Mathematical Institute
University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*

Particle systems and stochastic PDEs

Analysis of models from hydrodynamics and neuroscience

Andrea Clini
The Queen's College
University of Oxford

Supervised by
Prof. José A. Carrillo & Prof. Benjamin J. Fehrman

Trinity 2024

Abstract

This work investigates the interplay between partial differential equations (PDEs) and interacting particle systems through the study of two classes of models. The first class is the so called *generalized Dean–Kawasaki equation*, a stochastic PDE coming from diverse areas like fluctuating hydrodynamics, mean-field theory and stochastic geometry. The second class consists of a system of stochastic differential equations (SDEs) describing a network of interacting neurons which covers many models from computational neuroscience, in particular in the study of *grid cells*. For the first model we focus on the PDE side, which is where we gave our contribution, and we briefly discuss the connection with particle systems and related relevant results. For the second we concentrate instead on the particle side, where we contributed the most, and again mention important results at the PDE level. For both models we identify directions of current and future investigation. The work is organized so that the two studies mirror each other. This highlights the interplay between particle systems and PDEs and illustrates a range of techniques used in this area.

Acknowledgments

Many people have played a crucial role during the years of my PhD, and they deserve a huge "thank you". I will tell them in private. Here, I just want to say something to my advisors, simply because they deserve for everybody to know how amazing they have been.

José and Ben, thank you, just "thank you for everything". You have been the best advisors ever. I hope you understand how much I mean it.

I have two regrets: not showing my true affection, simply because I wanted to show respect and not cross boundaries, and letting you down every now and then. The help I have received from you, in every respect, is just inconceivable. I hope one day I can repay you.

José, you have been like a father to me. The kind of father you love and brag about. Thank you for welcoming me into your family.

Ben, you have been the older brother every PhD student – actually, everyone in general – would kill to have. Thank you.

Just two more people. Antonio, ti devo moltissimo. Sei stato amico e fratello maggiore, sempre a disposizione, una guida. Grazie.

László, brodi, you know. Can I close the window?

Statement of Originality

I declare that the contents of this thesis are, to the best of my knowledge, original and my own work, except where otherwise indicated, cited, or commonly known, nor has any part of this thesis been submitted for a degree at another university.

The work in Chapter 2 has been published in [Cli23b]. The work in Chapter 3 is joint work with Benjamin Fehrman and is a preprint [CF23a] submitted for publication. The work in Chapter 5 is joint work with José A. Carrillo and Susanne Solem, and has been published in [CCS23]. The work in Chapter 6 has been published in [Cli23a].

Contents

Introduction

Partial Differential Equations (PDEs) and Interacting Particle Systems are fundamental mathematical modellings of real world phenomena. They allow to characterize the behavior and the intrinsic randomness of the physical systems considered with an accuracy appropriate for practical applications. The mathematics used mainly draws from analysis and probability. Applications cover several sciences: statistical mechanics, fluid dynamics, biology, neuroscience, population dynamics, math finance, statistics and machine learning.

Almost every physical phenomenon can be modelled via PDEs or particle systems, often both. Particle systems consider one or more equations for each microscopic constituent of the phenomenon described, all coupled together. PDEs provide instead a macroscopic average description involving only few continuum equations. In this sense, PDEs should be regarded as the continuum average analogue of particle systems. It is indeed often the case that, as the number of 'microscopic constituents' becomes very large, the average behavior of these particles is described by suitable PDEs. To account for the randomness of the system considered, stochastic versions of both modellings include suitable noise terms into the equations.

Understanding the interplay between these two descriptions and the effect of randomness yields a deeper comprehension of the physical phenomenon, with immediate applications, and great advances in the mathematics. Features of interest typically include the aforementioned convergence of the microscopic particle system toward an average PDE description, the occurrence and quantification of small fluctuations or even rare events drastically deviating from this average and resulting in catastrophes, the long time behavior of the models, the numerical aspects and simulations.

The aim of this thesis is to address some of these questions and investigate the interaction between particle systems and PDEs through the study of two specific models. The first model is the so called *generalized Dean–Kawasaki equation*

$$\partial_t \rho = \Delta\phi(\rho) - \nabla \cdot \nu(\rho) - \nabla \cdot (\rho V * \rho) - \sqrt{\epsilon}\,\nabla \cdot \left(\sigma(\rho) \circ \dot{\xi}\right) \quad \text{in } Q \times (0,T), \tag{1}$$

a class of stochastic PDEs coming from diverse areas like fluctuating hydrodynamics, mean-field theory, biology and stochastic geometry. Here Q is a suitable space domain, ξ a space-time noise, σ a possibly irregular noise coefficient, ν and V local and nonlocal interactions respectively, and ϕ a possibly degenerate diffusion.

The second model is a system of Stochastic Differential Equations (SDEs) – a particle system – describing a *spatially extended* network of interacting neurons

$$\begin{aligned} &du_{ik}(t) = b(x_i, t, u_{ik}, f_{N,M})\,dt + \sigma(x_i, t, u_{ik}, f_{N,M})\,dW_{ik}(t) - d\ell_{ik}(t), \\ &\text{for } i = 1 \dots N,\ k = i \dots M \ \text{ and } \ f_{N,M}(t) = \frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M} \delta_{(x_j, u_{jm}(t))} \in \mathcal{P}\left(Q \times \mathbb{R}\right). \end{aligned} \tag{2}$$

Here u_{ik} denotes the activity level (i.e. essentially the firing rate) of the k^{th} neuron in the cluster at location $x_i \in Q$ in the brain cortex Q, b and σ are suitable drift and diffusion terms possibly depending on the cortex location x_i and the empirical measure $f_{M,N}$ of the network (i.e. on the behavior of the other neurons), W_{ik} are Brownian motions and ℓ_{ik} are terms enforcing reflecting boundary conditions to keep activity levels non-negative. This SDE system covers several models commonly used in computational neuroscience and it arises from the study of *grid cells*, a specific type of neuron discovered in 2005 and worth the 2014 Nobel prize in Medicine.

For the first model we focus on the PDE side, which is where we gave our contribution through the works [Cli23b, CF23a], and we briefly discuss the connection with particle systems and related relevant results. We also identify directions of current and future investigation. For the second we concentrate instead on the particle system side, where we contributed with the works [CCS23, Cli23a], and again mention important results on the continuum counterpart. Similarly, we discuss further directions of research. The work is organized so that the two studies mirror each other. We hope that this highlights the interplay between particle systems and PDEs, and showcases a range of techniques used in this area.

Structure of the thesis

The work is organized as follows. In Chapter 1 we first discuss the state of the art on the Dean–Kawasaki equation (1) and then present the main contributions of the thesis. Chapter 2 contains the work from [Cli23b] and study the wellposedness and further properties of rough path approximations to the Dean–Kawasaki equation. Chapter 3 contains the work from [CF23a] and analyzes the small noise fluctuations of space-correlated versions of the Dean–Kawasaki equation.

In Chapter 4 we similarly discuss the current theory on mathematical models in neuroscience akin to (2) and then present the main contributions of the thesis. Chapter 5 contains the work from [CCS23]: it establishes the wellposedness of the neuron network (2) and the associated Fokker–Planck and McKean–Vlasov equations, and it rigorously proves the passage to the mean field limit. Chapter 6 contains the work from [Cli23a] and analyzes the fluctuations of the network in the thermodynamic limit.

Part I

The generalized Dean–Kawasaki equation

Chapter 1

Introduction

1.1 State of the art and open questions

The first part of this thesis is devoted to the study of the so called *generalized Dean–Kawasaki* equation. Fokker–Planck type equations are evolution equations for the expected value of macroscopic observables of particle systems (e.g. density or velocity or other properties of particles) and stochastic variants of these equations aim to describe the random observable itself.

In this context, 'generalized Dean–Kawasaki equation' is an umbrella term for a class of conservative stochastic PDEs like

$$\partial_t \rho^n = \Delta\phi(\rho^n) - \nabla \cdot (\nu(\rho^n) + \rho^n \, V * \rho^n) + \frac{1}{\sqrt{n}} \nabla \cdot \left(\sigma(\rho^n) \, \dot{\xi} \right), \tag{1.1.1}$$

where n is to be regarded as the number of particles. It allows for degenerate diffusion ϕ, local ν and nonlocal, possibly singular $V * \rho^n$ particle interaction, and noise term with σ irregular (e.g. $\sigma(\rho) = \sqrt{\rho}$ in the prototypical Dean–Kawasaki case) and ξ a cylindrical Wiener process.

In this generality, it describes, in principle exactly, the evolution of the density field of a wide variety of particle systems. Variants of this equation and systems can account for other relevant observables such as momentum density or multiple species.

Informal derivation from particle systems.

The original form of the equation is

$$\partial_t \rho(x,t) = \nabla \cdot \left(\rho(x,t) \nabla \frac{\delta F[\rho]}{\delta \rho(x,t)} \right) + \nabla \cdot \left(\sqrt{\rho(x,t)} \dot{\xi}(x,t) \right)$$

where $F[\rho]$ is the free energy functional, and it is due independently to Dean [Dea96] and Kawasaki [Kaw98].

Kawasaki's argument proceeds from Dynamical Density Functional Theory (cf. the survey [tVLW20]) as sketched below. Consider the microscopic description of a system of n interacting particles. Particles' position and momentum are denoted by $\{x_i\}$ and $\{p_i\}$. The Liouville equation governs the evolution of the phase space distribution function $\rho_n(\{x_i\},\{p_i\},t)$:

$$\partial_t \rho_n + \{\rho_n, H\} = 0,$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket and H is the Hamiltonian of the system.

Momentum variables are then integrated out to obtain a description in terms of positions only

$$\rho_n(\{x_i\},t) = \int \rho_n(\{x_i\},\{p_i\},t)\, d\{p_i\}.$$

This is justified because the momentum relaxation time is much shorter than the position relaxation time in dense fluids.

Within the framework of DDFT, we introduce a free energy functional $F[\rho]$ that depends on the density field $\rho(x,t)$. The free energy functional typically includes an ideal gas term and a term accounting for particle interactions. The time evolution of the density field is governed by a continuity equation with velocity field expressed through the functional derivative of the free energy:

$$\partial_t \rho(x,t) = \nabla \cdot \left(\rho(x,t) \nabla \frac{\delta F[\rho]}{\delta \rho(x,t)} \right).$$

To account for thermal fluctuations, a stochastic term $\dot{\xi}(x,t)$ representing Gaussian white noise is introduced, leading the following SPDE for the density field

$$\partial_t \rho(x,t) = \nabla \cdot \left(\rho(x,t) \nabla \frac{\delta F[\rho]}{\delta \rho(x,t)} \right) + \nabla \cdot \left(\sqrt{\rho(x,t)}\dot{\xi}(x,t) \right).$$

Dean's derivation is instead based on Itô formula for the evolution of the density and a smart replacement of the martingale term with a different process formally exhibiting the same probabilistic behavior.

Specifically, consider a particle system in the overdamped regime obeying Langevin dynamics

$$dX_i(t) = \sqrt{2}\, dB_i(t) + \frac{1}{n} \sum_{j=1}^{n} V\left(X_i(t) - X_j(t)\right) dt, \tag{1.1.2}$$

for an interaction kernel V and independent Brownian motions B_i. Applying Itô formula to the empirical measure $\mu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$, one obtains the following equation in the distributional sense

$$\partial_t \mu_n = \Delta \mu_n - \nabla \cdot (\mu_n V * \mu_n) + \frac{\sqrt{2}}{n} \nabla \cdot \left(\sum_{i=1}^{n} \delta_{X_i} \dot{B}_i \right).$$

To obtain a closed equation for μ_n, Dean replaces the martingale term with the following one which formally exhibits the same same quadratic variation

$$\partial_t \mu_n = \Delta \mu_n - \nabla \cdot (\mu_n V * \mu_n) + \sqrt{\frac{2}{n}} \nabla \cdot \left(\sqrt{\mu_n}\, \xi \right), \tag{1.1.3}$$

where ξ is now a space time white noise.

As already mentioned, analogous informal derivations apply to other situations like inertial models [CSZ19b, CSZ19a]. In fact, the Dean–Kawasaki formalism originates an effective framework capable of including further details on the modelling. For example, generalized Dean–Kawasaki equations accounting for the effect of reflecting boundaries are considered by Bressloff in [Bre23] and the effect of stochastic resetting of particle positions is considered in [Bre24]. Dean–Kawasaki equations for multiple interacting species, including the possibility of species switching driven by Poisson noise, are studied in [SM24]. The case of singular interactions (e.g. Keller–Segel like), together with other features, is considered in [Cha10]. In Chapter 4 we will discuss the possibility of applying this formalism to the neuron network studied.

Wellposedness of the equation.

Already in its original form, the wellposedness of the Dean–Kawasaki equation is highly debated and several negative results are known. Indeed, in the prototypical form (1.1.3), the noise term $\nabla \cdot \left(\sqrt{\rho}\,\dot{\xi}\right)$ requires applying a singular nonlinearity to a solution which is in principle only distribution-valued. Even in dimension 1, the equation is super-critical in the frameworks of regularity structures [Hai14] and of paracontrolled calculus [GIP15], developed to handle singular stochastic PDEs, and other theories are needed to make sense of the equation.

A natural weak formulation of the equation is to consider the associated martingale problem. Namely, a measure valued process ρ_t, with initial mass one, is considered a solution to the equation

$$\partial_t \rho = \frac{\alpha}{2}\Delta\rho + \nabla \cdot \left(\rho \nabla \frac{\delta F}{\delta \rho}[\rho]\right) + \nabla \cdot (\sqrt{\rho}\,\dot{\xi}) \tag{1.1.4}$$

if, for each smooth test function φ, the process

$$M_t(\varphi) := \langle \varphi, \rho_t \rangle - \int_0^t \left[\frac{\alpha}{2} \langle \Delta\varphi, \rho_s \rangle + \left\langle \nabla\varphi \cdot \nabla \frac{\delta F}{\delta \rho_s}\left[\rho_s\right], \rho_s \right\rangle \right] ds$$

is a martingale with quadratic variation

$$[M_{\cdot}(\varphi)]_t = \int_0^t \left\langle \rho_s, |\nabla\varphi|^2 \right\rangle \mathrm{d}s. \tag{1.1.5}$$

The series of works [KLvR19b, KLvR19a, KM23] shows that, for smooth potentials F, solutions exist (and are unique) only when $\alpha \in \mathbb{N}$ and in that case they are precisely given by the empirical measure μ_n associated to the particle system (1.1.2) undergoing Langevin dynamics.

Beyond the above notion of martingale solution, at the current state of the theory some coloring of the noise seems necessary to make sense of the equation. Even in this case, the well-posedness of space-correlated noise versions of the Dean–Kawasaki equation has been a long standing open problem. The main difficulties in applying a classical concept of weak solution are the nonlinear noise coefficient, possibly only 1/2-Hölder, and singular terms which are not even known to be locally integrable.

The seminal paper [FG21b] proves the well-posedness of equation (1.1.1), with Stratonovich noise, in the weak sense of *stochastic kinetic solutions* (cf. Defintion 3.2.19) for any porous media diffusion $\phi(\rho) = \rho^m$, $m \in (0, \infty)$ and rough noise coefficient including the prototypical case $\sigma(\rho) = \sqrt{\rho}$. The result is then extended to the case of nonlocal interactions, possibly singular, up to the newtonian kernel excluded, in [WWZ24]. In particular, stochastic kinetic solutions satisfy the following pathwise contration property:

$$\|\rho_1(\cdot, t) - \rho_2(\cdot, t)\|_{L^1_x} \leqslant \|\rho_1(\cdot, 0) - \rho_2(\cdot, 0)\|_{L^1_x} \quad \forall\, t \geqslant 0, \quad \text{almost surely.} \tag{1.1.6}$$

Exploiting this property, in [FGG22] they show that (1.1.1) generates of a random dynamical system.

This notion of solution addresses the aforementioned issues by passing to the *kinetic formulation* of the equation [CP03]: an equation in the original space and time variables and in a new additional *velocity* variable, corresponding to the magnitude of the solution. Then, a renormalization away from zero and infinity is introduced: solutions are required to satisfy the PDE only after cutting out small and large values in order to enforce the local integrability and further regularity of the nonlinear terms. In fact, when the equation coefficients are smooth enough, this renormalization is not even needed and stochastic kinetic solutions satisfy the equation in the usual weak sense.

Open question. In the case of very singular interactions, such as the newtonian potential in the parabolic-elliptic Keller–Segel case

$$\partial_t \rho = \Delta\phi(\rho) - \chi\, \nabla \cdot (\rho \nabla K * \rho) + \frac{1}{\sqrt{n}} \nabla \cdot \big(\sigma(\rho)\, \dot{\xi}\,\big), \quad -\Delta K = \delta, \tag{1.1.7}$$

wellposedness, even locally in time, is completely open.

In this context, it is also worth mentioning the works [AvR10, vRS09], where nontrivial (i.e. not empirical measures) martingale solutions (1.1.4)-(1.1.5) are obtained in some specific cases with singular potential.

We finally mention that, for the aforementioned inertial model also tracking momentum density and again with regularized noise, a different well-posedness theory is developed via the semigroup approach in [CSZ21], where solutions are shown to exists with arbitrarily high probability.

Rigorous connection with particle systems and small noise analysis.

Beyond the previous informal derivations, the connection with particle systems can be made rigorous both for the case of full white noise, necessarily with numerical discretizations removing the illposedness issues, and for colored noise. In both cases, the equation indeed captures the fluctuations of the particle system asymptotically in the limit of large number n of particles.

In the case of white noise, numerical discretization cures the singularities associated with the stochastic term and the equation can be treated as if it admitted classical solutions. The seminal works by Cornalba et al. [CFIR23, CF21] show that the discretized equation is indeed capable of capturing mean and fluctuations of the associated particle system, with an error dominated by machine precision. Similar results, but considering the martingale formulation of the equation (1.1.4)-(1.1.5), are obtained by Perkowski et al. [DKP22].

curacy of the Dean–Kawasaki approach has been analyzed with numerical experiments for several models: for example, in the context of fluid dynamics [DFVE14], social dynamics [DCKD22, HDCD$^+$21] and random matrix theory [TLDS23]. Other numerical experiments include [JDI22, DCG$^+$18, DOL$^+$16].

Colored-noise versions of the equation are not artificial either and indeed can be analytically shown to provide continuous approximations to suitable particle systems.

The study of the small noise $n \to \infty$ behaviour of (1.1.1) has been initiated in [FG23] and extended to nonlocal interactions in [WR23]. In the $n \to \infty$ limit, stochastic kinetic solutions converge to the solution of the corresponding noiseless equation

$$\partial_t \bar{\rho} = \Delta\phi(\bar{\rho}) - \nabla \cdot (\nu(\bar{\rho}) + \bar{\rho}\, V * \bar{\rho}) \,. \tag{1.1.8}$$

Furthermore they satisfy a large deviation principle in $L^1_t L^1_x$ with rate function

$$I_{\rho_0}(\rho) = \inf \left\{ \|g\|^2_{L^2_{t,x}} \, : \, \partial_t \rho = \Delta\phi(\rho) - \nabla \cdot (\nu(\rho) + \rho V * \rho) + \nabla \cdot (\sigma(\rho) g) \, , \, \rho(\cdot, 0) = \rho_0 \right\} . \tag{1.1.9}$$

The non-equilibrium fluctuations are analyzed in [DFG20, CF23a]. The rescaled fluctuations $v^n := \sqrt{n}\,(\rho_n - \bar{\rho})$ converge to the solution of the Langevin equation obtained by linearizing (1.1.1) around the zero noise limit

$$\partial_t v = \Delta\big(\dot{\phi}(\bar{\rho})\, v\big) - \nabla \cdot (\dot{\nu}(\bar{\rho})\, v + vV * \bar{\rho} + \bar{\rho} V * v) + \nabla \cdot \Big(\sigma(\bar{\rho})\, \dot{\xi}\Big) \,. \tag{1.1.10}$$

The connection with particle systems is then made rigorous by comparing these results with the corresponding statements for particles. In general, in the thermodynamic limit the empirical measure μ_n of particle systems can be shown to converge to the solution $\bar{\rho}$ of a diffusion equation like (1.1.8) and, in turn, the rescaled fluctuations $\sqrt{n}\,(\mu_n - \bar{\rho})$ can be shown to converge to the solutions v of the associated Langevin SPDE (1.1.10), for suitable coefficients ϕ, V, ν, σ. Finally, the particle system can be shown to satisfy a large deviation principle with a rate function I that can often be identified with the expression (1.1.9).

For example, results of this kind are available for the zero-range process [FPV88, BKL95, FG23], for the simple exclusion process in the symmetric [Rav92, QRV99] and asymmetric case [JM18], and for interacting systems like (1.1.2) in [WZZ23, Seo17, CG22, HHMT24].

Combining the results for the SPDE (1.1.1) and the particle system in the $n \to \infty$ limit gives the expansion

$$\mu_n = \rho_n \, dx + \mathrm{o}\left(\frac{1}{\sqrt{n}}\right), \tag{1.1.11}$$

which furnishes a continuum approximation to the particle system correct to order $\frac{1}{\sqrt{n}}$ and exhibiting the same rare events because both μ_n and ρ_n obey the same LDP (1.1.9).

Finally, we point out that one could also consider the simpler expansion

$$\mu^n = \bar{\rho} \, \mathrm{d}x + \frac{1}{\sqrt{n}} v + \mathrm{o}\left(\frac{1}{\sqrt{n}}\right). \tag{1.1.12}$$

However $\bar{\rho}^n := \bar{\rho} \, \mathrm{d}x + \frac{1}{\sqrt{n}} v$ satisfies an LDP with rate function different from (1.1.9), given by

$$\begin{aligned} \bar{I}_{\rho_0}(\rho) = \inf \Big\{ \|g\|^2_{L^2_{t,x}} \, : \partial_t (\rho - \bar{\rho}) = \Delta\big(\dot{\phi}(\bar{\rho})(\rho - \bar{\rho})\big) - \nabla \cdot (\dot{\nu}(\bar{\rho}) \, (\rho - \bar{\rho})) \\ - \nabla \cdot ((\rho - \bar{\rho}) \, V * \bar{\rho} + \bar{\rho} V * (\rho - \bar{\rho})) + \nabla \cdot (\sigma(\bar{\rho}) \, g) \, , \quad \rho(\cdot, 0) = \rho_0 \Big\}, \end{aligned} \tag{1.1.13}$$

as follows from Schilder's theorem and a formal application of the contraction principle, and hence it would predict incorrect rare events for the particle system.

In conclusion, we mention the recent result [GWZ24] which analyzes the higher order corrections to (1.1.1) in the small noise limit $n \to \infty$ and establishes the expansion

$$\rho_n = \sum_{k=0}^{K} \rho^{(k)} \, n^{-k/2} + o\left(n^{-K/2}\right), \tag{1.1.14}$$

where $\rho^{(k)}$ are solutions to a recursive set of Langevin SPDEs.

Open question. Similarly, on the particle system side, there have been recent advances [HCR23, CS21, ACR21] on higher order corrections to the empirical measure μ_n, which yield an expansion similar to (1.1.14). It would be interesting to see whether the Dean–Kawasaki approximation holds beyond the first order and is capable of capturing these higher fluctuations.

The connection with other regularized Dean–Kawasaki type equations is also available. For example, the above mentioned regularized *inertial* Dean–Kawasaki equation is shown in [CSZ19b, CSZ19a, CSZ21] to be solved by a mollified version of the empirical measure of the particle system, where essentially particles are assigned a finite size $\epsilon > 0$. Similarly, particle approximations to the regularized martingale problem (1.1.4) in $1D$ are considered in [Din22].

Long-time behavior and stationary states.

Under suitable assumptions on the interaction potential V, one might expect ergodicity and convergence to equilibrium for the particle system (1.1.2). At the SPDE level (1.1.1), in the case of colored noise, the long time behavior has been analyzed in [FGG22]. Building upon the stochastic kinetic formulation [FG21b] and the contraction property (1.1.6), in the absence of nonlocal interactions they prove existence and uniqueness of an invariant measure and strong mixing.

Even more interesting is the effect of noise in the presence of *singular* nonlocal interactions possibly leading to finite time blow-up. The underlying idea of *regularization by noise* [Fla11] is that the addition of a stochastic term might regularize solutions and prevent, or at least delay, the blow-up.

Considering for example the prototypical Keller–Segel case (1.1.7), in [FGL21] they show that a stochastic transport term $\nabla \cdot (\rho\, \eta)$ with η a divergence-free space-time noise, formally arising from common noise at the particle level [CF16], guarantees existence up to an arbitrary time with arbitrarily high probability upon having strong enough noise. A similar stochastic term is considered in [MT23], extending the results for space independent η from [MST22], again coming from common noise at the particle level. They establish blow-up criteria in terms of the interaction strength χ and the initial solution mass akin to the deterministic case, which seem indeed to suggest a delayed blow-up effect by noise.

Open question. The addition of a rougher noise term like the prototypical Dean–Kawasaki case $\nabla \cdot \left(\sqrt{\rho}\, \xi\right)$, arising instead from the independent noise terms sensed by each particle in (1.1.2), is formally discussed in [Cha10]. The effect of such a term on the blow-up scenario is still unclear. Preliminary estimates however suggest it should not prevent blow-up beyond the classical Keller–Segel regime.

More generally, deterministic aggregation-diffusion equations, that is equation (1.1.1) without noise, have received lots of attention [CCY19, GC24]. Even when the diffusion dominates and global existence is known, the long time behavior is often unclear, where multiple steady states or metastable behavior can occur. The effect of noise on these phenomena is even more obscure.

Open question. We also mention the connection with the aforementioned large deviation principles in the small noise limit. For example, in the setting of scalar conservation laws, the seminal work [Mar08] shows that vanishing noise can furnish a selection principle for solutions. It is unclear whether the same phenomenon can occur for second order equations: that is esentially an LDP identifying the physically meaningful solutions as those where the LDP rate function is finite. Similarly, one might expect to relate the values of the rate function to the metastability of the corresponding states.

Further properties of solutions and numerics.

Further properties of the solutions have also been considered. For example, although the continuity of the noise-to-solution map is not to be expected in general, being typically false even at the SDE level, suitable regularizations of the equation might recover this feature. The rough path approximation to the Dean–Kawasaki equation discussed in [Cli23b] does indeed satisfy this property and this has immediate consequences on the large deviations and support properties of solutions (cf. Section 1.2 and Chapter 2).

Another important property expected from the physical interpretation of the equation is preservation of mass and positivity of solutions. In the colored case, both properties are shown for stochastic kinetic solutions in [FG21b].

Open question. When starting with a strictly positive initial data, strict positivity is proved asymptotically with high probability in the small noise limit in [CF23a], but it is still not clear whether the noise can actually drag the solution down to zero in some region.

Positivity is also crucial for numerical schemes. In the full white noise case, necessarily at the discretized level where the equation singularity is resolved, this is however not satisfied in general. Numerical discretizations of the equation are considered in the already mentioned works [CFIR23, CF21]. For the inertial Dean–Kawasaki model, in the regularized case where particles are assigned a finite size, a numerical scheme is discussed in [CS23]. In both cases, even for strictly positive initial data, the discrete solutions considered might get negative after long times and suitable stopping times need to be introduced.

Multilevel Monte Carlo methods to reduce the computational cost of simulations have been proposed and analyzed in [CF23b]. With the availability of large deviation principles, another promising variance reduction technique is that of importance sampling. In the infinite dimensional case, the approach has been recently considered for linear SPDEs with additive noise or multiplicative noise with Lipschitz coefficient [SS17, BT17, GSS23]. In particular, in the multiplicative noise case with nonlinear singular coefficients, the approach is yet to be explored.

1.2 The contribution of this thesis

1.2.1 Chapter 2: wellposedness of rough path approximations

The main contribution of Chapter 2, containing the work [Cli23b], is the wellposedness of stochastic porous media equations with Dirichlet boundary conditions of the form

$$\begin{cases} \partial_t u = \Delta(|u|^{m-1}u) + \nabla \cdot (A(x,u) \circ dz_t) & \text{in } Q \times (0,\infty), \\ u = 0 & \text{on } \partial Q \times (0,\infty), \\ u = u_0 & \text{on } Q \times \{0\}, \end{cases} \tag{1.2.1}$$

for a smooth bounded domain in $Q \subset \mathbb{R}^d$, for any diffusion exponent $m \in (0,\infty)$, initial data $u_0 \in L^2(Q)$, and an n-dimensional α-Hölder geometric rough path z, covering in particular the case of a Brownian motion. The matrix valued nonlinearity

$$A(x,u) : Q \times \mathbb{R} \to \mathbb{R}^{d\times n}$$

is assumed to be regular, with regularity dictated by the regularity of the rough path z.

In the context of this thesis, equation (1.2.1) is regarded as an approximation to generalized Dean–Kawasaki equations

$$\partial_t u = \Delta(|u|^{m-1}u) + \nabla \cdot (\sigma(x,u) \circ \dot{\xi}), \tag{1.2.2}$$

for a d-dimensional space-time white noise ξ and a possibly irregular nonlinearity $\sigma(x,u)$, for example $\sigma(x,u) = \sqrt{u}$. Indeed, for each $m \in \mathbb{N}$ we can take $n = md$, truncate the cylindrical expansion $\xi = \sum_{i=1}^{\infty} e_i(x) B_t^i$, for independent d-dimensional Brownian motions B_t^i and an orthonormal basis $e_i(x)$ of $L^2(Q)$, and set

$$A(x,u) = \sigma_m(x,u)\Big[e_1(x)\mathrm{I}_d \mid e_2(x)\mathrm{I}_d \mid \cdots \mid e_m(x)\mathrm{I}_d\Big], \quad z_t = (B_t^1, \ldots, B_t^m),$$

where I_d denotes the identity matrix of dimension d and σ_m is a smooth approximation to σ. Then $A(x,u) \circ dz_t = \sigma_m(x,u) \sum_{i=1}^{m} e_i(x) \circ d\dot{B}_t^i$ is indeed converging to $\sigma(x,u) \circ \dot{\xi}$ as $m \to \infty$.

The wellposedness of (1.2.1) is obtained in the sense of *pathwise kinetic solutions* (cf. Definition 2.2.4), a notion based on the kinetic formulation of the equation [CP03] and akin to the aforementioned *stochastic kinetic solutions* (cf. Definition 3.2.19), but which addresses the problem in a completely pathwise analytic way thanks to the rough path approach. The specific methods are discussed in Section 2.1.

The work extends to Dirichlet boundary conditions the periodic case presented in [FG19]. The imposition of vanishing boundary conditions, corresponding to absorbing boundaries at the particle level, is a challenging problem. In the stochastic case, even in the simple linear one-dimensional case $\partial_t u = \partial_{xx} u + (\partial_x u) \circ \dot{z}$, it is nontrivial [Kry03] and in fact not always possible to enforce them in a classical sense. In our specific case, we obtain H^1-regularity for powers u^m of the solution and impose their boundary trace to vanish.

The main results are summarized in the following theorems.

Theorem 1.2.1 (Theorem 2.1.2 and 2.1.3)**.** For any $m \in (0, \infty)$, for an α-Hölder geometric rough path z and a suitably smooth noise coefficient A, for any $u_0 \in L^2(Q)$ there exists a unique *pathwise kinetic solution* to the Cauchy problem (1.2.1). If $u_0 \geqslant 0$, then $u(t) \geqslant 0$ for every $t \geqslant 0$. Finally, pathwise kinetic solutions satisfy the following pathwise contraction property

$$\|u^1 - u^2\|_{L^\infty([0,\infty);L^1(Q))} \leqslant \|u_0^1 - u_0^2\|_{L^1(Q)}. \tag{1.2.3}$$

Despite the subsequent result [FG21b], which still requires smoothing of the white noise, but not of the rough noise coefficient σ, to obtain wellposedness of (1.2.2), the interest in the approximation (1.2.1) is nonetheless retained as exemplified by the following results.

Theorem 1.2.2 (Theorem 2.1.4)**.** Under the above assumptions, equation (1.2.1) defines a continuous random dynamical system on $L^2_+(Q)$. Namely, we have almost surely

$$u(u_0, s, t, z.(\omega)) = u(u_0, 0, t-s, z_{\cdot+s}(\omega)) \quad \forall\, 0 \leqslant s \leqslant t \;\; \forall\, u_0 \in L^2_+(Q),$$

where $u(u_0, s, t, z.(\omega))$ denotes the solution at time t started at time s with initial data $u_0 \in L^2(Q)$ and driving signal $z.(\omega)$. Moreover, the contraction principle (1.2.3) implies that the dynamical system is continuous in $L^1(Q)$.

Theorem 1.2.3 (Theorem 2.1.5)**.** Under the above assumption, let $\{z^k\}_{k\in\mathbb{N}}$ be a sequence of geometric rough paths such that

$$\lim_{k\to\infty} d_\alpha(z^k, z) = 0$$

in the α-Hölder metric d_α (cf. Section 2.5) and let $\{u^k\}_{k\in\mathbb{N}}$ and u denote the corresponding pathwise kinetic solutions with the same initial data u_0. Then we have

$$\lim_{k\to\infty} \|u^k - u\|_{L^1([0,T];L^1(Q))} = 0.$$

In the case z_t arises from the sample paths of a Brownian motion B_t enhanced with its iterated Stratonovich integrals, the continuity of the noise-to-solution map has immediate consequences on the large deviations and support properties of solutions. For example, because of Schilder's Theorem (in its version for rough paths [FV10, Theorem 13.42]) and the contraction principle, the solutions of (1.2.1) satisfy a large deviation principle in $L^1_t L^1_x$ with rate function

$$I_{u_0}(u) = \inf\left\{ \frac{1}{2}\int_0^T |\dot g(t)|^2 dt \;:\; \partial_t u = \Delta u^m + \nabla\cdot(A(x,u)\dot g(t)),\; u(\cdot,0) = u_0 \right\}.$$

Similarly, one has that the random solution u is arbitrarily close with positive probability to any solution u_g of the Cauchy problem (1.2.1) where the rough path z induced by the Brownian motion is replaced by a generic smooth path g (cf. Remark 2.1.6).

1.2.2 Chapter 3: fluctuations in the small noise limit

The main contribution of Chapter 3, containing the work [CF23a], is to analyze the small noise $n \to \infty$ fluctuations of the colored-noise Dean–Kawasaki equation (1.1.1) as sketched in the discussion (1.1.8)-(1.1.11).

The main results are summarized in the following theorem and proposition.

Theorem (Theorem 3.3.3 and 3.3.10)**.** Under suitable assumptions on the coefficients and along a suitable scaling regime where the colored noise $\xi^n \to \xi$ converges to space-time white noise, the rescaled small noise fluctuations $\sqrt{n}(\rho^n - \bar\rho)$ of the Dean–Kawasaki equation (1.1.1) around the zero noise limit (1.1.8) converge to the solution v of the Langevin SPDE (1.1.10). Namely,

$$v^n := \sqrt{n}(\rho^n - \bar\rho) \to v \quad \text{in } L^\tau([0,T]; H^{-\beta}(\mathbb{T}^d)) \text{ in probability},$$

for any $T > 0$, for $\tau = 2$ or $\tau = \infty$, for any $\beta > \frac{d}{2}$ or $\beta > 1 + \frac{d}{2}$ respectively, with an explicit rate of convergence which depends on β, the equation coefficients and the space regularity of the noise sequence $\xi^n \to \xi$.

A by-product of the techniques employed is the following strict positivity result with high probability in the small noise limit.

Proposition 1.2.4 (Proposition 3.3.8)**.** Under the same assumptions, stochastic kinetic solutions to (1.1.1) with strictly positive initial data ρ_0 satisfy

$$\forall \delta > 0 \quad \mathbb{P}\left(\rho^n > \inf \rho_0 - \delta\right) \geqslant 1 - C_n \quad \text{with} \quad C_n \xrightarrow{n \to \infty} 0, \tag{1.2.4}$$

for a sequence C_n depending on δ, $\inf \rho_0$, the equation coefficients and the space regularity of the noise ξ^n. In particular, solutions ρ^n are bounded away from zero with increasing probability as $n \to \infty$.

We refer to Section 3.1.1 for a thorough discussion of the strategy adopted. We simply mention that the main difficulties are due to the nonlinear coefficients, possibly singular or degenerate, and the renormalization prescription for stochastic kinetic solutions (cf. Definition 3.2.19), which make it difficult to exploit equation (1.1.1) explicitly in the arguments. To address this issue, we argue with regularized versions of the equation with coefficients smoothed only near their singularities, which admit more classical solutions. Thanks to the asymptotic strict positivity result (1.2.4) and the pathwise uniqueness (1.1.6), smoothed and true solutions coincide with increasing probability in the small noise limit and we can pass our arguments to the true equation.

Chapter 2

Wellposedness of rough path approximations

2.1 Introduction and main results

In this chapter we consider stochastic porous media and fast diffusion equations with nonlinear, conservative noise of the form

$$\begin{cases} \partial_t u = \Delta(|u|^{m-1}u) + \nabla \cdot (A(x,u) \circ dz_t) & \text{in } Q \times (0,\infty), \\ u = 0 & \text{on } \partial Q \times (0,\infty), \\ u = u_0 & \text{on } Q \times \{0\}, \end{cases} \tag{2.1.1}$$

for a diffusion exponent $m \in (0,\infty)$, initial data $u_0 \in L^2(Q)$, and an n-dimensional, α-Hölder continuous geometric rough path z, which in particular covers the case where z is an n-dimensional Brownian motion. The domain Q is a smooth bounded domain in $\mathbb{R}^d$. The matrix valued nonlinearity

$$A(x,\xi) : Q \times \mathbb{R} \to \mathbb{R}^{d\times n}$$

is assumed to be regular, with regularity dictated by the regularity of the rough path z.

We establish path-by-path existence and uniqueness of (2.1.1) for positive initial data $u_0 \in L^2_+(Q)$ in the full regime $m \in (0,\infty)$ (cf. Theorems 2.1.2 and 2.1.3), in terms of the central notion of *pathwise kinetic solution* (cf. Definition 2.2.4). Moreover, we show the solution map to (2.1.1) generates a random dynamical system associated to the equation (cf. Theorem 2.1.4) and is a continuous function of the noise (cf. Theorem 2.1.5). This has consequences on the support properties of the solution to (2.1.1) around the solution of its deterministic version (2.1.7) (cf. Remark 2.1.6). Finally, the existence of solutions to (2.1.1) is in fact proved for any signed data $u_0 \in L^2(Q)$ and in the full regime $m \in (0,\infty)$ (cf. Theorem 2.1.3). In particular, when $m = 1$ or $m \in (2,\infty)$, all the aforementioned results extend to general signed data (cf. Theorem 2.1.7).

Stochastic porous media equations of this type arise, for example, as a continuum limit of mean-field stochastic differential equations with common noise [CG19, KX99], with notable relation to the theory of mean field games [LL07, LL06]; or in the graph formulation of the stochastic mean curvature and curve shortening flow [ESR12, DLN01, SY04]. Moreover, equation (2.1.1) can be regarded as an approximate model for the fluctuating hydrodynamics of the zero-range particle process around its hydrodynamic limit [DSZ15, FG23, FPV88], as an approximation to the Dean-Kawasaki equation arising in fluid dynamics [MT99, Dea96, Kaw98, CSZ19b, FG21b], and as a model for thin films of Newtonian fluids with negligible surface tension [GMR06].

Our approach to (2.1.1) is crucially based on the work [FG19], where Fehrman and Gess proved analogous results for this equation posed on the d-dimensional torus with periodic boundary conditions. In turn, this was motivated by the works of Lions, Perthame and Souganidis [LPS13, LPS14], and Gess and Souganidis [GS14b, GS14a, GS16] on stochastic conservation laws and simpler versions of (2.1.1). The method is essentially based on passing to the equation's kinetic formulation, introduced by Chen and Perthame [CP03, Per02], for which the noise enters as a linear transport, and then on analytic techniques and rough path analysis.

The main difficulty with respect to the periodic case [FG19] is the handling of the Dirichlet boundary conditions. The imposition of boundary conditions is a challenging problem. In the stochastic case, even in the simple linear one-dimensional case $\partial_t u = \partial_{xx} u + (\partial_x u) \circ \dot{z}$, it is nontrivial [Kry03] and in fact not always possible to enforce them in a classical sense. The conditions can be recast in a weak sense, but even in the deterministic case, this is notoriously difficult for nonlinear equations.

In our case, depending on the diffusion exponent $m \in (0, \infty)$, we obtain H^1-regularity for the power $|u|^{m-1}u$ and impose the zero boundary conditions by requiring the trace of $|u|^{m-1}u$ to vanish. The behaviour of the noise at the boundary is controlled in terms of the Dirichlet conditions imposed on the solution. This requires a sharp ad-hoc treatment depending on the particular exponent $m \in (0, \infty)$ and the observation of crucial cancellations among the boundary error terms.

The well-posedness of (2.1.1) has been an open question for long time, even in the *probabilistic* (i.e. non patwhise) setting and even in the case z_t is given by a Brownian motion B_t. Generalized stochastic porous media equations of the form

$$d\, u = \Delta\phi(u)\, dt + \sigma(x, u)\, dB_t$$

have attracted considerable interest and their well-posedness has been obtained for several classes of nonlinearities ϕ, noise coefficients $\sigma(x, u)$ and boundary conditions. We refer to the monographs [BPR16, LR15], and to [FG21a, GH16, DHV16, BR17, BR14, BVW15] for recent contributions. To some extent, the case of a *linear* gradient noise, that is $A(x,u) = h(x)u$ in (2.1.1), could be treated with similar methods (cf. [DG17, RM16, Töl18]). However, the nonlinear structure of the gradient noise in (2.1.1) requires entirely different techniques.

The aforementioned works [LPS13, LPS14, GS14b, GS14a, GS16] started developing a kinetic approach to tackle scalar conservation laws and simplified versions of (2.1.1). The work of Fehrman and Gess [FG19] was the first to prove similar results for (2.1.1) in the pathwise context, at the cost of high regularity assumptions on the noise coefficient needed to overcome the roughness of the signal (cf. assumption (2.1.3) below). Dareiotis and Gess [DG20] then managed to lower the regularity assumptions up to $A \in C_b^3(\mathbb{T}^d \times \mathbb{R})$ and proved well-posedness of (2.1.1) with periodic boundary conditions, in the probabilistic sense. Furthermore, Fehrman and Gess [FG21b] proved probabilistic well-posedness of (2.1.1) with periodic boundary conditions, when z_t is a Brownian motion and, for example, when $A(x,u) = f(x)\sqrt{u}$ with $f \in C_b^2(\mathbb{T}^d)$.

Despite requiring higher regularity, the interest in pathwise results is nonetheless retained and twofold. First, it is well-known that solutions to stochastic differential equations do not depend continuously on the driving noise (see for instance [Lyo91]), and even more so if the noise coefficient is a nonlinear function of the solution itself. However, the continuity of the solution can be recovered by means of a finer rough path topology [Lyo98]. In the same fashion, Theorem 2.1.5 below establishes the continuous dependence in the noise of the solution map to (2.1.1).

Secondly, the pathwise nature of the existence and uniqueness Theorems 2.1.2 and 2.1.3 immediately implies the existence of a random dynamical system associated to (2.1.1), stated in Theorem 2.1.4. This is a notoriously difficult problem (see e.g. [Fla95, MZZ08, Ges12]). Even in the linear case $m = 1$, the existence of a random dynamical system for a nonlinear SPDE with nonlinear x-dependent noise could not be proved before [FG19]. Indeed, all the aforementioned works on (2.1.1) in the probabilistic setting could not obtain these two consequences.

Structure of the chapter

The material is organized as follows. We first give an overview of the methods and arguments employed in this chapter. In Section 2.1.1 we introduce our hypotheses and notations, and in Section 2.1.2 we present the main results. In Section 2.2 we bring forward the kinetic formulation of the equation; after analyzing the associated system of characteristics, we motivate and present the definition of pathwise kinetic solution. In Section 2.3 we prove the uniqueness of solutions to (2.1.1). Section 2.4 is devoted to the proof of existence of solutions to (2.1.1) and of their continuous dependence on the driving noise. In Section 2.5 we present some stability results from the theory of rough paths and we gather some estimates needed throughout the chapter.

Overview of the methods

The methods of this work build upon the kinetic approach put forward in [LPS13, LPS14, GS14b, GS14a, GS16, FG19, FG21a] for stochastic conservation laws and variants of the stochastic porous media equation. The aim of Section 2.2 is to explain the pathwise and kinetic approach to the (2.1.1), and to derive the central notion of *pathwise kinetic solution*. First, we pass to the kinetic formulation of the PDE [CP03]. This is an equation in $d+2$ variables: the time and space variables t and x, and an additional *velocity* variable ξ, which corresponds to the magnitude of the solution. The interest in such formulation of (2.1.1) is that now the noise enters the equation as a linear transport. The transport is well-defined for rough driving signals, when the underlying system is interpreted as a rough differential equation.

Following the strategy used in [LPS13, LPS14, GS14b, GS14a, GS16, FG19, FG21a], and previously put forward in the theory of stochastic viscosity solutions [LS98b, LS98a, LS00a, LS00b, LS02], we test the kinetic equation against a restricted class of test functions only, precisely test functions transported by an underlying conservative system of stochastic characteristics. When testing against these functions, the terms involving the noise automatically cancel out. Such test functions are simply obtained with the method of characteristics: namely by flowing an arbitrary initial data back in time along the characteristic curves of the rough differential equations prescribing the transport. Informally, this is of course the same as flowing the kinetic solution forward in time along the characteristics. This restricted class of test functions is nonetheless large enough to provide a comprehensive characterization of the solutions to (2.1.1), or better to its kinetic version, which is sufficient to prove its well-posedness. The precise notion of pathwise kinetic solution is given in Definition 2.2.4.

In comparison to [LPS13] and [FG21a], owing to the x-dependent gradient structure of the noise, the characteristics' equations cannot be solved explicitly. Therefore, the solutions need to be controlled with the rough path estimates from Section 2.5. Moreover, as time passes, the characteristics also move in space (cf. [FG19, LPS14]). With respect to [FG19, LPS14], the introduction of Dirichlet boundary conditions further complicates the picture: as time evolves the space characteristics might indeed escape the domain Q, and the resulting test functions flowed along these characteristics would not be compactly supported within Q if too much time has passed. To overcome this difficulty, we make further assumptions on the noise coefficient $A(x,u)$ defining the system of rough characteristics (2.2.21) so as to directly prevent the characteristics from escaping the domain Q (cf. (2.1.5) and (2.2.17)). These assumptions simplify the handling of the boundary conditions and are fundamental in our proofs. They are justified since we have in mind equation (2.1.1) as an approximate model for a noise term $\nabla \cdot (\sigma(x,u)\eta)$, with η a space-time white noise. Indeed, it is always possible to find a cylindrical expansion of η, which we then truncate at some order, such that the aforementioned conditions are satisfied. See Remark 2.1.1 for details.

Section 2.3 is devoted to the proof of uniqueness. The formal proof of uniqueness follows the same outline presented in [CP03] in the deterministic case. However, to justify the formal computation, care must be taken to avoid the product of δ-distributions. This is achieved with a regularization in the space and velocity variables. Moreover, to cancel the noise from the equation, we are allowed to use transported test functions only. Additional error terms arise due to the transport of test functions along characteristics, which are handled using a time-splitting argument that relies crucially on the conservative structure of the equation. In this setting, the interaction between the x-dependent characteristics and the nonlinear diffusion term further complicates the arguments: the error terms generated by the regularization procedure need to be controlled with sharp estimates (cf. Proposition 2.3.6 and 2.3.9) for singular moments of the parabolic defect measure (cf. Definition 2.2.4), especially in the case of small diffusion exponents $m \in (0,1) \cup (1,2]$.

The imposition of Dirichlet boundary conditions makes the analysis even more difficult: to keep everything compactly supported within the domain Q it is necessary to introduce a cutoff function. This, combined with the displacement of the space characteristics, generates new boundary layer error terms which are controlled with the zero-trace conditions imposed on the solution u. Concretely, the decay at the boundary of u is quantified by means of the H^1-regularity of $|u|^{m-1}u$ prescribed by the nonlinear diffusion. In turn, this forces us to choose the cutoff function $\phi = \phi(m)$ as a function of the particular exponent $m \in (0,\infty)$. Finally, besides the above estimates, the handling of both the interior and boundary error terms relies on some fundamental cancellations among them. Indeed, the gradient structure of the noise implies that the characteristics preserve the underlying Lebesgue measure, and this is crucial both for the above mentioned time-splitting argument and to observe these cancellations.

Finally, in Section 2.4 we prove the existence of pathwise kinetic solutions and their continuous dependence on the noise in the rough path topology. This is obtained by proving stable estimates for the solutions of smoothed PDEs approximating (2.1.1), and then using weak convergence and compactness arguments to prove these solutions converge to a solution of (2.1.1). The aforementioned estimates hold for the limiting solution as well. Furthermore, they hold uniformly for solutions of (2.1.1) corresponding to rough signals $\{z_t^n\}_{n\in\mathbb{N}}$ close to each other in rough path topology. Then, repeating the same weak convergence and compactness arguments used to show existence, and combining them with the uniqueness of solutions, we prove that the solution map to (2.1.1) depends continuously on the noise.

Concretely, we first prove stable estimates for singular moments of the parabolic defect measure of the approximating PDEs (2.2.2) totally akin to those used in the uniqueness proof (cf. Proposition 2.4.3 and Proposition 2.3.6). Then we exploit these estimates to show that the corresponding kinetic solutions are uniformly bounded in suitable fractional Sobolev spaces $W^{\ell,1}_{x,\xi}$ for the space and velocity variables (cf. Proposition 2.4.5). Moreover, they have enough regularity in time (cf. Proposition 2.4.4) to invoke the Aubin-Lions-Simons Lemma [Sim86] and obtain strong convergence.

In this regard, it is worth mentioning that the weak convergence arguments developed in [GS14b] for scalar conservation laws with x-dependent noise do not apply in the parabolic case (2.1.1). Indeed, the corresponding class of pathwise entropy solutions is not closed under weak convergence, since the second order structure of the equation cannot ensure the weak limiting object actually is a solution. Therefore, in [GS16] the same authors introduced a strong convergence method based on uniform BV-estimates to tackle the parabolic case. However, these arguments are probably restricted to the x-independent case as a uniform BV-estimate for solutions to (2.1.1) does not seem available.

2.1.1 Hypotheses and notations

The domain Q is a smooth bounded domain in $\mathbb{R}^d$ with $d \geqslant 1$. The diffusion exponent is $m \in (0,\infty)$, and for the signed power we shall use the shorthand

$$u^{[m]} := |u|^{m-1}u.$$

We shall denote by $L^2_+(Q)$ the closed subspace of L^2-integrable functions which are a.e. nonnegative. The noise is a geometric rough path: for $n \geqslant 1$ and a Hölder exponent $\alpha \in (0,1)$, for each $T > 0$

$$z_t = (z^1_t, \dots, z^n_t) \in C^{0,\alpha}\left([0,T]; G^{\lfloor \frac{1}{\alpha} \rfloor}(\mathbb{R}^n)\right), \tag{2.1.2}$$

where $C^{0,\alpha}\big([0,T]; G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^n)\big)$ is the space of n-dimensional α-Hölder geometric rough paths on $[0,T]$. We denote by d_α the α-Hölder metric defined on this space. See Section 2.5 for references on rough path theory.

As regards the noise coefficient $A(x,\xi)$, we assume it is smooth with bounded derivatives. Precisely, for some $\gamma > \frac{1}{\alpha}$ we assume

$$D_x A(x,\xi) \in C^{\gamma+2}_b(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^{d\times d\times n}) \quad \text{and} \quad \partial_\xi A(x,\xi) \in C^{\gamma+2}_b(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^{d\times n}). \tag{2.1.3}$$

This regularity is necessary in order to obtain the rough path estimates of Proposition 2.5.1. In particular, as the regularity of the noise decreases, more regularity is required for the coefficients.

We also need to impose conditions on A so as to control the direction of the characteristics.

To start with, we assume the nonlinearity $A(x,\xi)$ satisfies

$$\nabla_x \cdot A(x,0) = \sum_{i=1}^{d} \partial_{x_i} A_{i,\cdot}(x,0) = 0 \quad \text{for each } x \in \mathbb{R}^d. \tag{2.1.4}$$

This assumption guarantees that the underlying stochastic characteristics preserve the sign of the velocity variable. Even in the case of smooth driving signal, this condition is necessary to ensure that the evolution of (2.1.1) does not increase the mass of the initial condition.

Next, we impose two conditions to govern the behaviour of the space characteristics near the boundary. These assumptions are *crucial* in our arguments to impose Dirichlet boundary conditions and are justified by Remark 2.1.1 below. Namely, we impose that $A(x,\xi)$ satisfies

$$\partial_\xi A(x,\xi) = D_x \partial_\xi A(x,\xi) = 0 \quad \text{on } \partial Q \times \mathbb{R}. \tag{2.1.5}$$

As already mentioned in Section 2.1, the assumption $\partial_\xi A_{|\partial Q\times\mathbb{R}} = 0$ ensures that, as time evolves, the space characteristics never leave the domain Q. This condition is necessary to guarantee that the transport of a test function along the characteristics stays compactly supported in Q, and thus it is an admissible test function in the kinetic formulation of (2.1.1). The further hypothesis $D_x\partial_\xi A_{|\partial Q\times\mathbb{R}} = 0$ is slightly more technical and is needed to effectively exploit the condition $\partial_\xi A_{|\partial Q\times\mathbb{R}} = 0$ and further quantify the informal fact that the space characteristics move slower as they start closer to the boundary. Intuitively, it means that the strength of the transport term $\partial_\xi A(x,u)\nabla u \circ dz_t$ featuring in (2.1.1) decreases more than linearly in terms of the distance $\operatorname{dist}(x,\partial Q)$.

Remark 2.1.1. We stress that our main interest in (2.1.1) is to consider it as a space correlated approximation to the, possibly ill-posed, stochastic porous media equation

$$\partial_t u = \Delta u^{[m]} + \nabla \cdot (\sigma(x,u) \circ \eta),$$

where η is a d-dimensional space-time white noise and $\sigma(x,\xi)$ is a possibly nonsmooth nonlinearity with $\sigma(x,0) = 0$ for every $x \in Q$, for example $\sigma(x,\xi) = \sqrt{\xi}$. We observe that the assumptions (2.1.4)-(2.1.5) are perfectly compatible with this strategy. Indeed, a standard construction of the space-time white noise on $Q \times [0,\infty)$ is $\eta = \sum_{i=1}^{\infty} \rho_i(x) dB_t^i$, where $\{B_t^i\}_{i\in\mathbb{N}}$ are independent d-dimensional Brownian motions and $\{\rho_i(x)\}_{i\in\mathbb{N}}$ is an orthonormal basis of $L^2(Q)$. Now consider a basis such that $\rho_i \in C^\infty(\bar{Q})$ and such that $\rho_i = D_x\rho_i = 0$ on ∂Q for each $i \in \mathbb{N}$ (for example, consider a spectral basis of $L^2(Q)$, contained in $H_0^2(Q)$, for the symmetric compact operator Δ^{-2}). For each $m \in \mathbb{N}$, we can take $n = md$ and set

$$A(x,\xi) = \sigma_m(x,\xi)\Big[\rho_1(x)\mathrm{I}_d \mid \rho_2(x)\mathrm{I}_d \mid \cdots \mid \rho_m(x)\mathrm{I}_d\Big], \quad z_t = (B_t^1,\ldots,B_t^m),$$

where I_d denote the identity matrix of dimension d, and σ_m is a smooth approximation to σ with $\sigma_m(x,0) = 0$ for every $x \in Q$. Then $A(x,u)\circ dz_t = \sigma_m(x,u)\sum_{i=1}^m \rho_i(x)\circ dB_t^i$ is indeed converging to $\sigma(x,u)\circ\eta$ as $m \to \infty$, and elementary computations show that $A(x,\xi)$ satisfies the assumptions (2.1.4)-(2.1.5).

2.1.2 Main results

We now present the main results of the chapter. The precise notion of *pathwise kinetic solution* is given in Definition 2.2.4 below. Our first result, proved in Section 2.3, is a contraction principle for pathwise kinetic solutions with nonnegative initial data, which in particular implies their uniqueness.

Theorem 2.1.2. Let $m \in (0, \infty)$ and let $u_0^1, u_0^2 \in L^2_+(Q)$. Under the assumptions (2.1.3)-(2.1.5), pathwise kinetic solutions u^1 and u^2 of (2.1.1) with initial data u_0^1 and u_0^2 satisfy

$$\|u^1 - u^2\|_{L^\infty([0,\infty);L^1(Q))} \leqslant \|u_0^1 - u_0^2\|_{L^1(Q)}. \tag{2.1.6}$$

In particular, pathwise kinetic solutions with nonnegative initial data are unique.

As already mentioned in Section 2.1, our uniqueness result heavily relies on sharp a priori estimates needed to tackle the nonlinear diffusion and take care of the zero boundary conditions. In turn, these are coupled with other stable estimates for smoothed equations approximating (2.1.1) both in space and time. Then, using weak convergence and compactness arguments, in Section 2.4 we prove the existence of pathwise kinetic solutions with a limit procedure.

Theorem 2.1.3. Let $m \in (0, \infty)$ and let $u_0 \in L^2(Q)$. Under the assumptions (2.1.3)-(2.1.5), there exists a pathwise kinetic solution of (2.1.1) with initial data u_0. Furthermore, if $u_0 \in L^2_+(Q)$, the corresponding solution stays nonnegative, that is

$$u(x, t) \geqslant 0 \quad \text{almost everywhere in } Q \times [0, \infty).$$

The pathwise nature of Theorem 2.1.2 and 2.1.3 immediately implies the existence of a *random dynamical systems* associated to (2.1.1). Precisely, suppose that our driving noise $[0, \infty) \ni t \mapsto z_t = z_t(\omega)$ arises from the sample paths of a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in such a way that $z_t(\omega)$ is indeed an α-Hölder geometric rough path for almost every $\omega \in \Omega$. Then we have the following result.

Theorem 2.1.4. Assume the hypotheses (2.1.3)-(2.1.5) and let $m \in (0, \infty)$. When interpreted in the sense of pathwise kinetic solutions, equation (2.1.1) defines a random dynamical system on $L^2_+(Q)$. Let $u(u_0, s, t, z.(\omega))$ denote the solution at time t to (2.1.1) started at time s with initial data $u_0 \in L^2_+(Q)$ and driving signal $z.(\omega)$. Then, for almost every $\omega \in (\Omega, \mathcal{F}, \mathbb{P})$, we have

$$u(u_0, s, t, z.(\omega)) = u(u_0, 0, t - s, z_{\cdot+s}(\omega)) \quad \forall\, 0 \leqslant s \leqslant t \;\forall\, u_0 \in L^2_+(Q).$$

Moreover, the contraction principle (2.1.6) implies this dynamical system is continuous when considered with values in $L^1(Q)$.

Next we present a result stating that the solution map is continuous with respect to the driving noise. Indeed, all the aforementioned estimates needed for the existence theorem are pathwise estimates and depend on the driving noise z_t cosindered as a geometric rough path; in particular, they are uniform for geometric rough paths close to each other in the α-Hölder metric d_α. In Section 2.4, using the same compactness arguments as for the existence proof and exploiting the uniqueness of solutions to (2.1.1), we prove the following continuity result. Unfortunately, this method does not yield an explicit estimate quantifying the convergence of solutions in terms of the convergence of the driving signals.

Theorem 2.1.5. Assume the hypotheses (2.1.3)-(2.1.5). Let $m \in (0, \infty)$ and $u_0 \in L^2_+(Q)$. For any $T > 0$, let $\{z^k\}_{k \in \mathbb{N}}$ and z be a sequence of n-dimensional α-Hölder geometric rough paths on $[0, T]$ such that

$$\lim_{k \to \infty} d_\alpha(z^k, z) = 0.$$

Let $\{u^k\}_{k \in \mathbb{N}}$ and u denote the pathwise kinetic solutions to (2.1.1) on $[0, T]$ with initial data u_0 and driving signals $\{z^k\}_{k \in \mathbb{N}}$ and z respectively. Then we have

$$\lim_{k \to \infty} \|u^k - u\|_{L^1([0,T];L^1(Q))} = 0.$$

Remark 2.1.6. Theorem 2.1.5 has immediate consequences on the support properties of the solutions to (2.1.1). For simplicity, let us suppose that our driving noise $t \mapsto z_t$ arises from the sample paths of an n-dimensional Brownian motion $B_t(\omega)$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in the sense that the mapping $t \mapsto \mathbb{B}_t^{\text{Strat}}(\omega)$ defines $\mathbb{P}$ almost surely an α-Hölder geometric rough path $z_t(\omega)$, where $\mathbb{B}_t^{\text{Strat}}$ denotes the Brownian motion enhanced with its iterated Stratonovich integrals (see e.g. [FH14]). For $\mathbb{P}$ a.e. $\omega \in \Omega$, let $u(\omega)$ denote the pathwise kinetic solution of (2.1.1) with initial data $u_0 \in L^2_+(Q)$ and driving signal $t \mapsto \mathbb{B}_t^{\text{Strat}}(\omega)$. Given any smooth path $g : [0, \infty) \to \mathbb{R}^n$, let $\bar{u}_g$ denote the pathwise kinetic solution, with the same initial data u_0, of the deterministic porous media equation with convective term

$$\partial_t \bar{u}_g = \Delta \bar{u}_g^{[m]} + \nabla \cdot \big(A(x, \bar{u}_g)\, \dot{g}\big), \tag{2.1.7}$$

that is of equation (2.1.1) driven by the smooth path $z_t := g(t)$ (cf. 2.5). In this setting, Theorem 2.1.5 implies that the probability of u being arbitrarily close to the deterministic solution $\bar{u}_g$ in L^1-norm is always nonzero. Indeed, properties of the Stratonovich enhanced Brownian motion ensure that $\mathbb{P}\left(d_\alpha\left(\mathbb{B}_\cdot^{\text{Strat}}, g\right) \leqslant \epsilon\right) > 0$ for any $T \geqslant 0$ and any $\epsilon > 0$ (cf. [FV10, Chapter 13]). Theorem 2.1.5 immediately implies that, for any $T \geqslant 0$,

$$\mathbb{P}\left(\|u - \bar{u}_g\|_{L^1([0,T];L^1(Q))} \leqslant \epsilon\right) > 0 \quad \forall \epsilon > 0.$$

That is to say, the support of the law of u in $L^1([0,T]; L^1(Q))$ contains the solution of (2.1.7) for every smooth path g. In fact, since almost every sample path $\mathbb{B}_\cdot(\omega)$ arises as d_α-limit of smooth paths, we actually have that the support of the law of u is the closure of $\{\bar{u}_g \mid g \text{ smooth path}\}$ in $L^1([0,T]; L^1(Q))$.

Finally, we notice that the methods of this chapter apply to general initial data in $L^2(Q)$ provided the diffusion exponent satisfies $m = 1$ or $m > 2$. Indeed, the nonnegativity of the solution, and thus of the initial data, is only required in the a-priori estimate presented in Proposition 2.3.9. In turn, this estimate is only needed to tackle the case of small diffusion exponents $m \in (0,1) \cup (1,2]$ (cf. Remark 2.3.2). As a consequence we get the following result.

Theorem 2.1.7. Let $m = 1$ or $m > 2$. Under the assumptions (2.1.3)-(2.1.5), for every $u_0 \in L^2(Q)$ there exists a unique pathwise kinetic solution of (2.1.1) and the analogous results of Theorem 2.1.2 and Theorems 2.1.5 and 2.1.4 hold.

2.2 Definition and motivation of pathwise kinetic solutions

The aim of this section is to understand equation (2.1.1), and motivate and present the notion of *pathwise kinetic solution* given in Definition 2.2.4. For this purpose, we shall first consider a uniformly elliptic regularization of (2.1.1) driven by smooth noise. The assumption (2.1.2) ensures that there exists a sequence of smooth paths $\{z^\epsilon : [0,\infty) \to \mathbb{R}^n\}_{\epsilon\in(0,1)}$ such that, for each $T > 0$,

$$\lim_{\epsilon\to 0} d_\alpha\left(z, z^\epsilon\right) = 0, \tag{2.2.1}$$

where d_α denotes the α-Hölder metric on the space $C^{0,\alpha}\left([0,T]; G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^n)\right)$ of geometric rough paths. In what follows, for $\epsilon \in (0,1)$, we will denote by $\dot{z}^\epsilon$ the derivative of the smooth path.

Furthermore, it is necessary to introduce an η-perturbation by the Laplacian, for $\eta \in (0,1)$, in order to remove the degeneracy of the porous media diffusion. Therefore, for each $\eta \in (0,1)$ and $\epsilon \in (0,1)$, we consider the equation

$$\begin{cases} \partial_t \mathrm{u}^{\eta,\epsilon} = \Delta(\mathrm{u}^{\eta,\epsilon})^{[m]} + \eta\Delta \mathrm{u}^{\eta,\epsilon} + \nabla \cdot (A(x,\mathrm{u}^{\eta,\epsilon})\dot{z}^\epsilon_t) & \text{in } Q \times (0,\infty), \\ \mathrm{u}^{\eta,\epsilon} = 0 & \text{on } \partial Q \times (0,\infty), \\ \mathrm{u}^{\eta,\epsilon} = u_0 & \text{on } Q \times \{0\}. \end{cases} \tag{2.2.2}$$

We shall derive a formulation of the equation that is well-defined for singular driving signals. Namely, we shall pass to the kinetic form of (2.2.2), where the noise enters as a linear transport, and then derive a formulation that is well-defined even after passing to the limit with respect to the regularization.

The following proposition establishes the well-posedness of (2.2.2) in the classical sense. The proof is a small modification of [FG19, Proposition A.1], and thus is omitted.

Proposition 2.2.1. For each $\eta \in (0,1)$, each $\epsilon \in (0,1)$, and each $u_0 \in L^2(Q)$, there exists a classical solution $\mathrm{u}^{\eta,\epsilon}$ of equation (2.2.2) such that

$$u \in H^1\left([0,T]; H^1_0(Q), H^{-1}(Q)\right), \quad \text{and} \quad u^{[m]}, u^{[m+1/2]} \in L^2\left([0,T]; H^1_0(Q)\right).$$

We now pass to the kinetic form of (2.2.2), complete details of the following derivation are given in [FG19, Appendix A]. This formulation is obtained by introducing the kinetic function $\bar{\chi} : \mathbb{R}^2 \to \{-1, 0, 1\}$ defined by

$$\bar{\chi}(v, \xi) := \begin{cases} 1 & \text{if } 0 < \xi < v, \\ -1 & \text{if } v < \xi < 0, \\ 0 & \text{else.} \end{cases}$$

We then define, for each $\eta \in (0,1)$ and $\epsilon \in (0,1)$, for $\mathrm{u}^{\eta,\epsilon}$ the solution of (2.2.2), the composition

$$\chi^{\eta,\epsilon}(x, \xi, t) := \bar{\chi}(\mathrm{u}^{\eta,\epsilon}(x,t), \xi). \tag{2.2.3}$$

Proposition 2.2.2 below states that the kinetic function $\chi^{\eta,\epsilon}$ is a distributional solution of the equation

$$\begin{aligned} \partial_t \chi^{\eta,\epsilon} = m|\xi|^{m-1} \Delta_x \chi^{\eta,\epsilon} + \eta \Delta_x \chi^{\eta,\epsilon} + (\partial_\xi A(x,\xi) \dot{z}_t^\epsilon) \cdot \nabla_x \chi^{\eta,\epsilon} \\ - (\nabla_x \cdot A(x,\xi) \dot{z}_t^\epsilon)\, \partial_\xi \chi^{\eta,\epsilon} + \partial_\xi (\mathrm{p}^{\eta,\epsilon} + \mathrm{q}^{\eta,\epsilon}) \end{aligned} \tag{2.2.4}$$

in $Q \times \mathbb{R} \times (0, \infty)$, with Dirichlet boundary conditions and initial data $\bar{\chi}(u_0(x), \xi)$. Here, the measure $\mathrm{p}^{\eta,\epsilon}$ is the *entropy defect measure*

$$\mathrm{p}^{\eta,\epsilon}(x, \xi, t) := \delta_0(\xi - \mathrm{u}^{\eta,\epsilon}(x,t))\, \eta\, |\nabla \mathrm{u}^{\eta,\epsilon}(x,t)|^2, \tag{2.2.5}$$

and the measure $\mathrm{q}^{\eta,\epsilon}$ is the *parabolic defect measure*

$$\mathrm{q}^{\eta,\epsilon}(x, \xi, t) := \delta_0(\xi - \mathrm{u}^{\eta,\epsilon}(x,t)) \frac{4m}{(m+1)^2} \left| \nabla (\mathrm{u}^{\eta,\epsilon})^{\left[\frac{m+1}{2}\right]}(x,t) \right|^2, \tag{2.2.6}$$

where δ_0 denotes the one-dimensional Dirac mass centered at the origin. The sense in which the kinetic function satisfies (2.2.4) is made precise by the following proposition. The proof is again a small modification of [FG19, Proposition A.2] and is omitted.

Proposition 2.2.2. For each $\eta \in (0,1)$, $\epsilon \in (0,1)$, and $u_0 \in L^2(Q)$, let $\mathrm{u}^{\eta,\epsilon}$ be the solution of (2.2.2) from Proposition 2.2.1. The associated kinetic function $\chi^{\eta,\epsilon}$ defined in (2.2.3) is a distributional solution of equation (2.2.4) in the sense that, for every $t_1 \leqslant t_2 \in [0, \infty)$ and for every $\psi \in C_c^\infty(Q \times \mathbb{R} \times [t_1, t_2])$, we have

$$\begin{aligned} &\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}(x,\xi,r)\psi(x,\xi,r)\, dx\, d\xi \Big|_{r=t_1}^{r=t_2} \\ &\quad = \int_{t_1}^{t_2}\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}\, \partial_t \psi\, dx\, d\xi\, dr \\ &\quad + \int_{t_1}^{t_2}\int_{Q\times\mathbb{R}} \left(m|\xi|^{m-1} + \eta\right) \chi^{\eta,\epsilon}\, \Delta_x \psi\, dx\, d\xi\, dr \\ &\quad - \int_{t_1}^{t_2}\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon} (\nabla_x \cdot ((\partial_\xi A(x,\xi) \dot{z}_t^\epsilon)\psi) - \partial_\xi ((\nabla_x \cdot A(x,\xi) \dot{z}_t^\epsilon)\psi))\, dx\, d\xi\, dr \\ &\quad - \int_{t_1}^{t_2}\int_{Q\times\mathbb{R}} (\mathrm{p}^{\eta,\epsilon}(x,\xi,r) + \mathrm{q}^{\eta,\epsilon}(x,\xi,r))\, \partial_\xi \psi\, dx\, d\xi\, dr. \end{aligned} \tag{2.2.7}$$

The purpose of this section is to remove the dependency of equation (2.2.7) on the derivative of the noise, so as to get a formulation of (2.1.1) which is well-defined even for rough driving signals. To achieve this, rather than testing equation (2.2.7) against arbitrary test functions ψ, we shall limit ourselves to only use the solutions of the following transport equations, for any $t_0 \in [0, t_1]$ and any initial data $\rho_0 \in C_c^\infty(Q \times \mathbb{R})$:

$$\begin{cases} \partial_t \rho = \partial_\xi A(x,\xi)\dot{z}_t^\epsilon \cdot \nabla_x \rho - (\nabla_x \cdot A(x,\xi))\dot{z}_t^\epsilon \, \partial_\xi \rho & \text{in } \mathbb{R}^d \times \mathbb{R} \times (t_0, \infty), \\ \rho = \rho_0 & \text{on } \mathbb{R}^d \times \mathbb{R} \times \{t_0\}. \end{cases} \tag{2.2.8}$$

Indeed, using such test functions $\rho(x, \xi, t)$ in (2.2.7), the first and third term on the right-hand side automatically cancel out.

Owing to assumption (2.1.3) on A and to the smoothness of the paths z^ϵ, equation (2.2.8) is a standard first-order PDE with smooth coefficients, and its solution is represented using the associated characteristics, which we now analyze. The forward characteristic $(X_{t_0,t}^{x_0,\xi_0,\epsilon}, \Xi_{t_0,t}^{x_0,\xi_0,\epsilon})$ associated to (2.2.8) beginning at time $t_0 \geqslant 0$ from $(x_0, \xi_0) \in \mathbb{R}^d \times \mathbb{R}$ is defined as the solution of the system

$$\begin{cases} \dot{X}_{t_0,t}^{x_0,\xi_0,\epsilon} = -\partial_\xi A(X_{t_0,t}^{x_0,\xi_0,\epsilon}, \Xi_{t_0,t}^{x_0,\xi_0,\epsilon})\dot{z}_t^\epsilon & \text{in } (t_0, \infty), \\ \dot{\Xi}_{t_0,t}^{x_0,\xi_0,\epsilon} = (\nabla_x \cdot A(X_{t_0,t}^{x_0,\xi_0,\epsilon}, \Xi_{t_0,t}^{x_0,\xi_0,\epsilon}))\dot{z}_t^\epsilon & \text{in } (t_0, \infty), \\ (X_{t_0,t_0}^{x_0,\xi_0,\epsilon}, \Xi_{t_0,t_0}^{x_0,\xi_0,\epsilon}) = (x_0, \xi_0). \end{cases} \tag{2.2.9}$$

The associated backward characteristic $(Y_{t_0,s}^{x_0,\xi_0,\epsilon}, \Pi_{t_0,s}^{x_0,\xi_0,\epsilon})$ beginning at time 0 from $(x_0, \xi_0) \in \mathbb{R}^d \times \mathbb{R}$ and terminating at t_0, is simply defined as the forward characteristic $(X_{t_0,t}^{x_0,\xi_0,\epsilon}, \Xi_{t_0,t}^{x_0,\xi_0,\epsilon})$ run backward in time. Namely, we define

$$Y_{t_0,s}^{x_0,\xi_0,\epsilon} := X_{t_0,t_0-s}^{x_0,\xi_0,\epsilon} \quad \text{and} \quad \Pi_{t_0,s}^{x_0,\xi_0,\epsilon} := \Xi_{t_0,t_0-s}^{x_0,\xi_0,\epsilon} \qquad \text{for } s \in [0, t_0]. \tag{2.2.10}$$

The forward and backward characteristics are mutually inverse, in the sense that, for each $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}$, for each $t_0 \geqslant 0$ and $t \geqslant t_0$, and for each $s_0 \geqslant 0$ and $s \in [0, s_0]$, we have the relation

$$\left(X_{t_0,t}^{Y_{t,t-t_0}^{x,\xi,\epsilon}, \Pi_{t,t-t_0}^{x,\xi,\epsilon}, \epsilon}, \Xi_{t_0,t}^{Y_{t,t-t_0}^{x,\xi,\epsilon}, \Pi_{t,t-t_0}^{x,\xi,\epsilon}, \epsilon} \right) = \left(Y_{s_0,s}^{X_{s_0-s,s}^{x,\xi,\epsilon}, \Xi_{s_0-s,s}^{x,\xi,\epsilon}, \epsilon}, \Pi_{s_0,s}^{X_{s_0-s,s}^{x,\xi,\epsilon}, \Xi_{s_0-s,s}^{x,\xi,\epsilon}, \epsilon} \right) = (x, \xi). \tag{2.2.11}$$

Furthermore, for each $t_0 \geqslant 0$, we define the reversed path

$$z_{t_0,t}^\epsilon := z_{t-t_0}^\epsilon \qquad \text{for } t \in [0, t_0].$$

It is then easy to check that the backward characteristic $(Y_{t_0,s}^{x_0,\xi_0,\epsilon}, \Pi_{t_0,s}^{x_0,\xi_0,\epsilon})$ coincides with the solution of the system

$$\begin{cases} \dot{Y}_{t_0,t}^{x_0,\xi_0,\epsilon} = -\partial_\xi A(Y_{t_0,t}^{x_0,\xi_0,\epsilon}, \Pi_{t_0,t}^{x_0,\xi_0,\epsilon})\dot{z}_{t_0,t}^\epsilon & \text{in } (0, t_0), \\ \dot{\Pi}_{t_0,t}^{x_0,\xi_0,\epsilon} = (\nabla_x \cdot A(Y_{t_0,t}^{x_0,\xi_0,\epsilon}, \Pi_{t_0,t}^{x_0,\xi_0,\epsilon}))\dot{z}_{t_0,t}^\epsilon & \text{in } (0, t_0), \\ (Y_{t_0,t_0}^{x_0,\xi_0,\epsilon}, \Pi_{t_0,t_0}^{x_0,\xi_0,\epsilon}) = (x_0, \xi_0). \end{cases} \tag{2.2.12}$$

The solution of (2.2.8) is the transport of the initial data along the backward characteristics (2.2.10). Precisely, for each $\rho_0 \in C_c^\infty(Q \times \mathbb{R})$, a direct computation proves that $\rho_0(Y^{x,\xi,\epsilon}_{t,t-t_0}, \Pi^{x,\xi,\epsilon}_{t,t-t_0})$ is indeed the solution of (2.2.8). For each $t_0 \geqslant 0$ and $\rho_0 \in C_c^\infty(Q \times \mathbb{R})$, we shall use the notation

$$\rho^\epsilon_{t_0,t}(x,\xi) := \rho_0(Y^{x,\xi,\epsilon}_{t,t-t_0}, \Pi^{x,\xi,\epsilon}_{t,t-t_0}) \tag{2.2.13}$$

for the solution of the transport equation (2.2.8).

Furthermore, as a consequence of (2.2.8), the characteristics preserve the Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}$. That is, for every $0 \leqslant t_0 < t_1$ and $0 \leqslant s_1 < s_0$, for every $\psi \in L^1(\mathbb{R}^d \times \mathbb{R})$,

$$\int_{R^d\times\mathbb{R}} \psi(x,\xi)\,dx\,d\xi = \int_{R^d\times\mathbb{R}} \psi(X^{x,\xi,\epsilon}_{t_0,t_1}, \Xi^{x,\xi,\epsilon}_{t_0,t_1})\,dx\,d\xi = \int_{R^d\times\mathbb{R}} \psi(Y^{x,\xi,\epsilon}_{s_0,s_1}, \Pi^{x,\xi,\epsilon}_{s_0,s_1})\,dx\,d\xi. \tag{2.2.14}$$

This property in turn comes from the conservative structure of the noise term in (2.2.2), and it is essential to the proof of uniqueness in Section 2.3.

Finally, we analyze the consequences for the characteristics of the assumptions (2.1.4) and (2.1.5) on the noise coefficient $A(x,\xi)$. It is immediate from the second line of (2.2.9) and (2.2.12) that the hypothesis $\nabla_x \cdot A(x,0) \equiv 0$ implies the characteristics preserve the sign of the velocity variable. That is, for each $(x,\xi) \in \mathbb{R}^d \times \mathbb{R}$, for each $t_0 \geqslant 0$ and $t \geqslant t_0$, and for each $s_0 \geqslant 0$ and $s \in [0,s_0]$, we have

$$\Xi^{x,\xi,\epsilon}_{t_0,t} = \Pi^{x,\xi,\epsilon}_{s_0,s} = 0 \text{ if and only if } \xi = 0, \text{ and } \operatorname{sgn}(\xi) = \operatorname{sgn}(\Xi^{x,\xi,\epsilon}_{t_0,t}) = \operatorname{sgn}(\Pi^{x,\xi,\epsilon}_{s_0,s}) \text{ if } \xi \neq 0. \tag{2.2.15}$$

As regards the assumption $\partial_\xi A(x,\xi)_{|\partial Q\times\mathbb{R}} \equiv 0$, the first line of (2.2.9) and (2.2.12) ensure that space characteristics starting from the boundary do not move. That is, for any $t_0 \geqslant 0$,

$$\text{if } (x,\xi) \in \partial Q \times \mathbb{R}, \text{ then } X^{x,\xi,\epsilon}_{t_0,t} = Y^{x,\xi,\epsilon}_{t_0,s} = x \text{ for all } t \geqslant 0 \text{ and all } s \in [0,t_0]. \tag{2.2.16}$$

The uniqueness of solutions then implies that the space characteristics cannot cross the boundary ∂Q and thus, when starting within Q, they never leave the domain. That is, for any $t_0 \geqslant 0$,

$$\text{if } (x,\xi) \in Q \times \mathbb{R}, \text{ then } X^{x,\xi,\epsilon}_{t_0,t} \in Q \text{ for all } t \geqslant 0, \text{ and } Y^{x,\xi,\epsilon}_{t_0,s} \in Q \text{ for all } s \in [0,t_0]. \tag{2.2.17}$$

In fact, owing to the smoothness hypothesis (2.1.3), more is true: the closer to ∂Q is the space initial data $x \in \mathbb{R}^d$, the slower the associated space characteristic moves. Rigorously speaking, for any $T > 0$ we have

$$\left|X^{x,\xi,\epsilon}_{t_0,t} - x\right| + \left|Y^{x,\xi,\epsilon}_{t_0,s} - x\right| \leqslant C\operatorname{dist}(x,\partial Q) \quad \forall (x,\xi) \in \mathbb{R}^d \times \mathbb{R} \quad \forall t_0, t \in [0,T] \quad \forall s \in [0,t_0], \tag{2.2.18}$$

for a constant $C = C(T,A,z)$ depending on the time T, the noise coefficient $A(x,\xi)$ defining the systems (2.2.9) and (2.2.12), and the rough signal z, but uniform for all the paths $\{z^\epsilon\}_{\epsilon\in(0,1)}$, which are close to the rough path z in the metric d_α. This and other estimates are proved in Appendix 2.5.

Finally, the assumption $\partial_\xi A(x,\xi)_{|\partial Q\times\mathbb{R}} \equiv 0$ has fundamental consequences on the transported functions (2.2.13). Indeed estimate (2.2.18) ensures that

$$\text{if } \rho_0 \in C_c(Q\times\mathbb{R}), \text{ then } \rho^\epsilon_{t_0,t}(x,\xi) = \rho_0(Y^{x,\xi,\epsilon}_{t,t-t_0}, \Pi^{x,\xi,\epsilon}_{t,t-t_0}) \in C_c(Q\times\mathbb{R}\times[t_0,T]) \;\forall T \geqslant t_0 \geqslant 0. \tag{2.2.19}$$

That is, the transport along characteristics of a functions which is compactly supported within $Q \times \mathbb{R}$ stays compactly supported, and thus, in particular, is an admissible test function in equation (2.2.7).

Now we go back to equation (2.2.7). The following corollary makes precise the idea of testing the equation against functions transported along the inverse characteristics, so as to get rid of the noise in the equation. The proof is an immediate consequence of the discussion above, in particular equation (2.2.8), the representation (2.2.13) and property (2.2.19), and Proposition 2.2.2.

Corollary 2.2.3. Let $\eta, \epsilon \in (0,1)$ and $u_0 \in L^2(Q)$. The kinetic function $\chi^{\eta,\epsilon}$ from Proposition 2.2.2 satisfies, for each $t_0 \leqslant t_1 \in [0,\infty)$ and $\rho_0 \in C_c^\infty(Q \times \mathbb{R})$, for the solution $\rho^\epsilon_{t_0,r}(x,\xi) := \rho(Y^{x,\xi,\epsilon}_{r,r-t_0}, \Pi^{x,\xi,\epsilon}_{r,r-t_0})$ of (2.2.8),

$$\begin{aligned}\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}(x,\xi,r)\rho^\epsilon_{t_0,r}(x,\xi)\,dx\,d\xi\Big|_{r=t_0}^{r=t_1} \\ = \int_{t_0}^{t_1}\int_{Q\times\mathbb{R}} \left(m|\xi|^{m-1}+\eta\right)\chi^{\eta,\epsilon}(x,\xi,r)\Delta_x\rho^\epsilon_{t_0,r}(x,\xi)\,dx\,d\xi\,dr \\ - \int_{t_0}^{t_1}\int_{Q\times\mathbb{R}} (\mathrm{p}^{\eta,\epsilon}(x,\xi,r)+\mathrm{q}^{\eta,\epsilon}(x,\xi,r))\,\partial_\xi\rho^\epsilon_{t_0,r}(x,\xi)\,dx\,d\xi\,dr.\end{aligned} \tag{2.2.20}$$

The essential observation is that, in the passage to the singular limit $\epsilon \to 0$, the system of characteristics (2.2.9) is well-posed for rough noise when interpreted as a rough differential equation. In view of the representation (2.2.13), this implies the well-posedness of equation (2.2.8) for rough signals as well. Precisely, for each $(x_0,\xi_0) \in \mathbb{R}^d \times \mathbb{R}$ and $t_0 \geqslant 0$, we define the rough forward characteristic $(X^{x_0,\xi_0}_{t_0,t}, \Xi^{x_0,\xi_0}_{t_0,t})$ as the solution of the rough differential equation

$$\begin{cases} dX^{x_0,\xi_0}_{t_0,t} = -\partial_\xi A(X^{x_0,\xi_0}_{t_0,t}, \Xi^{x_0,\xi_0}_{t_0,t}) \circ dz_t & \text{in } (t_0,\infty), \\ d\Xi^{x_0,\xi_0}_{t_0,t} = (\nabla_x \cdot A(X^{x_0,\xi_0}_{t_0,t}, \Xi^{x_0,\xi_0}_{t_0,t})) \circ dz_t & \text{in } (t_0,\infty), \\ (X^{x_0,\xi_0}_{t_0,t_0}, \Xi^{x_0,\xi_0}_{t_0,t_0}) = (x_0,\xi_0). \end{cases} \tag{2.2.21}$$

The continuity properties of the Itô–Lyons map, summarized in Proposition 2.5.1, and the smoothness assumptions (2.1.3) on $A(x,\xi)$ imply that, for each $T>0$,

$$\lim_{\epsilon\to 0}\left(X^{x,\xi,\epsilon}_{t_0,t}, \Xi^{x,\xi,\epsilon}_{t_0,t}\right) = \left(X^{x,\xi}_{t_0,t}, \Xi^{x,\xi}_{t_0,t}\right) \text{ uniformly for } t_0 \leqslant t \in [0,T] \text{ and } (x,\xi) \in \mathbb{R}^d \times \mathbb{R}.$$

In analogy to the smooth case, we now list the properties of the rough characteristics. As before, the rough inverse characteristic $(Y^{x_0,\xi_0}_{t_0,s}, \Pi^{x_0,\xi_0}_{t_0,s})$ is defined as the forward characteristic run backward in time. Namely, we define

$$Y^{x_0,\xi_0}_{t_0,s} := X^{x_0,\xi_0}_{t_0,t_0-s} \quad \text{and} \quad \Pi^{x_0,\xi_0}_{t_0,s} := \Xi^{x_0,\xi_0}_{t_0,t_0-s} \qquad \text{for } s \in [0,t_0].$$

As before, for each $t_0 \geqslant 0$, we define the reversed path $z_{t_0,t} := z_{t-t_0}$ for $t \in [0,t_0]$, and the backward rough characteristic $(Y^{x_0,\xi_0,\epsilon}_{t_0,s}, \Pi^{x_0,\xi_0,\epsilon}_{t_0,s})$ coincides with the solution of the rough differential equation

$$\begin{cases} dY^{x_0,\xi_0}_{t_0,t} = -\partial_\xi A(Y^{x_0,\xi_0}_{t_0,t}, \Pi^{x_0,\xi_0}_{t_0,t}) \circ dz_{t_0,t} & \text{in } (0,t_0), \\ d\Pi^{x_0,\xi_0}_{t_0,t} = (\nabla_x \cdot A(Y^{x_0,\xi_0}_{t_0,t}, \Pi^{x_0,\xi_0}_{t_0,t})) \circ dz_{t_0,t} & \text{in } (0,t_0), \\ (Y^{x_0,\xi_0}_{t_0,t_0}, \Pi^{x_0,\xi_0}_{t_0,t_0}) = (x_0,\xi_0). \end{cases} \tag{2.2.22}$$

All the properties (2.2.11)-(2.2.19) of the smooth forward and backward characteristics continue to hold for the limiting rough characteristics, just by deleting the symbol ϵ form the formulas. The rough forward and backward characteristics are inverse of each other in the sense of formula (2.2.11), and they preserve the Lebesgue measure as prescribed by (2.2.14). The assumptions $\nabla_x \cdot A(x,0) \equiv 0$ still implies that the rough characteristics preserve the sign of the velocity variable as in (2.2.15). Similarly, the assumption $\partial_\xi A(x,\xi)_{|\partial Q\times\mathbb{R}} \equiv 0$ guarantees that the rough space characteristics do not move when starting from the boundary, and that, when starting within Q, cannot leave the domain and move slower the closer they get to the boundary, as specified by (2.2.16)-(2.2.18). These concepts are discussed in Section 2.5.

Finally, it follows from the well-posedness of the characteristic systems (2.2.21) and (2.2.22) that the rough transport equation

$$\begin{cases} \partial_t \rho = (\partial_\xi A(x,\xi) \circ dz_t) \cdot \nabla_x \rho - ((\nabla_x \cdot A(x,\xi)) \circ dz_t)\ \partial_\xi \rho & \text{in } \mathbb{R}^d \times \mathbb{R} \times (t_0,\infty), \\ \rho = \rho_0 & \text{on } \mathbb{R}^d \times \mathbb{R} \times \{t_0\}, \end{cases} \tag{2.2.23}$$

is well-posed for any $t_0 \geqslant 0$ and any initial data $\rho_0 \in C^\infty(\mathbb{R}^d \times \mathbb{R})$. Indeed, in analogy with (2.2.13), the solution is represented by the transport of the initial data along the inverse rough characteristics (2.2.22). That is, for each $t_0 \geqslant 0$ and $\rho_0 \in C^\infty(\mathbb{R}^d \times \mathbb{R})$, the solution of (2.2.23) can be written as

$$\rho_{t_0,t}(x,\xi) := \rho_0(Y^{x,\xi}_{t,t-t_0}, \Pi^{x,\xi}_{t,t-t_0}). \tag{2.2.24}$$

Then, as in the smooth case, the assumption $\partial_\xi A(x,\xi)_{|\partial Q\times\mathbb{R}} \equiv 0$ is fundamental to ensure that the transport along rough characteristics of a function that is compactly supported within $Q \times \mathbb{R}$ stays compactly supported. Namely, estimate (2.2.18) implies that

$$\text{if } \rho_0 \in C_c(Q\times\mathbb{R}), \text{ then } \rho_{t_0,t}(x,\xi) = \rho_0(Y^{x,\xi}_{t,t-t_0}, \Pi^{x,\xi}_{t,t-t_0}) \in C_c(Q\times\mathbb{R}\times[t_0,T])\ \forall\, T \geqslant t_0 \geqslant 0.$$

We are now prepared to present the definition of pathwise kinetic solution. Proposition 2.4.3 below proves that the solutions $\mathrm{u}^{\eta,\epsilon}$ of (2.2.2) satisfy, for each $T > 0$, for $C = C(m,Q,T,A,z)$,

$$\| \mathrm{u}^{\eta,\epsilon} \|^2_{L^\infty([0,T];L^2(Q))} + \left\| \nabla (\mathrm{u}^{\eta,\epsilon})^{[m+1/2]} \right\|^2_{L^2([0,T];L^2(Q))} + \left\| \eta^{\frac{1}{2}} \nabla (\mathrm{u}^{\eta,\epsilon}) \right\|^2_{L^2([0,T];L^2(Q))} \leqslant C\left(1 + \|u_0\|^2_{L^2(Q)}\right),$$

uniformly for $\eta \in (0,1)$ and $\epsilon \in (0,1)$. Using similar estimates, and weak convergence and compactness arguments, we would expect the smooth solutions $\mathrm{u}^{\eta,\epsilon}$ to converge to some limiting function u as we let $\eta,\epsilon \to 0$. In turn, we would expect $\chi^{\eta,\epsilon} = \bar{\chi}(\mathrm{u}^{\eta,\epsilon},\xi)$ to converge to

$\bar{\chi}(u,\xi)$. Moreover, the definitions (2.2.5)-(2.2.6) of the entropy and parabolic defect measures and the above estimate on their total mass immediately imply that $\mathrm{p}^{\eta,\epsilon}$ and $\mathrm{q}^{\eta,\epsilon}$ converge weakly in the $\eta,\epsilon \to 0$ limit. Finally, for each $T > 0$, Proposition 2.5.1 ensures that, as $\epsilon \to 0$,

$$Y^{x,\xi,\epsilon}_{t,s} \to Y^{x,\xi}_{t,s} \quad \text{and} \quad \Pi^{x,\xi,\epsilon}_{t,s} \to \Pi^{x,\xi}_{t,s} \quad \text{uniformly for } s \leqslant t \in [0,T] \text{ and } (x,\xi) \in \mathbb{R}^d \times \mathbb{R}.$$

As a consequence, we would expect equation (2.2.20) to pass to the $\eta,\epsilon \to 0$ limit as well. This informal argument motivates the following definition.

Definition 2.2.4. Let $u_0 \in L^2(Q)$. A *pathwise kinetic solution* with initial data u_0 is a function $u \in L^\infty_{\text{loc}}([0,\infty); L^2(Q))$ that satisfies the following properties.

(i) For each $T > 0$,

$$u^{\left[\frac{m+1}{2}\right]} \in L^2([0,T]; H^1_0(Q)).$$

In particular, for each $T > 0$, the parabolic defect measure

$$q(x,\xi,t) := \frac{4m}{(m+1)^2}\delta_0(\xi - u(x,t)) \left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|^2 \tag{2.2.25}$$

is finite on $Q \times \mathbb{R} \times [0,T]$.

(ii) For the kinetic function $\chi(x,\xi,t) = \bar{\chi}(u(x,t),\xi)$, there exists a nonnegative entropy defect measure p on $Q \times \mathbb{R} \times [0,\infty)$, which is finite on $Q \times \mathbb{R} \times [0,T]$ for each $T > 0$, and a subset $\mathcal{N} \subset (0,\infty)$ of Lebesgue measure zero such that, for every $t_0 < t_1 \in [0,\infty)\backslash\mathcal{N}$ and every $\rho_0 \in C^\infty_c(Q \times \mathbb{R})$, for $\rho_{t_0,t}(x,\xi)$ defined by (2.2.24),

$$\begin{aligned}\int_{Q\times\mathbb{R}} \chi(x,\xi,r)\rho_{t_0,r}(x,\xi)\,dx\,d\xi \Big|_{r=t_0}^{r=t_1} &= \int_{t_0}^{t_1}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi(x,\xi,r)\,\Delta_x\rho_{t_0,r}(x,\xi)\,dx\,d\xi\,dr \\ &\quad - \int_{t_0}^{t_1}\int_{Q\times\mathbb{R}} (p(x,\xi,r) + q(x,\xi,r))\;\partial_\xi\rho_{t_0,r}(x,\xi)\,dx\,d\xi\,dr.\end{aligned} \tag{2.2.26}$$

Moreover, the initial condition is enforced in the sense that, when $t_0 = 0$,

$$\int_{Q\times\mathbb{R}} \chi(x,\xi,0)\,\rho_{0,0}(x,\xi)\,dx\,d\xi = \int_{Q\times\mathbb{R}} \bar{\chi}(u_0(x),\xi)\,\rho_0(x,\xi)\,dx\,d\xi.$$

Remark 2.2.5. We observe that in fact, as $\eta \to 0$, the entropy defect measures $\mathrm{p}^{\eta,\epsilon}$ converge weakly to 0, owing to the regularity implied by the parabolic defect measures $\mathrm{q}^{\eta,\epsilon}$. However, due to the weak lower semicontinuity of the L^2-norm, along a subsequence, the weak limit of the parabolic defect measures $\mathrm{q}^{\eta,\epsilon}$ may lose mass in the limit, since the gradients $\nabla(\mathrm{u}^{\eta,\epsilon})^{[m+1/2]}$ will, in general, converge only weakly. The entropy defect measure appearing in Definition 2.2.4 is therefore necessary to account for this potential loss of mass. The complete details of this phenomenon are given in the proof of Theorem 2.1.3 from Section 2.4.

Remark 2.2.6. We remark that (2.2.26) is equivalent to requiring that the kinetic function χ satisfies, for each $\varphi \in C_c^\infty([0,\infty))$, each $t_0 \leqslant t_1 \in [0,\infty)\backslash\mathcal{N}$ and each $\rho_0 \in C_c^\infty(Q\times\mathbb{R})$, for the solution $\rho_{t_0,t}(x,\xi)$ of (2.2.23),

$$\begin{aligned}\int_{t_0}^{t_1}\int_{Q\times\mathbb{R}}\chi(x,\xi,r)\,\rho_{t_0,r}(x,\xi)\,\dot\varphi(r)\,dx\,d\xi\,dr &= \int_{Q\times\mathbb{R}}\chi(x,\xi,r)\rho_{t_0,r}(x,\xi)\varphi(r)\,dx\,d\xi\Big|_{r=t_0}^{r=t_1}\\ &\quad-\int_{t_0}^{t_1}\int_{Q\times\mathbb{R}}m|\xi|^{m-1}\chi(x,\xi,r)\,\Delta_x\rho_{t_0,r}(x,\xi)\,\varphi(r)\,dx\,d\xi\,dr\\ &\quad+\int_{t_0}^{t_1}\int_{Q\times\mathbb{R}}(p(x,\xi,r)+q(x,\xi,r))\,\partial_\xi\rho_{t_0,r}(x,\xi)\,\varphi(r)\,dx\,d\xi\,dr.\end{aligned}$$

Remark 2.2.7. An easy approximation argument shows that in fact we can test equation (2.2.26) against the transport along characteristics of any $\rho_0 \in C^2(Q\times\mathbb{R})$ as long as it has bounded first and second derivatives and it satisfies $\rho_0 \in C_c(Q\times[-M,M])$ for each $M > 0$, i.e. it is compactly supported in Q locally in ξ.

Finally, we observe that the regularity and zero-trace condition imposed by Definition 2.2.4.(i) imply that every pathwise kinetic solution satisfies the following integration by parts formula. The proof is a small modification of [FG19, Lemma 3.6] and is omitted.

Lemma 2.2.8. Let u be a pathwise kinetic solution of (2.1.1) in the sense of Defintion 2.2.4 and let $\chi(x,\xi,t)$ be the associated kinetic function. For each $\psi \in C_c^\infty(Q\times\mathbb{R}\times[0,\infty))$ and each $t_0 \geqslant 0$, we have

$$\int_0^{t_0}\int_{Q\times\mathbb{R}}\frac{m+1}{2}|\xi|^{\frac{m-1}{2}}\chi(x,\xi,r)\nabla\psi(x,\xi,r)\,dx\,d\xi\,dr = -\int_0^{t_0}\int_Q\nabla u^{[\frac{m+1}{2}]}\,\psi(x,u(x,r),r)\,dx\,dr. \tag{2.2.27}$$

2.3 Uniqueness of pathwise kinetic solutions

In this section we prove that pathwise kinetic solutions are unique. In order to motivate and give an overview of the actual proof, we sketch the uniqueness argument in the case of the deterministic porous media equation

$$\begin{cases}\partial_t u = \Delta u^{[m]} & \text{in } Q\times(0,\infty),\\ u = 0 & \text{on } \partial Q\times(0,\infty),\\ u = u_0 & \text{on } Q\times 0.\end{cases} \tag{2.3.1}$$

The corresponding kinetic formulation is

$$\begin{cases}\partial_t\chi = m|\xi|^{m-1}\Delta_x\chi + \partial_\xi(p+q) & \text{in } Q\times\mathbb{R}\times(0,\infty),\\ \chi = 0 & \text{on } \partial Q\times\mathbb{R}\times(0,\infty),\\ \chi = \bar\chi(u_0;\xi) & \text{on } Q\times\mathbb{R}\times\{0\},\end{cases} \tag{2.3.2}$$

where $p \geqslant 0$ is the nonnegative entropy defect measure and q is the parabolic defect measure defined by (2.2.25).

In this setting, the following proof of uniqueness is due to Chen and Perthame [CP03]. Suppose that u^1 and u^2 are two kinetic solutions of (2.3.1) in the sense that the associated kinetic functions χ^1 and χ^2 solve (2.3.2). Properties of the kinetic function yield the identity

$$\begin{aligned}\int_Q |u^1 - u^2|\, dx &= \int_{\mathbb{R}} \int_Q |\chi^1 - \chi^2|^2\, dx\, d\xi = \int_{\mathbb{R}} \int_Q |\chi^1| + |\chi^2| - 2\chi^1\chi^2\, dx\, d\xi \\ &= \int_{\mathbb{R}} \int_Q \chi^1 \mathrm{sgn}(\xi) + \chi^2 \mathrm{sgn}(\xi) - 2\chi^1\chi^2\, dx\, d\xi.\end{aligned} \tag{2.3.3}$$

Now we formally take the derivative in time of this equation and apply equation (2.3.2). Then we integrate by parts in space. Using the distributional equalities

$$\partial_\xi \chi^i(x,\xi,t) = \delta_0(\xi) - \delta_0(\xi - u^i(x,t)) \text{ and } \nabla_x \chi^i = \delta_0(\xi - u^i(x,t))\nabla u^i(x,t), \tag{2.3.4}$$

for $i = 1, 2$, and exploiting the zero boundary conditions we eventually get

$$\begin{aligned}\partial_t \int_Q |u^1 - u^2|\, dx =& \frac{16m}{(m+1)^2} \int_{\mathbb{R}} \int_Q \delta_0(\xi - u^1(x,t))\delta_0(\xi - u^2(x,t))\nabla (u^1)^{\left[\frac{m+1}{2}\right]} \nabla (u^2)^{\left[\frac{m+1}{2}\right]} \\ &- 2\int_{\mathbb{R}} \int_Q \delta_0(\xi - u^1(x,t)) \left(p^2(x,\xi,t) + q^2(x,\xi,t)\right) \\ &- 2\int_{\mathbb{R}} \int_Q \delta_0(\xi - u^2(x,t)) \left(p^1(x,\xi,t) + q^1(x,\xi,t)\right).\end{aligned} \tag{2.3.5}$$

Applications of Hölder's inequality and Young's inequality, together with the definition of the parabolic defect measure and the nonnegativity of the entropy defect measure, prove the right-hand side of (2.3.5) is nonpositive. Integrating in time then completes the proof of uniqueness.

The formal argument leading to (2.3.5) provides the outline for the proof of Theorem 2.1.2. However, even to justify the formal computation, care must be taken to avoid the product of δ-distributions and exploit the boundary conditions. This is achieved by regularizing the sgn and kinetic functions in the space and velocity variables. Additional error terms arise due to the transport of test functions by the inverse characteristics, which are handled using a time-splitting argument that relies crucially on the conservative structure of the equation.

In this setting, the main difficulty with respect to the case with periodic boundary conditions comes from the class of admissible test functions for which equation (2.2.26) actually holds. Indeed, equation (2.2.26) shall play the same role as equation (2.3.2) in the outline of the proof sketched above. Imposing Dirichlet boundary conditions implies that in (2.2.26) the admissible test functions need to be compactly supported within the domain Q. In turn, when writing down the rigorous version of the argument above, we shall need to introduce a cutoff function so as to effectively exploit (2.2.26). To counter the effect of the transport along characteristics imposed by (2.2.26), we consider a cutoff function that has already been transported along the backward characteristics. Having introduced a cutoff, new boundary error terms arise to avoid the product of δ-distributions. We shall have to choose the speed the cutoff function is approaching the boundary with according to the diffusion regime $m \in (0, \infty)$.

This is needed to exploit the zero boundary conditions and the Sobolev regularity of $u^{[m]}$ (cf. Lemma 2.3.7 below) in order to control the derivatives of the cutoff.

We can now prove the uniqueness of pathwise kinetic solutions. Besides the aforementioned regularization, time-splitting and cutoff procedures, the rigorous justification of the argument above relies on sharp estimates for singular moments of the parabolic and entropy defect measures. These estimates are postponed to Proposition 2.3.6 and 2.3.9 below.

Remark 2.3.1. For the remainder of the chapter, after applying the integration by parts formula (2.2.27), we will frequently encounter derivatives of functions $f(x,\xi,r) : Q \times \mathbb{R} \times [0,\infty) \to \mathbb{R}$ evaluated at $\xi = u(x,r)$. In order to simplify the notation, we make the convention that

$$\nabla_x f(x,u(x,r),r) = \nabla_x f(x,\xi,r)_{|\xi=u(x,r)}, \tag{2.3.6}$$

and analogous conventions for all possible derivatives. That is, in every case, the notation indicates the derivative of f evaluated at $(x,u(x,r),r)$ as opposed to the derivative of the full composition.

Proof of Theorem 2.1.2.
The proof will proceed in 10 Steps. In Step 0 we introduce the notation and the machinery needed in the proof. Step 1 presents the scheme for the rigorous justification of the formal proof (2.3.3)-(2.3.5) outlined above, which consists of a time-splitting argument, a cutoff procedure in space, and a regularization in the space and velocity variables. Step 2 analyzes the terms of (2.3.3) involving the sgn function, and Step 3 considers the mixed term. In Step 4 we rigorously observe the cancellation coming from the parabolic defect measure we formally noticed in (2.3.5). In Step 5 we justify the cancellation coming from the integration by parts and the application of the zero boundary conditions formally performed in (2.3.4)-(2.3.5). In Step 6 we make the point and gather the error terms produced by the rigorous computations performed up to this point. In Step 7 we analyze the internal error terms coming from the transport along characteristics of the sgn and kinetic functions. Step 8 tackles the boundary error terms produced by the rigorous application of the integration by parts and the zero boundary conditions. Finally, in Step 9, we pass to the limit with respect to the space and velocity regularization first, the space cutoff second, and the time splitting third, and we conclude the proof.

Step 0: Set up. Let u^1 and u^2 be two pathwise kinetic solutions with initial data $u_0^1, u_0^2 \in L^2_+(Q)$ respectively. For each $i = 1,2$, we shall write χ^i for the corresponding kinetic function, and p^i and q^i for the corresponding entropy and parabolic defect measures. In order to simplify the notation, for each $(x,\xi,r) \in Q \times \mathbb{R} \times [0,\infty)$ we shall use

$$\chi_r^i(x,\xi) := \chi(x,\xi,r), \quad p_r^i(x,\xi) := p^i(x,\xi,r), \quad q_r^i(x,\xi) := q^i(x,\xi,r).$$

Moreover, we shall consider the transported kinetic functions

$$\tilde{\chi}^i_{t,s}(x,\xi) := \chi^i_s(X^{x,\xi}_{t,s}, \Xi^{x,\xi}_{t,s}) \quad \text{for each } (x,\xi,t,s) \in Q \times \mathbb{R} \times [0,\infty)^2,$$

for the rough characteristics (2.2.21).

The proof will start with a time-splitting argument made possible by the conservative structure of the equation. For $i = 1, 2$, let $\mathcal{N}^i \subseteq (0,\infty)$ be the null set corresponding to u^i from Definition 2.2.4, and define $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2$. Let $T \in [0,\infty)\backslash\mathcal{N}$ be fixed, but arbitrary. We shall consider arbitrary partitions $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ of $[0,T]$, for any $N \in \mathbb{N}$, with the constraint that $\mathcal{P} \subseteq [0,T]\backslash\mathcal{N}$. We shall denote by $|\mathcal{P}| = \max_i |t_{i+1} - t_i|$ the diameter of the partition.

The imposition of the zero boundary conditions requires us to test the kinetic equation (2.2.26) against compactly supported test functions only, and we thus need to introduce a cutoff function. First, define the signed distance function $d_{\partial Q} : \mathbb{R}^d \to \mathbb{R}$ by

$$d_{\partial Q}(x) = \begin{cases} \operatorname{dist}(x,\partial Q) & \text{if } x \in \bar{Q}, \\ -\operatorname{dist}(x,\partial Q) & \text{if } x \in Q^c. \end{cases} \tag{2.3.7}$$

Since the domain Q is smooth, this defines a smooth function. Moreover, for $x \in \mathbb{R}^d$ close enough to ∂Q, we have the identity

$$\nabla d_{\partial Q}(x) = -\tilde{n}(x), \tag{2.3.8}$$

where $\tilde{n}(x)$ denotes the extended unit outward normal to ∂Q. Namely, for any $x_0 \in \partial Q$, let $n(x_0)$ denote the actual unit outward normal to ∂Q. For any $x \in \mathbb{R}^d$ close enough to ∂Q there exists a unique pair $(x^*, t) \in \partial Q \times \mathbb{R}$ such that $x = x^* - tn(x^*)$, precisely given by $t = d_{\partial Q}(x)$ and x^* the (unique) closest point on ∂Q, and we set $\tilde{n}(x) := n(x^*)$. For each $\ell \in (0,1)$, we shall denote

$$Q^\ell := \{x \in Q \mid d_{\partial Q}(x) \leqslant c_1\ell\} \quad \text{and} \quad Q_\ell := \{x \in Q \mid d_{\partial Q}(x) > c_1\ell\}, \tag{2.3.9}$$

for a constant $c_1 = c_1(T, A, z, Q, d, m) > 0$, possibly changing throughout the proof, but only depending on the parameters specified. It is then immediate from the definition that

$$\operatorname{meas}\left(Q^\ell\right) \leqslant C\ell, \tag{2.3.10}$$

for a constant $C = C(T, A, z, Q, m)$. Moreover, we point out the following property, which is a simple consequence of estimate (2.5.8): for any $(x,\xi) \in Q \times \mathbb{R}$ and any $s \leqslant t_0 \in [0,T]$, possibly after enlargening the constant c_1 in (2.3.9),

$$\text{if } Y^{x,\xi}_{t_0,s} \in Q^\ell, \text{ then } x \in Q^\ell, \text{ and viceversa.} \tag{2.3.11}$$

Now, for each $\beta \in (0,1)$, we define a cutoff function $\phi_\beta = \phi_{\beta,m}$ within Q, whose shape depends on the particular diffusion exponent $m \in (0,\infty)$. Precisely, set $\gamma_m = (m+2) \wedge 3$ and consider the piecewise linear function $\psi_\beta^\circ : \mathbb{R} \to [0,1]$ given by

$$\psi_\beta^\circ(s) = \begin{cases} 0 & \text{if } s \leqslant \beta^{\gamma_m}, \\ \beta^{-1}\left(s - \beta^{\gamma_m}\right) & \text{if } s \in \left[\beta^{\gamma_m}, \beta + \beta^{\gamma_m}\right], \\ 1 & \text{if } s \geqslant \beta + \beta^{\gamma_m}. \end{cases} \tag{2.3.12}$$

Let $\rho_1 : \mathbb{R} \to [0,1]$ be a standard 1-dimensional positive radial mollifier supported in the unitary ball, and set

$$\psi_\beta(s) := \rho_1^{\frac{1}{2}\beta^{\gamma_m}} * \psi_\beta^\circ(s) \quad \text{ for } s \in \mathbb{R}, \tag{2.3.13}$$

where $\rho_1^{\frac{1}{2}\beta^{\gamma_m}}$ denotes the mollifier rescaled at order $\frac{1}{2}\beta^{\gamma_m}$. Finally, we define a cutoff function $\phi_\beta : \mathbb{R}^d \to [0,1]$ by setting

$$\phi_\beta(x) := \psi_\beta(d_{\partial Q}(x)). \tag{2.3.14}$$

This definition guarantees that $\phi_\beta \in C_c^\infty(Q)$ and it satisfies

$$\phi_\beta(y) = \begin{cases} 1 & \text{if } d_{\partial Q}(y) > \beta + \dfrac{3}{2}\beta^{\gamma_m}, \\ 0 & \text{if } d_{\partial Q}(y) < \dfrac{1}{2}\beta^{\gamma_m}. \end{cases}$$

Since $\beta + \frac{3}{2}\beta^{\gamma_m} \simeq \beta$, recalling the notation (2.3.9), we shall write that $\phi_\beta \equiv 1$ in Q_β and $\phi_\beta \equiv 0$ in $Q^{\beta^{\gamma_m}}$. We also point out that, for $k \in \mathbb{N}$, for a constant $C = C(Q)$,

$$D^k\phi_\beta \equiv 0 \quad \text{in } Q_\beta \qquad \text{and} \qquad |D^k\phi_\beta| \leqslant C\beta^{-k} \quad \text{in } Q^\beta. \tag{2.3.15}$$

Moreover, we observe the following expression for the laplacian $\Delta\phi_\beta$. The distributional derivatives of ψ_β° and the mollification imply that

$$\ddot{\psi}_\beta(s) = \frac{1}{\beta}\rho_1^{\frac{1}{2}\beta^{\gamma_m}}\big(s - \beta^{\gamma_m}\big) - \frac{1}{\beta}\rho_1^{\frac{1}{2}\beta^{\gamma_m}}\big(s - (\beta + \beta^{\gamma_m})\big). \tag{2.3.16}$$

Then, the definition of ϕ_β and formula (2.3.8) yield

$$\Delta\phi_\beta = \ddot{\psi}_\beta(d_{\partial Q}(x)) - \dot{\psi}_\beta(d_{\partial Q}(x))\nabla \cdot \tilde{n}(x). \tag{2.3.17}$$

Finally, we shall need a regularization procedure. Let ρ_1 and ρ_d be standard 1-dimensional and d-dimensional positive radial mollifiers supported in the unitary ball. For each $\epsilon \in (0,1)$, denote by ρ_1^ϵ and ρ_d^ϵ their rescaling of order ϵ. For $i = 1,2$, with the notation (2.3.9), for each $(y,\eta) \in Q_\epsilon \times \mathbb{R}$ and each $t \leqslant r \in [0,\infty)$, we define the smoothed transported kinetic functions

$$\begin{aligned} \tilde{\chi}_{t,r}^{i,\epsilon}(y,\eta) := \left(\tilde{\chi}_{t_i,r}^i * \rho_d^\epsilon\rho_1^\epsilon\right)(y,\eta) &= \int_{Q\times\mathbb{R}} \chi_r^i(X_{t,r}^{x,\xi}, \Xi_{t,r}^{x,\xi})\rho_d^\epsilon(x-y)\rho_1^\epsilon(\xi-\eta)\,dx\,d\xi \\ &= \int_{Q\times\mathbb{R}} \chi_r^i(x,\xi)\rho_d^\epsilon(Y_{r,r-t}^{x,\xi} - y)\rho_1^\epsilon(\Pi_{r,r-t}^{x,\xi} - \eta)\,dx\,d\xi, \end{aligned} \tag{2.3.18}$$

where, in the last inequality, we used the inverse eqrefformula/inverse relation for smooth characteristics and conservative (2.2.14) properties of characteristics. In particular, for each $(y, \eta) \in Q \times \mathbb{R}$ and $t \leqslant r \in [0, \infty)$,

$$\lim_{\epsilon \to 0} \tilde{\chi}^{i,\epsilon}_{t,r}(y, \eta) = \tilde{\chi}^{i}_{t,r}(y, \eta).$$

To simplify the notation, for $(x, y, \xi, \eta) \in Q^2 \times \mathbb{R}^2$ and $t \leqslant r \in [0, \infty)$, we shall write

$$\rho^{\epsilon}_{t,r}(x, y, \xi, \eta) := \rho^{\epsilon}_{d}(Y^{x,\xi}_{r,r-t} - y)\rho^{\epsilon}_{1}(\Pi^{x,\xi}_{r,r-t} - \eta). \tag{2.3.19}$$

According to (2.2.24), this represents the solution to the rough differential equation (2.2.23) beginning at time t with initial data $\rho^{\epsilon}_{d}(\cdot - y)\rho^{\epsilon}_{1}(\cdot - \eta)$. We observe that, for each fixed $(y, \eta) \in Q_\epsilon \times \mathbb{R}$ and $t \leqslant t' \in [0, \infty)\backslash\mathcal{N}$, the kinetic equation (2.2.26) can be applied to (2.3.18).

In conclusion, we shall also consider the regularized transported sgn function, for $(y, \eta) \in Q_\epsilon \times \mathbb{R}$ and $t \leqslant r \in [0, \infty)$,

$$\tilde{\text{sgn}}^{\epsilon}_{t,r}(y, \eta) := \int_{Q \times \mathbb{R}} \text{sgn}(\Xi^{x,\xi}_{t,r})\rho^{\epsilon}_{d}(x - y)\rho^{\epsilon}_{1}(\xi - \eta)\,dx\,d\xi = \int_{Q \times \mathbb{R}} \text{sgn}(\xi)\rho^{\epsilon}_{t,r}(x, y, \xi, \eta)\,dx\,d\xi. \tag{2.3.20}$$

In fact, formula (2.2.15) ensures that $\tilde{\text{sgn}}_{t,r}(x, \xi) := \text{sgn}(\Xi^{x,\xi}_{t,r}) = \text{sgn}(\xi)$, so that we find

$$\tilde{\text{sgn}}^{\epsilon}_{t,r}(y, \eta) = \int_{Q \times \mathbb{R}} \text{sgn}(\xi)\rho^{\epsilon}_{d}(x - y)\rho^{\epsilon}_{1}(\xi - \eta)\,dx\,d\xi = \int_{\mathbb{R}} \text{sgn}(\xi)\rho^{\epsilon}_{1}(\xi - \eta)\,d\xi. \tag{2.3.21}$$

Thus $\tilde{\text{sgn}}^{\epsilon}_{t,r}(y, \eta) = \text{sgn}^{\epsilon}(\eta)$ is just the standard ϵ-mollification of the sgn function and is independent of $t \leqslant r \in [0, \infty)$ and $y \in Q$. It will nevertheless be useful to consider this regularized transported expression, since it will clarify important cancellations in the arguments to follow.

Step 1: Time splitting, cutoff and regularization. Let $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ be a partition of $[0, T]$ with $\mathcal{P} \subseteq [0, T]\backslash\mathcal{N}$. Properties of the kinetic function, the definition of the cutoff ϕ_β and property (2.2.17), and the conservative (2.2.14) and inverse (2.2.11) property of the characteristics imply that

$$\begin{aligned}
&\int_Q \left|u^1(y, r) - u^2(y, r)\right|\,dy\Big|_{r=0}^{r=T} \\
&= \sum_{i=0}^{N-1} \int_Q \left|u^1(y, r) - u^2(y, r)\right|\,dy\Big|_{r=t_i}^{r=t_{i+1}} \\
&= \sum_{i=0}^{N-1} \int_{Q \times \mathbb{R}} \left|\chi^1_r(y, \eta) - \chi^2_r(y, \eta)\right|^2\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&= \lim_{\beta \to 0} \sum_{i=0}^{N-1} \int_{Q \times \mathbb{R}} \left|\chi^1_{t_i,r}(y, \eta) - \chi^2_{t_i,r}(y, \eta)\right|^2 \phi_\beta(Y^{y,\eta}_{r,r-t_i})\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&= \lim_{\beta \to 0} \sum_{i=0}^{N-1} \int_{Q \times \mathbb{R}} \left|\tilde{\chi}^1_{t_i,r}(y, \eta) - \tilde{\chi}^2_{t_i,r}(y, \eta)\right|^2 \phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}.
\end{aligned} \tag{2.3.22}$$

Moreover, unfolding the square and introducing the ϵ-regularization, we get

$$
\begin{aligned}
&\int_Q \left|u^1(y,r)-u^2(y,r)\right| dy\Big|_{r=0}^{r=T} \\
&\quad = \lim_{\beta\to 0}\sum_{i=0}^{N-1}\int_{Q\times\mathbb{R}}\left(\tilde{\chi}^1_{t_i,r}\tilde{\operatorname{sgn}}_{t_i,r}+\tilde{\chi}^2_{t_i,r}\tilde{\operatorname{sgn}}_{t,i,r}-2\tilde{\chi}^1_{t_i,r}\tilde{\chi}^2_{t_i,r}\right)\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&\quad = \lim_{\beta\to 0}\lim_{\epsilon\to 0}\sum_{i=0}^{N-1}\int_{Q\times\mathbb{R}}\left(\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}+\tilde{\chi}^{2,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}-2\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}\right)\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}.
\end{aligned}
\tag{2.3.23}
$$

The next four steps are devoted to the analysis of the difference

$$
\int_{Q\times\mathbb{R}}\left(\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}+\tilde{\chi}^{2,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}-2\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}\right)\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}},
\tag{2.3.24}
$$

for any $i=0,\ldots,N-1$ and any $\epsilon,\beta\in(0,1)$, with $\epsilon\ll\beta$ and $\beta\ll|\mathcal{P}|$. In Step 3 we consider the sgn term $\tilde{\chi}^{j,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}$ for $j=1,2$, and in Step 4 we consider the mixed term $\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}$. Finally, in Step 5 and 6 we shall observe crucial cancellations among these terms.

Step 2: The sgn terms. We will first analyze the term involving the sgn function in (2.3.24). For the convolution kernel (2.3.19), we shall write $(x,\xi)\in Q\times\mathbb{R}$ for the integration variables defining $\tilde{\chi}^{1,\epsilon}_{t_i,r}$ and we shall write $\rho^{1,\epsilon}_{t_i,r}$ for the corresponding convolution kernel. The kinetic equation (2.2.26) and (2.3.21) imply that, with the notation from (2.3.18) and (2.3.19),

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}}\tilde{\chi}^{1,\epsilon}_{t_i,r}(y,\eta)\,\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&\qquad = \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\,\Delta_x\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\right)\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\phi_\beta(y)\,dy\,d\eta\,dr \\
&\qquad\quad -\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p^1_r+q^1_r)\,\partial_\xi\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\right)\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\phi_\beta(y)\,dy\,d\eta\,dr.
\end{aligned}
\tag{2.3.25}
$$

The first and second term in (2.3.25) will be handled separately. Observe that, from (2.3.19), for each $(x,y,\xi,\eta,r)\in\mathbb{R}^{2d}\times\mathbb{R}^2\times[t_i,\infty)$,

$$
\nabla_x\rho^{1,\epsilon}_{t_i,r}=-\nabla_y\rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\,D_xY^{x,\xi}_{r,r-t_i}-\partial_\eta\rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\,\nabla_x\Pi^{x,\xi}_{r,r-t_i},
\tag{2.3.26}
$$

and

$$
\partial_\xi\rho^{1,\epsilon}_{t_i,r}=-\nabla_y\rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\,\partial_\xi Y^{x,\xi}_{r,r-t_i}-\partial_\eta\rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\,\partial_\xi\Pi^{x,\xi}_{r,r-t_i}.
\tag{2.3.27}
$$

For the first term on the right-hand side of (2.3.25), we use formula (2.3.26) and then integrate by parts in the (y,η) variables to get

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\,\Delta_x\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\right)\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\phi_\beta(y)\,dy\,d\eta\,dr \\
&=\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\nabla_x\rho^{1,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\phi_\beta(y)\,dy\,d\eta\right)dx\,d\xi\,dr \\
&=\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\left(\nabla_y\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}D_xY^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_x\Pi^{x,\xi}_{r,r-t_i}\right)\phi_\beta\,dy\,d\eta\right)dx\,d\xi\,dr \\
&\quad+\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\right)dx\,d\xi\,dr.
\end{aligned}
\tag{2.3.28}
$$

For the first term on the right-hand side of (2.3.28), it follows from definition (2.3.20) and the computation (2.3.26) that, after adding and subtracting the terms $D_{x'}Y^{x',\xi'}_{r,r-t_i}$ and $\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}$,

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\left(\nabla_y\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}D_xY^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_x\Pi^{x,\xi}_{r,r-t_i}\right)\phi_\beta\,dy\,d\eta\right)dx\,d\xi\,dr\\&=IE^{\mathrm{sgn}1,1}_i-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr,\end{aligned}\tag{2.3.29}$$

for the internal error term relative to the sgn term, omitting the integration variables,

$$\begin{aligned}IE^{\mathrm{sgn}1,1}_i:=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\phi_\beta(y)\,\nabla_y\rho^{2,\epsilon}_{t_i,r}\left(D_xY^{x,\xi}_{r,r-t_i}-D_{x'}Y^{x',\xi'}_{r,r-t_i}\right)\right)\\&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\phi_\beta(y)\,\partial_\eta\rho^{2,\epsilon}_{t_i,r}\left(\nabla_x\Pi^{x,\xi}_{r,r-t_i}-\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}\right)\right),\end{aligned}\tag{2.3.30}$$

and where the last term of (2.3.29) vanishes after integrating by parts in the x'-variable. That is,

$$\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr=0.\tag{2.3.31}$$

For the second term of (2.3.28), we first rewrite

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\,\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\right)dx\,d\xi\,dr\\&=BE^{\mathrm{sgn}1,1}_i+\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2} m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dx\,dy\,d\xi\,d\eta\,dr,\end{aligned}\tag{2.3.32}$$

for the boundary error term relative to the sgn term

$$BE^{\mathrm{sgn}1,1}_i=\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\,\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_y\phi_\beta(y)\Delta_xY^{x,\xi}_{r,r-t_i}dx\,dy\,d\xi\,d\eta\,dr.\tag{2.3.33}$$

For the second term on the right-hand side of (2.3.32), arguing as in (2.3.28)-(2.3.30), we use formula (2.3.26), add and subtract the terms $D_{x'}Y^{x',\xi'}_{r,r-t_i}$ and $\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}$, and then integrate by parts in the (y,η) variables to get

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2} m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\tilde{\operatorname{sgn}}^\epsilon_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\,dx\,d\xi\,dr\\&=\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\ \operatorname{tr}\left(\left(D_xY^{x,\xi}_{r,r-t_i}\right)^TD^2_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\right)dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr\\&-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr+BE^{\mathrm{sgn}1,2}_i,\end{aligned}\tag{2.3.34}$$

for the boundary error term, omitting the integration variables,

$$\begin{aligned}BE^{\mathrm{sgn}1,2}_i:=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,\nabla_y\rho^{2,\epsilon}_{t_i,r}\left(D_xY^{x,\xi}_{r,r-t_i}-D_{x'}Y^{x',\xi'}_{r,r-t_i}\right)\\&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\partial_\eta\rho^{2,\epsilon}_{t_i,r}\left(\nabla_x\Pi^{x,\xi}_{r,r-t_i}-\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}\right),\end{aligned}\tag{2.3.35}$$

and where the second term in (2.3.34) vanishes after integrating by parts in the x'-variable,

$$\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\operatorname{sgn}(\xi')\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr=0.\tag{2.3.36}$$

We now consider the second term in (2.3.25). Using formula (2.3.27) and integrating by parts in the (y, η) variables, we obtain

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p_r^1+q_r^1)\,\partial_\xi\rho_{t_i,r}^{1,\epsilon}\,dx\,d\xi\right)\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\phi_\beta(y)\,dy\,d\eta\,dr\\
&= BE_i^{\operatorname{sgn}1,3}\\
&\quad+\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p_r^1+q_r^1)\rho_{t_i,r}^{1,\epsilon}\,\phi_\beta(y)\big(\nabla_y\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\partial_\xi Y_{r,r-t_i}^{x,\xi}+\partial_\eta\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\partial_\xi\Pi_{r,r-t_i}^{x,\xi}\big)\,dx\,dy\,d\xi\,d\eta\,dr,
\end{aligned}
\tag{2.3.37}
$$

for the boundary error term

$$
BE_i^{\operatorname{sgn}1,3}:=\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\nabla_y\phi_\beta(y)\partial_\xi Y_{r,r-t_i}^{x,\xi}dx\,dy\,d\xi\,d\eta\,dr. \tag{2.3.38}
$$

For the second term on the right-hand side of (2.3.37), we use definition (2.3.20), add and subtract the derivatives $\partial_{\xi'}Y_{r,r-t_i}^{x',\xi'}$ and $\partial_{\xi'}\Pi_{r,r-t_i}^{x',\xi'}$, and recall formula (2.3.27) to write

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p_r^1+q_r^1)\rho_{t_i,r}^{1,\epsilon}\,\phi_\beta(y)\big(\nabla_y\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\partial_\xi Y_{r,r-t_i}^{x,\xi}+\partial_\eta\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\partial_\xi\Pi_{r,r-t_i}^{x,\xi}\big)\,dx\,dy\,d\xi\,d\eta\,dr\\
&= IE_i^{\operatorname{sgn}1,2}-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\partial_{\xi'}\rho_{t_i,r}^{2,\epsilon}\operatorname{sgn}(\xi')\,\phi_\beta(y)\,dxdx'\,dy\,d\xi\,d\xi'\,d\eta\,dr,
\end{aligned}
\tag{2.3.39}
$$

for the internal error term, omitting the integration variables,

$$
\begin{aligned}
IE_i^{\operatorname{sgn}1,2}:=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\operatorname{sgn}(\xi')\,\phi_\beta(y)\,\nabla_y\rho_{t_i,r}^{2,\epsilon}\big(\partial_\xi Y_{r,r-t_i}^{x,\xi}-\partial_{\xi'}Y_{r,r-t_i}^{x',\xi'}\big)\\
&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\operatorname{sgn}(\xi')\,\phi_\beta(y)\,\partial_\eta\rho_{t_i,r}^{2,\epsilon}\big(\partial_\xi\Pi_{r,r-t_i}^{x,\xi}-\partial_{\xi'}\Pi_{r,r-t_i}^{x',\xi'}\big).
\end{aligned}
\tag{2.3.40}
$$

Moreover, after integrating by parts in the ξ'-variable and using the distributional equality $\partial_{\xi'}\operatorname{sgn}(\xi')=2\delta_0(\xi')$, the second term in (2.3.39) becomes

$$
\begin{aligned}
&-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\partial_{\xi'}\rho_{t_i,r}^{2,\epsilon}\operatorname{sgn}(\xi')\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr\\
&\qquad= 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr.
\end{aligned}
\tag{2.3.41}
$$

Returning to (2.3.25), it follows from (2.3.28), (2.3.29), (2.3.31), (2.3.32), (2.3.34), (2.3.36), (2.3.37), (2.3.39) and (2.3.41) that

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}}\tilde{\chi}_{t_i,r}^{1,\epsilon}(y,\eta)\,\tilde{\operatorname{sgn}}_{t_i,r}^{\epsilon}\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}\\
&\quad= IE_i^{\operatorname{sgn}1,1}-IE_i^{\operatorname{sgn}1,2}+BE_i^{\operatorname{sgn}1,1}+BE_i^{\operatorname{sgn}1,2}-BE_i^{\operatorname{sgn}1,3}\\
&\qquad+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi_r^1\,\rho_{t_i,r}^{1,\epsilon}\rho_{t_i,r}^{2,\epsilon}\operatorname{sgn}(\xi')\\
&\qquad\qquad\cdot\operatorname{tr}\Big(\big(D_xY_{r,r-t_i}^{x,\xi}\big)^TD_y^2\phi_\beta(y)D_xY_{r,r-t_i}^{x,\xi}\Big)\,dx\,dx'dyd\xi d\xi'd\eta dr\\
&\qquad-2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr.
\end{aligned}
\tag{2.3.42}
$$

Furthermore, identical considerations with χ^1 replaced by χ^2 prove that, after swapping the roles of (x,ξ) and (x',ξ'),

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}} \tilde{\chi}^{2,\epsilon}_{t_i,r}(y,\eta)\,\tilde{\mathrm{sgn}}^{\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&\quad = IE_i^{\mathrm{sgn}2,1} - IE_i^{\mathrm{sgn}2,2} + BE_i^{\mathrm{sgn}2,1} + BE_i^{\mathrm{sgn}2,2} - BE_i^{\mathrm{sgn}2,3} \\
&\qquad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}\mathrm{sgn}(\xi) \\
&\qquad\qquad \mathrm{tr}\Big(\big(D_{x'}Y^{x',\xi'}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y) D_{x'}Y^{x',\xi'}_{r,r-t_i}\Big)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\
&\qquad - 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^2_r+q^2_r)\,\rho^{2,\epsilon}_{t_i,r}\,\rho^{1,\epsilon}_{t_i,r}(x,y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi'\,d\eta\,dr,
\end{aligned}
\tag{2.3.43}
$$

for internal error terms $IE_i^{\mathrm{sgn}2,1}$, $IE_i^{\mathrm{sgn}2,2}$, and boundary error terms $BE_i^{\mathrm{sgn}2,1}$, $BE_i^{\mathrm{sgn}2,2}$ and $BE_i^{\mathrm{sgn}2,3}$, defined in exact analogy with (2.3.30), (2.3.40), and (2.3.33), (2.3.35) and (2.3.38) respectively.

Step 3: The mixed term. We shall now analyze the mixed term in (2.3.24). For the convolution kernel (2.3.19), we shall write $(x,\xi)\in Q\times\mathbb{R}$ for the integration variables defining $\tilde{\chi}^{1,\epsilon}_{t_i,r}$, and we shall write $\rho^{1,\epsilon}_{t_i,r}$ for the corresponding convolution kernel. We shall write $(x',\xi')\in Q\times\mathbb{R}$ for the integration variables defining $\tilde{\chi}^{2,\epsilon}_{t_i,r}$, and $\rho^{2,\epsilon}_{t_i,r}$ for the corresponding convolution kernel. The kinetic equation (2.2.26) implies that

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}} \tilde{\chi}^{1,\epsilon}_{t_i,r}(y,\eta)\,\tilde{\chi}^{2,\epsilon}_{t_i,r}(y,\eta)\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&\quad = \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\Big(\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\,\Delta_x\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\Big)\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr \\
&\qquad - \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\Big(\int_{Q\times\mathbb{R}} (p^1_r+q^1_r)\,\partial_\xi\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\Big)\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr \\
&\qquad + \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\Big(\int_{Q\times\mathbb{R}} m|\xi'|^{m-1}\chi^2_r\,\Delta_{x'}\rho^{2,\epsilon}_{t_i,r}\,dx'd\xi'\Big)\,\tilde{\chi}^{1,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr \\
&\qquad - \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\Big(\int_{Q\times\mathbb{R}} (p^2_r+q^2_r)\,\partial_{\xi'}\rho^{2,\epsilon}_{t_i,r}\,dx'd\xi'\Big)\,\tilde{\chi}^{1,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr.
\end{aligned}
\tag{2.3.44}
$$

The first term in (2.3.44) is studied in analogy to (2.3.28)-(2.3.36). Precisely, we use formula (2.3.26) and then integrate by parts in the (y,η) variables to write

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\Big(\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\,\Delta_x\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\Big)\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr \\
&\quad = \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\Big(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\big(\nabla_y\tilde{\chi}^{2,\epsilon}_{t_i,r}D_xY^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\chi}^{2,\epsilon}_{t_i,r}\nabla_x\Pi^{x,\xi}_{r,r-t_i}\big)\phi_\beta(y)\,dy\,d\eta\Big)\,dx\,d\xi\,dr \\
&\qquad + \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\Big(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\Big)\,dx\,d\xi\,dr.
\end{aligned}
\tag{2.3.45}
$$

For the first term on the right-hand side of (2.3.45), it follows from definition (2.3.18) and formula (2.3.26) that, after adding and subtracting the terms $D_{x'}Y^{x',\xi'}_{r,r-t_i}$ and $\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}$,

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\left(\nabla_y\tilde{\chi}^{2,\epsilon}_{t_i,r}D_xY^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\chi}^{2,\epsilon}_{t_i,r}\nabla_x\Pi^{x,\xi}_{r,r-t_i}\right)\phi_\beta(y)\,dy\,d\eta\right)dx\,d\xi\,dr\\&\quad=IE^{\text{mix}1,1}_i-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\chi^2_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr,\end{aligned}\tag{2.3.46}$$

for the internal error term relative to the mixed term, omitting the integration variables,

$$\begin{aligned}IE^{\text{mix}1,1}_i:=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\rho^{1,\epsilon}_{t_i,r}\chi^2_r\phi_\beta(y)\,\nabla_y\rho^{2,\epsilon}_{t_i,r}\left(D_xY^{x,\xi}_{r,r-t_i}-D_{x'}Y^{x',\xi'}_{r,r-t_i}\right)\right)\\&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\rho^{1,\epsilon}_{t_i,r}\chi^2_r\phi_\beta(y)\,\partial_\eta\rho^{2,\epsilon}_{t_i,r}\left(\nabla_x\Pi^{x,\xi}_{r,r-t_i}-\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}\right)\right).\end{aligned}\tag{2.3.47}$$

For the second term of (2.3.45), we first rewrite

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}m|\xi|^{m-1}\chi^1_r\nabla_x\cdot\left(\int_{Q\times\mathbb{R}}\rho^{1,\epsilon}_{t_i,r}\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\right)dx\,d\xi\,dr\\&\quad=BE^{\text{mix}1,1}_i+\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dx\,dy\,d\xi\,d\eta\,dr,\end{aligned}\tag{2.3.48}$$

for the boundary error term relative to the mixed term

$$BE^{\text{mix}1,1}_i=\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}m|\xi|^{m-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)\,\Delta_xY^{x,\xi}_{r,r-t_i}\,dx\,dy\,d\xi\,d\eta\,dr.\tag{2.3.49}$$

For the second term on the right-hand side of (2.3.48), arguing as in (2.3.45)-(2.3.47), we use formula (2.3.26), add and subtract the terms $D_{x'}Y^{x',\xi'}_{r,r-t_i}$ and $\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}$, and then integrate by parts in the (y,η) variables to get

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}m|\xi|^{m-1}\chi^1_r\nabla_x\rho^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dy\,d\eta\,dx\,d\xi\,dr\\&=\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}\,\mathrm{tr}\left(\left(D_xY^{x,\xi}_{r,r-t_i}\right)^TD^2_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\right)dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr\\&\quad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr+BE^{\text{mix}1,2}_i,\end{aligned}\tag{2.3.50}$$

for the boundary error term, omitting the integration variables,

$$\begin{aligned}BE^{\text{mix}1,2}_i:=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\rho^{1,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\nabla_y\rho^{2,\epsilon}_{t_i,r}\left(D_xY^{x,\xi}_{r,r-t_i}-D_{x'}Y^{x',\xi'}_{r,r-t_i}\right)\\&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\partial_\eta\rho^{2,\epsilon}_{t_i,r}\left(\nabla_x\Pi^{x,\xi}_{r,r-t_i}-\nabla_{x'}\Pi^{x',\xi'}_{r,r-t_i}\right).\end{aligned}\tag{2.3.51}$$

In conclusion, using (2.3.45), (2.3.46), (2.3.48) and (2.3.50), we rewrite the first term on the right-hand side of (2.3.44) as

$$\begin{aligned}&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}m|\xi|^{m-1}\chi^1_r\,\Delta_x\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\right)\tilde{\chi}^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr\\&\quad=IE^{\text{mix}1,1}_i+BE^{\text{mix}1,1}_i+BE^{\text{mix}1,2}_i\\&\quad+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\,\mathrm{tr}\left(\left(D_xY^{x,\xi}_{r,r-t_i}\right)^TD^2_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\right)dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr\\&\quad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr\\&\quad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr.\end{aligned}\tag{2.3.52}$$

Furthermore, after swapping the roles of χ^1 and χ^2, identical considerations allow us to rewrite the third term on the right-hand side of (2.3.44) as

$$\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}} m|\xi'|^{m-1}\chi^2_r\,\Delta_{x'}\rho^{2,\epsilon}_{t_i,r}\,dx'd\xi'\right)\tilde{\chi}^{1,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\quad= IE^{\mathrm{mix2},1}_i + BE^{\mathrm{mix2},1}_i + BE^{\mathrm{mix2},2}_i\\
&\quad+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\,\mathrm{tr}\Big(\big(D_{x'}Y^{x',\xi'}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}\Big)\,dx dx' dy d\xi d\xi' d\eta dr\\
&\quad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}\,dx dx' dy d\xi d\xi' d\eta dr\\
&\quad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)\,dx dx' dy d\xi d\xi' d\eta dr,
\end{aligned} \tag{2.3.53}$$

for internal and boundary error terms $IE^{\mathrm{sgn2},1}_i$, $BE^{\mathrm{mix2},1}_i$ and $BE^{\mathrm{mix2},2}_i$ defined in exact analogy to (2.3.47), (2.3.49) and (2.3.51) respectively.

We now treat the second and fourth term in (2.3.44) in analogy to (2.3.37)-(2.3.40). Using formula (2.3.27) and integrating by parts in the (y,η) variables, we obtain

$$\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p^1_r+q^1_r)\,\partial_\xi\rho^{1,\epsilon}_{t_i,r}\,dx\,d\xi\right)\tilde{\chi}^{2,\epsilon}_{t_i,r}\phi_\beta(y)\,dy\,d\eta\,dr\\
&\quad= BE^{\mathrm{mix1},3}_i+\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p^1_r+q^1_r)\rho^{1,\epsilon}_{t_i,r}\,\phi_\beta(y)\big(\nabla_y\tilde{\chi}^{2,\epsilon}_{t_i,r}\partial_\xi Y^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\chi}^{2,\epsilon}_{t_i,r}\partial_\xi\Pi^{x,\xi}_{r,r-t_i}\big)\,dx\,dy\,d\xi\,d\eta\,dr,
\end{aligned} \tag{2.3.54}$$

for the boundary error term

$$BE^{\mathrm{mix1},3}_i := \int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\tilde{\chi}^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta(y)\partial_\xi Y^{x,\xi}_{r,r-t_i}dx\,dy\,d\xi\,d\eta\,dr. \tag{2.3.55}$$

For the second term on the right-hand side of (2.3.54), we use definition (2.3.18), add and subtract the derivatives $\partial_{\xi'}Y^{x',\xi'}_{r,r-t_i}$ and $\partial_{\xi'}\Pi^{x',\xi'}_{r,r-t_i}$, and recall formula (2.3.27) to write

$$\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2}(p^1_r+q^1_r)\rho^{1,\epsilon}_{t_i,r}\,\phi_\beta(y)\big(\nabla_y\tilde{\chi}^{2,\epsilon}_{t_i,r}\partial_\xi Y^{x,\xi}_{r,r-t_i}+\partial_\eta\tilde{\chi}^{2,\epsilon}_{t_i,r}\partial_\xi\Pi^{x,\xi}_{r,r-t_i}\big)\,dx\,dy\,d\xi\,d\eta\,dr\\
&\quad= IE^{\mathrm{mix1},2}_i-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\partial_{\xi'}\rho^{2,\epsilon}_{t_i,r}\,\chi^2_r\,\phi_\beta(y)\,dxdx'\,dy\,d\xi\,d\xi'\,d\eta\,dr,
\end{aligned} \tag{2.3.56}$$

for the internal error term, omitting the integration variables,

$$\begin{aligned}
IE^{\mathrm{mix1},2}_i :=& \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\chi^2_r\,\phi_\beta(y)\,\nabla_y\rho^{2,\epsilon}_{t_i,r}\big(\partial_\xi Y^{x,\xi}_{r,r-t_i}-\partial_{\xi'}Y^{x',\xi'}_{r,r-t_i}\big)\\
&+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\chi^2_r\,\phi_\beta(y)\,\partial_\eta\rho^{2,\epsilon}_{t_i,r}\big(\partial_\xi\Pi^{x,\xi}_{r,r-t_i}-\partial_{\xi'}\Pi^{x',\xi'}_{r,r-t_i}\big).
\end{aligned} \tag{2.3.57}$$

For the second term on the right-hand side of (2.3.56), the distributional equality

$$\partial_{\xi'}\chi^2(x',\xi',r)=\delta_0(\xi')-\delta_0(u^2(x',r)-\xi')\quad \text{for } (x',\xi',r)\in Q\times\mathbb{R}\times[0,\infty),$$

implies that

$$
\begin{aligned}
&-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\partial_{\xi'}\rho_{t_i,r}^{2,\epsilon}\,\chi_r^2\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr\\
&\quad=\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr\\
&\qquad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,u^2(x',r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr.
\end{aligned}
\tag{2.3.58}
$$

Hence, returning to (2.3.54), it follows from (2.3.56) and (2.3.58) that

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p_r^1+q_r^1)\,\partial_\xi\rho_{t_i,r}^{1,\epsilon}\,dx\,d\xi\right)\tilde{\chi}_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\quad=IE_i^{\mathrm{mix}1,2}+BE_i^{\mathrm{mix}1,3}\\
&\qquad+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr\\
&\qquad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,u^2(x',r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr.
\end{aligned}
\tag{2.3.59}
$$

Furthermore, after swapping the roles of χ^1 and χ^2, virtually identical arguments prove that

$$
\begin{aligned}
&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p_r^2+q_r^2)\,\partial_{\xi'}\rho_{t_i,r}^{2,\epsilon}\,dx'd\xi'\right)\tilde{\chi}_{t_i,r}^{1,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\quad=IE_i^{\mathrm{mix}2,2}+BE_i^{\mathrm{mix}2,3}\\
&\qquad+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^2+q_r^2)\,\rho_{t_i,r}^{2,\epsilon}\,\rho_{t_i,r}^{1,\epsilon}(x,y,0,\eta)\,\phi_\beta(y)\,dx\,dx'dy\,d\xi'd\eta\,dr\\
&\qquad-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^2+q_r^2)\,\rho_{t_i,r}^{2,\epsilon}\,\rho_{t_i,r}^{1,\epsilon}(x,y,u^1(x,r),\eta)\,\phi_\beta(y)\,dx\,dx'dy\,d\xi'd\eta\,dr,
\end{aligned}
\tag{2.3.60}
$$

for internal and boundary errors $IE_i^{\mathrm{mix}2,2}$ and $BE_i^{\mathrm{mix}2,3}$ defined in exact analogy to (2.3.57) and (2.3.55) respectively.

We finally go back to (2.3.44), to complete the analysis of the mixed term:

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}}\tilde{\chi}_{t_i,r}^{1,\epsilon}(y,\eta)\,\tilde{\chi}_{t_i,r}^{2,\epsilon}(y,\eta)\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}\\
&\quad=\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}m|\xi|^{m-1}\chi_r^1\,\Delta_x\rho_{t_i,r}^{1,\epsilon}\,dx\,d\xi\right)\tilde{\chi}_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\qquad-\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p_r^1+q_r^1)\,\partial_\xi\rho_{t_i,r}^{1,\epsilon}\,dx\,d\xi\right)\tilde{\chi}_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\qquad+\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}m|\xi'|^{m-1}\chi_r^2\,\Delta_{x'}\rho_{t_i,r}^{2,\epsilon}\,dx'd\xi'\right)\tilde{\chi}_{t_i,r}^{1,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr\\
&\qquad-\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\left(\int_{Q\times\mathbb{R}}(p_r^2+q_r^2)\,\partial_{\xi'}\rho_{t_i,r}^{2,\epsilon}\,dx'd\xi'\right)\tilde{\chi}_{t_i,r}^{1,\epsilon}\,\phi_\beta(y)\,dy\,d\eta\,dr.
\end{aligned}
$$

Using (2.3.52), (2.3.53), (2.3.59) and (2.3.60), and rearranging the terms gives

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}} \tilde{\chi}^{1,\epsilon}_{t_i,r}(y,\eta)\,\tilde{\chi}^{2,\epsilon}_{t_i,r}(y,\eta)\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&= \sum_{j=1}^{2} IE_i^{\text{mix}j,1} - IE_i^{\text{mix}j,2} + \sum_{j=1}^{2} BE_i^{\text{mix}j,1} + BE_i^{\text{mix}j,2} - BE_i^{\text{mix}j,3} \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\ \mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\ \mathrm{tr}\Big(\big(D_{x'}Y^{x',\xi'}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y) D_{x'}Y^{x',\xi'}_{r,r-t_i}\Big)\,dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad - \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad - \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}dx\,dx'dyd\xi d\xi'd\eta dr \qquad (2.3.61)\\
&\quad - \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m\big(|\xi|^{m-1}+|\xi'|^{m-1}\big)\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad - \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr \\
&\quad - \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^2_r+q^2_r)\,\rho^{2,\epsilon}_{t_i,r}\,\rho^{1,\epsilon}_{t_i,r}(x,y,0,\eta)\,\phi_\beta(y)\,dx\,dx'dy\,d\xi'd\eta\,dr \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}(x',y,u^2(x',r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^2_r+q^2_r)\,\rho^{2,\epsilon}_{t_i,r}\,\rho^{1,\epsilon}_{t_i,r}(x,y,u^1(x,r),\eta)\,\phi_\beta(y)\,dx\,dx'dy\,d\xi'd\eta\,dr.
\end{aligned}
$$

Step 4: Internal term cancellation. Using (2.3.42), (2.3.43) and (2.3.61) in (2.3.24) gives

$$
\begin{aligned}
&\int_{Q\times\mathbb{R}} \big(\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\mathrm{sgn}}^{\epsilon}_{t_i,r} + \tilde{\chi}^{2,\epsilon}_{t_i,r}\tilde{\mathrm{sgn}}^{\epsilon}_{t_i,r} - 2\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}\big)\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}} \\
&= \sum_{j=1}^{2} IE_i^{\text{sgn}j,1} - IE_i^{\text{sgn}j,2} - 2IE_i^{\text{mix}j,1} + 2IE_i^{\text{mix}j,2} \\
&\quad + \sum_{j=1}^{2} BE_i^{\text{sgn}j,1} + BE_i^{\text{sgn}j,2} - BE_i^{\text{sgn}j,3} - 2BE_i^{\text{mix}j,1} - 2BE_i^{\text{mix}j,2} + 2BE_i^{\text{mix}j,3} \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\big(\chi^1_r\mathrm{sgn}(\xi')-2\chi^1_r\chi^2_r\big)\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r} \\
&\qquad\qquad \cdot \mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y) D_xY^{x,\xi}_{r,r-t_i}\Big)dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\big(\chi^2_r\mathrm{sgn}(\xi)-2\chi^1_r\chi^2_r\big)\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r} \qquad (2.3.62)\\
&\qquad\qquad \cdot \mathrm{tr}\Big(\big(D_{x'}Y^{x',\xi'}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y) D_{x'}Y^{x',\xi'}_{r,r-t_i}\Big)dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad + 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\,\chi^2_r\,\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad + 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\,\nabla_y\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad + 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m\big(|\xi|^{m-1}+|\xi'|^{m-1}\big)\chi^1_r\,\chi^2_r\,\nabla_x\rho^{1,\epsilon}_{t_i,r}\,\nabla_{x'}\rho^{2,\epsilon}_{t_i,r}\,\phi_\beta(y)dx\,dx'dyd\xi d\xi'd\eta dr \\
&\quad - 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^1_r+q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}(x',y,u^2(x',r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr \\
&\quad - 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^2_r+q^2_r)\,\rho^{2,\epsilon}_{t_i,r}\,\rho^{1,\epsilon}_{t_i,r}(x,y,u^1(x,r),\eta)\,\phi_\beta(y)\,dx\,dx'dy\,d\xi'd\eta\,dr,
\end{aligned}
$$

where we point out that the terms involving the convolution kernels evaluated at $\xi = 0$ or $\xi' = 0$ in equations (2.3.42) and (2.3.43) and equation (2.3.61) cancel out with each other. The aim of this step is to observe an additional cancellation among the residual internal terms in (2.3.62). Namely, among the last three lines of (2.3.62) denoted by

$$\begin{aligned} IR_i :=& 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m\big(|\xi|^{m-1}+|\xi'|^{m-1}\big)\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\ &- 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}(x',y,u^2(x',r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr \\ &- 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^2+q_r^2)\,\rho_{t_i,r}^{2,\epsilon}\,\rho_{t_i,r}^{1,\epsilon}(x,y,u^1(x,r),\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi'\,d\eta\,dr. \end{aligned} \tag{2.3.63}$$

This cancellation is an effect of the integration by parts formula (2.2.27). The elementary equality

$$\Big(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\Big)^2+2|\xi|^{\frac{m-1}{2}}|\xi'|^{\frac{m-1}{2}}=|\xi|^{m-1}+|\xi'|^{m-1}\quad\text{for }\xi,\xi'\in\mathbb{R}$$

allows us to rewrite the first term of (2.3.63) as

$$\begin{aligned} &2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m\big(|\xi|^{m-1}+|\xi'|^{m-1}\big)\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\ &\quad= 4m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}|\xi|^{\frac{m-1}{2}}|\xi'|^{\frac{m-1}{2}}\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\ &\quad\quad+ 2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}\Big(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\Big)^2\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr. \end{aligned} \tag{2.3.64}$$

For the first term on the right-hand side of (2.3.64), after applying the integration by parts formula (2.2.27) both in the x variable and the x' variable, we have

$$\begin{aligned} &4m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}|\xi|^{\frac{m-1}{2}}|\xi'|^{\frac{m-1}{2}}\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\ &\quad=\frac{16m}{(m+1)^2}\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}}\nabla(u^1)^{[\frac{m+1}{2}]}\nabla(u^2)^{[\frac{m+1}{2}]}\bar\rho_{t_i,r}^{1,\epsilon}(x,y,\eta)\,\bar\rho_{t_i,r}^{2,\epsilon}(x',y,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\eta\,dr, \end{aligned}$$

where, for $j=1,2$ and $(x,y,\eta,r)\in Q^2\times\mathbb{R}\times[t_i,\infty)$, we denote

$$\bar\rho_{t_i,r}^{j,\epsilon}(x,y,\eta):=\rho_{t_i,r}^{\epsilon}(x,u^j(x,r),y,\eta). \tag{2.3.65}$$

In turn, Cauchy's inequality, the definition of the parabolic defect measure, and the nonnegativity of the entropy defect measure prove that

$$\begin{aligned} &4m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}|\xi|^{\frac{m-1}{2}}|\xi'|^{\frac{m-1}{2}}\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr \\ &\quad\leqslant 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^1+q_r^1)\,\rho_{t_i,r}^{1,\epsilon}(x,y,\xi,\eta)\,\bar\rho_{t_i,r}^{2,\epsilon}(x',y,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr \\ &\quad\quad+ 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}(p_r^2+q_r^2)\,\bar\rho_{t_i,r}^{1,\epsilon}(x,y,\eta)\,\rho_{t_i,r}^{2,\epsilon}(x',y,\xi',\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi'\,d\eta\,dr. \end{aligned} \tag{2.3.66}$$

Therefore, it follows from (2.3.64) and (2.3.66) that, for the internal residual term IR_i defined in (2.3.63),

$$\limsup_{\epsilon\to 0} IR_i \leqslant \limsup_{\epsilon\to 0} IE_i^{\text{canc}}, \tag{2.3.67}$$

for the internal cancellation error term

$$IE_i^{\text{canc}} := 2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}\Big(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\Big)^2\,\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)dxdx'dyd\xi d\xi' d\eta dr. \tag{2.3.68}$$

Step 5: Boundary term cancellation. The aim of this step is to observe a crucial cancellation between the residual boundary terms on the right-hand side of (2.3.62). Namely, between the third and fourth lines of (2.3.62), denoted by

$$\begin{aligned}BR_i^A =&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\big(\chi_r^1\text{sgn}(\xi')-2\chi_r^1\chi_r^2\big)\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}\ \text{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^TD_y^2\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\Big)\\ &+\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\big(\chi_r^2\text{sgn}(\xi)-2\chi_r^1\chi_r^2\big)\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}\ \text{tr}\Big(\big(D_{x'}Y^{x',\xi'}_{r,r-t_i}\big)^TD_y^2\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}\Big)\,,\end{aligned} \tag{2.3.69}$$

and the fifth and sixth lines of (2.3.62), denoted by

$$\begin{aligned}BR_i^B =&2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi_r^1\,\chi_r^2\,\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dxdx'dyd\xi d\xi'd\eta dr\\ &+2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi'|^{m-1}\chi_r^1\,\chi_r^2\,\nabla_x\rho_{t_i,r}^{1,\epsilon}\,\rho_{t_i,r}^{2,\epsilon}\,\nabla_y\phi_\beta(y)D_{x'}Y^{x',\xi'}_{r,r-t_i}dxdx'dyd\xi d\xi'd\eta dr.\end{aligned} \tag{2.3.70}$$

We begin by analyzing the first term of BR_i^B. First, we add and subtract $|\xi'|^{m-1}$ and then use the integration by parts formula (2.2.27) in the x' variable and formula (2.3.181) from Lemma 2.3.7 below to get

$$\begin{aligned}&2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi_r^1\,\chi_r^2\,\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dxdx'dyd\xi d\xi'd\eta dr\\ &=BE_i^{\text{canc}B1}-2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}\nabla\big(u^2\big)^{[m]}\,\chi_r^1(x,\xi)\rho_{t_i,r}^{1,\epsilon}\bar\rho_{t_i,r}^{2,\epsilon}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dxdx'dyd\xi d\eta dr,\end{aligned} \tag{2.3.71}$$

with the notation (2.3.65) for $\bar\rho_{t_i,r}^{2,\epsilon}$, for the boundary cancellation error term

$$\begin{aligned}BE_i^{\text{canc}B1}:=&2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m\left(|\xi|^{m-1}-|\xi'|^{m-1}\right)\\ &\cdot\chi_r^1\,\chi_r^2\,\rho_{t_i,r}^{1,\epsilon}\,\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dxdx'dyd\xi d\xi'd\eta dr.\end{aligned} \tag{2.3.72}$$

For the second term in (2.3.71), recalling the definition (2.3.19) of the convolution kernels $\rho_{t_i,r}^{j,\epsilon}$ and the inverse property (2.2.11) of characteristics, we observe

$$\begin{aligned}&\lim_{\epsilon\to 0}2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}\nabla\big(u^2\big)^{[m]}\,\chi_r^1(x,\xi)\rho_{t_i,r}^{1,\epsilon}(x,y,\xi,\eta)\bar\rho_{t_i,r}^{2,\epsilon}(x',y,\eta)\,\nabla_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}dxdx'dyd\xi d\eta dr\\ &\qquad=2\int_{t_i}^{t_{i+1}}\int_Q\nabla(u^2)^{[m]}\,\chi_r^1(x',u^2(x',r))\,\nabla_y\phi_\beta\big(Y^{x',u^2}_{r,r-t_i}\big)\,D_{x'}Y^{x',u^2}_{r,r-t_i}dx'dr,\end{aligned} \tag{2.3.73}$$

where we recall the convention (2.3.6). The analysis of the second line of (2.3.70) is virtually identical tp (2.3.71)-(2.3.73), simply swapping the roles of u^1 and u^2, and it produces a boundary error term $BE_i^{\mathrm{canc}B2}$ defined in exact analogy to (2.3.72). Then, using (2.3.71) and (2.3.73), and the analogous formulas for the second line of (2.3.70), for the boundary residue BR_i^B defined in (2.3.70) we conclude that

$$\begin{aligned}\limsup_{\epsilon\to 0} BR_i^B \leqslant\ & \limsup_{\epsilon\to 0}\ BE_i^{\mathrm{canc}B1}+BE_i^{\mathrm{canc}B2}\\ & -2\int_{t_i}^{t_{i+1}}\int_Q \nabla(u^2)^{[m]}\,\chi_r^1(x',u^2(x',r))\,\nabla_y\phi_\beta\big(Y^{x',u^2}_{r,r-t_i}\big)\,D_{x'}Y^{x',u^2}_{r,r-t_i}dx'\,dr \\ & -2\int_{t_i}^{t_{i+1}}\int_Q \nabla(u^1)^{[m]}\,\chi_r^2(x,u^1(x,r))\,\nabla_y\phi_\beta\big(Y^{x,u^1}_{r,r-t_i}\big)\,D_{x}Y^{x,u^1}_{r,r-t_i}dx\,dr.\end{aligned} \tag{2.3.74}$$

We now consider the residual term BR_i^A, defined in (2.3.69). First of all, we notice that, upon swapping the roles of the variables (x,ξ) and (x',ξ') in the second integral, this term is rewritten as

$$\begin{aligned}BR_i^A=&\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\big(\chi_r^1(x,\xi)\,\mathrm{sgn}(\xi')+\chi_r^2(x,\xi)\,\mathrm{sgn}(\xi')-2\chi_r^1(x,\xi)\chi_r^2(x',\xi')-2\chi_r^1(x',\xi')\chi_r^2(x,\xi)\big)\\ &\qquad\cdot\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}\ \mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D_y^2\phi_\beta(y) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,dx'\,dy\,d\xi\,d\xi'\,d\eta\,dr.\end{aligned}$$

Then, recalling the definition (2.3.19) of the convolution kernels and the inverse property (2.2.11) of the characteristics, and using properties of the kinetic function, we observe that

$$\begin{aligned}\lim_{\epsilon\to 0} BR_i^A =& \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\big(\chi_r^1(x,\xi)\mathrm{sgn}(\xi)+\chi_r^2(x,\xi)\mathrm{sgn}(\xi)-4\chi_r^1(x,\xi)\chi_r^2(x,\xi)\big)\\ &\qquad\cdot\mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D_y^2\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,d\xi\,dr\\ =& \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\,\big|\chi_r^1(x,\xi)-\chi_r^2(x,\xi)\big|^2\\ &\qquad\cdot\mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D_y^2\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,d\xi\,dr\\ &-2\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi_r^1(x,\xi)\chi_r^2(x,\xi)\\ &\qquad\cdot\mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D_y^2\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,d\xi\,dr.\end{aligned} \tag{2.3.75}$$

For the second term on the right-hand side of (2.3.75), performing an elementary computation, we obtain

$$\begin{aligned}&-2\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi_r^1(x,\xi)\chi_r^2(x,\xi)\ \mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D_y^2\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big) D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,d\xi\,dr\\ &= BE_i^{\mathrm{canc}A,1}-2\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi_r^1(x,\xi)\chi_r^2(x,\xi)\,\nabla_x\cdot\Big(\nabla_y\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big)D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,d\xi\,dr,\end{aligned} \tag{2.3.76}$$

for the boundary cancellation error term

$$BE_i^{\mathrm{canc}A,1} := 2\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi_r^1(x,\xi)\chi_r^2(x,\xi)\,\nabla_y\phi_\beta\big(Y^{x,\xi}_{r,r-t_i}\big)\Delta_xY^{x,\xi}_{r,r-t_i}dx\,d\xi\,dr. \tag{2.3.77}$$

Furthermore, for the second term in (2.3.76), using the integration by parts formula (2.2.27), Lemma 2.3.7 below and the product rule for derivatives, we have

$$\begin{aligned}
-2\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} & m|\xi|^{m-1}\chi_r^1(x,\xi)\chi_r^2(x,\xi)\,\nabla_x\cdot\Big(\nabla_y\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)D_xY_{r,r-t_i}^{x,\xi}\Big)\,dx\,d\xi\,dr \\
&=2\int_{t_i}^{t_{i+1}}\int_Q \nabla\left(u^1\right)^{[m]}\chi_r^2(x,u^1(x,r))\,\nabla_y\phi_\beta\big(Y_{r,r-t_i}^{x,u^1}\big)D_xY_{r,r-t_i}^{x,u^1}dx\,dr \\
&\quad+2\int_{t_i}^{t_{i+1}}\int_Q \nabla\left(u^2\right)^{[m]}\chi_r^1(x,u^2(x,r))\,\nabla_y\phi_\beta\big(Y_{r,r-t_i}^{x,u^2}\big)D_xY_{r,r-t_i}^{x,u^2}dx\,dr,
\end{aligned} \tag{2.3.78}$$

which is the exact opposite of the last two terms on the right-hand side of (2.3.74).

Finally, we analyze the first term on the right-hand side of (2.3.75). After adding and subtracting the identity matrix I_d to $D_xY_{r,r-t_i}^{x,\xi}$ twice, we obtain

$$\begin{aligned}
\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} & m|\xi|^{m-1}\left|\chi_r^1(x,\xi)-\chi_r^2(x,\xi)\right|^2\operatorname{tr}\Big(\big(D_xY_{r,r-t_i}^{x,\xi}\big)^TD_y^2\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)D_xY_{r,r-t_i}^{x,\xi}\Big)\,dx\,d\xi\,dr \\
&= BE_i^{\mathrm{canc}A,2}+\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1(x,\xi)-\chi_r^2(x,\xi)\right|^2\Delta_y\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)dx\,d\xi\,dr,
\end{aligned} \tag{2.3.79}$$

for the boundary error term

$$\begin{aligned}
BE_i^{\mathrm{canc}A,2}:=&\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2 \\
&\cdot\operatorname{tr}\Big(\big(D_xY_{r,r-t_i}^{x,\xi}-\mathrm{I}_d\big)^TD_y^2\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)D_xY_{r,r-t_i}^{x,\xi}+D_y^2\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)\big(D_xY_{r,r-t_i}^{x,\xi}-\mathrm{I}_d\big)\Big)dx\,d\xi\,dr.
\end{aligned} \tag{2.3.80}$$

Heuristically, the second term in (2.3.79) should be mostly negative as it consists of positive terms multiplied by the laplacian of a cutoff function, and this is constantly equal to 1 within the domain Q and bends downward near the boundary decreasing up to 0. Rigorously, recalling the explicit choice (2.3.14) of the cutoff, we use formulas (2.3.16)-(2.3.17) and the nonnegativity of the convolution kernel ρ_1 in definition (2.3.13) to compute

$$\begin{aligned}
\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} & m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\Delta_y\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)dx\,d\xi\,dr \\
&=-\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\dot\psi_\beta\big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)\big)\nabla_x\cdot\tilde n\big(Y_{r,r-t_i}^{x,\xi}\big)dx\,d\xi\,dr \\
&\quad+\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\ddot\psi_\beta\big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)\big)dx\,d\xi\,dr \\
&= BE_i^{\mathrm{canc}A,3} \\
&\quad+\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2 \\
&\qquad\cdot\left(\frac{1}{\beta}\,\rho_1^{\frac12\beta^{\gamma_m}}\Big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)-\beta^{\gamma_m}\Big)-\frac{1}{\beta}\,\rho_1^{\frac12\beta^{\gamma_m}}\Big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)-(\beta+\beta^{\gamma_m})\Big)\right)dx\,d\xi\,dr \\
&\leqslant BE_i^{\mathrm{canc}A,3}+BE_i^{\mathrm{canc}A,4},
\end{aligned} \tag{2.3.81}$$

for the boundary error terms

$$BE_i^{\mathrm{canc}A,3}:=-\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\ \dot\psi_\beta\big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)\big)\,\nabla_x\cdot\tilde n\big(Y_{r,r-t_i}^{x,\xi}\big)\,dx\,d\xi\,dr, \tag{2.3.82}$$

and

$$BE_i^{\mathrm{canc}A,4}:=\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\ \frac{1}{\beta}\,\rho_1^{\frac12\beta^{\gamma_m}}\Big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)-\beta^{\gamma_m}\Big)\,dx\,d\xi\,dr. \tag{2.3.83}$$

In conclusion, it follows from (2.3.74), (2.3.75), (2.3.76), (2.3.78), (2.3.79) and (2.3.81) that, for the residual boundary terms BR_i^A and BR_i^B defined in (2.3.69) and (2.3.70) respectively,

$$\limsup_{\epsilon\to 0}\left(BR_i^A + BR_i^B\right) \leqslant \limsup_{\epsilon\to 0}\left(\sum_{k=1}^{4} BE_i^{\mathrm{canc}A,k} + \sum_{j=1}^{2} BE_i^{\mathrm{canc}Bj}\right). \tag{2.3.84}$$

Step 6: The final inequality. We now go back go (2.3.62). Using (2.3.63) and (2.3.67), and (2.3.69) (2.3.70) and (2.3.84), we obtain that

$$\begin{aligned}
&\limsup_{\epsilon\to 0}\int_{Q\times\mathbb{R}}\left(\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\mathrm{sgn}}^{\epsilon}_{t_i,r} + \tilde{\chi}^{2,\epsilon}_{t_i,r}\tilde{\mathrm{sgn}}^{\epsilon}_{t_i,r} - 2\tilde{\chi}^{1,\epsilon}_{t_i,r}\tilde{\chi}^{2,\epsilon}_{t_i,r}\right)\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}\\
&\leqslant \limsup_{\epsilon\to 0}\Bigg(\sum_{j=1}^{2} IE_i^{\mathrm{sgn}j,1} - IE_i^{\mathrm{sgn}j,2} - 2IE_i^{\mathrm{mix}j,1} + 2IE_i^{\mathrm{mix}j,2}\\
&\qquad + IE_i^{\mathrm{canc}}\\
&\qquad + \sum_{j=1}^{2} BE_i^{\mathrm{sgn}j,1} + BE_i^{\mathrm{sgn}j,2} - BE_i^{\mathrm{sgn}j,3} - 2BE_i^{\mathrm{mix}j,1} - 2BE_i^{\mathrm{mix}j,2} + 2BE_i^{\mathrm{mix}j,3}\\
&\qquad + \sum_{k=1}^{4} BE_i^{\mathrm{canc}A,k} + \sum_{j=1}^{2} BE_i^{\mathrm{canc}Bj}\Bigg).
\end{aligned}$$

In turn, from this and (2.3.23) it follows that

$$\begin{aligned}
&\int_Q |u^1(y,r) - u^2(y,r)|\,dy\Big|_{r=0}^{r=T}\\
&\leqslant \limsup_{\beta\to 0}\limsup_{\epsilon\to 0}\sum_{i=0}^{N-1}\Bigg(\sum_{j=1}^{2} IE_i^{\mathrm{sgn}j,1} - IE_i^{\mathrm{sgn}j,2} - 2IE_i^{\mathrm{mix}j,1} + 2IE_i^{\mathrm{mix}j,2}\\
&\qquad + IE_i^{\mathrm{canc}}\\
&\qquad + \sum_{j=1}^{2} BE_i^{\mathrm{sgn}j,1} + BE_i^{\mathrm{sgn}j,2} - BE_i^{\mathrm{sgn}j,3} - 2BE_i^{\mathrm{mix}j,1} - 2BE_i^{\mathrm{mix}j,2} + 2BE_i^{\mathrm{mix}j,3}\\
&\qquad + \sum_{k=1}^{4} BE_i^{\mathrm{canc}A,k} + \sum_{j=1}^{2} BE_i^{\mathrm{canc}Bj}\Bigg),
\end{aligned} \tag{2.3.85}$$

for the internal, displacement and boundary error terms defined in the previous steps, and where we recall that $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = T\} \subseteq [0,T]\backslash\mathcal{N}$ is an arbitrary partition. Incidentally, we mention that, according to the relevant definitions, some of these error terms are actually independent of $\epsilon \in (0,1)$. The aim of the next two steps is to provide estimates for the internal and boundary errors respectively. These estimates shall depend on $\epsilon, \beta \in (0,1)$ and on the size $|\mathcal{P}|$ of the arbitrary partition, in such a way that, as we let $\epsilon \to 0$ first, $\beta \to 0$ then, and finally $|\mathcal{P}| \to 0$, the right-hand side of (2.3.85) vanishes, thus proving the theorem.

Step 7: The internal errors. In this step we analyze the internal error terms in (2.3.85). We begin with the internal errors from Step 2 and Step 3. These are easily handled using the stability estimates on the characteristics from Section 2.5 and the next crucial observation, which follows immediately from the definition (2.3.19) of the convolution kernels. For every (x,ξ), (x',ξ') and $(y,\eta)\in Q\times\mathbb{R}$ and for every $t_i\in\mathcal{P}$ and $r\in[t_i,\infty)$,

$$\text{if } \rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}(x',y,\xi',\eta)\neq 0, \text{ then } \left|Y^{x,\xi}_{r,r-t_i}-Y^{x',\xi'}_{r,r-t_i}\right|+\left|\Pi^{x,\xi}_{r,r-t_i}-\Pi^{x',\xi'}_{r,r-t_i}\right|\leqslant 2\epsilon. \quad (2.3.86)$$

For the internal error $IE_i^{\text{sgn1},1}$ defined in (2.3.30), using the integration by parts formula (2.2.27) and Lemma 2.3.7 we obtain, with the notation (2.3.65),

$$\begin{aligned} IE_i^{\text{sgn1},1} = &-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}\nabla\left(u^1\right)^{[m]}\bar{\rho}^{1,\epsilon}_{t_i,r}\text{sgn}(\xi')\phi_\beta(y)\,\nabla_y\rho^{2,\epsilon}_{t_i,r}\left(D_xY^{x,u^1}_{r,r-t_i}-D_{x'}Y^{x',\xi'}_{r,r-t_i}\right) \\ &-\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2}\nabla\left(u^1\right)^{[m]}\bar{\rho}^{1,\epsilon}_{t_i,r}\text{sgn}(\xi')\phi_\beta(y)\,\partial_\eta\rho^{2,\epsilon}_{t_i,r}\left(D_x\Pi^{x,u^1}_{r,r-t_i}-D_{x'}\Pi^{x',\xi'}_{r,r-t_i}\right). \end{aligned}$$

The error term $IE_i^{\text{mix1},1}$ defined in (2.3.64) is transformed identically, integrating by parts and using Lemma 2.3.7, and the specular terms $IE_i^{\text{sgn2},1}$ and $IE_i^{\text{mix2},1}$ are handled in exact analogy, swapping the roles of χ^1 and χ^2. Since the functions ϕ_β, sgn and χ^j are bounded, and since there exists a constant C, only depending on the the standard mollifier ρ chosen in Step 0, such that

$$\int_{Q\times\mathbb{R}}|\nabla_y\rho^{j,\epsilon}_{t_i,r}(x',y,\xi',\eta)|+|\partial_\eta\rho^{j,\epsilon}_{t_i,r}(x',y,\xi',\eta)|\,dx'\,d\xi'\leqslant C\epsilon^{-1}, \quad (2.3.87)$$

integrating the convolution kernels over the variables (y,η) and (x',ξ'), using (2.3.86) combined with the estimates (2.5.6), we get

$$|IE_i^{\text{sgn}j,1}|+|IE_i^{\text{mix}j,1}|\leqslant C|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_Q\left|\nabla\left(u^j\right)^{[m]}\right|\,dx\,dr, \quad (2.3.88)$$

for $j=1,2$, for a constant $C=C(T,A,z)$.

The remaining error terms $IE_i^{\text{sgn1},2}$, defined in (2.3.40), and $IE_i^{\text{mix1},2}$, defined in (2.3.57), and the specular terms $IE_i^{\text{sgn2},2}$ and $IE_i^{\text{mix2},2}$, are treated in analogy to (2.3.88). Indeed, we use the boundedness of ϕ_β, sgn and χ^j, formula (2.3.87), and formula (2.3.86) combined with (2.5.6), and integrate the convolution kernels to estimate

$$|IE_i^{\text{sgn}j,2}|+|IE_i^{\text{mix}j,2}|\leqslant C|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}p^j_r(x,\xi)+q^j_r(x,\xi)\;\,dx\,d\xi\,dr, \quad (2.3.89)$$

for $j=1,2$, for a constant $C=C(T,A,z)$. Finally, using (2.3.88) and (2.3.89), and summing over the partition, we obtain the following estimate, for $C=C(T,A,z)$,

$$\begin{aligned} &\limsup_{\beta\to0}\limsup_{\epsilon\to0}\sum_{i=0}^{N-1}\sum_{j=1}^{2}\sum_{k=1}^{2}|IE_i^{\text{sgn}j,k}|+|IE_i^{\text{mix}j,k}| \\ &\qquad\leqslant C|\mathcal{P}|^\alpha\sum_{j=1}^{2}\left(\int_0^T\int_Q\left|\nabla\left(u^j\right)^{[m]}\right|\,dx\,dr+\int_0^T\int_{Q\times\mathbb{R}}p^j_r+q^j_r\,dx\,d\xi\,dr\right). \end{aligned} \quad (2.3.90)$$

We now consider the error term coming from the internal cancellation in Step 4, namely the internal error IE_i^{canc} defined in (2.3.68). The analysis is broken down in three cases: $m=1$, $m\in(2,\infty)$ and $m\in(0,1)\cup(1,2]$.

Case $m = 1$. This case is trivial. Indeed, if $m = 1$, it follows automatically from definition (2.3.68) that $IE_i^{\text{canc}} = 0$.

Case $m \in (2, \infty)$. We form a velocity decomposition of the integral. For each $M > 1$, let $K_M : \mathbb{R} \to [0, 1]$ be a smooth function satisfying

$$K_M(\xi) = \begin{cases} 1 & \text{if } |\xi| \leqslant M, \\ 0 & \text{if } |\xi| \leqslant M+1. \end{cases}$$

Then, for each $M > 1$, we split

$$\begin{aligned} IE_i^{\text{canc}} = 2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(1-K_M(\xi))\left(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\right)^2\chi_r^1\chi_r^2\nabla_x\rho_{t_i,r}^{1,\epsilon}\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\,\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr \\ +2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}K_M(\xi)\left(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\right)^2\chi_r^1\chi_r^2\nabla_x\rho_{t_i,r}^{1,\epsilon}\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr \end{aligned} \tag{2.3.91}$$

For the first term on the right-hand side of (2.3.91) we write

$$\begin{aligned} &\left|2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}K_M(\xi)\left(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\right)^2\chi_r^1\chi_r^2\nabla_x\rho_{t_i,r}^{1,\epsilon}\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr\right| \\ &\qquad\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3\cap\{|\xi-\xi'|\leqslant c\epsilon\}}|\xi-\xi'|^{(m-1)\wedge 2}\left|\nabla_x\rho_{t_i,r}^{1,\epsilon}\right|\left|\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\right|dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr \\ &\qquad\leqslant C\epsilon^{-2}\int_{t_i}^{t_{i+1}}\int_Q\int_{-c\epsilon}^{c\epsilon}|\theta|^{(m-1)\wedge 2}d\theta\,dy\,dr \leqslant C|t_{i+1}-t_i|\epsilon^{(3\wedge m)-2}, \end{aligned} \tag{2.3.92}$$

for a constant $C = C(M, m, Q, T, A, z)$. In the first passage we used the boundedness of χ^j, observation (2.3.86) combined with estimate (2.5.4), and the local Lipschitz continuity, if $m \geqslant 3$, or the Hölder continuity, if $m \in (2, 3)$, of the map $\mathbb{R} \ni \xi \mapsto |\xi|^{\frac{m-1}{2}}$. In the second passage we exploited formula (2.3.26) for the derivatives of the convolution kernels, the boundedness of the derivatives of the characteristics from Proposition 2.5.1 and formula (2.3.87), and we changed variables by setting $\theta = \xi - \xi'$. In the last passage we simply used the boundedness of the domain Q.

For the second term on the right-hand side of (2.3.91) we use the following elementary inequality

$$\left(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\right)^2 = \left|\int_{\xi'}^{\xi}\frac{m-1}{2}\theta^{\left[\frac{m-3}{2}\right]}d\theta\right| \leqslant \left|\frac{m-1}{2}\right|^2\left(|\xi|^{m-3}+|\xi'|^{m-3}\right)|\xi-\xi'|^2. \tag{2.3.93}$$

Then we estimate

$$\begin{aligned} &\left|2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(1-K_M(\xi))\left(|\xi|^{\frac{m-1}{2}}-|\xi'|^{\frac{m-1}{2}}\right)^2\chi_r^1\chi_r^2\nabla_x\rho_{t_i,r}^{1,\epsilon}\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\phi_\beta(y)\,dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr\right| \\ &\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3\cap\{|\xi-\xi'|\leqslant c\epsilon\}}(1-K_M(\xi))\left(|\xi|^{m-3}+|\xi'|^{m-3}\right)\epsilon^2\left|\chi_r^1\right|\left|\chi_r^2\right|\left|\nabla_x\rho_{t_i,r}^{1,\epsilon}\right|\left|\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\right|dx\,dx'dy\,d\xi\,d\xi'd\eta\,dr \\ &\leqslant C\left(\int_{t_i}^{t_{i+1}}\int_{Q\cap\{|u^1|\geqslant M\}}\left|u^1\right|^{m-2}dx\,dr+\int_{t_i}^{t_{i+1}}\int_{Q\cap\{|u^2|\geqslant M-c\epsilon\}}\left|u^2\right|^{m-2}dx'dr\right), \end{aligned} \tag{2.3.94}$$

for constants $C = C(m, T, A, z)$ and $c = c(T, A, z)$, independent of $M \geqslant 1$. In the first passage we used (2.3.93), and (2.3.86) combined with (2.5.4). In the second passage we exploited properties of the kinetic function, and formula (2.3.26) combined with Proposition 2.5.1 and formula (2.3.87).

Finally, combining (2.3.91) with (2.3.92) and (2.3.94), and summing over the partition $\mathcal{P}$, we obtain

$$\sum_{i=0}^{N-1} |IE_i^{\text{canc}}| \leqslant C_1\, \epsilon^{(3\wedge m)-2} + C_2 \left(\int_0^T\!\!\int_{Q\cap\{|u^1|\geqslant M\}} \left|u^1\right|^{m-2} dx\, dr + \int_0^T\!\!\int_{Q\cap\{|u^2|\geqslant M-c\epsilon\}} \left|u^2\right|^{m-2} dx' dr \right), \tag{2.3.95}$$

for constants $C_1 = C_1(M, m, Q, T, A, z)$, $C_2 = C_2(m, T, A, z)$ and $c = c(T, A, z)$. Since $m \in (2, \infty)$ and by Definition 2.2.4 of kinetic solution we have $u^j \in L^{m+1}([0, T]; L^{m+1}(Q))$, and since the constant C_2 is independent of M, Hölder's inequality and the dominated convergence theorem prove that the last two terms on the right-hand side of (2.3.95) vanish in the limit $M \to \infty$, uniformly for $\epsilon \in (0, 1)$. Therefore, passing first to the limit $\epsilon \to 0$ and second to the limit $M \to \infty$, formula (2.3.95) yields

$$\limsup_{\epsilon\to 0} \sum_{i=0}^{N-1} |IE_i^{\text{canc}}| = 0. \tag{2.3.96}$$

Case $m \in (0, 1) \cup (1, 2]$. For this case the idea is to remove the singularity at the origin and to exploit the full regularity of the solution implied by Proposition 2.3.9 below. Using the integration by parts formula (2.2.27), which is justified exploiting an approximation argument and Proposition 2.3.6 below, both in the (x, ξ) and the (x', ξ') variables, we write

$$\begin{aligned} IE_i^{\text{canc}} = \frac{4m}{(m+1)^2} \int_{t_i}^{t_{i+1}}\!\!\int_{Q^3\times\mathbb{R}} &\sigma_m(u^1, u^2) \left|u^1\right|^{-1/2} \nabla \left(u^1\right)^{\left[\frac{m+1}{2}\right]} \left|u^2\right|^{-1/2} \nabla \left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\cdot \bar{\rho}^{1,\epsilon}_{t_i,r}(x, y, \eta) \bar{\rho}^{2,\epsilon}_{t_i,r}(x', y, \eta) \phi_\beta(y)\, dx\, dx'\, dy\, d\eta\, dr, \end{aligned} \tag{2.3.97}$$

where we have defined

$$\sigma_m(\xi, \xi') := |\xi|^{\frac{2-m}{2}} |\xi'|^{\frac{2-m}{2}} \left(|\xi|^{\frac{m-1}{2}} - |\xi|^{\frac{m-1}{2}} \right)^2 \quad \text{for } \xi, \xi' \in \mathbb{R}. \tag{2.3.98}$$

We notice that definition (2.3.65), observation (2.3.86) and estimate (2.5.4) imply that,

$$\text{if } \bar{\rho}^{1,\epsilon}_{t_i,r}(x, y, \eta) \bar{\rho}^{2,\epsilon}_{t_i,r}(x', y, \eta) \neq 0, \text{ then } \left|x - x'\right| + \left|u^1 - u^2\right| \leqslant c_1 \epsilon, \tag{2.3.99}$$

for a constant $c_1 = c_1(T, A, z)$. Moreover, we observe that

$$\text{if } |\xi - \xi'| \leqslant c_1\epsilon, \text{ then } \left|\sigma_m(\xi, \xi')\right| \leqslant C\, \epsilon, \tag{2.3.100}$$

for a constant $C = C(c_1, m)$. Indeed,if $\max\{|\xi|, |\xi'|\} \leqslant 2c_1\epsilon$, then, recalling $m \in (0, 1) \cup (1, 2]$, a direct computation yields

$$\sigma_m(\xi, \xi') = |\xi|^{\frac{m}{2}} |\xi'|^{\frac{2-m}{2}} + 2|\xi|^{\frac{1}{2}} |\xi'|^{\frac{1}{2}} + |\xi|^{\frac{2-m}{2}} |\xi'|^{\frac{m}{2}} \leqslant C\, \epsilon.$$

Conversely, assume without loss of generality that $|\xi| > 2c_1\epsilon$ with $|\xi| \geqslant |\xi'|$ and $|\xi - \xi'| \leqslant c_1\epsilon$. Thus, in particular, ξ and ξ' have the same sign and $|\xi'| \geqslant \frac{1}{2}|\xi|$. Then we compute

$$\sigma_m(\xi,\xi') = |\xi|^{\frac{2-m}{2}}|\xi'|^{\frac{2-m}{2}}\left(\int_{\xi'}^{\xi}\frac{m-1}{2}\theta^{\left[\frac{m-3}{2}\right]}d\theta\right)^2 \leqslant C|\xi|^{-1}\epsilon^2 \leqslant C\epsilon.$$

We now form a velocity decomposition of the integral. For each $\delta \in (0,1)$, let $K^\delta : \mathbb{R} \to [0,1]$ denote a smooth cutoff function satisfying

$$K^\delta(\xi) = \begin{cases} 1 & \text{if } |\xi| \leqslant \delta \text{ or } \dfrac{2}{\delta} \leqslant |\xi|, \\ 0 & \text{if } 2\delta \leqslant |\xi| \leqslant \dfrac{1}{\delta}. \end{cases} \tag{2.3.101}$$

Returning to (2.3.97), consider the decomposition

$$\begin{aligned} IE_i^{\text{canc}} = &\frac{4m}{(m+1)^2}\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}} \sigma_m^\delta(u^1,u^2)\,|u^1|^{-1/2}\,\nabla\left(u^1\right)^{\left[\frac{m+1}{2}\right]}|u^2|^{-1/2}\,\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\qquad \cdot \bar{\rho}^{1,\epsilon}_{t_i,r}(x,y,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\phi_\beta(y)\,dx\,dx'\,dy\,d\eta\,dr \\ &+ \frac{4m}{(m+1)^2}\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}} \tilde{\sigma}_m^\delta(u^1,u^2)\,|u^1|^{-1/2}\,\nabla\left(u^1\right)^{\left[\frac{m+1}{2}\right]}|u^2|^{-1/2}\,\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\qquad \cdot \bar{\rho}^{1,\epsilon}_{t_i,r}(x,y,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\phi_\beta(y)\,dx\,dx'\,dy\,d\eta\,dr, \end{aligned} \tag{2.3.102}$$

where, for each $\delta \in (0,1)$, the functions $\sigma_m^\delta, \tilde{\sigma}_m^\delta : \mathbb{R}^2 \to \mathbb{R}$ are defined by

$$\sigma_m^\delta(\xi,\xi') = \left(K^\delta(\xi) + K^\delta(\xi') - K^\delta(\xi)K^\delta(\xi')\right)\sigma_m(\xi,\xi'), \tag{2.3.103}$$

and

$$\tilde{\sigma}_m^\delta(\xi,\xi') = \left(1 - K^\delta(\xi)\right)\left(1 - K^\delta(\xi')\right)\sigma_m(\xi,\xi').$$

We start with the second term in (2.3.102). It follows from (2.3.98), (2.3.101) and the local Lipschitz continuity of the map $\mathbb{R} \ni \xi \mapsto |\xi|^{\frac{m-1}{2}}$ on the set $\{\delta \leqslant |\xi| \leqslant \frac{2}{\delta}\}$ that, for $C = C(m,\delta)$,

$$\left|\tilde{\sigma}_m^\delta(\xi,\xi')\right| \leqslant C\,|\xi - \xi'|^2. \tag{2.3.104}$$

Moreover, recalling observation (2.3.99) and the definition (2.3.19) and (2.3.65) of the kernels $\bar{\rho}^{j,\epsilon}_{t_i,r}$, we point out that

$$\int_{Q^2\times\mathbb{R}} \bar{\rho}^{1,\epsilon}_{t_i,r}(x,y,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\,dx'dy\,d\eta = \int_{\{x'\in Q|\,|x-x'|\leqslant c_1\epsilon\}\times Q\times\mathbb{R}} \bar{\rho}^{1,\epsilon}_{t_i,r}(x,y,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\,dx'dy\,d\eta \leqslant C\,\epsilon^{-1}, \tag{2.3.105}$$

for a constant $C = C(d,c_1)$ depending on the constant $c_1 = c_1(T,A,z)$ from (2.3.99). Then, for the second term of (2.3.102), we compute, for a constant $C = C(\delta, m, d, T, A, z)$,

$$
\begin{aligned}
&\left|\frac{4m}{(m+1)^2}\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}}\tilde{\sigma}^{\delta}_m(u^1,u^2)\left|u^1\right|^{-1/2}\nabla(u^1)^{\left[\frac{m+1}{2}\right]}\left|u^2\right|^{-1/2}\nabla(u^2)^{\left[\frac{m+1}{2}\right]}\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta\,dx\,dx'dy\,d\eta\,dr\right| \\
&\quad\leqslant C\epsilon^2\left(\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}}\left|u^1\right|^{-1}\left|\nabla\left(u^1\right)^{\left[\frac{m+1}{2}\right]}\right|^2\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta(y)\,dx\,dx'dy\,d\eta\,dr\right)^{\frac{1}{2}} \\
&\qquad\cdot\left(\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}}\left|u^2\right|^{-1}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|^2\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta(y)\,dx\,dx'dy\,d\eta\,dr\right)^{\frac{1}{2}} \qquad (2.3.106)\\
&\quad\leqslant C\epsilon\left(\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}|\xi|^{-1}q^1_r(x,\xi)\,dx\,d\xi\,dr\right)^{\frac{1}{2}}\left(\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}|\xi'|^{-1}q^2_r(x',\xi')\,dx'd\xi'dr\right)^{\frac{1}{2}} \\
&\quad\leqslant C\,\epsilon\sum_{j=1}^{2}\left(1+\|u^j_0\|^2_{L^2(Q)}\right).
\end{aligned}
$$

In the first passage we used Hölder's inequality and (2.3.99) and (2.3.104). In the second passage we exploited (2.3.105) and the definition of parabolic defect measure. The last passage follows from Proposition 2.3.9 below.

We now consider to the first term in (2.3.102). It follows immediately from (2.3.100), (2.3.101) and (2.3.103) that

$$
\text{if } |\xi-\xi'|\leqslant c_1\epsilon,\ \text{ then } \left|\sigma^{\delta}_m(\xi,\xi')\right|\leqslant C\,\epsilon, \qquad (2.3.107)
$$

for $C = C(m, c_1)$, and that, for $\epsilon < \delta \in (0,1)$, with ϵ small enough depending on c_1 and δ,

$$
\text{if } |\xi-\xi'|\leqslant c_1\epsilon \ \text{ and } \sigma^{\delta}_m(\xi,\xi')\neq 0,\ \text{ then } 0<|\xi|,|\xi'|<3\delta \ \text{ or } \ \frac{1}{2\delta}<|\xi|,|\xi'|. \qquad (2.3.108)
$$

Then, for the first term of (2.3.102), first using Hölder's inequality and observations (2.3.99),(2.3.107) and (2.3.108), and then using observation (2.3.105) and the definition of parabolic defect measure, we compute

$$
\begin{aligned}
&\left|\frac{4m}{(m+1)^2}\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}}\sigma^{\delta}_m(u^1,u^2)\left|u^1\right|^{-1/2}\nabla(u^1)^{\left[\frac{m+1}{2}\right]}\left|u^2\right|^{-1/2}\nabla(u^2)^{\left[\frac{m+1}{2}\right]}\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta\,dx\,dx'dy\,d\eta\,dr\right| \\
&\quad\leqslant C\epsilon\left(\int_{t_i}^{t_{i+1}}\int_{U^1_\delta\times U^2_\delta\times Q\times\mathbb{R}}\left|u^1\right|^{-1}\left|\nabla\left(u^1\right)^{\left[\frac{m+1}{2}\right]}\right|^2\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta(y)\,dx\,dx'dy\,d\eta\,dr\right)^{\frac{1}{2}} \\
&\qquad\cdot\left(\int_{t_i}^{t_{i+1}}\int_{U^1_\delta\times U^2_\delta\times Q\times\mathbb{R}}\left|u^2\right|^{-1}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|^2\bar{\rho}^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\phi_\beta(y)\,dx\,dx'dy\,d\eta\,dr\right)^{\frac{1}{2}} \qquad (2.3.109)\\
&\quad\leqslant C\sum_{j=1}^{2}\left(\int_{t_i}^{t_{i+1}}\int_{U^j_\delta\times\mathbb{R}}|\xi|^{-1}q^j_r(x,\xi)\,dx\,d\xi\,dr\right),
\end{aligned}
$$

for a constant $C = C(m, d, T, A, z)$, where we have defined, for $r \in [0,\infty)$,

$$
U^j_\delta = U^j_\delta(r) := \left\{x\in Q\ \middle|\ 0<|u^j(x,r)|<3\delta \text{ or } \frac{1}{2\delta}<|u^j(x,r)|\right\}. \qquad (2.3.110)
$$

In conclusion, summing over the partition $\mathcal{P}$, the splitting (2.3.102) and estimates (2.3.106) and (2.3.109) imply that, for each $\delta \in (0,1)$, for $C = C(m, d, T, A, z)$,

$$
\limsup_{\epsilon\to 0}\sum_{i=0}^{N-1}|IE^{\text{canc}}_i|\leqslant C\sum_{j=1}^{2}\int_0^T\int_{U^j_\delta\times\mathbb{R}}|\xi|^{-1}q^j_r(x,\xi)\,dx\,d\xi\,dr. \qquad (2.3.111)
$$

The dominated convergence theorem, Proposition 2.3.9 below and (2.3.110) imply that the right-hand side of (2.3.111) vanishes in the $\delta \to 0$ limit. Therefore, we deduce

$$\limsup_{\epsilon\to 0} \sum_{i=0}^{N-1} |IE_i^{\text{canc}}| = 0. \tag{2.3.112}$$

This concludes the analysis of the internal error terms.

Step 8: The boundary errors. In this step we analyze the boundary error terms in (2.3.85). We begin with the errors produced in Step 2 and Step 3. These are handled using the estimates on the behaviour of the characteristics near the boundary from Section 2.5, properties (2.3.15) of the gradient $\nabla\phi_\beta$, observation (2.3.86), and the following crucial fact. Namely, with the notation (2.3.9), it follows from (2.3.11), (2.3.15) and (2.3.19) that, for any $\epsilon < \beta \in (0,1)$, for any $(x,y,\xi,\eta) \in Q^2 \times \mathbb{R}^2$ and any $t_i \leqslant r \in [0,T]$, for $j = 1,2$,

$$\text{if } \nabla_y\phi_\beta(y)\rho^{j,\epsilon}_{t_i,r}(x,y,\xi,\eta) \neq 0, \text{ then } y, Y^{x,\xi}_{r,r-t_i}, x \in Q^\beta. \tag{2.3.113}$$

For the boundary error term $BE_i^{\text{sgn}1,1}$ defined in (2.3.33), for a constant $C = C(T,A,z,Q)$, we compute

$$\begin{aligned}
\left|BE_i^{\text{sgn}1,1}\right| &\leqslant \int_{t_i}^{t_{i+1}}\int_{Q^2\times\mathbb{R}^2} m|\xi|^{m-1}\left|\chi^1_r\right|\left|\rho^{1,\epsilon}_{t_i,r}\right|\left|\tilde{\text{sgn}}^\epsilon_{t_i,r}\right|\left|\nabla_y\phi_\beta\right|\left|\Delta_x Y^{x,\xi}_{r,r-t_i}\right|\,dx\,dy\,d\xi\,d\eta\,dr \\
&\leqslant C|t_{i+1}-t_i|^\alpha\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta\times\mathbb{R}} m|\xi|^{m-1}\left|\chi^1_r\right|\,dx\,d\xi\,dr \\
&\leqslant C|t_{i+1}-t_i|^\alpha\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\left(u^1\right)^{[m]}\right|\,dx\,dr \\
&\leqslant C|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|\,dx\,dr.
\end{aligned}$$

In the second inequality we used observation (2.3.113), estimates (2.3.15) and (2.5.5), and the boundedness of $\tilde{\text{sgn}}^\epsilon$, and we integrated the mollifier. In the third passage we simply used properties of the kinetic function. In the last passage we exploited the Sobolev regularity $(u^1)^{[m]} \in W^{1,p_m}_0(Q)$ from Lemma 2.3.7 below, which in particular implies that $(u^1)^{[m]}$ vanishes on the boundary, and we used the mean value theorem applied to points $x \in Q^\beta$, which therefore satisfy $\text{dist}(x,\partial Q) \leqslant C\beta$.

An identical analysis holds for $BE_i^{\text{sgn}2,1}$, simply replacing χ^1 and u^1 with χ^2 and u^2, and for the analogous error terms $BE_i^{\text{mix}j,1}$ defined in (2.3.49), for $j = 1,2$, simply replacing $\tilde{\text{sgn}}^\epsilon_{t_i,r}$ with $\tilde{\chi}^{\cdot,\epsilon}_{t_i,r}$, which is bounded as well. Precisely, we have

$$\left|BE_i^{\text{sgn}j,1}\right| + \left|BE_i^{\text{mix}j,1}\right| \leqslant C|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\nabla\left(u^j\right)^{[m]}\right|\,dx\,dr, \tag{2.3.114}$$

for $j = 1,2$, for a constant $C = C(T,A,z,Q)$.

Next we analyze the boundary error terms $BE_i^{\text{sgn}1,2}$ and $BE_i^{\text{mix}1,2}$, defined in (2.3.35) and (2.3.51) respectively, and the specular terms $BE_i^{\text{sgn}2,2}$ and $BE_i^{\text{mix}2,2}$. Observations (2.3.86) and (2.3.113) combined with estimate (2.5.5) imply that, for $C = C(T, A, z)$,

$$\text{if } \nabla_y \phi_\beta\, \rho^{1,\epsilon}_{t_i,r}\, \rho^{2,\epsilon}_{t_i,r} \neq 0, \text{ then } |D_x Y^{x,\xi}_{r,r-t_i} - D_{x'} Y^{x',\xi'}_{r,r-t_i}| + |\nabla_x \Pi^{x,\xi}_{r,r-t_i} - \nabla_{x'} \Pi^{x',\xi'}_{r,r-t_i}| \leqslant C|t_{i+1} - t_i|^\alpha \epsilon. \quad (2.3.115)$$

Then, for $BE^{\text{sgn}1,2}_i$, for a constant $C = C(T, A, z, Q)$, we compute

$$\begin{aligned}\left|BE^{\text{sgn}1,2}_i\right| &\leqslant C\,|t_{i+1} - t_i|^\alpha \epsilon \beta^{-1} \int_{t_i}^{t_{i+1}} \int_{(Q^\beta)^3 \times \mathbb{R}^3} m|\xi|^{m-1}\,|\chi^1_r|\,\left|\rho^{1,\epsilon}_{t_i,r}\right| \left(\left|\nabla_y \rho^{2,\epsilon}_{t_i,r}\right| + \left|\partial_\eta \rho^{2,\epsilon}_{t_i,r}\right|\right) \\ &\leqslant C\,|t_{i+1} - t_i|^\alpha \beta^{-1} \int_{t_i}^{t_{i+1}} \int_{Q^\beta} \left|\left(u^1\right)^{[m]}\right| dx\,dr \\ &\leqslant C\,|t_{i+1} - t_i|^\alpha \int_{t_i}^{t_{i+1}} \int_{Q^\beta} \left|\nabla\left(u^1\right)^{[m]}\right| dx\,dr.\end{aligned}$$

In the first passage we used the boundedness of $D_x Y^{x,\xi}_{r,r-t_i}$ from Proposition 2.5.1 and of the sgn function, estimate (2.3.15), and observations (2.3.113) and (2.3.115). In the second passage we integrated the mollifiers over the variables (y, η) and (x', ξ'), recalling estimate (2.3.87), and then used properties of the kinetic function. In the last passage we used the Sobolev regularity $(u^1)^{[m]} \in W^{1,p_m}_0(Q)$ from Lemma 2.3.7 and applied the mean value theorem to points $x \in Q^\beta$.

Virtually identical analyses and estimates hold for $BE^{\text{sgn}2,2}_i$ and for $BE^{\text{mix}j,2}_i$, for $j = 1, 2$. Precisely, we have

$$\left|BE^{\text{sgn}j,2}_i\right| + \left|BE^{\text{mix}j,2}_i\right| \leqslant C\,|t_{i+1} - t_i|^\alpha \int_{t_i}^{t_{i+1}} \int_{Q^\beta} \left|\nabla\left(u^j\right)^{[m]}\right| dx\,dr, \quad (2.3.116)$$

for $j = 1, 2$ and $C = C(Q, T, A, z)$.

In conclusion, we sum estimates (2.3.114) and (2.3.116) over the partition $\mathcal{P}$ to conclude that, for $C = C(Q, T, A, z)$,

$$\sum_{j=1}^{2} \sum_{k=1}^{2} \sum_{i=0}^{N-1} \left|BE^{\text{sgn}j,k}_i\right| + \left|BE^{\text{mix}j,k}_i\right| \leqslant C\,|\mathcal{P}|^\alpha \sum_{j=1}^{2} \int_0^T \int_{Q^\beta} \left|\nabla\left(u^j\right)^{[m]}\right| dx\,dr. \quad (2.3.117)$$

Finally, we analyze the boundary errors $BE^{\text{sgn}1,3}_i$ and $BE^{\text{mix}1,3}_i$, defined in (2.3.38) and (2.3.55) respectively, and the specular terms $BE^{\text{sgn}2,3}_i$ and $BE^{\text{mix}2,3}_i$, simply obtained by swapping the roles of u^1 and u^2. For $BE^{\text{sgn}1,3}_i$, we use the boundedness of $\tilde{\text{sgn}}^\epsilon_{t_i,r}$, we exploit estimate (2.3.15), and estimate (2.5.11) combined with observation (2.3.86), and then we integrate the mollifier to obtain, for $C = C(T, A, z)$,

$$\begin{aligned}\left|BE^{\text{sgn}1,3}_i\right| &\leqslant \int_{t_i}^{t_{i+1}} \int_{Q^2 \times \mathbb{R}^2} (p^1_r + q^1_r)\,\left|\rho^{1,\epsilon}_{t_i,r}\right|\,|\tilde{\text{sgn}}^\epsilon_{t_i,r}|\,\left|\nabla_y \phi_\beta(y) \partial_\xi Y^{x,\xi}_{r,r-t_i}\right| dx\,dy\,d\xi\,d\eta\,dr \\ &\leqslant C|t_{i+1} - t_i|^\alpha \int_{t_i}^{t_{i+1}} \int_{Q^\beta \times \mathbb{R}} (p^j_r + q^j_r)\,dx\,d\xi\,dr.\end{aligned} \quad (2.3.118)$$

Virtually identical analyses hold for $BE^{\text{sgn}2,3}_i$ and for $BE^{\text{mix}j,3}_i$, for $j = 1, 2$. Summing over the partition, we conclude that, for $C = C(T, A, z)$,

$$\sum_{j=1}^{2} \sum_{i=0}^{N-1} \left|BE^{\text{sgn}j,3}_i\right| + \left|BE^{\text{mix}j,3}_i\right| \leqslant C\,|\mathcal{P}|^\alpha \sum_{j=1}^{2} \int_0^T \int_{Q^\beta \times \mathbb{R}} (p^j_r + q^j_r)\,dx\,d\xi\,dr. \quad (2.3.119)$$

In conclusion, combining estimates (2.3.117) with the integrability of $\nabla\left(u^j\right)^{[m]}$ from Lemma 2.3.7, and estimate (2.3.119) with the finiteness of the parabolic and entropy defect measures over $Q \times \mathbb{R} \times [0,T]$, and exploiting the dominated convergence theorem, for the boundary error terms produced in Step 2 and Step 3, we conclude that

$$\limsup_{\beta\to 0}\limsup_{\epsilon\to 0}\sum_{j=1}^{2}\sum_{k=1}^{3}\sum_{i=0}^{N-1}\left|BE_i^{\mathrm{sgn}j,k}\right|+\left|BE_i^{\mathrm{mix}j,k}\right|=0. \tag{2.3.120}$$

Now we analyze the boundary cancellation error terms produced in Step 5. First we observe that (2.3.11) combined with (2.3.15), for $k=1$ or $k=2$, implies that

$$\left|D_y^k\phi_\beta\right|\leqslant C\beta^{-k},\ \text{ and that, if }\ D_y^k\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)\neq 0,\ \text{ then } x, Y_{r,r-t_i}^{x,\xi}\in Q^\beta. \tag{2.3.121}$$

Then, for the error term $BE_i^{\mathrm{canc}A,1}$ defined in (2.3.77), we compute for $C=C(T,A,z,Q)$

$$\begin{aligned}\left|BE_i^{\mathrm{canc}A,1}\right| &\leqslant C|t_{i+1}-t_i|^\alpha\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\left(u^1\right)^{[m]}\right|\,dx\,dr\\ &\leqslant C|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|\,dx\,dr\end{aligned} \tag{2.3.122}$$

In the first passage we used use observation (2.3.121) and estimate (2.5.5), and then we exploited the boundedness of $|\chi^2|$ and integrated $|\chi^1|$ (or viceversa). In the second passage we used $\left(u^1\right)^{[m]}\in W_0^{1,p_m}(Q)$ and applied the mean value theorem to points $x\in Q^\beta$.

The error term $BE_i^{\mathrm{canc}A,2}$ defined in (2.3.80) is handled almost identically. Indeed, we use the boundedness of $D_xY_{r,r-t_i}^{x,\xi}$, and observation (2.3.121) for $k=2$ combined with estimate (2.5.12), then we integrate the kinetic functions, and finally we exploit again $\left(u^j\right)^{[m]}\in W_0^{1,p_m}(Q)$ and the mean value theorem to compute, for $C=C(T,A,z,Q)$,

$$\begin{aligned}\left|BE_i^{\mathrm{canc}A,2}\right| &\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left(|\chi_r^1|+|\chi_r^2|\right)\left|D_y^2\phi_\beta\big(Y_{r,r-t_i}^{x,\xi}\big)\right|\left|D_xY_{r,r-t_i}^{x,\xi}-\mathrm{I}_d\right|\,dx\,d\xi\,dr\\ &\leqslant C\,|t_{i+1}-t_i|^\alpha\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\left(u^1\right)^{[m]}\right|+\left|\left(u^2\right)^{[m]}\right|\,dx\,dr\\ &\leqslant C\,|t_{i+1}-t_i|^\alpha\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|+\left|\nabla\left(u^2\right)^{[m]}\right|\,dx\,dr.\end{aligned} \tag{2.3.123}$$

In conclusion, we sum estimates (2.3.122) and (2.3.123) over the partition $\mathcal{P}$ to conclude that, for $C=C(Q,T,A,z)$,

$$\sum_{i=0}^{N-1}\left|BE_i^{\mathrm{canc}A,1}\right|+\left|BE_i^{\mathrm{canc}A,2}\right|\leqslant C\,|\mathcal{P}|^\alpha\int_0^T\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|+\left|\nabla\left(u^2\right)^{[m]}\right|\,dx\,dr. \tag{2.3.124}$$

We now consider the error term $BE_i^{\mathrm{canc}A,3}$ defined in (2.3.82). First, for the function $d_{\partial Q}$ defined in (2.3.7) and ψ_β defined in (2.3.13), recalling (2.3.11), we observe that

$$\left|\dot\psi_\beta(s)\right|\leqslant\beta^{-1}\ \forall s\in\mathbb{R}\quad\text{and, if }\ \dot\psi_\beta\big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)\big)\neq 0,\ \text{ then } x, Y_{r,r-t_i}^{x,\xi}\in Q^\beta. \tag{2.3.125}$$

We also mention that we have a uniform bound $|\nabla\cdot\tilde{n}|\leqslant C$ for some constant $C=C(Q)$, for the extended unit outward normal to Q defined in (2.3.8). Then we compute

$$\begin{aligned}\left|BE_i^{\mathrm{canc}A,3}\right| &\leqslant \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\left|\dot{\psi}_\beta\big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)\big)\right||\nabla\cdot\tilde{n}|\,dx\,d\xi\,dr\\ &\leqslant C\,\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta\times\mathbb{R}} m|\xi|^{m-1}\left(\left|\chi_r^1\right|+\left|\chi_r^2\right|\right)dx\,d\xi\,dr\\ &= C\,\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\left(u^1\right)^{[m]}\right|+\left|\left(u^2\right)^{[m]}\right|dx\,dr\\ &\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|+\left|\nabla\left(u^2\right)^{[m]}\right|dx\,dr,\end{aligned}$$

for a constant $C=C(Q,T,A,z)$. In the second inequality we used observation (2.3.125), the boundedness of $\nabla\cdot\tilde{n}$ and properties of the kinetic function. In the last inequality we used $\left(u^j\right)^{[m]}\in W_0^{1,p_m}(Q)$ and the mean value theorem. In conclusion, we sum over the partition to estimate, for $C=C(Q,T,A,z)$,

$$\sum_{i=0}^{N-1}\left|BE_i^{\mathrm{canc}A,3}\right|\leqslant C\sum_{j=1}^{2}\int_0^T\int_{Q^\beta}\left|\nabla\left(u^1\right)^{[m]}\right|+\left|\nabla\left(u^2\right)^{[m]}\right|dx\,dr. \tag{2.3.126}$$

Finally we consider the error term $BE_i^{\mathrm{canc}A,4}$ defined in (2.3.83). First, we make the following observation. Namely, for the standard 1-dimensional convolution kernel $\rho_1^{\frac{1}{2}\beta^{\gamma_m}}$ introduced in (2.3.12)-(2.3.13), rescaled at order $\frac{1}{2}\beta^{\gamma_m}$ with $\gamma_m=(m+2)\vee 3$, recalling (2.3.11), we notice that

$$\left|\rho_1^{\frac{1}{2}\beta^{\gamma_m}}\right|\leqslant 2\,\beta^{-\gamma_m}\quad\text{and, if }\ \rho_1^{\frac{1}{2}\beta^{\gamma_m}}\Big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)-\beta^{\gamma_m}\Big)\neq 0,\ \text{ then } x,Y_{r,r-t_i}^{x,\xi}\in Q^{\beta^{\gamma_m}}. \tag{2.3.127}$$

Then we compute

$$\begin{aligned}\left|BE_i^{\mathrm{canc}A,4}\right| &= \int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\left|\chi_r^1-\chi_r^2\right|^2\beta^{-1}\rho_1^{\frac{1}{2}\beta^{\gamma_m}}\Big(d_{\partial Q}\big(Y_{r,r-t_i}^{x,\xi}\big)-\beta^{\gamma_m}\Big)dx\,d\xi\,dr\\ &\leqslant C\beta^{-1}\beta^{-\gamma_m}\int_{t_i}^{t_{i+1}}\int_{Q^{\beta^{\gamma_m}}\times\mathbb{R}} m|\xi|^{m-1}\left(\left|\chi_r^1\right|+\left|\chi_r^2\right|\right)dx\,d\xi\,dr\\ &= C\beta^{-(1+\gamma_m)}\int_{t_i}^{t_{i+1}}\int_{Q^{\beta^{\gamma_m}}}\left|\left(u^1\right)^{[m]}\right|+\left|\left(u^2\right)^{[m]}\right|dx\,dr\\ &\leqslant C\beta^{-1}\int_{t_i}^{t_{i+1}}\int_{Q^{\beta^{\gamma_m}}}\left|\nabla\left(u^1\right)^{[m]}\right|+\left|\nabla\left(u^2\right)^{[m]}\right|dx\,dr,\end{aligned} \tag{2.3.128}$$

for a constant $C=C(T,A,z)$. In the first inequality we used observation (2.3.127) and properties of the kinetic function. In the last inequality we used the Sobolev regularity $\left(u^j\right)^{[m]}\in W_0^{1,p_m}(Q)$ and the mean value theorem applied to points $x\in Q^{\beta^{\gamma_m}}$, which satisfy $\mathrm{dist}(x,\partial Q)\leqslant C\beta^{\gamma_m}$.

Next, we sum estimate (2.3.128) over the partition, then we apply Hölder's inequality combined with Lemma 2.3.7, prescribing the higher integrability $\nabla\left(u^j\right)^{[m]}\in L^{p_m}([0,T];L^{p_m}(Q))$, and observation (2.3.10), and we recall that $\gamma_m = (m+2) \vee 3$ and $p_m = \frac{m+1}{m} \wedge 2$, to estimate

$$
\begin{aligned}
\sum_{i=0}^{N-1} \left| BE_i^{\text{canc}A,4} \right| &\leqslant C\beta^{-1} \sum_{j=1}^{2} \int_0^T \int_{Q^{\beta\gamma_m}} \left| \nabla \left(u^j\right)^{[m]} \right| dx\, dr \\
&\leqslant C\beta^{-1} \left(\beta^{\gamma_m}\right)^{\frac{p_m-1}{p_m}} \sum_{j=1}^{2} \left\| \nabla \left(u^j\right)^{[m]} \right\|_{L^{p_m}\left([0,T];L^{p_m}\left(Q^{\beta\gamma_m}\right)\right)} \qquad (2.3.129) \\
&= C\beta^{\frac{1}{2\vee(m+1)}} \sum_{j=1}^{2} \left\| \nabla \left(u^j\right)^{[m]} \right\|_{L^{p_m}\left([0,T];L^{p_m}\left(Q^{\beta\gamma_m}\right)\right)},
\end{aligned}
$$

for a constant $C = C(Q, T, A, z)$.

In conclusion, combining estimates (2.3.124), (2.3.126) and (2.3.129) with Lemma 2.3.7, it follows from the dominated convergence theorem that

$$
\limsup_{\beta \to 0} \limsup_{\epsilon \to 0} \sum_{k=1}^{4} \sum_{i=0}^{N-1} \left| BE_i^{\text{canc}A,k} \right| = \limsup_{\beta \to 0} \sum_{k=1}^{4} \sum_{i=0}^{N-1} \left| BE_i^{\text{canc}A,k} \right| = 0. \qquad (2.3.130)
$$

We conclude this step by analyzing the last boundary error terms, namely the boundary cancellation errors $BE_i^{\text{canc}Bj}$ defined in (2.3.72), for $j = 1, 2$. The analysis is broken down in three cases: $m = 1$, $m \in (1, \infty)$ and $m \in (0, 1)$.

Case $m = 1$. This case is trivial. Indeed, if $m = 1$, it follows automatically from definition (2.3.72) that $BE_i^{\text{canc}Bj} = 0$.

Case $m \in (1, \infty)$. We form a velocity decomposition of the integral. For each $M > 1$, let $K_M : \mathbb{R} \to [0, 1]$ be a smooth function satisfying

$$
K_M(\xi) = \begin{cases} 1 & \text{if } |\xi| \leqslant M, \\ 0 & \text{if } |\xi| \leqslant M+1. \end{cases}
$$

Then, for each $M > 1$, we split

$$
\begin{aligned}
BE_i^{\text{canc}B1} = {} & 2m \int_{t_i}^{t_{i+1}} \int_{Q^3 \times \mathbb{R}^3} K_M(\xi) \left(|\xi|^{m-1} - |\xi'|^{m-1}\right) \chi_r^1 \chi_r^2 \, \rho_{t_i,r}^{1,\epsilon} \nabla_{x'} \rho_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi} \\
& + 2m \int_{t_i}^{t_{i+1}} \int_{Q^3 \times \mathbb{R}^3} (1 - K_M(\xi)) \left(|\xi|^{m-1} - |\xi'|^{m-1}\right) \chi_r^1 \chi_r^2 \, \rho_{t_i,r}^{1,\epsilon} \nabla_{x'} \rho_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi}.
\end{aligned} \qquad (2.3.131)
$$

For the first term on the right-hand side of (2.3.131) we write

$$
\begin{aligned}
& \left| 2m \int_{t_i}^{t_{i+1}} \int_{Q^3 \times \mathbb{R}^3} K_M(\xi) \left(|\xi|^{m-1} - |\xi'|^{m-1}\right) \chi_r^1 \chi_r^2 \, \rho_{t_i,r}^{1,\epsilon} \nabla_{x'} \rho_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi} \, dx\, dx'\, dy\, d\xi\, d\xi'\, d\eta\, dr \right| \\
& \qquad \leqslant C\beta^{-1} \int_{t_i}^{t_{i+1}} \int_{(Q^\beta)^3 \times \mathbb{R}^3 \cap \{|\xi-\xi'| \leqslant c\epsilon\}} |\xi - \xi'|^{(m-1)\wedge 1} \, \rho_{t_i,r}^{1,\epsilon} \left| \nabla_{x'} \rho_{t_i,r}^{2,\epsilon} \right| dx\, dx'\, dy\, d\xi\, d\xi'\, d\eta\, dr \qquad (2.3.132) \\
& \qquad \leqslant C\beta^{-1}\epsilon^{-1} \int_{t_i}^{t_{i+1}} \int_{Q^\beta} \int_{-c\epsilon}^{c\epsilon} |\theta|^{(m-1)\wedge 1} d\theta\, dy\, dr \leqslant C \left|t_{i+1} - t_i\right| \epsilon^{(m-1)\wedge 1},
\end{aligned}
$$

for a constant $C = C(M, m, Q, T, A, z)$. In the first passage we used the boundedness of K_M, χ^j and $D_x Y_{r,r-t_i}^{x,\xi}$, observation (2.3.15) and (2.3.113), observation (2.3.86) combined with estimate (2.5.4), and the local Lipschitz continuity, if $m \in (2, \infty)$, or the Hölder continuity, if $m \in (1,2)$, of the map $\mathbb{R} \ni \xi \mapsto |\xi|^{m-1}$. In the second passage we exploited formula (2.3.26) for the derivatives of the convolution kernels, the boundedness of the derivatives of the characteristics from Proposition 2.5.1, and then we integrated the convolution kernels, recalling observation (2.3.87) and changing variables by setting $\theta = \xi - \xi'$. In the last passage we simply used formula (2.3.10).

For the second term on the right-hand side of (2.3.131) we use the following elementary inequality

$$\left||\xi|^{m-1} - |\xi'|^{m-1}\right| = \left|\int_{\xi'}^{\xi}(m-1)\theta^{[m-2]}d\theta\right| \leqslant (m-1)\left(|\xi|^{m-2} + |\xi'|^{m-2}\right)|\xi - \xi'|. \tag{2.3.133}$$

Then we estimate

$$\begin{aligned}
&\left|2m\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3}(1-K_M)\left(|\xi|^{m-1} - |\xi'|^{m-1}\right)\chi_r^1\chi_r^2\,\rho_{t_i,r}^{1,\epsilon}\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\nabla_y\phi_\beta D_x Y_{r,r-t_i}^{x,\xi}dxdx'dyd\xi d\xi' d\eta dr\right| \\
&\leqslant C\int_{t_i}^{t_{i+1}}\int_{(Q^\beta)^3\times\mathbb{R}^3\cap\{|\xi-\xi'|\leqslant c\epsilon\}}(1-K_M(\xi))\left(|\xi|^{m-2} + |\xi'|^{m-2}\right)\epsilon\left|\chi_r^1\right|\left|\chi_r^2\right|\rho_{t_i,r}^{1,\epsilon}\left|\nabla_{x'}\rho_{t_i,r}^{2,\epsilon}\right|dxdx'dyd\xi d\xi' d\eta dr \\
&\leqslant C\left(\int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u^1|\geqslant M\}}\left|u^1\right|^{m-1}dx\,dr + \int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u^2|\geqslant M-c\epsilon\}}\left|u^2\right|^{m-1}dx'dr\right),
\end{aligned} \tag{2.3.134}$$

for constants $C = C(\beta, m, T, A, z)$ and $c = c(T, A, z)$, independent of $M \geqslant 1$. In the first passage we used (2.3.133), (2.3.86) combined with (2.5.4), and (2.3.15). In the second passage we used observation (2.3.87) and integrated the convolution kernels in the (x', ξ') and (y, η) variables, when hitting $|\xi|^{m-1}$, and in the (x, ξ) and (y, η) variables, when hitting $|\xi'|^{m-1}$, and then we used properties of the kinetic function.

Finally, combining (2.3.131) with (2.3.132) and (2.3.134), and summing over the partition $\mathcal{P}$, we obtain

$$\sum_{i=0}^{N-1}\left|BE_i^{\mathrm{canc}B1}\right| \leqslant C_1\,\epsilon^{(m-1)\wedge 1} + C_2\left(\int_0^T\int_{Q\cap\{|u^1|\geqslant M\}}\left|u^1\right|^{m-1}dxdr + \int_0^T\int_{Q\cap\{|u^2|\geqslant M-c\epsilon\}}\left|u^2\right|^{m-1}dx'dr\right), \tag{2.3.135}$$

for constants $C_1 = C_1(M, m, Q, T, A, z)$, $C_2 = C_2(\beta, m, Q, T, A, z)$ and $c = c(T, A, z)$. An inequality virtually identical to (2.3.135) holds for $BE_i^{\mathrm{canc}B2}$, swapping u^1 and u^2. Since $m \in (1, \infty)$ and by Definition 2.2.4 of kinetic solution we have $u^j \in L^{m+1}([0,T]; L^{m+1}(Q))$, and since the constant C_2 is independent of $M \geqslant 1$, Hölder's inequality and the dominated convergence theorem prove that the last two terms on the right-hand side of (2.3.135) vanish in the limit $M \to \infty$, uniformly for $\epsilon \in (0,1)$. Therefore, passing first to the limit $\epsilon \to 0$ and second to the limit $M \to \infty$, formula (2.3.135) yields

$$\limsup_{\epsilon\to 0}\sum_{i=0}^{N-1}\sum_{j=1}^{2}\left|BE_i^{\mathrm{canc}Bj}\right| = 0. \tag{2.3.136}$$

Case $m \in (0,1)$. For this case the idea is to remove the singularity at the origin and to exploit the full regularity of the solution implied by Proposition 2.3.9 below. Using the integration by parts formula (2.2.27) in the (x', ξ') variable, which is justified exploiting an approximation argument and Proposition 2.3.6 below, we write

$$\begin{aligned} BE_i^{\mathrm{canc}B1} &= 2m \int_{t_i}^{t_{i+1}} \int_{Q^3 \times \mathbb{R}^3} \sigma_m(\xi, \xi') |\xi|^{\frac{m}{2}-1} |\xi'|^{\frac{m}{2}-1} \chi_r^1 \chi_r^2 \rho_{t_i,r}^{1,\epsilon} \nabla_{x'} \rho_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta D_x Y_{r,r-t_i}^{x,\xi} dx\, dx' dy\, d\xi\, d\xi' d\eta\, dr \\ &= -\frac{4m}{(m+1)} \int_{t_i}^{t_{i+1}} \int_Q |u^2|^{-\frac{1}{2}} \nabla \left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\quad \cdot \left(\int_{Q^2 \times \mathbb{R}^2} \sigma_m(\xi, u^2) |\xi|^{\frac{m}{2}-1} \chi_r^1 \rho_{t_i,r}^{1,\epsilon} \bar{\rho}_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi} dx\, dy\, d\xi\, d\eta \right) dx' dr, \end{aligned} \tag{2.3.137}$$

where we recall the notation (2.3.65) for $\bar{\rho}_{t_i,r}^{2,\epsilon}(x', y, \eta)$, and where we have defined

$$\sigma_m(\xi, \xi') := |\xi|^{\frac{m}{2}} |\xi'|^{1-\frac{m}{2}} - |\xi'|^{\frac{m}{2}} |\xi|^{1-\frac{m}{2}} \quad \text{for } \xi, \xi' \in \mathbb{R}. \tag{2.3.138}$$

Observation (2.3.86) and estimate (2.5.4) prove that, for $c_1 = c_1(T, A, z)$,

$$\text{if } \rho_{t_i,r}^{1,\epsilon}(x, y, \xi, \eta) \bar{\rho}_{t_i,r}^{2,\epsilon}(x', y, \eta) \neq 0, \text{ then } |x - x'| + |\xi - u^2| \leqslant c_1 \epsilon. \tag{2.3.139}$$

Moreover, we observe that

$$\text{if } |\xi - \xi'| \leqslant c_1 \epsilon, \text{ then } |\sigma_m(\xi, \xi')| \leqslant C\, \epsilon, \tag{2.3.140}$$

for a constant $C = C(c_1, m)$. Indeed,if $\max\{|\xi|, |\xi'|\} \leqslant 2c_1\epsilon$, recalling $m \in (0, 1)$, the result is immediate from (2.3.138). Conversely, assume without loss of generality that $|\xi| > 2c_1\epsilon$ with $|\xi| \geqslant |\xi'|$ and $|\xi - \xi'| \leqslant c_1\epsilon$. Thus, in particular, ξ and ξ' have the same sign and $|\xi'| \geqslant \frac{1}{2}|\xi|$. Then, using $\xi \simeq \xi'$, we compute

$$\sigma_m(\xi, \xi') = |\xi|^{1-\frac{m}{2}} |\xi'|^{1-\frac{m}{2}} \int_{\xi'}^{\xi} (m-1) \theta^{[m-2]}\, d\theta \leqslant C|\xi - \xi'| \leqslant C\epsilon.$$

We now form a velocity decomposition of the integral. For each $\delta \in (0, 1)$, let $K^\delta : \mathbb{R} \to [0, 1]$ denote a smooth cutoff function satisfying

$$K^\delta(\xi) = \begin{cases} 1 & \text{if } |\xi| \leqslant \delta, \\ 0 & \text{if } |\xi| \geqslant 2\delta. \end{cases}$$

Returning to (2.3.137), consider the decomposition

$$\begin{aligned} BE_i^{\mathrm{canc}B1} = &-\frac{4m}{(m+1)} \int_{t_i}^{t_{i+1}} \int_Q |u^2|^{-\frac{1}{2}} \nabla \left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\quad \cdot \left(\int_{Q^2 \times \mathbb{R}^2} \sigma_m^\delta\left(\xi, u^2\right) |\xi|^{\frac{m}{2}-1} \chi_r^1 \rho_{t_i,r}^{1,\epsilon} \bar{\rho}_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi} dx\, dy\, d\xi\, d\eta \right) dx' dr \\ &-\frac{4m}{(m+1)} \int_{t_i}^{t_{i+1}} \int_Q |u^2|^{-\frac{1}{2}} \nabla \left(u^2\right)^{\left[\frac{m+1}{2}\right]} \\ &\quad \cdot \left(\int_{Q^2 \times \mathbb{R}^2} \tilde{\sigma}_m^\delta\left(\xi, u^2\right) |\xi|^{\frac{m}{2}-1} \chi_r^1 \rho_{t_i,r}^{1,\epsilon} \bar{\rho}_{t_i,r}^{2,\epsilon} \nabla_y \phi_\beta(y) D_x Y_{r,r-t_i}^{x,\xi} dx\, dy\, d\xi\, d\eta \right) dx' dr, \end{aligned} \tag{2.3.141}$$

where, for each $\delta \in (0, 1)$, the functions $\sigma_m^\delta, \tilde{\sigma}_m^\delta : \mathbb{R}^2 \to \mathbb{R}$ are defined by

$$\sigma_m^\delta(\xi, \xi') = \left(K^\delta(\xi) + K^\delta(\xi') - K^\delta(\xi) K^\delta(\xi')\right) \sigma_m(\xi, \xi'), \tag{2.3.142}$$

and

$$\tilde{\sigma}^{\delta}_{m}(\xi,\xi') = \left(1 - K^{\delta}(\xi)\right)\left(1 - K^{\delta}(\xi')\right)\sigma_m(\xi,\xi'). \tag{2.3.143}$$

We start with the second term in (2.3.141). It follows from (2.3.139), (2.3.140) and (2.3.143) that, for a constant $C = C(m,T,A,z)$, for any $(x,x',y,\xi,\eta,r) \in Q^3 \times \mathbb{R}^2 \times [t_i,\infty)$,

$$\text{if } \rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\tilde{\sigma}^{\delta}_{m}\left(\xi,u^2\right) \neq 0, \text{ then } |\xi|,\left|u^2\right| \geqslant \delta \text{ and } \left|\tilde{\sigma}^{\delta}_{m}\left(\xi,u^2\right)\right| \leqslant C\epsilon. \tag{2.3.144}$$

Then, for a constant $C = C(\delta,\beta,m,T,A,z)$, we compute

$$\begin{aligned}
&\left|\frac{4m}{(m+1)}\int_{t_i}^{t_{i+1}}\!dr\int_Q dx' \left|u^2\right|^{-\frac{1}{2}}\nabla(u^2)^{\left[\frac{m+1}{2}\right]}\int_{Q^2\times\mathbb{R}^2}\tilde{\sigma}^{\delta}_{m}\left(\xi,u^2\right)|\xi|^{\frac{m}{2}-1}\chi^1_r\rho^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta D_xY^{x,\xi}_{r,r-t_i}\,dx\,dy\,d\xi\,d\eta\right| \\
&\leqslant C\int_{t_i}^{t_{i+1}}\!\int_Q \left|u^2\right|^{-\frac{1}{2}}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|\left(\int_{Q^2\times\mathbb{R}^2\cap\{|\xi|\geqslant\delta\}}\epsilon\,|\xi|^{\frac{m}{2}-1}\chi^1_r\,\rho^{1,\epsilon}_{t_i,r}\,\bar{\rho}^{2,\epsilon}_{t_i,r}\,dx\,dy\,d\xi\,d\eta\right)dx'\,dr \\
&\leqslant C\epsilon\int_{t_i}^{t_{i+1}}\!\int_Q \left|u^2\right|^{-\frac{1}{2}}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|dx'\,dr.
\end{aligned} \tag{2.3.145}$$

In the first passage we used (2.3.144), (2.3.15) and the boundedness of χ^1 and $D_xY^{x,\xi}_{r,r-t_i}$. In the second passage we recalled $m \in (0,1)$, so that $\frac{m}{2} - 1 < 0$ and $|\xi|^{\frac{m}{2}-1} \leqslant \delta^{\frac{m}{2}-1}$, and integrated the convolution kernels.

We now consider the first term in (2.3.141). It follows from (2.3.139), (2.3.140) and (2.3.142) that, for a constant $C = C(m,T,A,z)$, for any $\epsilon < \delta \in (0,1)$, with ϵ small enough depending on C and δ, and any $(x,x',y,\xi,\eta,r) \in Q^3 \times \mathbb{R}^2 \times [t_i,\infty)$,

$$\text{if } \rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\sigma^{\delta}_{m}\left(\xi,u^2\right) \neq 0, \text{ then } |\xi|,\left|u^2\right| \leqslant 3\delta \text{ and } \left|\sigma^{\delta}_{m}\left(\xi,u^2\right)\right| \leqslant C\epsilon. \tag{2.3.146}$$

Moreover, the definition (2.3.19) of the convolution kernels implies that, for any $(x',r) \in Q \times [t_i,\infty)$,

$$\int_{Q^2\times\mathbb{R}} \rho^{1,\epsilon}_{t_i,r}(x,y,\xi,\eta)\,\bar{\rho}^{2,\epsilon}_{t_i,r}(x',y,\eta)\;dx\,dy\,d\eta \leqslant \epsilon^{-1}. \tag{2.3.147}$$

Then, for the first term of (2.3.141), we compute

$$\begin{aligned}
&\left|\frac{4m}{(m+1)}\int_{t_i}^{t_{i+1}}\!dr\int_Q dx' \left|u^2\right|^{-\frac{1}{2}}\nabla(u^2)^{\left[\frac{m+1}{2}\right]}\int_{Q^2\times\mathbb{R}^2}\sigma^{\delta}_{m}\left(\xi,u^2\right)|\xi|^{\frac{m}{2}-1}\chi^1_r\rho^{1,\epsilon}_{t_i,r}\bar{\rho}^{2,\epsilon}_{t_i,r}\nabla_y\phi_\beta D_xY^{x,\xi}_{r,r-t_i}\,dx\,dy\,d\xi\,d\eta\right| \\
&\leqslant C\int_{t_i}^{t_{i+1}}\!\int_Q \left|u^2\right|^{-\frac{1}{2}}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|\int_{-3\delta}^{3\delta}|\xi|^{\frac{m}{2}-1}\left(\int_{Q^2\times\mathbb{R}}\epsilon\,\rho^{1,\epsilon}_{t_i,r}\,\bar{\rho}^{2,\epsilon}_{t_i,r}\,dx\,dy\,d\eta\right)d\xi\,dx'\,dr \\
&\leqslant C\delta^{\frac{m}{2}}\int_{t_i}^{t_{i+1}}\!\int_Q \left|u^2\right|^{-\frac{1}{2}}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|dx'\,dr,
\end{aligned} \tag{2.3.148}$$

for a constant $C = C(\beta,m,T,A,z)$, independent of $\delta \in (0,1)$. In the first passage we used observation (2.3.146), (2.3.15) and the boundedness of χ^1 and $D_xY^{x,\xi}_{r,r-t_i}$. In the second passage we applied observation (2.3.147).

Finally, combining the splitting (2.3.141) with estimates (2.3.145) and (2.3.148), summing over the partition and applying Hölder's inequality, we obtain

$$\sum_{i=0}^{N-1}\left|BE^{\mathrm{canc}B1}_i\right| \leqslant \left(C_1\,\epsilon + C_2\,\delta^{\frac{m}{2}}\right)\left(\int_0^T\!\!\int_Q \left|u^2\right|^{-1}\left|\nabla\left(u^2\right)^{\left[\frac{m+1}{2}\right]}\right|^2 dx'\,dr\right)^{\frac{1}{2}}, \tag{2.3.149}$$

for constants $C_1 = C_1(\delta, \beta, m, Q, T, A, z)$ and $C_2 = C_2(\beta, m, Q, T, A, z)$. An inequality virtually identical to (2.3.149) holds for the specular errors $BE_i^{\mathrm{canc}B2}$, simply swapping the roles of u^1 and u^2. Proposition 2.3.9 below ensures that the right-hand side of (2.3.149) is finite. Since the constant C_2 is independent of δ and ϵ, passing first to the limit $\epsilon \to 0$ and then to the limit $\delta \to 0$, formula (2.3.149) and its analogue version for $j = 2$ yield

$$\limsup_{\epsilon \to 0} \sum_{i=0}^{N-1} \sum_{j=1}^{2} \left| BE_i^{\mathrm{canc}Bj} \right| = 0. \tag{2.3.150}$$

This concludes the analysis of the boundary error terms.

Step 9: The conclusion. We are finally ready to conclude the proof. Returning to inequality (2.3.85), we use estimates (2.3.90), (2.3.96) and (2.3.112), (2.3.120), (2.3.130), and (2.3.136) and (2.3.150), to deduce

$$\int_Q \left| u^1(x,r) - u^2(x,r) \right| dx \Big|_{r=0}^{r=T} \leqslant C\, |\mathcal{P}|^\alpha \sum_{j=1}^{2} \left(\int_0^T \int_Q \left| \nabla \left(u^j\right)^{[m]} \right| dx\, dr + \int_0^T \int_{Q\times\mathbb{R}} p_r^j + q_r^j \, dx\, d\xi\, dr \right), \tag{2.3.151}$$

for a constant $C = C(T, A, z)$. Definition 2.2.4 of pathwise kinetic solution and Lemma 2.3.7 guarantee that the right-hand side of (2.3.151) is finite. Recalling from Step 0 that the partition $\mathcal{P} \subseteq [0,T] \backslash \mathcal{N}$ is arbitrary, we let $|\mathcal{P}| \to 0$ and from (2.3.151) we conclude that

$$\int_Q \left| u^1(x,T) - u^2(x,T) \right| dx \leqslant \int_Q \left| u_0^1(x) - u_0^2(x) \right| dx.$$

This completes the uniqueness proof.

□

Remark 2.3.2. We observe that in the proof of Theorem 2.1.2 the positivity of the initial data was exploited only in Step 7, to handle the error terms IE_i^{canc} in the case $m \in (0,1) \cup (1,2]$, and in Step 8, to handle the error terms $BE_i^{\mathrm{canc}Bj}$ in the case $m \in (0,1)$. Namely, we relied on the positivity of the initial data through the application of Proposition 2.3.4 and 2.3.9 below. The remaining arguments of this chapter are obtained for general initial data in $L^2(Q)$. This completes the proof of Theorem 2.1.7.

Repeating the proof of Theorem 2.1.2 with $u_2 = 0$, we get the following estimate for the L^1-norm of pathwise kinetic solutions with any arbitrary initial data.

Corollary 2.3.3. Let $m \in (0, \infty)$ and let $u_0 \in L^2(Q)$ be arbitrary. Suppose that u is a pathwise kinetic solution of (2.1.1) with initial data u_0, then

$$\|u\|_{L^\infty([0,T];L^1(Q))} \leqslant \|u_0\|_{L^1(Q)}.$$

Proof. Let $u_0 \in L^2(Q)$ be arbitrary and let u be a pathwise kinetic solution of (2.1.1) with initial data u_0. Repeating the proof of Theorem 2.1.2 with $\chi^2 = 0$ implies that

$$\|u\|_{L^\infty([0,T];L^1(Q))} = \|u - 0\|_{L^\infty([0,T];L^1(Q))} \leqslant \|u_0 - 0\|_{L^1(Q)} = \|u_0\|_{L^1(Q)}.$$

Indeed, if $\chi^2 = 0$, in the final inequality (2.3.85) the error terms IE_i^{canc}, defined in (2.3.68), and $BE_i^{\text{canc}Bj}$, for $j = 1, 2$, defined in (2.3.72), are identically zero. In particular, as observed in Remark 2.3.2, these errors are the only terms requiring the positivity of the initial data to tackle the regime $m \in (0, 1) \cup (1, 2]$. Thus, when $u^2 = 0$, the proof works in the full regime $m \in (0, \infty)$ regardless of the sign of u. $\square$

We conclude this section with a few auxiliary results. The following proposition ensures that kinetic solutions with positive initial data stay positive, and it is crucial to obtain Proposition 2.3.9 below.

Proposition 2.3.4. Let u be a pathwise kinetic solution with nonnegative initial data $u_0 \in L^2_+(Q)$. Then we have $u(x, t) \geqslant 0$ almost everywhere in $Q \times [0, \infty)$. Moreover, for almost every $t \in [0, \infty)$ we have that $\|u(\cdot, t)\|_{L^1(Q)} = \|u_0\|_{L^1(Q)}$.

Proof. Suppose $u_0 \in L^2_+(Q)$ and let u be a pathwise kinetic solution to (2.1.1) with initial data u_0, kinetic function χ and exceptional set $\mathcal{N}$. Let $\mathcal{P} = \{0 = t_0 < t_1 < \cdots < t_N = T\}$ be an arbitrary partition of $[0, T]$ with $\mathcal{P} \subseteq [0, T] \backslash \mathcal{N}$. By repeating the same argument leading from (2.3.22) to (2.3.23), it follows that

$$\int_{Q\times\mathbb{R}} \chi_r(y,\eta)\mathrm{sgn}_-(\eta)\,dy\,d\eta\Big|_{r=0}^{r=T} = \lim_{\beta\to 0}\lim_{\epsilon\to 0}\sum_{i=0}^{N-1}\int_{Q\times\mathbb{R}} \tilde{\chi}^{\epsilon}_{t_i,r}(y,\eta)\,\widetilde{(\mathrm{sgn}_-)}^{\epsilon}_{t_i,r}\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}, \tag{2.3.152}$$

where $\mathrm{sgn}_-(\xi) := \mathrm{sgn}(\xi) \wedge 0$, and $\widetilde{(\mathrm{sgn}_-)}^{\epsilon}_{t_i,r}$ is defined as in (2.3.20)-(2.3.21) with sgn_- replacing sgn. As regards the second term in the sum, the same reasoning leading from (2.3.25) to (2.3.42), with $\mathrm{sgn}_-(\xi')$ replacing $\mathrm{sgn}(\xi')$, show that

$$\begin{aligned}
&\int_{Q\times\mathbb{R}} \tilde{\chi}^{\epsilon}_{t_i,r}(y,\eta)\,\widetilde{(\mathrm{sgn}_-)}^{\epsilon}_{t_i,r}\,\phi_\beta(y)\,dy\,d\eta\Big|_{r=t_i}^{r=t_{i+1}}\\
&= \overline{IE}_i^{\mathrm{sgn}1,1} - \overline{IE}_i^{\mathrm{sgn}1,2} + \overline{BE}_i^{\mathrm{sgn}1,1} + \overline{BE}_i^{\mathrm{sgn}1,2} - \overline{BE}_i^{\mathrm{sgn}1,3}\\
&\quad + \int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}\mathrm{sgn}_-(\xi')\\
&\qquad\qquad \cdot \mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\Big)\,dx\,dx'dyd\xi d\xi' d\eta dr\\
&\quad - 2\int_{t_i}^{t_{i+1}}\int_{Q^3\times\mathbb{R}^2} (p^1_r + q^1_r)\,\rho^{1,\epsilon}_{t_i,r}\,\rho^{2,\epsilon}_{t_i,r}(x',y,0,\eta)\,\phi_\beta(y)\,dx\,dx'\,dy\,d\xi\,d\eta\,dr,
\end{aligned} \tag{2.3.153}$$

for error terms $\overline{IE}_i^{\mathrm{sgn}1,1}$, $\overline{IE}_i^{\mathrm{sgn}1,2}$, $\overline{BE}_i^{\mathrm{sgn}1,1}$, $\overline{BE}_i^{\mathrm{sgn}1,2}$ and $\overline{BE}_i^{\mathrm{sgn}1,3}$ defined simply by setting $u^1 = u$ and replacing $\mathrm{sgn}(\xi')$ with $\mathrm{sgn}_-(\xi')$ in definition (2.3.30), (2.3.40), (2.3.33), (2.3.35) and (2.3.38) respectively. For the second to last term in (2.3.153), we first notice that

$$\begin{aligned}
&\lim_{\epsilon\to 0}\int_{t_i}^{t_{i+1}} dr \int_{Q^3\times\mathbb{R}^3} m|\xi|^{m-1}\chi^1_r\rho^{1,\epsilon}_{t_i,r}\rho^{2,\epsilon}_{t_i,r}\mathrm{sgn}_-(\xi')\;\mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D^2_y\phi_\beta(y)D_xY^{x,\xi}_{r,r-t_i}\Big)\,dxdx'dyd\xi d\xi' d\eta\\
&= \int_{t_i}^{t_{i+1}} dr \int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi^1_r\mathrm{sgn}_-(\xi)\;\mathrm{tr}\Big(\big(D_xY^{x,\xi}_{r,r-t_i}\big)^T D^2_y\phi_\beta(Y^{x,\xi}_{r,r-t_i})D_xY^{x,\xi}_{r,r-t_i}\Big)\,dxd\xi\,.
\end{aligned} \tag{2.3.154}$$

Then, observing that $|\xi|^{m-1}\chi^1_r\mathrm{sgn}_-(\xi)$ is always nonnegative and mimicking the arguments leading from (2.3.79) to (2.3.83), we conclude that

$$\begin{aligned}\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\chi_r^1\mathrm{sgn}_-(\xi)\ \mathrm{tr}\Big(\big(D_xY_{r,r-t_i}^{x,\xi}\big)^T D_y^2\phi_\beta(Y_{r,r-t_i}^{x,\xi})D_xY_{r,r-t_i}^{x,\xi}\Big)\,dx\,d\xi\,dr \\ \leqslant \overline{BE}_i^{\mathrm{canc}A,2}+\overline{BE}_i^{\mathrm{canc}A,3}+\overline{BE}_i^{\mathrm{canc}A,4},\end{aligned} \tag{2.3.155}$$

for error terms $\overline{BE}_i^{\mathrm{canc}A,2}$, $\overline{BE}_i^{\mathrm{canc}A,3}$ and $\overline{BE}_i^{\mathrm{canc}A,4}$ defined simply by replacing $|\xi|^{m-1}|\chi_r^1-\chi_r^2|^2$ with $|\xi|^{m-1}\chi_r\mathrm{sgn}_-(\xi)$ in definition (2.3.80), (2.3.82) and (2.3.83) respectively.

We now go back to (2.3.152). Owing to the nonnegativity of the entropy and parabolic defect measure, we drop the last term in (2.3.153), and then we exploit (2.3.154) combined with (2.3.155) to obtain

$$\begin{aligned}\int_{Q\times\mathbb{R}}\chi_r(y,\eta)\mathrm{sgn}_-(\eta)\,dy\,d\eta\Big|_{r=0}^{r=T} \leqslant \limsup_{\beta\to 0}\limsup_{\epsilon\to 0}\sum_{i=0}^{N-1}\Big(&\overline{IE}_i^{\mathrm{sgn}1,1}-\overline{IE}_i^{\mathrm{sgn}1,2}\\ &+\overline{BE}_i^{\mathrm{sgn}1,1}+\overline{BE}_i^{\mathrm{sgn}1,2}-\overline{BE}_i^{\mathrm{sgn}1,3}\\ &+\overline{BE}_i^{\mathrm{canc}A,2}+\overline{BE}_i^{\mathrm{canc}A,3}+\overline{BE}_i^{\mathrm{canc}A,4}\Big).\end{aligned}$$

Arguments virtually identical to those in Step 8 ensure that the sum of the boundary terms $\overline{BE}_i^{\mathrm{sgn}1,1}$, $\overline{BE}_i^{\mathrm{sgn}1,2}$, $\overline{BE}_i^{\mathrm{sgn}1,3}$, $\overline{BE}_i^{\mathrm{canc}A,2}$, $\overline{BE}_i^{\mathrm{canc}A,3}$ and $\overline{BE}_i^{\mathrm{canc}A,4}$ vanishes in the limit as $\epsilon\to 0$ first and $\beta\to 0$ next. Finally, recalling that the partition $\mathcal{P}\subseteq[0,T]\backslash\mathcal{N}$ is arbitrary, arguments completely analogous to those in Step 7 and Step 9 prove that the sum of the internal errors $\overline{IE}_i^{\mathrm{sgn}1,1}$ and $\overline{IE}_i^{\mathrm{sgn}1,2}$ goes to 0 as we let $\epsilon\to 0$ first, $\beta\to 0$ next, and then also $|\mathcal{P}|\to 0$. Putting everything together, we conclude that

$$0\leqslant\int_{Q\times\mathbb{R}}\chi(x,\xi,T)\mathrm{sgn}_-(\xi)\,dx\,d\xi\leqslant\int_{Q\times\mathbb{R}}\bar\chi(u_0(x),\xi)\mathrm{sgn}_-(\xi)\,dx\,d\xi\leqslant 0.$$

Here, the first inequality follows from the definition of the kinetic function and the last inequality follows from the nonnegativity of u_0. Therefore we deduce that, if $u_0\in L^2_+(Q)$, then $u\geqslant 0$ almost everywhere on $Q\times[0,\infty)$.

To prove the second assertion, we test the kinetic equation against the cutoff ϕ_β, we then let $\beta\to 0$ and we exploit the nonnegativity of the solution. An approximation argument shows that we can take $\rho_0(x,\xi):=\phi_\beta(x)$ in equation (2.2.26). For any $t\in[0,\infty)\backslash\mathcal{N}$, applying the integration by parts formula (2.2.27), we obtain

$$\begin{aligned}\int_{Q\times\mathbb{R}}\chi(x,\xi,r)\phi_\beta(Y_{r,r}^{x,\xi})\,dx\,d\xi\Big|_{r=0}^{r=t} = &-\frac{2m}{m+1}\int_0^t\int_Q|u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\nabla_y\phi_\beta(Y_{r,r}^{x,u})D_xY_{r,r}^{x,u}\,dx\,dr\\ &-\int_0^t\int_{Q\times\mathbb{R}}(p_r+q_r)\nabla_y\phi_\beta(Y_{r,r}^{x,\xi})\partial_\xi Y_{r,r}^{x,\xi}\,dx\,d\xi\,dr.\end{aligned} \tag{2.3.156}$$

The first term on the right-hand side of (2.3.156) vanishes as we let $\beta\to 0$. Indeed, we exploit observation (2.3.121), the Sobolev regularity $u^{\left[\frac{m+1}{2}\right]}\in H_0^1(Q)$ and the mean value theorem applied to points $x \in Q^\beta$, and the definition of the parabolic defect measure to estimate

$$\begin{aligned}\left|\int_0^t\int_Q |u|^{\frac{m-1}{2}} \nabla u^{\left[\frac{m+1}{2}\right]} \nabla_y \phi_\beta(Y^{x,u}_{r,r}) \nabla_x Y^{x,u}_{r,r} dx dr\right| &\leqslant C\beta^{-1} \int_0^t \int_{Q^\beta} |u|^{-1} \left|u^{\left[\frac{m+1}{2}\right]}\right| \left|\nabla u^{\left[\frac{m+1}{2}\right]}\right| dx dr \\ &\leqslant C \int_0^t \int_{Q^\beta} |u|^{-1} \left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|^2 dx dr \\ &= C \int_0^t \int_{Q^\beta \times \mathbb{R}} |\xi|^{-1} q(x,\xi,r)\, dx d\xi dr,\end{aligned} \tag{2.3.157}$$

for a constant $C = C(Q, m, T, A, z)$. Since we know that u is nonnegative, Proposition 2.3.9 below and the dominated convergence theorem ensure that, as $\beta \to 0$, the last line of (2.3.157) vanishes. Returning to (2.3.156), virtually the same estimate as (2.3.118)-(2.3.119) shows that the second term on the right-hand side of (2.3.156) vanishes in the $\beta \to 0$ limit. Finally, the dominated convergence theorem implies that, as $\beta \to 0$, the left-hand side of (2.3.156) converges to $\int_{Q\times\mathbb{R}} \chi(x,\xi,r)\, dx d\xi \Big|_{r=0}^{r=t} = \|u(\cdot,t)\|_{L^1(Q)} - \|u_0\|_{L^1(Q)}$. This completes the proof. □

Next, we present a result on the higher integrability of the parabolic and entropy defect measures in a neighbourhood of the origin. In turn, this result will help us to improve the H^1-regularity of $u^{\left[\frac{m+1}{2}\right]}$ prescribed by the definition of pathwise kinetic solution, and obtain Sobolev regularity for $u^{[m]}$ in Lemma 2.3.7. We shall need the following estimate, which follows immediately from Poincaré inequality.

Lemma 2.3.5. Let $v : Q \to \mathbb{R}$ be a measurable function such that $v^{[m+1/2]} \in H^1_0(Q)$. Then, for $C = C(Q)$, we have

$$\|v\|^{m+1}_{L^{m+1}(Q)} = \left\|v^{\left[\frac{m+1}{2}\right]}\right\|^2_{L^2(Q)} \leqslant C \left\|\nabla v^{\left[\frac{m+1}{2}\right]}\right\|^2_{L^2(Q)}.$$

Proposition 2.3.6. Let $u_0 \in L^2(Q)$ and $\delta \in (0,1]$ be arbitrary. Suppose that u is a pathwise kinetic solution of (2.1.1) with initial data u_0. Then, for each $T > 0$, for a constant $C = C(m, Q, T, A, z)$,

$$\|u\|^{1+\delta}_{L^\infty([0,T];L^{1+\delta}(Q))} + \delta \int_0^T \int_{Q\times\mathbb{R}} |\xi|^{\delta-1}(p+q)\, dx d\xi dr \leqslant C\left(\|u_0\|^{1+\delta}_{L^{1+\delta}(Q)} + \|u_0\|^2_{L^2(Q)}\right). \tag{2.3.158}$$

Proof. Let $\delta \in (0,1]$ be arbitrary. Let u be a pathwise kinetic solution with initial data $u_0 \in L^2(Q)$. We shall write χ for its kinetic function, p and q for its entropy and parabolic defect measure respectively, and $\mathcal{N} \subseteq (0,\infty)$ for its null set. The proof is based on an iterative argument along a suitably fine partition of $[0,T]$ and on the idea of formally testing the kinetic equation (2.2.26) against the function $\mathbb{R} \ni \xi \mapsto \xi^{[\delta]}$. Consider a partition $\mathcal{P} = \{t_0 = 0 < t_1 \cdots < t_N = T\}$ with the constraints that $\mathcal{P} \subseteq [0,T]\backslash\mathcal{N}$ and that the diameter $|\mathcal{P}| = \max_i |t_{i+1} - t_i|$ is suitably small, as specified by conditions (2.3.174) and (2.3.178) below. For each $\beta \in (0,1)$, consider the cutoff $\phi_\beta \in C^\infty_c(Q)$ introduced in (2.3.14). We stress that we consider this particular cutoff only to ease the referencing to follow, but its peculiar shape is not needed in this proof and any cutoff of scale β would work. According to Lemma 2.5.3, there exists a positive constant, which we denote by $C_T > 1$, such that

$$C_T^{-1}|\xi| \leqslant \left|\Pi_{t,s}^{x,\xi}\right| \leqslant C_T|\xi|, \quad \text{for each } (x,\xi) \in \mathbb{R}^d \times \mathbb{R} \text{ and } s \leqslant t \in [0,T]. \tag{2.3.159}$$

For each $\theta \in (0,1)$, we shall consider an approximation to the function $\mathbb{R} \ni \xi \to \xi^{[\delta]}$ with bounded derivatives. Namely, we introduce the piecewise C^1 function $g_\theta : \mathbb{R} \to \mathbb{R}$ defined by

$$g_\theta(\xi) := \int_0^\xi \dot{g}_\theta(\eta)\, d\eta, \qquad \text{and} \qquad \dot{g}_\theta(\xi) = \begin{cases} \delta\,(C_T\theta)^{\delta-1} & \text{if } |\xi| \leqslant C_T\theta, \\ \delta|\xi|^{\delta-1} & \text{if } C_T\theta \leqslant |\xi| \leqslant C_T\theta^{-1}, \\ 0 & \text{else.} \end{cases} \tag{2.3.160}$$

The following properties follow immediately

$$0 \leqslant \dot{g}_\theta(\xi) \leqslant \delta\,(C_T\theta)^{\delta-1} \text{ for each } \theta \in (0,1), \text{ and } \dot{g}_\theta(\xi) \uparrow \delta|\xi|^{\delta-1} \text{ as } \theta \to 0, \ \forall \xi \in \mathbb{R}. \tag{2.3.161}$$

Similarly, we have

$$|g_\theta(\xi)| \leqslant C_T^\delta \theta^{-\delta} \text{ for each } \theta \in (0,1), \text{ and } \lim_{\theta \to 0} g_\theta(\xi) = \xi^{[\delta]}, \ \forall \xi \in \mathbb{R}. \tag{2.3.162}$$

Now, for each $i = 0, \ldots, N-1$ and each $\theta, \beta \in (0,1)$, we test the kinetic equation (2.2.26) of u against the transport along characteristics, started from time $t_i \in \mathcal{P}$, of the function $\phi_\beta(x) g_\theta(\xi) \in C(Q \times \mathbb{R})$. An approximation argument shows that $\phi_\beta(x) g_\theta(\xi)$ is indeed an admissible test function. For any $t \in [t_i, t_{i+1}] \backslash \mathcal{N}$, after using the integration by parts formula (2.2.27), the kinetic equation becomes

$$\begin{aligned} 0 = & \int_{Q\times\mathbb{R}} \chi_r(x,\xi) \phi_\beta(Y_{r,r-t_i}^{x,\xi}) g_\theta(\Pi_{r,r-t_i}^{x,\xi})\, dx\, d\xi \Big|_{r=t_i}^{r=t} \\ & + \int_{t_i}^t \int_{Q\times\mathbb{R}} \phi_\beta(Y_{r,r-t_i}^{x,\xi}) \dot{g}_\theta(\Pi_{r,r-t_i}^{x,\xi}) \partial_\xi \Pi_{r,r-t_i}^{x,\xi} \,(p_r + q_r)\, dx\, d\xi\, dr + BE_i^{\text{vel}} \\ & + \frac{2m}{m+1} \int_{t_i}^t \int_Q |u|^{\frac{m-1}{2}} \nabla u^{\left[\frac{m+1}{2}\right]} \phi_\beta(Y_{r,r-t_i}^{x,u}) \dot{g}_\theta(\Pi_{r,r-t_i}^{x,u}) \nabla_x \Pi_{r,r-t_i}^{x,u}\, dx\, dr + BE_i^{\text{space}}, \end{aligned} \tag{2.3.163}$$

for the boundary velocity error relative to the interval $[t_i, t_{i+1}]$

$$BE_i^{\text{vel}} := \int_{t_i}^t \int_{Q\times\mathbb{R}} \nabla_y \phi_\beta(Y_{r,r-t_i}^{x,\xi}) \partial_\xi Y_{r,r-t_i}^{x,\xi} g_\theta(\Pi_{r,r-t_i}^{x,\xi})\,(p_r + q_r)\, dx\, d\xi\, dr, \tag{2.3.164}$$

and the boundary space error relative to the interval $[t_i, t_{i+1}]$

$$BE_i^{\text{space}} := \frac{2m}{m+1} \int_{t_i}^t \int_Q |u|^{\frac{m-1}{2}} \nabla u^{\left[\frac{m+1}{2}\right]} \nabla_y \phi_\beta(Y_{r,r-t_i}^{x,u}) D_x Y_{r,r-t_i}^{x,u} g_\theta(\Pi_{r,r-t_i}^{x,u})\, dx\, dr. \tag{2.3.165}$$

Now the idea is to let $\beta \to 0$ first and $\theta \to 0$ next, and to show that (2.3.163) yields the desired inequality (2.3.158) in the interval $[t_i, t_{i+1}]$. Then, an iteration argument along the partition $\mathcal{P}$ completes the proof.

Let us first show that the boundary errors vanish in the $\beta \to 0$ limit. The boundary velocity error term is handled with the same arguments (2.3.118)-(2.3.119) as for the error terms $BE_i^{\mathrm{sgn}1,3}$ in the uniqueness proof. Namely, using (2.3.121) with $k = 1$, estimate (2.5.11), and the bound (2.3.162) for g_θ with θ fixed, we conclude that

$$\lim_{\beta\to 0} \left|BE_i^{\mathrm{vel}}\right| = 0. \tag{2.3.166}$$

For the boundary space error, we start by splitting the integral according to the parameter $\theta \in (0,1)$:

$$\begin{aligned} BE_i^{\mathrm{space}} =& \frac{2m}{m+1}\int_{t_i}^{t}\int_{Q\cap\{|u|\geqslant\theta\}} |u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\nabla_y\phi_\beta(Y^{x,u}_{r,r-t_i})D_xY^{x,u}_{r,r-t_i}g_\theta(\Pi^{x,u}_{r,r-t_i})\,dx\,dr \\ &+\frac{2m}{m+1}\int_{t_i}^{t}\int_{Q\cap\{|u|<\theta\}} |u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\nabla_y\phi_\beta(Y^{x,u}_{r,r-t_i})D_xY^{x,u}_{r,r-t_i}g_\theta(\Pi^{x,u}_{r,r-t_i})\,dx\,dr. \end{aligned} \tag{2.3.167}$$

For the first term in (2.3.167) we compute

$$\begin{aligned} &\left|\frac{2m}{m+1}\int_{t_i}^{t}\int_{Q\cap\{|u|\geqslant\theta\}} |u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\nabla_y\phi_\beta(Y^{x,u}_{r,r-t_i})D_xY^{x,u}_{r,r-t_i}g_\theta(\Pi^{x,u}_{r,r-t_i})\,dx\,dr\right| \\ &\quad\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u|\geqslant\theta\}} \theta^{-1}\left|u^{\left[\frac{m+1}{2}\right]}\right|\left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|\beta^{-1}\,dx\,dr \\ &\quad\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u|\geqslant\theta\}} \left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|^2\,dx\,dr, \end{aligned} \tag{2.3.168}$$

for a constant $C = C(\theta,\delta,m,Q,T,A,z)$ independent of $\beta \in (0,1)$. In the first passage we used (2.3.121), (2.3.162) and $|u|^{\frac{m-1}{2}} \leqslant \theta^{-1}|u|^{\frac{m+1}{2}}$. In the second passage we used the Sobolev regularity $u^{\left[\frac{m+1}{2}\right]} \in H_0^1(Q)$ and the mean value theorem applied to points $x \in Q^\beta$. For the second term in (2.3.167), we preliminarily observe that, for any $(x,\xi) \in Q \times [t_i,T]$,

$$\text{if } |u| \leqslant \theta, \text{ then } \left|g_\theta(\Pi^{x,u}_{r,r-t_i})\right| \leqslant \delta C_T^\delta\theta^{\delta-1}|u|. \tag{2.3.169}$$

Indeed, when $|u| \leqslant \theta$, Lemma 2.5.3 guarantees that $\left|\Pi^{x,u}_{r,r-t_i}\right| \leqslant C_T|u| \leqslant C_T\theta$, and in turn definition (2.3.160) implies that $\left|g_\theta(\Pi^{x,u}_{r,r-t_i})\right| \leqslant \delta C_T^{\delta-1}\theta^{\delta-1}\left|\Pi^{x,u}_{r,r-t_i}\right|$. Then we calculate

$$\begin{aligned} &\left|\frac{2m}{m+1}\int_{t_i}^{t}\int_{Q\cap\{|u|<\theta\}} |u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\nabla_y\phi_\beta(Y^{x,u}_{r,r-t_i})D_xY^{x,u}_{r,r-t_i}g_\theta(\Pi^{x,u}_{r,r-t_i})\,dx\,dr\right| \\ &\quad\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u|<\theta\}} \left|u^{\left[\frac{m+1}{2}\right]}\right|\left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|\beta^{-1}\,dx\,dr \\ &\quad\leqslant C\int_{t_i}^{t_{i+1}}\int_{Q^\beta\cap\{|u|<\theta\}} \left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|^2\,dx\,dr, \end{aligned} \tag{2.3.170}$$

for a constant $C = C(\theta,\delta,m,Q,T,A,z)$ independent of $\beta \in (0,1)$. In the first passage we used observation (2.3.121) and observation (2.3.169). In the second passage we used $u^{\left[\frac{m+1}{2}\right]} \in H_0^1(Q)$ and the mean value theorem applied to points $x \in Q^\beta$. In conclusion, combining the splitting (2.3.167) with (2.3.168) and (2.3.170), the Sobolev regularity $u^{\left[\frac{m+1}{2}\right]} \in L^2([0,T];H_0^1(Q))$ and the dominated convergence theorem yield

$$\lim_{\beta\to 0}|BE_i| = 0. \tag{2.3.171}$$

This concludes the analysis of the error terms.

We now consider the remaining terms in (2.3.163). For the first term on the right-hand side of (2.3.163), owing to the integrability $u \in L^\infty([0,T];L^2(Q))$ from Definition 2.2.4, properties of the kinetic function, definition (2.3.160) and the proportionality (2.3.159), we apply the dominated convergence theorem and then use again formula (2.3.159) to get

$$\begin{aligned}\lim_{\theta\to 0}\lim_{\beta\to 0}\int_{Q\times\mathbb{R}}\chi_r\,\phi_\beta(Y^{x,\xi}_{r,r-t_i})g_\theta(\Pi^{x,\xi}_{r,r-t_i})\,dx\,d\xi\Big|_{r=t_i}^{r=t} &= \int_{Q\times\mathbb{R}}\chi_r\,(\Pi^{x,\xi}_{r,r-t_i})^{[\delta]}\,dx\,d\xi\Big|_{r=t_i}^{r=t} \\ &\geqslant C_T^{-1}\int_Q|u(x,t)|^{1+\delta}dx - C_T\int_Q|u(x,t_i)|^{1+\delta}dx.\end{aligned} \tag{2.3.172}$$

For the second term in (2.3.163), we first make the following observation. Namely, Proposition 2.5.1 and the equality $\partial_\xi\Pi^{x,\xi}_{r,0}\equiv 1\ \forall(x,\xi,r)\in\mathbb{R}^d\times\mathbb{R}\times[0,\infty)$ imply there exists a suitably small value, which we denote by $t^*>0$, such that, for any $(x,\xi)\in\mathbb{R}^d\times\mathbb{R}$ and any $s\leqslant r\in[0,T]$,

$$\text{if }\ s\leqslant t^*,\ \text{ then }\ \partial_\xi\Pi^{x,\xi}_{r,s}\geqslant\frac{3}{4}. \tag{2.3.173}$$

Therefore, if the partition $\mathcal{P}$ satisfies the condition

$$|\mathcal{P}| = \max_i|t_{i+1}-t_i|\leqslant t^*, \tag{2.3.174}$$

then observation (2.3.173) and the nonnegativity of the terms integrated yield

$$\begin{aligned}&\int_{t_i}^t\int_{Q\times\mathbb{R}}\phi_\beta(Y^{x,\xi}_{r,r-t_i})\,\dot g_\theta(\Pi^{x,\xi}_{r,r-t_i})\,\partial_\xi\Pi^{x,\xi}_{r,r-t_i}\,(p_r+q_r)\ dx\,d\xi\,dr \\ &\qquad\geqslant\frac{3}{4}\int_{t_i}^t\int_{Q\times\mathbb{R}}\phi_\beta(Y^{x,\xi}_{r,r-t_i})\,\dot g_\theta(\Pi^{x,\xi}_{r,r-t_i})\,(p_r+q_r)\ dx\,d\xi\,dr.\end{aligned} \tag{2.3.175}$$

For the fourth term on the right-hand side of (2.3.163), using (2.3.159) and (2.3.161), we preliminarily notice that

$$0\leqslant\dot g_\theta(\Pi^{x,u}_{r,r-t_i})\leqslant\delta\,C_T^{\delta-1}\,|u|^{\delta-1}\quad\forall(x,r)\in Q\times[t_i,T]. \tag{2.3.176}$$

Then we compute, for a constant $C^*=C^*(m,Q,T,A,z)$,

$$\begin{aligned}&\left|\frac{2m}{m+1}\int_{t_i}^t\int_Q|u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\phi_\beta(Y^{x,u}_{r,r-t_i})\dot g_\theta(\Pi^{x,u}_{r,r-t_i})\nabla_x\Pi^{x,u}_{r,r-t_i}\,dx\,dr\right| \\ &\leqslant C^*\int_{t_i}^t\int_Q|u|^{\frac{m-1}{2}}\left|\nabla_x\Pi^{x,u}_{r,r-t_i}\right|\dot g_\theta(\Pi^{x,u}_{r,r-t_i})^{\frac12} \\ &\qquad\qquad\cdot\phi_\beta(Y^{x,u}_{r,r-t_i})^{\frac12}\left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|\dot g_\theta(\Pi^{x,u}_{r,r-t_i})^{\frac12}\phi_\beta(Y^{x,u}_{r,r-t_i})^{\frac12}\,dx\,dr \\ &\leqslant C^*|t_{i+1}-t_i|^\alpha\int_{t_i}^t\int_Q|u|^{\frac{m+\delta}{2}}\delta^{\frac12}\phi_\beta(Y^{x,u}_{r,r-t_i})^{\frac12}\left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|\dot g_\theta(\Pi^{x,u}_{r,r-t_i})^{\frac12}\phi_\beta(Y^{x,u}_{r,r-t_i})^{\frac12}\,dx\,dr \\ &\leqslant C^*|\mathcal{P}|^\alpha\left(\delta\int_{t_i}^t\int_Q|u|^{m+\delta}dx\,dr+\int_{t_i}^t\int_{Q\times\mathbb{R}}\dot g_\theta(\Pi^{x,\xi}_{r,r-t_i})\phi_\beta(Y^{x,\xi}_{r,r-t_i})q_r(x,\xi)\,dx\,dr\right) \\ &\leqslant C^*|\mathcal{P}|^\alpha\left(\delta\left(\int_{t_i}^t\int_{Q\times\mathbb{R}}q_r\,dx\,d\xi\,dr\right)^{\frac{m+\delta}{m+1}}+\int_{t_i}^t\int_{Q\times\mathbb{R}}\dot g_\theta(\Pi^{x,\xi}_{r,r-t_i})\phi_\beta(Y^{x,\xi}_{r,r-t_i})q_r(x,\xi)\,dx\,dr\right).\end{aligned} \tag{2.3.177}$$

In the second passage we used (2.3.176) and (2.5.3). The third passage follows from Hölder's inequality and the definition of parabolic defect measure. In the last passage we controlled the first term with Hölder's inequality and Lemma 2.3.5.

Therefore, if the partition $\mathcal{P}$ satisfies the further condition, for the constant C^* from (2.3.177),

$$C^*|\mathcal{P}| < \frac{1}{4}, \tag{2.3.178}$$

then, for the second and fourth term in (2.3.163), combining (2.3.175) and (2.3.177), the monotone convergence theorem, property (2.3.161) and the proportionality (2.3.159) yield

$$\begin{aligned}
\liminf_{\theta\to 0}\liminf_{\beta\to 0} & \int_{t_i}^{t}\int_{Q\times\mathbb{R}} \phi_\beta(Y^{x,\xi}_{r,r-t_i})\dot{g}_\theta(\Pi^{x,\xi}_{r,r-t_i})\partial_\xi \Pi^{x,\xi}_{r,r-t_i}\,(p_r+q_r)\;dx\,d\xi\,dr \\
& + \frac{2m}{m+1}\int_{t_i}^{t}\int_{Q}|u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\phi_\beta(Y^{x,u}_{r,r-t_i})\dot{g}_\theta(\Pi^{x,u}_{r,r-t_i})\nabla_x \Pi^{x,u}_{r,r-t_i}\;dx\,dr \\
\geqslant \lim_{\theta\to 0}\lim_{\beta\to 0} & \frac{2}{4}\int_{t_i}^{t}\int_{Q\times\mathbb{R}}\phi_\beta(Y^{x,\xi}_{r,r-t_i})\dot{g}_\theta(\Pi^{x,\xi}_{r,r-t_i})\,(p_r+q_r)\;dx\,d\xi\,dr - \frac{\delta}{4}\left(\int_{t_i}^{t}\int_{Q\times\mathbb{R}} q_r\,dx\,d\xi\,dr\right)^{\frac{m+\delta}{m+1}} \\
\geqslant \frac{1}{2}\,\delta\, C_T^{\delta-1} & \int_{t_i}^{t}\int_{Q\times\mathbb{R}}|\xi|^{\delta-1}\,(p_r+q_r)\;dx\,d\xi\,dr - \frac{\delta}{4}\left(\int_{t_i}^{t}\int_{Q\times\mathbb{R}} q_r\,dx\,d\xi\,dr\right)^{\frac{m+\delta}{m+1}}.
\end{aligned} \tag{2.3.179}$$

We now conclude the proof with an iterative argument along the partition $\mathcal{P}$. Let us first consider the case $\delta = 1$. Passing to the limit $\beta \to 0$ first and $\theta \to 0$ next in equation (2.3.163), using (2.3.166), (2.3.171), (2.3.172), and (2.3.179) with $\delta = 1$, and recalling that $t \in [t_i, t_{i+1}]\backslash\mathcal{N}$ is arbitrary, we obtain

$$C_T^{-1}\|u\|^2_{L^\infty([t_i,t_{i+1}];L^2(Q))} + \frac{1}{4}\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}(p_r+q_r)\;dx\,d\xi\,dr \leqslant C_T\|u(\cdot,t_i)\|^2_{L^2(Q)}.$$

Then, arguing iteratively first in the interval $[0, t_1]$, then in $[t_1, t_2]$, and so forth up to $[t_{N-1}, T]$, we conclude that

$$\|u\|^2_{L^\infty([0,T];L^2(Q))} + \int_0^T\int_{Q\times\mathbb{R}}(p_r+q_r)\;dx\,d\xi\,dr \leqslant C\|u_0\|^2_{L^2(Q)}, \tag{2.3.180}$$

for a constant $C = C(m, Q, T, A, z)$ depending on the constant C_T from (2.3.159) and on the cardinality of the partition $\mathcal{P}$, that is on the conditions (2.3.174) and (2.3.178).

Finally, let us consider any arbitrary $\delta \in (0, 1]$. Taking again the limit $\beta \to 0$ first and $\theta \to 0$ next in equation (2.3.163), using (2.3.166), (2.3.171), (2.3.172) and (2.3.179), and recalling that $t \in [t_i, t_{i+1}]\backslash\mathcal{N}$ is arbitrary, we obtain

$$\begin{aligned}
C_T^{-1}\|u\|^{1+\delta}_{L^\infty([t_i,t_{i+1}];L^{1+\delta}(Q))} &+ \frac{\delta}{2}C_T^{\delta-1}\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}|\xi|^{\delta-1}\,(p_r+q_r)\;dx\,d\xi\,dr \\
&\leqslant C_T\|u(\cdot,t_i)\|^{1+\delta}_{L^{1+\delta}(Q)} + \frac{\delta}{4}\left(\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}} q_r\,dx\,d\xi\,dr\right)^{\frac{m+\delta}{m+1}}.
\end{aligned}$$

In conclusion, arguing iteratively over the intervals $[t_i, t_{i+1}]$, and then using Young's inequality and the newly found estimate (2.3.180), we deduce

$$\|u\|^{1+\delta}_{L^\infty([0,T];L^{1+\delta}(Q))}+\delta\int_0^T\!\!\int_{Q\times\mathbb{R}}|\xi|^{\delta-1}(p_r+q_r)\,dxd\xi dr \leqslant C\left(\|u_0\|^{1+\delta}_{L^{1+\delta}(Q)}+\delta\left(\int_0^T\!\!\int_{Q\times\mathbb{R}}q_r\,dxd\xi dr\right)^{\frac{m+\delta}{m+1}}\right)$$
$$\leqslant C\left(\|u_0\|^{1+\delta}_{L^{1+\delta}(Q)}+\|u_0\|^2_{L^2(Q)}\right),$$

for a constant $C=C(m,Q,T,A,z)$ depending on C_T from (2.3.159) and on the cardinality of $\mathcal{P}$. ☐

As anticipated, we now exploit the above proposition to prove that, if u is a pathwise kinetic solution, the power $u^{[m]}$ lies in a suitable Sobolev space $W^{1,p_m}_0(Q)$, for an exponent p_m depending on the diffusion regime m. In particular, we can pass the vanishing boundary conditions at the level of $u^{[m]}$. We used this property in the proof of Theorem 2.1.2 to handle the error terms coming from the cutoff procedure. Namely, Proposition 2.3.6 and a straightforward modification of [FG21a, Lemma A.1] prove the following.

Lemma 2.3.7. Let $u_0\in L^2(Q)$ and suppose that u is a pathwise kinetic solution of (2.1.1) with initial data u_0. Then, for $p_m:=(\frac{m+1}{m}\wedge 2)$, for any $T>0$ we have

$$u^{[m]}\in L^{p_m}\left([0,T];W^{1,p_m}_0(Q)\right)\quad\text{with}\quad\nabla u^{[m]}=\frac{2m}{m+1}|u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}. \tag{2.3.181}$$

In particular $u^{[m]}$ has vanishing trace, and we have the estimate, for $C=C(m,Q,T)$,

$$\|u^{[m]}\|_{L^{p_m}([0,T];W^{1,p_m}_0(Q))}\leqslant C\left(1+\|u_0\|^2_{L^2(Q)}\right).$$

Finally, we extend Proposition 2.3.6 to the case $\delta=0$ and establish a bound on the first singular moment of the defect measures. This result is used in the proof of uniqueness for diffusion exponents $m\in(0,1)\cup(1,2]$. Informally, it implies the local L^2-integrability of $\nabla u^{\frac{m}{2}}$.

Remark 2.3.8. We require the nonnegativity of the initial data, and indeed Proposition 2.3.9 is false for signed data. Consider for simplicity the case $d=m=1$ and $A(x,\xi)=0$. Suppose that $u_0(x)=x$ in a neighbourhood of the origin. Then, since the heat flow preserves the linear behaviour of the initial data locally in time, the failure of Proposition 2.3.9 manifests as the non-integrability of the map $\mathbb{R}\ni x\mapsto 1/|x|$ in a neighbourhood of the origin.

Proposition 2.3.9. Suppose that $u_0\in L^2_+(Q)$ and let u be a pathwise kinetic solution of (2.1.1) with initial data u_0. Then, for each $T>0$, there exists $C=C(m,Q,T,A,z)$ such that

$$\int_0^T\int_{\mathbb{R}\times Q}|\xi|^{-1}(p+q)\,dx\,d\xi\,dr\leqslant C\left(1+\|u_0\|^2_{L^2(Q)}\right). \tag{2.3.182}$$

Proof. The proof is essentially the same as for Proposition 2.3.6, except that now $\delta=0$ and the idea is to formally test the kinetic equation against the function $\xi\mapsto\log(\xi)$. We stress that in order for this to work, that is in order to find an approximation of $\xi \mapsto \log(\xi)$ whose derivative is bounded below and increases to the function $\xi \mapsto |\xi|^{-1}$, so as to justify the analogous of the monotone convergence passage (2.3.179), it is crucial to consider positive values of ξ only. In other words, it is crucial that the kinetic solution u is nonnegative, as ensured by Proposition 2.3.4.

We now sketch the relevant modifications. With the notation from the proof of Proposition 2.3.6, fix a partition $\mathcal{P}$ satisfying conditions (2.3.174) and (2.3.178). The approximation $g_\theta : \mathbb{R} \to \mathbb{R}$ is now defined by

$$g_\theta(\xi) := \int_1^\xi \dot{g}_\theta(\eta)\, d\eta, \quad \text{and} \quad \dot{g}_\theta(\xi) = \begin{cases} C_T^{-1}\theta^{-1} & \text{if } 0 \leqslant \xi \leqslant C_T\theta, \\ |\xi|^{-1} & \text{if } C_T\theta \leqslant \xi \leqslant C_T\theta^{-1}, \\ 0 & \text{if } \xi \leqslant 0 \text{ or } C_T\theta^{-1} < \xi. \end{cases} \tag{2.3.183}$$

The properties (2.3.161) and (2.3.162) are replaced respectively by

$$0 \leqslant \dot{g}_\theta(\xi) \leqslant C_T^{-1}\theta^{-1} \text{ for each } \theta \in (0,1), \text{ and } \dot{g}_\theta(\xi) \uparrow |\xi|^{-1} \text{ as } \theta \to 0, \ \forall \xi \geqslant 0,$$

and by

$$|g_\theta| \leqslant C\log(C_T\theta^{-1}) \text{ for each } \theta \in (0,1), \text{ and } \lim_{\theta\to 0} g_\theta(\xi) = \log(\xi), \ \forall \xi \geqslant 0. \tag{2.3.184}$$

Testing the kinetic equation against $\phi_\beta(x)g_\theta(\xi) \in C(Q \times \mathbb{R})$ we still get equation (2.3.163), with the only difference that g_θ is now defined by (2.3.183). The error term BE_i^{vel}, still defined by (2.3.164), is handled exactly as before, simply using (2.3.184) in place of (2.3.162). For the error term BE_i^{space}, still defined by (2.3.165), we consider again the splitting (2.3.167). The first term in (2.3.167) is handled exactly as in (2.3.168), simply using (2.3.184) in place of (2.3.162). The only difference is that, for the second term in (2.3.167), we now compute

$$\begin{aligned} &\left| \frac{2m}{m+1} \int_{t_i}^t \int_{Q\cap\{|u|<\theta\}} |u|^{\frac{m-1}{2}} \nabla u^{\left[\frac{m+1}{2}\right]} \nabla_y \phi_\beta(Y_{r,r-t_i}^{x,u}) D_x Y_{r,r-t_i}^{x,u} g_\theta(\Pi_{r,r-t_i}^{x,u})\, dx\, dr \right| \\ &\leqslant C \int_{t_i}^{t_{i+1}} \int_{Q^\beta\cap\{|u|<\theta\}} |u|^{\frac{m-1}{m+1}} \left| u^{\left[\frac{m+1}{2}\right]} \right| \left| \nabla u^{\left[\frac{m+1}{2}\right]} \right| \beta^{-1}\, dx\, dr \\ &\leqslant C \int_{t_i}^{t_{i+1}} \int_{Q^\beta\cap\{|u|<\theta\}} |u|^{\frac{m-1}{m+1}} \left| \nabla u^{\left[\frac{m+1}{2}\right]} \right|^2 dx\, dr \\ &\leqslant C \int_{t_i}^{t_{i+1}} \int_{Q^\beta\times\mathbb{R}\cap\{|\xi|<\theta\}} |\xi|^{\frac{m-1}{m+1}\wedge 0} \theta^{\frac{m-1}{m+1}\vee 0} q_r(x,\xi)\, dx\, d\xi\, dr, \end{aligned} \tag{2.3.185}$$

for a constant $C = C(\theta, m, Q, T, A, z)$ independent of $\beta \in (0,1)$. In the first passage we used observation (2.3.121) and (2.3.184). In the second passage we used the Sobolev regularity $u^{\left[\frac{m+1}{2}\right]} \in H_0^1(Q)$ and the mean value theorem. The last passage follows from the definition of parabolic defect measure. Now, formulas (2.3.167), (2.3.168), and (2.3.185) coupled with Proposition 2.3.6, if $m \in (0,1)$, or simply with the finiteness of the measure q, if $m \in [1,\infty)$, and the dominated convergence theorem yield

$$\lim_{\beta\to 0}|BE_i| = 0. \tag{2.3.186}$$

For the first term on the right-hand side of (2.3.163), the integrability $u \in L^\infty([0,T];L^2(Q))$, the nonnegativity of u, which follows from Proposition 2.3.4, the integrability of $\xi \mapsto \log(\xi)$ near 0 and its growth at infinity, properties of the kinetic function, the proportionality (2.3.159) and the dominated convergence theorem imply that

$$\begin{aligned}\lim_{\theta\to 0}\lim_{\beta\to 0}\int_{Q\times\mathbb{R}}\chi_r(x,\xi)\,\phi_\beta(Y^{x,\xi}_{r,r-t_i})g_\theta(\Pi^{x,\xi}_{r,r-t_i})\,dx\,d\xi\Big|_{r=t_i}^{r=t_{i+1}} &= \int_{Q\times\mathbb{R}}\chi_r(x,\xi)\,\log(\Pi^{x,\xi}_{r,r-t_i})\,dx\,d\xi\Big|_{r=t_i}^{r=t_{i+1}} \\ &\leqslant C\left(1+\|u\|^2_{L^\infty([t_i,t_{i+1}];L^2(Q))}\right),\end{aligned} \tag{2.3.187}$$

for a constant $C = C(T,A,z)$. The second and fourth term in (2.3.163) are handled with the same arguments as in (2.3.175)-(2.3.179), except that g_θ is now given by (2.3.183) and estimate (2.3.176) is replaced by

$$0 \leqslant \dot{g}_\theta(\Pi^{x,u}_{r,r-t_i}) \leqslant C_T^{-1}|u|^{-1}\;\; \forall (x,r)\in Q\times[t_i,T].$$

In turn, the final inequality becomes

$$\begin{aligned}&\liminf_{\theta\to 0}\liminf_{\beta\to 0}\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}\phi_\beta(Y^{x,\xi}_{r,r-t_i})\,\dot{g}_\theta(\Pi^{x,\xi}_{r,r-t_i})\,\partial_\xi\Pi^{x,\xi}_{r,r-t_i}\,(p_r+q_r)\;dx\,d\xi\,dr \\ &\qquad + \frac{2m}{m+1}\int_{t_i}^{t_{i+1}}\int_Q |u|^{\frac{m-1}{2}}\nabla u^{\left[\frac{m+1}{2}\right]}\phi_\beta(Y^{x,u}_{r,r-t_i})\dot{g}_\theta(\Pi^{x,u}_{r,r-t_i})\nabla_x\Pi^{x,u}_{r,r-t_i}\,dx\,dr \\ &\geqslant \frac{1}{2}C_T^{-1}\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}|\xi|^{-1}\,(p_r+q_r)\;dx\,d\xi\,dr - \frac{1}{4}\left(\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}q_r\,dx\,d\xi\,dr\right)^{\frac{m}{m+1}}.\end{aligned} \tag{2.3.188}$$

In conclusion, passing to the limit as $\beta \to 0$ first and $\theta \to 0$ then in (2.3.163), using (2.3.166), (2.3.186), (2.3.187) and (2.3.188), we obtain that

$$\int_{t_i}^{t_{i+1}}dr\int_{Q\times\mathbb{R}}|\xi|^{-1}\,(p_r+q_r)\;dx\,d\xi \leqslant C\left(1+\|u\|^2_{L^\infty([t_i,t_{i+1}]L^2(Q))}+\left(\int_{t_i}^{t_{i+1}}\int_{Q\times\mathbb{R}}q_r\,dx\,d\xi\,dr\right)^{\frac{m}{m+1}}\right), \tag{2.3.189}$$

for a constant $C = C(T,A,z)$. Iterating (2.3.189) over the partition $\mathcal{P}$, using Young's inequality and estimate (2.3.180), we obtain formula (2.3.182). □

2.4 Existence of pathwise kinetic solutions

In this section we establish the existence of pathwise kinetic solutions to equation (2.1.1). For this, we consider the setting outlined in Section 2.2 and derive stable estimates for the regularized equation (2.2.2), defined for each $\eta \in (0,1)$ and $\epsilon \in (0,1)$.

Let $u_0 \in L^2(Q)$ and let $\eta,\epsilon \in (0,1)$ be fixed but arbitrary. Recall that $z^\epsilon : [0,\infty) \to \mathbb{R}^n$ are smooth paths converging to z with respect to the α-Hölder metric on the space of geometric rough paths $C^{0,\alpha}\big([0,T];G^{\lfloor 1/\alpha\rfloor}(\mathbb{R}^n)\big)$, for each $T>0$. Let $\mathrm{u}^{\eta,\epsilon}$ be the solution of (2.2.2) from

Proposition 2.2.1 and let $\chi^{\eta,\epsilon}(x,\xi,t) := \bar{\chi}(\mathrm{u}^{\eta,\epsilon}(x,t),\xi)$ be its kinetic function and $\mathrm{p}^{\eta,\epsilon}$ and $\mathrm{q}^{\eta,\epsilon}$ the associated entropy and parabolic defect measures. In Proposition 2.2.2 we established the kinetic function is a distributional solution of equation (2.2.4). In Corollary 2.2.3 we got rid of the noise in the equation by the testing it against test functions transported along the smooth backward characteristics (2.2.12) and showed the kinetic function solves equation (2.2.20).

The idea is now to establish stable estimates for the solutions $\mathrm{u}^{\eta,\epsilon}$ and the associated kinetic functions $\chi^{\eta,\epsilon}$, that allow us to pass to the limit $\eta,\epsilon \to 0$ and find a coherent limit for the functions $\mathrm{u}^{\eta,\epsilon}$ and the equations (2.2.20).

As in Section 4, consider for motivation to the kinetic formulation (2.3.2) of the deterministic porous media equation. Following [CP03], estimates for the solution are obtained by testing the equation against the maps $\mathbb{R} \ni \xi \mapsto \operatorname{sgn}(\xi)$ and $\mathbb{R} \ni \xi \mapsto \xi$. In the first case, owing to the positivity of the parabolic and entropy defect measures, we informally get

$$\|u\|_{L^\infty([0,\infty);L^1(Q))} = \|\chi\|_{L^\infty([0,\infty);L^1(Q\times\mathbb{R}))} \leqslant \|\chi(u_0,\xi)\|_{L^1(Q\times\mathbb{R})} = \|u_0\|_{L^1(Q)}.$$

In the second case, we informally observe the estimate

$$\frac{1}{2}\|u\|^2_{L^\infty([0,\infty);L^2(Q))} + \int_0^\infty \int_{\mathbb{R}\times Q} (p(x,\xi,r) + q(x,\xi,r))\, dx\, d\xi\, dr \leqslant \frac{1}{2}\|u_0\|^2_{L^2(Q)}.$$

Remark 2.4.1. In the discussion to follow, we will first establish estimates and the existence of pathwise kinetic solutions for initial data $u_0 \in C_c^\infty(Q)$. The general results will follow by density, repeating the arguments presented.

In Proposition 2.4.2 we obtain the analogue of the L^1-estimate, and in Proposition 2.4.3 we obtain the analogue of the L^2-estimate and the estimate for the defect measures. The argument for Proposition 2.4.2 is just a small modification of the relevant details of Corollary 2.3.3. The proof of Proposition 2.4.3 is essentially identical to that of Proposition 2.3.6. We therefore omit the details.

Proposition 2.4.2. For each $u_0 \in L^2(Q)$, $\eta \in (0,1)$ and $\epsilon \in (0,1)$, the solution $\mathrm{u}^{\eta,\epsilon}$ of (2.2.2) from Proposition 2.2.1 satisfies

$$\| \mathrm{u}^{\eta,\epsilon} \|_{L^\infty([0,\infty);L^1(Q))} \leqslant \|u_0\|_{L^1(Q)}.$$

Proposition 2.4.3. For each $u_0 \in L^2(Q)$, $\eta \in (0,1)$ and $\epsilon \in (0,1)$, let $\mathrm{u}^{\eta,\epsilon}$ be the solution of (2.2.2) from Proposition 2.2.1. Let $\delta \in (0,1)$. For each $T > 0$, there exists $C = C(m,Q,T,A,z)$ such that

$$\| \mathrm{u}^{\eta,\epsilon} \|^2_{L^\infty([0,T];L^2(Q))} + \int_0^T \int_{\mathbb{R}\times Q} \left(1 + \delta|\xi|^{\delta-1}\right)(\mathrm{p}^{\eta,\epsilon} + \mathrm{q}^{\eta,\epsilon})\, dx d\xi dr \leqslant C\left(1 + \|u_0\|^2_{L^2(Q)}\right).$$

In general, we do not expect to obtain a stable estimate in time for the solutions $\mathrm{u}^{\eta,\epsilon}$. However, we can obtain some regularity for the time derivative of the transported kinetic functions $\tilde{\chi}^{\eta,\epsilon}$, defined by

$$\tilde{\chi}^{\eta,\epsilon}(x,\xi,t) := \chi^{\eta,\epsilon}(X^{x,\xi,\epsilon}_{0,t}, \Xi^{x,\xi,\epsilon}_{0,t}, t) \quad \text{for} \quad (x,\xi,t) \in Q \times \mathbb{R}^d \times [0,\infty). \tag{2.4.1}$$

In practice, the transport cancels the oscillations introduced by the noise. The following proposition proves the functions $\partial_t \tilde{\chi}^{\eta,\epsilon}$ are uniformly bounded in the negative Sobolev space H^{-s}, for s big enough.

Proposition 2.4.4. For each $u_0 \in L^2(Q)$, $\eta \in (0,1)$ and $\epsilon \in (0,1)$, consider the transported kinetic function (2.4.1). For each $T > 0$, for any Sobolev exponent $s > \frac{d}{2} + 1$, there exists $C = C(m, z, Q, T, s)$ such that

$$\|\partial_t \tilde{\chi}^{\eta,\epsilon}\|_{L^1([0,T];H^{-s}(Q\times\mathbb{R}))} \leqslant \left(1 + \|u_0\|^2_{L^2(Q)}\right).$$

Proof. For any $\rho_0 \in C^\infty_c(Q \times \mathbb{R})$ denote $\rho^\epsilon_{0,r}(x,\xi) := \rho_0\left(Y^{x,\xi,\epsilon}_{r,r}, \Pi^{x,\xi,\epsilon}_{r,r}\right)$. Take any $\varphi \in C^\infty_c([0,\infty))$. Testing equation (2.2.7) against $\psi(x,\xi,t) := \rho^\epsilon_{0,r}(x,\xi)\varphi(t)$ and exploiting the cancellations coming from the transport of ρ_0 along the characteristics, for any $a \leqslant b \in [0,\infty)$, we get

$$\begin{aligned}\int_a^b\!\!\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}(x,\xi,t)\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\dot{\varphi}(t)\,dt = &\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}(x,\xi,t)\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\varphi(t)\Big|_{t=a}^{t=b} \\ &- \int_a^b\!\!\int_{Q\times\mathbb{R}} \left(m|\xi|^{m-1} + \eta\right)\chi^{\eta,\epsilon}(x,\xi,t)\Delta_x\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\varphi(t)\,dt \\ &+ \int_a^b\!\!\int_{Q\times\mathbb{R}} (\mathrm{p}^{\eta,\epsilon}(x,\xi,t) + \mathrm{q}^{\eta,\epsilon}(x,\xi,t))\,\partial_\xi\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\varphi(t)\,dt.\end{aligned}$$

Using the conservative property (2.2.14) of the characteristics and integrating by parts the second term on the right-hand side of the equation, we obtain

$$\begin{aligned}\int_a^b\!\!\int_{Q\times\mathbb{R}} \tilde{\chi}^{\eta,\epsilon}(x,\xi,t)\rho_0(x,\xi)\,dx\,d\xi\,\dot{\varphi}(t)\,dt = &\int_{Q\times\mathbb{R}} \chi^{\eta,\epsilon}(x,\xi,t)\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\varphi(t)\Big|_{t=a}^{t=b} \\ &+ \int_a^b\!\!\int_Q \left(\frac{2m}{m+1}\,|\mathrm{u}^{\eta,\epsilon}|^{\frac{m-1}{2}}\,\nabla\,(\mathrm{u}^{\eta,\epsilon})^{\left[\frac{m+1}{2}\right]} + \eta\nabla\,\mathrm{u}^{\eta,\epsilon}\right)\nabla_x\rho^\epsilon_{0,t}(x,\mathrm{u}^{\eta,\epsilon})dx\,\varphi(t)\,dt \\ &+ \int_a^b\!\!\int_{Q\times\mathbb{R}} (\mathrm{p}^{\eta,\epsilon}(x,\xi,t) + \mathrm{q}^{\eta,\epsilon}(x,\xi,t))\,\partial_\xi\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi\,\varphi(t)\,dt.\end{aligned}$$

Since $\varphi \in C^\infty_c([0,\infty))$ is arbitrary, this shows that, for any $\rho_0 \in C^\infty_c(Q \times \mathbb{R})$, the mapping

$$[t_0,\infty) \ni t \mapsto \int_{Q\times\mathbb{R}} \tilde{\chi}^{\eta,\epsilon}(x,\xi,t)\rho_0(x,\xi)\,dx\,d\xi$$

has a weak derivative given by

$$\begin{aligned}[t_0,\infty) \ni t \mapsto &-\int_Q \left(\frac{2m}{m+1}\,|\mathrm{u}^{\eta,\epsilon}|^{\frac{m-1}{2}}\,\nabla\,(\mathrm{u}^{\eta,\epsilon})^{\left[\frac{m+1}{2}\right]} + \eta\nabla\,\mathrm{u}^{\eta,\epsilon}\right)\nabla_x\rho^\epsilon_{0,t}(x,\mathrm{u}^{\eta,\epsilon})\,dx \\ &- \int_{Q\times\mathbb{R}} (\mathrm{p}^{\eta,\epsilon}(x,\xi,t) + \mathrm{q}^{\eta,\epsilon}(x,\xi,t))\,\partial_\xi\rho^\epsilon_{0,t}(x,\xi)\,dx\,d\xi.\end{aligned}$$

The embedding theorem for Sobolev spaces ensures that for any $s > \frac{d}{2} + 1$, for $C = C(Q,s)$, $\|\rho\|_{W^{1,\infty}(Q\times\mathbb{R})} \leqslant C\|\rho_0\|_{H^s(Q\times\mathbb{R})}$. Then we use this bound, Cauchy's inequality and the definition of the parabolic defect measure to estimate

$$|\langle \partial_t \tilde{\chi}^{\eta,\epsilon}, \rho_0\rangle| \leqslant C\left\{\eta + \int_Q |\mathrm{u}^{\eta,\epsilon}|^{(m-1)\vee 0}\, dx + \int_{Q\times\mathbb{R}} \left(\mathrm{p}^{\eta,\epsilon} + \left(1+|\xi|^{m-1\wedge 0}\right)\mathrm{q}^{\eta,\epsilon}\right) dx\, d\xi\right\} \cdot \|\rho_0\|_{H^s(Q\times\mathbb{R})},$$

for $C = C(m, z, Q, s)$. Finally, we use the density of $C_c(Q \times \mathbb{R})$ in $H^s_0(Q \times \mathbb{R})$ and then we integrate in time to get

$$\begin{aligned}\|\partial_t \tilde{\chi}^{\eta,\epsilon}\|_{L^1([0,T];H^{-s}(Q\times\mathbb{R}))} &\leqslant C\int_0^T dt \left\{\eta + \int_Q |\mathrm{u}^{\eta,\epsilon}|^{(m-1)\vee 0}\, dx + \int_{Q\times\mathbb{R}} \left(\mathrm{p}^{\eta,\epsilon} + \left(1+|\xi|^{m-1\wedge 0}\right)\mathrm{q}^{\eta,\epsilon}\right) dx\, d\xi\right\}\\ &\leqslant C\left(1 + \|u_0\|_{L^2(Q)}\right),\end{aligned}$$

for $C = C(m, z, Q, T, s)$. In the last inequality we used Lemma 2.3.5, Proposition 2.4.3, and Young's and Hölder's inequalities. □

It remains to establish the regularity of the kinetic functions with respect to the spatial and velocity variables. Actually, we shall consider the transported kinetic functions, since these are the ones we proved regularity in time for. Straightforward modifications to [FG19, Corollary 5.5] prove the following estimate in the fractional Sobolev space $W^{\ell,1}(Q \times \mathbb{R})$, for ℓ suitably small.

Proposition 2.4.5. For each $u_0 \in L^2(Q)$, $\eta \in (0, 1)$ and $\epsilon \in (0, 1)$, consider the transported kinetic function (2.4.1). For each $T > 0$ and for each $\ell \in (0, \frac{2}{m+1} \wedge 1)$, there exists $C = C(m, Q, T, \ell)$ such that

$$\|\tilde{\chi}^{\eta,\epsilon}\|_{L^1([0,T];W^{\ell,1}(Q\times\mathbb{R}))} \leqslant C\left(1 + \|u_0\|^2_{L^2(Q)}\right).$$

We can now establish the existence of pathwise kinetic solutions with initial data $u_0 \in L^2(Q)$. The proof is a consequence of Propositions 2.4.3, 2.4.4 and 2.4.5, and the Aubin-Lions-Simon Lemma [Sim86]. We remark that the necessity of an entropy defect measure in Definition 2.2.4 arises from the fact that, after passing to a subsequence, the gradients $\nabla(\mathrm{u}^{\eta,\epsilon})^{[m+1/2]}$ will converge only weakly in the $\epsilon, \eta \to 0$ limit. Due to the weak lower semicontinuity of the norm, the limit of the parabolic defect measures $\mathrm{q}^{\eta,\epsilon}$ may therefore overestimate the energy of the signed power of the limiting solution. The total mass of the entropy defect measure quantifies this loss.

Proof of Theorem 2.1.3.
Let $u_0 \in L^2(Q)$ be arbitrary. For any $\eta, \epsilon \in (0, 1)$, let $\mathrm{u}^{\eta,\epsilon}$ be the solution of the regularized equation (2.2.2) with initial data u_0, with transported kinetic function $\tilde{\chi}^{\eta,\epsilon}$, entropy defect measure $\mathrm{p}^{\eta,\epsilon}$ and parabolic defect measure $\mathrm{q}^{\eta,\epsilon}$. We recall that, for each $\ell \in (0, \frac{2}{m+1} \wedge 1)$ and each $R > 0$, the embedding of $W^{\ell,1}(Q \times [-R, R])$ into $L^1(Q \times [-R, R])$ is compact, and that $L^1(Q \times [-R, R])$ embeds continuously into $H^{-s}(Q \times \mathbb{R})$ for $s > \frac{d}{2} + 1$. Then, Propositions 2.4.4 and 2.4.5, the Aubin–Lions–Simon Lemma [Sim86] and a diagonal argument needed to control the tail of $Q \times \mathbb{R}$ show that, for each $T > 0$, the family

$$\{\tilde{\chi}^{\eta,\epsilon}\}_{\eta,\epsilon\in(0,1)} \text{ is precompact in } L^1([0,T]; L^1(Q \times \mathbb{R})).$$

The conservative property of the characteristics (2.2.14) then implies that, for each $T > 0$, the family

$$\{\chi^{\eta,\epsilon}\}_{\eta,\epsilon\in(0,1)} \text{ is precompact in } L^1([0,T];L^1(Q\times\mathbb{R})).$$

In turn, the definition of kinetic function immediately shows that, for each $T > 0$, the family

$$\{\mathrm{u}^{\eta,\epsilon}\}_{\eta,\epsilon\in(0,1)} \text{ is precompact in } L^1([0,T];L^1(Q)).$$

Furthermore, Proposition 2.4.3 and the Riesz–Markov Theorem imply that the sequence of measures

$$\{(\mathrm{p}^{\eta,\epsilon},\mathrm{q}^{\eta,\epsilon})\}_{\eta,\epsilon\in(0,1)} \text{ is weakly precompact in } C_c(Q\times\mathbb{R}\times[0,T])^*,$$

and that the family

$$\{(\mathrm{u}^{\eta,\epsilon})^{\left[\frac{m+1}{2}\right]}\}_{\eta,\epsilon\in(0,1)} \text{ is weakly precompact in } L^2([0,T];H^1_0(Q)).$$

After passing to a subsequence $\{(\eta_k,\epsilon_k)\}_{k\in\mathbb{N}}$ with $\lim_{k\to\infty}(\eta_k,\epsilon_k) = (0,0)$, it follows that there exists a function u such that $u \in L^1([0,T];L^1(Q))$ for any $T > 0$, and such that as $k \to \infty$

$$u^{\eta_k,\epsilon_k} \to u \quad \text{strongly in } L^1([0,T];L^1(Q)) \tag{2.4.2}$$

and

$$(u^{\eta_k,\epsilon_k})^{[m+1/2]} \rightharpoonup u^{[m+1/2]} \text{ weakly in } L^2([0,T];H^1_0(Q)). \tag{2.4.3}$$

Furthermore, there exist positive measures (p', q') such that as $k \to \infty$, for any $T > 0$,

$$(\mathrm{p}^{\eta,\epsilon},\mathrm{q}^{\eta,\epsilon}) \rightharpoonup (p',q') \quad \text{weakly in } C_c(Q\times\mathbb{R}\times[0,T])^*. \tag{2.4.4}$$

Recalling the definition (2.2.6) of the parabolic defect measure, it follows from the strong convergence (2.4.2) and the weak lower semicontinuity of the Sobolev norm that, in the sense of measures,

$$\delta_0(\xi - u(x,t))\frac{4m}{(m+1)^2}\left|\nabla u^{\left[\frac{m+1}{2}\right]}(x,t)\right|^2 \leqslant q'(x,\xi,t) \quad \text{for } (x,\xi,t)\in Q\times\mathbb{R}\times[0,\infty). \tag{2.4.5}$$

To see this, let $\psi \in C_c(Q\times\mathbb{R}\times[0,T])$ be an arbitrary nonnegative function. The strong convergence (2.4.2) implies that as $k \to \infty$

$$\sqrt{\psi(x,u^{\eta_k,\epsilon_k}(x,t),t)} \to \sqrt{\psi(x,u(x,t),t)} \quad \text{strongly in } L^2(Q\times[0,T]).$$

In turn, this and the weak convergence (2.4.3) yield

$$\sqrt{\psi(x,u^{\eta_k,\epsilon_k},t)}\,\nabla\,(u^{\eta_k,\epsilon_k})^{\left[\frac{m+1}{2}\right]} \rightharpoonup \sqrt{\psi(x,u,t)}\,\nabla u^{\left[\frac{m+1}{2}\right]} \quad \text{weakly in } L^2(Q\times[0,T]).$$

Therefore, the weak convergence (2.4.4), the definition of the measures q^{η_k,ϵ_k} and the weak lower semicontinuity of the L^2-norm prove that

$$\begin{aligned}\frac{4m}{(m+1)^2}\int_0^T\int_Q \psi(x,u,t)\left|\nabla u^{\left[\frac{m+1}{2}\right]}\right|^2 &\leqslant \liminf_{k\to\infty}\frac{4m}{(m+1)^2}\int_0^T\int_Q \psi(x,u^{\eta_k,\epsilon_k},t)\left|\nabla\left(u^{\eta_k,\epsilon_k}\right)^{\left[\frac{m+1}{2}\right]}\right|^2\\ &= \liminf_{k\to\infty}\int_0^T\int_Q\int_{\mathbb{R}}\psi(x,\xi,t)\,q^{\eta_k,\epsilon_k}\\ &= \int_0^T\int_Q\int_{\mathbb{R}}\psi(x,\xi,t)\,q'.\end{aligned}$$

Since ψ was arbitrary, this establishes the inequality (2.4.5).

Now we define the parabolic defect measure by the usual formula (2.2.25), and, since (2.4.5) implies that $q'-q$ is nonnegative, we define a positive entropy defect measure

$$p := p' + q' - q \geqslant 0 \quad \text{on } Q\times\mathbb{R}\times[0,\infty).$$

Finally, the regularity assumptions (2.1.3), the convergence (2.2.1) of the paths z^ϵ and Proposition 2.5.1 implies that, for each $T>0$ and each $k=0,1,2$,

$$\lim_{\epsilon\to 0}\left|D^k_{(x,\xi)}Y^{x,\xi,\epsilon}_{t_0,t} - D^k_{(x,\xi)}Y^{x,\xi}_{t_0,t}\right| + \left|D^k_{(x,\xi)}\Pi^{x,\xi,\epsilon}_{t_0,t} - D^k_{(x,\xi)}\Pi^{x,\xi}_{t_0,t}\right| = 0, \tag{2.4.6}$$

uniformly for $(x,\xi)\in\mathbb{R}^d\times\mathbb{R}$ and $t_0\leqslant t\in[0,T]$.

For the kinetic function χ of u, the convergence (2.4.2) proves that, for a subset $\mathcal{N}\subset(0,\infty)$ of measure zero, for each $t\in[0,\infty)\backslash\mathcal{N}$,

$$\lim_{k\to\infty}\left\|u^{\eta_k,\epsilon_k}(\cdot,t)-u(\cdot,t)\right\|_{L^1(Q)} = 0.$$

This and the additional convergences (2.4.3), (2.4.4) and (2.4.6) imply that, for every $t_0\leqslant t_1\in[0,\infty)\backslash\mathcal{N}$ and every $\rho_0\in C^\infty_c(Q\times\mathbb{R})$, we can pass to the limit in equation (2.2.20) and get

$$\begin{aligned}\int_{Q\times\mathbb{R}}\chi(x,\xi,r)\rho_{t_0,r}(x,\xi)\,dx\,d\xi\Big|_{r=t_0}^{r=t_1} &= \int_{t_0}^{t_1}\int_{Q\times\mathbb{R}} m|\xi|^{m-1}\,\chi^{\eta,\epsilon}(x,\xi,r)\,\Delta_x\rho_{t_0,r}(x,\xi)\,dx\,d\xi\,dr\\ &\quad-\int_{t_0}^{t_1}\int_{Q\times\mathbb{R}}\left(p(x,\xi,r)+q(x,\xi,r)\right)\,\partial_\xi\rho_{t_0,r}(x,\xi)\,dx\,d\xi\,dr,\end{aligned}$$

where $\rho_{t_0,t}$, given by (2.2.24), is the solution of (2.2.23) with initial data ρ_0. Moreover, when $t_0=0$,

$$\int_{Q\times\mathbb{R}}\chi(x,\xi,0)\,\rho_{0,0}(x,\xi)\,dx\,d\xi = \lim_{k\to\infty}\int_{Q\times\mathbb{R}}\chi^{\eta_k,\epsilon_k}(x,\xi,0)\,\rho_{0,0}(x,\xi)\,dx\,d\xi = \int_{Q\times\mathbb{R}}\bar{\chi}(u_0(x),\xi)\,\rho_0(x,\xi)\,dx\,d\xi.$$

This completes the proof that u is a pathwise kinetic solution. $\square$

Finally, we show that pathwise kinetic solutions depend continuously on the driving noise. The proof will follow from a compactness argument relying on the estimates used in the proof of Theorem 2.1.3, the rough path estimates of Proposition 2.5.1, and the uniqueness of solutions from Theorem 2.1.2. Unfortunately, we remark that these methods do not yield an explicit estimate quantifying the convergence of the solutions in terms of the convergence of the noise.

Proof of Theorem 2.1.5.

Let $u_0 \in L^2_+(Q)$ and $T > 0$. Let $\{z^n\}_{n\in\mathbb{N}}$ and z be α-Hölder continuous geometric rough paths on $[0, T]$ satisfying

$$\lim_{n\to\infty} d_\alpha(z^n, z) = 0.$$

This ensures that we can find $R_0 \geqslant 0$ such that condition (2.5.2) from Section A holds.

For each $n \in \mathbb{N}$, let u^n denote the solution of (2.1.1) constructed in Theorem 2.1.3 with initial data u_0 and driving signal z^n respectively. It follows from (2.5.2) and Proposition 2.5.1 that the solutions u^n satisfy the estimates of Propositions 2.4.2, 2.4.3, 2.4.4 and 2.4.5 on the interval $[0, T]$ for a constant that is independent of $n \in \mathbb{N}$.

A repetition of the proof of Theorem 2.1.3 proves that, after passing to a subsequence $\{u^{n_k}\}_{k\in\mathbb{N}}$, there exists a pathwise kinetic solution u of (2.1.1) with initial condition u_0 and driving noise z such that

$$\lim_{k\to\infty} \|u^{n_k} - u\|_{L^1([0,T];L^1(Q))} = 0.$$

However, since it follows from Theorem 2.1.2 that u is the unique solution of (2.1.1) with initial condition u_0 and driving noise z, we conclude that $\lim_{n\to\infty} \|u^n - u\|_{L^1([0,T];L^1(Q))} = 0$ along the full sequence. □

2.5 Rough path estimates

In this section we present some stability results for rough differential equations. Then we apply these results to the systems of characteristics (2.2.9) and (2.2.21) to obtain several properties and estimates needed throughout the chapter. We refer the reader to Friz and Hairer [FH14] and Friz and Victoir [FV10] for detailed expositions of the theory of rough paths, originally introduced by Lyons [Lyo98].

For $d \geqslant 1$, $T \geqslant 0$ and $\beta \in (0, 1)$, we denote by $C^{0,\beta}\big([0, T]; G^{\lfloor 1/\beta \rfloor}(\mathbb{R}^n)\big)$ the space of β-Hölder continuous geometric rough paths, and by d_β the associated β-Hölder metric defined on this space (cf. [FV10, Definition 7.41]). For each $x \in \mathbb{R}^d$ and $z \in C^{0,\beta}\big([0, T]; G^{\lfloor 1/\beta \rfloor}(\mathbb{R}^n)\big)$, let $X^{x,z}$ be the solution of the rough differential equation

$$\begin{cases} dX^{x,z}_t = V(X^{x,z}_t) \circ dz_t & \text{on } (0, \infty), \\ X^{x,z}_0 = x. \end{cases} \tag{2.5.1}$$

The collection (2.5.1) defines a flow map $\psi^z : \mathbb{R}^d \times [0, T] \to \mathbb{R}^d$ by the rule

$$\psi^z_t(x) = X^{x,z}_t \quad \text{for } (x, t) \in \mathbb{R}^d \times [0, T].$$

The next proposition encodes the regularity of the flow map with respect to the initial condition and the driving signal. The regularity is inherited from the nonlinearity V, which must be sufficiently regular to overcome the roughness of the noise. A proof of the proposition is given in [CDFO13, Lemma 13]. In the statement below, we shall write $e = 1 \oplus 0 \oplus \cdots \oplus 0$ to denote the signature of the zero path.

Proposition 2.5.1. Fix $T \geqslant 0$, $\beta \in (0,1)$, $\gamma > \frac{1}{\beta} \geqslant 1$, and $k \in \mathbb{N}$. Assume $V \in C_b^{k+\gamma}(\mathbb{R}^d; \mathbb{R}^{d\times n})$. For any $R \geqslant 0$ there exist constants $C = C(R, \|V\|_{C_b^{k+\gamma}}) > 0$ and $M = M(R, \|V\|_{C_b^{k+\gamma}}) > 0$ such that, for any $z^1, z^2 \in C^{0,\beta}\big([0,T]; G^{\lfloor 1/\beta \rfloor}(\mathbb{R}^n)\big)$ with

$$d_\beta(z^i, e) \leqslant R \quad \text{for } i = 1, 2,$$

the following properties hold. Here $\|\cdot\|_\beta$ is the standard Hölder norm in $C^\beta([0,T]; \mathbb{R}^N)$ for some $N \in \mathbb{N}$.

i) For every $0 \leqslant j \leqslant k$,

$$\sup_{x\in\mathbb{R}^d} \left\| D_x^j\Big(\psi_t^{z^1} - \psi_t^{z^2}\Big)(x)\right\|_\beta + \left\| D_x^j\Big(\big(\psi_t^{z^1}\big)^{-1} - \big(\psi_t^{z^2}\big)^{-1}\Big)(x)\right\|_\beta \leqslant C d_\beta(z^1, z^2).$$

ii) For every $0 \leqslant j \leqslant k$,

$$\sup_{x\in\mathbb{R}^d} \left\| D_x^j\Big(\psi_t^{z^1}\Big)(x)\right\|_\beta + \left\| D_x^j\Big(\big(\psi_t^{z^1}\big)^{-1}\Big)(x)\right\|_\beta \leqslant M.$$

Now we consider the setting outlined in Sections 2.1.1 and 2.2, and apply this regularity result to the rough differential equations (2.2.9) and (2.2.21) defining the characteristics.

Remark 2.5.2. In the setting of Section 2.1.1 and 2.2, the assumption that, for each $T > 0$, the smooth paths z^ϵ converge to z in the metric d_α on $C^{0,\alpha}\big([0,T]; G^{\lfloor 1/\alpha \rfloor}(\mathbb{R}^n)\big)$, ensures that we can find $R_0 \geqslant 0$ such that

$$d_\alpha(z, e) + \sup_{\epsilon\in(0,1)} d_\alpha(z^\epsilon, e) \leqslant R_0. \tag{2.5.2}$$

Therefore, all the consequences of Proposition 2.5.1 holds uniformly for z^ϵ, for any $\epsilon \in (0,1)$, and for z, with the same constants. That is, they hold uniformly for the smooth characteristics (2.2.9), for any $\epsilon \in (0,1)$, and for the rough characteristics (2.2.21). In the remainder of the section we shall consider the path z and the system (2.2.21). We remark that the exact same arguments work for z^ϵ and the system (2.2.9), just inserting ϵ where needed.

First we present a lemma which asserts that the velocity characteristics are locally in time comparable to their initial condition. The proof is given in [FG19, Lemma B.2].

Lemma 2.5.3. For each $T > 0$ there exists $C = C(T, A, R_0) \geqslant 1$ such that, for each $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}$ and $t \leqslant t_0 \in [0,T]$,

$$C^{-1}|\xi| \leqslant \left|\Pi_{t_0,t}^{x,\xi}\right| \leqslant C|\xi|.$$

Furthermore, there exists $C = C(T, A, R_0) > 0$ such that, for each $(x,\xi) \in \mathbb{R}^d \times \mathbb{R}$ and $t \leqslant t_0 \in [0,T]$,

$$\left|\nabla_x \Pi^{x,\xi}_{t_0,t}\right| \leqslant C\, t^{\alpha}(|\xi| \wedge 1). \tag{2.5.3}$$

We now exploit Proposition 2.5.1 to study the continuity of the characteristics with respect to the initial data. First, recalling (2.2.11), we observe that for each $(x,\xi), (x',\xi') \in \mathbb{R}^d \times \mathbb{R}$, each $T \geqslant 0$ and each $s \leqslant t_0 \in [0,T]$,

$$\begin{aligned}|x - x'| &= \left| X^{Y^{x,\xi}_{t_0,s}, \Pi^{x,\xi}_{t_0,s}}_{t_0-s,t_0} - X^{Y^{x',\xi'}_{t_0,s}, \Pi^{x',\xi'}_{t_0,s}}_{t_0-s,t_0} \right| \\ &\leqslant \sup_{(y,\eta)\in\mathbb{R}^d\times\mathbb{R}} \left|D_y X^{y,\eta}_{t_0-s,t_0}\right| \left|Y^{x,\xi}_{t_0,s} - Y^{x',\xi'}_{t_0,s}\right| + \sup_{(y,\eta)\in\mathbb{R}^d\times\mathbb{R}} \left|\partial_\eta X^{y,\eta}_{t_0-s,t_0}\right| \left|\Pi^{x,\xi}_{t_0,s} - \Pi^{x',\xi'}_{t_0,s}\right|.\end{aligned}$$

An identical estimate holds for $|\xi - \xi'|$. Therefore, assumption (2.1.3) and Proposition 2.5.1 with $k = 1$ imply that, for a constant $C = C(T, A, R_0)$,

$$|x - x'| + |\xi - \xi'| \leqslant C \left(\left|Y^{x,\xi}_{t_0,s} - Y^{x',\xi'}_{t_0,s}\right| + \left|\Pi^{x,\xi}_{t_0,s} - \Pi^{x',\xi'}_{t_0,s}\right| \right). \tag{2.5.4}$$

In turn, we can use this bound to estimate the difference between derivatives of the characteristics starting from distinct points in terms of the characteristics themselves. First, notice that the equalities $D^2_{(y,\eta)} Y^{y,\eta}_{t_0,0} = 0$ and $D^2_{(y,\eta)} \Pi^{y,\eta}_{t_0,0} = 0$, which follow immediately from the initial conditions, and Proposition 2.5.1 with $k = 2$ imply that, for $C = C(T, A, R_0)$,

$$\sup_{x\in\mathbb{R}^d, \xi\in\mathbb{R}} \left|D^2_{(x,\xi)} Y^{x,\xi}_{t_0,s}\right| + \left|D^2_{(x,\xi)} \Pi^{x,\xi}_{t_0,s}\right| \leqslant C|s|^{\alpha} \quad \forall s \leqslant t_0 \in [0,T]. \tag{2.5.5}$$

Then, for the same constant C, we compute

$$\begin{aligned}|D_x Y^{x,\xi}_{t_0,s} - D_{x'} Y^{x',\xi'}_{t_0,s}| &\leqslant \sup_{y\in\mathbb{R}^d, \eta\in\mathbb{R}} \left(\left|D^2_y Y^{y,\eta}_{t_0,s}\right| + \left|\partial_\eta D_y Y^{y,\eta}_{t_0,s}\right|\right)(|x - x'| + |\xi - \xi'|) \\ &\leqslant C|s|^{\alpha}(|x - x'| + |\xi - \xi'|) \\ &\leqslant C|s|^{\alpha} \left(\left|Y^{x,\xi}_{t_0,s} - Y^{x',\xi'}_{t_0,s}\right| + \left|\Pi^{x,\xi}_{t_0,s} - \Pi^{x',\xi'}_{t_0,s}\right|\right).\end{aligned}$$

Identical computations holds for the other derivatives of $Y^{x,\xi}_{t_0,s}$ and $\Pi^{x,\xi}_{t_0,s}$ and we conclude that, for $C = C(T, A, R_0)$,

$$\begin{aligned}&|D_x Y^{x,\xi}_{t_0,s} - D_{x'} Y^{x',\xi'}_{t_0,s}| + |\partial_\xi Y^{x,\xi}_{t_0,s} - \partial_{\xi'} Y^{x',\xi'}_{t_0,s}| + |\nabla_x \Pi^{x,\xi}_{t_0,s} - \nabla_{x'} \Pi^{x',\xi'}_{t_0,s}| + |\partial_\xi \Pi^{x,\xi}_{t_0,s} - \partial_{\xi'} \Pi^{x',\xi'}_{t_0,s}| \\ &\qquad \leqslant C|s|^{\alpha} \left(\left|Y^{x,\xi}_{t_0,s} - Y^{x',\xi'}_{t_0,s}\right| + \left|\Pi^{x,\xi}_{t_0,s} - \Pi^{x',\xi'}_{t_0,s}\right|\right).\end{aligned} \tag{2.5.6}$$

We conclude this section by examining the consequences of the conditions $\partial_\xi A_{|\partial Q\times\mathbb{R}} \equiv 0$ and $D_x \partial_\xi A_{|\partial Q\times\mathbb{R}} \equiv 0$. Then we exploit the continuity properties from Proposition 2.5.1 to obtain information on the behaviour of the space characteristics near the boundary. These estimates are crucial in the proof of Theorem 2.1.2 to tackle the boundary error terms coming from the introduction of the cutoff and the transport along characteristics.

The assumption $\partial_\xi A(x,\xi)_{|\partial Q\times\mathbb{R}} \equiv 0$, combined with the first line of (2.2.21) and (2.2.22), ensures that space characteristics starting from the boundary do not move. That is, for each $t_0 \geqslant 0$,

$$\text{if} \quad (x,\xi) \in \partial Q \times \mathbb{R}, \text{ then } X^{x,\xi}_{t_0,t} = Y^{x,\xi}_{t_0,s} = x \text{ for all } t \geqslant 0 \text{ and all } s \in [0,t_0]. \tag{2.5.7}$$

The uniqueness of solutions then guarantees that space characteristics cannot cross the boundary. Thus, when starting from $x \in Q$ they never leave the domain. That is, for every $t_0 \geqslant 0$,

$$\text{if} \quad (x,\xi) \in Q \times \mathbb{R}, \quad \text{then } X^{x,\xi}_{t_0,t} \in Q \text{ for all } t \geqslant 0, \quad \text{and } Y^{x,\xi}_{t_0,s} \in Q \text{ for all } s \in [0,t_0].$$

In fact, owing to the smoothness hypothesis (2.1.3) and Proposition 2.5.1 with $k = 1$, more is true: the closer to ∂Q is the space initial data $x \in \mathbb{R}^d$, the slower the associated space characteristic moves. Rigorously, given any $x \in \mathbb{R}^d$, let $x^* \in \partial Q$ be such that $|x - x^*| = \operatorname{dist}(x, \partial Q)$. Then, for any $\xi \in \mathbb{R}$, any $T \geqslant 0$ and any $s \leqslant t_0 \in [0,T]$, for $C = C(T, A, R_0)$, we compute

$$\begin{aligned}
\left|Y^{x,\xi}_{t_0,s} - x\right| \leqslant \left|Y^{x,\xi}_{t_0,s} - x^*\right| + |x^* - x| = \left|Y^{x,\xi}_{t_0,s} - Y^{x^*,\xi}_{t_0,s}\right| + |x^* - x| \\
\leqslant \sup_{(y,\eta)\in\mathbb{R}^d\times\mathbb{R}} \left|D_x Y^{y,\eta}_{t_0,s}\right| |x - x^*| + |x - x^*| \\
\leqslant C\,|x - x^*| = C \operatorname{dist}(x, \partial Q).
\end{aligned} \tag{2.5.8}$$

Here we used $Y^{x^*\xi}_{t_0,s} \equiv x^*$ from (2.5.7). An identical computation holds for $X^{x,\xi}_{t_0,t}$.

The constancy of the space characteristics along the boundary has consequences on their velocity and space derivatives. Indeed, the smoothness of characteristics combined with (2.5.7) implies that the ξ-derivative of space characteristics vanishes on the boundary. Namely, for each $t_0 \in [0, \infty)$,

$$\text{if } (x,\xi) \in \partial Q \times \mathbb{R}, \text{ then } \partial_\xi X^{x,\xi}_{t_0,t} = \partial_\xi Y^{x,\xi}_{t_0,s} = 0 \quad \text{for all } t \geqslant 0 \text{ and all } s \in [0,t_0]. \tag{2.5.9}$$

Moreover, Proposition 2.5.1 allows us to derive (2.2.22) with respect to x, and to show that, for any $(x_0,\xi_0) \in \mathbb{R}^d \times \mathbb{R}$, the derivative $D_x Y^{x_0,\xi_0}_{t_0,s}$ solves the rough differential equation

$$\begin{cases}
dD_x Y^{x_0,\xi_0}_{t_0,s} = -\Big(D_x\partial_\xi A(Y^{x_0,\xi_0}_{t_0,s}, \Pi^{x_0,\xi_0}_{t_0,s}) D_x Y^{x_0,\xi_0}_{t_0,s} + \partial^2_\xi A(Y^{x_0,\xi_0}_{t_0,s}, \Pi^{x_0,\xi_0}_{t_0,s}) \nabla_x \Pi^{x_0,\xi_0}_{t_0,s}\Big) \circ dz_{t_0,s} \text{ in } (0,t_0), \\
D_x Y^{x_0,\xi_0}_{t_0,0} = \mathrm{I}_d.
\end{cases}$$

When $x_0 \in \partial Q$, we have $Y^{x_0,\xi_0}_{t_0,s} = x_0$ for all $s \in [0,t_0]$. Then, using $\partial_\xi A_{|\partial Q\times\mathbb{R}} \equiv 0$, which trivially implies also $\partial^2_\xi A_{|\partial Q\times\mathbb{R}} \equiv 0$, and the further assumption $D_x\partial_\xi A_{|\partial Q\times\mathbb{R}} \equiv 0$ from (2.1.5), the previous equation simply reduces to

$$dD_x Y^{x_0,\xi_0}_{t_0,s} = 0 \text{ in } (0,t_0), \quad D_x Y^{x_0,\xi_0}_{t_0,0} = \mathrm{I}_d.$$

We conclude that, for each $t_0 \geqslant 0$,

$$\text{if } (x_0,\xi_0) \in \partial Q \times \mathbb{R}, \text{ then } D_x Y^{x_0,\xi_0}_{t_0,s} \equiv \mathrm{I}_d \quad \text{for every } s \in [0,t_0]. \tag{2.5.10}$$

Finally, we use these results and the stability properties from Proposition 2.5.1 with $k = 2$ to obtain estimates on the derivatives of the space characteristics near the boundary. Given any $x \in \mathbb{R}^d$, let $x^* \in \partial Q$ be such that $|x - x^*| = \operatorname{dist}(x, \partial Q)$. Then, using (2.5.5) and (2.5.9), for any $\xi \in \mathbb{R}$, any $T \geqslant 0$ and any $s \leqslant t_0 \in [0, T]$, we compute for the the ξ-derivative

$$\begin{aligned} \left|\partial_\xi Y_{t_0,s}^{x,\xi}\right| = \left|\partial_\xi Y_{t_0,s}^{x,\xi} - \partial_\xi Y_{t_0,s}^{x^*,\xi}\right| &\leqslant \sup_{y \in \mathbb{R}^d} \left|D_x \partial_\xi Y_{t_0,s}^{y,\xi}\right| |x - x^*| \\ &\leqslant C\, |s|^\alpha \operatorname{dist}(x, \partial Q), \end{aligned} \tag{2.5.11}$$

for a constant $C = C(T, A, R_0)$. As regards the x-derivative, the same argument using (2.5.5) and (2.5.10) implies that, for $C = C(T, A, R_0)$,

$$\left|D_x Y_{t_0,s}^{x,\xi} - \mathrm{I}_d\right| = \left|D_x Y_{t_0,s}^{x,\xi} - D_x Y_{t_0,s}^{x^*,\xi}\right| \leqslant C\, |s|^\alpha \operatorname{dist}(x, \partial Q). \tag{2.5.12}$$

Chapter 3

Fluctuations in the small noise limit

3.1 Introduction and main results

In this chapter we study the small noise fluctuations of nonnegative solutions ρ^ϵ to conservative, stochastic PDEs of the type

$$\begin{cases} \partial_t\rho^\epsilon = \Delta\phi(\rho^\epsilon) - \nabla\cdot\nu(\rho^\epsilon) - \sqrt{\epsilon}\nabla\cdot\Big(\sigma(\rho^\epsilon)\circ\dot{\xi}^\epsilon\Big) \;\text{ in } \mathbb{T}^d\times(0,T), \\ \rho^\epsilon(\cdot,0) = \rho_0 \;\text{ in } \mathbb{T}^d\times\{0\}, \end{cases} \tag{3.1.1}$$

about their zero noise limit $\bar\rho$ solving the equation

$$\partial_t\bar\rho = \Delta\phi(\bar\rho) - \nabla\cdot\nu(\bar\rho) \;\text{ in } \mathbb{T}^d\times(0,T), \quad \bar\rho(\cdot,0) = \rho_0 \;\text{ in } \mathbb{T}^d\times\{0\}. \tag{3.1.2}$$

We prove that the fluctuations $\epsilon^{-\frac{1}{2}}(\rho^\epsilon - \bar\rho)$ converge in probability, in a space of distributions, to the solution of the linearized Langevin equation

$$\partial_t v = \Delta\big(\dot\phi(\bar\rho)\, v\big) - \nabla\cdot\Big(\dot\nu(\bar\rho)\, v + \sigma(\bar\rho)\circ\dot\xi\Big) \;\text{ in } \mathbb{T}^d\times(0,T), \quad v(\cdot,0) = 0 \;\text{ in } \mathbb{T}^d\times\{0\}, \tag{3.1.3}$$

which is the linearization of (3.1.1) around $\bar\rho$.

The assumptions on the noise sequence ξ^ϵ, which can converge to a space-time white noise or some other cylindrical noise ξ as $\epsilon\to 0$, on the initial data ρ_0 and on the nonlinearities ϕ, ν and σ are given in Section 3.2.1, and include a range of relevant stochastic PDEs (see Example 3.2.11). In particular, they apply to the full range of porous media diffusions, that is $\phi(z) = z^m$ for every $m\in(0,\infty)$, and to degenerate convective terms σ and ν, including the square root $\sigma(z) = \sqrt{z}$, and hence to the nonlinear version of the Dean–Kawasaki equation with correlated noise

$$\partial_t\rho^\epsilon = \Delta(\rho^\epsilon)^m - \sqrt{\epsilon}\,\nabla\cdot\Big(\sqrt{\rho^\epsilon}\circ\dot\xi^\epsilon\Big).$$

Equations of type (3.1.2) and (3.1.1) arise in the fluctuating hydrodynamics of particle systems, to describe respectively the limiting behavior and the nonequilibrium fluctuations of particle systems, in the context of mean field theory with common noise and in the area of stochastic geometric PDEs. Applications to particle systems are discussed in Section 3.1.1.

The wellposedness of (3.1.1) has been a long-standing open problem. In the case the noise ξ^ϵ has sufficient space correlation, existence and uniqueness for (3.1.1) have been established only recently in [FG21b], through the concept of stochastic kinetic solution (cf. Definition 3.2.19 and Theorem 3.2.21 below).

The study of the small noise behaviour of (3.1.1) has been initiated in [FG23]. The authors proved that, in the $\epsilon \to 0$ limit, solutions of (3.1.1) converge to solutions of the zero-noise limit (3.1.2):

$$\lim_{\epsilon \to 0} \|\rho^\epsilon - \bar{\rho}\|_{L^1([0,T];L^1(\mathbb{T}^d))} = 0 \text{ in probability.} \tag{3.1.4}$$

Furthermore they satisfy a large deviation principle in $L^1([0,T];L^1(\mathbb{T}^d))$ with rate function

$$I_{\rho_0}(\rho) = \inf\left\{ \|g\|^2_{L^2(\mathbb{T}^d\times[0,T])^d} \,\middle|\, \partial_t \rho = \Delta\phi(\rho) - \nabla\cdot\nu(\rho) - \nabla\cdot(\sigma(\rho)g)\,,\ \rho(\cdot,0) = \rho_0 \right\}. \tag{3.1.5}$$

Finally the authors rigorously identified the rate function (3.1.5) with that governing the large deviations of the zero range particle process (cf. [FG23, Theorem 6.8 and 8.6]), first studied in [BKL95], formalizing the connection between the particle system and the above SPDEs.

The contribution of this work is to complete the picture above by proving a central limit theorem characterizing the fluctuations of (3.1.1) around its deterministic limit (3.1.2). See Section 3.1.1 for an overview of the relevant literature and related results. Precisely, consider the normalized fluctuations $v^\epsilon := \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho})$, which are readily seen to solve

$$\begin{cases} \partial_t v^\epsilon = \Delta\left(\epsilon^{-1/2}\left(\phi(\rho^\epsilon) - \phi(\bar{\rho})\right)\right) - \nabla\cdot\left(\epsilon^{-1/2}\left(\nu(\rho^\epsilon) - \nu(\bar{\rho})\right) + \sigma(\rho^\epsilon)\circ\dot{\xi}^\epsilon\right) \text{ in } \mathbb{T}^d\times(0,T), \\ v^\epsilon(\cdot,0) = 0 \text{ on } \mathbb{T}^d\times\{0\}. \end{cases} \tag{3.1.6}$$

Since $\rho^\epsilon \to \bar{\rho}$ and $\xi^\epsilon \to \xi$ (cf. Assumption 3.2.4) as $\epsilon \to 0$, one formally deduces that v^ϵ should converge to the solution of the Langevin equation (3.1.3). The main result of the present work is to make this ansatz rigorous.

Theorem (Theorem 3.3.3 and 3.3.10 below)**.** Let $(\xi^\epsilon)_{\epsilon>0}$ satisfy Assumption 3.2.4. Let ϕ, ν, σ satisfy Assumption 3.2.6 and 3.2.8. Let $\rho_0 \in L^\infty(\Omega;(0,\infty))$ be an $\mathcal{F}_0$-measurable random constant bounded away from zero, so that $\bar{\rho}(x,t) \equiv \rho_0$ solves (3.1.2). For each $\epsilon > 0$, let ρ^ϵ be the stochastic kinetic solution to (3.1.1) with initial data ρ_0, in the sense of Definition 3.2.19. Let $v \in L^2(\Omega\times[0,T];H^{-\alpha}(\mathbb{T}^d))$ be the solution of (3.1.3) with noise $\xi = \lim_{\epsilon\to 0}\xi^\epsilon$, for any $\alpha > \frac{d}{2}$. For any $T > 0$, for $\tau = 2$ or $\tau = \infty$, for any $\beta > \frac{d}{2}$ or $\beta > 1 + \frac{d}{2}$ respectively, we have the following results.

i) Along a suitable scaling regime where $\epsilon \to 0$ and $\xi^\epsilon \to \xi$, explicitly given in (3.3.11), the nonequilibrium fluctuations satisfy

$$v^\epsilon := \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho}) \to v \quad \text{in } L^\tau([0,T];H^{-\beta}(\mathbb{T}^d)) \text{ in probability,} \tag{3.1.7}$$

with an explicit rate of convergence, given in (3.3.47), which depends on β, T, the coefficients ϕ, ν, σ, the initial data ρ_0, and the space regularity of the noise sequence ξ^ϵ.

ii) In addition, if the coefficients ϕ, ν, σ also satisfy Assumption 3.2.9, along the same scaling regime (3.3.11), we have

$$v^\epsilon \to v \quad \text{in } L^2(\Omega; L^\tau([0,T]; H^{-\beta}(\mathbb{T}^d))), \tag{3.1.8}$$

with an explicit rate of convergence, given in (3.3.10), which depends on β, T, the coefficients ϕ, ν, σ, the initial data ρ_0, and the space regularity of the noise sequence ξ^ϵ.

The results of this work and [FG23] have immediate consequences on the simulation of several particle systems. Indeed, upon choosing the right coefficients ϕ, ν and σ according to the hydrodynamics of the particle process considered, equation (3.1.1) provides a continuum model which correctly captures the fluctuations of the particle system up to order one and exhibits the same nonequilibrium large deviations. This is discussed in detail in Section 3.1.1.

Structure of the chapter

The work is organized as follows. We end this section with a discussion of the applications to particle systems, followed by an overview of our methods and of the relevant literature. In Section 3.2 we first lay out our notations and assumptions; then we present the solution theory for the equations involved: the zero noise limit (3.1.2), the stochastic PDE (3.1.3) for the asymptotic fluctuations and the original SPDE (3.1.1). In Section 3.3 we finally prove our central limit theorem. First we prove the stronger version (3.1.8) for coefficients satisfying the extra Assumption 3.2.9; then we extend this to rougher coefficients only satisfying Assumption 3.2.8 and prove the central limit theorem in probability (3.1.7).

3.1.1 Applications to particle systems, methods and relevant literature

Applications to particle systems.

A byproduct of the above results is that the SPDE (3.1.1) provides a continuum model correctly simulating the fluctuations of several conservative particle systems up to order one and exhibiting the same nonequilibrium large deviations. Consider for example a symmetric zero range process on the torus with mean local jump rate ϕ and slowly varying initial state ρ_0. As the number n of particles increases, the parabolically rescaled empirical measure μ^n converges in probability, weakly in the sense of measure, to its hydrodinamic limit: the deterministic measure $\bar{\rho}\, \mathrm{d}x$ solving

$$\partial_t \bar{\rho} = \Delta\phi(\bar{\rho}) \;\text{ in } \mathbb{T}^d \times (0,T), \quad \bar{\rho}(\cdot, 0) = \rho_0 \;\text{ in } \mathbb{T}^d \times \{0\}. \tag{3.1.9}$$

In this way, solutions to the porous media equation (3.1.9) describe the dynamics of the particle process up to order zero (see e.g. [KL99, Chapter 5]).

In order to reach higher order continuum approximations, it is necessary to incorporate the fluctuations present in μ^n. A detailed study of the nonequilibrium central limit fluctuations is presented in [KL99, Chapter 11], where it is shown that the measures

$$\mathrm{m}^n := \sqrt{n}\,(\mu^n - \bar{\rho}\,dx) \tag{3.1.10}$$

converge as $n \to \infty$ to the solution of the linear stochastic PDE

$$\partial_t \mathrm{m} = \Delta\big(\dot{\phi}(\bar{\rho})\,\mathrm{m}\big) - \nabla \cdot \Big(\phi^{1/2}(\bar{\rho}) \circ \dot{\xi}\Big) \text{ in } \mathbb{T}^d \times (0,T), \quad \mathrm{m}(\cdot,0) = 0 \text{ in } \mathbb{T}^d \times \{0\}, \tag{3.1.11}$$

for ξ a space-time white noise and $\bar{\rho}$ solving (3.1.9). Since solutions to (3.1.11) are not function-valued the convergence is in the sense of distributions. Furthermore, the results of [BKL95] and [FG23] prove that the non-equilibrium large deviations of μ^n are described in terms of the rate function

$$I_{\rho_0}(\rho) = \inf\Big\{\|g\|^2_{L^2(\mathbb{T}^d\times[0,T])^d} : \partial_t\rho = \Delta\phi(\rho) - \nabla\cdot\Big(\phi^{1/2}(\rho)\,g\Big) \text{ and } \rho(\cdot,0) = \rho_0\Big\}. \tag{3.1.12}$$

Consider now the the nonlinear stochastic PDE

$$\partial_t \rho^n = \Delta\phi(\rho^n) - \frac{1}{\sqrt{n}}\nabla\cdot\Big(\phi^{1/2}(\rho^n)\circ\dot{\xi}^n\Big), \quad \rho^n(\cdot,0) = \rho_0, \tag{3.1.13}$$

for suitable approximations $\{\xi^n\}_{n\in\mathbb{N}}$ of an $\mathbb{R}^d$-valued space-time white noise (cf. Assumption 3.2.4). Our main result Theorem 3.3.10 ensures that, as $n \to \infty$, the small noise fluctuations $\sqrt{n}(\rho^n - \bar{\rho})$ converge to the solution m of (3.1.11). Therefore, this and (3.1.10)-(3.1.11) give us the expansion

$$\mu^n = \rho^n\,dx + \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

which now correctly captures the fluctuations of μ^n up to order one. Furthermore ρ^n also features the same large deviation rate function (3.1.12) of μ^n, as demonstrated by (3.1.5) from [FG23, Theorem 6.8].

Equation (3.1.13) is an approximation for the formal nonlinear SPDE

$$\partial_t \rho^n = \Delta\phi(\rho^n) - \frac{1}{\sqrt{n}}\nabla\cdot\Big(\phi^{1/2}(\rho^n)\circ\dot{\xi}\Big). \tag{3.1.14}$$

The approximation is needed because of the irregularity of the space-time white noise ξ. Indeed, equation (3.1.14) is supercritical in the language of regularity structures [Hai14]; its intrinsic ill-posedness and negative results have been discussed in the seminal work [KLvR19a]. In fact, it can also be argued (see e.g. [GLP99, Section 3]) that the microscopic particle system comes with a typical correlation length for the noise, like the grid size, which leads to consider equation (3.1.13) for some space correlated noise ξ^n instead of equation (3.1.14).

Finally, it is worth pointing out that one could also consider the simpler first order expansion

$$\mu^n = \bar{\rho}\,\mathrm{d}x + \frac{1}{\sqrt{n}}\mathrm{m} + \mathrm{o}\left(\frac{1}{\sqrt{n}}\right),$$

which follows immediately from (3.1.10)-(3.1.11). However $\bar{\rho}^n := \bar{\rho}\,\mathrm{d}x + \frac{1}{\sqrt{n}}\mathrm{m}$ exhibits a rate function different from (3.1.12), given by

$$\bar{I}_{\rho_0}(\rho) = \inf\left\{ \|g\|^2_{L^2(\mathbb{T}^d\times[0,T])^d} \,\Big|\, \partial_t\,(\rho-\bar{\rho}) = \Delta\left(\dot{\phi}(\bar{\rho})(\rho-\bar{\rho})\right) - \nabla\cdot\left(\phi^{1/2}(\bar{\rho})\,g\right),\ \rho(\cdot,0)=\rho_0 \right\},$$

as follows from Schilder's theorem and a formal application of the contraction principle. Therefore this expansion does not capture the large deviations of μ^n and yields an imprecise approximation.

In conclusion we remark that completely analogous results hold for more general zero range processes and exclusion processes. Essentially the results hold for all the particle processes whose fluctuations can be described through SPDEs falling in the framework above. We refer the reader to the monographs [KL99, Spo12] and the surveys [GLP99] for results in this sense.

Overview of the methods.
As mentioned above, a concept of solution for (3.1.1) with an actual white noise seems to be out of hand. If the noise ξ^ϵ has some space correlation (cf. Assumption 3.2.2), the well-posedness of equation (3.1.1) is proved in [FG21b]. The main difficulties in applying a classical concept of weak solutions are due to nonlinearities that are possibly only $1/2$-Hölder continuous and singular terms which are not even known to be locally integrable. These issues are addressed by the the notion of *stochastic kinetic solution* (cf. Definition 3.2.19). After rewriting the Stratonovich equation (3.1.1) in the equivalent Itô form, the equation is recast in its *kinetic formulation*: an equation in the original space and time variables and in a new additional *velocity* variable, corresponding to the magnitude of the solution. Then a renormalization away from zero and infinity is introduced: solutions are required to satisfy the PDE only after cutting out small and large values in order to enforce the local integrability and further regularity of the nonlinear terms. In fact, when the equation coefficients are nice enough, this renormalization is not even needed and stochastic kinetic solutions satisfy the equation in a classical weak sense (cf. Definition 3.2.17 and Proposition 3.2.25). Furthermore, stochastic kinetic solutions depend continuously on the equation coefficients in a suitable sense (cf. Proposition 3.2.23). These two properties will be crucial in our arguments. We refer the reader to [FG21b, Section 1.1] for more details on their methods.

The small noise large deviations of (3.1.1) are analyzed in [FG23]. Schilder's Theorem and a formal application of the contraction principle lead to guess the rate function (3.1.5). After a careful analysis of the *energy critical* PDE featuring in the definition of I_{ρ_0}, this ansatz is made rigourous via the so-called *weak approach to large deviations* [DE97, BDM08, BD19, BDS19].

We finally come to the criticalities and the methods of this work. A first difficulty is that the compactness of the noise-to-solution map of (3.1.1) is not enough to show convergence of the fluctuations $v^\epsilon = \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho})$. Indeed, although it is true that $\rho^\epsilon \to \bar{\rho}$ in a suitable sense, the previous arguments yield no information about the convergence rate; whereas we need to quantify this to compensate for the blowing up factor $\epsilon^{-1/2}$. Moreover, the high nonlinearity and degeneracy of the equation hinder the application of Fourier analysis. Finally, the renormalization away from small and large values in the kinetic formulation of the equation and the possible lack of regularity coming from the degeneracy of the diffusion make it difficult to exploit the equation for ρ^ϵ in our computations. The picture is further complicated by the irregularity and unboundedness of the noise coefficient σ.

To address these issues, we consider approximating versions of equation (3.1.1) with smoothed coefficients $\phi^\eta, \nu^\eta, \sigma^\eta$ indexed by a parameter $\eta \in (0, 1)$, for which a stronger notion of solution is available (cf. Definition 3.2.17). We always work with the corresponding solutions $\rho^{\epsilon,\eta}$, which have enough regularity to justify our computations, and then pass the results to the true solutions ρ^ϵ either with Fatou's Lemma or with probabilistic arguments based on analytical estimates.

We begin with the stronger version (3.1.8) of the CLT for coefficients satisfying the additional Assumption 3.2.9. This already applies to the model case $\phi(z) = z^m$, $\sigma(z) = z^{\frac{m}{2}}$ in the regime $m \in [2, \infty)$, or to the linear case $\phi(z) = z$ for suitable noise coefficients σ.

First we obtain estimates on the L^h-norm of the fluctuations v^ϵ, for some h big enough depending on the nonlinearities involved. Since v^ϵ is converging to v, which is only distribution valued, we expect these estimates to blow up as $\epsilon \to 0$ and we want to quantify the explosion rate in terms of the small parameter ϵ. Formally, this is achieved by applying Itô formula to the power $|v^\epsilon|^h$ and exploiting the equation satisfied by v^ϵ. Equipped with these moment estimates, we estimate the Fourier coefficients of $v^\epsilon - v$ and show that $\|v^\epsilon - v\|_{L^\infty([0,T];H^{-\beta}(\mathbb{T}^d))}$ vanishes in $L^2(\Omega)$ as $\epsilon \to 0$.

To consider even rougher coefficients, including $\sigma(z) = \sqrt{z}$ and $\phi(z) = z^m$ for every $m \in (0, \infty)$, some extra work is needed. With a Moser iteration argument, adapted from [DFG20], we show that the solutions ρ^ϵ of (3.1.1) are bounded from below by the minimum of the *positive* initial data ρ_0 with increasing probability as $\epsilon \to 0$ (cf. Proposition 3.3.8 and Corollary 3.3.9). In particular, solutions ρ^ϵ stay away from the irregularities at zero of ϕ, ν, σ on the events ω of a certain subset $\Omega^\epsilon \subseteq \Omega$ where they satisfy $\rho^\epsilon(x, t, \omega) \in [\inf \rho_0 - \delta, \infty)$ for a.e. x and t. A refinement of the pathwise uniqueness result for (3.1.1) shows that on Ω^ϵ the solution ρ^ϵ must coincide with the solution $\rho^{\epsilon,\eta}$ of the smoothed version of equation (3.1.1) with coefficients $\phi^\eta, \nu^\eta, \sigma^\eta$ that match the true coefficients ϕ, σ, η sufficiently away from the irregularity points (cf. Propositition 3.3.6). For such solutions $\rho^{\epsilon,\eta}$ we have the above CLT in $L^2(\Omega)$ at our disposal and with standard probabilistic arguments we pass this to a CLT in probability for the true solutions ρ^ϵ.

Overview of the literature.

Linear and nonlinear diffusion equation with different kinds of noise terms have received a lot of attention, from several point of view. We refer to [FG21b] for a detailed overview of the literature and we mention only the most recent results on the equation considered here [CFIR23, FGG22, DKP22, Cli23b, WWZ24]. Similarly, we refer to the introduction of [FG23] for the literature on large deviations principles on related SPDEs.

We collect here instead some results concerning central limit theorems for SPDEs. Namely, central limit theorems for stochastic heat equations or stochastic wave equations with Lipschitz continuous noise coefficients have been obtained in [CKNP22, HNV20, HNVZ19] and in [DVNZ20] respectively. A central limit theorem for the heat equation driven by nonlinear gradient noise with Hölder continuous coefficient has been proved by Dirr, Gess and the second author in [DFG20]. Finally, central limit theorems for primitive equations, a specific type of semilinear evolution equation, in low dimension have been established in [Sla21, ZZG19, HLW18].

3.2 Preliminaries and solution theory

3.2.1 Hypotheses and notations

In this section we present our assumptions and notations, and we remark that they apply to a range of relevant cases (see Example 3.2.5 and 3.2.11). We first take care of the randomness in the equation. We fix a probability space $\left(\Omega, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P}\right)$ with a complete right-continuous filtration $\mathcal{F}_t$ and supporting a countable sequence of $\mathbb{R}^d$-valued independent Brownian motions $\left(B_t^k\right)_{k \in \mathbb{N}}$ and the random initial condition ρ_0.

As regards the initial condition we require the following.

Assumption 3.2.1 (Assumptions on the initial data)**.** The initial data $\rho_0 \in L^1(\Omega; L^1(\mathbb{T}^d))$ is nonnegative, $\mathcal{F}_0$-measurable and satisfies one of the following hypotheses, for $p \geqslant 2$ and $m \geqslant 1$ given in Assumption 3.2.6 below:

(i) $\rho_0 \in L^p(\Omega; L^p(\mathbb{T}^d)) \cap L^{p+m-1}(\Omega; L^1(\mathbb{T}^d))$;

(ii) $\rho_0 \in L^{p+m-1}(\Omega; (0, \infty))$ with $\rho \geqslant r > 0$ a.e. for some $r \in (0, \infty)$, i.e. the initial data is a random positive constant.

We now define the noise sequence ξ^ϵ. For each $\epsilon > 0$, let $F^\epsilon = (f_k^\epsilon)_{k \in \mathbb{N}}$ be a family of continuously differentiable functions on $\mathbb{T}^d$ and define the noise

$$\xi^\epsilon = \sum_{k \in \mathbb{N}} f_k^\epsilon B_t^k.$$

It then follows that the Stratonovich equation (3.1.1) is formally equivalent to the Itô equation

$$\mathrm{d}\rho^\epsilon = \Delta\phi(\rho^\epsilon)\mathrm{d}t - \nabla\cdot\nu(\rho^\epsilon) - \sqrt{\epsilon}\,\nabla\cdot(\sigma(\rho^\epsilon)\,\mathrm{d}\xi^\epsilon) + \frac{\epsilon}{2}\sum_{k=1}^{\infty}\nabla\cdot\Big(f_k^\epsilon\dot{\sigma}(\rho^\epsilon)\nabla\big(f_k^\epsilon\sigma(\rho^\epsilon)\big)\Big)\mathrm{d}t, \quad (3.2.1)$$

which can be written in the form

$$\mathrm{d}\rho^\epsilon = \Delta\phi(\rho^\epsilon)\mathrm{d}t - \nabla\cdot\nu(\rho^\epsilon) - \sqrt{\epsilon}\,\nabla\cdot(\sigma(\rho^\epsilon)\,\mathrm{d}\xi^\epsilon) + \frac{\epsilon}{2}\nabla\cdot\Big(F_1^\epsilon\,(\dot{\sigma}(\rho^\epsilon))^2\,\nabla\rho^\epsilon + \dot{\sigma}(\rho^\epsilon)\sigma(\rho^\epsilon)F_2^\epsilon\Big)\,\mathrm{d}t, \quad (3.2.2)$$

for $F_1^\epsilon:\mathbb{T}^d\to\mathbb{R}$ and $F_2^\epsilon:\mathbb{T}^d\to\mathbb{R}^d$ defined by

$$F_1^\epsilon(x) = \sum_{k=1}^{\infty}(f_k^\epsilon)^2(x) \;\text{ and }\; F_2^\epsilon(x) = \sum_{k=1}^{\infty}f_k^\epsilon(x)\nabla f_k^\epsilon(x).$$

We make the following assumptions on the noise.

Assumption 3.2.2 (Assumptions on the noise)**.** For each $\epsilon > 0$, we assume that the sums $\{F_i^\epsilon\}_{i=1,2,3}$ defined by

$$F_1^\epsilon = \sum_{k=1}^{\infty}(f_k^\epsilon)^2 \;\text{ and }\; F_2^\epsilon = \frac{1}{2}\sum_{k=1}^{\infty}\nabla(f_k^\epsilon)^2 \;\text{ and }\; F_3^\epsilon = \sum_{k=1}^{\infty}|\nabla f_k^\epsilon|^2$$

are continuous on $\mathbb{T}^d$ – where the finiteness of F_1^ϵ and F_3^ϵ implies the absolute convergence of F_2^ϵ – and assume that the divergence of F_2^ϵ vanishes:

$$\nabla\cdot F_2^\epsilon = \frac{1}{2}\Delta F_1^\epsilon = 0. \quad (3.2.3)$$

Remark 3.2.3. Condition (3.2.3) is equivalent to the noise being probabilistically stationary in the sense that it has the same law at every point in space, a property satisfied by space-time white noise and all of its standard approximations, such as those presented in Example 3.2.5 below.

As mentioned above, in our analysis of the small noise behaviour of (3.1.1) we let the noise terms ξ^ϵ converge to some noise ξ as $\epsilon\to 0$. Precisely, we make the following assumption.

Assumption 3.2.4 (Assumptions on the noise sequence)**.** For each $\epsilon > 0$ the family F^ϵ satisfies Assumption 3.2.2 and there exists a family of functions $F = F^0 = (f_m)_{m\in\mathbb{N}} \subseteq L^2(\mathbb{T}^d)$ such that, for some $C > 0$,

$$\sum_{m=1}^{\infty}(f_m^\epsilon, u)^2_{L^2(\mathbb{T}^d)} \leqslant C\,\|u\|^2_{L^2(\mathbb{T}^d)} \quad \forall u\in L^2(\mathbb{T}^d) \quad \forall\epsilon\geqslant 0,$$

and such that

$$\lim_{\epsilon\to 0}\sum_{m=1}^{\infty}(f_m - f_m^\epsilon, u)^2_{L^2(\mathbb{T}^d)} = 0 \quad \forall u\in L^2(\mathbb{T}^d).$$

Example 3.2.5 (Applicability of the noise assumptions)**.** Assumption 3.2.4 serves to deal with cylindrical Wiener processes for which the space $L^2(\mathbb{T}^d)$ is too small to live in (see e.g.

[DPZ92, Chapter 4]) and for which the Dean–Kawasaki equation (3.1.1) might be ill-posed. For example, this is the case for a space-time white noise ξ on $L^2(\mathbb{T}^d)$, which admits the spectral representation

$$\xi = \sum_{m=1}^{\infty} f_m(x) B_t^m,$$

for any orthonormal basis $(f_m)_{m\in\mathbb{N}}$ of $L^2(\mathbb{T}^d)$. In this setting, important examples of the approximating sequence of noise terms include spatial convolutions $\xi^\epsilon = \xi * \varphi^\epsilon$, that is $f_m^\epsilon = f_m * \varphi^\epsilon$ for a suitable mollifier φ; ultraviolet cut-offs like

$$\xi^\epsilon = \sum_{k\in\mathbb{Z}^d,\, |k|\leqslant M_\epsilon} e^{i2\pi k\cdot x}\, B_t^k,$$

for any arbitrary sequence $(M_\epsilon)_{\epsilon>0}$ increasing to infinity as $\epsilon \to 0$, where for simplicity we used the orthonormal basis of complex exponentials, and we have $f_k^\epsilon = e^{i2\pi k\cdot x}$ if $|k| \leqslant M_\epsilon$ and $f_k^\epsilon = 0$ otherwise; and weighted expansions like

$$\xi^{a^\epsilon} = \sum_{k\in\mathbb{Z}^d} a_k^\epsilon\, e^{i2\pi k\cdot x}\, B_t^k,$$

for coefficients $a = (a_k^\epsilon)_{k\in\mathbb{Z}^d}$ satisfying $\sum_{k\in\mathbb{Z}^d} |k|^2 |a_k^\epsilon|^2 < \infty$ and $\lim_{\epsilon\to 0} a_k^\epsilon = 1$. Indeed, an explicit computation gives, for the mollification,

$$F_1^\epsilon = \frac{1}{\epsilon^d}\int_{\mathbb{T}^d} |\varphi(y)|^2\, dy,\ F_2^\epsilon = 0,\ F_3^\epsilon = \frac{1}{\epsilon^{d+2}}\int_{\mathbb{T}^d} |\nabla\varphi(y)|^2\, dy \text{ and } \sum_{|m|=0}^{\infty} (f_m - f_m^\epsilon, u)_{L^2}^2 = \|u - u * \varphi^\epsilon\|_{L^2};$$

for the ultraviolet cut-off,

$$F_1^\epsilon(x) = \sum_{|k|=0}^{M_\epsilon} f_k(x)^2,\ F_2^\epsilon(x) = \frac{1}{2}\sum_{|k|=0}^{M_\epsilon} \nabla\,(f_k)^2,\ F_3^\epsilon(x) = \sum_{|k|=0}^{M_\epsilon} |\nabla f_k(x)|^2 \text{ and } \sum_{|k|=0}^{\infty} (f_k - f_k^\epsilon, u)_{L^2}^2 = \sum_{|k|>M_\epsilon}^{\infty} (f_k, u)_{L^2}^2;$$

and for the weighted expansion,

$$F_1^{a^\epsilon} = \sum_{m\in\mathbb{Z}^d} |a_m^\epsilon|^2,\ F_2^{a^\epsilon} = 0,\ F_3^{a^\epsilon} = \sum_{m\in\mathbb{Z}^d} |m|^2 |a_m^\epsilon|^2 \text{ and } \sum_{m\in\mathbb{Z}^d} (f_m - f_m^\epsilon, u)_{L^2}^2 = \sum_{m\in\mathbb{Z}^d} |a_m^\epsilon - 1|^2 (f_m, u)_{L^2}^2.$$

Finally we collect our assumptions on the coefficients ϕ, ν, σ, and discuss their meaning and applicability. The following set of assumptions coincides exactly with [FG21b, Assumption 4.1 and 5.2]. Together with Assumption 3.2.1 and 3.2.2 on the initial data and the noise, it guarantees the well-posedness of equation (3.2.1) in the sense of stochastic kinetic solutions (see Theorem 3.2.21).

Assumption 3.2.6 (Assumptions on the coefficients for the well-posedness of the equation)**.** Let $\phi, \sigma \in C([0,\infty)) \cap C^{1,1}_{\text{loc}}((0,\infty))$ and $\nu \in C([0,\infty))^d \cap C^{1,1}_{\text{loc}}((0,\infty))^d$, let $p \in [2,\infty)$ and $m \in [1,\infty)$, and for every $q \geqslant 2$ let $\Theta_{\phi,q} \in C([0,\infty)) \cap C^1((0,\infty))$ be the unique function satisfying

$$\Theta_{\phi,q}(0) = 0 \quad \text{and} \quad \dot{\Theta}_{\phi,q}(z) = z^{\frac{q-2}{2}} \left(\dot{\phi}(z)\right)^{\frac{1}{2}}. \tag{3.2.4}$$

Assume that the following eight conditions are satisfied.

(i) We have $\phi(0) = \sigma(0) = 0$ and $\dot{\phi} > 0$ on $(0, \infty)$.

(ii) There exists $c \in (0, \infty)$ such that

$$\phi(z) \leqslant c\,(1 + z^m) \ \text{ for every } \ z \in [0, \infty).$$

(iii) There exists $c \in (0, \infty)$ such that

$$\limsup_{z \to 0^+} \frac{\sigma^2(z)}{z} \leqslant c,$$

which in particular implies that $\sigma(0) = 0$.

(iv) There exists $c \in (0, \infty)$ such that

$$\sup_{z' \in [0,z]} \sigma^2(z') \leqslant c\,(1 + z + \sigma^2(z)) \ \text{ for every } \ z \in [0, \infty).$$

(v) There exists $c \in (0, \infty)$ such that

$$\sup_{z' \in [0,z]} \nu^2(z') \leqslant c\,(1 + z + \nu^2(z)) \ \text{ for every } \ z \in [0, \infty).$$

(vi) For $\Theta_{\phi,p}$ defined in (3.2.4), either there exists $c \in (0, \infty)$ and $\gamma \in [0, 1/2]$ such that

$$\left(\dot{\Theta}_{\phi,p}(z)\right)^{-1} \leqslant c\, z^{\gamma} \ \text{ for every } \ z \in (0, \infty), \tag{3.2.5}$$

or there exists $c \in (0, \infty)$ and $q \in [1, \infty)$ such that, for every $z, z' \in [0, \infty)$,

$$\left| z - z' \right|^q \leqslant c \left| \Theta_{\phi,p}(z) - \Theta_{\phi,p}(z') \right|^2. \tag{3.2.6}$$

(vii) For $\Theta_{\phi,2}$ and $\Theta_{\phi,p}$ defined in (3.2.4), there exists $c \in (0, \infty)$ such that, for every $z \in [0, \infty)$,

$$\sigma^2(z) \leqslant c\,\left(1 + z + \Theta^2_{\phi,2}(z)\right) \ \text{ and } \ z^{p-2}\sigma^2(z) \leqslant c\left(1 + z + \Theta^2_{\phi,p}(z)\right).$$

(viii) For every $\delta \in (0, 1)$ there exists $c_\delta \in (0, \infty)$ such that, for every $z \in (\delta, \infty)$,

$$\frac{[\dot{\sigma}(z)]^4}{\dot{\phi}(z)} + (\sigma(z)\dot{\sigma}(z))^2 + |\nu(z)| + \dot{\phi}(z) \leqslant c_\delta\,\left(1 + z + \Theta^2_{\phi,p}(z)\right).$$

Remark 3.2.7. With regards to Assumption 3.2.6, in Condition (i) the assumption $\phi(0) = 0$ is just a normalization, whereas the assumption $\sigma(0) = 0$ is crucial to prevent the noise from dragging solution ρ^ϵ towards negative values. Condition (ii) amounts to polynomial growth of ϕ at infinity. Condition (iii) requires that $\sigma(z)$ goes to zero as $z \to 0$ at least as fast as $\sqrt{z}$. Condition (iv) amounts to an assumption on the magnitude of the oscillations of σ at infinity, regardless of their frequency (see Example 3.2.11 below). For example, it is satisfied if σ^2 is monotone or if σ^2 grows linearly at infinity or if the oscillations of σ grow linearly at infinity. Identical considerations holds for the analogous condition (v) on ν. Condition (vi), both in the form (3.2.5) or (3.2.6), corresponds to a regularity assumption on ϕ: specifically, Hölder continuity of the inverse of the resulting function $\Theta_{\phi,p}$. Condition (vii) amounts to a growth condition at infinity on σ, with the aim that one of the convective terms, or better one of the resulting Itô correction convective terms, is somehow dominated by the diffusion, namely by $\Theta_{\phi,p}$. Similarly, Condition (viii) is a hypothesis on the growth away from zero of the convective terms, either the deterministic term ν or other Itô correction terms involving σ.

Assumption 3.2.6 on the coefficients, together with Assumption 3.2.1 and 3.2.2 on the initial data and the noise, is sufficient to guarantee existence and uniqueness for equation (3.2.2). To establish the CLT in probability (3.1.7) we just need some more control on the coefficients at infinity, regardless of their behaviour near the irregularity at zero.

Assumption 3.2.8 (Assumptions on the coefficients for the CLT in probability)**.** We assume that ϕ, ν, σ satisfy the following, for $p \in [2, \infty)$ given in Assumption 3.2.6.

(i) For some $k \in [0, 0 \vee \frac{p-4}{4}]$, for every $\delta > 0$ there exists $c_\delta \in (0, \infty)$ such that

$$|\sigma(z)| \leqslant c_\delta\,(1 + z^{k+1}) \quad \text{and} \quad |\dot{\sigma}(z)| \leqslant c_\delta\,(1 + z^{k}) \quad \forall\ z \in (\delta, \infty).$$

(ii) We have $\phi, \nu \in C^2_{\text{loc}}((0,\infty))$ and there exists $g \in [0, 0 \vee \frac{p}{2(k+1)} - 2]$ such that, for every $\delta > 0$ there exists $c_\delta \in (0, \infty)$ such that

$$|\ddot{\phi}(z)| + |\ddot{\nu}(z)| \leqslant c_\delta\,(1 + z^{g}) \quad \forall\ z \in (\delta, \infty).$$

To establish the stronger version (3.1.8) of the CLT, we will replace Assumption 3.2.8 with the following stronger version. It is essentially the requirement that Assumption 3.2.8 holds uniformly on $(0, \infty)$ up to the irregularity at zero. We stress that *this assumption is NOT needed* for the general CLT in probability (3.1.7).

Assumption 3.2.9 (Assumptions on the coefficients for the CLT in $L^2(\Omega)$)**.** We assume that ϕ, ν, σ satisfy the following, for some constant $c \geqslant 0$, for $p \in [2, \infty)$ given in Assumption 3.2.6.

(i) For some $\theta \in (0, \frac{1}{2})$ and some $k \in [0, 0 \vee \frac{p-4}{4}]$, for all $z \in (0, \infty)$,

$$|\sigma(z)| \leqslant c\,(1 + z^{k+1}), \quad |\dot{\sigma}(z)| \leqslant c\,(1 + z^{-\theta} + z^{k}) \quad \text{and} \quad |\sigma(z)\dot{\sigma}(z)| \leqslant c\,(1 + z^{2k+1}).$$

(ii) We have $\phi, \nu \in C^2((0,\infty))$ and there exists $g \in [0, 0 \vee \frac{p}{2(k+1)} - 2]$ such that, for all $z \in (0, \infty)$,

$$|\ddot{\phi}(z)| + |\ddot{\nu}(z)| \leqslant c\,(1 + z^{g}).$$

Remark 3.2.10. Assumption 3.2.8(i) or 3.2.9(i) serves to control the convective term $\nabla\cdot(\sigma(\rho^\epsilon)\circ\xi^\epsilon)$, so that it is dominated by the diffusion, and indeed it might be replaced by the somewhat more explicit conditions $|\sigma(z)|\leqslant c\,(1+\phi^{1/2}(z))$ and $|\dot\sigma(z)|\leqslant c\,(1+(\phi^{1/2})'(z))$ for all $z\in(0,\infty)$. Assumption 3.2.8(ii) or 3.2.9(ii) corresponds to polynomial growth at infinity, and also regularity near zero for 3.2.9(ii), of the diffusion nonlinearity ϕ and of the deterministic convective term ν. It is needed to recast the diffusive term $\Delta\big(\phi(\rho^\epsilon)-\phi(\bar\rho)\big)$, or the convective term respectively, of the nonequilibrium fluctuation v^ϵ in (3.1.6) in terms of the fluctuation v^ϵ itself.

Example 3.2.11 (Applicability of the coefficient assumptions)**.** In the porous media case $\phi(z)=z^{m_0}$ all the conditions involving ϕ in Assumption 3.2.6 and 3.2.8 are verified in the full regime $m_0\in(0,\infty)$. In this case, the function $\Theta_{\phi,p}$ defined in (3.2.4) is given for a constant $c_{p,m_0}\in(0,\infty)$ by

$$\Theta_{\phi,p}(z)=c_{p,m_0}\,z^{\frac{m_0+p-1}{2}}.$$

The conditions are satisfied by taking $m=1$, $p=2$, $g=k=0$ and choosing option (3.2.5) with $\gamma=\frac{1-m_0}{2}$ when $m_0\in(0,1)$, and by taking $m=m_0$, any $p\geqslant 4\vee m_0^2$, $k=0\vee\frac{m_0-2}{2}$, $g=0\vee m_0-2$ and choosing option (3.2.6) with $q=m_0+p-1$ when $m_0\in[1,\infty)$.

In the important model case $\phi(z)=z^{m_0}$, $\sigma(z)=\phi^{1/2}(z)=z^{\frac{m_0}{2}}$, with $\nu=0$ for simplicity, all the conditions in Assumption 3.2.6 and 3.2.8 are verified for any $m_0\in[1,\infty)$ with the choices of m, p, k, g given above.

In fact, Assumption 3.2.8 and all the conditions in Assumption 3.2.6 except for condition (iii) are verified also for $m_0\in(0,1)$, with the corresponding choices of m, p, k, g, γ. Only condition (iii) in Assumption 3.2.6 breaks down. Indeed, already to prove well-posedness of (3.1.1) we need the noise coefficient σ to be $\frac{1}{2}$-Hölder near the irregularity point zero.

We also comment on Assumption 3.2.6(iv) on the oscillations of σ. An interesting example satisfying all the conditions in Assumption 3.2.6 and 3.2.8, is given by $\sigma^2(z)=z^{m_0}+z\sin\left(z^h\right)$ for every $m_0,h\in[1,\infty)$. That is, condition (iv) imposes a condition on the growth of the magnitude of the oscillations of σ^2 at infinity, but not on the growth of the frequency of the oscillations.

Finally the stronger Assumption 3.2.9 for the CLT in $L^2(\Omega)$, together with Assumption 3.2.6, applies to the model case $\phi(z)=z^{m_0}$, $\sigma(z)=z^{\frac{m_0}{2}}$ in the regime $m_0\in[2,\infty)$, or to the linear case $\phi(z)=z$ provided we consider a different noise coefficient σ satisfying the assumptions, for example $\sigma(z)=z^{\frac{1}{s}}$ for some $s\in(0,2)$.

3.2.2 The zero noise limit and the generalized Ornstein–Uhlenbeck process

In this section we collect some well-established results on the well-posedness of the limiting equations (3.1.2) and (3.1.3). As regards the deterministic limit (3.1.2) we have the following classical theory. We refer the reader to standard monographs like [Vaz07] or [Lie96].

Definition 3.2.12. A weak solution of (3.1.2) is a function $\bar{\rho} \in L^1([0,T];L^1(\mathbb{T}^d))$ such that $\phi(\bar{\rho}), \nu(\bar{\rho}) \in L^1([0,T];L^1(\mathbb{T}^d))$ and such that, for any $t \in [0,T]$ and any $\psi \in C^\infty(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d} \bar{\rho}(x,t)\,\psi(x)\,dx = \int_{\mathbb{T}^d} \rho_0(x)\,\psi(x)\,dx + \int_0^t\int_{\mathbb{T}^d} \phi(\bar{\rho})(x,s)\,\Delta\psi(x)\,dx\,ds + \int_0^t\int_{\mathbb{T}^d} \nu(\bar{\rho})(x,s)\,\nabla\psi(x)\,dx\,ds.$$

Theorem 3.2.13. Under Assumption 3.2.6, for any nonnegative $\rho_0 \in L^2(\mathbb{T}^d)$ bounded away from zero, there exists a unique weak solution $\bar{\rho}$ of (3.1.2) in the sense of Definition 3.2.12. Furthermore the solution satisfies $\bar{\rho} \in L^2([0,T];L^2(\mathbb{T}^d))$ and $\phi(\bar{\rho}) \in L^2([0,T];H^1(\mathbb{T}^d))$.

Remark 3.2.14. Except for the initial data ρ_0, which can be random, equation (3.1.2) is purely deterministic. Furthermore, one immediately see that for $\rho_0 \in (0,\infty)$ a (possibly random) constant, the solution is simply the constant $\bar{\rho}(x,t,\omega) \equiv \rho_0(\omega)$.

We now consider the stochastic equation

$$\partial_t v = \Delta\,(a\,v) - \nabla\cdot\left(b\,v + c\,\dot{\xi}\right) \;\text{ in } \mathbb{T}^d \times (0,T), \quad v(\cdot,0) = 0 \;\text{ in } \mathbb{T}^d \times \{0\}, \tag{3.2.7}$$

for some $\mathcal{F}_0$-measurable random variables $a, b \in L^0(\Omega;\mathbb{R})$ with $a \geqslant r > 0$ a.e. for some $r \in (0,\infty)$, and $c \in L^2(\Omega;\mathbb{R})$, and for the noise $\xi = \lim_{\epsilon\to 0}\xi^\epsilon$ specified in Assumption 3.2.4. First of all, we make precise our notion of solution, here $\alpha \geqslant 0$ depends on the regularity of the noise ξ, which satisfies Assumption 3.2.4.

Definition 3.2.15. A weak solution in $H^{-\alpha}(\mathbb{T}^d)$ of equation (3.2.7) is a predictable process $v \in L^2(\Omega \times (0,T);H^{-\alpha}(\mathbb{T}^d))$ such that, for every $t \in [0,T]$ and every $w \in H^{\alpha+2}(\mathbb{T}^d)$, $\mathbb{P}$-almost surely we have

$$\langle v(s), w(s)\rangle\Big|_{s=0}^{s=t} = \int_0^t \langle v(r), a\,\Delta w(r) + b\nabla w(r)\rangle\,dr + \int_0^t\int_{\mathbb{T}^d} c\,\nabla w(x,r)\,\mathrm{d}\xi.$$

We have the following well-posedness result. This can be proved for example via semigroup methods (see e.g. [DPZ92, Chapter 5]) or Fourier analysis (see e.g. [DFG20]).

Theorem 3.2.16. Let ξ satisfies Assumption 3.2.4. There exists a unique weak solution $v \in L^2(\Omega \times (0,T);H^{-\alpha}(\mathbb{T}^d))$ of equation (3.2.7) for any $\alpha > \frac{d}{2}$. Furthermore, if $a_n \to a$ and $b_n \to b$ almost surely, with $a_n \geqslant r > 0$ for every $n \in \mathbb{N}$, and $c_n \to c$ in $L^2(\Omega;\mathbb{R})$, as $n \to \infty$, and we denote by v_n and v the corresponding solutions of (3.2.7), then we have $v_n \to v$ in $L^2(\Omega \times (0,T);H^{-\alpha}(\mathbb{T}^d))$ for any $\alpha > \frac{d}{2}$.

3.2.3 The generalized Dean–Kawasaki equation

In this section we summarize the solution theory to equation (3.1.1), interpreted in its Itô formulation (3.2.2), as put forward in [FG21b]. In particular, we introduce the notions of weak solution and of stochastic kinetic solution and we collect some of their properties.

The notion of weak solution is classical.

Definition 3.2.17. Let $\rho_0 \in L^1(\Omega; L^1(\mathbb{T}^d))$ be nonnegative and $\mathcal{F}_0$-measurable. A weak solution to (3.2.2) with initial data ρ_0 is an $\mathcal{F}_t$-adapted nonnegative continuous $L^1(\mathbb{T}^d)$-valued process $\rho \in L^1(\Omega \times [0,T]; L^1(\mathbb{T}^d))$ such that almost surely

$$\rho, \phi^{1/2}(\rho), \sigma(\rho) \in L^2([0,T]; H^1(\mathbb{T}^d)),\ \dot{\sigma}(\rho) \in L^2([0,T]; L^2(\mathbb{T}^d)) \text{ and } \nu(\rho) \in L^1([0,T]; L^1(\mathbb{T}^d)),$$

and such that almost surely, for every $\psi \in C^\infty(\mathbb{T}^d)$ and every $t \in [0,T]$,

$$\begin{aligned}\int_{\mathbb{T}^d} \rho(x,t)\psi(x)\,dx = &\int_{\mathbb{T}^d} \rho_0(x)\,\psi(x)\,dx \\ &+ \int_0^t \int_{\mathbb{T}^d} \Delta\psi(x)\,\phi(\rho)(x,s)\,dx\,ds \\ &+ \int_0^t \int_{\mathbb{T}^d} \nabla\psi(x)\,\nu(\rho)(x,s)\,dx\,ds \\ &+ \sqrt{\epsilon}\int_0^t \int_{\mathbb{T}^d} \nabla\psi(x)\,\sigma(\rho)(x,s)\,d\xi^\epsilon \\ &- \frac{\epsilon}{2}\int_0^t \int_{\mathbb{T}^d} \nabla\psi(x)(F_1^\epsilon \dot{\sigma}(\rho)\nabla\sigma(\rho) + \dot{\sigma}(\rho)\sigma(\rho)F_2^\epsilon)\,dx\,ds.\end{aligned} \tag{3.2.8}$$

When the coefficients of the equation are sufficiently smooth, weak solutions exist and are unique.

Theorem 3.2.18 (Theorem 5.20 in [FG21b])**.** Let ρ_0 satisfy Assumption 3.2.1(i), let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ, ν and σ satisfy Assumption 3.2.6. Suppose furthermore that ϕ, σ satisfy the additional assumptions:

(i) $\dot{\phi}(z) \geqslant c > 0$ for all $z \in (0,\infty)$, for some $c \in (0,\infty)$;

(ii) $\sigma \in C([0,\infty)) \cap C^\infty((0,\infty))$ with $\dot{\sigma} \in C_c^\infty([0,\infty))$.

There exists a unique weak solution to equation (3.2.2) in the sense of Definition 3.2.17.

Despite being quite natural, the notion of weak solution is often doomed to fail when σ is not smooth enough. Indeed, as discussed in Section 3.1.1, already in the prototypical Dean–Kawasaki case $\sigma(z) = \sqrt{z}$, it is not clear how to interpret some of the terms featuring in (3.2.8). To handle rougher coefficients a wider solution theory is needed.

Namely, given a nonnegative function $\rho : \mathbb{T}^d \times [0,T] \to \mathbb{R}$, one introduces an additional variable $z \in \mathbb{R}$ and consider the *kinetic function* $\chi : \mathbb{R} \times \mathbb{T}^d \times [0,T] \to \{0,1\}$ of ρ defined by

$$\chi(z,x,t) = \bar{\chi}(z,\rho(x,t)) = \mathbf{1}_{\{0<z<\rho(x,t)\}},$$

for $\bar{\chi}(z,v) := \mathbf{1}_{\{0<z<v\}} - \mathbf{1}_{\{v<z<0\}}$. The *velocity* variable z essentially corresponds to the magnitude of the solution ρ. A formal computation presented in [FG21b], based on the fundamental identity

$$S(\rho(x,t)) = \int_{\mathbb{R}} \dot{S}(z)\chi(z,x,t)\,dz,$$

for every smooth function $S:\mathbb{R}\to\mathbb{R}$ with $S(0)=0$, motivates the notion of stochastic kinetic solution.

Definition 3.2.19. Let $\rho_0 \in L^1(\Omega; L^1(\mathbb{T}^d))$ be nonnegative and $\mathcal{F}_0$-measurable. A *stochastic kinetic solution* of (3.2.2) is a nonnegative, a.s. continuous $L^1(\mathbb{T}^d)$-valued $\mathcal{F}_t$-predictable function $\rho \in L^1(\Omega\times[0,T]; L^1(\mathbb{T}^d))$ that satisfies the following three properties.

(i) *Preservation of mass*: a.s. for every $t\in[0,T]$,

$$\|\rho(\cdot,t)\|_{L^1(\mathbb{T}^d)} = \|\rho_0\|_{L^1(\mathbb{T}^d)}.$$

(ii) *Integrability of the flux*: we have that

$$\sigma(\rho)\in L^2(\Omega; L^2(\mathbb{T}^d\times[0,T])) \text{ and } \nu(\rho)\in L^1(\Omega; L^1(\mathbb{T}^d\times[0,T];\mathbb{R}^d)).$$

(iii) *Local regularity*: for every $K\in\mathbb{N}$,

$$[(\rho\wedge K)\vee 1/K]\in L^2(\Omega; L^2([0,T]; H^1(\mathbb{T}^d))).$$

Furthermore, there exists a *kinetic measure* q, that is a map q from Ω to the space of nonnegative, locally finite measures on $\mathbb{T}^d\times(0,\infty)\times[0,T]$ such that, for every $\psi\in\mathrm{C}^\infty(\mathbb{T}^d\times(0,\infty))$, the process

$$\Omega\times[0,T]\ni(\omega,t)\mapsto\int_0^t\int_{\mathbb{R}}\int_{\mathbb{T}^d}\psi(x,\xi)\,\mathrm{d}q(\omega) \text{ is } \mathcal{F}_t\text{-predictable},$$

that satisfies the following three properties.

(iv) *Regularity*: almost surely as nonnegative measures,

$$\delta_0(z-\rho)\dot{\phi}(z)|\nabla\rho|^2\leqslant q \text{ on } \mathbb{T}^d\times(0,\infty)\times[0,T].$$

(v) *Vanishing at infinity*: we have that

$$\lim_{M\to\infty}\mathbb{E}\left[q(\mathbb{T}^d\times[M,M+1]\times[0,T])\right]=0. \tag{3.2.9}$$

(vi) *The equation*: almost surely, for every $\psi\in\mathrm{C}_c^\infty(\mathbb{T}^d\times(0,\infty))$ and every $t\in[0,T]$,

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{T}^d}\chi(x,z,t)\psi(x,z) &= \int_{\mathbb{R}}\int_{\mathbb{T}^d}\overline{\chi}(\rho_0)\psi(x,z) - \int_0^t\int_{\mathbb{T}^d}\dot{\phi}(\rho)\nabla\rho\cdot(\nabla\psi)(x,\rho) \\ &- \frac{\epsilon}{2}\int_0^t\int_{\mathbb{T}^d}F_1^\epsilon[\dot{\sigma}(\rho)]^2\nabla\rho\cdot(\nabla\psi)(x,\rho) - \frac{\epsilon}{2}\int_0^t\int_{\mathbb{T}^d}\sigma(\rho)\dot{\sigma}(\rho)F_2^\epsilon\cdot(\nabla\psi)(x,\rho) \\ &- \int_0^t\int_{\mathbb{R}}\int_{\mathbb{T}^d}\partial_z\psi(x,z)\,\mathrm{d}q + \frac{\epsilon}{2}\int_0^t\int_{\mathbb{T}^d}\left(\sigma(\rho)\dot{\sigma}(\rho)\nabla\rho\cdot F_2^\epsilon + F_3^\epsilon\sigma^2(\rho)\right)(\partial_z\psi)(x,\rho) \\ &- \int_0^t\int_{\mathbb{T}^d}\psi(x,\rho)\nabla\cdot\nu(\rho)\mathrm{d}t - \sqrt{\epsilon}\int_0^t\int_{\mathbb{T}^d}\psi(x,\rho)\nabla\cdot(\sigma(\rho)\,\mathrm{d}\xi^\epsilon)\,.\end{aligned} \tag{3.2.10}$$

Remark 3.2.20. (i) We write $(\nabla\psi)(x,\rho(x,t)) = \nabla\psi(x,z)|_{z=\rho(x,t)}$ to mean the gradient of $\nabla_x\psi(x,z)$ evaluated at the point $(x,\rho(x,t))$ as opposed to the full gradient of the composition $\psi(x,\rho(x,t))$.

(ii) For $\mathcal{F}_t$-adapted processes $g_t \in L^2(\Omega\times[0,T];L^2(\mathbb{T}^d))$ and $h_t \in L^2(\Omega\times[0,T];H^1(\mathbb{T}^d))$ and for $t\in[0,T]$, we write

$$\int_0^t\int_{\mathbb{T}^d} g_s\nabla\cdot(h_s\,\mathrm{d}\xi^\epsilon) = \sum_{k=1}^\infty\left(\int_0^t\int_{\mathbb{T}^d} g_s f_k^\epsilon \nabla h_s\cdot \mathrm{d}B_s^k + \int_0^t\int_{\mathbb{T}^d} g_s h_s\nabla f_k^\epsilon\cdot \mathrm{d}B_s^k\right),$$

where the integrals are interpreted in the Itô sense.

The upshot of the kinetic formulation of (3.2.2) is of course that stochastic kinetic solutions exist and are unique for a much wider class of coefficients ϕ and σ.

Theorem 3.2.21 (Theorem 5.29 in [FG21b]). Let ρ_0 satisfy Assumption 3.2.1(i), let ξ^ϵ satisfy Assumption 3.2.2, and let ϕ, ν and σ satisfy Assumption 3.2.6. There exists a unique stochastic kinetic solution ρ to (3.2.2) in the sense of Definition 3.2.19. Furthermore, for $p\geqslant 2$ and $m\geqslant 1$ given in Assumption 3.2.6, for $n=2$ and $n=p$, it satisfies

$$\begin{aligned}\sup_{t\in[0,T]}\mathbb{E}\left[\int_{\mathbb{T}^d}|\rho(x,t)|^n\,dx\right] + \inf_z \dot\phi(z)\,\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}|\rho|^{n-2}|\nabla\rho|^2\,dx\,dt\right] \qquad (3.2.11)\\ \leqslant C\left(1+\epsilon\,\|F_3^\epsilon\|_\infty\right)^{\frac{d}{2}(p+m)}\left(1+\mathbb{E}\left[\|\rho_0\|_{L^1(\mathbb{T}^d)}^{m+n-1}+\|\rho_0\|_{L^n(\mathbb{T}^d)}^n\right]\right),\end{aligned}$$

for a constant $C=C(T,\phi,\nu,\sigma)$ depending on ϕ, ν and σ only through the constants c appearing in Assumption 3.2.6.

Weak and stochastic kinetic solutions are obviously related. First of all, when they exist, weak solutions are stochastic kinetic solutions.

Proposition 3.2.22 (Proposition 5.21 in [FG21b]). Let ρ_0 satisfy Assumption 3.2.1(i), let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ, ν, σ satisfy Assumption 3.2.6 and the additional assumptions (i)-(ii) in Theorem 3.2.18. Let ρ be the unique weak solution of (3.2.2) in the sense of Definition 3.2.17, then ρ is also a stochastic kinetic solution in the sense of Definition 3.2.19 and its kinetic measure is given by $q=\delta_0(z-\rho)\dot\phi(z)|\nabla\rho|^2dx\,dt$.

The following statement asserts that stochastic kinetic solutions, and thus also weak solutions thanks to the previous proposition, depend continuously on the coefficients of the equation, and it will be crucial in our arguments. The proof is a straightforward adaptation of the final part of the proof of [FG21b, Theorem 5.29].

Proposition 3.2.23. Let ρ_0 satisfy Assumption 3.2.1(i), let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ,ν,σ satisfy Assumption 3.2.6. For $\eta\in(0,1)$, let $\phi^\eta,\nu^\eta,\sigma^\eta$ be a sequence of coefficients such that:

(i) $\phi^\eta, \nu^\eta, \sigma^\eta$ satisfy Assumption 3.2.6 uniformly in $\eta \in (0,1)$;

(ii) $\phi^\eta \to \phi$, $\nu^\eta \to \nu$ and $\sigma^\eta \to \sigma$ in $C_{\mathrm{loc}}([0,\infty)) \cap C^1_{\mathrm{loc}}((0,\infty))$ as $\eta \to 0$.

For each $\eta \in (0,1)$, let $\rho^{\epsilon,\eta}$ be the unique stochastic kinetic solution of

$$\begin{cases} \mathrm{d}\rho^{\epsilon,\eta} = \Delta\phi^\eta(\rho^{\epsilon,\eta})\mathrm{d}t - \nabla\cdot\nu^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}t - \sqrt{\epsilon}\,\nabla\cdot(\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon) \\ \qquad\qquad + \frac{\epsilon}{2}\nabla\cdot\left(F_1^\epsilon\,(\dot\sigma^\eta(\rho^{\epsilon,\eta}))^2\,\nabla\rho^{\epsilon,\eta} + \dot\sigma^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon\right)\mathrm{d}t, \\ \rho^{\epsilon,\eta}(\cdot,0) = \rho_0, \end{cases} \tag{3.2.12}$$

and let ρ^ϵ be the stochastic kinetic solution of (3.2.2) with initial data ρ_0, in the sense of Definition 3.2.19. Then we have

$$\lim_{\eta\to 0}\|\rho^{\epsilon,\eta} - \rho^\epsilon\|_{L^1([0,T];L^1(\mathbb{T}^d))} = 0 \quad \text{in probability.}$$

For the sake of completeness, we end this section by establishing that, if the stochastic convective term σ is more regular near zero and somewhat dominated by the diffusion, we have the converse of Proposition 3.2.22: stochastic kinetic solutions are weak solutions. Since we will not need this fact in the following, we only sketch the main details. The proof is based on the following proposition, which establishes the H^1-regularity of the nonlinear term $\phi^{1/2}(\rho)$ imposed by the diffusion.

Proposition 3.2.24 (Corollary 5.31 in [FG21b]). Let ρ_0 satisfy Assumption 3.2.1, let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ, ν and σ satisfy Assumption 3.2.6. Let ρ be the unique stochastic kinetic solution to equation (3.2.2). Assume furthermore the following:

(i) there exists $c \in (0,\infty)$ such that $|\sigma(z)| \leqslant c\,\phi^{1/2}(z)$ for every $z \in [0,\infty)$;

(ii) there exists $c \in (0,\infty)$ such that

$$|\nu(z)| + \dot\phi(z) \leqslant c\,(1 + z + \phi(z)) \text{ for every } z \in [0,\infty);$$

(iii) the map $z \mapsto \log(\phi(z))$ is locally integrable on $[0,\infty)$.

Then for the unique function $\Psi_\phi \in C([0,\infty);\mathbb{R}) \cap C^1((0,\infty);\mathbb{R})$ satisfying $\Psi_\phi(0) = 0$ and $\dot\Psi_\Phi(z) = \log(\phi(z))$, for a constant $C = C(T,\phi,\nu,\sigma)$ depending on ϕ, ν and σ only through the constants c appearing in Assumption 3.2.6 and in the further hypotheses (i)-(iii) above, we have

$$\begin{aligned} &\mathbb{E}\left[\sup_{t\in[0,T]}\int_{\mathbb{T}^d}\Psi_\phi(\rho(x,t))\right] + \mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}\left|\nabla\phi^{1/2}(\rho)\right|^2\right] + \inf_{z>0}\dot\phi(z)\,\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}\mathbf{1}_{\{\rho>0\}}\frac{\phi'(\rho)}{\phi(\rho)}|\nabla\rho|^2\right] \\ &\leqslant C\left(1 + \|\epsilon F_1^\epsilon\|_\infty + \|\epsilon F_3^\epsilon\|_\infty^{\frac{d}{2}(m+1)}\right)\left(1 + \mathbb{E}\left[\|\rho_0\|^m_{L^1(\mathbb{T}^d)} + \int_{\mathbb{T}^d}\Psi_\phi(\rho_0)\right]\right). \end{aligned} \tag{3.2.13}$$

We can now prove the following.

Proposition 3.2.25. In the setting of Proposition 3.2.24, assume furthermore that, for some $c > 0$,

$$|\dot{\sigma}(z)| \leqslant c \left(\dot{\phi^{1/2}}\right)(z) \quad \text{and} \quad |\dot{\sigma}(z)|^2 \leqslant c\,(1 + z + \phi(z)) \quad \forall\, z \in (0, \infty). \tag{3.2.14}$$

Then the stochastic kinetic solution ρ of (3.2.2) satisfies the weak formulation (3.2.8) of the equation.

Proof. The idea is to test the kinetic formulation (3.2.10) of the equation against functions of the form $\psi(x, z) = \varphi(x) K_\delta(z)$ for arbitrary $\varphi \in C^\infty(\mathbb{T}^d)$ and for a smooth cut-off $K_\delta : \mathbb{R} \to [0, 1]$ such that

$$K_\delta(z) = \begin{cases} 1 & \text{if } z \in [2\delta, 1/\delta], \\ 0 & \text{if } z \in [0, \delta] \cup [2/\delta, \infty), \end{cases}$$

and then let $\delta \to 0$ and show that, along a suitable subsequence $\delta_n \to 0$, the resulting expression coincides with the weak formulation (3.2.8).

Except for the term involving the kinetic measure, each of the integrals in (3.2.10) is shown to converge to the corresponding quantity in (3.2.8) or to zero by dominated convergence. This is justified using Assumption 3.2.6, the regularity estimates (3.2.11) and (3.2.13), the further hypotheses (i)-(iii) in Proposition 3.2.24, and the hypothesis (3.2.14), which together with (3.2.13) implies that $\dot{\sigma}(\rho) \in L^2([0, T]; L^2(\mathbb{T}^d))$ and $\sigma(\rho) \in L^2([0, T]; H^1(\mathbb{T}^d))$ with

$$|\nabla \sigma(\rho)| \leqslant C \left| \nabla \phi^{1/2}(\rho) \right|.$$

The term involving the kinetic measure is shown to converge to zero along a subsequence $\delta_n \to 0$ using property (3.2.9) and the following property, which is proved in [FG21b, Proposition 4.6],

$$\liminf_{\delta \to 0} \left(\delta^{-1} q(\mathbb{T}^d \times [\delta/2, \delta] \times [0, T]) \right) = 0.$$

□

3.3 The central limit theorem for fluctuations

3.3.1 The central limit theorem in $L^2(\Omega)$ for more regular coefficients

This section is devoted to the proof of our central limit theorems: Theorem 3.3.3 and 3.3.10. In the first part we obtain the stronger version (3.1.8) of the CLT, for coefficients satisfying Assumption 3.2.6 and the stronger Assumption 3.2.9. In the second part we extend the result to rougher coefficients satisfying Assumption 3.2.6 and Assumption 3.2.8 only.

The nonlinearity of the diffusion ϕ makes it difficult to apply standard Fourier analysis and we have to accompany this with some moment estimates. Our ansatz is that v^ϵ converge to v, solving (3.1.3), which can be distribution valued. Therefore we expect the norm of v^ϵ

in any function space to blow up as $\epsilon \to 0$. After an approximation lemma, the following proposition quantifies this explosion.

Lemma 3.3.1. Let $\phi, \nu, \sigma \in W^{1,1}_{\mathrm{loc}}([0,\infty)) \cap C^1((0,\infty))$ with $\phi(0) = \sigma(0) = 0$ and $\dot{\phi}(z) > 0$ for every $z \in (0,\infty)$. There exists a sequence of approximations $(\phi^\eta, \nu^\eta, \sigma^\eta)_{\eta \in (0,1)}$ such that

(i) $\phi^\eta \to \phi$, $\nu^\eta \to \nu$ and $\sigma^\eta \to \sigma$ in $C^1_{\mathrm{loc}}(0,\infty) \cap C_{\mathrm{loc}}([0,\infty))$ as $\eta \to 0$;

(ii) $\phi^\eta, \nu^\eta, \sigma^\eta$ satisfy Assumption 3.2.6, 3.2.8, 3.2.9 and Assumption (i)-(ii) in Theorem 3.2.18.

Furthermore, if ϕ, ν and σ satisfy Assumption 3.2.6, 3.2.8, 3.2.9 or Assumption (i) or (ii) in Theorem 3.2.18, then $\phi^\eta, \nu^\eta, \sigma^\eta$ satisfy Assumption 3.2.6, 3.2.8, 3.2.9 or Assumption (i) or (ii) in Theorem 3.2.18, respectively, uniformly in $\eta \in (0,1)$.

Proof. A standard smoothing procedure by convolution and cut-off. □

Proposition 3.3.2. Let ρ_0 satisfy Assumption 3.2.1(ii), let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ, ν, σ satisfy Assumption 3.2.6 and 3.2.9, for some $p \geqslant 2$ and $m \geqslant 1$. Let ρ^ϵ be the stochastic kinetic solution to (3.2.2) with initial data ρ_0, let $\bar{\rho} \equiv \rho_0$ be the solution of (3.1.2) and define $v^\epsilon = \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho})$. Then, for any $T \in [0,\infty)$ and any $h \in \left[1, \frac{p}{(k+1)}\right]$, we have

$$\mathbb{E}\left[\|v^\epsilon\|^h_{L^h([0,T];\mathbb{T}^d)}\right] \leqslant C\,\|F_3^\epsilon\|_\infty^{h/2}(1+\epsilon\,\|F_3^\epsilon\|_\infty)^{\frac{d}{2}(p+m)}\left(1+\mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right), \tag{3.3.1}$$

for a constant $C = C(T, \phi, \nu, \sigma, p)$ depending on ϕ, ν and σ only through the constants c featuring in Assumption 3.2.6 and 3.2.9.

Proof. Consider the sequence $(\phi^\eta, \nu^\eta, \sigma^\eta)_{\eta \in (0,1)}$ of approximations given by Lemma 3.3.1. Let $\rho^{\epsilon,\eta}$ be the corresponding weak solution of (3.2.12), whose existence is guaranteed by Theorem 3.2.18, and let $\bar{\rho} \equiv \rho_0$ be the unique solution of $\partial_t \bar{\rho} = \Delta\phi^\eta(\bar{\rho}) - \nabla\cdot\nu^\eta(\bar{\rho})$ with initial data ρ_0. It follows that the fluctuations $v^{\epsilon,\eta} = \epsilon^{-1/2}(\rho^{\epsilon,\eta} - \bar{\rho})$ are a weak solution, in the sense of Definition 3.2.17, to

$$\begin{cases} \mathrm{d}v^{\epsilon,\eta} = \Delta\left(\epsilon^{-1/2}\big(\phi^\eta(\rho^{\epsilon,\eta}) - \phi^\eta(\bar{\rho})\big)\right)\mathrm{d}t - \nabla\cdot\left(\epsilon^{-1/2}\big(\nu^\eta(\rho^{\epsilon,\eta}) - \nu^\eta(\bar{\rho})\big)\right)\mathrm{d}t - \nabla\cdot(\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon) \\ \qquad\qquad + \frac{\sqrt{\epsilon}}{2}\nabla\cdot\left(F_1^\epsilon\dot{\sigma}^\eta(\rho^{\epsilon,\eta})\nabla\sigma^\eta(\rho^{\epsilon,\eta}) + \dot{\sigma}^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon\right)\mathrm{d}t, \\ v^{\epsilon,\eta}(\cdot,0) = 0. \end{cases}$$

Applying Itô's formula, in the version proved in Krylov [Kry12], yields:

$$\begin{aligned}
\int_{\mathbb{T}^d} |v^{\epsilon,\eta}(x,t)|^h\,dx = &-\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\,\nabla v^{\epsilon,\eta}\nabla\left(\epsilon^{-1/2}\left(\phi^\eta(\rho^{\epsilon,\eta})-\phi^\eta(\bar\rho)\right)\right)dx\,ds\\
&+\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\,\nabla v^{\epsilon,\eta}\epsilon^{-1/2}\big(\nu^\eta(\rho^{\epsilon,\eta})-\nu^\eta(\bar\rho)\big)\,dx\,ds\\
&-\int_0^t\int_{\mathbb{T}^d} h\,|v^{\epsilon,\eta}(x,t)|^{h-2}\,v^{\epsilon,\eta}\nabla\cdot\left(\sigma^\eta(\rho^{\epsilon,\eta})\mathrm{d}\xi^\epsilon\right)\\
&-\frac{\sqrt{\epsilon}}{2}\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\,\nabla v^{\epsilon,\eta}\\
&\qquad\cdot\left(F_1^\epsilon\dot\sigma^\eta(\rho^{\epsilon,\eta})\nabla\sigma^\eta(\rho^{\epsilon,\eta})+\dot\sigma^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon\right)dx\,ds\\
&+\frac{1}{2}\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\\
&\qquad\cdot\left(|\nabla\sigma^\eta(\rho^{\epsilon,\eta})|^2F_1^\epsilon+2\sigma^\eta(\rho^{\epsilon,\eta})\nabla\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon+\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\right)dx\,ds.
\end{aligned}$$

Rearranging the terms, and using that $\nabla v^{\epsilon,\eta}=\epsilon^{-1/2}\nabla\rho^{\epsilon,\eta}$ since $\bar\rho\equiv$ constant, the smoothness of ϕ^η, ν^η, σ^η, and the regularity properties of $\rho^{\epsilon,\eta}$ guaranteed by Theorem 3.2.18, we obtain

$$\begin{aligned}
&\int_{\mathbb{T}^d} |v^{\epsilon,\eta}(x,t)|^h\,dx+\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,\nabla v^{\epsilon,\eta}\nabla\left(\epsilon^{-1/2}\left(\phi^\eta(\rho^{\epsilon,\eta})-\phi^\eta(\bar\rho)\right)\right)dx\,ds\\
&=\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,\nabla v^{\epsilon,\eta}\epsilon^{-\frac{1}{2}}\big(\nu^\eta(\rho^{\epsilon,\eta})-\nu^\eta(\bar\rho)\big)\,dx\,ds\\
&\quad-\int_0^t\int_{\mathbb{T}^d} h\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,v^{\epsilon,\eta}\nabla\cdot\left(\sigma^\eta(\rho^{\epsilon,\eta})\mathrm{d}\xi^\epsilon\right)\\
&\quad+\frac{1}{2}\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,\sigma^\eta(\rho^{\epsilon,\eta})\nabla\sigma^\eta(\rho^{\epsilon,\eta})\cdot F_2^\epsilon\,dx\,ds\\
&\quad+\frac{1}{2}\int_0^t\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\,dx\,ds.
\end{aligned}\tag{3.3.2}$$

The first term on the right-hand side of (3.3.2) is identically zero because of integration by parts and the rewriting

$$\begin{aligned}
&\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,\nabla v^{\epsilon,\eta}\,\epsilon^{-1/2}\big(\nu^\eta(\rho^{\epsilon,\eta})-\nu^\eta(\bar\rho)\big)\,dx\\
&\quad=\int_{\mathbb{T}^d} h(h-1)\nabla\cdot\left(\int_0^{\rho^{\epsilon,\eta}}\left|\frac{z-\bar\rho}{\sqrt{\epsilon}}\right|^{h-2}\epsilon^{-1/2}\big(\nu^\eta(z)-\nu^\eta(\bar\rho)\big)\,dz\right)dx=0.
\end{aligned}\tag{3.3.3}$$

Similarly, the third term on the right-hand side of (3.3.2) vanishes because of the assumption $\nabla\cdot F_2^\epsilon=0$ and the integration by parts

$$\int_{\mathbb{T}^d}|v^{\epsilon,\eta}(x,s)|^{h-2}\,\sigma^\eta(\rho^{\epsilon,\eta})\nabla\sigma^\eta(\rho^{\epsilon,\eta})\cdot F_2^\epsilon\,dx=-\int_{\mathbb{T}^d}\left(\int_0^{\rho^{\epsilon,\eta}}\left|\frac{z-\bar\rho}{\sqrt{\epsilon}}\right|^{h-2}\sigma^\eta(z)\,\dot\sigma^\eta(z)\,dz\right)\nabla\cdot F_2^\epsilon\,dx.\tag{3.3.4}$$

Furthermore, the second term on the right-hand side of (3.3.2) is a true martingale vanishing at time zero, thanks to Remark 3.2.20(ii), Assumption 3.2.2 and Assumption (i)-(ii) in Theorem 3.2.18, and the regularity of $\rho^{\epsilon,\eta}$. Therefore, applying the expectation to (3.3.2) and further integrating over $t\in[0,T]$ yields:

$$\begin{aligned}
&\mathbb{E}\left[\|v^{\epsilon,\eta}\|^h_{L^h([0,T];\mathbb{T}^d)}\right]+\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\,|\nabla v^{\epsilon,\eta}|^2\,\dot\phi^\eta(\rho^{\epsilon,\eta})\,(T-s)\,dx\,ds\right]\\
&\qquad=\frac{1}{2}\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,t)|^{h-2}\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\,(T-s)\,dx\,ds\right].
\end{aligned}\tag{3.3.5}$$

We now estimate the right-hand side of (3.3.5) as follows. Assumption 3.2.9(i) on σ and Lemma 3.3.1 ensure that

$$|\sigma^\eta(z)| \leqslant C\left(1+|z|^{k+1}\right), \tag{3.3.6}$$

for a constant $C = C(\sigma)$ independent of $\eta \in (0,1)$. Thus, applying Hölder's inequality, Young's inequality and formula (3.3.6), noticing that $h(k+1) \leqslant p$, yields:

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d} h(h-1)\,|v^{\epsilon,\eta}(x,s)|^{h-2}\,(\sigma^\eta(\rho^{\epsilon,\eta}))^2\,F_3^\epsilon(T-s)\,dx\,ds\right]\\
&\leqslant C\,\|F_3^\epsilon\|_\infty\,\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}|v^{\epsilon,\eta}|^h\,dx\,ds\right]^{\frac{h-2}{h}}\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}(\sigma^\eta(\rho^{\epsilon,\eta}))^h\,dx\,ds\right]^{\frac{2}{h}}\\
&\leqslant \frac{1}{2}\mathbb{E}\left[\|v^{\epsilon,\eta}\|^h_{L^h([0,T];\mathbb{T}^d)}\right]+C\,\|F_3^\epsilon\|_\infty^{h/2}\left(1+\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}(\rho^{\epsilon,\eta})^p dx\,ds\right]\right),
\end{aligned} \tag{3.3.7}$$

for a constant $C = C(\sigma,p,T)$ independent of $\eta \in (0,1)$. Going back to (3.3.5), we use (3.3.7) and we absorb the first term on the right hand side of (3.3.7) into the corresponding term on the left hand side of (3.3.5), thanks to the factor $\frac{1}{2}$ in front, to obtain

$$\mathbb{E}\left[\|v^{\epsilon,\eta}\|^h_{L^h([0,T];\mathbb{T}^d)}\right] \leqslant C\,\|F_3^\epsilon\|_\infty^{h/2}\left(1+\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}(\rho^{\epsilon,\eta})^p dx\,ds\right]\right), \tag{3.3.8}$$

for a constant $C = C(\sigma,p,T)$ independent of η.

Now we recall that Lemma 3.3.1 ensures that $\phi^\eta,\nu^\eta,\sigma^\eta$ satisfy Assumption 3.2.6 and 3.2.9 uniformly in $\eta \in (0,1)$, since these are satisfied by ϕ,ν,σ. Therefore estimate (3.2.11) holds for a constant $C = C(\phi,\nu,\sigma,p,T)$ independent of $\eta \in (0,1)$. Using estimate (3.2.11) with $n = p$ in (3.3.8), we obtain

$$\mathbb{E}\left[\|v^{\epsilon,\eta}\|^h_{L^h([0,T];\mathbb{T}^d)}\right] \leqslant C\,\|F_3^\epsilon\|_\infty^{h/2}\,(1+\epsilon\,\|F_3^\epsilon\|_\infty)^{\frac{d}{2}(p+m)}\left(1+\mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)}+\|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right), \tag{3.3.9}$$

for a constant $C = C(\sigma,\nu,\phi,p,T)$ independent of η. Finally, Proposition 3.2.23 ensures that

$$\lim_{\eta\to 0}\|\rho^{\epsilon,\eta}-\rho^\epsilon\|_{L^1([0,T];L^1(\mathbb{T}^d))} = 0 \quad \text{in probability.}$$

Therefore, upon passing to a subsequence $\eta_n \to 0$, we have $\rho^{\epsilon,\eta_n} \to \rho^\epsilon$ for a.e. $(x,t,\omega) \in \mathbb{T}^d \times [0,T] \times \Omega$. Letting $\eta_n \to 0$ in (3.3.9) and applying Fatou's Lemma eventually yield formula (3.3.1). □

We are now ready to estimate the difference between the actual fluctuations v^ϵ and their asymptotic description v, and obtain the CLT in $L^2(\Omega)$.

Theorem 3.3.3 (Central limit theorem in $L^2(\Omega)$)**.** Let ρ_0 satisfy Assumption 3.2.1(ii) and let ϕ,ν,σ satisfy Assumption 3.2.6 and 3.2.9, for some $p \geqslant 2$ and $m \geqslant 1$. Let $\xi = \lim_{\epsilon\to 0}\xi^\epsilon$, where $(\xi^\epsilon)_{\epsilon>0}$ satisfy Assumption 3.2.4, and let v be the corresponding solution of the Langevin equation (3.1.3). For any $\epsilon > 0$, let ρ^ϵ be the stochastic kinetic solution to the generalized

Dean–Kawasaki equation (3.2.2) with initial data ρ_0, let $\bar{\rho} \equiv \rho_0$ be the solution of the zero noise limit(3.1.2) and let $v^\epsilon = \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho})$. Then, for any $T \in [0, \infty)$, for $\tau = 2$ or $\tau = \infty$, for any $\beta > \frac{d}{2}$ or $\beta > \frac{d}{2} + 1$ respectively, we have

$$\begin{aligned}
&\mathbb{E}\left[\|v^\epsilon(t) - v(t)\|^2_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))}\right] \\
&\leqslant C\Big(\epsilon\,(|F_1^\epsilon|_\infty + |F_3^\epsilon|_\infty)^2 + \epsilon^{1/2}|F_3^\epsilon|_\infty^{1/2}\Big)(1 + \epsilon\,|F_3^\epsilon|_\infty)^{g+\frac{d}{2}(p+m)}\Big(1 + \mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\Big) \\
&\quad + C\sum_{n\in\mathbb{Z}^d} n^{(2-\frac{4}{\tau})-2\beta}\sum_k \int_0^T \left|\int_{\mathbb{T}^d} e^{i2\pi nx}\,(f_k^\epsilon - f_k)\,dx\right|^2 ds,
\end{aligned} \tag{3.3.10}$$

for a constant $C = C(T, \bar{\rho}, \phi, \nu, \sigma, p, \beta)$ depending on ϕ, ν and σ only through the constants c featuring in Assumption 3.2.6 and 3.2.9.

In particular, along a scaling regime where $\epsilon \to 0$ and $\xi^\epsilon \to \xi$ such that

$$\lim_{\epsilon\to 0} \sqrt{\epsilon}\left(\|F_1^\epsilon\|_\infty + \|F_2^\epsilon\|_\infty + \|F_3^\epsilon\|_\infty\right) = 0, \tag{3.3.11}$$

the nonequilibrium fluctuations v^ϵ converge to v in $L^2(\Omega; L^\tau_{\mathrm{loc}}([0,\infty); H^{-\beta}(\mathbb{T}^d)))$ with rate of convergence (3.3.10).

Proof. Consider the sequence of approximations $(\phi^\eta, \nu^\eta, \sigma^\eta)_{\eta\in(0,1)}$ given in Lemma 3.3.1, let $\rho^{\epsilon,\eta}$ be the solution of (3.2.12), let $\bar{\rho} \equiv \rho_0$ be the solution of (3.1.2) and let $v^{\epsilon,\eta} = \epsilon^{-1/2}(\rho^{\epsilon,\eta} - \bar{\rho})$ be the corresponding fluctuations. Finally let v^η be the solution of

$$\mathrm{d}v^\eta = \Delta\big(\dot{\phi}^\eta(\bar{\rho})\,v^\eta\big)\,\mathrm{d}t - \nabla\cdot(\dot{\nu}(\bar{\rho})\,v^\eta)\;\mathrm{d}t - \nabla\cdot(\sigma^\eta(\bar{\rho})\,\mathrm{d}\xi), \quad v^\eta(\cdot,0) = 0,$$

as given in Theorem 3.2.16. The difference $v^{\epsilon,\eta} - v^\eta$ is a weak solution of

$$\begin{cases}
\mathrm{d}(v^{\epsilon,\eta} - v^\eta) = \Delta\left(\epsilon^{-1/2}\big(\phi^\eta(\rho^{\epsilon,\eta}) - \phi^\eta(\bar{\rho})\big) - \dot{\phi}^\eta(\bar{\rho})v^\eta\right)\mathrm{d}t, \\
\qquad - \nabla\cdot\left(\epsilon^{-1/2}\big(\nu^\eta(\rho^{\epsilon,\eta}) - \nu^\eta(\bar{\rho})\big) - \dot{\nu}^\eta(\bar{\rho})v^\eta\right)\mathrm{d}t - \nabla\cdot\big(\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon - \sigma^\eta(\bar{\rho})\,\mathrm{d}\xi\big) \\
\qquad + \dfrac{\epsilon^{1/2}}{2}\nabla\cdot\left(F_1^\epsilon\,(\dot{\sigma}^\eta(\rho^{\epsilon,\eta}))^2\,\nabla\rho^{\epsilon,\eta} + \dot{\sigma}^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon\right)\mathrm{d}t, \\
v^{\epsilon,\eta}(\cdot,0) = 0.
\end{cases}$$

Using the fundamental theorem of calculus, this is rewritten as

$$\begin{aligned}
\mathrm{d}(v^{\epsilon,\eta} - v^\eta) =& \Delta\left(\dot{\phi}^\eta(\bar{\rho})(v^{\epsilon,\eta} - v^\eta) + \int_0^1\int_0^1 \ddot{\phi}^\eta(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho}))\,\lambda\,d\lambda d\mu\,(v^{\epsilon,\eta})^2\,\epsilon^{1/2}\right)\mathrm{d}t \\
&- \nabla\cdot\left(\dot{\nu}^\eta(\bar{\rho})(v^{\epsilon,\eta} - v^\eta) + \int_0^1\int_0^1 \ddot{\nu}^\eta(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho}))\,\lambda\,d\lambda d\mu\,(v^{\epsilon,\eta})^2\,\epsilon^{1/2}\right)\mathrm{d}t \\
&+ \frac{\epsilon^{1/2}}{2}\nabla\cdot\left(F_1^\epsilon\,\nabla\int_0^{\rho^{\epsilon,\eta}} (\dot{\sigma}^\eta(z))^2\,dz + \dot{\sigma}^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})F_2^\epsilon\right)\mathrm{d}t \\
&- \nabla\cdot\big((\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho}))\,\mathrm{d}\xi^\epsilon + \sigma^\eta(\bar{\rho})\,\mathrm{d}(\xi^\epsilon - \xi)\big).
\end{aligned} \tag{3.3.12}$$

Using (3.3.12), recalling that $\nabla F_1^\epsilon = 2F_2^\epsilon$ and integrating by parts several times, we compute the Fourier coefficients of $v^{\epsilon,\eta} - v^\eta$. Namely, for each $n \in \mathbb{Z}^d$, for $e_n(x) := e^{i2\pi n\cdot x}$

and $\hat{v}_n^{\epsilon,\eta} - \hat{v}_n^{\eta} := \langle v^{\epsilon,\eta} - v^{\eta}, e_n \rangle$,

$$\begin{aligned}
\mathrm{d}(\hat{v}_n^{\epsilon,\eta} - \hat{v}_n^{\eta}) = &\int_{\mathbb{T}^d} \Delta e_n(x) \left(\dot{\phi}^{\eta}(\bar{\rho})(v^{\epsilon,\eta} - v^{\eta}) + \int_0^1 \int_0^1 \ddot{\phi}^{\eta}(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho}))\, \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta})^2\, \epsilon^{1/2} \right) dx\ \mathrm{d}t \\
&+ \int_{\mathbb{T}^d} \nabla e_n(x) \left(\dot{\nu}^{\eta}(\bar{\rho})(v^{\epsilon,\eta} - v^{\eta}) + \int_0^1 \int_0^1 \ddot{\nu}^{\eta}(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho}))\, \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta})^2\, \epsilon^{1/2} \right) dx\ \mathrm{d}t \\
&+ \frac{\epsilon^{1/2}}{2} \int_{\mathbb{T}^d} (\Delta e_n F_1^{\epsilon} + \nabla e_n 2F_2^{\epsilon}) \int_0^{\rho^{\epsilon,\eta}} (\dot{\sigma}^{\eta}(z))^2\, dz\, - \nabla e_n F_2^{\epsilon} \dot{\sigma}^{\eta}(\rho^{\epsilon,\eta}) \sigma^{\eta}(\rho^{\epsilon,\eta})\, dx\ \mathrm{d}t \\
&+ \int_{\mathbb{T}^d} \nabla e_n(x) \big((\sigma^{\eta}(\rho^{\epsilon,\eta}) - \sigma^{\eta}(\bar{\rho}))\, \mathrm{d}\xi^{\epsilon} + \sigma^{\eta}(\bar{\rho})\, \mathrm{d}(\xi^{\epsilon} - \xi) \big).
\end{aligned}$$

This is an SDE for $\hat{v}_n^{\epsilon,\eta} - \hat{v}_n^{\eta}$, which is readily solved by variation of constants:

$$\begin{aligned}
&\hat{v}_n^{\epsilon,\eta}(t) - \hat{v}_n^{\eta}(t) \\
&= \epsilon^{1/2} \int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta}(\bar{\rho}) + i2\pi n \dot{\nu}^{\eta}(\bar{\rho}))(t-s)} \\
&\quad \cdot \left(\int_{\mathbb{T}^d} \Delta e_n \int_0^1 \int_0^1 \big(\ddot{\phi}^{\eta}(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho})) + \ddot{\nu}^{\eta}(\bar{\rho} + \lambda\mu(\rho^{\epsilon,\eta} - \bar{\rho})) \big) \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta}(x,s))^2 dx \right) \mathrm{d}s \\
&+ \frac{\epsilon^{1/2}}{2} \int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta}(\bar{\rho}) + i2\pi n \dot{\nu}^{\eta}(\bar{\rho}))(t-s)} \qquad (3.3.13) \\
&\quad \cdot \left(\int_{\mathbb{T}^d} (\Delta e_n F_1^{\epsilon} + \nabla e_n 2F_2^{\epsilon}) \int_0^{\rho^{\epsilon,\eta}} (\dot{\sigma}^{\eta}(z))^2\, dz\, - \nabla e_n F_2^{\epsilon} \dot{\sigma}^{\eta}(\rho^{\epsilon,\eta}) \sigma^{\eta}(\rho^{\epsilon,\eta})\, dx \right) \mathrm{d}s \\
&+ \int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta}(\bar{\rho}) + i2\pi n \dot{\nu}^{\eta}(\bar{\rho}))(t-s)} \int_{\mathbb{T}^d} \nabla e_n(x) \big((\sigma^{\eta}(\rho^{\epsilon,\eta}) - \sigma^{\eta}(\bar{\rho}))\, \mathrm{d}\xi^{\epsilon} + \sigma^{\eta}(\bar{\rho})\, \mathrm{d}(\xi^{\epsilon} - \xi) \big).
\end{aligned}$$

We now estimate each term on the right hand side of (3.3.13) separately. We first consider the case $\tau = 2$. For the first term we compute, for a constant $C = C(T, p, \phi, \sigma, \nu, \bar{\rho})$,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \left| \int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta} + i2\pi n \dot{\nu}^{\eta})(t-s)}\, \epsilon^{1/2} \left(\int_{\mathbb{T}^d} \Delta e_n \int_0^1 \int_0^1 (\ddot{\phi}^{\eta} + \ddot{\nu}^{\eta}) \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta})^2 dx \right) \mathrm{d}s \right|^2 dt \right] \\
&\leqslant \epsilon\, \mathbb{E}\left[\int_0^T \left(\int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta}(\bar{\rho}) + i2\pi n \dot{\nu}^{\eta}(\bar{\rho}))(t-s)} ds \right) \right. \\
&\qquad \left. \cdot \left(\int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^{\eta} + i2\pi n \dot{\nu}^{\eta})(t-s)} \left(\int_{\mathbb{T}^d} \Delta e_n \int_0^1 \int_0^1 (\ddot{\phi}^{\eta} + \ddot{\nu}^{\eta}) \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta})^2 dx \right)^2 ds \right) dt \right] \\
&\leqslant C\, \frac{\epsilon}{n^2} \mathbb{E}\left[\int_0^T \left(\int_s^T e^{(-4\pi^2 n^2 \dot{\phi}^{\eta}(\bar{\rho}) + i2\pi n \dot{\nu}^{\eta}(\bar{\rho}))(t-s)} dt \right) \right. \qquad (3.3.14) \\
&\qquad \left. \cdot \left(\int_{\mathbb{T}^d} \Delta e_n \int_0^1 \int_0^1 (\ddot{\phi}^{\eta} + \ddot{\nu}^{\eta}) \lambda\, d\lambda d\mu\, (v^{\epsilon,\eta}(x,s))^2 dx \right)^2 \mathrm{d}s \right] \\
&\leqslant C\, \frac{\epsilon}{n^4} \mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} |\Delta e_n|^2 \int_0^1 \int_0^1 (\ddot{\phi}^{\eta} + \ddot{\nu}^{\eta})^2 \lambda^2\, d\lambda\, d\mu\, (v^{\epsilon,\eta})^4\, \mathrm{d}x\, \mathrm{d}s \right] \\
&\leqslant C\, \epsilon\, \mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} \left(1 + |\bar{\rho}| + \epsilon^{\frac{1}{2}} |v^{\epsilon,\eta}| \right)^{2g} (v^{\epsilon,\eta})^4\, \mathrm{d}x\, \mathrm{d}s \right] \\
&\leqslant C\, \epsilon\, |F_3^{\epsilon}|_{\infty}^2\, (1 + \epsilon |F_3^{\epsilon}|_{\infty})^{g + \frac{d}{2}(p+m)} \left(1 + \mathbb{E}\left[\|\rho_0\|_{L^1(\mathbb{T}^d)}^{m+p-1} + \|\rho_0\|_{L^p(\mathbb{T}^d)}^p \right] \right).
\end{aligned}$$

In the first passage we used Hölder's inequality and in the second passage we estimated the first time integral and used Fubini's theorem to swap remaining time integrals. In the third passage we estimated again the time integral of the exponential and then applied Hölder's inequality several times. The fourth passage follows from $\nabla e_n = i2\pi n e_n$ and Assumption 3.2.9(ii). The last passage follows from Hölder's inequality, Proposition 3.3.2 and $2g+4 \leqslant \frac{p}{k+1}$ in Assumption 3.2.9(ii).

For the second term on the right hand side of (3.3.13) we compute, for a constant $C = C(T, p, \phi, \sigma, \nu, \bar{\rho})$,

$$
\begin{aligned}
&\mathbb{E}\Bigg[\int_0^T \Bigg(\int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^\eta(\bar{\rho}) + i2\pi n \dot{\nu}^\eta(\bar{\rho}))(t-s)} \\
&\qquad \cdot \frac{\epsilon^{1/2}}{2}\left(\int_{\mathbb{T}^d} (\Delta e_n F_1^\epsilon + \nabla e_n 2F_2^\epsilon)\int_0^{\rho^{\epsilon,\eta}} (\dot{\sigma}^\eta(z))^2\, dz - \nabla e_n F_2^\epsilon \dot{\sigma}^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})\, dx\right) \mathrm{d}s\Bigg)^2 dt\Bigg] \\
&\leqslant C\,\frac{\epsilon}{n^4}\mathbb{E}\Bigg[\int_0^T \int_{\mathbb{T}^d} (|\Delta e_n| + |\nabla e_n|)^2 \left(|F_1^\epsilon|_\infty + |F_2^\epsilon|_\infty\right)^2 \\
&\qquad \cdot \left(\int_0^{\rho^{\epsilon,\eta}} (\dot{\sigma}^\eta(z))^2\, dz + |\dot{\sigma}^\eta(\rho^{\epsilon,\eta})\sigma^\eta(\rho^{\epsilon,\eta})|\right)^2 dx\, ds\Bigg] \\
&\leqslant C\,\epsilon\, \left(|F_1^\epsilon|_\infty + |F_2^\epsilon|_\infty\right)^2 \mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} 1 + |\rho^{\epsilon,\eta}|^{2-4\theta} + |\rho^{\epsilon,\eta}|^{4k+2} dx\, ds\right] \\
&\leqslant C\,\epsilon\, \left(|F_1^\epsilon|_\infty + |F_2^\epsilon|_\infty\right)^2 \left(1 + \epsilon|F_3^\epsilon|_\infty\right)^{\frac{d}{2}(p+m)} \left(1 + \mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right).
\end{aligned}
\tag{3.3.15}
$$

In the first passage we used Hölder's inequality and Fubini's theorem for the time integrals, and then estimated the integrals involving the exponential. The second passage follows from $\nabla e_n = i2\pi n e_n$ and Assumption 3.2.9(i) on σ. The last passage follows from Hölder's inequality, estimate (3.2.11), and $\theta \in [0, 1/2)$ and $4k+4 \leqslant p$ in Assumption 3.2.9(i).

We finally consider the last term on the right hand side of (3.3.13). We first use Itô isometry, then Fubini's theorem and finally estimate the time integral of the exponential and use Assumption 3.2.4 to obtain, for a constant $C = C(T, \phi, \bar{\rho})$,

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^T \left(\int_0^t e^{(-4\pi^2 n^2 \dot{\phi}^\eta + i2\pi n \dot{\nu}^\eta (t-s)} \int_{\mathbb{T}^d} \nabla e_n(x)\left((\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho}))\, \mathrm{d}\xi^\epsilon + \sigma^\eta(\bar{\rho})\, \mathrm{d}(\xi^\epsilon - \xi)\right)\right)^2 dt\right] \\
&= \mathbb{E}\Bigg[\int_0^T \int_0^t e^{(-8\pi^2 n^2 \dot{\phi}^\eta(\bar{\rho}) + i4\pi n \dot{\nu}^\eta(\bar{\rho}))(t-s)} \\
&\qquad \cdot \sum_m \left(\left|\int_{\mathbb{T}^d} \nabla e_n (\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho})) f_m^\epsilon\, dx\right|^2 + \left|\int_{\mathbb{T}^d} \nabla e_n \sigma^\eta(\bar{\rho})(f_m^\epsilon - f_m)\, dx\right|^2\right) ds\, dt\Bigg] \\
&= \mathbb{E}\Bigg[\int_0^T \left(\int_s^T e^{(-8\pi^2 n^2 \dot{\phi}^\eta(\bar{\rho}) + i4\pi n \dot{\nu}^\eta(\bar{\rho}))(t-s)}\; 4\pi^2 n^2\; dt\right) \\
&\qquad \cdot \sum_m \left(\left|\int_{\mathbb{T}^d} e_n (\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho})) f_m^\epsilon\, dx\right|^2 + \left|\int_{\mathbb{T}^d} e_n \sigma^\eta(\bar{\rho})(f_m^\epsilon - f_m)\, dx\right|^2\right) ds\Bigg] \\
&\leqslant C\,\mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} |\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho})|^2 dx + \sum_m \left|\int_{\mathbb{T}^d} e_n \sigma^\eta(\bar{\rho})(f_m^\epsilon - f_m)\, dx\right|^2 ds\right].
\end{aligned}
\tag{3.3.16}
$$

In turn, for the first term on the right hand side of (3.3.16), for a constant $C = C(T, p, \phi, \sigma, \nu, \bar{\rho})$,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} |\sigma^\eta(\rho^{\epsilon,\eta}) - \sigma^\eta(\bar{\rho})|^2 dx ds\right] \\
&\leqslant C\mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} |\rho^{\epsilon,\eta} - \bar{\rho}|^2 dx\, ds\right]^{1/2} \mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} \left(\int_{\bar{\rho}}^{\rho^{\epsilon,\eta}} (\dot{\sigma}^\eta(z))^2 dz\right)^2 dxds\right]^{1/2} \\
&\leqslant C\,\epsilon^{1/2}\mathbb{E}\left[\|v^{\epsilon,\eta}\|^2_{L^2_{t,x}}\right]^{1/2} \left(1 + \mathbb{E}\left[\int_0^T \int_{\mathbb{T}^d} |\rho^{\epsilon,\eta}|^{(2k+1)2} dx\, ds\right]\right)^{1/2} \\
&\leqslant C\,\epsilon^{1/2}|F_3^\epsilon|_\infty^{1/2} \left(1 + \epsilon|F_3^\epsilon|_\infty\right)^{\frac{d}{2}(p+m)} \left(1 + \mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right).
\end{aligned} \tag{3.3.17}$$

The first passage follows from Hölder's inequality and the fundamental theorem of calculus. In the second passage we used the definition of $v^{\epsilon,\eta}$, Assumption 3.2.9(i) on $\dot{\sigma}$ and Young's inequality. In the last passage we used Proposition 3.3.2, Hölder's inequality with $2(2k+1) \leqslant p$ from Assumption 3.2.9 and the L^p-estimate (3.2.11).

In conclusion, combining (3.3.13) with the estimates (3.3.14), (3.3.15), (3.3.16) and (3.3.17), we obtain, for a constant $C = C(T, \bar{\rho}, p, \phi, \sigma, \nu)$,

$$\begin{aligned}
&\mathbb{E}\left[\|v_n^{\epsilon,\eta}(t) - v_n^\eta(t)\|^2_{L^2([0,T];H^{-\beta}(\mathbb{T}^d))}\right] \\
&= \mathbb{E}\left[\int_0^T \sum_{n\in\mathbb{Z}^d} |\hat{v}_n^{\epsilon,\eta}(t) - \hat{v}_n^\eta(t)|\, n^{-2\beta} dt\right] \\
&\leqslant\ C \sum_{n\in\mathbb{Z}^d} n^{-2\beta} \sum_m \int_0^T \left|\int_{\mathbb{T}^d} e_n\left(f_m^\epsilon - f_m\right) dx\right|^2 ds \\
&\quad + C\left(\sum_{n\in\mathbb{Z}^d} n^{-2\beta}\right)\left(\epsilon\left(|F_1^\epsilon|_\infty + |F_3^\epsilon|_\infty\right)^2 + \epsilon^{1/2}|F_3^\epsilon|_\infty^{1/2}\right) \\
&\qquad \cdot \left(1 + \epsilon\,|F_3^\epsilon|_\infty\right)^{g+\frac{d}{2}(p+m)} \left(1 + \mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right).
\end{aligned} \tag{3.3.18}$$

We now let $\eta \to 0$, keeping $\epsilon \in (0,1)$ fixed, and use Fatou's Lemma in (3.3.18) to obtain formula (3.3.10) in the case $\tau = 2$.

The estimate for $\tau = \infty$ is obtained similarly, estimating $\mathbb{E}\left[\sup_{t\in[0,T]} |\hat{v}_n^{\epsilon,\eta}(t) - \hat{v}_n^\eta(t)|\, n^{-2\beta}\right]$ for each Fourier mode via the expression (3.3.13) and using computations completely analogous to (3.3.14)-(3.3.16), where we simply replace $\int_0^T dt$ by $\sup_{t\in[0,T]}$. The only differences are that in the first passage of (3.3.16) we use Doob's maximal inequality for stochastic convolutions (cf. [DPZ92]) instead of the Itô isometry, and that in the respective computations (3.3.14)-(3.3.16) we pick up a factor n^2 on the right hand side of each estimate since this is no more compensated by the time integral of the exponential. This forces us to require $\beta > \frac{d}{2} + 1$ in this case. Then we conclude identically to (3.3.17)-(3.3.18).

Finally we argue that $v^\epsilon \to v$ along the scaling regime (3.3.11). Ideed, since $\beta > \frac{d}{2}$, or $\beta > \frac{d}{2} + 1$ respectively, the first term on the right hand side of (3.3.10) vanishes along the prescribed scaling regime. The second term vanishes thanks to Assumption 3.2.4 on the noise sequence and dominated convergence. □

3.3.2 The central limit theorem in probability for rougher coefficients

In this subsection we extend the CLT to rougher coefficients satisfying Assumption 3.2.6 and Assumption 3.2.8 only.

In the following, for $\eta \in (0,1)$, we consider smoothed coefficients $\phi^\eta, \nu^\eta, \sigma^\eta$ satisfying further assumptions, but obtained by smoothing the original coefficients ϕ, ν, σ *only near zero*. Namely, a standard smoothing procedure yields the following lemma.

Lemma 3.3.4. Consider coefficients ϕ, ν, σ satisfying Assumption 3.2.6 and 3.2.8. There exists a sequence of approximating coefficients $\{\phi^\eta, \nu^\eta, \sigma^\eta\}_{\eta \in (0,1)}$ that satisfies Assumption 3.2.6 and 3.2.8 uniformly in $\eta \in (0,1)$, and Assumption 3.2.9 and Assumption (i)-(ii) in Theorem 3.2.18, not necessarily uniformly in η, and such that, defining

$$\delta_\eta := \inf\{\delta \geqslant 0 \mid \phi^\eta(z) = \phi(z),\, \nu^\eta(z) = \nu(z),\, \sigma^\eta(z) = \sigma(z)\, \forall z \in [\delta, \infty)\}, \tag{3.3.19}$$

we have that δ_η decreases to zero as η decreases to zero. That is $\lim_{\eta\to 0} \delta_\eta = 0$.

Given the stochastic kinetic solution $\rho^{\epsilon,\eta}$ to equation (3.2.12) with these smoothed coefficients ϕ^η, ν^η, σ^η, we denote

$$\Omega^{\epsilon,\eta} := \Big\{\omega \in \Omega \,\Big|\, \operatorname*{ess-inf}_{\mathbb{T}^d \times [0,T]} \rho^{\epsilon,\eta} > \delta_\eta\Big\}.$$

Furthermore, given any initial data ρ_0 satisfying Assumption 3.2.1(ii), we denote

$$\ell := \operatorname*{ess-inf}_{\Omega \times \mathbb{T}^d} \rho_0.$$

The following two results establish stronger uniqueness both for the zero-noise limit and for the generalized Dean–Kawasaki equation. The first is an immediate application of the maximum principle and the uniqueness for the deterministic equation (3.1.2).

Lemma 3.3.5. Let $\bar{\rho}$ and $\bar{\rho}^\eta$ be the solutions of equation (3.1.2) and of its smoothed version $\partial_t \bar{\rho}^\eta = \Delta \phi^\eta(\bar{\rho}^\eta) - \nabla \cdot \nu^\eta(\bar{\rho}^\eta)$, both with initial data ρ_0 satisfying Assumption 3.2.1(ii), i.e. a random positive constant. Let $\eta \in (0,1)$ be small enough so that $\delta_\eta < \ell$, that is so that the smoothed coefficients match the true coefficients on $[\ell, \infty)$. Then we have $\bar{\rho} = \bar{\rho}^\eta$ for a.e. $(\omega, x, t) \in \Omega \times \mathbb{T}^d \times [0, \infty)$.

A straightforward adaptation of the uniqueness proof in [FG21b, Theorem 4.7], which just amounts to restricting all the arguments to a smaller probability subset $\Omega_0 \subseteq \Omega$, establishes the following.

Lemma 3.3.6 (Enhanced pathwise uniqueness)**.** For $i = 1, 2$, let ρ^i be the stochastic kinetic solution to the equation

$$\partial_t \rho^i = \Delta \phi_i(\rho^i) - \nabla \cdot \nu_i(\rho^i) - \sqrt{\epsilon} \nabla \cdot \left(\sigma_i(\rho^i) \circ \dot{\xi}^\epsilon\right) \text{ in } \mathbb{T}^d \times (0,T), \quad \rho^i(\cdot, 0) = \rho_0^i \text{ in } \mathbb{T}^d \times \{0\},$$

with noise ξ^ϵ satisfying Assumption 3.2.2, coefficients ϕ_i, ν_i and σ_i satisfying Assumption 3.2.6, and initial data ρ_0^i satisfying Assumption 3.2.1. Let $\Omega_0 \subseteq \Omega$ be a measurable subset such that

i) $\rho_0^1(x,\omega) = \rho_0^2(x,\omega)$ for every $x \in \mathbb{T}^d$, for a.e. $\omega \in \Omega_0$;

ii) for some $\delta > 0$, for a.e. $\omega \in \Omega_0$,

$$\phi_1(z) = \phi_2(z),\ \nu_1(z) = \nu_2(z),\ \sigma_1(z) = \sigma_2(z)\ \ \forall z \in \left(\operatorname*{ess\,inf}_{\mathbb{T}^d \times [0,T] \times \Omega_0} \rho^1 - \delta,\ \operatorname*{ess\text{-}sup}_{\mathbb{T}^d \times [0,T] \times \Omega_0} \rho^1 + \delta \right).$$

Then we have that

$$\sup_{t \in [0,T]} \|\rho^1(\cdot,t) - \rho^2(\cdot,t)\|_{L^1(\mathbb{T}^d)} \leqslant \|\rho_0^1 - \rho_0^2\|_{L^1(\mathbb{T}^d)} \quad \text{almost surely on } \Omega_0.$$

The two previous Lemmas, the convergence (3.1.4) from [FG23, Theorem 6.6] and Theorem 3.3.3 from the previous section immediately yield the following.

Corollary 3.3.7. Let ρ_0, $(\xi^\epsilon)_{\epsilon>0}$ and ϕ, ν, σ satisfy Assumption 3.2.1(ii), Assumption 3.2.4, and Assumption 3.2.6 and 3.2.8 respectively. Consider the smoothed coefficients from Lemma 3.3.4 and let $\eta \in (0,1)$ be small enough so that $\delta_\eta < \ell$. Let $\rho^{\epsilon,\eta}$ be the stochastic kinetic solution to the smoothed equation (3.2.12). We have, keeping η fixed,

$$\lim_{\epsilon \to 0} \|\rho^{\epsilon,\eta} - \bar{\rho}\|_{L^1([0,T];L^1(\mathbb{T}^d))} = 0 \text{ in probability.}$$

Furthermore, keeping η fixed and letting $\epsilon \to 0$ along the scaling regime (3.3.11), we have

$$v^{\epsilon,\eta} = \epsilon^{-1/2}(\rho^{\epsilon,\eta} - \bar{\rho}) \to v \quad \text{in } L^2(\Omega; L^\tau_{\text{loc}}([0,\infty); H^{-\beta}(\mathbb{T}^d))),$$

for $\tau = 2$ or $\tau = \infty$, for any $\beta > \frac{d}{2}$ or $\beta > 1 + \frac{d}{2}$ respectively, with rate of convergence given in formula (3.3.10). Finally, for every $\epsilon \in (0,1)$ and every $\eta \in (0,1)$ small enough, we have

$$\rho^\epsilon = \rho^{\epsilon,\eta} \quad \text{and} \quad v^\epsilon = v^{\epsilon,\eta} \quad \text{for a.e. } (x,t) \in \mathbb{T}^d \times [0,T], \text{ for every } \omega \in \Omega^{\epsilon,\eta}.$$

We now present the main ingredient to establish the central limit theorem in probability. Corollary 3.3.7 implies that the sets $\Omega^{\epsilon,\eta}$ satisfy $\Omega^{\epsilon,\eta_1} \supseteq \Omega^{\epsilon,\eta_2}$ for every $\eta_1 < \eta_2 \in (0,1)$ and every $\epsilon \in (0,1)$. The following proposition essentially establishes that the measure of $\Omega^{\epsilon,\eta} \subseteq \Omega$ increases to 1 as $\epsilon \to 0$.

The proposition is adapted from [DFG20] to the nonlinear diffusion case and we state it on its own, since the result is interesting per se and it holds for any stochastic kinetic solution of equation (3.2.1) provided the coefficients satisfy Assumption 3.2.6 and the diffusion is nondegenerate.

Proposition 3.3.8 (Moser iteration argument)**.** Fix $\epsilon > 0$. Let ρ_0 satisfy Assumption 3.2.1, let ξ^ϵ satisfy Assumption 3.2.2 and let ϕ, ν, σ satisfy Assumption 3.2.6. Furthermore, suppose that $\inf \dot{\phi} > 0$. Let ρ^ϵ be the stochastic kinetic solution to (3.2.2) and let $\ell := \operatorname*{ess\text{-}inf}_{\Omega \times \mathbb{T}^d} \rho_0$. For every $T \geqslant 0$, we have

$$\mathbb{E}\Big[\|(\rho^\epsilon - \ell)_-\|_{L^\infty(\mathbb{T}^d \times [0,T])}\Big] \leqslant c^* \sum_{j=1}^{\infty} \frac{1}{j^2} R_\epsilon^{\,\zeta\,(1+2/d)^{-j}},$$

where c^* is a *numeric* constant, and R_ϵ and the exponent ζ are given by

$$R_\epsilon := C_0\,(\inf \dot{\phi})^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty \|F_3^\epsilon\|_\infty + \epsilon\,\|F_3^\epsilon\|_\infty\big), \quad \zeta := \begin{cases} \frac{d}{2}(1 + 2/d)^2 & \text{if } R_\epsilon \geqslant 1, \\ e^{-d(d+1)}(1/2 + 1/d) & \text{if } R_\epsilon \in (0,1), \end{cases}$$

for a constant $C_0 = C_0(T, d, \ell, \phi, \sigma, \nu)$ depending on the coefficients ϕ, ν, σ only through the constants c featuring in Assumption 3.2.6.

Proof. For $\eta \in (0,1)$ consider smoothed coefficients $\phi^\eta, \nu^\eta, \sigma^\eta$ obtained from the original coefficients ϕ, ν, σ thanks to Lemma 3.3.1. They satisfy Assumption 3.2.6 and also Assumption (i) in Theorem 3.2.18, since ϕ is nondegenerate, uniformly in η. In particular we require $\inf \dot{\phi}^\eta \geqslant \inf \dot{\phi} > 0$. Let $\rho^{\epsilon,\eta}$ be the weak solution to equation (3.2.12) with coefficients $\phi^\eta, \nu^\eta, \sigma^\eta$ and the same noise ξ^ϵ and initial data ρ_0. Finally let $\ell := \operatorname*{ess\text{-}inf}_{\Omega \times \mathbb{T}^d} \rho_0$.

For any $\alpha \in [1, \infty)$, consider the function $f(z) := (z - \ell)_-^{\alpha+1} = (0 \vee (\ell - z))^{\alpha+1}$. Applying Itô's formula in the form proved in Krylov [Kry12] yields

$$\begin{aligned}
\int_{\mathbb{T}^d} (\rho_r^{\epsilon,\eta} - \ell)_-^{\alpha+1} dx\Big|_{r=0}^{r=t} = &- \int_0^t \int_{\mathbb{T}^d} \alpha(\alpha+1)(\rho_r^{\epsilon,\eta} - \ell)_-^{\alpha-1} \dot{\phi}^\eta(\rho^{\epsilon,\eta}) |\nabla \rho_r^{\epsilon,\eta}|^2 \, dx\, dr \\
&- \int_0^t \int_{\mathbb{T}^d} (\alpha+1)(\rho_r^{\epsilon,\eta} - \ell)_-^\alpha \nabla \cdot (\nu^\eta(\rho^{\epsilon,\eta}))\, dx\, dr \\
&- \epsilon^{1/2} \int_0^t \int_{\mathbb{T}^d} (\alpha+1)(\rho_r^{\epsilon,\eta} - \ell)_-^\alpha\, \nabla \cdot (\sigma^\eta(\rho^{\epsilon,\eta})\, \mathrm{d}\xi^\epsilon) \qquad (3.3.20)\\
&+ \frac{\epsilon}{2} \int_0^t \int_{\mathbb{T}^d} \alpha(\alpha+1)(\rho_r^{\epsilon,\eta} - \ell)_-^{\alpha-1} \sigma^\eta(\rho^{\epsilon,\eta}) \nabla \sigma^\eta(\rho^{\epsilon,\eta}) F_2^\epsilon \, dx\, dr \\
&+ \frac{\epsilon}{2} \int_0^t \int_{\mathbb{T}^d} \alpha(\alpha+1)(\rho_r^{\epsilon,\eta} - \ell)_-^{\alpha-1} \left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2 F_3^\epsilon \, dx\, dr.
\end{aligned}$$

We note that that the integral at time $r = 0$ on the left hand side of (3.3.20) vanishes because $(\rho_0 - \ell)_- \equiv 0$ by definition of ℓ. Furthermore, arguing as in (3.3.3)-(3.3.4), using integration by parts and the periodic boundary conditions or the assumption $\nabla \cdot F_2^\epsilon = 0$, shows that the second and the fourth term on the right hand side of (3.3.20) are identically zero. We integrate by parts the third term on the right hand side of (3.3.20), then we use the identity

$$\nabla \left(\rho^{\epsilon,\eta}\right)^{\frac{\alpha+1}{2}} = \frac{\alpha+1}{2} \left(\rho^{\epsilon,\eta}\right)^{\frac{\alpha-1}{2}} \nabla \rho^{\epsilon,\eta}$$

to rewrite the first and the third term on the right hand side, and finally we rearrange the terms to obtain

$$\begin{aligned}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx+\int_0^t\int_{\mathbb{T}^d}\frac{4\alpha}{\alpha+1}\dot{\phi}^\eta(\rho_r^{\epsilon,\eta})\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\right|^2\,dx\,dr\\ = \epsilon^{1/2}\int_0^t\int_{\mathbb{T}^d}2\alpha\,(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha-1}{2}}\,\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\,\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon\\ +\frac{\epsilon}{2}\int_0^t\int_{\mathbb{T}^d}\alpha(\alpha+1)(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\,dx\,dr.\end{aligned} \tag{3.3.21}$$

Since $\rho^{\epsilon,\eta}\geqslant 0$ always, we have $|(\rho^{\epsilon,\eta}-\ell)_-|_\infty\leqslant\ell$. Moreover, the assumption $\sigma\in C_{\text{loc}}([0,\infty))$ implies that there exists $C=C(\ell,\sigma)$ independent of $\eta\in(0,1)$ such that

$$\left|\sigma^\eta(\rho^{\epsilon,\eta})\,\mathbf{1}_{\{(\rho^{\epsilon,\eta}-\ell)_-\neq 0\}}\right|\leqslant C. \tag{3.3.22}$$

Hence, for the second term on the right hand side of (3.3.21), we have, for a constant $C=C(\ell,\sigma)$ independent of $\eta,\epsilon\in(0,1)$ and of $\alpha\geqslant 1$, almost surely for every $t\in[0,T]$,

$$\frac{\epsilon}{2}\int_0^t\int_{\mathbb{T}^d}\alpha(\alpha+1)(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\,dx\,dr\leqslant C\,\alpha^2\,\epsilon\,\|F_3^\epsilon\|_\infty\int_0^t\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\,dx\,dr. \tag{3.3.23}$$

Given an arbitrary bounded stopping time $\tau\leqslant T$, we now evaluate (3.3.21) at $t=\tau$. After taking the expectation, so that the martingale term vanishes, using estimate (3.3.23) and multiplying both sides by $(\inf\dot{\phi})^{-1}$, we obtain, for a constant $C=C(\ell,\sigma)$ independent of $\eta,\epsilon\in(0,1)$ and $\alpha\geqslant 1$,

$$\mathbb{E}\left[\int_0^\tau\int_{\mathbb{T}^d}\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)^{\frac{\alpha+1}{2}}\right|^2\,dx\,dr\right]\leqslant C\,\alpha^2\,(\inf\dot{\phi})^{-1}\,\epsilon\,\|F_3^\epsilon\|_\infty\mathbb{E}\left[\int_0^\tau\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\,dx\,dr\right]. \tag{3.3.24}$$

For the martingale term in (3.3.21), given an arbitrary bounded stopping time $\tau\leqslant T$, we compute, for a constant $C=C(\ell,T,\sigma)$ independent of $\eta,\epsilon\in(0,1)$ and $\alpha\geqslant 1$, for any

$\delta_\alpha \in (0,\ell)$,

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,\tau]}\Big|\epsilon^{1/2}\int_0^t\int_{\mathbb{T}^d} 2\alpha\,(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha-1}{2}}\,\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\,\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon\Big|\Big]\\
&\leqslant C\epsilon^{1/2}\alpha\,\mathbb{E}\left[\left(\int_0^\tau\sum_{k=1}^\infty\Big(\int_{\mathbb{T}^d}2(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha-1}{2}}\,\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\,\sigma^\eta(\rho^{\epsilon,\eta})\,f_k(x)\,dx\Big)^2dr\right)^{1/2}\right]\\
&\leqslant C\epsilon^{1/2}\alpha\,\mathbb{E}\left[\left(\int_0^\tau\Big(\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\sigma^\eta(\rho^{\epsilon,\eta})^2F_1^\epsilon\,dx\Big)\Big(\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\Big)\,dr\right)^{1/2}\right]\\
&\leqslant C\alpha\,\epsilon^{1/2}\|F_1^\epsilon\|_\infty^{1/2}\mathbb{E}\left[\sup_{t\in[0,\tau]}\left(\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha-1}\big(\mathbf{1}_{\{\rho^{\epsilon,\eta}>\ell-\delta_\alpha\}}+\mathbf{1}_{\{\rho^{\epsilon,\eta}\leqslant\ell-\delta_\alpha\}}\big)dx\right)^{1/2}\right.\\
&\qquad\qquad\qquad\qquad\qquad\qquad\left.\cdot\left(\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\right)^{1/2}\right]\\
&\leqslant (C\alpha^2\,\epsilon\|F_1^\epsilon\|_\infty)^{1/2}\left(\delta_\alpha^{\alpha-1}+\delta_\alpha^{-2}\,\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx\Big]\right)^{1/2}\\
&\qquad\cdot\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\Big]^{1/2}\\
&\leqslant (C\alpha^2\,\epsilon\|F_1^\epsilon\|_\infty)^{1/2}\delta_\alpha^{\frac{\alpha-1}{2}}\,\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\Big]^{1/2}\\
&\quad+(C\alpha^2\,\epsilon\|F_1^\epsilon\|_\infty)^{1/2}\delta_\alpha^{-1}\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx\Big]^{1/2}\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\Big]^{1/2}\\
&\leqslant (C\alpha^2\,\epsilon\|F_1^\epsilon\|_\infty)^{1/2}\left(\frac{1}{2}\delta_\alpha^{\alpha-1}+\frac{1}{2}\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\Big]\right)\\
&\quad+\frac{1}{2}\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx\Big]\\
&\quad+\frac{1}{2}(C\alpha^2\,\epsilon\|F_1^\epsilon\|_\infty)\delta_\alpha^{-2}\,\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}\Big|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\Big|^2dx\,dr\Big]\\
&\leqslant\frac{1}{2}\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx\Big]\\
&\quad+(C\alpha^2\epsilon\|F_1^\epsilon\|_\infty)^{1/2}\,\delta_\alpha^{\alpha-1}\\
&\quad+\Big(C\alpha^4\,\delta_\alpha^{-2}(\inf\dot\phi)^{-1}\big(\epsilon^2\|F_1^\epsilon\|_\infty+\epsilon^{3/2}\|F_1^\epsilon\|_\infty^{1/2}\big)\|F_3^\epsilon\|_\infty\Big)\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}dx\,dr\Big].
\end{aligned}
\tag{3.3.25}
$$

In the first passage we used the Burkholder-Davis-Gundy inequality and the form of the noise term. The second passage follows from Hölder's inequality. In the third passage we used (3.3.22) and we took the supremum in time and then split the first space integral according to the values of $\rho^{\epsilon,\eta}$. The fourth passage follows again from Hölder's inequality and the splitting introduced. The fifth and sixth passage follows from convexity and Hölder's inequality. The last passage follows from (3.3.24).

Now, given an arbitrary bounded stopping time $\tau\leqslant T$, combining (3.3.21) with (3.3.23) and (3.3.25) yields, for a constant $C=C(\ell,T,\sigma)$ independent of $\eta,\epsilon\in(0,1)$ and $\alpha\geqslant 1$, for any $\delta_\alpha \in (0,\ell)$,

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx+\int_0^\tau\int_{\mathbb{T}^d}\frac{4\alpha}{\alpha+1}\dot\phi^\eta(\rho_r^{\epsilon,\eta})\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\right|^2\,dx\,dr\Big]\\
&\quad\leqslant\ \mathbb{E}\Big[\sup_{t\in[0,\tau]}\Big|\epsilon^{1/2}\int_0^t\int_{\mathbb{T}^d}2\alpha\,(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha-1}{2}}\,\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\,\sigma^\eta(\rho^{\epsilon,\eta})\,\mathrm{d}\xi^\epsilon\Big|\Big]\\
&\qquad+\mathbb{E}\Big[\frac{\epsilon}{2}\int_0^\tau\int_{\mathbb{T}^d}\alpha(\alpha+1)(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}\left(\sigma^\eta(\rho^{\epsilon,\eta})\right)^2F_3^\epsilon\,dx\,dr\Big]\\
&\quad\leqslant\ \frac12\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx\Big]\\
&\qquad+\Big(C\alpha^4\,\delta_\alpha^{-2}(\inf\dot\phi)^{-1}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty+\epsilon\,\|F_3^\epsilon\|_\infty\big)\Big)\\
&\qquad\quad\cdot\Big(\delta_\alpha^{\alpha-1}\ +\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}dx\,dr\Big]\Big).
\end{aligned}\tag{3.3.26}$$

In turn, absorbing the first term on the right hand side of (3.3.26) into the left hand side, thanks to the factor $\frac12$ in front, and then multiplying both sides by $\max\{\inf(\dot\phi)^{-1},1\}$ yields, for a *fixed* constant $\bar c=\bar c(T,\sigma,\ell)$ independent of $\eta,\epsilon\in(0,1)$ and $\alpha\geqslant 1$,

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,\tau]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx+\int_0^\tau\int_{\mathbb{T}^d}\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\right|^2\,dx\,dr\Big]\\
&\leqslant\Big(\bar c\,\alpha^4\,\delta_\alpha^{-2}(\inf\dot\phi)^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty+\epsilon\,\|F_3^\epsilon\|_\infty\big)\Big)\Big(\delta_\alpha^{\alpha-1}+\mathbb{E}\Big[\int_0^\tau\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}dx\,dr\Big]\Big).
\end{aligned}\tag{3.3.27}$$

In particular we notice that, upon possibly enlarging the constant $\bar c$ by a factor $(1+T)$, when $\alpha=1$ formula (3.3.27) reduces to

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{2}dx+\int_0^\tau\int_{\mathbb{T}^d}\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)\right|^2\,dx\,dr\Big]\\
&\qquad\leqslant\Big(\bar c\,\alpha^4\,\delta_\alpha^{-2}(\inf\dot\phi)^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty+\epsilon\,\|F_3^\epsilon\|_\infty\big)\Big).
\end{aligned}\tag{3.3.28}$$

Since the stopping time τ in (3.3.27) is arbitrary, it follows from [RY99, Chapter IV, Proposition 4.7 and Exercise 4.30] that, for any $\alpha\geqslant 1$, for the same constant $\bar c=\bar c(T,\sigma,\ell)$ introduced in (3.3.27) and independent of $\eta,\epsilon\in(0,1)$ and of $\alpha\geqslant 1$, for any $\delta_\alpha\in(0,1)$,

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\sup_{t\in[0,T]}\int_{\mathbb{T}^d}(\rho_t^{\epsilon,\eta}-\ell)_-^{\alpha+1}dx+\int_0^T\int_{\mathbb{T}^d}\left|\nabla(\rho_r^{\epsilon,\eta}-\ell)_-^{\frac{\alpha+1}{2}}\right|^2\,dx\,dr\Big)^{\frac{1}{\alpha+1}}\Big]\\
&\quad\leqslant\frac{(\alpha+1)^{\frac{1}{\alpha+1}}}{1-\frac{1}{\alpha+1}}\,\big(\alpha^4\delta_\alpha^{-2}\big)^{\frac{1}{\alpha+1}}\Big(\bar c\,(\inf\dot\phi)^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty+\epsilon\,\|F_3^\epsilon\|_\infty\big)\Big)^{\frac{1}{\alpha+1}}\\
&\qquad\qquad\cdot\mathbb{E}\Big[\Big(\delta_\alpha^{\alpha-1}\ +\int_0^T\int_{\mathbb{T}^d}(\rho_r^{\epsilon,\eta}-\ell)_-^{\alpha-1}dx\,dr\Big)^{\frac{1}{\alpha+1}}\Big].
\end{aligned}\tag{3.3.29}$$

We are now ready to conclude the proof using a Moser iteration argument. To ease the notation, let us define $\psi:=(\rho^{\epsilon,\eta}-\ell)_-$. Using that $(x+y)^{1/h}\leqslant x^{1/h}+y^{1/h}$, for $h\geqslant 1$ and $x,y\geqslant 0$, and Hölder's inequality with exponent $\frac{\alpha+1}{\alpha-1}$ in (3.3.29) gives, for the constant $\bar c$ introduced in (3.3.27), for any $\alpha\geqslant 1$ and any $\delta_\alpha\in(0,1)$,

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\|\psi\|^{\alpha+1}_{L_t^\infty L_x^{\alpha+1}}+\|\nabla\psi^{\frac{\alpha+1}{2}}\|^2_{L_t^2L_x^2}\Big)^{\frac{1}{\alpha+1}}\Big]\\
&\quad\leqslant\frac{(\alpha+1)^{\frac{1}{\alpha+1}}}{1-\frac{1}{\alpha+1}}\,\big(\alpha^4\delta_\alpha^{-2}\big)^{\frac{1}{\alpha+1}}\Big(\bar c\,(\inf\dot\phi)^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty+\epsilon\,\|F_3^\epsilon\|_\infty\big)\Big)^{\frac{1}{\alpha+1}}\\
&\qquad\qquad\cdot\Big(\delta_\alpha^{\frac{\alpha-1}{\alpha+1}}+\mathbb{E}\big[\|\psi\|_{L_t^{\alpha-1}L_x^{\alpha-1}}\big]^{\frac{\alpha-1}{\alpha+1}}\Big).
\end{aligned}\tag{3.3.30}$$

We first use interpolation and Sobolev inequalities to deduce that for

$$\lambda = \frac{2}{2+d} \text{ and } q = \frac{(2+d)(\alpha+1)}{d},$$

we have that, for a *fixed* constant $\tilde{c} = \tilde{c}(T,d)$ independent of $\eta, \epsilon \in (0,1)$ and $\alpha \geqslant 1$,

$$\begin{aligned} \|\psi\|_{L^q(\mathbb{T}^d\times[0,T])} &\leqslant \|\psi\|^{\lambda}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} \|\psi\|^{1-\lambda}_{L^{\alpha+1}([0,T];L^{\frac{2^*}{2}(\alpha+1)}(\mathbb{T}^d))} \\ &= \|\psi\|^{\lambda}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} \left\|\psi^{\frac{\alpha+1}{2}}\right\|^{\frac{2(1-\lambda)}{\alpha+1}}_{L^2([0,T];L^{2^*}(\mathbb{T}^d))}. \\ &\leqslant \|\psi\|^{\lambda}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} \left(\tilde{c}\left(\|\psi\|^{\frac{\alpha+1}{2}}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} + \left\|\nabla\psi^{\frac{\alpha+1}{2}}\right\|_{L^2([0,T];L^2(\mathbb{T}^d))}\right)\right)^{\frac{2(1-\lambda)}{\alpha+1}}. \end{aligned} \tag{3.3.31}$$

Now Hölder's inequality, the inequality $(x+y)^2 \leqslant 2(x^2+y^2)$ for all $x, y \in [0,\infty)$, the fact that $\lambda \in (0,1)$, $\alpha \in [1,\infty)$, and (3.3.30) prove that, for the constant $C_0 := \bar{c}\tilde{c}$, for $\bar{c}$ and $\tilde{c}$ introduced in (3.3.27) and (3.3.31), independent of $\epsilon, \eta \in (0,1)$ and $\alpha \geqslant 1$,

$$\begin{aligned} &\mathbb{E}\Big[\,\|\psi\|_{L^q(\mathbb{T}^d\times[0,T])}\Big] \\ &\leqslant \mathbb{E}\Big[\|\psi\|_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))}\Big]^{\lambda} \mathbb{E}\left[\left(\tilde{c}\,\|\psi\|^{\alpha+1}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} + \tilde{c}\left\|\nabla\psi^{\frac{\alpha+1}{2}}\right\|^2_{L^2([0,T];L^2(\mathbb{T}^d))}\right)^{\frac{1}{\alpha+1}}\right]^{1-\lambda} \\ &\leqslant \mathbb{E}\left[\left(\tilde{c}\,\|\psi\|^{\alpha+1}_{L^\infty([0,T];L^{\alpha+1}(\mathbb{T}^d))} + \tilde{c}\left\|\nabla\psi^{\frac{\alpha+1}{2}}\right\|^2_{L^2([0,T];L^2(\mathbb{T}^d))}\right)^{\frac{1}{\alpha+1}}\right] \\ &\leqslant \frac{(\alpha+1)^{\frac{1}{\alpha+1}}}{1-\frac{1}{\alpha+1}}\left(\alpha^4\delta_\alpha^{-2}\right)^{\frac{1}{\alpha+1}}\left(C_0\,(\inf\dot{\phi})^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty + \epsilon\,\|F_3^\epsilon\|_\infty\big)\right)^{\frac{1}{\alpha+1}} \\ &\qquad\cdot\left(\delta_\alpha^{\frac{\alpha-1}{\alpha+1}} + \mathbb{E}\big[\|\psi\|_{L_t^{\alpha-1}L_x^{\alpha-1}}\big]^{\frac{\alpha-1}{\alpha+1}}\right). \end{aligned} \tag{3.3.32}$$

Furthermore, thanks to (3.3.28), when $\alpha = 1$ we can replace the last factor on the right hand side of (3.3.32) simply with 1.

We now iterate inequality (3.3.32) along a suitable sequence of exponents α_k tending to infinity and small parameters $\delta_k = \delta_{\alpha_k}$ for $k \in \mathbb{N}$. Namely, mimicking the relation between the exponents q, α and $\alpha - 1$ in formula (3.3.32), which holds for any $\alpha \in [1,\infty)$, we define

$$\alpha_k = \frac{2+d}{d}(\alpha_{k-1}+2) \quad \text{and} \quad \beta_k = \alpha_{k-1}+1 \quad \forall\, k \in \mathbb{N}^*, \quad \alpha_0 = 0. \tag{3.3.33}$$

This gives

$$\alpha_k = (d+2)\big((1+{}^2\!/\!_d)^k - 1\big) \quad \text{and} \quad \beta_k = (d+2)\big((1+{}^2\!/\!_d)^{k-1} - 1\big) + 1 \quad \forall\, k \in \mathbb{N}^*. \tag{3.3.34}$$

We also introduce the shorthand

$$u_{\beta_k} := \frac{1}{\beta_k+1} \quad \forall\, k \in \mathbb{N}^*.$$

Let us now define

$$\gamma := \inf_{k\geqslant 2}\prod_{i=2}^{k}\frac{\beta_i-1}{\beta_i+1} = \lim_{k\to\infty}\prod_{i=2}^{k}\frac{\beta_i-1}{\beta_i+1} \in (0,1). \tag{3.3.35}$$

It is shown in (3.3.39) below that indeed $\gamma > 0$. In turn we define

$$\delta_k := k^{-2/\gamma} \quad \forall\, k \in \mathbb{N}^*. \tag{3.3.36}$$

Finally, let us also introduce the shorthand

$$R_\epsilon := C_0\,(\inf \dot{\phi})^{-2}\big(\epsilon^2\, \|F_1^\epsilon\|_\infty \|F_3^\epsilon\|_\infty + \epsilon\, \|F_3^\epsilon\|_\infty\big). \tag{3.3.37}$$

Iterating inequality (3.3.32) over the sequence α_k, using the inequality $(x+y)^{1/h} \leqslant x^{1/h} + y^{1/h}$ for $h \geqslant 1$ and $x, y \geqslant 0$ at each step, yields

$$\begin{aligned}
\mathbb{E}\Big[\|\psi\|_{L^{\alpha_k}(\mathbb{T}^d\times[0,T])}\Big] &\leqslant \frac{u_{\beta_k}^{-u_{\beta_k}}}{1-u_{\beta_k}} \beta_k^{4u_{\beta_k}} \delta_k^{-2u_{\beta_k}} R_\epsilon^{u_{\beta_k}} \left(\delta_k^{\frac{\beta_k-1}{\beta_k+1}} + \mathbb{E}\Big[\|\psi\|_{L^{\alpha_{k-1}}(\mathbb{T}^d\times[0,T])}\Big]^{\frac{\beta_k-1}{\beta_k+1}} \right) \\
&\leqslant \frac{u_{\beta_k}^{-u_{\beta_k}}}{1-u_{\beta_k}} \beta_k^{4u_{\beta_k}} \delta_k^{-2u_{\beta_k}} R_\epsilon^{u_{\beta_k}} \delta_k^{\frac{\beta_k-1}{\beta_k+1}} \\
&\quad + \frac{u_{\beta_k}^{-u_{\beta_k}}}{1-u_{\beta_k}} \beta_k^{4u_{\beta_k}} \delta_k^{-2u_{\beta_k}} R_\epsilon^{u_{\beta_k}} \left(\frac{u_{\beta_{k-1}}^{-u_{\beta_{k-1}}}}{1-u_{\beta_{k-1}}} \beta_{k-1}^{4u_{\beta_{k-1}}} \delta_{k-1}^{-2u_{\beta_{k-1}}} R_\epsilon^{u_{\beta_{k-1}}} \right)^{\frac{\beta_k-1}{\beta_k+1}} \\
&\qquad\qquad \cdot \left(\delta_{k-1}^{\frac{\beta_{k-1}-1}{\beta_{k-1}+1}} + \mathbb{E}\Big[\|\psi\|_{L^{\alpha_{k-2}}(\mathbb{T}^d\times[0,T])}\Big]^{\frac{\beta_{k-1}-1}{\beta_{k-1}+1}} \right)^{\frac{\beta_k-1}{\beta_k+1}} \\
&\leqslant \sum_{j=1}^{k} \delta_j^{\prod_{i=j}^k \frac{\beta_i-1}{\beta_i+1}} \times \prod_{l=j}^{k} \left(\frac{u_{\beta_l}^{-u_{\beta_l}}}{1-u_{\beta_l}}\, \beta_l^{4u_{\beta_l}}\, \delta_l^{-2u_{\beta_l}} \right)^{\prod_{i=l+1}^k \frac{\beta_i-1}{\beta_i+1}} \times R_\epsilon^{\sum_{l=j}^k \left(u_{\beta_l} \prod_{i=l+1}^k \frac{\beta_i-1}{\beta_i+1}\right)}.
\end{aligned}$$

We remark that in the last passage of the iteration, from α_1 to α_0, we do not pick up any integral of ψ since by construction $\alpha_0 = 0$ (cf. formula (3.3.28) and the comment after formula (3.3.32)). Finally, recalling (3.3.35)-(3.3.36), we estimate

$$\begin{aligned}
\mathbb{E}\Big[\|\psi\|_{L^{\alpha_k}(\mathbb{T}^d\times[0,T])}\Big] &\leqslant \left(\sum_{j=1}^{k} \delta_j^\gamma\, R_\epsilon^{\sum_{l=j}^k \left(u_{\beta_l} \prod_{i=l+1}^k \frac{\beta_i-1}{\beta_i+1}\right)} \right) \\
&\quad \times \sup_{k\in\mathbb{N}^*} \prod_{l=1}^{k} \left(\frac{u_{\beta_l}^{-u_{\beta_l}}}{1-u_{\beta_l}}\, \beta_l^{4u_{\beta_l}} \right)^{\prod_{i=l+1}^k \frac{\beta_i-1}{\beta_i+1}} \times \sup_{k\in\mathbb{N}^*} \prod_{l=1}^{k} \delta_l^{-2u_{\beta_l} \prod_{i=l+1}^k \frac{\beta_i-1}{\beta_i+1}}.
\end{aligned} \tag{3.3.38}$$

We now analyze the limit of the infinite sum and products in (3.3.38). As anticipated in (3.3.35) we have $\gamma \in (0,1)$. Indeed, passing to logarithms, using that $\log(1-z) \geqslant -(1+d)z$ for $z \leqslant \frac{d}{1+d}$ and the expression (3.3.34), we find

$$\log\left(\prod_{i=2}^{k} \frac{\beta_i - 1}{\beta_i + 1} \right) = \sum_{i=2}^{k} \log\left(1 - \frac{2}{\beta_i+1}\right) \geqslant -2(1+d) \sum_{i=1}^{\infty} \frac{1}{\beta_i+1} > -\infty. \tag{3.3.39}$$

In particular the series $\sum_{j=1}^\infty \delta_j^\gamma < \infty$ converges since $\delta_j = j^{-2/\gamma}$. Furthermore, for every $l \leqslant k \in \mathbb{N}^*$, we have

$$\gamma\, u_{\beta_l} \leqslant u_{\beta_l} \prod_{i=l+1}^{k} \frac{\beta_i-1}{\beta_i+1} \leqslant u_{\beta_l}.$$

Therefore, by (3.3.34) and (3.3.39) and dominated convergence, for every $j \in \mathbb{N}^*$ the following limit exists and satisfies the inequality

$$\gamma \sum_{l=j}^{\infty} u_{\beta_l} \leqslant \lim_{k\to\infty} \sum_{l=j}^{k} \left(u_{\beta_l} \prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1} \right) = \sum_{l=j}^{\infty} \left(u_{\beta_l} \prod_{i=l+1}^{\infty} \frac{\beta_i - 1}{\beta_i + 1} \right) \leqslant \sum_{l=j}^{\infty} u_{\beta_l} < \infty.$$

In particular we have the estimate, for every $j \in \mathbb{N}^*$,

$$R_\epsilon^{\sum_{l=j}^{k} \left(u_{\beta_l} \prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1} \right)} \leqslant \max\left\{ 1, R_\epsilon^{\sum_{l=1}^{\infty} u_{\beta_l}} \right\}. \tag{3.3.40}$$

Furthermore, the first supremum on the right hand side of (3.3.38) is finite since we compute, recalling the expression (3.3.34),

$$\begin{aligned} &\log\left(\prod_{l=1}^{k} \left(\frac{u_{\beta_l}^{-u_{\beta_l}}}{1 - u_{\beta_l}} \beta_l^{4u_{\beta_l}} \right)^{\prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1}} \right) \\ &= \sum_{l=1}^{k} \prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1} \left(\frac{1}{\beta_l + 1} \log(\beta_l + 1) + \log(1 + 1/\beta_l) + \frac{4}{\beta_l + 1} \log(\beta_l) \right) \\ &\lesssim \sum_{l=1}^{\infty} \frac{1 + \log(\beta_l)}{\beta_l} < \infty. \end{aligned}$$

Similarly, the second supremum on the right hand side of (3.3.38) is finite since we compute, recalling also (3.3.35)-(3.3.36),

$$\log\left(\prod_{l=1}^{k} \delta_l^{-2u_{\beta_l} \prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1}} \right) = \sum_{l=1}^{k} 2u_{\beta_l} \prod_{i=l+1}^{k} \frac{\beta_i - 1}{\beta_i + 1} \log(\delta_l^{-1}) \leqslant \sum_{l=1}^{k} \frac{4}{\gamma} \frac{\log(l)}{\beta_l + 1} < \infty. \tag{3.3.41}$$

We now let $k \to \infty$ in (3.3.38) and, by formulas (3.3.39)-(3.3.41) and dominated convergence, we obtain, for a *numeric* constant $c^* \geqslant 0$ only depending on the sequences chosen in (3.3.33)-(3.3.36),

$$\begin{aligned} \mathbb{E}\Big[\|(\rho^{\epsilon,\eta} - \ell)_-\|_{L^\infty(\mathbb{T}^d \times [0,T])} \Big] &= \lim_{k\to\infty} \mathbb{E}\Big[\|(\rho^{\epsilon,\eta} - \ell)_-\|_{L^{\alpha_k}(\mathbb{T}^d \times [0,T])} \Big] \\ &\leqslant c^* \sum_{j=1}^{\infty} j^{-2} R_\epsilon^{\sum_{l=j}^{\infty} \left(u_{\beta_l} \prod_{i=l+1}^{\infty} \frac{\beta_i - 1}{\beta_i + 1} \right)}. \end{aligned} \tag{3.3.42}$$

In particular, thanks to (3.3.40), if $R_\epsilon \to 0$ then the right hand side of (3.3.42) vanishes by dominated convergences. In fact, formula (3.3.33)-(3.3.36) and tedious, but elementary, estimates with geometric series and logarithms allow to further quantify the exponent hitting R_ϵ in (3.3.42), and yield the estimate, for the same numeric constant c^*,

$$\mathbb{E}\Big[\|(\rho^{\epsilon,\eta} - \ell)_-\|_{L^\infty(\mathbb{T}^d \times [0,T])} \Big] \leqslant c^* \sum_{j=1}^{\infty} j^{-2} \begin{cases} R_\epsilon^{d/2(1+2/d)^{2-j}} & \text{if } R_\epsilon \geqslant 1, \\ R_\epsilon^{\exp(-d(d+1))1/2(1+2/d)^{1-j}} & \text{if } R_\epsilon \in (0,1]. \end{cases} \tag{3.3.43}$$

In conclusion, we note that the constant C_0 introduced in (3.3.32) and featuring in the definition (3.3.37) of R_ϵ is independent of $\eta \in (0,1)$. Therefore we pass to the limit $\eta \to 0$ and use Proposition 3.2.23 and Fatou's Lemma in (3.3.43) to obtain an identical estimate for the solution ρ^ϵ of the original equation. This concludes the proof. □

We now go back to the the central limit theorem. Applying Proposition 3.3.8 to the solution $\rho^{\epsilon,\eta}$ of (3.2.12) with the smoothed coefficients ϕ^η, ν^η, σ^η given in Lemma 3.3.4 and using Markov's inequality we readily obtain the following.

Corollary 3.3.9. In the setting of Corollary 3.3.7, for every $\epsilon, \eta \in (0,1)$ and every $\delta \in [0,\ell)$, we have

$$\mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-inf}}\,\rho^{\epsilon,\eta} < \ell\right) = \mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,(\rho^{\epsilon,\eta}-\ell)_- > \ell-\delta\right) \leqslant \frac{c^*}{\ell-\delta}\sum_{j=1}^{\infty}\frac{1}{j^2}\,R_{\epsilon,\eta}^{\zeta\,(1+2/d)^{-j}},$$

where $c^* \in (0,\infty)$ is a *numeric* constant and where

$$\begin{aligned} R_{\epsilon,\eta} &:= C_0\,(\inf\dot{\phi}^\eta)^{-2}\big(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty + \epsilon\,\|F_3^\epsilon\|_\infty\big), \\ \zeta &:= \begin{cases} \frac{d}{2}(1+2/d)^2 & \text{if } R_\epsilon \geqslant 1, \\ e^{-d(d+1)}(1/2+1/d) & \text{if } R_\epsilon \in (0,1), \end{cases} \end{aligned} \tag{3.3.44}$$

for a constant $C_0 = C_0(T,d,\phi,\sigma,\nu,\ell)$ independent of $\epsilon,\eta \in (0,1)$.

In turn, Corollary 3.3.7 and the definition of $\Omega^{\epsilon,\eta}$ yield that, for every $\epsilon \in (0,1)$ and every $\eta \in (0,1)$ small enough,

$$\begin{aligned} \mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|v^\epsilon - v^{\epsilon,\eta}| \neq 0\right) &= \mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|\rho^\epsilon - \rho^{\epsilon,\eta}| \neq 0\right) \\ &\leqslant \mathbb{P}\Big((\Omega^{\epsilon,\eta})^c\Big) \leqslant \frac{c^*}{\ell-\delta_\eta}\sum_{j=1}^{\infty}\frac{1}{j^2}\,R_{\epsilon,\eta}^{\zeta\,(1+2/d)^{-j}}. \end{aligned} \tag{3.3.45}$$

In particular we remark that, as we keep $\eta \in (0,1)$ fixed and we let $\epsilon \to 0$ along the scaling regime (3.3.11), the right hand side of (3.3.45) vanishes by dominated convergence.

We finally establish our main result.

Theorem 3.3.10 (Central limit theorem in probability)**.** Let ρ_0 satisfy Assumption 3.2.1(ii), i.e. a random positive constant, let $(\xi^\epsilon)_{\epsilon>0}$ satisfy Assumption 3.2.4, and let ϕ,ν,σ satisfy Assumption 3.2.6 and 3.2.8, for some $p \geqslant 2$ and $m \geqslant 1$. For every $\epsilon > 0$, let ρ^ϵ be the stochastic kinetic solution to the generalized Dean–Kawasaki equation (3.2.2) with initial data ρ_0 and let $\bar{\rho} \equiv \rho_0$ be the solution to the zero noise limit (3.1.2). Let v be the solution to the Langevin equation (3.1.3) with noise $\xi = \lim_{\epsilon\to 0}\xi^\epsilon$. Along a scaling regime where $\epsilon \to 0$ and $\xi^\epsilon \to \xi$ such that

$$\lim_{\epsilon\to 0}\,\sqrt{\epsilon}\left(\|F_1^\epsilon\|_\infty + \|F_2^\epsilon\|_\infty + \|F_3^\epsilon\|_\infty\right) = 0,$$

for the nonequilibrium fluctuations we have

$$v^\epsilon = \epsilon^{-1/2}(\rho^\epsilon - \bar{\rho}) \to v \quad \text{in } L^\tau_{\text{loc}}([0,\infty);H^{-\beta}(\mathbb{T}^d)) \text{ in probability},$$

for $\tau = 2$ or $\tau = \infty$, for any $\beta > \frac{d}{2}$ or $\beta > 1 + \frac{d}{2}$ respectively, with explicit rate of convergence given in (3.3.47) below.

Proof. We fix $\eta_0 \in (0,1)$ small enough. For any $a > 0$, using formula (3.3.45) and Markov's inequality we estimate

$$\begin{aligned}
&\mathbb{P}\Big(\|v^\epsilon - v\|_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))} > a\Big) \\
&\leqslant \mathbb{P}\left(\|v^\epsilon - v\|_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))} > a,\ \underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|v^\epsilon - v^{\epsilon,\eta_0}| \neq 0\right) \\
&\quad + \mathbb{P}\left(\|v^\epsilon - v\|_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))} > a,\ \underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|v^\epsilon - v^{\epsilon,\eta_0}| = 0\right) \\
&\leqslant \mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|v^\epsilon - v^{\epsilon,\eta_0}| \neq 0\right) + \mathbb{P}\left(\|v^{\epsilon,\eta_0} - v\|_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))} > a\right) \\
&\leqslant \mathbb{P}\left(\underset{\mathbb{T}^d\times[0,T]}{\text{ess-sup}}\,|v^\epsilon - v^{\epsilon,\eta_0}| \neq 0\right) + \frac{1}{a}\mathbb{E}\left[\|v^{\epsilon,\eta_0} - v\|^2_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))}\right]^{\frac{1}{2}}.
\end{aligned} \tag{3.3.46}$$

We keep η_0 fixed and we let $\epsilon \to 0$. The first term on the right hand side of (3.3.46) vanishes because of formula (3.3.45), the scaling regime chosen and dominated convergence. The second term vanishes because of the scaling regime and Corollary 3.3.7. Since $a > 0$ is arbitrary, we conclude that $v^\epsilon \to v$ in probability.

To obtain an explicit convergence rate, we consider again the last line of formula (3.3.46). The constant δ_η defined in (3.3.19) decreases to zero as $\eta \to 0$, and therefore is eventually smaller than $\ell/2$ for any $\eta \in (0,\eta_0]$, for a suitable η_0 which only depends on ℓ, the original coefficients ϕ,η,σ and the approximation procedure employed in Lemma 3.3.4. Fixed such an $\eta_0 \in (0,1)$, we can explicitly compute the constant C appearing in formula (3.3.10). Indeed, the constant depends on this fixed η_0 only through the original coefficients ϕ,ν,σ and the approximation procedure used in Lemma 3.3.4. Furthermore, the definition of R_{ϵ,η_0} in (3.3.44) depends on this fixed η_0 only through $\inf \dot{\phi}^{\eta_0}$, and we can require this to be greater than some positive number, say $\ell/3$. Applying this reasoning and plugging (3.3.10) and (3.3.45), with this η_0, into the last line of (3.3.46) we obtain, for arbitrary $a > 0$,

$$\begin{aligned}
&\mathbb{P}\Big(\|v^\epsilon - v\|_{L^\tau([0,T];H^{-\beta}(\mathbb{T}^d))} > a\Big) \\
&\leqslant C\sum_{j=1}^{\infty}\frac{1}{j^2}\left(\epsilon^2\,\|F_1^\epsilon\|_\infty\|F_3^\epsilon\|_\infty + \epsilon\,\|F_3^\epsilon\|_\infty\right)^{e^{-d(d+1)}(1+2/d)^{1-j}} \\
&\quad + \frac{C}{a}\left(\epsilon\left(|F_1^\epsilon|_\infty + |F_3^\epsilon|_\infty\right)^2 + \epsilon^{1/2}|F_3^\epsilon|_\infty^{1/2}\right) \\
&\qquad\qquad \cdot\left(1+\epsilon\,|F_3^\epsilon|_\infty\right)^{g+\frac{d}{2}(p+m)}\left(1+\mathbb{E}\left[\|\rho_0\|^{m+p-1}_{L^1(\mathbb{T}^d)} + \|\rho_0\|^p_{L^p(\mathbb{T}^d)}\right]\right) \\
&\quad + \frac{C}{a}\sum_{n\in\mathbb{Z}^d} n^{(2-\frac{4}{\tau})-2\beta}\sum_k\int_0^T\left|\int_{\mathbb{T}^d} e^{i2\pi nx}\left(f_k^\epsilon - f_k\right)dx\right|^2 ds,
\end{aligned} \tag{3.3.47}$$

for a constant $C = C(\ell,\phi,\nu,\sigma,T,d,p,\beta,\bar{\rho})$ depending only on the initial data and the original coefficients, which can be computed explicitly in terms of $\ell = \text{ess-inf}\rho_0$ and the constants c featuring in Assumption 3.2.6 and 3.2.8. □

Part II

An SDE system for interacting neurons

Chapter 4

Introduction

4.1 State of the art and open questions

Grid cells and mathematical models in neuroscience

Grid cells are a particular type of neuron in the brain of mammals discovered in 2005 [HFM+05], see also the review [MMM17, RRMM16]. These neurons fire at regular intervals as an animal moves across an area, storing information such as position, direction and velocity, and thus enabling it to understand its movement in space. In a nutshell, the regular firing of grid cells creates characteristic activation patterns in the brain cortex, typically hexagonal lattices for 2D [CWZ+13, RRMM16] and Face-Centered-Cubic lattices for 3D navigation [KM19, YU13], encoding the mapping of the space.

Since their discovery, there has been extensive research to understand the precise behavior of grid cells (cf. [RRMM16, MMM17, Bre11] and the references therein). Mathematically, their network is commonly described by deterministic continuous attractor-network dynamics through a system of neural field models [ET10, MBJ+06, BF09, CWZ+13], based on the seminal papers [WC72, WC73, Ama77] and [PBS+96] (see also the review [BSH18]).

In particular, as the brain is inherently noisy [RD10, ET10, Bre14], understanding how the grid cell network is affected by noise is one of the currently open challenges in the field [RRMM16]. This question has recently been addressed from several directions [Bre19, BF12, BAC19, KE13, MB20, Tou12, CHS22, CRS23a].

In this second part of the thesis we study the following SDE system, based on the neuron network models from [BF09, CWZ+13] (see also the reviews in [Bre14, BSH18]). Given space points $x_1, \ldots, x_N \in Q$ in a region Q of the neural cortex, we have the interaction among NM neurons stacked in N columns at locations x_i with M neurons each. The *activity level* u_{ik} (a quantity of use in neuroscience akin to the firing rate and *defined* by the following equation) of the k^{th} neuron at location x_i evolves according to

$$\begin{aligned} \mathrm{d}u_{ik}(t)\,\tau_i =& \sigma \mathrm{d}W_{ik}(t) - \mathrm{d}\ell_{ik}(t) \\ &+ \left(-u_{ik}(t) + \phi\Big(B(x_i,t) + \frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M} K(x_i - x_j) u_{jm}(t) \Big) \right) dt, \end{aligned} \tag{4.1.1}$$

where ℓ_{ik} is a reflecting term (Skorokhod problem) ensuring the activity level stays positive, as physically meaningful, rigorously defined as the finite variation process

$$\ell_{ik}(t) = -\big|\ell_{ik}\big|(t), \quad \big|\ell_{ik}\big|(t) = \int_0^t 1_{\{u_{ik}(r)=0\}}\, d\big|\ell_{ik}\big|(r).$$

For simplicity we consider $Q = [0,1]^d$, but most of the results extend to any bounded open subset $Q \subseteq \mathbb{R}^d$, for any $d \geqslant 1$. For integers $k = 1, \ldots, M$, we have i.i.d. families of random initial conditions $\{u_{ik}(0)\}_{i=1,\ldots,N}$ for each space point x_i in the cortex Q.

For integers $i = 1, \ldots, N$ and $k = 1, \ldots, M$, the Brownian motions W_{ik}, possibly correlated, represent the noise on the system. The modelling choice of Brownian noise instead of Poisson is formally justified in [ET10, Bre14] and is essentialy due to the wide variety of uncorrelated random stimula received by neurons (opening/closing of ion channels, release of neurotransmitters, asynchronous firing of other neurons etc).

The nonlinear function $\phi : \mathbb{R} \to \mathbb{R}$ represents the *firing rate* of neurons in the network, typically a Lipschitz function like a ReLU function. The external input $B : Q \times [0,\infty) \to \mathbb{R}$ and the interaction kernel $K : \mathbb{R}^d \to \mathbb{R}$ are only required to be locally bounded functions and Hölder continuous in the x variable. The interaction kernel takes into account the inhibitory/excitatory effect on nearby neurons and a typical choice in computational neuroscience is the Mexican hat function. The typical relaxation times are indicated by τ_i and for simplicity we will take $\tau_i = 1$ in this introduction.

In the modelling of grid cells, one often considers families of neurons with orientation preference $\beta = 1, \ldots, d_v$, typically $d_v = 4$ or 6 for the spatial directions in 2D or 3D navigation respectively. The above equations extend to a system where the vector $u_{ik} = (u_{ik}^1, \ldots, u_{ik}^{\beta})$ tracks the activity level for the various orientations and one can also consider orientation dependent coefficients B^β, ϕ^β, W_{ik}^β etc. We use vectorial notation in this exposition but avoid the index β for simplicity.

We will sometime refer to the network (4.1.1) as a *spatially extended* SDE system. Indeed, it is more accurate to regard (4.1.1) as a collection of N mean field systems with M particles each, one for each column of neurons, interacting among themselves, rather than a classical mean field system with MN particles. The main consequence of this dependence on $x \in Q$ in the system is the *loss of exchangeability*: neurons u_{ik} are exchangeable in the index k, tracking the position within the column, but not on the index i, tracking the location in the cortex. This feature reflects into specific scaling regimes between the number N of clusters and the number M of neurons per cluster needed to observe coherent effects in the thermodynamic limit.

We note that the argument of the firing rate ϕ in (4.1.1) can be rewritten as

$$\frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M} K(x_i - x_j)u_{jm}(t) = \int_{Q\times\mathbb{R}^{d_v}} K(x_i - y)\, u\, f_{N,M}(t, dy, du),$$

for the empirical measure $f_{N,M}$ of the particle system

$$f_{N,M}(t, dy, du) = \frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M} \delta_{(x_j, u_{jm}(t))} \quad \text{regarded as a measure on } Q \times \mathbb{R}^{d_v}. \tag{4.1.2}$$

Applying Ito formula to the empirical measure and formally passing to the thermodynamic limit $M, N \to \infty$, the probability distribution f of the activity level $u \in \mathbb{R}^{d_v}$ at location x should satisfy the Fokker–Planck equation

$$\begin{aligned}\partial_t f(t,x,u) + \nabla_u \cdot \left(f(t,x,u) \left(-u + \phi\Big(B(x,t) + \int_{Q\times\mathbb{R}^{d_v}} K(x-y)\, v\, f(t,dy,dv)\Big) \right)\right) \\ = \frac{\sigma^2}{2}\Delta_u f(t,x,u),\end{aligned} \tag{4.1.3}$$

with the no-flux boundary conditions, coming from the reflecting boundaries at the SDE level,

$$f(t,x,u) \left(-u + \phi\Big(B(x,t) + \int_{Q\times\mathbb{R}^{d_v}} K(x-y)\, v\, f(t,dy,dv)\Big) \right) - \frac{\sigma^2}{2}\nabla_u f(t,x,u)\Big|_{u=0} = 0. \tag{4.1.4}$$

In particular, we remark the meaning of the modelling: at each time $t \in \mathbb{R}$, space patterns $Q \ni x \mapsto f(t,x,u)$ correspond to the animal's mapping of the space in the brain. However, the specific correspondence between points $x \in Q$ in the cortex and points in physical space is not clear yet [BF09, SSS$^+$12].

Similarly, formally passing to the thermodynamic limit in (4.1.1) gives the following McKean–Vlasov equation

$$\begin{aligned}\mathrm{d}\bar{u}^\epsilon(x,t) &= \sigma\, \mathrm{d}W^\epsilon(x,t) - \mathrm{d}\bar{\ell}(x,t) \\ &\quad + \left(-\bar{u}^\epsilon(x,r) + \phi\Big(B(x,r) + \frac{1}{4}\int_{Q\times\mathbb{R}^{d_v}} K(x-y)\, v\, f(t,dy,dv)\Big)\right) \mathrm{d}t,\end{aligned} \tag{4.1.5}$$

where $f(t,dx,du) = \mathrm{Law}(\bar{u}(t))$ solves the above Fokker–Planck equation and is independent of ϵ. The parameter ϵ, related to the space correlation of the noise, deserves particular care and is discussed in detail in the next subsection.

In this thesis, the SDE system (4.1.1) is generalized and rewritten in the form

$$\mathrm{d}u_{ik}(t) = b(x_i, t, u_{ik}(t), f_{N,M}(t))\, \mathrm{d}t + \sigma(x_i, t, u_{ik}(t), f_{N,M}(t))\, \mathrm{d}W_{ik}(t) - \mathrm{d}\ell_{ik}(t) \tag{4.1.6}$$

for general drift and diffusion terms, Lipschitz in the activity level variable u and α-Hölder in the space variable x, depending on the empirical measure (4.1.2) in the form

$$b(x,t,u,f) = b_0(x,t,u) + \phi\left(\int_{Q\times\mathbb{R}^{d_v}} b_1(x,y,t,u,v)\, f(dy,dv) \right) \tag{4.1.7}$$

for suitable coefficients b_0, b_1 and nonlinearity ϕ, and similarly for σ.

Again a formal passage to the thermodynamic limit $M, N \to \infty$ yields the McKean–Vlasov equation

$$\mathrm{d}\bar{u}^\epsilon(x,t) = b(x,t,\bar{u}^\epsilon(x,t), f(t))\,\mathrm{d}t + \sigma(x,t,\bar{u}^\epsilon(x,t), f(t))\,\mathrm{d}W^\epsilon(x,t) - \mathrm{d}\bar{\ell}(x,t), \qquad (4.1.8)$$

and the associated Fokker–Planck equation

$$\partial_t f(t,x,u) + \nabla_u \cdot \Big(b(x,t,u,f(t)) f(t,x,u)\Big) = \frac{1}{2}\Delta_u \Big(\sigma(x,t,u,f(t))^2 f(t,x,u)\Big), \qquad (4.1.9)$$

again with no-flux boundary conditions at $u = 0$ analogous to (4.1.4).

In this generality the system covers other models for grid cells [AB20, BF12] and, more broadly, several other network models for interacting neurons [BFFT12, THF12, FTC09, Tou11, CT18]. In particular, simply removing the noise $\sigma = 0$, it covers the classical formulation of the Amari model [Ama77, FI15]

$$\partial_t u(t,x) = -u(t,x) + \int_Q K(x,y)\, G(u(t,y))\, dy.$$

We also mention the growing interest in stochastic versions of continuous neural field equations (cf. [FI15, Bre14, Bre11] and the references therein). Many of these are covered by model (4.1.8) after introducing environmental noise, as discussed later on.

More generally, beyond computational neuroscience, there is increasing interest in these *spatially extended* SDE systems analogous to (4.1.6), for example in PDEs on graphs [Fan21, EPSS21], where space locations correspond to nodes of the graph, or stochastic gradient descent algorithms in machine learning [SS20, GGK22, WPC$^+$20], where space points track the sub-batches of data considered for parameter optimization.

The modelling of the noise

The parameter ϵ featuring in (4.1.5) and (4.1.8) deserves particular care and we discuss the main aspects. First, we stress that we will always keep $\epsilon > 0$ fixed in the following, although letting $\epsilon \to 0$ in the thermodynamic limit can also be explored and we quickly discuss this below.

The noise $W^\epsilon(x,t)$ featuring in (4.1.8) is a suitable mollified *rescaled* version of space-time white noise on $Q \times [0,\infty)$. From the modelling point of view, at any discrete step $M, N \in \mathbb{N}$ we want each neuron of the network (4.1.6) to be perturbed by a standard Brownian motion W_{ik}. To formalize this mathematically, we consider a sequence of actual space-time white noise terms $\{W_k(x,t)\}_{k\in\mathbb{N}}$ in $Q \times [0\infty)$. The white noise terms W_k are taken independent of each other, but other modelling choices like a constant sequence $W_k \equiv W$, essentially resulting in environmental noise, are possible (cf. the next subsection and Remarks 5.1.3 and 6.1.2).

Then, for any $\epsilon > 0$, we define

$$W_k^\epsilon(x,t) := C_\rho^{\frac{1}{2}}\, \epsilon^{\frac{d}{2}}\, \langle W_t, \rho_\epsilon(\cdot - x)\rangle \quad \text{with } C_\rho = \left(\int_{\mathbb{R}^d} \rho(z)^2\, dz\right)^{-1}, \tag{4.1.10}$$

where ρ_ϵ is the standard ϵ-rescaling of a mollifier $\rho : \mathbb{R}^d \to [0,1]$ supported in the unitary ball. Equivalently, $W_k^\epsilon : Q \times [0,\infty)$ is a Gaussian random field with zero mean and covariance

$$\mathbb{E}\left[W_k^\epsilon(x,t) W_k^\epsilon(y,s)\right] = (t \wedge s)\, C_\rho\, \epsilon^d \int_{\mathbb{R}^d} \rho_\epsilon(z-x)\rho_\epsilon(z-y)\, dz.$$

Setting $W_{ik}(t) := W_k^\epsilon(x_i, t)$ gives the desired Brownian motions.

The choice of the modelling is dictated by the following considerations. First, we want each neuron in (4.1.1) to sense noise with a given fixed strength modulated only by σ, whereas considering a mollified white noise $\tilde{W}^\epsilon := W * \rho_\epsilon$ without the $\epsilon^{d/2}$ rescaling would lead to stronger and stronger noise as ϵ decreases. Similarly, the goal of the noise perturbed network (4.1.1) introduced in [CCS23] was the recover the diffusion enhanced Fokker–Planck equation (4.1.3) first proposed in [CHS22], whereas considering the noise $\tilde{W}^\epsilon$ would give increasing diffusion $\epsilon^{-d}\sigma^2$.

The law $f(\cdot, x, du) := \mathrm{Law}_{C([0,T];\mathbb{R}^{d_v})}(\bar{u}^\epsilon(x,\cdot))$ of a single 'McKean–Vlasov particle' at location $x \in Q$, and thus the Fokker–Planck equation (4.1.9), is independent of the noise correlation radius ϵ because in (4.1.7)-(4.1.8) the interaction between two neurons $\bar{u}^\epsilon(x,t)$ and $\bar{u}^\epsilon(y,t)$ happens only through their law f as single random variables, so that we do not get a BBGKY hierarchy. Other models, like the stochastic Amari model (4.1.12) below, do not exhibit this feature and the one particle distribution does depend on the noise correlation. In fact, the dependence on the correlation radius is recovered as soon as we consider the joint law $f^\epsilon(x,y,du,dv) = \mathrm{Law}(\bar{u}^\epsilon(x,\cdot), \bar{u}^\epsilon(y,\cdot))$ of two or more particles. For example, this aspect will be apparent when analyzing the fluctuations of the network (4.1.17)-(4.1.18) and will also give issues when trying to apply a 'Dean–Kawasaki approach' to approximate the system.

Next we discuss fixing $\epsilon > 0$ versus letting $\epsilon \to 0$ along a suitable scaling regime as as $M, N \to 0$. The choice of fixing $\epsilon > 0$ is dictated by the modelling consideration that closeby neurons $x \sim y$ in the cortex should sense correlated noise, with a typical correlation radius ϵ. Indeed, the choice (4.1.10) implies that, although constant at microscopic scales, the strength of the noise on macroscopical scales – i.e. testing with 'space observables' of the form $\langle \varphi, W^\epsilon \rangle = \int_Q \varphi(x)\, W^\epsilon(x,t)\, dx$ – vanishes as the correlation $\epsilon \to 0$.

Similarly to what happens for the law of McKean–Vlasov particles, the effect of the correlation radius is not seen at order zero in the mean field limit. As stated in (4.2.1) (cf. the theorems in Section 4.2), the convergence $f_{MN} \to f$ in the mean field limit holds also if we let $\epsilon \to 0$ at the same time.

As anticipated, the effect of ϵ is instead visible when analyzing the fluctuations of the network (4.1.17)-(4.1.18). Indeed, if we let $\epsilon \to 0$ as $M, N \to \infty$, the martingale term M_t^{MN}

featuring in (4.1.16) vanishes in the limit and the Langevin equation (4.1.17) giving the first order corrections simply becomes a linear PDE with random initial data η_0 (which vanishes too in case of deterministic initial data, leading to trivial first order correction).

Finally, we briefly discuss the option of considering a mollified white noise $\tilde{W}^\epsilon := W * \rho_\epsilon$ *not* rescaled by the extra factor $\epsilon^{d/2}$, ignoring the aforementioned modelling choices. For brevity we simply consider the effect of letting $\epsilon = 0$ directly in the McKean–Vlasov equation.

For white noise, considering the case (4.1.8) or the stochastic Amari model (4.1.12) below to fix ideas, even for additive noise $\sigma \equiv$ constant or σ a bounded Lipschitz nonlinearity, the typical drift term considered in neuroscience does not provide enough smoothing for the stochastic integral to exist in a function space and obtain function valued solutions [DPZ92], and we would need to make sense of the nonlinear part of the drift applied to the distributional solution.

On the other hand, in the specific case of (4.1.5) or the Amari model (4.1.12) with $G(u) = u$ respectively, with additive noise, this is doable and one has the weak formulation for the McKean–Vlasov equation $u(x,t) \in L^2\big(\Omega \times [0,T]; H_x^{-d/2^-}\big)$, for test functions ψ,

$$\langle \psi, u(t)\rangle - \langle \psi, u(0)\rangle = \int_0^t \langle \psi, -u(s) + \phi\left(B(x,s) + \mathbb{E}[(u(s) * K)(x)]\right)\rangle \, ds + \sigma \langle \psi, W_t\rangle, \qquad (4.1.11)$$

or the same with $\mathbb{E}[(u(s) * K)(x)]$ replaced by $(u(s) * K)(x)$ respectively.

In this case, considering a suitable scaling regime between $M, N \to \infty$ and $\epsilon \to 0$ and adapting the arguments from Chapter 5, it might actually be possible to prove that the network (4.1.1) with noise $W_{ik} := \tilde{W}_k^\epsilon(x_i)$ converges to the solution of (4.1.11).

Wellposedness, mean field limit and environmental noise.

The wellposedness of the generalized network (4.1.6) is obtained in [CCS23], adapting classical arguments for the treatment of reflecting boundary conditions (Skorokhod problem) [LS84, Szn84b].

As for the Fokker–Planck equations, systems of the form (4.1.3) are challenging to analyse because of the non-local nonlinear drift and the nonlinear Robin-like boundary condition, so that standard well-posedness theory for parabolic equations with smooth initial data does not apply. The wellposedness of the generalized McKean–Vlasov equation (4.1.8) and Fokker–Planck PDE (4.1.9), in the weak sense, is instead obtained with a contraction argument in suitable functional spaces in the work [CCS23] , again adapting the approach from [Szn84b].

For the specific Fokker–Planck equation (4.1.3), wellposedness of classical solutions have later been proved in [CRS23b]: building on methods developed for other nonlinear Fokker–Planck PDEs arising in neuroscience, a change of variables transforms the equation into a Stefan-like free boundary problem for which solutions can be represented via Green functions and local or global-in-time existence can be probed with explicit estimates.

Concerning the thermodynamic limit $M, N \to \infty$, as already mentioned, it is more accurate to regard the spatially extended network (4.1.1) as a collection of N mean field systems with M particles each, one for each column of neurons, interacting among themselves, rather than a classical mean field system with MN particles. This feature makes particles u_{ik} *not exchangeable* in the index i for space location $x \in Q$ and in turn reflects into the specific scaling regimes between the number N of clusters and the number M of neurons per clusters needed to observe coherent effects(cf. the statements in Section 4.2).

The passage to the mean field limit is rigorously proved in [CCS23]: the network (4.1.6) converges to the McKean–Vlasov SDE (4.1.8) and its empirical measure converges to the solution of the Fokker–Planck PDE (4.1.9).

Finally, we note that from the modelling point of view it makes sense that neurons in the same cluster also sense similar noise, possibly besides their own specific noise, and this similar noise adds a coherent stochastic term at the PDE level. There is indeed growing interest in considering stochastic versions of continuous neural field equations (cf. [FI15, Bre14, Bre11] and the references therein), for example stochastic Amari models

$$d\,u(x,t) = \left(-u(x,t) + \int_Q K(x,y)\,G(u(y,t))\,dy\right)\,dt + \theta(x,u(x,t))\,d\xi(x,t), \tag{4.1.12}$$

where ξ is a suitable cylindrical Wiener process in $Q \times [0,\infty)$, studied in [FI15, KT19, KR14].

Stochastic PDEs of this kind essentially arise as mean field limit of networks subjected to common noise and we quickly discuss this. Consider for example the following network, akin to (4.1.6) and further subjected to environmental noise,

$$du_{ik}(t) = b(x_i, t, u_{ik}(t), f_{MN})\,dt + \sigma\,dW_{ik}(t) + \theta(x_i, u_{ik}(t))\,d\xi(x_i,t), \tag{4.1.13}$$

where for simplicity we removed the reflecting term ℓ_{ik} and where $\xi(x,t)$ is the above Wiener process (or a suitable regularization). Following classical arguments [CF16], the empirical measure conditioned with respect to the noise ξ

$$\bar{f}_{MN} = \mathbb{E}\left[\frac{1}{MN}\sum_{i=1}^{N}\sum_{h=1}^{M}\delta_{(x_i,u_{ih})}\middle|\mathcal{F}_t^{\xi}\right]$$

formally solves the stochastic Fokker–Planck

$$\begin{aligned}\partial_t \bar{f}(t,x,u) &+ \nabla_u\cdot\Big(b(x,t,u,\bar{f}(t))\,\bar{f}(t,x,u)\Big)\\ &= \frac{1}{2}\Delta_u\big((\sigma^2+\theta(x,u)^2)\bar{f}(t,x,u)\big) - \nabla_u\cdot\Big(\theta(x,u)\bar{f}(t,x,u)\,\xi(x,t)\Big),\end{aligned}$$

and converges, in the $M, N \to \infty$ limit, to the conditional law $\bar{f} = \mathrm{Law}_{\mathbb{R}^d}\left(u(x,t)\,|\,\mathcal{F}_t^{\xi}\right)$ of the McKean–Vlasov SDE

$$du(x,t) = b(x,t,u(x,t),\bar{f})\,dt + \sigma\,dW^{\epsilon}(x,t) + \theta(x,u(x,t))\,d\xi(x,t).$$

In particular, taking $\sigma = 0$ and $b(x,t,u,f) := -u + \int_{Q\times\mathbb{R}} K(x,y)\, G(v)\, f(dy,dv)$ recovers the Amari model (4.1.12). Extensions with common noise depending also on the activity level $u \in \mathbb{R}$ are of course possible and implemented simply by replacing the last term in (4.1.13) with cylindrical expansions in $Q \times \mathbb{R} \times [0,\infty)$ like $\sum_{l=1}^{\infty} \theta_l(x_i, u_{ik})\, \xi_l(x_i,t)$ for suitable sequences θ_l and ξ_l.

Long time behavior and phase transition.

We remark that space patterns $Q \ni x \mapsto f(t,x,u)$ in the distribution of activity levels correspond to mapping of space in the mammal's brain. Stationary states are therefore of relevant importance, where space homogeneous ones correspond to complete absence of any understanding of space navigation. Intuitively, we expect these space homogeneous states to become more and more attractive as the noise strength σ increases.

For the discrete network (4.1.1), under suitable assumptions on the coefficients and essentially regardless of the noise strength $\sigma > 0$, one obtains existence and uniqueness of stationary states by solving the associated linear Fokker–Planck equation. Furthermore, one can often prove convergence to equilibrium as $t \to \infty$. For example, in the simplest case where $\|\phi\|_{\mathrm{Lip}}\|K\|_\infty < 1$, the evolution semigroup $\mu \mapsto \mathcal{P}_t^* \mu$ associated to (4.1.1) is readily seen be an exponential contraction in Wasserstein 2 distance.

In the continuum limit, where the Fokker–Planck equation is nonlinear, the situation is more complex. The long time behavior of (4.1.3) has been investigated in the series of papers [CHS22, CRS23a, CRS23b].

Considering the the 1D case for simplicity (no preferred neuron orientations), under the assumption of constant external input B, setting the left hand side of (4.1.3) to zero, integrating in u and using the no-flux boundary conditions (4.1.4), the stationary states satisfy

$$\sigma \partial_u f(x,u) = -(u - \Phi(K * \bar{f}(x) + B)) f(x,u),$$

where we denoted the average activity level

$$\bar{f}(x) := \int_0^{+\infty} u f(x,u)\, \mathrm{d}u.$$

Therefore the stationary states must solve

$$f(x,u) = \frac{1}{Z_f} \exp\left(-\frac{(u - \Phi(K * \bar{f}(x) + B))^2}{2\sigma} \right) \tag{4.1.14}$$

for the normalization

$$Z_f = \mathrm{meas}(Q) \int_0^{+\infty} \exp\left(-\frac{(u - \Phi(K * \bar{f}(x) + B))^2}{2\sigma} \right) \mathrm{d}u.$$

Under suitable assumptions on the coefficients, in [CHS22] they prove existence and uniqueness of *space homogeneous* stationary states for any noise strength $\sigma > 0$ and that these states are linearly asymptotically stable for noise strong enough depending on the coefficients. As σ decreases, multiple continuous bifurcations of other stationary states arise from the space homogeneous branch and their form is characterized locally in [CRS23a]. Finally, under suitable assumptions, in [CRS23b] the above linear stability is upgraded to full asymptotic stability for any stationary state.

Numerical simulations from [CHS22, CRS23a] confirm the analytical results and highlight new phenomena. There is numerical evidence that the PDE admits the characteristic hexagonal-pattern stationary states, more and more blurred as the noise increases, besides the always present space homogeneous states. The space homogeneous states eventually lose asymptotic stability as the noise σ decreases. Both continuous and discontinuous phase transition away from the homogeneous state are observed: stationary states continuously bifurcating from the homogeneous one (as demonstrated analytically) do not feature hexagonal patterns (as indeed expected from heuristic arguments), while hexagonal-pattern steady states arise discontinuously as the noise σ decreases, giving rise to hysteresis.

Open question. A full characterization of stationary states to (4.1.3) is not available beyond the fixed point problem (4.1.14). Moreover, the commutation of the thermodynamic limit $M, N \to \infty$ and the long-time $t \to \infty$ limit is not clear. One might expect the absence of phase transition at the continuum level to be related to propagation of chaos uniformly in time for the particle system (cf. the results for classical particle systems e.g. with gradient structure [DGPS23, CD22]). In particular, analyzing the thermodynamic limit of the discrete network's steady states could shed light on the picture at the PDE level and the bifurcation phenomena.

Fluctuations, large deviations and 'Dean–Kawasaki framework'.

The aforementioned mean field limit (cf. Section 4.2) yields the formal expansion, as elements of $\mathscr{P}(Q \times \mathbb{R}^{d_v})$,

$$f_{MN} = f + O\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right) \quad \text{as } M, N \to \infty. \tag{4.1.15}$$

The effect of the noise at the particle level, beyond the simple diffusion, has been analyzed in [Cli23a]. The rescaled fluctuations $\eta_t^{MN} := \sqrt{M}(f_{MN} - f)$ admit a semimartingale decomposition

$$\eta_t^{MN} = \eta_0^{MN} + \int_0^t \mathscr{L}_r(f_{MN}, f)^*[\eta_r^{MN}]\, dr + M_t^{MN}, \tag{4.1.16}$$

and in the thermodynamic limit they converge to the solution of the Langevin SPDE

$$\eta_t^\epsilon = \eta_0 + \int_0^t \mathscr{L}_t(f, f)^*[\eta_r^\epsilon]\, dr + G_t^\epsilon, \tag{4.1.17}$$

where the operator $\mathscr{L}(f, f)^*$ is essentially the adjoint of the linearization of the generator of the McKean–Vlasov SDE around the mean field limit f, and η_0 and G_t^ϵ are a suitable

Gaussian r.v. and Gaussian process, uniquely characterized by their covariance and quadratic variation respectively.

In the spirit of (4.1.15), this yields the first order expansion, in the sense of distributions,

$$f_{MN} = f + \frac{1}{\sqrt{M}}\eta_t^\epsilon + o\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right) \quad \text{as } M, N \to \infty \text{ and } \sqrt{M}N^{-\alpha/d} \to 0, \tag{4.1.18}$$

where we remind that the specific scaling regime between M and N is due to the 'spatial structure' of the network.

Open question. As already mentioned in Section 1.1 when discussing the small noise fluctuations of the Dean–Kawasaki equation, we point out the recent advances on higher order corrections to the empirical measure μ_n of particle systems in the thermodynamic limit, beyond the first order Gaussian fluctuations [HCR23, CS21, ACR21]. In the same spirit of (4.1.18), these give an expansion in the number n of particles

$$\mu_n = \sum_{k=0}^{K} \mu^{(k)}\, n^{-k/2} + o\left(n^{-K/2}\right),$$

where $\mu^{(k)}$ are solutions to a recursive set of Langevin equations. It would be interesting to develop the same framework for spatially dependent SDE systems (4.1.2).

The noise induced rare events in the neuron network can also be studied. Considering additive noise for simplicity, the 'weak approach to large deviations' developed in [BDF12] allows to formally identify the following large deviation principle for the empirical measure $f_{M,N} \in C([0,T]; \mathcal{P}(Q \times \mathbb{R}^d))$ of (4.1.6):

$$\mathbb{P}\left(f_{MN} = g\right) \simeq e^{-M\mathcal{I}(g)} \quad \text{as } M, N \to \infty,\ \sqrt{M}N^{-\frac{\alpha}{d}} \to 0,$$

for the rate function

$$\begin{aligned} \mathcal{I}(g) = \inf_w \Bigg\{ \frac{1}{2}\mathbb{E}_{\tilde{\Omega}}\int_0^T\int_Q |w(x,t)|^2 dx\, dt \ &: \text{the McKean-Vlasov SDE} \\ \mathrm{d}\tilde{u}^\epsilon(x,t) &= b(x,t,\tilde{u}^\epsilon, g(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W^\epsilon(x,t) \\ &\quad + \sigma C_\rho^{\frac{1}{2}}\epsilon^{\frac{d}{2}}\rho_\epsilon * w(x,t)\,\mathrm{d}t - \mathrm{d}\tilde{\ell}(x,t) \\ &\text{is such that } g(t,x,du) = \mathrm{Law}(\tilde{u}(x,t)) \Bigg\}, \end{aligned} \tag{4.1.19}$$

where $w \in L^2(\tilde{\Omega} \times Q \times [0,T])$ varies over all the possible stochastic bases $\tilde{\Omega}$ supporting such a 'controlled' McKean–Vlasov SDE. That is, the cost of a given profile g deviating from the mean field limit f is (the exponential of the infimum of) the L^2 energy of a stochastic drift w forcing the profile g instead of f.

Open question. As in Section 1.1, we mention again the connection between large deviation principles and stationary or metastable states. In particular, because of their modelling relevance, in the concrete case (4.1.3) it would be worth investigating whether these can furnish a selection criterion for the aforementioned multiple steady states or characterize metastability.

In conclusion, we discuss the difficulties in applying a 'Dean–Kawasaki framework' as discussed in Chapter 1: that is finding a suitable stochastic PDE enclosing all the information (4.1.18)-(4.1.19) gathered on the system (4.1.6), whose solution ρ_M captures the fluctuations

$$f_{MN} = \rho_M + o\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\frac{\alpha}{d}}}\right) \quad \text{as } M, N \to \infty \text{ and } \sqrt{M}N^{-\frac{\alpha}{d}} \to 0,$$

and exhibits the same large deviations of the network in the small noise limit.

As already discussed in (1.1.12)-(1.1.13) in Section 1.1, the naive approximation $f_{MN} \simeq f + \frac{1}{\sqrt{M}}\eta^\epsilon$ furnished by the fluctuation analysis (4.1.18) would miss the rare events (4.1.19) of the network.

On the other hand, the large deviation principle as conjectured formally in (4.1.19) does not seem to immediately suggest a suitable stochastic PDE exhibiting the same rare events. Indeed, extracting a Fokker–Planck equation for the test profile g from the controlled McKean–Vlasov equation in (4.1.19) involves in principle the joint law of the controlled McKean–Vlasov particle $\tilde{u}$ and the control w. In this sense, it would be worth trying to express a large deviation principle via the classical approach by Dawson and Gärtner [DG87].

Finally, Dean's trick (1.1.2)-(1.1.3) – i.e. replacing the martingale term in Ito formula for the empirical measure – does not seem to help either because of the nonzero space correlation $\epsilon > 0$ retained in the thermodynamic limit. Namely, one would conjecture the following stochastic Fokker–Planck

$$\begin{aligned} \partial_t \rho_M(t,x,u) + \nabla_u \cdot \Big(b(x,t,u,\rho_M(t))\, \rho_M(t,x,u)\Big) \\ = \frac{\sigma^2}{2}\Delta_u \rho_M(t,x,u) - \frac{1}{\sqrt{M}}\nabla_u \cdot \Big(\sigma \sqrt{\rho_M(t,x,u)}\, \xi^\epsilon(t,x,u)\Big), \end{aligned}$$

where the noise ξ^ϵ is white in the time t and the activity level u, and has the same rescaling and mollification as in (4.1.10) in the x variable. However, regardless of rare events, in this case we have problems with the fluctuations. Indeed a formal computation shows that the fluctuations of the solution $\sqrt{M}(\rho_M - f)$ in the small noise limit $M \to \infty$ converge to a Langevin SPDE

$$\mathrm{d}\tilde{\eta}^\epsilon_t = \mathscr{L}_t(f,f)^*[\tilde{\eta}^\epsilon_t]\,\mathrm{d}t + \mathrm{d}\tilde{G}^\epsilon_t,$$

with a Gaussian process $\tilde{G}^\epsilon_t$ different (different quadratic variation) from the process G^ϵ_t featuring in (4.1.17).

Numerical aspects

We finally quickly discuss simulations and numerical aspects. In general, numerical simulations of neural networks and continuous neural fields are widespread. We refer to the monographs [Izh07, GKNP14] for an overview of the models and related references and to the reviews [GMM11, RRMM16] specifically for grid cells.

The network (4.1.1) has been simulated e.g. in [BF09, CWZ$^+$13, MBJ$^+$06], giving numerical evidence that the model is indeed able to reproduce the characteristic hexagonal activation patterns observed in the mammalian brain cortex.

In the presence of noise, the Fokker–Planck equation (4.1.3) and its long time behavior have been studied numerically in [CHS22, CRS23a]. As discussed in the subsection on the long time behavior, the simulations confirm the analytical results available and highlight a whole range of new phenomena, such as continuous and discontinuous phase transition, yet to be rigorously understood.

Open question. As regards simulations of the noisy neuron network (4.1.1), variance reduction techniques are crucial for rare event sensitive observables. In particular, there has been recent advances in importance sampling techniques for SDE systems [dRST23, RHAPT22b, RHAPT22a] and it would be worth extending these to the spatially extended case.

4.2 The contribution of this thesis

4.2.1 Chapter 5: wellposedness and mean field limit

The main contribution of Chapter 5 is to establish the well-posedness framework for the particle system (4.1.6), the McKean–Vlasov SDE (4.1.8) and Fokker–Planck PDE (4.1.9), and to rigorously prove the mean field limit as discussed previously.

The approach is essentially an adaptation of Sznitman's coupling method [Szn91], accounting for the reflecting boundary conditions [Szn84b, Szn84a] and most importantly for the lack of exchangeability of particles due to the spatial dependence of the SDE system (4.1.6). For the convergence of the empirical measure in the mean field limit we also use optimal results for the approximation of empirical measures by i.i.d. samples obtained in [FG13].

The results are summarized in the following two theorems.

Theorem (Theorems 5.2.3, 5.2.4 and 5.2.5)**.** Consider coefficients b, σ Lipschitz in $u \in \mathbb{R}^{d_v}$ and α-Hölder in $x \in Q$. There exists a pathwise unique solution of the SDE system (4.1.6). For any initial data $f_0 \in C^\alpha(Q; \mathscr{P}_2(\mathbb{R}^{d_v}))$, there exists a unique weak solution $f \in C^\alpha\left(Q; \mathscr{P}_2\left(C([0,T]; \mathbb{R}^{d_v})\right)\right)$ of the Fokker–Planck PDE (4.1.9). For any correlation radius $\epsilon > 0$ and for initial data $u(\cdot, 0) \in C^\alpha(Q; L^2(\Omega))$, there exists a pathwise unique solution $\bar{u}^\epsilon \in C^\alpha(Q; L^2(\Omega; C([0,T]; \mathbb{R}^{d_v})))$ of the McKean–Vlasov equation (4.1.8) with colored noise W^ϵ. Furthermore, the single particle law $f(\cdot, x, du) := \mathrm{Law}_{C([0,T];\mathbb{R}^{d_v})}(\bar{u}^\epsilon(x, \cdot))$ is independent of $\epsilon > 0$ and is uniquely characterized as the solution of the Fokker–Planck PDE.

Theorem (Theorems 5.2.6 and 5.2.7)**.** Under the same assumptions, we have the following estimate for the difference between the network (4.1.6), with N clusters at locations $x_1, \ldots, x_N \in Q$ with M neurons each, and the 'McKean–Vlasov neurons' (4.1.8) at those locations x_i

$$\mathbb{E}\left[\sup_{r\in[0,t]} |u^\epsilon_{ik}(r) - \bar{u}^\epsilon_{ik}(r)|^2\right]^{1/2} \leqslant C\,t\left(\frac{1}{N^{\frac{\alpha}{d}}} + \frac{1}{M^{\frac{1}{2}}}\right)\left(1 + \sup_{x\in Q}\mathbb{E}\left[|u_k(x,0)|^2\right]^{1/2}\right),$$

for any $i = 1, \ldots, N$, $k = 1, \ldots, M$ and $t \in [0, T]$, where $C = C(T, b, \sigma, [u(\cdot, 0)]_\alpha)$ and $[u(\cdot, 0)]_\alpha$ denotes the Hölder seminorm of the initial data.

Furthermore, as $M, N \to \infty$, and possibly but not necessarily as $\epsilon \to 0$, the empirical measure (4.1.2) of the network converges to the solutions of the Fokker–Planck equation (4.1.9) with the following rate

$$\sup_{t\in[0,T]} \mathbb{E}\left[\mathcal{W}_1(Q \times \mathbb{R}^{d_v})(f^\epsilon_{N,M}(t), f(t))\right] \leqslant C\left(\frac{1}{N^{\frac{\alpha}{d}}} + \frac{1}{M^{\frac{1}{2}}} + \frac{1}{M^{\frac{1}{4}}}\right), \tag{4.2.1}$$

for any $T > 0$, for a constant $C = C(T, b, \sigma, u(\cdot, 0), Q)$.

We remark that the decay rate $\left(1/N^{\frac{\alpha}{d}} + 1/M^{\frac{1}{2}}\right)$ instead of the usual $1/(MN)^{1/2}$, as typically expected for MN particles, is due to the spatial dependence of the SDE system (4.1.6) and the lack of exchangeability. The same phenomenon yields the rate (4.2.1), which is in general optimal [FG13], and the specific scaling regime between M and N needed to observe nontrivial fluctuations discussed below.

4.2.2 Chapter 6: fluctuations in the large number of particles limit

The main contribution of Chapter 6, containing the work [Cli23a], is to analyze the fluctuations of the generalized neuron network (4.1.6) in the thermodynamic limit as sketched in the discussion (4.1.15)-(4.1.18).

The results are summarized in the following theorem.

Theorem (Theorem 6.3.12)**.** Fix a correlation radius for the noise $\epsilon > 0$, along a scaling regime $M, N \to \infty$ such that $\sqrt{M}N^{-\alpha/d} \to 0$, where α is the Hölder spatial regularity of the coefficients, the fluctuations $\eta^{MN}_t = \sqrt{M}(f_{MN} - f)$ of the empirical measure f_{MN} of the particle system (4.1.6) around the deterministic limit f, that is the solution of the Fokker–Planck equation (4.1.9), satisfy

$$\eta^{MN}_t \to \eta^\epsilon_t \quad \text{in law in } C\left([0, T]; \mathcal{H}\right).$$

Here $\mathcal{H}$ is a suitable Hilbert space of distributions on $Q \times \mathbb{R}^{d_v}$ and η^ϵ_t is the unique solution of the Langevin SDPE

$$\eta^\epsilon_t = \eta_0 + \int_0^t \mathscr{L}_r(f, f)^*[\eta^\epsilon_r]\,dr + G^\epsilon_t,$$

where the operator $\mathscr{L}(f,f)^*$ is essentially the adjoint of the linearization of the generator of the McKean–Vlasov SDE (4.1.8) around the mean field limit f, and η_0 and G_t^ϵ are suitable Gaussian r.v. and process, uniquely characterized by their covariance and quadratic variation respectively.

The methods adopted are deeply based on the approach to fluctuations put forward in Fernandez and Méléard [FM97], Bezandry, Ferland, Fernique and Giroux [BFG93, FFG92] and Hitsuda and Mitoma [HM86]. Our situation presents additional difficulties: unboundedness of the coefficients, the reflecting boundary conditions and of course the 'space extension', i.e. interaction between different mean-field families of coupled neurons across the cortex, ultimately resulting in the lack of exchangeability.

The fluctuations $\eta_t^{MN} = \sqrt{M}(f_{MN} - f)$ are random signed measures in the path space over $Q \times \mathbb{R}^{d_v}$. A first problem is to find a distribution space $C([0,T];\mathcal{H})$ in which both η_t^{MN} and its limit belong, where $\mathcal{H}$ is essentially a vector valued Sobolev space taking values in another weighted Sobolev space of functions satisfying suitable no-flux boundary conditions.

The strategy then revolves around the semimartingale equation (4.1.16) satisfied by η_t^{MN}. Suitable estimates on η_t^{MN} and its martingale part, in different nested distribution spaces, are obtained with martingale arguments. These estimates are then translated into tightness in path space of the processes involved thanks to compact embeddings of the above distribution spaces and to Aldous-like criteria.

A fine analysis based on Sznitman's coupling method [Szn84a] shows that the quadratic variation of the martingale term M_t^{MN} is converging to a deterministic limit, thus uniquely identifying a limiting Gaussian process G_t^ϵ. An adaptation of the classical Lévy central limit theorem proves the convergence of the initial data η_0^{MN}. We conclude passing to the limit in equation (4.1.16) and using the well-posedness of the Langevin SPDE (4.1.17).

Chapter 5

Wellposedness and mean field limit

5.1 Introduction

The discovery of a type of neurons in the brain named grid cells in 2005 [HFM$^+$05] led to a breakthrough in the understanding of the navigational system in mammalian brains, see [MMM17] for an extensive review. These neurons fire as an animal moves around in an open area, enabling the animal to understand its position in space. The grid cell network has commonly been described by deterministic continuous attractor network dynamics through a system of neural field models [ET10, MBJ$^+$06, BF09, CWZ$^+$13], which are based on the classical papers [WC72, WC73, Ama77]. The models can fairly accurately predict what can be observed in experiments. However, the question of how the grid cell network is affected by noise, posed as a challenge in [RRMM16], has been left open.

In [BF12] fundamental limits on how information dissipates in attractor networks of noisy neurons were derived. A different direction pursuing further understanding of the effect of noise on grid cell networks was made in [CHS22] by studying a system of Fokker–Planck-like partial differential equations (PDEs). The system of PDEs was derived by adding noise to the attractor network models in [BF09, CWZ$^+$13] and formally taking the mean field limit. In the present work this limit is rigorously proved. In addition, we derive the limit for more general noise terms, which covers the models considered in [BF12, AB20].

The mean field limit of interacting particle systems has lately received lots of attention in mathematical biology [BCnC11, FI15, CS18], see [JW17] for a survey. The closest result to the analysis presented in this work, shows the mean field limit of a stochastic delayed set of interacting neurons [Tou12]. The system of stochastic differential equations (SDEs) describing interacting grid cells in this work introduces different challenges: boundary conditions imposing positivity of the activity level of the neurons, non-linearity of the firing rate, and coupling between different families of neurons.

The neural model under consideration, which is based on the model in [BF09], can be described as follows. Given space points $x_1, \ldots, x_N \in Q$ in a region Q of the neural cortex, we will consider the following model for the interaction among NM neurons stacked in N columns at locations x_i with M neurons each, where u^β_{ik} represents the activity level with orientation β of the k^{th} neuron at location x_i:

$$
\begin{cases}
u^\beta_{ik}(t)\tau^\beta_i = \tau^\beta_i u^\beta_{ik}(0) + \sigma W^\beta_{ik}(t) - \ell^\beta_{ik}(t) \\
\qquad + \int_0^t \left(-u^\beta_{ik}(r) + \phi\Big(B^\beta(x_i, r) + \frac{1}{4NM} \sum_{\gamma=1}^{4} \sum_{j=1}^{N} \sum_{m=1}^{M} K^\gamma(x_i - x_j) u^\gamma_{jm}(r)\Big) \right) dr, & (5.1.1a) \\
\ell^\beta_{ik}(t) = -\big|\ell^\beta_{ik}\big|(t), \quad \big|\ell^\beta_{ik}\big|(t) = \int_0^t 1_{\{u^\beta_{ik}(r)=0\}} d\big|\ell^\beta_{ik}\big|(r) \quad \text{for } \beta = 1, 2, 3, 4. & (5.1.1b)
\end{cases}
$$

For simplicity, we consider $Q = [0, 1]^d$. The results in this work are easily extended to any bounded open subset $Q \subseteq \mathbb{R}^d$, for any $d \geqslant 1$. Here, for integers $k = 1, \ldots, M$, we have i.i.d. families of random initial conditions $\{u_{ik}(0)\}_{i=1,\ldots,N}$ for each space point x_i in the cortex Q. Moreover, for integers $i = 1, \ldots, N$ and $k = 1, \ldots, M$, we have 4-dimensional Brownian motions $(W^\beta_{ik})_{\beta=1,2,3,4}$, which can also be correlated.

The nonlinear function $\phi : \mathbb{R} \to \mathbb{R}$, representing the firing rate of neurons in the network, is globally Lipschitz, whereas the external inputs $B^\beta : Q \times \mathbb{R} \to \mathbb{R}$ and the interaction kernels $K^\beta : \mathbb{R}^{d_Q} \to \mathbb{R}$ for $\beta = 1, 2, 3, 4$ are only required to be locally bounded functions and α-Hölder continuous in the x variable for some $\alpha \in (0, 1]$. The interaction kernels takes into account the inhibitory/excitatory effect on nearby neurons. A typical choice of the interaction kernel in computational neuroscience [BF09] is given by the so-called Mexican hat function. The relaxation times τ^β_i satisfy the condition $0 < \inf_{i,\beta} \tau^\beta_i \leqslant \sup_{i,\beta} \tau^\beta_i < +\infty$.

Finally, for each i, k and β, the term ℓ^β_{ik} is a finite variation process defined by (5.1.1b) which prevents the activity level u^β_{ik} from taking negative values. Namely, as we can see in its definition, at each time t this process equals the opposite of its total variation $\ell^\beta_{ik}(t) = -\big|\ell^\beta_{ik}\big|(t)$. In turn, the total variation stays constant when $u^\beta_{ik} > 0$ and it increases in the form $\big|\ell^\beta_{ik}\big|(t) = \int_0^t 1_{\{u^\beta_{ik}(r)=0\}} d\big|\ell^\beta_{ik}\big|(r)$ when $u^\beta_{ik} = 0$, so as to push u^β_{ik} away from zero which is being dragged by the other terms at the right hand side of (5.1.1a). The introduction of such terms and constraints is therefore known as imposing *reflecting boundary conditions* and ℓ^β_{ik} is called a *reflection term.* The existence and uniqueness of such a term need of course to be proved and this process is often referred to as the *Skorokhod problem.* Precise details concerning the well-posedness and the construction of the reflection term in our setting are all presented in the seminal papers [LS84, Szn84b] by Lions and Sznitman.

Going back to (5.1.1a), we notice that the argument of ϕ in (5.1.1a) can be rewritten as

$$
\frac{1}{4NM} \sum_{\gamma=1}^{4} \sum_{j=1}^{N} \sum_{m=1}^{M} K^\gamma(x_i - x_j) u^\gamma_{jm}(r) = \int_{Q \times \mathbb{R}^4} \frac{1}{4} \sum_{\gamma=1}^{4} K^\gamma(x_i - y) u^\gamma f_{N,M}(r, dy, du),
$$

by considering the empirical measure associated to these particles, that is

$$f_{N,M}(r,dy,du) = \frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M}\delta_{(x_j,u_{jm}(r))} \quad \text{regarded as a measure on } Q\times\mathbb{R}^4. \tag{5.1.2}$$

Concerning the initial conditions and the form of the noise term in (5.1.1a), from a modelling point of view it is reasonable to assume that, for $k\in\mathbb{N}$, we have i.i.d. families of initial conditions $\big(u_k(x,0)\big)_{x\in Q}$ for each space point x in the cortex Q. Similarly, we assume that, for $k\in\mathbb{N}$, we have independent 4-dimensional space-time white noise terms $\big(W_k(x,t)\big)_{t\geqslant 0,x\in Q}$. Naively, $W_k^{\beta}(x,t)$ is a centered Gaussian random field indexed by $k\in\mathbb{N}$, $\beta=1,2,3,4$, $x\in Q$ and $t\in[0,\infty)$ with covariance

$$\mathbb{E}\Big[W_k^{\beta}(x,t)W_h^{\gamma}(y,s)\Big] = (t\wedge s)\,\delta_0(k-h)\,\delta_0(\beta-\gamma)\,\delta_0(x-y). \tag{5.1.3}$$

Then we can just choose points $x_1,\dots,x_N\in Q$ and set $u_{ik}(0):=u_k(x_i,0)$ and $W_{ik}(t):=W_k(x_i,t)$. As long as we work in a countable setting, this naive construction can be made rigorous upon taking a suitable modification of the W_{ik}'s via the Kolmogorov continuity theorem. We also point out that the way we choose the cloud of points $x_1,\dots x_N\in Q$ is not that important if we are only concerned with the discrete model for fixed M and N. However, to get a nice limiting behaviour as $N,M\to\infty$, it is useful to take these points to be the nodes of a grid of Q whose mesh tends to zero. Precise details on this are given in Section 5.5.

Remark 5.1.1. One should not expect the initial data $(u_k(x,0))_{x\in Q}$ to be independent for different values of x, nor to be equidistributed. Indeed, from the point of view of modelling in neuroscience, $u_k(x,0)$ should be close to $u_k(y,0)$ for x close to y. This fact will have consequences both on the exchangeability properties of the particles u_{ik}, which are expected to be exchangeable in the index k only, and on the rate of convergence towards the limiting behavior.

As we let $M,N\to\infty$ the limiting behaviour should be described by independent copies, in the column index k, of solutions to an associated mean field McKean–Vlasov equation. Namely, the activity level of any neuron located at a point $x\in Q$ should satisfy an equation like

$$\begin{cases} \bar u^{\beta}(x,t)\tau^{\beta}(x) = \tau^{\beta}(x)u^{\beta}(x,0)+\sigma W^{\beta}(x,t)-\bar\ell^{\beta}(x,t) \\ \qquad\qquad + \displaystyle\int_0^t\bigg(-\bar u^{\beta}(x,r)+\phi\Big(B^{\beta}(x,r)+\frac{1}{4}\sum_{\gamma=1}^{4}\int_{Q\times\mathbb{R}^4}K^{\gamma}(x-y)u^{\gamma}f(r,y,du)dy\Big)\bigg)\,dr, \\ \bar\ell^{\beta}(x,t) = -\big|\bar\ell^{\beta}(x,\cdot)\big|(t), \quad \big|\bar\ell^{\beta}(x,\cdot)\big|(t) = \displaystyle\int_0^t 1_{\{\bar u^{\beta}(x,r)=0\}}d\big|\bar\ell^{\beta}(x,\cdot)\big|(r) \quad \text{for } \beta=1,2,3,4, \end{cases} \tag{5.1.4}$$

where we have set $f(t,y,du):=\mathrm{Law}_{\mathbb{R}^4}(\bar u(y,t))$ considered as a measure on $\mathbb{R}^4$ depending on $t\in[0,\infty)$ and $y\in Q$. Notice that in turn this induces a probability measure $f(t,dx,du)$ on $Q\times\mathbb{R}^4$ defined by integration as

$$\int_{Q\times\mathbb{R}^4}\varphi(x,u)f(t,dx,du) := \int_Q\int_{\mathbb{R}^4}\varphi(x,u)f(t,x,du)\,dx \qquad \text{for any } \varphi\in C_b(Q\times\mathbb{R}^4). \tag{5.1.5}$$

For each fixed $x \in Q$ and $\beta = 1, 2, 3, 4$, the finite variation process $\bar{\ell}^\beta(x,t)$ is again the reflection term coming from the Skorokhod problem (see the explanation after equation (5.1.1)) and it ensures that $\bar{u}^\beta(x,t) \geqslant 0$ for every x, t and β. We refer the reader to [Szn84b] for the details about such a process in the context of a classical McKean–Vlasov equation.

Remark 5.1.2. The McKean–Vlasov equation (5.1.4) suffers from a major technical issue. Indeed, formula (5.1.3) does define an $\mathbb{R}$-valued Gaussian random field. However, it is well-known that such a random field cannot be jointly measurable in the x variable and the sample ω. This reflects into lack of x-measurability of the particles $\bar{u}^\beta(x,t)$ and, in turn, into that of the law $f(t,x,du)$, which we need to be Lebesgue integrable. In this work, we resolve this issue by considering ϵ-correlated noise.

Another approach, coming from the theory of mean field games, is to address the issue by introducing a "Fubini extension" of the product probability space $Q \times \Omega$. We refer the reader to [ACL21] and the references therein. However, this approach did not seem to fit our modelling purposes. It allows to regain the x-measurability only with respect to a bigger σ-algebra, *strictly* containing the Lebesgue measurable sets. In turn, the space integral in (5.1.4) would not be taken with respect to the Lebesgue measure, but instead with respect to some exotic extension of this.

A formal application of the Itô formula shows that f, the joint distribution of the activity levels u^β in the four directions β, satisfies the nonlinear Fokker–Planck equation

$$\begin{aligned}&\partial_t f(t,x,u) - \frac{\sigma^2}{2}\sum_{\beta=1}^{4} \frac{1}{\tau^\beta(x)^2}\partial^2_{u^\beta u^\beta} f(t,x,u)\\ &= \sum_{\beta=1}^{4} \frac{1}{\tau^\beta(x)}\partial_{u^\beta}\left(f(t,x,u)\Big(-u^\beta+\phi\Big(B^\beta(x,t)+\frac{1}{4}\sum_{\gamma=1}^{4}\int_{Q\times\mathbb{R}^4} K^\gamma(x-y)v^\gamma f(t,y,dv)dy\Big)\Big)\right),\end{aligned} \tag{5.1.6}$$

in the weak sense, with initial condition $f(0,x,du) = \mathrm{Law}_{\mathbb{R}^4}(u(x,0))$ and subjected to the no-flux boundary conditions, for $\beta = 1, 2, 3, 4$,

$$\phi\Big(B^\beta(x,t)+\frac{1}{4}\sum_{\gamma=1}^{4}\int_{Q\times\mathbb{R}^4} K^\gamma(x-y)v^\gamma f(t,y,dv)dy\Big)f(t,x,u) - \frac{\sigma^2}{2}\frac{1}{\tau^\beta(x)}\frac{\partial}{\partial u^\beta}f(t,x,u)\Big|_{u^\beta=0} = 0,$$

which come from the reflecting boundary conditions at the SDE level.

Remark 5.1.3. It is worth pointing out that equation (5.1.6) would arise as the law of $\bar{u}(x,t)$ even if we set $W(x,t) \equiv B_t$ for every $x \in Q$ for a single Brownian motion B_t, that is if all the particles were affected by the same noise. The same holds for many other choices of $W(x,t)$, and follows immediately from the Itô formula: the effect of the term $W(x,t)$ is only to generate diffusion in the u variable, for fixed x. The choice of noise to consider in (5.1.1a) and (5.1.4) is therefore dictated by modelling purposes only.

Remark 5.1.4. We notice that for each $\beta = 1, 2, 3, 4$, integrating equation (5.1.6) in $\mathbb{R}^3_+$ over the remaining variables u^γ for $\gamma \neq \beta$ and exploiting the boundary conditions, we get the equation satisfied by the marginal distribution $f^\beta(r,y,du^\beta) = \mathrm{Law}_{\mathbb{R}}(\bar{u}^\beta(y,r))$. Namely, we obtain

$$\begin{aligned}
&\partial_t f^\beta(t,x,u^\beta) - \frac{\sigma^2}{2}\frac{1}{\tau^\beta(x)^2}\frac{\partial^2 f^\beta}{(\partial u^\beta)^2}(t,x,u^\beta) \\
&\quad = \frac{1}{\tau^\beta(x)}\partial_{u^\beta}\bigg(f^\beta(t,x,u^\beta)\Big(-u^\beta + \phi\Big(B^\beta(x,t) + \frac{1}{4}\sum_{\gamma=1}^{4}\int_{Q\times\mathbb{R}^4} K^\gamma(x-y)v^\gamma f^\gamma(t,y,dv^\gamma)dy\Big)\Big)\bigg).
\end{aligned} \tag{5.1.7}$$

In particular, we stress the fact that each marginal f^β satisfies an equation involving only the other marginals f^γ, and not the full joint distribution f. On the other hand, if we sum equation (5.1.7) over $\beta = 1,2,3,4$, then we get back equation (5.1.6) above for the decoupled distribution $\tilde{f} := \Pi_{\beta=1}^4 f^\beta$. Thus equation (5.1.6) and the system of equations (5.1.7) for $\beta = 1,2,3,4$ are completely equivalent, at least for decoupled initial data $f_0 = \Pi_{\beta=1}^4 f_0^\beta$. Finally, Theorem 5.2.5 below asserts we have existence and uniqueness for equation (5.1.6). The previous argument then shows that, if we start with decoupled initial data, this structure is preserved: the corresponding solution satisfies $f(t) = \Pi_\beta f^\beta(t)$ for all $t \geqslant 0$. Notice that (5.1.7) is the model formally introduced in [CHS22].

The structure of the work is as follows. The next section introduces the notation and the setting needed for the results. We finish the section by stating the main theorems concerning the mean field limit of (5.1.1) and its extensions. Sections 5.3 and 5.4 focus on the existence and uniqueness of the particle systems, and the associated McKean–Vlasov equations and Fokker–Planck type PDEs. The main core of this work is found in Section 5.5, where we rigorously prove the mean field limit. Section 5.6 adapts previous results on empirical measure error estimates [FG13] to the present setting to provide rates of convergence for the associated empirical measure.

5.2 Preliminaries and main results

5.2.1 Hypotheses and notation

In this section we introduce the hypotheses we assume for our problem. First, we point out that the results presented extend to the more general particle system

$$\begin{cases}
u_{ik}(t) = u_{ik}(0) + \displaystyle\int_0^t b(x_i, r, u_{ik}(r), f_{N,M}(r))\,dr \\
\qquad\qquad + \displaystyle\int_0^t \sigma(x_i, r, u_{ik}(r), f_{N,M}(r))\,dW_{ik}(r) - \ell_{ik}(t), \\
\ell_{ik}^\beta(t) = -|\ell_{ik}^\beta|(t), \quad |\ell_{ik}^\beta|(t) = \displaystyle\int_0^t 1_{\{u_{ik}^\beta(r)=0\}} d|\ell_{ik}^\beta|(r) \quad \text{for } \beta = 1,2,3,4,
\end{cases} \tag{5.2.1}$$

for N columns of M neurons each, located at $x_1, \ldots, x_N$, with general drift term b and diffusion term σ. Here $f_{N,M}(r,dy,du)$ is again the empirical measure associated to the particles

(5.2.1), given by (5.1.2). As before, ℓ_{ik}^{β} is the the reflection term coming from the Skorokhod problem [LS84] forcing $u_{ik}^{\beta}(t) \geqslant 0$ for every $t \geqslant 0$.

The precise details on the shape and hypotheses on b and σ are given here below and they are simply deduced from the properties of the concrete model (5.1.1).

Let $\mathscr{P}(Q \times \mathbb{R}^4)$ denote the set of probability measures on $Q \times \mathbb{R}^4$, for $\beta = 1, 2, 3, 4$ we assume that $b_\beta, \sigma_\beta : Q \times \mathbb{R}^+ \times \mathbb{R}^4 \times \mathscr{P}(Q \times \mathbb{R}^4) \to \mathbb{R}$ take the forms

$$b_\beta(x, r, u, f) = b_0^\beta(x, r, u) + \phi_{b_\beta}\left(\int_{Q\times\mathbb{R}^4} b_1^\beta(x, y, r, u, v)\, f(dy, dv)\right), \tag{5.2.2}$$

$$\sigma_\beta(x, r, u, f) = \sigma_0^\beta(x, r, u) + \phi_{\sigma_\beta}\left(\int_{Q\times\mathbb{R}^4} \sigma_1^\beta(x, y, r, u, v)\, f(dy, dv)\right). \tag{5.2.3}$$

Having in mind the concrete model (5.1.1a), we suppose $b_0^\beta, \sigma_0^\beta : Q \times \mathbb{R}^+ \times \mathbb{R}^4 \to \mathbb{R}$ are measurable, locally bounded, Lipschitz in $u \in \mathbb{R}^4$ uniformly in $x, r \in Q \times \mathbb{R}^+$, and α-Hölder in $x \in Q$ uniformly in $u, r \in \mathbb{R}^4 \times \mathbb{R}^+$. That is

$$\left|b_0^\beta(x, r, u) - b_0^\beta(x', r, u')\right| + \left|\sigma_0^\beta(x, r, u) - \sigma_0^\beta(x', r, u')\right| \leqslant L\left(|x - x'|^\alpha + |u - u'|\right), \tag{5.2.4}$$

$$\left|b_0^\beta(x, r, u)\right| + \left|\sigma_0^\beta(x, r, u)\right| \leqslant C\left(1 + |u|\right), \tag{5.2.5}$$

for all $x, x', u, u', r \in Q^2 \times \left(\mathbb{R}^4\right)^2 \times \mathbb{R}^+$, for suitable constants L and C. Furthermore we take the functions $\phi_{b_\beta}, \phi_{\sigma_\beta} : \mathbb{R} \to \mathbb{R}$ to be globally Lipschitz functions, and thus with sublinear growth. Similarly, the mappings $b_1^\beta, \sigma_1^\beta : Q \times Q \times \mathbb{R}^+ \times \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ are measurable, locally bounded, Lipschitz in $u, v \in \mathbb{R}^4$ uniformly in $x, y, r \in Q^2 \times \mathbb{R}^+$, and α-Hölder in $x, y \in Q$ uniformly in $u, v, r \in \left(\mathbb{R}^4\right)^2 \times \mathbb{R}^+$. That is,

$$\begin{aligned}&\left|b_1^\beta(x, y, r, u, v) - b_1^\beta(x, y, r, u', v')\right| \\ &\quad + \left|\sigma_1^\beta(x, y, r, u, v) - \sigma_1^\beta(x, y, r, u', v')\right| \leqslant L\left(|x - x'|^\alpha + |y - y'|^\alpha + |u - u'| + |v - v'|\right),\end{aligned} \tag{5.2.6}$$

$$\left|b_1^\beta(x, y, r, u, v)\right| + \left|\sigma_1^\beta(x, y, r, u, v)\right| \leqslant C\big(1 + |u| + |v|\big), \tag{5.2.7}$$

for all $x, y, x', y', u, v, u', v', r \in Q^4 \times \left(\mathbb{R}^4\right)^4 \times \mathbb{R}^+$, for suitable constants L and C.

Remark 5.2.1. With the notation just introduced, the starting model (5.1.1a) is recovered by setting

$$\tau^\beta(x) b^\beta(x, r, u, f) = -u^\beta + \phi\left(B^\beta(x, r) + \frac{1}{4}\sum_{\gamma=1}^{4}\int_{Q\times\mathbb{R}^4} K^\gamma(x - y) u^\gamma f(dy, du)\right),\ \tau^\beta(x)\sigma(x, r, u, f) \equiv \sigma.$$

We now consider the limiting McKean–Vlasov system. Taking into account the measurability issues pointed out in Remark 5.1.2, we consider instead equation (5.1.4) with a suitably rescaled ϵ-correlated noise, for some $\epsilon > 0$. In the setting of the general particle system (5.2.1), the equation reads:

$$\begin{cases}\bar{u}^\epsilon(x, t) = u(x, 0) + \displaystyle\int_0^t b(x, r, \bar{u}^\epsilon(x, r), f(r))\, dr + \int_0^t \sigma(x, r, \bar{u}^\epsilon(x, r), f(r))\, dW^\epsilon(x, r) - \bar{\ell}(x, t), \\ \bar{\ell}^\beta(x, t) = -\,|\bar{\ell}^\beta(x, \cdot)|(t), \quad |\bar{\ell}^\beta(x, \cdot)|(t) = \displaystyle\int_0^t 1_{\{(\bar{u}^\epsilon)^\beta(x, r) = 0\}}\, d|\bar{\ell}^\beta(x, \cdot)|(r) \quad \text{for } \beta = 1, 2, 3, 4,\end{cases} \tag{5.2.8}$$

where $f(r, y, du) = \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(y, r))$ is viewed as a measure on $\mathbb{R}^4$, and $f(r) = f(r, dx, du)$ the induced probability measure defined by (5.1.5) on $Q \times \mathbb{R}^4$. Similarly, the reflection term $\bar{\ell}(x, t)$ still ensures $(\bar{u}^\epsilon)^\beta(x, t) \geqslant 0$ for each x, t and β (see again [Szn84b]). Here $W^\epsilon : \Omega \times \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^4$ is a 4-dimensional Gaussian random field with independent components $\beta = 1, 2, 3, 4$, zero mean and covariance

$$\mathbb{E}\left[W^{\epsilon,\beta}(x,t)W^{\epsilon,\beta}(y,s)\right] = (t \wedge s)\, C_\rho\, \epsilon^d \int_{\mathbb{R}^d} \rho_\epsilon(z-x)\rho_\epsilon(z-y)\, dz, \text{ for } C_\rho = \left(\int_{\mathbb{R}^d} \rho(z)^2\, dz\right)^{-1}, \quad (5.2.9)$$

where $\rho : \mathbb{R}^d \to [0, 1]$ is a radial mollifier supported in the unitary ball, and ρ_ϵ the ϵ-rescaled version. Such a process $W^{\epsilon,\beta}$ can for example be obtained by convolution and rescaling from a "mathematically rigorous" space-time white noise (see e.g. [DPZ92]). That is, a distribution valued process $W : \Omega \times \mathbb{R}^+ \to \mathcal{S}'(\mathbb{R}^d)$ such that, for $\varphi \in \mathcal{S}(\mathbb{R}^d)$, the processes $\langle W_t, \varphi\rangle$ are jointly Gaussian with covariance

$$\mathbb{E}\left[\langle W_t, \varphi\rangle \langle W_s, \psi\rangle\right] = (t \wedge s) \int_{\mathbb{R}^d} \varphi(z)\, \psi(z)\, dz.$$

Then, for $\beta = 1, 2, 3, 4$ and independent copies of such a white noise, one defines

$$W^{\epsilon,\beta}(x,t) := C_\rho^{\frac{1}{2}}\, \epsilon^{\frac{d}{2}}\, \langle W_t, \rho_\epsilon(\cdot - x)\rangle. \quad (5.2.10)$$

For future reference, we highlight some of the properties of W^ϵ. First, from (5.2.9) we have that $\mathbb{E}[W^{\epsilon,\beta}(x,t)W^{\epsilon,\gamma}(x,s)] = \delta_0(\beta - \gamma)\, t \wedge s$. Thus, for fixed x, the process $t \mapsto W^\epsilon(x, t)$ is a 4-dimensional Brownian motion. Similarly, from $\sup(\rho) \subseteq B(0, 1)$ it follows that

$$\mathbb{E}\left[W^\epsilon(x,t)W^\epsilon(y,s)\right] = 0 \text{ if } |x - y| > 2\epsilon.$$

Hence the processes $W^\epsilon(x, t)$ and $W^\epsilon(y, t)$ are independent for $|x - y| > 2\epsilon$. Furthermore, using (5.2.9) one computes

$$\begin{aligned}
\mathbb{E}\left[|W^\epsilon(x,t) - W^\epsilon(y,s)|^2\right] &\leqslant C\, \mathbb{E}\left[|W^\epsilon(x,t) - W^\epsilon(x,s)|^2 + |W^\epsilon(x,s) - W^\epsilon(y,s)|^2\right] \\
&\leqslant C\left(|t-s| + s\, C_\rho\, \epsilon^d \int_{\mathbb{R}^d} \left(\rho_\epsilon(z-x) - \rho_\epsilon(z-y)\right)^2\, dz\right) \\
&\leqslant C\left(|t-s| + \frac{|x-y|^2}{\epsilon^2}\right),
\end{aligned}$$

for a constant $C = C(\rho)$. Similar estimates hold for any higher moment $p \geqslant 2$ and the Kolmogorov continuity theorem ensures the existence of a suitable modification of W^ϵ with continuous trajectories in both x and t. In particular, we have that W^ϵ is jointly measurable in $(x, t) \in \mathbb{R}^d \times \mathbb{R}^+$ and in the sample path $\omega \in \Omega$. Finally, for any $x, y \in \mathbb{R}^d$, a direct computation shows that the quadratic variation of the martingale $W^{\epsilon,\beta}(x, t) - W^{\epsilon,\beta}(y, t)$ satisfies

$$\begin{aligned}
\left[W^{\epsilon,\beta}(x,\cdot) - W^{\epsilon,\beta}(y,\cdot)\right]_t &= t\, C_\rho\, \epsilon^d \int_{\mathbb{R}^d} \left(\rho_\epsilon(z-x) - \rho_\epsilon(z-y)\right)^2\, dz, \\
&\leqslant t\, C\, \frac{|x-y|^2}{\epsilon^2},
\end{aligned} \quad (5.2.11)$$

for a constant $C = C(\rho)$.

Finally, $f(r, y, du) = \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(y, r))$ should solve the associated nonlinear Fokker–Planck equation with no-flux boundary conditions,

$$\begin{cases} \partial_t f(t,x,u) + \nabla_u \cdot \Big(b(x,t,u,f(t)) f(t,x,u)\Big) = \dfrac{1}{2} \displaystyle\sum_{\beta=1}^{4} \frac{\partial^2}{\partial u^\beta \partial u^\beta} \Big(\sigma_\beta(x,t,u,f(t))^2 f(t,x,u)\Big), \\ b_\beta(t,x,u,f(t)) f(t,x,u) - \dfrac{1}{2} \dfrac{\partial}{\partial u^\beta} \Big(\sigma_\beta(x,t,u,f(t))^2 f(t,x,u)\Big) \Big|_{u^\beta = 0} = 0 \quad \text{for } \beta = 1,2,3,4, \end{cases} \tag{5.2.12}$$

in the weak sense.

Let us now see how (5.2.2)–(5.2.3) and the assumptions (5.2.4)–(5.2.7) translate into Hölder, Lipschitz and sublinear growth properties of the actual drift and diffusion terms. First, notice that for any fixed $f \in \mathscr{P}(Q \times \mathbb{R}^4)$ the mappings

$$Q \times \mathbb{R}^+ \times \mathbb{R}^4 \to \mathbb{R} \quad \Big| \quad x, r, u \mapsto b_\beta(x,r,u,f),\ \sigma_\beta(x,r,u,f), \tag{5.2.13}$$

are easily seen to be α-Hölder in x, Lipschitz and with sublinear growth in u, uniformly in r. Next, given a Banach space X with norm $|\cdot|_X$ and a positive integer $m \in \mathbb{N}$, let us denote by $\mathscr{P}_m(X)$ the space of probability measures on X with finite mth moments, endowed with the mth order Wasserstein distance (see e.g. [Vil03]),

$$\mathcal{W}_m(X)(P_1, P_2) := \inf_{\pi \in \Pi(P_1, P_2)} \left\{ \int_X |x - y|_X^m \, \pi(dx, dy) \right\}^{\frac{1}{m}}, \tag{5.2.14}$$

where $\Pi(P_1, P_2)$ denotes the set of probability measures on $X \times X$ with marginals P_1 and P_2. When X is clear from the context we write $\mathcal{W}_m(P_1, P_2)$. Let $L^\infty(Q; \mathscr{P}_m(X))$ be the space of measurable functions $f : Q \to \mathscr{P}_m(X)$ such that

$$|f|_{L^\infty(Q;\mathscr{P}_m(X))} = \sup_{y \in Q} \Big(\int_X |x|^m \, f(y, dx) \Big)^{\frac{1}{m}} < \infty,$$

endowed with the distance

$$d_{L^\infty(Q;\mathscr{P}_m(X))}(f, g) = \sup_{y \in Q} \mathcal{W}_m(X)\big(f(y, dx), g(y, dx)\big).$$

Assume $f, g \in L^\infty(Q; \mathscr{P}_m(\mathbb{R}^4))$. Then we can identify them as elements in $\mathscr{P}(Q \times \mathbb{R}^4)$ by their actions on test functions

$$\int_{Q \times \mathbb{R}^4} \psi(y, s) \, f(dy, ds) := \int_Q \int_{\mathbb{R}^4} \psi(y, s) \, f(y, ds) \, dy \qquad \forall \psi \in C_b(Q \times \mathbb{R}^4),$$

and similarly for g. Now, using the structure (5.2.2)–(5.2.3), the Hölder, Lipschitz and sublinear growth properties of b_i^β and σ_i^β for $i = 0, 1$, and Hölder's inequality, it is straightforward to prove the following lemma.

Lemma 5.2.2. In the setting outlined above, and under the assumptions on b and σ,

$$\begin{aligned} \big|b_\beta(x,r,u,f) - b_\beta(x',r,u',f')\big| + \big|\sigma_\beta(x,r,u,f) - \sigma_\beta(x',r,u',f')\big| & \\ \leqslant L \left(|x - x'|^\alpha + |u - u'| + d_{L^\infty(Q;\mathscr{P}_m(\mathbb{R}^4))}(f, f') \right), & \\ \big|b_\beta(x,r,u,f)\big| + \big|\sigma_\beta(x,r,u,f)\big| \leqslant C \left(1 + |u| + |f|_{L^\infty(Q;\mathscr{P}_m(\mathbb{R}^4))} \right), & \end{aligned}$$

for all $x, x', u, u', r \in Q^2 \times (\mathbb{R}^4)^2 \times \mathbb{R}^+$ and all $f, f' \in L^\infty(Q; \mathscr{P}_m(\mathbb{R}^4))$, for suitable constants L and C.

5.2.2 Main results

We now present the main results of this work. The theorems are stated for the general models (5.2.1), (5.2.8), and (5.2.12). First we present a result on existence and uniqueness of the particle systems, which is proved in Section 5.3.

Theorem 5.2.3 (Strong existence and uniqueness for the particle systems)**.** Assume that the initial data satisfies $\sup_{1\leqslant i\leqslant N}\sup_{1\leqslant k\leqslant M}\mathbb{E}[|u_{ik}(0)|^2] < +\infty$. Then, under assumptions (5.2.2)–(5.2.7) on the coefficients, there exists a pathwise unique solution of the particle system (5.2.1).

Next we state the theorems on well-posedness of the McKean–Vlasov equations and the associated PDE. The following two results are proved in Section 5.4.

Theorem 5.2.4 (Strong existence and uniqueness of the McKean–Vlasov equation)**.** Under the assumptions (5.2.2)–(5.2.7) on the coefficients, for any initial data $u(\cdot, 0) \in L^\infty(Q; L^2(\Omega))$, and for any $\epsilon > 0$, there exists a pathwise unique solution $u^\epsilon \in L^\infty(Q; L^2(\Omega; C([0,T]; \mathbb{R}^4)))$ of the McKean–Vlasov equation (5.2.8). Moreover, for every $T < \infty$,

$$\sup_{x\in Q}\mathbb{E}\left[\sup_{t\in[0,T]}|u^\epsilon(x,t)|^2\right] \leqslant C\left(1+\sup_{x\in Q}\mathbb{E}[|u(x,0)|^2]\right), \tag{5.2.15}$$

for a constant $C = C(T, b, \sigma)$. Finally, if the initial data satisfies $u(\cdot, 0) \in C^\alpha(Q; L^2(\Omega))$ for some $\alpha \in (0,1)$, then $u^\epsilon \in C^\alpha(Q; L^2(\Omega; C([0,T]; \mathbb{R}^4)))$.

Theorem 5.2.5 (Well-posedness of the non-linear Fokker–Planck equation)**.** Under the assumptions (5.2.2)–(5.2.7) on the coefficients, for any initial data $f_0(x, du) \in L^\infty(Q; \mathscr{P}_2(\mathbb{R}^4))$, there exists a weak solution $f(x, t, du) \in L^\infty(Q; C([0,\infty); \mathscr{P}_2(\mathbb{R}^4)))$ of the non-linear Fokker–Planck equation (5.2.12). If $|\sigma(x,t,u,g)| \geqslant c > 0$ for every x, t, u and g, the solution is also unique. The map f is uniquely characterized as $f(t, x, du) = \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(x,t))$ for any arbitrary $\epsilon > 0$. Moreover, for each fixed $x \in Q$ and for any time $T > 0$, the restriction $f(x,t,du)|_{t\in[0,T]}$ can be seen as a probability measure on the space $C([0,T]; \mathbb{R}^4)$ of continuous paths, and it satisfies

$$\sup_{x\in Q}\int_{C([0,T];\mathbb{R}^4)}\sup_{t\in[0,T]}|v(t)|^2 f(x,dv) \leqslant C\left(1+\sup_{x\in Q}\mathbb{E}[|u(x,0)|^2]\right), \tag{5.2.16}$$

for a constant $C = C(T, b, \sigma)$, where $v \in C([0,T]; \mathbb{R}^4)$. Finally, if $f_0 \in C^\alpha(Q; \mathscr{P}_2(\mathbb{R}^4))$, then $f \in C^\alpha\left(Q; \mathscr{P}_2\left(C([0,T]; \mathbb{R}^4)\right)\right)$, that is to say

$$\mathcal{W}_2\left(C([0,T];\mathbb{R}^4)\right)(f(x,\cdot), f(y,\cdot)) \leqslant C\,|x-y|^\alpha \quad \forall x, y \in Q, \tag{5.2.17}$$

for a constant $C = C(T, b, \sigma, f_0)$.

We finally present two statements concerning the convergence of the particle system towards the limiting model as $M, N \to \infty$. The setting is the following. For $k \in \mathbb{N}$, let $(W_k(x,t))_{x\in Q, t\geqslant 0}$ be independent 4-dimensional space-time white noise terms over $Q \times [0,\infty)$, which we then convolve and rescale to obtain W_k^ϵ as in formula (5.2.10). Similarly, let $u_k(\cdot, 0) \in C^\alpha(Q; L^2(\Omega))$ be i.i.d. families of random initial conditions along the cortex Q, and let them be independent of all the white noise terms. Finally, let X_i for $i = 1, \dots, N$ be points on a equispaced grid on Q, with squares of sidelength $N^{\frac{1}{d}}$. More details on the setting and the proofs of the results are given in Sections 5.5 and 5.6.

Theorem 5.2.6 (Mean squared error estimates for actual particles vs. McKean–Vlasov particles)**.** In the setting outlined above and in Theorems 5.2.3 and 5.2.4, for any $N, M \in \mathbb{N}$, let $u_{ik}^\epsilon(\cdot)$ be the solution of the particle system (5.2.1), with initial data $u_{ik}(0) := u_k(X_i, 0)$ and noise terms $W_{ik}(t) := W_k^\epsilon(X_i, t)$. For each $k \in \mathbb{N}$, let $\bar{u}_k^\epsilon(x,t)$ be the solution of the McKean–Vlasov equation (5.2.8), with initial data $(u_k(x,0))_{x\in Q}$ and rescaled ϵ-correlated noise $(W_k^\epsilon(x,t))_{x\in Q, t\geqslant 0}$. For each $i \in \mathbb{N}$, denote $\bar{u}_{ik}^\epsilon(t) := \bar{u}_k^\epsilon(X_i, t)$. Then, for any $T > 0$,

$$\mathbb{E}\left[\sup_{r\in[0,t]} |u_{ik}^\epsilon(r) - \bar{u}_{ik}^\epsilon(r)|^2\right]^{1/2} \leqslant C\, t \left(\frac{1}{N^{\frac{\alpha}{d}}} + \frac{1}{M^{\frac{1}{2}}}\right)\left(1 + \sup_{x\in Q} \mathbb{E}\left[|u_k(x,0)|^2\right]^{1/2}\right), \quad (5.2.18)$$

for any $i = 1, \dots, N$, $k = 1, \dots, M$ and $t \in [0,T]$, where $C = C(T, \rho, b, \sigma, [u(\cdot,0)]_\alpha)$ and $[u(\cdot,0)]_\alpha$ denotes the Hölder seminorm of $u(\cdot, 0)$.

We notice that the decay has rate $\left(1/N^{\frac{\alpha}{d}} + 1/M^{\frac{1}{2}}\right)$ instead of the usual $1/(MN)^{1/2}$, as we might expect according to classical results in mean field theory [Szn91] since we have MN particles. As anticipated in Remark 5.1.1, this phenomenon goes back to the fact that the particles u_{ik} are exchangeable in the second index only. Hence, what we will get is a mean field limit in the column index k, but a Riemann sum type convergence in the space index i. This phenomenon will be made clear when we perform the computations in Section 5.5.

We also remark that the ratio between ϵ and $N^{-1/d}$ in Theorem 5.2.6, that is between the correlation radius of the noise and the distance among the neuron locations x_i, is completely arbitrary and the decay rate in (5.2.18) is independent of this. The choice of this ratio is purely dictated by modelling arguments, namely by the correlation strength we want for the noise sensed by two nearby neurons, which for example can be taken to be zero.

Finally we translate the previous result about convergence of particles to the level of laws.

Theorem 5.2.7 (Rate of convergence for the empirical measure)**.** In the setting of Theorem 5.2.6, let

$$f_{N,M}^\epsilon(t, dx, du) = \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \delta_{\left(X_j, u_{jm}^\epsilon(t)\right)}$$

be the empirical measure on $Q \times \mathbb{R}^4$ associated with the particle system (5.2.1). Let $f(t,x,du)$ be the unique solution of the Fokker–Planck equation (5.2.12) and consider the induced probability measure $f(t,dx,du)$ on $Q \times \mathbb{R}^4$ given by (5.1.5). Then, as $M, N \to \infty$, and possibly but not necessarily as $\epsilon \to 0$, $f^{\epsilon}_{N,M}(t,dx,du)$ converges to $f(t,dx,du)$ in the Wasserstein distance in the sense

$$\sup_{t\in[0,T]} \mathbb{E}\left[\mathcal{W}_1(Q\times\mathbb{R}^4)(f^{\epsilon}_{N,M}(t),f(t))\right] \leqslant C\left(1+\sup_{x\in Q}\mathbb{E}\left[|u_k(x,0)|^2\right]^{\frac{1}{2}}\right)\left(\frac{1}{N^{\frac{\alpha}{d}}}+\frac{1}{M^{\frac{1}{2}}}+\frac{1}{M^{\frac{1}{4}}}\right),$$

for any $T>0$, for $C = C(T,\rho,b,\sigma,[u(\cdot,0)]_{\alpha},Q)$.

5.3 Strong existence and uniqueness for the particle systems

In this section we establish strong existence and uniqueness for the particle system (5.2.1), thus proving Theorem 5.2.3. The proof is based on a classical contraction argument and a crucial observation about the reflection term ℓ_{ik}.

Proof of Theorem 5.2.3. Fix $N,M\in\mathbb{N}$. Take a probability space $(\Omega,\mathcal{F},\mathbb{P})$ supporting the initial conditions $u_{ik}(0):\Omega\to\mathbb{R}^4$ and the 4-dimensional Brownian motions $W_{ik}(t)$ for $i=1,\dots,N$ and $k=1,\dots,M$. Suppose $\mathbb{E}\left[|u_{ik}(0)|^2\right]<\infty$ for all i and k. For any $T>0$ let us define the Banach space

$$H^2_T := \left\{\text{continuous adapted processes } Y_t:\Omega\to\mathbb{R}^4 \quad \text{with} \quad \mathbb{E}\left[\sup_{t\in[0,T]}|Y_t|^2\right]^{\frac{1}{2}}<\infty\right\}, \tag{5.3.1}$$

endowed with the norm $\|Y_\cdot\| := \mathbb{E}\big[\sup_{t\in[0,T]}|Y_t|^2\big]^{\frac{1}{2}}$, and then consider the product space $(H^2_T)^{NM}$ equipped with the product norm.

Define $F:(H^2_T)^{NM}\to(H^2_T)^{NM}$ by sending an element $v_{ik}\in(H^2_T)^{NM}$ to the pathwise solutions $\tilde{v}_{ik}$ of the SDEs with reflecting boundary conditions, for $i=1,\dots,N$ and $k=1,\dots,M$,

$$\begin{cases} \tilde{v}_{ik}(t) = u_{ik}(0) + \displaystyle\int_0^t b(x_i,r,v_{ik}(r),f^{v}_{N,M}(r))\,dr \\ \qquad\qquad + \displaystyle\int_0^t \sigma(x_i,r,v_{ik}(r),f^{v}_{N,M}(r))\,dW_{ik}(r) - \ell^{v}_{ik}(t), \\ \ell^{v,\beta}_{ik}(t) = -\,|\ell^{v,\beta}_{ik}|(t), \quad |\ell^{v,\beta}_{ik}|(t) = \displaystyle\int_0^t 1_{\{\tilde{v}_{ik}(r)=0\}}d|\ell^{v,\beta}_{ik}|(r) \quad \text{for } \beta=1,2,3,4. \end{cases} \tag{5.3.2}$$

where we define

$$f^{v}_{N,M}(t) = \frac{1}{NM}\sum_{j=1}^{N}\sum_{m=1}^{M}\delta_{(x_j,v_{jm}(t))}.$$

Under the hypotheses (5.2.2)–(5.2.7) on b and σ, and by straightforward modifications of the setting and the proofs in [Szn84b, LS84], strong existence and uniqueness can be established for the SDEs (5.3.2) with initial data with bounded second moments. Moreover, for initial data with $\mathbb{E}\big[|u_{ik}(0)|^2\big]<\infty$ and data $v_{ik}\in H^2_T$, the solutions $\tilde{v}_{ik}$ belong to H^2_T.

We want to find T small enough so that the map F is a contraction. Take two elements $u_{ik},v_{ik}\in(H^2_T)^{NM}$, and consider $\tilde{u}_{ik}=F(u_{ik})$ and $\tilde{v}_{ik}=F(v_{ik})$. We apply Itô formula to

$|\tilde u_{ik}-\tilde v_{ik}|^2$ and exploit the respective equations (5.3.2) to get

$$\begin{aligned}
&|\tilde u_{ik}(t)-\tilde v_{ik}(t)|^2\\
&= 2\int_0^t (\tilde u_{ik}(r)-\tilde v_{ik}(r))\big(b(x_i,r,u_{ik}(r),f^u_{N,M}(r))-b(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big)\,dr\\
&\quad+2\int_0^t (\tilde u_{ik}(r)-\tilde v_{ik}(r))\big(\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big)\,dW_{ik}(r)\\
&\quad+2\int_0^t (\tilde u_{ik}-\tilde v_{ik})\big(d\ell^v_{ik}(r)-d\ell^u_{ik}(r)\big)+\int_0^t \big(\sigma(x_i,r,u_{ik},f^u_{N,M})-\sigma(x_i,r,v_{ik},f^v_{N,M})\big)^2\,dr,
\end{aligned}\tag{5.3.3}$$

Exploiting the very definition of the reflection terms ℓ^u_{ik} and ℓ^v_{ik} shows that the third term on the right hand side of (5.3.3) is negative. Indeed, we use the second line in (5.3.2) to expand this term as

$$\begin{aligned}
\int_0^t (\tilde u_{ik}-\tilde v_{ik})\big(d\ell^v_{ik}(r)-d\ell^u_{ik}(r)\big) = \sum_{\beta=1}^4 \bigg(&\int_0^t (\tilde u^\beta_{ik}-\tilde v^\beta_{ik})1_{\{\tilde u_{ik}(r)=0\}}d|\ell^{u,\beta}_{ik}|(r)\\
&+\int_0^t (\tilde v^\beta_{ik}-\tilde u^\beta_{ik})1_{\{\tilde v_{ik}(r)=0\}}d|\ell^{v,\beta}_{ik}|(r)\bigg).
\end{aligned}$$

Since the reflecting boundary conditions ensure that $\tilde u^\beta_{ik},\tilde v^\beta_{ik}\geqslant 0$, we see that all the integrals in the sum on the right hand side are negative, since the integrands are.

Now we drop the third term in (5.3.3), take the supremum in time and apply the expectation to get

$$\begin{aligned}
&\mathbb{E}\bigg[\sup_{t\in[0,T]}|\tilde u_{ik}(t)-\tilde v_{ik}(t)|^2\bigg]\\
&\leqslant\bigg(\int_0^T \mathbb{E}\big[\,|\tilde u_{ik}-\tilde v_{ik}|\,\big|b(x_i,r,u_{ik}(r),f^u_{N,M}(r))-b(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big|\,\big]\,dr\\
&\quad+\mathbb{E}\bigg[\sup_{t\in[0,T]}\Big(\int_0^t(\tilde u_{ik}-\tilde v_{ik})\big(\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big)\,dW_{ik}(r)\Big)\bigg]\\
&\quad+\int_0^T \mathbb{E}\Big[\,\big|\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big|^2\,\Big]\,dr\bigg).
\end{aligned}\tag{5.3.4}$$

The second term on the right hand side is handled with the Burkholder–Davis–Gundy inequality and with Hölder's inequality:

$$\begin{aligned}
&\mathbb{E}\bigg[\sup_{t\in[0,T]}\Big(\int_0^t(\tilde u_{ik}-\tilde v_{ik})\big(\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big)\,dW_{ik}(r)\Big)\bigg]\\
&\leqslant\mathbb{E}\bigg[\Big(\int_0^T|\tilde u_{ik}-\tilde v_{ik}|^2\,\big|\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big|^2\,dr\Big)^{\frac12}\bigg]\\
&\leqslant\mathbb{E}\bigg[\sup_{t\in[0,T]}|\tilde u_{ik}-\tilde v_{ik}|\Big(\int_0^T\big|\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big|^2\,dr\Big)^{\frac12}\bigg]\\
&\leqslant\frac12\mathbb{E}\bigg[\sup_{t\in[0,T]}|\tilde u_{ik}-\tilde v_{ik}|^2\bigg]\\
&\qquad+\frac12\mathbb{E}\bigg[\int_0^T\big|\sigma(x_i,r,u_{ik}(r),f^u_{N,M}(r))-\sigma(x_i,r,v_{ik}(r),f^v_{N,M}(r))\big|^2\,dr\bigg].
\end{aligned}\tag{5.3.5}$$

Then we absorb the first term on the right hand side of (5.3.5) into the left hand side of

(5.3.4) to get, for C a numeric constant,

$$
\begin{aligned}
&\mathbb{E}\Bigg[\sup_{t\in[0,T]} |\tilde{u}_{ik}(t) - \tilde{v}_{ik}(t)|^2\Bigg] \\
&\leqslant C\Bigg(\int_0^T \mathbb{E}\big[\,|\tilde{u}_{ik} - \tilde{v}_{ik}|\,\big|b(x_i, r, u_{ik}(r), f^u_{N,M}(r)) - b(x_i, r, v_{ik}(r), f^v_{N,M}(r))\big|\,\big]\,dr \\
&\quad + \int_0^T \mathbb{E}\Big[\,\big|\sigma(x_i, r, u_{ik}(r), f^u_{N,M}(r)) - \sigma(x_i, r, v_{ik}(r), f^v_{N,M}(r))\big|^2\,\Big]\,dr\Bigg).
\end{aligned}
\tag{5.3.6}
$$

Now, we use the structure (5.2.2)–(5.2.3) and the Lipschitz properties (5.2.4)–(5.2.6) of b and σ, the definition of f^u_{MN} and f^v_{MN}, and applications of Hölder's inequality to get, for $C = C(b, \sigma)$:

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{u}_{ik}(t) - \tilde{v}_{ik}(t)|^2\right] \leqslant C\Bigg(&\int_0^T \mathbb{E}\left[|\tilde{u}_{ik}(r) - \tilde{v}_{ik}(r)|^2\right]\,dr \\
&+ \frac{1}{MN}\sum_{j=1}^N \sum_{m=1}^M \int_0^T \mathbb{E}\left[|u_{jm}(r) - v_{jm}(r)|^2\right]\,dr\Bigg).
\end{aligned}
$$

Then, we exploit Grönwall's lemma to get, for $C = C(T, b, \sigma)$:

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{u}_{ik}(t) - \tilde{v}_{ik}(t)|^2\right] &\leqslant C\frac{1}{MN}\sum_{j=1}^N \sum_{m=1}^M \int_0^T \mathbb{E}\left[|u_{jm}(r) - v_{jm}(r)|^2\right]\,dr \\
&\leqslant TC\frac{1}{MN}\sum_{j=1}^N \sum_{m=1}^M \mathbb{E}\left[\sup_{t\in[0,T]} |u_{jm}(t) - v_{jm}(t)|^2\right].
\end{aligned}
$$

Finally we sum over $i = 1, \ldots, N$ and $k = 1, \ldots, M$. In conclusion, by taking another time $T^* < T$ small enough with respect to $C = C(T, b, \sigma)$, we find that the map $F : (H^2_{T^*})^{NM} \to (H^2_{T^*})^{NM}$ is indeed a contraction. The unique fixed point $u_{ik} = F(u_{ik}) \in H^2_{T^*}$ is then the (pathwise unique) solution on $[0, T^*]$. We conclude by gluing solutions on subsequent intervals $[nT^*, (n+1)T^*]$ up to $[0, \infty)$. □

5.4 Well-posedness of the limiting McKean–Vlasov SDEs and PDE

In this section we analyze the limiting model for the particle system (5.2.1), that is the McKean–Vlasov equation (5.2.8) and the nonlinear Fokker–Planck equation (5.2.12). In particular, Theorems 5.2.4 and 5.2.5 about existence and uniqueness for these equations will be proved using a contraction argument.

Let us define the functional setting for the contraction argument. The Banach space $H^2_T := L^2(\Omega$ is defined as in (5.3.1). For any $T > 0$ we shall also consider the complete metric space $C^2_T := C([0, T]; \mathscr{P}_2(\mathbb{R}^4))$ of continuous functions with values in the complete metric space $(\mathscr{P}_2(\mathbb{R}^4), \mathcal{W}_2(\mathbb{R}^4))$, where $\mathcal{W}_2$ is the Wasserstein distance (5.2.14), endowed with the supremum distance $d_{C_T^2}(f,g) = \sup_{t\in[0,T]} \mathcal{W}_2(f(t),g(t))$. Finally, we will employ the Banach space $\mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})$ of bounded measurable maps $Q \to H^2_T$ endowed with the norm

$$|u|_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})} := \sup_{x\in Q} |u(x,\cdot)|_{H^2_T} = \sup_{x\in Q} \mathbb{E}\left[\sup_{t\in[0,T]} |u(x,t)|^2\right]^{\frac{1}{2}}.$$

Similarly, we also make use of the space $\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$. Notice that despite $C^2_T = C([0,T];\mathscr{P}_2(\mathbb{R}^4))$ not being a vector space, it still makes sense to say that a function $f : Q \to C^2_T$ is bounded by taking an arbitrary point $P_0 \in C^2_T$ and imposing $\sup_{x\in Q} d_{C^2_T}(f(x),P_0) < \infty$. For simplicity, we take $P_0(t) \equiv \delta_0$ the function (in t) identically equal to $\delta_0 \in \mathscr{P}_2(\mathbb{R}^4)$ — the Dirac mass centered at zero. With abuse of notation, we denote

$$|f|_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})} := \sup_{x\in Q} d_{C^2_T}(f(x),\delta_0) = \sup_{x\in Q}\sup_{t\in[0,T]} \Big(\int_{\mathbb{R}^4} |v|^2 f(t,x,dv)\Big)^{\frac{1}{2}}.$$

Then $\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ is a complete metric space endowed with the distance $d_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})}(f,g) := \sup_{x\in Q} d_{C^2_T}(f(x),g(x))$.

Let us now introduce the maps yielding the contraction. We are interested in the composition

$$\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T}) \xrightarrow{S^\epsilon} \mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T}) \xrightarrow{L} \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T}). \tag{5.4.1}$$

The map L sends an element $u \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})$ to its bounded-in-space and continuous-in-time law on $\mathbb{R}^4$. That is to say $L[u](x,\cdot) \in C([0,T];\mathscr{P}_2(\mathbb{R}^4))$ is given by $L[u](x,t) = \mathrm{Law}_{\mathbb{R}^4}(u(x,t))$ for each $x \in Q$ and $t \in [0,T]$. A direct computation gives

$$\begin{aligned}\sup_{x\in Q} d_{C^2_T}(L[u](x),\delta_0) &= \sup_{x\in Q}\sup_{t\in[0,T]} \mathcal{W}_2(L[u](x,t),\delta_0)\\ &= \sup_{x\in Q}\sup_{t\in[0,T]} \mathbb{E}\left[|u(x,t)|^2\right]^{\frac{1}{2}} \leqslant \sup_{x\in Q}\mathbb{E}\left[\sup_{t\in[0,T]}|u(x,t)|^2\right]^{\frac{1}{2}} = |u|_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})} < \infty,\end{aligned}$$

and $L[u]$ is indeed an element of $\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$.

The map S^ϵ is defined by sending an element $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ to the solutions $(S^\epsilon[f](x,t))_{t\geqslant 0}$ of the following SDEs with reflecting boundary conditions: for each fixed $x \in Q$

$$\begin{cases} S^\epsilon[f](x,t) = u(x,0) + \displaystyle\int_0^t b(x,r,S^\epsilon[f](x,r),f(r))\,dr \\ \qquad\qquad + \displaystyle\int_0^t \sigma(x,r,S^\epsilon[f](x,r),f(r))\,dW^\epsilon(x,t) - \ell[f](x,t), \\ \ell^\beta[f](x,t) = -\,|\ell^\beta[f](x,\cdot)|(t), \\ |\ell^\beta[f](x,\cdot)|(t) = \displaystyle\int_0^t 1_{\{S^\epsilon[f]^\beta(x,r)=0\}} d|\ell^\beta[f](x,\cdot)|(r), \;\; \beta = 1,2,3,4, \end{cases} \tag{5.4.2}$$

where $u(\cdot,0) \in L^\infty(Q;L^2(\Omega)$ is the initial condition for the McKean–Vlasov equation. Notice that we slightly abuse notation since we identify an element $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ with the time dependent probability measure $f(t,dx,du)$ on $Q \times \mathbb{R}^4$ defined by

$$\int_{Q\times\mathbb{R}^4} \varphi(x,u) f(t,dx,du) := \int_Q\int_{\mathbb{R}^4} \varphi(x,u) f(t,x,du)\,dx \qquad \text{for any} \quad \varphi \in C_b(Q\times\mathbb{R}^4).$$

Standard theory of SDEs with reflecting boundary conditions [Szn84b] ensures that, for each fixed $x \in Q$, equation (5.4.2) has a pathwise unique solution. Indeed, owing to the conditions (5.2.2)–(5.2.7) on b and σ, for fixed $x \in Q$ and $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ the drift $\tilde{b}(r,u) := b(x,r,u,f(r))$ and diffusion $\tilde{\sigma}(r,u) := \sigma(x,r,u,f(r))$ terms can be verified to satisfy the needed assumptions. The measurability of $x \mapsto S^\epsilon[f](x,\cdot) \in H^2_T$ then immediately follows from that of the initial data $u(x,0)$ and of the noise $W^\epsilon(x,t)$, using a Picard iteration representation of the solution of the SDEs (5.4.2). The fact that the map S is well-defined, i.e. that $S^\epsilon[f](x,\cdot)$ is indeed an element of H^2_T uniformly bounded in $x \in Q$, is the subject of Lemma 5.4.1.

By definition, for every $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ we have $(L \circ S^\epsilon)[f](x,t) = \mathrm{Law}_{\mathbb{R}^4}(S^\epsilon[f](x,t))$ for all $x \in Q$ and $t \in [0,T]$, but in fact we can say that $(L \circ S^\epsilon)[f](x,\cdot) = \mathrm{Law}_{C([0,T];\mathbb{R}^4)}(S^\epsilon[f](x,\cdot))$, since it is the law of the SDE (5.4.2). That is, $(L \circ S^\epsilon)[f](x,\cdot)$ can be seen as a probability measure on the space of continuous paths $C([0,T];\mathbb{R}^4)$. Furthermore, if $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ is a fixed point of $L \circ S^\epsilon$, namely $f(x,t) = \mathrm{Law}_{\mathbb{R}^4}(S^\epsilon[f](x,t))$ for all $x \in Q$ and $t \in [0,T]$, then $f(x,\cdot) = \mathrm{Law}_{C([0,T];\mathbb{R}^4)}(S^\epsilon[f](x,\cdot))$.

Lemma 5.4.1 (A priori estimates on moments)**.** Given $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$, the pathwise unique solution $(S^\epsilon[f](x,t))_{x\in Q,\, t\geqslant 0}$ to (5.4.2) satisfies

$$\begin{aligned} |S^\epsilon[f]|^2_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})} &= \sup_{x\in Q} \mathbb{E}\left[\sup_{t\in[0,T]} |S^\epsilon[f](x,t)|^2\right] \\ &\leqslant C\left(1 + \sup_{x\in Q}\mathbb{E}\left[|u(x,0)|^2\right] + \int_0^T \sup_{x\in Q}\int_{\mathbb{R}^4} |u|^2\, f(t,x,du)\,dt\right) \\ &\leqslant C\left(1 + \sup_{x\in Q}\mathbb{E}\left[|u(x,0)|^2\right] + |f|^2_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})}\right), \end{aligned} \tag{5.4.3}$$

for a constant $C = C(T,b,\sigma)$. In particular, $S^\epsilon[f] \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})$, and the map S^ϵ and the composition $L \circ S^\epsilon$ are well defined. Moreover, if $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ is a fixed point of $L \circ S^\epsilon$, then $f(x,\cdot)$, as a probability measure on the space of continuous paths $C([0,T];\mathbb{R}^4)$, satisfies the stronger bound

$$\sup_{x\in Q}\int_{C([0,T];\mathbb{R}^4)} \sup_{t\in[0,T]} |v(t)|^2\, f(x,dv) = |S^\epsilon[f]|^2_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{H}^2_\mathrm{T})} \leqslant C\Big(1 + \sup_{x\in Q}\mathbb{E}\left[|u(x,0)|^2\right]\Big). \tag{5.4.4}$$

Proof. Fix any $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$, we want to estimate $|S^\epsilon[f](x,t)|^2$. Owing to the structure (5.2.2)–(5.2.3) and the sublinear growth properties (5.2.5)–(5.2.7) of the drift and diffusion terms, we have

$$|b(x,r,u,f(r))| + |\sigma(x,r,u,f(r))| \leqslant C\left(1 + |u| + \sup_{y\in Q}\int_{\mathbb{R}^4} |v|\, f(r,y,dv)\right). \tag{5.4.5}$$

Then, using Hölder's inequality one gets

$$\begin{aligned}\sup_{y\in Q}\int_{\mathbb{R}^4}|v|\,f(r,y,dv) &\leqslant \sup_{y\in Q}\Big(\int_{\mathbb{R}^4}|v|^2\,f(r,y,dv)\Big)^{\frac{1}{2}}\\ &\leqslant \sup_{y\in Q}\sup_{r\in[0,T]}\Big(\int_{\mathbb{R}^4}|v|^2\,f(r,y,dv)\Big)^{\frac{1}{2}} = |f|_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_{\mathrm{T}})}.\end{aligned} \tag{5.4.6}$$

Moreover, the explicit details in [Szn84b] on the construction of the reflection term ℓ in the SDE (5.4.2) imply that we can control it as follows:

$$|\ell[f](x,t)| \leqslant \sup_{\tau\in[0,t]}\left|u(x,0)+\int_0^\tau b(x,r,S^\epsilon[f](x,r),f(r))\,dr+\int_0^\tau \sigma(x,r,S^\epsilon[f],f\,dW^\epsilon(x,r)\right|. \tag{5.4.7}$$

Squaring both sides of the SDE (5.4.2), controlling the reflection term with the estimate (5.4.7), applying convexity inequalities, taking the supremum over $t\in[0,T]$ and then the expectation, and finally handling the deterministic integral with Hölder's inequality and the stochastic integral with Itô isometry, we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}|S^\epsilon[f](x,t)|^2\right] \leqslant C\Bigg(&\mathbb{E}\left[u(x,0)^2\right]+\mathbb{E}\left[\int_0^T b(x,r,S^\epsilon[f](x,r),f(r))^2\,dr\right]\,dt\\ &+\mathbb{E}\left[\int_0^T \sigma(x,r,S^\epsilon[f](x,r),f(r))^2\,dr)\right]\Bigg),\end{aligned}$$

for a numeric constant C. In turn, using the sublinear growth estimates (5.4.5)–(5.4.6), we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}|S^\epsilon[f](x,t)|^2\right] \leqslant C\Bigg(&1+\mathbb{E}\left[u(x,0)^2\right]\\ &+\int_0^T\mathbb{E}\left[\sup_{r\in[0,t]}|S^\epsilon[f](x,r)|^2\right]\,dt+\int_0^T\sup_{y\in Q}\int_{\mathbb{R}^4}|u|^2\,f(t,y,du)\,dt\Bigg),\end{aligned} \tag{5.4.8}$$

for a constant $C = C(T,b,\sigma)$. Then we exploit Grönwall's Lemma to get rid of the third term on the right hand side of the inequality in (5.4.8). Eventually, by taking the supremum over $x\in Q$ and using (5.4.6) again, we deduce the inequalities (5.4.3).

Suppose now that f is a fixed point of $L\circ S^\epsilon$. Since $f(t,y) = \mathrm{Law}_{\mathbb{R}^4}(S^\epsilon[f](y,t))$, we readily verify that

$$\int_0^T\sup_{y\in Q}\int_{\mathbb{R}^4}|u|^2\,f(t,y,du)\,dt = \int_0^T\sup_{y\in Q}\mathbb{E}\left[|S^\epsilon[f](y,t)|^2\right]\,dt \leqslant \int_0^T\sup_{y\in Q}\mathbb{E}\left[\sup_{r\in[0,t]}|S^\epsilon[f](y,r)|^2\right]\,dt.$$

Then we exploit this bound and again use Grönwall's Lemma in the first inequality (5.4.3) to get rid of the third term at the right hand side and obtain (5.4.4) in the statement. □

Now, assuming that the composition $L\circ S^\epsilon$ is a contraction in $\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_{\mathrm{T}})$, we first show how to conclude the strong existence and uniqueness for the McKean–Vlasov equation (5.2.8). Let $f\in\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_{\mathrm{T}})$ be the unique fixed point of $L\circ S^\epsilon$: since $S^\epsilon[f]$ solves (5.4.2) and $L\circ S^\epsilon[f] = f$, we obtain that $S^\epsilon[f]$ solves the McKean–Vlasov equation (5.2.8) on our stochastic basis with initial data $(u(x,0))_{x\in Q}$. Conversely, let $(u(x,t))_{x\in Q,\,t\geqslant 0}$ be a strong solution of (5.2.8) on our stochastic basis with these initial data, then $L[u] \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$ is a fixed point of $L\circ S^\epsilon$ and thus we must have $L[u] = f$, the unique fixed point; but then, since we have strong uniqueness for the SDEs (5.4.2) defining the map S^ϵ and since $u(x,t)$ solves these SDEs with this data f, we conclude that $(u(x,t))_{x\in Q,\,t\geqslant 0} = S^\epsilon[f]$.

Proof of Theorem 5.2.4. We show that the mapping $L\circ S^\epsilon$ is a strict contraction and then apply the Banach fixed point theorem. Take $f,g \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$. By definition of the map S^ϵ we have

$$S^\epsilon[f](x,t) = u(x,0) + \int_0^t b(x,r,S^\epsilon[f],f)\,dr + \int_0^t \sigma(x,r,S^\epsilon[f],f)\,dW^\epsilon(x,r) - \ell[f](x,t), \tag{5.4.9}$$

$$S^\epsilon[g](x,t) = u(x,0) + \int_0^t b(x,r,S^\epsilon[g],g)\,dr + \int_0^t \sigma(x,r,S^\epsilon[g],g)\,dW^\epsilon(x,r) - \ell[g](x,t). \tag{5.4.10}$$

Then, by definition of the map L, we have that $(S^\epsilon[f](x,t), S^\epsilon[g](x,t))$ is an admissible coupling for $(L\circ S^\epsilon[f](x,t), L\circ S^\epsilon[g](x,t))$ and we can use it to estimate $\mathcal{W}_2(L\circ S^\epsilon[f](x,t), L\circ S^\epsilon[g](x,t))$.

We take the difference of equations (5.4.9) and (5.4.10) and use the Itô formula to get

$$\begin{aligned}
&|S^\epsilon[f](x,t) - S^\epsilon[g](x,t)|^2 \\
&\quad = 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[g](x,r)\big)\big(b(x,r,S^\epsilon[f],f) - b(x,r,S^\epsilon[g],g)\big)\,dr \\
&\qquad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[g](x,r)\big)\big(\sigma(x,r,S^\epsilon[f],f) - \sigma(x,r,S^\epsilon[g],g)\big)\,dW^\epsilon(x,r) \\
&\qquad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[g](x,r)\big)\big(d\ell[g](x,r) - d\ell[f](x,r)\big) \\
&\qquad + \int_0^t \big(\sigma(x,r,S^\epsilon[f](x,r),f(r)) - \sigma(x,r,S^\epsilon[g](x,r),g(r))\big)^2\,dr.
\end{aligned} \tag{5.4.11}$$

We now argue analogously to (5.3.3)–(5.3.6) in the proof of Theorem 5.2.3. First, as in (5.3.4) the third the second term on the right hand side of (5.4.11) is negative, and we drop it. Then we take the supremum in time and apply the expectation, we control the first deterministic integral with Hölder's inequality and the stochastic integral with the Burkholder–Davis–Gundy and Hölder's inequality, and finally we absorb the necessary terms on the left hand side of (5.4.11) to get

$$\begin{aligned}
\mathbb{E}\left[\sup_{r\in[0,t]} |S^\epsilon[f](x,r) - S^\epsilon[g](x,r)|^2\right] \leqslant C\,\mathbb{E}\Bigg[&\int_0^t \big(b(x,r,S^\epsilon[f],f) - b(x,r,S^\epsilon[g],g)\big)^2\,dr \\
&+ \int_0^t \big(\sigma(x,r,S^\epsilon[f],f) - \sigma(x,r,S^\epsilon[g],g)\big)^2\,dr\Bigg],
\end{aligned}$$

for a numeric constant C. Now we exploit the Lipschitz properties of the drift and diffusion terms stated in Lemma 5.2.2 and we obtain, for $C = C(b,\sigma)$,

$$\mathbb{E}\left[\sup_{r\in[0,t]} |S^\epsilon[f](x,r) - S^\epsilon[g](x,r)|^2\right] \leqslant C\bigg(\int_0^t \mathbb{E}\left[|S^\epsilon[f](x,r) - S^\epsilon[g](x,r)|^2\right] dr + \int_0^t \sup_{y\in Q} \mathcal{W}_2(f(r,y,du), g(r,y,du))^2\, dr\bigg).$$

Using Grönwall's Lemma we get rid of the first term on the right hand side at the expense of a larger constant $C = C(T,b,\sigma)$. Moreover, we have $\mathcal{W}_2(f(r,y,du), g(r,y,du)) \leqslant \sup_{r\in[0,T]} \mathcal{W}_2(f(r,y,du), g(r,y,du))$ for any $r \in [0,T]$, and we conclude that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |S^\epsilon[f](x,t) - S^\epsilon[g](x,t)|^2\right] \leqslant C\,T \sup_{y\in Q}\sup_{r\in[0,T]} \mathcal{W}_2(f(r,y,du), g(r,y,du))^2, \quad (5.4.12)$$

for a constant $C = C(T,b,\sigma)$. Finally, since the right hand side is independent of x, we take the supremum over $x \in Q$ on the left hand side of (5.4.12).

In conclusion, recalling that $(S^\epsilon[f](x,t), S^\epsilon[g](x,t))$ is a coupling for $(L \circ S^\epsilon[f](x,t), L \circ S^\epsilon[g](x,t))$, we obtain

$$\begin{aligned} d_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})}(L\circ S^\epsilon[f], L\circ S^\epsilon[g])^2 &= \sup_{x\in Q}\sup_{t\in[0,T]} \mathcal{W}_2(L\circ S^\epsilon[f](x,t), L\circ S^\epsilon[g](x,t))^2 \\ &\leqslant \sup_{x\in Q}\sup_{t\in[0,T]} \mathbb{E}\left[|S^\epsilon[f](x,t) - S^\epsilon[g](x,t)|^2\right] \\ &\leqslant \sup_{x\in Q}\mathbb{E}\left[\sup_{t\in[0,T]} |S^\epsilon[f](x,t) - S^\epsilon[g](x,t)|^2\right] \\ &\leqslant C\,T \sup_{x\in Q}\sup_{t\in[0,T]} \mathcal{W}_2(f(x,t), g(x,t))^2 = C\,T\, d_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})}(f,g)^2, \end{aligned}$$

for $C = C(T,b,\sigma)$. This constant C is increasing in T. Therefore, upon possibly working in $L^\infty(Q; C^2_{T^*})$ for some smaller $T^* < T$, we can assume that $CT < 1$. That is to say, if $T > 0$ is small enough, we have a contraction in $\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$. In turn this implies that we have a pathwise unique solution to (5.2.8) over $[0,T]$. Repeating the same argument over $[T,2T]$, $[2T,3T]$ and so on, and exploiting the uniqueness, we can show there exists a pathwise unique solution defined over all $[0,\infty)$.

Now, assume in addition that $u(x,0) \in C^\alpha(Q; L^2(\Omega))$. The following argument proves that in this case, for any $f \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$, we have $S^\epsilon[f] \in C^\alpha(Q; L^2(\Omega; C([0,T];\mathbb{R}^4)))$. In particular, the solution of the McKean equation (5.2.8) satisfies $u^\epsilon(x,t) \in C^\alpha(Q; L^2(\Omega; C([0,T];\mathbb{R}^4)))$.

Given $x, y \in Q$, we manipulate the equations (5.4.2) for $S^\epsilon[f](x)$ and $S^\epsilon[f](y)$ to write

$$\begin{aligned} S^\epsilon[f](x,t) - S^\epsilon[f](y,t) = &\big(u(x,0) - u(y,0)\big) \\ &+ \int_0^t \big(b(x,r,S^\epsilon[f](x,r),f(r)) - b(x,r,S^\epsilon[f](y,r),f(r))\big)\,dr \\ &+ \int_0^t \big(\sigma(x,r,S^\epsilon[f](x,r),f(r)) - \sigma(y,r,S^\epsilon[f](y,r),f(r))\big)\,dW^\epsilon(x,r) \\ &+ \int_0^t \sigma(y,r,S^\epsilon[f](y,r),f(r))\,(dW^\epsilon(x,r) - dW^\epsilon(y,r)) \\ &+ \big(\ell[f](y,r) - \ell[f](x,r)\big). \end{aligned}$$

Applying the Itô formula to the squared power yields

$$\begin{aligned} &|S^\epsilon[f](x,t) - S^\epsilon[f](y,t)|^2 \\ &\quad + (u(x,0) - u(y,0))^2 \\ &\quad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)\big(b(x,r,S^\epsilon[f],f) - b(x,r,S^\epsilon[f],f)\big)\,dr \\ &\quad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)\big(\sigma(x,r,S^\epsilon[f],f) - \sigma(y,r,S^\epsilon[f],f)\big)\,dW^\epsilon(x,r) \\ &\quad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)\,\sigma(y,r,S^\epsilon[f],f)\,(dW^\epsilon(x,r) - dW^\epsilon(y,r)) \\ &\quad + 2\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)\big(d\ell[f](y,r) - d\ell[f](x,r)\big) \\ &\quad + \int_0^t \big(\sigma(x,r,S^\epsilon[f](x,r),f(r)) - \sigma(y,r,S^\epsilon[f](y,r),f(r))\big)^2\,dr \\ &\quad + \int_0^t \sigma(y,r,S^\epsilon[f](y,r),f(r))^2\,d\left[W^\epsilon(x,r) - W^\epsilon(y,r)\right]. \end{aligned} \tag{5.4.13}$$

For the first stochastic integral, the Burkholder–Davis–Gundy inequality and Hölder's inequality yield

$$\begin{aligned} &\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)\big(\sigma(x,r,S^\epsilon[f],f) - \sigma(y,r,S^\epsilon[f],f)\big)\,dW^\epsilon(x,r)\right|\right] \\ &\quad \leqslant \mathbb{E}\left[\left(\int_0^T \big(S^\epsilon[f](x,r) - S^\epsilon[f](y,r)\big)^2\big(\sigma(x,r,S^\epsilon[f],f) - \sigma(y,r,S^\epsilon[f],f)^2\,dr\right)^{\frac{1}{2}}\right] \\ &\quad \leqslant \mathbb{E}\left[\sup_{t\in[0,T]}\big|S^\epsilon[f](x,t) - S^\epsilon[f](y,t)\big|\left(\int_0^T \big(\sigma(x,r,S^\epsilon[f],f) - \sigma(y,r,S^\epsilon[f],f)^2\,dr\right)^{\frac{1}{2}}\right] \\ &\quad \leqslant \delta\,\mathbb{E}\left[\sup_{t\in[0,T]}\big|S^\epsilon[f](x,t) - S^\epsilon[f](y,t)\big|^2\right] \\ &\qquad + \frac{1}{\delta}\,\mathbb{E}\left[\int_0^T \big(\sigma(x,r,S^\epsilon[f],f) - \sigma(y,r,S^\epsilon[f],f)^2\,dr\right], \end{aligned} \tag{5.4.14}$$

where $\delta > 0$ shall be chosen small enough so as to absorb the first term on the right hand side. Similarly, for the second stochastic integral we find

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \big(S^\epsilon[f](x,r)-S^\epsilon[f](y,r)\big)\,\sigma(y,r,S^\epsilon[f](y,r),f(r))\,(dW^\epsilon(x,r)-dW^\epsilon(y,r)\right|\right]\\
&\quad\leqslant \mathbb{E}\left[\left(\int_0^T \big(S^\epsilon[f](x,r)-S^\epsilon[f](y,r)\big)^2\,\sigma(y,r,S^\epsilon[f],f)^2\,d\left[W^\epsilon(x,r)-W^\epsilon(y,r)\right]\right)^{\frac{1}{2}}\right]\\
&\quad\leqslant \delta\,\mathbb{E}\left[\sup_{t\in[0,T]}\left|S^\epsilon[f](x,t)-S^\epsilon[f](y,t)\right|^2\right]\\
&\qquad+\frac{1}{\delta}\,\mathbb{E}\left[\int_0^T \sigma(y,r,S^\epsilon[f],f)^2 d\left[W^\epsilon(x,r)-W^\epsilon(y,r)\right]\right],
\end{aligned}
\tag{5.4.15}
$$

where again $\delta > 0$ shall be chosen small enough to absorb the first term on the right hand side.

We now go back to (5.4.13). As in (5.3.3), the third term on the right hand side is always negative and we drop it. Then we take the supremum in time and we apply the expectation, we handle the first deterministic integral with Hölder's inequality and we use estimates (5.4.14) and (5.4.15) for the stochastic integrals, absorbing the necessary terms on the left hand side by choosing δ small enough. We obtain, for a numeric constant C,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|S^\epsilon[f](x,t)-S^\epsilon[f](y,t)|^2\right]\leqslant C\Bigg(&\mathbb{E}\left[|u(x,0)-u(y,0)|^2\right]\\
&+\int_0^T \mathbb{E}\left[|S^\epsilon[f](x,r)-S^\epsilon[f](y,r)|^2\right]dr\\
&+\int_0^T \mathbb{E}\left[\left|b(x,r,S^\epsilon[f],f)-b(y,r,S^\epsilon[f],f)\right|^2\right]dr\\
&+\int_0^T \mathbb{E}\left[|\sigma(x,r,S^\epsilon[f],f)-\sigma(y,r,S^\epsilon[f],f)|^2\right]dr\\
&+\mathbb{E}\left[\int_0^T \sigma(y,r,S^\epsilon[f](y,r),f(r))^2\,d\left[W^\epsilon(x,r)-W^\epsilon(y,r)\right]\right]\Bigg).
\end{aligned}
$$

Now we recall formula (5.2.11) for the quadratic variation of $W^\epsilon(x,t)-W^\epsilon(y,t)$, we use the Lipschitz and Hölder properties (5.2.13) of b and σ and convexity inequalities to get

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}|S^\epsilon[f](x,t)-S^\epsilon[f](y,t)|^2\right]\\
&\quad\leqslant C\Bigg(\mathbb{E}\left[|u(x,0)-u(y,0)|^2\right]+\int_0^T \mathbb{E}\left[|S^\epsilon[f](x,r)-S^\epsilon[f](y,r)|^2\right]dr+|x-y|^{2\alpha}\\
&\qquad+\frac{|x-y|^2}{\epsilon^2}\int_0^T \mathbb{E}\left[|\sigma(y,r,S^\epsilon[f],f)|^2\right]dr\Bigg),
\end{aligned}
\tag{5.4.16}
$$

for a constant $C = C(T,b,\sigma,\rho)$. We get rid of the second term on the right hand side of (5.4.16) with Grönwall's Lemma, at the price of a larger constant $C = C(T,b,\sigma)$. The first term is handled with the assumption $u(\cdot,0) \in C^\alpha(Q;L^2(\Omega))$. We control the last term with the sublinear growth property (5.2.13) of σ and the a priori estimate (5.4.3). In conclusion we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]}|S^\epsilon[f](x,t)-S^\epsilon[f](y,t)|^2\right]\leqslant C\,|x-y|^{2\alpha} + C\,\frac{|x-y|^2}{\epsilon^2}\left(1+\sup_{z\in Q}\mathbb{E}\left[|u(z,0)|^2\right]+|f|^2_{\mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})}\right), \tag{5.4.17}$$

for a constant $C = C(T, b, \sigma, \rho, [u(\cdot,0)]_\alpha)$. Since $x, y \in Q$ are arbitrary, this concludes the proof that $S^\epsilon[f] \in C^\alpha(Q; L^2(\Omega; C([0,T];\mathbb{R}^4)))$. □

We end this section by proving the existence and uniqueness of solutions to the associated Fokker–Planck equation.

Proof of Theorem 5.2.5. The result is a consequence of the Itô formula, the same fixed point argument as for the McKean–Vlasov equation and the uniqueness statement for the *linear* version of the Fokker–Planck type equation. Given any admissible initial condition $f_0(x,du) \in L^\infty(Q;\mathscr{P}_2(\mathbb{R}^4))$, standard probability theory ensures that we can find a probability space $(\Omega,\mathcal{F},\mathbb{P})$ supporting a 4-dimensional space-time white noise $(W(x,t))_{x\in Q,t\geqslant 0}$ and a family of random variables $u(x,0) \in L^\infty(Q;L^2(\Omega))$ independent of the noise $W(x,t)$ with $\mathrm{Law}_{\mathbb{R}^4}(u(x,0)) = f_0(x,du)$ for every $x \in Q$. Given any $\epsilon > 0$, we convolve and rescale the white noise to obtain $W^\epsilon(x,t)$ as in (5.2.10). With this stochastic basis and initial data, let $(\bar{u}^\epsilon(x,t) \in L^\infty(Q;H^2_T)$ be the solution of the ϵ-correlated McKean–Vlasov equation (5.2.8), whose existence is guaranteed by Theorem 5.2.4, and let us denote $f^\epsilon(x,t,du) = \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(x,t)) \in \mathrm{L}^\infty(\mathrm{Q};\mathrm{C}^2_\mathrm{T})$. We claim that f^ϵ is a weak solution of equation (5.2.12).

Take any $\phi \in C^2_c(\mathbb{R}_+\times\mathbb{R}^4)$ satisfying the Neumann boundary condition

$$\nabla_u\phi(t,u)\cdot n_{\partial(\mathbb{R}^4_+)}(u) = 0 \quad \text{for all } t,u \in \mathbb{R}^+\times\partial(\mathbb{R}^4_+), \tag{5.4.18}$$

where $n_{\partial(\mathbb{R}^4_+)}(u)$ denotes the unit outward normal at u. An application of the Itô formula yields

$$\begin{aligned}\phi(t,\bar{u}^\epsilon(x,t)) =&\phi(0,\bar{u}(x,0))\\ &+\int_0^t\partial_t\phi(r,\bar{u}^\epsilon(x,r))\,dr+\int_0^t\nabla_u\phi(r,\bar{u}^\epsilon(x,r))\cdot b(x,r,\bar{u}^\epsilon(x,r),f^\epsilon(r))\,dr\\ &+\int_0^t\nabla_u\phi(r,\bar{u}^\epsilon(x,r))\cdot\sigma(x,r,\bar{u}^\epsilon(x,r),f^\epsilon(r))\,dW^\epsilon(x,r)\\ &+\int_0^t\sum_{\beta=1}^4\partial_{u^\beta}\phi(r,\bar{u}^\epsilon(x,r))\,1_{\{\bar{u}^{\epsilon,\beta}(x,r)=0\}}\,d|\ell^\beta(x,\cdot)|(r)\\ &+\int_0^t\frac{1}{2}\sum_{\beta=1}^4\partial^2_{u^\beta u^\beta}\phi(r,\bar{u}^\epsilon(x,r))\,\left(\sigma_\beta(x,r,\bar{u}^\epsilon(x,r),f^\epsilon(r))\right)^2dr.\end{aligned} \tag{5.4.19}$$

The fifth term on the right hand side is identically zero thanks to the condition (5.4.18) on ϕ. Now we apply the expectation on both sides. The fourth term on the right hand side vanishes by the martingale property of the stochastic integral. Recalling that $\bar{u}^\epsilon(x,t)$ takes values in $\mathbb{R}^4_+$ only, we get

$$\begin{aligned}\int_{\mathbb{R}^4_+} \phi(t,u) f^\epsilon(t,x,du) &= \int_{\mathbb{R}^4_+} \phi(0,u) f_0(x,du) + \int_0^t \int_{\mathbb{R}^4_+} \partial_t \phi(r,u) f^\epsilon(r,x,du)\, dr \\ &+ \int_0^t \int_{\mathbb{R}^4_+} \nabla_u \phi(r,u) \cdot b(x,r,u,f^\epsilon(r)) f^\epsilon(r,x,du)\, dr \\ &+ \int_0^t \int_{\mathbb{R}^4_+} \frac{1}{2} \sum_{\beta=1}^4 \partial^2_{u^\beta u^\beta} \phi(r,u) \, \left(\sigma_\beta(x,r,u,f^\epsilon(r))\right)^2 f^\epsilon(r,x,du)\, dr.\end{aligned} \tag{5.4.20}$$

This is nothing but the weak formulation of (5.2.12) subjected to the no-flux boundary conditions. Since for every $T > 0$ we have $\bar{u}^\epsilon(x,\cdot) \in H^2_T$ and since it satisfies the bound (5.2.15), we conclude that $f^\epsilon(x,t) = \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(x,t))$ is a weak solution of (5.2.12) with initial condition $f_0(x,du)$, that it lies in the space $L^\infty(Q; C([0,\infty); \mathscr{P}_2(\mathbb{R}^4)))$ and that it actually satisfies the stronger bound (5.2.16).

Conversely, let $g \in L^\infty(Q; C([0,\infty); \mathscr{P}_2(\mathbb{R}^4)))$ be a weak solution of the non-linear Fokker–Planck equation (5.2.12) with the same initial data f_0. We claim that $g = f^\epsilon$. First, we can solve the family of standard SDEs with reflecting boundary conditions for the chosen g, for $x \in Q$:

$$\begin{cases} S^\epsilon[g](x,t) = u(x,0) + \displaystyle\int_0^t b(r,x,S^\epsilon[g](x,r),g(r))\, dr + \int_0^t \sigma(r,x,S^\epsilon[g], g\, dW^\epsilon(x,r) - \ell[g](x,t), \\ \ell^\beta[g](x,t) = -\, |\ell^\beta[g](x,\cdot)|(t)\,, \\ |\ell^\beta[g](x,\cdot)|(t) = \displaystyle\int_0^t 1_{\{S^\epsilon[g]^\beta(x,r)=0\}} d|\ell^\beta[g](x,\cdot)|(r) \quad \text{for } \beta = 1,2,3,4. \end{cases}$$

Arguing as in (5.4.19)–(5.4.20), we see that $h := L \circ S^\epsilon[g]$ now solves the *linear* Fokker–Planck equation with this fixed g and with the same initial data f_0:

$$\begin{cases} \partial_t h(t,x,u) + \nabla_u \cdot \big(b(x,t,u,g(t)) h(t,x,u)\big) = \dfrac{1}{2} \displaystyle\sum_{\beta=1}^4 \partial^2_{u^\beta u^\beta} \big(\sigma_\beta(x,t,u,g(t))^2 h(t,x,u)\big), \\ b^\beta(x,t,u,g(t)) h(t,x,u) - \dfrac{1}{2} \dfrac{\partial}{\partial u^\beta} \big(\sigma_\beta(x,t,u,g(t))^2 h(t,x,u)\big)\Big|_{u^\beta=0} = 0 \quad \text{for } \beta = 1,2,3,4. \end{cases} \tag{5.4.21}$$

This linear equation is readily verified to satisfy uniqueness by a duality argument: indeed, for fixed $x \in Q$, it suffices to test it against arbitrary functions $\varphi(t,u)$ satisfying the so-called *backward Kolmogorov equation* with Neumann boundary conditions on $\mathbb{R}^4_+$. That is,

$$\begin{cases} \partial_t \varphi + \nabla_u \varphi \cdot b(x,t,u,g(t)) + \dfrac{1}{2} \displaystyle\sum_{\beta=1}^4 \big(\sigma_\beta(x,t,u,g(t))\big)^2 \partial^2_{u^\beta u^\beta} \varphi = 0 \quad \text{on } (0,t_0) \times \mathbb{R}^4_+, \\ \nabla_u \varphi(t,u) \cdot n_{\partial(\mathbb{R}^4_+)}(u) = 0 \quad \text{on } (0,t_0) \times \partial(\mathbb{R}^4_+)\ , \\ \varphi(t_0,u) = \Phi(u) \quad \text{on } \{t_0\} \times \mathbb{R}^4_+, \end{cases}$$

where we let $t_0 \in \mathbb{R}^+$ and $\Phi \in C^2_c(\mathbb{R}^4_+)$ be arbitrary. Such an equation is always solvable since we have the right sign of the diffusion term (see [HR96] for details). Going back to

(5.4.21), we know that g as well is a solution of this equation and thus we must conclude that $g = L \circ S^\epsilon[g]$. Now let $T > 0$ be small enough so that the composition map $L \circ S^\epsilon$ is a contraction in $L^\infty(Q; C_T^2)$. This implies that $g \in L^\infty(Q; C_T^2)$ is a fixed point of $L \circ S^\epsilon$ and hence it must coincide with f^ϵ over $[0, T]$. Applying the same argument over subsequent intervals $[T, 2T]$, $[2T, 3T]$ and so forth proves the uniqueness statement.

In particular, given any two $\epsilon, \tilde{\epsilon} > 0$, we take $g = f^{\tilde{\epsilon}}$ and we conclude that $f^\epsilon = f^{\tilde{\epsilon}}$. That is to say $f(x,t) := \mathrm{Law}_{\mathbb{R}^4}(\bar{u}^\epsilon(x,t))$ is independent of ϵ and is the unique solution of the nonlinear Fokker–Planck equation.

Finally, we assume that $f_0 \in C^\alpha(Q; \mathscr{P}_2(\mathbb{R}^4))$ and we show that the corresponding solution satisfies $f \in C^\alpha(Q; \mathscr{P}_2(C[0,T]; \mathbb{R}^4))$. The theory of Wasserstein distances (see e.g. [Vil03]) ensures that we can find a stochastic basis supporting the white noise W and random variables $u(x, 0) \in C^\alpha(Q; L^2(\Omega))$ such that $\mathrm{Law}_{\mathbb{R}^4}(u(x, 0)) = f_0(x, du)$ for every $x \in Q$. We fix $\epsilon = 1$ and we consider the iteration maps (5.4.1) defined via this stochastic basis and with these initial data. In particular we have $L \circ S^\epsilon[f] = f$, and thus for any $x, y \in Q$ we obtain

$$\mathcal{W}_2\left(C\left([0,T]; \mathbb{R}^4\right)\right)(f(x,\cdot), f(y,\cdot)) \leqslant \mathbb{E}\left[\sup_{t\in[0,T]} |S^\epsilon[f](x,t) - S^\epsilon[f](y,t)|^2\right].$$

This and formula (5.4.17) with $\epsilon = 1$ show that $f \in C^\alpha(Q; \mathscr{P}_2(C[0,T]; \mathbb{R}^4))$. □

5.5 Error estimates between the particle system and the limiting model

In this section we rigorously show that the limiting behaviour of the particle system (5.2.1) as $M, N \to \infty$ is described by the McKean–Vlasov equation (5.2.8) as stated in Theorem 5.2.6 by obtaining an error estimate. We will use the so-called Sznitman coupling method (cf. [Szn91]).

First, we lay out the right setting so as to get the convergence result. We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assume it supports all the random variables listed below. First, for each $k \in \mathbb{N}$, let $\{W_k(x,t)\}_{k\in\mathbb{N}}$ be independent 4-dimensional space-time white noise terms over $Q \times [0, \infty)$.

For any $\epsilon > 0$ we then convolve and rescale the noise terms to obtain the ϵ-correlated noise W_k^ϵ as in formula (5.2.9). For $h \in \mathbb{N}$, we assume i.i.d. families of random initial conditions $u_h(x, 0) \in C^\alpha(Q; L^2(\Omega))$ on the sheet Q. Moreover, we require them to be independent of the white noise terms $\{W_k(x,t)\}_{k\in\mathbb{N}}$. Finally, as noted in Section 5.1, we take points $X_1, \ldots, X_N \in Q$ in the center of the squares of an equispaced grid on $Q = [0,1]^d$ with side length $N^{-\frac{1}{d}}$. We denote by Q_i^N the square with center X_i, and we notice that $\mathrm{meas}(Q_i^N) = \frac{1}{N}$ and $\mathrm{diam}(Q_i^N) = \sqrt{d} N^{-\frac{1}{d}}$.

We finally introduce the particles for the coupling method. For $i = 1, \dots, N$ and $k = 1, \dots, M$, let $u_{ik}^\epsilon(t)$ be the solution of the particle system (5.2.1) with initial data $u_{ik}(0) := u_k(X_i, 0)$ and Brownian motions $W_{ik}^\epsilon(t) := W_k^\epsilon(X_i, t)$. Let $\bar{u}_k^\epsilon(x,t)$ be the solution of the McKean–Vlasov equation with initial data $u_k(x,0)$ and correlated noise $W_k^\epsilon(x,t)$, and for $i = 1, \dots, N$ define $\bar{u}_{ik}^\epsilon(x,t) := \bar{u}_k^\epsilon(X_i, t)$.

Owing the i.i.d. properties of the initial data and the noise terms, we have the following.

Lemma 5.5.1. For fixed i, the particles $u_{ik}^\epsilon(t)$ are exchangeable for $k = 1, \dots, M$. Moreover, for fixed i, the particles $\bar{u}_{ik}^\epsilon(t)$ are i.i.d. for $k \in \mathbb{N}$.

We point out that this is not the case for the index i, both for the particles u_{ik}^ϵ and $\bar{u}_{ik}^\epsilon$. Indeed, the laws of u_{ik}^ϵ and u_{jk}^ϵ, or $\bar{u}_{ik}^\epsilon$ and $\bar{u}_{ik}^\epsilon$ respectively, might differ as a result of the x dependence of their defining equations. Furthermore, even if the points $X_i, X_j \in Q$ are far from each other, namely if $|X_i - X_j| > 2\epsilon$, so that their noise terms $W_k^\epsilon(X_i, t)$ and $W_k^\epsilon(X_j, t)$ are independent, the particles might still be correlated as a result of their initial data. In fact, from the point of view of modelling in neuroscience, we expect $u_k(x, 0)$ to be close to $u_k(y, 0)$ for x close to y.

We are finally ready to prove the convergence result of Theorem 5.2.6. We first stress the following.

Remark 5.5.2. As mentioned in Section 5.2.2, we point out that we do not need to impose any constraint on the ratio between the correlation radius ϵ of the noise and the minimum distance $\sqrt{d}N^{-1/d}$ between two grid points $X_i, X_j \in Q$. The choice of the scaling regime (ϵ, N), with $N \to \infty$ and $\epsilon \to 0$ or possibly also $\epsilon \equiv \epsilon_0$ a constant, is purely arbitrary and dictated by modelling arguments only. One might impose $\epsilon N^{\frac{1}{d}} < \sqrt{d}$ so that all the particles sense independent noise, or choose to impose a certain ratio $\epsilon N^{\frac{1}{d}} > \sqrt{d}$, so that neurons at locations close enough to each other sense correlated noise. The results and the proof of Theorem 5.2.6 are unchanged.

Proof of Theorem 5.2.6. For any $i = 1, \dots, N$ and $k = 1, \dots, M$, take the difference $|u_{ik}^\epsilon - \bar{u}_{ik}^\epsilon|$ between actual particles and McKean–Vlasov particles. Applying the Itô formula and exploiting the respective equations (5.2.1) and (5.2.8), we get

$$\begin{aligned}
|u_{ik}^\epsilon(t) - \bar{u}_{ik}^\epsilon(t)|^2 = {} & 2\int_0^t \big(u_{ik}^\epsilon(r) - \bar{u}_{ik}^\epsilon(r)\big)\big(b(X_i, r, u_{ik}^\epsilon(r), f_{MN}^\epsilon(r)) - b(X_i, r, \bar{u}_{ik}^\epsilon(r), f(r))\big)\, dr \\
& + 2\int_0^t \big(u_{ik}^\epsilon(r) - \bar{u}_{ik}^\epsilon(r)\big)\big(\sigma(X_i, r, u_{ik}^\epsilon, f_{MN}^\epsilon) - \sigma(X_i, r, \bar{u}_{ik}^\epsilon, f)\big)\, dW^\epsilon(X_i, r) \\
& + 2\int_0^t \big(u_{ik}^\epsilon(r) - \bar{u}_{ik}^\epsilon(r)\big)\big(d\bar{\ell}_{ik}(r) - d\ell_{ik}(r)\big) \\
& + \int_0^t \big(\sigma(X_i, r, u_{ik}^\epsilon(r), f_{MN}^\epsilon(r)) - \sigma(X_i, r, \bar{u}_{ik}^\epsilon(r), f(r))\big)^2\, dr.
\end{aligned} \tag{5.5.1}$$

Now we argue as in (5.3.3)–(5.3.6). First we drop the third term in (5.5.1), which is always negative owing to the definition of the reflection terms ℓ_{ik} and $\bar{\ell}_{ik}$. Then we take the supremum in $t \in [0,\tau]$ and apply the expectation. Next we use the Burkholder–Davis–Gundy and Hölder's inequality, we absorb the necessary terms into the left hand side and finaly we exploit Grönwall's Lemma. We eventually obtain, for $C = C(T)$,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,\tau]} |u^\epsilon_{ik}(t) - \bar{u}^\epsilon_{ik}(t)|^2\right] \leqslant C\Bigg(&\int_0^\tau \mathbb{E}\left[|b(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f)|^2\right] dr \\ &+ \int_0^\tau \mathbb{E}\left[|\sigma(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - \sigma(X_i, r, \bar{u}^\epsilon_{ik}, f)|^2\right] dr\Bigg).\end{aligned} \tag{5.5.2}$$

In order to split the terms on the right hand side of the inequality (5.5.2) and exploit the particular structure of the drift and diffusion terms, we introduce the following probability measure on $Q \times \mathbb{R}^4$:

$$\bar{f}^\epsilon_{MN}(t, dy, dv) = \frac{1}{MN}\sum_{j=1}^N\sum_{m=1}^M \delta_{(X_j, \bar{u}^\epsilon_{jm}(t))} \in \mathscr{P}(Q \times \mathbb{R}^4). \tag{5.5.3}$$

This measure is just the empirical measure associated to the collection of McKean–Vlasov particles $(X_j, \bar{u}^\epsilon_{jm}(t))$. We have

$$\begin{aligned}|b(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f)| \leqslant &|b(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f^\epsilon_{MN})| \\ &+ |b(X_i, , r, \bar{u}^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, \bar{f}^\epsilon_{MN})| \\ &+ |b(X_i, , r, \bar{u}^\epsilon_{ik}, \bar{f}^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f)|.\end{aligned} \tag{5.5.4}$$

Due to the structure of the drift term (5.2.2) and its Lipschitz properties (5.2.4)–(5.2.6), and owing to the definition of $\bar{f}^\epsilon_{MN}(r)$ in (5.5.3), we get the following estimates for terms on the right hand side of (5.5.4):

$$\begin{aligned}\left|b(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f^\epsilon_{MN})\right| &\leqslant C\left|u^\epsilon_{ik}(r) - \bar{u}^\epsilon_{ik}(r)\right|, \\ \left|b(X_i, r, \bar{u}^\epsilon_{ik}, f^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, \bar{f}^\epsilon_{MN})\right| &\leqslant C\left|\int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v) f^\epsilon_{MN}(r, dy, dv)\right. \\ &\qquad \left. - \int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v)\bar{f}^\epsilon_{MN}(r, dy, dv)\right| \\ &\leqslant C\,\frac{1}{MN}\sum_{j=1}^N\sum_{m=1}^M \left|u^\epsilon_{jm}(r) - \bar{u}^\epsilon_{jm}(r)\right|, \\ \left|b(X_i, r, \bar{u}^\epsilon_{ik}, \bar{f}^\epsilon_{MN}) - b(X_i, r, \bar{u}^\epsilon_{ik}, f)\right| &\leqslant C\left|\int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v)\bar{f}^\epsilon_{MN}(r, dy, dv)\right. \\ &\qquad \left. - \int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v) f(r, dy, dv)\right| \\ &= C\left|\frac{1}{MN}\sum_{j=1}^N\sum_{m=1}^M\Bigg(b_1(X_i, X_j, r, \bar{u}^\epsilon_{ik}(r), \bar{u}^\epsilon_{jm}(r))\right. \\ &\qquad \left. - \int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v) f(r, dy, dv)\Bigg)\right|,\end{aligned} \tag{5.5.5}$$

for a constant $C = C(b)$ only depending on the Lipschitz constants of b. An identical splitting (5.5.4) holds for the term $(\sigma(X_i, r, u^\epsilon_{ik}, f^\epsilon_{MN}) - \sigma(X_i, r, \bar{u}^\epsilon_{ik}, f))$ and using the Lipschitz properties (5.2.4)–(5.2.6) of σ we obtain analogous estimates to (5.5.5).

Going back to (5.5.2), we exploit (5.5.4) and (5.5.5). After standard convexity inequalities we obtain, for $C = C(T, b, \sigma)$,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,\tau]} |u^\epsilon_{ik}(t) - \bar{u}^\epsilon_{ik}(t)|^2\right] \leqslant C\Bigg\{&\int_0^\tau \mathbb{E}\big[|u^\epsilon_{ik}(r) - \bar{u}^\epsilon_{ik}(r)|^2\big]\,dr \\ &+ \int_0^\tau \frac{1}{MN}\sum_{j,m=1}^{N,M} \mathbb{E}\big[|u^\epsilon_{jm}(r) - \bar{u}^\epsilon_{jm}(r)|^2\big]\,dr \\ &+ \int_0^\tau \mathbb{E}\Bigg[\Bigg|\frac{1}{MN}\sum_{j,m=1}^{N,M}\Big(b_1(X_i, X_j, r, \bar{u}^\epsilon_{ik}, \bar{u}^\epsilon_{jm}) \\ &\qquad - \int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v) f(r, dy, dv)\Big)\Bigg|^2\Bigg]dr \\ &+ \int_0^\tau \mathbb{E}\Bigg[\Bigg|\frac{1}{MN}\sum_{j,m=1}^{N,M}\Big(\sigma_1(X_i, X_j, r, \bar{u}^\epsilon_{ik}, \bar{u}^\epsilon_{jm}) \\ &\qquad - \int_{Q\times\mathbb{R}^4} \sigma_1(X_i, y, r, \bar{u}^\epsilon_{ik}, v) f(r, dy, dv)\Big)\Bigg|^2\Bigg]dr\Bigg\}.\end{aligned} \tag{5.5.6}$$

Averaging (5.5.6) over $i = 1, \dots, N$ and $k = 1, \dots, M$, and then using Grönwall's Lemma to get rid of the first two terms on the right hand side, we obtain

$$\frac{1}{MN}\sum_{i=1}^N\sum_{k=1}^M \mathbb{E}\left[\sup_{t\in[0,\tau]} |u^\epsilon_{ik}(t) - \bar{u}^\epsilon_{ik}(t)|^2\right] \leqslant C\frac{1}{MN}\sum_{i=1}^N\sum_{k=1}^M \int_0^\tau R^b_{ik}(t) + R^\sigma_{ik}(t)\ dt, \tag{5.5.7}$$

for another constant $C = C(T, b, \sigma)$. Here we have defined

$$R^b_{ik}(t) = \mathbb{E}\left[\left|\frac{1}{MN}\sum_{j=1}^N\sum_{m=1}^M \Big(b_1(X_i, X_j, t, \bar{u}^\epsilon_{ik}(t), \bar{u}^\epsilon_{jm}(t)) - \int_{Q\times\mathbb{R}^4} b_1(X_i, y, t, \bar{u}^\epsilon_{ik}(t), v) f(t, y, dv)\,dy\Big)\right|^2\right],$$

$$R^\sigma_{ik}(t) = \mathbb{E}\left[\left|\frac{1}{MN}\sum_{j=1}^N\sum_{m=1}^M \Big(\sigma_1(X_i, X_j, t, \bar{u}^\epsilon_{ik}(t), \bar{u}^\epsilon_{jm}(t)) - \int_{Q\times\mathbb{R}^4} \sigma_1(X_i, y, t, \bar{u}^\epsilon_{ik}(t), v) f(t, y, dv)\,dy\Big)\right|^2\right],$$

which are the arguments of the last two integrals on the right hand side of (5.5.6). Heuristically, the error terms R^b_{ik} and R^σ_{ik} should be small in view of the weak law of large numbers. Indeed, upon conditioning on $\bar{u}^\epsilon_{ik}$, for each fixed $j = 1, \dots, N$, we are essentially taking the average of the i.i.d. terms $b_1(X_i, X_j, t, \bar{u}^\epsilon_{ik}(t), \bar{u}^\epsilon_{jm}(t))$ for $m = 1, \dots, M$, and then subtracting their common expectation $\int_{Q\times\mathbb{R}^4} b_1(X_i, y, r, \bar{u}^\epsilon_{ik}(t), v) f(t, y, dv)\,dy$.

In order to control the term to the right in (5.5.7), we need the following estimate whose proof is postponed for the sake of the reader. For any $T > 0$, we have

$$\sup_{t\in[0,T]} |R^b_{ik}(t)| + |R^\sigma_{ik}(t)| \leqslant C\left(1 + \sup_{x\in Q}\mathbb{E}\left[|u_k(x, 0)|^2\right]\right)\left(\frac{1}{M} + \frac{1}{N^{\frac{\alpha}{d}}}\right), \tag{5.5.8}$$

for a constant $C = C(T, b, \sigma, \rho, [u(\cdot, 0)]_\alpha)$, for every $i = 1, \ldots, N$ and $k = 1, \ldots, M$. Plugging (5.5.8) into (5.5.7) we obtain, for $C = C(T, b, \sigma, \rho, [u(\cdot, 0)]_\alpha)$,

$$\frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \mathbb{E}\left[\sup_{t \in [0,\tau]} |u_{ik}^{\epsilon}(t) - \bar{u}_{ik}^{\epsilon}(t)|^2 \right] \leqslant C \left(1 + \sup_{x \in Q} \mathbb{E}\left[|u_k(x,0)|^2 \right] \right) \left(\frac{1}{M} + \frac{1}{N^{\frac{\alpha}{d}}} \right). \tag{5.5.9}$$

We can now finally prove Theorem 5.2.6. We go back to (5.5.6), and get rid of the first term on the right hand side with Grönwall's Lemma. We control the second term on the right hand side with (5.5.9) and the last two terms with (5.5.8). This yields formula (5.2.18) and concludes the proof.

□

Proof of estimate (5.5.8). We prove the estimate for R_{ik}^b. Identical computations replacing b with σ prove the analogous result for R_{ik}^σ. Recalling that $\text{meas}(Q_j^N) = \frac{1}{N}$, we split

$$\begin{aligned}
R_{ik}^b(t) = \mathbb{E}\Bigg[\Bigg| \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \Big(& b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), \bar{u}_{jm}^{\epsilon}(t)) \\
& - \int_{Q \times \mathbb{R}^4} b_1(X_i, y, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, y, dv)\, dy \Big) \Bigg|^2 \Bigg] \\
\leqslant \mathbb{E}\Bigg[\Bigg| \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \Big(& b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), \bar{u}_{jm}^{\epsilon}(t)) \\
& - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, X_j, dv) \Big) \Bigg|^2 \Bigg] \\
+ \mathbb{E}\Bigg[\Bigg| \sum_{j=1}^{N} \Big(\int_{Q_j^N} \int_{\mathbb{R}^4} & b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), v)\, f(t, X_j, dv) \\
& - \int_{\mathbb{R}^4} b_1(X_i, y, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, y, dv)\, dy \Big) \Bigg|^2 \Bigg].
\end{aligned} \tag{5.5.10}$$

For the first term of (5.5.10), the estimate is proved similarly to the weak law of large numbers. Indeed, for $C = C(T, b, \sigma)$, we compute

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg| \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \Big(b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), \bar{u}_{jm}^{\epsilon}(t)) - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, X_j, dv) \Big) \Bigg|^2 \Bigg] \\
&\leqslant \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\Bigg[\Bigg| \frac{1}{M} \sum_{m=1}^{M} \Big(b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), \bar{u}_{jm}^{\epsilon}(t)) \\
&\qquad - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, X_j, dv) \Big) \Bigg|^2 \Bigg] \\
&= \frac{1}{N} \sum_{j=1}^{N} \frac{1}{M^2} \sum_{\substack{m_1=1 \\ m_2=1}}^{M} \mathbb{E}\Bigg[\Big(b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, \bar{u}_{jm_1}^{\epsilon}) - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, v) f(t, X_j, dv) \Big) \\
&\qquad \cdot \Big(b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, \bar{u}_{jm_2}^{\epsilon}) - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, v) f(t, X_j, dv) \Big) \Bigg] \\
&= \frac{1}{N} \sum_{j=1}^{N} \frac{1}{M^2} \sum_{m=1}^{M} \mathbb{E}\Bigg[\Big(b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, \bar{u}_{jm}^{\epsilon}) - \int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}, v) f(t, X_j, dv) \Big)^2 \Bigg] \\
&\leqslant \frac{1}{M} C \left(1 + \sup_{x \in Q} \mathbb{E}\left[|u_k(x,0)|^2 \right] \right).
\end{aligned} \tag{5.5.11}$$

In the first passage we used a convexity inequality. In the last passage we used the sublinear growth properties (5.2.7) of b_1 and the a priori estimate (5.2.15) for McKean–Vlasov particles. In the second passage we unfolded the square, and in the third we noticed that, after conditioning with respect to $\bar{u}^{\epsilon}_{ik}(t)$, only the "diagonal terms" survive in the sum, i.e. those with $m_1 = m_2$. Namely, when $m_1 \neq m_2$ the corresponding term in (5.5.11) is identically zero. Indeed, under this condition, assuming by symmetry $m_1 \neq k$, we have that $\bar{u}^{\epsilon}_{jm_1}(t)$ is independent of $\bar{u}^{\epsilon}_{jm_2}(t)$ and $\bar{u}^{\epsilon}_{ik}(t)$. Hence we compute

$$
\begin{aligned}
&\mathbb{E}\Bigg[\Big(b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},\bar{u}^{\epsilon}_{jm_1}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},v)f(t,X_j,dv)\Big) \\
&\qquad\qquad \cdot \Big(b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},\bar{u}^{\epsilon}_{jm_2}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},v)f(t,X_j,dv)\Big)\Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{E}\Bigg[\Big(b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},\bar{u}^{\epsilon}_{jm_1}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},v)f(t,X_j,dv)\Big) \\
&\qquad\qquad \cdot \Big(b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},\bar{u}^{\epsilon}_{jm_2}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},v)f(t,X_j,dv)\Big) \Bigg| \, \bar{u}^{\epsilon}_{ik} \Bigg]\Bigg] \qquad (5.5.12) \\
&= \mathbb{E}\Bigg[\mathbb{E}\Bigg[\Big(b_1(X_i,X_j,t,u,\bar{u}^{\epsilon}_{jm_1}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,u,v)f(t,X_j,dv)\Big)\Bigg]_{u=\bar{u}^{\epsilon}_{ik}} \\
&\qquad\qquad \cdot \mathbb{E}\Bigg[\Big(b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},\bar{u}^{\epsilon}_{jm_2}) - \int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik},v)f(t,X_j,dv)\Big) \Bigg| \, \bar{u}^{\epsilon}_{ik} \Bigg]\Bigg] \\
&= 0.
\end{aligned}
$$

In the second passage we conditioned on $\bar{u}^{\epsilon}_{ik}$ and in the third passage we used standard properties of the conditional expectation (see e.g. [DPZ92, Chapter 2]). Finally we used that $E[b_1(X_i,X_j,t,u,\bar{u}^{\epsilon}_{jm})] = \int_{\mathbb{R}^4} b_1(X_i,X_j,t,u,v)f(t,X_j,dv)$ by definition of $f(t,X_j,dv)$.

For the second term second term on the right hand side of (5.5.10), we first compute

$$
\begin{aligned}
&\Big|\int_{\mathbb{R}^4} b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik}(t),v)\, f(t,X_j,dv) - \int_{\mathbb{R}^4} b_1(X_i,y,t,\bar{u}^{\epsilon}_{ik}(t),v) f(t,y,dv)\Big| \\
&\qquad \leqslant \int_{(\mathbb{R}^4)^2} |b_1(X_i,X_j,t,\bar{u}^{\epsilon}_{ik}(t),v) - b_1(X_i,y,t,\bar{u}^{\epsilon}_{ik}(t),w)| \; \pi_0(X_j,y,dv,dw) \\
&\qquad \leqslant C \int_{(\mathbb{R}^4)^2} |X_j - y|^{\alpha} + |v-w| \; \pi_0(X_j,y,dv,dw) \qquad (5.5.13) \\
&\qquad \leqslant C\,|X_j - y|^{\alpha} + \mathcal{W}_1\left(\mathbb{R}^4\right)\left(f(t,X_j,dv), f(t,y,dv)\right) \\
&\qquad \leqslant C\,|X_j - y|^{\alpha}
\end{aligned}
$$

for a constant $C = C(T,b,\sigma,\rho,[u(\cdot,0)]_{\alpha})$. In the second passage we took any optimal pairing $\pi_0(X_j,y,dv,dw)$ for $\mathcal{W}_1(f(t,X_j,dv),f(t,y,dv))$, in the third we used the Lipschitz and Hölder properties (5.2.6) of b_1, and in the last we used the ordering $\mathcal{W}_1 \leqslant \mathcal{W}_2$ of Wasserstein distances and the Hölder continuity (5.2.17) of f in $\mathcal{W}_2$. Then, using (5.5.13) and recalling that $\text{meas}(Q_j^N) = \frac{1}{N}$ and $\text{diam}(Q_j^N) = N^{-\frac{1}{d}}$, we compute

$$\begin{aligned}
&\mathbb{E}\left[\Big|\sum_{j=1}^{N}\Big(\int_{Q_j^N}\int_{\mathbb{R}^4} b_1(X_i, X_j, t, \bar{u}_{ik}^{\epsilon}(t), v)\, f(t, X_j, dv) - \int_{\mathbb{R}^4} b_1(X_i, y, t, \bar{u}_{ik}^{\epsilon}(t), v) f(t, y, dv)\; dy\Big)\Big|^2\right] \\
&\leqslant \mathbb{E}\left[\Big|\sum_{j=1}^{N}\int_{Q_j^N} C|X_j - y|^{\alpha}\, dy\Big|^2\right] \\
&\leqslant C\,\text{diam}(Q_j^N)^{2\alpha} \\
&= C\, N^{-\frac{2\alpha}{d}},
\end{aligned} \tag{5.5.14}$$

for a constant $C = C(T, b, \sigma, \rho, [u(\cdot, 0)]_{\alpha})$.

In conclusion, combining (5.5.10) with estimates (5.5.12) and (5.5.14) we obtain the estimate (5.5.8). □

5.6 Convergence of empirical measures

In this last section, we further analyze the limiting behaviour of the particle system as we let $M, N \to \infty$ and prove Theorem 5.2.7. In the same setting outlined in Section 5.5, we show that the time dependent empirical measure

$$f_{MN}^{\epsilon}(t, dx, du) = \frac{1}{MN}\sum_{j=1}^{N}\sum_{m=1}^{M}\delta_{(X_j, u_{jm}^{\epsilon}(t))} \in \mathscr{P}(Q \times \mathbb{R}^4),$$

associated to the particle system (5.2.1), located at the grid points $X_1, \dots, X_N$, converges in Wasserstein distance $\mathcal{W}_1(Q \times \mathbb{R}^4)$ to the measure $f(t, dx, du)$, obtained from the solution of the Fokker–Planck equation (5.2.12) via formula (5.1.5). The key step towards the result is to split the Wasserstein distance:

$$\mathcal{W}_1(Q\times\mathbb{R}^4)(f_{MN}^{\epsilon}(t), f(t)) \leqslant \mathcal{W}_1(f_{MN}^{\epsilon}(t), \bar{f}_{MN}^{\epsilon}(t)) + \mathcal{W}_1(\bar{f}_{MN}^{\epsilon}(t), \bar{f}_N(t)) + \mathcal{W}_1(\bar{f}_N(t), f(t)). \tag{5.6.1}$$

Here $\bar{f}_{MN}^{\epsilon} = \frac{1}{MN}\sum_{j=1}^{N}\sum_{m=1}^{M}\delta_{(X_j, \bar{u}_{jm}^{\epsilon}(t))}$ is the empirical measure of the associated McKean–Vlasov particles as in Section 5.5, and

$$\bar{f}_N(t, dx, du) := \frac{1}{N}\sum_{j=1}^{N}\delta_{X_j}\otimes f(t, X_j, du),$$

an auxiliary measure, can be viewed as a Riemann sum approximation for the measure $f(t, dx, du)$. Then Theorem 5.2.7 is an immediate consequence of the splitting (5.6.1) and Lemma 5.6.1, 5.6.4 and 5.6.5 below. The first term in (5.6.1) is readily handled with Theorem 5.2.6 as follows.

Lemma 5.6.1. In the setting above, for any $T > 0$ we have

$$\sup_{t\in[0,T]}\mathbb{E}\left[\mathcal{W}_1(Q\times\mathbb{R}^4)(f_{MN}^{\epsilon}(t), \bar{f}_{MN}^{\epsilon}(t))\right] \leqslant C\left(1 + \sup_{x\in Q}\mathbb{E}\left[|u_k(x,0)|^2\right]^{\frac{1}{2}}\right)\left(\frac{1}{M^{\frac{1}{2}}} + \frac{1}{N^{\frac{\alpha}{d}}}\right),$$

for a constant $C = C(T, \rho, b, \sigma, [u(\cdot, 0)]_{\alpha})$.

Proof. It suffices to notice that

$$\pi_0 = \frac{1}{MN} \sum_{j=1}^{N} \sum_{k=1}^{M} \delta_{(X_j, u^\epsilon_{jk}(t), X_j, \bar{u}^\epsilon_{jk}(t))},$$

is an admissible pairing for $f^\epsilon_{MN}(t)$ and $\bar{f}^\epsilon_{MN}(t)$. Then, by definition of the Wasserstein distance,

$$\mathcal{W}_1(f^\epsilon_{MN}(t), \bar{f}^\epsilon_{MN}(t)) \leqslant \int_{(Q\times\mathbb{R}^4)^2} |x-y| + |u-v|\, d\pi_0 = \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} |u^\epsilon_{jm}(t) - \bar{u}^\epsilon_{jm}(t)|.$$

Next we apply $\mathbb{E}$ at both sides of the inequality and use Hölder's inequality to get

$$\begin{aligned} \mathbb{E}\big[\, \mathcal{W}_1(f^\epsilon_{MN}(t), \bar{f}^\epsilon_{MN}(t))\big] &\leqslant \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \mathbb{E}\big[|u^\epsilon_{jm}(t) - \bar{u}^\epsilon_{jm}(t)|\big] \\ &\leqslant \frac{1}{MN} \sum_{j=1}^{N} \sum_{m=1}^{M} \mathbb{E}\left[|u^\epsilon_{jm}(t) - \bar{u}^\epsilon_{jm}(t)|^2\right]^{\frac{1}{2}}. \end{aligned} \tag{5.6.2}$$

Now we conclude by plugging (5.2.18) into (5.6.2). □

Let us now turn to the second term in the splitting (5.6.1). To start with, in a weak law of large numbers manner, we get the following lemma.

Lemma 5.6.2. In the setting above, for every $N \in \mathbb{N}$ and every $\epsilon > 0$, for any $t \geqslant 0$,

$$\lim_{M\to\infty} \mathcal{W}_1(Q\times\mathbb{R}^4)(\bar{f}^\epsilon_{MN}(t), \bar{f}_N(t)) = 0 \quad \text{in probability.}$$

Proof. The key observation is the following: if $\varphi(x,v) \in C(Q\times\mathbb{R}^4)$ has linear growth in v, that is $|\varphi(x,v)| \leqslant L_\varphi(1+|v|)$ for some constant $L_\varphi \geqslant 0$, then we have $\langle \varphi, \bar{f}_{MN}(t)\rangle \to \langle \varphi, \bar{f}_N(t)\rangle$ in $L^2(\Omega)$ as $M \to \infty$, uniformly in $N \in \mathbb{N}$. Indeed, with the same arguments as in the proof of Lemma 5.5.8, we have, for a constant $C = C(T, b, \sigma)$ independent of φ,

$$\begin{aligned} &\mathbb{E}\left[|\langle \varphi, \bar{f}^\epsilon_{MN}(t)\rangle - \langle \varphi, \bar{f}_N(t)\rangle|^2\right] \\ &\quad\leqslant \frac{1}{N}\sum_{j=1}^{N} \mathbb{E}\left[\left|\frac{1}{M}\sum_{m=1}^{M}\Big(\varphi(X_j, \bar{u}^\epsilon_{jm}(t)) - \int_{\mathbb{R}^4}\varphi(X_j, v) f(t, X_j, dv)\Big)\right|^2\right] \\ &\quad= \frac{1}{N}\sum_{j=1}^{N}\frac{1}{M^2}\sum_{m_1\neq m_2}\mathbb{E}\Bigg[\Big(\varphi(X_j, \bar{u}^\epsilon_{jm_1}(t)) - \int_{\mathbb{R}^4}\varphi(X_j, v) f(t, X_j, dv)\Big) \\ &\qquad\qquad\cdot\Big(\varphi(X_j, \bar{u}^\epsilon_{jm_2}(t)) - \int_{\mathbb{R}^4}\varphi(X_j, v) f(t, X_j, dv)\Big)\Bigg] \\ &\quad\leqslant \frac{1}{M} L^2_\varphi\, C(T, b, \sigma)\left(1 + \sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^2\big]\right). \end{aligned} \tag{5.6.3}$$

We now collect some auxiliary facts and then use these observations to complete the proof. Since $Q\times\mathbb{R}^4$ is a Polish space, it embeds continuously in the compact space $[0,1]^{\mathbb{N}}$ endowed with the distance $\eta(x,y) := \sum_{k=1}^{\infty} \frac{1}{2^k}|x_k - y_k|$. Let $\overline{Q\times\mathbb{R}^4}$ denote the closure of (the image of) $Q\times\mathbb{R}^4$ in $[0,1]^{\mathbb{N}}$. Let $U_\eta(Q\times\mathbb{R}^4)$ denote the space of function $\psi : Q\times\mathbb{R}^4 \to \mathbb{R}$ which are bounded and uniformly continuous with respect to the distance η restricted to $Q\times\mathbb{R}^4$. By the continuous extension theorem we have that $U_\eta(Q\times\mathbb{R}^4) = C_b(\overline{Q\times\mathbb{R}^4})$, that is to say each bounded uniformly continuous function on $Q\times\mathbb{R}^4$ extends uniquely to a bounded continuous function on $\overline{Q\times\mathbb{R}^4}$ and conversely each such function restricts to a bounded uniformly continuous function on $Q\times\mathbb{R}^4$. The space $U_\eta(Q\times\mathbb{R}^4)$ is separable, since $\overline{Q\times\mathbb{R}^4}$ is compact. Let $\{\varphi_n\}_{n\in\mathbb{N}}$ be a dense countable subset and set $\psi_n := \frac{1}{\|\varphi_n\|_\infty+1}\varphi_n$ for every n.

Given measures $\mu_j,\mu\in\mathscr{P}(Q\times\mathbb{R}^4)$, the Portmanteau theorem implies that $\mu_j\rightharpoonup\mu$ as $j\to\infty$ if and only if $\langle\varphi,\mu_j-\mu\rangle\to 0$ for every $\varphi\in U_\eta(Q\times\mathbb{R}^4)$. Defining the distance δ on $\mathscr{P}(Q\times\mathbb{R}^4)$ by

$$\delta(\mu,\nu) := \sum_{n\in\mathbb{N}}\frac{1}{2^n}|\langle\psi_n,\mu-\nu\rangle|,$$

we immediately see that, as $j\to\infty$,

$$\mu_j\rightharpoonup\mu \iff \delta(\mu_j,\mu)\to 0.$$

Finally, we recall that the convergence in Wasserstein distance of order 1 is equivalent to weak convergence combined with convergence of first moments (see e.g. [Vil03, Chapter 7]), i.e.

$$\mathcal{W}_1(Q\times\mathbb{R}^4)(\mu_j,\mu)\to 0 \iff \begin{cases}\delta(\mu_j,\mu)\to 0,\\ \int_{Q\times\mathbb{R}^4}|x|+|u|\,d\mu_j\to\int_{Q\times\mathbb{R}^4}|x|+|u|\,d\mu.\end{cases} \tag{5.6.4}$$

We now conclude the proof. Using the definitions of the measures $\bar f_{MN}$ and $\bar f_N$, convexity inequalities and (5.6.3), we compute, for every $M,N\in\mathbb{N}$ and $\epsilon>0$,

$$\begin{aligned}\mathbb{E}\left[\delta(\bar f_{MN},\bar f_N)^2\right] &\leqslant \sum_{k=1}^{\infty}\frac{1}{2^k}\mathbb{E}\left[|\langle\psi_k,\bar f_{MN}-\bar f_N\rangle|^2\right]\\ &\leqslant\frac{1}{M}C(T,b,\sigma)\left(1+\sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^2\big]\right).\end{aligned} \tag{5.6.5}$$

Analogously we have, for every $M,N\in\mathbb{N}$ and $\epsilon>0$,

$$\mathbb{E}\left[\left|\int_{Q\times\mathbb{R}^4}|x|+|u|\,d\bar f_{MN}-\int_{Q\times\mathbb{R}^4}|x|+|u|\,d\bar f_N\right|^2\right]\leqslant\frac{1}{M}C(T,b,\sigma)\left(1+\sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^2\big]\right). \tag{5.6.6}$$

Given any arbitrary subsequence $M_k\to\infty$, using (5.6.5)–(5.6.6) and a diagonal argument, we find a sub-subsequence $M_{k_j}\to\infty$ such that

$$\forall N\in\mathbb{N}\quad \delta(\bar f_{NM_{k_j}},\bar f_N)\to 0\quad\text{and}\quad\int_{Q\times\mathbb{R}^4}|x|+|u|\;\;(d\bar f_{NM_{k_j}}-d\bar f_N)\;\to 0\quad\text{almost surely.}$$

The result now follows from (5.6.4) and the relation between almost sure convergence and convergence in probability. ☐

Remark 5.6.3. It is possible to improve the result of the previous lemma with elementary cut-off techniques and show that

$$\lim_{M\to\infty}\mathbb{E}[\mathcal{W}_1(Q\times\mathbb{R}^4)(f^\epsilon_{MN}(t),\bar f_N(t))]=0\quad\text{for every } t\geqslant 0,\ N\in\mathbb{N}\text{ and }\epsilon>0.$$

However, this method does not retain any information about the precise rate of convergence to 0, which in principle also depends on N. To keep track of this, we need to rely on a more sophisticated result by Fournier and Guillin [FG13] about the convergence of empirical laws of i.i.d particles towards their actual law.

Lemma 5.6.4. In the setting above, for any $T > 0$ and any $N \in \mathbb{N}$,

$$\sup_{t\in[0,T]} \mathbb{E}\left[\mathcal{W}_1(Q\times\mathbb{R}^4)(\bar{f}^{\epsilon}_{MN}(t), \bar{f}_N(t))\right] \leqslant C\left(1+\sup_{x\in Q}\mathbb{E}\left[|u_k(x,0)|^2\right]^{\frac{1}{2}}\right)\frac{1}{M^{\frac{1}{4}}},$$

for a constant $C = C(T, b, \sigma, Q)$.

Proof. The explicit expressions for $\bar{f}^{\epsilon}_{MN}$ and $\bar{f}_N$ and the convexity of the Wasserstein distance yield

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{W}_1(Q\times\mathbb{R}^4)\big(f^{\epsilon}_{MN}(t), \bar{f}_N(t)\big)\right]\\
&\leqslant \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathcal{W}_1(Q\times\mathbb{R}^4)\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{(X_j,\bar{u}^{\epsilon}_{jm})}, \delta_{X_j}\otimes f(t,X_j,dv)\right)\right]\\
&= \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}\left[\mathcal{W}_1(\mathbb{R}^4)\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{\bar{u}^{\epsilon}_{jm}}, f(t,X_j,dv)\right)\right].
\end{aligned} \tag{5.6.7}$$

Now, observe that for each fixed j, the particles $\bar{u}^{\epsilon}_{jm}(t)$ for $m = 1, \dots, M$ are i.i.d. with common law $f(t, X_j, dv)$. A direct application of Theorem 1 in [FG13], with $p = 1$, $q = 2$ and $d = 4$, implies

$$\mathbb{E}\left[\mathcal{W}_1(\mathbb{R}^4)\left(\frac{1}{M}\sum_{m=1}^{M}\delta_{\bar{u}^{\epsilon}_{jm}(t)}, f(t,X_j,dv)\right)\right] \leqslant C\,\mathbb{E}\left[|\bar{u}^{\epsilon}_{jm}(t)|^2\right] M^{-\frac{1}{4}}. \tag{5.6.8}$$

We conclude using the a priori estimates (5.2.15) and plugging formula (5.6.8) into (5.6.7). ☐

Finally we consider the last term in (5.6.1). For this *deterministic* term we have the following.

Lemma 5.6.5. In the setting above, for any $T \geqslant 0$,

$$\sup_{t\in[0,T]} \mathcal{W}_1(Q\times\mathbb{R}^4)(\bar{f}_N(t), f(t)) \leqslant C\frac{1}{N^{\frac{\alpha}{d}}},$$

for a constant $C = C(T, \rho, b, \sigma, [u(\cdot, 0)]_{\alpha})$.

Proof. For every $t \in [0, T]$, we consider the following pairing $\pi(t)$ defined by integration as

$$\begin{aligned}
&\int_{(Q\times\mathbb{R}^4)^2}\varphi(x,y,u,v)\,\pi(t,dx,dy,du,dv)\\
&\qquad = \sum_{j=1}^{N}\int_{Q^N_j}\int_{(\mathbb{R}^4)^2}\varphi(X_j,y,u,v)\,\pi_0(t,X_j,y,du,dv)\,dy \qquad \forall\varphi\in C_b\left((Q\times\mathbb{R}^4)^2\right),
\end{aligned}$$

where $\pi_0(t, X_j, y, du, dv)$ is a chosen optimal pairing for $\mathcal{W}_1(\mathbb{R}^4)(f(t, X_j, du), f(t, y, dv))$. Recalling that $\text{meas}(Q^N_j) = 1/N$, an elementary check shows that $\pi(t)$ is indeed a pairing. Taking $\varphi(x, y, u, v) = |x-y|+|u-v|$, using the definition of $\pi(t)$ and $\pi_0(t, X_j, y)$, and recalling formula

(5.2.17) we compute

$$\begin{aligned}
\mathcal{W}_1(\bar{f}_N(t), f(t)) &\leqslant \int_{(Q\times\mathbb{R}^4)^2} |x-y| + |u-v|\,\pi(t,dx,dy,du,dv) \\
&= \sum_{j=1}^{N} \int_{Q_j^N} \int_{(\mathbb{R}^4)^2} |X_j - y| + |u-v|\,\pi_0(t,X_j,y,du,dv)\,dy \\
&\leqslant \operatorname{diam}(Q_j^N) + \sum_{j=1}^{N} \int_{Q_j^N} \int_{(\mathbb{R}^4)^2} |u-v|\,\pi_0(t,X_j,y,du,dv)\,dy \\
&= \operatorname{diam}(Q_j^N) + \sum_{j=1}^{N} \int_{Q_j^N} \mathcal{W}_1(\mathbb{R}^4)(f(t,X_j,du), f(t,y,dv))\,dy \\
&\leqslant \operatorname{diam}(Q_j^N) + \sum_{j=1}^{N} \int_{Q_j^N} C\,|X_j - y|^{\alpha}\,dy \\
&\leqslant \operatorname{diam}(Q_j^N) + C \operatorname{diam}(Q_j^N)^{\alpha},
\end{aligned}$$

for a constant $C = C(T, \rho, b, \sigma, [u(\cdot, 0)]_\alpha)$. Recalling that $\operatorname{diam}(Q_j^N) = N^{-\frac{1}{d}}$ concludes the proof. □

Chapter 6

Fluctuations in the large number of particles limit

6.1 Introduction and main results

In this chapter we continue the analysis of the particle system introduced in Chapter 5 to model the grid cell network and we study the fluctuations around the mean field limit. Grid cells are a particular type of neuron in the brain of mammals discovered in 2005 [HFM+05] (see also the review [MMM17]). These neurons fire at regular intervals as an animal moves across an area, storing information such as position, direction and velocity, and thus enabling it to understand its movement in space.

Since their discovery, there has been extensive research to understand the precise behavior of grid cells, see [RRMM16, MMM17, Bre11] and the references therein. Mathematically, their network is commonly described by deterministic continuous attractor network dynamics through a system of neural field models [ET10, MBJ+06, BF09, CWZ+13], based on the classical papers [WC72, WC73, Ama77]. In particular, as the brain is inherently noisy, understanding how the grid cell network is affected by noise is one of the currently open challenges in the field. This question has recently been addressed from several directions [Bre19, BF12, BAC19, KE13, MB20, Tou12, CHS22, CRS23a].

The work presented here, published in [Cli23a], is the continuation of [CCS23] (discussed in Chapter 5), along the direction initiated in [CHS22]. In [CHS22] the authors studied a system of Fokker–Planck PDEs derived by adding noise to the attractor network models in [BF09, CWZ+13] and formally taking the mean field limit. The PDE has been further analyzed in [CRS23a]. The passage to the mean field limit has been rigorously proved in [CCS23], where the authors derive the limit for a generalized model, covering diverse concrete models commonly proposed (cf. [AB20, BF09, BF12, CWZ+13] and the review [Bre11]). The aim of the present work is to analyze the fluctuations of the empirical measure of the generalized noisy network model around its deterministic mean field limit and prove a central limit theorem (cf. Theorem 6.3.12 below).

Precisely, we consider the following model (already introduced in Chpater 5) for the interaction among NM neurons stacked along N columns, with M neurons each, at the locations $x_1, \ldots, x_N \in Q$ in a region Q of the neural cortex.

The behavior of neurons is characterized via their *activity level*, a quantity similar to the firing rate and derived from this through a defining equation (6.1.1). In concrete cases, the firing rate would explicitly appear on the right hand side of (6.1.1) (see example (6.2.3) below).

The k^{th} neuron at location x_i has an activity level u_{ik}^β for each possible orientation $\beta = 1, \ldots, d_v$, corresponding to spatial directions (typically $d_v = 4$). The total activity level $u_{ik} = (u_{ik}^1, \ldots, u_{ik}^{d_v})$ evolves according to the system

$$\begin{cases} u_{ik}(t) = u_{ik}(0) + \displaystyle\int_0^t b(x_i, r, u_{ik}(r), f_{N,M}(r))\, dr & (6.1.1a) \\ \qquad\qquad + \displaystyle\int_0^t \sigma(x_i, r, u_{ik}(r), f_{N,M}(r))\, dW_{ik}(r) - \ell_{ik}(t), & \\ \ell_{ik}^\beta(t) = -|\ell_{ik}^\beta|(t), \quad |\ell_{ik}^\beta|(t) = \displaystyle\int_0^t 1_{\{u_{ik}^\beta(r)=0\}} d|\ell_{ik}^\beta|(r) \quad \text{for } \beta = 1, \ldots, d_v. & (6.1.1b) \end{cases}$$

For simplicity, we take the cortex to be $Q = [0,1]^d$ and we extend everything periodically out of Q, since the network of neurons is commonly considered to have toroidal connectivity. The results in this work are easily extended to any bounded open subset $Q \subseteq \mathbb{R}^d$, for any $d \geqslant 1$.

In particular, we stress that we make no assumptions on the way the brain cortex is mapped onto the physical space i.e. the way points $x \in Q$ correspond to points in the area the animal is moving into. One can take Q to be a domain mapped via the identity function to physical space as in the original work [BF09] or any other domain mapped via some nontrivial function. Our result are indeed independent of this modelling choice.

For integers $k = 1, \ldots, M$, we have i.i.d. families $\{u_{ik}(0)\}_{i=1,\ldots,N}$ of random initial conditions for each space point x_i in the cortex Q. Moreover, for integers $i = 1, \ldots, N$ and $k = 1, \ldots, M$, we have d_v-dimensional Brownian motions $(W_{ik}^\beta)_{\beta=1,\ldots,d_v}$,which can also be correlated.

The properties and the modelling interpretation of the coefficients b and σ are given in Section 6.2.1 and are deduced from those of the concrete models considered, briefly reported in Section 6.2.1 (see also the introduction of [CCS23]). We also denoted $f_{N,M}$ the empirical measure associated to these particles, that is

$$f_{N,M}(r, dy, du) = \frac{1}{NM} \sum_{j=1}^{N} \sum_{m=1}^{M} \delta_{(x_j, u_{jm}(r))} \quad \text{regarded as a measure on } Q \times \mathbb{R}^{d_v}.$$

Finally, for each i, k and β, the term ℓ_{ik}^β is a finite variation process defined by (6.1.1b) which prevents the activity level u_{ik}^β from taking negative values. Namely, as we can see in its definition, at each time t this process equals the opposite of its total variation $\ell_{ik}^{\beta}(t) = -|\ell_{ik}^{\beta}|(t)$. In turn, the total variation stays constant when $u_{ik}^{\beta} > 0$ and it increases in the form $|\ell_{ik}^{\beta}|(t) = \int_0^t 1_{\{u_{ik}^{\beta}(r)=0\}} d|\ell_{ik}^{\beta}|(r)$ when $u_{ik}^{\beta} = 0$, so as to push u_{ik}^{β} away from zero which is being dragged by the other terms on the right hand side of (6.1.1a). The introduction of such terms and constraints is therefore known as imposing *reflecting boundary conditions* and ℓ_{ik}^{β} is called a *reflection term*. The existence and uniqueness of such a term need of course to be proved and this process is often referred to as the *Skorokhod problem*. Precise details concerning the well-posedness and the construction of the reflection term in our setting are all presented in the seminal papers [LS84, Szn84a] by Lions and Sznitman.

Remark 6.1.1. Rather than being a classical mean field system with MN particles, it is more accurate to regard the model (6.1.1) as a collection of N mean field systems with M particles, one for each column of neurons, interacting among themselves. Different neurons in the same column are subjected to independent sources of noise and are initiated with i.i.d. data. However, due to our continuum formalism in the space variable, we have to require some continuity of initial data and noise sources in the variable $x \in Q$ to avoid technical issues. So that two 'extremely close' columns of neurons could sense correlated noise. This will have consequences in some technical arguments: we will not be able to use independence arguments and will rely instead on the continuity of the quantities involved.

However, this is not a fault in the modelling: we can let this correlation radius be smaller than any finite size, e.g. smaller than the size of a neuron, after which our continuum formalism loses any meaning (see also Remark 6.1.3). From the point of view of the modelling, one should think of this continuous quantities as some interpolation of the actual quantities corresponding to the *space discrete* columns of neurons.

This feature is well-analyzed in [CCS23] (see Chapter 5) and results in the convergence rate $1/\sqrt{M} + 1/N^{\alpha/d}$ for the mean field limit $M, N \to \infty$ (cf. Theorem 6.2.8 below) in contrast to the usual decay rate $1/\sqrt{MN}$ for a mean field problem with MN particles. In turn, this aspect will have consequences on the correct scaling regime needed to see a nontrivial behavior of the fluctuations (cf. Remark 6.3.5 and Theorem 6.3.12 below).

Remark 6.1.2. We point out that another good modelling choice is to have the same noise hitting all the neurons in a given column, and possibly varying in different columns. That is, having terms dW_i (without the 'column' index k) on the right hand side of (6.1.1a). In this case, the empirical measure of the neuron network still converges to the law of the McKean–Vlasov SDE (6.1.2), but it also directly solves a *stochastic version* of the Fokker–Planck equation (6.1.4) below and the noise does not cancel out in the mean field limit. We refer to the seminal paper [CF16] for a prototypical case of this situation with no space interaction.

Under suitable assumptions, strong existence and uniqueness for (6.1.1) is proved in

[CCS23, Theorem 2.3] (see Chapter 5). In turn, the authors rigorously analyze the mean field limit of (6.1.1) as the number of columns N and the number of neurons per column M go to infinity. It is shown (cf. Theorem 6.2.8 below) that the particles converge to the solution of the family of McKean–Vlasov SDEs, for $x \in Q$,

$$\begin{cases} \bar{u}^\epsilon(x,t) = u(x,0) + \int_0^t b(x,r,\bar{u}^\epsilon(x,r),f(r))\,dr & (6.1.2a) \\ \qquad\qquad + \int_0^t \sigma(x,r,\bar{u}^\epsilon(x,r),f(r))\,dW^\epsilon(x,r) - \bar{\ell}(x,t), \\ \bar{\ell}^\beta(x,t) = -\,|\bar{\ell}^\beta(x,\cdot)|(t),\ \ |\bar{\ell}^\beta(x,\cdot)|(t) = \int_0^t 1_{\{(\bar{u}^\epsilon)^\beta(x,r)=0\}}\,d|\bar{\ell}^\beta(x,\cdot)|(r) \ \ \beta = 1,\dots,d_v. & (6.1.2b) \end{cases}$$

Here we have a family of random initial data $(u(x,0))_{x\in Q}$ for each point x in the cortex. The term $W^\epsilon(x,t)$ denotes a Gaussian noise term, white in time and ϵ-correlated in space. The space correlation is needed since equation (6.1.2a) would be ill-posed for an actual space-time white noise, and in fact it is also meaningful from the modelling point of view (cf. [CCS23, Remark 1.2] and Remark 6.1.3 below). The precise definition and further properties of $W^\epsilon(x,t)$ are given in Section 6.2.1.

The coefficients b and σ are the same as in (6.1.1), but here we have set $f(t,y,du) := \mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}(y,t))$, considered as a measure on $\mathbb{R}^{d_v}$ depending on $t \in [0,\infty)$ and $y \in Q$. This in turn induces a probability measure $f(t,dx,du)$ on $Q \times \mathbb{R}^{d_v}$ defined by integration as

$$\int_{Q\times\mathbb{R}^{d_v}} \varphi(x,u) f(t,dx,du) := \int_Q \int_{\mathbb{R}^{d_v}} \varphi(x,u) f(t,x,du)\,dx \qquad \text{for any } \varphi \in C_b(Q\times\mathbb{R}^{d_v}). \tag{6.1.3}$$

Finally, for each fixed $x \in Q$ and $\beta = 1,\dots,d_v$, the finite variation process $\bar{\ell}^\beta(x,t)$ defined in (6.1.2b) is again the reflection term coming from the Skorokhod problem (see the explanation after equation (6.1.1)) and it ensures that $(\bar{u}^\epsilon)^\beta(x,t) \geqslant 0$ for each x, t and β. We refer the reader to [Szn84a] for the details about such a process in the context of a classical McKean–Vlasov equation.

A formal application of Itô formula shows that the law $f(t,x,du) := \mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}^\epsilon(x,t))$ is a weak solution of the Fokker–Planck equation with no-flux boundary conditions

$$\begin{cases} \partial_t f(t,x,u) + \nabla_u \cdot \Big(b(x,t,u,f(t))f(t,x,u)\Big) = \frac{1}{2}\sum_{\beta=1}^{d_v} \frac{\partial^2}{\partial u^\beta \partial u^\beta}\Big(\sigma_\beta(x,t,u,f(t))^2 f(t,x,u)\Big), & (6.1.4a) \\ b_\beta(t,x,u,f(t))f(t,x,u) - \frac{1}{2}\frac{\partial}{\partial u^\beta}\Big(\sigma_\beta(x,t,u,f(t))^2 f(t,x,u)\Big)\Big|_{u^\beta=0} = 0 \ \text{ for } \beta = 1,\dots,d_v. & (6.1.4b) \end{cases}$$

The meaning of the boundary conditions (6.1.4b) is that the PDE (6.1.4a) is satisfied in the weak sense when tested against any sufficiently smooth function ψ with $\nabla_u\psi(u)\cdot \mathbf{n}_{\partial\mathbb{R}^{d_v}_+} \equiv 0$.

Under suitable assumptions, Theorems 2.4 and 2.5 in [CCS23] (see Chapter 5) ensure strong existence and uniqueness for the McKean–Vlasov equation (6.1.2) and that the law $f(t,x,du) := \mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}(x,t))$ is the unique weak solution of the Fokker–Planck PDE (6.1.4).

One of the key features of the model is that equation (6.1.4) has a unique solution independent of $\epsilon > 0$. That is to say, the law $f(t,x,du) = \mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}^\epsilon(y,t))$ of a single McKean–Vlasov particle is independent of the noise correlation radius ϵ. This is because in (6.1.2) the interaction between two neurons $\bar{u}^\epsilon(x,t)$ and $\bar{u}^\epsilon(y,t)$ happens only through their law f as single random variables. In fact we will see that the dependence on the correlation radius is recovered as soon as we consider the joint law $f(x,y,t,,du,dv) = \mathrm{Law}_{\mathbb{R}^{2d_v}}(\bar{u}^\epsilon(x,t),\bar{u}^\epsilon(y,t))$ of two or more particles (cf. Remark 6.2.9).

We also remark that the Fokker–Planck PDE (6.1.4), in the case of diffusion term $\sigma \equiv \sigma_0$ constant, has been further analyzed in [CHS22, CRS23a]. Under particular assumptions on the drift term, it is shown that for every noise strength $\sigma_0 > 0$ there exists a unique stationary space homogeneous solution (from the modelling point of view this corresponds to the animal completely loosing track of its position in space). This state is asymptotically stable for noise σ_0 big enough, but phase transition with multiple bifurcations, whose shape is also characterized locally, occurs as σ_0 decreases.

Finally, in [CCS23] (cf. Chapter 5) it is shown that the empirical measure f_{MN} of the particle system (6.1.1) converges in a suitable sense (see Theorem 6.2.11 below) to the weak solution f of the Fokker–Planck equation (6.1.4). Informally, this yields the zeroth order expansion, as elements of $\mathscr{P}(Q \times \mathbb{R}^{d_v})$,

$$f_{MN} = f + O\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right) \quad \text{as } M,N \to \infty. \tag{6.1.5}$$

The aim of the present work is to sharpen this picture by proving a central limit theorem for the fluctuations of the empirical measure around the deterministic limit f. Precisely, we consider the rescaled fluctuations $\eta_t^{MN} := \sqrt{M}(f_{MN} - f)$. Proposition 6.3.1 below proves that, for suitable test functions ψ, the fluctuations admits the semimartingale expression

$$\langle \eta_t^{MN}, \psi\rangle = \langle \eta_0^{MN}, \psi\rangle + \int_0^t \langle \eta_r^{MN}, \mathscr{L}_r(f, f_{MN})[\psi]\rangle\, dr + M_t^{MN}(\psi), \tag{6.1.6}$$

where

$$M_t^{MN}(\psi) = \frac{\sqrt{M}}{MN}\sum_{i=i}^{N}\sum_{k=1}^{M}\int_0^t \nabla_u\psi(x_i, u_{ik}(r))\sigma(x_i, r, u_{ik}(r), f_{MN})\, dW_{ik}(r)$$

is a martingale term coming from the stochasticity of f_{MN} and $\mathscr{L}(f, f_{MN})$ is a suitable differential operator, essentially the linearization of the generator of the SDE (6.1.1) around f. The aforementioned result already ensures that $f_{MN} \to f$ as $M,N \to \infty$. A formal central limit theorem argument suggests also that $\eta_0^{MN} \to \eta_0$ for some Gaussian random variable in a suitable distribution space. Finally, by analyzing the quadratic variation of the martingale M_t^{MN} (cf. Proposition 6.3.2 below) we expect this to converge to some Gaussian process G_t^ϵ in a suitable distribution space. In turn, from (6.1.6), we conjecture that η_t^{MN} converges to the solution of the Langevin SPDE (in integral form)

$$\eta_t^\epsilon = \eta_0 + \int_0^t \mathscr{L}_r(f,f)^*[\eta_r^\epsilon]\, dr + G_t^\epsilon, \tag{6.1.7}$$

where the operator $\mathscr{L}_r(f, f)$ is basically the linearization of the generator of the McKean–Vlasov SDE (6.1.2) around f, and η_0 and G^ϵ_t are the above mentioned Gaussian terms. The aim of this paper is to make this ansatz rigorous. Here $\alpha \in (0, 1]$ denotes the Hölder regularity in space of the coefficients b and σ in (6.1.1) (cf. Section 6.2.1).

Theorem (Theorem 6.3.12 below)**.** For every $\epsilon > 0$ fixed, along a scaling regime $M, N \to \infty$ such that $\sqrt{M}N^{-\alpha/d} \to 0$, the fluctuations $\eta^{MN}_t = \sqrt{M}(f_{MN} - f)$ of the empirical measure f_{MN} of the particle system (6.1.1) around its deterministic limit, the solution f of the Fokker–Planck equation (6.1.4), satisfy

$$\eta^{MN}_t \to \eta^\epsilon_t \quad \text{in law in } C\left([0, T]; \mathcal{H}\right),$$

where η^ϵ_t is the unique solution of the Langevin SDPE (6.1.7) and $\mathcal{H}$ is a suitable Hilbert space of distributions on $Q \times \mathbb{R}^{d_v}_+$.

The decay involving $M^{-1/2}$, related to the mean field interaction along columns, and $N^{-\alpha/d}$, related to the correlation between different columns, already appears in the mean field expansion (6.1.5). The rescaling factor $\sqrt{M}$ in front of the fluctuations corresponds to the classical convergence rate for a mean field problem with M particles subjected to independent sources of noise and is needed so that the fluctuations do not to vanish in the limit (cf. [HM86, FFG92, BFG93, FM97]). However, as discussed in Remark 6.1.1, we actually have N correlated 'M-particle mean field clusters' interacting among themselves with strength $N^{-\alpha/d}$. The scaling regime $\sqrt{M}N^{-\alpha/d} \to 0$ serves to prevent these clusters from resonating and resulting in yet unbounded fluctuations (cf. Remark 6.3.5).

In the same fashion as (6.1.5), the theorem above yields the first order expansion

$$f_{MN} = f + \frac{1}{\sqrt{M}}\eta^\epsilon_t + o\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right) \quad \text{as } M, N \to \infty \text{ and } \sqrt{M}N^{-\alpha/d} \to 0. \tag{6.1.8}$$

This is particularly relevant from the simulation point of view: in order to understand the behavior of a large number of grid cells obeying the model (6.1.1), simulating the continuous model $f + M^{-1/2}\eta^\epsilon_t$ is numerically much cheaper than simulating MN interacting SDEs. On the other hand, our main interest is the effect of noise on grid cells. The zero order term f alone completely loses track of the noise present at the particle level, which is instead retained by the first order expansion (6.1.8) thus furnishing a sharper description.

Remark 6.1.3. We point out that we could also let $\epsilon = \epsilon(M, N)$ vary with the number of neurons and consider a joint scaling regime $M, N \to \infty$ and $\epsilon(M, N) \to \bar{\epsilon}$. As long as $\bar{\epsilon} > 0$, all the results of the paper hold with virtually identical proofs. Some results and estimates also extend to the case $\epsilon(M, N) \to 0$, possibly after imposing some constraint on the scaling regime between $\epsilon(M, N) \to 0$ and $N \to \infty$. However, the point is that in the case $\epsilon(M, N) \to 0$, the limiting objects for the fluctuations and, even before, the supposed mean field limit (6.1.2) for the SDE system (6.1.1) is mathematically ill-posed. For $\epsilon = 0$ the noise $W^\epsilon(x,t)$ featuring in the McKean–Vlasov equation (6.1.2) makes sense only as a distribution (see e.g. [DPZ92, Chapter 4]) and in turn the equation, even in the simpler case $\sigma(x,t,u,f) \equiv \sigma_0$ constant, involves applying the drift nonlinearity $b(\dots)$ to a distribution. It is not clear whether it is even possible to give a meaning to equation (6.1.2) for $\epsilon = 0$ (see also [CCS23, Remark 1.2]) and in turn to the joint law $f^{2,\epsilon}$, introduced in (6.2.20) below, featuring in the definition (6.3.10) of the quadratic variation of the fluctuation limit η_t^ϵ.

Overview of the methods.

The methods of this paper are deeply based on the approach to fluctuations put forward in Fernandez and Méléard [FM97], Bezandry, Ferland, Fernique and Giroux [BFG93, FFG92] and Hitsuda and Mitoma [HM86]. Our situation presents additional difficulties: unboundedness of the coefficients, boundary conditions imposing positivity of the activity level of the neurons, and interaction between different mean-field families of coupled neurons across the cortex.

The fluctuations $\eta_t^{MN} = \sqrt{M}(f_{MN} - f)$ are random signed measures in the path space over $Q \times \mathbb{R}^{d_v}$, whose space part is independent both of the time and the stochasticity. The first problem is to find a distribution space in which both η_t^{MN} and its limit belong. In Theorem 6.3.12 we prove that, along the scaling regime $\sqrt{M}N^{-\alpha/d} \to 0$, the fluctuations converge in law in $C\big([0,T]; H_x^{-\,\mathrm{e}_\alpha^+} \mathscr{H}^{-\,\kappa_2,\theta_2}\big)$. The spaces $H_x^{\mathrm{e}} \mathscr{H}^{\kappa,\theta}$, introduced in Section 6.2.1, are vector valued Sobolev spaces taking values in other weighted Sobolev spaces with suitable boundary conditions needed to account for the reflecting boundary conditions at the SDE level, which reduce the class of admissible test functions, and the space interaction among neurons. The particular exponents e, κ and θ, given in (6.2.17), depend on the dimension of the spaces $Q \subseteq \mathbb{R}^d$ and $\mathbb{R}^{d_v}$ to furnish suitable Sobolev embeddings. It will appear from the proofs that the space $H_x^{-\,\mathrm{e}_\alpha^+} \mathscr{H}^{-\,\kappa_2,\theta_2}$ is optimal and provides the minimal Hilbert space in which to embed the fluctuations (cf. Remark 6.2.4).

Our strategy revolves around the semimartingale equation (6.1.6) for η_t^{MN}. Estimates on the norm of M_t^{MN} are readily obtained with martingale arguments (cf. Proposition 6.3.3). Here the main difficulty is that the operator $\mathscr{L}(f, f_{MN})$ featuring in (6.1.6) is a second order differential operator and thus not bounded in any particular space of the form $H_x^{\mathrm{e}} \mathscr{H}^{\kappa,\theta}$, since it reduces the regularity of the test functions. Hence we cannot use equation (6.1.6) and Grönwall's Lemma to estimate the norm of η_t^{MN}. To circumvent this problem, we work between two nested spaces $H_x^{\mathrm{e}_\alpha^+} \mathscr{H}^{\kappa_2,\theta_2} \subseteq H_x^{\mathrm{e}_\alpha} \mathscr{H}^{\kappa_1,\theta_1}$ where the operator $\mathscr{L}(f, f_{MN})$ results bounded (cf. Lemma 6.3.6) and we obtain the needed estimates on η_t^{MN} (cf. Lemma 6.3.4 and Proposition 6.3.7). It is the combination of the estimates for M_t^{MN} and η_t^{MN} that imposes the particular scaling regime $\sqrt{M}N^{-\alpha/d} \to 0$ (cf. Remark 6.3.5).

These estimates are readily translated into tightness of the processes M_t^{MN} and η_t^{MN}

thanks to the compactness of the embedding $H_x^{e_\alpha^+}\mathscr{H}^{\kappa_2,\theta_2} \subseteq H_x^{e_\alpha}\mathscr{H}^{\kappa_1,\theta_1}$ and martingale methods (cf. Proposition 6.3.8). A fine analysis based on Sznitman's coupling method [Szn84a] shows that the quadratic variation of the martingale term M_t^{MN} is converging to a deterministic limit (cf. Proposition 6.3.2) and this uniquely identifies a Gaussian process G_t^ϵ such that $M_t^{MN} \to G_t^\epsilon$ (cf. Proposition 6.3.10). An adaptation of the standard Lévy central limit theorem shows the convergence of the initial data η_0^{MN} (cf. Lemma 6.3.11). Theorem 6.3.12 follows from passing to the limit in equation (6.1.6) and the well-posedness of the Langevin SPDE (6.1.7).

Structure of the work.
The work is structured as follows. In the next section we first present our hypotheses and notation and deduce some immediate consequences for later convenience. Then we lay out the precise setting for our results and we collect some auxiliary results needed in our arguments. In Section 6.3 we establish our central limit theorem following the strategy discussed above. In Section 6.4 we collect the proofs of the results.

6.2 Preliminaries and auxiliary results

6.2.1 Hypotheses and notation

In this section we introduce our hypotheses and notations. We first take care of the random objectes: the initial data and the noise terms in the McKean–Vlasov equation (6.1.2). The specific filtered probability space supporting all of these random variables is introduced rigorously in Setting 2.4 below. The initial data and the noise in the particle system (6.1.1) will then be given by the corresponding quantities at the locations $x_i \in Q$ (cf. Setting 6.2.5). As regards the initial data, we will consider random families $(u(x,0))_{x\in Q}$ such that

$$u(\cdot,0) \in C^\alpha\big(Q;L^4(\Omega)\big) \cap L^\infty\big(Q;L^{4\,\theta_1}(\Omega)\big) \text{ for some } \alpha \in (0,1] \text{ and } \theta_1 > 1+\frac{d_v}{2}+\max\left\{2,\frac{d_v}{2}\right\}. \quad (6.2.1)$$

The Hölder continuity in space is needed to control the interaction among columns at different locations and is motivated by the modelling, as we expect the randomness to be somewhat continuous along the cortex. The moments will serve to obtain bounds for the fluctuations in the dual space of suitably regular function spaces (cf. (6.2.16)).

As regards the noise, we consider a Gaussian random field $W^\epsilon : \Omega \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}^{d_v}$ with independent components $\beta = 1\ldots d_v$, zero mean and covariance

$$\mathbb{E}\left[W^{\epsilon,\beta}(x,t)W^{\epsilon,\beta}(y,s)\right] = (t\wedge s)\,C_\rho\,\epsilon^d \int_{\mathbb{R}^d}\rho_\epsilon(z-x)\rho_\epsilon(z-y)\,dz, \quad \text{for } C_\rho = \left(\int_{\mathbb{R}^d}\rho(z)^2\,dz\right)^{-1}, \quad (6.2.2)$$

where $\rho : \mathbb{R}^d \to [0,1]$ is a radial mollifier supported in the unitary ball and ρ_ϵ is its the ϵ-rescaled version. Such a process can be obtained for example from convolution and rescaling of independent copies W^β of an actual distribution valued space-time white noise (see e.g.

[DPZ92, Chapter 4]), by setting

$$W^{\epsilon,\beta}(x,t) := C_\rho^{\frac{1}{2}}\,\epsilon^{\frac{d}{2}}\,\langle W_t, \rho_\epsilon(\cdot - x)\rangle.$$

As mentioned in Remark 6.1.3, the ϵ-correlation in space is needed to avoid mathematical issues and in fact, as for the initial data, is strongly motivated by the modelling.

For future reference, we highlight some of the properties of W^ϵ. First, from (6.2.2) we have that $\mathbb{E}[W^{\epsilon,\beta}(x,t)W^{\epsilon,\gamma}(x,s)] = \delta_{\beta\gamma}\; t \wedge s$. Thus, for fixed x, the process $t \mapsto W^\epsilon(x,t)$ is a d_v-dimensional Brownian motion. Similarly, from $\sup(\rho) \subseteq B(0,1)$ it follows that

$$\mathbb{E}\left[W^\epsilon(x,t)W^\epsilon(y,s)\right] = 0 \;\text{ if }\; |x-y| > 2\epsilon.$$

Hence the processes $W^\epsilon(x,t)$ and $W^\epsilon(y,t)$ are independent for $|x-y| > 2\epsilon$. Furthermore, using (6.2.2), we compute

$$\begin{aligned}\mathbb{E}\Big[|W^\epsilon(x,t) - W^\epsilon(y,s)|^2\Big] &\leqslant C\,\mathbb{E}\Big[|W^\epsilon(x,t) - W^\epsilon(x,s)|^2 + |W^\epsilon(x,s) - W^\epsilon(y,s)|^2\Big] \\ &\leqslant C\left(|t-s| + s\,C_\rho\,\epsilon^d \int_{\mathbb{R}^d} \left(\rho_\epsilon(z-x) - \rho_\epsilon(z-y)\right)^2\, dz\right) \\ &\leqslant C\left(|t-s| + \frac{|x-y|^2}{\epsilon^2}\right),\end{aligned}$$

for a constant $C = C(\rho)$. Similar estimates hold for any higher moment $p \geqslant 2$ and the Kolmogorov continuity theorem ensures the existence of a suitable modification of W^ϵ with continuous trajectories in both x and t. In particular, we have that W^ϵ is jointly measurable in $(x,t) \in \mathbb{R}^d \times \mathbb{R}^+$ and in the sample path $\omega \in \Omega$. Finally, for any $x, y \in \mathbb{R}^d$, a direct computation shows that the quadratic variation of the martingale $W^{\epsilon,\beta}(x,t) - W^{\epsilon,\beta}(y,t)$ satisfies

$$\left\langle W^{\epsilon,\beta}(x,t - W^{\epsilon,\beta}(y,t)\right\rangle = t\,C_\rho\,\epsilon^d \int_{\mathbb{R}^d} \left(\rho_\epsilon(z-x) - \rho_\epsilon(z-y)\right)^2\, dz \leqslant t\,C\,\frac{|x-y|^2}{\epsilon^2},$$

for a constant $C = C(\rho)$.

Next we present the assumptions on the coefficients of the equations. We remark that the specific shape and properties of the generalized coefficients are deduced from those of the concrete model analyzed in [CCS23]

$$\begin{aligned}\mathrm{d}u_{ik}^\beta(t) =&\sigma\,\mathrm{d}W_{ik}^\beta(t) - \mathrm{d}\ell_{ik}^\beta(t) \\ &+ \left(-u_{ik}^\beta(t) + \phi^\beta\left(B^\beta(x_i,r) + \frac{1}{4NM}\sum_{\gamma=1}^{4}\sum_{j=1}^{N}\sum_{m=1}^{M} K^\gamma(x_i - x_j)u_{jm}^\gamma(t)\right)\right)\,dt.\end{aligned} \tag{6.2.3}$$

The nonlinearities $\phi^\beta : \mathbb{R} \to \mathbb{R}$ represent the firing rate of neurons in the network, typically a sigmoid or a ReLU function. The external input is represented by $B^\beta : Q \times \mathbb{R} \to \mathbb{R}$. The interaction kernels $K^\beta : \mathbb{R}^{d_Q} \to \mathbb{R}$ take into account the inhibitory/excitatory effect on nearby neurons. A typical choice for these [BF09] is the so-called Mexican hat function.

We consider coefficients $b, \sigma : Q \times \mathbb{R}^+ \times \mathbb{R}^{d_v} \times \mathscr{P}(Q \times \mathbb{R}^{d_v}) \to \mathbb{R}^{d_v}$ of the form

$$\begin{aligned}b^\beta(x,r,u,\mu) &= b_0^\beta(x,r,u) + \phi^\beta\left(\int_{Q\times\mathbb{R}^{d_v}} b_1^\beta(x,y,r,u,v)\,\mu(dy,dv)\right), \\ \sigma^\beta(x,r,u,\mu) &= \sigma_0^\beta(x,r,u) + \phi^\beta\left(\int_{Q\times\mathbb{R}^{d_v}} \sigma_1^\beta(x,y,r,u,v)\,\mu(dy,dv)\right),\end{aligned} \tag{6.2.4}$$

for suitable functions $b_0, \sigma_0 : Q \times \mathbb{R}^+ \times \mathbb{R}^{d_v} \to \mathbb{R}^{d_v}$ and $b_1, \sigma_1 : Q^2 \times \mathbb{R}^+ \times \mathbb{R}^{2d_v} \to \mathbb{R}^{d_v}$, and $\phi : \mathbb{R} \to \mathbb{R}^{d_v}$.

Remark 6.2.1. In order simplify the exposition and the computations, we shall assume the coefficients are the same for every direction. That is to say, for every $\beta = 1, \ldots, d_v$,

$$b^\beta(x, r, u, \mu) = b_0(x, r, u) + \phi\left(\int_{Q\times\mathbb{R}^{d_v}} b_1(x, y, r, u, v)\,\mu(dy, dv)\right),$$

and similarly for the coefficients σ, for suitable functions b_0, b_1, σ_0, σ_1 and ϕ satisfying the assumptions below. We stress that all the results hold identically with identical proofs for general direction dependent coefficients b^β, σ^β and ϕ_β satisfying the assumptions below for each single β. Furthermore, to ease the notation, we will also use the shorthand

$$\begin{aligned} b_1(x, t, u, \mu) &:= \int_{Q\times\mathbb{R}^{d_v}} b_1(x, y, t, u, v)\,\mu(dy, dv), \\ \sigma_1(x, t, u, \mu) &:= \int_{Q\times\mathbb{R}^{d_v}} \sigma_1(x, y, t, u, v)\,\mu(dy, dv). \end{aligned} \tag{6.2.5}$$

We assume the nonlinearity ϕ is smooth with bounded derivatives. Precisely we assume that

$$\frac{d^j\phi}{dz^j}(z) \in L^\infty(\mathbb{R}) \quad \forall\, j \geqslant 1. \tag{6.2.6}$$

In particular ϕ is globally Lipschitz and with sublinear growth. We assume that b_0 and σ_0 are smooth in the x and u variables, uniformly α-Hölder continuous in the x-variable (for $\alpha \in (0, 1]$ given in (6.2.1)) and with u-derivatives bounded uniformly in x and t. That is to say, for every $h, k \in \mathbb{N}$, for suitable constants C,

$$\begin{aligned} |D_x^h D_u^{k+1} b_0(x, t, u)| &\leqslant C, \\ \left|D_x^h b_0(x, t, u) - D_x^h b_0(x', t, u)\right| &\leqslant C|x - x'|^\alpha, \end{aligned} \qquad \forall x, x', t, u \in Q^2 \times \mathbb{R}_+ \times \mathbb{R}^{d_v}, \tag{6.2.7}$$

and analogously for σ_0. In particular this implies Lipschitzianity and sublinear growth in u. That is, for every $h, k \in \mathbb{N}$, for suitable constants C,

$$\begin{aligned} \left|D_x^h b_0(x, t, u)\right| &\leqslant C(1 + |u|), \\ \left|D_x^h D_u^k b_0(x, t, u) - D_x^h D_u^k b_0(x, t, u')\right| &\leqslant C|u - u'|, \end{aligned} \qquad \forall x, t, u, u' \in Q \times \mathbb{R}_+ \times \mathbb{R}^{2d_v},$$

and analogously for σ_0. Similarly, we assume that b_1 and σ_1 are smooth in the x, y and u, v variables, uniformly α-Hölder in the x, y-variable and with u, v-derivatives bounded uniformly in x, y and t. That is, for every $h, k \in \mathbb{N}$, for suitable constants C, for every $x, x', y, y', t, u, v \in Q^4 \times \mathbb{R}_+ \times \mathbb{R}^{2d_v}$,

$$\begin{aligned} |D_{x,y}^h D_{u,v}^{k+1} b_1(x, y, t, u, v)| &\leqslant C, \\ \left|D_{x,y}^h b_1(x, y, t, u, v) - D_{x,y}^h b_0(x', y', t, u, v)\right| &\leqslant C|x - x'|^\alpha + |y - y'|^\alpha, \end{aligned}$$

and analogously for σ_1. In particular this implies Lipschitzianity and sublinear growth in u and v. That is, for every $h, k \in \mathbb{N}$, for suitable constants C, for every $x, y, t, u, u', v, v' \in Q^2 \times \mathbb{R}_+ \times \mathbb{R}^{4d_v}$,

$$\begin{aligned} \left| D^h_{x,y} b_1(x,y,t,u,v) \right| &\leqslant C(1 + |u| + |v|), \\ \left| D^h_{x,y} D^k_{u,v} b_1(x,y,t,u,v) - D^h_{x,y} D^k_{u,v} b_1(x,y,t,u',v') \right| &\leqslant C(|u-u'| + |v-v'|), \end{aligned} \tag{6.2.8}$$

and analogously for σ_1.

The shape (6.2.4) of the coefficients and assumptions (6.2.6)-(6.2.8) immediately translate into the following properties of b and σ.

Lemma 6.2.2. The coefficients $b, \sigma : Q \times \mathbb{R}^+ \times \mathbb{R}^{d_v} \times \mathscr{P}(Q \times \mathbb{R}^{d_v}) \to \mathbb{R}$ are smooth in the x and u variables. Furthermore we have, for every $h, k \in \mathbb{N}$, for suitable constants $C \geqslant 0$, for every $x, x', t, u, u', \mu, \mu' \in Q^2 \times \mathbb{R}_+ \times \mathbb{R}^{2d_v} \times \mathscr{P}(Q \times \mathbb{R}^{d_v})^2$,

$$\begin{aligned} &|D^h_x D^{k+1}_u b(x,t,u,\mu)| \leqslant C, \\ &|D^h_x b(x,t,u,\mu)| \leqslant C \left(1 + |u| + \int_{Q \times \mathbb{R}^{d_v}} |v| \, \mu(dy, dv) \right), \\ &|D^h_x D^k_u b(x,t,u,\mu) - D^h_x D^k_u b(x',t,u',\mu')| \\ &\qquad \leqslant C \Big(|x-x'|^\alpha + |u-u'| + \mathcal{W}_1(Q \times \mathbb{R}^{d_v})(\mu,\mu')^\alpha + \mathcal{W}_1(Q \times \mathbb{R}^{d_v})(\mu,\mu') \Big), \end{aligned} \tag{6.2.9}$$

and analogously for σ, where $\mathcal{W}$ denotes the Wasserstein distance (see formula (6.2.10) below).

We now introduce the function spaces needed to state our results. First of all, given a Banach space X, we consider the space $\mathscr{P}_m(X)$ of probability measures on X with finite m^{th} moment, which is a Polish space when endowed with the m^{th} order Wasserstein distance

$$\mathcal{W}_m(X)(P,Q) := \inf \left\{ \int_{X^2} |x-y|^m \, d\pi \mid \pi \text{ pairing between } P \text{ and } Q \right\}^{1/m}. \tag{6.2.10}$$

Next we introduce the following weighted Sobolev spaces with no flux boundary conditions. We denote by $W^{k,p}_{loc}$ the space of distributions having the first k derivatives locally p-integrable. For any integer $k \geqslant 2$ and any $\theta \geqslant 0$, we define

$$\mathscr{H}^{k,\theta} := \left\{ \psi \in W^{k,2}_{loc}(\mathbb{R}^{d_v}_+) \mid \nabla\psi(u) \cdot \mathbf{n}_{\partial\mathbb{R}^{d_v}_+} \equiv 0, \quad \sum_{j=0}^k \int_{\mathbb{R}^{d_v}_+} \frac{|D^j_u \psi(u)|^2}{1+|u|^{2\theta}} \, du < \infty \right\},$$

which are Hilbert spaces when endowed with the norm $\|\psi\|_{\mathscr{H}^{k,\theta}} = \left(\sum_{j=0}^k \int_{\mathbb{R}^{d_v}_+} \frac{|D^j_u \psi(u)|^2}{1+|u|^{2\theta}} \, du \right)^{1/2}$. Similarly, we will also consider the Banach spaces

$$\mathscr{W}^{k,\infty,\theta} := \left\{ \psi \in W^{k,\infty}_{loc}(\mathbb{R}^{d_v}_+) \mid \nabla\psi(u) \cdot \mathbf{n}_{\partial\mathbb{R}^{d_v}_+} \equiv 0, \quad \sup_{u \in \mathbb{R}^{d_v}_+} \sum_{j=0}^k \frac{|D^j_u \psi(u)|}{1+|u|^\theta} < \infty \right\},$$

endowed with the analogous norm $\|\psi\|_{\mathscr{W}^{k,\infty,\theta}} = \sup_{u \in \mathbb{R}^{d_v}_+} \sum_{j=0}^k \frac{|D^j_u \psi(u)|}{1+|u|^\theta}$.

Remark 6.2.3. These spaces are particularly adapted to our needs because of the no-flux condition $\nabla\psi(u)\cdot \mathbf{n}_{\partial\mathbb{R}^{d_v}_+} \equiv 0$. This ensures that ψ is an admissible test function for the Fokker-Planck equation (6.1.4) according to our definition of weak solution. Furthermore it ensures that, when applying Itô formula to $\psi(u_{ik}(t))$ for u_{ik} solving (6.1.1) or to $\psi(\bar{u}^\epsilon(x,t))$ for $\bar{u}^\epsilon(x,t)$ solving (6.1.2) respectively, the resulting reflection term

$$\int_0^t \nabla\psi(u_{ik}(r))\, d\ell_{ik}(r) = -\int_0^t \nabla\psi(u_{ik}(r))\cdot \mathbf{n}_{\partial\mathbb{R}^{d_v}_+}\, 1_{\{u_{ik}\in\partial\mathbb{R}^{d_v}_+\}}\, d|\ell_{ik}|(r) = 0,$$

or $\int_0^t \nabla\psi(\bar{u}^\epsilon(x,r))d\ell(x,r) = 0$ respectively, is identically zero.

Classical embedding results extend to these particular Sobolev spaces (see e.g. [AF03]). In particular, we will make use of the following *continuous* inclusions

$$\mathscr{H}^{k+j,\theta} \hookrightarrow \mathscr{W}^{j,\infty,\theta} \quad \forall\, k > \frac{d_v}{2} \;\; \forall\, j \geqslant 0 \;\; \forall\, \theta \geqslant 0, \tag{6.2.11}$$

and the following *compact* inclusions

$$\mathscr{H}^{k+j,\theta} \hookrightarrow\hookrightarrow \mathscr{H}^{j,\theta+\beta} \quad \forall\, k > d_v/2 \;\; \forall\, \beta > d_v/2 \;\; \forall\, j \geqslant 0 \;\; \forall\, \theta \geqslant 0.$$

Given a Hilbert space V we will also consider the vector valued Hölder and Sobolev spaces, for some $e \geqslant 0$,

$$C^\alpha_x V := C^\alpha(Q;V), \quad H^e_x V := W^{e,2}(Q;V).$$

In particular, $H^e_x V$ is a Hilbert spaces with the usual Sobolev norm (cf. Remark 6.2.4 below). We will need the following vector valued versions of Sobolev embeddings (see e.g. [Ama00]). For any Hilbert space V we have the continuous embedding

$$H^e_x V \hookrightarrow C^\alpha_x V \quad \forall\, e > \alpha + d/2.$$

For any compact embedding $V \hookrightarrow\hookrightarrow V_0$ of Hilbert spaces, we have the compact embedding

$$H^{e^+}_x V \hookrightarrow\hookrightarrow H^e_x V_0 \quad \forall\, e \geqslant 0, \tag{6.2.12}$$

where we denoted $e^+ := e + \gamma$ for any arbitrary $\gamma > 0$.

In particular, combining (6.2.11)-(6.2.12) we obtain the following embeddings, which we collect here for later convenience,

$$H^e_x\, \mathscr{H}^{k+j,\theta} \hookrightarrow C^\alpha_x \mathscr{W}^{j,\infty,\theta} \text{ continuously} \quad \forall e > \frac{d}{2} + \alpha \;\; \forall\, k > \frac{d_v}{2} \;\; \forall j,\theta \geqslant 0, \tag{6.2.13}$$

$$H^{e^+}_x\, \mathscr{H}^{k+j,\theta} \hookrightarrow\hookrightarrow H^e_x\, \mathscr{H}^{j,\theta+\beta} \text{ compactly} \quad \forall e \geqslant 0 \;\; \forall\, k,\beta > \frac{d_v}{2} \;\; \forall j,\theta \geqslant 0. \tag{6.2.14}$$

Remark 6.2.4. The spaces $H^e_x\, \mathscr{H}^{\kappa,\theta}$ are introduced for the sole reason of working with subspaces of $C^\alpha(Q;\mathscr{H}^{\kappa,\theta})$ possessing a Hilbert space structure so as to simplify many arguments in the following. In fact, results completely analogous to those presented could be obtained by working directly in the Banach spaces $C^\alpha(Q;\mathscr{H}^{\kappa,\theta})$. However, having no Hilbert space structure at disposal, the methods of this paper need to be carefully adapted by introducing several technicalities which, in the author's opinion, would not improve drastically the results stated. To start with, even to state the results, we would have to introduce the theory of stochastic integration in Banach spaces, put forward only in recent years (see e.g. [vNVW15]).

Finally, we introduce the following 'evaluation operators' and we estimate their norm for later use. Namely, for any $j \in \mathbb{N}$ and any $(x,u),(y,v) \in Q \times \mathbb{R}^{d_v}_+$, for $a = \binom{d_v+j-1}{j}$, we define

$$
\begin{aligned}
&V^j_{(x,u)} : C^\alpha_x \mathscr{W}^{j,\infty,\theta} \to \mathbb{R}^a \,\big|\, \psi \mapsto D^j_u \psi(x,u), \\
&V^{j,\mathrm{dif}}_{(x,u),(y,v)} : C^\alpha_x \mathscr{W}^{j+1,\infty,\theta} \to \mathbb{R}^a \,\big|\, \psi \mapsto \left(D^j_u \psi(x,u) - D^j_u \psi(y,v)\right).
\end{aligned}
\tag{6.2.15}
$$

Using the definition of the spaces $\mathscr{W}^{j,\infty,\theta}$ we compute

$$
|V^j_{(x,u)}\psi| = |D^j_u\psi(x,u)| \leqslant \sup_{y,v} \frac{|D^j_u\psi(y,v)|}{1+|v|^\theta}\,(1+|u|^\theta) \lesssim \|\psi\|_{C^\alpha_x\mathscr{W}^{j,\infty,\theta}}\,(1+|u|^\theta),
$$

and similarly

$$
\begin{aligned}
|V^{j,\mathrm{dif}}_{(x,u),(y,v)}\psi| &\leqslant |D^j_u\psi(x,u) - D^j_u\psi(y,u)| + |D^j_u\psi(y,u) - D^j_u\psi(y,v)| \\
&\lesssim \sup_w \frac{|D^j_u\psi(x,w) - D^j_u\psi(y,w)|}{1+|w|^\theta}\,(1+|u|^\theta) \\
&\quad + \sup_{w\in\overrightarrow{uv}} \frac{|D^{j+1}_u\psi(y,w)|}{1+|w|^\theta}\,(1+|u|^\theta+|v|^\theta)\,|u-v| \\
&\lesssim \|\psi\|_{C^\alpha_x\mathscr{W}^{j+1,\infty,\theta}}|x-y|^\alpha(1+|u|^\theta) + \|\psi\|_{C^\alpha_x\mathscr{W}^{j+1,\infty,\theta}}(1+|u|^\theta+|v|^\theta)|u-v|.
\end{aligned}
$$

That is, we have bounds for $\|V^j_{(x,u)}\|_{(C^\alpha_x\mathscr{W}^{j,\infty,\theta})^*}$ and $\|V^{j,\mathrm{dif}}_{(x,u),(y,v)}\|_{(C^\alpha_x\mathscr{W}^{j+1,\infty,\theta})^*}$. Combining this with the embeddings (6.2.13) immediately yields the estimates, for suitable constants C,

$$
\begin{aligned}
\|V^j_{(x,u)}\|_{H^{-e}_x\mathscr{H}^{-(k+j),\theta}} &\leqslant C\,(1+|u|)^\theta, \\
\|V^{j,\mathrm{dif}}_{(x,u),(y,v)}\|_{H^{-e}_x\mathscr{H}^{-(k+j+1),\theta}} &\leqslant C\,(1+|u|+|v|)^\theta \qquad \forall e > \frac{d}{2}+\alpha\ \forall k > \frac{d_v}{2}\ \forall j,\theta \geqslant 0. \\
&\quad \cdot \left(|x-y|^\alpha + |u-v|\right),
\end{aligned}
\tag{6.2.16}
$$

6.2.2 Setting and auxiliary results

In this section we first lay out the precise setting for our results and then we collect some auxiliary results for later convenience.

Notation. To ease the formulas for the function spaces involved, we will use the shorthands:

$$
\begin{aligned}
&\kappa_1 := \left(1+\frac{d_v}{2}\right)^+, \qquad \kappa_2 := \kappa_1 + \max\left\{2, \left(\frac{d_v}{2}\right)^+\right\}, \qquad \mathrm{e}_\alpha = \alpha + \frac{d}{2}, \\
&\theta_1 := \theta_2 + \max\left\{2, \left(\frac{d_v}{2}\right)^+\right\}, \quad \theta_2 := \left(1+\frac{d_v}{2}\right)^+,
\end{aligned}
\tag{6.2.17}
$$

where for any $\beta \in \mathbb{R}$ we denote $\beta^+ := \beta + \gamma$ for any arbitrary $\gamma > 0$.

This choice of Sobolev exponents and weights will give us the needed embeddings of type (6.2.13) and (6.2.14), and the needed bounds of type (6.2.16).

Setting 6.2.5. *We fix a filtered probability space* $(\Omega, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ *satisfying the* usual conditions *and we assume it support the following random variables. First, for each* $k \in \mathbb{N}$*, it supports independent adapted* d_v*-dimensional space time white noise terms* $\{W_k(x,t)\}_{k\in\mathbb{N}}$*. Secondly, for* $h \in \mathbb{N}$*, it supports i.i.d. families of* $\mathcal{F}_0$*-measurable random initial conditions* $u_h(x,0) \in C^\alpha\big(Q; L^4(\Omega)\big) \cap L^\infty\big(Q; L^{4\,\theta_1}(\Omega)\big)$ *on the sheet* Q*, which are also independent of the noise terms* $\{W_k(x,t)\}_{k\in\mathbb{N}}$*. Indeed, we can take* $\mathcal{F}_t$ *to be the filtration generated by the initial data and the noise. For each fixed* $\epsilon > 0$ *we can then convolve the white-noise terms* W_k *and obtain the Gaussian processes* W_k^ϵ *discussed in formula* (6.2.2)*.*

Next, for each $N \in \mathbb{N}$ *we take points* $x_1, \ldots, x_N \in Q$ *in the center of the squares of an equispaced grid on* $Q = [0,1]^d$ *with side length* $N^{-\frac{1}{d}}$*. We denote by* Q_i^N *the square with center* x_i*, and we notice that* $\operatorname{meas}(Q_i^N) = \frac{1}{N}$ *and diam*$(Q_i^N) = \sqrt{d}N^{-1/d}$*. Finally, for every* $N, M \in \mathbb{N}$*, we introduce the particles for the coupling method. For* $i = 1, \ldots, N$ *and* $k = 1, \ldots, M$*, let* $u_{ik}^\epsilon(t)$ *be the solution of the particle system* (6.1.1) *with initial data* $u_{ik}(0) := u_k(x_i, 0)$ *and Brownian motions* $W_{ik}^\epsilon(t) := W_k^\epsilon(x_i, t)$*. For each* $k \in \mathbb{N}$*, let* $\bar{u}_k^\epsilon(x,t)$ *be the solution of the McKean–Vlasov equation* (6.1.2) *with initial data* $u_k(x,0)$ *and noise* $W_k^\epsilon(x,t)$*, and for* $i = 1, \ldots, N$ *define* $\bar{u}_{ik}^\epsilon(x,t) := \bar{u}_k^\epsilon(x_i, t)$*.*

In this setting, Theorems 2.3, 2.4 and 2.5 in [CCS23] ensure the systems (6.1.1) and (6.1.2) are well-posed and that $f(x,t,du) = \operatorname{Law}_{\mathbb{R}^{d_v}}(\bar{u}_k^\epsilon(x,t))$ is the unique weak solution of the Fokker–Planck PDE (6.1.4) with initial data $f_0(x,du) = \operatorname{Law}_{\mathbb{R}^{d_v}}(u_k(x,0))$.

We now collect some auxiliary results either new or taken from [CCS23] we will need in the following. First of all, owing the i.i.d. properties of the initial data and the noise terms, we have the following.

Lemma 6.2.6. The collections of particles $\{u_{ik}^\epsilon(t)\}_{i=1,\ldots,N}$ are exchangeable families for $k = 1, \ldots, M$. Moreover, the collections of particles $\{\bar{u}_{ik}^\epsilon(t)\}_{i=1,\ldots,N}$ are i.i.d. for $k \in \mathbb{N}$.

We point out that this is not the case for the index i, both for the particles u_{ik}^ϵ and $\bar{u}_{ik}^\epsilon = \bar{u}_k^\epsilon(x_i, \cdot)$. Indeed, the laws of u_{ik}^ϵ and u_{jk}^ϵ, or $\bar{u}_{ik}^\epsilon$ and $\bar{u}_{ik}^\epsilon$ respectively, might differ as a result of the x dependence of their defining equations. Furthermore, even if the points $x_i, x_j \in Q$ are far from each other, namely if $|x_i - x_j| > 2\epsilon$, so that their noise terms $W_k^\epsilon(x_i,t)$ and $W_k^\epsilon(x_j,t)$ are independent, these particles might still be correlated as a result of their initial data. In fact, from the point of view of modelling in neuroscience, we expect $u_k(x,0)$ to be close to $u_k(y,0)$ for x close to y. This is essentially the issue we pointed out in Remark 6.1.1.

We have the following uniform moment estimates both for the true particles u_{ik}^ϵ and for the McKean–Vlasov particles $\bar{u}^{\epsilon}_{ik}$. The case $p = 2$ is proved in [CCS23, Theorem 2.3-4] (see Chapter 5) and the case of a general $p \geqslant 1$ is proved almost identically.

Proposition 6.2.7 (Uniform moment estimates)**.** In the Setting 6.2.5 above, for every $T \geqslant 0$ and every $p \geqslant 1$, we have

$$\sup_{i,k} \mathbb{E}\left[\sup_{t\in[0,T]} |u^{\epsilon}_{ik}|^p\right] + \sup_{x\in Q} \mathbb{E}\left[\sup_{t\in[0,T]} |\bar{u}^{\epsilon}_{k}(x,t)|^p\right] \leqslant C\left(1 + \sup_{x\in Q} \mathbb{E}[|u(x,0)|^p]\right), \tag{6.2.18}$$

for a constant $C = C(T, p, b, \sigma)$ independent of M, N and ϵ.

Furthermore, we have the following estimates for the error between the true and the McKean–Vlasov particles. The case $p = 2$ is proved in [CCS23, Theorem 2.6] (see Chapter 5) and the case $p = 4$ is a straightforward adaptation.

Theorem 6.2.8 (Mean squared error estimates for actual particles vs. McKean–Vlasov particles)**.** In the Setting 6.2.5 above, for $p = 2, 4$ and for any $T > 0$, we have

$$\sup_{i,k} \mathbb{E}\left[\sup_{t\in[0,T]} |u^{\epsilon}_{ik}(t) - \bar{u}^{\epsilon}_{ik}(t)|^p\right]^{1/p} \leqslant C\left(\frac{1}{N^{\frac{\alpha}{d}}} + \frac{1}{M^{\frac{1}{2}}}\right)\left(1 + \sup_{x\in Q} \mathbb{E}\left[|u_k(x,0)|^p\right]^{1/p}\right), \tag{6.2.19}$$

for a constant $C = C(T, b, \sigma, [u(\cdot,0)]_\alpha)$, where $[u(\cdot,0)]_\alpha$ denotes the Hölder seminorm of $u(\cdot,0) \in C^\alpha\big(Q; L^p(\Omega)\big)$.

We now consider empirical and actual measures associated to the true and McKean–Vlasov particles. In [CCS23] (see Chapter 5) it is established that the solution $f(x,t,du) = \mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}^\epsilon(x,t))$ of the Fokker–Planck PDE (6.1.4) is independent of ϵ and it is α-Hölder continuous as a function of x with values in $\mathscr{P}_2\left(C([0,T];\mathbb{R}^{d_v})\right)$ endowed with the 2-Wasserstein metric whenever $u(\cdot,0) \in C^\alpha(Q; L^2(\Omega))$. Furthermore it is proved that $f_{MN} \to f$ in Wasserstein distance of order 1, in expectation.

In the following we will need analogous results for the joint law of two McKean–Vlasov particles

$$f^{2,\epsilon}(x,y,t,du,dv) := \mathrm{Law}_{\mathbb{R}^{2d_v}}\left(\bar{u}^\epsilon(x,t), \bar{u}^\epsilon(y,t)\right), \tag{6.2.20}$$

which is a measure on $\mathbb{R}^{2d_v}$ depending on $x, y \in Q$, $t \in \mathbb{R}^+$ and $\epsilon > 0$, inducing a measure on $Q^2 \times \mathbb{R}^{2d_v}$ via integration as in (6.1.3), and for the joint empirical measure associated to the actual particles

$$f^2_{MN}(t,dx,dy,du,dv) := \frac{1}{MN^2}\sum_{i,j=1}^{N}\sum_{k=1}^{M} \delta_{(x_i,x_j,u^\epsilon_{ik},u^\epsilon_{jk})} \in \mathscr{P}\left(Q^2 \times \mathbb{R}^{2d_v}\right).$$

Remark 6.2.9. As anticipated in the introduction, the joint law $f^{2,\epsilon}(x,y,t,du,dv)$ does depend on the correlation radius $\epsilon > 0$, in contrast with the case of the single law $f(x,t,du)$.

Indeed, the independence on ϵ of f is an artefact of the specific type of interaction among particles, which takes the form, highlighting the drift term only,

$$
\begin{aligned}
d\bar{u}^\epsilon(x,t) &= \cdots + \phi\left(\int_{Q\times\mathbb{R}^{d_v}} b_1(x,y,t,\bar{u}^\epsilon(x,t),v) f(y,t,dv)\,dy\right) dt + \ldots \\
&= \cdots + \phi\left(\int_Q \mathbb{E}\left[b_1(x,y,t,u,\bar{u}^\epsilon(y,t))\right]_{u=\bar{u}^\epsilon(x,t)}\,dy\right) dt + \ldots
\end{aligned}
$$

That is to say, the particle $\bar{u}^\epsilon(x,t)$ interacts with the particle $\bar{u}^\epsilon(y,t)$ only through its law $f(y,t,dv)$. Therefore the correlation radius of the noise sources $W^\epsilon(x,t)$ and $W^\epsilon(y,t)$ affecting the particles plays no role since each of these terms behaves individually as a Brownian motion, regardless of the value of ϵ. Accordingly, the effect of ϵ is recovered as soon as we consider the *joint* law of $\bar{u}^\epsilon(x,t)$ and $\bar{u}^\epsilon(y,t)$.

We mention that other commonly proposed models (see e.g. [Ama77, KR14, FI15] and the review [Bre11, Section 6]) consider a different type of interaction, where $\bar{u}^\epsilon(x,t)$ does interact with $\bar{u}^\epsilon(y,t)$ directly and not only through its law, so that the dependence on ϵ is retained already at the level of the single law $\mathrm{Law}_{\mathbb{R}^{d_v}}(\bar{u}^\epsilon(x,t))$.

We promote f and $f^{2,\epsilon}$ to measures on $C([0,T];\mathbb{R}^{d_v})$ and $C([0,T];\mathbb{R}^{2d_v})$ respectively by setting $f(x,\cdot) = \mathrm{Law}_{C([0,T];\mathbb{R}^{d_v})}(\bar{u}_k(x,\cdot))$ and $f^{2,\epsilon}(x,y,\cdot) = \mathrm{Law}_{C([0,T];\mathbb{R}^{2d_v})}(\bar{u}_k(x,\cdot),\bar{u}_k(y,\cdot))$. Then, from [CCS23, Theorem 2.5](see Chapter 5), or straightforward adaptations of its arguments in the cases of $p=4$ and of the joint measure, we have the following.

Proposition 6.2.10 (Hölder continuity of the single and joint empirical measure)**.** In the Setting 6.2.5 above, for $p = 2,4$, we have $f \in C^\alpha\left(Q;\mathscr{P}_p\left(C([0,T];\mathbb{R}^{d_v})\right)\right)$ and $f^{2,\epsilon} \in C^\alpha\left(Q^2;\mathscr{P}_p\left(C([0,T];\mathbb{R}^{2d_v})\right)\right)$ with the following estimates

$$
\begin{aligned}
&\mathcal{W}_p\Big(C([0,T];\mathbb{R}^{d_v})\Big)(f(x,\cdot),f(x',\cdot)) \leqslant C\left(1+\sup_{x\in Q}\mathbb{E}[|u(x,0)|^p]^{1/p}\right)|x-x'|^\alpha \quad \forall x,x'\in Q,\\
&\mathcal{W}_p\Big(C([0,T];\mathbb{R}^{2d_v})\Big)(f(x,y,\cdot),f(x',y',\cdot))\\
&\quad\leqslant C\left(1+\sup_{x\in Q}\mathbb{E}[|u(x,0)|^p]^{1/p}\right)\left(|x-x'|^\alpha+|y-y'|^\alpha+\frac{|x-x'|+|y-y'|}{\epsilon}\right) \quad \forall x,x',y,y'\in Q,
\end{aligned}
$$

for a constant $C = C(T,\rho,b,\sigma,[u(\cdot,0)]_\alpha)$, where $[u(\cdot,0)]_\alpha$ denotes the Hölder seminorm of $u(\cdot,0)\in C^\alpha\big(Q;L^p(\Omega)\big)$.

Finally, we establish the convergence for the empirical measure. We recall that a space dependent measure $g(x,du)\in\mathscr{P}(\mathbb{R}^{d_v})$ induces a measure on $Q\times\mathbb{R}^{d_v}$ via formula (6.1.3) and similarly for a measure $g(x,y,du,dv)\in\mathscr{P}(\mathbb{R}^{2d_v})$ inducing a measure on $Q^2\times\mathbb{R}^{2d_v}$. A straightforward adaptation of [CCS23, Theorem 2.7] (see Chapter 5) proves the following.

Theorem 6.2.11 (Convergence of the single and joint empirical measure)**.** In the Setting 6.2.5 above, we have $f_{MN}\to f$ and $f^2_{MN}\to f^{2,\epsilon}$ in Wasserstein distance $\mathcal{W}_2$ in expectation with convergence rates, for every $T\geqslant 0$ and every $q>2$, for a constant

$C = C(T, q, \rho, b, \sigma, [u(\cdot, 0)]_\alpha)$,

$$
\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}\Big[\mathcal{W}_2 \left(Q \times \mathbb{R}^{d_v}\right) (f_{MN}(t), f(t))^2 \Big] \\
&\quad \leqslant C \left(1 + \sup_{x\in Q} \mathbb{E}[|u(x,0)|^2]\right) \left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right)^2 \\
&\quad + C \left(1 + \sup_{x\in Q} \mathbb{E}[|u(x,0)|^q]^{2/q}\right) \begin{cases} M^{-1/2} + M^{-(q-2)/q} & \text{if } 2 > d_v/2 \text{ and } q \neq 4, \\ M^{-1/2} \log(1+M) + M^{-(q-2)/q} & \text{if } 2 = d_v/2 \text{ and } q \neq 4, \\ M^{-2/d_v} + M^{-(q-2)/q} & \text{if } 2 \in (0, d_v/2) \text{ and } q \neq \frac{d_v}{d_v-2}, \end{cases} \\
&\sup_{t\in[0,T]} \mathbb{E}\Big[\mathcal{W}_2 \left(Q^2 \times \mathbb{R}^{2d_v}\right) (f^2_{MN}(t), f^{2,\epsilon}(t))^2 \Big] \\
&\quad \leqslant C \left(1 + \sup_{x\in Q} \mathbb{E}[|u(x,0)|^2]\right) \left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}} + \frac{1}{\epsilon N^{1/d}}\right)^2 \\
&\quad + C \left(1 + \sup_{x\in Q} \mathbb{E}[|u(x,0)|^q]^{2/q}\right) \begin{cases} M^{-1/2} + M^{-(q-2)/q} & \text{if } 2 > d_v \text{ and } q \neq 4, \\ M^{-1/2} \log(1+M) + M^{-(q-2)/q} & \text{if } 2 = d_v \text{ and } q \neq 4, \\ M^{-1/d_v} + M^{-(q-2)/q} & \text{if } 2 \in (0, d_v) \text{ and } q \neq \frac{d_v}{d_v-1}. \end{cases}
\end{aligned}
$$

6.3 The central limit theorem

In this section we prove our main result following the strategy sketched in the introduction. We will *always* consider the Setting 6.2.5.

We want to find an equation satisfied by the rescaled fluctuations $\eta_t^{MN} = \mathrm{c_{MN}}(f_{MN} - f)$, for some scaling parameter $\mathrm{c_{MN}}$, and pass this to the limit as $M, N \to \infty$. We anticipate that the right scaling for the fluctuations is $\mathrm{c_{MN}} = \sqrt{M}$ along a suitable scaling regime $M, N \to \infty$, namely $\sqrt{M}N^{-\alpha/d} \to 0$. However, we choose to keep the scaling factor $\mathrm{c_{MN}}$ implicit so as to highlight why this is the only possible choice in order to obatin a nontrivial limit for the fluctuations (cf. Remark 6.3.5 below).

Let $\psi : Q \times \mathbb{R}^{d_v} \to \mathbb{R}$ be sufficiently smooth with $\nabla_u \psi(x, u) \cdot \mathbf{n}_{\partial \mathbb{R}^{d_v}_+} \equiv 0$, for example $\psi \in H^{\mathrm{e}_\alpha}_x \mathscr{H}^{\kappa_1, \theta_1}$. Testing f_{MN} against ψ, applying Itô formula and using the boundary conditions on ψ we obtain

$$
\begin{aligned}
\langle f_{MN}(t), \psi \rangle =& \frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \psi(x_i, u_{ik}^{\epsilon}(0)) \\
&+ \int_0^t \frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \nabla_u \psi(x_i, u_{ik}^{\epsilon}(r))\, b(x_i, u_{ik}^{\epsilon}, r, f_{MN})\, dr \\
&+ \int_0^t \frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \Delta_u \psi(x_i, u_{ik}^{\epsilon}(r))\, \sigma(x_i, u_{ik}^{\epsilon}, r, f_{MN})^2 \ dr \\
&+ \frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \int_0^t \nabla_u \psi(x_i, u_{ik}^{\epsilon}(r))\, \sigma(x_i, u_{ik}^{\epsilon}, r, f_{MN}) \ dW_{ik}^{\epsilon}(r) \\
&+ \frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} \int_0^t \nabla_u \psi(x_i, u_{ik}^{\epsilon}(r)) \cdot \mathbf{n}_{\partial \mathbb{R}_+^{d_v}} \ d\ell_{ik}^{\epsilon}(r) \\
=& \langle f_{MN}(0), \psi \rangle + \int_0^t \langle f_{MN}(r), L_r(f_{MN}(r))[\psi] \rangle\, dr + \frac{1}{\mathrm{c_{MN}}} M_t^{MN}(\psi),
\end{aligned}
\tag{6.3.1}
$$

for the martingale term

$$
M_t^{MN}(\psi) = \frac{\mathrm{c_{MN}}}{MN} \sum_{i=i}^{N} \sum_{k=1}^{M} \int_0^t \nabla_u \psi(x_i, u_{ik}(r))\, \sigma(x_i, r, u_{ik}(r), f_{MN})\, dW_{ik}(r), \tag{6.3.2}
$$

and the differential operator $L_r(f_{MN}(r))$, depending on $f_{MN}(r, dx, du) \in \mathscr{P}(Q \times \mathbb{R}^{d_v})$ and on the time r, defined for any $\psi : Q \times \mathbb{R}^{d_v} \to \mathbb{R}$ sufficiently smooth by

$$
L_r(\mu)[\psi](x, u) := \nabla_u \psi(x, u)\, b(x, r, u, \mu) + \Delta_u \psi(x, u)\, \sigma(x, r, u, \mu)^2 \quad \forall \mu \in \mathscr{P}(Q \times \mathbb{R}^{d_v}). \tag{6.3.3}
$$

Similarly we test ψ, which is admissible because of the boundary conditions, against the weak solution $f(t, x, du)$ of the Fokker–Planck PDE (6.1.4), which is regarded as an element of $\mathscr{P}(Q \times \mathbb{R}^{d_v})$ via formula (6.1.3). We obtain

$$
\langle f(t), \psi \rangle = \langle f(0), \psi \rangle + \int_0^t \langle f(r, dx, du), L_r(f(r))[\psi] \rangle\, dr, \tag{6.3.4}
$$

where $L_r(f(r))$ is the differential operator (6.3.3) evaluated on $f(r, dx, du) \in \mathscr{P}(Q \times \mathbb{R}^{d_v})$.

Combining (6.3.1) and (6.3.4) yields

$$
\langle \eta_t^{MN}, \psi \rangle = \langle \eta_0^{MN}, \psi \rangle + \int_0^t \mathrm{c_{MN}} \left(\langle f_{MN}(r), L_r(f_{MN})[\psi] \rangle - \langle f(r), L_r(f)[\psi] \rangle \right) \, dr + M_t^{MN}(\psi). \tag{6.3.5}
$$

Finally we rewrite

$$
\begin{aligned}
&\mathrm{c_{MN}} \Big(\langle f_{MN}(r), L_r(f_{MN})[\psi] \rangle - \langle f(r), L_r(f)[\psi] \rangle \Big) \\
&\qquad\qquad = \langle \eta_r^{MN}(dy, dv), \mathscr{L}_r(f_{MN}(r), f(r))[\psi](y, v) \rangle
\end{aligned}
\tag{6.3.6}
$$

for the linear differential operator $\mathscr{L}_t(\mu,\nu)$, depending on $\mu,\nu\in\mathscr{P}(Q\times\mathbb{R}^{d_v})$, defined by

$$
\begin{aligned}
&\mathscr{L}_t(\mu,\nu)[\psi](y,v)\\
&= L_t(\mu)[\psi](y,v)\\
&\quad+\Big\langle \nu(dx,du),\nabla_u\psi(x,u)b_1(x,y,t,u,v)\Big(\int_0^1\dot\phi\left((1-\lambda)b_1(x,t,u,\nu)+\lambda b_1(x,t,u,\mu)\right)d\lambda\Big)\Big\rangle\\
&\quad+\langle \nu(dx,du),\Delta_u\psi(x,u)\sigma_1(x,y,t,u,v)2\sigma_0(x,t,u)\\
&\qquad\qquad\cdot\Big(\int_0^1\dot\phi\left((1-\lambda)\sigma_1(x,t,u,\nu)+\lambda\sigma_1(x,t,u,\mu)\right)d\lambda\Big)\Big\rangle\\
&\quad+\Big\langle \nu(dx,du),\Delta_u\psi(x,u)\sigma_1(x,y,t,u,v)\big(\phi(\sigma_1(x,t,u,\nu))+\phi(\sigma_1(x,t,u,\mu))\big)\\
&\qquad\qquad\cdot\Big(\int_0^1\dot\phi\left((1-\lambda)\sigma_1(x,t,u,\nu)+\lambda\sigma_1(x,t,u,\mu)\right)d\lambda\Big)\Big\rangle,
\end{aligned}
\tag{6.3.7}
$$

where we used the shorthand (6.2.5) for $b_1(x,t,u,\mu)$ and the analogous terms. Equation (6.3.6) is justified by direct computation simply by plugging in the expression (6.3.7) for $\mathscr{L}(f_{MN},f)$, which is essentially the linearization of $L(f_{MN})-L(f)$ around f. Combining (6.3.5) and (6.3.6) we have proved the following.

Proposition 6.3.1. For any $\psi\in H_x^{e_\alpha}\mathscr{H}^{\kappa_1,\theta_1}$ the process $\langle\eta_t^{MN},\psi\rangle$ is a real valued continuous semimartingale with decomposition

$$
\langle\eta_t^{MN},\psi\rangle=\langle\eta_0^{MN},\psi\rangle+\int_0^t\langle\eta_r^{MN},\mathscr{L}_r(f_{MN},f)[\psi]\rangle\,dr+M_t^{MN}(\psi).\tag{6.3.8}
$$

The next step is to understand the behavior of the martingale term $M_t^{MN}(\psi)$ in the $M,N\to\infty$ limit. We start by analyzing its quadratic variation. For any $\varphi,\psi\in H_x^{e_\alpha}\mathscr{H}^{\kappa_1,\theta_1}$, from the expressions (6.3.2) and (6.2.2), we compute:

$$
\begin{aligned}
&\langle M_t^{MN}(\varphi),M_t^{MN}(\psi)\rangle\\
&=\frac{{\mathrm{c_{MN}}}^2}{M}\int_0^t\frac{1}{MN^2}\sum_{i,j=1}^N\sum_{k=1}^M\left(\nabla_u\varphi(x_i,u_{ik}^\epsilon)\sigma(x_i,r,u_{ik}^\epsilon,f_{MN})\right)\left(\nabla_u\psi(x_j,u_{jk}^\epsilon)\sigma(x_j,r,u_{jk}^\epsilon,f_{MN})\right)\\
&\qquad\qquad\cdot\Big(\int_{\mathbb{R}^d}\rho(z+\epsilon^{-1}(x_i-x_j))\rho(z)\,dz\,C_\rho\Big)\,dr\\
&=\frac{{\mathrm{c_{MN}}}^2}{M}\int_0^t\int_{Q^2\times\mathbb{R}^{2d_v}}(\nabla_u\varphi(x,u)\sigma(x,r,u,f_{MN}))\\
&\qquad\qquad\cdot(\nabla_u\psi(y,v)\sigma(y,r,v,f_{MN}))\,R^\epsilon(x-y)\,df_{MN}^2(r,dx,dy,du,dv)\,dr\\
&=\frac{{\mathrm{c_{MN}}}^2}{M}\int_0^t\Big\langle(\nabla_u\varphi(x,u)\sigma(x,r,u,f_{MN}))\,(\nabla_u\psi(y,v)\sigma(y,r,v,f_{MN}))\,R^\epsilon(x-y),\\
&\qquad\qquad df_{MN}^2(r,dx,dy,du,dv)\Big\rangle\,dr,
\end{aligned}
\tag{6.3.9}
$$

where we have defined

$$
R^\epsilon(x):=C_\rho\int_{\mathbb{R}^d}\rho(z+\epsilon^{-1}x)\rho(z)\,dz.
$$

The expression (6.3.9) and the convergence $f_{MN}\to f$ and $f_{MN}^2\to f^{2,\epsilon}$ from Theorem 6.2.11 suggest that the quadratic variation $\langle M_t^{MN}(\varphi),M_t^{MN}(\psi)\rangle$ converges to the *deterministic* finite variation function

$$
\begin{aligned}
g_t^\epsilon(\varphi,\psi) &:= \int_0^t \int_{Q^2\times\mathbb{R}^{2d_v}} (\nabla_u\varphi(x,u)\sigma(x,r,u,f))\,(\nabla_u\psi(y,v)\sigma(y,r,v,f)) \\
&\qquad\qquad \cdot R^\epsilon(x-y)\,df^{2,\epsilon}(r,dx,dy,du,dv)\,dr \\
&= \int_0^t \langle(\nabla_u\varphi(x,u)\sigma(x,r,u,f))\,(\nabla_u\psi(y,v)\sigma(y,r,v,f))\,R^\epsilon(x-y), \\
&\qquad\qquad df^{2,\epsilon}(r,dx,dy,du,dv)\rangle\,dr.
\end{aligned}
\tag{6.3.10}
$$

The next proposition establishes this rigorously.

Proposition 6.3.2. For any $\varphi,\psi\in H_x^{\mathrm{e}\alpha}\mathscr{H}^{\kappa_2,\theta_2}$, for any $T\geqslant 0$ we have

$$
\begin{aligned}
\lim_{M,N\to\infty}\sup_{r\in[0,T]}\mathbb{E}\Big[\Big|\Big\langle&(\nabla_u\varphi(x,u)\sigma(x,r,u,f_{MN}))\,(\nabla_u\psi(y,v)\sigma(y,r,v,f_{MN}))\,R^\epsilon(x-y), \\
&\qquad f^2_{MN}(r,dx,dy,du,dv)\Big\rangle \\
&-\Big\langle(\nabla_u\varphi(x,u)\sigma(x,r,u,f))\,(\nabla_u\psi(y,v)\sigma(y,r,v,f))\,R^\epsilon(x-y), \\
&\qquad f^{2,\epsilon}(r,dx,dy,du,dv)\Big\rangle\Big|\Big]=0.
\end{aligned}
$$

In particular we have

$$
\lim_{M,N\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\langle M_t^{MN}(\varphi),M_t^{MN}(\psi)\rangle - g^\epsilon(\varphi,\psi)(t)\right|\right]=0. \tag{6.3.11}
$$

Given an admissible test function ψ, standard probability theory ensures there exists a (unique in law) Gaussian martingale $G_t^\epsilon(\psi)$ having the increasing function $g_t^\epsilon(\psi,\psi)$ as its quadratic variation. Proposition 6.3.2 hints that $M_t^{MN}(\psi)$ converges to the process $G_t^\epsilon(\psi)$. In turn, assuming that η_t^{MN} has some limit η_t^ϵ, the convergence $f_{MN}\to f$ suggests that we pass to the limit in equation (6.3.8) and obtain the following expression

$$
\langle\eta_t^\epsilon,\psi\rangle = \langle\eta_0,\psi\rangle + \int_0^t\langle\eta_r^\epsilon,\mathscr{L}_r(f(r),f(r))[\psi]\rangle\,dr + G_t^\epsilon(\psi). \tag{6.3.12}
$$

The remaining part of this section is devoted to making this argument rigorous. The missing ingredient is essentially the tightness of the processes involved.

First of all, we pass everything to the level of distributions. We start with the martingale term. Indeed, it is not hard to see that, for a fixed time t, the association $H_x^{\mathrm{e}\alpha}\mathscr{H}^{\kappa_1,\theta_1}\ni\psi\mapsto M_t^{MN}(\psi)$ defines an $H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$-valued random variable M_t^{MN}. We have the following result.

Proposition 6.3.3. The process M_t^{MN} is an $H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$-valued continuous square integrable martingale. Its quadratic variation is the $H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}\otimes H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$-valued process $\langle M_t^{MN}\rangle$ defined by

$$
\langle M_t^{MN}\rangle(\varphi,\psi) := \langle M_t^{MN}(\varphi),M_t^{MN}(\psi)\rangle.
$$

For every $T \geqslant 0$, for a constant $C = C(T, b, \sigma)$, we have

$$\sup_{M,N} \mathbb{E}\left[\sup_{t\in[0,T]} \|M_t^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha} \mathscr{H}^{-\,\kappa_1,\theta_1}}\right] \leqslant C \sup_{M,N} \frac{\mathrm{c_{MN}}^2}{M}\left(1+\sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^{2\,\theta_1+2}\big]\right). \tag{6.3.13}$$

Next we want to study the fluctuation process $\eta_t^{MN} \in \mathscr{P}(Q \times \mathbb{R}^{d_v}) \subseteq H_x^{-\,\mathrm{e}_\alpha} \mathscr{H}^{-\,\kappa_1,\theta_1}$, obtain bounds on its norm and pass (6.3.8) to an equation at the distributional level. This will be done in two steps, working between the nested spaces $H_x^{\mathrm{e}_\alpha^+} \mathscr{H}^{\kappa_2,\theta_2} \subseteq H_x^{\mathrm{e}_\alpha} \mathscr{H}^{\kappa_1,\theta_1}$.

Lemma 6.3.4. For every $T \geqslant 0$, for a constant $C = C(T, b, \sigma)$, we have

$$\sup_{MN}\sup_{t\in[0,T]} \mathbb{E}\left[\|\eta_t^{MN}\|^2_{H^{-\,\mathrm{e}_\alpha} \mathscr{H}^{-\,\kappa_1,\theta_1}}\right] \leqslant C \sup_{M,N}\left(\frac{\mathrm{c_{MN}}^2}{M} + \frac{\mathrm{c_{MN}}^2}{N^{2\alpha/d}}\right)\left(1+\sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^{4\,\theta_1}\big]\right). \tag{6.3.14}$$

Remark 6.3.5. Lemma 6.3.4 implies that, in order to keep the fluctuations bounded in a suitable distribution space, we have to impose the constraint $\limsup_{M,N\to\infty}\left(\frac{\mathrm{c_{MN}}^2}{M} + \frac{\mathrm{c_{MN}}^2}{N^{2\alpha/d}}\right) < \infty$. On the other hand, estimate (6.3.13) implies that, in order for the martingale term not to vanish, which would imply $\eta_\cdot^{MN} \to 0$ and result in a suboptimal expansion (6.1.8) for the empirical measure, we have to impose $\mathrm{c_{MN}} \gtrsim \sqrt{M}$. The requirements on the scaling regime are therefore $\mathrm{c_{MN}} \simeq \sqrt{M}$ and $\limsup_{M,N\to\infty} \sqrt{M}N^{-\alpha/d} < \infty$. In fact, we will see that the convergence of the initial data η_0^{MN} further requires $\lim_{M,N\to\infty} \sqrt{M}N^{-\alpha/d} = 0$.

We can make sense of these constraints as follows. As explained in Remark 6.1.1, our situation essentially corresponds to N classical mean field problems with M particles each, one for each column of M neurons at location x_i, interacting among themselves. The decay rate $\sqrt{M}$ is the usual decay rate for a standard mean field setting with M particles subjected to independent noise sources $W_k^\epsilon(x_i, t)$ for $k = 1, \ldots, M$. Therefore, if we want to see a nontrivial behavior of the rescaled fluctuations $\mathrm{c_{MN}}(f_{MN} - f)$ and also to keep these bounded, we have to choose $\mathrm{c_{MN}} = \sqrt{M}$.

However, we have N such mean field clusters interacting with each other and two different columns at nearby locations x_i, x_j are initiated with possibly correlated initial data and sense correlated sources of noise $\{W_k^\epsilon(x_i, t)\}_{k=1\ldots M}$ and $\{W_k^\epsilon(x_j, t)\}_{k=1\ldots M}$ with correlation strength proportional to $|x_i - x_j|^\alpha \simeq N^{-\alpha/d}$. The condition $\limsup \sqrt{M}N^{-\alpha/d} < \infty$ serves therefore to prevent these correlated random clusters from interacting constructively and making the $\sqrt{M}$-rescaled fluctuations yet unbounded.

We can obtain better estimates for the process η_t^{MN} by exploiting the semimartingale decomposition (6.3.8). For this, we first need the following lemma on the norm of the operators $\mathscr{L}_t(\mu, \nu)$ defined in (6.3.7).

Lemma 6.3.6. The linear operators $\mathscr{L}_t(f_{MN}, f)$ and $\mathscr{L}_t(f, f)$ define two maps $H_x^{e} \mathscr{H}^{\kappa_2,\theta_2} \to$

$H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}$ and two maps $H_x^{\mathrm{e}}\mathscr{H}^{\kappa_2+2,\theta_2-2}\to H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}$ satisfying the bounds,

$$\begin{aligned}\sup_{t\in[0,T]}\sup_{M,N}\operatorname*{ess-sup}_{\omega\in\Omega}&\|\mathscr{L}_t(f_{MN},f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}\\ &+\sup_{t\in[0,T]}\|\mathscr{L}(f,f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}\leqslant C\left(1+\sup_{x\in Q}\mathbb{E}\left[|u(x,0)|^{4+2\theta_1}\right]\right)^{1/2},\\ \sup_{t\in[0,T]}\sup_{M,N}\operatorname*{ess-sup}_{\omega\in\Omega}&\|\mathscr{L}_t(f_{MN},f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2+2,\theta_2-2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}}\\ &+\sup_{t\in[0,T]}\|\mathscr{L}(f,f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2+2,\theta_2-2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}}\leqslant C\left(1+\sup_{x\in Q}\mathbb{E}\left[|u(x,0)|^{4+2\theta_1}\right]\right)^{1/2},\end{aligned}\tag{6.3.15}$$

for a constant $C=C(T,b,\sigma,d_v,d,\alpha)$ independent of the event $\omega\in\Omega$.

Using these bounds, we now obtain a semimartingale expression for η_t^{MN} in the bigger space $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_2,\theta_2}\supseteq H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$.

Proposition 6.3.7. The process η_t^{MN} is an $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_2,\theta_2}$-valued continuous square integrable semimartingale with decomposition

$$\eta_t^{MN}=\eta_0^{MN}+\int_0^t\mathscr{L}_r(f_{MN},f)^*[\eta_r^{MN}]\,dr+M_t^{MN},\tag{6.3.16}$$

where $\mathscr{L}_r(f_{MN},f)^*:H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}\to H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_2,\theta_2}$ is the adjoint operator of $\mathscr{L}_r(f_{MN},f)$. For every $T\geqslant 0$, for a constant $C=C(T,b,\sigma,d,d_v,\alpha)$, we have

$$\sup_{MN}\mathbb{E}\left[\sup_{t\in[0,T]}\|\eta_t^{MN}\|^2_{H^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_2,\theta_2}}\right]\leqslant C\sup_{M,N}\left(\frac{\mathrm{c_{MN}}^2}{M}+\frac{\mathrm{c_{MN}}^2}{N^{2\alpha/d}}\right)\left(1+\sup_{x\in Q}\mathbb{E}\big[|u(x,0)|^{4\theta_1}\big]\right).\tag{6.3.17}$$

In order to pass to the limit in the expression (6.3.16) we exploit the tightness of the terms involved. This is proved with the previous estimates, the compactness of the embedding $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}\subseteq H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_2,\theta_2}$ and the Aldous criterion (see e.g. [JM86]).

Proposition 6.3.8. For $M,N\in\mathbb{N}$ the random variables $M_t^{MN}\in C\big([0,\infty];H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}\big)$ are tight. Furthermore, the random variables $\eta_t^{MN}\in C\big([0,\infty];H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}\big)$ are tight along a scaling regime such that $\limsup_{M,N\to\infty}\sqrt{M}N^{-\alpha/d}<\infty$.

Next we use Prokhorov's theorem and the convergence (6.3.11) of the quadratic variation $\langle M_t^{MN}\rangle(\varphi,\psi)$ to establish that M_t^{MN} does have a unique limit in $C\big([0,T];H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}\big)$. The first step is to identify uniquely this limit. The following lemma is proved similarly to estimate (6.3.13).

Lemma 6.3.9. The association $H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}\ni\varphi,\psi\mapsto g_t^\epsilon(\varphi,\psi)$ defines a *deterministic* increasing positive definite function g_t^ϵ with values in $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}\otimes H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$.

Thus g_t^ϵ is an admissible covariation function in $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$ and standard probability theory (see e.g. [DPZ92, Chapter 3]) ensures there exists a unique-in-law $H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}$-valued Gaussian martingale G_t^ϵ with quadratic variation g_t^ϵ.

The uniform bound (6.3.13), Proposition 6.3.8 and Prokhorov's theorem imply that, along subsequences, the process M_t^{MN} converges both in law in $C\big([0,T]; H_x^{-\,\mathrm{e}_\alpha^+}\mathscr{H}^{-\,\kappa_2,\theta_2}\big)$ and weakly in the Hilbert space $\mathcal{M}_T^2\big(H_x^{\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}\big)$ of square integrable continuous martingales to (possibly different) limit martingales. In addition, Proposition 6.3.2 implies that along any such subsequence we have $\langle M_t^{MN}\rangle(\varphi,\psi) \to g_t^\epsilon(\varphi,\psi)$ for every $\varphi,\psi \in H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}$. Hence, along any subsequence, the limit martingale must be G_t^ϵ and a sub-subsequence argument readily proves that the whole sequence M_t^{MN} is converging to G_t^ϵ. Hence we have proved the following.

Proposition 6.3.10. We have $\lim_{M,N\to\infty} M_t^{MN} = G_t^\epsilon$ in law in $C\big([0,T]; H_x^{-\,\mathrm{e}_\alpha^+}\mathscr{H}^{-\,\kappa_2,\theta_2}\big)$, with no constraints on the scaling regime for $M,N\to\infty$.

Finally we exploit the tightness of η_t^{MN} and the convergence $f_{MN}\to f$ and $M_t^{MN}\to G_t^\epsilon$ to pass to the limit in equation (6.3.16) and show that the fluctuations η_t^{MN} have a unique limit. We first establish the convergence of the initial data η_0^{MN}. This is obtained with adaptations of the standard argument for Lévy's classical central limit theorem.

Lemma 6.3.11. Consider a scaling regime $M,N\to\infty$ such that $\lim \sqrt{M}N^{-\alpha/d} = 0$, then we have

$$\eta_0^{MN} \to \eta_0 \quad \text{in law in } H_x^{-\,\mathrm{e}_\alpha^+}\mathscr{H}^{-\,\kappa_2,\theta_2},$$

where $\eta_0 \sim \mathcal{N}(0,\mathcal{Q})$ is a Gaussian r.v. in $H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}$ with mean zero and covariance

$$\begin{aligned}\mathcal{Q}(\varphi,\psi) &= \mathbb{E}\Big[\Big(\int_Q \varphi(x,u_k(x,0)) - \mathbb{E}[\varphi(x,u_k(x,0))]\,dx\Big)\\&\qquad\cdot\Big(\int_Q \psi(y,u_k(y,0)) - \mathbb{E}[\psi(y,u_k(y,0))]\,dy\Big)\Big]\\&= \int_{Q^2\times\mathbb{R}^{2d_v}}\Big(\varphi(x,u) - \int_{\mathbb{R}^{d_v}}\varphi(x,w)f_0(x,dw)\Big)\\&\qquad\cdot\Big(\psi(y,v) - \int_{\mathbb{R}^{d_v}}\psi(y,w)f_0(y,dw)\Big)f_0^2(x,y,du,dv)\,dx\,dy.\end{aligned} \tag{6.3.18}$$

In particular η_0 is independent of ϵ since $f_0(x,du) = \mathrm{Law}_{\mathbb{R}^{d_v}}(u_k(x,0))$ and $f_0^2(x,y,du,dv) = \mathrm{Law}_{\mathbb{R}^{2d_v}}(u_k(x,0),u_k(y,0))$ are.

Finally we establish our main result: a central limit theorem for the whole fluctuation process η_t^{MN}. Here below, the well-posedness of equation (6.3.19) is guaranteed by classical stochastic analysis in infinite dimension (see e.g. [DPZ92, Chapter 5]).

Theorem 6.3.12. Consider the Setting 6.2.5. For every $\epsilon > 0$ fixed, as $M,N\to\infty$ with scaling regime $\lim\sqrt{M}N^{-\alpha/d} = 0$, we have

$$\eta_t^{MN} \to \eta_t^\epsilon \quad \text{in law in } C\big([0,T]; H_x^{-\,\mathrm{e}_\alpha^+}\mathscr{H}^{-\,\kappa_2,\theta_2}\big),$$

where η^ϵ is the unique weak solution in $H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-(\kappa_2+2),\theta_2-2}$ of the Langevin SPDE

$$\eta_t^\epsilon = \eta_0 + \int_0^t \mathscr{L}_r(f,f)^*[\eta_r^\epsilon]\,dr + G_t^\epsilon. \tag{6.3.19}$$

6.4 Proof of the results

Proof of Proposition 6.3.2. By polarization it is sufficient to prove the result when $\varphi = \psi$. We introduce the following random time dependent probability measures on $Q^2 \times \mathbb{R}^{2d_v}$:

$$\bar{f}^2_{MN}(t) = \frac{1}{MN^2} \sum_{i,j=1}^{N} \sum_{k=1}^{M} \delta_{\left(x_i, x_j, \bar{u}^\epsilon_{ik}(t), \bar{u}^\epsilon_{jk}(t)\right)} \quad \text{and} \quad \bar{f}^2_N(t) = \frac{1}{N^2} \sum_{i,j=1}^{N} \delta_{(x_i,x_j)} \otimes f^{2,\epsilon}(x_i, x_j, t, du, dv).$$

We consider the splitting

$$\begin{aligned}
&\Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f_{MN}))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f_{MN}))\, R^\epsilon(x-y),\, df^2_{MN}(r,dx,dy,du,dv) \Big\rangle \\
&\quad - \Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f))\, R^\epsilon(x-y),\, f^{2,\epsilon}(r,dx,dy,du,dv) \Big\rangle \\
&= \Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f_{MN}))\, (\nabla_u \psi(y,v)(\sigma(y,r,v,f_{MN}) - \sigma(y,r,v,f)))\, R^\epsilon(x-y),\, f^2_{MN} \Big\rangle \\
&\quad + \Big\langle (\nabla_u \psi(x,u)(\sigma(x,r,u,f_{MN}) - \sigma(x,r,u,f)))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f))\, R^\epsilon(x-y),\, f^2_{MN} \Big\rangle \\
&\quad + \Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f))\, R^\epsilon(x-y),\, f^2_{MN} - \bar{f}^2_{MN} \Big\rangle \\
&\quad + \Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f))\, R^\epsilon(x-y),\, \bar{f}^2_{MN} - \bar{f}^2_N \Big\rangle \\
&\quad + \Big\langle (\nabla_u \psi(x,u)\sigma(x,r,u,f))\, (\nabla_u \psi(y,v)\sigma(y,r,v,f))\, R^\epsilon(x-y),\, \bar{f}^2_N - f^{2,\epsilon} \Big\rangle \\
&= E^1_r + E^2_r + E^3_r + E^4_r + E^5_r.
\end{aligned} \tag{6.4.1}$$

We consider each term E^i_r separately and we show that they satisfy $\sup_{r\in[0,T]} \mathbb{E}\big[|E^i_r|\big] \to 0$.

For the first term we compute, for a constant $C = C(T, b, \sigma)$,

$$\begin{aligned}
&\sup_{r\in[0,T]} \mathbb{E}\big[|E^1_r|\big] \\
&\leqslant \sup_{r\in[0,T]} \mathbb{E}\bigg[\int_{Q^2\times\mathbb{R}^{2d_v}} |\nabla_u \psi(x,u)|\, |\sigma(x,r,u,f_{MN})|\, |\nabla_u \psi(y,v)| \\
&\qquad\qquad \cdot |\sigma(y,r,v,f_{MN}) - \sigma(y,r,v,f))|\, |R^\epsilon(x-y)|\; df^2_{MN} \bigg] \\
&\leqslant C \sup_{r\in[0,T]} \mathbb{E}\bigg[\int_{Q^2\times\mathbb{R}^{2d_v}} \|\psi\|^2_{H^{e_\alpha}_x \mathscr{H}^{\kappa_2,\theta_2}} (1+|u|^{\theta_2+1})(1+|v|^{\theta_2}) \big(\mathcal{W}_1(Q\times\mathbb{R}^{d_v})(f_{MN}, f)\big)^\alpha\, df^2_{MN} \bigg] \\
&\leqslant C\, \|\psi\|^2_{H^{e_\alpha}_x \mathscr{H}^{\kappa_2,\theta_2}} \sup_{r\in[0,T]} \mathbb{E}\bigg[\int_{Q^2\times\mathbb{R}^{2d_v}} 1 + |u|^{4\theta_2+2} + |v|^{4\theta_2}\; df^2_{MN} \bigg]^{1/2} \\
&\qquad \cdot \sup_{r\in[0,T]} \mathbb{E}\Big[\big(\mathcal{W}_1(Q\times\mathbb{R}^{d_v})(f_{MN}, f)\big)^2 \Big]^{\alpha/2} \\
&\leqslant C\, \|\psi\|^2_{H^{e_\alpha}_x \mathscr{H}^{\kappa_2,\theta_2}} \Big(1 + \sup_{x\in Q} \mathbb{E}\big[|u_k(x,0)|^{4\theta_2+2}\big] \Big)^{1/2} \sup_{r\in[0,T]} \mathbb{E}\Big[\big(\mathcal{W}_2(Q\times\mathbb{R}^{d_v})(f_{MN}, f)\big)^2 \Big]^{\alpha/2}.
\end{aligned} \tag{6.4.2}$$

In the second passage we used the embedding (6.2.13) to bound $|\nabla_u \psi|$, the linear growth and Lipschitz properties (6.2.9) of σ and the boundedness of R^ϵ. The third passage follows from several application of Hölder's and Young's inequality. In the last passage we used the definition of f^2_{MN} and the moment estimates (6.2.18).

Swapping the roles of (x, u) and (y, v) and of f_{MN} and f in (6.4.2), we obtain an identical estimate for the second term E^2_r. In particular, we notice that the right hand side of (6.4.2) vanishes as $M, N \to \infty$ thanks to Theorem 6.2.11.

For the third term in (6.4.1) we compute, for a constant $C = C(T, b, \sigma)$,

$$
\begin{aligned}
|E_r^3| \leqslant \frac{C}{MN^2} \sum_{i,j=1}^{N} \sum_{k=1}^{M} \Big(& |\nabla_u \psi(x_i, u_{ik}^\epsilon) - \nabla_u \psi(x_i, \bar{u}_{ik}^\epsilon)| \\
& \qquad \cdot |\sigma(x_i, u_{ik}^\epsilon, r, f) \nabla_u \psi(x_j, u_{jk}^\epsilon \sigma(x_j, u_{jk}^\epsilon, r, f)| \\
& + |\nabla_u \psi(x_i, \bar{u}_{ik}^\epsilon)| |\sigma(x_i, u_{ik}^\epsilon, r, f) - \sigma(x_i, \bar{u}_{ik}^\epsilon, r, f)| \\
& \qquad \cdot |\nabla_u \psi(x_j, u_{jk}^\epsilon) \sigma(x_j, u_{jk}^\epsilon, r, f)| \\
& + |\nabla_u \psi(x_i, \bar{u}_{ik}^\epsilon)| |\sigma(x_i, \bar{u}_{ik}^\epsilon, r, f)| \\
& \qquad \cdot |\nabla_u \psi(x_j, u_{jk}^\epsilon) - \nabla_u \psi(x_j, \bar{u}_{jk}^\epsilon)| |\sigma(x_j, u_{jk}^\epsilon, r, f)| \\
& + |\nabla_u \psi(x_i, \bar{u}_{ik}^\epsilon)| |\sigma(x_i, \bar{u}_{ik}^\epsilon, r, f)| \\
& \qquad \cdot |\nabla_u \psi(x_j, \bar{u}_{jk}^\epsilon)| |\sigma(x_j, u_{jk}^\epsilon, r, f) - \sigma(x_j, \bar{u}_{jk}^\epsilon, r, f)| \Big) \\
\leqslant \frac{C}{MN^2} \sum_{i,j=1}^{N} \sum_{k=1}^{M} & \big(1 + \|\psi\|^3_{H_x^{e\alpha} \mathscr{H}^{\kappa_2, \theta_2}}\big) \\
& \cdot \big(1 + |u_{ik}^\epsilon| + |\bar{u}_{ik}^\epsilon| + |u_{jk}^\epsilon| + |\bar{u}_{jk}^\epsilon|\big)^{3\,\theta_2 + 1} \big(|u_{ik}^\epsilon - \bar{u}_{ik}^\epsilon| + |u_{jk}^\epsilon - \bar{u}_{jk}^\epsilon|\big).
\end{aligned}
\tag{6.4.3}
$$

In the first passage we unfolded the term E_r^3 using the definition of f_{MN}^2 and $\bar{f}_{MN}^2$, and then we added and subtracted several mixed term. In the second passage we used the Lipschitz and linear growth properties (6.2.9) of σ, the definition of the operators $V_{(x,u)}^1$ and $V_{(x,u),(y,v)}^{1,\mathrm{dif}}$ and the estimates (6.2.16) on their norms, and several applications of Hölder's and Young's inequality.

We now take the expectation and the supremum in time of (6.4.3), and then we apply the Cauchy–Schwarz inequality and the moment estimates (6.2.18) to obtain, for a constant $C = C(T, b, \sigma)$,

$$
\sup_{r \in [0,T]} \mathbb{E}\big[|E_r^3|\big] \leqslant C \big(1 + \|\psi\|^3_{H_x^{e\alpha} \mathscr{H}^{\kappa_2, \theta_2}}\big) \left(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x, 0)|^{6\,\theta_2 + 2}\big]^{1/2}\right) \sup_{i,k} \sup_{r \in [0,T]} \mathbb{E}\big[|u_{ik}^\epsilon(r) - \bar{u}_{ik}^\epsilon(r)|^2\big]^{1/2}.
$$

The right-hand side of this inequality vanishes in the $M, N \to \infty$ limit thanks to Theorem 6.2.8.

For the fourth term in (6.4.1) we compute, for a constant $C = C(T, b, \sigma)$,

$$
\begin{aligned}
& \sup_{r \in [0,T]} \mathbb{E}\big[|E_r^4|^2\big] \\
& \leqslant C \frac{1}{N^2} \sum_{i,j=1}^{N} \sup_{r \in [0,T]} \mathbb{E}\left[\left|\frac{1}{M} \sum_{k=1}^{M} \Big((\nabla_u \psi(x, u) \sigma(x, r, u, f)) \, (\nabla_u \psi(y, v) \sigma(y, r, v, f)) \right.\right. \\
& \qquad \left.\left. - \int_{\mathbb{R}^{2d_v}} (\nabla_u \psi(x, u) \sigma(x, r, u, f)) \, (\nabla_u \psi(y, v) \sigma(y, r, v, f)) \; f^{2,\epsilon}(r, x_i, x_j, du, dv) \Big) \right|^2 \right] \\
& \leqslant C \|\psi\|^4_{H_x^{e\alpha} \mathscr{H}^{\kappa_2, \theta_2}} \frac{1}{N^2 M^2} \sum_{i,j=1}^{N} \sum_{k=1}^{M} \sup_{r \in [0,T]} \mathbb{E}\Big[1 + |\bar{u}_{ik}^\epsilon|^{4\,\theta_2 + 4} + |\bar{u}_{jk}^\epsilon|^{4\,\theta_2 + 4} \\
& \qquad + \int_{\mathbb{R}^{2d_v}} |u|^{4\,\theta_2 + 2} + |v|^{4\,\theta_2 + 4} df^{2,\epsilon}(r, x_i, x_j) \Big] \\
& \leqslant \frac{C}{M} \|\psi\|^4_{H_x^{e\alpha} \mathscr{H}^{\kappa_2, \theta_2}} \left(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x, 0)|^{4\,\theta_2 + 4}\big]\right).
\end{aligned}
$$

In particular, we note that the last line vanishes in the $M, N \to \infty$ limit. In the first passage we used convexity inequalities. In the second passage we used that for fixed indices $i, j = 1, \dots, N$ the couples $\{(\bar{u}^\epsilon_{ik}, \bar{u}^\epsilon_{jk})\}_{k=1,\dots,M}$ are i.i.d. with law $f^{2,\epsilon}(x_i, x_j, t, du, dv)$, so that only diagonal terms survive when expanding the square of the sum over the index k. In the third passage we used the definition of the operator $V^1_{(x,u)}$ and the estimate (6.2.16) on its norms, the linear growth properties (6.2.9) of σ and several applications of Hölder's and Young's inequality. The last line follows from the moment estimates (6.2.18).

We finally consider the fifth term in (6.4.1), which is *deterministic*. We start with the splitting:

$$
\begin{aligned}
E^5_r = & \sum_{i,j=1}^N \int_{Q_i^N \times Q_j^N} \Bigg(R^\epsilon(x_i - x_j) \int_{\mathbb{R}^{2d_v}} \nabla_u \psi(x_i, u) \sigma(x_i, u) \nabla_u \psi(x_j, v) \sigma(x_j, v) f^{2,\epsilon}(r, x_i, x_j, du, dv) \\
& \qquad - R^\epsilon(x - y) \int_{\mathbb{R}^{2d_v}} \nabla_u \psi(x, u) \sigma(x, u) \nabla_u \psi(y, v) \sigma(y, v) f^{2,\epsilon}(r, x, y, du, dv) \Bigg) \, dx \, dy \\
= & \sum_{i,j=1}^N \int_{Q_i^N \times Q_j^N} (R^\epsilon(x_i - x_j) - R^\epsilon(x - y)) \\
& \qquad \cdot \int_{\mathbb{R}^{2d_v}} \nabla_u \psi(x_i, u) \sigma(x_i, u) \nabla_u \psi(x_j, v) \sigma(x_j, v) f^{2,\epsilon}(r, x_i, x_j, du, dv) \, dx \, dy \\
& + \sum_{i,j=1}^N \int_{Q_i^N \times Q_j^N} R^\epsilon(x - y) \Bigg(\int_{\mathbb{R}^{2d_v} \times \mathbb{R}^{2d_v}} \nabla_u \psi(x_i, u) \sigma(x_i, u) \nabla_u \psi(x_j, v) \sigma(x_j, v) \\
& \qquad - \nabla_u \psi(x, u) \sigma(x, u) \nabla_u \psi(y, v) \sigma(y, v) \ \pi_r(x_i, x_j, x, y, du, dv, du', dv') \Bigg) \, dx \, dy \\
= & E^{5.1}_r + E^{5.2}_r,
\end{aligned} \tag{6.4.4}
$$

where $\pi_r(x_i, x_j, x, y, du, dv, du', dv')$ denotes an optimal pairing in $\mathcal{W}_2(\mathbb{R}^{2d_v})$ between the measures $f^{2,\epsilon}(r, x_i, x_j, du, dv)$ and $f^{2,\epsilon}(r, x, y, du, dv)$, and where we have used the shorthand $\sigma(z, w) = \sigma(z, r, w, f)$.

For the first term in (6.4.4) we compute, for a constant $C = C(T, b, \sigma, \rho)$,

$$
\begin{aligned}
\left| E^{5.1}_r \right| \leqslant & C \sum_{i,j=1}^N \int_{Q_i^N \times Q_j^N} |R^\epsilon(x_i - x_j) - R^\epsilon(x - y)| \\
& \qquad \cdot \int_{\mathbb{R}^{2d_v}} \|\psi\|^2_{H^{e\alpha}_x \mathscr{H}^{\kappa_2, \theta_2}} \left(1 + |u|^{2\,\theta_2 + 2} + |v|^{2\,\theta_2 + 2}\right) f^{2,\epsilon}(r, x_i, x_j, du, dv) \, dx \, dy \\
\leqslant & C \|\psi\|^2_{H^{e\alpha}_x \mathscr{H}^{\kappa_2, \theta_2}} \left(1 + \sup_{x \in Q} \mathbb{E}\left[|u_k(x, 0)|^{2\,\theta_2 + 2} \right] \right) N^{-\alpha/d} \epsilon^{-1}.
\end{aligned} \tag{6.4.5}
$$

In the first inequality we used the estimate (6.2.16) on $V^1_{(x,u)}$, the linear growth (6.2.9) of σ and several applications of Hölder's and Young's inequality. In the second passage we used the moment estimates (6.2.18), the Lipschitzianity of R^ϵ, namely

$$
R^\epsilon(x) - R^\epsilon(y) = C_\rho \int_{\mathbb{R}^d} \left(\rho\left(x/\epsilon + z\right) - \rho\left(y/\epsilon + z\right) \right) \rho(z) \, dz \lesssim \min\left\{ \frac{x - y}{\epsilon}, 1 \right\},
$$

and the fact that $\operatorname{diam}(Q_i^N) \simeq N^{1/d}$ and $\operatorname{meas}(Q_i^N) = \frac{1}{N}$.

For the second term in (6.4.4) we compute, for a constant $C = C(T, b, \sigma)$,

$$
\begin{aligned}
\left|E_r^{5.2}\right| \leqslant & \sum_{i,j=1}^{N} \int_{Q_i^N \times Q_j^N} dx\, dy \int_{\mathbb{R}^{4d_v}} d\pi_r(x_i, x_j, x, y) \\
& \cdot \Big(|\nabla_u \psi(x_i, u) - \nabla_u \psi(x, u')| |\sigma(x_i, u, r, f) \nabla_u \psi(x_j, v) \sigma(x_j, v, r, f)| \\
& + |\nabla_u \psi(x, u')| |\sigma(x_i, u, r, f) - \sigma(x, u', r, f)| |\nabla_u \psi(x_j, v) \sigma(x_j, v, r, f)| \\
& + |\nabla_u \psi(x, u') \sigma(x, u', r, f)| |\nabla_u \psi(x_j, v) - \nabla_u \psi(y, v')| |\sigma(x_j, v, r, f)| \\
& + |\nabla_u \psi(x, u') \sigma(x, u', r, f) \nabla_u \psi(y, v)| |\sigma(x_j, v, r, f) - \sigma(y, v', r, f)| \Big) \\
\leqslant & \; C \sum_{i,j=1}^{N} \int_{Q_i^N \times Q_j^N} dx\, dy \int_{\mathbb{R}^{4d_v}} \big(\|\psi\|^2_{H_x^{\mathrm{e}\alpha} \mathscr{H}^{\kappa_2, \theta_2}} + 1 \big) \big(1 + |u| + |v| + |u'| + |v'| \big)^{2\theta_2 + 2} \\
& \cdot \big(|x_i - x|^\alpha + |x_j - y|^\alpha + |u - u'| + |v - v'| \big) \, \pi_r(x_i, x_j, x, y, du, dv, du', dv') \\
\leqslant & \; C \big(1 + \|\psi\|^2_{H_x^{\mathrm{e}\alpha} \mathscr{H}^{\kappa_2, \theta_2}} \big) \left(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x, 0)|^{4\theta_2 + 4} \big] \right)^{1/2} \\
& \cdot \sum_{i,j=1}^{N} \int_{Q_i^N \times Q_j^N} |x_i - x|^\alpha + |x_j - y|^\alpha + \mathcal{W}_2(\mathbb{R}^{2d_v})(f^{2,\epsilon}(r, x_i, x_j), f^{2,\epsilon}(r, x, y)) \, dx\, dy \\
\leqslant & \; C \big(1 + \|\psi\|^2_{H_x^{\mathrm{e}\alpha} \mathscr{H}^{\kappa_2, \theta_2}} \big) \left(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x, 0)|^{4\theta_2 + 4} \big] \right) \big(N^{-\alpha/d} + N^{-1/d} \epsilon^{-1} \big).
\end{aligned} \tag{6.4.6}
$$

In the first line we simply used the boundedness of R^ϵ and added and subtracted several mixed terms. In the second passage we used the Lipschitz and linear growth properties (6.2.9) of σ, the operators $V^1_{(x,u)}$ and $V^{1,\mathrm{dif}}_{(x,u),(y,v)}$ and the estimates (6.2.16) on their norms, and several applications of Hölder's and Young's inequality. In the third passage we used Hölder's inequality, the fact that π is an optimal pairing and the moment estimates (6.2.18). In the last passage we used Proposition 6.2.10 and that $\mathrm{diam}(Q_i^N) \simeq N^{1/d}$ and $\mathrm{meas}(Q_i^N) = \frac{1}{N}$.

Combining (6.4.4), (6.4.5) and (6.4.6), we have that $\sup_{r \in [0,T]} |E_r^5| \to 0$ as $M, N \to \infty$. This concludes the proof. □

Proof of Proposition 6.3.3. Consider a Hilbert basis $\{\psi_p\}_{p \geqslant 1}$ of $H_x^{\mathrm{e}\alpha} \mathscr{H}^{\kappa_1, \theta_1}$, we compute, for a constant $C = C(T, b, \sigma)$,

$$
\begin{aligned}
& \sup_{M,N} \mathbb{E}\left[\sup_{t \in [0,T]} \|M_t^{MN}\|^2_{H_x^{-\mathrm{e}\alpha} \mathscr{H}^{-\kappa_1, \theta_1}} \right] \\
& \leqslant \sup_{M,N} \mathbb{E}\left[\sum_{p \geqslant 1} \sup_{t \in [0,T]} |M_t^{MN}(\psi_p)|^2 \right] \\
& \leqslant C \sup_{M,N} \mathbb{E}\left[\sum_{p \geqslant 1} \langle M_T^{MN}(\psi_p) \rangle \right] \\
& \leqslant C \sup_{M,N} \frac{C_{MN}^2}{M} \frac{1}{MN^2} \sum_{k=1}^{M} \sum_{i,j=1}^{N} \int_0^T \mathbb{E}\left[\sum_{p \geqslant 1} |\nabla_u \psi_p(x_i, u_{ik}^\epsilon)|^2 |\sigma(x_i, u_{ik}^\epsilon, r, f_{MN})|^2 \right] dr \, R^\epsilon(x_i - x_j) \\
& \leqslant C \sup_{M,N} \frac{C_{MN}^2}{M} \frac{1}{MN} \sum_{k=1}^{M} \sum_{i=1}^{N} \int_0^T 1 + \mathbb{E}\big[|u_{ik}^\epsilon|^{2\theta_1 + 2} \big] \, dr \\
& \leqslant C \sup_{M,N} \frac{C_{MN}^2}{M} \left(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x, 0)|^{2\theta_1 + 2} \big] \right).
\end{aligned} \tag{6.4.7}
$$

In the first passage we simply used Parseval identity and in the second passage we used the Burkholder-Davis-Gundy inequality. In the third passage we used the definition (6.3.9) of the quadratic variation and the Cauchy–Schwarz inequality. In the fourth passage we used the Parseval identity

$$\|V^1_{x_i,u^\epsilon_{ik}}\|_{H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}} = \sum_{p\geqslant 1} |\nabla_u \psi_p(x_i, u^\epsilon_{ik})|^2$$

together with the estimates (6.2.16), the sublinear growth (6.2.9) of σ and the boundedness of R^ϵ. In the last passage we used the uniform moment estimates (6.2.18).

We now show the continuity of the trajectories. By (6.4.7) and Chebyshev's inequality we find a subset of full probability $\Omega_0 \subset \Omega$ such that

$$\forall \omega \in \Omega_0 \ \forall M, N \in \mathbb{N} \ \forall \delta > 0 \ \exists K = K(M, N, \omega, \delta) \text{ s.t. } \sum_{p\geqslant K} \sup_{t\in[0,T]} |M^{MN}_t(\psi_p)(\omega)|^2 < \delta, \quad (6.4.8)$$

and such that $r \mapsto M^{MN}_r(\psi_p)(\omega)$ is continuous for each $p \in \mathbb{N}$. Now for any $\omega \in \Omega_0$ and any $s, t \in [0, T]$, we compute

$$\begin{aligned} \|M^{MN}_t - M^{MN}_s\|_{H_x^{-\mathrm{e}\alpha}\mathscr{H}^{-\kappa_1,\theta_1}} &= \sum_{p\geqslant 1} |M^{MN}_t(\psi_p) - M^{MN}_s(\psi_p)|^2 \\ &\leqslant \sum_{p=1}^{K-1} |M^{MN}_t(\psi_p) - M^{MN}_s(\psi_p)|^2 + 2\sum_{p\geqslant K} \sup_{r\in[0,T]} |M^{MN}_r(\psi_p)|^2 \\ &\leqslant \sum_{p=1}^{K-1} |M^{MN}_t(\psi_p) - M^{MN}_s(\psi_p)|^2 + 2\delta. \end{aligned} \quad (6.4.9)$$

Continuity follows since $r \mapsto M^{MN}_r(\psi_p)$ is continuous for each $p \in \mathbb{N}$ and $\omega \in \Omega_0$. □

Proof of Proposition 6.3.4. We start with the splitting

$$\eta^{MN}_t = \mathrm{c}_{\mathrm{MN}}(f_{MN} - \bar{f}_{MN}) + \mathrm{c}_{\mathrm{MN}}(\bar{f}_{MN} - \bar{f}_N) + \mathrm{c}_{\mathrm{MN}}(\bar{f}_N - f),$$

for the measures

$$\bar{f}_{MN}(t, dx, du) := \frac{1}{MN} \sum_{j=1}^N \sum_{k=1}^M \delta_{(x_j, \bar{u}^\epsilon_{jk})} \ \text{ and } \ \bar{f}_N(t, dx, du) := \frac{1}{N} \sum_{j=1}^N \delta_{x_j} \otimes f(t, x_j, du).$$

For a Hilbert basis of $\{\psi_p\}_{p\geqslant 1}$ of $H_x^{\mathrm{e}\alpha}\mathscr{H}^{\kappa_1,\theta_1}$, Parseval identity and convexity inequalities yield, for a numeric constant C,

$$\begin{aligned} \|\eta^{MN}_t\|_{H_x^{\mathrm{e}\alpha}\mathscr{H}^{\kappa_1,\theta_1}} \leqslant C \sum_{p\geqslant 1} &\langle \mathrm{c}_{\mathrm{MN}}(f_{MN} - \bar{f}_{MN}), \psi_p\rangle^2 \\ &+ \langle \mathrm{c}_{\mathrm{MN}}(\bar{f}_{MN} - \bar{f}_N), \psi_p\rangle^2 + \langle \mathrm{c}_{\mathrm{MN}}(\bar{f}_N - f), \psi_p\rangle^2. \end{aligned} \quad (6.4.10)$$

We handle each sum in (6.4.10) separately. The result will follow from (6.4.10) and estimates (6.4.11), (6.4.12) and (6.4.13) below. For the first one we compute, for a constant

$C = C(T, b, \sigma)$,

$$
\begin{aligned}
&\mathbb{E}\Big[\sum_{p\geqslant 1}\langle \mathrm{c_{MN}}(f_{MN} - \bar{f}_{MN}), \psi_p\rangle^2\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{MN}\sum_{i=1}^{N}\sum_{k=1}^{M}\mathbb{E}\Big[\sum_{p\geqslant 1}|\psi_p(x_i, u^\epsilon_{ik}) - \psi_p(x_i, \bar{u}^\epsilon_{ik})|^2\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{MN}\sum_{i=1}^{N}\sum_{k=1}^{M}\mathbb{E}\Big[\|V^{0,\mathrm{dif}}_{(x_i,u^\epsilon_{ik}),(x_i,\bar{u}^\epsilon_{ik})}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\,\mathscr{H}^{-\,\kappa_1,\theta_1}}\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{MN}\sum_{i=1}^{N}\sum_{k=1}^{M}\mathbb{E}\big[|u^\epsilon_{ik} - \bar{u}^\epsilon_{ik}|^4\big]^{1/2}\mathbb{E}\big[1 + |u^\epsilon_{ik}|^{4\,\theta_1} + |\bar{u}^\epsilon_{ik}|^{4\,\theta_1}\big]^{1/2} \\
&\quad\leqslant C\,\mathrm{c_{MN}}^2\left(\frac{1}{\sqrt{M}} + \frac{1}{N^{\alpha/d}}\right)^2\left(1 + \sup_{x\in Q}\mathbb{E}\big[|u_k(x,0)|^{4\,\theta_1}\big]^{1/2}\right).
\end{aligned}
\tag{6.4.11}
$$

In the first passage we used the definition of f_{MN} and $\bar{f}_{MN}$ and convexity properties. In the second passage we used the definition (6.2.15) of $V^{0,\mathrm{dif}}_{(x_i,u^\epsilon_{ik}),(x_i,\bar{u}^\epsilon_{ik})}$ and Parseval identity. In the third passage we exploited the estimate (6.2.16) on its norm and used the Cauchy–Schwarz inequality. Finally in the last passage we used the error estimates (6.2.19) and the moment estimates (6.2.18).

We now consider the second term in (6.4.10). For a constant $C = C(T, b, \sigma)$, we compute

$$
\begin{aligned}
&\mathbb{E}\Big[\sum_{p\geqslant 1}\langle \mathrm{c_{MN}}(\bar{f}_{MN} - \bar{f}_{N}), \psi_p\rangle^2\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}}{N}\sum_{i=1}^{N}\frac{1}{M^2}\mathbb{E}\Big[\sum_{p\geqslant 1}\Big|\sum_{k=1}^{M}\psi_p(x_i, \bar{u}^\epsilon_{ik}) - \int_{\mathbb{R}^{d_v}}\psi_p(x_i, v) f(x_i, t, dv)\Big|^2\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{N}\sum_{i=1}^{N}\frac{1}{M}\mathbb{E}\Big[\sum_{p\geqslant 1}\Big|\psi_p(x_i, \bar{u}^\epsilon_{ik}) - \int_{\mathbb{R}^{d_v}}\psi_p(x_i, v) f(x_i, t, dv)\Big|^2\Big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{M}\,\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\big[\sum_{p\geqslant 1}|\psi_p(x_i, \bar{u}^\epsilon_{ik})|^2\big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{M}\,\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\big[\|V^0_{x_i,\bar{u}^\epsilon_{ik}}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\,\mathscr{H}^{-\,\kappa_1,\theta_1}}\big] \\
&\quad\leqslant C\,\frac{\mathrm{c_{MN}}^2}{M}\left(1 + \sup_{x\in Q}\mathbb{E}\big[|u_k(x,0)|^{2\,\theta_1}\big]\right).
\end{aligned}
\tag{6.4.12}
$$

In the first passage we used the definition of $\bar{f}_{MN}$ and $\bar{f}_N$ and convexity. In the second passage we used that for each fixed $i = 1, \ldots, N$ the particles $\{\bar{u}^\epsilon_{ik}\}_{k=1,\ldots,M}$ are i.i.d. with law $f(x_i, t, dv)$, so that only diagonal terms survive when expanding the square of the sum over the index k. In the third passage we used $(a-b)^2 \leqslant 2a^2 + 2b^2$, Hölder's inequality and again that the $\bar{u}^\epsilon_{ik}$ are i.i.d. with law $f(x_i, t, dv)$. In the fourth passage we used the definition (6.2.15) of $V^0_{x_i,\bar{u}^\epsilon_{ik}}$ and Parseval identity. In the fifth passage we exploited the estimate (6.2.16) on its norm and the moment estimates (6.2.18).

Finally we consider the third term in (6.4.10). For every $x, y \in Q$, let $\pi_t(x_i, y, du, dv)$ denote an optimal pairing between $f(x, t, u)$ and $f(y, t, v)$ for the Wasserstein distance $\mathcal{W}_4(\mathbb{R}^{d_v})$.

For this *deterministic* term we compute, for a constant $C = C(T, b, \sigma)$,

$$
\begin{aligned}
&\sum_{p\geqslant 1} \langle \mathrm{c_{MN}}(\bar{f}_N - f), \psi_p \rangle^2 \\
&\quad = \sum_{p\geqslant 1} \mathrm{c_{MN}}^2 \Big(\sum_{i=1}^{N} \int_{Q_i^N} \int_{\mathbb{R}^{2d_v}} \psi_p(x_i, u) - \psi_p(y, v)\, \pi_t(x_i, y, du, dv)\, dy \Big)^2 \\
&\quad \leqslant C\, \mathrm{c_{MN}}^2 \sum_{i=1}^{N} \int_{Q_i^N} \int_{\mathbb{R}^{2d_v}} \sum_{p\geqslant 1} |\psi_p(x_i, u) - \psi_p(y, v)|^2\, \pi_t(x_i, y, du, dv)\, dy \\
&\quad \leqslant C\, \mathrm{c_{MN}}^2 \sum_{i=1}^{N} \int_{Q_i^N} \int_{\mathbb{R}^{2d_v}} (1 + |u| + |v|)^{2\,\theta_1} (|x_i - y|^{2\alpha} + |u - v|^2)\, \pi_t(x_i, y, du, dv)\, dy \\
&\quad \leqslant C\, \mathrm{c_{MN}}^2 \sum_{i=1}^{N} \int_{Q_i^N} \Big(\int_{\mathbb{R}^{d_v}} 1 + |u|^{4\,\theta_1} + |v|^{4\,\theta_1} \pi_t(x_i, y, du, dv) \Big)^{1/2} \\
&\qquad\qquad \cdot \Big(|x_i - y|^{4\alpha} + \mathcal{W}_4(\mathbb{R}^{d_v})(f(x_i, t, du), f(y, t, dv)) \Big)^{1/2} dy \\
&\quad \leqslant C\, \mathrm{c_{MN}}^2 \Big(1 + \sup_{x\in Q} \mathbb{E}\big[|u_k(x, 0)|^{4\,\theta_1} \big]^{1/2} \Big) N^{-2\alpha/d}.
\end{aligned}
\tag{6.4.13}
$$

In the first passage we simply used the definitions of $\bar{f}_N$, f and $\pi_t(x_i, y)$ and that $\mathrm{meas}(Q_i^N) = 1/N$. In the second passage we used Parseval identity for the quantity $\|V^{0,\mathrm{dif}}_{(x_i,u),(y,v)}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\,\mathscr{H}^{-\,\kappa_1,\theta_1}}$ and estimate (6.2.16). In the third inequality we used Hölder's inequality and the choice of π_t. In the last passage we used Proposition 6.2.10, the moment estimates (6.2.18) and the fact that $\mathrm{diam}(Q_i^N) \simeq N^{-1/d}$. □

Proof of Proposition 6.3.6. We prove the first bound in (6.3.15), the other bounds are proved almost identically. To streamline the proof, we consider two prototypical terms in the expression (6.3.7) for the operator $\mathscr{L}(f_{MN}, f)$, namely the second and the fourth. The remaining terms are treated analogously, and are simpler since they involve less derivatives in (x, u) or (y, v) and less 'unbounded products' of terms involving b, σ and ψ.

For the second term in (6.3.7), using the linear growth properties (6.2.9) of the coefficient σ and its derivative, we compute, for a constant $C = C(T, \sigma, d, d_v)$,

$$
\begin{aligned}
&\big\| \Delta_v \psi(y, v)\, \sigma(y, v, t, f_{MN})^2 \big\|^2_{H_x^{\mathrm{e}_\alpha}\,\mathscr{H}^{\kappa_1,\theta_1}} \\
&\quad := \sum_{j=0}^{\mathrm{e}_\alpha} \sum_{h=0}^{\kappa_1} \int_Q \int_{\mathbb{R}^{d_v}} \big| D_y^j D_v^h (\Delta_v\, \psi \sigma^2) \big|^2 (1 + |v|)^{-2\,\theta_1} dy\, dv \\
&\quad \leqslant C \int_Q \int_{\mathbb{R}^{d_v}} \Big(\sum_{j=0}^{\mathrm{e}_\alpha} \sum_{h=0}^{\kappa_1} \Big| D_y^j D_v^{h+2} \psi \Big|^2 \Big) \Big(\sum_{j=0}^{\mathrm{e}_\alpha} \sum_{h=0}^{\kappa_1} \Big| D_y^j D_v^h \sigma^2 \Big|^2 \Big) (1 + |v|)^{-2\,\theta_1} dy\, dv \\
&\quad \leqslant C \sum_{j=0}^{\mathrm{e}_\alpha} \sum_{h=0}^{\kappa_1+2} \int_Q \int_{\mathbb{R}^{d_v}} \Big| D_y^j D_v^h \psi \Big|^2 (1 + |v|)^{-2(\theta_1 - 2)} dv\, dy \\
&\quad =: C \|\psi\|_{H_x^{\mathrm{e}_\alpha}\,\mathscr{H}^{\kappa_1+2,\theta_1-2}} \leqslant C \|\psi\|_{H_x^{\mathrm{e}_\alpha}\,\mathscr{H}^{\kappa_2,\theta_2}}.
\end{aligned}
$$

Similarly for the fourth term in (6.3.7), we compute, for a constant $C = C(T, \sigma, d_v, d)$,

$$
\begin{aligned}
&\Big\| \int_{Q\times\mathbb{R}^{d_v}} \Delta_u \psi(x,u)\sigma_1(x,y,t,u,v) 2\sigma_0(x,t,u) \\
&\qquad\qquad \cdot \Big(\int_0^1 \dot\phi\big((1-\lambda)\sigma_1(x,t,u,f) + \lambda\sigma_1(x,t,u,f_{MN})\big) d\lambda \Big) f(t,dx,du) \Big\|^2_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}} \\
&\leqslant C \sum_{j=0}^{\mathrm{e}_\alpha}\sum_{h=0}^{\kappa_1} \int_{Q\times\mathbb{R}^{d_v}}\int_{Q\times\mathbb{R}^{d_v}} |\Delta_u\psi(x,u)|^2 |D_y^j D_v^h \sigma_1(x,y,t,u,v)|^2 \\
&\qquad\qquad\qquad \cdot |\sigma_0(x,t,u)|^2 |\dot\phi|^2_\infty (1+|v|)^{-2\,\theta_1} f(t,dx,du)\,dy\,dv \\
&\leqslant C \int_{Q\times\mathbb{R}^{d_v}}\int_{Q\times\mathbb{R}^{d_v}} \|\psi\|^2_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}} (1+|u|)^{4+2\,\theta_2}(1+|v|)^{2-2\,\theta_1} f(t,dx,du)\,dy\,dv \\
&\leqslant C\,\|\psi\|^2_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}} \Big(1 + \sup_{x\in Q}\mathbb{E}\big[|u_k(x,0)|^{4+2\,\theta_2}\big]\Big).
\end{aligned}
$$

In the first passage we simply used the definition of the norm $\|\cdot\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}$, differentiated under the integral sign and applied Hölder's inequality. In the second passage we used the boundedness of $\dot\phi$, the linear growth assumptions (6.2.7)-(6.2.8) on σ_0 and σ_1 and the embedding (6.2.13) to bound $\Delta_u\psi$. In the last passage, the integral in the variable v is finite since $2(\theta_1 - 1) > d_v$ and we applied the moment estimates (6.2.18) in the u variable. □

Proof of Proposition 6.3.7. We already know that η_t^{MN} takes values in $H^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1} \subseteq H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}$ and that for every $\psi \in H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}$ the process $\langle \eta_t^{MN}, \psi\rangle$ is a real coninuous semimartingale with decomposition (6.3.12). We are left with showing the bound (6.3.17) and the continuity of the trajectories in $H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}$. The decomposition (6.3.16) at the distribution level then follows immediately from the decomposition (6.3.12) and from the estimate (6.3.15), which implies that the integral in (6.3.16) is well-defined. For the bound (6.3.17), we consider a Hilbert basis $\{\psi_p\}_{p\geqslant 1}$ of $H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}$. For any $M, N \in \mathbb{N}$ we compute, for a constant $C = C(T, b, \sigma, d_v, d)$,

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \|\eta_t^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}}\Big] \\
&\leqslant \mathbb{E}\Big[\sum_{p\geqslant 1}\sup_{t\in[0,T]} |\langle \eta_t^{MN}, \psi_p\rangle|^2\Big] \\
&\leqslant C\Big(\mathbb{E}\Big[\|\eta_0^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}}\Big] + \mathbb{E}\Big[\|M_T^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}}\Big] \\
&\qquad\qquad + \int_0^T \mathbb{E}\Big[\|\mathscr{L}_r(f_{MN},f)^*[\eta_r^{MN}]\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}}\Big]\,dr\Big) \\
&\leqslant C\Big(\mathbb{E}\Big[\|\eta_0^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}}\Big] + \mathbb{E}\Big[\|M_T^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}}\Big] \\
&\quad + \sup_{r\in[0,T]}\operatorname*{ess\text{-}sup}_{\omega} \|\mathscr{L}_r(f_{MN},f)^*\|_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}, H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_2,\theta_2}} \sup_{r\in[0,T]}\mathbb{E}\big[\|\eta_r^{MN}\|^2_{H_x^{-\,\mathrm{e}_\alpha}\mathscr{H}^{-\,\kappa_1,\theta_1}}\big]\Big) \\
&\leqslant C\sup_{M,N}\left(\frac{\mathrm{c_{MN}}^2}{M} + \frac{\mathrm{c_{MN}}^2}{N^{2\alpha/d}}\right)\Big(1 + \sup_{x\in Q}\mathbb{E}\Big[|u(x,0)|^{4\,\theta_1}\Big]\Big).
\end{aligned}
\tag{6.4.14}
$$

In the first passage we used Parseval identity. In the second passage we exploited the decomposition (6.3.12) for each ψ_p, convexity and Hölder's inequality, and again several applications of the Parseval identity. In the last passage we simply inserted the estimates (6.3.13), (6.3.14) and (6.3.15).

The continuity of the trajectories in $H^e_x \mathscr{H}^{-\kappa_2,\theta_2}$ now follows from (6.4.14) using the exact same argument as in(6.4.8)-(6.4.9), with η^{MN}_t and $H^e_x \mathscr{H}^{-\kappa_2,\theta_2}$ replacing M^{MN}_t and $H^e_x \mathscr{H}^{-\kappa_1,\theta_1}$. $\square$

Proof of Proposition 6.3.8. We first take care of the martingale term. Thanks to the Ascoli–Arzelà criterion for compactness, we have to show that M^{MN}_t satisfies the following conditions.

(C1) For every t in a dense subset of $\mathbb{R}^+$ the laws of the r.v. M^{MN}_t are tight in $H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}$.

(C2) For every $T \geqslant 0$ and every $\beta, \lambda > 0$, there exists $\delta > 0$ such that

$$\limsup_{M,N\to\infty} \mathbb{P}\left(\sup_{t,s\in[0,T],\,|t-s|\leqslant\delta} \|M^{MN}_t - M^{MN}_s\|_{H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}} > \beta \right) < \lambda.$$

Condition (C1) is readily verified. Indeed, the compact embedding $H^{-e}_x \mathscr{H}^{-\kappa_1,\theta_1} \hookrightarrow H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}$, which follows from (6.2.14), ensures that balls of $H^{-e}_x \mathscr{H}^{-\kappa_1,\theta_1}$ are compact in $H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}$. In turn, the uniform estimate (6.3.13) and Chebyshev's inequality yield, for $C = C(T, b, \sigma)$,

$$\sup_{M,N} \mathbb{P}\left(\|M^{MN}_t\|_{H^{-\mathrm{e}_\alpha}_x \mathscr{H}^{-\kappa_1,\theta_1}} \geqslant R \right) \leqslant \frac{C}{R^2} \sup_{M,N} \mathbb{E}\left[\sup_{t\in[0,T]} \|M^{MN}_t\|^2_{H^{-\mathrm{e}_\alpha}_x \mathscr{H}^{-\kappa_1,\theta_1}} \right] \leqslant \frac{C}{R^2} \sup_{M,N} \frac{\mathrm{c_{MN}}^2}{M}.$$

Condition (C2) is verified using the well-known Aldous criterion (see e.g. [JM86, Theorem 2.2.2]). In our context, it establishes that condition (C2) is implied by the following condition.

(A) For every $T \geqslant 0$ and every $\beta, \lambda > 0$, there exist $\delta > 0$ and $K \in \mathbb{N}$ such that, for every sequence τ_{MN} of $\mathcal{F}_t$-stopping times satisfying $\tau_{MN} \leqslant T$ almost surely, we have

$$\sup_{M,N\geqslant K} \sup_{\theta\leqslant\delta} \mathbb{P}\left(\|M^{MN}_{\tau_{MN}+\theta} - M^{MN}_{\tau_{MN}}\|_{H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}} > \beta \right) < \lambda.$$

Condition (A) follows immediately from the next inequality. Let $\{\psi_p\}_{p\geqslant1}$ be a Hilbert basis of $H^{\mathrm{e}_\alpha}_x \mathscr{H}^{\kappa_2,\theta_2}$. We compute, for a constant $C = C(T, b, \sigma, d_v, d)$,

$$\begin{aligned}
&\mathbb{P}\Big(\|M^{MN}_{\tau_{MN}+\theta} - M^{MN}_{\tau_{MN}}\|_{H^{-\mathrm{e}^+_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}} > \beta \Big) \\
&\leqslant \frac{1}{\beta^2} \mathbb{E}\Big[\sum_{p\geqslant1} \big| M^{MN}_{\tau_{MN}+\theta}(\psi_p) - M^{MN}_{\tau_{MN}}(\psi_p) \big|^2 \Big] \\
&\leqslant \frac{C}{\beta^2} \frac{\mathrm{c_{MN}}^2}{M} \mathbb{E}\Big[\sum_{p\geqslant1} \int_{\tau_{MN}}^{\tau_{MN}+\theta} \int_{Q^2\times\mathbb{R}^{2d_v}} \big(|\nabla_u \psi_p(x,u)|^2 + |\nabla_v \psi_p(y,v)|^2 \big) \\
&\qquad\qquad \cdot (1+|u|)(1+|v|)\, f^2_{MN}(r, dx, dy, du, dv)\, dr \Big] \\
&\leqslant \frac{C}{\beta^2} \frac{\mathrm{c_{MN}}^2}{M} \mathbb{E}\Big[\delta \sup_{r\in[0,T]} \int_{Q^2\times\mathbb{R}^{2d_v}} 1 + |u|^{2\,\theta_2+2} + |v|^{2\,\theta_2+2}\; f^2_{MN}(r, dx, dy, du, dv) \Big] \\
&\leqslant \frac{\delta}{\beta^2} C \sup_{M,N} \frac{\mathrm{c_{MN}}^2}{M} \Big(1 + \sup_{x\in Q} \mathbb{E}\big[|u_k(x,0)|^{2\,\theta_2+2} \big] \Big).
\end{aligned}$$

In the first line we used Parseval identity and Chebyshev's inequality. In the second line we used the Burkholder-Davis-Gundy inequality and the expression (6.3.9) for the quadratic variation, and then the Cauchy–Schwarz inequality, the linear growth (6.2.9) of σ and the boundedness of R^ϵ. In the third passage we used Parseval identity for the norm $\|V^1_{x,u}\|_{H^{-\mathrm{e}_\alpha}_x \mathscr{H}^{-\kappa_2,\theta_2}}$

and the bound (6.2.16) and Young's and Hölder's inequality. In the last passage we used the definition of f^2_{MN} and the moment estimates (6.2.18).

We now prove the tightness of the fluctuations η_t^{MN}. For this, we prove that each term of the semimartingale decomposition (6.3.16) is tight. We have just proved that the term M_t^{MN} is tight in $C([0,\infty); H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2})$. Furthermore, as in (6.4.8), condition (C1) holds for η_t^{MN} thanks to the compact embedding $H_x^{-\mathrm{e}}\mathscr{H}^{-\kappa_1,\theta_1} \hookrightarrow H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}$ and to estimate (6.3.14). In particular, condition (C1) holds for η_0^{MN} and in turn also for the integral term in (6.3.16). Since the initial data is independent of time, condition (C2) is automatically satisfied and η_0^{MN} is tight. To conclude, we just have to show that the integral term in (6.3.16) satisfies condition (C2). For this we compute, for a constant $C = C(T,b,\sigma)$,

$$
\begin{aligned}
&\mathbb{P}\Big(\sup_{|s-t|\leqslant\delta}\Big\|\int_s^t \mathscr{L}_r(f_{MN},f)^*[\eta_r^{MN}]\,dr\Big\|_{H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}} > \beta\Big)\\
&\leqslant \mathbb{P}\left(\sup_{|s-t|\leqslant\delta}\int_s^t \|\mathscr{L}_r(f_{MN},f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}\|\eta_r^{MN}\|_{H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}}\,dr > \beta\right)\\
&\leqslant \mathbb{P}\left(\delta^{1/2}\sup_{t\in[0,T]}\operatorname*{ess-sup}_{\omega\in\Omega}\|\mathscr{L}_t(f_{MN},f)\|_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}\Big(\int_0^T\|\eta_r^{MN}\|^2_{H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}}\,dr\Big)^{1/2} > \beta\right)\\
&\leqslant \frac{\delta}{\beta^2}\sup_{t\in[0,T]}\operatorname*{ess-sup}_{\omega\in\Omega}\|\mathscr{L}_t(f_{MN},f)\|^2_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}\,\mathbb{E}\Big[\int_0^T\|\eta_r^{MN}\|^2_{H_x^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}}\,dr\Big]\\
&\leqslant \frac{\delta}{\beta^2}\sup_{MN}\left(\sup_{t\in[0,T]}\operatorname*{ess-sup}_{\omega\in\Omega}\|\mathscr{L}_t(f_{MN},f)\|^2_{H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2},H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_1,\theta_1}}T\sup_{r\in[0,T]}\mathbb{E}\Big[\|\eta_r^{MN}\|^2_{H^{-\mathrm{e}_\alpha}\mathscr{H}^{-\kappa_1,\theta_1}}\Big]\right).
\end{aligned}
$$

Now condition (C2) follows by using the estimates (6.3.15) and (6.3.14) and choosing δ suitably small. □

Proof of Proposition 6.3.11. By Proposition 6.3.8 we know that $\eta_0^{MN} = \mathrm{c_{MN}}(f_{MN}(0) - f(0))$ is tight in $H_x^{-\mathrm{e}_\alpha^+}\mathscr{H}^{-\kappa_2,\theta_2}$. For any $\psi \in H_x^{\mathrm{e}_\alpha}\mathscr{H}^{\kappa_2,\theta_2}$, we consider the splitting

$$
\begin{aligned}
\langle\eta_0^{MN},\psi\rangle =& \mathrm{c_{MN}}\left(\frac{1}{MN}\sum_{i=1}^N\sum_{k=1}^M\psi(x_i,u_k(x_i,0)) - \int_Q\int_{\mathbb{R}^{d_v}}\psi(x,v)\,f_0(x,dv)\right)\\
=&\frac{\mathrm{c_{MN}}}{M}\sum_{k=1}^M\left(\frac{1}{N}\sum_{i=1}^N\psi(x_i,u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}}\psi(x_i,v)\,f_0(x_i,dv)\right)\\
&+\mathrm{c_{MN}}\sum_{i=1}^N\int_{Q_i^N}\left(\int_{\mathbb{R}^{d_v}}\psi(x_i,v)\,f_0(x_i,dv) - \int_{\mathbb{R}^{d_v}}\psi(x,v)\,f_0(x,dv)\right)dx\\
=&Y_1+Y_2.
\end{aligned}
\tag{6.4.15}
$$

We claim that $Y_2 \to 0$ as $M,N\to\infty$ along the scaling $\mathrm{c_{MN}}\,N^{-\alpha/d}\to 0$. Indeed we compute, for $\pi(x_i,x,du,dv)$ an optimal pairing in $\mathcal{W}_2(\mathbb{R}^{d_v})$ between $f_0(x_i,du)$ and $f_0(x,dv)$, for a constant $C = C(f_0)$,

$$
\begin{aligned}
|Y_2| &\leqslant c_{\mathrm{MN}} \sum_{i=1}^{N} \int_{Q_i^N} \int_{\mathbb{R}^{2d_v}} |\psi(x_i,u) - \psi(x,v)|\, \pi(x_i,x,du,dv)\, dx \\
&\leqslant C\|\psi\|_{H_x^{e_\alpha^+} \mathscr{H}^{\kappa_2,\theta_2}}\, c_{\mathrm{MN}} \sum_{i=1}^{N} \int_{Q_i^N} \int_{\mathbb{R}^{2d_v}} (1+|u|+|v|)^{\theta_2} (|x_i - x|^\alpha + |u-v|)\, \pi(x_i,x,du,dv)\, dx \\
&\leqslant C\|\psi\|_{H_x^{e_\alpha^+} \mathscr{H}^{\kappa_2,\theta_2}} \Big(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x,0)|^{2\,\theta_2}\big]\Big)^{1/2} \\
&\qquad \cdot c_{\mathrm{MN}} \sum_{i=1}^{N} \int_{Q_i^N} |x_i - x|^\alpha + \mathcal{W}_2(\mathbb{R}^{d_v})(f_0(x_i), f_0(x))\, dx \\
&\leqslant C\|\psi\|_{H_x^{e_\alpha^+} \mathscr{H}^{\kappa_2,\theta_2}} \Big(1 + \sup_{x \in Q} \mathbb{E}\big[|u_k(x,0)|^{2\,\theta_2}\big]\Big)^{1/2} c_{\mathrm{MN}}\, N^{-\alpha/d}.
\end{aligned}
\tag{6.4.16}
$$

In the second passage we used the estimate (6.2.16) on the operator $V^{0,\mathrm{dif}}_{(x_i,u),(x,v)}$. In the third passage we used Hölder's inequality, the definition of π and the moment estimates (6.2.18). In the last passage we used the Hölder continuity of f_0 and that $\mathrm{diam}(Q_i^N) \simeq N^{-1/d}$ and $\mathrm{meas}(Q_i^N) = \frac{1}{N}$.

We now consider the first term in (6.4.15). We claim that its characteristic function satisfies

$$
\lim_{M,N\to\infty} \mathbb{E}[\exp(itY_1)] = e^{-\frac{t^2}{2}\mathcal{Q}(\psi,\psi)}. \tag{6.4.17}
$$

Combining this with (6.4.15) and (6.4.16) implies that, along a scaling regime such that $\sqrt{M} N^{-\alpha/d} \to 0$,

$$
\langle \eta_0^{MN}, \psi \rangle \to \mathcal{N}\big(0, \mathcal{Q}(\psi,\psi)\big) \quad \text{in law for every } \psi \in H_x^{e_\alpha} \mathscr{H}^{\kappa_2,\theta_2}.
$$

Hence, using the tightness of η_0^{MN} in $H_x^{-e_\alpha^+} \mathscr{H}^{-\kappa_2,\theta_2}$, we conclude that

$$
\eta_0^{MN} \to \mathcal{N}\big(0, \mathcal{Q}\big) \quad \text{in law in } H_x^{-e_\alpha^+} \mathscr{H}^{-\kappa_2,\theta_2}.
$$

We prove (6.4.17). Recalling that the collections $\{u_k(x_i,0)\}_{i=1,\dots,N}$ are i.i.d. for $k = 1,\dots,M$, we compute

$$
\begin{aligned}
\mathbb{E}[\exp(itY_1)] &= \left(\mathbb{E}\Big[\frac{it}{\sqrt{M}} \frac{1}{N} \sum_{i=1}^{N} \psi(x_i, u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}} \psi(x_i,v)\, f_0(x_i,dv)\Big]\right)^M \\
&= \exp\left(-\frac{t^2}{2}\mathbb{E}\Big[\Big(\tfrac{1}{N}\textstyle\sum_{i=1}^{N} \psi(x_i,u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}} \psi(x_i,v)\, f_0(x_i,dv)\Big)^2\Big] + O\big(M^{-1/2}\big)\right),
\end{aligned}
$$

where the expression $O\big(M^{-1/2}\big)$ depends on ψ and $(u_k(x,0))_{x\in Q}$, but is uniform in $N \in \mathbb{N}$ since

$$
\Big|\frac{1}{N}\sum_{i=1}^{N} \psi(x_i,u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}} \psi(x_i,v)\, f_0(x_i,dv)\Big|^p \leqslant \sup_{x\in Q}\Big|\psi(x,u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}} \psi(x,v)\, f_0(x,dv)\Big|^p \quad \forall p \geqslant 1.
$$

Therefore formula (6.4.17) will follow if we prove that

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\Big[\Big(\tfrac{1}{N}\textstyle\sum_{i=1}^{N} \psi(x_i,u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}} \psi(x_i,v)\, f_0(x_i,dv)\Big)^2\Big] \\
= \mathbb{E}\Big[\Big(\textstyle\int_Q \psi(x,u_k(x,0)) - \int_{\mathbb{R}^{d_v}} \psi(x,v)\, f_0(x,dv)\, dx\Big)^2\Big] = \mathcal{Q}(\psi,\psi).
\end{aligned}
$$

This follows from the next inequality. For $\pi(x_i, x, du, dv)$ an optimal pairing in $\mathcal{W}_2(\mathbb{R}^{d_v})$ between $f_0(x_i, du)$ and $f_0(x, dv)$, for a constant $C = C(f_0, d, d_v)$, we compute

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{1}{N}\sum_{i=1}^{N}\Big(\psi(x_i, u_k(x_i,0)) - \int_{\mathbb{R}^{d_v}}\psi(x_i,v)\, f_0(x_i,dv)\Big) - \Big(\int_Q \psi(x, u_k(x,0)) - \int_{\mathbb{R}^{d_v}}\psi(x,v)\, f_0(x,dv)\, dx\Big)\Big|^2\Big] \\
&\leqslant \sum_{i=1}^{N}\int_{Q_i^N} dx\ \mathbb{E}\Big[\Big| V^{0,\mathrm{dif}}_{(x_i,u_k(x_i,0)),(x,u_k(x,0))}(\psi) + \int_{\mathbb{R}^{2d_v}} V^{0,\mathrm{dif}}_{(x_i,u_k(x_i,0)),(x,u_k(x,0))}(\psi)\ \pi(x_i,x,du,dv)\Big|^2\Big] \\
&\leqslant C\|\psi\|_{H_x^{\mathrm{e}_\alpha}\,\mathscr{H}^{\kappa_2,\theta_2}}\Big(1+\sup_{x\in Q}\mathbb{E}\big[|u_k(x,0)|^{2\,\theta_2}\big]^{1/2}\Big) \\
&\qquad\qquad \cdot\sum_{i=1}^{N}\int_{Q_i^N}|x_i - x|^\alpha + \mathbb{E}\big[|u_k(x_i,0)-u_k(x,0)|^2\big]^{1/2} + \mathcal{W}_2(\mathbb{R}^{d_v})(f_0(x_i), f_0(x))\ dx \\
&\leqslant C\|\psi\|_{H_x^{\mathrm{e}_\alpha}\,\mathscr{H}^{\kappa_2,\theta_2}}\Big(1+\sup_{x\in Q}\mathbb{E}\big[|u_k(x,0)|^{2\,\theta_2}\big]^{1/2}\Big)\, N^{-\alpha/d}.
\end{aligned}$$

The first inequality follows from convexity. In the second passage we first used the estimate (6.2.16) on the operators $V^{0,\mathrm{dif}}$ and then used Hölder's inequality, the definition of π and the moment estimates (6.2.18). In the last passage we used the Hölder continuity of $f_0 \in C^\alpha(Q; \mathbb{P}_2(\mathbb{R}^{d_v})$ and $u_k(\cdot,0) \in C(Q; L^2(\Omega))$, and that $\mathrm{diam}(Q_i^N) \simeq N^{-1/d}$ and $\mathrm{meas}(Q_i^N) = \frac{1}{N}$. □

Proof of Theorem 6.3.12. By Proposition 6.3.8 the fluctuations η_t^{MN} form a tight sequence in $C([0,T]; H_x^{-\,\mathrm{e}_\alpha^+}\,\mathscr{H}^{-\,\kappa_2,\theta_2})$. Consider any subsequence (M_k, N_k), which we relabel (M, N), such that η_t^{MN} is converging in law to some limit element η_t^∞. By Theorem 6.2.11 and Proposition 6.3.10 and 6.3.11, the sequences η_0^{MN}, η_t^{MN}, M_t^{MN}, f_{MN} and f^2_{MN} are converging in law to η_0, η_t^∞, G_t^ϵ, f and $f^{2,\epsilon}$ in the space

$$\begin{aligned}
\mathbf{X} =& H_x^{-\,\mathrm{e}_\alpha^+}\,\mathscr{H}^{-\,\kappa_2,\theta_2} \times C\big([0,T]; H_x^{-\,\mathrm{e}_\alpha^+}\,\mathscr{H}^{-\,\kappa_2,\theta_2}\big)^2 \\
&\times C\big([0,T]; \mathscr{P}_2(Q\times\mathbb{R}^{d_v})\big) \times C\big([0,T]; \mathscr{P}_2(Q^2\times\mathbb{R}^{2d_v})\big).
\end{aligned}$$

Using Skorokhod's representation theorem, we find another probability space $\tilde{\Omega}$ and random sequences and limit elements supported in there such that

$$(\tilde{\eta}_0^{MN}, \tilde{\eta}_t^{MN}, \tilde{M}_t^{MN}, \tilde{f}_{MN}, \tilde{f}^2_{MN}) \to (\tilde{\eta}_0, \tilde{\eta}_t^\infty, \tilde{G}_t^\epsilon, f, f^{2,\epsilon}) \text{ in } \mathbf{X}, \text{ almost surely in } \tilde{\Omega}, \tag{6.4.18}$$

and such that each term has the same law as the corresponding term supported in Ω. In particular, $\tilde{\eta}_0$ is a Gaussian r.v. with mean zero and covariance $\mathcal{Q}$ given in (6.3.18), $\tilde{G}_t^\epsilon$ is a Gaussian process with mean zero and covariance function g_t^ϵ given in (6.3.10) and f, $f^{2,\epsilon}$ are the previous *deterministic* measures.

Fatou's lemma and estimate (6.3.17) yield $\mathbb{E}\Big[\sup_{r\in[0,T]}\|\tilde{\eta}_r^\epsilon\|^2_{H_x^{-\,\mathrm{e}_\alpha^+}\,\mathscr{H}^{-\,\kappa_2,\theta_2}}\Big] < \infty$. In particular, thanks to (6.3.15), the integral $\int_0^t \mathscr{L}_r(f,f)^*\tilde{\eta}_r^\epsilon\, dr$ is well-defined. Furthermore, we still have the equation, written in weak form for $\psi \in H^{\mathrm{e}_\alpha^+}\,\mathscr{H}^{\kappa_2+2,\theta_2-2}$, for every $t\in[0,T]$, almost surely in $\tilde{\Omega}$,

$$\langle \tilde{\eta}_t^{MN}, \psi\rangle - \int_0^t \langle \tilde{\eta}_r^{MN}, \mathscr{L}_r(\tilde{f}_{MN}, f)[\psi]\,\rangle\, dr = \langle \tilde{\eta}_0^{MN}, \psi\rangle + \langle \tilde{M}_t^{MN}, \psi\rangle. \tag{6.4.19}$$

We claim that, along sub-subsequences, equation (6.4.19) passes to the limit $M, N \to \infty$ and converges to

$$\tilde{\eta}^{\epsilon}_t - \int_0^t \mathscr{L}_r(f,f)^*[\tilde{\eta}^{\epsilon}_r]\, dr = \tilde{\eta}^{\epsilon}_0 + \tilde{G}^{\epsilon}_t. \tag{6.4.20}$$

That is, it converges to the Langevin SPDE (6.3.19) written in $\tilde{\Omega}$.

Assume the claim (6.4.19)-(6.4.20). It follows that $\tilde{\eta}^{\epsilon}_t$ must be the unique-in-law weak solution in $C([0,T]; H_x^{-\mathrm{e}_\alpha^+} \mathscr{H}^{-(\kappa_2+2),\theta_2-2})$ of the Langevin SPDE (6.4.20). Since by construction η^{∞}_t and $\tilde{\eta}^{\epsilon}_t$ have the same law in $C([0,T]; H_x^{-\mathrm{e}_\alpha^+} \mathscr{H}^{-\kappa_2,\theta_2})$ and since $\eta^{MN}_t \to \eta^{\infty}_t$ in law, we conclude that along the initial subsequence η^{MN}_t is converging in law to the solution of the SPDE (6.3.19). Since the converging subsequence (M_k, N_k) was arbitrary, we conclude that the whole sequence η^{MN}_t is converging to the solution of (6.3.19) and the theorem is proved.

We now prove the claim (6.4.19)-(6.4.20). By (6.4.18) we know that $\tilde{\eta}^{MN}_0 \to \tilde{\eta}_0$ and $\tilde{M}^{MN}_t \to \tilde{G}^{\epsilon}_t$ almost surely. Hence the right-hand side of (6.4.19) does converge to the right-hand side of (6.4.20).

We consider the left-hand side of the equations (6.4.19)-(6.4.20). Thanks to estimate (6.3.15), for every $\psi \in H_x^{\mathrm{e}_\alpha^+} \mathscr{H}^{\kappa_2+2,\theta_2-2}$ and $t \in [0,T]$, we have a continuous map

$$F^t_\psi : C([0,T]; H_x^{-\mathrm{e}_\alpha^+} \mathscr{H}^{-\kappa_2,\theta_2}) \to \mathbb{R} \mid F^t_\psi(\zeta) = \langle \zeta, \psi \rangle - \int_0^t \langle \zeta, \mathscr{L}_r(f,f)[\psi] \rangle\, dr.$$

Since we know that $\tilde{\eta}^{MN}_\cdot \to \tilde{\eta}^{\epsilon}_\cdot$ almost surely in $C([0,T]; H_x^{-\mathrm{e}_\alpha^+} \mathscr{H}^{-\kappa_2,\theta_2})$ and F^t_ψ is continuous, we obtain that $F^t_\psi(\tilde{\eta}^{MN}_\cdot) \to F^t_\psi(\tilde{\eta}^{\epsilon}_\cdot)$ almost surely. Hence the claim is proved upon showing that, almost surely along sub-subsequences,

$$\int_0^t \langle \tilde{\eta}^{MN}_r, \mathscr{L}_r(\tilde{f}_{MN}, f)[\psi] \rangle\, dr - \int_0^t \langle \tilde{\eta}^{MN}_r, \mathscr{L}_r(f,f)[\psi] \rangle\, dr \to 0.$$

For this, we compute

$$\begin{aligned} &\mathbb{E}\left[\left| \int_0^t \langle \tilde{\eta}^{MN}_r, \mathscr{L}_r(\tilde{f}_{MN}, f)[\psi] \rangle\, dr - \int_0^t \langle \tilde{\eta}^{MN}_r, \mathscr{L}_r(f,f)[\psi] \rangle\, dr \right|\right] \\ &\leqslant \mathbb{E}\left[\sup_{r\in[0,T]} \|\eta^{MN}_r\|^2_{H_x^{-\mathrm{e}_\alpha} \mathscr{H}^{-\kappa_2,\theta_2}} \right]^{1/2} \int_0^t \mathbb{E}\left[\left\| (\mathscr{L}_r(f_{MN}, f) - \mathscr{L}_r(f,f))[\psi] \right\|^2_{H_x^{\mathrm{e}_\alpha} \mathscr{H}^{\kappa_2,\theta_2}} \right]^{1/2} dr, \end{aligned} \tag{6.4.21}$$

where we can drop the $\tilde{\cdot}$ symbols and go back to working in Ω because of the equality in law. The first term on the right-hand side of (6.4.21) is bounded uniformly in M, N by the estimates (6.3.17) and we finally show that integrand in the second term vanishes uniformly in time as $M, N \to \infty$.

This fact is proved almost identically to Lemma 6.3.6. We show how to handle one of the terms in $\left\| (\mathscr{L}_r(\tilde{f}_{MN}, f) - \mathscr{L}_r(f,f))[\psi] \right\|^2_{H_x^{\mathrm{e}_\alpha} \mathscr{H}^{\kappa_2,\theta_2}}$, the others are treated analogously. Namely, for the difference of the second terms in the expression (6.3.7) for $\mathscr{L}_r(f_{MN}, f)$ and $\mathscr{L}_r(f,f)$, using the linear growth and Lipschitz properties (6.2.9), we compute, for a constant

$C = (T, b, \sigma, d, d_v)$,

$$\begin{aligned}
&\mathbb{E}\Big[\big\|\Delta_v\psi(y,v)\,\sigma(y,v,t,f_{MN})^2 - \Delta_v\psi(y,v)\,\sigma(y,v,t,f)^2\big\|^2_{H^{\mathrm{e}_\alpha}_x\,\mathscr{H}^{\kappa_2,\theta_2}}\Big] \\
&\quad\leqslant C\mathbb{E}\Big[\int_Q\int_{\mathbb{R}^{d_v}}\Big(\sum_{j=0}^{\mathrm{e}_\alpha}\sum_{h=0}^{\kappa_2}\big|D_y^jD_v^{h+2}\psi\big|^2\Big) \\
&\qquad\qquad\cdot\Big(\sum_{j=0}^{\mathrm{e}_\alpha}\sum_{h=0}^{\kappa_2}\big|D_y^jD_v^h(\sigma(y,v,t,f_{MN})^2-\sigma(y,v,t,f))^2\big|^2\Big)(1+|v|)^{-2\,\theta_2}dy\,dv\Big] \\
&\quad\leqslant C\mathbb{E}\Big[\int_Q\int_{\mathbb{R}^{d_v}}\Big(\sum_{j=0}^{\mathrm{e}_\alpha}\sum_{h=0}^{\kappa_2}\big|D_y^jD_v^{h+2}\psi\big|^2\Big)\big(\mathcal{W}_1(Q\times\mathbb{R}^{d_v})(f_{MN}(t),f(t))\big)^2(1+|v|)^{2-2\,\theta_2}dy\,dv\Big] \\
&\quad\leqslant C\|\psi\|_{H^{\mathrm{e}_\alpha}_x\,\mathscr{H}^{\kappa_2+2,\theta_2-1}}\mathbb{E}\Big[\big(\mathcal{W}_2(Q\times\mathbb{R}^{d_v})(f_{MN}(t),f(t))\big)^2\Big].
\end{aligned}$$

Now Theorem 6.2.11 ensures the last line vanishes uniformly in $t \in [0, T]$ as $M, N \to \infty$. □

Bibliography

[AB20] Haggai Agmon and Yoram Burak. A theory of joint attractor dynamics in the hippocampus and the entorhinal cortex accounts for artificial remapping and grid cell field-to-field variability. *eLife*, 9:e56894, 2020.

[ACL21] Alexander Aurell, Rene Carmona, and Mathieu Lauriere. Stochastic graphon games: Ii. the linear-quadratic case. Preprint, 2021.

[ACR21] Mario Ayala, Gioia Carinci, and Frank Redig. Higher order fluctuation fields and orthogonal duality polynomials. *Electronic Journal of Probability*, 26(none):1 – 35, 2021.

[AF03] R. Adams and J. Fournier. *Sobolev Spaces*. Elsevier, 2003.

[Ama77] Shun-ichi Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybernet.*, 27(2):77–87, 1977.

[Ama00] Herbert Amann. Compact embeddings of vector valued sobolev and besov spaces. *Glasnik Matematicki; Vol.35 No.1*, 35, 01 2000.

[AvR10] Sebastian Andres and Max-K von Renesse. Particle approximation of the wasserstein diffusion. *Journal of Functional Analysis*, 258(11):3879–3905, 2010.

[BAC19] Áine Byrne, Daniele Avitabile, and Stephen Coombes. Next-generation neural field model: The evolution of synchrony within patterns and waves. *Physical Review E*, 99(1):012313, 2019.

[BCnC11] F. Bolley, J. A. Cañizo, and J. A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–2210, 2011.

[BD19] Amarjit Budhiraja and Paul Dupuis. *Analysis and approximation of rare events*, volume 94 of *Probability Theory and Stochastic Modelling*. Springer, New York, 2019. Representations and weak convergence methods.

[BDF12] Amarjit Budhiraja, Paul Dupuis, and Markus Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *The Annals of Probability*, 40(1):74 – 102, 2012.

[BDM08] Amarjit Budhiraja, Paul Dupuis, and Vasileios Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability*, 36(4):1390 – 1420, 2008.

[BDS19] Amarjit Budhiraja, Paul Dupuis, and Michael Salins. Uniform large deviation principles for Banach space valued stochastic evolution equations. *Trans. Amer. Math. Soc.*, 372(12):8363–8421, 2019.

[BF09] Y. Burak and I.R. Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS Comput. Biol.*, 5(2):e1000291, 2009.

[BF12] Y. Burak and I. Fiete. Fundamental limits on persistent activity in networks of noisy neurons. *PNAS*, 109:17645–17650, 2012.

[BFFT12] Javier Baladron, Diego Fasoli, Olivier Faugeras, and Jonathan Touboul. Mean-field description and propagation of chaos in networks of hodgkin-huxley and fitzhugh-nagumo neurons. *The Journal of Mathematical Neuroscience*, 2:1–50, 2012.

[BFG93] Paul Hubert Bezandry, Xavier Fernique, and Gaston Giroux. A functional central limit theorem for a nonequilibrium model of interacting particles with unbounded intensity. *Journal of Statistical Physics*, 72(1-2):329–353, July 1993.

[BKL95] O. Benois, C. Kipnis, and C. Landim. Large deviations from the hydrodynamical limit of mean zero asymmetric zero range processes. *Stochastic Processes and their Applications*, 55(1):65–89, January 1995.

[BPR16] Viorel Barbu, Giuseppe Da Prato, and Michael Röckner. *Stochastic Porous Media Equations*, volume 2163 of *Lecture Notes in Mathematics*. Springer, 2016.

[BR14] Viorel Barbu and Michael Röckner. An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *Journal of the European Mathematical Society*, 17, 02 2014.

[BR17] Viorel Barbu and Michael Röckner. Nonlinear fokker-planck equations driven by gaussian linear multiplicative noise. *Journal of Differential Equations*, 08 2017.

[Bre11] Paul Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45:033001, 12 2011.

[Bre14] Paul C Bressloff. Stochastic neural field theory. *Neural Fields: Theory and Applications*, pages 235–268, 2014.

[Bre19] Paul Bressloff. Stochastic neural field model of stimulus-dependent variability in cortical neurons. *PLoS Comput Biology*, 01 2019.

[Bre23] Paul C Bressloff. A generalized dean-kawasaki equation for an interacting brownian gas in a partially absorbing medium. *arXiv preprint arXiv:2312.06294*, 2023.

[Bre24] Paul C Bressloff. Global density equations for interacting particle systems with stochastic resetting: From overdamped brownian motion to phase synchronization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 34(4), 2024.

[BSH18] Daniel Bush and Christoph Schmidt-Hieber. Computational models of grid cell firing. *Hippocampal Microcircuits: A Computational Modeler's Resource Book*, pages 585–613, 2018.

[BT17] Evelyn Buckwar and Andreas Thalhammer. Importance sampling techniques for stochastic partial differential equations. 2017.

[BVW15] Caroline Bauzet, Guy Vallet, and Petra Wittbold. A degenerate parabolic-hyperbolic Cauchy problem with a stochastic force. *Journal of Hyperbolic Differential Equations*, 12(03):501–533, September 2015.

[CCS23] José A. Carrillo, Andrea Clini, , and Susanne Solem. The mean field limit of stochastic differential equation systems modeling grid cells. *SIAM Journal on Mathematical Analysis*, 55(4):3602–3634, link, 2023.

[CCY19] José A. Carrillo, Katy Craig, and Yao Yao. *Aggregation-Diffusion Equations: Dynamics, Asymptotics, and Singular Limits*, pages 65–108. Springer International Publishing, 08 2019.

[CD22] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. 2. applications. *Kinetic and Related Models*, 15(6):1017, 2022.

[CDFO13] D. Crisan, J. Diehl, P. K. Friz, and H. Oberhauser. Robust filtering: Correlated noise and multidimensional observation. *The Annals of Applied Probability*, 23(5), Oct 2013.

[CF16] Michele Coghi and Franco Flandoli. Propagation of chaos for interacting particles subject to environmental noise. *The Annals of Applied Probability*, 26(3):1407 – 1442, 2016.

[CF21] Federico Cornalba and Julian Fischer. The dean-kawasaki equation and the structure of density fluctuations in systems of diffusing particles, 2021.

[CF23a] Andrea Clini and Benjamin Fehrman. A central limit theorem for nonlinear conservative SPDEs. *arXiv preprint*, link, 2023.

[CF23b] Federico Cornalba and Julian Fischer. Multilevel monte carlo methods for the dean-kawasaki equation from fluctuating hydrodynamics. *arXiv preprint arXiv:2311.08872*, 2023.

[CFIR23] Federico Cornalba, Julian Fischer, Jonas Ingmanns, and Claudia Raithel. Density fluctuations in weakly interacting particle systems via the dean-kawasaki equation, 2023.

[CG19] Michele Coghi and Benjamin Gess. Stochastic nonlinear fokker–planck equations. *Nonlinear Analysis*, 2019.

[CG22] Chenyang Chen and Hao Ge. Sample-path large deviation principle for a 2-d stochastic interacting vortex dynamics with singular kernel. *arXiv preprint arXiv:2205.11013*, 2022.

[Cha10] Pierre-Henri Chavanis. A stochastic keller–segel model of chemotaxis. *Communications in Nonlinear Science and Numerical Simulation*, 15(1):60–70, 2010.

[CHS22] José A. Carrillo, Helge Holden, and Susanne Solem. Noise-driven bifurcations in a neural field system modelling networks of grid cells. *J. Math. Biol.*, 85(4):Paper No. 42, 30, 2022.

[CKNP22] Le Chen, Davar Khoshnevisan, David Nualart, and Fei Pu. Central limit theorems for parabolic stochastic partial differential equations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 2022.

[Cli23a] Andrea Clini. On the fluctuations of an SDE system modelling grid cells. *to appear on Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, arXiv preprint*, link, 2023.

[Cli23b] Andrea Clini. Porous media equations with nonlinear gradient noise and dirichlet boundary conditions. *Stochastic Processes and their Applications*, 159:428–498, link, 2023.

[CP03] Gui-Qiang Chen and Benoit Perthame. Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations. *Annales de l'Institut Henri Poincaré, Analyse Non Linéaire*, 20:645–668, 07 2003.

[CRS23a] José A. Carrillo, Pierre Roux, and Susanne Solem. Noise-driven bifurcations in a nonlinear fokker–planck system describing stochastic neural fields. *Physica D: Nonlinear Phenomena*, 449:133736, 2023.

[CRS23b] José Antonio Carrillo, Pierre Roux, and Susanne Solem. Well-posedness and stability of a stochastic neural field in the form of a partial differential equation. *arXiv preprint arXiv:2307.08077*, 2023.

[CS18] Young-Pil Choi and Samir Salem. Propagation of chaos for aggregation equations with no-flux boundary conditions and sharp sensing zones. *Math. Models Methods Appl. Sci.*, 28(2):223–258, 2018.

[CS21] Joe P. Chen and Federico Sau. Higher-order hydrodynamics and equilibrium fluctuations of interacting particle systems. *Markov Processes and Related Fields*, 27(3):339–380, 2021.

[CS23] Federico Cornalba and Tony Shardlow. The regularised inertial dean–kawasaki equation: discontinuous galerkin approximation and modelling for low-density regime. *ESAIM: Mathematical Modelling and Numerical Analysis*, 57(5):3061–3090, 2023.

[CSZ19a] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. From weakly interacting particles to a regularised dean–kawasaki model, 2019.

[CSZ19b] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. A regularized dean-kawasaki model: Derivation and analysis. *SIAM J. Math. Anal.*, 51:1137–1187, 2019.

[CSZ21] Federico Cornalba, Tony Shardlow, and Johannes Zimmer. Well-posedness for a regularised inertial dean–kawasaki model for slender particles in several space dimensions. *Journal of Differential Equations*, 284:253–283, 2021.

[CT18] Tanguy Cabana and Jonathan D Touboul. Large deviations for randomly connected neural networks: I. spatially extended systems. *Advances in applied probability*, 50(3):944–982, 2018.

[CWZ+13] Jonathan J Couey, Aree Witoelar, Sheng-Jia Zhang, Kang Zheng, Jing Ye, Benjamin Dunn, Rafal Czajkowski, May-Britt Moser, Edvard I Moser, Yasser

Roudi, et al. Recurrent inhibitory circuitry as a mechanism for grid formation. *Nature neuroscience*, 16(3):318–324, 2013.

[DCG+18] Pierre-Michel Déjardin, Yann Cornaton, Pierre Ghesquière, Cyril Caliot, and R Brouzet. Calculation of the orientational linear and nonlinear correlation factors of polar liquids from the rotational dean-kawasaki equation. *The Journal of Chemical Physics*, 148(4), 2018.

[DCKD22] Nataša Djurdjevac Conrad, Jonas Köppl, and Ana Djurdjevac. Feedback loops in opinion dynamics of agent-based models with multiplicative noise. *Entropy*, 24(10):1352, 2022.

[DE97] Paul Dupuis and Richard Ellis. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1997.

[Dea96] David S Dean. Langevin equation for the density of a system of interacting langevin processes. *Journal of Physics A: Mathematical and General*, 29(24):L613–L617, Dec 1996.

[DFG20] Nicolas Dirr, Benjamin Fehrman, and Benjamin Gess. Conservative stochastic pde and fluctuations of the symmetric simple exclusion process, 2020.

[DFVE14] Aleksandar Donev, Thomas G Fai, and Eric Vanden-Eijnden. A reversible mesoscopic model of diffusion in liquids: from giant fluctuations to fick's law. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(4):P04004, 2014.

[DG87] Donald A Dawsont and Jürgen Gärtner. Large deviations from the mckean-vlasov limit for weakly interacting diffusions. *Stochastics: An International Journal of Probability and Stochastic Processes*, 20(4):247–308, 1987.

[DG17] Konstantinos Dareiotis and Benjamin Gess. Supremum estimates for degenerate, quasilinear stochastic partial differential equations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55, 12 2017.

[DG20] Konstantinos Dareiotis and Benjamin Gess. Nonlinear diffusion equations with nonlinear gradient noise. *Electronic Journal of Probability*, 25, 02 2020.

[DGPS23] Matías G Delgadino, Rishabh S Gvalani, Grigorios A Pavliotis, and Scott A Smith. Phase transitions, logarithmic sobolev inequalities, and uniform-in-time propagation of chaos for weakly interacting diffusions. *Communications in Mathematical Physics*, 401(1):275–323, 2023.

[DHV16] Arnaud Debussche, Martina Hofmanova, and Julien Vovelle. Degenerate Parabolic Stochastic Partial Differential Equations: Quasilinear case. *Annals of Probability*, 44(3):1916–1955, 2016.

[Din22] Hao Ding. A new particle approximation to the diffusive dean-kawasaki equation with colored noise. *arXiv preprint arXiv:2204.11309*, 2022.

[DKP22] Ana Djurdjevac, Helena Kremp, and Nicolas Perkowski. Weak error analysis for a nonlinear spde approximation of the dean-kawasaki equation, 2022.

[DLN01] Nicolas Dirr, Stephan Luckhaus, and Matteo Novaga. A stochastic selection principle in case of fattening for curvature flow. *Calculus of Variations*, 13:405–425, 12 2001.

[DOL+16] Jean-Baptiste Delfau, Hélène Ollivier, Cristóbal López, Bernd Blasius, and Emilio Hernández-García. Pattern formation with repulsive soft-core interactions: Discrete particle dynamics and dean-kawasaki equation. *Physical Review E*, 94(4):042120, 2016.

[DPZ92] Guiseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[dRST23] Gonçalo dos Reis, Greig Smith, and Peter Tankov. Importance sampling for mckean-vlasov sdes. *Applied Mathematics and Computation*, 453:128078, 2023.

[DSZ15] Nicolas Dirr, Marios Georgios Stamatakis, and Johannes Zimmer. Entropic and gradient flow formulations for nonlinear diffusion. *Journal of Mathematical Physics*, 57, 08 2015.

[DVNZ20] Francisco Delgado-Vences, David Nualart, and Guangqu Zheng. A Central Limit Theorem for the stochastic wave equation with fractional noise. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56(4):3020 – 3042, 2020.

[EPSS21] Antonio Esposito, Francesco S Patacchini, André Schlichting, and Dejan Slepčev. Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit. *Archive for Rational Mechanics and Analysis*, 240(2):699–760, 2021.

[ESR12] Abdelhadi Es-Sarhir and Max Renesse. Ergodicity of stochastic curve shortening flow in the plane. *SIAM J. Math. Analysis*, 44:224–244, 01 2012.

[ET10] G. Bard Ermentrout and David H. Terman. *Mathematical Foundations of Neuroscience*, volume 35 of *Interdisciplinary Applied Mathematics*. Springer, New York, 2010.

[Fan21] Wai-Tong Louis Fan. Stochastic pdes on graphs as scaling limits of discrete interacting systems. *Bernoulli*, 27(3):1899–1941, 2021.

[FFG92] Reńe Ferland, Xavier M. Fernique, and Gaston Giroux. Compactness of the fluctuations associated with some generalized nonlinear boltzmann equations. *Canadian Journal of Mathematics*, 44:1192 – 1205, 1992.

[FG13] N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162:707–738, 2013.

[FG19] Benjamin Fehrman and Benjamin Gess. Well-posedness of nonlinear diffusion equations with nonlinear, conservative noise. *Archive for Rational Mechanics and Analysis*, 233, 07 2019.

[FG21a] Benjamin Fehrman and Benjamin Gess. Path-by-path well-posedness of nonlinear diffusion equations with multiplicative noise. *Journal de Mathématiques Pures et Appliquées*, 148, 01 2021.

[FG21b] Benjamin Fehrman and Benjamin Gess. Well-posedness of the Dean–Kawasaki and the nonlinear Dawson–Watanabe equation with correlated noise, 2021.

[FG23] Benjamin Fehrman and Benjamin Gess. Non-equilibrium large deviations and parabolic-hyperbolic pde with irregular drift. *Inventiones mathematicae*, pages 1–64, 2023.

[FGG22] Benjamin Fehrman, Benjamin Gess, and Rishabh S. Gvalani. Ergodicity and random dynamical systems for conservative spdes, 2022.

[FGL21] Franco Flandoli, Lucio Galeati, and Dejun Luo. Delayed blow-up by transport noise. *Communications in Partial Differential Equations*, 46(9):1757–1788, 2021.

[FH14] Peter K. Friz and Martin Hairer. *A Course on Rough Paths*. Universitext. Springer, 2014.

[FI15] O. Faugeras and J. Inglis. Stochastic neural field equations: a rigorous footing. *J. Math. Biol.*, 71(2):259–300, 2015.

[Fla95] Franco Flandoli. *Regularity theory and stochastic flows for parabolic SPDEs*. Volume 9 of Stochastic Monographs. Gordon and Breach Science Publishers, 1995.

[Fla11] Franco Flandoli. *Random Perturbation of PDEs and Fluid Dynamic Models: École d'été de Probabilités de Saint-Flour XL–2010*, volume 2015. Springer Science & Business Media, 2011.

[FM97] Begoña Fernandez and Sylvie Méléard. A hilbertian approach for fluctuations on the mckean-vlasov model. *Stochastic Processes and their Applications*, 71(1):33–53, 1997.

[FPV88] Pablo A. Ferrari, Errico Presutti, and M. E. Vares. Non equilibrium fluctuations for a zero range process. *Annales de l'I.H.P. Probabilités et statistiques*, 24(2):237–268, 1988.

[FTC09] Olivier D Faugeras, Jonathan D Touboul, and Bruno Cessac. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in computational neuroscience*, 3:323, 2009.

[FV10] Peter Friz and Nicolas Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge Studies in Advanced Mathematics (120). Cambridge University Press, 2010.

[GC24] David Gómez-Castro. Beginner's guide to aggregation-diffusion equations. *SeMA Journal*, pages 1–57, 2024.

[Ges12] Benjamin Gess. Random attractors for degenerate stochastic partial differential equations. *Journal of Dynamics and Differential Equations*, 25, 06 2012.

[GGK22] Benjamin Gess, Rishabh S Gvalani, and Vitalii Konarovskyi. Conservative spdes as fluctuating mean field limits of stochastic gradient descent. *arXiv preprint arXiv:2207.05705*, 2022.

[GH16] Benjamin Gess and Martina Hofmanová. Well-posedness and regularity for quasilinear degenerate parabolic-hyperbolic spde. *Annals of Probability*, 46, 11 2016.

[GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular pdes. In *Forum of Mathematics, Pi*, volume 3, page e6. Cambridge University Press, 2015.

[GKNP14] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014.

[GLP99] Giambattista Giacomin, Joel Lebowitz, and Errico Presutti. Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In *Mathematical Surveys and Monographs*, pages 107–152. American Mathematical Society, nov 1999.

[GMM11] Lisa M Giocomo, May-Britt Moser, and Edvard I Moser. Computational models of grid cells. *Neuron*, 71(4):589–603, 2011.

[GMR06] Günther Grün, Klaus Mecke, and Markus Rauscher. Thin-film flow influenced by thermal noise. *Journal of Statistical Physics*, 122:1261–1291, 01 2006.

[GS14a] Benjamin Gess and Panagiotis Souganidis. Long-time behavior, invariant measures, and regularizing effects for stochastic scalar conservation laws. *Communications on Pure and Applied Mathematics*, 70, 11 2014.

[GS14b] Benjamin Gess and Panagiotis Souganidis. Scalar conservation laws with multiple rough fluxes. *Communications in Mathematical Sciences*, 13, 06 2014.

[GS16] Benjamin Gess and Panagiotis E. Souganidis. Stochastic non-isotropic degenerate parabolic–hyperbolic equations. *Stochastic Processes and their Applications*, 127:2961–3004, 2016.

[GSS23] Ioannis Gasteratos, Michael Salins, and Konstantinos Spiliopoulos. Importance sampling for stochastic reaction–diffusion equations in the moderate deviation regime. *Stochastics and Partial Differential Equations: Analysis and Computations*, pages 1–70, 2023.

[GWZ24] Benjamin Gess, Zhengyan Wu, and Rangrang Zhang. Higher order fluctuation expansions for nonlinear stochastic heat equations in singular limits. *arXiv preprint arXiv:2406.17892*, 2024.

[Hai14] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, mar 2014.

[HCR23] Elias Hess-Childs and Keefer Rowan. Higher-order propagation of chaos in $l2$ for interacting diffusions. *arXiv preprint arXiv:2310.09654*, 2023.

[HDCD+21] Luzie Helfmann, Nataša Djurdjevac Conrad, Ana Djurdjevac, Stefanie Winkelmann, and Christof Schütte. From interacting agents to density-based modeling with stochastic pdes. *Communications in Applied Mathematics and Computational Science*, 16(1):1–32, 2021.

[HFM+05] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436:801–806, 2005.

[HHMT24] Jasper Hoeksema, Thomas Holding, Mario Maurelli, and Oliver Tse. Large deviations for singularly interacting diffusions. *Ann. Inst. H. Poincaré Probab. Statist.*, 60(1):492–548, 2024.

[HLW18] Shulan Hu, Ruinan Li, and Xinyu Wang. Central limit theorem and moderate deviations for a class of semilinear spdes. *arXiv: Probability*, 2018.

[HM86] Masuyuki Hitsuda and Itaru Mitoma. Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. *Journal of Multivariate Analysis*, 19(2):311–328, 1986.

[HNV20] Jingyu Huang, David Nualart, and Lauri Viitasaari. A central limit theorem for the stochastic heat equation. *Stochastic Processes and their Applications*, 130(12):7170–7184, 2020.

[HNVZ19] Jingyu Huang, David Nualart, Lauri Viitasaari, and Guangqu Zheng. Gaussian fluctuations for the stochastic heat equation with colored noise. *Stochastics and Partial Differential Equations: Analysis and Computations*, 8:402–421, 2019.

[HR96] T. Frank H. Risken. *The Fokker-Planck equation: methods of solution and applications.* Springer series in synergetics. Springer, 2nd edition, 1996.

[Izh07] Eugene M Izhikevich. *Dynamical systems in neuroscience.* MIT press, 2007.

[JDI22] Marie Jardat, Vincent Dahirel, and Pierre Illien. Diffusion of a tracer in a dense mixture of soft particles connected to different thermostats. *Physical Review E*, 106(6):064608, 2022.

[JM86] Anatole Joffe and Michel Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18:20 – 65, 1986.

[JM18] Milton Jara and Otávio Menezes. Non-equilibrium fluctuations of interacting particle systems. *arXiv preprint arXiv:1810.09526*, 2018.

[JW17] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and appli-*

cations, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017.

[Kaw98] Kyozi Kawasaki. Microscopic analyses of the dynamical density functional equation of dense fluids. *Journal of Statistical Physics*, 93:527–546, 11 1998.

[KE13] Zachary P. Kilpatrick and Bard Ermentrout. Wandering bumps in stochastic neural fields. *SIAM J. Appl. Dyn. Syst.*, 12(1):61–94, 2013.

[KL99] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems.* Springer, 1999.

[KLvR19a] Vitalii Konarovskyi, Tobias Lehmann, and Max von Renesse. On dean–kawasaki dynamics with smooth drift potential. *Journal of Statistical Physics*, 178(3):666–681, nov 2019.

[KLvR19b] Vitalii Konarovskyi, Tobias Lehmann, and Max-K. von Renesse. Dean-Kawasaki dynamics: ill-posedness vs. triviality. *Electronic Communications in Probability*, 24(none):1 – 9, 2019.

[KM19] Misun Kim and Eleanor A Maguire. Can we study 3d grid codes non-invasively in the human brain? methodological considerations and fmri findings. *NeuroImage*, 186:667–678, 2019.

[KM23] Vitalii Konarovskyi and Fenna Müller. Dean-kawasaki equation with initial condition in the space of positive distributions. *arXiv preprint arXiv:2311.10006*, 2023.

[KR14] Christian Kuehn and Martin Riedler. Large deviations for nonlocal stochastic neural fields. *Journal of mathematical neuroscience*, 4:1, 04 2014.

[Kry03] N. V. Krylov. Brownian trajectory is a regular lateral boundary for the heat equation. *SIAM J. Math. Anal.*, 34(5):1167–1182, 2003.

[Kry12] Nicolai Krylov. A relatively short proof of itô's formula for spdes and its applications. *Stochastic Partial Differential Equations: Analysis and Computations*, 1, 08 2012.

[KT19] Christian Kuehn and Jonas M Tölle. A gradient flow formulation for the stochastic amari neural field model. *Journal of mathematical biology*, 79(4):1227–1252, 2019.

[KX99] Thomas Kurtz and Jie Xiong. Particle representations for a class of nonlinear spdes. *Stochastic Processes and their Applications*, 83, 05 1999.

[Lie96] Gary M. Lieberman. *Second order parabolic differential equations.* World Scientific, 1996.

[LL06] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. i – le cas stationnaire. *Comptes Rendus Mathematique - C R MATH*, 343:619–625, 11 2006.

[LL07] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2:229–260, 03 2007.

[LPS13] Pierre Louis Lions, Benoît Perthame, and Panagiotis E. Souganidis. Scalar conservation laws with rough (stochastic) fluxes. *Stochastics and Partial Differential Equations: Analysis and Computations*, 1(4):664–686, November 2013.

[LPS14] Pierre Louis Lions, Benoît Perthame, and Panagiotis E. Souganidis. Scalar conservation laws with rough (stochastic) fluxes; the spatially dependent case. *Stochastics and Partial Differential Equations: Analysis and Computations*, 2(4):517–538, December 2014.

[LR15] Wei Liu and Michael Röckner. *Stochastic Partial Differential Equations: an Introduction.* Universitext. Springer, 2015.

[LS84] P. Lions and A.S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37:511–537, 1984.

[LS98a] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic partial differential equations. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 326:1085–1092, 1998.

[LS98b] Pierre-Louis Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic partial differential equations: non-smooth equations and applications. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 327:735–741, 1998.

[LS00a] Pierre-Louis Lions and Panagiotis Souganidis. Fully nonlinear stochastic pde with semilinear stochastic dependence. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 331:617–624, 2000.

[LS00b] Pierre-Louis Lions and Panagiotis Souganidis. Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 331:783–790, 2000.

[LS02] P. Lions and Panagiotis Souganidis. Viscosity solutions of fully nonlinear stochastic partial differential equations. *Sūrikaisekikenkyūsho Kōkyūroku*, 1287, 01 2002.

[Lyo91] T. Lyons. On the non-existence of path integrals. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 432:281 – 290, 1991.

[Lyo98] Terry J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.

[Mar08] Mauro Mariani. Large deviations principle for stochastic scalar conservation laws. *Probab. Theory Related Fields*, 147:607–648, 05 2008.

[MB20] James N. MacLaurin and Paul C. Bressloff. Wandering bumps in a stochastic neural field: a variational approach. *Phys. D*, 406:132403, 9, 2020.

[MBJ+06] Bruce L. McNaughton, Francesco P. Battaglia, Ole Jensen, Edvard I Moser, and May-Britt Moser. Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, 2006.

[MMM17] B.L. McNaughton, E.I. Moser, and M.-B. Moser. Spatial representation in the hippocampal formation: a history. *Nat. Neurosci.*, 20:1448–1464, 2017.

[MST22] Oleksandr Misiats, Oleksandr Stanzhytskyi, and Ihsan Topaloglu. On global existence and blowup of solutions of stochastic keller–segel type equation. *Nonlinear Differential Equations and Applications NoDEA*, 29(1):3, 2022.

[MT99] Umberto Marini Bettolo Marconi and Pedro Tarazona. Dynamic density functional theory of fluids. *The Journal of Chemical Physics*, 110(16):8032–8044, Apr 1999.

[MT23] Avi Mayorcas and Milica Tomašević. Blow-up for a stochastic model of chemotaxis driven by conservative noise on r 2. *Journal of Evolution Equations*, 23(3):57, 2023.

[MZZ08] Salah-Eldin Mohammed, Tusheng Zhang, and Huaizhong Zhao. The stable manifold theorem for semilinear stochastic evolution equations and stochastic partial differential equations ii: Existence of stable and unstable manifolds. *Memoirs of the American Mathematical Society*, 196:1–105, 2008.

[PBS+96] David J Pinto, Joshua C Brumberg, Daniel J Simons, G Bard Ermentrout, and Roger Traub. A quantitative population model of whisker barrels: re-examining the wilson-cowan equations. *Journal of computational neuroscience*, 3:247–264, 1996.

[Per02] B. Perthame. *Kinetic Formulation of Conservation Laws.* Volume 21 of Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, 2002.

[QRV99] Jeremy Quastel, Fraydoun Rezakhanlou, and S. R. S. Varadhan. Large deviations for the symmetric simple exclusion process in dimensions $d \geqslant 3$. *Probability Theory and Related Fields*, 113:1–84, 1999.

[Rav92] K. Ravishankar. Fluctuations from the hydrodynamical limit for the symmetric simple exclusion in zd. *Stochastic Processes and their Applications*, 42(1):31–37, 1992.

[RD10] Edmund T Rolls and Gustavo Deco. *The noisy brain: stochastic dynamics as a principle of brain function.* Oxford university press, 2010.

[RHAPT22a] Nadhir Ben Rached, Abdul-Lateef Haji-Ali, Shyam Mohan Subbiah Pillai, and Raúl Tempone. Multilevel importance sampling for mckean-vlasov stochastic differential equation. *arXiv preprint arXiv:2208.03225*, 2022.

[RHAPT22b] Nadhir Ben Rached, Abdul-Lateef Haji-Ali, Shyam Mohan Subbiah Pillai, and Raúl Tempone. Single level importance sampling for mckean-vlasov stochastic differential equation. *arXiv preprint arXiv:2207.06926*, 2022.

[RM16] Michael Rockner and Ionuţ Munteanu. The total variation flow perturbed by gradient linear multiplicative noise. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 21, 05 2016.

[RRMM16] D. C. Rowland, Y. Roudi, M.-B. Moser, and E. I. Moser. Ten years of grid cells. *Annu. Rev. Neurosci.*, 39:19–40, 2016.

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion.* Number 293 in Grundlehren der mathematischen Wissenschaften. Springer, Berlin [u.a.], 3. ed edition, 1999.

[Seo17] Insuk Seo. Large-deviation principle for interacting brownian motions. *Communications on Pure and Applied Mathematics*, 70(2):203–288, 2017.

[Sim86] Jacques Simon. *Compact sets in the space $L^p(0,T;B)$*, volume 146. Annali di Matematica Pura ed Applicata, 01 1986.

[Sla21] Jakub Slav'ik. Large and moderate deviations principles and central limit theorem for the stochastic 3d primitive equations with gradient-dependent noise. *Journal of Theoretical Probability*, 2021.

[SM24] Richard E Spinney and Richard G Morris. A dean-kawasaki equation for reaction diffusion systems driven by poisson noise. *arXiv preprint arXiv:2404.02487*, 2024.

[Spo12] H. Spohn. *Large Scale Dynamics of Interacting Particles.* Springer, 2012.

[SS17] Michael Salins and Konstantinos Spiliopoulos. Rare event simulation via importance sampling for linear spde's. *Stochastics and Partial Differential Equations: Analysis and Computations*, 5:652–690, 2017.

[SS20] Justin Sirignano and Konstantinos Spiliopoulos. Mean field analysis of neural networks: A law of large numbers. *SIAM Journal on Applied Mathematics*, 80(2):725–752, 2020.

[SSS+12] Hanne Stensola, Tor Stensola, Trygve Solstad, Kristian Froland, May Britt Moser, and Edvard Moser. The entorhinal grid map is discretized. *Nature*, 492(7427):72–78, 2012.

[SY04] Panagiotis Souganidis and Nung Kwan Yip. Uniqueness of motion by mean curvature perturbed by stochastic noise. *Annales de l'Institut Henri Poincaré, Analyse Non Linéaire*, 21:1–23, 02 2004.

[Szn84a] Alain-Sol Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.*, 56(3):311–336, 1984.

[Szn84b] A.S. Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *Journal of Functional Analysis*, 56(3):311–336, 1984.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[THF12] Jonathan Touboul, Geoffroy Hermann, and Olivier Faugeras. Noise-induced behaviors in neural mean field dynamics. *SIAM Journal on Applied Dynamical Systems*, 11(1):49–81, 2012.

[TLDS23] Léo Touzo, Pierre Le Doussal, and Grégory Schehr. Interacting, running and tumbling: The active dyson brownian motion. *Europhysics Letters*, 142(6):61004, 2023.

[Töl18] J.M. Tölle. Estimates for nonlinear stochastic partial differential equations with gradient noise via dirichlet forms. *Springer Proceedings in Mathematics and Statistics*, 229:249–262, 2018.

[Tou11] Jonathan Touboul. Mean-field equations for stochastic neural fields with spatio-temporal delays. *arXiv preprint arXiv:1108.2414*, 2011.

[Tou12] J. Touboul. Mean-field equations for stochastic firing-rate neural fields with delays: Derivation and noise-induced transitions. *Phys. D*, 241(15):1223–1244, 2012.

[tVLW20] Michael te Vrugt, Hartmut Löwen, and Raphael Wittkowski. Classical dynamical density functional theory: from fundamentals to applications. *Advances in Physics*, 69(2):121–247, 2020.

[Vaz07] Juan Luis Vazquez. *The Porous Medium Equation: Mathematical Theory*. Oxford Science Publications, 2007.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[vNVW15] Jan van Neerven, Mark Veraar, and Lutz Weis. Stochastic integration in banach spaces – a survey. In Robert C. Dalang, Marco Dozzi, Franco Flandoli, and Francesco Russo, editors, *Stochastic Analysis: A Series of Lectures*, pages 297–332, Basel, 2015. Springer Basel.

[vRS09] Max-K. von Renesse and Karl-Theodor Sturm. Entropic measure and wasserstein diffusion. *The Annals of Probability*, 37(3):1114–1191, 2009.

[WC72] H. Wilson and J. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J.*, 12:1–24, 1972.

[WC73] H. Wilson and J. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cybern.*, 13:55–80, 1973.

[WPC+20] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

[WR23] Xingyu Wang and Chang-Han Rhee. Large deviations and metastability analysis for heavy-tailed dynamical systems, 2023.

[WWZ24] Likun Wang, Zhengyan Wu, and Rangrang Zhang. Dean-kawasaki equation with singular interactions and applications to dynamical ising-kac model, 2024.

[WZZ23] Zhenfu Wang, Xianliang Zhao, and Rongchan Zhu. Gaussian fluctuations for interacting particle systems with singular kernels. *Archive for Rational Mechanics and Analysis*, 247(5):101, 2023.

[YU13] Michael M Yartsev and Nachum Ulanovsky. Representation of three-dimensional space in the hippocampus of flying bats. *Science*, 340(6130):367–372, 2013.

[ZZG19] Rangrang Zhang, Guoli Zhou, and Boling Guo. Stochastic 2d primitive equations: Central limit theorem and moderate deviation principle. *Computers & Mathematics with Applications*, 77(4):928–946, 2019.