int_{|x| \geq 1} \mathbb{E}_{s,x,y} \left[ 2(\tilde{p}^ε_{s-} + \sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x) + (\sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x)^2 \right] \nu(dx)ds \\
\leq \tilde{C} \int_0^t \left( \int_{|x| \geq 1} |x| \nu(dx) \mathbb{E}_{s,x,y} \left[ |\tilde{p}^ε_s|^2 \right] + \int_{|x| \geq 1} |x|^2 \nu(dx) \right) ds.
\]

For the last term, we have

\[
\mathbb{E}_{s,x,y} \left[ \int_0^t \int_{|x| < 1} ( (\tilde{p}^ε_{s-} + \sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x)^2 - (\tilde{p}^ε_{s-})^2 \right. \\
\left. - 2(\sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x, \tilde{p}^ε_{s-}) ) \nu(dx)ds \right] \\
= \mathbb{E}_{s,x,y} \left[ \int_0^t \int_{|x| < 1} (\sigma(s, q^ε_{s-}, \tilde{p}^ε_{s-} + y)x)^2 \nu(dx)ds \right] \leq \tilde{C} \int_0^t \int_{|x| < 1} |x|^2 \nu(dx)ds.
\]

Taking (5.23)–(5.28) together with $\sup_{t \in [0,T]} \mathbb{E}[|p^ε_t|] = O(\varepsilon^{1/2})$, we deduce that

\[
\sup_{t \in [0,T]} \mathbb{E}_{s,y} \left[ (\tilde{p}^ε_t)^2 \right] < \tilde{C} (1 + \varepsilon^{-1/2}|y|) T.
\]

Hence, by Markov inequality, we have

\[
\mathbb{P}_{s,y} (|p^{s+\delta} - y| > \lambda) \leq \frac{\sup_{t \in [0,\delta]} \mathbb{E}_{s,y} \left[ (\tilde{p}^ε_t)^2 \right]}{\lambda^2} \leq \frac{\tilde{C} (1 + \varepsilon^{-1/2}|y|) \delta}{\lambda^2}.
\]
Combining the estimates with Markov property, we obtain
\[
\mathbb{P}(|p_{t+\delta}^\varepsilon - p_t^\varepsilon| > \lambda) = \mathbb{P}(|p_{t+\delta}^\varepsilon - p_t^\varepsilon| > \lambda, |p_t^\varepsilon| > \varepsilon^{1/2} R) \\
+ \mathbb{P}(|p_{t+\delta}^\varepsilon - p_t^\varepsilon| > \lambda, |p_t^\varepsilon| \leq \varepsilon^{1/2} R) \\
\leq \mathbb{P}(|p_t^\varepsilon| > \varepsilon^{1/2} R) + \mathbb{E}\left[\mathbb{P}_s, y \left(|p_{t+\delta}^\varepsilon - y| > \lambda\right) 1_{\{|p_t^\varepsilon| \leq \varepsilon^{1/2} R\}}\right] \\
\leq \sup_{t \in [0, T]} \mathbb{E}\left[|p_t^\varepsilon|^2\right] \frac{\hat{C}(1 + R)\delta}{\lambda^2} \\
\leq \frac{\hat{C} T}{\lambda^2} + \frac{\hat{C}(1 + R)\delta}{\lambda^2}.
\]

Letting \( \delta \to 0 \) first, and then \( R \to +\infty \), one can obtain (5.20). This completes the proof. \( \square \)

**Lemma 5.4** Under Assumptions 1–6 with \( K(\varepsilon, t, p) = \frac{|p|^2}{2\varepsilon} \). The family of stochastic process \( (p_t^\varepsilon)_a(p_t^\varepsilon)_b \) and \((y^{-1})(t, q_t^\varepsilon)p_t^\varepsilon \) converge to zero in probability with respect to Skorokhod topology.

**Proof** By the definition of the Skorokhod metric, we have
\[
\mathbb{P}(d((p_t^\varepsilon)_a(p_t^\varepsilon)_b, 0) > \delta) = \mathbb{P}\left(\inf_{\lambda \in \Lambda} \left(\|\lambda\|_\infty \sup_t |(p_t^\varepsilon)_a(p_t^\varepsilon)_b|\right) > \delta\right) \\
\leq \mathbb{P}\left(\inf_{\lambda \in \Lambda} \left(\|\lambda\|_\infty \sup_t |(p_t^\varepsilon)_a| \sup_t |(p_t^\varepsilon)_b|\right) > \delta, \sup_t |(p_t^\varepsilon)_a| \leq K\right) \\
+ \mathbb{P}\left(\sup_t |(p_t^\varepsilon)_a| > K\right) \\
\leq \mathbb{P}\left(\inf_{\lambda \in \Lambda} \left(\|\lambda\|_\infty \sup_t |(p_t^\varepsilon)_a| \sup_t |(p_t^\varepsilon)_b|\right) > \delta, \sup_t |(p_t^\varepsilon)_a| \leq K\right) \\
+ \mathbb{P}\left(\sup_t |(p_t^\varepsilon)_a| > K\right) \\
= \mathbb{P}\left(d(((p_t^\varepsilon)_b, 0) > \frac{\delta}{K}\right) + \mathbb{P}\left(\sup_t |(p_t^\varepsilon)_a| > K\right).
\]

By Markov inequality and (3.21), we have
\[
\sup_{0 < \varepsilon < 1} \mathbb{P}\left(\sup_{t \in [0, T]} |(p_t^\varepsilon)_a| > K\right) \leq \sup_{0 < \varepsilon < 1} \frac{\mathbb{E}\left[\sup_{t \in [0, T]} |p_t^\varepsilon|^2\right]}{K^2} \leq \frac{\hat{C}(T, F, \sigma, y, \nu)}{K^2}.
\]
Hence, for every $\delta > 0$, the limit $\lim_{\varepsilon \to 0} \mathbb{P} \left( d \left( (\gamma^{-1})(t, q_t^\varepsilon) p_t^\varepsilon, 0 \right) > \delta \right) = 0$ follows by letting $\varepsilon \to 0$ first and then $K \to +\infty$ in (5.29). Similarly, for every $\delta > 0$, we have
\[
\lim_{\varepsilon \to 0} \mathbb{P} \left( d \left( (\gamma^{-1})(t, q_t^\varepsilon) p_t^\varepsilon, 0 \right) > \delta \right) = \mathbb{P} \left( \inf_{\lambda \in \Lambda} \left( ||\lambda||^0 \vee \sup_t |(\gamma^{-1})(t, q_t^\varepsilon) p_t^\varepsilon| \right) > \delta \right)
\leq \lim_{\varepsilon \to 0} \mathbb{P} \left( \inf_{\lambda \in \Lambda} \left( ||\lambda||^0 \vee ||\gamma^{-1}||_\infty \sup_t |p_t^\varepsilon| \right) > \delta \right)
\leq \lim_{\varepsilon \to 0} \mathbb{P} \left( \inf_{\lambda \in \Lambda} \left( ||\lambda||^0 \vee \sup_t |p_t^\varepsilon| \right) > \frac{\delta}{1 + ||\gamma^{-1}||_\infty} \right) = 0.
\]

The proof is complete. \hfill \Box

Finally, we give the following lemma to illustrate the stochastic boundedness for total variation of process $(p_t^\varepsilon)(p_t^\varepsilon)^*$, which is a component of process $Y_t^\varepsilon$.

**Lemma 5.5** Under Assumptions 1–6 with $K(\varepsilon, t, p) = \frac{|p|^2}{2\varepsilon}$. The total variation of the increasing process in the Doob–Meyer decomposition of process $(p_t^\varepsilon)(p_t^\varepsilon)^*$ is stochastically bounded for each $t \in [0, T]$.

**Proof** Applying Itô product formula to $(p_t^\varepsilon)(p_t^\varepsilon)^*$, we have the following Doob–Meyer decomposition $(p_t^\varepsilon)(p_t^\varepsilon)^* = M_t^\varepsilon + A_t^\varepsilon$, where
\[
A_t^\varepsilon = \int_0^t \left( -\frac{1}{\varepsilon} \gamma(s, q_s^\varepsilon) p_s^\varepsilon + F(s, q_s^\varepsilon, p_s^\varepsilon) \right)^* ds
+ \int_0^t \int_{|x| \geq 1} (p_s^\varepsilon)(\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x)^* N(ds, dx)
+ \int_0^t (p_t^\varepsilon)^* \left[ -\frac{1}{\varepsilon} \gamma(s, q_s^\varepsilon) p_s^\varepsilon + F(s, q_s^\varepsilon, p_s^\varepsilon) \right] ds
+ \int_0^t \int_{|x| \geq 1} (p_s^\varepsilon)^* (\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x) N(ds, dx)
+ \int_0^t (\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x)(\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x)^* N(ds, dx)
+ \int_0^t \int_{|x| < 1} (\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x)(\sigma(s, q_s^\varepsilon, p_s^\varepsilon)x)^* n(dx) ds,
\]
and $M_t^\varepsilon$ is a local martingale. Actually, total variations of $A_t^\varepsilon$ are stochastically bounded uniformly in $\varepsilon$ since $\sup_{t \in [0, T]} \mathbb{E}[(p_t^\varepsilon)^2] = O(\varepsilon)$ and boundedness of the functions $\gamma, \sigma, F$. \hfill \Box

**Remark 5.1** It is important to note that we cannot prove convergence to the limiting equation for the general kinetic energy $K$ of the growth order more than quadratic satisfying Assumptions 1–6. This is because that the total variations of $A_t^\varepsilon$ corresponding
to \((p^ε_t, p^ε_t)\)

\[ A^ε_t = \int_0^t (p^ε_s) \left[ -\gamma(s, q^ε_s) \nabla_p K^ε(s, p^ε_s) + F(s, q^ε_s, p^ε_s) \right]^* ds \]

\[ + \int_0^t \int_{|x| \geq 1} (p^ε_{s-})(\sigma(s, q^ε_{s-}, p^ε_{s-})x)^* N(ds, dx) \]

\[ + \int_0^t (p^ε_t)^* \left[ -\gamma(s, q^ε_s) \nabla_p K^ε(s, p^ε_s) + F(s, q^ε_s, p^ε_s) \right] ds \]

\[ + \int_0^t \int_{|x| \geq 1} (\sigma(s, q^ε_{s-}, p^ε_{s-})x)(\sigma(s, q^ε_{s-}, p^ε_{s-})x)^* N(ds, dx) \]

\[ + \int_0^t \int_{|x| < 1} (\sigma(s, q^ε_s, p^ε_s)x)(\sigma(s, q^ε_s, p^ε_s)x)^* v(dx) ds, \]

are not stochastically bounded uniformly in \(\varepsilon\). Since by the growth assumptions for \(\nabla_x K\), we have

\[
\mathbb{E}[|p^ε_s||\gamma(s, q^ε_s) \nabla_p K^ε(s, p^ε_s)|] \\
\leq ||\gamma||_\infty \mathbb{E} \left[ C_2 \frac{|p^ε_s|}{\sqrt{\varepsilon}} \left( 1 + \frac{|p^ε_s|}{\sqrt{\varepsilon}} \right)^{-\eta - 1} \right] \leq \tilde{C} \left( \varepsilon^{-1/2} \mathbb{E} \left[ |p^ε_s|^\eta \right] + \varepsilon^{-\eta/2} \mathbb{E} \left[ |p^ε_s|^{2\eta} \right] \right),
\]

which has order of \(O(\varepsilon^{1-\eta/2})\). It will diverge for \(\eta > 2\) and is bounded uniformly in \(\varepsilon\) for \(\eta = 2\) according to Proposition 3.3.

Now, we are going to prove that the stochastic Hamiltonian system (5.1) converges to the limiting Eq. (5.2). Note that we can only demonstrate the convergence under Assumptions 1–6 with classical Hamiltonian \(K(\varepsilon, t, p) = \frac{|p|^2}{2\varepsilon}\); therefore, we have the following result by concluding these conditions again.

**Theorem 5.1** Let \((q^ε_t, p^ε_t)\) be the solution of stochastic Hamiltonian systems (5.1) with initial condition \((p^0_0, q^0_0)\) and \(q_t\) be the solution of stochastic differential equation (5.2) with initial condition \(q_0\). Moreover, assume that the initial kinetic energy is bounded by a positive number \(C_0\), that is, \(\frac{(p^0_0)^2}{2\varepsilon} \leq C_0\). Further, suppose the following conditions hold:

1. The dissipative coefficient \(\gamma\) is bounded, independent of \(p\) and \(C^1\) in \((t, q)\). Besides, it is symmetric with eigenvalues bounded below by a constant \(\lambda > 0\).
2. The noise intensity coefficient \(\sigma\) and the external force \(F\) are bounded and locally Lipschitz.
3. There exists a constant \(\theta_0 \geq 1\) such that the Lévy measure \(v\) satisfies

\[
\int_{|x| \geq 1} |x|^{4\theta_0} v(dx) < \infty.
\]
4. The functions $\partial_t \gamma$ and $\partial_{q_i} \gamma$ are bounded.

5. The initial condition $(q_0^\varepsilon, p_0^\varepsilon)$ converges to $(q_0, p_0)$ in probability, i.e., for every $\delta > 0$,

$$\lim_{\varepsilon \to 0} \mathbb{P} \left( |p_0^\varepsilon - p_0| + |q_0^\varepsilon - q_0| > \delta \right) = 0.$$  

Then, for arbitrary positive $T$, $\delta$, we have

$$\lim_{\varepsilon \to 0} \mathbb{P} \left( d(q_t^\varepsilon, q_t) > \delta \right) = 0. \quad (5.31)$$

**Remark 5.2** When dissipative coefficient $\gamma$ is independent of $q$, the noise-induced drift (5.3) vanishes, and the limiting equation becomes

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

This result coincides with that in Zhang (2008).

**Proof** We can check that the coefficients in the limiting Eq. (5.2) satisfy the conditions in Lemma (B.1). Hence, there exists a unique global solution $q_t$ to (5.2). To finish the proof of Theorem 5.1, all that remains is to verify the conditions (C1) and (C2) in Lemma 5.1. On the one hand, applying Markov inequality together with the fact that the uniform metric is finer than Skorokhod metric, we see that for every $\delta > 0$,

$$\mathbb{P} \left( d(\tilde{R}_t^\varepsilon, 0) > \delta \right) \leq \mathbb{E} \left( \sup_{t \in [0, T]} \left| \tilde{R}_t^\varepsilon \right| > \delta \right) \leq \frac{\mathbb{E} \left( \sup_{t \in [0, T]} \left| \tilde{R}_t^\varepsilon \right|^2 \right)}{\delta^2}. $$

Then, Lemma 5.2 implies that $\tilde{R}_t^\varepsilon$ converge to 0 as $\varepsilon \to 0$ in probability with respect to Skorokhod topology. This together with Lemmas 5.3 and 5.4 shows that $(U_t^\varepsilon, Y_t^\varepsilon)$ converges to $(0, Y_t)$ in probability with respect to Skorokhod topology, that is, (C1) holds.

On the other hand, we have the Doob–Meyer decomposition of the process $Y_t^\varepsilon := M_t^\varepsilon + A_t^\varepsilon$ with a local martingale $M_t^\varepsilon$, and an increasing process $A_t^\varepsilon$ as following

$$A_t^\varepsilon = \begin{pmatrix} t \\ \int_0^t \int_{|x| \geq 1} xN(t, dx) \\ \int_0^t \int_{|x| \leq 1} x_1 xN(ds, dx) + \int_0^t \int_{|x| < 1} x_1 x_1 \nu(dx)(dx)ds \\ \vdots \\ \int_0^t \int_{|x| \geq 1} x_d xN(ds, dx) + \int_0^t \int_{|x| < 1} x_d x_1 \nu(dx)(dx)ds \\ (A_t^\varepsilon)_1 \\ \vdots \\ (A_t^\varepsilon)^d \\ 0 \end{pmatrix}.$$
where \((A^ε_1, ..., A^ε_d)\) is defined by (5.30). According to Lemma 5.5, the total variation of \(A^ε_t\) is stochastically bounded uniformly in \(\varepsilon\). This implies that (C2) holds. Therefore, Theorem 5.1 is proved.

\[\square\]

5.2 Convergence Under Unbounded Assumptions

As a matter of fact, we can extend the result in Theorem 5.1 by using the global existence of the limiting Eq. (5.2). Now, we present a stronger result by weakening the external force \(F\) and diffusion coefficient \(\sigma\) to be unbounded as below.

**Theorem 5.2** Let \((q^ε_t, p^ε_t)\) be the solution of stochastic Hamiltonian system (5.1) with initial condition \((p^ε_0, q^ε_0)\) and \(q_t\) be the solution of the limiting stochastic differential equation (5.2) with the same initial condition \(q_0\). Moreover, assume that the initial kinetic energy is bounded by a positive number \(C_0\), that is, \(\frac{(p^ε_0)^2}{2\varepsilon} \leq C_0\). Further, suppose the following conditions hold:

1. The dissipative coefficient \(\gamma\) is independent of \(p\) and \(C_1\) in \((t, q)\). Besides, it is symmetric with eigenvalues bounded below by a constant \(\lambda > 0\).
2. The external force \(F\) and the noise intensity coefficient \(\sigma\) are locally Lipschitz continuous; And there exists a constant \(C\) such that 
   \[|F(t, q, 0)|^2 \leq C(1+|q|^2), \quad ||\sigma(t, q, 0)||^2 \leq C(1+|q|).\]
3. There exists a constant \(θ_0 \geq 1\) such that the Lévy measure \(ν\) satisfies 
   \[\int_{|x| \geq 1} |x|^{4θ_0} ν(dx) < ∞.\]
4. The function \(\partial q^ε_t \gamma\) is bounded.
5. The initial condition \((q^ε_0, p^ε_0)\) converges to \((q_0, p_0)\) in probability, i.e., for every \(δ > 0\),
   \[\lim_{\varepsilon \to 0} P(\{|p^ε_0 - p_0| + |q^ε_0 - q_0| > δ\}) = 0.\]

Then, for arbitrary positive \(T, δ\), we have
\[\lim_{\varepsilon \to 0} P(d(q^ε_t, q_t) > δ) = 0.\]  
(5.32)

**Proof** Under above assumptions, we can check that the coefficients in the limiting Eq. (5.2) satisfy the conditions in Lemma (B.1). Hence, there exists a unique global solution \(q_t\) to (5.2). Let \(χ : \mathbb{R}^n → [0, 1]\) be a \(C^∞\) bump function which equal to 1 on \(B_1(0) := \{|q| < 1\}\) and equal to zero outside \(B_2(0)\). Let \(χ_r(q) = χ(q/r)\), and we define following functions
\[F_r(t, x) = χ_r(q)F(t, x), \sigma_r(t, x) = χ_r(q)σ(t, x), γ_r(t, q)\]
where the first term on the right hand side converges to 0 as $\varepsilon \to 0$. We focus on the second term, (5.2) with these coefficients. Then, Theorem 5.1 implies that, for every $q_{t}$ the solution to the corresponding stochastic Hamiltonian systems satisfying the conditions in Theorem 5.1. Let $(q^{f,e}_{t}, p^{f,e}_{t})$ be solution to the corresponding stochastic Hamiltonian systems. Similarly, let $q^{r}_{t}$ be the solution to the corresponding limiting stochastic differential equation (5.2) with these coefficients. Then, Theorem 5.1 implies that, for every $T > 0$, $\delta > 0$,

$$\mathbb{P}(d(q^{r,e}_{t}, q^{f}_{t}) > \delta) \to 0, \text{ as } \varepsilon \to 0. \quad (5.33)$$

We will use this result to prove that $q^{e}_{t}$ converges to $q_{t}$ in probability.

Define stopping times $\tau^{e}_{t} = \inf\{t : |q^{e}_{t}| \geq r\}$ and $\tau^{r}_{t} = \inf\{t : |q^{r}_{t}| \geq r\}$. The drift coefficients and diffusion coefficients of the modified and unmodified stochastic Hamiltonian systems agree on the ball $\{|q| < r\}$. Hence, we have

$$q^{e}_{t,\wedge t} = q^{r,e}_{t,\wedge t}, \quad q^{r}_{t,\wedge t} = q^{r}_{t,\wedge t} \text{ for all } t \geq 0 \text{ a.s.} \quad (5.34)$$

For every $T > 0$, $\delta > 0$ and $t \in [0, T]$, we can deduce that

$$\mathbb{P}(d(q^{e}_{t}, q_{t}) > \delta)$$

$$= \mathbb{P}(\tau^{e}_{t} > T, d(q^{e}_{t,\wedge t}, q^{r}_{t,\wedge t}) > \delta) + \mathbb{P}(\tau^{e}_{t} \leq T, d(q^{e}_{t}, q_{t}) > \delta)$$

$$= \mathbb{P}(\tau^{r}_{t} > T, d(q^{r,e}_{t,\wedge t}, q^{e}_{t,\wedge t}) > \delta) + \mathbb{P}(\tau^{e}_{t} \leq T, d(q^{e}_{t}, q_{t}) > \delta)$$

$$\leq \mathbb{P}(d(q^{r,e}_{t,\wedge t}, q^{e}_{t,\wedge t}) > \delta) + \mathbb{P}(\tau^{e}_{t} \leq T), \quad (5.35)$$

where the first term on the right hand side converges to 0 as $\varepsilon \to 0$ by (5.33). Then, we focus on the second term,

$$\mathbb{P}(\tau^{e}_{t} \leq T)$$

$$= \mathbb{P}(\tau^{e}_{t} \leq T) + \mathbb{P}(\tau^{e}_{t} > T)$$

$$\leq \mathbb{P}(\tau^{e}_{t} \leq T) + \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| > 1)$$

$$+ \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| \leq 1)$$

$$\leq \mathbb{P}(\tau^{e}_{t} \leq T) + \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{r}_{t,\wedge T}| > 1)$$

$$+ \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| \leq 1)$$

$$\leq \mathbb{P}(\sup_{t \in [0, T]} |q_{t}| \geq r) + \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| > 1)$$

$$+ \mathbb{P}(\tau^{e}_{t} > T, |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| \leq 1), \quad (5.36)$$

where we used the equalities (5.34) in the last two inequalities. Note that when $\tau^{e}_{t} \leq T$, we have $|q^{e}_{t,\wedge T}| \geq r$. Hence, by $|q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| \leq 1$, we can deduce

$$|q^{e}_{t,\wedge T}| \geq |q^{r,e}_{t,\wedge T} - q^{e}_{t,\wedge T}| \geq r - 1.$$
This implies that
\[
\mathbb{P}\left( \tau_{T}^{\varepsilon} \leq T, |q_{T}^{\varepsilon \wedge T} - q_{T}^{\varepsilon \wedge T}| \leq 1 \right) \leq \mathbb{P}\left( \tau_{T}^{\varepsilon} \leq T, |q_{T}^{\varepsilon \wedge T}| \geq r - 1 \right)
\]
\[
\leq \mathbb{P}\left( \sup_{t \in [0, T]} |q_{t}^{\varepsilon}| \geq r - 1 \right). \tag{5.37}
\]
Combining (5.35), (5.36) and (5.37) together, we have
\[
\limsup_{\varepsilon \to 0} \mathbb{P}\left( d(q_{t}^{\varepsilon}, q_{t}) > \delta \right)
\]
\[
\leq \limsup_{\varepsilon \to 0} \mathbb{P}\left( d(q_{t-1}^{\varepsilon}, q_{t-1}^{\varepsilon}) > \delta \right) + \limsup_{\varepsilon \to 0} \mathbb{P}\left( |q_{T}^{\varepsilon \wedge T} - q_{T}^{\varepsilon \wedge T}| > 1 \right)
\]
\[
+ \mathbb{P}\left( \sup_{t \in [0, T]} |q_{t}^{\varepsilon}| \geq r \right) + \mathbb{P}\left( \sup_{t \in [0, T]} |q_{t}| \geq r - 1 \right). \tag{5.38}
\]
Non-explosion of \( q_{t} \) implies that \( \mathbb{P}\left( \sup_{t \in [0, T]} |q_{t}^{\varepsilon}| \geq r \right) \) and \( \mathbb{P}\left( \sup_{t \in [0, T]} |q_{t}| \geq r - 1 \right) \) → 0 as \( r \to +\infty \). The first two terms converge to 0 as \( \varepsilon \to 0 \) due to the equality (5.33). Hence, we have
\[
\limsup_{\varepsilon \to 0} \mathbb{P}\left( d(q_{t}^{\varepsilon}, q_{t}) > \delta \right) \to 0 \text{ as } \varepsilon \to 0. \tag{5.39}
\]
Thus, we complete the proof of Theorem 5.2. \( \square \)

6 An Example

In this section, we present a one-dimensional example.

**Example 6.1** We consider the following one dimensional Hamiltonian system with Lévy noise
\[
d q_{t}^{m} = \frac{1}{m} q_{t}^{m} \, dt,
\]
\[
d p_{t}^{m} = \left( -\frac{1}{m} \gamma(q_{t}^{m}) p_{t}^{m} + F(q_{t}^{m}) \right) \, dt + \sigma(q_{t}^{m}) dL_{t}, \tag{6.1}
\]
where the potential \( V \) vanishes and \( \gamma, \sigma, F \) depend only on position \( q \). Then, the limiting equation becomes
\[
d q_{t} = \frac{F(q_{t})}{\gamma(q_{t})} \, dt + \frac{\sigma(q_{t})}{\gamma(q_{t})} dL_{t} - \int_{\mathbb{R} \setminus \{0\}} \frac{\gamma'(q_{s})}{2\gamma^{3}(q_{s})} \sigma^{2}(q_{s}) \xi_{2} N(ds, dz). \tag{6.2}
\]
Let us compare with the same system with Brownian noise. Recall the corresponding Langevin equation driven by Brownian noise (Eqs. (95) and (97) in Hottovy et al. (2015))
Fig. 1 Positions $q_{tm}$, $q_t$ and $\tilde{q}_t$ are solutions of (6.1), (6.2) and (6.5) with $F = 0.1, \gamma(q) = 0.2q, \sigma(q) = 0.2q, m = 0.001$, respectively. a Trajectory of $q_{tm}$ (blue), $q_t$ (red) and $\tilde{q}_t$ (yellow); b difference of $q_{tm} - q_t$ (red) and $q_{tm} - \tilde{q}_t$ (yellow) (Color figure online)

Fig. 2 Positions $q_{tm}$, $q_t$ and $\tilde{q}_t$ are solutions of (6.1), (6.2) and (6.5) with $F = 0.1, \gamma(q) = 0.2q, \sigma(q) = 0.2q, m = 0.01$, respectively. a Trajectory of $q_{tm}$ (blue), $q_t$ (red) and $\tilde{q}_t$ (yellow); b difference of $q_{tm} - q_t$ (red) and $q_{tm} - \tilde{q}_t$ (yellow) (Color figure online)

\begin{align*}
\text{d}q_{tm} &= \frac{1}{m} p_t^m \, dt, \\
\text{d}p_t^m &= \left(-\frac{1}{m} \gamma(q_t^m) p_t^m + F(q_t^m)\right) \, dt + \sigma(q_t^m) \, dB_t, \\
\text{and its limiting equation}
\end{align*}

\begin{align*}
\text{d}q_t &= \frac{F(q_t)}{\gamma(q_t)} \, dt + \frac{\sigma(q_t)}{\gamma(q_t)} \, dB_t - \frac{\gamma'(q_t)}{2\gamma^3(q_t)} \sigma^2(q_t) \, dt.
\end{align*}

Thus, we discover that the noise-induced drift term, i.e., the last terms in (6.2) and (6.4) have the same intensity. Nevertheless, in the Brownian case, the noise-induced drift is continuous in time while the noise-induced drift is a jump process in the Lévy case, on account of the discontinuity of Lévy process.

Finally, we numerically simulate the trajectory $q_{tm}$ in (6.1) and $q_t$ in (6.2). We also simulate (6.2) without noise-induced drift term for comparison, that is, we numerically simulate $\tilde{q}_t$ satisfying the following equation
We notice from Fig. 1 that $q_t$ is a better approximation for the position $q^m_t$ than $\tilde{q}_t$. Therefore, the noise-induced drift emergent in the limiting Eq. (6.2) is an essential term which may play a significant role in microscopic systems. In addition, comparing Fig. 1 with Fig. 2, we find that when the mass $m$ is smaller, the limiting equation we obtained is more accurate, which is consistent with our theoretic deduction.

**Acknowledgements** The authors would like to thank Min Dai, Lingyu Feng, Jianyu Hu, Xianming Liu, Wei Wang, Pingyuan Wei, Wei Wei, Shenglan Yuan and Yanjie Zhang for helpful discussions. The authors also want to thank the anonymous reviewers for their valuable comments, which considerably improved the presentation of this paper. This work was partly supported by NSFC Grants 11771449 and 11531006.

**Data Availability** The data that support the findings of this study are available in Github https://github.com/zibowanghust/Levy-Smoluchowski-Kramers-limit

**Appendix**

**A. Non-explosion of the Solution to Dissipative Hamiltonian Equation**

**Lemma A.1** Under Assumption 1–3, there exists a unique non-explosive solution to (2.4) 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 Applebaum 2009) in every bounded cylinder $I \times U(R)$, where $U(R)$ is a ball with radius $R$. Hence, there exists a unique local solution $x^\varepsilon = (x^\varepsilon_t, 0 \leq t < \tau_\infty)$ to (2.4), where $\tau_n$ be the first exit time of $x^\varepsilon_t$ from the ball $B(0, n)$. Then, we will prove that there is no explosion. From the right-continuity of the process $x^\varepsilon_t$, we infer that

$$|x^\varepsilon_{\tau_n}| \geq n. \quad (A.1)$$

Define a function $U^\varepsilon(t, x) = |q|^2 + K^\varepsilon(t, x) + C$. By Itô formula, we have

$$e^{at} U^\varepsilon(t, x^\varepsilon_0) - U^\varepsilon(0, x^\varepsilon_0)$$

$$= \int_0^t a e^{as} U^\varepsilon(s, x^\varepsilon_s) ds + \int_0^t e^{as} \left( 2q_s^\varepsilon + \nabla_q K^\varepsilon(s, x^\varepsilon_s), \nabla_p K^\varepsilon(s, x^\varepsilon_s) \right) ds$$

$$+ \int_0^t e^{as} \left( \nabla_p K^\varepsilon(s, x^\varepsilon_s), -\gamma(s, x^\varepsilon_s) \nabla_p K^\varepsilon(s, x^\varepsilon_s) - \nabla_q K^\varepsilon(s, x^\varepsilon_s) + F(s, x^\varepsilon_s) \right) ds$$

$$+ \int_{|x| \geq 1} e^{as} [K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) + \sigma(s, x^\varepsilon_{s-}) x] - K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) ] N(ds, ds)$$

$$+ \int_{|x| < 1} e^{as} [K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) + \sigma(s, x^\varepsilon_{s-}) x] - K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) ] \tilde{N}(ds, ds)$$

$$+ \int_{|x| < 1} e^{as} [K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) + \sigma(s, x^\varepsilon_{s-}) x] - K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-})$$

$$- (\nabla_p K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}), \sigma(s, x^\varepsilon_{s-}) x)\nu(dx) ds. \quad (A.2)$$
Taking expectation on both sides and applying mean valued theorem, there exists some \( \xi \in [0, 1] \) such that

\[
\mathbb{E} \left[ e^{at} U^\varepsilon(t, x^\varepsilon) \right] - \mathbb{E} \left[ U^\varepsilon(0, x^\varepsilon_0) \right] = \mathbb{E} \left[ \int_0^t a e^{as} U^\varepsilon(s, x^\varepsilon) \, ds + \int_0^t e^{as} \left( 2q^\varepsilon + \nabla_p K^\varepsilon(s, x^\varepsilon) \right) \, ds \right] + \mathbb{E} \left[ \int_0^t \int_{|x| \geq 1} e^{as} \mathbb{E} \left[ \left( \nabla_p K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) - \gamma(s, x^\varepsilon) \nabla_p K^\varepsilon(s, x^\varepsilon) + F(s, x^\varepsilon) \right) \right] \nu(dx) \, ds \right]
\]

\[
+ \mathbb{E} \left[ \int_0^t \int_{|x| < 1} e^{as} \mathbb{E} \left[ \left( \sigma(s, x^\varepsilon_{s-}) x, \nabla_p K^\varepsilon(s, q^\varepsilon_{s-}, p^\varepsilon_{s-}) \right) \right] \nu(dx) \, ds \right].
\]

(A.3)

According to Assumption 1 and using Lemma 3.1 with \( \theta = 0 \), we have

\[
\mathbb{E} \left[ e^{at} U^\varepsilon(t, x^\varepsilon) \right] - \mathbb{E} \left[ U^\varepsilon(0, x^\varepsilon_0) \right] \leq \mathbb{E} \left[ \int_0^t a e^{as} U^\varepsilon(s, x^\varepsilon) \, ds + \int_0^t e^{as} \left( \frac{|q^\varepsilon|^2}{\lambda} + \lambda |\nabla_p K^\varepsilon(s, x^\varepsilon)|^2 \right) \, ds \right]
\]

\[
- \lambda |\nabla_p K^\varepsilon(s, x^\varepsilon)|^2 + ||F||_\infty |\nabla_p K^\varepsilon(s, x^\varepsilon)| \, ds \]

\[
+ \int_0^t \int_{|x| < 1} e^{as} \frac{1}{\sqrt{\varepsilon}} \mathbb{E} \left[ \left( K^\varepsilon(s, x^\varepsilon) + \frac{\epsilon \sigma(s, x^\varepsilon) x}{\sqrt{\varepsilon}} \right)^{\eta-1} + 1 \right] |\sigma(s, x^\varepsilon)| \, ds
\]

\[
+ \int_0^t \int_{|x| < 1} e^{as} \frac{1}{\sqrt{\varepsilon}} \mathbb{E} \left[ \left( K^\varepsilon(s, x^\varepsilon) + \frac{\epsilon \sigma(s, x^\varepsilon) x}{\sqrt{\varepsilon}} \right)^{\eta-2} + 1 \right] |\sigma(s, x^\varepsilon)|^2 \, ds
\]

\[
\leq \int_0^t a e^{as} \mathbb{E} \left[ \int_{|x| \geq 1} |x|^{|F||_\infty + ||\sigma||_\infty |x|^{|F||_\infty + ||\sigma||_\infty \int_{|x| \geq 1} |x|^2 \nu(dx) + \frac{||\sigma||_2^2}{\varepsilon} \int_{|x| \geq 1} |x|^2 \nu(dx) \right] \, ds
\]

\[
\mathbb{E} \left[ K^\varepsilon(s, x^\varepsilon) \right] \, ds
\]

\[
+ \int_0^t C e^{as} \left( \frac{||F||_\infty + ||\sigma||_\infty}{\sqrt{\varepsilon}} \int_{|x| \geq 1} |x|^2 \nu(dx) \right)
\]

\[
+ \frac{||\sigma||_2^2}{\varepsilon} \int_{|x| < 1} |x|^2 \nu(dx) \right) \, ds.
\]

(A.4)

For fixed \( \varepsilon \), we can choose a negative constant \( a \) and a positive constant \( C' \) such that \( \mathbb{E}[e^{at} U^\varepsilon(t, x^\varepsilon)] - \mathbb{E}[U^\varepsilon(0, x^\varepsilon_0)] \leq 0 \). Moreover, we have

\[
\mathbb{E} \left[ U^\varepsilon(t \wedge \tau_n, x^\varepsilon_{\tau_n \wedge t}) \right] = \mathbb{E} \left[ U^\varepsilon(\tau_n \wedge t, x^\varepsilon_{\tau_n \wedge t} 1_{\tau_n < t}) \right] + \mathbb{E} \left[ U^\varepsilon(\tau_n \wedge t, x^\varepsilon_{\tau_n \wedge t} 1_{\tau_n \geq t}) \right]
\]

\[
= \mathbb{E} \left[ U^\varepsilon(\tau_n, x^\varepsilon_{\tau_n}) 1_{\tau_n < t} \right] + \mathbb{E} \left[ U^\varepsilon(t, x^\varepsilon_{\tau_n}) 1_{\tau_n \geq t} \right].
\]
\[ \geq \inf_{x \geq n, t \geq 0} U^\varepsilon(t, x) \mathbb{P}(\tau_n < t). \]  

(A.5)

Hence, we obtain
\[ \mathbb{P}(\tau_n < t) \leq \frac{\mathbb{E} U^\varepsilon(t, x^\varepsilon_n)}{\inf_{x \geq n, t \geq 0} U^\varepsilon(t, x)} \leq \frac{e^{-at} U^\varepsilon(0, x_0^\varepsilon)}{\inf_{x \geq n, t \geq 0} U^\varepsilon(t, x)}. \]  

(A.6)

Therefore, we have
\[ \lim_{n \to \infty} \mathbb{P}(\tau_n < t) = 0 \text{ for all } t. \]  

(A.7)

That is the desired assertion, as required.

\[ \square \]

**B. Non-explosion of the Solution to the Limiting Equation**

Using the notations in Sect. 5, the limiting Eq. (5.2) can be written as
\[ dq_t = \tilde{F}(t, q_t, 0)dt + \tilde{\sigma}(t, q_t, 0)dL_t + \tilde{S}^{kl}(t, q_t, 0) \int_{|x| < 1} x_k x_l N(dt, dx). \]  

(B.1)

Since \( \int_{|x| < 1} |x|^2 v(dx) < \infty \), the last integral of Eq. (B.1) can be decomposed as a compensated Poisson integral and an integral against jump measure \( v \). Then, (B.1) can be written as
\[ dq_t = \tilde{F}(t, q_t)dt + \tilde{S}^{kl}(t, q_t) \int_{|x| < 1} x_k x_l v(dx)dt 
\]
\[ + \int_{|x| \geq 1} [\tilde{\sigma}(t, q_t)x + \tilde{S}^{kl}(t, q_t)x_k x_l] N(dt, dx) \]
\[ + \int_{|x| < 1} [\tilde{\sigma}(t, q_t)x + \tilde{S}^{kl}(t, q_t)x_k x_l] \tilde{N}(dt, dx), \]  

(B.2)

where we omit zero \( p \) value in the coefficients for convenience. We now state a similar result for non-explosion of the limiting Eq. (B.1) as Albeverio et al. (2010).

**Lemma B.1** Assume that the coefficients \( \tilde{F}, \tilde{\sigma}, \tilde{S} \) are locally Lipschitz with linear growth, i.e.,

(i) For every \( R > 0 \), there exists a constant \( L > 0 \) such that
\[ |\tilde{F}(t, q_1) - \tilde{F}(t, q_2)|^2 + ||\tilde{\sigma}(t, q_1) - \tilde{\sigma}(t, q_2)||^2 + ||\tilde{S}(t, q_1) - \tilde{S}(t, q_2)||^2 \leq L|q_1 - q_2|^2, |q_1|, |q_2| \leq R. \]

(ii) There exists a constant \( C > 0 \) such that
\[ |\tilde{F}(t, q)|^2 + ||\tilde{\sigma}(t, q)||^2 + ||\tilde{S}(t, q)||^2 \leq C(1 + |q|^2). \]

Then, there exists a unique global solution \( q_t \) to (B.1).
Proof Using the same argument as in the proof of Lemma A.1, there exists a unique local solution $q = (q_t, 0 \leq t < \tau_\infty)$ to (B.1), where $\tau_n$ is the first exit time of $q_t$ from the ball $B(0,n)$. Utilizing Itô formula with $\varphi(q) = |q|^2$, we have

$$
|q_{t \wedge \tau_n}|^2 = |q_0|^2 + 2 \int_0^{t \wedge \tau_n} \langle \bar{F}(s, q_s) + \bar{S}^{kl}(s, q_s) \int_{|x|<1} x_k x_l \nu(dx), q_s \rangle ds \\
+ \int_0^{t \wedge \tau_n} \int_{|x|<1} \left[ |q_{s^-} + \bar{\sigma}(s, q_{s^-}) x + \bar{S}^{kl}(s, q_{s^-}) x_k x_l|^2 - |q_{s^-}|^2 \right] \bar{N}(ds, dx) \\
+ \int_0^{t \wedge \tau_n} \int_{|x| \geq 1} \left[ |q_{s^-} + \bar{\sigma}(s, q_{s^-}) x + \bar{S}^{kl}(s, q_{s^-}) x_k x_l|^2 - |q_{s^-}|^2 \right] N(ds, dx) \\
+ \int_0^{t \wedge \tau_n} \int_{|x|<1} \left| \bar{\sigma}(s, q_{s^-}) x + \bar{S}^{kl}(s, q_{s^-}) x_k x_l \right|^2 \nu(dx) ds.
$$

(B.3)

Taking expectation on both sides and using Cauchy–Schwarz inequality, we have

$$
\mathbb{E}[|q_{t \wedge \tau_n}|^2] \\
\leq \mathbb{E}[|q_0|^2] + \mathbb{E} \left[ \int_0^{t \wedge \tau_n} |q_s|^2 + \left| \bar{F}(s, q_s) + \bar{S}^{kl}(s, q_s) \int_{|x|<1} x_k x_l \nu(dx) \right|^2 ds \right] \\
+ \mathbb{E} \left[ \int_0^{t \wedge \tau_n} \int_{|x| \geq 1} \left| q_{s^-} \right|^2 + 2 \left| \bar{\sigma}(s, q_{s^-}) x + \bar{S}^{kl}(s, q_{s^-}) x_k x_l \right|^2 \nu(dx) ds \right] \\
+ \mathbb{E} \left[ \int_0^{t \wedge \tau_n} \int_{|x|<1} \left| \bar{\sigma}(s, q_{s^-}) x + \bar{S}^{kl}(s, q_{s^-}) x_k x_l \right|^2 \nu(dx) ds \right] \\
\leq \mathbb{E}[|q_0|^2] + \tilde{C} \int_0^{t \wedge \tau_n} 1 + \mathbb{E}[|q_{t \wedge \tau_n}|^2] ds,
$$

(B.4)

where we use the linear growth condition in the second inequality. Then, by Gronwall inequality, we obtain that

$$
\mathbb{E}[|q_{t \wedge \tau_n}|^2] \leq \left( 1 + \mathbb{E}[|q_0|^2] \right) e^{\tilde{C} t} - 1.
$$

(B.5)

On the other hand, we have

$$
\mathbb{E}[|q_{t \wedge \tau_n}|^2] = \mathbb{E}[|q_{t \wedge \tau_n}|^2 1_{\{\tau_n < t\}}] + \mathbb{E}[|q_{t \wedge \tau_n}|^2 1_{\{\tau_n \geq t\}}] \geq \mathbb{E}[|q_{t \wedge \tau_n}|^2 1_{\{\tau_n < t\}}] \geq n^2 \mathbb{P}(\tau_n < t).
$$

(B.6)

Therefore, for all $n$,

$$
\mathbb{P}(\tau_n < t) \leq n^{-2} \mathbb{E}[|p_{t \wedge \tau_n}|^2] \leq n^{-2} \left( 1 + \mathbb{E}[|q_0|^2] \right) e^{\tilde{C} t}.
$$

(B.7)

Hence, $\lim_{n \to \infty} \mathbb{P}(\tau_n < t) = 0$ for all $t$. This finishes the proof.  

\(\square\)
References

Albeverio, S., Brzeźniak, Z., Wu, J.-L.: Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients. J. Math. Anal. Appl. 371(1), 309–322 (2010)

Al-Talibi, H., Hilbert, A., Kolokoltsov, V.: Nelson-type limit for a particular class of Lévy processes. In: AIP Conference Proceedings, vol. 1232, pp. 189–193. American Institute of Physics (2010)

Applebaum, D.: Lévy Processes and Stochastic Calculus. Cambridge University Press (2009)

Arnol’d, V.I.: Mathematical Methods of Classical Mechanics, vol. 60. Springer (2013)

Billingsley, P.: Convergence of Probability Measures. Wiley (2013)

Birrell, J., Wehr, J.: Homogenization of dissipative, noisy, Hamiltonian dynamics. Stoch. Process. Appl. 128(7), 2367–2403 (2018)

Birrell, J., Wehr, J.: A homogenization theorem for Langevin systems with an application to Hamiltonian dynamics. In: Sojourns in Probability Theory and Statistical Physics-I, pp. 89–122. Springer (2019)

Birrell, J., Hottovy, S., Volpe, G., Wehr, J.: Small mass limit of a Langevin equation on a manifold. In: Annales Henri Poincaré, vol. 18, pp. 707–755. Springer (2017)

Cerrai, S., Freidlin, M.: On the Smoluchowski–Kramers approximation for a system with an infinite number of degrees of freedom. Probab. Theory Relat. Fields 135(3), 363–394 (2006)

Doering, C.: Modeling complex systems: stochastic processes, stochastic differential equations, and Fokker-Planck equations. In: 1990 Lectures In Complex Systems, pp. 3–51. Addison-Wesley (1990)

Duan, J.: An Introduction to Stochastic Dynamics, vol. 51. Cambridge University Press (2015)

Dybicz, B., Kleczkowski, A., Gilligan, C.A.: Modelling control of epidemics spreading by long-range interactions. J. R. Soc. Interface 6(39), 941–950 (2009)

Dybicz, B., Parrondo, J.M., Gudowska-Nowak, E.: Fluctuation-dissipation relations under Lévy noises. EPL (Europhys. Lett.) 98(5), 50006 (2012)

Freidlin, M.: Some remarks on the Smoluchowski–Kramers approximation. J. Stat. Phys. 117(3), 617–634 (2004)

Hanggi, P.: Nonlinear fluctuations: the problem of deterministic limit and reconstruction of stochastic dynamics. Phys. Rev. A 25(2), 1130 (1982)

Hottovy, S., McDaniel, A., Volpe, G., Wehr, J.: The Smoluchowski–Kramers limit of stochastic differential equations with arbitrary state-dependent friction. Commun. Math. Phys. 336(3), 1259–1283 (2015)

Jacod, J., Shiryaev, A.: Limit Theorems for Stochastic Processes, vol. 288. Springer (2013)

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

Kurtz, T.G., Protter, P.: Weak limit theorems for stochastic integrals and stochastic differential equations. Ann. Probab. 19, 1035–1070 (1991)

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

Nelson, E.: Dynamical Theories of Brownian Motion, vol. 106. Princeton University Press (2020)

Ortega, J.M.: Matrix Theory: A Second Course. Springer (2013)

Pavliotis, G., Stuart, A.: Multiscale Methods: Averaging and Homogenization. Springer (2008)

Sánchez, R., Newman, D.E., Lebœuf, J.-N., Decyk, V., Carreras, B.A.: Nature of transport across sheared zonal flows in electrostatic ion-temperature-gradient gyrokinetic plasma turbulence. Phys. Rev. Lett. 101(20), 205002 (2008)

Smoluchowski, M.: Drei vortrage über diffusion, brownsche bewegung und koagulation von kolloidteilchen. Z. Angew. Phys. 17, 557–585 (1916)

Volpe, G., Helden, L., Brettscheider, T., Wehr, J., Bechinger, C.: Influence of noise on force measurements. Phys. Rev. Lett. 104(17), 170602 (2010)

Wei, P., Chao, Y., Duan, J.: Hamiltonian systems with Lévy noise: symplecticity, Hamilton’s principle and averaging principle. Physica D 398, 69–83 (2019)

Wu, L.: Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. Stoch. Process. Appl. 91(2), 205–238 (2001)

Xu, Y., Li, Y., Zhang, H., Li, X., Kurths, J.: The switch in a genetic toggle system with Lévy noise. Sci. Rep. 6(1), 1–11 (2016)

Yuan, S., Blömker, D.: Modulation and amplitude equations on bounded domains for nonlinear spdes driven by cylindrical α-stable Lévy processes. SIAM J. Appl. Dyn. Syst. 21, 1748–1777 (2022)
Zhang, S.: Smoluchowski-Kramers approximation for stochastic equations with Lévy-noise. Ph.D. thesis, Purdue University (2008)
Zhang, Q., Duan, J.: Linear response theory for nonlinear stochastic differential equations with $\alpha$-stable Lévy noises. J. Stat. Phys. 182(2), 1–28 (2021)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.