LUISS

IRIS - Luiss Guido Carli Research Institutional Repository | Catalogo istituzionale della ricerca Luiss Guido Carli

A mean-field model with discontinuous coefficients for neurons with spatial interaction

Post-print version of the following publication: | Versione post-print della seguente pubblicazione:

Original Citation/Citazione:
Flandoli, F.; Priola, E.; Zanco, Giovanni Alessandro. (2019). A mean-field model with discontinuous coefficients for neurons with spatial interaction. DISCRETE AND CONTINUOUS DYNAMICAL SYSTEMS, (ISSN: 1078-0947), 39:6, 3037-3067. Doi: 10.3934/dcds.2019126.

Availability/Disponibilità:
This version is available at: 11385/191543 since: 2020-01-09T18:25:32Z - Questa versione è disponibile alla pagina: 11385/191543 dal: 2020-01-09T18:25:32Z

Publisher/Casa editrice:

Published version/Pubblicato:
DOI: https://dx.doi.org/10.3934/dcds.2019126

License/Licenza:
DRM (Digital rights management) non definiti

Availability/Termini d'uso:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. Works made available under a Creative Commons license can be used according to the terms and conditions of said license. For all terms of use and more information see the publisher's website. | I termini e le condizioni relativi al riutilizzo della presente versione della pubblicazione sono disciplinati dalla politica editoriale. Le opere messe a disposizione con licenze Creative Commons possono essere utilizzate conformemente ai termini e alle condizioni previste da tali licenze. Per l'insieme delle condizioni di utilizzo e per ulteriori informazioni si rinvia al sito web dell'editore.

This item was downloaded from IRIS Luiss (https://iris.luiss.it/). When citing, please refer to the published version. | Questo documento è stato scaricato da IRIS Luiss (https://iris.luiss.it/). Per la citazione, fare riferimento alla versione pubblicata sul sito dell'editore.

(Article begins on next page | Il contributo inizia nella pagina successiva)

10 October 2026

Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

doi:10.3934/xx.xx.xx.xx

pp. **X–XX**

A MEAN-FIELD MODEL WITH DISCONTINUOUS COEFFICIENTS FOR NEURONS WITH SPATIAL INTERACTION

FRANCO FLANDOLI

Scuola Normale Superiore
Piazza dei Cavalieri 7, Pisa, Italy

ENRICO PRIOLA

Dipartimento di Matematica, Università di Torino
Via Carlo Alberto 10, Torino, Italy

GIOVANNI ZANCO*

Institute of Science and Technology Austria (IST Austria)
Am Campus 1, Klosterneuburg, Austria
LUISS Università Guido Carli
Viale Romania 32, Roma, Italy

(Communicated by the associate editor name)

ABSTRACT. Starting from a microscopic model for a system of neurons evolving in time which individually follow a stochastic integrate-and-fire type model, we study a mean-field limit of the system. Our model is described by a system of SDEs with discontinuous coefficients for the action potential of each neuron and takes into account the (random) spatial configuration of neurons allowing the interaction to depend on it. In the limit as the number of particles tends to infinity, we obtain a nonlinear Fokker-Planck type PDE in two variables, with derivatives only with respect to one variable and discontinuous coefficients. We also study strong well-posedness of the system of SDEs and prove the existence and uniqueness of a weak measure-valued solution to the PDE, obtained as the limit of the laws of the empirical measures for the system of particles.

1. **Introduction.** We propose a model for the action potential of N neurons, with positions fixed in time, that follow integrate-and-fire type dynamics subject to noise and interact with each other through their spikes. The interaction we consider depends also on the positions of the neurons and is of mean-field type. Therefore, in the limit as N tends to infinity each neuron interacts with infinitely many other neurons.
The presence of noise in the neuronal dynamics is experimentally confirmed and has been considered by various authors (see the monographs [9], [19]). Integrate-and-fire (IF) models describe a simplified dynamics in which such effect can be studied in detail. Considering large networks of interacting neurons, each one having a membrane potential that evolves following a IF dynamic, leads to modeling the mean-firing rate of the network as

2010 *Mathematics Subject Classification.* Primary: 60H10, 35Q84; Secondary: 92C20, 62M45.

Key words and phrases. Mean-field model for neurons, spatial interaction, SDEs with discontinuous coefficients, nonlinear Fokker-Planck PDEs, McKean-Vlasov equation.

The second author has been partially supported by INdAM through the GNAMPA Research Project (2017) "Sistemi stocastici singolari: buona posizione e problemi di controllo". The third author was partly funded by the Austrian Science Fund (FWF) project F 65.

* Corresponding author: Giovanni Zanco.

the solution to a nonlinear partial differential equation that, at least for mean-field type interactions, is of Fokker-Planck type.
Fokker-Planck PDEs for neural networks have been studied recently in [14], [7], [1], [2], [4], based on an IF model for the potential of each neuron given in [12]. As pointed out in [4], not much attention is paid in the literature to how the Fokker-Planck PDE is obtained; in particular one expects that the empirical measures of a network with N neurons converge as $N \to \infty$ to the solution of the PDE. This has been rigorously shown only in [3], proving convergence to a McKean-Vlasov stochastic differential equation, and in [5], where the hydrodynamic limit is considered.
The Fokker-Planck PDE obtained in the cited works exhibits blow-up in finite time, thus there is no global well-posedness, for certain ranges of parameters, due to the interaction term.
The model we propose here is simpler but it incorporates two additional aspects: a refractory period after the spike and a localized version of the interaction term, that is, an explicit dependence of the interaction on the positions of the neurons. The refractory period accounts for the fact that after emitting a spike, each neuron is inhibited from interaction. The dependence on a space variable allows to precisely prescribe the interation between different parts of the network; it can also describe the subdivision of the network in subpopulations, whose interaction with each other is of particular interest in neuroscience (see e.g. [11], [16]). This leads to a description of finite speed signal propagation along the network.
More precisely, our mean-field interaction term has two main features: first, it depends both on the positions and on the voltage of the neurons, unlike many models available in the literature; second, it contains indicator functions of suitable intervals in $\mathbb{R}$, thus requiring us to study a system of SDEs and a Fokker-Planck type PDE with irregular coefficients and dependence on the positions of the neurons that we treat as stochastic parameters.
We allow for great generality in the choice of the law of the positions of the neurons, only requiring finite first moment. Hence one can prescribe the geometry of the neural network choosing the law accordingly.
We study the limit behaviour of the empirical measures of the network and prove that the limit measure-valued function is the unique weak solution to a nonlinear PDE of Fokker-Planck type and that it exists up to any fixed time T, thus not exhibiting blow-up. From the technical point of view, to study the limit of the empirical measures we will also use some ideas of [15].
Our model includes discontinuous coefficients, and is therefore a first step in the study of stochastic interacting particle systems with irregular coefficients. Some of the results we obtain can be immediately generalized to the case of SDEs with measurable and bounded coefficients, but we are able to study the limit PDE only when the coefficients are discontinuous on a set with 0 Lebesgue measure (see, in particular, Lemma 5.4). Therefore we stick in the main part of the paper to the particular coefficients coming from the model, and mention some possible generalizations in Appendix.

The potential V of each neuron is modeled, as a function of time, with a stochastic differential equation whose solution is projected on a torus given by the interval $[0, 2]$ with the identification $0 \equiv 2$. This choice allows to model the cycle of spikes of each neuron as we describe below. It is important to notice that, similarly to what is done in most IF models, we do not give a precise description of the spike phenomenon, but we model only the charging phase from the resting potential $v_R = 0$ to the firing threshold $v_F = 1$ and

the refractory period after the spike; moreover we assume that there are no external input currents.

Consider a single cycle, that is $0 \leq V \leq 2$. As $0 \leq V < 1$ the neuron charges, subject to spikes by nearby neurons (i.e., to interaction), to randomness and to the effect of discharge with constant rate (that corresponds to the fact that if no spikes happen in the connected neurons then some charge is lost as time passes); when V reaches the threshold value 1 the neuron fires and emits a spike into the network. On a real neuron this would have two effects: the potential would rapidly decrease below 0 and then be restored to 0, and the neuron would be at rest, inhibited from interacting and spiking for a small amount of time (the refractory period). We model this effect "switching off" the interaction term when $V > 1$ and letting V evolve as $\mathrm{d}V = \mathrm{d}t$ until it reaches the value 2, where it is restored to 0 (through the equivalence relation that defines the torus) and the charging cycle begins again. Therefore the values of V between 1 and 2 do not correspond to a real life situation but are only a tool we resort to in order to have a convenient mathematical description of the phenomenon.

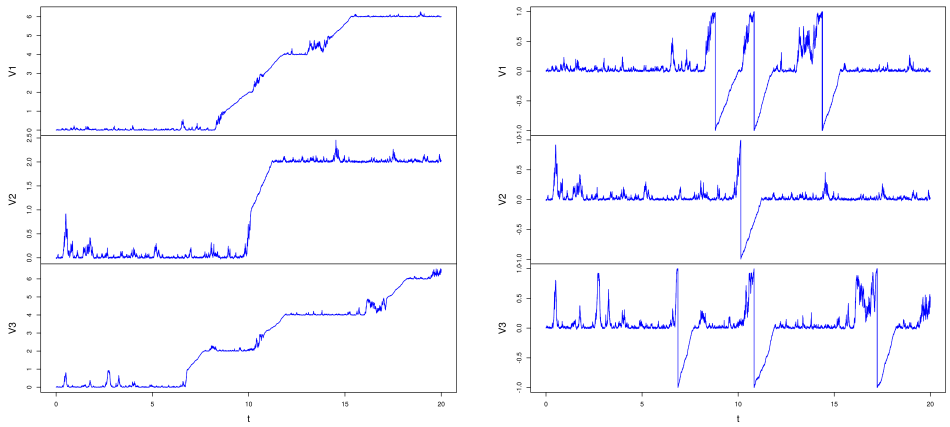
To consider the interaction between N neurons we deal with three factors (see also equation (2) below). Indeed if we consider the voltage $V^{i,N}$ and position X^i of the i-th neuron, following the description above, a factor $\theta(X^i, X^j)$ accounts for the neuron being connected to some of the other neurons with positions $(X^j)_j$; a factor $\mathbb{1}_{[0,1]}(V^i)$ is due to the fact that the neuron feels the interaction only if it is in the charging phase; finally a factor $\mathbb{1}_{[1,1+\delta]}(V^j)$ is due to the fact that the interaction considers contributions to the charging process only from neurons that have just had their spike ($\delta \in (0,1)$). The choice of the values 0, 1 and 2 is completely arbitrary, and is just used for our mathematical description; we also do not specify explicitly the form of some of the functions involved, since we only need to make assumptions on their regularity.

A possible more accurate model of the inhibition phase could require that also the noise term be switched off during the refractory period, that is, in our setting, as V becomes larger than 1. We are forced to include a small noise also in the inhibition phase, for mathematical reasons (i.e., we need ϵ below to be strictly positive, see in particular Theorems 2.2 and 3.2 and Lemma 5.1). The effect of oscillations due to the noise at the transition between the active phase and the inhibition phase appears to be negligible on macroscopic scales, thanks to the drift (see for example figure 1b). On the other hand, the analysis of a model in which noise contributes only to the charging phase is mathematically extremely interesting, and we will face it in a future work.

Now we will introduce precisely the equations describing the model and will give an account of our results and of the following sections. We also include some figures obtained simulating our model for a finite number of interacting neurons, showing that, even if simple, the model we propose gives a realistic description of single neurons and networks.

1.1. **The model.** For a Borel set A in an euclidean space, we will denote by $\mathscr{L}_A$ the Lebesgue measure restricted to A.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let D be an open connected domain in $\mathbb{R}^3$ and $[0,T] \subset \mathbb{R}$ a time interval. The microscopic model is as follows: for each $N \in \mathbb{N}$ consider N neurons, each identified by

i. its position $X_0^i = \xi^i$, where ξ^i, $i \in \mathbb{N}$ are i.i.d. random variables with finite first moment and such that $\forall i\ \mathbb{P}\left(\xi^i \in D\right) = 1$. We denote by ν the law of each ξ^i. Since



(A) Solutions $V^{i,3}$, $i = 1, 2, 3$ (B) Solutions $V^{i,3} \pmod 2$, $i = 1, 2, 3$

FIGURE 1. Simulation of the system (2) for 3 neurons with no spatial interaction.

the neurons do not move, their position is modeled by the system of trivial equations

$$\mathrm{d}X_t^i = 0\,, \quad X_0^i = \xi^i\,; \tag{1}$$

ii. its action potential given by $V_t^{i,N} \pmod 2 \in [0, 2)$, where $V_t^{i,N} \in \mathbb{R}$ is the strong solution to

$$\begin{aligned}\mathrm{d}V_t^{i,N} &= \lambda\left(V_t^{i,N} \pmod 2\right)\mathrm{d}t \\ &+ \frac{1}{N}\sum_{j=1}^{N}\theta\left(\xi^i,\xi^j\right)\mathbb{1}_{[1,1+\delta]}\left(V_t^{j,N} \pmod 2\right)\mathbb{1}_{[0,1]}\left(V_t^{i,N} \pmod 2\right)\mathrm{d}t \\ &+ \sigma^\epsilon\left(V_t^{i,N} \pmod 2\right)\mathrm{d}B_t^i\,, \end{aligned} \tag{2}$$

with initial condition $V_0^{i,N} = \eta^i \in [0, 2)$. We assume that all random variables η^i are i.i.d with law $\tilde{\rho}_0\mathscr{L}_{[0,2)}$ and $\tilde{\rho}_0 \in L^2(0, 2)$. Moreover, we assume that $\{\xi^1, ..., \xi^N, \eta^1, ...\eta^N\}$ are independent for any $N \in \mathbb{N}$.

The functions appearing above are given by

$$\lambda(v) = -\hat{\lambda}v\mathbb{1}_{[0,1]}(v) + \mathbb{1}_{(1,2)}(v), \quad \text{with } \hat{\lambda} > 0\,;$$

$\theta(x, y)$ is a bounded uniformly continuous function on $D \times D$;

$$\begin{gathered}\sigma^\epsilon(v) \text{ is a } C_b^1([0, 2])\text{-function such that } \sigma^\epsilon \geq \sqrt{2\epsilon} > 0 \\ \text{and } \sigma^\epsilon(v) = \sqrt{2\epsilon} \text{ on } [1, 2],\ \sigma^\epsilon(2) = \sigma^\epsilon(0) = \sqrt{2\epsilon}, \\ \frac{d\sigma^\epsilon}{dx}(0) = \frac{d\sigma^\epsilon}{dx}(2) = 0\,, \text{ with } \epsilon \text{ fixed.}\end{gathered}$$

For each $i \in \mathbb{N}$ the processes $\left(B_t^i\right)$ are independent real-valued Brownian motions, independent of $\left(\xi^i\right)$ and $\left(\eta^i\right)$, and δ is a fixed real number in $(0, 1)$.

One could use as λ any bounded function on $[0, 2)$ that has a jump discontinuity in $v = 1$ and is continuous elsewhere; all the results herein apply in this case with no modifications

in the arguments, therefore we stick to the simple case given just above.
We will show that for each N the system of equations (2) has a unique solution $\left(V^{i,N}\right)_{i=1,\dots,N}$ with $V^{i,N}$ having continuous trajectories in $\mathbb{R}$. This forces the trajectories of $V^{i,N} \pmod 2$ to have jump discontinuities at every t such that $V_t^{i,N} \in 2\mathbb{Z}$. However continuity is easily restored seeing $V^{i,N} \pmod 2$ as a process with values on the torus $\mathbb{T} := \mathbb{R}/2\mathbb{Z}$ (where $2\mathbb{Z}$ is seen as a subgroup of translations on $\mathbb{R}$). This corresponds to considering the interval $[0,2]$ with the identification $0 \equiv 2$. Moreover $\mathbb{T}$ is homeomorphic to $1/\pi S^1 \subset \mathbb{R}^2$, the circle with radius $1/\pi$.

On $\mathbb{T}$ we consider the metric

$$d_{\mathbb{T}}(v_1, v_2) = \min\{(v_1 - v_2) \pmod 2, (v_2 - v_1) \pmod 2\} \tag{3}$$

where on the right-had side v_1 and v_2 are seen as elements in $[0,2) \subset \mathbb{R}$; this corresponds to the shortest-path (or geodesic) metric, which is the arc-length on S^1. This metric induces the quotient topology on $\mathbb{T}$.
We will always consider the Euclidean metric on D and endow $D \times \mathbb{T}$ with the product metric, denoted by $d_{D\times\mathbb{T}}$.
The choice to represent solutions on the torus is natural since the coefficients we introduced above are clearly 2-periodic. To stress periodicity and also to lighten notation for $v \in \mathbb{R}$ and $x, y \in D$ we define the functions

$$\begin{aligned} \lambda_2(v) &:= \lambda\big(v \pmod 2\big)\,, \\ g_2(x,v,y,w) &:= \theta(x,y)\mathbb{1}_{[1,1+\delta]}\big(w \pmod 2\big)\mathbb{1}_{[0,1]}\big(v \pmod 2\big)\,, \\ \sigma_2^\epsilon(v) &:= \sigma^\epsilon\big(v \pmod 2\big) \end{aligned}$$

so that equation (2) takes the more readable form

$$\mathrm{d}V_t^{i,N} = \lambda_2\left(V_t^{i,N}\right)\mathrm{d}t + \frac{1}{N}\sum_{j=1}^N g_2\left(\xi^i, V_t^{i,N}, \xi^j, V_t^{j,N}\right)\mathrm{d}t + \sigma_2^\epsilon\left(V_t^{i,N}\right)\mathrm{d}B_t^i\,. \tag{4}$$

Since $\mathbb{T}$ is homeomorphic to $1/\pi S^1 \subset \mathbb{R}^2$, one can define the Lebesgue measure on $\mathbb{T}$ as the push-forward of the Lebesgue measure on $[0,2)$ through the map $t \mapsto (\cos(\pi t), \sin(\pi t))$; since $\mathbb{T}$ is endowed with the quotient topology, any measure on the Borel sets of $\mathbb{T}$ can be obtained in this way. Therefore we can interpret a Borel measure on $D \times \mathbb{T}$ as a Borel measure on $D \times [0,2)$, and we will do so henceforth. Notice that any Borel measure on $\mathbb{T}$ defines a Borel measure on the whole $\mathbb{R}$ by 2-periodic replication; we will not distinguish between the two in the sequel.
We will show that the solution to (4) has a density which is 2-periodic, thanks to the form of the coefficients; hence this density can be identified with a Borel measure on $D \times \mathbb{T}$.

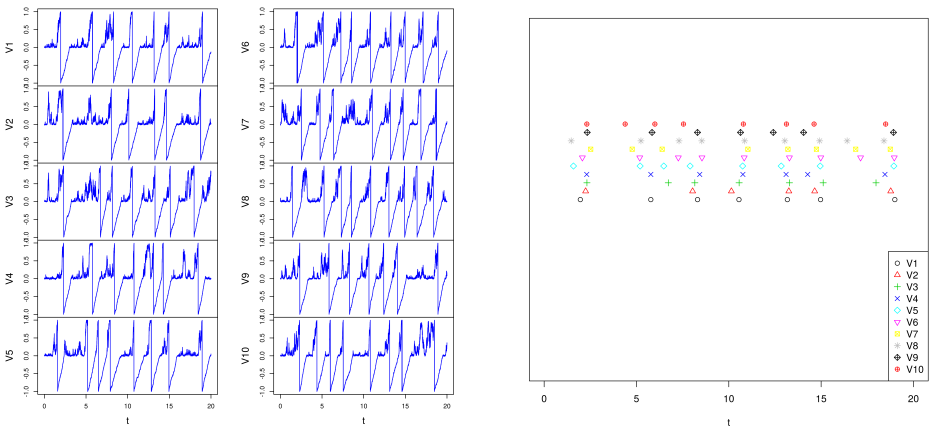
Let S_t^N denote the empirical measure

$$S_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{\left(\xi^i, V_t^{i,N} \pmod 2\right)} = \frac{1}{N}\sum_{i=1}^N \delta_{\left(\xi^i, V_t^{i,N} \pmod 2\right)} \tag{5}$$

and set for every $x, y \in D$ and $v \in \mathbb{R}$

$$\sigma_2^\epsilon(x,v) := \sigma_2^\epsilon(v)\,, \quad \theta(x,y,v) := \theta(x,y) \quad \text{and} \quad \lambda_2(x,v) := \lambda_2(v)$$

for later use (see for instance (8)). To any function ϕ on $D \times \mathbb{T}$ which is continuous corresponds a unique continuous function on $D \times \mathbb{R}$ that is 2-periodic with respect to its second variable, given by $(x,v) \mapsto \phi(x, v \pmod 2)$; in the sequel we will identify these



(A) Solutions $V^{i,10}$ (B) Spikes of each neuron in time

FIGURE 2. Solution of system (2) for 10 neurons with strong interaction uniform in space. A spike from one neuron propagates to all other neurons in the network

two functions and denote by ϕ also its 2-periodic representation on $D \times \mathbb{R}$. With this convention, for any smooth and compactly supported function ϕ on $D \times \mathbb{T}$ we have

$$\begin{aligned}
\mathrm{d}\langle S_t^N, \phi\rangle &= \mathrm{d}\frac{1}{N}\sum_{i=1}^N \phi\left(X_0^i, V_t^{i,N}\right) \\
&= \frac{1}{N}\sum_{i=1}^N \partial_v\phi\left(X_0^i, V_t^{i,N}\right)\left(\lambda_2\left(V_t^{i,N}\right) + \frac{1}{N}\sum_{j=1}^N g_2\left(X_0^i, V_t^{i,N}, X_0^j, V_t^{j,N}\right)\right)\mathrm{d}t \\
&+ \frac{1}{N}\sum_{i=1}^N \sigma_2^\epsilon\left(V_t^{i,N}\right)\partial_v\phi\left(X_0^i, V_t^{i,N}\right)\mathrm{d}B_t^i + \frac{1}{N}\sum_{i=1}^N \frac{\sigma_2^\epsilon\left(V_t^{i,N}\right)^2}{2}\partial_v^2\phi\left(X_0^i, V_t^{i,N}\right)\mathrm{d}t \\
&= \langle S_t^N, \lambda_2\partial_v\phi\rangle\,\mathrm{d}t + \langle S_t^N, \langle S_t^N, g_2(x,v,\cdot,\cdot)\rangle\partial_v\phi\rangle\,\mathrm{d}t + \langle S_t^N, \frac{(\sigma_2^\epsilon)^2}{2}\partial_v^2\phi\rangle\,\mathrm{d}t + \mathrm{d}M_t^{N,\phi}
\end{aligned} \tag{6}$$

where we use the notation

$$\begin{aligned}
&\langle S_t^N, \langle S_t^N, g_2(x,v,\cdot,\cdot)\rangle\partial_v\phi\rangle \\
&\qquad = \int_{D\times\mathbb{T}} \partial_v\phi(x,v) S_t^N(\mathrm{d}x,\,\mathrm{d}v)\int_{D\times\mathbb{T}} g_2(x,v,y,w) S_t^N(\mathrm{d}y,\,\mathrm{d}w)
\end{aligned} \tag{7}$$

and

$$M_t^{N,\phi} = \int_0^t \frac{1}{N}\sum_{i=1}^N \sigma_2^\epsilon\left(V_s^{i,N}\right)\partial_v\phi\left(X^i, V_s^{i,N}\right)\mathrm{d}B_s^i$$

is a martingale such that $\mathbb{E}\big[\sup_t \left|M_t^{N,\phi}\right|^2\big] \to 0$ as $N \to \infty$ (due to the stochastic integrals being uncorrelated).

If we suppose that the sequence of random measures $S_t^N(\mathrm{d}x,\,\mathrm{d}v)$ converges in probability

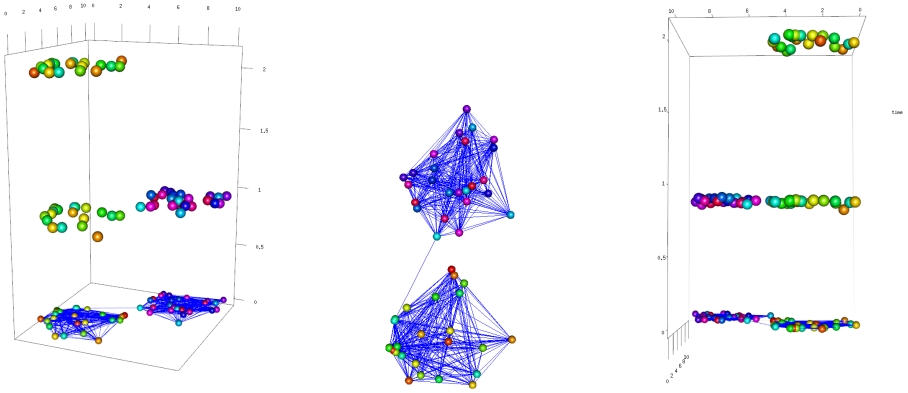


FIGURE 3. A network made of two subnetworks with localized strong interaction. Each blue line corresponds to the presence of interaction; only one neuron of the first subnetwork interacts with a single neuron of the second one. The vertical axis represents time; the network is drawn on the $t = 0$ plane and to each spike of a neuron corresponds a sphere above it, with the same color as the neuron. Near time $t = 1$ many neurons in the first network spike and the signal is propagated to the second network, while at time $t = 2$ the signal does not propagate to the second network.

(in a suitable space) to a probability measure $\rho_t(x, v)\,\mathrm{d}x\,\mathrm{d}v$ on $D \times \mathbb{T}$, then, heuristically, a passage to the limit in N suggests that ρ_t solves weakly the partial differential equation of Fokker-Planck type

$$\frac{\partial}{\partial t}\rho_t(x,v) = \frac{1}{2}\frac{\partial^2}{\partial v^2}\left(\left(\sigma_2^\epsilon\right)^2 \rho_t\right)(x,v) - \frac{\partial}{\partial v}\left(\lambda_2 \rho_t\right)(x,v) - \frac{\partial}{\partial v}\left(\rho_t \int g_2(\cdot,\cdot,y,w)\rho_t(y,w)\,\mathrm{d}y\,\mathrm{d}w\right)(x,v)\,. \quad (8)$$

In the sequel we will prove rigorously a similar assertion involving measures μ_t instead of densities ρ_t.

1.2. **Main results.** The main aim of the paper is to show that the empirical measure actually converges in a weak sense to a limiting probability measure $\mu_t(\,\mathrm{d}x,\,\mathrm{d}v)$ such that the marginal with respect to v has a L^2-density and which is the unique solution to the above PDE (8). This is the content of Theorem 5.6.

The paper is organized as follows. In Section 2 we will prove strong well-posedness for the system of SDEs (2); a modification of the standard theory for finite-dimensional SDEs with bounded and measurable drift is needed here to deal with the dependence of the equations on the random variables X_0^i. In Section 3 we define a weak measure-valued solution and show that the PDE (8) has at most one such solution.

To show existence of a solution to the PDE we first prove that the laws $\mathbb{Q}^N$ of the empirical measures of $(X^i, V^{i,N} \pmod 2)$ (see (5)) are tight as probability measures on the space of continuous measure-valued functions of time (Section 4). Then we prove that any limit point $\mathbb{Q}$ of $\mathbb{Q}^N$ gives full measure to the set of functions with values that are continuous measures with marginal with respect to v having a L^2-density, that $\mathbb{Q}$ is supported by the set of weak measure-valued solutions to the Fokker-Planck PDE and that actually all the

sequence $\mathbb{Q}^N$ converges to the same limit. This provides existence of a solution and is discussed in Section 5.
Section 6 briefly shows that well-posedness of the Fokker-Planck PDE implies existence of a unique strong solution to the McKean-Vlasov SDE associated with the particle system. We conclude with an appendix giving some immediate generalizations of our results together with some indications on how the proofs have to be adapted to this more general setting.

2. **The system of particles.** Consider independent Brownian motions B^i, $i \in \mathbb{N}$ and assume that the random variables η^i introduced in the previous section are i.i.d. and independent from all B^i; denote by $(\mathcal{F}_t^0)_t$ the filtration $\mathcal{F}_t^0 = \sigma\left(B_s^i, 0 \le s \le t, i \in \mathbb{N}\right) \vee \sigma(\eta^i, i \in \mathbb{N})$ augmented with the $\mathbb{P}$-null sets.
We also denote by $\mathcal{G}$ the σ-algebra $\sigma\left(\xi^i, i \in \mathbb{N}\right)$ and assume that for any $t \ge 0$, $\mathcal{G}$ and $\mathcal{F}_t^0$ are independent.
Finally we introduce the filtration $(\mathcal{F}_t)$ where $\mathcal{F}_t$ is the completion of $\mathcal{F}_t^0 \vee \mathcal{G}$.

Let us write our system of equations in vector form: we fix $N \in \mathbb{N}$ and introduce, for the variables $\mathbf{v} = (v^1, \dots, v^N)^\top \in \mathbb{R}^N$, $\mathbf{x} = (x^1, \dots, x^N)^\top \in D^N$, the functions

$$\bar{\lambda}_2 : \mathbb{R}^N \to \mathbb{R}^N$$
$$\bar{\lambda}_2(\mathbf{v}) = \left(\lambda_2(v^1), \dots, \lambda_2(v^N)\right)^\top ,$$

$$\bar{g}_2 : (D \times \mathbb{R})^{N+1} \to \mathbb{R}^N$$
$$\bar{g}_2(\mathbf{x}, \mathbf{v}, y, w) = \left(g_2(x^1, v^1, y, w), \dots, g_2(x^N, v^N, y, w)\right)^\top ,$$

$$\bar{\sigma}_2^\epsilon : \mathbb{R}^N \to \mathbb{R}^{N \times N}$$
$$\bar{\sigma}_2^\epsilon(\mathbf{v}) = \operatorname{diag}\left(\sigma_2^\epsilon(v^1), \dots, \sigma_2^\epsilon(v^N)\right).$$

Setting $\Xi = \left(\xi^1, \dots, \xi^N\right)$, $\Psi = \left(\eta^1, \dots, \eta^N\right)$, we want to show existence and uniqueness of strong solutions to

$$\begin{cases} \mathrm{d}\mathbf{V}_t = \bar{\lambda}_2\left(\mathbf{V}_t\right) \mathrm{d}t + \frac{1}{N} \sum_{j=1}^N \bar{g}_2\left(\Xi, \mathbf{V}_t, \Xi^j, V_t^j\right) \mathrm{d}t + \bar{\sigma}_2^\epsilon\left(\mathbf{V}_t\right) \mathrm{d}\mathbf{B}_t \\ \mathbf{V}_0 = \Psi \, . \end{cases} \tag{9}$$

The classical reference for existence of a strong solution for SDEs with bounded measurable drift is [20]. However, the results proved therein do not apply directly to equation (9), because they do not guarantee the measurable dependence of the solution V on the stochastic parameter Ξ. Therefore we introduce the following definition.

Let us denote by $\mathcal{S}$ the Banach space of all continuous paths from $[0, T]$ into $\mathbb{R}^N$ endowed with the supremum norm $|\cdot|_{\mathcal{S}}$. We also denote by ν_N the law of the random variable Ξ on the Borel σ-algebra of D^N.

Definition 2.1 (Strong solution)**.** A *strong solution* to (9) is a family of continuous $\mathbb{R}^N$-valued $\mathcal{F}_t^0$-adapted processes $(\mathbf{V}_t^{\mathbf{x}})$, $\mathbf{x} \in D^N$, such that

(i) the mapping $(\mathbf{x}, \omega) \mapsto \mathbf{V}_\cdot^{\mathbf{x}}(\omega) \in \mathcal{S}$ is measurable from $(D^N \times \Omega, \mathcal{B}(D^N) \times \mathcal{F})$ to $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$;
(ii) for ν_N-almost every $\mathbf{x} \in D^N$, $(\mathbf{V}_t^{\mathbf{x}})$ satisfies equation (9) in the strong sense when $\Xi = \mathbf{x}$.

The above definition is motivated by the fact that if $(\mathbf{V}_t^{\mathbf{x}})$ is a strong solution, then the $\mathcal{F}_t$-adapted process $(\mathbf{V}_t^{\boldsymbol{\Xi}})$ satisfies equation (9) $\mathbb{P}$-almost surely for any $t \in [0,T]$. In fact, for Ξ random variable as above, the process $(\mathbf{V}_t^{\Xi})$ is well-defined with values in $\mathbb{R}^N$, has continuous paths and is $\mathcal{F}_t$-adapted.
To prove that $\left(\mathbf{V}_t^{\Xi}\right)$ satisfies (9) it is enough to compute, using conditional expectation with respect to $\mathcal{G}$ and independence,

$$\begin{aligned}
&\mathbb{E}\left[\left|\Psi + \int_0^t \bar{\lambda}_2\left(\mathbf{V}_r^{\Xi}\right)\,\mathrm{d}r + \frac{1}{N}\sum_{j=1}^N \int_0^t \bar{g}_2\left(\Xi, \mathbf{V}_r^{\Xi}, \Xi^j, V_r^{\Xi,j}\right)\,\mathrm{d}r \right.\right.\\
&\qquad\qquad \left.\left. + \int_0^t \bar{\sigma}_2^{\epsilon}\left(\mathbf{V}_r^{\Xi}\right)\,\mathrm{d}\mathbf{B}_r - \mathbf{V}_t^{\Xi}\right|\right]\\
&= \mathbb{E}\mathbb{E}\left[\left|\Psi + \int_0^t \bar{\lambda}_2\left(\mathbf{V}_r^{\Xi}\right)\,\mathrm{d}r + \frac{1}{N}\sum_{j=1}^N \int_0^t \bar{g}_2\left(\Xi, \mathbf{V}_r^{\Xi}, \Xi^j, V_r^{\Xi,j}\right)\,\mathrm{d}r \right.\right.\\
&\qquad\qquad \left.\left. + \int_0^t \bar{\sigma}_2^{\epsilon}\left(\mathbf{V}_r^{\Xi}\right)\,\mathrm{d}\mathbf{B}_r - \mathbf{V}_t^{\Xi}\right| \,\middle|\, \mathcal{G}\right]\\
&= \mathbb{E}\mathbb{E}\left[\left|\Psi + \int_0^t \bar{\lambda}_2\left(\mathbf{V}_r^{\mathbf{x}}\right)\,\mathrm{d}r + \frac{1}{N}\sum_{j=1}^N \int_0^t \bar{g}_2\left(\mathbf{x}, \mathbf{V}_r^{\mathbf{x}}, x^j, V_r^{\mathbf{x},j}\right)\,\mathrm{d}r \right.\right.\\
&\qquad\qquad \left.\left. + \int_0^t \bar{\sigma}_2^{\epsilon}\left(\mathbf{V}_r^{\mathbf{x}}\right)\,\mathrm{d}\mathbf{B}_r - \mathbf{V}_t^{\mathbf{x}}\right|_{\mathbf{x}=\Xi}\right]\\
&= 0.
\end{aligned}$$

Theorem 2.2. *For every $N \in \mathbb{N}$ there exists a strong solution to (4). Two strong solutions on the same probability space associated to the same initial condition Ψ are indistinguishable for ν_N-almost every $\mathbf{x} \in D^N$.*

Proof. Existence. First fix $\mathbf{x} \in D^N$; the SDE

$$\begin{cases} \mathrm{d}\mathbf{V}_t^{\mathbf{x}} = \bar{\lambda}_2\left(\mathbf{V}_t^{\mathbf{x}}\right)\,\mathrm{d}t + \frac{1}{N}\sum_{j=1}^N \bar{g}_2\left(\mathbf{x}, \mathbf{V}_t^{\mathbf{x}}, x^j, V_t^{\mathbf{x},j}\right)\,\mathrm{d}t + \bar{\sigma}_2^{\epsilon}\left(\mathbf{V}_t^{\mathbf{x}}\right)\,\mathrm{d}\mathbf{B}_t \\ \mathbf{V}_0^{\mathbf{x}} = \Psi\ . \end{cases} \tag{10}$$

admits a unique strong solution $\mathbf{V}^{\mathbf{x}}$ by the results proved in [20]. We now clarify the measurability of $\mathbf{V}^{\mathbf{x}}$ with respect to $\mathbf{x}$. One difficulty is that the proof of the main well-posedness results in [20] is based on the Yamada-Watanabe theorem and is indeed abstract and non-contructive. This is why to prove such a measurability property we follow the approach in [8]. Fix a sequence of partitions $\{\pi^n\}_{n\in\mathbb{N}}$ of $[0,T]$, where each π^n is given by points $0 = t_0^n < t_1^n < \cdots < t_n^n = T$, set $\kappa^n(t) = \sum_{i=0}^{n-1} t_i^n \mathbb{1}_{[t_i^n, t_{i+1}^n)}(t)$ and consider Euler's approximations to equation (10) given by

$$\begin{cases} \mathrm{d}\mathbf{V}_t^{\mathbf{x},n} = \bar{\lambda}_2\left(\mathbf{V}^{\mathbf{x},n}(\kappa^n(t))\right)\,\mathrm{d}t + \frac{1}{N}\sum_{j=1}^N \bar{g}_2\left(\mathbf{x}, \mathbf{V}^{\mathbf{x},n}(\kappa^n(t)), x^j, V^{\mathbf{x},n,j}(\kappa^n(t))\right)\,\mathrm{d}t \\ \qquad\qquad + \bar{\sigma}_2^{\epsilon}\left(\mathbf{V}^{\mathbf{x},n}(\kappa^n(t))\right)\,\mathrm{d}\mathbf{B}_t, \\ \mathbf{V}_0^{\mathbf{x},n} = \Psi\ . \end{cases} \tag{11}$$

In the next steps, we will use that, for any $n \geq 1$, the processes $\mathbf{V}^{\mathbf{x},\mathbf{n}}$ enjoy all the measurability properties we need.

By Theorem 2.8 in [8] we know that, for any $\mathbf{x} \in D^N$ and $\delta > 0$

$$\mathbb{P}\left(|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}} > \delta\right) \to 0, \quad \text{as } n \to \infty. \tag{12}$$

Note that, for any $p \geq 2$ (using the boundedness of the coefficients of the SDEs) there exists $C_p > 0$ (independent of n and $\mathbf{x} \in \mathbf{D}$) such that, for any $n \geq 1$, $\mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^p\right] \leq C_p$. Writing

$$\begin{aligned}
&\mathbb{E}\,|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^2 \\
&\quad = \mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^2 \, \mathbb{1}_{\{|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}} > \delta\}}\right] + \mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^2 \, \mathbb{1}_{\{|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}} \leq \delta\}}\right]
\end{aligned}$$

and using also the Hölder inequality, we easily deduce that, for any $\mathbf{x} \in D^N$,

$$\mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^2\right] \to 0, \quad \text{as } n \to \infty\,. \tag{13}$$

Hence, for any $\mathbf{x} \in D^N$, $\mathbf{V}^{\mathbf{x},n}$ converges to $\mathbf{V}^{\mathbf{x}}$ in the Banach space $L^2(\Omega;\mathcal{S})$ and so the mapping $\mathbf{x} \mapsto \mathbf{V}^{\mathbf{x}} \in L^2(\Omega;\mathcal{S})$ is Borel measurable on $\left(D^N, \mathcal{B}(D^N)\right)$. By the dominated convergence theorem we infer

$$\int_{D^N} \mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x}}|_{\mathcal{S}}^2\right] \nu_N(\,\mathrm{d}x) \to 0, \quad \text{as } n \to \infty\,, \tag{14}$$

i.e., $(\mathbf{V}^{\cdot,n})$ converges to $\mathbf{V}^{\cdot}$ in $\mathcal{Z} = L^2\left((D^N, \nu_N); L^2(\Omega;\mathcal{S})\right)$. It follows that $(\mathbf{V}^{\mathbf{x},n})$ is a Cauchy sequence in $\mathcal{Z}$. Using that, for any $n, m \geq 1$,

$$\int_{D^N} \mathbb{E}\left[|\mathbf{V}^{\mathbf{x},n} - \mathbf{V}^{\mathbf{x},m}|_{\mathcal{S}}^2\right] \nu_N(\,\mathrm{d}x) = \int_{D^N \times \Omega} |\mathbf{V}^{\mathbf{x},n}(\omega) - \mathbf{V}^{\mathbf{x},m}(\omega)|_{\mathcal{S}}^2 \, \nu_N(\,\mathrm{d}x)\mathbb{P}(\,\mathrm{d}\omega)\,, \tag{15}$$

we get that $(\mathbf{V}^{\cdot,n})$ converges to some $\tilde{\mathbf{V}}^{\cdot}$ in $L^2\left((D^N \times \Omega, \nu_N \times \mathbb{P}); \mathcal{S}\right)$. In particular, $\tilde{\mathbf{V}}^{\cdot}$ is measurable on $(D^N \times \Omega, \mathcal{B}(D^N) \times \mathcal{F})$ with values in $\mathcal{S}$.
It follows that, for a.e. $\mathbf{x} \in D^N$, we have $\tilde{\mathbf{V}}^{\mathbf{x}} = \mathbf{V}^{\mathbf{x}}$ in $\mathcal{S}$, $\mathbb{P}$-a.s. (we have obtained a version of the strong solution which has the required measurability properties with respect to $\mathbf{x}$).
Uniqueness. It follows directly from the celebrated Veretennikov result. □

Remark 1. To show existence and uniqueness of a solution one could weaken the assumptions on the regularity of σ^ϵ, similarly to what is done in the references [20] and [8]. What is needed above is that pathwise uniqueness holds for equation (10), and there are many well-known conditions assuring that this happens. However at a later stage in the paper (Section 5) we will need to assume that σ^ϵ is differentiable with bounded derivative. This does not seem to be a limitation on the model, since there is no reason to assume that the diffusion coefficient be particularly rough.

3. **The limit PDE: uniqueness of measure-valued solutions.** Let $\mathrm{Pr}(D \times \mathbb{T})$ be the space of Borel probability measures over $D \times \mathbb{T}$. Let $\mathrm{Pr}_1(D \times \mathbb{T}) \subset \mathrm{Pr}(D \times \mathbb{T})$ be the space of probability measures over $D \times \mathbb{T}$ with finite first moment, endowed with the 1-Wasserstein metric $\mathcal{W}_1$.
Suppose as above that the empirical measures S_t^N converge in a weak sense to a probability measure μ_t on $D \times \mathbb{T}$. Without assuming that μ_t has a density, we expect that it solve the PDE

$$\frac{\partial}{\partial t}\mu_t = \frac{1}{2}\frac{\partial^2}{\partial v^2}\left(\left(\sigma_2^\epsilon\right)^2 \mu_t\right) - \frac{\partial}{\partial v}\left(\lambda_2 \mu_t\right) - \frac{\partial}{\partial v}\left(\mu_t \int g_2(\cdot,\cdot,y,w)\mu_t(\,\mathrm{d}y,\,\mathrm{d}w)\right)\,, \tag{16}$$

with initial condition $\nu \times \tilde{\rho}_0 \mathscr{L}_{\mathbb{T}}$, meaning that $\mu_0(\,\mathrm{d}x,\,\mathrm{d}v) = \nu(\,\mathrm{d}x)\tilde{\rho}_0(v)\mathscr{L}_{\mathbb{T}}\,(\,\mathrm{d}v)$. Fix $T > 0$; we will denote by $\mathscr{C}$ the space

$$\mathscr{C} := C\left([0,T]; \mathrm{Pr}_1(D \times \mathbb{T})\right)$$

and for a measure $\zeta \in \mathrm{Pr}(D \times \mathbb{T})$ we will adopt the notation

$$b(\zeta)(x,v) := \lambda_2(v) + \int_{D\times\mathbb{T}} g_2(x,v,y,w)\zeta(\,\mathrm{d}y,\,\mathrm{d}w), \quad x \in D,\, v \in \mathbb{T}\,, \tag{17}$$

throughout the rest of the paper.
In the sequel we will often use that if $g : \mathbb{R} \to \mathbb{R}$ is 2-periodic and differentiable on $\mathbb{R}$ then its derivative is also 2-periodic (thus g can be identified with a differentiable function on $\mathbb{T}$). Recall the Banach space $B_b(D \times \mathbb{T})$ consisting of all Borel and bounded functions $f : D \times \mathbb{T} \to \mathbb{R}$ endowed with the supremum norm $\|\cdot\|_\infty$. We will also consider $C_b(D \times \mathbb{T}) \subset B_b(D \times \mathbb{T})$ consisting of all bounded continuous functions. We introduce the space of test functions

$$\mathcal{T} \subset C_b(D \times \mathbb{T})$$

defined as the space of all $\phi \in C_b(D \times \mathbb{T})$ such that there exist the partial derivatives $\partial_v\phi$ and $\partial_v^2\phi$ which both belong to $C_b(D \times \mathbb{T})$.

Definition 3.1. We say that $\mu \in \mathscr{C}$ is a weak measure-valued solution of the nonlinear Fokker-Planck equation (16), with initial condition $\mu^0 \in \mathrm{Pr}_1\,(D \times \mathbb{T})$, if

$$\langle \mu_t, \phi\rangle = \langle \mu^0, \phi\rangle + \int_0^t \langle \mu_s, b(\mu_s)\partial_v\phi\rangle\,\mathrm{d}s + \int_0^t \left\langle \mu_s, \frac{(\sigma_2^\epsilon)^2}{2}\partial_v^2\phi \right\rangle \mathrm{d}s, \tag{18}$$

for every test function $\phi \in \mathcal{T}$.

Consider now the total variation distance on Borel probability measures on $D \times \mathbb{T}$

$$\begin{aligned} d_{\mathrm{TV}}(\nu^1,\nu^2) :&= \sup\left\{\left|\langle\nu^1,\phi\rangle - \langle\nu^2,\phi\rangle\right| : \phi \in B_b(D\times\mathbb{T}), \|\phi\|_\infty \le 1\right\} \\ &= \sup\left\{\left|\langle\nu^1,\phi\rangle - \langle\nu^2,\phi\rangle\right| : \phi \in C_b(D\times\mathbb{T}), \|\phi\|_\infty \le 1\right\}\,. \end{aligned}$$

Remark 2. Let $\mu^1, \mu^2 \in \mathscr{C}$. One can show that the mapping

$$t \mapsto d_{\mathrm{TV}}(\mu_t^1, \mu_t^2) \quad \text{is Borel and bounded on } [0,T].$$

To this purpose we first remark that, for given probability measures ν^1 and ν^2 on Borel sets of $D \times \mathbb{T}$, one has

$$d_{\mathrm{TV}}(\nu^1,\nu^2) = \sup\left\{\left|\langle\nu^1,\phi\rangle - \langle\nu^2,\phi\rangle\right| : \phi \in C_c^\infty(D\times\mathbb{T}), \|\phi\|_\infty \le 1\right\}\,. \tag{19}$$

(recall that $\phi \in C_c^\infty(D \times \mathbb{T})$ if $\phi \in C^\infty(D \times \mathbb{T})$ and has compact support). As before if $f : D \times \mathbb{T} \to \mathbb{R}$ we identify the function $(x,v) \mapsto f(x, v \pmod 2) \in \mathbb{R}$ defined on $D \times \mathbb{R}$ with f.

To prove (19) let $f \in C_b(D \times \mathbb{T})$ with $\|f\|_\infty \le 1$, by truncating f and by considering standard mollifiers (defined on $\mathbb{R}^4$) we can find a sequence $(f_n) \subset C_c^\infty(D \times \mathbb{T})$ such that $\|f_n\|_\infty \le 1$ and $f_n(z) \to f(z)$, as $n \to \infty$, for any $z = (x,v) \in D \times \mathbb{T}$. By using

$$\left|\langle\nu^1,f_n\rangle - \langle\nu^2,f_n\rangle\right| \le \sup\left\{\left|\langle\nu^1,\phi\rangle - \langle\nu^2,\phi\rangle\right| : \phi \in C_c^\infty, \|\phi\|_\infty \le 1\right\}, \quad n \ge 1,$$

and the Lebesgue convergence theorem we obtain that the previous formula holds even when f_n is replaced by f; this leads to (19).

Then we show that there exists a countable set $K_\infty \subset C_c^\infty$ such that for any $f \in C_c^\infty$, we can find a sequence $(f_k) \subset K_\infty$ satisfying

$$\lim_{k\to\infty} \|f - f_k\|_\infty = 0. \tag{20}$$

As a simple consequence we get that

$$d_{\mathrm{TV}}(\nu^1,\nu^2)=\sup\left\{\left|\langle\nu^1,\phi\rangle-\langle\nu^2,\phi\rangle\right|:\phi\in K_\infty,\|\phi\|_\infty\le 1\right\}\ . \tag{21}$$

To prove assertion (20) set $F_n=\{f\in C_c^\infty$ with support$(f)\subset B_n\}$ where $B_n=\{(x,v)\in D\times\mathbb{T}\ :\ \mathrm{dist}(x,\partial D)\ge 1/n\ \ \text{and}\ |x|\le n\}$. Each F_n is separable: indeed $F_n\subset C(B_n)$ and so there exists a countable set $K_n\subset F_n$ which is dense in F_n. To finish we define $K_\infty=\cup_{n\ge 1}K_n$. □

Theorem 3.2. *Let $\mu^0\in\mathrm{Pr}_1(D\times\mathbb{T})$. There exists at most a unique weak measure-valued solution to equation (16), with initial condition μ^0, in $\mathscr{C}$.*

Proof. Given a function $f:\mathbb{T}\to\mathbb{R}$ we still denote by f its 2-periodic version defined on $\mathbb{R}$. Let $\psi\in\mathcal{T}$ and define the operator

$$A\psi(x,v)=A\psi(x,\cdot)(v)=\frac{\sigma_2^\epsilon(v)}{2}\frac{\partial^2\psi}{\partial v^2}(x,v)\ .$$

It is well known that A is the infinitesimal generator of a diffusion semigroup $T_t:B_b(D\times\mathbb{R})\to B_b(D\times\mathbb{R})$:

$$T_t\zeta(x,v)=\int_{\mathbb{R}}\zeta(x,v')p_t(v,v')dv',\ \ x\in D,\ v\in\mathbb{R},\ \zeta\in B_b(D\times\mathbb{R}),\ t\ge 0, \tag{22}$$

where the density $p_t(\cdot,v')\in C^2(\mathbb{R})$, for any $v'\in\mathbb{R}$, $t>0$ (see, for instance, Chapter 6 in [6]). Moreover, for $t>0$, $p_t(v,v')$, $\partial_v p_t(v,v')$ and $\partial_v^2 p_t(v,v')$ are continuous functions on $\mathbb{R}^2$. In addition, for any $g\in C_b(D\times\mathbb{R})$, $t\in(0,T)$, we have

$$\|\partial_v T_t g\|_\infty\le\frac{c}{\sqrt{t}}\|g\|_\infty,\quad \|\partial_v^2 T_t g\|_\infty\le\frac{c}{t}\|g\|_\infty. \tag{23}$$

Finally, if $f\in C_b^2(D\times\mathbb{R})$, we have $T_tf\in C_b^2(D\times\mathbb{R})$, $t\ge 0$, and $\partial_t T_tf(x,v)=T_tAf(x,v)=AT_tf(x,v)$, $t\ge 0$, $x\in D$, $v\in\mathbb{R}$.

Since in our case σ_2 is also 2-periodic, it is not difficult to prove that, for $t>0$, p_t is 2-periodic in both variables, i.e., $p_t(v+2,v'+2)=p_t(v,v')$, $v,v'\in\mathbb{R}$. It follows that if $\psi\in B_b(D\times\mathbb{R})$ is 2-periodic in the v-variable then also $T_t\psi$ is 2-periodic in the v-variable. Differentiating, we obtain that $\partial_v p_t(v,v')$ and $\partial_v^2 p_t(v,v')$ are 2-periodic continuous functions in both variables. Hence, in particular, $T_t\psi\in\mathcal{T}$ if $\psi\in\mathcal{T}$, $t\ge 0$. One can prove that $\mu\in\mathscr{C}$ is a weak solution to (16) if and only if it is a mild solution, i.e., if and only if

$$\langle\mu_t,\phi\rangle=\langle\mu^0,T_t\phi\rangle+\int_0^t\langle\mu_s,b(\mu_s)\frac{\partial}{\partial v}T_{t-s}\phi\rangle\,\mathrm{d}s,\quad \phi\in\mathcal{T},t\in[0,T]. \tag{24}$$

We only show that any weak solution is a mild solution (this is the part we need to prove our uniqueness claim). We fix ϕ and $t>0$. Differentiating with respect to $s\in(0,t)$ the mapping

$$s\mapsto\langle T_{t-s}\phi,\mu_s\rangle=\int_{D\times\mathbb{R}}\mu_s(dx,dv)\int_{\mathbb{R}}\phi(x,v')p_{t-s}(v,v')dv'$$

we get

$$\frac{d}{ds}(\langle T_{t-s}\phi,\mu_s\rangle)=-\langle T_{t-s}A\phi,\mu_s\rangle)+\langle\mu_s,b(\mu_s)\partial_v T_{t-s}\phi\rangle+\langle\mu_s,T_{t-s}A\phi\rangle\ .$$

Integrating with respect to s on $[0,t]$ we find the assertion.
Now we prove the claim of the theorem. Let $\mu^1,\mu^2\in\mathscr{C}$ be two solutions to (16) with the

same initial condition μ^0; then for every t (in the sequel we can consider the supremum over $\phi \in K_\infty \subset C_c^\infty$ such that $\|\phi\|_\infty \le 1$ as in the previous remark)

$$\begin{aligned} d_{\mathrm{TV}}(\mu_t^1, \mu_t^2) &= \sup_{\|\phi\|_\infty \le 1} \left| \int_0^t \left[\langle \mu_s^1, b(\mu_s^1) \partial_v T_{t-s}\phi \rangle - \langle \mu_s^2, b(\mu_s^2) \partial_v T_{t-s}\phi \rangle \right] \mathrm{d}s \right| \\ &\le \sup_{\|\phi\|_\infty \le 1} \left| \int_0^t \langle \mu_s^1, \left[b(\mu_s^1) - b(\mu_s^2) \right] \partial_v T_{t-s}\phi \rangle \, \mathrm{d}s \right| \qquad (25) \\ &\quad + \sup_{\|\phi\|_\infty \le 1} \left| \int_0^t \langle \mu_s^1 - \mu_s^2, b(\mu_s^2) \partial_v T_{t-s}\phi \rangle \, \mathrm{d}s \right| . \qquad (26) \end{aligned}$$

The function $b(\mu_s^2)\partial_v T_{t-s}\phi$ is bounded and measurable and we have the estimate (cf. (23))

$$\|\partial_v T_{t-s}\phi\|_\infty \le \frac{C}{\sqrt{t-s}} \|\phi\|_\infty ;$$

we can thus bound the term (26) by

$$\sup_{\|\phi\|_\infty \le 1} \int_0^t d_{\mathrm{TV}}(\mu_s^1, \mu_s^2) \left\| b(\mu_s^2) \partial_v T_{t-s}\phi \right\|_\infty \mathrm{d}s \le (\|\theta\|_\infty + \|\lambda\|_\infty) \int_0^t \frac{C}{\sqrt{t-s}} d_{\mathrm{TV}}(\mu_s^1, \mu_s^2) \, \mathrm{d}s .$$

Similarly, (25) is bounded by

$$\sup_{\|\phi\|_\infty \le 1} \int_0^t \left\| \left(b(\mu_s^1) - b(\mu_s^2) \right) \partial_v T_{t-s}\phi \right\|_\infty \mathrm{d}s \le \|\theta\|_\infty \int_0^t \frac{C}{\sqrt{t-s}} d_{\mathrm{TV}}(\mu_s^1, \mu_s^2) \, \mathrm{d}s .$$

An application of a generalized version of Gronwall's lemma (see, for instance, [10, Section 1.2.1]) yields that $\mu_t^1 = \mu_t^2$ for every t. □

4. **The laws of the empirical measures.** We denote by $\mathbb{Q}^N$ the law of S^N on $\mathscr{C}$ (we are considering each S^N as a r.v. with values in $\mathscr{C}$). As explained in the introduction, we need to show tightness of the family $\mathbb{Q}^N$.

Theorem 4.1. *The sequence* $\{\mathbb{Q}^N\}_{N \in \mathbb{N}}$ *is tight in* $\mathscr{C}$.

Proof. Fix any $(x_0, v_0) \in D \times \mathbb{T}$ and consider the set

$$\mathcal{K}_{M,R} = \Big\{ \mu \in \mathscr{C} : \sup_{t \in [0,T]} \int_{D \times \mathbb{T}} d_{D \times \mathbb{T}} \big((x_0, v_0), (x, v) \big) \mu_t \left(\mathrm{d}x, \mathrm{d}v \right) \le M, \int_0^T \int_0^T \frac{\mathcal{W}_1 \left(\mu_t, \mu_s \right)^p}{|t-s|^{1+\alpha p}} \, \mathrm{d}t \, \mathrm{d}s \le R \Big\}$$

where we choose $\alpha \in (0,1)$ and $p \ge 1$ such that $\alpha p > 1$.
We show that $\mathcal{K}_{M,R}$ is relatively compact in $\mathscr{C}$. Let $B_{(x_0,v_0)}(r)$ denote the open ball with radius r and center (x_0, v_0) in $D \times \mathbb{T}$. Then for $\mu \in \mathcal{K}_{M,R}$ and $t \in [0,T]$

$$\mu_t \left(B_{(x_0,v_0)}(r)^{\complement} \right) \le \frac{1}{r} \int_{D \times \mathbb{T}} d_{D \times \mathbb{T}} \big((x_0, v_0), (x, v) \big) \mu_t(\mathrm{d}x, \mathrm{d}v) \le \frac{M}{r} .$$

Therefore for every $e > 0$ and for every $t \in [0,T]$ we can find $r = r(e,t)$ such that

$$\mu_t \left(B_{(x_0,v_0)}(r) \right) > 1 - e \qquad (27)$$

for every $\mu \in \mathcal{K}_{M,R}$.
By the Sobolev embedding theorem, if $\beta < (\alpha p - 1)/p$, we have that, for any Lipschitz continuous function ϕ on $D \times \mathbb{T}$ and any $t, s \in [0,T]$,

$$|\langle \mu_t, \phi\rangle - \langle \mu_s, \phi\rangle| \le C|t-s|^\beta \left(\int_0^T \int_0^T \frac{|\langle \mu_t, \phi\rangle - \langle \mu_s, \phi\rangle|^p}{|t-s|^{1+\alpha p}} \,\mathrm{d}t \,\mathrm{d}s \right)^{\frac{1}{p}}$$

so that, thanks to Kantorovich-Rubinstein characterization of the 1-Wasserstein distance, we can take the supremum over Lipschitz functions on $D \times \mathbb{T}$ with Lipschitz seminorm bounded by 1 on both sides of the previous inequality, obtaining

$$\sup_{t \neq s} \frac{\mathcal{W}_1(\mu_t, \mu_s)}{|t-s|^\beta} \le C \left(\int_0^T \int_0^T \frac{\mathcal{W}_1(\mu_t, \mu_s)^p}{|t-s|^{1+\alpha p}} \,\mathrm{d}t \,\mathrm{d}s \right)^{\frac{1}{p}} .$$

Therefore the collection of measures $\mathcal{K}_{M,R}$ is equicontinuous; this together with (27) implies relative compactness by the Ascoli-Arzelà theorem.

To show tightness we now compute

$$\begin{aligned}
\mathbb{Q}^N \left(\mathcal{K}_{M,R}^{\complement} \right) &= \mathbb{P}\left(S^N \in \mathcal{K}_{M,R}^{\complement} \right) \\
&\le \mathbb{P}\left(\sup_{t\in[0,T]} \int_{D\times\mathbb{T}} d_{D\times\mathbb{T}}\left((x_0,v_0),(x,v)\right) S_t^N(\,\mathrm{d}x,\,\mathrm{d}v) > M \right) \\
&\qquad + \mathbb{P}\left(\int_0^T \int_0^T \frac{\mathcal{W}_1\left(S_t^N, S_s^N\right)^p}{|t-s|^{1+\alpha p}} \,\mathrm{d}t \,\mathrm{d}s > R \right).
\end{aligned}$$

For the first term we have

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]} \int_{D\times\mathbb{T}} d_{D\times\mathbb{T}}\left((x_0,v_0),(x,v)\right) S_t^N(\,\mathrm{d}x,\,\mathrm{d}v) > M \right) \\
&\quad \le \frac{1}{M}\mathbb{E}\left[\left| \sup_{t\in[0,T]} \int_{D\times\mathbb{T}} d_{D\times\mathbb{T}}\left((x_0,v_0),(x,v)\right) S_t^N(\,\mathrm{d}x,\,\mathrm{d}v) \right| \right] \\
&\quad \le \frac{1}{MN}\mathbb{E}\left[\sup_{t\in[0,T]} \sum_{j=1}^N d_{D\times\mathbb{T}}\left((x_0,v_0), \left(X_0^i, V_t^{i,N} \pmod 2\right)\right) \right] \\
&\quad \le \frac{1}{M} + \frac{1}{MN}\sum_{i=1}^N \mathbb{E}\left[\sup_{t\in[0,T]} \left| \left(X_0^i - x_0, V_t^{i,N} \right) \right| \right] \\
&\quad \le \frac{1}{M} + \frac{C}{MN}\sum_{i=1}^N \mathbb{E}\Bigg[\left| \left(X_0^i - x_0, V_0^{i,N} \right) \right| \\
&\qquad + \int_0^t \sup_{t\in[0,T]} \left| \lambda_2\left(V_s^{i,N}\right) + \frac{1}{N}\sum_{j=1}^N g_2\left(X_0^i, V_s^{i,N}, X_0^j, V_s^{j,N}\right) \right| \mathrm{d}s \\
&\qquad + \sup_{t\in[0,T]} \left| \int_0^t \sigma_2^\epsilon\left(V_s^{i,N}\right) \,\mathrm{d}B_s^i \right| \Bigg] \le \frac{C}{M}
\end{aligned}$$

for a certain constant $C = C(\hat{\lambda}, \theta, \sigma^\epsilon, T)$, thanks to the Burkholder-Davis-Gundy inequality, the boundedness of λ_2, g_2 and σ_2^ϵ and the fact that $\nu \times \tilde{\rho}_0 \mathscr{L}_{\mathbb{T}}$ has finite first moment.

For the second term

$$\mathbb{P}\left(\int_0^T\int_0^T \frac{\mathcal{W}_1\left(S_t^N,S_s^N\right)^p}{|t-s|^{1+\alpha p}}\,\mathrm{d}t\,\mathrm{d}s>R\right)\le\frac{1}{R}\int_0^T\int_0^T\frac{\mathbb{E}\left[\mathcal{W}_1\left(S_t^N,S_s^N\right)^p\right]}{|t-s|^{1+\alpha p}}\,\mathrm{d}t\,\mathrm{d}s.$$

Let ϕ be a Lipschitz function on $D\times\mathbb{T}$ with Lipschitz constant $K_\phi\le 1$. Then

$$\left|\left\langle S_t^N,\phi\right\rangle-\left\langle S_s^N,\phi\right\rangle\right|\le\frac{1}{N}\sum_{i=1}^N d_{\mathbb{T}}\left(V_t^{i,N},V_s^{i,N}\right)$$

so that, by the Kantorovich characterization of the 1-Wasserstein distance and by Hölder's inequality,

$$\mathbb{E}\left[\mathcal{W}_1\left(S_t^N,S_s^N\right)^p\right]\le\frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[d_{\mathbb{T}}\left(V_t^{i,N},V_s^{i,N}\right)^p\right].$$

Recalling (3), (4) and notation (17), we can write

$$d_{\mathbb{T}}\left(V_t^{i,N},V_s^{i,N}\right)\le\left|V_t^{i,N}-V_s^{i,N}\right|\le\int_s^t\left|b(S_r^N)(X_0^i,V_r^{i,N})\right|\,\mathrm{d}r+\left|\int_s^t\sigma_2^\epsilon(V_r^{i,N})\,\mathrm{d}B_r^i\right|$$

so that

$$\mathbb{E}\left[d_{\mathbb{T}}\left(V_t^{i,N},V_s^{i,N}\right)^p\right|\le C'|t-s|^{p/2}$$

for a suitable constant $C'=C'(\hat{\lambda},\theta,\sigma^\epsilon)$, again by boundedness of the coefficients and the Burkholder-Davis-Gundy inequality. Choosing $p>2$ and α such that $\alpha p<{}^p/_2-1$ we find

$$\mathbb{P}\left(\int_0^T\int_0^T\frac{\mathcal{W}_1\left(S_t^N,S_s^N\right)^p}{|t-s|^{1+\alpha p}}\,\mathrm{d}t\,\mathrm{d}s>R\right)\le\frac{C'}{R}.$$

For any $e>0$ we can now choose M and R so that $\mathbb{Q}^N\left(\mathcal{K}_{M,R}^{\complement}\right)<e$, concluding the proof. □

An alternative approach to prove theorem 4.1 could be based on tightness results from [18, Chapters I and II], using the boundedness of the coefficients and the interchangeability of the $V^{i,N}$. However, the above direct proof can be applied to more general situations as well.

5. **The limit PDE: existence and convergence.**

5.1. **Density Estimates.**

Recall that the empirical measure

$$S_t^N=\frac{1}{N}\sum_{i=1}^N\delta_{\left(\xi^i,V_t^{i,N}\pmod 2\right)}$$

satisfies

$$\begin{aligned}\left\langle S_t^N,\phi\right\rangle=\left\langle S_0^N,\phi\right\rangle&+\int_0^t\left\langle S_r^N,\lambda_2\partial_v\phi\right\rangle\,\mathrm{d}r\\&+\int_0^t\left\langle S_r^N,\left\langle S_r^N,g_2\left(x,v,\cdot,\cdot\right)\right\rangle\partial_v\phi\right\rangle\,\mathrm{d}r+\int_0^t\left\langle S_r^N,\frac{\left(\sigma_2^\epsilon\right)^2}{2}\partial_v^2\phi\right\rangle\,\mathrm{d}r+M_t^{N,\phi}\end{aligned}$$

where

$$M_t^{N,\phi}=\frac{1}{N}\sum_{i=1}^N\int_0^t\sigma_2^\epsilon\left(V_r^{i,N}\right)\partial_v\phi\left(X_0^i,V_r^{i,N}\right)\,\mathrm{d}B_r^i.$$

We consider a smooth probability density $\gamma : \mathbb{T} \to \mathbb{R}$ defined as follows:

$$\gamma(v) = \begin{cases} c\ \exp\left(-\frac{1}{\frac{1}{4}-(d_{\mathbb{T}}(v,0))^2}\right) \cdot (\frac{1}{4} - d_{\mathbb{T}}(v,0)^2)^2, & \text{if } d_{\mathbb{T}}(v,0) < 1/2, \\ 0 \ \text{ otherwise.} \end{cases}$$

and introduce a correspondent family of mollifiers $\gamma_N(v) = \alpha_N^{-1}\gamma\left(\alpha_N^{-1}v\right)$ on $\mathbb{T}$. Note that there exists $C > 0$ such that

$$|\gamma'(w)\, w| \le C\gamma(w), \quad w \in \mathbb{T}.$$

Concerning the positive scaling factor, we assume that $\alpha_N \to 0$ as $N \to \infty$ and

$$\alpha_N^{-3} \le N.$$

Consider the empirical density

$$\begin{aligned} u_t^N(v) :&= \frac{1}{N}\sum_i \int_{D\times\mathbb{T}} \gamma_N(v-v')\, \delta_{\left(X_0^i, V_t^{i,N} \pmod 2\right)}(\mathrm{d}x', \mathrm{d}v') \\ &= \frac{1}{N}\sum_i \gamma_N\left(v - V_t^{i,N} \pmod 2\right) \\ &= \int_{D\times\mathbb{T}} \gamma_N(v-v')\, S_t^N(\mathrm{d}x', \mathrm{d}v') \end{aligned}$$

(where sums and differences are understood on $\mathbb{T}$, i.e. for $v_1, v_2 \in \mathbb{T}$, $v_1 \pm v_2 = (v_1 \pm_{\mathbb{R}} v_2) \pmod 2 \in \mathbb{T}$). It satisfies

$$\begin{aligned} \mathrm{d}u_t^N(v) = &-\langle S_t^N, \lambda_2 \partial_v \gamma_N(v-\cdot)\rangle\, \mathrm{d}t \\ &- \left\langle S_t^N, \langle S_t^N, g_2(x', v', \cdot, \cdot)\rangle \partial_v \gamma_N(v-v')\right\rangle \mathrm{d}t \\ &+ \langle S_t^N, \frac{(\sigma_2^\epsilon)^2}{2}\partial_v^2 \gamma_N(v-\cdot)\rangle\, \mathrm{d}t + \mathrm{d}\overline{M}_t^N(v) \end{aligned}$$

where

$$\overline{M}_t^N(v) = -\frac{1}{N}\sum_{i=1}^N \int_0^t \partial_v \gamma_N\left(v - V_r^{i,N}\right) \sigma_2^\epsilon\left(V_r^{i,N}\right) \mathrm{d}B_r^i$$

and according to (7) we write

$$\begin{aligned} &\left\langle S_t^N, \left(\left\langle S_t^N, g_2(x', v', \cdot, \cdot)\right\rangle\right) \partial_v \gamma_N(v-v')\right\rangle \\ &= \int_{D\times\mathbb{T}} \partial_v \gamma_N(v-v') \int_{D\times\mathbb{T}} g_2(x', v', y, w) S_t^N(\mathrm{d}y, \mathrm{d}w) S_t^N(\mathrm{d}x', \mathrm{d}v'). \end{aligned}$$

In the next lemma we will use that $\sigma_2^\epsilon(v)$ is differentiable with bounded derivative and

$$0 < \epsilon \le \frac{(\sigma_2^\epsilon(v))^2}{2} \le C.$$

Lemma 5.1. *There exists a constant $C_\epsilon > 0$ such that*

$$\sup_{t\in[0,T]} E\int_{\mathbb{T}} \left|u_t^N(v)\right|^2 \mathrm{d}v + E\int_0^T \int_{\mathbb{T}} \left|\partial_v u_t^N(v)\right|^2 \mathrm{d}v\, \mathrm{d}t \le C_\epsilon$$

for every $N \in \mathbb{N}$.

Proof. **Step 1** (energy identity). One has by Itô's formula, integrating by parts,

$$\begin{aligned}\frac{1}{2}\,\mathrm{d}\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\,\mathrm{d}v &= \int_{\mathbb{T}}\frac{\left(\sigma_2^\epsilon(v)\right)^2}{2}u_t^N(v)\,\partial_v^2 u_t^N(v)\,\mathrm{d}v\,\mathrm{d}t\\ &+\int_{\mathbb{T}}\left\langle S_t^N,\left(\lambda_2+\left\langle S_t^N,g_2(x',v',\cdot,\cdot)\right\rangle\right)\gamma_N(v-\cdot)\right\rangle\partial_v u_t^N(v)\,\mathrm{d}v\,\mathrm{d}t\\ &-\int_{\mathbb{T}}R_t^N(v)\,\partial_v u_t^N(v)\,\mathrm{d}v\,\mathrm{d}t+\int_{\mathbb{T}}u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v+\frac{1}{2}\int_{\mathbb{T}}\mathrm{d}\left[\overline{M}^N(v)\right]_t\,\mathrm{d}v\end{aligned}$$

where

$$R_t^N(v)=\partial_v\int_{D\times\mathbb{T}}\gamma_N(v-v')\left[-\frac{\left(\sigma_2^\epsilon(v)\right)^2}{2}+\frac{\left(\sigma_2^\epsilon(v')\right)^2}{2}\right]S_t^N(\mathrm{d}x',\mathrm{d}v'),$$

and we write $\int_{\mathbb{T}}u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v$ for

$$\int_{\mathbb{T}}u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v=\frac{1}{N}\sum_{i=1}^N\left(\int_{\mathbb{T}}\partial_v\gamma_N\left(v-V_t^{i,N}\right)u_t^N(v)\,\mathrm{d}v\right)\sigma_2^\epsilon\left(V_t^{i,N}\right)\mathrm{d}B_t^i$$

and $\int_{\mathbb{T}}\mathrm{d}\left[\overline{M}^N(v)\right]_t\,\mathrm{d}v$ for

$$\int_{\mathbb{T}}\mathrm{d}\left[\overline{M}^N(v)\right]_t\,\mathrm{d}v=\frac{1}{N^2}\sum_{i=1}^N\int_{\mathbb{T}}\left|\partial_v\gamma_N\left(v-V_t^{i,N}\right)\right|^2\mathrm{d}v\left|\sigma_2^\epsilon\left(V_t^{i,N}\right)\right|^2\mathrm{d}t.$$

Step 2 (deterministic terms). Using the assumptions on $\sigma_2^\epsilon(v)$, one has integrating by parts

$$\int_{\mathbb{T}}\frac{\left(\sigma_2^\epsilon(v)\right)^2}{2}u_t^N(v)\,\partial_v^2u_t^N(v)\,\mathrm{d}v\le-\frac{\epsilon}{2}\int_{\mathbb{T}}\left|\partial_v u_t^N(v)\right|^2\mathrm{d}v+C_\epsilon\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v$$

Since (due to the boundedness of λ_2 and g_2)

$$\left|\lambda_2+\left\langle S_t^N,g_2(x,v,\cdot,\cdot)\right\rangle\right|\le C$$

one has

$$\left|\left\langle S_t^N,\left(\lambda_2+\left\langle S_t^N,g_2(x',v',\cdot,\cdot)\right\rangle\right)\gamma_N(v-\cdot)\right\rangle\right|\le Cu_t^N(v).$$

Therefore, $\mathbb{P}$-a.s.,

$$\begin{aligned}\int_{\mathbb{T}}\left\langle S_t^N,\left(\lambda_2+\left\langle S_t^N,g_2(x',v',\cdot,\cdot)\right\rangle\right)\gamma_N(v-\cdot)\right\rangle\partial_v u_t^N(v)\,\mathrm{d}v&\\ \le\frac{\epsilon}{4}\int_{\mathbb{T}}\left|\partial_v u_t^N(v)\right|^2\mathrm{d}v+C_\epsilon\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v.&\end{aligned}$$

We have got that $\mathbb{P}$-a.s.

$$\begin{aligned}&\frac{1}{2}\,\mathrm{d}\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v+\frac{\epsilon}{4}\int_{\mathbb{T}}\left|\partial_v u_t^N(v)\right|^2\mathrm{d}v\,\mathrm{d}t\le C_\epsilon\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v\\ &-\int_{\mathbb{T}}R_t^N(v)\,\partial_v u_t^N(v)\,\mathrm{d}v\,\mathrm{d}t+\int_{\mathbb{T}}u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v+\frac{1}{2}\int_{\mathbb{T}}\mathrm{d}\left[\overline{M}^N(v)\right]_t\,\mathrm{d}v.\end{aligned}$$

Finally, using also the assumption $|\gamma'(w)\,w| \le C\gamma(w)$,

$$\begin{aligned}\left|R_t^N(v)\right| &\le C\int_{D\times\mathbb{T}} \left[|\partial_v\gamma_N(v-v')|\,|v-v'| + \gamma_N(v-v')\right] S_t^N(\mathrm{d}x',\mathrm{d}v')\\ &= C\int_{D\times\mathbb{T}} \left[\alpha_N^{-1}\left|\gamma'\left(\alpha_N^{-1}(v-v')\right)\right|\alpha_N^{-1}|v-v'| + \gamma_N(v-v')\right] S_t^N(\mathrm{d}x',\mathrm{d}v')\\ &\le C'\int_{D\times\mathbb{T}} \alpha_N^{-1}\gamma\left(\alpha_N^{-1}(v-v')\right) S_t^N(\mathrm{d}x',\mathrm{d}v')\\ &= C'u_t^N(v)\end{aligned}$$

which yields

$$\left|\int_{\mathbb{T}} R_t^N(v)\,\partial_v u_t^N(v)\,\mathrm{d}v\right| \le \frac{\epsilon}{8}\int_{\mathbb{T}}\left|\partial_v u_t^N(v)\right|^2\mathrm{d}v + C_\epsilon\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v.$$

This implies

$$\begin{aligned}&\frac{1}{2}\,\mathrm{d}\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v + \frac{\epsilon}{8}\int_{\mathbb{T}}\left|\partial_v u_t^N(v)\right|^2\mathrm{d}v\,\mathrm{d}t\\ &\qquad\le C_\epsilon\int_{\mathbb{T}}\left|u_t^N(v)\right|^2\mathrm{d}v + \int_{\mathbb{T}} u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v + \frac{1}{2}\int_{\mathbb{T}}\mathrm{d}\left[\overline{M}^N(v)\right]_t\mathrm{d}v\,,\end{aligned}$$

$\mathbb{P}$ almost surely.

Step 3 (martingale terms and conclusion). It remain to handle the sum

$$\int_{\mathbb{T}} u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v + \frac{1}{2}\,\mathrm{d}\left[\overline{M}^N(v)\right]_t\,.$$

The term $\int_{\mathbb{T}} u_t^N(v)\,\mathrm{d}\overline{M}_t^N(v)\,\mathrm{d}v$ is a martingale, hence it has mean zero. Indeed, for every N and $i = 1, ..., N$,

$$\begin{aligned}&\mathbb{E}\int_0^T\left|\left(\int_{\mathbb{T}}\partial_v\gamma_N\left(v-V_t^{i,N}\right)u_t^N(v)\,\mathrm{d}v\right)\sigma_2^\epsilon\left(V_t^{i,N}\right)\right|^2\mathrm{d}t\\ &\qquad\le C\mathbb{E}\int_0^T\left|\int_{\mathbb{T}}\partial_v\gamma_N\left(v-V_t^{i,N}\right)u_t^N(v)\,\mathrm{d}v\right|^2\mathrm{d}t\\ &\qquad\le C_N\mathbb{E}\int_0^T\left|\int_{\mathbb{T}}u_t^N(v)\,\mathrm{d}v\right|^2\mathrm{d}t = C_NT\end{aligned}$$

(because $\int_{\mathbb{T}} u_t^N(v)\,\mathrm{d}v = 1$). As to the corrector, we have

$$\begin{aligned}\int_{\mathbb{T}}\left|\partial_v\gamma_N\left(v-V_r^{i,N}\right)\right|^2\mathrm{d}v &= \int_{\mathbb{T}}\left|\partial_v\gamma_N(v)\right|^2\mathrm{d}v\\ &= \alpha_N^{-2}\alpha_N^{-2}\int_{\mathbb{T}}\left|\gamma'\left(\alpha_N^{-1}v\right)\right|^2\mathrm{d}v = C\alpha_N^{-3}\end{aligned}$$

where $C = \int_{\mathbb{T}}\left|\gamma'(v)\right|^2\mathrm{d}v$. Hence, $\mathbb{P}$-a.s.,

$$\begin{aligned}\int_{\mathbb{T}}\left[\overline{M}^N(v)\right]_t dv &= \frac{1}{N^2}\sum_{i=1}^N\int_0^t\int_{\mathbb{T}}\left|\partial_v\gamma_N\left(v-V_r^{i,N}\right)\right|^2\mathrm{d}v\left|\sigma_2^\epsilon\left(V_r^{i,N}\right)\right|^2\mathrm{d}r\\ &\le C\frac{1}{N^2}\sum_{i=1}^N\int_0^t\left(\int_{\mathbb{T}}\left|\partial_v\gamma_N\left(v-V_r^{i,N}\right)\right|^2\mathrm{d}v\right)\mathrm{d}r = C\frac{\alpha_N^{-3}}{N}t \le Ct\end{aligned}$$

under the assumption $\alpha_N^{-3} \le N$.

Using the assumption that the law of the initial data η^i has an L^2 density, it is not difficult

to show that the L^2 norm of u_0^N is bounded uniformly with respect to N. To this purpose let us recall that we denote by $\tilde{\rho}$ the density of each η^i. Using also standard property of convolutions we get:

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{T}} |u_0^N(v)|^2 \,\mathrm{d}v = \int_{\mathbb{T}} \mathbb{E}\Big|\frac{1}{N}\sum_i \gamma_N\left(v-\eta^i\right)\Big|^2 \,\mathrm{d}v \\
&= \frac{1}{N^2}\frac{1}{\alpha_N^2}\int_{\mathbb{T}} \sum_i \mathbb{E}|\gamma\left(\frac{v-\eta^i}{\alpha_N}\right)|^2 \,\mathrm{d}v = \frac{1}{N}\frac{1}{\alpha_N^2}\int_{\mathbb{T}} \,\mathrm{d}v \int_{\mathbb{T}} \left|\gamma\left(\frac{v-w}{\alpha_N}\right)\right|^2 \tilde{\rho}(w)\,\mathrm{d}w \\
&\quad \le \|\tilde{\rho}\|_{L^1(\mathbb{T})}\frac{1}{N}\frac{1}{\alpha_N^2}\int_{\mathbb{T}} \left|\gamma\left(\frac{v}{\alpha_N}\right)\right|^2 \mathrm{d}v = \|\tilde{\rho}\|_{L^1(\mathbb{T})}\frac{1}{N}\frac{1}{\alpha_N}\int_{\mathbb{T}} \left|\gamma\left(v'\right)\right|^2 \mathrm{d}v' \le C,
\end{aligned}$$

where $C>0$ is independent of N. We can therefore take expectation and apply Gronwall's lemma, thus deducing the claim from the results of the two previous steps. □

Lemma 5.2. *There exists $\alpha>0$ small enough such that*

$$\mathbb{E}\int_0^T\int_0^T \frac{\left\|u_t^N-u_s^N\right\|_{H^{-2}}^2}{|t-s|^{1+2\alpha}}\,\mathrm{d}s\,\mathrm{d}t \le C_\epsilon.$$

where $H^{-2}=H^{-2}(\mathbb{T})$.

Proof. Arguing as in Lemma 5.1, we have, $\mathbb{P}$-a.s., for any $0\le s\le t\le T$, $\phi\in H^2(\mathbb{T})$,

$$\begin{aligned}
\int_{\mathbb{T}}[u_t^N(v)-u_s^N(v)]\phi(v)dv &= \int_s^t dr\int_{\mathbb{T}} \frac{\left(\sigma_2^\epsilon(v)\right)^2}{2}u_r^N(v)\,\partial_v^2\phi(v)\,\mathrm{d}v \\
&+\int_s^t dr\int_{\mathbb{T}} \left\langle S_r^N,\left(\lambda_2+\left\langle S_r^N,g_2\left(x',v',\cdot,\cdot\right)\right\rangle\right)\gamma_N\left(v-v'\right)\right\rangle \partial_v\phi(v)\,\mathrm{d}v \\
&-\int_s^t dr\int_{\mathbb{T}} R_r^N(v)\,\partial_v\phi(v)\,\mathrm{d}v+\int_{\mathbb{T}}[\overline{M}_t^N(v)-\overline{M}_s^N(v)]\phi(v)dv
\end{aligned}$$

where $R_t^N(v)$ and $\overline{M}_t^N(v)$ are given in Step 1 of the previous lemma. Then (using the same inequalities proved above in Step 2 of the previous lemma)

$$\left\|u_t^N-u_s^N\right\|_{H^{-2}}^2 \le 3\times C(t-s)\int_s^t \left\|u_r^N\right\|_{L^2}^2\,\mathrm{d}r + C\left\|\overline{M}_t^N-\overline{M}_s^N\right\|_{H^{-2}}^2.$$

It is sufficient (because of the claim of the previous lemma) to prove that

$$\mathbb{E}\int_0^T\int_0^T \frac{\left\|\overline{M}_t^N-\overline{M}_s^N\right\|_{H^{-2}}^2}{|t-s|^{1+2\alpha}}\,\mathrm{d}s\,\mathrm{d}t \le C_\epsilon.$$

Recall that

$$\overline{M}_t^N(v) = \partial_v\left(\frac{1}{N}\sum_{i=1}^N\int_0^t \gamma_N\left(v-V_r^{i,N}\right)\sigma_2^\epsilon\left(V_r^{i,N}\right)\,\mathrm{d}B_r^i\right).$$

Then

$$\begin{aligned}
\mathbb{E}\left[\left\|\overline{M}_t^N-\overline{M}_s^N\right\|_{H^{-2}}^2\right] &\leq C\mathbb{E}\left[\left\|\frac{1}{N}\sum_{i=1}^N\int_s^t\gamma_N\left(v-V_r^{i,N}\right)\sigma_2^\epsilon\left(V_r^{i,N}\right)\,\mathrm{d}B_r^i\right\|_{L^2}^2\right]\\
&=\frac{C}{N^2}\int_{\mathbb{T}}\mathbb{E}\left[\left|\sum_{i=1}^N\int_s^t\gamma_N\left(v-V_r^{i,N}\right)\sigma_2^\epsilon\left(V_r^{i,N}\right)\,\mathrm{d}B_r^i\right|^2\right]\mathrm{d}v\\
&=\frac{C}{N^2}\int_{\mathbb{T}}\sum_{i=1}^N\mathbb{E}\left[\int_s^t\left|\gamma_N\left(v-V_r^{i,N}\right)\sigma_2^\epsilon\left(V_r^{i,N}\right)\right|^2\mathrm{d}r\right]\mathrm{d}v\\
&\leq\frac{C}{N^2}\int_{\mathbb{T}}\sum_{i=1}^N\mathbb{E}\left[\int_s^t\left|\gamma_N\left(v-V_r^{i,N}\right)\right|^2dr\right]\mathrm{d}v\\
&=\frac{C}{N^2}\sum_{i=1}^N\int_s^t\mathbb{E}\left[\int_{\mathbb{T}}\left|\gamma_N\left(v-V_r^{i,N}\right)\right|^2\mathrm{d}v\right]\mathrm{d}r\\
&=\frac{C}{N^2}\sum_{i=1}^N\int_s^t\int_{\mathbb{T}}\left|\gamma_N\left(v\right)\right|^2\mathrm{d}v\,\mathrm{d}r\leq C\frac{\alpha_N^{-1}}{N}\left(t-s\right)\leq C\left(t-s\right)
\end{aligned}$$

where we have used the estimate $\int_{\mathbb{T}}\left|\gamma_N\left(v\right)\right|^2\,\mathrm{d}v\leq C\alpha_N^{-1}$ and the assumption $\alpha_N^{-3}\leq N$. The proof is complete. □

Now let $\mathbb{Q}_{u^N}$ denote the law of the process u^N. From the previous two lemmas, we deduce that the family $(\mathbb{Q}_{u^N})$ is tight in

$$L^2\left(0,T;L^2\left(\mathbb{T}\right)\right)$$

due to a generalized version of Aubin-Lions lemma, which claims that the space

$$L^2\left(0,T;W^{1,2}\left(\mathbb{T}\right)\right)\cap W^{\alpha,2}\left(0,T;H^{-2}\left(\mathbb{T}\right)\right)$$

is relatively compact in $L^2\left(0,T;L^2\left(\mathbb{T}\right)\right)$, for $\alpha>0$ (cf. [17]).

Remark 3. Introducing the mollifiers $\gamma_n\left(v\right)=\alpha_n^{-1}\gamma\left(\alpha_n^{-1}v\right)$, $n\geq1$, with $\left|\gamma'\left(w\right)w\right|\leq C\gamma\left(w\right),\quad w\in\mathbb{T}$, $\alpha_n\to0$ as $n\to\infty$, and $\alpha_n^{-3}\leq N$, and following the proof of Lemma 5.1, we can obtain that there exists a constant $C_\epsilon>0$ such that

$$\sup_{t\in[0,T]}\mathbb{E}\int_{\mathbb{T}}\left|u_t^{n,N}\left(v\right)\right|^2\mathrm{d}v+\mathbb{E}\int_0^T\int_{\mathbb{T}}\left|\partial_v u_t^{n,N}\left(v\right)\right|^2\mathrm{d}v\,\mathrm{d}t\leq C_\epsilon,\tag{28}$$

for every $n,\ N\in\mathbb{N}$, where $u_t^{n,N}(v)=\int_{D\times\mathbb{T}}\gamma_n\left(v-v'\right)S_t^N\left(\mathrm{d}x',\,\mathrm{d}v'\right)$.

Now note that given a Borel probability measure ν on $\mathbb{T}$, if there exists $c>0$, such that, for any $n\geq1$,

$$\|\nu*\gamma_n\|_{L^2(\mathbb{T})}\leq c,\tag{29}$$

then $\nu\in L^2(\mathbb{T})$ and $\|\nu\|_{L^2(\mathbb{T})}\leq c$. Indeed, by (29), for any $\phi\in L^2(\mathbb{T})$, we have

$$\left|\int_{\mathbb{T}}\phi(y)dy\int_{\mathbb{T}}\gamma_n(y-y')\nu(dy')\right|=\left|\int_{\mathbb{T}}\nu(dy')\int_{\mathbb{T}}\phi(y)\gamma_n(y-y')dy\right|\leq c\|\phi\|_{L^2(\mathbb{T})}.$$

Passing to the limit as $n\to\infty$ and using the Riesz theorem we get the assertion.

Estimate (28) and the previous argument could be used to prove existence of solutions to (16) in $\tilde{X}$ (see the next section) avoiding the previous Aubin-Lions lemma.

5.2. **Convergence and existence of solutions.** Set for notational convenience

$$\widetilde{X} = \left\{ \mu \in \mathscr{C} \colon \pi_v \mu_t \ll \mathscr{L}_{\mathbb{T}} \text{ with } \frac{\mathrm{d}\left(\pi_v \mu_t\right)}{\mathrm{d}\mathscr{L}_{\mathbb{T}}} \in L^2(\mathbb{T}), \text{ for a.e. } t \in [0,T] \right\} \tag{30}$$

where $\pi_v \mu_t$ is the marginal on the v-component of μ_t:

$$\int_{\mathbb{T}} f(v) \left(\pi_v \mu_t\right)(dv) := \int_{D \times \mathbb{T}} f(v) \, \mu_t(dx, dv), \quad f \in C(\mathbb{T}).$$

Lemma 5.3. *The space $\tilde{X}$ is a Borel subset of $\mathscr{C}$.*

Proof. It is enough to show that

$$\Lambda = \left\{ \mu \in \mathrm{Pr}_1(D \times \mathbb{T}) : \pi_v \mu \ll \mathscr{L}_{\mathbb{T}} \text{ with } \frac{\mathrm{d}\left(\pi_v \mu\right)}{\mathrm{d}\mathscr{L}_{\mathbb{T}}} \in L^2(\mathbb{T}) \right\}$$

is a Borel subset of $\mathrm{Pr}_1(D \times \mathbb{T})$.

We consider the continuous mapping $\mathcal{J} : \mathrm{Pr}_1(D \times \mathbb{T}) \to \mathrm{Pr}_1(\mathbb{T})$ given by $\mathcal{J}\mu = \pi_v \mu$, for any $\mu \in \mathrm{Pr}_1(D \times \mathbb{T})$. If we prove that

$$\Gamma = \left\{ \mu \in \mathrm{Pr}_1(\mathbb{T}) : \mu \ll \mathscr{L}_{\mathbb{T}} \text{ with } \frac{\mathrm{d}\mu}{\mathrm{d}\mathscr{L}_{\mathbb{T}}} \in L^2(\mathbb{T}) \right\}$$

is Borel in $\mathrm{Pr}_1(\mathbb{T})$ then we get that $\Lambda = \mathcal{J}^{-1}(\Gamma)$ is Borel and this finishes the proof. Let us check the assertion on Γ.

Let $\mu \in \mathrm{Pr}_1(\mathbb{T})$. Using the Riesz theorem and the fact that $C(\mathbb{T})$ is dense in $L^2(\mathbb{T})$, we know that $\mu \in \Gamma$ if and only if there exists $c > 0$ such that

$$\left| \int_{\mathbb{T}} f(y) \mu(dy) \right| \le c \|f\|_{L^2}, \quad \text{for all } f \in C(\mathbb{T}) \tag{31}$$

(indeed if (31) holds for $\mu \in \mathrm{Pr}_1(\mathbb{T})$ then μ can be uniquely extended to a linear functional on $L^2(\mathbb{T})$). Let us define, for integers $N \ge 1$,

$$\Gamma_N = \{ \mu \in \mathrm{Pr}_1(\mathbb{T}) : \ (31) \text{ holds with } c \text{ replaced by } N \}$$

It is easy to check that each Γ_N is closed in $\mathrm{Pr}_1(\mathbb{T})$. We have $\Gamma = \bigcup_{N \ge 1} \Gamma_N$ and this shows that Γ is Borel. □

For any test function $\phi \in \mathcal{T}$ and any $\mu^0 \in \mathrm{Pr}(D \times \mathbb{T})$ define on $\mathscr{C}$ the functional

$$\begin{aligned} &\Phi_\phi^{\mu^0}(\mu) \\ &= \sup_{t \in [0,T]} \left| \langle \mu_t, \phi \rangle - \langle \mu^0, \phi \rangle - \int_0^t \langle \mu_s, b(\mu_s) \partial_v \phi \rangle \, \mathrm{d}s - \int_0^t \left\langle \mu_s, \frac{(\sigma_2^\epsilon)^2}{2} \partial_v^2 \phi \right\rangle \mathrm{d}s \right| \wedge 1 \end{aligned} \tag{32}$$

where

$$b(\mu_s)(x,v) = \lambda_2(v) + \langle \mu_s, g_2(x,v,\cdot,\cdot) \rangle \ .$$

Lemma 5.4. *For every $\phi \in \mathcal{T}$ and every $\mu^0 \in \mathrm{Pr}(D \times \mathbb{T})$, the bounded Borel measurable functional $\Phi_\phi^{\mu^0} : \mathscr{C} \to \mathbb{R}$ is continuous at every point of $\widetilde{X}$.*

Therefore, if $\{Q^N\}_{N \in \mathbb{N}}$ and Q are probability measures on $\mathscr{C}$ such that $Q^N \to Q$ weakly and $Q\left(\widetilde{X}\right) = 1$, then

$$\int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}Q^N \to \int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}Q \ \text{ as } N \to \infty \, .$$

Proof. Since in the definition of $\Phi_\phi^{\mu^0}$ we can consider the supremum over rational numbers in $[0,T]$, to prove the measurability of $\Phi_\phi^{\mu^0}$ we can fix $t \in [0,T]$ and study separately the measurability of three functionals:

$$\Phi_1(\mu) = \langle \mu_t, \phi \rangle - \langle \mu_0, \phi \rangle ,$$
$$\Phi_2(\mu) = \int_0^t \langle \mu_s, b(\mu_s)\, \partial_v \phi \rangle \, \mathrm{d}s,$$
$$\Phi_3(\mu) = \int_0^t \left\langle \mu_s, \frac{(\sigma_2^\epsilon)^2}{2} \partial_v^2 \phi \right\rangle \mathrm{d}s,$$

for $\mu \in \mathscr{C}$. Note that Φ_1 and Φ_3 are even continuous mappings on $\mathscr{C}$. Concerning the measurability of Φ_2 we first approximate pointwise the functions λ_2 and g_2 by regular functions λ_2^n and g_2^n (indeed $\lambda_2(\cdot)$ and $g_2(x,\cdot,y,\cdot)$ have only simple discontinuities) and then consider the corresponding functions b^n given by

$$b^n(\mu_s)(x,v) = \lambda_2^n(v) + \langle \mu_s, g_2^n(x,v,\cdot,\cdot) \rangle . \tag{33}$$

It is not difficult to prove that for each n the functional $\Phi_2^n : \mathscr{C} \to \mathbb{R}$,

$$\Phi_2^n(\mu) = \int_0^t \langle \mu_s, b^n(\mu_s)\, \partial_v \phi \rangle \, \mathrm{d}s$$

is continuous on $\mathscr{C}$. By the dominated convergence theorem we deduce that $\Phi_2^n(\mu) \to \Phi_2(\mu)$ as $n \to \infty$, for any $\mu \in \mathscr{C}$. This shows that also Φ_2 is measurable.

Let now $\mu \in \widetilde{X}$ and $\mu^n \in \mathscr{C}$ be given with $\mu^n \to \mu$ in $\mathscr{C}$. This implies $\mu_t^n \to \mu_t$ in weak sense, hence $\langle \mu_t^n, \phi \rangle \to \langle \mu_t, \phi \rangle$, uniformly in $t \in [0,T]$. The convergence of $\left\langle \mu_s^n, \frac{(\sigma_2^\epsilon)^2}{2} \partial_v^2 \phi \right\rangle$ to $\left\langle \mu_s, \frac{(\sigma_2^\epsilon)^2}{2} \partial_v^2 \phi \right\rangle$ for every $s \in [0,T]$ is similar and, by Lebesgue dominated convergence theorem, the last integral in the definition of $\Phi_\phi^{\mu^0}$ converges, uniformly in $t \in [0,T]$. It remains to prove that the first integral converges. Again by Lebesgue dominated convergence theorem, the problem is reduced to prove that, for a.e. $s \in [0,T]$,

$$\langle \mu_s^n, b(\mu_s^n)\, \partial_v \phi \rangle \to \langle \mu_s, b(\mu_s)\, \partial_v \phi \rangle .$$

This is more difficult since λ_2 and g_2 contain discontinuities. Since $\mu \in \widetilde{X}$, we know that $\pi_v \mu_s \ll \mathscr{L}_{\mathbb{T}}$ for a.e. $s \in [0,T]$, thus in the sequel we restrict to such values of s.

Let us first explain why

$$\langle \mu_s^n, \lambda_2 \partial_v \phi \rangle \to \langle \mu_s, \lambda_2 \partial_v \phi \rangle . \tag{34}$$

The function $(x,v) \mapsto \lambda_2(v)\, \partial_v \phi(x,v)$ is bounded; and it is continuous except on the set

$$S = D \times \{0,1\} \subset D \times \mathbb{T} .$$

These sets are exceptional for the measure μ_s:

$$\int_S \mu_s(\mathrm{d}x, \mathrm{d}v) = \int_{D\times\{v=0\}} \mu_s(\mathrm{d}x, \mathrm{d}v) + \int_{D\times\{v=1\}} \mu_s(\mathrm{d}x, \mathrm{d}v)$$
$$= \int_{\{v=0\}\cup\{v=1\}} (\pi_v \mu_s)(\mathrm{d}v) = 0 .$$

Now, the following fact is known: if a sequence of probability measures ρ_n on a Polish space Y converges weakly to a probability measures ρ and $f : Y \to \mathbb{R}$ is a bounded Borel measurable function, continuous on a set $\widetilde{Y} \subset Y$ with $\rho\left(\widetilde{Y}\right) = 1$, then $\int_Y f d\rho_n \to$

$\int_Y f d\rho$. The proof is easy using Skorohod representation theorem. We apply this fact with $Y = D \times \mathbb{T}$, $\rho_n = \mu_s^n$, $\rho = \mu_s$, $\widetilde{Y} = S^c$, $f = \lambda_2 \partial_v \phi$ and deduce (34).

Finally, let us explain why

$$\langle \mu_s^n, \langle \mu_s^n, g_2(x, v, \cdot, \cdot) \rangle \partial_v \phi \rangle - \langle \mu_s, \langle \mu_s, g_2(x, v, \cdot, \cdot) \rangle \partial_v \phi \rangle \to 0. \tag{35}$$

The previous difference can be rewritten as the sum of two terms:

$$\langle \mu_s^n, \langle (\mu_s^n - \mu_s), g_2(x, v, \cdot, \cdot) \rangle \partial_v \phi \rangle$$

and

$$\langle (\mu_s^n - \mu_s), \langle \mu_s, g_2(x, v, \cdot, \cdot) \rangle \partial_v \phi \rangle .$$

The convergence to zero of the second term is similar to (34), because the function

$$\begin{aligned} D \times \mathbb{T} \ni (x, v) \mapsto \langle \mu_s, g_2(x, v, \cdot, \cdot) \rangle &= \int_{D \times \mathbb{T}} g_2(x, v, x', v') \mu_s(\mathrm{d}x', \mathrm{d}v') \\ &= \mathbb{1}_{[0,1]}(v) \int_{D \times \mathbb{T}} \theta(x, x') \mathbb{1}_{[1,1+\delta]}(v') \mu_s(\mathrm{d}x', \mathrm{d}v') \end{aligned}$$

is continuous on S^c. To treat the first term in the sum, we first fix $\tau > 0$. By the weak convergence, we know that (μ_s^n) is tight and so there exists a compact set $K_\tau \subset D$ such that

$$\mu_s^n((K_\tau \times \mathbb{T})^c) < \tau, \quad \mu_s((K_\tau \times \mathbb{T})^c) < \tau, \quad n \geq 1.$$

We have

$$\begin{aligned} |\langle \mu_s^n, \langle (\mu_s^n - \mu_s), g_2(x, v, \cdot, \cdot) \rangle \partial_v \phi \rangle| &\leq \tau \|\partial_v \phi\|_\infty \|\theta\|_\infty \\ + \|\partial_v \phi\|_\infty \int_{K_\tau \times \mathbb{T}} |\langle (\mu_s^n - \mu_s), g_2(x, v, \cdot, \cdot) \rangle| \mu_s^n(dx, dv). \end{aligned}$$

Now, for any $(x, v) \in K_\tau \times \mathbb{T}$, $n \geq 1$,

$$|\langle (\mu_s^n - \mu_s), g_2(x, v, \cdot, \cdot) \rangle| \leq |g_n(x) - g(x)|$$

where

$$\begin{aligned} g_n(x) &:= \int_{D \times \mathbb{T}} \theta(x, x') \mathbb{1}_{[1,1+\delta]}(v') \mu_s^n(\mathrm{d}x', \mathrm{d}v'), \\ g(x) &:= \int_{D \times \mathbb{T}} \theta(x, x') \mathbb{1}_{[1,1+\delta]}(v') \mu_s(\mathrm{d}x', \mathrm{d}v'). \end{aligned}$$

To check (35) we have to prove that $g_n \to g$ uniformly on K_ϵ. We know it converges pointwise, by the same argument used above for (34), because the function $\mathbb{T} \ni v' \mapsto \mathbb{1}_{[1,1+\delta]}(v')$ is continuous apart in $v' = 1$ and $v' = 1 + \delta$. Uniform convergence then follows from the fact that the family $\{g_n\}$ is equi-bounded and equi-uniformly continuous; the last fact is a consequence of the assumption that θ is uniformly continuous on $D \times D$.

This completes the proof of the first claim of the lemma. The second claim is a simple consequence using the convergence criterion recalled above in this proof, applied with $Y = X$, $\widetilde{Y} = \widetilde{X}$, $\rho_n = Q^n$, $\rho = Q$, $f = \Phi_\phi^{\mu^0}$. □

Lemma 5.5. *Recall that $\mathbb{Q}^N$ are the laws on $\mathscr{C}$ of the empirical process S^N. If $\mathbb{Q}$ is a weak limit point of any subsequence of $\{\mathbb{Q}^N\}$ then*

$$\mathbb{Q}\left(\widetilde{X}\right) = 1.$$

Proof. **Step 1**. We have already proved not only tightness of the family $\{\mathbb{Q}^N\}$ in $\mathscr{C}$ (see Section 4) but also tightness of the family of laws of u^N in $\mathcal{H} := L^2([0,T]\times\mathbb{T})$, where $u_t^N(v) = \int_{D\times\mathbb{T}} \gamma_N(v-v')\, S_t^N(\mathrm{d}x', \mathrm{d}v')$ (see the end of Section 5.1). Consider the pair (S^N, u^N) with values in $\mathscr{C}\times\mathcal{H}$; their laws $\rho^n = \mathcal{L}(S^N, u^N)$ form a tight family in $\mathscr{C}\times\mathcal{H}$. Given a weak limit point $\mathbb{Q}$ of $\{\mathbb{Q}^N\}$, there is thus a subsequence N_k such that ρ^{N_k} converges weakly to a probability measure ρ on $\mathscr{C}\times\mathcal{H}$, with marginal $\mathbb{Q}$ on $\mathscr{C}$.

By the Skorohod embedding theorem there exists a new probability space $\left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}\right)$, $\mathscr{C}\times\mathcal{H}$-valued random variables $\left(\widehat{S}^{N_k}, \widehat{u}^{N_k}\right)$ and $\left(\widehat{S}, \widehat{u}\right)$, with laws ρ^{N_k} and ρ respectively, such that $\left(\widehat{S}^{N_k}, \widehat{u}^{N_k}\right) \to \left(\widehat{S}, \widehat{u}\right)$ in the strong topology of $\mathscr{C}\times\mathcal{H}$, with $\widehat{\mathbb{P}}$-probability one. Notice that the law of $\widehat{S}$ is $\mathbb{Q}$.

Step 2. Let $\mathcal{D}$ be a countable dense family in $C(\mathbb{T})$. Let $\mathscr{L}_{[0,T]}$ be the Lebesgue measure on $[0,T]$. We claim that, given $\phi\in C(\mathbb{T})$, $\left(\widehat{\mathbb{P}}\otimes\mathscr{L}_{[0,T]}\right)$-almost everywhere,

$$\int_{\mathbb{T}} \phi(v)\left(\pi_v \widehat{S}_t\right)(\mathrm{d}v) = \int_{\mathbb{T}} \phi(v)\,\widehat{u}_t(v)\,\mathrm{d}v. \tag{36}$$

To prove this claim, let us start from the definition

$$u_t^{N_k}(v) := \int_{D\times\mathbb{T}} \gamma_{N_k}(v-v')\, S_t^{N_k}(\mathrm{d}x', \mathrm{d}v'), \quad t\in[0,T],\ v\in\mathbb{T}. \tag{37}$$

Note that this implies that

$$0 = \mathbb{E}\int_0^T dt \int_{\mathbb{T}} \left| u_t^{N_k}(v) - \int_{D\times\mathbb{T}} \gamma_{N_k}(v-v')\, S_t^{N_k}(\mathrm{d}x', \mathrm{d}v') \right|^2 dv$$

$$= \widehat{\mathbb{E}}\int_0^T dt \int_{\mathbb{T}} \left| \widehat{u}_t^{N_k}(v) - \int_{D\times\mathbb{T}} \gamma_{N_k}(v-v')\, \widehat{S}_t^{N_k}(\mathrm{d}x', \mathrm{d}v') \right|^2 dv.$$

It follows that given N_k, with $\widehat{\mathbb{P}}$-probability one,

$$\widehat{u}_t^{N_k}(v) = \int_{D\times\mathbb{T}} \gamma_{N_k}(v-v')\, \widehat{S}_t^{N_k}(\mathrm{d}x', \mathrm{d}v')$$

as an identity in $\mathcal{H}$. We can also say that $\left(\widehat{\mathbb{P}}\otimes\mathscr{L}_{[0,T]}\right)$-almost everywhere, for any $k\in\mathbb{N}$ the previous identity holds in $L^2(\mathbb{T})$. Therefore given $\phi\in\mathcal{D}$, we have

$$\int_{\mathbb{T}} \phi(v)\,\widehat{u}_t^{N_k}(v)\,\mathrm{d}v = \int_{D\times\mathbb{T}} \left(\int_{\mathbb{T}} \gamma_{N_k}(v-v')\,\phi(v)\,\mathrm{d}v \right) \widehat{S}_t^{N_k}(\mathrm{d}x', \mathrm{d}v').$$

Up to passing to a subsequence N_k' we can assume that $\widehat{u}^{N_k'}$ converges to $\widehat{u}$ in the strong topology of $L^2(\mathbb{T})$, for $\left(\widehat{\mathbb{P}}\otimes\mathscr{L}_{[0,T]}\right)$-a.e. $(\widehat{\omega}, t)\in\widehat{\Omega}\times[0,T]$. Therefore $\int_{\mathbb{T}} \phi(v)\,\widehat{u}_t^{N_k'}(v)\,\mathrm{d}v$ converges to $\int_{\mathbb{T}} \phi(v)\,\widehat{u}_t(v)\,dv$ for $\left(\widehat{\mathbb{P}}\otimes\mathscr{L}_{[0,T]}\right)$-a.e. $(\widehat{\omega}, t)\in\widehat{\Omega}\times[0,T]$. And, for $\widehat{\mathbb{P}}$-a.e. $\widehat{\omega}\in\widehat{\Omega}$, for every $t\in[0,T]$, $\widehat{S}_t^{N_k'}$ converges in the weak topology of probability measures to $\widehat{S}_t$, so

$$\lim_{k\to\infty} \int_{D\times\mathbb{T}} \left(\int_{\mathbb{T}} \gamma_{N_k'}(v-v')\,\phi(v)\,\mathrm{d}v \right) \widehat{S}_t^{N_k'}(\mathrm{d}x', \mathrm{d}v') = \int_{D\times\mathbb{T}} \phi(v')\,\widehat{S}_t(\mathrm{d}x', \mathrm{d}v')$$

$$= \int_{\mathbb{T}} \phi(v)\left(\pi_v \widehat{S}_t\right)(\mathrm{d}v)$$

where we have used the property

$$\lim_{k\to\infty}\int_{\mathbb{T}}\gamma_{N'_k}(v-v')\,\phi(v)\,\mathrm{d}v=\phi(v')\,,\quad\text{uniformly in } v'\in\mathbb{T}\ ,$$

because ϕ is continuous and γ_N are mollifiers on the torus. This proves (36).

Step 3. Since $\mathcal{D}$ is countable, property (36) holds true uniformly in $\phi\in\mathcal{D}$. Hence we can say that, for $\left(\widehat{\mathbb{P}}\otimes\mathscr{L}_{[0,T]}\right)$-a.e. $(\widehat{\omega},t)\in\widehat{\Omega}\times[0,T]$, $\pi_v\widehat{S}_t(\widehat{\omega})\ll\mathscr{L}_{\mathbb{T}}$ with density in $L^2(\mathbb{T})$. This implies that for $\widehat{\mathbb{P}}$-a.e. $\widehat{\omega}\in\widehat{\Omega}$, we have the property that $\pi_v\widehat{S}_t(\widehat{\omega})\ll\mathscr{L}_{\mathbb{T}}$ with density in $L^2(\mathbb{T})$ for $\mathscr{L}_{[0,T]}$-a.e. $t\in[0,T]$. Hence $\widehat{\mathbb{P}}\left(\widehat{\omega}\in\widehat{\Omega}:\widehat{S}(\widehat{\omega})\in\widetilde{X}\right)=1$, which implies $\mathbb{Q}\left(\widetilde{X}\right)=1$ (recall that the law of $\widehat{S}$ is $\mathbb{Q}$). The proof is complete. □

Let us eventually study the nonlinear Fokker-Planck equation (16), that is

$$\partial_t\mu_t+\partial_v\left(\mu_t b(\mu_t)\right)=\partial_v^2\left(\frac{(\sigma_2^\epsilon)^2}{2}\mu_t\right),$$

with initial condition μ^0, where $b(\mu_t)$ is given by (17).

Theorem 5.6. *Let $\mu^0=\nu\times\tilde{\rho}_0\mathscr{L}_{\mathbb{T}}$ with $\nu\in\mathrm{Pr}_1(D)$ and $\tilde{\rho}\in L^2(0,2)$ (cf. Section 1.1). Then:*

- *i) the nonlinear Fokker-Planck equation (16), with initial condition μ^0, has one and only one weak measure-valued solution $\mu\in\mathscr{C}$; this measure belongs to the space $\widetilde{X}$ (see (30));*
- *ii) let $\mathbb{Q}^N$ be the laws on $\mathscr{C}$ of the empirical process S^N: then $\mathbb{Q}^N$ converges weakly to δ_μ. Moreover, S^N converges in probability to μ, in the topology of $\mathscr{C}$.*

Proof. Let us consider the sequence (γ_N) of Section 5.1. Note that the L^2 norm of

$$u_0^N(v):=\int_{D\times\mathbb{T}}\gamma_N(v-v')\mu_0(\mathrm{d}x',\mathrm{d}v')=\int_{\mathbb{T}}\gamma_N(v-v')(\pi_v\mu_0)(\mathrm{d}v')$$

is bounded uniformly in N.

From Section 4 we know that the family $\left\{\mathbb{Q}^N\right\}$ is tight. Let $\left\{\mathbb{Q}^{N_k}\right\}$ be a weakly convergence subsequence, with limit $\mathbb{Q}$. We are going to prove below that $\mathbb{Q}$ is supported by the set of weak measure-valued solutions of equation (16), with initial condition μ^0. This implies existence of at least one such solution. Uniqueness has been proved in Section 3; recalling Lemma 5.5, we then immediately have claim (i). Moreover, we also have $\mathbb{Q}=\delta_\mu$, by the uniqueness of weak measure-valued solutions; therefore, since any weak limit point of $\left\{\mathbb{Q}^N\right\}$ is the same measure δ_μ, it follows that the full sequence $\left\{\mathbb{Q}^N\right\}$ converges weakly to δ_μ. Since S^N converges in law to a constant, it also converges in probability (in the topology of $\mathscr{C}$).

It remains to prove the claim made above that $\mathbb{Q}$ is supported by the set of weak measure-valued solutions of equation (16) with initial condition μ^0, i.e., the following set (cf. (32))

$$\Sigma=\left\{\mu\in\mathscr{C}:\Phi_\phi^{\mu^0}(\mu)=0\text{ for all }\phi\in\mathcal{T}\right\}$$

is a Borel subset of $\mathcal{C}$ and $\mathbb{Q}(\Sigma)=1$. Arguing as in Remark 2 it is not difficult to show that given $\phi\in\mathcal{T}$ there exists a sequence $(\phi_n)\subset C_c^2(D\times\mathbb{T})$ (i.e. ϕ_n is a C^2 function with compact support) such that $\|\phi_n\|_\infty+\|\partial_v\phi_n\|_\infty+\|\partial_{vv}^2\phi_n\|_\infty\le M$ (for some $M>0$ independent of n) and $\phi_n(z)\to\phi(z)$, $\partial_v\phi_n(z)\to\partial_v\phi(z)$, $\partial_{vv}^2\phi_n(z)\to\partial_{vv}^2\phi(z)$ as

$n \to \infty$, for any $z = (x, v) \in D \times \mathbb{T}$. Hence by the dominated convergence theorem we have

$$\Sigma = \left\{ \mu \in \mathscr{C} : \Phi_\phi^{\mu^0}(\mu) = 0 \text{ for all } \phi \in C_c^2(D \times \mathbb{T}) \right\}.$$

Moreover, since there exists a countable set $H_0 \subset C_c^2(D \times \mathbb{T})$ such that for any $\phi \in C_c^2(D \times \mathbb{T})$ we can find a sequence $(\phi_k) \subset H_0$ satisfying

$$\lim_{k\to\infty} \left(\|\phi - \phi_k\|_\infty + \|\partial_v \phi_k - \partial_v \phi\|_\infty + \|\partial_{vv}^2 \phi_k - \partial_{vv}^2 \phi\|_\infty \right) = 0,$$

we obtain that $\Sigma = \{\mu \in \mathscr{C} : \Phi_\phi(\mu) = 0 \text{ for all } \phi \in H_0\}$ which is a Borel subset of $\mathcal{C}$. To finish the proof we need to prove that

$$\mathbb{Q}\left(\mu \in \mathscr{C} : \Phi_\phi^{\mu^0}(\mu) = 0\right) = 1\,,$$

for every $\phi \in H_0$. Since $\Phi_\phi^{\mu^0} \geq 0$, it is equivalent to prove $\int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}\mathbb{Q} = 0$, for every $\phi \in H_0$. Due to Lemma 5.4, it is enough to show

$$\lim_{N\to\infty} \int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}\mathbb{Q}^N = 0 \tag{38}$$

for every $\phi \in H_0$. Using identity (6) we have

$$\begin{aligned}\int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}\mathbb{Q}^N &= \mathbb{E}\Bigg[\sup_{t\in[0,T]} \Bigg| \left\langle S_t^N, \phi \right\rangle - \left\langle \mu^0, \phi \right\rangle - \int_0^t \left\langle S_s^N, b\left(S_s^N\right) \partial_v \phi \right\rangle \, \mathrm{d}s \\ &\qquad - \int_0^t \left\langle S_s^N, \frac{(\sigma_2^\epsilon)^2}{2} \partial_v^2 \phi \right\rangle \, \mathrm{d}s \Bigg| \wedge 1 \Bigg] \\ &= \mathbb{E}\left[\sup_{t\in[0,T]} \left| M_t^{N,\phi} \right| \wedge 1 \right] \leq \mathbb{E}\left[\sup_{t\in[0,T]} \left| M_t^{N,\phi} \right|^2 \right]^{1/2}\end{aligned}$$

where

$$M_t^{N,\phi} = \int_0^t \frac{1}{N} \sum_{i=1}^N \sigma_2^\epsilon\left(V_s^{i,N}\right) \partial_v \phi\left(X_0^i, V_s^{i,N}\right) \mathrm{d}B_s^i.$$

Therefore, from Doob's inequality and the boundedness of σ_2^ϵ and $\partial_v \phi$, for some constants generically denoted by $C > 0$, we have

$$\left(\int_{\mathscr{C}} \Phi_\phi^{\mu^0} \, \mathrm{d}\mathbb{Q}^N \right)^2 \leq C \int_0^T \frac{1}{N^2} \sum_{i=1}^N \mathbb{E}\left[\left| \sigma_2^\epsilon\left(V_s^{i,N}\right) \partial_v \phi\left(X_0^i, V_s^{i,N}\right) \right|^2 \right] \mathrm{d}s \leq \frac{C}{N}$$

which implies (38) and completes the proof. □

6. **The McKean-Vlasov SDE.** Similarly to what is done in Section 2, let B be a standard real Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$, let η be a $\mathbb{T}$-valued random variable independent of B with density $\tilde{\rho}_0 \in L^2(0,2)$ and denote by $\mathcal{G}_t^0$ the augmentation of $\sigma(B_s, 0 \leq s \leq t) \vee \sigma(\eta)$ with the $\mathbb{P}$-null sets. Let then ξ be a random variable with values in D and law ν, having finite first moment, independent of $\mathcal{G}_t^0$ for every t. Finally denote by $\mathcal{G}_t$ the completion of $\mathcal{G}_t^0 \vee \sigma(\xi)$.
Let us consider the so-called McKean-Vlasov SDE associated with the system (10)

$$\begin{cases} \mathrm{d}V_t = \lambda_2(V_t)\,\mathrm{d}t + \int_{D\times\mathbb{T}} g_2(\xi, V_t, y, w)\mu_t(\,\mathrm{d}y,\,\mathrm{d}w)\,\mathrm{d}t + \sigma_2^\epsilon(V_t)\,\mathrm{d}B_t \\ \mathrm{d}X_t = 0 \\ (X_0, V_0) = (\xi, \eta) \\ \mu_t = \mathcal{L}(X_t, V_t)\,. \end{cases} \tag{39}$$

In analogy with Definition 2.1, we say that a strong solution to equation (39) is a family of continuous $\mathcal{G}_t^0$-adapted processes (V_t^x, μ_t^x), $x \in D$, with values in $\mathbb{R} \times \mathrm{Pr}_1(D \times \mathbb{T})$ such that the mapping: $(x, \omega) \mapsto (V_\cdot^x(\omega), \mu_\cdot^x(\omega))$ is measurable $(D \times \Omega, \mathcal{B}(D) \times \mathcal{F}) \to (\mathcal{S} \times \mathscr{C}, \mathcal{B}(\mathcal{S}) \times \mathcal{B}(\mathscr{C}))$ and for ν-almost every $x \in D$ the process (V^x, μ^x) satisfies (39) for $\xi = x$. If (V^x, μ^x) is a strong solution then $(\xi, V_t^\xi, \mu_t^\xi = \mathcal{L}(\xi, V_t^\xi))$ is a well-defined continuous $\mathcal{G}_t$-adapted process with values in $D \times \mathbb{R} \times \mathrm{Pr}_1(D \times \mathbb{T})$ which satisfies equations (39) $\mathbb{P}$-almost surely for any $t \in [0, T]$.

Theorem 6.1. *There exist a unique strong solution (V, μ) to equation (39).*

Proof. Fix a measure $\tilde{\mu}_\cdot \in \mathscr{C}$ and set

$$b(\tilde{\mu}_t)(x, v) := \lambda_2(v) + \int_{\mathbb{R}^3} \int_{\mathbb{R}} g_2(x, v, y, w) \tilde{\mu}_t(\mathrm{d}w, \mathrm{d}y) \, .$$

By the same arguments as in Section 2 it can be proved that there exist a unique strong solution $(V^{\tilde{\mu}, x})_x$ to the SDE

$$\begin{cases} \mathrm{d}V_t^{\tilde{\mu}, x} = b(\tilde{\mu}_t)(x, V_t^{\tilde{\mu}, x}) \, \mathrm{d}t + \sigma_2^\epsilon(V_t^{\tilde{\mu}, x}) \, \mathrm{d}B_t \, , \\ V_0^{\tilde{\mu}, x} = \eta \, . \end{cases} \tag{40}$$

Hence for X_0 as in (39) we have that $V^{\tilde{\mu}} := V^{\tilde{\mu}, X_0}$ satisfies

$$\begin{cases} \mathrm{d}V_t^{\tilde{\mu}} = b(\tilde{\mu}_t)(X_0, V_t^{\tilde{\mu}}) \, \mathrm{d}t + \sigma_2^\epsilon(V_t^{\tilde{\mu}}) \, \mathrm{d}B_t \, , \\ V_0^{\tilde{\mu}} = \eta \, . \end{cases} \tag{41}$$

Now choose as $\tilde{\mu}$ the unique weak solution in $\mathscr{C}$ to the nonlinear Fokker-Planck PDE (16) with initial condition $\tilde{\mu}^0 = \nu \times \tilde{\rho}_0 \mathscr{L}_{\mathbb{T}}$; given the associated process $V^{\tilde{\mu}}$ as above, denote by μ_t the law of the vector $(\xi, V_t^{\tilde{\mu}})$. Then μ is a solution in $\mathscr{C}$ to the linear PDE

$$\begin{cases} \frac{\partial}{\partial t} \mu_t = \frac{1}{2} \frac{\partial^2}{\partial v^2} \left((\sigma_2^\epsilon)^2 \, \mu_t \right) - \frac{\partial}{\partial v} \left(\mu_t b(\tilde{\mu}_t) \right) \\ \mu_0 = \tilde{\mu}^0 \, . \end{cases}$$

Since there is uniqueness of measure-valued solution to the latter (the proof being a simplified version of that of uniqueness in the nonlinear case; see Theorem 3.2), and clearly also $\tilde{\mu}$ is a solution, we necessarily have $\tilde{\mu} = \mu$ and (V, μ) is a strong solution to (39).

Let now $(\overline{V}, \bar{\mu})$ be another solution. Then $\bar{\mu}^x$ solves the Fokker-Planck equation (16) with initial condition $\delta_x \times \tilde{\rho}_0 \mathscr{L}_{\mathbb{T}}$, for ν-almost every $x \in D$, and therefore $\bar{\mu}^x = \mu^x$. Finally $V_t^x = \overline{V}_t^x$ a.s. for every t and ν-almost every $x \in D$, since they both satisfy a SDE like (41) with $\tilde{\mu} = \mu^x$ for which strong uniqueness holds. Hence (V, μ) is the unique solution to (39). □

Appendix A. **Appendix: Extension of some results.** We state here some further results that are easy generalizations of what we presented so far; we will comment on the proofs when needed.

Notice that in Section 2 neither the particular form of the coefficients nor the fact that they are 2-periodic plays any role. Since the cited results we built our proof on apply to multidimensional SDE with bounded and measurable drift, we immediately obtain the following theorem, with proof identical to that of Theorem 2.2. Here, similarly to what done previously, for $k, l, m, n \in \mathbb{N}$ we consider a m-dimensional Brownian motions $\mathbf{W}$ and independent random variables $\Xi = \left(\Xi^j\right)_{j=1,\dots,n}$ and $\Psi = \left(\Psi^j\right)_{j=1,\dots,n}$ (independent from $\mathbf{W}$ as well) with values in $E^n \subseteq \left(\mathbb{R}^k\right)^n$ and $\left(\mathbb{R}^l\right)^n$, respectively ($E$ is an open subset of $\mathbb{R}^k$), all defined on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with finite first moment.

We assume that the Ξ^j's are identically distributed with law $\boldsymbol{\nu}$ and the Ψ^j's are identically distributed with absolutely continuous law $\boldsymbol{\rho}_0 \mathscr{L}_{\mathbb{R}^l}$, with $\boldsymbol{\rho}_0 \in L^2(\mathbb{R}^l)$. We define the σ-algebra $\mathcal{E} = \sigma(\Xi)$, the filtration $(\mathcal{A}_t^0)_t$ as the augmentation of $\sigma(\mathbf{W}_s, 0 \le s \le t) \vee \sigma(\Phi)$ with the $\mathbb{P}$-null sets and the filtration $(\mathcal{A}_t)_t$ as the completion of $\mathcal{A}_t^0 \vee \mathcal{E}$. Finally we fix $T > 0$.

Theorem A.1. *Let $\mathbf{b} : [0,T] \times E^n \times \mathbb{R}^{l\times n} \times E \times \mathbb{R}^l \to \mathbb{R}^{l\times n}$ and $\boldsymbol{\sigma} : [0,T] \times \mathbb{R}^{l\times n} \to \mathbb{R}^{l\times n\times m}$ be bounded and Borel measurable functions. Assume that $\boldsymbol{\sigma}\boldsymbol{\sigma}^\top(t,\cdot)$ is uniformly elliptic and that $\boldsymbol{\sigma}(t,\cdot)$ is Lipschitz; both properties have to be satisfied uniformly in t. Then there exists a unique strong solution $\mathbf{Y} = \left(Y^j\right)_{j=1,\dots,n}$, in the sense of Definition 2.1 and Theorem 2.2, to the SDE*

$$\begin{cases} \mathrm{d}\mathbf{Y}_t = \left\langle \mathbf{b}\left(t, \Xi, \mathbf{Y}_t, \cdot, \cdot\right), \overline{S}_t^n \right\rangle \mathrm{d}t + \boldsymbol{\sigma}(t, \mathbf{Y}_t)\, \mathrm{d}\mathbf{W}_t\,, \qquad t \in [0,T]\,, \\ \mathbf{Y}_0 = \Psi \end{cases} \tag{42}$$

where

$$\overline{S}_t^n = \frac{1}{n}\sum_{j=1}^n \delta_{(\Xi^j, Y^j)}\,.$$

Finer refinements are possible: for example one can treat the cases when $T = \infty$ and the SDE is to be solved on a domain $U \subset \mathbb{R}^{l\times n}$, and the assumptions on $\boldsymbol{\sigma}$ can be weakened (cf. Remark 1). We refer to [8] for details.

Also when studying our Fokker-Planck PDE the periodicity of the coefficients plays no particular role, nor does the compactness of the torus $\mathbb{T}$.
We need anyway some more assumptions on the coefficients $\mathbf{b}(t,x,y,x',y')$, $t \in [0,T]$, $x \in E^n$, $y \in \mathbb{R}^{t\times n}$, $x' \in E$, $y' \in \mathbb{R}^l$, and $\boldsymbol{\sigma}(t,y)$ above, namely:

(E1) $\mathbf{b}$ does not depend on t, is bounded and uniformly continuous in x and x' (uniformly in y and y') and the set of discontinuities of the map

$$(y,y') \mapsto \mathbf{b}(x,y,x',y')$$

has Lebesgue measure 0 in $\mathbb{R}^{l\times n} \times \mathbb{R}^l$, for any x, x';

(E2) $\boldsymbol{\sigma}$ does not depend on t and belongs to $C_b^1(\mathbb{R}^{l\times n}; \mathbb{R}^{l\times n\times m})$.

One can repeat the arguments of Sections 3, 4, 5 with minor modifications, working in the space $\mathrm{Pr}_1(E \times \mathbb{R}^l)$ with the 1-Wasserstein metric, using the Euclidean norm in place of the metric $d_{D\times\mathbb{T}}$ and choosing all the test functions accordingly. If we solve equation (42) above for $\mathbf{Y}$ and consider the empirical measure $\overline{S}^n$, we can define the empirical density $\bar{u}^n$ as

$$\bar{u}_t^n(y) = \int_{E\times\mathbb{R}^l} \bar{\gamma}_n(y-y')\overline{S}_t^n(\,\mathrm{d}x',\,\mathrm{d}y')$$

where $(\bar{\gamma}_n)$ is a family of compactly supported mollifiers in $\mathbb{R}^l$.
We say that $f \in L^2_{\mathrm{loc}}(\mathbb{R}^l)$ if $f \in L^2(K)$ for every compact set $K \subset \mathbb{R}^l$. Consider a strictly increasing sequence $(P_j)_{j\in\mathbb{N}}$ of compact sets in $\mathbb{R}^l$ such that $\mathbb{R}^l = \cup_j P_j$; then

$$d(f,g) = \sum_{j\ge 1} \frac{1}{2^j} \frac{\|f-g\|_{L^2(P_j)}}{1+\|f-g\|_{L^2(P_j)}}$$

is a metric on $L^2_{\mathrm{loc}}(\mathbb{R}^l)$. Lemmata 5.1 and 5.2 apply to $\bar{u}^n$'s as well thanks to assumption (E2), and imply tightness of their laws in the space $L^2\left(0,T; L^2_{\mathrm{loc}}\left(\mathbb{R}^l\right)\right)$ due to a generalized version of Aubin-Lions lemma, which claims that the space

$$L^2\left(0,T; W^{1,2}\left(\mathbb{R}^l\right)\right) \cap W^{\alpha,2}\left(0,T; H^{-2}\left(\mathbb{R}^l\right)\right)$$

is relatively compact in $L^2\left(0,T;L^2_{loc}\left(\mathbb{R}^l\right)\right)$.
Let $\mathcal{C} := C\left([0,T];\mathrm{Pr}_1\left(E\times\mathbb{R}^l\right)\right)$. The space $\widetilde{X}$ of Subsection 5.2 has to be consequently substituted with

$$\overset{\diamond}{X} = \left\{\mu\in\mathcal{C}\ :\ \pi_v\mu_t \ll \mathscr{L}_{\mathbb{R}^l} \text{ with } \frac{\mathrm{d}\left(\pi_v\mu_t\right)}{\mathrm{d}\mathscr{L}^l_{\mathbb{R}}}\in L^2(\mathbb{R}^l),\ \text{ for a.e. } t\in[0,T]\right\}.$$

To show that $\overset{\diamond}{X}$ is Borel it is enough to repeat the argument in the proof of Lemma 5.3 using the density of $C_c(\mathbb{R}^l)$ in $L^2(\mathbb{R}^l)$. The functionals $\Phi^{\mu^0}_\phi$ are now continuous on $\overset{\diamond}{X}$, because the proof of Lemma 5.4 can be repeated thanks to assumption (E1). To show that any limit point of the family of laws of the $\overline{S}^n$ gives full measure to $\overset{\diamond}{X}$ one can repeat the proof of Lemma 5.5, noting that it is enough to check identity (36) for $\phi\in C_b(\mathbb{R}^l)$.
Now we can repeat verbatim the arguments in the proof of Theorem 5.6 obtaining:

Theorem A.2. *Let $\zeta^0 = \boldsymbol{\nu}\times\boldsymbol{\rho_0}\mathscr{L}_{\mathbb{R}^l}$. Then:*

i) the nonlinear Fokker-Planck equation

$$\partial_t\zeta_t + \mathrm{div}_y\left(\zeta_t\langle\mathbf{b},\zeta_t\rangle\right) = \mathrm{Tr}\left[D^2_y\left(\frac{\boldsymbol{\sigma}\boldsymbol{\sigma}^{\boldsymbol{T}}}{2}\zeta_t\right)\right] \tag{43}$$

with initial condition ζ^0, has one and only one weak measure-valued solution ζ; this measure belongs to the space $\overset{\diamond}{X}$.

ii) Let $\mathbb{Q}^n$ be the laws on $\mathcal{C}$ of the empirical process S^n; then $\mathbb{Q}^n$ converges weakly to δ_ζ. Further S^n converges in probability to ζ, in the topology of $\mathcal{C}$.

The extension fo the well-posedness result for the McKean-Vlasov equation given in Section 6 is then straightforward. For related results on strong well-posedness for McKean-Vlasov equations (without dependence on stochastic parameters) see [13] and references therein.

References

[1] M. J. Cáceres, J. A. Carrillo and B. Perthame, Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states, *J. Math. Neurosci.*, **1(1)** (2011), 7–33.

[2] M. J. Cáceres and B. Perthame, Beyond blow-up in excitatory integrate and fire neuronal networks: refractory period and spontaneous activity, *J. Theoret. Biol.*, **350** (2014), 81–89.

[3] F. Delarue, J. Inglis, S. Rubenthaler and E. Tanré, Particle systems with a singular mean-field self-excitation. Application to neuronal networks, *Stochastic Processes and their Applications*, **125(6)** (2015), 2451–2492.

[4] F. Delarue, J. Inglis, S. Rubenthaler and E. Tanré, Global solvability of a networked integrate-and-fire model of McKean-Vlasov type, *Ann. Appl. Probab.*, **25(4)** (2015), 2096–2133.

[5] A. De Masi, A. Galves, E. Löcherbach and E. Presutti, Hydrodynamic limit for interacting neurons, *J. Stat. Phys.*, **158(4)** (2015), 866–902.

[6] A. Friedman, **Stochastic differential equations and applications. Vol. 1**, Probability and Mathematical Statistics, Vol. 28, Academic Press, New York-London, 1975.

[7] N. Fourcaud-Trocmé and N. Brunel, Dynamics of the instantaneous firing rate in response to changes in input statistics, *J. Comput. Neurosci.*, **18(3)** (2005), 311–321.

[8] I. Gyöngy and N. Krylov, Existence of strong solutions for Itô's stochastic equations via approximations, *Prob. Theory Relat. Fields*, **105** (1996), 143–158.

[9] W. Gerstner and W. M. Kistler. *Spiking neuron models*. Cambridge University Press, Cambridge, 2002.

[10] Daniel Henry. *Geometric theory of semilinear parabolic equations*, vol. 840 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1981.

[11] B. W. Knight, D. Manin and L. Sirovich, Dynamical models of interacting neuron populations in visual cortex, *Robot Cybern*, **54**(1996), 4–8.

[12] T. J. Lewis and J. Rinzel, Dynamics of spiking neurons connected by both inhibitory and electrical coupling. *J. Comput. Neurosci.*, **14(3)** (2003), 283–309.

[13] Y. S. Mishura and A-Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations, arXiv:1603.02212.
[14] S. Ostojic, N. Brunel and V. Hakim, Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities, *J. Comput. Neurosci.*, **26(3)** (2009), 369–392.
[15] Karl Oelschläger, On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes, *Probability Theory and Related Fields*, **82(4)** (1989), 565–586.
[16] T. Schwalger, M. Deger and W. Gerstner, Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size, *PLoS Comput. Biol.*, **13(4)** (2017).
[17] J. Simon, Compact sets in the space $L(0,T;B)$, *Ann. Mat. Pura. Appl.*, **146(4)** (1987), 65–96.
[18] A.-S. Sznitman, *Topics in propagation of chaos*, vol. 1464 of *Lecture Notes in Mathematics.*, Springer-Verlag, Berlin, 1991.
[19] H. C. Tuckwell, *Introduction to theoretical neurobiology. Vol. 1*, vol. 8 of *Cambridge Studies in Mathematical Biology*, Cambridge University Press, Cambridge, 1988.
[20] A. J. Veretennikov, Strong solutions and explicit formulas for solutions of stochastic integral equations, *Math. USSR-Sb. (N.S.)*, **39(3)** (1981), 487–403.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: franco.flandoli@sns.it
E-mail address: enrico.priola@unito.it
E-mail address: gzanco@luiss.it