Asymptotic behavior of a class of multiple time scales stochastic kinetic equations

Charles-Edouard Bréhier, Shmuel Rakotonirina–Ricquebourg

Asymptotic behavior of a class of multiple time scales stochastic kinetic equations

Charles-Edouard Bréhier and Shmuel Rakotonirina-Ricquebourg*

June 11, 2021

Abstract

We consider a class of stochastic kinetic equations, depending on two time scale separation parameters ε and δ: the evolution equation contains singular terms with respect to ε, and is driven by a fast ergodic process which evolves at the time scale t/δ^2. We prove that when $(\varepsilon, \delta) \to (0, 0)$ the density converges to the solution of a linear diffusion PDE. This is a mixture of diffusion approximation in the PDE sense (with respect to the parameter ε) and of averaging in the probabilistic sense (with respect to the parameter δ). The proof employs stopping times arguments and a suitable perturbed test functions approach which is adapted to consider the general regime $\varepsilon \neq \delta$.

1 Introduction

Multiscale and/or stochastic models are popular in all fields of science and engineering. In this paper, we consider a stochastic kinetic partial differential equation of the type

$$\partial_t f^{\varepsilon,\delta} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon,\delta} + b(v) \cdot \nabla_x f^{\varepsilon,\delta} + \sigma(m^\delta(t,x)) f^{\varepsilon,\delta} = \frac{1}{\varepsilon^2} L f^{\varepsilon,\delta}, \tag{1}$$

with initial condition $f^{\varepsilon,\delta}(0) = f_0^{\varepsilon,\delta}$. The unknow $f^{\varepsilon,\delta}$ is a function of time $t \geqslant 0$, position $x \in \mathbb{T}^d$ (the flat d-dimensional torus) and velocity $v \in V$. Assumptions on the velocity fields a and b and on the mapping σ are given below (see Section 2). Note that $f^{\varepsilon,\delta}(t,x,v)$ may be interpreted as a density of particles with position x and velocity v at time t; the system is not conservative (the integral of $f^{\varepsilon,\delta}$ is not constant) due to the source term $\sigma(m^\delta) f^{\varepsilon,\delta}$. In addition, the linear operator L describes interactions between the particles: in this paper, we assume that L is the Bhatnagar-Gross-Krook operator, given by

$$Lf = \rho \mathcal{M} - f,$$

where μ is a σ-finite measure on V, the spatial density is defined by $\rho \doteq \langle f \rangle \doteq \int_V f d\mu$, and where $\mathcal{M} \in L^1(V, d\mu)$ is a density function, often called the Maxwellian (see Assumption 1).

The evolution equation (1) depends on the so-called driving process m^δ, defined as follows: one has $m^\delta(t,x) = m(t/\delta^2, x)$.

*Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France,brehier@math.univ-lyon1.fr,rakoto@math.univ-lyon1.fr

The stochastic evolution equation (1) depends on two parameters ε and δ. In this paper, we are interested in the asymptotic behavior when $(\varepsilon, \delta) \to (0, 0)$. We prove the following result: $\rho^{\varepsilon,\delta} = \left\langle f^{\varepsilon,\delta} \right\rangle$ converges to the solution $\overline{\rho}$ of the following partial differential equation:

$$\partial_t \overline{\rho} + J \cdot \nabla \overline{\rho} + \overline{\sigma}\, \overline{\rho} = \operatorname{div}(K \nabla \overline{\rho}), \tag{2}$$

where J, $\overline{\sigma}$ and K are defined below, with initial condition $\overline{\rho}(0) = \overline{\rho}_0 = \lim \rho^{\varepsilon,\delta}(0)$. We refer to the main results of this paper, Theorems 3.2 and 3.3 for rigorous statements, in particular concerning the mode of convergence. Let us describe how the form of the limit equation (2) arises in the asymptotic regime.

On the one hand, the parameter ε drives the behavior of the deterministic part of the evolution. Under appropriate assumptions (including a centering condition for the velocity field a), the spatial density $\rho^{\varepsilon,\delta} = \langle f \rangle^{\varepsilon,\delta}$ converges when $\varepsilon \to 0$ to the solution of a diffusion partial differential equation, where the velocity variable v has been eliminated. In the literature, such convergence results are referred to as diffusion approximation results, see for instance [DGP00]. The result is also partly an averaging result, since the term $b(v) \cdot \nabla_x$ is replaced in the limit equation by $J \cdot \nabla_x$, where J is the average of b (with respect to an appropriate measure).

On the other hand, the parameter δ is a time-scale separation parameter which determines the random part of the evolution. When the process m such that $m^\delta(t, x) = m(t/\delta^2, x)$ is assumed to be ergodic (see Section 2, for instance one may consider an Ornstein-Uhlenbeck process), the randomness may be eliminated when $\delta \to 0$: only the average $\overline{\sigma}$ of $\sigma(m)$ with respect to the invariant distribution remains in the limit evolution equation (it is a law of large numbers effect). In the literature, such convergence results are referred to as averaging principle results.

In this paper, we thus prove the mixture of an diffusion approximation result in the PDE sense, and of an averaging principle result in the probability sense, when simultaneously $\varepsilon \to 0$ and $\delta \to 0$. To the best of our knowledge, this regime has not been considered in the literature so far. Note that one of the major tasks in the analysis is to consider the general case when ε and δ go to 0 independently. Indeed, the analysis would be simpler if $\varepsilon = \delta$. Note also that it would be simpler if $\varepsilon = 0$ and $\delta \to 0$, or if $\varepsilon \to 0$ and $\delta = 0$, *i.e.* if the limits are taken successively. The latter case is not included in the analysis but may be handled with similar techniques. The limit equation (2) is the same in all regimes.

For deterministic problems ($\sigma = 0$), diffusion approximation results have been extensively studied. We refer for instance to [LK74, BLP79]. Kinetic models with small parameters appear in various situations, for example when studying semi-conductors [GP92] and discrete velocity models [LT97], or as limits for description of systems of particles, either with a single particle [GR09] or multiple particles [PV03]. The asymptotic behavior of stochastic kinetic multiscale problems have also been recently studied: we refer to the seminal article [DV12], and the more recent contributions [DV20, DRV20, RR20]. In those works, the authors have obtained diffusion approximation results both in the PDE and the probabilistic senses: the limit equation is a stochastic linear diffusion PDE driven by a Wiener process (with Stratonovich interpretation). Indeed, in those works $\sigma(m^\delta)$ (with $\delta = \varepsilon$) is replaced by m^δ/δ in (1), and the authors assume that the driving process m^δ satisfies an appropriate centering condition. In the present article, we consider a law of large numbers regime (hence the averaging principle result), instead of a central limit theorem regime. In spite of this fundamental difference, the setting is very close to [RR20]: in particular a major technical difficulty which is solved in this paper is to avoid boundedness assumptions on the driving process m, using only moment conditions, which allow us to encompass for instance Ornstein-Uhlenbeck processes.

The literature concerning the averaging principle for stochastic differential equations and stochastic partial differential equations is huge. The averaging principle in the SDE case has been introduced in the seminal reference [Kha68], see also the monograph [PS08] and references

therein. In the SPDE case, authors have mainly studied the averaging principle for parabolic semilinear SPDE systems, see for instance [Cer09, CF09, Bré12, Bré20, RXY20] and references therein. Let us also mention the recent preprints [CX20, XY21] where diffusion approximation results are proved for such systems. The list of references above is not exhaustive.

The first main result of this manuscript is Theorem 3.2: the convergence of $\rho^{\varepsilon,\delta} = \langle f^{\varepsilon,\delta} \rangle$ to $\overline{\rho}$ is understood as convergence in distribution (in the probabilistic sense), in the space $\mathcal{C}^0([0,T], H^{-\varsigma}(\mathbb{T}^d))$, for all arbitrarily small positive ς. Under an additional assumption (which allows us to employ an averaging lemma), the convergence holds in the space $L^2([0,T], L^2(\mathbb{T}^d))$, see the second main result of this manuscript, Theorem 3.3. The functional spaces above are the standard spaces where convergence holds in the deterministic case. The convergence in distribution is the natural mode of the convergence for the probabilistic variable. However, since the limit equation (2) is deterministic, if the limit initial condition $\overline{\rho}_0$ is also deterministic, then the convergence holds in probability.

Let us now describe the main tools for the proof of the main results of this manuscript. We follow a martingale problem approach combined with the perturbed test functions method, as in the classical article [PSV77] (see also [Kus84, EK86, FGPSl07, PS08, dBG12]). Perturbed test functions in the context of PDEs with diffusive limits applies in various situations, for instance in the context of viscosity solutions [Eva89], nonlinear Schrödinger equations [dBG12], a parabolic PDE [PP03] or, as in this article, kinetic SPDEs [DV12, DV20, DRV20, RR20].

The idea of the perturbed test function is to identify the limit generator $\mathcal{L}$ of the limit equation (2) as

$$\mathcal{L}\varphi = \lim_{(\varepsilon,\delta)\to(0,0)} \mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta},$$

where $\varphi = \varphi(\rho)$ is an arbitrary test function (depending only on the spatial density variable $\rho = \langle f \rangle$ appearing in the limit equation), and $\varphi^{\varepsilon,\delta} = \varphi^{\varepsilon,\delta}(f,m)$ is the perturbed test function given by

$$\varphi^{\varepsilon,\delta} = \varphi + \varepsilon\varphi_{1,0} + \varepsilon^2\varphi_{2,0} + \delta^2\varphi_{0,2} + \varepsilon\delta^2\varphi_{1,2}.$$

We refer to Proposition 4.3 for a rigorous statement. Note that due to the assumption that in general $\varepsilon \neq \delta$, the construction of the perturbed test function requires to define the corrector $\varphi_{1,2}$ (corresponding to the term of the order $\varepsilon\delta^2$). This corrector does not appear in the analysis if $\varepsilon = \delta$. The construction of $\varphi_{1,2}$ is one of the novelties of this manuscript.

In addition, like in the preprint [RR20], the driving process is not assumed to be bounded (like in [DV12] for instance), and only moment conditions are satisfied. This generalization allows us for instance to consider Ornstein-Uhlenbeck processes. The introduction of stopping times arguments is required, hence the need to control the asymptotic behavior of the stopping times τ^δ when $\delta \to 0$. Note that we prove below that $\tau^\delta \to \infty$ in probability: the arguments to prove convergence are thus simpler than in [RR20].

The main result of this manuscript is the convergence of $\rho^{\varepsilon,\delta}$ to $\overline{\rho}$. In future works, it may be interesting to study the fluctuations, *i.e.* to prove that $\rho^{\varepsilon,\delta} - \overline{\rho}$, properly rescaled, converges in distribution to a Gaussian process, solution of a linear stochastic evolution equation. Again the use of the perturbed test functions approach may be a suitable approach. It would also be interesting to investigate rates of convergence, both in the strong and weak senses, in the spirit of [Bré12, Bré20] concerning parabolic systems, using Kolmogorov equations techniques. Finally, the validity of diffusion approximation and averaging principle results is fundamental for the efficient numerical simulation of the systems. In the deterministic setting, there has been a lot of activity to develop asymptotic preserving and uniformly accurate numerical methods, see for instance [Jin99, Jin12]. An asymptotic preserving scheme has been proposed in [AF19] for a class of kinetic stochastic equations in the diffusion approximation regime. In a future work [BHRR21],

we plan to investigate the generalization of the asymptotic preserving schemes introduced and analyzed in the recent preprint [BRR20] for stochastic differential equations.

The manuscript is organized as follows. The setting is described in Section 2 (in particular precise assumptions concerning the driving process m are provided). The main results, Theorems 3.2 and 3.3, are stated and discussed in Section 3. Section 4 is devoted to the proofs of the main results, using martingale problem formulations, tightness arguments and identification of the limit. Auxiliary fundamental results are stated there: first, Proposition 4.1 concerning the asymptotic behavior of the stopping time; second, Proposition 4.2 providing an a priori estimate in an appropriate weighted L^2 norm, uniformly with respect to ε, δ; third, Proposition 4.3 giving the details on the perturbed test functions. The proofs of those three auxiliary results are given in Section 5. Note that these three results are essential, and their proofs are given in a separate section since they are the most original technical contributions of this manuscript (compared with the more standard strategy described in Section 4).

2 Setting

2.1 Notation

The solution $f^{\varepsilon,\delta}$ of the stochastic kinetic problem considered in this article is a process $f^{\varepsilon,\delta} : (t,x,v) \in [0,\infty) \times \mathbb{T}^d \times V \to \mathbb{R}$, where V is a measurable space, equipped with a σ-finite measure μ, and $\mathbb{T}^d$ denotes the flat d-dimensional torus.

Let us introduce the standard Hilbert spaces $L^2_x = L^2(\mathbb{T}^d, \mathbb{R})$ and $L^2 = L^2(\mathbb{T}^d \times V, \mathbb{R})$ of real-valued functions, with inner products defined as follows:

$$(h,k)_{L^2_x} = \int_{\mathbb{T}^d} h(x)k(x)dx, \quad (h,k)_{L^2} = \int_{\mathbb{T}^d \times V} h(x,v)k(x,v)dxd\mu(v).$$

The associated norms are denoted by $\|\cdot\|_{L^2_x}$ and $\|\cdot\|_{L^2}$ respectively.

The following notation is used in the sequel: for all $f \in L^1(\mathbb{T}^d \times V)$, let $\rho \in L^1(\mathbb{T}^d)$ be defined by

$$\rho \doteq \langle f \rangle \doteq \int_V f d\mu.$$

In addition, for all $T \in (0,\infty)$, $\varsigma \in (0,1]$, $i \in \mathbb{N}_0$ and $p \in [1,\infty]$, introduce the Banach spaces $\mathcal{C}^i_x \doteq \mathcal{C}^i(\mathbb{T}^d, \mathbb{R})$, $\mathcal{C}^0_T H^{-\varsigma}_x \doteq \mathcal{C}^0([0,T], H^{-\varsigma}(\mathbb{T}^d, \mathbb{R}))$ and $L^p_T L^2_x \doteq L^p([0,T], L^2(\mathbb{T}^d, \mathbb{R}))$.

Finally, the state space E of the driving stochastic process is assumed to be a Banach space, with norm denoted by $\|\cdot\|_E$.

2.2 Assumptions on the coefficients

Let us now state the assumptions concerning the linear operator L and the mappings a, b and σ.

Assumption 1.

1. Let the mapping $\mathcal{M} \in L^1(V,\mu)$, be such that $\mathcal{M}(v) > 0$ for all $v \in V$, and normalized such that $\int_V \mathcal{M}(v)d\mu(v) = 1$.
2. The linear operator L is defined as follows: for all $f \in L^1(V,\mu)$
$$Lf = \rho\mathcal{M} - f = \langle f \rangle \mathcal{M} - f. \tag{3}$$
3. The mappings $a : V \to \mathbb{R}^d$ and $b : V \to \mathbb{R}^d$ are bounded. In addition a satisfies the centering condition
$$\int_V a(v)\mathcal{M}(v)d\mu(v) = 0.$$

4. The linear operators A and B are defined by

$$Af(x,v) \doteq a(v) \cdot \nabla_x f(x,v),$$
$$Bf(x,v) \doteq b(v) \cdot \nabla_x f(x,v).$$

5. The mapping $\sigma : E \to \mathcal{C}_x^{\lfloor d/2 \rfloor + 2}$ is Lipschitz continuous.

Let us also introduce the weighted Hilbert space $L^2(\mathcal{M}^{-1})$, with the inner product

$$(h,k)_{L^2(\mathcal{M}^{-1})} = \int_{\mathbb{T}^d \times V} h(x,v)k(x,v)dx\frac{d\mu(v)}{\mathcal{M}(v)}.$$

The associated norm is denoted by $\|\cdot\|_{L^2(\mathcal{M}^{-1})}$. Observe that applying the Cauchy-Schwarz inequality yields the following results: for all $f \in L^2(\mathcal{M}^{-1})$, one has $f \in L^1(\mathbb{T}^d \times V)$, $\rho \in L^2_x$, and $Lf \in L^2(\mathcal{M}^{-1})$, with

$$\iint_{\mathbb{T}^d \times V} |f(x,v)|d\mu(v)dx \leqslant \|f\|_{L^2(\mathcal{M}^{-1})}$$
$$\|\rho\mathcal{M}\|_{L^2(\mathcal{M}^{-1})} = \|\rho\|_{L^2_x} \leqslant \|f\|_{L^2(\mathcal{M}^{-1})}.$$

In addition, if $f_1 \in L^2(\mathcal{M}^{-1}) \cap L^2$ and $f_2 \in L^2(\mathcal{M}) \cap L^2$, then

$$|(f_1, f_2)_{L^2}| \leqslant \|f_1\|_{L^2(\mathcal{M}^{-1})} \|f_2\|_{L^2(\mathcal{M})}$$

as a consequence of the Cauchy-Schwarz inequality.

Example 2.1. The conditions in Assumption 1 above are satisfied in the following two examples:

1. Continuous velocities
   - The space $V = \mathbb{R}^d$ is equipped with the Lebesgue measure $d\mu(v) = dv$.
   - The function $\mathcal{M}(v) = (2\pi)^{-d/2} \exp(-\|v\|^2/2)$ for all $v \in \mathbb{R}^d$ is the standard Maxwellian.
   - The velocity a is a bounded odd function. For instance, relativistic particles satisfy $a(v) = \frac{v}{\sqrt{1+\|v\|^2}}$ in convenient units.
2. Discrete velocities
   - The space $V = \{\pm 1\}^d$ is equipped with the counting measure.
   - The function $\mathcal{M}$ is constant: $\mathcal{M}(v) = \frac{1}{2d}$ for all $v \in \{\pm 1\}^d$
   - The velocity a is an odd function. For instance, the isotropic discrete velocity is given by $a(v) = v$ for all $v \in \{\pm 1\}^d$.

In both examples, the function σ is defined either as

- $\sigma(\ell)(x) = \sigma_1(\ell(x))$ for all $\ell \in E = \mathcal{C}_x^{\lfloor d/2 \rfloor + 2}$ and $x \in \mathbb{T}^d$, with a mapping $\sigma_1 : \mathbb{R} \to \mathbb{R}$,
- or as $\sigma(\ell)(x) = \sigma_2(x, \ell)$ for $\ell \in E = \mathbb{R}$ and $x \in \mathbb{T}^d$, with a mapping $\sigma_2 : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}$,

where σ_1, σ_2 are of class $\mathcal{C}^{\lfloor d/2 \rfloor + 3}$ with bounded derivatives of all orders.

2.3 Assumptions on the driving process

The driving process is a Markov process, which satisfies the conditions below.

Assumption 2. The family $(m_\ell(t))_{\ell\in E}$ defines a E-valued Markov process, where one has the initial condition $m_\ell(0) = \ell$. Let $\mathcal{L}_m$ denote its infinitesimal generator, with domain denoted by $D(\mathcal{L}_m)$.

We assume that this Markov process is ergodic, and that its unique invariant distribution, denoted by ν, is integrable: $\int_E \|\ell\|\, d\nu(\ell) < \infty$.

The following notation is used throughout the article: for all Lipschitz continuous mappings $\theta : E \to \mathbb{R}$, set

$$\overline{\theta} = \int_E \theta(\ell) d\nu(\ell).$$

Let $\ell_0 \in E$ be a given initial condition, in the sequel the value of ℓ_0 is omitted to simplify notation: for all $t \geqslant 0$, let

$$m(t) = m_{\ell_0}(t).$$

Assumption 2 is sufficient to state the main convergence results below. However, the analysis of the asymptotic behavior of $f^{\varepsilon,\delta}$ when $\delta \to 0$ requires additional technical assumptions. Since they are not needed to state the convergence results below, they may be skipped by the reader, until they are used to prove auxiliary results.

Assumption 3. The Markov process introduced in Assumption 2 satisfies the appropriate moment bounds: there exists $\gamma \in (2, \infty)$ such that

$$\sup_{i\in\mathbb{N}_0} \mathbb{E}\left[\sup_{t\in[i,i+1]} \|m(t)\|_E^\gamma\right] < \infty. \tag{4}$$

The assumption that $\gamma > 2$ is crucial in the analysis. Observe that Assumption 3 implies the following results:

$$\int \|\ell\|^\gamma \, d\nu(\ell) < \infty$$
$$\sup_{t\geqslant 0} \mathbb{E}\left[\|m(t)\|_E^\gamma\right] < \infty.$$

A mixing assumption is employed below to have quantitative information on the large time behavior of the driving process.

Assumption 4. The Markov process introduced in Assumption 2 satisfies a mixing property: there exists a nonnegative function $\gamma_{\text{mix}} \in L^1(\mathbb{R}^+)$ such that, for all initial conditions $\ell_1, \ell_2 \in E$, there exists a coupling $(m^*_{\ell_1}, m^*_{\ell_2})$ of (m_{ℓ_1}, m_{ℓ_2}), satisfying the inequality

$$\mathbb{E}\left[\left\|m^*_{\ell_1}(t) - m^*_{\ell_2}(t)\right\|_E\right] \leqslant \gamma_{\text{mix}}(t) \left\|\ell_1 - \ell_2\right\|_E$$

for all $t \geqslant 0$.

Let us recall that a $E \times E$-valued random process $(m^*_{\ell_1}, m^*_{\ell_2})$ is a coupling of (m_{ℓ_1}, m_{ℓ_2}) if the marginals safisty $m^*_{\ell_1} \overset{d}{=} m_{\ell_1}$ and $m^*_{\ell_2} \overset{d}{=} m_{\ell_2}$ (where equality is understood in distribution). As a consequence of Assumption 4, it is straightforward to obtain the following result: if $\theta : E \to \mathbb{R}$ is Lipschitz continuous, then for all $\ell \in E$ and all $t \geqslant 0$, one has

$$\left|\mathbb{E}[\theta(m_\ell(t))] - \overline{\theta}\right| \leqslant \max(1, \int_E \left\|\ell'\right\|_E d\nu(\ell'))\text{Lip}(\theta)\gamma_{\text{mix}}(t)(1 + \|\ell\|_E),$$

where $\mathrm{Lip}(\theta)$ denotes the Lipschitz constant of θ. Due to this consequence of the mixing property (Assumption 4), the resolvent operator R_0 introduced below is well-defined.

Definition 2.1 (Resolvent operator)**.** Let $\left(\mathcal{C}_x^{\lfloor d/2 \rfloor+2}\right)^*$ be the set of continuous linear forms on $\mathcal{C}_x^{\lfloor d/2 \rfloor+2}$ and let $E^*(\sigma) \doteq \left\{ u \circ \sigma \mid u \in \left(\mathcal{C}_x^{\lfloor d/2 \rfloor+2}\right)^* \right\} \subset \mathrm{Lip}(E, \mathbb{R})$. The resolvent operator R_0 is defined as follows: for all $\theta \in E^*(\sigma)$

$$R_0(\theta - \overline{\theta})(\ell) = \int_0^\infty \mathbb{E}\left[\theta(m_\ell(t)) - \overline{\theta}\right] dt.$$

The function $\psi_\theta \doteq R_0(\theta - \overline{\theta})$ is the unique solution of the Poisson equation

$$-\mathcal{L}_m \psi_\theta = \theta - \overline{\theta}, \tag{5}$$

satisfying the condition $\overline{\psi}_\theta = 0$.

Note that the functions ψ_θ satisfy the following bound: there exists $C \in (0, \infty)$, such that for all $\theta \in E^*(\sigma)$ and for all $\ell \in E$, one has

$$|\psi_\theta(\ell)| \leqslant C \mathrm{Lip}(\theta)(1 + \|\ell\|_E), \tag{6}$$

The remaining assumption deals with the infinitesimal generator $\mathcal{L}_m$ of the driving process.

Assumption 5. For all $\theta_1, \theta_2 \in E^*(\sigma)$, assume that $\psi_{\theta_1}\psi_{\theta_2}$ is in the domain $D(\mathcal{L}_m)$ of the generator $\mathcal{L}_m$ of the driving process, and that $\mathcal{L}_m(\psi_{\theta_1}\psi_{\theta_2})$ has at most polynomial growth:

$$\sup_{\ell \in E} \frac{\left|\mathcal{L}_m\left(\psi_{\theta_1}\psi_{\theta_2}\right)(\ell)\right|}{1 + \|\ell\|_E^2} < \infty.$$

The assumptions above are satisfied if the driving process is a E-valued Ornstein-Uhlenbeck process, as explained below.

Example 2.2. Let $(m_\ell(t))_{\ell \in E}$ be defined by

$$dm_{\ell,t} = -(m_{\ell,t} - \overline{m})dt + dW_t, \quad m_{\ell,0} = \ell, \tag{7}$$

where W is an E-valued Wiener process. It satisfies Assumption 2 since it is ergodic and its unique invariant distribution is a normal distribution, hence integrable. Moreover, we have

$$m_\ell(t) = \ell e^{-t} + \overline{m}\left(1 - e^{-t}\right) + \int_0^t e^{s-t} dW_s. \tag{8}$$

Assumption 3 is satisfied for any $\gamma \in (2, \infty)$. The coupling $(m^*_{\ell_1}, m^*_{\ell_2})$ of Assumption 4 is obtained by driving both processes by the same Wiener process W. Indeed, (8) becomes

$$m^*_{\ell_1}(t) - m^*_{\ell_2}(t) = (\ell_1 - \ell_2)e^{-t},$$

and Assumption 4 is satisfied with $\gamma_{\mathrm{mix}}(t) = e^{-t}$. Finally, with the notation of Assumption 5, we have

$$\psi_\theta(\ell) = \int_0^\infty \left(\theta(\ell) - \theta(\overline{m})\right) e^{-t} dt = \theta(\overline{m}) - \theta(\ell).$$

Since the infinitesimal generator is given by $\mathcal{L}_m \varphi(\ell) = D\varphi(\ell) \cdot (\overline{m} - \ell) + \frac{1}{2} \mathrm{Tr}\left(D^2\varphi(\ell)\right)$, Assumption 5 is also satisfied.

3 Main result

3.1 Description of the model and of the limit problem

The multiscale stochastic problem considered in this article depends on two parameters ε and δ. Since the objective of this work is to prove a convergence result when $(\varepsilon,\delta)\to(0,0)$, without loss of generality it is assumed that $\varepsilon\in(0,\varepsilon_0]$ and $\delta\in(0,\delta_0]$, where ε_0,δ_0 are fixed – a precise condition is stated below. To simplify notation, we use the following convention: $\left(X^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ stands for the family of random variables $\left(X^{\varepsilon,\delta}\right)_{\varepsilon\in(0,\varepsilon_0],\delta\in(0,\delta_0]}$.

First, for all $\delta\in(0,\delta_0]$, the fast driving process m^δ is defined as follows: for all $t\geqslant 0$, set

$$m^\delta(t)=m(t/\delta^2), \tag{9}$$

where m is the driving process given by Assumption 2, with the initial condition $m^\delta(0)=m(0)=\ell_0$, which is assumed to be independent of δ.

We study the asymptotic behavior when $(\varepsilon,\delta)\to(0,0)$ of the solution $f^{\varepsilon,\delta}$ of the following stochastic kinetic problem

$$\partial_t f^{\varepsilon,\delta}+\left(\frac{1}{\varepsilon}a(v)+b(v)\right)\cdot\nabla_x f^{\varepsilon,\delta}+\sigma(m^\delta)f^{\varepsilon,\delta}=\frac{1}{\varepsilon^2}Lf^{\varepsilon,\delta}. \tag{10}$$

with initial condition $f^{\varepsilon,\delta}(0)=f_0^{\varepsilon,\delta}$.

For any fixed $\varepsilon\in(0,\varepsilon_0]$, $\delta\in(0,\delta_0]$, the problem (10) is globally well-posed in the following sense.

Proposition 3.1. *Introduce the linear operator $A^\varepsilon=\varepsilon^{-1}A+B$, with domain $D(A^\varepsilon)\doteq\left\{f\in L^2(\mathcal{M}^{-1})\mid(x,v)\mapsto\left(\varepsilon^{-1}a(v)+b(v)\right)\cdot\nabla_x f(x,v)\in L^2(\mathcal{M}^{-1})\right\}$, for all $\varepsilon\in(0,\varepsilon_0]$.*

Let $T\in(0,\infty)$, $\varepsilon\in(0,\varepsilon_0]$ and $\delta\in(0,\delta_0]$. For any $f_0^{\varepsilon,\delta}\in L^2(\mathcal{M}^{-1})$, there exists, almost surely, a unique mild solution $f^{\varepsilon,\delta}$ of (10) in $\mathcal{C}^0([0,T];L^2(\mathcal{M}^{-1}))$, in the sense that, almost surely, for all $t\in[0,T]$, one has

$$f^{\varepsilon,\delta}(t)=e^{-tA^\varepsilon}f_0^{\varepsilon,\delta}+\int_0^t e^{-(t-s)A^\varepsilon}\left(\frac{1}{\varepsilon^2}Lf^{\varepsilon,\delta}(s)-\sigma(m^\delta(s))f^{\varepsilon,\delta}(s)\right)ds.$$

The proof of Proposition 3.1 is based on a standard fixed point argument, combined with the following observation:

$$\sup_{t\in[0,T]}\left\|m^\delta(t)\right\|_E\leqslant\sup_{i\leqslant T\delta^{-2}+1}\sup_{t\in[i,i+1]}\|m(t)\|_E<\infty$$

owing to Assumption 3 on the moments of the driving process m. The proof of Proposition 3.1 is thus omitted.

Note that the statement of Proposition 3.1 is given for fixed $\varepsilon>0$ and $\delta>0$, and does not provide uniform estimates of the solution $f^{\varepsilon,\delta}$ with respect to these parameters. Proving such estimates needs extra arguments, which are not needed to state the main results of the article. We refer to Proposition... below for the statement of an appropriate a priori estimate in the $L^2(\mathcal{M}^{-1})$ norm, which requires the introduction of a stopping time τ^δ defined below.

Let us now introduce the so-called averaged equation. First, set

$$\begin{aligned}K&=\int_V a(v)\otimes a(v)\mathcal{M}(v)d\mu(v)\in\mathrm{Sym}^+(d)\\ J&=\int_V b(v)\mathcal{M}(v)d\mu(v)\in\mathbb{R}^d\\ \overline{\sigma}&=\int_E\sigma(\ell)d\nu(\ell)\in\mathcal{C}_x^{\lfloor d/2\rfloor+2}.\end{aligned} \tag{11}$$

Note that K and J are well-defined since a and b are assumed to be bounded (see Assumption 1). In addition, $\overline{\sigma}$ is well-defined since σ is globally Lipschitz continuous from E to $\mathcal{C}_x^{\lfloor d/2 \rfloor + 2}$ (see Assumption 1) and since the probability distribution ν is integrable (see Assumption 2).

The unknown $\overline{\rho}$ of the averaged equation is a mapping defined on $[0, \infty) \times \mathbb{T}^d$. We are finally in position to write the averaged equation:

$$\partial_t \overline{\rho} + J \cdot \nabla \overline{\rho} + \overline{\sigma}\, \overline{\rho} = \operatorname{div}(K \nabla \overline{\rho}), \tag{12}$$

with initial condition $\overline{\rho}(0) = \overline{\rho}_0$. In (12) above, ∇ and div are the gradient and divergence operators with respect to the variable x respectively. We consider solutions of (12) in the weak sense, see Definition 3.1 below. Note that the solution may be a random process, even if the evolution is deterministic: it may happen that the initial condition $\overline{\rho}_0$ is random.

Definition 3.1. Let $T \in (0, \infty)$ and $\overline{\rho}_0$ be a L_x^2-valued random variable. A stochastic process $\overline{\rho}$ is a weak solution of (12) in L_x^2 if $\overline{\rho} \in L_T^\infty L_x^2$ almost surely and if, for all $\xi \in H_x^2$ and $t \in [0, T]$, almost surely,

$$(\overline{\rho}(t), \xi)_{L_x^2} = (\overline{\rho}_0, \xi)_{L_x^2} + \int_0^t (\overline{\rho}(s), \operatorname{div}(K \nabla \xi) + J \cdot \nabla \xi - \overline{\sigma} \xi)_{L_x^2}\, ds. \tag{13}$$

For any L_x^2-valued random variable, the averaged equation (12) admits a unique weak solution in the sense of Definition 3.1.

3.2 Convergence results

Let us now state the main results of this article, concerning the asymptotic behavior of $\rho^{\varepsilon,\delta}$ and $f^{\varepsilon,\delta}$ when $(\varepsilon, \delta) \to (0, 0)$. In the sequel, it is convenient to impose the following non restrictive conditions on the parameters ε_0 and δ_0 (such that $\varepsilon \in (0, \varepsilon_0]$ and $\delta \in (0, \delta_0]$):

$$\varepsilon_0 < \min\left(1, \left(4(\|a\|_{L^\infty} + \|b\|_{L^\infty})(1 + T\, \|\overline{\sigma}\|_{\mathcal{C}_x^1})\right)^{-1}\right), \quad \delta_0 < \min\left(1, \|\ell_0\|_E^{-1}\right). \tag{14}$$

Note that the condition for δ_0 depends on the initial condition ℓ_0 of the driving process. This is one of the reasons why it is convenient to assume that it is deterministic and that it does not depend on δ. Extensions to more general initial conditions for the driving process would require extra technical assumptions and computations, which are omitted to simplify the setting and focus on the main aspects of the analysis of the asymptotic behavior of the stochastic multiscale problem (10) when $(\varepsilon, \delta) \to (0, 0)$.

The initial condition $f_0^{\varepsilon,\delta}$ of (10) is assumed to satisfy the following conditions.

Assumption 6. The family of initial conditions $\left(f_0^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ satisfies the following moment bound:

$$\sup_{\varepsilon \in (0,\varepsilon_0], \delta \in (0,\delta_0]} \mathbb{E}\left[\left\| f_0^{\varepsilon,\delta} \right\|_{L^2(\mathcal{M}^{-1})}^{12}\right] < \infty.$$

In addition, the initial density $\rho_0^{\varepsilon,\delta} \in L_x^2$ converges in distribution when $(\varepsilon, \delta) \to (0, 0)$ in L_x^2 to $\overline{\rho}_0 \in L_x^2$: recall that this means that for any bounded continuous mapping $\Phi : L_x^2 \to \mathbb{R}$, one has

$$\mathbb{E}\left[\Phi(\rho_0^{\varepsilon,\delta})\right] \xrightarrow[\varepsilon,\delta \to 0]{} \mathbb{E}\left[\Phi(\overline{\rho}_0)\right].$$

Observe that in general $\overline{\rho}_0$ is a L^2_x-valued random variable.

We are now in position to state the two main convergence results of this article. First, see Theorem 3.2, $\rho^{\varepsilon,\delta} = \left\langle f^{\varepsilon,\delta} \right\rangle$ converges in distribution to the unique solution $\overline{\rho}$ of the averaged equation (12), in the space $\mathcal{C}^0_T H^{-\varsigma}_x$ for all $\varsigma > 0$. Second, under an additional technical assumption which allows to apply an averaging lemma, one obtains a stronger result, see Theorem 3.3: $\rho^{\varepsilon,\delta}$ converges in distribution to $\overline{\rho}$ in the space $L^2_T L^2_x$, and $f^{\varepsilon,\delta}$ converges in distribution to $\overline{\rho}\mathcal{M}$ in the space $L^2_T L^2(\mathcal{M}^{-1})$. Moreover, the convergence results hold in probability if the initial condition $\overline{\rho}_0$ (given by Assumption 6) of the averaged equation (12) is deterministic.

Let us state the first main result of this article.

Theorem 3.2. *Let Assumptions 1 to 6 be satisfied. Let $f^{\varepsilon,\delta}$ be the solution of (10). Then, when $(\varepsilon,\delta) \to (0,0)$, the random variable $\rho^{\varepsilon,\delta} = \left\langle f^{\varepsilon,\delta} \right\rangle$ converges in distribution to the unique weak solution $\overline{\rho}$ of (12), in $\mathcal{C}^0_T H^{-\varsigma}_x$.*

In addition, if the initial condition $\overline{\rho}_0 = \lim_{(\varepsilon,\delta)\to(0,0)} \rho_0^{\varepsilon,\delta} \in L^2_x$ is deterministic, then the convergence of $\rho^{\varepsilon,\delta}$ to $\overline{\rho}$ holds in probability.

An additional assumption is required to state Theorem 3.3, as in [RR20] (Assumption 7 below). It allows us to apply a so-called averaging lemma (see [BD99, Theorem 2.3]), developed for the study of kinetic PDEs, and to obtain convergence in the space L^2_x (in Theorem 3.3) instead of $H^{-\varsigma}$ (in Theorem 3.2).

Assumption 7.

- The space for the velocity variable is given by $(V, d\mu) = (\mathbb{R}^n, \frac{d\mu}{dv}(v)dv)$, with a Radon-Nikodym derivative (with respect to Lebesgue measure) satisfying $\frac{d\mu}{dv} \in H^1(\mathbb{R}^n)$.
- The mapping a is locally Lipschitz continuous
- There exist $C_{a,\mu} \in (0,\infty)$ and $\varsigma_{a,\mu} \in (0,1]$ such that, for all $u \in S^{d-1}$, all $\lambda \in \mathbb{R}$ and all $\eta \in (0,\infty)$, one has

$$\int_{\lambda < a(v)\cdot u < \lambda+\eta} \left(\left| \frac{d\mu}{dv}(v) \right|^2 + \left| \nabla \frac{d\mu}{dv}(v) \right|^2 \right) dv \leqslant C_{a,\mu} \eta^{\varsigma_{a,\mu}}.$$

We are now in position to state the second main result of this article.

Theorem 3.3. *Let Assumptions 1 to 6 and Assumption 7 be satisfied.*

Then, when $(\varepsilon,\delta) \to (0,0)$, the random variable $\rho^{\varepsilon,\delta} = \left\langle f^{\varepsilon,\delta} \right\rangle$ converges in distribution to the unique weak solution $\overline{\rho}$ of (12), in the space $L^2_T L^2_x$. Moreover, when $(\varepsilon,\delta) \to (0,0)$, the random variable $f^{\varepsilon,\delta}$ converges in distribution to $\overline{\rho}\mathcal{M}$, in the space $L^2_T L^2(\mathcal{M}^{-1})$.

In addition, if the initial condition $\overline{\rho}_0 = \lim_{(\varepsilon,\delta)\to(0,0)} \rho_0^{\varepsilon,\delta} \in L^2_x$ is deterministic, then the convergence of $\rho^{\varepsilon,\delta}$ to $\overline{\rho}$ and of $f^{\varepsilon,\delta}$ to $\overline{\rho}\mathcal{M}$ hold in probability.

3.3 Discussion

The main results, Theorems 3.2 and 3.3, state that diffusion approximation (in the PDE sense) and averaging (in the probabilistic sense) results hold, when $(\varepsilon,\delta) \to (0,0)$. These results are natural generalizations of previously obtained results, either in the deterministic case ($\varepsilon \to 0$, $\sigma = 0$), or in the probabilistic case ($\delta \to 0$, $\varepsilon > 0$ fixed). In fact, using the same arguments as in Section 4 below, one may obtain the following results, where the limits $\varepsilon \to 0$ and $\delta \to 0$ are taken successively.

On the one hand, if $\delta > 0$ is fixed, then $\rho^{\varepsilon,\delta} = \langle f \rangle^{\varepsilon,\delta}$ converges when $\varepsilon \to 0$, to the solution $\rho^{0,\delta}$ of the evolution equation

$$\partial_t \rho^{0,\delta} + \sigma(m^\delta)\rho^{0,\delta} = \mathrm{div}(K\nabla\rho^{0,\delta}) - J \cdot \nabla\rho^{0,\delta}.$$

That result is a standard diffusion approximation result in the PDE sense. Then, when $\delta \to 0$, $\rho^{0,\delta}$ converges to the solution $\overline{\rho}$ of the limit equation (12), owing to the standard averaging principle for stochastic problems.

On the other hand, if $\varepsilon > 0$ is fixed, then owing to the standard averaging principle, $f^{\varepsilon,\delta}$ converges when $\delta \to 0$, to the solution $f^{\varepsilon,0}$ of the evolution equation

$$\partial_t f^{\varepsilon,0} + \frac{1}{\varepsilon} a(v) \cdot \nabla_x f^{\varepsilon,0} + b(v) \cdot \nabla_x f^{\varepsilon,0} + \overline{\sigma} f^{\varepsilon,0} = \frac{1}{\varepsilon^2} L f^{\varepsilon,0}.$$

Then, when $\varepsilon \to 0$, $\rho^{\varepsilon,0} = \langle f \rangle^{\varepsilon,0}$ converges to the solution $\overline{\rho}$ of the limit equation (12), owing to the standard diffusion approximation result in the PDE sense.

To the best of our knowledge, the results above have not been rigorously proved in the literature, however they are variants of well-studied results. The proofs of Theorems 3.2 and 3.3 do not encompass those regimes when either $\varepsilon = 0$ or $\delta = 0$: essentially this would require to adapt the construction of the perturbed test function. Indeed, below we directly focus on the behavior of $\rho^{\varepsilon,\delta}$, thus the convergence $f^{\varepsilon,\delta} \to f^{\varepsilon,0}$ when $\delta \to 0$ cannot be covered directly, for instance. In addition, one of the arguments of the proofs is the convergence $\tau^\delta \to \infty$ when $\delta \to 0$ (where τ^δ are stopping times defined below), thus the convergence $\rho^{\varepsilon,\delta} \to \rho^{0,\delta}$ when $\varepsilon \to 0$ cannot be covered without substantial modifications.

Still, one obtains the following result:

$$\lim_{\delta\to 0} \lim_{\varepsilon\to 0} \rho^{\varepsilon,\delta} = \lim_{\varepsilon\to 0} \lim_{\delta\to 0} \rho^{\varepsilon,\delta} = \lim_{(\varepsilon,\delta)\to(0,0)} \rho^{\varepsilon,\delta} = \overline{\rho},$$

where the convergence is understood in the appropriate sense. The analysis presented in this article thus departs from the setting of [DV12, DV20, DRV20, RR20]. where in all cases there is only one small parameter $\varepsilon = \delta$. Our result is expected but important: it shows that the diffusion approximation and the averaging principle can be decoupled. As already mentioned in the introduction, considering the general case $\varepsilon \neq \delta$ requires new arguments, in particular the construction of the perturbed test functions needs an additional corrector.

In the setting, it is assumed that the initial condition $m^\delta(0)$ of the fast driving process is deterministic and independent of δ: $m^\delta(0) = m_0$. It would be possible to extend the results to more general initial conditions, under appropriate modified moment conditions. This would for instance allow us to include the case where $m^\delta(0)$ is random and distributed following the ergodic invariant distribution ν. Note also that if either the mapping σ or the process m would have a bounded support, the analysis would be simplified.

As explained in the introduction, we have left open several questions for future works: first, the analysis of fluctuations, second, the identification of (strong and weak) rates of convergence.

4 Description of the proof

In the sequel, the following convention is employed: given variables u and parameters λ, the notation $X_1(u) \lesssim X_2(u)$ means that for all parameters λ, there exists $C(\lambda) \in (0, \infty)$ such that one has $X_1(u) \leqslant C(\lambda)X_2(u)$ for all u. From the context the identification of variables (typically $\varepsilon, \delta, f, \ell$) and parameters (typically T, φ) will be clear.

First, we introduce two of the most important tools of the proofs of the main results: the stopping time τ^δ (see Section 4.1), and the perturbed test function $\varphi^{\varepsilon,\delta} = \varphi + \varepsilon\varphi_{1,0} + \varepsilon^2\varphi_{2,0} + \delta^2\varphi_{0,2} + \varepsilon\delta^2\varphi_{1,2}$ (see Section 4.2), constructed for a class of admissible functions φ such that $\varphi(f,\ell) = \varphi(\langle f\rangle) = \varphi(\rho)$. Sections 4.1 and 4.2 contain important auxiliary results, which require technical arguments in their proofs: the proofs are therefore postponed to Section 5.

The use of the stopping time τ^δ is instrumental to obtain appropriate moment bounds of the solutions, uniformly with respect to ε, δ. We prove that $\tau^\delta \to \infty$ in probability when $\delta \to 0$. In the arguments, as a consequence of Slutsky's lemma (see more details below), it is then sufficient to consider the stopped processes defined by $f^{\varepsilon,\delta,\tau^\delta} = f^{\varepsilon,\delta}(\cdot \wedge \tau^\delta)$, $\rho^{\varepsilon,\delta,\tau^\delta} = \rho^{\varepsilon,\delta}(\cdot \wedge \tau^\delta)$ and $m^{\delta,\tau^\delta} = m^\delta(\cdot \wedge \tau^\delta)$.

The use of the perturbed test function method is standard in the analysis of multiscale stochastic problems, see for instance [DV12, DV20, DRV20, RR20] in a similar context of stochastic kinetic equations. Compared to those references, note that it is necessary to consider two parameters ε, δ which may be independent. When $\delta = \varepsilon$, the construction of the corrector $\varphi_{1,2}$ is not needed.

We then proceed to the proof of the main results of this article. The arguments are standard. We first check a tightness property for $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta,\tau^\delta}$ in the appropriate function space. We then check that the Markov process $(f^{\varepsilon,\delta,\tau^\delta}, m^{\delta,\tau^\delta})$ is the solution of a martingale problem for all $\varepsilon > 0, \delta > 0$, and letting $(\varepsilon,\delta) \to (0,0)$, using the perturbed test function, we prove that any limit point of the family $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$ is a weak solution of the averaged equation, in the sense of (13). Since $\tau^\delta \to \infty$ in probability when $\delta \to 0$, a uniqueness argument for the averaged equation then concludes the proof of Theorem 3.2. Theorem 3.3 is then obtained by the application of the averaging lemma. The fact that convergence holds in probability when $\overline{\rho}_0$ is deterministic is a straightforward well-known consequence of Portmanteau Theorem (and is not specific to the PDE framework of this paper): in that case the solution $\overline{\rho}$ of the averaged equation is also deterministic.

4.1 Stopping times and a priori estimates

In this section, first we give the definition of the stopping time τ^δ, second we state that $\tau^\delta \to \infty$ in probability when $\delta \to 0$. Finally, we state an a priori estimate for $f^{\varepsilon,\delta,\tau^\delta}$ in $L^2(\mathcal{M}^{-1})$, which is uniform with respect to ε, δ. The proofs are postponed to Section 5.

For all $\delta \in (0, \delta_0]$ and all $t \geqslant 0$, set

$$\zeta^\delta(t) = \frac{1}{\delta}\int_0^t \left(\sigma(m^\delta(s)) - \overline{\sigma}\right) ds \in \mathcal{C}_x^{\lfloor d/2\rfloor+2}. \tag{15}$$

Definition 4.1. Let $\alpha \in (\frac{2}{\gamma}, 1)$, then for all $\delta \in (0, \delta_0]$, define

$$\tau^\delta = \tau_m^\delta \wedge \tau_\zeta^\delta,$$

where

$$\tau_m^\delta = \inf\left\{t \in [0,T] \mid \left\|m^\delta(t)\right\|_E \geqslant \delta^{-\alpha}\right\},$$
$$\tau_\zeta^\delta = \inf\left\{t \in [0,T] \mid \left\|\zeta^\delta(t)\right\|_{\mathcal{C}_x^1} \geqslant \delta^{-1}\right\}.$$

Note that τ^δ is a stopping time for the filtration $\left(\mathcal{F}_t^\delta\right)_{t\in\mathbb{R}^+}$ generated by the driving process m^δ, with $\mathcal{F}_t^\delta = \sigma\left(m^\delta(s)\right)_{0\leqslant s\leqslant t}$. In the definition above, γ is the parameter introduced in Assumption 3.

The initial conditions of the stopped processes m^{δ,τ^δ} and $\zeta^{\delta,\tau^\delta}$ satisfy $\left\|m^{\delta,\tau^\delta}(0)\right\|_E = \|\ell_0\|_E \leqslant \delta_0^{-1} \leqslant \delta^{-\alpha}$ (since $\delta \leqslant \delta_0 \leqslant 1$ and $\alpha < 1$), and $\zeta^{\delta,\tau^\delta}(0) = 0$. As a consequence, almost surely $\tau^\delta > 0$, and the following estimates for m^δ and ζ^δ hold: for all $t \geqslant 0$ and all $\delta \in (0, \delta_0]$, one has

$$\left\|m^{\delta,\tau^\delta}(t)\right\|_E = \left\|m^\delta(t \wedge \tau^\delta)\right\|_E \leqslant \delta^{-\alpha}, \tag{16}$$

and

$$\left\|\zeta^{\delta,\tau^\delta}(t)\right\|_{\mathcal{C}_x^1} = \left\|\zeta^\delta(t \wedge \tau^\delta)\right\|_{\mathcal{C}_x^1} \leqslant \delta^{-1}. \tag{17}$$

Let us now study the behavior of the stopping time τ^δ when $\delta \to 0$.

Proposition 4.1. *When $\delta \to 0$, $\tau^\delta \to \infty$ in probability: for all $T > 0$, one has*

$$\mathbb{P}\left(\tau^\delta < T\right) \xrightarrow[(\varepsilon,\delta)\to 0]{} 0.$$

The proof of Proposition 4.1 is postponed to Section 5.

Finally, let us state the following a priori estimate in the $L^2(\mathcal{M}^{-1})$ norm for $f^{\varepsilon,\delta,\tau^\delta}(t)$, in an almost sure sense.

Proposition 4.2. *For $T \in (0,\infty)$, there exists $C(T) \in (0,\infty)$, such that for all $t \in [0,T]$, $\varepsilon \in (0,\varepsilon_0]$ and $\delta \in (0,\delta_0]$, almost surely one has*

$$\left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|^2_{L^2(\mathcal{M}^{-1})} + \frac{1}{2\varepsilon^2}\int_0^{t\wedge\tau^\delta} \left\|Lf^{\varepsilon,\delta,\tau^\delta}(s)\right\|^2_{L^2(\mathcal{M}^{-1})} ds \leqslant C(T)\left\|f_0^{\varepsilon,\delta}\right\|^2_{L^2(\mathcal{M}^{-1})}. \tag{18}$$

Let us emphasize that the constant $C(T)$ appearing on the right-hand side of (18) is deterministic and does not depend on ε and δ.

The proof of Proposition 4.2 is posponed to Section 5. Note that the a priori estimate $\left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})} \lesssim \left\|f_0^{\varepsilon,\delta}\right\|_{L^2(\mathcal{M}^{-1})}$ is instrumental in all the arguments of the proof of Theorem 3.2, whereas the upper bound for the integral term on the left-hand side of (18) is used only in the proof of Theorem 3.3.

4.2 Perturbed test functions

In this section, we describe the construction of a perturbed test function $\varphi^{\varepsilon,\delta}$, such that the following properties are satisfied:

$$\begin{aligned} \varphi^{\varepsilon,\delta} &\xrightarrow[\varepsilon,\delta\to 0]{} \varphi \\ \mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} &\xrightarrow[\varepsilon,\delta\to 0]{} \mathcal{L}\varphi, \end{aligned}$$

where φ is any sufficiently smooth function such that $\varphi(f,\ell) = \varphi(\langle f\rangle)$, and $\mathcal{L}^{\varepsilon,\delta}$ and $\mathcal{L}$ are the infinitesimal generators associated with the Markov processes $(f^{\varepsilon,\delta}, m^\delta)$ and $\overline{\rho}$, solving (10) and (12) respectively.

To state more rigorously and more precisely the properties mentioned above, we first introduce two appropriate classes of test functions, such that $\varphi \in \Theta_{\text{lim}}$ and $\varphi^{\varepsilon,\delta} \in \Theta$ for all ε, δ, and such that the errors $|\varphi^{\varepsilon,\delta} - \varphi|$ and $|\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} - \mathcal{L}\varphi|$ are quantified in terms of ε, δ.

Let us first describe the class of functions Θ_{lim}.

Definition 4.2. Let $\Theta_{\lim}$ be the class of real-valued test functions φ such that for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, one has

$$\varphi(f,\ell) = \varphi(\rho) = \chi((\rho,\xi)_{L^2_x}), \tag{19}$$

where we recall that $\rho = \langle f \rangle = \int_V f d\mu$, with arbitrary $\chi \in \mathcal{C}^3_b(\mathbb{R},\mathbb{R})$ and $\xi \in \mathcal{C}^3_x$.

Recall that K, J and $\overline{\sigma}$ are defined by (11). The infinitesimal generator $\mathcal{L}$ associated with the limit problem (12) is defined by

$$\mathcal{L}\varphi(\rho) = D\varphi(\rho) \cdot (\mathrm{div}_x(K\nabla_x\rho) - J \cdot \nabla_x\rho - \overline{\sigma}\rho)\,, \tag{20}$$

for all $\rho \in L^2_x$ and all $\varphi \in D(\mathcal{L})$, with the domain $D(\mathcal{L})$ given by

$$D(\mathcal{L}) = \left\{\varphi \in \mathcal{C}^0(L^2_x) \mid \mathcal{L}\varphi \in \mathcal{C}^0(L^2_x)\right\}.$$

Note that $\Theta_{\lim} \subset D(\mathcal{L})$. In addition, for all $\varphi \in \Theta_{\lim}$, of the form (19), one has for all $\rho \in L^2_x$

$$\mathcal{L}\varphi(\rho) = \chi'((\rho,\xi)_{L^2_x})\,(\rho, \mathrm{div}_x(K\nabla_x\xi) + J \cdot \nabla_x\xi - \overline{\sigma}\xi)_{L^2_x}\,. \tag{21}$$

Let us now describe the class of functions Θ.

Definition 4.3. Let Θ be the class of test functions $\varphi : L^2(\mathcal{M}^{-1}) \times E \to \mathbb{R}$ satisfying the following conditions.

- For all $\ell \in E$, $\varphi(\cdot,\ell) \in \mathcal{C}^1(L^2(\mathcal{M}^{-1}))$.
- For all $f \in L^2(\mathcal{M}^{-1})$, $\varphi(f,\cdot) \in \mathcal{C}^0(E)$.
- For all $i \in \{1,2\}$ and all $f \in L^2(\mathcal{M}^{-1})$, $\varphi(f,\cdot)^i \in D(\mathcal{L}_m)$ and $\mathcal{L}_m(\varphi^i) \in \mathcal{C}^0(L^2(\mathcal{M}^{-1}) \times E)$.
- For all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, denote by $\nabla_f\varphi(f,\ell) \in L^2(\mathcal{M}^{-1})$ the gradient of φ at (f,ℓ), which is defined defined such that

  $$\forall h \in L^2(\mathcal{M}^{-1}), (\nabla_f\varphi(f,\ell),h)_{L^2(\mathcal{M}^{-1})} = D_f\varphi(f,\ell)\cdot h.$$

  Then, for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, one has

  $$\int\int \|\nabla_x\nabla_f\varphi(f,\ell)(x,v)\|^2\, dx\frac{d\mu(v)}{\mathcal{M}(v)} < \infty.$$

- There exists $C_\varphi(0,\infty)$ such that, for all $f,h \in L^2(\mathcal{M}^{-1})$ and $\ell_1,\ell_2 \in E$,

  $$\begin{aligned}|\varphi(f,\ell_1)| + |D_f\varphi(f,\ell_1)(Ah)| + |D_f\varphi(f,\ell_1)(Bh)| + |D_f\varphi(f,\ell_1)(\ell_2 h)| + |D_f\varphi(f,\ell_1)(Lh)| \\ \leqslant C_\varphi\left(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})} + \|h\|^3_{L^2(\mathcal{M}^{-1})}\right)\left(1 + \|\ell_1\|^2_E + \|\ell_2\|^2_E\right).\end{aligned}$$

The infinitesimal generator $\mathcal{L}^{\varepsilon,\delta}$ associated with the stochastic problem (10) with driving process m^δ, satisfies the following multiscale expansion in terms of ε,δ:

$$\mathcal{L}^{\varepsilon,\delta} = \mathcal{L}_0 + \varepsilon^{-1}\mathcal{L}_1 + \varepsilon^{-2}\mathcal{L}_2 + \delta^{-2}\mathcal{L}_m \tag{22}$$

where the infinitesimal generator of the driving process $\mathcal{L}_m$ is introduced in Assumption 2 above, and $\mathcal{L}_0$, $\mathcal{L}_1$ and $\mathcal{L}_2$ are defined as follows: for all $f \in L^2(\mathcal{M}^{-1})$ and all $\ell \in E$, set

$$\begin{aligned}\mathcal{L}_0\varphi(f,\ell) &= -D_f\varphi(f,\ell)\cdot(\sigma(\ell)f + Bf),\\ \mathcal{L}_1\varphi(f,\ell) &= -D_f\varphi(f,\ell)\cdot(Af),\\ \mathcal{L}_2\varphi(f,\ell) &= D_f\varphi(f,\ell)\cdot Lf,\end{aligned}$$

for real-valued functions $\varphi \in D(\mathcal{L}^{\varepsilon,\delta})$ where

$$D(\mathcal{L}^{\varepsilon,\delta}) = \left\{\varphi \in \mathcal{C}^0(L^2(\mathcal{M}^{-1}) \times E) \mid \mathcal{L}^{\varepsilon,\delta}\varphi \in \mathcal{C}^0(L^2(\mathcal{M}^{-1}) \times E)\right\}$$

is the domain of the unbounded linear operator $\mathcal{L}^{\varepsilon,\delta}$. Recall that $A = a(v) \cdot \nabla_x$ and $B = b(v) \cdot \nabla_x$ (see Assumption 1).

Note that the following property is satisfied: for all $\varphi \in \Theta$, one has $\varphi \in D(\mathcal{L}^{\varepsilon,\delta})$ and $\varphi^2 \in D(\mathcal{L}^{\varepsilon,\delta})$. In addition, observe that $\Theta_{\lim} \subset \Theta$.

We are now in position to state the main result of this section.

Proposition 4.3. *For all $\varphi \in \Theta_{\lim}$, there exists functions $\varphi_{1,0}$, $\varphi_{2,0}$, $\varphi_{0,2}$ and $\varphi_{1,2}$ such that the following properties hold.*

- *Set*
$$\varphi^{\varepsilon,\delta} = \varphi + \varepsilon\varphi_{1,0} + \varepsilon^2\varphi_{2,0} + \delta^2\varphi_{0,2} + \varepsilon\delta^2\varphi_{1,2}. \tag{23}$$
Then $\varphi^{\varepsilon,\delta} \in \Theta$.

- *There exists $C(\varphi) \in (0,\infty)$, such that for all $\varepsilon \in (0,\varepsilon_0]$, $\delta \in (0,\delta_0]$ and all $(f,\ell) \in L^2(\mathcal{M}^{-1}) \times E$, one has*
$$\left|\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} - \mathcal{L}\varphi\right|(f,\ell) \leqslant C(\varphi)(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})})\Big(\varepsilon(1 + \|\ell\|_E) + \delta^2(1 + \|\ell\|^2_E)\Big). \tag{24}$$

- *One has the following upper bounds: there exists $C(\varphi) \in (0,\infty)$, such that for all $(f,\ell) \in L^2(\mathcal{M}^{-1}) \times E$,*
$$|\varphi_{1,0}(f)| \leqslant C(\varphi)\,\|f\|_{L^2(\mathcal{M}^{-1})}, \tag{25}$$
$$|\varphi_{2,0}(f)| \leqslant C(\varphi)(1 + \|f\|^2_{L^2(\mathcal{M}^{-1})}), \tag{26}$$
$$|\varphi_{0,2}(f,\ell)| \leqslant C(\varphi)\,\|f\|_{L^2(\mathcal{M}^{-1})}\,(1 + \|\ell\|_E), \tag{27}$$
$$|\varphi_{1,2}(f,\ell)| \leqslant C(\varphi)(1 + \|f\|^2_{L^2(\mathcal{M}^{-1})})(1 + \|\ell\|_E). \tag{28}$$

Note that the following error estimates holds: for all $(f,\ell) \in L^2(\mathcal{M}^{-1}) \times E$, one has

$$\left|\varphi^{\varepsilon,\delta}(f,\ell) - \varphi(\langle f\rangle)\right| \lesssim_\varphi (1 + \|f\|^2_{L^2(\mathcal{M}^{-1})})(\varepsilon + \delta^2(1 + \|\ell\|_E)). \tag{29}$$

The proof of Proposition 4.3 is postponed to Section 5. Using the standard terminology, the functions $\varphi_{1,0}$, $\varphi_{2,0}$, $\varphi_{0,2}$ and $\varphi_{1,2}$ are referred to as the correctors in the sequel.

Note that one of the novelties of the result above is the construction of the corrector $\varphi_{1,2}$, which is not required in the case $\varepsilon = \delta$ which is treated in other contributions, see [DV12, DV20, DRV20, RR20]. More precisely, if ε, δ, it suffices to construct a perturbed test function of the type $\varphi^\varepsilon = \varphi + \varepsilon\varphi_1 + \varepsilon^2\varphi_2$. In fact, $\varphi_1 = \varphi_{1,0}$ and $\varphi_2 = \varphi_{2,0} + \varphi_{0,2}$.

4.3 Martingale property

The proofs of Theorems 3.2 and 3.3 is based on an interpretation in terms of solutions of martingale problems. Indeed, combined with the perturbed test function approach described above, this formulation is convenient to identify limit points when $(\varepsilon,\delta) \to (0,0)$.

Several arguments in the proofs of the auxiliary results below employ the following auxiliary result.

Proposition 4.4. *Let $\varphi \in \Theta$. For all $t \in \mathbb{R}^+$, set*

$$M_\varphi^{\varepsilon,\delta}(t) \doteq \varphi(f^{\varepsilon,\delta}(t), m^\delta(t)) - \varphi(f^{\varepsilon,\delta}(0), m^\delta(0)) - \int_0^t \mathcal{L}^{\varepsilon,\delta}\varphi(f^{\varepsilon,\delta}(s), m^\delta(s))ds. \tag{30}$$

Then, $M_\varphi^{\varepsilon,\delta,\tau^\delta}$ is a càdlàg $\left(\mathcal{F}_t^\delta\right)_{t\in\mathbb{R}^+}$-martingale. In addition, for all $t \in \mathbb{R}^+$, one has

$$\begin{aligned}\mathbb{E}\left[\left|M_\varphi^{\varepsilon,\delta,\tau^\delta}(t)\right|^2\right] &= \mathbb{E}\left[\int_0^{t\wedge\tau^\delta} \left(\mathcal{L}^{\varepsilon,\delta}(\varphi^2) - 2\varphi\mathcal{L}^{\varepsilon,\delta}\varphi\right)(f^{\varepsilon,\delta}(s), m^\delta(s))ds\right] \\ &= \frac{1}{\delta^2}\mathbb{E}\left[\int_0^{t\wedge\tau^\delta} \left(\mathcal{L}_m(\varphi^2) - 2\varphi\mathcal{L}_m\varphi\right)(f^{\varepsilon,\delta}(s), m^\delta(s))ds\right].\end{aligned}$$

In the statement of Proposition 4.4 above, note that the process $M_\varphi^{\varepsilon,\delta}$ is stopped, where the stopping time τ^δ is given by Definition 4.1. Considering the stopped process allows us to use the estimate (18) of Proposition 4.2 in the sequel. Moreover, note that φ is assumed to belong to the class of functions Θ introduced in Definition 4.3: in fact Proposition 4.4 is the justification of the requirements on φ in Definition 4.3.

The proof of Proposition 4.4 is standard and is omitted.

4.4 Proofs of Theorems 3.2 and 3.3

In order to prove the convergence in distribution results of $\rho^{\varepsilon,\delta}$ to $\overline{\rho}$ when $(\varepsilon,\delta) \to (0,0)$, it suffices to prove that for any arbitrary sequence $(\varepsilon_i,\delta_i)_{i\geqslant 1}$ such that $(\varepsilon_i,\delta_i) \to (0,0)$ when $i \to \infty$, $\rho^{\varepsilon_i,\delta_i}$ converges in distribution to $\overline{\rho}$. To simplify the notation, we fix such a sequence and in the sequel one should interpret $\varepsilon = \varepsilon_i$ and $\delta = \delta_i$. Moreover, in Proposition 4.6 and in the proofs, $(\varepsilon_i,\delta_i)_{i\geqslant 1}$ may also denote a subsequence of the original sequence.

4.4.1 Two technical results

Two additional technical results are required for the proof of the main results of this article.

First, let us state a tightness result.

Proposition 4.5. *Let Assumptions 1 to 6 be satisfied.*

The family of processes $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ is tight in the space $\mathcal{C}_T^0 H_x^{-\varsigma}$, for all arbitrarily small $\varsigma \in (0,1]$.

Moreover, if Assumption 7 is also satisfied, then the family of processes $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ is tight in the space $L_T^2 L_x^2$.

Second, let us state a result which allows us to identify limit points of the family $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$.

Proposition 4.6. *Assume that ρ_∞ is a $\mathcal{C}_T^0 H_x^{-\varsigma}$-valued random variable, such that $\rho^{\varepsilon_i,\delta_i} \xrightarrow[i\to\infty]{d} \rho_\infty$, in distribution in $\mathcal{C}_T^0 H_x^{-\varsigma}$, for some sequence $(\varepsilon_i,\delta_i) \xrightarrow[i\to\infty]{} 0, \varepsilon_i \in (0,\varepsilon_0], \delta_i \in (0,\delta_0]$.*

Then, for all $\varphi \in \Theta_{\lim}$, almost surely, for all $t \in [0,T]$, one has

$$\varphi(\rho_\infty(t)) - \varphi(\rho_\infty(0)) - \int_0^t \mathcal{L}\varphi(\rho_\infty(s))ds = 0.$$

The proofs of Propositions 4.5 and 4.6 are technical and are given below.

4.4.2 Proofs of the main results

We are now in position to prove the main results of this article.

Proof of Theorem 3.2. Let $\varsigma \in (0,1)$ be fixed. Owing to Proposition 4.5, the family of processes $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ is tight in the space $\mathcal{C}^0_T H^{-\varsigma}_x$. Due to Prohorov Theorem, there exist a $\mathcal{C}^0_t T H^{-\varsigma}_x$-valued random variable ρ_∞ and sequences $(\varepsilon_i)_{i\in\mathbb{N}}$ and $(\delta_i)_{i\in\mathbb{N}}$, such that $\varepsilon_i \to 0$ and $\delta_i \to 0$, and $\rho^{\varepsilon_i,\delta_i} \xrightarrow[i\to\infty]{d} \rho_\infty$. Let us prove that ρ_∞ is a weak solution of (12): this requires to prove that (13) holds and that $\rho_\infty \in L^\infty_T L^2_x$ almost surely.

To obtain (13), it suffices to use Proposition 4.6 and a localization argument. Introduce the auxiliary functions χ^r, for all $r \geqslant 0$, such that $\chi^r \in \mathcal{C}^3_b(\mathbb{R})$, χ^r is odd and

$$\forall u \in [0,r], \chi^r(u) = u,$$
$$\forall u \in [r+1,\infty), \chi^r(u) = r+1.$$

Let $\xi \in H^2_x$ and define the test function $\varphi^r \in \Theta_{\mathrm{lim}}$ by

$$\varphi^r(\rho) = \chi^r\left((\rho,\xi)_{L^2_x}\right).$$

Since $\rho_\infty \in \mathcal{C}^0_T H^{-\varsigma}_x$ almost surely, the random variable

$$S \doteq \sup_{t\in[0,T]} \left|(\rho_\infty(t),\xi)_{L^2_x}\right|,$$

is finite almost surely. On the one hand, by the definition of S, one has $\varphi^S(\rho_\infty(t)) = (\rho_\infty(t),\xi)_{L^2_x}$ and similarly $\mathcal{L}\varphi^S(\rho_\infty(t)) = (\overline{\rho}_\infty(t), \mathrm{div}(K\nabla\xi) + J\cdot\nabla\xi - \overline{\sigma}\xi)_{L^2_x}$, for all $t \geqslant 0$. On the other hand, owing to Proposition 4.6, one has, for all $t \geqslant 0$,

$$\varphi^S(\rho_\infty(t)) = \varphi^S(\rho_\infty(0)) + \int_0^t \mathcal{L}\varphi^S(\rho_\infty(s))ds.$$

Combining the arguments proves that (13) holds, for all $\xi \in H^2_x$.

Let us now prove that $\rho_\infty \in L^\infty_T L^2_x$. Consider the self-adjoint operator $\mathcal{S}$ on L^2_x defined by

$$D(\mathcal{S}) = H^2_x, \quad \mathcal{S}\rho = \mathrm{div}(K\nabla\rho) - \overline{\sigma}\rho.$$

Since K is a positive symmetric matrix, the operator $\mathcal{S} - \Lambda\,\mathrm{id}$ is invertible when $\Lambda > \|\overline{\sigma}\|_{\mathcal{C}^0_x}$. Its inverse is a compact operator, owing to the compact embedding $H^2_x \subset L^2_x$. Therefore, there exists a complete orthonormal system $(e_i)_{i\in\mathbb{N}_0}$ of L^2_x composed of eigenvectors for $(\mathcal{S} - \Lambda\,\mathrm{id})^{-1}$. For $i \in \mathbb{N}_0$, let $\lambda_i \in \mathbb{R}$ be the eigenvalue of $\mathcal{S}$ associated to the eigenvector e_i: $\mathcal{S}e_i = \lambda_i$. Note that $\lambda_i \leqslant \Lambda$ for all $i \in \mathbb{N}_0$.

Fix $i \in \mathbb{N}_0$. Since $e_i \in H^2_x$, (13) reads for $t \in [0,T]$

$$(\rho_\infty(t),e_i)_{L^2_x} = (\overline{\rho}_0,e_i)_{L^2_x} + \lambda_i \int_0^t (\rho_\infty(s),e_i)_{L^2_x}\,ds + \int_0^t (\rho_\infty(s), J\cdot\nabla_x e_i)_{L^2_x}\,ds.$$

Therefore, one has, for all $t \in [0,T]$

$$(\rho_\infty(t,\cdot+tJ),e_i)_{L^2_x} = e^{\lambda_i t}\,(\overline{\rho}_0,e_i)_{L^2_x}.$$

Since $\lambda_i t \leqslant \Lambda T$ and $\overline{\rho}_0 \in L^2_x$, $\left|(\rho_\infty(t, \cdot + tJ), e_i)_{L^2_x}\right|^2$ is summable. We thus have $\rho_\infty(t) \in L^2_x$ and

$$\|\rho_\infty(t)\|^2_{L^2_x} = \|\rho_\infty(t, \cdot + tJ)\|^2_{L^2_x} \leqslant e^{2\Lambda T} \|\overline{\rho}_0\|^2_{L^2_x}. \tag{31}$$

This proves that $\rho_\infty \in L^\infty_T L^2_x$, and that ρ_∞ is a weak solution of (12) in the sense of Definition 3.1.

To prove the convergence in distribution stated in Theorem 3.2, it only remains to prove that the weak solution of (12) is unique. Since the evolution equation is linear, it is sufficient to prove that, if $\overline{\rho}_0 = 0$, then any weak solution $\overline{\rho}$ of (12) satisfies $\overline{\rho}(t) = 0$ for all $t \in [0, T]$. This claim is a straightforward consequence of (31) (satisfied by any function $\overline{\rho}$ satisfying (13)).

As a consequence, any limit point ρ_∞ of the tight family $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ in the space $\mathcal{C}^0_T H^{-\varsigma}_x$ is the unique solution $\overline{\rho}$ of (12). Therefore $\rho^{\varepsilon,\delta}$ converges in distribution to the unique weak solution $\overline{\rho}$ of (12).

When the initial condition $\overline{\rho}_0$ is deterministic, the solution $\overline{\rho}$ of (12) is also deterministic. Then, using Portmanteau Theorem, in that case $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ converges to $\overline{\rho}$ in probability.

This concludes the proof of Theorem 3.2.

□

Proof of Theorem 3.3. Let Assumption 7 be satisfied. Owing to Proposition 4.5, the family of processes $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ is tight in $L^2_T L^2_x$. Therefore, there exist a $L^2_T L^2_x$-valued random variable ρ_∞ and sequences $(\varepsilon_i)_{i\in\mathbb{N}}$ and $(\delta_i)_{i\in\mathbb{N}}$, such that $\varepsilon_i \to 0$ and $\delta_i \to 0$, and $\rho^{\varepsilon_i,\delta_i} \xrightarrow[i\to\infty]{d} \rho_\infty$. On the one hand, convergence in $L^2_T L^2_x$ implies convergence in $L^2_T H^{-\varsigma}_x$. On the other hand, owing to Theorem 3.2, $\rho^{\varepsilon_i,\delta_i}$ converges to $\overline{\rho}$ in $\mathcal{C}^0_T H^{-\varsigma}_x$, thus in $L^2_T H^{-\sigma}_x$. Therefore ρ_∞ is equal to $\overline{\rho}$ in distribution. By uniqueness of the limit points, one thus obtains the convergence of $\left(\rho^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ to $\overline{\rho}$ in $L^2_T L^2_x$.

It remains to establish the convergence of $f^{\varepsilon,\delta}$ to $\overline{\rho}\mathcal{M}$.

Note first that the mapping $h \in L^2_T L^2_x \mapsto h\mathcal{M} \in L^2_T L^2(\mathcal{M}^{-1})$ is continuous (it is a bounded linear operator). Thus, $\rho^{\varepsilon,\delta}\mathcal{M} \xrightarrow[\varepsilon,\delta\to 0]{} \overline{\rho}\mathcal{M}$ in distribution in $L^2_T L^2(\mathcal{M}^{-1})$.

Owing to Slutsky's Lemma (since [Bil99, Theorem 4.1]) and to the identity

$$f^{\varepsilon,\delta} = -Lf^{\varepsilon,\delta} + \rho^{\varepsilon,\delta}\mathcal{M},$$

it only remains to prove that $Lf^{\varepsilon,\delta} \xrightarrow[\varepsilon,\delta\to 0]{} 0$ in probability in $L^2_T L^2(\mathcal{M}^{-1})$.

Owing to Proposition 4.2, almost surely one has

$$\int_0^{T\wedge\tau^\delta} \left\|Lf^{\varepsilon,\delta}(s)\right\|^2_{L^2(\mathcal{M}^{-1})} ds \leqslant \varepsilon^2 C(T) \left\|f_0^{\varepsilon,\delta}\right\|^2_{L^2(\mathcal{M}^{-1})}.$$

Therefore, on the event $\tau^\delta \geqslant T$, one has $\left\|Lf^{\varepsilon,\delta}\right\|_{L^2_T L^2(\mathcal{M}^{-1})} \leqslant \varepsilon\sqrt{C(T)} \left\|f_0^{\varepsilon,\delta}\right\|_{L^2(\mathcal{M}^{-1})}$. As a consequence, we have for $\eta \in (0,1)$

$$\begin{aligned}
\mathbb{P}\left(\left\|Lf^{\varepsilon,\delta}\right\|_{L^2_T L^2(\mathcal{M}^{-1})} > \eta\right) &= \mathbb{P}\left(\tau^\delta < T, \left\|Lf^{\varepsilon,\delta}\right\|_{L^2_T L^2(\mathcal{M}^{-1})} > \eta\right) \\
&\quad + \mathbb{P}\left(\tau^\delta \geqslant T, \left\|Lf^{\varepsilon,\delta}\right\|_{L^2_T L^2(\mathcal{M}^{-1})} > \eta\right) \\
&\leqslant \mathbb{P}\left(\tau^\delta < T\right) + \mathbb{P}\left(\varepsilon\sqrt{C(T)} \left\|f_0^{\varepsilon,\delta}\right\|_{L^2(\mathcal{M}^{-1})} > \eta\right) \\
&\leqslant \mathbb{P}\left(\tau^\delta < T\right) + \eta^{-1}\varepsilon\sqrt{C(T)}\mathbb{E}\left[\left\|f_0^{\varepsilon,\delta}\right\|_{L^2(\mathcal{M}^{-1})}\right],
\end{aligned}$$

owing to the Markov inequality. Using Proposition 4.1 and Assumption 6, one obtains the convergence of $Lf^{\varepsilon,\delta}$ to 0 in probability in $L^2_T L^2(\mathcal{M}^{-1})$ when $(\varepsilon,\delta)\to(0,0)$. This concludes the proof of the convergence in distribution of $f^{\varepsilon,\delta}$ to $\overline{\rho}\mathcal{M}$.

When $\overline{\rho}_0$ is deterministic, $\overline{\rho}$ is deterministic, and the convergence results hold in probability, using Portmanteau Theorem.

This concludes the proof of Theorem 3.3.

□

4.4.3 Proofs of the two technical results

The following lemma is used in the proofs of Propositions 4.5 and 4.6.

Lemma 4.7. *Let $\varphi : L^2(\mathcal{M}^{-1})\times E\to\mathbb{R}$ be a function such that*

$$\sup_{f\in L^2(\mathcal{M}^{-1}),\ell\in E}\frac{|\varphi(f,\ell)|}{(1+\|f\|^6_{L^2(\mathcal{M}^{-1})})(1+\|\ell\|_E)}<\infty.$$

Then, for all $T\in(0,\infty)$, $\varepsilon\in(0,\varepsilon_0]$ and $\delta\in(0,\delta_0]$ and for all random times τ_1 and τ_2 satisfying almost surely $0\leqslant\tau_1\leqslant\tau_2\leqslant T$, one has

$$\sup_{\varepsilon,\delta}\mathbb{E}\left[\int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta}\left|\varphi(f^{\varepsilon,\delta,\tau^\delta}(t),m^{\delta,\tau^\delta}(t))\right|dt\right]\lesssim\mathbb{E}\left[\tau_2-\tau_1\right]^{\frac12-\frac1\gamma}. \tag{32}$$

Recall that $\gamma>2$ is given by Assumption 3.

Proof of Lemma 4.7. Since $m^{\delta,\tau^\delta}(t)=m^\delta(t)$ for all $t\in[\tau_1\wedge\tau^\delta,\tau_2\wedge\tau^\delta]$, the estimate (18) from Proposition 4.2 yields

$$\mathbb{E}\left[\int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta}\left|\varphi(f^{\varepsilon,\delta,\tau^\delta}(t),m^{\delta,\tau^\delta}(t))\right|dt\right]\lesssim_{T,\varphi}\mathbb{E}\left[\int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta}\left(1+\left\|f_0^{\varepsilon,\delta}\right\|^6_{L^2(\mathcal{M}^{-1})}\right)\left(1+\|m^\delta(t)\|\right)dt\right].$$

Let $p^*\in[1,\infty)$ defined by $\frac1\gamma+\frac12+\frac{1}{p^*}=1$. Using the moment estimates for m^δ (see Assumption 3) and for $f^{\varepsilon,\delta}$ (see Assumption 6), and applying Hölder inequality, one obtains

$$\begin{aligned}
&\mathbb{E}\left[\int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta}\left|\varphi(f^{\varepsilon,\delta,\tau^\delta}(t),m^{\delta,\tau^\delta}(t))\right|dt\right]\\
&\lesssim\int_0^T\mathbb{E}\left[\mathbb{1}_{\{[\tau_1,\tau_2]\}}(t)\left(1+\left\|f_0^{\varepsilon,\delta}\right\|^6_{L^2(\mathcal{M}^{-1})}\right)\left(1+\|m^\delta(t)\|\right)\right]dt\\
&\lesssim\int_0^T\mathbb{E}\left[\mathbb{1}_{\{[\tau_1,\tau_2]\}}(t)^{p^*}\right]^{1/p^*}\mathbb{E}\left[\left(1+\left\|f_0^{\varepsilon,\delta}\right\|^6_{L^2(\mathcal{M}^{-1})}\right)^2\right]^{1/2}\mathbb{E}\left[\left(1+\|m^\delta(t)\|\right)^\gamma\right]^{1/\gamma}dt\\
&\lesssim\int_0^T\mathbb{E}\left[\mathbb{1}_{\{[\tau_1,\tau_2]\}}(t)\right]^{1/p^*}dt\\
&\lesssim\left(\int_0^T\mathbb{E}\left[\mathbb{1}_{\{[\tau_1,\tau_2]\}}(t)\right]dt\right)^{1/p^*}\\
&\lesssim\mathbb{E}\left[\int_{\tau_1}^{\tau_2}dt\right]^{1/p^*}\\
&\lesssim\mathbb{E}\left[\tau_2-\tau_1\right]^{1/p^*}.
\end{aligned}$$

All the estimates above are uniform with respect to $\varepsilon \in (0, \varepsilon_0]$ and $\delta \in (0, \delta_0]$. This concludes the proof of Lemma 4.7. □

Before proceeding, let us recall some useful results concerning tightness.

Introduce the Skorokhod space $D_T H_x^{-\varsigma}$, which is the space of $H_x^{-\varsigma}$-valued càdlàg functions on $[0,T]$. For all $X \in \mathcal{C}_T^0 H_x^{-\varsigma}$ and all $\Delta \in [0,T]$, set

$$w_X(\Delta) \doteq \sup_{0 \leqslant t \leqslant s \leqslant t+\Delta \leqslant T} \|X(s) - X(t)\|$$
$$w'_X(\Delta) \doteq \sup_{(t_i)_i} \max_i \sup_{t_i \leqslant t \leqslant s < t_{i+1}} \|X(s) - X(t)\|,$$

where $(t_i)_i$ denotes any finite subdivision of $[0,T]$. The moduli of continuity ω_X and ω'_X satisfy the following inequality (see [Bil99, equation (14.11)]): for all $\Delta \in [0,T]$ and all $X \in \mathcal{C}_T^0 H_x^{-\varsigma}$, one has

$$w_X(\Delta) \leqslant 2 w'_X(\Delta).$$

We refer to [Bil99, Theorems 8.2 and 15.2] for tightness criteria in the spaces $\mathcal{C}_T^0 H_x^{-\varsigma}$ and $D_T H^{-\varsigma}$. As a consequence, tightness in $D_T H_x^{-\varsigma}$ of a family a family of processes $\left(X^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ implies its tightness in $\mathcal{C}_T^0 H_x^{-\varsigma}$.

Observe that tightness in $D_T H^{-\varsigma}$ is easier to prove than tightness in $\mathcal{C}_T^0 H_x^{-\varsigma}$, owing to [Jak86, Theorem 3.1]. More precisely, since the class of functions $\Theta_{\lim}$ is closed under addition and separates points, tightness of a family $\left(X^{\varepsilon,\delta}\right)_{\varepsilon,\delta}$ in the Skorokhod space $D_T H^{-\varsigma}$ is equivalent to the following claims:

(i) For all $\eta \in (0,1]$, there exists a compact set $K_\eta \subset H_x^{-\varsigma}$ such that, for all $\varepsilon \in (0, \varepsilon_0]$ and $\delta \in (0, \delta_0]$,
$$\mathbb{P}\left(\forall t \in [0,T], X^{\varepsilon,\delta}(t) \in K_\eta\right) > 1 - \eta.$$

(ii) For all $\varphi \in \Theta_{\lim}$, $\left(\varphi(X^{\varepsilon,\delta})\right)_{\varepsilon,\delta}$ is tight in the Skorohod space $D([0,T],\mathbb{R})$ of càdlàg real-valued functions defined on the interval $[0,T]$.

To check that (ii) is satisfied, we employ Aldous's criterion, see [JS03, Theorem 4.5 p356] (note that the criterion is simplified since here φ is bounded): it suffices to prove that for all $\eta \in (0,\infty)$, one has

$$\lim_{\Delta \to 0} \limsup_{(\varepsilon,\delta) \to (0,0)} \sup_{\tau_1 \leqslant \tau_2 \leqslant \tau_1 + \Delta} \mathbb{P}\left(\left|\varphi(X^{\varepsilon,\delta}(\tau_2)) - \varphi(X^{\varepsilon,\delta}(\tau_1))\right| > \eta\right) = 0, \tag{33}$$

where $\sup_{\tau_1 \leqslant \tau_2 \leqslant \tau_1 + \Delta}$ denotes the supremum with respect to all $\left(\mathcal{F}_t^\delta\right)_{t \in \mathbb{R}^+}$-stopping times τ_1 and τ_2 satisfying a.s. $\tau_1, \tau_2 \in [0,T]$ and $\tau_1 \leqslant \tau_2 \leqslant \tau_1 + \Delta$.

Proof of Proposition 4.5. First, recall that $\tau^\delta \to \infty$ in probability, when $\delta \to 0$ (see Proposition 4.1. Owing to Slutsky's Lemma [Bil99, Theorem 4.1], it thus suffices to prove the tightness of the family $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$.

Let us first establish the tightness in the space $\mathcal{C}_T^0 H_x^{-\varsigma}$, with an arbitrarily small parameter $\varsigma > 0$. As explained above, in fact we establish the tightness in $D_T H^{-\varsigma}$, using the criteria stated above. First, (i) is satisfied: indeed the embedding $L_x^2 \subset H_x^{-\varsigma}$ is compact, and the a priori estimate (18) (see Proposition 4.2) yields the uniform moment bound

$$\sup_{\varepsilon,\delta} \mathbb{E}\Big[\sup_{0 \leqslant t \leqslant T} \left\|\rho^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L_x^2}^2\Big] \leqslant \sup_{\varepsilon,\delta} \mathbb{E}\Big[\sup_{0 \leqslant t \leqslant T} \left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}^2\Big] \leqslant \sup_{\varepsilon,\delta} \mathbb{E}\big[\left\|f^{\varepsilon,\delta}(0)\right\|_{L_x^2}^2\big] < \infty$$

owing to Assumption 6. Then (i) is a straightforward consequence of Markov inequality.

It remains to establish (ii), using Aldous criterion 33. Let $\varphi^{\varepsilon,\delta}$ be the perturbed test function given by (23), see Proposition 4.3, and let $M^{\varepsilon,\delta}_{\varphi^{\varepsilon,\delta}}$ be defined as by (30) (in Proposition 4.4). For all $t \geqslant 0$, set

$$\begin{aligned}\theta^{\varepsilon,\delta}(t) &= \varphi(\rho^{\varepsilon,\delta}(0)) + \varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta}(t), m^{\delta}(t)) - \varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta}(0), m^{\delta}(0)) \\ &= \varphi(\rho^{\varepsilon,\delta}(0)) + \int_0^t \mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta}(s), m^{\delta}(s))ds + M^{\varepsilon,\delta}_{\varphi^{\varepsilon,\delta}}(t), \end{aligned} \tag{34}$$

For all stopping times τ_1, τ_2, one has the equality

$$\begin{aligned}\varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_2)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_1)) &= \left(\theta^{\varepsilon,\delta,\tau^\delta}(\tau_2) - \theta^{\varepsilon,\delta,\tau^\delta}(\tau_1)\right) \\ &\quad - \left(\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta,\tau^\delta}(\tau_2), m^{\delta,\tau^\delta}(\tau_2)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_2))\right) \\ &\quad + \left(\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta,\tau^\delta}(\tau_1), m^{\delta,\tau^\delta}(\tau_1)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_1))\right).\end{aligned}$$

On the one hand, owing to the error estimate (29), the estimates (16) and (18) for $\left\|m^{\delta,\tau^\delta}(t)\right\|_E$ and $\left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}$ yield

$$\left|\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta,\tau^\delta}(t), m^{\delta,\tau^\delta}(t)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(t))\right| \lesssim_\varphi (1 + \left\|f_0^{\varepsilon,\delta}\right\|^2_{L^2(\mathcal{M}^{-1})})(\varepsilon + \delta^2(1+\delta^{-\alpha})).$$

Since $\alpha < 1$ in Definition 4.1, we get

$$\mathbb{E}\left[\left|\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta,\tau^\delta}(\tau_i), m^{\delta,\tau^\delta}(\tau_i)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_i))\right|\right] \xrightarrow[\varepsilon,\delta\to 0]{} 0,$$

for $i = 1, 2$.

On the other hand, we claim that

$$\sup_{\varepsilon,\delta}\sup_{\tau_1\leqslant\tau_2\leqslant\tau_1+\Delta} \mathbb{E}\left[\left|\theta^{\varepsilon,\delta,\tau^\delta}(\tau_2) - \theta^{\varepsilon,\delta,\tau^\delta}(\tau_1)\right|\right] \xrightarrow[\Delta\to 0]{} 0. \tag{35}$$

To prove that this claim holds, note that

$$\left|\theta^{\varepsilon,\delta,\tau^\delta}(\tau_2) - \theta^{\varepsilon,\delta,\tau^\delta}(\tau_1)\right| \leqslant \int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta} \left|\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}(f^{\varepsilon,\delta}(s), m^{\delta}(s))\right| ds + \left|M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_2) - M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_1)\right|. \tag{36}$$

To treat the first term on the right-hand side of (36), observe that for $(f,\ell) \in L^2(\mathcal{M}^{-1}) \times E$, one has

$$\begin{aligned}\left|\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}(f,\ell)\right| &\lesssim \varepsilon\left(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})}\right)(1 + \|\ell\|) + \delta^2\left(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})}\right)\left(1 + \|\ell\|^2\right) \\ &\quad + |\mathcal{L}\varphi(\rho)| \\ &\lesssim \left(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})}\right)(1 + \|\ell\|) + \delta^2\left(1 + \|f\|^3_{L^2(\mathcal{M}^{-1})}\right)\left(1 + \|\ell\|^2\right),\end{aligned}$$

owing to the error estimate (24) for $\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} - \mathcal{L}\varphi$ (see Proposition 4.3) and to the expression (20) of $\mathcal{L}\varphi$. Using Lemma 4.7, one obtains

$$\sup_{\varepsilon,\delta}\sup_{\tau_1\leqslant\tau_2\leqslant\tau_1+\Delta} \mathbb{E}\left[\int_{\tau_1\wedge\tau^\delta}^{\tau_2\wedge\tau^\delta} \left|\left(1 + \left\|f^{\varepsilon,\delta}(s)\right\|^3_{L^2(\mathcal{M}^{-1})}\right)\left(1 + \|m^{\delta}(s)\|\right)\right| ds\right] \lesssim \Delta^{1/2-1/\gamma} \xrightarrow[\Delta\to 0]{} 0,$$

since $\gamma > 2$ (see Assumption 3. In addition, recall that $\alpha < 1$: using the estimate (16), one obtains

$$\sup_{\varepsilon,\delta} \sup_{\tau_1 \leqslant \tau_2 \leqslant \tau_1 + \Delta} \mathbb{E}\left[\int_{\tau_1 \wedge \tau^\delta}^{\tau_2 \wedge \tau^\delta} \left| \delta^2 \left(1 + \left\| f^{\varepsilon,\delta}(s) \right\|^3_{L^2(\mathcal{M}^{-1})}\right) \left(1 + \left\| m^\delta(s) \right\|^2\right) \right| ds \right] \lesssim \Delta \xrightarrow[\Delta \to 0]{} 0.$$

To treat the second term on the right-hand side of (36), note that $M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}$ is a square-integrable martingale, owing to Proposition 4.4. In addition, since $\tau_1 \leqslant \tau_2$ are stopping times, one has

$$\begin{aligned}
\mathbb{E}\left[\left| M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_2) - M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_1) \right|^2 \right] &= \mathbb{E}\left[\left| M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_2) \right|^2 - \left| M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_1) \right|^2 \right] \\
&= \frac{1}{\delta^2} \mathbb{E}\left[\int_{\tau_1 \wedge \tau^\delta}^{\tau_2 \wedge \tau^\delta} \left(\mathcal{L}_m((\varphi^{\varepsilon,\delta})^2) - 2\varphi^{\varepsilon,\delta} \mathcal{L}_m \varphi^{\varepsilon,\delta} \right) (f^{\varepsilon,\delta}(s), m^\delta(s)) ds \right].
\end{aligned}$$

Recall the expression (23) of the perturbed test function $\varphi^{\varepsilon,\delta}$. Since φ, $\varphi_{1,0}$ and $\varphi_{2,0}$ do not depend on $\ell \in E$, one has

$$\begin{aligned}
&\mathbb{E}\left[\left| M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_2) - M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_1) \right|^2 \right] \\
&= \delta^2 \mathbb{E}\left[\int_{\tau_1 \wedge \tau^\delta}^{\tau_2 \wedge \tau^\delta} \left(\mathcal{L}_m((\varphi_{0,2} + \varepsilon\varphi_{1,2})^2) - 2(\varphi_{0,2} + \varepsilon\varphi_{1,2}) \mathcal{L}_m(\varphi_{0,2} + \varepsilon\varphi_{1,2}) \right) (f^{\varepsilon,\delta}(s), m^\delta(s)) ds \right].
\end{aligned}$$

The correctors $\varphi_{0,2}$ and $\varphi_{1,2}$ satisfy the properties (27) and (28) respectively. In addition, $f^{\varepsilon,\delta}$ and m^δ satisfy the estimates (18) and (16) respectively. Using Assumption 5 and the condition $0 \leqslant \tau_2 - \tau_1 \leqslant \Delta$, one finally obtains

$$\begin{aligned}
\mathbb{E}\left[\left| M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_2) - M^{\varepsilon,\delta,\tau^\delta}_{\varphi^{\varepsilon,\delta}}(\tau_1) \right|^2 \right] &\lesssim \Delta \delta^2 \mathbb{E}\left[(1 + \left\| f_0^{\varepsilon,\delta} \right\|^2_{L^2(\mathcal{M}^{-1})})(1 + \delta^{-2\alpha}) \right] \\
&\lesssim \Delta \xrightarrow[\Delta \to 0]{} 0,
\end{aligned}$$

since $\alpha < 1$ and using Assumption 6.

Gathering the estimates for the two terms on the right-hand side of (36), the claim (35) is proved. One finally obtains

$$\sup_{\varepsilon,\delta} \sup_{\tau_1 \leqslant \tau_2 \leqslant \tau_1 + \Delta} \mathbb{E}\left[\left| \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_2)) - \varphi(\rho^{\varepsilon,\delta,\tau^\delta}(\tau_1)) \right| \right] \xrightarrow[\Delta \to 0]{} 0,$$

hence (33) holds. This concludes the proof of the tightness of the family $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$ is tight in the space $\mathcal{C}^0_T H^{-\varsigma}_x$.

It remains to prove the tightness of $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$ is tight in the space $L^2_T L^2_x$, if Assumption 7 is satisfied. It suffices to establish the following claims: for all $\eta \in (0,1)$, there exists $R \in (0,\infty)$ and $\varsigma'' \in (0,1)$, such that

$$\lim_{\Delta \to 0} \limsup_{(\varepsilon,\delta) \to (0,0)} \mathbb{P}\left(w_{\rho^{\varepsilon,\delta,\tau^\delta}}(\Delta) > \eta \right) = 0, \tag{37}$$

and

$$\sup_{\varepsilon,\delta} \mathbb{P}\left(\left\| \rho^{\varepsilon,\delta,\tau^\delta} \right\|_{L^2_T H^{\varsigma''}_x} > R \right) < \eta. \tag{38}$$

Indeed, for all $R>0$, $\varsigma''>0$ and $\eta:(0,\infty)\to[0,\infty)$ such that $\eta(\Delta)\xrightarrow[\Delta\to 0]{}0$, the set

$$K_{R,\eta}\doteq\left\{\rho\in L^2_TL^2_x\mid \|\rho\|_{L^2_TH^{\varsigma''}_x}\leqslant R \text{ and } \forall\Delta\in(0,1), w_\rho(\Delta)<\eta(\Delta)\right\}$$

is compact in $L^2_TL^2_x$, see [Sim87, Theorem 5]

The first claim is a consequence of the tightness of $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$ in $\mathcal{C}^0_TH^{-\varsigma}_x$ proved above, see [Bil99, Theorem 8.2].

Owing to Markov inequality, it suffices to prove the following inequality:

$$\sup_{\varepsilon,\delta}\mathbb{E}\left[\left\|\rho^{\varepsilon,\delta,\tau^\delta}\right\|_{L^2_TH^{\varsigma''}_x}\right]\lesssim 1, \tag{39}$$

Let $g^{\varepsilon,\delta}=\varepsilon\partial_t f^{\varepsilon,\delta}+a(v)\cdot\nabla_x f^{\varepsilon,\delta}+\varepsilon b(v)\cdot\nabla_x f^{\varepsilon,\delta}$. By Assumption 7, we are in position to apply an averaging lemma , precisely [BD99, Theorem 2.3] (with $f(t)=f^{\varepsilon,\delta}(\varepsilon t)$, $g(t)=g^{\varepsilon,\delta}(\varepsilon t)$ and $h=0$ until time $T\wedge\tau^\delta$). After rescaling the time $t\mapsto t/\varepsilon$, one obtains the inequality

$$\begin{aligned}\left\|\rho^{\varepsilon,\delta,\tau^\delta}\right\|^2_{L^2_TH^{\varsigma'/4}_x}&=\int_0^{T\wedge\tau^\delta}\left\|\rho^{\varepsilon,\delta,\tau^\delta}(t)\right\|^2_{H^{\varsigma'/4}_x}dt\\&\lesssim\varepsilon\left\|f^{\varepsilon,\delta}_0\right\|^2_{L^2_x}+\int_0^{T\wedge\tau^\delta}\left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|^2_{L^2(\mathcal{M}^{-1})}dt+\int_0^{T\wedge\tau^\delta}\left\|g^{\varepsilon,\delta,\tau^\delta}(t)\right\|^2_{L^2(\mathcal{M}^{-1})}dt.\end{aligned}$$

Applying the Cauchy-Schwarz inequality gives

$$\begin{aligned}\left\|g^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}&=\left\|\varepsilon f^{\varepsilon,\delta,\tau^\delta}(t)\sigma(m^{\delta,\tau^\delta}(t))+\frac{1}{\varepsilon}Lf^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}\\&\leqslant\varepsilon\left\|f^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}\left\|\sigma(m^{\delta,\tau^\delta}(t))\right\|_{\mathcal{C}^0_x}+\frac{1}{\varepsilon}\left\|Lf^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}.\end{aligned}$$

It is now crucial to use the a priori estimate (18) (see Proposition 4.2) to control the intergral term

$$\int_0^{T\wedge\tau^\delta}\frac{1}{\varepsilon}\left\|Lf^{\varepsilon,\delta,\tau^\delta}(t)\right\|_{L^2(\mathcal{M}^{-1})}dt$$

uniformly with respect to ε,δ. Using Assumption 6, the estimates (16) and (18) for $m^{\delta,\tau^\delta}(t)$ and $f^{\varepsilon,\delta,\tau^\delta}(t)$, and Lemma 4.7, the claim (39) is proved, with $\varsigma''=\frac{\varsigma'}{4}$.

Since we proved (37) and (38), the family $\left(\rho^{\varepsilon,\delta,\tau^\delta}\right)_{\varepsilon,\delta}$ is tight in $L^2_TL^2_x$, when Assumption 7 is satisfied.

This concludes the proof of Proposition 4.5. □

Proof of Proposition 4.6. Let ρ_∞ be such that $\rho_\infty=\lim_{i\to\infty}\rho^{\varepsilon_i,\delta_i}$, for some sequence $(\varepsilon_i,\delta_i)\to(0,0)$.

For all $\varphi\in\Theta_{\lim}$, define the stochastic process M_φ as follows: for all $t\geqslant 0$,

$$M_\varphi(t)=\varphi(\rho_\infty(t))-\varphi(\rho_\infty(0))-\int_0^t\mathcal{L}\varphi(\rho_\infty(s))ds.$$

Let us start by proving that, for all $\varphi\in\Theta_{\lim}$, M_φ is a square integrable martingale adapted to the filtration generated by ρ_∞.

Let the test function $\varphi \in \Theta_{\lim}$ be fixed. Since φ is bounded, to prove that the process M_φ is square-integrable, it suffices to prove that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|\mathcal{L}\varphi(\rho_\infty(t))\right|^2\right] < \infty. \tag{40}$$

Observe that, for all $t \in [0,T]$, the mapping $\rho \in \mathcal{C}_T^0 H_x^{-\varsigma} \mapsto |\mathcal{L}\varphi(\rho(t))|^2 \in \mathbb{R}$ is continuous. Thus one has the convergence in distribution $\left|\mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t))\right|^2 \xrightarrow[i\to\infty]{d} |\mathcal{L}\varphi(\rho_\infty(t))|^2$ (see [Bou04, Proposition IX.5.7]). In addition, $\left(\left|\mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t))\right|^2\right)_{\varepsilon_i,\delta_i}$ is uniformly integrable: one has

$$\sup_{i\in\mathbb{N}} \mathbb{E}\left[\left|\mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t))\right|^4\right] \lesssim \sup_{\varepsilon,\delta} \mathbb{E}\left[\left\|f_0^{\varepsilon,\delta}\right\|^4_{L^2(\mathcal{M}^{-1})}\right] < \infty,$$

using the estimates from Proposition 4.2 and Assumption 6, and the expression (21) of $\mathcal{L}\varphi$ when $\varphi \in \Theta_{\lim}$. Using [Bil99, Theorem 5.4], the convergence in distribution and uniform integrability property give

$$\mathbb{E}\left[\left|\mathcal{L}\varphi(\rho_\infty(t))\right|^2\right] = \lim_{i\to\infty} \mathbb{E}\left[\left|\mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t))\right|^2\right] \lesssim \sup_{\varepsilon,\delta} \mathbb{E}\left[\left\|f_0^{\varepsilon,\delta}\right\|^2_{L^2(\mathcal{M}^{-1})}\right] < \infty.$$

This yields the square integrability property (40).

The next step is to prove that M_φ is a martingale. Let $0 \leqslant s_1 \leqslant \ldots \leqslant s_j \leqslant s \leqslant t$ and let $g \in \mathcal{C}_b^0((H_x^{-\varsigma})^j)$ be a continuous bounded function. Define the mapping

$$\Phi : \rho \in \mathcal{C}_T^0 H_x^{-\varsigma} \mapsto \left(\varphi(\rho(t)) - \varphi(\rho(s)) - \int_s^t \mathcal{L}\varphi(\rho(u))du\right) g(\rho(s_1),\ldots,\rho(s_j)).$$

To prove that M_φ is a martingale, it suffices to prove the following claim: $\mathbb{E}\left[\Phi(\rho_\infty)\right] = 0$.

Let us first check that $\mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right] \to \mathbb{E}\left[\Phi(\rho_\infty)\right]$ when $i \to \infty$. This claim follows from the definition of ρ_∞ as the limit in distribution of $\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}$ and straightforward arguments. The mapping Φ is continuous on $\mathcal{C}_T^0 H_x^{-\varsigma}$ and $\left(\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right)_{\varepsilon_i,\delta_i}$ is uniformly integrable. Using the same arguments as for the proof of (40), one obtains the following convergence result:

$$\mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right] \xrightarrow[i\to\infty]{} \mathbb{E}\left[\Phi(\rho_\infty)\right].$$

Let us now check that $\mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right] \to 0$ when $i \to \infty$. This claim follows from a martingale property and the perturbed test function to take the limit $i \to \infty$. Let $\varphi^{\varepsilon,\delta}$ be the perturbed test function given by (23), see Proposition 4.3. Using the martingale property from Proposition 4.4, one has

$$\mathbb{E}\left[\left(\varphi^{\varepsilon_i,\delta_i}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t), m^{\delta_i,\tau^{\delta_i}}(t)) - \varphi^{\varepsilon_i,\delta_i}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s), m^{\delta_i,\tau^{\delta_i}}(s))\right.\right.$$
$$\left.\left. - \int_{s\wedge\tau^{\delta_i}}^{t\wedge\tau^{\delta_i}} \mathcal{L}^{\varepsilon_i,\delta_i}\varphi^{\varepsilon_i,\delta_i}(f^{\varepsilon_i,\delta_i}(u), m^{\delta_i}(u))du\right) g(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s_1),\ldots,\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s_j))\right] = 0.$$

Using the expression $\varphi^{\varepsilon,\delta} = \varphi + \varepsilon\varphi_{1,0} + \varepsilon^2\varphi_{2,0} + \delta^2\varphi_{0,2} + \varepsilon\delta^2\varphi_{1,2}$ (see (23), and the boundedness of g, one obtains the upper bound

$$\left|\mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right]\right| \lesssim \sum_{j=1}^{6} \mathbb{E}\left[|r_j|\right],$$

with

$$\begin{aligned}
r_1 &= \varepsilon_i(\varphi_{1,0}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t)) - \varphi_{1,0}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s)))\\
r_2 &= \varepsilon_i^2(\varphi_{2,0}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t)) - \varphi_{2,0}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s)))\\
r_3 &= \delta_i^2(\varphi_{0,2}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t), m^{\delta_i,\tau^{\delta_i}}(t)) - \varphi_{0,2}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s), m^{\delta_i,\tau^{\delta_i}}(s)))\\
r_4 &= \varepsilon_i\delta_i^2(\varphi_{1,2}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(t), m^{\delta_i,\tau^{\delta_i}}(t)) - \varphi_{1,2}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(s), m^{\delta_i,\tau^{\delta_i}}(s)))\\
r_5 &= \int_{s\wedge\tau^{\delta_i}}^{t\wedge\tau^{\delta_i}} \left(\mathcal{L}^{\varepsilon_i,\delta_i}\varphi^{\varepsilon_i,\delta_i}(f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(u), m^{\delta_i,\tau^{\delta_i}}(u)) - \mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(u))\right) du\\
r_6 &= \int_{t\wedge\tau^{\delta_i}}^{t} \mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(u))du - \int_{s\wedge\tau^{\delta_i}}^{s} \mathcal{L}\varphi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}}(u))du.
\end{aligned}$$

Using estimates (25), (26), (27) and (28) of Proposition 4.3 and estimates (16) and (18) on $m^{\delta_i,\tau^{\delta_i}}$ and $f^{\varepsilon_i,\delta_i,\tau^{\delta_i}}$, one has

$$\begin{aligned}
|r_1| &\lesssim \varepsilon_i \left\|f_0^{\varepsilon_i,\delta_i}\right\|_{L^2(\mathcal{M}^{-1})},\\
|r_2| &\lesssim \varepsilon_i^2(1 + \left\|f_0^{\varepsilon_i,\delta_i}\right\|^2_{L^2(\mathcal{M}^{-1})}),\\
|r_3| &\lesssim \delta_i^2(1 + \left\|f_0^{\varepsilon_i,\delta_i}\right\|_{L^2(\mathcal{M}^{-1})})(1 + \delta_i^{-\alpha}),\\
|r_4| &\lesssim \varepsilon_i\delta_i^2(1 + \left\|f_0^{\varepsilon_i,\delta_i}\right\|^2_{L^2(\mathcal{M}^{-1})})(1 + \delta_i^{-\alpha}).
\end{aligned}$$

Since $\alpha < 1$ (see Definition 4.1), using Assumption 6, one obtains $\mathbb{E}\left[|r_k|\right] \xrightarrow[i\to\infty]{} 0$, for $k \in [\![1, 4]\!]$.

To treat the next term r_5, it is necessary to use Lemma 4.7: using the estimate (24), one obtains

$$\mathbb{E}\left[|r_5|\right] \lesssim \varepsilon_i(t-s)^{\frac{1}{2}-\frac{1}{\gamma}} + \delta_i^2(1 + \delta_i^{-2\alpha}) \xrightarrow[i\to\infty]{} 0,$$

using the condition $\alpha < 1$.

The last term r_6 is treated as follows: using the Cauchy-Schwarz inequality, the expression (21) for $\mathcal{L}\varphi$ and the estimate 4.2, one obtains

$$\begin{aligned}
\mathbb{E}\left[|r_6|\right]^2 &\lesssim \mathbb{E}\left[\left\|f_0^{\varepsilon,\delta}\right\|^2_{L^2(\mathcal{M}^{-1})}\right] \mathbb{E}\left[\left(\left|t - t\wedge\tau^{\delta_i}\right| + \left|s - s\wedge\tau^{\delta_i}\right|\right)^2\right]\\
&\lesssim \mathbb{P}\left(\tau^{\delta_i} < T\right) \xrightarrow[i\to\infty]{} 0,
\end{aligned}$$

owing to Assumption 6 and to Proposition 4.1 which gives the convergence in probability $\tau^\delta \to \infty$.

Gathering the results, one obtains $\mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right] \to 0$ when $i \to \infty$, hence $\mathbb{E}\left[\Phi(\rho_\infty)\right] = \lim_{i\to\infty} \mathbb{E}\left[\Phi(\rho^{\varepsilon_i,\delta_i,\tau^{\delta_i}})\right] = 0$. This concludes the proof that, for all $\varphi \in \Theta_{\lim}$, M_φ is a square-integrable martingale adapted to the filtration generated by ρ_∞.

The final step is to prove that $\mathbb{E}\left[|M_\varphi(t)|^2\right] = 0$, for all $t \in [0,T]$ and $\varphi \in \Theta_{\lim}$.

Let $\Delta \in (0,\infty)$ be an arbitrarily small real-number and let $0 = t_0 < t_1 < ... < t_j = t$ be a subdivision of $[0,t]$ such that $\max_{k\in[\![0,j-1]\!]} |t_{k+1} - t_k| \leqslant \Delta$. Since M_φ is a centered martingale, with $M_\varphi(0) = 0$, one has

$$\begin{aligned}\mathbb{E}\left[|M_\varphi(t)|^2\right] &= \sum_{i=0}^{n-1} \mathbb{E}\left[|M_\varphi(t_{k+1}) - M_\varphi(t_k)|^2\right] \\ &\leqslant 2\sum_{i=0}^{n-1} \mathbb{E}\left[|\varphi(\rho_\infty(t_{k+1})) - \varphi(\rho_\infty(t_k))|^2\right] + 2\sum_{i=0}^{n-1} \mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} \mathcal{L}\varphi(\rho_\infty(s))ds\right|^2\right]. \quad (41)\end{aligned}$$

On the one hand, for $0 \leqslant t \leqslant t' \leqslant T$, one has the identity

$$\begin{aligned}\left|\varphi(\rho_\infty(t')) - \varphi(\rho_\infty(t))\right|^2 &= M_{\varphi^2}(t') - M_{\varphi^2}(t) - 2\varphi(\rho_\infty(t))\left(M_\varphi(t') - M_\varphi(t)\right) \\ &\quad + \int_t^{t'} \mathcal{L}(\varphi^2)(\rho_\infty(s))ds - 2\varphi(\rho_\infty(t))\int_t^{t'} \mathcal{L}(\varphi)(\rho_\infty(s))ds.\end{aligned}$$

The test functions φ and φ^2 belong to the class $\Theta_{\lim}$, therefore, owing to the first part of the proof, M_φ and M_{φ^2} are centered martingales for the filtration generated by ρ_∞. As a consequence,

$$\mathbb{E}\left[\left|\varphi(\rho_\infty(t')) - \varphi(\rho_\infty(t))\right|^2\right] = \mathbb{E}\left[\int_t^{t'} \mathcal{L}(\varphi^2)(\rho_\infty(s))ds - 2\varphi(\rho_\infty(t))\int_t^{t'} \mathcal{L}\varphi(\rho_\infty(s))ds\right]. \quad (42)$$

Since $\mathcal{L}$ is a first order derivative operator (see (20)), it is straightforward to check that $\mathcal{L}(\varphi^2) = 2\varphi\mathcal{L}\varphi$. One thus obtains

$$\begin{aligned}\mathbb{E}\left[|\varphi(\rho_\infty(t_{k+1})) - \varphi(\rho_\infty(t_k))|^2\right] &= 2\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \left(\varphi(\rho_\infty(s)) - \varphi(\rho_\infty(t_k))\right)\mathcal{L}\varphi(\rho_\infty(s))ds\right] \\ &\leqslant 2\int_{t_k}^{t_{k+1}} \mathbb{E}\left[|\varphi(\rho_\infty(s)) - \varphi(\rho_\infty(t_k))|^2\right]^{1/2} \mathbb{E}\left[|\mathcal{L}\varphi(\rho_\infty(s))|^2\right]^{1/2} ds,\end{aligned}$$

owing to the Cauchy-Schwarz inequality. Since $\varphi \in \Theta_{\lim}$, one may use the inequality (40), and φ is bounded. Thus, (42) gives (with $t = t_k$ and $t' = s$)

$$\mathbb{E}\left[|\varphi(\rho_\infty(s)) - \varphi(\rho_\infty(t_k))|^2\right] \lesssim s - t_k.$$

Using (40), one obtains

$$\mathbb{E}\left[|\varphi(\rho_\infty(t_{k+1})) - \varphi(\rho_\infty(t))|^2\right] \lesssim \int_{t_k}^{t_{k+1}} (s - t_k)^{1/2} ds \lesssim \Delta^{3/2}. \quad (43)$$

On the other hand, (40) yields

$$\mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} \mathcal{L}\varphi(\rho_\infty(s))ds\right|^2\right] \lesssim \Delta^2. \quad (44)$$

Finally, (41), (43) and (44) yield

$$\mathbb{E}\left[|M_\varphi(t)|^2\right] \lesssim \Delta^{1/2} + \Delta \xrightarrow[\Delta\to 0]{} 0.$$

This concludes the proof that $\mathbb{E}\left[|M_\varphi(t)|^2\right] = 0$. We deduce that, for all $t \in [0,T]$, almost surely, $M_\varphi(t) = 0$. Since $\rho_\infty \in \mathcal{C}^0_T H^{-\varsigma}$, M_φ is a continuous process. As a consequence, almost surely, for all $t \in [0,T]$, $M_\varphi(t) = 0$, which concludes the proof of Proposition 4.6.

□

5 Proof of the auxiliary results

5.1 Asymptotic behavior of the stopping time

The goal of this section is to provide the proof of Proposition 4.1: $\tau^\delta \to \infty$ in probability when $\delta \to 0$.

Proof of Proposition 4.1. Let us first prove the first claim: $\tau^\delta_m \to \infty$ in probability.

Introduce the following sequence of real-valued random variables: for all $i \in \{0, 1, \dots\}$, set

$$S_i \doteq \sup_{t\in[i,i+1]} \|m(t)\|_E .$$

The random variables S_i are almost surely finite, since $\mathbb{E}[S_i^\gamma] < \infty$ for all $i \geqslant 0$ owing to Assumption 3.

Owing to the condition $\alpha > 2/\gamma$, there exists α' such that $2/\gamma < \alpha' < \alpha$. Using Markov inequality and Assumption 3, one obtains

$$\sum_{i=1}^{\infty} \mathbb{P}\left(S_i \geqslant i^{\frac{\alpha'}{2}}\right) \leqslant \sum_{i=1}^{\infty} \frac{\mathbb{E}\left[S_i^\gamma\right]}{i^{\frac{\alpha'\gamma}{2}}} = \sup_{i\geqslant 1} \mathbb{E}\left[S_i^\gamma\right] \sum_{i=1}^{\infty} \frac{1}{i^{\frac{\alpha'\gamma}{2}}} < \infty.$$

Using Borel-Cantelli's lemma, there exists a $\mathbb{N}$-valued random variable I_0, such that almost surely,

$$\forall i > I_0, S_i < i^{\frac{\alpha'}{2}}.$$

Define the random variable $Z = \sup_{0\leqslant i\leqslant I_0} S_i$. Since I_0 is almost surely finite, Z is also an almost surely finite random variable. Observe that almost surely, for all $t \geqslant 0$, one has

$$\left\|m^\delta(t)\right\|_E \leqslant S_{\lfloor t\delta^{-2}\rfloor} \leqslant Z + \left\lfloor t\delta^{-2}\right\rfloor^{\frac{\alpha'}{2}} \leqslant Z + \left(t\delta^{-2}\right)^{\frac{\alpha'}{2}}.$$

We are now in position to conclude the proof of the first claim: for all $T \in (0,\infty)$, one obtains

$$\mathbb{P}\left(\tau^\delta_m < T\right) = \mathbb{P}\left(\sup_{t\in[0,T]} \left\|m^\delta(t)\right\|_E > \delta^{-\alpha}\right) \leqslant \mathbb{P}\left(Z + T^{\frac{\alpha'}{2}}\delta^{-\alpha'} > \delta^{-\alpha}\right) \xrightarrow[\delta\to 0]{} 0,$$

since $\alpha' < \alpha$. Thus $\tau^\delta_m \to \infty$ in probability.

It remains to prove the second claim: $\tau^\delta_\varsigma \to \infty$ in probability. Let $p \in (\alpha, 1)$ be an arbitrary real number. Owing to the Markov inequality and to the continuous embedding $H^{\lfloor d/2\rfloor+2}_x \subset \mathcal{C}^1_x$,

for all $T \in (0,\infty)$ and for all $\delta \in (0,\varepsilon_0]$, one has

$$\begin{aligned}
\mathbb{P}\left(\tau_\zeta^\delta < T\right) &= \mathbb{P}\left(\tau_\zeta^\delta < T, \tau_\zeta^\delta < \tau_m^\delta\right) + \mathbb{P}\left(\tau_\zeta^\delta < T, \tau_\zeta^\delta \geqslant \tau_m^\delta\right) \\
&\leqslant \mathbb{P}\left(\delta \sup_{t\in[0,T]} \left\|\zeta^{\delta,\tau^\delta}(t)\right\|_{\mathcal{C}_x^1} \geqslant 1\right) + \mathbb{P}\left(\tau_m^\delta < T\right) \\
&\leqslant \delta^2 \mathbb{E}\left[\sup_{t\in[0,T]} \left\|\zeta^{\delta,\tau^\delta}(t)\right\|_{\mathcal{C}_x^1}^2\right] + \mathbb{P}\left(\tau_m^\delta < T\right) \\
&\leqslant \delta^{2(1-p)} \sup_\delta \mathbb{E}\left[\delta^{2p} \sup_{t\in[0,T]} \left\|\zeta^{\delta,\tau^\delta}(t)\right\|_{H_x^{\lfloor d/2\rfloor+2}}^2\right] + \mathbb{P}\left(\tau_m^\delta < T\right).
\end{aligned}$$

Owing to the first claim, it thus remains to prove that

$$\sup_\delta \mathbb{E}\left[\delta^{2p} \sup_{t\in[0,T]} \left\|\zeta^{\delta,\tau^\delta}(t)\right\|_{H_x^{\lfloor d/2\rfloor+2}}^2\right] < \infty,$$

for all $T \in (0,\infty)$.

Let β be an arbitrary multi-index of size $|\beta| \leqslant \lfloor d/2\rfloor + 2$. For all $\delta \in (0,\delta_0]$ and $T \in (0,\infty)$, one has

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left\|\frac{\partial^{|\beta|}\zeta^{\delta,\tau^\delta}(t)}{\partial x^\beta}\right\|_{L_x^2}^2\right] \leqslant \int_{\mathbb{T}^d} \mathbb{E}\left[\sup_{t\in[0,T]} \left|\frac{\partial^{|\beta|}\zeta^{\delta,\tau^\delta}(t,x)}{\partial x^\beta}\right|^2\right] dx. \tag{45}$$

Below an upper bound for the expectation on the right-hand side of (45) is obtained using properties of a well-chosen martingale. For all $x \in \mathbb{T}^d$ and all multi-indices β such that $|\beta| \leqslant \lfloor d/2\rfloor + 2$, introduce the auxiliary function θ_x^β defined by

$$\theta_x^\beta(\ell) = \frac{\partial^{|\beta|}\sigma(\ell)}{\partial x^\beta}(x),$$

for all $n \in E$. Observe that θ_x^β is an element of the set $E^*(\sigma)$, and one may define $\psi_x^\beta \doteq R_0(\theta_x^\beta - \theta_x^\beta(\overline{\sigma}))$, which solves of the Poisson equation $-\mathcal{L}_m\psi_x^\beta = \left(\theta_x^\beta - \theta_x^\beta(\overline{\sigma})\right)$, see Definition 2.1. Observe that one has the identities

$$\begin{aligned}
\frac{\partial^{|\beta|}\zeta^\delta(t,x)}{\partial x^\beta} &= \frac{1}{\delta}\int_0^t \frac{\partial^{|\beta|}(\sigma(m^\delta(s)) - \overline{\sigma})}{\partial x^\beta}(x) ds \\
&= \frac{1}{\delta}\int_0^t \left(\theta_x^\beta(m^\delta(s)) - \overline{\theta}_x^\beta\right) ds \\
&= -\frac{1}{\delta}\int_0^t \mathcal{L}_m\psi_x^\beta(m^\delta(s)) ds.
\end{aligned}$$

Finally, for all $x \in \mathbb{T}^d$, $\delta \in (0,\delta_0]$ and $t \geqslant 0$, set

$$\begin{aligned}
M^\delta_{\delta^{1+p}\psi_x^\beta}(t) &= \delta^{1+p}\psi_x^\beta(m^\delta(t)) - \delta^{1+p}\psi_x^\beta(m(0)) - \frac{1}{\delta^2}\int_0^t \delta^{1+p}\mathcal{L}_m\psi_x^\beta(m^\delta(s)) ds \\
&= \delta^{1+p}\psi_x^\beta(m^\delta(t)) - \delta^{1+p}\psi_x^\beta(m(0)) + \delta^p \frac{\partial^{|\beta|}\zeta^\delta(t,x)}{\partial x^\beta}
\end{aligned} \tag{46}$$

Then the stopped process $(M^{\delta}_{\delta^{1+p}})^{\tau^\delta} = M^{\delta}_{\delta^{1+p}}(\cdot \wedge \tau^\delta)$ is a martingale. In addition, one obtains the upper bound

$$\mathbb{E}\left[\delta^{2p}\sup_{t\in[0,T]}\left|\frac{\partial^{|\beta|}\zeta^{\delta,\tau^\delta}(t,x)}{\partial x^\beta}\right|^2\right] \leqslant 2\mathbb{E}\left[\sup_{t\in[0,T]}\left|\delta^{1+p}\psi_x^\beta(m^{\delta,\tau^\delta}(t))\right|^2\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|M^{\delta,\tau^\delta}_{\delta^{1+p}\psi_x^\beta}(t)\right|^2\right].$$

On the one hand, owing to estimates (6) and (16), one obtains

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\delta^{1+p}\psi_x^\beta(m^{\delta,\tau^\delta}(t))\right|^2\right] \lesssim \delta^{2+2p}\left(1+\delta^{-2\alpha}\right) \lesssim 1,$$

since $\alpha < 1$.

On the other hand, Doob's Maximal Inequality yields

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}\left|M^{\delta,\tau^\delta}_{\delta^{1+p}\psi_x^\beta}(t)\right|^2\right] &\leqslant 4\mathbb{E}\left[\left|M^{\delta,\tau^\delta}_{\delta^{1+p}\psi_x^\beta}(T)\right|^2\right] \\
&\leqslant \frac{4}{\delta^2}\mathbb{E}\left[(\delta^{1+p})^2\int_0^{T\wedge\tau^\delta}\left(\mathcal{L}_m\left(\left(\psi_x^\beta\right)^2\right) - 2\psi_x^\beta\mathcal{L}_m\psi_x^\beta\right)(m^\delta(s))ds\right] \\
&\lesssim_T \delta^{2p}\left(1+\delta^{-2\alpha}\right) \\
&\lesssim 1,
\end{aligned}$$

since $p \in (\alpha, 1)$, using Assumption 5 and the estimates (6) and (16).

Since all the bounds obtained above are uniform with respect to $x \in \mathbb{T}^d$, gathering the estimates one obtains (46) yields

$$\int_{x\in\mathbb{T}^d}\mathbb{E}\left[\delta^{2p}\sup_{t\in[0,T]}\left|\frac{\partial^{|\beta|}\zeta^{\delta,\tau^\delta}(t,x)}{\partial x^\beta}\right|^2\right]dx \lesssim 1.$$

The arguments explained above then yield the result: for all $T \in (0,\infty)$,

$$\mathbb{P}\left(\tau_\zeta^\delta < T\right) \lesssim \delta^{2(1-p)} + \mathbb{P}\left(\tau_m^\delta < T\right) \underset{\delta\to 0}{\to} 0,$$

therefore one has $\tau_\zeta^\delta \to \infty$ in probability.

Since $\tau^\delta = \tau_m^\delta \wedge \tau_\zeta^\delta$, the proof of Proposition 4.1 is completed. □

5.2 A priori estimate of the solution in $L^2(\mathcal{M}^{-1})$

The goal of this section is to provide the proof of Proposition 4.2, which gives an almost sure a priori estimate for $f^{\varepsilon,\delta,\tau^\delta}$ in the space $C_T^0L^2(\mathcal{M}^{-1})$ and for $Lf^{\varepsilon,\delta,\tau^\delta}$ in the space $L^2_{t\wedge\tau^\delta}L^2(\mathcal{M}^{-1})$, in terms of the norm of the initial condition $\left\|f_0^{\varepsilon,\delta}\right\|_{L^2(\mathcal{M}^{-1})}$.

Proof of Proposition 4.2. For all $\delta \in (0,\delta_0]$ and all $t \geqslant 0$, set

$$\eta^\delta(t,\cdot) = \int_0^t \sigma(m^\delta(s))(\cdot)ds = \delta\zeta^\delta(t,\cdot) + t\overline{\sigma}(\cdot) \in \mathcal{C}_x^{\lfloor d/2\rfloor+2},$$

where ζ^δ is defined by(15). Note that, owing to (17), for all $\delta \in (0, \delta_0]$, and $i = 0, 1$, one has

$$\sup_{t\in[0,\tau^\delta\wedge T]} \left\|\eta^\delta(t)\right\|_{\mathcal{C}^i_x} \leqslant 1 + T\left\|\overline{\sigma}\right\|_{\mathcal{C}^i_x}. \tag{47}$$

Let us now introduce the following family of weight functions $\mathcal{M}^\delta$ indexed by $\delta \in (0, \delta_0]$: for all $t \geqslant 0$, $x \in \mathbb{T}^d$ and $v \in V$, set

$$\mathcal{M}^\delta(t, x, v) = \exp\left(-2\eta^\delta(t, x)\right)\mathcal{M}(v),$$

The associated weighted L^2 norm is defined by

$$\|f\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} \doteq \iint \frac{|f(x, v)|^2}{\mathcal{M}^\delta(t, x, v)} dxd\mu(v).$$

Note that for all $h \in L^2(\mathcal{M}^{-1})$ and all $t \in [0, \tau^\delta \wedge T]$, the inequality (47) yields

$$\|h\|^2_{L^2(\mathcal{M}^{-1})} \leqslant \|h\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} e^{2+2T\|\overline{\sigma}\|_{\mathcal{C}^0_x}}. \tag{48}$$

It is thus sufficient to prove estimates in the weight norm $\|\cdot\|_{L^2(\mathcal{M}^\delta(t)^{-1})}$, to retrieve the a priori estimates in the space $L^2(\mathcal{M}^{-1})$, when the condition $t \leqslant \tau^\delta \wedge T$ is satisfied.

For all $t \geqslant 0$, one has

$$\begin{aligned}
\frac{1}{2}\partial_t \left\|f^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} &= \iint \frac{f^{\varepsilon,\delta}(t, x, v)}{\mathcal{M}^\delta(t, x, v)} \partial_t f^{\varepsilon,\delta}(t, x, v) dxd\mu(v) \\
&\quad - \iint \frac{\left|f^{\varepsilon,\delta}(t, x, v)\right|^2}{2\left|\mathcal{M}^\delta(t, x, v)\right|^2} \partial_t \mathcal{M}^\delta(t, x, v) dxd\mu(v) \\
&= \mathcal{A}_{\varepsilon,\delta}(t) + \mathcal{B}_{\varepsilon,\delta}(t) + \mathcal{C}_{\varepsilon,\delta}(t)
\end{aligned}$$

with

$$\begin{aligned}
\mathcal{A}_{\varepsilon,\delta}(t) &= \frac{1}{\varepsilon^2} \iint \frac{f^{\varepsilon,\delta}(t, x, v)}{\mathcal{M}^\delta(t, x, v)} Lf^{\varepsilon,\delta}(t, x, v) dxd\mu(v) \\
\mathcal{B}_{\varepsilon,\delta}(t) &= -\frac{1}{\varepsilon} \iint \frac{f^{\varepsilon,\delta}(t, x, v)}{\mathcal{M}^\delta(t, x, v)} (a(v) + \varepsilon b(v)) \cdot \nabla_x f^{\varepsilon,\delta}(t, x, v) dxd\mu(v) \\
\mathcal{C}_{\varepsilon,\delta}(t) &= -\iint \frac{\left|f^{\varepsilon,\delta}(t, x, v)\right|^2}{\mathcal{M}^\delta(t, x, v)} \left(\sigma(m^\delta) + \frac{\partial_t \mathcal{M}^\delta}{2\mathcal{M}^\delta}\right)(t, x, v) dxd\mu(v).
\end{aligned}$$

Note that the third term vanishes: $\mathcal{C}_{\varepsilon,\delta}(t) = 0$ for all $t \geqslant 0$. Indeed, the definition of the weight function $\mathcal{M}^\delta$ yields the identity $\sigma(m^\delta) + \frac{\partial_t \mathcal{M}^\delta}{2\mathcal{M}^\delta} = 0$.

Using the identity $f^{\varepsilon,\delta} = \rho^{\varepsilon,\delta}\mathcal{M} - Lf^{\varepsilon,\delta}$ and the property $\int_V Lf^{\varepsilon,\delta}(t, x, v)d\mu(v) = 0$, the first term $\mathcal{A}_{\varepsilon,\delta}(t)$ is written as follows: for all $t \geqslant 0$,

$$\begin{aligned}
\mathcal{A}_{\varepsilon,\delta}(t) &= \frac{1}{\varepsilon^2} \iint \frac{f^{\varepsilon,\delta}(t, x, v)}{\mathcal{M}^\delta(t, x, v)} Lf^{\varepsilon,\delta}(t, x, v) dxd\mu(v) \\
&= \frac{1}{\varepsilon^2} \int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)} \rho^{\varepsilon,\delta}(t, x) \int_V Lf^{\varepsilon,\delta}(t, x, v) d\mu(v) dx - \frac{1}{\varepsilon^2} \iint \frac{\left|Lf^{\varepsilon,\delta}(t, x, v)\right|^2}{\mathcal{M}^\delta(t, x, v)} d\mu(v) dx \\
&= -\frac{1}{\varepsilon^2} \left\|Lf^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}.
\end{aligned}$$

The treatment of the second term $\mathcal{B}_{\varepsilon,\delta}(t)$ requires technical computations. Using an integration by parts arguments, for all $t \geqslant 0$ one obtains

$$\begin{aligned}\mathcal{B}_{\varepsilon,\delta}(t) &= -\frac{1}{\varepsilon}\iint (a(v)+\varepsilon b(v))\cdot \frac{f^{\varepsilon,\delta}(t,x,v)\nabla_x f^{\varepsilon,\delta}(t,x,v)}{\mathcal{M}^\delta(t,x,v)}dxd\mu(v)\\ &= -\frac{1}{\varepsilon}\iint (a(v)+\varepsilon b(v))\cdot \frac{\frac{1}{2}\left|f^{\varepsilon,\delta}(t,x,v)\right|^2\nabla_x\mathcal{M}^\delta(t,x,v)}{\left|\mathcal{M}^\delta(t,x,v)\right|^2}dxd\mu(v)\\ &= \frac{1}{\varepsilon}\iint (a(v)+\varepsilon b(v))\cdot\nabla_x\eta^\delta(t,x)\frac{\left|f^{\varepsilon,\delta}(t,x,v)\right|^2}{\mathcal{M}^\delta(t,x,v)}d\mu(v)dx.\end{aligned}$$

Using again the identity $f^{\varepsilon,\delta} = \rho^{\varepsilon,\delta}\mathcal{M} - Lf^{\varepsilon,\delta}$ then gives, for all $t \geqslant 0$,

$$\begin{aligned}\mathcal{B}_{\varepsilon,\delta}(t) &= \frac{1}{\varepsilon}\int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\left|\rho^{\varepsilon,\delta}(t,x)\right|^2\nabla_x\eta^\delta(t,x)\cdot\left(\int_V (a(v)+\varepsilon b(v))\mathcal{M}(v)d\mu(v)\right)dx\\ &\quad+\frac{1}{\varepsilon}\iint (a(v)+\varepsilon b(v))\cdot\nabla_x\eta^\delta(t,x)\frac{\left|Lf^{\varepsilon,\delta}(t,x,v)\right|^2}{\mathcal{M}^\delta(t,x,v)}d\mu(v)dx\\ &\quad-\frac{2}{\varepsilon}\iint (a(v)+\varepsilon b(v))\cdot\nabla_x\eta^\delta(t,x)\rho^{\varepsilon,\delta}(t,x)Lf^{\varepsilon,\delta}(t,x,v)\frac{\mathcal{M}(v)}{\mathcal{M}^\delta(t,x,v)}d\mu(v)dx\\ &= \mathcal{B}^1_{\varepsilon,\delta}(t)+\mathcal{B}^2_{\varepsilon,\delta}(t)+\mathcal{B}^3_{\varepsilon,\delta}(t).\end{aligned}$$

Let us now treat successively the three terms appearing on the right-hand side above.

- Owing to Assumption 1 and to the definition of J, one has $\int_V (a(v)+\varepsilon b(v))\mathcal{M}(v)d\mu(v) = \varepsilon J$. Therefore

$$\begin{aligned}\left|\mathcal{B}^1_{\varepsilon,\delta}(t)\right| &= \left|\int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\left|\rho^{\varepsilon,\delta}(t,x)\right|^2 J\cdot\nabla_x\eta^\delta(t,x)dx\right|\\ &\leqslant \|b\|_{L^\infty}\left\|\nabla_x\eta^\delta(t)\right\|_{C_x}\int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\left|\rho^{\varepsilon,\delta}(t,x)\right|^2 dx.\end{aligned}$$

  Owing to the Cauchy-Schwarz inequality, one obtains

$$\begin{aligned}\int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\left|\rho^{\varepsilon,\delta}(t,x)\right|^2 dx &\leqslant \int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\int_V\frac{\left|f^{\varepsilon,\delta}(t,x,v)\right|^2}{\mathcal{M}^\delta(t,x,v)}d\mu(v)\int_V\mathcal{M}^\delta(t,x,v)d\mu(v)dx\\ &= \left\|f^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})},\end{aligned}\tag{49}$$

  since $\int_V \mathcal{M}^\delta(t,x,v)d\mu(v) = e^{-2\eta^\delta(t,x)}$ for all $t \geqslant 0$ and all $x \in \mathbb{T}^d$. As a consequence, using the inequality (47), for all $t \in [0, \tau^\delta \wedge T]$, one has

$$\left|\mathcal{B}^1_{\varepsilon,\delta}(t)\right| \leqslant \|b\|_{L^\infty}\left(1+T\|\overline{\sigma}\|_{\mathcal{C}^1_x}\right)\left\|f^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}.$$

- Using the condition $\varepsilon \leqslant \varepsilon_0$ with ε_0 satisfying (14), and the inequality (47), for $t \in [0, \tau^\delta \wedge T]$ one has

$$\left|\mathcal{B}^2_{\varepsilon,\delta}(t)\right| \leqslant \frac{1}{4\varepsilon^2}\left\|Lf^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}.$$

- Using Young's inequality, then using the inequalities (47) and (49), one obtains, for all $t \in [0, \tau^\delta \wedge T]$,

$$\begin{aligned}\left|\mathcal{B}^3_{\varepsilon,\delta}(t)\right| &\leqslant 4(\|a\|_{L^\infty}+\|b\|_{L^\infty})^2\left\|\nabla_x\eta^\delta(t)\right\|^2_{C_x}\int_{\mathbb{T}^d} e^{2\eta^\delta(t,x)}\left|\rho^{\varepsilon,\delta}(t,x)\right|^2 dx+\frac{1}{4\varepsilon^2}\left\|Lf^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}\\ &\leqslant 4(\|a\|_{L^\infty}+\|b\|_{L^\infty})^2(1+T\|\overline{\sigma}\|_{\mathcal{C}^1_x})^2\left\|f^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}+\frac{1}{4\varepsilon^2}\left\|Lf^{\varepsilon,\delta}(t)\right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}.\end{aligned}$$

Gathering the estimates, one obtains the following inequalities: for all $t \in [0, \tau^\delta \wedge T]$,

$$|\mathcal{B}_{\varepsilon,\delta}(t)| \leqslant C_0(T) \left\| f^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} + \frac{1}{2\varepsilon^2} \left\| Lf^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}$$

and

$$\frac{1}{2}\partial_t \left\| f^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} + \frac{1}{2\varepsilon^2} \left\| Lf^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} \leqslant C_0(T) \left\| f^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})}.$$

where $C_0(T) \in (0, \infty)$ is a deterministic real number, and does not depend on ε, δ. Applying Gronwall's inequality, for all $t \in [0, \tau^\delta \wedge T]$, one gets

$$\left\| f^{\varepsilon,\delta}(t) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} + \frac{1}{2\varepsilon^2} \int_0^t \left\| Lf^{\varepsilon,\delta}(s) \right\|^2_{L^2(\mathcal{M}^\delta(t)^{-1})} ds \leqslant C(T) \left\| f_0^{\varepsilon,\delta} \right\|^2_{L^2(\mathcal{M}^{-1})},$$

where $C(T) \in (0, \infty)$ is a deterministic real number, and does not depend on ε, δ. Finally, using (48), one obtains the a priori estimate (18), which concludes the proof of Proposition 4.2. □

5.3 Construction of the perturbed test function $\varphi^{\varepsilon,\delta}$

The objective of this section is to prove Proposition 4.3: more precisely, given a test function $\varphi \in \Theta_{\text{lim}}$ (which only depends on $\rho = \langle f \rangle$), we construct the four correctors $\varphi_{1,0}$, $\varphi_{2,0}$, $\varphi_{0,2}$ and $\varphi_{1,2}$, such that the perturbed test function $\varphi^{\varepsilon,\delta}$ defined by (23) satisfies the error estimates (24) and (29), and appropriate upper bounds.

Below, we first state auxiliary results concerning solutions of Poisson equations. The expression of $\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}$ is then expanded in powers of ε and δ, and the resulting equality yields a family of equations to be satisfied by the correctors in order to satisfy (24). The correctors are finally constructed successively as solutions of appropriate Poisson equations. Eventually, it only remains to check the required regularity properties and upper bounds, this step follows from straightforward computations.

5.3.1 Auxiliary results on Poisson equations

As will be clear below, the construction of the correctors requires to solve Poisson equations of the type $-\mathcal{L}_2\psi(f, \ell) = \vartheta(f, \ell)$ (where ℓ is considered as a fixed parameter) and $-\mathcal{L}_m\psi(f, \ell) = \vartheta(f, \ell)$ (where f is considered as a fixed parameter). We describe below the corresponding centering conditions which are needed for the solvability of those Poisson equations, and give the expressions of the solutions.

To solve the first class of Poisson equations, let us introduce the process $(g_f(t))_{t\in\mathbb{R}^+}$, associated with the infinitesimal generator $\mathcal{L}_2$, with the initial condition $g_f(0) = f$: for all $t \geqslant 0$, one has

$$g_f(t) = \rho\mathcal{M} + e^{-t}(f - \rho\mathcal{M}),$$

where $\rho = \langle f \rangle = \langle g \rangle_f(t)$ for all $t \geqslant 0$.

The solvability of the first class of Poisson equations $-\mathcal{L}_2\psi = \vartheta$ is ensured when the following centering condition is satisfied: for all $\rho \in L^2_x$ and all $\ell \in E$,

$$\vartheta(\rho\mathcal{M}, \ell) = 0. \tag{50}$$

If (50) is satisfied, then the function ψ defined by

$$\psi(f, \ell) = \int_0^\infty \vartheta(g_f(t), \ell)dt, \tag{51}$$

for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, is a solution of the Poisson equation $-\mathcal{L}_2\psi = \vartheta$. It is the unique solution such that $\psi(\rho\mathcal{M}, \ell) = 0$ for all $\rho \in L^2_x$ and $\ell \in E$.

The solvability of the second class of Poisson equations $-\mathcal{L}_m\psi = \vartheta$ is ensured when the following centering condition is satisfied: for all $f \in L^2(\mathcal{M}^{-1})$,

$$\int_E \vartheta(f, \ell) d\nu(\ell) = 0. \tag{52}$$

If (52) is satisfied, then the function ψ defined by

$$\psi(f, \ell) = \int_0^\infty \mathbb{E}\left[\vartheta(f, m_\ell(t))\right] dt. \tag{53}$$

for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, is a solution of the Poisson equation $-\mathcal{L}_m\psi = \vartheta$. It is the unique solution such that $\int \psi(f, \ell) d\nu(\ell) = 0$ for all $f \in L^2(\mathcal{M}^{-1})$.

In the sequel, the following class of functions ϑ is considered. For all $h, k \in L^2_x$, define $\theta_{h,k}(\ell) = (h\sigma(\ell), k)_{L^2_x}$ for all $\ell \in E$. Observe that $\theta_{h,k} \in E^*(\sigma)$ (see Definition 2.1). For all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, set $\vartheta_{h,k}(f, \ell) = \theta_{h,k}(n) - \overline{\theta}_{h,k}$, with $\overline{\theta}_{h,k} = \int \theta_{h,k}(\ell) d\nu(\ell) = (h\overline{\sigma}, k)_{L^2_x}$. For such functions $\vartheta_{h,k}$, the solution of the Poisson equation $-\mathcal{L}_m\psi_{h,k} = \vartheta_{h,k} = \theta_{h,k} - \overline{\theta}_{h,k}$ is given by

$$\psi_{h,k} = R_0\left(\theta_{h,k}(\ell) - \overline{\theta}_{h,k}\right),$$

where the resolvent operator R_0 is introduced in Definition 2.1. Using the estimate (6) and the to Riesz representation Theorem, for all $\ell \in E$, there exists a bounded linear operator $R_0(\ell) : L^2_x \to L^2_x$ such that, for all $\ell \in E$ and $h, k \in L^2_x$, $\psi_{h,k}(\ell) = (R_0(\ell)h, k)_{L^2}$. Let us state some useful properties of the operators $R_0(\ell)$.

(i) For all $\ell \in E$, $R_0(\ell)$ is self-adjoint.

(ii) For all $\ell \in E$ and $h \in L^2_x$, one has

$$\|R_0(\ell)h\|_{L^2} \lesssim \|h\|_{L^2}\left(1 + \|\ell\|_E\right).$$

(iii) For all $\ell \in E$ and all $h \in H^1_x$, one has $R_0(\ell)h \in H^1_x$, and

$$\|R_0(\ell)h\|_{H^1_x} \lesssim \|h\|_{H^1_x}\left(1 + \|\ell\|_E\right).$$

The proof of Claim (i) is straightforward: for all $h, k \in L^2_x$, $\theta_{h,k} = \theta_{k,h}$, thus $\psi_{h,k} = \psi_{k,h}$, which gives for all $\ell \in E$

$$(R_0(\ell)h, k)_{L^2} = (h, R_0(\ell)k)_{L^2}.$$

Claim (ii) follows from the estimate (6): one has $|\psi_{h,k}(\ell)| \lesssim \|h\|_{L^2}\|k\|_{L^2}(1 + \|\ell\|_E)$, since $\mathrm{Lip}(\theta_{h,k}) \leqslant \mathrm{Lip}(\sigma)\|h\|_{L^2}\|k\|_{L^2}$.

Finally, Claim (iii) is obtained as follows: since $E \in \mathcal{C}^1_x$, one has $\|\theta_{h,k}\|_{\mathrm{Lip}(E)} \leqslant \|h\|_{H^1_x}\|k\|_{H^{-1}_x}$ if $h \in H^1_x$ and $k \in L^2_x$. Using the estimate (6) then gives $|\psi_{h,k}(\ell)| \lesssim \|h\|_{H^1_x}\|k\|_{H^{-1}_x}(1 + \|\ell\|_E)$.

5.3.2 Multiscale expansion and family of Poisson equations

Let $\varphi^{\varepsilon,\delta}$ be of the form (23), where the correctors are not known at this stage. Then $\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}$ is expressed as follows:

$$\begin{aligned}\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} = {} & \varepsilon^{-2}\mathcal{L}_2\varphi + \delta^{-2}\mathcal{L}_m\varphi \\ & + \varepsilon^{-1}\left(\mathcal{L}_1\varphi + \mathcal{L}_2\varphi_{1,0}\right) + \varepsilon\delta^{-2}\mathcal{L}_m\varphi_{1,0} \\ & + \left(\mathcal{L}_0\varphi + \mathcal{L}_1\varphi_{1,0} + \mathcal{L}_2\varphi_{2,0} + \mathcal{L}_m\varphi_{0,2}\right) + \varepsilon^{-2}\delta^2\mathcal{L}_2\varphi_{0,2} + \varepsilon^2\delta^{-2}\mathcal{L}_m\varphi_{2,0} \\ & + \varepsilon\left(\mathcal{L}_0\varphi_{1,0} + \mathcal{L}_1\varphi_{2,0} + \mathcal{L}_m\varphi_{1,2}\right) + \varepsilon^{-1}\delta^2\left(\mathcal{L}_1\varphi_{0,2} + \mathcal{L}_2\varphi_{1,2}\right) \\ & + \varepsilon^2\mathcal{L}_0\varphi_{2,0} + \delta^2\mathcal{L}_0\varphi_{0,2} + \delta^2\mathcal{L}_1\varphi_{1,2} \\ & + \varepsilon\delta^2\mathcal{L}_0\varphi_{1,2},\end{aligned}$$

where each line on the right-hand side corresponds to expressions of degree $-2,\dots,3$ in terms of the variables ε,δ. The goal is to construct the correctors such that $\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta} - \mathcal{L}\varphi$ goes to 0 when $(\varepsilon,\delta)\to(0,0)$, more precisely such that (24) holds. The following family of conditions provide sufficient conditions on the correctors to satisfy (24): on the one hand, the correctors solve the following system of equations,

$$\mathcal{L}_2\varphi = \mathcal{L}_m\varphi = 0, \tag{54}$$
$$\mathcal{L}_1\varphi + \mathcal{L}_2\varphi_{1,0} = \mathcal{L}_m\varphi_{1,0} = 0, \tag{55}$$
$$\mathcal{L}_0\varphi + \mathcal{L}_1\varphi_{1,0} + \mathcal{L}_2\varphi_{2,0} + \mathcal{L}_m\varphi_{0,2} = \mathcal{L}\varphi, \tag{56}$$
$$\mathcal{L}_2\varphi_{0,2} = \mathcal{L}_m\varphi_{2,0} = 0, \tag{57}$$
$$\mathcal{L}_1\varphi_{0,2} + \mathcal{L}_2\varphi_{1,2} = 0, \tag{58}$$

on the other hand, the following estimates are satisfied,

$$|\mathcal{L}_0\varphi_{1,0}(f,\ell)| + |\mathcal{L}_1\varphi_{2,0}(f,\ell)| + |\mathcal{L}_m\varphi_{1,2}(f,\ell)| + |\mathcal{L}_0\varphi_{2,0}(f,\ell)| \lesssim (1+\|f\|^3_{L^2(\mathcal{M}^{-1})})(1+\|\ell\|_E), \tag{59}$$
$$|\mathcal{L}_0\varphi_{0,2}(f,\ell)| + |\mathcal{L}_1\varphi_{1,2}(f,\ell)| + |\mathcal{L}_0\varphi_{1,2}(f,\ell)| \lesssim (1+\|f\|^3_{L^2(\mathcal{M}^{-1})})(1+\|\ell\|^2_E), \tag{60}$$

for all $f\in L^2(\mathcal{M}^{-1})$ and $\ell\in E$. The estimates (25), (26), (26) and (28) are byproducts of the constructions of the correctors below.

Observe that the last equation in the system, Equation (58), would not appear if $\varepsilon=\delta$, or if a constraint of the type $\varepsilon^{-1}\delta^2\to 0$ is satisfied. In the expression of $\mathcal{L}^{\varepsilon,\delta}\varphi^{\varepsilon,\delta}$ above, the condition (58) corresponds to a contribution of the term of degree 1. One of the novelties of this work is to consider the general case, hence the need to construct the corrector $\varphi_{1,2}$. On the contrary, as will be clear below, the condition (57) on the correctors $\varphi_{0,2}$ and $\varphi_{2,0}$ is simpler to treat, it only means that $\varphi_{0,2}(f,\ell)=\varphi_{0,2}(\rho\mathcal{M},\ell)$ and $\varphi_{2,0}(f,\ell)=\varphi_{2,0}(f)$. In the case $\varepsilon=\delta$, this only consists in writing the corrector φ_2 of order 2 as a sum of two terms $\varphi_{2,0}+\varphi_{0,2}$.

Recall that the test function φ belongs to the class of test functions $\Theta_{\lim}$ (see Definition 4.2): $\varphi(f,\ell)=\chi((\rho,\xi)_{L^2_x})$. In the sequel, to simplify the expressions, we use the notation $\chi_\rho = \chi((\rho,\xi)_{L^2_x})$, $\chi'_\rho=\chi'((\rho,\xi)_{L^2_x})$, and similar notation for the higher order derivatives χ''_ρ and χ'''_ρ.

Let us explain how the rest of this section proceeds. We first check that (54) holds when φ is a function in the class $\Theta_{\lim}$. Then, we successively construct the correctors $\varphi_{1,0}$, $\varphi_{2,0}$ and $\varphi_{0,2}$, and $\varphi_{1,2}$, as solutions of Poisson equations, using the tools above. We finally check that (59) and (60) holds. The proof of Proposition 4.3 is concluded when all those steps are completed.

5.3.3 Verification of the condition (54)

The function φ is in the class $\Theta_{\lim}$: as a consequence $\varphi(f,\ell) = \varphi(\rho\mathcal{M})$ does not depend on ℓ, and only depends on f through $\rho = \langle f \rangle$. It is then straightforward to check that $\mathcal{L}_m\varphi(f,\ell) = \mathcal{L}_2\varphi(f,\ell) = 0$ for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$. Therefore (54) holds.

5.3.4 Construction of the corrector $\varphi_{1,0}$ using the condition (55)

First, using (53), the condition $\mathcal{L}_m\varphi_{1,0} = 0$ implies that $\varphi_{1,0}$ is independent of ℓ: $\varphi_{1,0}(f,\ell) = \varphi(f)$ for all $f \in L^2(\mathcal{M}^{-1}), \ell \in E$.

Second, $\varphi_{1,0}$ is the solution of the Poisson equation

$$-\mathcal{L}_2\varphi_{1,0}(f) = \mathcal{L}_1\varphi(f) = -\chi'_\rho\left(\langle Af\rangle, \xi\right)_{L^2_x}.$$

The centering condition (50) is satisfied for $\vartheta = \mathcal{L}_1\varphi$, indeed $\langle A\rho\mathcal{M}\rangle = 0$ for all $\rho \in L^2_x$. Using the expression (51), one thus defines the corrector $\varphi_{1,0}$ as follows: for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, set

$$\varphi_{1,0}(f,\ell) = \int_0^\infty \mathcal{L}_1\varphi(g_f(t))dt.$$

Recall that $\langle g_f(t)\rangle = \rho$ and $\langle Ag_f(t)\rangle = e^{-t}\langle Af\rangle$ for all $t \geqslant 0$, thus one has, for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$,

$$\begin{aligned}\varphi_{1,0}(f,\ell) &= -\int_0^\infty e^{-t}\chi'_\rho\left(\langle Af\rangle, \xi\right)_{L^2_x} dt = -\chi'_\rho\left(\langle Af\rangle, \xi\right)_{L^2_x} \\ &= \chi'_\rho\left(f, A\xi\right)_{L^2}, \end{aligned} \tag{61}$$

using an integration by parts argument in the last equality.

Using the expression (61), it is then straightforward to check that the estimate (25) is satisfied. This concludes the construction of the corrector $\varphi_{1,0}$.

5.3.5 Construction of the correctors $\varphi_{2,0}$ and $\varphi_{0,2}$ using the conditions (56)–(57)

It is required to combine the two conditions (56) and (57) in order to build the correctors $\varphi_{2,0}$ and $\varphi_{0,2}$.

First, using (53), the condition $\mathcal{L}_m\varphi_{2,0} = 0$ implies that $\varphi_{2,0}$ is independent of ℓ: $\varphi_{2,0}(f,\ell) = \varphi(f)$ for all $f \in L^2(\mathcal{M}^{-1}), \ell \in E$. Similarly, using (51), the condition $\mathcal{L}_2\varphi_{0,2} = 0$ implies that $\varphi_{0,2}(f,\ell) = \varphi_{0,2}(\rho\mathcal{M},\ell)$, with $\rho = \langle f\rangle$, for all $f \in L^2(\mathcal{M}^{-1}), n \in E$.

Second, observe that one has

$$\begin{aligned}\mathcal{L}_0\varphi(f,\ell) + \mathcal{L}_1\varphi_{1,0}(f,\ell) &= -\chi'_\rho\left(\sigma(\ell)\rho, \xi\right)_{L^2_x} - \chi'_\rho\left(\langle Bf\rangle, \xi\right)_{L^2_x} \\ &\quad - \chi''_\rho\left(\langle Af\rangle, \xi\right)_{L^2_x}\left(f, A\xi\right)_{L^2} - \chi'_\rho\left(Af, A\xi\right)_{L^2} \\ &= \vartheta_{0,2}(\rho\mathcal{M},\ell) + \vartheta_{2,0}(f),\end{aligned}$$

for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, where the auxiliary functions $\vartheta_{0,2}$ and $\vartheta_{2,0}$ are defined as follows:

$$\begin{aligned}\vartheta_{0,2}(\rho\mathcal{M},\ell) &= -\chi'_\rho\left(\sigma(\ell)\rho, \xi\right)_{L^2_x} \\ \vartheta_{2,0}(f) &= \chi'_\rho\left(f, B\xi\right)_{L^2} + \chi''_\rho\left(\langle Af\rangle, \xi\right)^2_{L^2_x} + \chi'_\rho\left(f, A^2\xi\right)_{L^2} \\ &= \chi''_\rho\left(\langle Af\rangle, \xi\right)^2_{L^2_x} + \chi'_\rho\left(f, A^2\xi + B\xi\right)_{L^2},\end{aligned}$$

Note that using (21), for all $\rho \in L^2_x$, one has

$$\begin{aligned}
\mathcal{L}\varphi(\rho) &= \chi'_\rho\left(\rho, \mathrm{div}_x(K\nabla_x\xi)\right)_{L^2_x} + \chi'_\rho\left(\rho, J\cdot\nabla_x\xi\right)_{L^2_x} - \chi'_\rho\left(\rho,\overline{\sigma}\xi\right)_{L^2_x} \\
&= \chi'_\rho\left(\rho, K:\nabla^2_x\xi\right)_{L^2_x} + \chi'_\rho\left(\rho, J\cdot\nabla_x\xi\right)_{L^2_x} - \chi'_\rho\left(\overline{\sigma}\rho,\xi\right)_{L^2_x} \\
&= \chi'_\rho\left(\rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} - \chi'_\rho\left(\overline{\sigma}\rho,\xi\right)_{L^2_x} \\
&= \vartheta_{2,0}(\rho\mathcal{M}) + \int_E \vartheta_{0,2}(\rho\mathcal{M},\cdot)d\nu \\
&= \vartheta_{2,0}(\rho\mathcal{M}) + \int_E \vartheta_{0,2}(\rho\mathcal{M},\cdot)d\nu.
\end{aligned}$$

Indeed, $\langle A\rho\mathcal{M}\rangle = 0$ for all $\rho \in L^2_x$.

As a consequence, the correctors $\varphi_{2,0}$ and $\varphi_{0,2}$ are constructed as the solutions of the Poisson equations

$$-\mathcal{L}_2\varphi_{2,0}(f) = \left(\vartheta_{2,0}(f) - \vartheta_{2,0}(\rho\mathcal{M})\right), \tag{62}$$

$$-\mathcal{L}_m\varphi_{0,2}(\rho\mathcal{M},\ell) = \left(\vartheta_{0,2}(\rho\mathcal{M},\ell) - \int_E \vartheta_{0,2}(\rho\mathcal{M},\cdot)d\nu\right). \tag{63}$$

Indeed, if (62) and (63) are satisfied, then the arguments above show that the conditions (56) and (57) hold. It remains to solve the two Poisson equations, and to check that the estimates (26) and (27) are satisfied.

On the one hand, the centering condition (50) is satisfied, and (51) gives the following definition for $\varphi_{2,0}$: for all $f \in L^2(\mathcal{M}^{-1})$, set

$$\varphi_{2,0}(f) = \int_0^\infty \left(\vartheta_{2,0}(g_f(t)) - \vartheta_{2,0}(\rho\mathcal{M})\right) dt \tag{64}$$

Using the properties $\langle g_f(t)\rangle = \rho$, $\langle Ag_f(t)\rangle = e^{-t}\langle Af\rangle$ and $g_f(t) - \rho\mathcal{M} = e^{-t}(f - \rho\mathcal{M})$, for all $t \geqslant 0$, one obtains the following expression for $\varphi_{2,0}$: for all $f \in L^2(\mathcal{M}^{-1})$,

$$\begin{aligned}
\varphi_{2,0}(f) &= \int_0^\infty \chi''_\rho\left(\langle Ag_f(t)\rangle,\xi\right)^2_{L^2_x} dt + \int_0^\infty \chi'_\rho\left(g_f(t) - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} dt \\
&= \int_0^\infty e^{-2t}\chi''_\rho\left(\langle Af\rangle,\xi\right)^2_{L^2_x} dt + \int_0^\infty e^{-t}\chi'_\rho\left(f - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} dt \\
&= \frac{1}{2}\chi''_\rho\left(\langle Af\rangle,\xi\right)^2_{L^2_x} + \chi'_\rho\left(f - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} \\
&= \frac{1}{2}\chi''_\rho\left(f, A\xi\right)^2_{L^2} + \chi'_\rho\left(f - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2}.
\end{aligned} \tag{65}$$

On the other hand, the centering condition (52) is satisfied, and (53) gives the following definition for $\varphi_{0,2}$: for all $\rho \in L^2_x$ and $\ell \in E$, set

$$\begin{aligned}
\varphi_{0,2}(f,\ell) &= \int_0^\infty \mathbb{E}\left[\vartheta_{0,2}(\rho\mathcal{M}, m_\ell(t)) - \int_E \vartheta_{0,2}(\rho\mathcal{M},\cdot)d\nu\right] dt \\
&= -\int_0^\infty \mathbb{E}\left[\chi'_\rho\left((\sigma(m_\ell(t)) - \overline{\sigma})\rho,\xi\right)_{L^2_x}\right] dt \\
&= -\chi'_\rho\left(\rho, R_0(\ell)\xi\right)_{L^2_x}.
\end{aligned} \tag{66}$$

Using the expressions (65) and (66), it is then straightforward to check that the estimates (26) and (27) are satisfied. This concludes the construction of the correctors $\varphi_{2,0}$ and $\varphi_{0,2}$.

5.3.6 Construction of the corrector $\varphi_{1,2}$ using the condition (58)

The last step in the construction of the correctors, is to define $\varphi_{1,2}$: owing to (58), it is constructed as the solution of the Poisson equation $-\mathcal{L}_2\varphi_{1,2} = \mathcal{L}_1\varphi_{0,2}$. Using the expression (66) of $\varphi_{0,2}$, one has

$$\mathcal{L}_1\varphi_{0,2}(f,\ell) = \chi_\rho''\left(\langle Af\rangle, \xi\right)_{L^2_x}\left(\rho, R_0(\ell)\xi\right)_{L^2_x} + \chi_\rho'\left(\langle Af\rangle, R_0(\ell)\xi\right)_{L^2_x},$$

for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$.

Note that the centering condition (50) is satisfied: for all $\rho \in L^2_x$ and $\ell \in E$, one has

$$\mathcal{L}_1\varphi_{0,2}(\rho\mathcal{M},\ell) = 0,$$

since $\langle A\rho\mathcal{M}\rangle = 0$. Therefore, $\varphi_{1,2}$ is defined using (51): for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$,

$$\varphi_{1,2}(f,\ell) = \int_0^\infty \mathcal{L}_1\varphi_{0,2}(g_f(t),\ell)dt.$$

Using the properties $\langle g_f(t)\rangle = \rho$ and $\langle Ag_f(t)\rangle = e^{-t}\langle Af\rangle$ for all $t \geqslant 0$, one obtains the following expression for $\varphi_{1,2}$: for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$,

$$\begin{aligned}\varphi_{1,2}(f,\ell) &= \int_0^\infty e^{-t}\left(\chi_\rho''\left(\langle Af\rangle, \xi\right)_{L^2_x}\left(\rho, R_0(\ell)\xi\right)_{L^2_x} + \chi_\rho'\left(\langle Af\rangle, R_0(\ell)\xi\right)_{L^2_x}\right)dt \\ &= \chi_\rho''\left(\langle Af\rangle, \xi\right)_{L^2_x}\left(\rho, R_0(\ell)\xi\right)_{L^2_x} + \chi_\rho'\left(\langle Af\rangle, R_0(\ell)\xi\right)_{L^2_x} \\ &= -\chi_\rho''\left(f, A\xi\right)_{L^2}\left(\rho, R_0(\ell)\xi\right)_{L^2_x} - \chi_\rho'\left(f, A(R_0(\ell)\xi)\right)_{L^2}. \end{aligned} \tag{67}$$

Using the expression (67), it is then straightforward to check that the estimate (28) is satisfied. This concludes the construction of the corrector $\varphi_{1,2}$.

5.3.7 Verification of the conditions (59)–(60)

Note that $\varphi^{\varepsilon,\delta}$ defined by (23) belongs to the class of functions Θ given in Definition 4.3. It thus only remains to check that the conditions (59) and (60) hold. This is done using the following expressions: for all $f \in L^2(\mathcal{M}^{-1})$ and $\ell \in E$, one has

$$\mathcal{L}_0\varphi_{1,0}(f,\ell) = -\chi_\rho''\left(\sigma(\ell)\rho + \langle Bf\rangle, \xi\right)_{L^2_x}\left(f, A\xi\right)_{L^2} - \chi_\rho'\left(\sigma(\ell)f, A\xi\right)_{L^2},$$

$$\begin{aligned}\mathcal{L}_1\varphi_{2,0}(f,\ell) = &\frac{1}{2}\chi_\rho'''\left(f, A\xi\right)^3_{L^2} + \chi_\rho''\left(f, A\xi\right)_{L^2}\left(f, A^2\xi\right)_{L^2} \\ &+ \chi_\rho''\left(f, A\xi\right)_{L^2}\left(f - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} + \chi_\rho'\left(f - \rho\mathcal{M}, A^3\xi + AB\xi\right)_{L^2},\end{aligned}$$

$$\mathcal{L}_m\varphi_{1,2}(f,\ell) = \chi_\rho''\left(f, A\xi\right)_{L^2}\left(\rho, \sigma(\ell)\xi\right)_{L^2_x} + \chi_\rho'\left(f, A(\sigma(\ell)\xi)\right)_{L^2},$$

$$\begin{aligned}\mathcal{L}_0\varphi_{2,0}(f,\ell) = &-\frac{1}{2}\chi_\rho'''\left(\sigma(\ell)\rho + \langle Bf\rangle, \xi\right)_{L^2_x}\left(f, A\xi\right)^2_{L^2} - \chi_\rho''\left(f, A\xi\right)_{L^2}\left(\sigma(\ell)f + Bf, A\xi\right)_{L^2} \\ &- \chi_\rho''\left(\sigma(\ell)\rho + \langle Bf\rangle, \xi\right)_{L^2_x}\left(f - \rho\mathcal{M}, A^2\xi + B\xi\right)_{L^2} \\ &- \chi_\rho'\left(\sigma(\ell)(f - \rho\mathcal{M}) + B(f - \rho\mathcal{M}), A^2\xi + B\xi\right)_{L^2},\end{aligned}$$

$$\mathcal{L}_0\varphi_{0,2}(f,\ell) = -\chi_\rho''\left(\sigma(\ell)\rho + \langle Bf\rangle, \xi\right)_{L^2_x}\left(R_0(\ell)\rho, \xi\right)_{L^2_x} - \chi_\rho'\left(\sigma(\ell)\rho + \langle Bf\rangle, R_0(\ell)\xi\right)_{L^2_x},$$

$$\begin{aligned}\mathcal{L}_1\varphi_{1,2}(f,\ell) = &-\chi_\rho'''\left(f,A\xi\right)_{L^2}^2\left(\rho,R_0(\ell)\xi\right)_{L^2_x} - \chi_\rho''\left(f,A^2\xi\right)_{L^2_x}\left(\rho,R_0(\ell)\xi\right)_{L^2_x}\\ &- \chi_\rho''\left(f,A\xi\right)_{L^2}\left(\rho,A(R_0(\ell)\xi)\right)_{L^2_x} - \chi_\rho''\left(f,A\xi\right)_{L^2}\left(f,A(R_0(\ell)\xi)\right)_{L^2}\\ &- \chi_\rho'\left(f,A^2(R_0(\ell)\xi)\right)_{L^2},\end{aligned}$$

$$\begin{aligned}\mathcal{L}_0\varphi_{1,2}(f,\ell) = &\chi_\rho''\left(\sigma(\ell)\rho + \langle Bf\rangle,\xi\right)_{L^2_x}\left(f,A\xi\right)_{L^2}\left(\rho,R_0(\ell)\xi\right)_{L^2_x} + \chi_\rho''\left(\sigma(\ell)fBf,A\xi\right)_{L^2_x}\left(\rho,R_0(\ell)\xi\right)_{L^2_x}\\ &+ \chi_\rho''\left(f,A\xi\right)_{L^2}\left(\sigma(\ell)\rho + \langle Bf\rangle,R_0(\ell)\xi\right)_{L^2_x} + \chi_\rho''\left(\sigma(\ell)\rho + \langle Bf\rangle,\xi\right)_{L^2_x}\left(f,A(R_0(\ell)\xi)\right)_{L^2}\\ &+ \chi_\rho'\left(\sigma(\ell)f + Bf,A(R_0(\ell)\xi)\right)_{L^2}.\end{aligned}$$

It is then straightforward to check that (59) and (60) hold. This concludes the proof of Proposition 4.3.

Acknowledgments

The authors would like to thank Julien Vovelle for his remarks on an early version of this manuscript. The work of C.-E. B. is partially supported by the following projects operated by the French National Research Agency: ADA (ANR-19-CE40-0019-02) and BORDS (ANR-16-CE40-0027-01).

References

[AF19] Nathalie Ayi and Erwan Faou. Analysis of an asymptotic preserving scheme for stochastic linear kinetic equations in the diffusion limit. *SIAM/ASA J. Uncertain. Quantif.*, 7(2):760–785, 2019.

[BD99] François Bouchut and Laurent Desvillettes. Averaging lemmas without time Fourier transform and application to discretized kinetic equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 129(1):19–36, 1999.

[BHRR21] Charles-Edouard Bréhier, Hélène Hivert, and Shmuel Rakotonirina-Ricquebourg. Asymptotic preserving schemes for stochastic kinetic equations in the diffusion approximation regime. in preparation, 2021.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BLP79] Alain Bensoussan, Jacques-L. Lions, and George C. Papanicolaou. Boundary layers and homogenization of transport processes. *Publ. Res. Inst. Math. Sci.*, 15(1):53–157, 1979.

[Bou04] Nicolas Bourbaki. *Integration. II. Chapters 7–9*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2004. Translated from the 1963 and 1969 French originals by Sterling K. Berberian.

[Bré12] Charles-Edouard Bréhier. Strong and weak orders in averaging for SPDEs. *Stochastic Processes and their Applications*, 122(7):2553–2593, 2012.

[Bré20] Charles-Edouard Bréhier. Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component. *Stochastic Process. Appli.*, 130(6):3325–3368, 2020.

[BRR20] Charles-Edouard Bréhier and Shmuel Rakotonirina-Ricquebourg. On asymptotic preserving schemes for a class of stochastic differential equations in averaging and diffusion approximation regimes. arXiv preprint arXiv:2011.02341, 2020.

[Cer09] Sandra Cerrai. A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *The Annals of Applied Probability*, 19(3):899–948, 2009.

[CF09] Sandra Cerrai and Mark Freidlin. Averaging principle for a class of stochastic reaction-diffusion equations. *Probability theory and related fields*, 144(1-2):137–177, 2009.

[CX20] Sandra Cerrai and Guangyu Xi. A smoluchowski-kramers approximation for an infinite dimensional system with state-dependent damping. arXiv preprint arXiv:2011.14236, 2020.

[dBG12] Anne de Bouard and Maxime Gazeau. A diffusion approximation theorem for a nonlinear PDE with application to random birefringent optical fibers. *Ann. Appl. Probab.*, 22(6):2460–2504, 2012.

[DGP00] Pierre Degond, Thierry Goudon, and Frédéric Poupaud. Diffusion limit for nonhomogeneous and non-micro-reversible processes. *Indiana Univ. Math. J.*, 49(3):1175–1198, 2000.

[DRV20] Arnaud Debussche, Angelo Rosello, and Julien Vovelle. Diffusion-approximation for a kinetic spray-like system with random forcing. arXiv preprint arXiv:2005.09374, 2020.

[DV12] Arnaud Debussche and Julien Vovelle. Diffusion limit for a stochastic kinetic problem. *Commun. Pure Appl. Anal.*, 11(6):2305–2326, 2012.

[DV20] Arnaud Debussche and Julien Vovelle. Diffusion-approximation in stochastically forced kinetic equations. *Tunisian Journal of Mathematics*, 3(1):1–53, 2020.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[Eva89] Lawrence C. Evans. The perturbed test function method for viscosity solutions of nonlinear PDE. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 111(3-4):359–375, 1989.

[FGPSl07] Jean-Pierre Fouque, Josselin Garnier, George C. Papanicolaou, and Knut Sø lna. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.

[GP92] François Golse and Frédéric Poupaud. Limite fluide des équations de Boltzmann des semi-conducteurs pour une statistique de Fermi-Dirac. *Asymptotic Anal.*, 6(2):135–160, 1992.

[GR09] Thierry Goudon and Mathias Rousset. Stochastic acceleration in an inhomogeneous time random force field. *Applied Mathematics Research Express*, 2009(1):1–46, 2009.

[Jak86] Adam Jakubowski. On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(3):263–285, 1986.

[Jin99] Shi Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.

[Jin12] Shi Jin. Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: a review. *Riv. Math. Univ. Parma (N.S.)*, 3(2):177–216, 2012.

[JS03] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[Kha68] Rafail Z. Khasminskii. On the principle of averaging the Itô's stochastic differential equations. *Kibernetika 4*, pages 260–279, 1968.

[Kus84] Harold J. Kushner. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*, volume 6 of *MIT Press Series in Signal Processing, Optimization, and Control*. MIT Press, Cambridge, MA, 1984.

[LK74] Edward W. Larsen and Joseph B. Keller. Asymptotic solution of neutron transport problems for small mean free paths. *J. Mathematical Phys.*, 15:75–81, 1974.

[LT97] Pierre-Louis Lions and Giuseppe Toscani. Diffusive limit for finite velocity Boltzmann kinetic models. *Rev. Mat. Iberoamericana*, 13(3):473–513, 1997.

[PP03] Étienne Pardoux and Andrey L. Piatnitski. Homogenization of a nonlinear random parabolic partial differential equation. *Stochastic Process. Appl.*, 104(1):1–27, 2003.

[PS08] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

[PSV77] George C. Papanicolaou, Daniel W. Stroock, and S. R. Srinivasa Varadhan. Martingale approach to some limit theorems. *Duke Univ. Math. Ser., Vol. III*, pages ii+120, 1977.

[PV03] Frédéric Poupaud and Alexis Vasseur. Classical and quantum transport in random media. *Journal de mathématiques pures et appliquées*, 82(6):711–748, 2003.

[RR20] Shmuel Rakotonirina-Ricquebourg. Diffusion limit for a stochastic kinetic problem with unbounded driving process. arXiv preprint arXiv:2009.10406, 2020.

[RXY20] Michael Röckner, Longjie Xie, and Li Yang. Asymptotic behavior of multiscale stochastic partial differential equations. arXiv preprint arXiv:2010.14897, 2020.

[Sim87] Jacques Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[XY21] Xie and Li Yang. Diffusion approximation for multi-scale stochastic reaction-diffusion equations. arXiv preprint arXiv:2101.03917, 2021.