Journal of Statistical Physics (2024) 191:22
https://doi.org/10.1007/s10955-024-03239-2

Averaging on Macroscopic Scales with Application to Smoluchowski–Kramers Approximation

Mengmeng Wang[1] · Dong Su[1] · Wei Wang[1]

Received: 26 September 2023 / Accepted: 22 January 2024 / Published online: 17 February 2024

Abstract

This paper develops an averaging approach on macroscopic scales to derive Smoluchowski–Kramers approximation for a Langevin equation with state dependent friction in d-dimensional space. In this approach we couple the microscopic dynamics to the macroscopic scales. The weak convergence rate is also presented.

Keywords Smoluchowski–Kramers approximation · State-dependent friction · Lyapunov equation · Averaging

1 Introduction

The Smoluchowski–Kramers (SK) approximation is useful to describe the motion of a particle with small mass which has been studied in lots of works beginning with Smoluchowski [20] and Kramers [17]. The motion of a particle with mass $0 < \epsilon \ll 1$ in $\mathbb{R}^d$ $(d \geq 1)$ is described by the following Langevin equation

$$\epsilon \ddot{x}_t^\epsilon + \alpha \dot{x}_t^\epsilon = F(x_t^\epsilon) + \sigma(x_t^\epsilon)\dot{B}_t\,, \quad x^\epsilon(0) = x_0\,, \quad \dot{x}^\epsilon(0) = v_0,$$

where constant friction $\alpha > 0$, $F(x) : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma(x) : \mathbb{R}^d \to \mathbb{R}^{d\times k}$ and $\{B_t\}$ is k-dimensional standard Wiener process. The classical SK approximation states that for every $T > 0$

$$\lim_{\epsilon\to 0} \mathbb{E} \sup_{0\leq t\leq T} \|x_t^\epsilon - x_t\|_{\mathbb{R}^d} = 0,$$

Communicated by Stefano Olla.

This work is supported in part by NSFC Grants No. 11771207 and 12371243.

✉ Mengmeng Wang
wmm3540@gmail.com

Dong Su
dsu224466@163.com

Wei Wang
wangweinju@nju.edu.cn

1 Department of Mathematics, Nanjing University, Nanjing 210093, People's Republic of China

with

$$\alpha \dot{x}_t = F(x_t) + \sigma(x_t)\dot{B}_t\,, \quad x(0) = x_0\,.$$

For more detail one can refer to [8]. The above limit equation, with just letting $\epsilon = 0$, is not surprising for such constant friction α. However a noise-induced drift was observed in experiment [21] for the case of state dependent friction, which implies the limit equation can not be obtained by letting $\epsilon = 0$. Recent work by Hottovy et al. [14] presented a mathematical explanation, but lack of some intuition, by a theory of the convergence of stochastic integral with respect to semimartingale, for such experimental observation.

In this paper we present a new approach which makes the limit equation more intuitively, although in a weak sense. We consider the following Langevin equation with state dependent friction,

$$\epsilon \ddot{x}_t^\epsilon + \alpha(x_t^\epsilon)\dot{x}_t^\epsilon = F(x_t^\epsilon) + \sigma(x_t^\epsilon)\dot{B}_t, \tag{1.1}$$

$$x_0^\epsilon = x_0 \quad , \dot{x}_0^\epsilon = v_0, \quad x_0, v_0 \in \mathbb{R}^d, \tag{1.2}$$

where $\alpha(x) : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is a $d \times d$ invertible matrix-valued function. Our idea is to consider the limit of ρ_t^ϵ, the law of x_t^ϵ, as $\epsilon \to 0$. For this we first write out the equations solved by ρ_t^ϵ (see (1.6)–(1.7)). However, these equations are not closed, we couple the equations (1.1)–(1.2) to (1.6)–(1.7). Then we pass the limit $\epsilon \to 0$ in equations (1.6)–(1.7) via an averaging approach.

Typically, write the equation (1.1) into the following equivalent form

$$\dot{x}_t^\epsilon = v_t^\epsilon, \tag{1.3}$$

$$\epsilon \dot{v}_t^\epsilon = -\alpha(x_t^\epsilon)v_t^\epsilon + F(x_t^\epsilon) + \sigma(x_t^\epsilon)\dot{B}_t. \tag{1.4}$$

First it is known that the law $f_t^\epsilon \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$, the set consisting of all probability measures on $\mathbb{R}^d \times \mathbb{R}^d$, of $(x_t^\epsilon, \dot{x}_t^\epsilon)$ satisfies the Fokker–Planck equation

$$\partial_t f_t^\epsilon + v \cdot \nabla_x f_t^\epsilon - \frac{1}{\epsilon}\nabla_v \cdot (\alpha(x)vf_t^\epsilon - F(x)f_t^\epsilon) = \frac{1}{\epsilon^2}\sum_{i=1}^d\sum_{j=1}^d \partial_{v_i}\partial_{v_j}\left(a_{ij}(x)f_t^\epsilon\right) \tag{1.5}$$

in the weak sense, where $a(x) = \sigma(x)\sigma^\top(x)$ and $\sigma^\top(x)$ is the transpose of $\sigma(x)$, that is for $\varphi \in C_0^\infty(\mathbb{R}^d \times \mathbb{R}^d)$,

$$\begin{aligned}
&\langle f_t^\epsilon, \varphi\rangle - \langle f_0^\epsilon, \varphi\rangle \\
&= \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big[v \cdot \nabla_x \varphi \\
&\quad - \frac{1}{\epsilon}(\alpha(x)v - F(x)) \cdot \nabla_v \varphi + \frac{1}{\epsilon^2}\sum_{i=1}^d\sum_{j=1}^d a_{ij}(x)\partial_{v_i}\partial_{v_j}\varphi\Big] f_s^\epsilon(dx, dv)ds.
\end{aligned}$$

The law of x_t^ϵ is

$$\rho_t^\epsilon(x) = \int_{\mathbb{R}^d} f_t^\epsilon(x, v)dv,$$

and define

$$Y_t^\epsilon(x) = \int_{\mathbb{R}^d} v f_t^\epsilon(x, v)dv.$$

Integrating both sides of the equation (1.5) with respect to v, and multiplying both sides of (1.5) by v, then integrating with respect to v, we get

$$\partial_t \rho_t^\epsilon(x) = -\nabla_x \cdot Y_t^\epsilon(x), \tag{1.6}$$

$$\partial_t Y_t^\epsilon(x) = -\left[\nabla_x \cdot \left(\int v \otimes v f_t^\epsilon dv\right)\right]^\top - \frac{1}{\epsilon}\alpha(x) Y_t^\epsilon(x) + \frac{1}{\epsilon} F(x)\rho_t^\epsilon(x). \tag{1.7}$$

Notice that

$$\int v \otimes v f_t^\epsilon dv = \rho_t^\epsilon(x) \int v \otimes v \frac{f_t^\epsilon}{\rho_t^\epsilon(x)} dv = \rho_t^\epsilon(x) \mathbb{E}^x(v_t^\epsilon \otimes v_t^\epsilon),$$

providing $\rho_t^\epsilon \neq 0$. Here $\mathbb{E}^x(v_t^\epsilon \otimes v_t^\epsilon)$ is the expectation with fixing $x^\epsilon = x$ in equation (1.4). Thus, we obtain the following closed system

$$\begin{cases} \partial_t \rho_t^\epsilon(x) = -\nabla_x \cdot Y_t^\epsilon(x), \\ \partial_t Y_t^\epsilon(x) = -\frac{1}{\epsilon}\alpha(x) Y_t^\epsilon(x) + \frac{1}{\epsilon} F(x)\rho_t^\epsilon(x) - \left[\nabla_x \cdot (\rho_t^\epsilon(x)\mathbb{E}^x(v_t^\epsilon \otimes v_t^\epsilon))\right]^\top, \\ \dot{x}_t^\epsilon = v_t^\epsilon, \\ \epsilon v_t^\epsilon = -\alpha(x_t^\epsilon) v_t^\epsilon + F(x_t^\epsilon) + \sigma(x_t^\epsilon)\dot{B}_t. \end{cases} \tag{1.8}$$

The above equations (1.8) is in the form of a slow-fast system with slow component $(\rho_t^\epsilon, x_t^\epsilon)$ and fast component $(Y_t^\epsilon, v_t^\epsilon)$. So an averaging method is applicable to pass the limit $\epsilon \to 0$ [7, 13, 18, 19, e.g.]. In fact, by the idea of averaging approach [7, Chapter 5], fixing ρ_t^ϵ to a probability measure ρ (see equation (4.1) and Remark 4.1), we have Y_t^ϵ, as $\epsilon \to 0$, converges weakly to (see Lemma 4.1)

$$Y^{*,\rho}(x) = \alpha^{-1}(x)F(x)\rho(x) - \alpha^{-1}(x)\left[\nabla_x \cdot (\rho(x)J(x))\right]^\top.$$

Then we formally derive the limit equation (2.4) by replacing Y_t^ϵ by $Y^{*,\rho}$ in the first equation of (1.8). We call the above an averaging principle on macroscopic scale.

There are a lot of literature about SK approximation in case of variable friction. Freidlin and Hu [9] considered the SK approximation for (1.1) by regularizing the noise. Freidlin, Hu and Wentzell [10], also by regularization method, considered the SK approximation with some degenerating friction. There are also some works on SK approximation of infinite dimensional system with constant damping [3] and state–dependent damping [5, 6, e.g.] and some related problem, large deviation e.g. [4].

The rest of this paper is organized as follows. In Sect. 2, we give some preliminaries, assumptions and the main result. The tightness of $\{\rho_t^\epsilon\}$ is shown in Sect. 3, then the averaging procedure is implemented in the last section. It should be clarified that the positive constant C and C_T may be different from line to line in the proofs.

2 Preliminaries and Main Result

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, and $\mathbb{E}$ denote the expectation with respect to $\mathbb{P}$. Denote by $|\cdot|$ the norm on $\mathbb{R}^d$ and $\langle\cdot,\cdot\rangle$ the inner product in space $L^2(\mathbb{R}^d)$.

We make the following assumptions.

(H$_1$) $\alpha(x) : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuous differentiable function. The smallest eigenvalue $\lambda_1(x)$ of $\frac{1}{2}(\alpha + \alpha^\top)$ is positive uniformly with respect to x, i.e. for some constant $C_{\lambda_\alpha} > 0$,

$$\lambda_1(x) \geq C_{\lambda_\alpha} > 0. \tag{2.1}$$

($\mathbf{H_2}$) $F(x)$ and $\sigma(x)$ are continuous differentiable and Lipschitz functions with Lipschitz constant C_F and C_σ respectively, i.e., for $x, y \in \mathbb{R}^d$,

$$|F(x) - F(y)| \leq C_F|x - y|, \tag{2.2}$$

$$|\sigma(x) - \sigma(y)| \leq C_\sigma|x - y|. \tag{2.3}$$

($\mathbf{H_3}$) There is a constant $C > 0$ such that $|\partial_{x_k}\alpha_{ij}(x)| \leq C$, for all $1 \leq i, j, k \leq d, x \in \mathbb{R}^d$.

Remark 2.1 Since the global Lipschitz condition implies linear growth, from (2.2), we have $|F(x)| \leq C_F(1 + |x|)$. Here we keep the same notation C_F for simplicity. In the following, we also use $|F(x)| \leq C_F\sqrt{1 + |x|^2}$. A similar bound holds for $\sigma(x)$.

Hottovy et al. [14] assumed that the solutions are tight, that they just needed some continuity property of F and σ, to pass the limit $\epsilon \to 0$. Here we pose the Lipschitz assumption on F and σ to show the tightness of the solutions.

Next we present our main result.

Theorem 2.1 *Under the assumptions of* ($\mathbf{H_1}$)–($\mathbf{H_3}$)*, for every* $t > 0$*,* ρ_t^ϵ*, the solution to equation* (1.8)*, converges weakly to* ρ_t *solving the following equation in weak sense*

$$\partial_t\rho_t(x) = -\nabla_x \cdot (\alpha^{-1}(x)F(x)\rho_t(x) + \alpha^{-1}(x)(\nabla_x \cdot (\rho_t(x)J(x)))^\top), \tag{2.4}$$

which corresponds to the following stochastic differential equation (SDE)

$$\dot{x}_t = \alpha^{-1}(x_t)F(x_t) + S(x_t) + \alpha^{-1}(x_t)\sigma(x_t)\dot{B}_t\,. \tag{2.5}$$

Here

$$S_i(x) = \frac{\partial}{\partial x_k}(\alpha^{-1}(x))_{ij}J_{jk}(x).$$

and $J(x)$ *is the solution of the Lyapunov equation*

$$J(x)\alpha^\top(x) + \alpha(x)J(x) = \sigma(x)\sigma^\top(x). \tag{2.6}$$

Furthermore, there is a constant $C_T > 0$ *such that for every* $t \in (0, T)$ *and* $\psi \in C_0^\infty(\mathbb{R}^d)$

$$|\langle\rho_t^\epsilon(x) - \rho_t(x), \psi\rangle| \leq \epsilon C_T\|\nabla\psi\|_{Lip}, \tag{2.7}$$

where $\|\cdot\|_{Lip}$ *denotes the Lipschitz norm defined by*

$$\|f\|_{Lip} = \|f\|_\infty + \sup_{x\neq y}\frac{|f(x) - f(y)|}{|x - y|}.$$

Remark 2.2 The above convergence rate in (2.7), from the estimate (4.11) in the proof, is sharp. So an interesting problem is the higher order correction to ρ_t^ϵ, that is what is the limit of

$$\frac{1}{\epsilon}(\rho_t^\epsilon - \rho_t)$$

as $\epsilon \to 0$. To determine the limit we have to give a more detail estimation than that in Lemma 4.1. This will be considered in our future work.

Remark 2.3 The unique solution to equation (2.6) has the following explicit expression [2, Page 179]

$$J(x)=\int_0^{\infty} e^{-\alpha(x)t}\sigma(x)\sigma^{\top}(x)e^{-\alpha^{\top}(x)t}dt.$$

In fact the matrix $J(x)$ is the limit, as $\epsilon\to 0$, of the covariance of $\sqrt{\epsilon}v_t^{\epsilon}$ with freezing $x^{\epsilon}=x$ (Lemma 3.3).

Remark 2.4 To see the relationship between (2.4) and (2.5), by Einstein summation notation, write (2.4) as

$$\begin{aligned}\partial_t\rho_t(x)&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x))+\partial_{x_i}(\alpha_{ik}^{-1}\partial_{x_j}(\rho_t(x)J_{jk}))\\&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x))+\partial_{x_i}(\alpha_{ik}^{-1}(\partial_{x_j}\rho_t(x)J_{jk}+\rho_t(x)\partial_{x_j}J_{jk}))\\&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x))+\partial_{x_i}(\partial_{x_j}\rho_t(x)\alpha_{ik}^{-1}J_{jk}+\rho_t(x)\alpha_{ik}^{-1}\partial_{x_j}J_{jk})\\&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x))+\partial_{x_i}(\partial_{x_j}(\rho_t(x)\alpha_{ik}^{-1}J_{jk})-\rho_t(x)J_{jk}\partial_{x_j}\alpha_{ik}^{-1}).\end{aligned}$$

From (2.6), we have

$$\alpha^{-1}(x)J(x)+J(x)[\alpha^{-1}(x)]^{\top}=\alpha^{-1}(x)\sigma(x)\sigma^{\top}(x)[\alpha^{-1}(x)]^{\top}. \tag{2.8}$$

Denote the right hand side of (2.8) by A, and extract the (i,j) element of both sides,

$$\alpha_{ik}^{-1}J_{kj}+J_{ik}\alpha_{jk}^{-1}=A_{ij}.$$

By the symmetry of $J(x)$, we get

$$\begin{aligned}\partial_t\rho_t(x)&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x))+\partial_{x_i}(\frac{1}{2}\partial_{x_j}(A_{ij}\rho_t(x))-\rho_t(x)J_{jk}\partial_{x_j}\alpha_{ik}^{-1})\\&=-\nabla_x\cdot(\alpha^{-1}(x)F(x)\rho_t(x)+S(x_t)\rho_t(x))+\frac{1}{2}\partial_{x_i}\partial_{x_j}(A_{ij}\rho_t(x)),\end{aligned}$$

which corresponds to SDE (2.5).

To prove Theorem 2.1, we first show the tightness of $\{x_t^{\epsilon}\}$ in Sect. 3, then for all sequences of $\{\rho_\cdot^{\epsilon}\}$ there exits a subsequence $\{\rho_\cdot^{\epsilon_k}\}$ converges weakly to $\{\rho_\cdot\}$ as $\epsilon_k\to 0$. Then, for the convergent subsequence $\rho_\cdot^{\epsilon_k}$, we apply the averaging approach (Sect. 4) to the slow-fast system (1.8) to derive the limit equation.

The following lemma is used to give an explicit representation for the covariance of $\sqrt{\epsilon}v_t^{\epsilon}$ with freezing $x^{\epsilon}=x$ in Lemma 3.3.

Lemma 2.1 [1, Theorem 2] *Let $I\subseteq\mathbb{R}$ be an open interval with $t_0\in I$, $A\in\mathbb{C}^{n\times n}$, $B\in\mathbb{C}^{m\times m}$, $C\in\mathcal{C}(I,\mathbb{C}^{n\times m})$ and $D\in\mathbb{C}^{m\times n}$. The Lyapunov differential equation*

$$\dot{X}(t)=AX(t)+X(t)B+C(t),\ \ X(t_0)=D,$$

has the unique solution

$$X(t)=e^{A(t-t_0)}De^{B(t-t_0)}+\int_{t_0}^{t}e^{A(t-s)}C(s)e^{B(t-s)}\mathrm{d}s.$$

The following lemma is important in the averaging approach in Sect. 4.

Lemma 2.2 [15, p. 120] *Let* $\mathbf{A} = (a_{ij})_{1\le i,j\le d}$ *and* $\mathbf{u} = (u_i)_{1\le i\le d}$ *be* $d \times d$ *matrix and* $d \times 1$ *vector respectively. Each element of* $\mathbf{A}$ *and* $\mathbf{u}$ *is a function of* $\mathbf{x} = (x_1, x_2, \dots, x_d)$*, then*

$$\nabla \cdot (\mathbf{A}\mathbf{u}) = (\nabla \cdot \mathbf{A})\mathbf{u} + \mathrm{tr}(\mathbf{A}\ \mathrm{grad}\ \mathbf{u}),$$

where grad $\mathbf{u} = \left(\frac{\partial u_i}{\partial x_j}\right)_{1\le i,j\le d}$ *and* $(\nabla \cdot \mathbf{A})_j = \nabla \cdot \mathbf{A}_j = \sum_{i=1}^{d} \frac{\partial a_{ij}}{\partial x_i}$ *where* $\mathbf{A}_j$ *is the* j*-th column of* $\mathbf{A}$.

3 Tightness of $\{x^\epsilon\}_\epsilon$

To prove the tightness of $\{x^\epsilon\}_\epsilon$ in space $C(0, T; \mathbb{R}^d)$, we need to show the boundedness in $C(0, T; \mathbb{R}^d)$ and the Hölder continuity of $\{x^\epsilon\}_\epsilon$.

Lemma 3.1 *Under assumptions* $(\mathbf{H_1})$ *and* $(\mathbf{H_2})$*, for all* $T > 0$*,*

$$\mathbb{E} \sup_{0\le t\le T} |x_t^\epsilon|^2 \le C_T, \tag{3.1}$$

and for $0 \le t_1, t_2 \le T$*,*

$$\mathbb{E}|x_{t_2}^\epsilon - x_{t_1}^\epsilon|^2 \le C|t_2 - t_1|. \tag{3.2}$$

Proof We intend to write an expression of x_t^ϵ in a mild formulation. Due to the state dependent friction, this is of some difficulty. For this we first consider the linear part of the v^ϵ-equation (1.4), that is the following equation

$$\dot{y}_t = -\frac{1}{\epsilon}\alpha(x_t^\epsilon)y_t,$$
$$y_0 = I_d.$$

Then

$$\begin{aligned}
\frac{d}{dt}(y_t^{-1}\dot{x}_t^\epsilon) &= \frac{d}{dt}y_t^{-1}\dot{x}_t^\epsilon + y_t^{-1}\ddot{x}_t^\epsilon \\
&= -y_t^{-1}\frac{dy_t}{dt}y_t^{-1}\dot{x}_t^\epsilon + y_t^{-1}\ddot{x}_t^\epsilon \\
&= \frac{1}{\epsilon}y_t^{-1}\alpha(x_t^\epsilon)\dot{x}_t^\epsilon + y_t^{-1}\left(-\frac{\alpha(x_t^\epsilon)}{\epsilon}\dot{x}_t^\epsilon + \frac{1}{\epsilon}F(x_t^\epsilon) + \frac{1}{\epsilon}\sigma(x_t^\epsilon)\dot{B}_t\right) \\
&= \frac{1}{\epsilon}y_t^{-1}F(x_t^\epsilon) + \frac{1}{\epsilon}y_t^{-1}\sigma(x_t^\epsilon)\dot{B}_t.
\end{aligned}$$

Integrating from 0 to τ yields

$$y_\tau^{-1}\dot{x}_\tau^\epsilon = v_0 + \frac{1}{\epsilon}\int_0^\tau y_s^{-1}F(x_s^\epsilon)ds + \frac{1}{\epsilon}\int_0^\tau y_s^{-1}\sigma(x_s^\epsilon)dB_s,$$

and

$$\dot{x}_\tau^\epsilon = y_\tau v_0 + \frac{1}{\epsilon}\int_0^\tau y_\tau y_s^{-1}F(x_s^\epsilon)ds + \frac{1}{\epsilon}\int_0^\tau y_\tau y_s^{-1}\sigma(x_s^\epsilon)dB_s. \tag{3.3}$$

Integrating from 0 to t for equation (3.3) yields

$$x_t^\epsilon = x_0 + \int_0^t y_\tau v_0 d\tau$$

$$+\frac{1}{\epsilon}\int_0^t\int_0^\tau y_\tau y_s^{-1}F(x_s^\epsilon)dsd\tau$$
$$+\frac{1}{\epsilon}\int_0^t\int_0^\tau y_\tau y_s^{-1}\sigma(x_s^\epsilon)dB_sd\tau \triangleq x_0+\sum_{i=1}^3 I_i(t).$$

Now define

$$z_\tau = y_\tau y_s^{-1}F(x_s^\epsilon),\ \ \tau\geq s, \tag{3.4}$$

then

$$z_s = F(x_s^\epsilon),$$

and

$$\begin{aligned}\frac{dz_\tau}{d\tau} &= \frac{dy_\tau}{d\tau}y_s^{-1}F(x_s^\epsilon)\\ &= -\frac{1}{\epsilon}\alpha(x_\tau^\epsilon)y_\tau y_s^{-1}F(x_s^\epsilon)\\ &= -\frac{1}{\epsilon}\alpha(x_\tau^\epsilon)z_\tau.\end{aligned}$$

Thus we have [12, Lemma 4.2 of Chpter IV],

$$|z_\tau|\leq |F(x_s^\epsilon)|e^{-\int_s^\tau\frac{1}{\epsilon}\lambda_1(x)dt}\leq |F(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}.$$

Similarly,

$$|y_\tau y_s^{-1}\sigma(x_s^\epsilon)|\leq|\sigma(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)},$$

and

$$|y_\tau v_0|\leq |v_0|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}\tau}.$$

Then we derive

$$\begin{aligned}\mathbb{E}\sup_{0\leq t\leq T}|I_1(t)|^2 &= \mathbb{E}\sup_{0\leq t\leq T}\left|\int_0^t y_\tau v_0 d\tau\right|^2\\ &\leq |v_0|^2\mathbb{E}\sup_{0\leq t\leq T}\left|\int_0^t e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}\tau}d\tau\right|^2\\ &\leq |v_0|^2\frac{\epsilon^2}{C_{\lambda_\alpha}^2}.\end{aligned} \tag{3.5}$$

By Hölder inequality and Fubini theorem

$$\begin{aligned}\mathbb{E}\sup_{0\leq t\leq T}|I_2(t)|^2 &= \mathbb{E}\sup_{0\leq t\leq T}\left|\frac{1}{\epsilon}\int_0^t\int_0^\tau z_\tau dsd\tau\right|^2\\ &\leq \frac{1}{\epsilon^2}\mathbb{E}\sup_{0\leq t\leq T}\left|\int_0^t\int_0^\tau \|F(x_t^\epsilon)\|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}dsd\tau\right|^2\\ &= \frac{1}{\epsilon^2}\mathbb{E}\sup_{0\leq t\leq T}\left|\int_0^t\int_s^t \|F(x_t^\epsilon)\|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}d\tau ds\right|^2\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{E} \sup_{0\leq t\leq T}\left|\int_0^t \|F(x_t^\epsilon)\|\frac{1}{C_{\lambda_\alpha}}\right|^2 ds\\
&\leq \frac{C_T}{C_{\lambda_\alpha}^2}\left(1+\int_0^T \mathbb{E}\sup_{0\leq r\leq s}|x_r^\epsilon|^2 ds\right).
\end{aligned}
\tag{3.6}
$$

By Doob's maximal inequality, Fubini theorem and Hölder inequality, we obtain

$$
\begin{aligned}
\mathbb{E}\sup_{0\leq t\leq T}|I_3(t)|^2 &= \mathbb{E}\sup_{0\leq t\leq T}\left|\frac{1}{\epsilon}\int_0^t\int_0^\tau y_\tau y_s^{-1}\sigma(x_s^\epsilon)dB_s d\tau\right|^2\\
&= \mathbb{E}\sup_{0\leq t\leq T}\left|\frac{1}{\epsilon}\int_0^t\int_s^t y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s\right|^2\\
&\leq \frac{4}{\epsilon^2}\mathbb{E}\left|\int_0^T\int_s^t y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s\right|^2\\
&\leq \frac{4}{\epsilon^2}\mathbb{E}\int_0^T\left(\int_s^t |\sigma(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)d\tau}\right)^2 ds\\
&\leq \frac{C_T}{C_{\lambda_\alpha}^2}\left(1+\int_0^T\mathbb{E}\sup_{0\leq r\leq s}|x_r^\epsilon|^2 ds\right).
\end{aligned}
\tag{3.7}
$$

Combining (3.5), (3.6) with (3.7) yields

$$
\mathbb{E}\sup_{0\leq t\leq T}|x_t^\epsilon|^2 \leq C\left(|x_0|^2+|v_0|^2\frac{\epsilon^2}{C_{\lambda_\alpha}^2}+\frac{C_T}{C_{\lambda_\alpha}^2}\left(1+\int_0^T\mathbb{E}\sup_{0\leq r\leq s}|x_r^\epsilon|^2 ds\right)\right).
$$

Then Gronwall inequality yields

$$
\mathbb{E}\sup_{0\leq t\leq T}|x_t^\epsilon|^2 \leq C\left(|x_0|^2+|v_0|^2\frac{\epsilon^2}{C_{\lambda_\alpha}^2}+\frac{C_T}{C_{\lambda_\alpha}^2}\right)e^{\frac{C_T}{C_{\lambda_\alpha}^2}} \leq C_T.
$$

Next, let $0\leq t_1<t_2\leq T$,

$$
\begin{aligned}
x_{t_2}^\epsilon - x_{t_1}^\epsilon &= \int_{t_1}^{t_2} y_\tau v_0 d\tau + \frac{1}{\epsilon}\int_{t_1}^{t_2}\int_0^\tau y_\tau y_s^{-1}F(x_s^\epsilon)ds d\tau\\
&\quad + \frac{1}{\epsilon}\int_{t_1}^{t_2}\int_0^\tau y_\tau y_s^{-1}\sigma(x_s^\epsilon)dB_s d\tau\\
&\triangleq \sum_{i=1}^3 J_i.
\end{aligned}
$$

First, Hölder inequality yields

$$
\mathbb{E}|J_1|^2 = \mathbb{E}\left|\int_{t_1}^{t_2} y_\tau v_0 d\tau\right|^2 \leq |v_0|^2\mathbb{E}\left|\int_{t_1}^{t_2} e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}\tau}d\tau\right|^2 \leq C|t_2-t_1|^2.
\tag{3.8}
$$

Further by Hölder inequality, Fubini theorem and integral median theorem we have

$$
\begin{aligned}
\mathbb{E}|J_2|^2 &= \frac{1}{\epsilon^2}\mathbb{E}\left|\int_{t_1}^{t_2}\int_0^\tau y_\tau y_s^{-1}F(x_s^\epsilon)ds d\tau\right|^2\\
&= \frac{1}{\epsilon^2}\mathbb{E}\left|\int_0^{t_1}\int_{t_1}^{t_2} y_\tau y_s^{-1}F(x_s^\epsilon)d\tau ds + \int_{t_1}^{t_2}\int_s^{t_2} y_\tau y_s^{-1}F(x_s^\epsilon)d\tau ds\right|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{2}{\epsilon^2}\mathbb{E}\left|\int_0^{t_1}\int_{t_1}^{t_2} y_\tau y_s^{-1}F(x_s^\epsilon)d\tau ds\right|^2 + \frac{2}{\epsilon^2}\left|\int_{t_1}^{t_2}\int_s^{t_2} y_\tau y_s^{-1}F(x_s^\epsilon)d\tau ds\right|^2 \\
&\leq \frac{2}{\epsilon^2}\mathbb{E}\left(\int_0^{t_1}\int_{t_1}^{t_2} |F(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}d\tau ds\right)^2 \\
&\quad + \frac{2}{\epsilon^2}\mathbb{E}\left(\int_{t_1}^{t_2}\int_s^{t_2} |F(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}d\tau ds\right)^2 \\
&= \frac{2}{\epsilon^2}\mathbb{E}\left(\int_0^{t_1} |F(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\xi-s)}|t_2-t_1|ds\right)^2 \\
&\quad + \frac{2}{C_{\lambda_\alpha}^2}\mathbb{E}\left(\int_{t_1}^{t_2} |F(x_s^\epsilon)|(1-e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(t_2-s)})ds\right)^2 \\
&\leq \frac{2}{\epsilon^2}|t_2-t_1|^2\mathbb{E}\int_0^{t_1}|F(x_s^\epsilon)|^2 ds\mathbb{E}\int_0^{t_1} e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\xi-s)}ds \\
&\quad + \frac{2}{C_{\lambda_\alpha}^2}\mathbb{E}\int_{t_1}^{t_2}|F(x_s^\epsilon)|^2 ds|t_2-t_1| \\
&\leq \frac{C_F^2}{\epsilon C_{\lambda_\alpha}}|t_2-t_1|^2\int_0^{t_1}(1+\mathbb{E}\sup_{0<u\leq s}|x_u^\epsilon|^2)ds(e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}(\xi-t_1)} - e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}\xi}) \\
&\qquad + \frac{2C}{C_{\lambda_\alpha}^2}\int_{t_1}^{t_2}(1+\mathbb{E}\sup_{0<u\leq s}|x_u^\epsilon|^2)ds|t_2-t_1| \\
&\leq \frac{C}{\epsilon}|t_2-t_1|^2(e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}(\xi-t_1)} + e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}\xi}) + C|t_2-t_1|^2 \\
&\leq C|t_2-t_1|^2.
\end{aligned} \tag{3.9}
$$

In the last step of (3.9), we have used the fact that $f(x) = xe^{-ax}$, $a > 0$, $x \in (0, +\infty)$ is bounded. Similarly, we have

$$
\begin{aligned}
&\mathbb{E}|J_3|^2 = \frac{1}{\epsilon^2}\mathbb{E}\left|\int_{t_1}^{t_2}\int_0^{\tau} y_\tau y_s^{-1}\sigma(x_s^\epsilon)dB_s d\tau\right|^2 \\
&= \frac{1}{\epsilon^2}\mathbb{E}\left|\int_0^{t_1}\int_{t_1}^{t_2} y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s + \int_{t_1}^{t_2}\int_s^{t_2} y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s\right|^2 \\
&\leq \frac{2}{\epsilon^2}\mathbb{E}\left|\int_0^{t_1}\int_{t_1}^{t_2} y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s\right|^2 + \frac{2}{\epsilon^2}\mathbb{E}\left|\int_{t_1}^{t_2}\int_s^{t_2} y_\tau y_s^{-1}\sigma(x_s^\epsilon)d\tau dB_s\right|^2 \\
&\leq \frac{2}{\epsilon^2}\int_0^{t_1}\mathbb{E}\left(\int_{t_1}^{t_2}|\sigma(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}d\tau\right)^2 ds + \frac{2}{\epsilon^2}\int_{t_1}^{t_2}\mathbb{E}\left(\int_s^{t_2}|\sigma(x_s^\epsilon)|e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(\tau-s)}d\tau\right)^2 ds \\
&\leq \frac{2}{\epsilon^2}\int_0^{t_1}\mathbb{E}|\sigma(x_s^\epsilon)|^2 e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}(\xi-s)}|t_2-t_1|^2 ds + \frac{2}{C_{\lambda_\alpha^2}}\int_{t_1}^{t_2}\mathbb{E}|\sigma(x_s^\epsilon)|^2(1-e^{-\frac{1}{\epsilon}C_{\lambda_\alpha}(t_2-s)})^2 ds \\
&\leq \frac{2C_\sigma^2}{\epsilon^2}|t_2-t_1|^2\int_0^{t_1}(1+\mathbb{E}\sup_{0<u\leq s}|x_u^\epsilon|^2)(e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}(\xi-s)})ds + \frac{2C_\sigma^2}{C_{\lambda_\alpha}^2}\int_{t_1}^{t_2}(1+\mathbb{E}\sup_{0<u\leq s}|x_u^\epsilon|^2)ds \\
&\leq \frac{C}{\epsilon}|t_2-t_1|^2(e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}(\xi-t_1)} - e^{-\frac{2}{\epsilon}C_{\lambda_\alpha}\xi}) + C|t_2-t_1| \\
&\leq C|t_2-t_1|.
\end{aligned} \tag{3.10}
$$

Now (3.8)–(3.10) yields (3.2). The proof is complete. □

Now by Lemma 3.1 and the Garcia–Rademich–Rumsey theorem [11], we have the tightness of solutions.

Lemma 3.2 *The process* $\{x^\epsilon\}_\epsilon$ *is tight in space* $C(0, T, \mathbb{R}^d)$ *for all* $T > 0$.

Remark 3.1 The above result is assumed by Hottovy et al. [14, Assumption 3].

Next we show the limit of the covariance of $\sqrt{\epsilon} v_t^\epsilon$ as $\epsilon \to 0$ with frozen $x^\epsilon = x$. For this we consider the following linear equation for $x \in \mathbb{R}^d$,

$$\epsilon \dot{v}_t^{\epsilon,x} = -\alpha(x) v_t^{\epsilon,x} + F(x) + \sigma(x)\dot{B}_t.$$

Then we have

Lemma 3.3 *Assume* $(\mathbf{H_1})$ *and* $(\mathbf{H_2})$ *hold, for* $x \in \mathbb{R}^d$

$$\epsilon \mathbb{E} v_t^{\epsilon,x} \otimes v_t^{\epsilon,x} = J(x) + \epsilon C(x,t),$$

where $|C(x,t)| \le C(1+|x|^2)$ *and* $J(x)$ *solves* (2.6).

Proof First by the Itô's formula,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(\epsilon v_t^{\epsilon,x} \otimes v_t^{\epsilon,x}) = & -\frac{\alpha(x)}{\epsilon}\mathbb{E}(\epsilon v_t^{\epsilon,x} \otimes v_t^{\epsilon,x}) \\ & + F(x) \otimes \mathbb{E} v_t^{\epsilon,x} - \frac{1}{\epsilon}\mathbb{E}(\epsilon v_t^{\epsilon,x} \otimes v_t^{\epsilon,x})\alpha^\top(x) \\ & + \mathbb{E} v_t^{\epsilon,x} \otimes F(x) + \frac{1}{\epsilon}\sigma(x)\sigma^\top(x),\end{aligned} \tag{3.11}$$

and

$$\frac{d}{dt}\mathbb{E} v_t^{\epsilon,x} = -\frac{1}{\epsilon}\alpha(x)\mathbb{E} v_t^{\epsilon,x} + \frac{1}{\epsilon}F(x). \tag{3.12}$$

Applying Lemme 2.1 to equation (3.11) and the Duhamel's principle to equation (3.12) respectively, yields

$$\begin{aligned}&\mathbb{E}(\epsilon v_t^{\epsilon,x} \otimes v_t^{\epsilon,x}) \\ &= e^{-\frac{\alpha(x)}{\epsilon}t}\mathbb{E}(\epsilon v_0 \otimes v_0) e^{-\frac{\alpha^\top(x)}{\epsilon}t} \\ &\quad + \int_0^t e^{-\frac{\alpha(x)}{\epsilon}(t-s)}\left(F(x) \otimes \mathbb{E} v_t^{\epsilon,x} + \mathbb{E} v_t^{\epsilon,x} \otimes F(x) + \frac{1}{\epsilon}\sigma(x)\sigma^\top(x)\right) e^{-\frac{\alpha^\top(x)}{\epsilon}(t-s)} ds,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E} v_t^{\epsilon,x} &= e^{-\frac{\alpha(x)}{\epsilon}t} v_0 + \frac{1}{\epsilon}\int_0^t e^{-\frac{\alpha(x)}{\epsilon}(t-s)} F(x) ds \\ &= e^{-\frac{\alpha(x)}{\epsilon}t} v_0 + \alpha^{-1}(x)(I - e^{-\frac{\alpha(x)}{\epsilon}t})F(x).\end{aligned}$$

Thus

$$\begin{aligned}|\mathbb{E} v_t^{\epsilon,x}| &\le |e^{-\frac{\alpha(x)}{\epsilon}t} v_0| + |\alpha^{-1}(x)(I - e^{-\frac{\alpha(x)}{\epsilon}t})F(x)| \\ &\le C + \frac{C_F}{C_{\lambda_\alpha}}(1+|x|) \\ &\le C(1+|x|),\end{aligned}$$

and then $|F(x) \otimes \mathbb{E}v_t^{\epsilon,x}| \le C(1+|x|^2)$. Now let $\tau = \frac{t-s}{\epsilon}$, we have

$$\begin{aligned}
&\int_0^t e^{-\frac{\alpha(x)}{\epsilon}(t-s)} \left(F(x) \otimes \mathbb{E}v_t^{\epsilon,x} + \mathbb{E}v_t^{\epsilon,x} \otimes F(x) + \frac{1}{\epsilon}\sigma(x)\sigma^\top(x) \right) e^{-\frac{\alpha^\top(x)}{\epsilon}(t-s)} ds \\
&= \int_0^{\frac{t}{\epsilon}} e^{-\alpha(x)\tau} (\epsilon F(x) \otimes \mathbb{E}v_t^{\epsilon,x} + \epsilon \mathbb{E}v_t^{\epsilon,x} \otimes F(x) + \sigma(x)\sigma^\top(x)) e^{-\alpha^\top(x)\tau} d\tau \\
&= J(x) + \epsilon \int_0^{\frac{t}{\epsilon}} e^{-\alpha(x)\tau} (F(x) \otimes \mathbb{E}v_t^{\epsilon,x} + \mathbb{E}v_t^{\epsilon,x} \otimes F(x)) e^{-\alpha^\top(x)\tau} d\tau \\
&\quad - \int_{\frac{t}{\epsilon}}^\infty e^{-\alpha(x)\tau} \sigma(x)\sigma^\top(x) e^{-\alpha^\top(x)\tau} d\tau \\
&= J(x) + \epsilon C(x,t)\,.
\end{aligned}$$

The proof is complete. □

Remark 3.2 Here we point out that an important step is to estimate $\epsilon\mathbb{E}|v_t^\epsilon|^2$ in the work of Hottovy et al. [14]. However, in our approach we need the estimate of $\epsilon\mathbb{E}v_t^{\epsilon,x} \otimes v_t^{\epsilon,x}$ instead with fixed x.

4 Averaging Approach

In this section we just consider a convergent subsequence $\rho_\cdot^{\epsilon_k}$ and for simplicity we still write it as $\rho_\cdot^\epsilon$. Let $\rho_\cdot$ be the limit of $\rho_\cdot^\epsilon$. Next we determine the equation for $\rho_\cdot$ by an averaging approach.

Averaging is effective to study the approximation for a slow-fast system [7, 13, 16]. Here we apply the Khasminskii's scheme [16] to (1.8). For small ϵ, ρ_t^ϵ evolves slow, so we can consider the fast part Y^ϵ by freezing the slow part ρ_t^ϵ to be some $\rho \in \mathcal{P}(\mathbb{R}^d)$ and fix $t = \tau$ in $\mathbb{E}^x(v_t^\epsilon \otimes v_t^\epsilon)$. For this we introduce $\tilde{Y}_t^{\epsilon,\rho,\tau}(x)$ the solution of the following equation

$$\partial_t \tilde{Y}_t^{\epsilon,\rho,\tau}(x) = -\frac{\alpha(x)}{\epsilon}\tilde{Y}_t^{\epsilon,\rho,\tau}(x) + \frac{1}{\epsilon}F(x)\rho(x) - [\nabla_x \cdot (\rho(x)\mathbb{E}^x(v_\tau^\epsilon \otimes v_\tau^\epsilon))]^\top, \quad (4.1)$$

with $\tilde{Y}_0^{\epsilon,\rho,\tau}(x) = Y_0$. The following lemma shows that the fast part converges uniformly in τ to some vector with frozen slow part as $\epsilon \to 0$.

Lemma 4.1 *For every fixed $t_* > 0$, under the assumptions $(\mathbf{H_1})$–$(\mathbf{H_3})$, fix $\rho_t^\epsilon = \rho \in \mathcal{P}(\mathbb{R}^d)$ with $\int |x|^2\rho(x)dx$ and $\||Y_0|\|_{L^1}$ bounded, there is a constant $C_T > 0$ such that for $\varphi \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$, and $t \ge t_*$,*

$$|\langle \tilde{Y}_t^{\epsilon,\rho,\tau}(x) - Y^{*,\rho}(x), \varphi\rangle| \le C_T\epsilon\|\varphi\|_{Lip},$$

where

$$Y^{*,\rho}(x) = \alpha^{-1}(x)F(x)\rho(x) - \alpha^{-1}(x)\,[\nabla_x \cdot (\rho(x)J(x))]^\top\,.$$

Proof Applying Duhamel's principle to equation (4.1) yields

$$\begin{aligned}
\tilde{Y}_t^{\epsilon,\rho,\tau}(x) &= e^{-\frac{\alpha(x)}{\epsilon}t}Y_0 + \frac{1}{\epsilon}\int_0^t e^{-\frac{\alpha(x)}{\epsilon}(t-s)}F(x)\rho(x)ds \\
&\quad - \frac{1}{\epsilon}\int_0^t e^{-\frac{\alpha(x)}{\epsilon}(t-s)}[\nabla_x \cdot (\rho(x)\mathbb{E}^x(\epsilon v_\tau^\epsilon \otimes v_\tau^\epsilon))]^\top ds. \quad (4.2)
\end{aligned}$$

Multiplying both sides of the equation (4.2) by the test function φ yields

$$\begin{aligned}\langle \tilde{Y}_t^{\epsilon,\rho,\tau}(x), \varphi\rangle &= \left\langle e^{-\frac{\alpha(x)}{\epsilon}t} Y_0, \varphi \right\rangle + \frac{1}{\epsilon}\int_0^t \left\langle e^{-\frac{\alpha(x)}{\epsilon}(t-s)} F(x)\rho(x), \varphi \right\rangle ds \\ &\quad - \frac{1}{\epsilon}\int_0^t \left\langle e^{-\frac{\alpha(x)}{\epsilon}(t-s)} [\nabla_x \cdot (\rho(x)\mathbb{E}^x(\epsilon v_\tau^\epsilon \otimes v_\tau^\epsilon))]^\top, \varphi \right\rangle ds \\ &\triangleq J_1 + J_2 + J_3.\end{aligned}$$

By Hölder inequality,

$$|J_1| = |\langle e^{-\frac{\alpha(x)}{\epsilon}t} Y_0, \varphi\rangle| \le e^{-\frac{C_{\lambda_\alpha}}{\epsilon}t_*} \| |Y_0| \|_{L^1} \|\varphi\|_{Lip} \le C\epsilon \|\varphi\|_{Lip}. \tag{4.3}$$

Next,

$$\begin{aligned}J_2 &= \left\langle \alpha^{-1}(x)\left(I - e^{-\frac{\alpha(x)}{\epsilon}t}\right) F(x)\rho(x), \varphi \right\rangle \\ &= \langle \alpha^{-1}(x)F(x)\rho(x), \varphi\rangle - \left\langle \alpha^{-1}(x)e^{-\frac{\alpha(x)}{\epsilon}t} F(x)\rho(x), \varphi \right\rangle,\end{aligned}$$

by $(\mathbf{H_1})$ and $(\mathbf{H_2})$,

$$|J_2 - \langle \alpha^{-1}(x)F(x)\rho(x), \varphi\rangle| \le C\epsilon \|F(x)\sqrt{\rho(x)}\|_{L^2} \|\varphi\sqrt{\rho(x)}\|_{L^2} \le C_T\epsilon\|\varphi\|_{Lip}. \tag{4.4}$$

At last, by Lemma 3.3,

$$\begin{aligned}J_3 &= -\left\langle \alpha^{-1}(x)\left(I - e^{-\frac{\alpha(x)}{\epsilon}t}\right)[\nabla_x \cdot (\rho(x)\mathbb{E}^x(v_\tau^\epsilon \otimes v_\tau^\epsilon))]^\top, \varphi \right\rangle \\ &= -\langle \alpha^{-1}(x)\left(I - e^{-\frac{\alpha(x)}{\epsilon}t}\right)[\nabla_x \cdot (\rho(x)(J(x) + \epsilon C(x,\tau)))]^\top, \varphi\rangle \\ &= -\langle \alpha^{-1}(x)(\nabla_x \cdot (\rho(x)J(x)))^\top, \varphi\rangle - \epsilon\langle \alpha^{-1}(x)[\nabla_x \cdot (\rho(x)C(x,\tau))]^\top, \varphi\rangle \\ &\quad + \langle \alpha^{-1}(x)e^{-\frac{\alpha(x)}{\epsilon}t}[\nabla_x \cdot (\rho(x)(J(x) + \epsilon C(x,\tau)))]^\top, \varphi\rangle.\end{aligned}$$

By Gaussian property and the definition of $J(x)$, $|\nabla_x \cdot (\rho(x)J(x))| \le C(1+|x|)\rho(x)$, then

$$|\langle \alpha^{-1}(x)e^{-\frac{\alpha(x)}{\epsilon}t}[\nabla_x \cdot (\rho(x)J(x))]^\top, \varphi\rangle| \le C\epsilon\|(1+|x|)\rho(x)\|_{L^1}\|\varphi\|_{Lip} \le C\epsilon\|\varphi\|_{Lip},$$

and

$$\begin{aligned}|\langle \alpha^{-1}(x)e^{-\frac{\alpha(x)}{\epsilon}t}[\nabla_x \cdot (\epsilon\rho(x)C(x,\tau))]^\top, \varphi\rangle| &= \epsilon|\langle \mathrm{tr}(\rho(x)C(x,\tau)\mathrm{grad}(\alpha^{-1}(x)e^{-\frac{\alpha(x)}{\epsilon}t}\varphi))\rangle| \\ &\le C\epsilon\|(1+|x|^2)\rho(x)\|_{L^1}\|\varphi\|_{Lip} \\ &\le C\epsilon\|\varphi\|_{Lip}.\end{aligned}$$

Thus

$$\begin{aligned}&|J_3 + \langle \alpha^{-1}(x)(\nabla_x \cdot (\rho(x)J(x)))^\top, \varphi\rangle| \\ &\le \epsilon|\langle \mathrm{tr}(\rho(x)C(x,\tau)\mathrm{grad}(\alpha^{-1}(x)\varphi))\rangle| + C\epsilon\|\varphi\|_{Lip} \\ &\le C\epsilon\|(1+|x|^2)\rho(x)\|_{L^1}\|\varphi\|_{Lip} \\ &\le C_T\epsilon\|\varphi\|_{Lip}.\end{aligned} \tag{4.5}$$

By (4.3), (4.4) and (4.5), the proof is complete. □

Remark 4.1 As we have mentioned in the Introduction, one can derive the limit equation formally for ρ_t by replacing Y_t^ϵ by $Y^{*,\rho}$ in the first equation of (1.8).

However the slow part ρ_t^ϵ does evolve, in order to approximate Y_t^ϵ we follow the Khasminskii's scheme. For this we restrict the system in a small time interval, for example $[t_k, t_{k+1}]$ and freeze the slow part to be $\rho_{t_k}^\epsilon$. We show that (Lemma 4.2) Y_t^ϵ is approximated well by $\hat{Y}_t^\epsilon$ with frozen $\rho_t^\epsilon = \rho_{t_k}^\epsilon$ if the length of time interval $[t_k, t_{k+1}]$ is small. For this we divide the time interval $[0, T]$ into small intervals of size $\delta > 0$, i.e. $0 = t_0 < t_1 < \ldots < t_{\lfloor T/\delta \rfloor} + 1 = T$, $t_k = k\delta$, $k = 0, 1, \ldots, \lfloor T/\delta \rfloor$. For $t \in [t_k, t_{k+1}]$, we define the auxiliary process $\{\hat{\rho}_t^\epsilon(x), \hat{Y}_t^\epsilon(x)\}_{0 \le t \le T}$ satisfying

$$
\begin{aligned}
&\partial_t \hat{\rho}_t^\epsilon(x) = -\nabla_x \cdot \hat{Y}_t^\epsilon(x), \\
&\partial_t \hat{Y}_t^\epsilon(x) = -\frac{1}{\epsilon}\alpha(x)\hat{Y}_t^\epsilon(x) + \frac{1}{\epsilon}F(x)\rho_{t_k}^\epsilon(x) - \left[\nabla_x \cdot (\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon \otimes v_{t_k}^\epsilon)\right]^\top, \\
&\hat{\rho}_{t_k}^\epsilon(x) = \rho_{t_k}^\epsilon(x), \quad \hat{Y}_0^\epsilon(x) = Y_0.
\end{aligned}
$$

Remark 4.2 One can see that $\hat{Y}_t^\epsilon = \tilde{Y}_t^{\epsilon, \rho_{t_k}^\epsilon, t_k}$.

Lemma 4.2 *Assume* $(\mathbf{H_1})$–$(\mathbf{H_3})$ *hold, for* $T > 0$ *and* $\varphi \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$,

$$
\sup_{0 \le t \le T} |\langle Y_t^\epsilon(x) - \hat{Y}_t^\epsilon(x), \varphi\rangle| \le C_T \left(\frac{\delta}{\epsilon^2} + \frac{\delta}{\epsilon}\right) \|\varphi\|_{Lip}.
$$

Proof Let $Z_t^\epsilon(x) = Y_t^\epsilon(x) - \hat{Y}_t^\epsilon(x)$. For all $t \in [t_k, t_{k+1}]$, we have

$$
\begin{aligned}
\partial_t Z_t^\epsilon(x) = &-\frac{1}{\epsilon}\alpha(x) Z_t^\epsilon(x) + \frac{1}{\epsilon}F(x)(\rho_t^\epsilon(x) - \rho_{t_k}^\epsilon(x)) \\
&- \left[\nabla_x \cdot (\rho_t^\epsilon(x)\mathbb{E}^x(v_t^\epsilon \otimes v_t^\epsilon) - \nabla_x \cdot (\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon \otimes v_{t_k}^\epsilon)\right]^\top.
\end{aligned}
$$

By Duhamel's principle,

$$
\begin{aligned}
Z_t^\epsilon(x) = &\frac{1}{\epsilon}\int_{t_k}^t e^{-\frac{1}{\epsilon}\alpha(x)(t-s)} F(x)(\rho_s^\epsilon(x) - \rho_{t_k}^\epsilon(x))ds \\
&- \int_{t_k}^t e^{-\frac{1}{\epsilon}\alpha(x)(t-s)}[\nabla_x \cdot (\rho_s^\epsilon(x)\mathbb{E}^x(v_s^\epsilon \otimes v_s^\epsilon) - \nabla_x \cdot (\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon \otimes v_{t_k}^\epsilon)]^\top ds.
\end{aligned}
$$

For $\varphi \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$, we obtain

$$
\begin{aligned}
&\langle Z_t^\epsilon(x), \varphi\rangle \\
&= \frac{1}{\epsilon}\int_{t_k}^t \langle e^{-\frac{1}{\epsilon}\alpha(x)(t-s)} F(x)(\rho_s^\epsilon(x) - \rho_{t_k}^\epsilon(x)), \varphi\rangle ds \\
&\quad - \int_{t_k}^t \langle e^{-\frac{1}{\epsilon}\alpha(x)(t-s)}[\nabla_x \cdot (\rho_s^\epsilon(x)\mathbb{E}^x(v_s^\epsilon \otimes v_s^\epsilon) - \nabla_x \cdot (\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon \otimes v_{t_k}^\epsilon)]^\top, \varphi\rangle ds \\
&\triangleq I_1 + I_2.
\end{aligned}
$$

First

$$
\begin{aligned}
|I_1| &= \frac{1}{\epsilon}\left|\int_{t_k}^t [\mathbb{E}(e^{-\frac{1}{\epsilon}\alpha(x_s^\epsilon)(t-s)} F(x_s^\epsilon)\varphi(x_s^\epsilon)) - \mathbb{E}(e^{-\frac{1}{\epsilon}\alpha(x_{t_k}^\epsilon)(t-s)} F(x_{t_k}^\epsilon)\varphi(x_{t_k}^\epsilon))]ds\right| \\
&\le \frac{1}{\epsilon}\int_{t_k}^t [|\mathbb{E}(e^{-\frac{1}{\epsilon}\alpha(x_s^\epsilon)(t-s)} F(x_s^\epsilon)\varphi(x_s^\epsilon))| + |\mathbb{E}(e^{-\frac{1}{\epsilon}\alpha(x_{t_k}^\epsilon)(t-s)} F(x_{t_k}^\epsilon)\varphi(x_{t_k}^\epsilon))|]ds
\end{aligned}
$$

$$\leq \frac{1}{\epsilon}\int_{t_k}^{t} C_F\|\varphi\|_{Lip}(1+\mathbb{E}|x_s|)ds + \frac{1}{\epsilon}\int_{t_k}^{t} C_F\|\varphi\|_{Lip}(1+\mathbb{E}|x_{t_k}|)ds\,.$$

Then, by Lemma 3.1, we have

$$|I_1| \leq C_T\frac{\delta}{\epsilon}\|\varphi\|_{Lip}. \tag{4.6}$$

Further by Lemma 2.2,

$$\begin{aligned} I_2 &= -\int_{t_k}^{t}\langle[\nabla_x\cdot(\rho_s^\epsilon(x)\mathbb{E}^x(v_s^\epsilon\otimes v_s^\epsilon) - \nabla_x\cdot(\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon\otimes v_{t_k}^\epsilon)], e^{-\frac{1}{\epsilon}\alpha^\top(x)(t-s)}\varphi\rangle ds \\ &= \int_{t_k}^{t}\int_{\mathbb{R}^d}\mathrm{tr}\left[(\rho_s^\epsilon(x)\mathbb{E}^x(v_s^\epsilon\otimes v_s^\epsilon) - \rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon\otimes v_{t_k}^\epsilon))\mathrm{grad}\left(e^{-\frac{1}{\epsilon}\alpha^\top(x)(t-s)}\varphi\right)\right]dxds\,. \end{aligned}$$

Let $g(x) = e^{-\frac{1}{\epsilon}\alpha^\top(x)(t-s)}\varphi(x)$, by the chain rule,

$$\frac{\partial}{\partial x_i}g(x) = \frac{\partial}{\partial x_i}(e^{-\frac{1}{\epsilon}\alpha^\top(x)(t-s)})\varphi(x) + e^{-\frac{1}{\epsilon}\alpha^\top(x)(t-s)}\frac{\partial}{\partial x_i}\varphi(x)\,.$$

Then, by assumptions ($\mathbf{H_1}$) and ($\mathbf{H_3}$),

$$|\frac{\partial}{\partial x_i}g(x)| \leq C(\frac{1}{\epsilon}+1)\|\varphi\|_{Lip}, \tag{4.7}$$

thus

$$|\mathrm{grad}(g(x))| \leq C(\frac{1}{\epsilon}+1)\|\varphi\|_{Lip}. \tag{4.8}$$

By Lemma 3.3 and (4.8),

$$\begin{aligned} &\left|\int_{\mathbb{R}^d}\mathrm{tr}(\rho_s^\epsilon(x)\mathbb{E}^x(v_s^\epsilon\otimes v_s^\epsilon)\mathrm{grad}(g(x)))dx\right| \\ &= \frac{1}{\epsilon}|\mathrm{tr}[\mathbb{E}(\mathbb{E}^{x_s^\epsilon}(\epsilon v_s^\epsilon\otimes v_s^\epsilon)\mathrm{grad}(g(x_s^\epsilon)))]| \\ &\leq \frac{C}{\epsilon}\mathbb{E}[(J(x_s^\epsilon)+\epsilon C(x_s^\epsilon, s))(\frac{1}{\epsilon}+1)\|\varphi\|_{Lip} \\ &\leq C_T(\frac{1}{\epsilon^2}+\frac{1}{\epsilon})\|\varphi\|_{Lip}. \end{aligned}$$

Similarly,

$$\left|\int_{\mathbb{R}^d}\mathrm{tr}(\rho_{t_k}^\epsilon(x)\mathbb{E}^x(v_{t_k}^\epsilon\otimes v_{t_k}^\epsilon)\mathrm{grad}(g(x)))dx\right| \leq C_T(\frac{1}{\epsilon^2}+\frac{1}{\epsilon})\|\varphi\|_{Lip}.$$

Then we have

$$|I_2| \leq C_T(\frac{\delta}{\epsilon^2}+\frac{\delta}{\epsilon})\|\varphi\|_{Lip}.$$

The proof is complete. □

Proof of Theorem 2.1. From the first equation of (1.8), for $\psi \in C_0^\infty(\mathbb{R}^d)$ and $\varphi = \nabla\psi$ we derive

$$\langle\rho_t^\epsilon(x), \psi\rangle = \langle\rho_{t_*}, \psi\rangle + \int_{t_*}^{t}\langle Y_s^\epsilon(x), \varphi\rangle ds$$

$$
\begin{aligned}
&= \langle \rho_{t_*}, \psi \rangle + \int_{t_*}^{t} \langle Y^{*,\rho_s^\epsilon}(x), \varphi \rangle ds \\
&\quad + \int_{t_*}^{t} \langle Y_s^\epsilon(x) - \hat{Y}_s^\epsilon(x), \varphi \rangle ds + \int_{t_*}^{t} \langle \hat{Y}_s^\epsilon(x) - Y^{*,\rho_s^\epsilon}(x), \varphi \rangle ds \\
&\triangleq \langle \rho_{t_*}, \psi \rangle + K_1 + K_2 + K_3.
\end{aligned}
$$

First, by the expression of $Y^{*,\rho}$,

$$
K_1 = \int_{t_0}^{t} \langle Y^{*,\rho_s^\epsilon}(x), \varphi \rangle ds \to \int_{t_0}^{t} \langle Y^{*,\rho_s}(x), \varphi \rangle ds. \tag{4.9}
$$

Next, by Lemma 4.2,

$$
|K_2| \le C_T \left(\frac{\delta}{\epsilon^2} + \frac{\delta}{\epsilon} \right) \|\varphi\|_{Lip} \to 0, \tag{4.10}
$$

by choosing $\delta = O(\epsilon^3)$.

Note that on the time interval $[t_k, t_{k+1}]$, $\{\hat{Y}_t^\epsilon\} = \{\tilde{Y}_t^{\epsilon,\rho_{t_k}^\epsilon,t_k}\}$, let $t_* \in [t_{i_0}, t_{i_0+1}]$,

$$
\begin{aligned}
K_3 &= \int_{t_*}^{t_{i_0+1}} \langle \tilde{Y}_s^{\epsilon,\rho_{t_{i_0}}^\epsilon,t_{i_0}} - Y^{*,\rho_{t_{i_0}}^\epsilon}, \varphi \rangle ds + \sum_{k=i_0+1}^{\lfloor t/\delta \rfloor - 1} \int_{t_k}^{t_{k+1}} \langle \tilde{Y}_s^{\epsilon,\rho_{t_k}^\epsilon,t_k} - Y^{*,\rho_{t_k}^\epsilon}, \varphi \rangle ds \\
&\quad + \int_{t_{\lfloor t/\delta \rfloor}}^{t} \langle \tilde{Y}_s^{\epsilon,\rho_{t_{\lfloor t/\delta \rfloor}}^\epsilon,t_{\lfloor t/\delta \rfloor}} - Y^{*,\rho_{t_{\lfloor t/\delta \rfloor}}^\epsilon}, \varphi \rangle ds + \int_{t_*}^{t_{i_0+1}} \langle Y^{*,\rho_{t_{i_0+1}}^\epsilon} - Y^{*,\rho_s^\epsilon}, \varphi \rangle ds \\
&\quad + \sum_{k=i_0+1}^{\lfloor t/\delta \rfloor - 1} \int_{t_k}^{t_{k+1}} \langle Y^{*,\rho_{t_k}^\epsilon} - Y^{*,\rho_s^\epsilon}, \varphi \rangle ds + \int_{t_{\lfloor t/\delta \rfloor}}^{t} \langle Y^{*,\rho_{t_{\lfloor t/\delta \rfloor}}^\epsilon} - Y^{*,\rho_s^\epsilon}, \varphi \rangle ds \\
&\triangleq K_{31} + K_{32} + K_{33} + K_{34} + K_{35} + K_{36}.
\end{aligned}
$$

By Lemma 4.1,

$$
|K_{31} + K_{32} + K_{33}| \le C_T \epsilon \|\varphi\|_{Lip}. \tag{4.11}
$$

By the defination of $Y^{*,\rho}$ and Lemma 3.1,

$$
\begin{aligned}
|K_{34} + K_{35} + K_{36}| &\le C_T \Bigg(\int_{t_*}^{t_{i_0+1}} \mathbb{E}|x_{t_{i_0+1}}^\epsilon - x_{t_*}^\epsilon|^2 ds + \sum_{k=i_0+1}^{\lfloor t/\delta \rfloor - 1} \int_{t_k}^{t_{k+1}} \mathbb{E}|x_{t_k}^\epsilon - x_s^\epsilon|^2 ds \\
&\quad + \int_{t_{\lfloor t/\delta \rfloor}}^{t} \mathbb{E}|x_{t_{\lfloor t/\delta \rfloor}}^\epsilon - x_s^\epsilon|^2 ds \Bigg) \|\varphi\|_{Lip} \\
&\le C_T \|\varphi\|_{Lip} \delta.
\end{aligned} \tag{4.12}
$$

By (4.9)–(4.12), passing the limit $\epsilon \to 0$ yields

$$
\langle \rho_t^\epsilon(x), \psi \rangle \to \langle \rho_{t_*}, \psi \rangle + \int_{t_*}^{t} \langle Y^{*,\rho_s}(x), \varphi \rangle ds.
$$

Since $\varphi = \nabla_x \psi$, it yields

$$
\langle \partial_t \rho_t(x), \psi \rangle = \langle Y^{*,\rho_t}(x), \nabla_x \psi \rangle = -\langle \nabla_x \cdot Y^{*,\rho_t}(x), \psi \rangle,
$$

which is the weak form of

$$\partial_t \rho_t(x) = -\nabla_x \cdot (\alpha^{-1}(x)F(x)\rho_t(x) + \alpha^{-1}(x)(\nabla_x \cdot (\rho_t(x)J(x)))^\top).$$

The proof is complete. □

Data Availability No data was used for the research described in the article.

Declarations

competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Behr, M., Benner, P., Heiland, J.: Solution formulas for differential Sylvester and Lyapunov equations. Calcolo **56**, 51 (2019)
2. Bellman, R.: Introduction to Matrix Analysis. Society for Industrial and Applied Mathematics (1997)
3. Cerrai, S., Freidlin, M.: On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom. Prob. Theory Relat. Fields **135**, 363–394 (2006)
4. Cerrai, S., Freidlin, M.: Large deviations for the Langevin equation with strong damping. J. Stat. Phys. **161**, 859–875 (2015)
5. Cerrai, S., Freidlin, M., Salins, M.: On the Smoluchowski-Kramers approximation for SPDEs and its interplay with large deviations and long time behavior. Discret. Contin. Dyn. Syst. **37**(1), 33–76 (2016)
6. Cerrai, S., Xi, G.: A Smoluchowski-Kramers approximation for an infinite dimensional system with state-dependent damping. Ann. Prob. **50**(3), 874–904 (2022)
7. Duan, J., Wang, W.: Effective Dynamics of Stochastic Partial Differential Equations. Elsevier, Amsterdam (2014)
8. Freidlin, M.: Some remarks on the Smoluchowski-Kramers approximation. J. Stat. Phys. **117**(3/4), 617–634 (2004)
9. Freidlin, M., Hu, W.: Smoluchowski-Kramers approximation in the case of variable friction. J. Math. Sci. **179**(1), 184–207 (2011)
10. Freidlin, M., Hu, W., Wentzell, A.: Small mass asymptotic for the motion with vanishing friction. Stoch. Proc. Appl. **123**(1), 45–75 (2013)
11. Garsia, A., Rademich, E., Rumsey, H.: A real variable lemma and the continuity of paths of some Gaussian processes. Indiana Univ. Math. J. **20**, 565–578 (1970/71)
12. Hartman, P.: Ordinary Differential Equations. Mathematics of Computation. Society for Industrial and Applied Mathematics (2002)
13. Hashemi, S.N., Heunis, A.J.: Averaging principle for diffusion processes. Stoch. Stoch. Rep. **62**, 201–216 (1998)
14. 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**, 1259–1283 (2015)
15. Kelly, P.: Mechanics Lecture Notes Part III: Foundations of Continuum Mechanics. https://www.homepages.engineering.auckland.ac.nz/~pkel015/SolidMechanicsBooks/index.html
16. Khasminskii, R.Z.: On the principle of averaging the Itô's stochastic differential equations (in Russian). Kibernetika **4**, 260–279 (1968)
17. Kramers, H.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**, 284–304 (1940)
18. Roberts, J.B., Spanos, P.D.: Stochastic averaging: an approximate method of solving random vibration, problems. Int. J. Non-Linear Mech. **21**(2), 111–134 (1986)
19. Pavliotis, G.A., Stuart, A.M.: Multiscale Methods: Averaging and Homogenization. Springer, Berlin (2008)
20. Smoluchowski, M.: Drei Vortrage über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. Phys. Z. **17**, 557–585 (1916)
21. Volpe, G., Helden, L., Brettschneider, T., Wehr, J., Bechinger, C.: Influence of noise on force measurements. Phys. Rev. Lett. **104**, 170602 (2010)