Stochastic interacting particle systems: from applications to theory

Hélène Guérin
Professor at Université du Québec à Montréal (UQAM)
heleneguerin.github.io

Lecture at Kyoto University

25 May – 5 June 2026

Picture credit: Romane Charpentier (kits.math)

Abstract

One of the objectives of this lecture is to build a bridge between probability theory and applications. We will start from various questions, mainly related to our environment or industrial questions, model them using random variables, and then use probability theory to study them.

This lecture does not aim to be exhaustive, as probabilistic models and applications are highly diverse, and there are numerous probabilistic methods for studying them. We will focus on certain models of interacting particles and on the concept of propagation of chaos. We will review some key concepts related to the convergence of processes in order to study two types of stochastic particle systems in detail.

- Lecture 1: Models of interacting particle systems in various fields.

  Exercise 1, page 10
- Lecture 2: Kac's propagation of chaos.
- Lecture 3: Convergence of stochastic processes: tightness, Aldous' criterion, and so on.

  Exercise 2, page 26
- Lecture 4: Analysis of a diffusive interacting particle system.

  Exercise 3, page 28
  Exercise 4, page 32
- Lecture 5: Analysis an interacting particle system with jumps.

  Exercise 5, page 37

For students who wish to receive credit for this course, the grade is divided into two parts:

- Attendance at class sessions: 10 points per lecture (please remember to sign the attendance sheet at each class);
- Submission of written assignments (at least 4 out of the 5 exercises): 50 points.

Assignments must be submitted by email to guerin.helene@uqam.ca (please make sure to include your **student ID number** both in your signature and in the names of your PDF files). The deadline for submitting assignments is **Tuesday, June 23, at 4:00 p.m.**

Please send your solutions, in English, as PDF files, either generated from a LaTeX source file or compiled from photos of your handwritten (legible) solutions. Only one PDF file per assignment is allowed (so if you use photos, please combine them into a single PDF file).

The grade will primarily reflect the quality of your writing and your ideas. Therefore, do not worry if you are unable to completely solve an exercise.

If you have any questions, please feel free to email me.

Contents

Acknowledgments: I would like to sincerely thank Kyoto University and the Research Institute for Mathematical Sciences for their wonderful welcome during my sabbatical year. In particular, I am deeply grateful to Professor Benoît Collins, Professor David Croydon, and Professor Naotaka Kajino.

A FEW NOTATIONS:

- $X \sim \mu$ means that the random variable follows the probability distribution μ: $\mathrm{Law}(X) = \mu$;
- Δ and ∇ represent, respectively, the Laplacian operator and the gradient;
- ∂_t and ∂_x denote the partial derivatives with respect to time t and space x, respectively;
- the indicator function $\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$;
- $\mathcal{P}(E)$ denotes the set of probability measures on a Polish space E;
- $\mathcal{C}_b(E)$ denotes the space of bounded and continuous functions on E;
- $\langle \mu, \varphi \rangle = \int_E \varphi \mathrm{d}\mu$ for $\mu \in \mathcal{P}(E)$ and φ a test function on E;
- $\|\varphi\|_\infty = \sup_{x \in E} |\varphi(x)|$ for a function defined on E;
- $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^d$;
- $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ denotes the space of continuous function defined on $\mathbb{R}_+$ with values in $\mathbb{R}^d$;
- $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ denotes the Skorokhod space, i.e. the space of *càdlàg* functions (right-continuous functions with left limits), defined on $\mathbb{R}_+$ with values in $\mathbb{R}^d$;
- $z(t-) = \lim\limits_{\substack{s \to t \\ s < t}} z(s)$ is the left limit of a function $z \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ in $t \in \mathbb{R}_+$;
- $t \wedge s = \min(t, s)$ for $t, s \in \mathbb{R}$.

1 Models of interacting particle systems in various fields

One of the objectives of this lecture is to build a bridge between probability theory and applications. We will start from various questions, mainly related to our environment or to industry issues, model them using random variables, and then use probability theory to study them.
This lecture does not aim to be exhaustive, as probabilistic models and applications are highly diverse, and there are numerous probabilistic methods for studying them. We will focus on certain models of interacting particles and on the concept of propagation of chaos. The term "particle" can refer to gas molecules, individuals, bacteria, animals, neurons, stock prices, and so on.

The terminology "propagation of chaos" comes from Kac and the idea is the following. We consider system of particles of size N that is initially chaotic, in the sense that particles are initially independent and identically distributed (i.i.d.), but then evolve through interaction. Thus, at each time, the particles are no more independent.
We say that there is propagation of chaos if, when we consider a fixed finite number of particles, and let the size of the system N go to infinity, those selected particles evolve independently in the limit, all with the same distribution. This implies that initially two particles are independent; then, due to the interaction, they become dependent, but when the size of the particle system is infinite, those two particles evolve independently ("chaotically"), according to a common distribution.

1.1 The McKean-Vlasov equation in physics

We start with the laboratory example described by Sznitman in his Saint Flour's lecture, entitled TOPICS IN PROPAGATION OF CHAOS, [Szn91].
The initial motivation for the subject was to try to investigate the connection between a detailed and a reduced description of particles' evolution, in a gas where interactions are pairwise such as in the Boltzmann equation and in the Landau equation.
The laboratory example proposed by Sznitmann is much simpler to study than the Boltzmann case. It is a model of interacting diffusions due to McKean. Let us consider a system of N particles in $\mathbb{R}^d$, with initial chaotic distribution $\mu_0^{\otimes N}$, satisfying the Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t^{i,N} = \sqrt{2}\mathrm{d}B_t^i + \frac{1}{N}\sum_{j=1}^N b(X_t^{i,N}, X_t^{j,N})\mathrm{d}t \tag{1.1}$$

for $1 \leqslant i \leqslant N$, where $B = (B^1, \ldots, B^N)$ is a standard N-dimensional Brownian motion and b is a regular compactly supported function (ensuring the existence and uniqueness of such a system). The particles evolve according to independent Brownian motions and interact through their drift terms. Introducing the empirical distribution $\overline{\mu}^N$ of the particle system: for $t \geqslant 0$,

$$\overline{\mu}_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}},$$

Equation (1.1) can be written

$$\mathrm{d}X_t^{i,N} = \sqrt{2}\mathrm{d}B_t^i + \int b(X_t^{i,N}, y)\mu_t^N(\mathrm{d}y).$$

Assuming that the initial particles are i.i.d. with a common distribution μ_0, as will be shown later, propagation of chaos holds as $N \to \infty$, and the common limiting distribution is the law of the solution, started from $X_0 \sim \mu_0$, to the *nonlinear* SDE

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}B_t + \int b(X_t, y)\mu_t(\mathrm{d}y)\mathrm{d}t,$$

where B is a standard one-dimensional Brownian motion and μ_t denotes the distribution of X_t. We say that the SDE is *nonlinear* because its coefficients depend on the distribution of the solution.
If $\mu_0(\mathrm{d}x) = u_0(x)\mathrm{d}x$ has a density with respect to the Lebesgue measure, then one can prove that μ_t also admits a density $u(t, \cdot)$ and that u is a weak solution to the *nonlinear* Partial Differential Equation (PDE)

$$\begin{cases} \partial_t u = \Delta u - \mathrm{div}\big(\int b(x, y)u(t, y)\mathrm{d}y\, u\big) \\ u(0, \cdot) = u_0(\cdot), \end{cases}$$

where $u(t, x)$ is the density of the particles in position x at time t. This system of interacting particles has been the subject of several studies, and in Section 4 we will investigate the probabilistic approach developed in [BRTV98, Mal03].

1.2 Atlas model in finance

The Atlas model has been introduced by Fernholz, [Fer02], to model equity markets in finance.
In the initial description of the model, the stock prices were modeled as Brownian motion except for the smaller stock price, which has an added positive drift that drives the market.
Formally, let us consider $B = (B^1, \ldots, B^N)$ a N-dimensional Brownian motion, the value of each stock price is solution to the SDE

$$\mathrm{d}X_t^{i,N} = \mathrm{d}B_t^i + b\mathbb{1}_{X_t^{i,N} = \min_{1\leqslant k\leqslant N}\left(X_t^{k,N}\right)}\mathrm{d}t,$$

with b a positive real number. As in the myth of Atlas, the lowest process pushes the others up. Because the growth of the whole portfolio is supported only by this smallest stock, see Figure 1, the model is named after the Titan Atlas, eternally holding up the sky. These rank-based stochas-

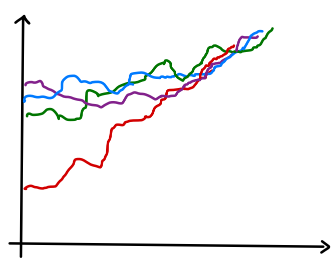

Figure 1: Illustration of trajectories of the Atlas model.

tic models successfully capture an empirical feature of real financial markets, namely the so-called small-cap premium. More precisely, in rank-based financial models, each company is ordered according to its market capitalization. The dynamics of a stock then depend on its rank in the market. The phenomenon of small-cap premium refers to the observation that, over long time periods, stocks of smaller companies tend to have higher average returns than stocks of larger companies.

This model has then been generalized under the following form

$$\mathrm{d}X_t^{i,N} = \sum_{k=1}^{N} \sigma_k \mathbb{1}_{X_t^{i,N}=X_t^{(k)}} \mathrm{d}B_t^i + \sum_{k=1}^{N} b_k \mathbb{1}_{X_t^{i,N}=X_t^{(k)}} \mathrm{d}t \tag{1.2}$$

where $\sigma = (\sigma_1, \dots, \sigma_N) \in \mathbb{R}_+^N$, $b = (b_1, \dots, b_N) \in \mathbb{R}^N$, and $X^{(1)}, X^{(2)}, \dots, X^{(N)}$ are the order statistics of the N-tuple $(X^{1,N}, X^{2,N}, \dots, X^{N,N})$: for all $t \geqslant 0$,

$$\min_{1 \leqslant i \leqslant N} X_t^{i,N} = X_t^{(1)} \leqslant X_t^{(2)} \leqslant \dots \leqslant X_t^{(N)} = \max_{1 \leqslant i \leqslant N} X_t^{i,N},$$

with i.i.d initial conditions $X_0^{1,N}, \dots, X_0^{N,N}$ of law μ_0. The existence and uniqueness of such a particle system have been studied in several articles. The weak existence and uniqueness when $\sigma \in (\mathbb{R}_+ \setminus \{0\})^N$ is a consequence of [BP87]. The strong existence and uniqueness were far more challenging to establish. They were first proved in the planar case ($N = 2$) in [FIKP13], and then extended to the high-dimensional setting up to the first triple collision in [IKS13], that is, before the time $\tau = \inf \left\{ t \geqslant 0 : \exists (i,j,k) \in \{1, \dots, N\}^3 \text{ distinct such that } X_t^{i,N} = X_t^{j,N} = X_t^{k,N} \right\}$. In [IKS13], a condition on σ is also provided to ensure that triple collisions cannot occur.

In two independent works, [JR13, KS18], propagation of chaos for the system (1.2) was proved under the following assumption:

There exist $\sigma : [0,1] \to (0, +\infty)$ and $b : [0,1] \to \mathbb{R}$ continuous functions such that

$$\sigma_k = \sigma\left(\frac{k}{N}\right) \quad \text{and} \quad b_k = b\left(\frac{k}{N}\right).$$

Introducing the empirical cumulative distribution

$$F_t^N(x) = \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{\left\{X_t^{k,N} \leqslant x\right\}},$$

we observe that Equation (1.2) can be written as

$$\mathrm{d}X_t^{i,N} = \sigma\Big(F_t^N(X_t^{i,N})\Big)\mathrm{d}B_t^i + b\Big(F_t^N(X_t^{i,N})\Big)\mathrm{d}t.$$

There is propagation of chaos when $N \to \infty$, and the common distribution is the law of the solution starting from $X_0 \sim \mu_0$ to the *nonlinear* SDE

$$\mathrm{d}X_t = \sigma(F_t(X_t))\mathrm{d}B_t + b(F_t(X_t))\mathrm{d}t,$$

where B is a standard one-dimensional Brownian motion and $F_t(x) = \mathbb{P}(X_t \leqslant x)$ is the cumulative distribution function of X_t. It was also proved that the cumulative distribution function $F(t, \cdot) :=$

$F_t(\cdot)$ of X_t is the weak solution of the nonlinear PDE

$$\begin{cases} \partial_t F = \frac{1}{2}\partial_x^2(A(F)) - \partial_x(B(F)) \\ F(0,x) = F_0(x) \end{cases}$$

where $A(u) = \int_0^u \sigma^2(v)\mathrm{d}v$, $B(u) = \int_0^u b(v)\mathrm{d}v$, and $F_0(x) = \mu_0((-\infty, x])$.

1.3 Neural dynamics in neurosciences

The neurons in our nervous system communicate with each other through an electrical signal that propagates along their axons. This specific electrical signal is characterized by rapid depolarization (spike), followed by hyperpolarization, before returning to the resting membrane potential.

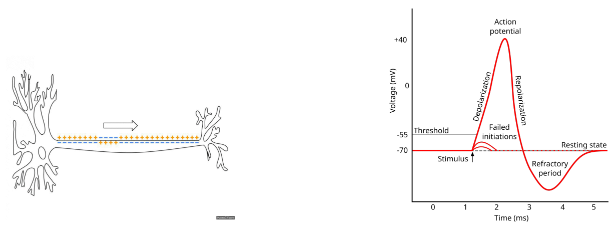

Figure 2: Action potential of a neuron (Source: Wikipedia).

We present the model introduced in [DMGLP15], which describes the time evolution of a population of N interacting neurons. Each neuron spikes randomly according to a point process whose rate depends on its membrane potential.
For $i \in \{1, \dots, N\}$, $X_t^{i,N}$ models the membrane potential of neuron i at time t. Neurons interact either by electrical, or by chemical synapses:

- Electrical synapses: they occur through gap-junctions which allow neurons in the brain to communicate directly. This induces an attraction between the values of the membrane potentials and, thus a deterministic drift of the system towards its average potential.
- Chemical synapses: each neuron spikes randomly following a point process with rate depending on the membrane potential of the neuron. At its spiking time, the membrane potential of the spiking neuron is reset to the value 0. At the same time, simultaneously, the other neurons, which do not spike, receive an additional amount of potential $\frac{1}{N}$, which is added to their membrane potential.

The generator of the particle system described above is thus given by, for any smooth test function $\varphi : \mathbb{R}_+^N \to \mathbb{R}$,

$$\mathcal{L}\varphi(x) = -\lambda \sum_{i=1}^N \left(\frac{\partial \varphi}{\partial x_i}(x)(x_i - \overline{x}) \right) + \sum_{i=1}^N f(x_i)\big(\varphi(x + \delta^i(x)) - \varphi(x)\big),$$

where $\delta^i(x) = (\delta^i_j(x))_{1 \leqslant j \leqslant N}$ with $\delta^i_j = \begin{cases} \frac{1}{N} & \text{if } j \neq i \\ -x_i & \text{if } j = i \end{cases}$, and $\overline{x} = \dfrac{1}{N}\displaystyle\sum_{i=1}^N x_i$. The quantities $\lambda \geqslant 0$ and the function f, called the spiking rate of the system, are parameters of the system.

The first term of the generator $\mathcal{L}$ describes the attraction of the neurons to the mean potential, and the second term describes jumps occurring at rate $f(x_i)$ due to the spiking of neurons with potential x_i. We assume that the function f is non-decreasing to model the phenomenon that the higher the potential, the more likely it is to jump to 0.
The evolution in time of the potential of each neuron can be also be described by a system of SDE driven by Poisson measures. Let $\left(Q^i(\mathrm{d}s, \mathrm{d}z)\right)_{i\geqslant N}$ be independent Poisson measures on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $\mathrm{d}s\mathrm{d}z$ modeling the spikes, and $(X_0^{i,N})_{1\leqslant i\leqslant N}$ be a sequence of i.i.d initial potentials, independent of the Poisson measures. Consequently, we write

$$\begin{aligned} X_t^{i,N} =& X_0^{i,N} - \lambda \int_0^t \left(X_s^{i,N} - \overline{X}_s^N\right)\mathrm{d}s \\ &- \int_0^t \int_{\mathbb{R}_+} X_{s-}^{i,N} \mathbb{1}_{\left\{z\leqslant f(X_{s-}^{i,N})\right\}} Q^i(\mathrm{d}s, \mathrm{d}z) + \frac{1}{N}\sum_{j\neq i} \int_0^t \int_{\mathbb{R}_+} \mathbb{1}_{\left\{z\leqslant f(X_{s-}^{j,N})\right\}} Q^j(\mathrm{d}s, \mathrm{d}z), \end{aligned}$$

where $\overline{X}_t^N = \frac{1}{N}\sum_{i=1}^N X_t^{i,N}$ is the mean potential of the system.
The asymptotic behavior this system is studied in [DMGLP15, FL16], and they state the propagation of chaos to the distribution of the solution X to the following nonlinear SDE

$$X_t = X_0 - \lambda \int_0^t (X_s - \mathbb{E}[X_s])\mathrm{d}s - \int_0^t \int_{\mathbb{R}_+} X_{s-}\mathbb{1}_{\{z\leqslant f(X_{s-})\}} Q(\mathrm{d}s, \mathrm{d}z) + \int_0^t \mathbb{E}[f(X_s)]\mathrm{d}s,$$

where $Q(\mathrm{d}s, \mathrm{d}z)$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $\mathrm{d}s\mathrm{d}z$.
From the PDE point of view, the density function $u(t, \cdot)$ of the distribution of X_t ($u(x,t)$ represents the density of neurons with membrane potential x at time t) satisfies the following equation:

$$\begin{cases} \partial_t u = (\lambda x - a_t)\partial_x u + (\lambda - f(x))u \\ u(t,0) = \frac{p_t}{a_t} \\ u(0,x) = u_0(x) \end{cases}$$

where $a_t = \int_0^\infty (f(x) + \lambda x)u(t,x)\mathrm{d}x$, $p_t = \int_0^\infty f(x)u(t,x)\mathrm{d}x$, and u_0 denotes the initial distribution. We study this system and its long lime behavior in detail in Section 5.

1.4 Epidemiologic models in medecine

We consider individual-based SIRS-type epidemiological models describing the spread of an infectious disease in which individuals can become infected, recover, and subsequently return to the susceptible state without acquiring total immunity. In such a population, susceptible individuals become infected through contact with infected individuals, while infected individuals recover at a given rate and immediately become susceptible again. We consider a closed population of size N (without immigration).
We assume that the interactions in the population are uniform between the individuals, and all the individuals behave in the same way to the disease.

Each individual is characterized by two random quantities that evolve over time: their infectivity, which determines the rate at which they transmit the disease, and their susceptibility, corresponding to the probability of being reinfected upon contact with an infected individual. The infectivity of

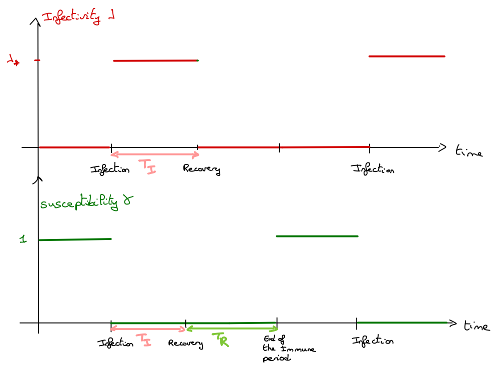

Figure 3: Evolution in time of the infectivity and susceptibility of an individual.

individual $i \in \{1, \dots, N\}$ at time t is denoted by $\lambda^{i,N}(t)$, while their susceptibility is denoted by $\gamma^{i,N}(t)$. Their value depends on the age $a^i(t)$ of individual i, namely the elapsed time since the last infection.

Each individual repeatedly alternates between two states: an infectious state (denoted by I), in which the individual cannot be infected but can spread the disease ($\lambda^{i,N}(t) \geqslant 0$ and $\gamma^{i,N}(t) = 0$), and a susceptible state (denoted by S), in which the individual does not spread the disease but can be reinfected ($\lambda^{i,N}(t) = 0$ and $\gamma^{i,N}(t) \geqslant 0$).

The individuals interact through the global infection $\mathfrak{F}^N(t)$ of the disease: individual i becomes infected at time t at rate $\gamma^{i,N}(t)\mathfrak{F}^N(t)$, where $\mathfrak{F}^N(t)$ is given by

$$\mathfrak{F}^N(t) = \frac{1}{N} \sum_{k=1}^{N} \lambda^{k,N}(t).$$

At each new infection, new infectivity and susceptibility curves are sampled for the infected individual, and its age is reset to 0, while the other individuals keep their own infectivity and susceptibility curves, as well as their ages.

In the classical SIRS model, the functions λ and γ are indicator functions of the age a since the last infection:

$$\lambda(a) = \lambda_* \mathbb{1}_{a \leqslant T_I}, \tag{1.3}$$

where T_I is the *random* duration of the infectious period and λ_* is a positive constant, and

$$\gamma(a) = \mathbb{1}_{a > T_R + T_I}, \tag{1.4}$$

where T_R is the *random* length of the immune period. Consequently, after recovery, the individual remains immune to the disease for a period of length T_R (and therefore cannot be infected), before becoming fully susceptible to the disease again (see Figure 3). The SIS model corresponds to the case where $T_R = 0$ almost surely. Models with waning immunity have recently been introduced and studied in [FPPZN25, GZN25].

■ **Exercise 1.** Consider a SIRS model for a population of size N.

The infectivity and susceptibility curves are given by (1.3) and (1.4), respectively. We assume that the *generic* distribution of the pair (T_I, T_R) admits a density f on $\mathbb{R}_+ \times \mathbb{R}_+$. Note that, in the

classical SIRS model, T_I and T_R are independent exponential random variables with respective rates $\nu > 0$ and $\delta > 0$.
The aim of this exercise is to write the system of SDEs describing the evolution of the disease, more precisely the infectious state of each individual. To this end, we introduce, for each individual, its age, infectious period, and immune period. For $i \in \{1, \dots, N\}$, $a^i(t)$ denotes the age of individual i at time $t \geqslant 0$, and $T^i(t) = (T^i_I(t), T^i_R(t))$ denotes the pair of infectious and immune durations associated with individual i at time t.
Let $a^i(0)$ and $T^i(0) = (T^i_I(0), T^i_R(0))$ denote their initial values.

1. Describe the evolution in time of $(a^i, T^i)_{1 \leqslant i \leqslant N}$ in your words.
2. Introducing independent Poisson measures $\left(Q^i\right)_{1 \leqslant i \leqslant N}$, write a system of SDEs and/or the infinitesimal generator satisfied by $(a^i, T^i)_{1 \leqslant i \leqslant N}$.
3. Without proof, give the SDE and/or the generator satisfied by the limiting distribution as the population size tends to infinity ($N \to \infty$).

1.5 A few interesting questions on such systems

One of the main objectives in the study of such systems is to analyze their long time behavior, that is, their behavior as $t \to \infty$: is there existence of an equilibrium in long time? For example, in epidemiology, a central question is to identify conditions ensuring the existence of an endemic equilibrium, that is, a regime in which the disease persists in the population (something that public health policies usually aim to prevent).
For a system of i.i.d. particles, studying a single particle is sufficient to characterize the long time behavior of the whole system. In contrast, when the particles interact, the analysis of the long time behavior becomes significantly more challenging, since the behavior of each particle is influenced by that of the others. Nevertheless, when propagation of chaos holds, the study of a large population can be reduced to the analysis of a *typical* particle, namely the solution of a nonlinear equation describing the limiting dynamics as the population size tends to infinity. This reduction is one of the main strengths of the propagation of chaos property.

2 Kac's propagation of chaos

This section is mainly based on the Saint-Flour's lecture of Sznitman [Szn91], and on the recent surveys of Chaintron and Diez [CD22a, CD22b]. Please note that we refer in this document to the arXiv versions of these recent papers.

Let $\mathcal{P}(E)$ be the set of probability measures on a Polish space E (separable completely metrizable topological space). The space of bounded and continuous functions on E is denoted by $\mathcal{C}_b(E)$. For $\mu \in \mathcal{P}(E)$ and φ a test function, $\langle \mu, \varphi \rangle = \int_E \varphi \mathrm{d}\mu$.

Let $N \geqslant 2$ and consider a system $\mathbf{X}^N = \left(X^{1,N}, \dots X^{N,N}\right) \in E^N$ of N random interacting particles. We say **interacting particles** by opposition to independent variables.

In the interacting particle systems presented in Section 1, the particles are random processes defined on $\mathbb{R}_+$: $X^{i,N} = (X_t^{i,N})_{t \geqslant 0}$, which are either continuous or càdlàg. Their state spaces are therefore either $E = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ or $E = \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$.

Definition 2.1. *A particle system $\mathbf{X}^N$ is said* **exchangeable** *(also called symmetric) when for any $N \geqslant 2$, the distribution of $\left(X^{i,N}\right)_{1 \leqslant i \leqslant N}$ is invariant by permutation of the indexes: for any π permutation of $\{1, \cdots, N\}$, $\left(X^{1,N}, \dots, X^{N,N}\right)$ and $\left(X^{\pi(1),N}, \dots, X^{\pi(N),N}\right)$ have the same distribution.*

The exchangeability implies in particular that $X^{1,N}$ has the same distribution as any particle $X^{k,N}$. The exchangeability assumption on the system $\mathbf{X}^N$ is crucial for establishing propagation of chaos for an interacting particle system, that is, for reducing the study of the system in the large-population limit to the study of a single particle under a suitable distribution.
Provided that their initial configuration $(X_0^{i,N})_{1 \leqslant i \leqslant N}$ is exchangeable, we observe that the interacting particle systems presented in Section 1 are exchangeable, since the interactions depend only on symmetric functionals of the particles, such as the empirical mean $\frac{1}{N}\sum_{i=1}^N X^{i,N}$, the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$, or the empirical cumulative distribution function $\frac{1}{N}\sum_{i=1}^N \mathbb{1}_{X^{i,N} \leqslant x}$.

Definition 2.2. *Let $\mu \in \mathcal{P}(E)$. For each $N \geqslant 2$, let μ^N denote the distribution of the system $\mathbf{X}^N$. We say that the system $\mathbf{X}^N$ is* **μ-chaotic** *if for any $k \geqslant 1$ and any $\varphi_1, \cdots, \varphi_k$ in $\mathcal{C}_b(E)$,*

$$\begin{aligned}
\lim_{N\to\infty} \left\langle \mu^N, \varphi_1 \otimes \dots \varphi_k \otimes 1 \otimes \dots 1 \right\rangle &= \lim_{N\to\infty} \mathbb{E}\left[\varphi_1(X^{1,N}) \dots \varphi_k(X^{k,N})\right] \\
&= \prod_{i=1}^k \langle \mu, \varphi_i \rangle.
\end{aligned}$$

We note that if $(Y_1, \dots, Y_k)$ are independent random variables with common distribution μ, then the system $\mathbf{X}^N$ is μ-chaotic if, for any $k \geqslant 1$ and any $\varphi_1, \cdots, \varphi_k$,

$$\lim_{N\to\infty} \mathbb{E}\left[\varphi_1(X^{1,N}) \dots \varphi_k(X^{k,N})\right] = \prod_{i=1}^k \mathbb{E}[\varphi_i(Y_i)] = \mathbb{E}[\varphi_1(Y_1) \dots \varphi_k(Y_k)].$$

Therefore, in the large-population limit ($N \to \infty$), we observe that k interacting particles of the system behave as k independent particles with common distribution μ.

From Definition 2.2, we easily deduce that if $\mathbf{X}^N$ is μ-chaotic, then the sequence $(X^{1,N})_{N \geqslant 1}$ converges in law to μ (take $k = 1$ in the definition). We can therefore identify the limiting distribution by studying only the asymptotic behavior of the first particle.

2.1 Criterion for propagation of chaos

We now present a useful criterion for proving that a system is chaotic.

Proposition 2.3. *Let $\mathbf{X}^N$ be an exchangeable system.*
The system $\mathbf{X}^N$ is μ-chaotic if and only if its empirical distribution $\overline{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ converges in law, as a $\mathcal{P}(E)$-valued random variable, to the deterministic probability measure μ when $N \to \infty$.

This result can be seen as a sort of Law of Large Numbers. If $\mathbf{X}^N$ is μ-chaotic, thus for any $\varphi \in \mathcal{C}_b(E)$,

$$\begin{aligned}\big\langle \overline{\mu}^N, \varphi \big\rangle - \mathbb{E}\big[\varphi(X^{1,N})\big] &= \frac{1}{N}\sum_{i=1}^N \varphi(X^{i,N}) - \mathbb{E}[\varphi(X^{1,N})] \\ &\underset{N\to\infty}{\longrightarrow} 0.\end{aligned}$$

Before proving Proposition 2.3, let us say a few words on the convergence in law of a sequence of random measures. The convergence in law is defined as the convergence against bounded continuous test functions. In the case of random measures, this representation is not intuitive.
However, we have the following useful result (proved in Section 2.2).

Proposition 2.4 (Corollary 1 in [CD22a])**.** *Let $\big(\mu^N\big)$ be a sequence of random probability measures. Consider a deterministic probability measure μ, i.e.* $\mathrm{Law}(\mu) = \delta_\mu$ *with $\mu \in \mathcal{P}(E)$. There is equivalence between the following assertions*

- *(i)* (μ^N) *converge in law to μ,*
- *(ii)* $\lim_{N\to\infty} \mathbb{E}\big[\big|\big\langle \mu^N - \mu, \varphi \big\rangle\big|\big] = 0$ *for any $\varphi \in \mathcal{C}_b(E)$.*

Note that, it is also equivalent to the convergence of $\mathbb{E}\Big[\big\langle \mu^N - \mu, \varphi \big\rangle^2\Big]$ to 0 for any $\varphi \in \mathcal{C}_b(E)$ uniformly continuous, since

$$\mathbb{E}\Big[\big\langle \mu^N - \mu, \varphi \big\rangle^2\Big] \leqslant 2\|\varphi\|_\infty \mathbb{E}\big[\big|\big\langle \mu^N - \mu, \varphi \big\rangle\big|\big] \leqslant 2\|\varphi\|_\infty \mathbb{E}\Big[\big\langle \mu^N - \mu, \varphi \big\rangle^2\Big]^{1/2}.$$

Proof of Proposition 2.3. This proposition is proved in [Szn91, Proposition 2.2] and [CD22a, Lemma 3.19].

1. Assume first that $\mathbf{X}^N$ is μ-chaotic. We prove the direct implication using Definition 2.2 with $k = 1, 2$.

   Let $\varphi \in \mathcal{C}_b(E)$. We have

$$\begin{aligned}\mathbb{E}\Big[\big\langle \overline{\mu}^N - \mu, \varphi \big\rangle^2\Big] &= \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^N \varphi(X^{i,N}) - \langle \mu, \varphi\rangle\right)^2\right] \\ &= \frac{1}{N^2}\sum_{i,j=1}^N \mathbb{E}\big[\varphi(X^{i,N})\varphi(X^{j,N})\big] - \frac{2}{N}\langle \mu, \varphi\rangle \sum_{i=1}^N \mathbb{E}\big[\varphi(X^{i,N})\big] + \langle \mu, \varphi\rangle^2.\end{aligned}$$

Using the exchangeability of the system, we deduce

$$\mathbb{E}\Big[\big\langle\overline{\mu}^N-\mu,\varphi\big\rangle^2\Big]=\frac{1}{N}\mathbb{E}\big[\varphi(X^{1,N})^2\big]\\ +\frac{N-1}{N}\mathbb{E}\big[\varphi(X^{1,N})\varphi(X^{2,N})\big]-2\langle\mu,\varphi\rangle\mathbb{E}\big[\varphi(X^{1,N})\big]+\langle\mu,\varphi\rangle^2,$$

which converges to 0 when $N\to\infty$ by Definition 2.2. We deduce the convergence in law of the random measure $\overline{\mu}^N$ to the constant measure μ (by Proposition 2.4).

2. Conversely, assume that the random probability measures $\overline{\mu}^N$ converges in law to the deterministic probability measure μ.

   Let $k\geqslant 1$ and $\varphi_1,\cdots,\varphi_k$ in $\mathcal{C}_b(E)$. We have

$$\left|\big\langle\overline{\mu}^N,\varphi_1\otimes\dots\varphi_k\otimes 1\otimes\dots 1\big\rangle-\prod_{i=1}^k\langle\mu,\varphi_i\rangle\right|=\left|\mathbb{E}\big[\varphi_1(X^{1,N})\dots\varphi_k(X^{k,N})\big]-\prod_{i=1}^k\langle\mu,\varphi_i\rangle\right|\\ \leqslant\left|\mathbb{E}\big[\varphi_1(X^{1,N})\dots\varphi_k(X^{k,N})\big]-\mathbb{E}\left[\prod_{i=1}^k\big\langle\overline{\mu}^N,\varphi_i\big\rangle\right]\right|+\left|\mathbb{E}\left[\prod_{i=1}^k\big\langle\overline{\mu}^N,\varphi_i\big\rangle\right]-\prod_{i=1}^k\langle\mu,\varphi_i\rangle\right|. \quad (2.1)$$

   By Proposition 2.4, since $\overline{\mu}^N$ converges in law to μ, and $\varphi_i\in\mathcal{C}_b(E)$, we have the convergence of $\big\langle\overline{\mu}^N,\varphi_i\big\rangle$ to $\langle\mu,\varphi_i\rangle$ in L^1 for any $i\in\{1,\dots,k\}$. As the functions φ_i are bounded, we deduce by induction on k the convergence of the second term in (2.1) to 0 when $N\to\infty$.

   We now study the first term of (2.1).

   Using the exchangeability of the system, we note that

$$\mathbb{E}\big[\varphi_1(X^{1,N})\dots\varphi_k(X^{k,N})\big]=\frac{1}{A(k,N)}\sum_{\substack{i_1,\dots,i_k\\ \text{pairwise distinct}}}\mathbb{E}\big[\varphi_1(X^{i_1,N})\dots\varphi_k(X^{i_k,N})\big],$$

   where $A(k,N)=\frac{N!}{(N-k)!}$ is the number of pairwise distinct tuples $(i_1,\dots,i_k)$ of integers between 1 and N. In addition, by definition of $\overline{\mu}^N$,

$$\prod_{i=1}^k\big\langle\overline{\mu}^N,\varphi_i\big\rangle=\frac{1}{N^k}\prod_{i=1}^k\left(\sum_{j=1}^N\varphi_i(X^{j,N})\right)=\frac{1}{N^k}\sum_{i_1,\dots,i_k}\varphi_1(X^{i_1,N})\dots\varphi_k(X^{i_k,N})\\ =\frac{1}{N^k}\sum_{\substack{i_1,\dots,i_k\\ \text{pairwise distinct}}}\varphi_1(X^{i_1,N})\dots\varphi_k(X^{i_k,N})+R(k,N),$$

   Denoting by $C=\sup_{1\leqslant i\leqslant k}\|\varphi_i\|_\infty$, we note that the remainder term $R(k,N)$ satisfies

$$|R(k,N)|\leqslant\frac{C^k}{N^k}\big(N^k-A(k,N)\big).$$

Consequently,

$$
\begin{aligned}
&\left|\mathbb{E}\big[\varphi_1(X^{1,N})\dots\varphi_k(X^{k,N})\big]-\mathbb{E}\left[\prod_{i=1}^{k}\langle\overline{\mu}^N,\varphi_i\rangle\right]\right|\\
&\leqslant\left(\frac{1}{A(k,N)}-\frac{1}{N^k}\right)A(k,N)C^k+\mathbb{E}[|R(k,N)|]\\
&\leqslant 2C^k\left(1-\frac{A(k,N)}{N^k}\right)\leqslant 2C^k\left(1-\left(1-\frac{k-1}{N}\right)^k\right)\leqslant 2C^k\frac{k(k-1)}{N},
\end{aligned}
$$

which goes to 0 when $N\to\infty$. Then $\mathbf{X}^N$ is μ-chaotic.

□

2.2 About the convergence in law of random measures

For completeness of this section, we now prove Proposition 2.4. Recall that weak convergence of measures in $\mathcal{P}(\mathcal{P}(E))$ is defined as the convergence against test functions in $\mathcal{C}_b(\mathcal{P}(E))$. An example of such functionals is given by the linear functions: $\mu\mapsto\langle\mu,\varphi\rangle$ for a given $\varphi\in\mathcal{C}_b(E)$. However, the weak convergence of measures is not simple to represent.

Note that for a random variable X on E, we usually denote by $\mu\in\mathcal{P}(E)$ its distribution: $\mu=\mathrm{Law}(X)$. For a $\mathcal{P}(E)$-valued random variable μ, we denote here by $\widehat{\mu}\in\mathcal{P}(\mathcal{P}(E))$ its distribution: $\widehat{\mu}=\mathrm{Law}(\mu)$.

We introduce various distances on $\mathcal{P}(E)$ and $\mathcal{P}(\mathcal{P}(E))$. First, the distance $\mathcal{D}$ on $\mathcal{P}(E)$, defined by

$$
\mathcal{D}(\mu,\nu)=\sum_{k=1}^{\infty}\frac{1}{2^k}|\langle\mu-\nu,\varphi_k\rangle|,
$$

where $(\varphi_k)_{k\in\mathbb{N}}$ is a family of continuous bounded functions on E with $\|\varphi_k\|_\infty\leqslant 1$, and satisfying the *separating property* (see [EK86, Section III-4]): $\forall\mu,\nu\in\mathcal{P}(E)$,

$$
\Big(\forall k\in\mathbb{N}\quad\langle\mu,\varphi_k\rangle=\langle\nu,\varphi_k\rangle\Big)\quad\Longrightarrow\quad\Big(\mu=\nu\Big),
$$

and such that $\mathcal{D}$ metrizes the topology of weak convergence on $\mathcal{P}(E)$. Since E is a Polish space such a family exists. When E is a compact Polish space, it is a consequence of the Stone-Weierstrass theorem, and when E is a general Polish space, we refer to [SV06, Theorem 1.1.2]).

Set $\mathcal{E}:=\mathcal{P}(E)$. For $\widehat{\mu}\in\mathcal{P}(\mathcal{E})$ and $\widehat{\nu}\in\mathcal{P}(\mathcal{E})$, let $\mathcal{W}_{\mathcal{D}}$ the Wasserstein distance on $\mathcal{P}(\mathcal{E})$ defined by

$$
\mathcal{W}_{\mathcal{D}}(\widehat{\mu},\widehat{\nu}):=\inf_{\mu\sim\widehat{\mu},\nu\sim\widehat{\nu}}\mathbb{E}[\mathcal{D}(\mu,\nu)]=\inf_{\widehat{\pi}\in\Pi(\widehat{\mu},\widehat{\nu})}\int_{\mathcal{E}\times\mathcal{E}}\mathcal{D}(\mu,\nu)\widehat{\pi}(\mathrm{d}\mu,\mathrm{d}\nu)
$$

with $\Pi(\widehat{\mu},\widehat{\nu})$ is the set of probability measures $\widehat{\pi}$ on $\mathcal{E}\times\mathcal{E}$ with marginals $\widehat{\mu}$ and $\widehat{\nu}$.

We first prove the following result.

Proposition 2.5 (Proposition 6 in [CD22a]). *Let $\left(\mu^N\right)_{N\in\mathbb{N}}$ be a sequence of random probability measures and μ be a random probability measure.*

(i) if $\lim_{N\to\infty}\mathcal{W}_{\mathcal{D}}\big(\mathrm{Law}(\mu^N),\mathrm{Law}(\mu)\big)=0$, then the sequence $(\mu^N)_{N\geqslant 0}$ of $\mathcal{P}(E)$-valued random variables converges in law towards μ.

(ii) if $\forall\varphi\in\mathcal{C}_b(E)$, $\lim_{N\to\infty}\mathbb{E}\big[\big|\langle\mu^N-\mu,\varphi\rangle\big|\big]=0$, then

$$\lim_{N\to\infty}\mathcal{W}_{\mathcal{D}}(\mathrm{Law}(\mu^N),\mathrm{Law}(\mu))=0.$$

(iii) If for any $\varphi\in\mathcal{C}_b(E)$, we have

$$\lim_{N\to\infty}\mathbb{E}\big[\big|\langle\mu^N-\mu,\varphi\rangle\big|\big]=0,$$

then $\left(\mu^N\right)_{N\in\mathbb{N}}$ converges in law to μ.

Proof. (i) Recall that the convergence in law is characterized by the convergence of integrals against test functions in the set of bounded uniformly continuous functions (Portmanteau theorem, see [Bil99, Theorem 2.1]). Let $\Phi\in\mathcal{C}_b(\mathcal{E})$ be a function uniformly continuous for the metric $\mathcal{D}$: for any $\varepsilon>0$, there exists $\delta:=\delta(\varepsilon)>0$ such that for any $\mu,\nu\in\mathcal{E}$,

$$\mathcal{D}(\mu,\nu)\leqslant\delta\quad\Rightarrow\quad|\Phi(\mu)-\Phi(\nu)|\leqslant\varepsilon.$$

Recall the well-known relation for a random variable X: $\langle Law(X),\Phi\rangle=\int\Phi\mathrm{d}\mathbb{P}_X=\mathbb{E}[\Phi(X)]$. Consequently, we have

$$\begin{aligned}\big|\langle\mathrm{Law}(\mu^N)-\mathrm{Law}(\mu),\Phi\rangle\big|&\leqslant\mathbb{E}\big[\big|\Phi(\mu^N)-\Phi(\mu)\big|\big]\\&\leqslant\varepsilon+2\|\Phi\|_\infty\mathbb{P}(\big|\Phi(\mu^N)-\Phi(\mu)\big|\geqslant\varepsilon)\\&\leqslant\varepsilon+2\|\Phi\|_\infty\mathbb{P}(\mathcal{D}\big(\mu^N,\mu\big)>\delta)\\&\leqslant\varepsilon+\frac{2\|\Phi\|_\infty}{\delta}\mathbb{E}\big[\mathcal{D}\big(\mu^N,\mu\big)\big],\end{aligned}\tag{2.2}$$

where the uniform continuity of Φ is used in the third line, and Markov's inequality in the last line. Taking the infimum on every pair of random measures (μ^N,μ) with distribution in $\Pi\big(\mathrm{Law}(\mu^N),\mathrm{Law}(\mu)\big)$, we deduce

$$\big|\langle\mathrm{Law}(\mu^N)-\mathrm{Law}(\mu),\Phi\rangle\big|\leqslant\varepsilon+\frac{2\|\Phi\|_\infty}{\delta}\mathcal{W}_{\mathcal{D}}(\mathrm{Law}(\mu^N),\mathrm{Law}(\mu)).$$

The result of (i) follows by taking first the limit when $N\to\infty$, and then when $\varepsilon\to 0$. We thus recover the fact that the topology induced by Wasserstein distance is stronger than the one induced by the weak convergence.

(ii) By definition of $\mathcal{W}_{\mathcal{D}}$ and the monotone convergence theorem, we have

$$\begin{aligned}\mathcal{W}_{\mathcal{D}}(\mathrm{Law}(\mu^N),\mathrm{Law}(\mu))&\leqslant\mathbb{E}[\mathcal{D}(\mu^N,\mu)]\\&\leqslant\sum_{k=1}^{\infty}\frac{1}{2^k}\mathbb{E}[\big|\langle\mu^N-\mu,\varphi_k\rangle\big|].\end{aligned}$$

Using the dominated convergence theorem, the conclusion of (ii) follows.

(iii) if $\forall\varphi \in \mathcal{C}_b(E)$ uniformly continuous, $\lim_{N\to\infty} \mathbb{E}\big[\big|\big\langle \mu^N - \mu, \varphi\big\rangle\big|\big] = 0$, (ii) implies $\lim_{N\to\infty} \mathcal{W}_{\mathcal{D}}(\mathrm{Law}(\mu^N), \mathrm{Law}(\mu)) = 0$, and thus, by (i), $(\mu^N)_{N\geqslant 0}$ converges in law to μ.

The proposition is proved. □

We are now able to prove Proposition 2.4.

Proof of Proposition 2.4. The implication $(ii) \Rightarrow (i)$ is given by Proposition 2.5.
To prove the direct implication $(i) \Rightarrow (ii)$, we observe that for a fixed $\varphi \in \mathcal{C}_b(E)$, when μ is deterministic the function $\Phi : \nu \mapsto |\langle \nu - \mu, \varphi\rangle|$ is a deterministic function on $\mathcal{P}(\mathcal{E})$. In addition, Φ is bounded by $2\|\varphi\|_\infty$ and continuous for the weak topology on $\mathcal{P}(E)$:

$$|\Phi(\nu_1) - \Phi(\nu_2)| = ||\langle \nu_1 - \mu, \varphi\rangle| - |\langle \nu_2 - \mu, \varphi\rangle|| \leqslant |\langle \nu_1 - \nu_2, \varphi\rangle|,$$

by the reverse triangle inequality. Consequently, it is a test function for the convergence in law, and since (μ^N) converges in law to μ, the result follows. □

2.3 Tightness

We first recall the definition of tightness.

Definition 2.6. *A sequence of probability measures $(\mu^N)_{N\geqslant 1}$ on E is* **tight** *if $\forall\varepsilon > 0$ there is a compact set K_ε of E such that $\sup\limits_{N\geqslant 1} \mu^N(K_\varepsilon^c) \leqslant \varepsilon$, where K_ε^c is the complement of K_ε.*

By abuse of notation, we say that a sequence of random varables $(Z^N)_{N\geqslant 1}$ is tight if the sequence of their distributions $(\mathrm{Law}(Z^N))_{N\geqslant 1}$ is tight.

The notion of tightness for a sequence of probability measures is related to the notion of relative compactness for the topology induced by weak convergence. The analogous notion for a numerical sequence $(u_N)_{N\geqslant 1}$ is boundedness. It is well known that every bounded sequence admits a convergent subsequence. Moreover, if the limit is unique, then the *whole* sequence $(u_N)_{N\geqslant 1}$ converges. Similar properties hold for tight sequences of probability measures, as we will see later.
When E is a polish space, we have the following result proved in [Bil99, Section 5] (only E metric space is needed for the direct implication).

Theorem 2.7 (Prokhorov theorem)**.** *A sequence of probability measures $\left(\mu^N\right)_{N\geqslant 1}$ on E is tight if and only if the sequence $\left(\mu^N\right)_{N\geqslant 1}$ is relatively compact.*

In the case of exchangeable interacting particle systems, we have the following result, which allows us to work with the distribution of a single particle rather than with the empirical distribution of the whole particle system.

Proposition 2.8. *Let $\mathbf{X}^N$ be an exchangeable system with values in E^N. The sequence of random measures $\overline{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ is tight if and only if the sequence $\left(\mathrm{Law}(X^{1,N})\right)_{N\geqslant 1}$ is tight.*

Let us introduced the **intensity measure** I_μ of a random measure μ as the probability measure in $\mathcal{P}(E)$ defined by

$$\langle I_\mu, \varphi\rangle = \mathbb{E}[\langle \mu, \varphi\rangle] = \int_{\mathcal{P}(E)} \langle \nu, \varphi\rangle \widehat{\mu}(\mathrm{d}\nu),$$

for $\varphi \in \mathcal{C}_b(E)$, where $\widehat{\mu} = \mathrm{Law}(\mu)$. We have the following result linking the tightness of random measures and the tightness of their intensity measures.

Lemma 2.9 (Lemma 3.15 in [CD22a])**.** *The tightness of a sequence of random measures $(\mu^N)_{N\geqslant 1}$, as $\mathcal{P}(E)$-valued random variables, is equivalent to the tightness of the sequence of their intensity measures $(I^N)_{N\geqslant 1}$.*

We easily see that Proposition 2.8 is a direct consequence of the above lemma, because the intensity measure I^N of $\overline{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$, by exchangeability of the system $\mathbf{X}^N$, is

$$\langle I^N, \varphi\rangle = \mathbb{E}[\varphi(X^{1,N})].$$

In this case, the intensity measure is just the law of $X^{1,N}$.
Note that the weak convergence of a sequence of distributions of random measures $(\mu^N)_{N\geqslant 1}$ in $\mathcal{P}(\mathcal{P}(E))$ is not equivalent to the weak convergence of their intensity measures $(I^N)_{N\geqslant 1}$ in $\mathcal{P}(E)$ (only the direct implication holds, see Lemma 3.14 in [CD22a]). Note that, in the case of a chaotic exchangeable system $\mathbf{X}^N$, the asymptotic measure μ is also the limiting distribution of the first particle $(X^{1,N})_{N\geqslant 1}$.
We know that $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d), \|\cdot\|_\infty)$ is a Polish space, therefore using [Bil99, Theorem 5.1] and its corollary, we deduce that the weak convergence of $(\mathrm{Law}(X^{1,N}))_{N\geqslant 0}$ to a measure μ in $\mathcal{P}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d))$ is equivalent to $(\mathbf{X}^N)_{N\geqslant 1}$ μ-chaotic (see also [Bil99, Examples 5.1 and 5.2]).
The situation is more complicated when working in the Skorokhod space $E = \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$, because the metric d that makes $(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d), \mathrm{d})$ a Polish space is weaker than the topology induced by the uniform norm on compact sets, and in particular the projection on the time marginals are not continuous, and therefore the limiting distributions have to be characterized with the limiting behavior of the empirical measure $\overline{\mu}^N$.
More details on the convergence of random sequences with values in $E = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ or $E = \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ are presented in the next section.

Proof of Lemma 3.5. We refer to the proofs of [Szn91, Proposition 2.2] and [CD22a, Lemma 3.15]. We note that the function $\mu \mapsto \langle \mu, \varphi\rangle$ is a continuous function, bounded by $\|\varphi\|_\infty$, for any $\varphi \in \mathcal{C}_b(E)$. Therefore, we easily deduce the continuity of $\mu \mapsto I_\mu$, and the direct implication follows. Consequently, we only need to prove that the tightness of I^N implies the tightness of μ^N.
Let $\varepsilon > 0$, and K_ε be a compact subset of E such that $\sup_{N\geqslant 1} I^N(K_\varepsilon^c) \leqslant \varepsilon$.
Let $\eta > 0$. We note that by the Markov inequality, for any $N \geqslant 1$,

$$\begin{aligned}\widehat{\mu}^N\big(\{\nu \in \mathcal{P}(E) : \nu(K_{\varepsilon\eta}^c) \geqslant \eta\}\big) &= \int \mathbb{1}_{\{\nu(K_{\varepsilon\eta}^c)\geqslant\eta\}}\widehat{\mu}^N(\mathrm{d}\nu) \leqslant \frac{1}{\eta}\int \nu\big(K_{\varepsilon\eta}^c\big)\widehat{\mu}^N(\mathrm{d}\nu)\\ &= \frac{1}{\eta}\int \Big\langle \nu, \mathbb{1}_{K_{\varepsilon\eta}^c}\Big\rangle \widehat{\mu}^N(\mathrm{d}\nu) = \frac{1}{\eta} I^N\big(K_{\varepsilon\eta}^c\big) \leqslant \varepsilon.\end{aligned}$$

We deduce that

$$\widehat{\mu}^N\left(\bigcup_{p\geqslant 1}\Big\{\nu \in \mathcal{P}(E) : \nu(K_{(\varepsilon 2^{-p})/p}^c) > 1/p\Big\}\right) \leqslant \varepsilon \sum_{p\geqslant 1} 2^{-p} = \varepsilon$$

Let $\mathcal{K}_\varepsilon = \bigcap_{p\geqslant 1}\Big\{\nu \in \mathcal{P}(E) : \nu(K_{\varepsilon 2^{-p}/p}^c) \leqslant 1/p\Big\}$. $\mathcal{K}_\varepsilon$ is a compact subset of $\mathcal{P}(E)$ and we deduce that $\big(\widehat{\mu}^N\big)_{N\geqslant 1}$ is tight. $\square$

From Proposition 2.8, and as we will see in the next section, we deduce that in order to prove that a propagation of chaos holds for an exchangeable system $\mathbf{X}^N$, it is sufficient to

- prove the tightness of the sequence of first particle's distributions $(X^{1,N})_{N\geqslant 1}$,
- identify the limiting distribution, and
- establish the uniqueness of this limit.

In Section 1, several interacting particle systems were introduced, where $\mathbf{X}^N = (X_t^{1,N}, \ldots, X_t^{N,N})_{t\geqslant 0}$ are continuous or càdlàg processes taking values in $\mathbb{R}^d$, $d \geqslant 1$. In the next section, we recall some properties of weak convergence and tightness for such processes. In Sections 4 and 5, we analyze two examples in detail and, whenever possible, provide an estimate of the rate of convergence to the asymptotic distribution. The long-time behavior of both particle systems and their limiting distributions is also briefly discussed.

3 Convergence of stochastic processes

Let us recall the definition of weak convergence of probability measures.

Definition 3.1. *A sequence of probability measures $(\mu^N)_{N\geqslant 1}$ on a Polish space E converges weakly to the probability measure μ if for any $f \in \mathcal{C}_b(E)$,*

$$\lim_{N\to\infty} \int f \mathrm{d}\mu^N = \int f \mathrm{d}\mu.$$

The objective of this section is not to provide an exhaustive treatment of the convergence of continuous or càdlàg processes, but only to highlight the main results that are useful in the study of several examples of interacting particle systems. We mainly refer to the books of Billingsley [Bil99], and of Jacod and Shiryaev [JS03] for more details on the convergence of stochastic processes. We can also mention the books of Ethier and Kurtz [EK86], and of Stroock and Varadhan [SV06].

For $d \geqslant 1$ and $T > 0$, let $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the space of continuous functions, and $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ denote the Skorokhod space, namely the space of *càdlàg* functions (right-continuous functions with left limits), defined on $\mathbb{R}_+$ with values in $\mathbb{R}^d$.
The Euclidean norm of a vector $x \in \mathbb{R}^d$ is denoted by $|x|$.

Endowed with suitable metrics, the spaces $E = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ and $E = \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ are Polish spaces (see [Bil99, JS03]). Note that the topology induced by the metric on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ that makes it a Polish space is weaker than the local uniform topology.

3.1 Tightness in $\mathcal{C}$

We start with a result ensuring the convergence for tight probability measures on $\mathcal{C}\big(\mathbb{R}_+, \mathbb{R}^d\big)$. This theorem is a consequence of Prokhorov's theorem (see Example 2.7+ Theorem 2.6 and Section 5 in [Bil99]).

Theorem 3.2. *Let $Z^N = (Z^N_t)_{t\geqslant 0}$ be a sequence of processes with values in $\mathcal{C}\big(\mathbb{R}_+, \mathbb{R}^d\big)$. Denote by μ^N the distribution of Z^N. Let μ a probability measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ and Z a random process with distribution μ.*
Then the sequence $(\mu^N)_{N\geqslant 1}$ converges weakly to μ if and only if it satisfies the two following conditions

- Tightness*: the sequence $\big(\mu^N\big)$ is tight in $\mathcal{C}\big(\mathbb{R}_+, \mathbb{R}^d\big)$;*
- Convergence of the finite-dimensional distributions*: for any $k \in \mathbb{N}$ and any $(t_1, \dots, t_k) \in \mathbb{R}^k_+$, the vector $(Z^N_{t_1}, \dots, Z^N_{t_k})$ converges in law to $(Z_{t_1}, \dots, Z_{t_k})$, as $N \to \infty$.*

We now present useful results on the tightness of continuous stochastic processes. To this end, we introduce the modulus of continuity of a function $z(\cdot)$ on $[0, T]$: for $\delta \in (0, T]$

$$w_T(z; \delta) := \sup_{\substack{|s-t|\leqslant\delta \\ \text{with } s,t\in[0,T]}} |z(s) - z(t)|.$$

Note that $z \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ if and only if $\lim_{\delta\downarrow 0} w_T(z; \delta) = 0$ for all $T \geqslant 0$ (Heine-Cantor theorem: a continuous function on a compact is uniformly continuous). We first recall a well-known result on compact sets of continuous functions.

Theorem 3.3 (Arzela-Ascoli theorem)**.** *A subset A of $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ is relatively compact (i.e. its closure $\overline{A}$ is compact) if and only if*

(i) $\sup_{z\in A} |z(0)| < \infty$*;*

(ii) Equicontinuity: $\lim_{\delta\to 0} \sup_{z\in A} w_T(z;\delta) = 0$ *for all* $T \geqslant 0$*.*

We deduce the following characterization of the tightness of probability measures on the space of continuous functions.

Theorem 3.4 (Theorem 7.3 in [Bil99])**.** *A sequence of continuous processes $(Z^N)_{N\geqslant 1}$ on $\mathbb{R}_+$ is tight if an only if the two following conditions are satisfied:*

(i) $\forall \eta > 0$*, there exist $a > 0$ and $N_0 \geqslant 1$ such that $\forall N \geqslant N_0$,*

$$\mathbb{P}\big(\big|Z_0^N\big| \geqslant a\big) \leqslant \eta,$$

(ii) $\forall a, T > 0$*,*

$$\lim_{\delta\to 0} \limsup_{N\to\infty} \mathbb{P}\big(w_T(Z^N;\delta) \geqslant a\big) = 0.$$

Proof. Let $T > 0$. We first prove the direct implication. Assume that (Z^N) is tight. Let $\eta > 0$, and K a compact set of $\mathcal{C}([0,T], E)$ such that $\mathbb{P}(Z^N \in K) > 1 - \eta$ for any $N \geqslant 1$. We have $K \subset \{z : |z(0)| \leqslant a\}$ for a large enough, and $K \subset \{z : w_T(z;\delta) \leqslant \varepsilon\}$ for δ small enough. Thus the conditions are satisfied for $N_0 = 1$.
Assume that Conditions (i) and (ii) are satisfied. As $\mathcal{C}([0,T], E)$ is separable and complete, a finite collection of measures is tight. Let $\eta > 0$, a given by (i), then $\mathbb{P}(Z^N \in B) \geqslant 1 - \eta$ where $B := \{z : |z(0) \leqslant a|\}$. By (ii), $\mathbb{P}(Z^N \in B^k) \geqslant 1 - \eta/2^k$ for all N, with $B^k = \{z : w_T(z;\delta_k) < 1/k\}$. Let K be the closure of $A = B \cap \bigcap_k B^k$, then $\mathbb{P}(Z^N \in K) \geqslant 1 - 2\eta$. By the Arzela-Ascoli theorem, A is relatively compact, and therefore K is compact. We conclude that Z^N is tight. ☐

We now give a sufficient condition to ensure that condition (ii) in the previous theorem is satisfied.

Theorem 3.5 (Theorem 7.3 + Condition (7.12) in [Bil99])**.** *A sequence of continuous processes $(Z^N)_{N\geqslant 1}$ on $\mathbb{R}_+$ is tight if the two following conditions are satisfied:*

(i) $\forall \eta > 0$*, there exist $a > 0$ and $N_0 \geqslant 1$ such that $\forall N \geqslant N_0$,*

$$\mathbb{P}\big(\big|Z_0^N\big| \geqslant a\big) \leqslant \eta,$$

(ii) $\forall a, \eta, T > 0$ *and $t \in [0,T]$, there exists $\delta \in (0,T)$ and $N_0 \geqslant 1$ such that for any $N \geqslant N_0$,*

$$\frac{1}{\delta}\mathbb{P}\left(\sup_{t\leqslant s\leqslant t+\delta} \big|Z_t^N - Z_s^N\big| \geqslant a\right) \leqslant \eta.$$

Proof. We only have to prove that Condition (ii) in Theorem 3.4 is satisfied. Let $a, \eta, T > 0$ and $t \in [0,T]$. If $t > T - \delta$, the supremum is thus taken on $s \in [t,T]$. We set $t_k = k\delta$ for $k < \lfloor T/\delta \rfloor$. We have $(t_k - t_{k-1}) = \delta$. For a continuous function z on $[0,T]$, we note that

$$w_T(z;\delta) \leqslant 3 \max_k \sup_{s\in[t_{k-1},t_k]} |z(s) - z(t_{k-1})|,$$

and therefore,

$$\mathbb{P}(w_T(Z^N;\delta) \geqslant 3a) \leqslant \sum_{k=1}^{\lfloor T/\delta \rfloor} \mathbb{P}\left(\sup_{s\in[t_{k-1},t_k]} \left| Z^N_s - Z_{t_{k-1}} \right| \geqslant a \right)$$
$$\leqslant T\eta.$$

Taking the limit when $\eta \to 0$, we deduce that Condition (ii) in Theorem 3.4 is satisfied. □

In many examples of exchangeable particle systems, the processes $X^{1,N}$ are in L^1, thus it is sufficient to obtain good control of the first moment of $\sup_{s\leqslant t\leqslant s+\delta} \left| X^{1,N}_t - X^{1,N}_s \right|$ to obtain tightness of $(X^{1,N})_{N\geqslant 1}$.

3.2 Tightness in $\mathbb{D}$

Prove tightness for sequences of càdlàg processes is a bit more challenging because the characterization of compact sets is more tricky and because the projections on time-marginal distributions are not continuous. More precisely, the projection $p_t : z \in\mapsto z_t$ defined on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ is continuous in z if and only if $s \mapsto z_s$ is continuous in t.

Theorem 3.6 (Theorem 13.1 in [Bil99])**.** *Let $Z^N = (Z^N_t)_{t\geqslant 0}$ be a sequence of processes in $\mathbb{D}([0,T], \mathbb{R}^d)$, and Z a process in $\mathbb{D}([0,T], \mathbb{R}^d)$. We note by μ^N the distribution of Z^N.*
We say that $t \in [0,T]$ is a continuous point of Z if $t \notin \{s \in [0,T] : \mathbb{P}(Z_s - Z_{s-} \neq 0) > 0\}$.
Then, if

- *$(\mu^N)_{N\geqslant 1}$ is tight, and*
- *for any $k \in \mathbb{N}$ and any $(t_1, \dots, t_k) \in [0,T]^k$ continuous points of Z a.s., the vector $(Z^N_{t_1}, \dots, Z^N_{t_k})$ converges in distribution to $(Z_{t_1}, \dots, Z_{t_k})$ when $n \to \infty$*

Then (Z^N) converges in law to Z.

We introduce the following modulus on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$, as defined in [JS03]:

$$w'_T(z;\delta) = \inf_\pi \Big\{ \max_{1\leqslant k\leqslant r} w(z;[t_{k-1},t_k)) \text{ with } \pi = (t_0,\dots,t_r) \text{ partition of } [0,T]$$
$$\text{such that } \min_{1\leqslant k< r} (t_k - t_{k-1}) > \delta \Big\},$$

where $w(z;A) = \sup_{s,t\in E} |z(t) - z(s)|$ for $A \subset [0,T]$.

Remark 3.7. *The modulus on $\mathbb{D}([0,T], \mathbb{R})$ is defined in [Bil99] by*

$$w'_T(z;\delta) = \inf_\pi \Big\{ \max_{1\leqslant k\leqslant r} w(z;[t_{k-1},t_k)) \textit{ with } \pi = (t_0,\dots,t_r) \textit{ partition of } [0,T]$$
$$\textit{such that } \min_{1\leqslant k\leqslant r} (t_k - t_{k-1}) > \delta \Big\},$$

The main difference with the previous definition is that [JS03] doesn't require $t_r - t_{r-1} > \delta$. The reason is that, in the topology of $\mathbb{D}([0,T], \mathbb{R}^d)$, the end point $t = T$ plays a specific role, while the points $T \geqslant 0$ should not play any specific role in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$. Consequently, the proofs are slightly more involved in $\mathbb{D}([0,T], \mathbb{R}^d)$ than in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$.

Remark 3.8. *We have $z \in \mathbb{D}(\mathbb{R}_+, \mathbb{R})$ if and only if for any $T > 0$, $\sup_{0\leqslant t\leqslant T} |z(t)| < \infty$ and $\lim_{\delta\downarrow 0} w'_T(z;\delta) = 0$. In addition, we note that*

$$w'_T(z;\delta) \leqslant w_T(z;2\delta) \quad \textit{and} \quad w_T(z;\delta) \leqslant 2w'_T(z,\delta) + \sup_{t\in[0,T]} |z(t) - z(t-)|.$$

Proof. We refer to [JS03, Lemma VI.1.11] for the equivalence. We only prove here the relation between w'_T and w_T.
For the first inequality, take a partition π such that $\delta < t_k - t_{k-1} \leqslant 2\delta$.
For the second inequality, recall that $w(z;\delta) = \sup_{|t-s|\leqslant\delta} |x(t) - x(s)|$. By the condition on the partition, for $|t-s| \leqslant \delta$, there is $k \geqslant 1$ such that

- $(s,t) \in [t_{k-1}, t_k)$, and thus $|z(t) - z(s)| \leqslant w(z;[t_{k-1}, t_k))$,
- or $s < t_{k-1} \leqslant t < t_k$, and since a jump can occur in t_{k-1}, we have

$$|z(t) - z(s)| \leqslant |z(s) - z(t_{k-1})| + |x(t_{k-1}) - x((t_{k-1})-)| + |z(t) - z(t_{k-1})|$$

- or $t_{k-1} < s < t_k \leqslant t$, and since a jump can occur in t_k, we have

$$|z(t) - z(s)| \leqslant |z(s) - z(t_k)| + |x(t_k) - x(t_k-)| + |z(t) - z(t_k)|$$

We deduce that

$$|z(t) - z(s)| \leqslant 2w(z;[t_{k-1}, t_k)) + \sup_{t\in[0,T]} |z(t) - z(t-)|.$$

□

Theorem 3.9 (Ascoli's Theorem)**.** *A subset $A \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ is relatively compact if and only if for any $T \geqslant 0$*

(i) $\sup_{z\in A} \sup_{t\in[0,T]} |z(t)| < \infty$*;*

(ii) $\lim_{\delta\to 0} \sup_{z\in A} w'_T(z;\delta) = 0$*.*

We deduce the following characterization of the tightness of distributions in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$.

Theorem 3.10 (Theorem 13.2 in [Bil99])**.** *Let $Z^N = (Z^N_t)_{t\geqslant 0}$ be a sequence of adapted càdlàg processes on $\mathbb{R}_+$. The sequence Z^N is tight if and only if for any $T \geqslant 0$*

(i) the sequence $\left(\sup_{t\in[0,T]} \left|Z^N_t\right|\right)_{N\geqslant 1}$ *is tight:*

$$\lim_{a\to\infty} \limsup_N \mathbb{P}\left(\sup_{t\in[0,T]} \left|Z^N_t\right| \geqslant a\right) = 0;$$

(ii) $\forall a > 0$,

$$\lim_{\delta\to 0} \limsup_N \mathbb{P}\big(w'_T(Z^N, \delta) > a\big) = 0.$$

The following criterion due to Aldous, [Ald78], is one of the most classical results to prove the tightness of the laws of sequences of processes in $\mathbb{D}([0,\infty), \mathbb{R})$.

Theorem 3.11 (Aldous criterion)**.**
Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space. Let $Z^N = (Z^N_t)_{t\geqslant 0}$ be a sequence of adapted càdlàg processes on $[0,\infty)$. Assume that the sequence of processes satisfies the following conditions: for any $T \geqslant 0$

i) $\forall \eta > 0$, there exist $a > 0$ and $N_0 \geqslant 1$ such that $\forall N \geq N_0$,

$$\mathbb{P}\left(\sup_{t\leq T} |Z^N_t| \geqslant a\right) \leqslant \eta$$

(in other words, $\lim_{a\to\infty} \limsup_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} |Z^N_t| \geqslant a\right) = 0$);

ii) $\forall a > 0$,

$$\lim_{\delta\downarrow 0} \limsup_{N\to\infty} \sup_{\substack{S,S' \text{stopping times:} \\ S\leqslant S'\leqslant S+\delta\leqslant T}} \mathbb{P}\big(|Z^N_{S'} - Z^N_S| \geqslant a\big) = 0.$$

Then the sequence of processes $\left(Z^N\right)_{N\geqslant 0}$ is tight on $\mathbb{D}([0,\infty), \mathbb{R})$.

Various proofs of the Aldous criterion can be found in the literature, see for e.g., [Bil99, Theorem 16.10], [JS03, Chapter VI, Section 4a], and [EK86, Chapter 3, Theorem 8.6].

Proof. We only need to prove that condition (ii) in Theorem 3.10 holds.
Let $T, a, \varepsilon > 0$. By assumptions, there exists $\delta(\varepsilon) > 0$ and $N(\varepsilon) \geqslant 1$ such that $\forall N \geqslant N(\varepsilon)$, for any $(\mathcal{F}^N)$- stopping times S, S' bounded by T with $S \leqslant S' \leqslant S + \delta(\varepsilon)$.

$$\mathbb{P}(\big|Z^N_{S'} - Z^N_S \geqslant a\big|) \leqslant \varepsilon. \tag{3.1}$$

We define by induction the following sequence of stopping times:

$$\begin{aligned}
&S^N_0 = 0, \\
&S^N_1 = \inf\left\{t > 0 : \left|Z^N_t - Z^N_0\right| > a\right\}, \\
&S^N_k = \inf\left\{t > S^N_{k-1} : \left|Z^N_t - Z^N_{S^N_{k-1}}\right| > a\right\}.
\end{aligned}$$

We notice that if $S^N_{k+1} < \infty$, then $\left|Z^N_{S^N_k} - Z^N_{S^N_{k-1}}\right| \geqslant a$. Consequently, for any $N \geqslant N_1 := N(\varepsilon)$, $k \geqslant 1$,

$$\mathbb{P}\big(S^N_k \leqslant T, S^N_k \leqslant S^N_{k-1} + \delta_1\big) \leqslant \varepsilon,$$

with $\delta_1 = \delta(\varepsilon)$.

Let $q \geqslant 1$ be such that $q\delta_1 > 2T$.

Since $S_q^N = \sum_{k=1}^q S_k^N - S_{k-1}^N$, we have for $N \geqslant N_1$,

$$
\begin{aligned}
T\mathbb{P}\big(S_q^N < T\big) &\geqslant \mathbb{E}\Big[S_q^N \mathbb{1}_{\{S_q^N \leqslant T\}}\Big] = \mathbb{E}\left[\sum_{k=1}^q \big(S_k^N - S_{k-1}^q\big)\mathbb{1}_{\{S_q^N \leqslant T\}}\right] \\
&\geqslant \sum_{k=1}^q \mathbb{E}\Big[\big(S_k^N - S_{k-1}^N\big)\mathbb{1}_{\{S_q^N \leqslant T, S_k^N - S_{k-1}^N > \delta_1\}}\Big] \\
&\geqslant \sum_{k=1}^q \mathbb{E}\Big[\delta_1\Big(\mathbb{1}_{\{S_q^N \leqslant T\}} - \mathbb{1}_{\{S_q^N \leqslant T, S_k^N - S_{k-1}^N \leqslant \delta_1\}}\Big)\Big] \\
&\geqslant \delta_1 q \mathbb{P}\big(S_q^N \leqslant T\big) - \delta_1 q \varepsilon.
\end{aligned}
$$

Since $q\delta_1 > 2T$, we deduce that for $N \geqslant N_1$,

$$
\frac{1}{2}\mathbb{P}\big(S_q^N < T\big) \geqslant \mathbb{P}\big(S_q^N < T\big) - \varepsilon
$$

and thus for $N \geqslant N_1$,

$$
\mathbb{P}\big(S_q^N < T\big) \leqslant 2\varepsilon.
$$

Using the same argument as above, we have for any $N \geqslant N_2 := \max\{N_1, N(\varepsilon/q)\}$, $k \geqslant 1$,

$$
\mathbb{P}\big(S_k^N \leqslant T, S_k^N \leqslant S_{k-1}^N + \delta_2\big) \leqslant \frac{\varepsilon}{q},
$$

with $\delta_2 = \delta(\varepsilon/q)$. We now consider the event

$$
A^N = \big\{S_q^N \geqslant T\big\} \cap \bigcap_{k=1}^q \big\{S_k^N > \inf\big\{T, S_{k-1}^N + \delta_2\big\}\big\}.
$$

We have, for $N \geqslant N_2$,

$$
\begin{aligned}
\mathbb{P}((A^N)^c) &\leqslant \mathbb{P}(S_q^N < T) + \sum_{k=1}^q \mathbb{P}\big(S_k^N \leqslant T \text{ and } S_k^N - S_{k-1}^N \leqslant \delta_2\big) \\
&\leqslant 2\varepsilon + \varepsilon = 3\varepsilon
\end{aligned}
$$

Let $\omega \in A^N$, and introduce $r = \inf\big\{k : S_k^N(\omega) \geqslant T\big\}$. Consider the subdivision of $[0,T]$: $0 = t_0 < t_1 < \ldots < t_r = T$ with $t_k = S_k^N(\omega)$ for $k < r$. By construction of the sequence (S_k^N), we have

$$
w(Z^N(\omega); [t_{k-1}, t_k)) = \sup_{s,t \in [t_{k-1}, t_k)} \big|Z_s^N(\omega) - Z_t^N(\omega)\big| \leqslant 2a,
$$

and $t_k - t_{k-1} \geqslant \delta_2$ for $k < r$ (by the definition of A^N and r). Consequently,

$$
w'_T(Z^N; \delta_2) \leqslant 2a.
$$

Then, the sequence is tight.

□

Definition 3.12 (Definition 3.25 in [JS03]). *A sequence $Z^N = (Z_t^N)_{t \geqslant 0}$ of processes in $\mathbb{D}\big(\mathbb{R}_+, \mathbb{R}^d\big)$ is said C-tight if it is tight and and all its limit points are supported on $\mathcal{C}\big([0,T], \mathbb{R}^d\big)$.*

Proposition 3.13 (Proposition 3.26 in [JS03]). *There is equivalence between*

(i) the sequence (Z^N) is $\mathcal{C}$-tight;

(ii) the sequence (Z^N) is tight and for all $T, a > 0$

$$\lim_{N \to \infty} \mathbb{P}\left(\sup_{t \leqslant T} \big| Z_t^N - Z_{t-}^N \big| > a \right) = 0.$$

Exercise 2. Let us recall a few properties on Poisson processes:

- $Z = (Z_t)_{t \geqslant 0}$ is a Poisson process with parameter λ if $Z_0 = 0$, Z has independent increments, and, for every $t \geqslant 0$, Z_t follows a Poisson distribution with parameter λt.
- It is also known that $M_t := Z_t - \lambda t$ and $N_t := M_t^2 - \lambda t$ are martingales with respect to the natural filtration of the Poisson process.

We consider Z^N a Poisson process with parameter $\lambda_N > 0$. By definition, $Z^N \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$. We assume that the sequence $(\lambda_N)_{N \geqslant 1}$ converges to $\lambda > 0$.

Using Aldous' criterion for the tightness, prove that $(Z^N)_{N \geqslant 1}$ converges in law to the Poisson process Z with parameter λ.

There are two main methods to prove the propagation of chaos of a symmetric interacting particle system:

- by coupling between the particle system and its asymptotic, providing a rate of convergence. It usually requires a good regularity of the coefficients, such as Lipschitz continuity.
- by tightness, consistency, and uniqueness of the nonlinear process. It can be applied to systems with nonregular coefficient, but doesn't provide any rate of convergence.

In the next section, we will detail the study of two interacting systems: the first is a diffusive interacting system and will be analyzed by coupling, and the second one is an interacting system with jumps, which will be studied by tightness method.

4 Analysis of a diffusive interacting particle system

We present in detail in this section a diffusive model of interacting particles, studied by Benachour et al. in [BRTV98] and by Malrieu in [Mal03], which is a particular case of the laboratory example of [Szn91] presented in Section 1.1. The propagation of chaos property will be proved using a coupling method.

We consider the following nonlinear equation

$$\partial_t u = \Delta u + \operatorname{div}(u \nabla W * u), \tag{4.1}$$

where $u(t, \cdot)$ is a time-dependent probability measure on $\mathbb{R}^d$, $W : \mathbb{R}^d \to \mathbb{R}^d$ is an interaction potential, and $*$ stands for the convolution operator:

$$\nabla W * u(x) = \int \nabla W(x-y) u(\mathrm{d}y).$$

This equation arises for example in the modeling of granular media with $W(x) = |x|$ when $d = 1$. The norm of $x \in \mathbb{R}^d$ is denoted by $|x|$.
We work under the following assumptions on the potential W.

Assumption 1. *1. The potential is symmetric: $\forall x \in \mathbb{R}^d$, $W(-x) = W(x)$.*

2. The potential is uniformly convex: there is $\lambda > 0$ such that $\forall x, v \in \mathbb{R}^d$

$$\langle \mathrm{Hess} W(x) v, v \rangle \geqslant \lambda |v|^2,$$

where $\mathrm{Hess} W$ is the hessian matrix of W.

3. The gradient of W is a locally Lipschitz-continuous function with polynomial growth: there exists a polynomial P such that $\forall x, y \in \mathbb{R}^d$,

$$|\nabla W(x) - \nabla W(y)| \leqslant |x-y||P(x) + P(y)|.$$

Contrary to the assumptions of Section 1.1, the function ∇W is neither bounded nor compactly supported.

The stochastic differential equation associated with (4.1) is the following. Let $\overline{X}$ be a Markov process solving the nonlinear SDE

$$\begin{cases} \mathrm{d}\overline{X}_t = \sqrt{2}\mathrm{d}B_t - \nabla W * u_t(\overline{X}_t)\mathrm{d}t, \\ \mathrm{Law}(\overline{X}_0) = u_0, \end{cases} \tag{4.2}$$

where $u_t := \mathrm{Law}(\overline{X}_t)$ is the distribution of $\overline{X}_t$ and B is a Brownian motion. In [BRTV98, Theorem 3.1], the strong existence and uniqueness of the solution to the nonlinear SDE (4.2), when u_0 has a finite moment of order $2(1+r)^2$, with r the degree of the polynomial growth of W, is established. Note that (4.2) can be also be written,

$$\begin{aligned} \overline{X}_t &= \overline{X}_0 + \sqrt{2}B_t - \int_0^t \nabla W * u_s(\overline{X}_s)\mathrm{d}s \\ &= \overline{X}_0 + \sqrt{2}B_t - \int_0^t \int_{\mathbb{R}^d} \nabla W(\overline{X}_s - y) u_s(\mathrm{d}y)\mathrm{d}s \end{aligned}$$

In addition, [BRTV98] proved that the moments of $\overline{X}$ are uniformly bounded in time.

Lemma 4.1 (Proposition 3.10 in [BRTV98], Lemma 5.2 in [Mal03])**.** *Let $q \geqslant 1$. If $\mathbb{E}\left[|X_0|^{2q}\right]$ is finite, then there exists $K_q > 0$ such that*

$$\sup_{t\geqslant 0} \mathbb{E}\left[\left|\overline{X}_t\right|^{2q}\right] \leqslant K_q.$$

■ **Exercise 3.** Using Itô's formula, show that the distribution of $\overline{X}$ is a weak solution to (4.1), in the sense that if $\mathrm{Law}(\overline{X}_t) = u_t(x)\mathrm{d}x$, we have $\forall \varphi \in \mathcal{C}_c^2(\mathbb{R}^d, \mathbb{R})$ test function of class $\mathcal{C}^2$ with compact support, $\forall t \geqslant 0$,

$$\int_{\mathbb{R}^d} \varphi(x) u_t(x)\mathrm{d}x = \int_{\mathbb{R}^d} \varphi(x) u_0(x)\mathrm{d}x + \int_0^t \int_{\mathbb{R}^d} \varphi(x)(\Delta u_s(x) + \nabla.(u_s \nabla W * u_s)(x))\mathrm{d}s.$$

4.1 A first particle system

A natural particle system $\mathbf{X}^N = (X^{1,N}, \dots, X^{N,N})$ associated with (4.2) is the following: for $i \in \{1, \dots, N\}$,

$$\begin{cases} \mathrm{d}X_t^{i,N} = \sqrt{2}\mathrm{d}B_t^i - \frac{1}{N}\sum_{j=1}^N \nabla W\left(X_t^{i,N} - X_t^{j,N}\right)\mathrm{d}t, \\ X_0^{i,N} = X_0^i, \end{cases} \tag{4.3}$$

where $(B^i)_{i\geqslant 1}$ and $(X_0^i)_{i\geqslant 1}$ are two independent collections of independent Brownian motions and independent random variables with common distribution u_0 respectively. The existence of a unique strong solution to (4.3) is established in [BRTV98, Proposition 5.1]. The proof is based on an approximation of ∇W by bounded Lipschitz continuous functions and a control of the moments of the particles.
However, as we will see later, this is not the most suitable system for studying convergence to equilibrium in the long-time regime.

Asymptotic in large population Let us study the system (4.3) in the large-population limit ($N \to \infty$). We prove that propagation of chaos holds, but not uniformly in time. The proof is based on a coupling argument.
To this end, we couple the particle system (4.3) with a collection $(\overline{X}^i)$ of independent nonlinear processes driven by the same Brownian motions and starting from the same initial conditions as the interacting particles.
Let $(\overline{X}^i)_{i\geqslant 1}$ be the solution to the *nonlinear* SDE

$$\begin{cases} \mathrm{d}\overline{X}_t^i = \sqrt{2}\mathrm{d}B_t^i - \nabla W * u_t\left(\overline{X}_t^i\right)\mathrm{d}t, \\ \overline{X}_0^i = X_0^i, \end{cases}$$

where u_t is the distribution of $\overline{X}_t^i$.

Theorem 4.2 (Benachour et al. [BRTV98])**.** *If u_0 has a finite 2-order moment, then there exists a constant $C > 0$, such that for any $T \geqslant 0$ and $N \geqslant 1$,*

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}\left[\left|X_t^{1,N} - \overline{X}_t^1\right|^2\right] \leqslant C\frac{T^2}{N}.$$

We deduce that $(X^{1,N})_{N\geqslant 1}$ converges in distribution to $\overline{X}$. By uniqueness of the solution to the nonlinear SDE (4.2) and Proposition 2.8, the system satisfies the propagation of chaos property.

Proof of Theorem 4.2. We detail the proof when $d=1$, the case $d\geqslant 1$ is similar. We observe that

$$X_t^{i,N}-\overline{X}_t^i=-\frac{1}{N}\sum_{j=1}^N\int_0^t\int\Big(W'\big(X_s^{i,N}-X_s^{j,N}\big)-W'(\overline{X}_s^i-y)\Big)u_s(\mathrm{d}y)\mathrm{d}s.$$

Then, by Itô's Formula,

$$\begin{aligned}
&\sum_{i=1}^N\Big(X_t^{i,N}-\overline{X}_t^i\Big)^2\\
&=-\frac{2}{N}\sum_{i,j=1}^N\int_0^t\int\Big(X_s^{i,N}-\overline{X}_s^i\Big)\Big(W'\big(X_s^{i,N}-X_s^{j,N}\big)-W'(\overline{X}_s^i-y)\Big)u_s(\mathrm{d}y)\mathrm{d}s\\
&=-\frac{2}{N}\sum_{i,j=1}^N\int_0^t(I_{i,j}^1+I_{i,j}^2)\mathrm{d}s,
\end{aligned}$$

with

$$\begin{aligned}
I_{i,j}^1&=\Big(X_s^{i,N}-\overline{X}_s^i\Big)\Big(W'\big(X_s^{i,N}-X_s^{j,N}\big)-W'(\overline{X}_s^i-\overline{X}_s^j)\Big)\\
I_{i,j}^2&=\Big(X_s^{i,N}-\overline{X}_s^i\Big)\int\Big(W'(\overline{X}_s^i-\overline{X}_s^j)-W'(\overline{X}_s^i-y)\Big)u_s(\mathrm{d}y)
\end{aligned}$$

As W is symmetric, we have $W'(0)=0$ and $W'(x)=-W'(-x)$. Therefore, the sum of the first term gives

$$\begin{aligned}
\sum_{i,j=1}^N I_{i,j}^1&=\sum_{i<j}(I_{i,j}^1+I_{j,i}^1)\\
&=\sum_{i<j}\Big(X_s^{i,N}-X_s^{j,N}-\overline{X}_s^i+\overline{X}_s^j\Big)\Big(W'\big(X_s^{i,N}-X_s^{j,N}\big)-W'(\overline{X}_s^i-\overline{X}_s^j)\Big).
\end{aligned}$$

As the potential is uniformly convex, we deduce $(x-y)(W'(x)-W'(y))\geqslant 0$. Consequently,

$$\sum_{i,j=1}^N I_{i,j}^1\geqslant 0,$$

and we deduce that

$$\sum_{j=1}^N\Big(X_t^{i,N}-\overline{X}_t^i\Big)^2\leqslant-\frac{2}{N}\sum_{i,j=1}^N\int_0^t I_{i,j}^2\mathrm{d}s. \tag{4.4}$$

On the other hand, by the Cauchy-Schwartz inequality,

$$\begin{aligned}
\left|\mathbb{E}\left[\sum_{j=1}^N I_{i,j}^2\right]\right|&=\left|\mathbb{E}\left[\Big(X_s^{i,N}-\overline{X}_s^i\Big)\sum_{j=1}^N\Big(W'(\overline{X}_s^i-\overline{X}_s^j)-W'*u_s(\overline{X}_s^i\Big)\right]\right|\\
&\leqslant\mathbb{E}\Big[\Big(X_s^{i,N}-\overline{X}_s^i\Big)^2\Big]^{1/2}h_i(s)^{1/2},
\end{aligned} \tag{4.5}$$

with

$$h_i(s) = \mathbb{E}\left[\left(\sum_{j=1}^{N}\Big(W'(\overline{X}_s^i - \overline{X}_s^j) - W' * u_s(\overline{X}_s^i)\Big)\right)^2\right]$$
$$= \sum_{j=1}^{N}\mathbb{E}\left[\Big(W'(\overline{X}_s^i - \overline{X}_s^j) - W' * u_s(\overline{X}_s^i)\Big)^2\right],$$

by independence of $\overline{X}^i, \overline{X}^j, \overline{X}^k$, and since $u_s = \mathrm{Law}(\overline{X}_s^j) = \mathrm{Law}(\overline{X}_s^k)$ implying that each term in the sum has a mean equal to zero. Since W' is locally Lipschitz-continuous with polynomial growth by assumption, there is $r \geqslant 0$ and $c > 0$ such that

$$|W'(x)| \leqslant c(1 + |x|^{1+r}).$$

By Lemma 4.1, the moments of $\overline{X}$ are uniformly bounded in time, i.e.

$$\sup_{t\geqslant 0}\mathbb{E}[|\overline{X}_t|^{2(1+r)}] < \infty.$$

Consequently, there is a constant $c > 0$ such that $h_i(s) \leqslant Nc^2$, and by symmetry of the system $\mathbb{E}\left[\left(X_s^{i,N} - \overline{X}_s^i\right)^2\right] = \mathbb{E}\left[\left(X_s^{1,N} - \overline{X}_s^1\right)^2\right]$. Therefore, by (4.5), we have

$$\left|\mathbb{E}\left[\sum_{j=1}^{N} I_{i,j}^2\right]\right| \leqslant cN^{1/2}\mathbb{E}\left[\left(X_s^{1,N} - \overline{X}_s^1\right)^2\right]^{1/2}.$$

Finally, using (4.4) and again the symmetry of the system, we obtain

$$N\mathbb{E}\left[\left(X_t^{1,N} - \overline{X}_t^1\right)^2\right] \leqslant 2cN^{1/2}\int_0^t \mathbb{E}\left[\left(X_s^{1,N} - \overline{X}_s^1\right)^2\right]^{1/2} \mathrm{d}s$$

and consequently, for $t \in [0, T]$,

$$N \sup_{0\leqslant t\leqslant T}\mathbb{E}\left[\left(X_t^{1,N} - \overline{X}_t^1\right)^2\right] \leqslant 2cN^{1/2}T\left(\sup_{0\leqslant t\leqslant T}\mathbb{E}\left[\left(X_t^{1,N} - \overline{X}_t^1\right)^2\right]\right)^{1/2}$$

and the result follows. □

Long time behavior of the particle system Since ∇W is odd, the empirical mean of the particle system (4.3), given by

$$\frac{1}{N}\sum_{k=1}^{N} X_t^{k,N} = \frac{1}{N}\sum_{k=1}^{N} X_0^k + \frac{\sqrt{2}}{N}\sum_{k=1}^{N} B_t^k,$$

is the sum of a random variable and a Gaussian process, with time marginal of distribution $\mathcal{N}(0, \frac{2t}{N})$. Consequently, the empirical measure does not converge when $t \to \infty$ to a probability measure. We deduce that the direction $(1, \dots, 1)$ as a bad influence on the long time behavior of the system.

However, since the potential W is strictly convex, it known that the nonlinear process has an invariant measure, and as proved in [CMV03], the solution of(4.1) converges exponentially fast to its equilibrium.
Consequently, the above particle system does not behave as well as the nonlinear process it approximates. Following [Mal03], we introduce a new particle system, which admits an invariant measure.

4.2 A new particle system

We note that the mean of the nonlinear process $\bar{X}$ satisfies

$$\begin{aligned}\mathbb{E}[\overline{X}_t] &= \mathbb{E}[X_0] - \int_0^t \mathbb{E}\big[\nabla W * u_s(\overline{X}_s)\big]\mathrm{d}s \\ &= \mathbb{E}[X_0] - \int_0^t \mathbb{E}\Big[\nabla W(\overline{X}_s - \overline{X}'_s)\Big]\mathrm{d}s,\end{aligned}$$

where $\overline{X}'$ is an independent copy of $\overline{X}$. Since ∇W is an odd function, we deduce that $\mathbb{E}[\overline{X}_t] = \mathbb{E}[X_0]$ for any $t \geqslant 0$. To construct a particle system that provides a good approximation of the nonlinear process, it seems reasonable to introduce a system whose empirical mean is equal to the mean of u_0.

Without lost of generality, we assume in the sequel that $\int x u_0(\mathrm{d}x) = \mathbb{E}[X_0] = 0$. Recall that the system $\mathbf{X}^N$ is defined in (4.3). We introduce the process $\mathbf{Y}^N = \left(Y^{1,N}, \dots, Y^{N,N}\right)$ on $\left(\mathbb{R}^d\right)^N$ defined by

$$Y_t^{i,N} = X_t^{i,N} - \frac{1}{N}\sum_{k=1}^N X_t^{k,N},$$

which is the projection of $\mathbf{X}^N$ on $\mathcal{H} = \left\{\mathbf{x} \in \left(\mathbb{R}^d\right)^N : \sum_{k=1}^N x^k = 0\right\}$. Then, $\forall t \geqslant 0$,

$$\frac{1}{N}\sum_{i=1}^N Y_t^{i,N} = 0.$$

We easily note that

$$X_t^{i,N} - X_t^{j,N} = Y_t^{i,N} - Y_t^{j,N},$$

and we deduce by (4.3),

$$Y_t^{i,N} = X_0^i - \frac{1}{N}\sum_{j=1}^N X_0^j + \sqrt{2}B_t^i - \frac{\sqrt{2}}{N}\sum_{j=1}^N B_t^j - \frac{1}{N}\sum_{j=1}^N \int_0^t \nabla W(Y_s^{i,N} - Y_s^{j,N})\mathrm{d}s.$$

Uniform propagation of chaos

Theorem 4.3 (Theorem 5.1 in [Mal03])**.** *There exists a constant $C > 0$ such that for every $N \geqslant 1$*

$$\sup_{t \geqslant 0} \mathbb{E}\left[\left|Y_t^{1,N} - \overline{X}_t^1\right|^2\right] \leqslant \frac{C}{N}.$$

Proof. We introduce $\overline{\mathbf{Y}}^N = \left(\overline{Y}^{1,N}, \dots, \overline{Y}^{N,N}\right)$ the projection of the nonlinear process $\overline{\mathbf{X}} = \left(\overline{X}^1, \dots, \overline{X}^N\right)$ on $\mathcal{H}$. We have

$$\begin{aligned}\overline{Y}_t^{i,N} &:= \overline{X}_t^i - \frac{1}{N}\sum_{i=1}^N \overline{X}_t^j \\ &= X_0^i - \frac{1}{N}\sum_{j=1}^N X_0^j + \sqrt{2}B_t^i - \int_0^t \nabla W * u_s\left(\overline{X}_s^i\right)\mathrm{d}s - \frac{\sqrt{2}}{N}\sum_{j=1}^N B_t^j + \frac{1}{N}\sum_{j=1}^N \int_0^t \nabla W * u_t(\overline{X}_t^j)\mathrm{d}s.\end{aligned}$$

Note that

$$\left\|Y_t^{i,N} - \overline{X}_t^i\right\|_{L^2} \leqslant \left\|Y_t^{i,N} - \overline{Y}_t^{i,N}\right\|_{L^2} + \left\|\overline{Y}_t^{i,N} - \overline{X}_t^i\right\|_{L^2}.$$

On one hand, by independence of the processes $(\overline{X}^i)_{i\geqslant 1}$, we have

$$\mathbb{E}\left[\left(\overline{Y}_t^{i,N} - \overline{X}_t^i\right)^2\right] = \frac{1}{N}\mathbb{E}\left[\left|\overline{X}_t^1\right|^2\right].$$

By lemma 4.1, the moments of $\overline{X}^1$ are uniformly bounded in time.
On the other hand,

$$\begin{aligned}Y_t^{i,N} - \overline{Y}_t^{i,N} = &-\frac{1}{N}\int_0^t \sum_{j=1}^N \left(\nabla W(Y_s^{i,N} - Y_s^{j,N}) - \nabla W * u_s(\overline{X}_s^i)\right)\mathrm{d}s \\ &- \frac{1}{N}\sum_{j=1}^N \int_0^t \nabla W * u_t(\overline{X}_t^j)\mathrm{d}s.\end{aligned}$$

Using Itô's formula, the same kind of decomposition as in the proof of Theorem 4.2, and that the sum of the coordinates of $\overline{Y}^{i,N}$ and $Y^{i,N}$ are equal to 0 since they belong to $\mathcal{H}$, we deduce that there is a constant $c > 0$ such that

$$\sum_{i=1}^N \mathbb{E}\left[\left|Y_t^{i,N} - \overline{Y}_t^{i,N}\right|^2\right] \leqslant -2\lambda \int_0^t \sum_{i=1}^N \mathbb{E}\left[\left|Y_s^{i,N} - \overline{Y}_s^{i,N}\right|^2\right]\mathrm{d}s + c\sqrt{N}\int_0^t \mathbb{E}\left[\left|Y_s^{i,N} - \overline{Y}_s^{i,N}\right|^2\right]^{1/2}\mathrm{d}s \tag{4.6}$$

By symmetry, we deduce

$$\mathbb{E}\left[\left|Y_t^{1,N} - \overline{Y}_t^{1,N}\right|^2\right] \leqslant -2\lambda \int_0^t \mathbb{E}\left[\left|Y_s^{1,N} - \overline{Y}_s^{1,N}\right|^2\right]\mathrm{d}s + \frac{c}{\sqrt{N}}\int_0^t \mathbb{E}\left[\left|Y_s^{1,N} - \overline{Y}_s^{1,N}\right|^2\right]^{1/2}\mathrm{d}s.$$

Introducing the function $\alpha(t) = \mathbb{E}\left[\left|Y_t^{1,N} - \overline{Y}_t^{1,N}\right|^2\right]^{1/2}$, using a similar proof as the one of the Gronwall lemma, we deduce

$$\mathbb{E}\left[\left(Y_t^{1,N} - \overline{Y}_t^{1,N}\right)^2\right]^{1/2} \leqslant \frac{c}{\lambda\sqrt{N}}(1 - \mathrm{e}^{-\lambda t}) \leqslant \frac{c}{\lambda\sqrt{N}},$$

which gives the expected upper bound. □

■ **Exercise 4.** Using the same kind of decomposition as in the proof of Theorem 4.2, prove Inequality (4.6).

Long time behavior We introduce the Wasserstein distance $\mathcal{W}_2$ between two probability measures ν and μ, with a finite second moment:

$$\mathcal{W}_2(\nu,\mu) := \left(\inf \int |x-y|^2 \pi(\mathrm{d}x,\mathrm{d}y)\right)^{1/2}$$

where the infimum is taken on the probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginales ν and μ.

The potential W being strongly convex by assumption, the nonlinear system admits an invariant measure, and Carrillo, McCann and Villani [CMV03] proves that u_t convergences exponentially fast to its equilibrium u_∞ along a well-adapted functional. More precisely, let us introduce the functional

$$\eta(u) = \int u(x)\log u(x)\mathrm{d}x + \frac{1}{2}\iint W(x-y)u(x)u(y)\mathrm{d}x\mathrm{d}y.$$

Carrillo, McCann and Villani [CMV03] established that there is a constant $K > 0$ such that

$$\eta(u_t) - \eta(u_\infty) \leqslant K\exp(-2\lambda t),$$

where u_∞ is the unique minimizer of η with the same mean of u_0.

Using the particle system and logarithmic Sobolev inequalities, [Mal03] obtain the speed of convergence in Wasserstein distance:

Theorem 4.4 (Theorem 1.4 in [Mal03])**.** *There is a constant $K > 0$ such that*

$$\mathcal{W}_2(u_t, u_\infty) \leqslant K\mathrm{e}^{-\lambda t}.$$

Indeed, [Mal03] proves in Proposition 4.3 that the invariant measure ($t \to \infty$) of the particle system $\mathbf{Y}^N$ is given by

$$u_\infty^N = \frac{1}{z_N}\mathbb{1}_{\mathcal{H}}(y)\exp\left(\frac{1}{2N}\sum_{i,j}^N W(y_i - y_j)\right)\mathrm{d}y,$$

with $z_N = \int_{\mathcal{H}}\exp\left(\frac{1}{2N}\sum_{i,j}^N W(y_i-y_j)\right)\mathrm{d}y$, which satisfies a logarithmic Sobolev inequality with constant $\frac{2}{\lambda}$, that is for every smooth function f

$$\mathrm{Ent}_{u_\infty^N}(f^2) \leqslant \frac{2}{\lambda}\int |\nabla f|^2 \mathrm{d}u_\infty^N,$$

with $\mathrm{Ent}_\mu(f^2) := \int f^2\log f^2\mathrm{d}\mu - \int f^2\mathrm{d}\mu\,\log\left(\int f^2\mathrm{d}\mu\right)$. They deduce (see [Mal03, corollary 4.4]) that the relative entropy satisfies for any $t \geqslant 0$

$$\mathrm{Ent}(u_t^N|u_\infty^N) \leqslant \mathrm{Ent}(u_0^N|u_\infty^N)\mathrm{e}^{-2\lambda t} \tag{4.7}$$

where u_t^N denotes the distribution of $\mathbf{Y}_t^N$ and the relative entropy is defined by $\mathrm{Ent}(\nu|\mu) = \int \log g\mathrm{d}\nu = \int g\log g\mathrm{d}\mu$ with g the density of ν with respect to μ: $\nu(\mathrm{d}x) = g\mu(\mathrm{d}x)$.

Ideas of the proof of Theorem 4.4. Let $N \geqslant 2$ be arbitrary. By the triangular inequality, one has

$$\mathcal{W}_2(u_t, u_\infty) \leqslant \mathcal{W}_2(u_t, u_t^{(1,N)}) + \mathcal{W}_2(u_t^{(1,N)}, u_\infty^{(1,N)}) + \mathcal{W}_2(u_\infty^{(1,N)}, u_\infty),$$

where $u^{(1,N)}$ is the distribution of $Y^{1,N}$. From the uniform propagation of chaos established in Theorem 4.3, we have

$$\mathcal{W}_2(u_t, u_t^{(1,N)}) + \mathcal{W}_2(u_\infty^{(1,N)}, u_\infty) \leqslant 2\sqrt{\sup_{t\geqslant 0} \mathbb{E}\left[\left|Y_t^{1,N} - \overline{X}_t^1\right|^2\right]} \leqslant 2\frac{C}{\sqrt{N}}.$$

Note that, in the absence of uniform propagation of chaos, we cannot deduce anything about $\mathcal{W}_2(u_\infty^{(1,N)}, u_\infty)$.

We thus only have to focus on the term $\mathcal{W}_2(u_t^{(1,N)}, u_\infty^{(1,N)})$. Note that, for μ and ν two probability measures on $\mathbb{R}^d$, we have

$$N\mathcal{W}_2(\mu, \nu)^2 \leqslant \mathcal{W}_2(\mu_N, \nu_N),$$

for any probability measures μ_N and ν_N on $\left(\mathbb{R}^d\right)^N$ with respective marginals $\mu, \ldots, \mu$ and $\nu \ldots, \nu$. In fact, for every $x = (x_1, \ldots, x_N)$ and $y = (y_1, \ldots, y_N)$ in $\left(\mathbb{R}^d\right)^N$, we easily observe that

$$|x - y|^2 = \sum_{i=1}^N |x_i - y_i|^2$$

and, therefore, for every measure $\pi_N = \pi^{\otimes N}$ on $\mathbb{R}^{Nd} \times \mathbb{R}^{Nd}$ with marginals μ_N and ν_N, we have

$$\int |x - y|^2 \pi_N(\mathrm{d}x, \mathrm{d}y) = N \int |x - y|^2 \pi(\mathrm{d}x, \mathrm{d}y)$$

where π is a measure on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν.
Consequently,

$$\mathcal{W}_2(u_t^{(1,N)}, u_\infty^{(1,N)}) \leqslant \frac{1}{\sqrt{N}} \mathcal{W}_2(u_t^N, u_\infty^N),$$

where u^N is the distribution of the particle system $\mathbf{Y}^N$.

We now use the following result (see also [BGL01, Corollary 3.1]).

Theorem 4.5 (Otto-Villani [OV00]). *Let μ be a absolutely continuous probability measure which satisfies a logarithmic Sobolev inequality with constant C.*
Then, for every probability measure ν absolutely continuous with respect to μ, we have

$$\mathcal{W}_2(\mu, \nu)^2 \leqslant \frac{C}{2} \mathrm{Ent}(\nu|\mu).$$

Using also (4.7), we thus deduce

$$\begin{aligned}\mathcal{W}_2(u_t^{(1,N)}, u_\infty^{(1,N)}) &\leqslant \sqrt{\frac{C}{2N} \mathrm{Ent}(u_t^N | u_\infty^N)} \\ &\leqslant \sqrt{\frac{C}{2N} \mathrm{Ent}(u_0^N | u_\infty^N)} \mathrm{e}^{-\lambda t}.\end{aligned}$$

By assumption, $u_0^N = u_0^{\otimes N}$ is the distribution of N independent random variables, and thus $\mathrm{Ent}(u_0^N|u_\infty^N)$ is of order N.
Consequently, we finally obtain the existence of a constant $C > 0$ such that

$$\mathcal{W}_2(u_t, u_\infty) \leqslant \frac{C}{\sqrt{N}} + C\mathrm{e}^{-\lambda t}.$$

We conclude by letting $N \to \infty$, since N was chosen arbitrarily. $\square$

5 Analysis of an interacting particle system with jumps

In the previous section we used a coupling to prove propagation of chaos, in this section the proof is based on tightness and uniqueness of the limiting distribution.

We focus on the toy model of interacting neurons introduced in Section 1.3. This section is based on the articles of De Masi et al. [DMGLP15] and of Fournier and Löcherbach [FL16]. Let us recall the interacting particle system and its nonlinear asymptotic.
Let us consider a family of neurons of size N. We denote by $X_t^{i,N}$ the membrane potential of neuron i at time $t \geqslant 0$.
We assume that initially the potentials are i.i.d: $(X_0^i)_{i\geqslant 0}$ is a sequence of nonnegative i.i.d initial potentials, with distribution g_0 on $\mathbb{R}_+$.
We introduce the two parameters of the model:

- the spiking rate of the system is modeled by the function $f : \mathbb{R}_+ \to \mathbb{R}_+$;
- $\lambda \geqslant 0$ the attraction force to the mean potential.

The evolution in time of the potentials of each neuron is given by the following SDEs:

$$\begin{aligned} X_t^{i,N} =& X_0^i - \lambda \int_0^t \left(X_s^{i,N} - \overline{X}_s^N \right) \mathrm{d}s \qquad (5.1) \\ & - \int_0^t \int_{\mathbb{R}_+} X_s^{i,N} \mathbb{1}_{\left\{z \leqslant f(X_{s-}^{i,N})\right\}} Q^i(\mathrm{d}s, \mathrm{d}z) + \frac{1}{N} \sum_{j\neq i} \int_0^t \int_{\mathbb{R}_+} \mathbb{1}_{\left\{z \leqslant f(X_{s-}^{j,N})\right\}} Q^j(\mathrm{d}s, \mathrm{d}z), \end{aligned}$$

with $\left(Q^i(\mathrm{d}s, \mathrm{d}z)\right)_{i\geqslant N}$ i.i.d. Poisson measures on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $\mathrm{d}s\mathrm{d}z$, independent of $(X_0^i)_{i\geqslant 0}$, and $\overline{X}_t^N = \frac{1}{N}\sum_{i=1}^N X_t^{i,N}$ is the empirical average potential of the system.
The first integral models the attraction to the mean value, the second integral models the jump to 0 of the potential when neuron i spikes, and finally the last integral models the additive potential of size $\frac{1}{N}$ at each spike of another neuron.

Strong well-posedness of the system (5.1) holds under the following assumption on the jump rate, see [FL16, Proposition 2].

Assumption 2. *f is a non-decreasing function of class $\mathcal{C}^1$, with $f(0) = 0$, $f(x) > 0$ for $x > 0$, and* $\lim_{x\to\infty} f(x) = \infty$.

Since f is non decreasing, the higher the potential, the bigger the probability of spiking.
The associated nonlinear SDE is the following

$$X_t = X_0 - \lambda \int_0^t (X_s - \mathbb{E}[X_s])\mathrm{d}s - \int_0^t \int_{\mathbb{R}_+} X_{s-} \mathbb{1}_{\{z\leqslant f(X_{s-})\}} Q(\mathrm{d}s, \mathrm{d}z) + \int_0^t \mathbb{E}[f(X_s)]\mathrm{d}s, \qquad (5.2)$$

where $Q(\mathrm{d}s, \mathrm{d}z)$ is a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $\mathrm{d}s\mathrm{d}z$, and X_0 is independent of Q, with distribution g_0. When g_0 is compactly supported, there is uniqueness of the solution to (5.2) satisfying $\int_0^t \mathbb{E}[X_s f(X_s)]\mathrm{d}s < \infty$ for any $t \geqslant 0$, under Assumption 2 (see [FL16, Theorem 4]).
We easily note that for any $t \geqslant 0$, $X_t^{i,N} \geqslant 0$ and $X_t \geqslant 0$, and that they are càdlàg functions (in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$).

We prensent here the proof of [FL16, Theorem 5] on the propagation of chaos for the system based on tightness arguments.

Theorem 5.1. *Assume that the initial potentials are integrable, i.e.* $\int x g_0 \mathrm{d}x < \infty$*. Then,*

1. *The sequence of empirical measures* $\bar{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ *is tight in* $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$*.*
2. *If the initial distribution is compactly supported, there is propagation of chaos to the distribution of the solution to* (5.2)*.*

The assumption to obtain propagation of chaos can be relaxed as explained in [FL16]. In addition, under stronger condition on the jump rate f, using a coupling method and a well adapted distance, [FL16, Theorem 7] gives a speed of convergence to the chaos. They also obtain the invariant measures (when $t \to \infty$) of the nonlinear SDE (5.2) under additional assumptions in [FL16, Theorem 8]: there are two possible equilibrium, either $g(\mathrm{d}x) = \delta_0$ or

$$
\begin{aligned}
g(\mathrm{d}x) &= \frac{p}{p + \lambda m - \lambda x} \exp\left(-\int_0^x \frac{f(y)}{p + \lambda(m-y)}\mathrm{d}y\right) \mathbb{1}_{0 \leqslant x \leqslant m + p/\lambda} \mathrm{d}x \quad \text{when } \lambda > 0, \\
g(\mathrm{d}x) &= \exp\left(-\frac{1}{p}\int_0^x f(y)\mathrm{d}y\right)\mathrm{d}x \quad \text{when } \lambda = 0,
\end{aligned}
$$

where $p, m > 0$ are constants determined by the constraints $\int g(\mathrm{d}x) = 1$ and $\int x g(\mathrm{d}x) = m$.

5.1 Tightness

By exchangeability of the system and by Proposition 2.8, the tightness of $\left(\bar{\mu}^N\right)_{N\geqslant 1}$ is equivalent to the tightness of $\left(\mathrm{Law}(X^{1,N})\right)_{N\geqslant 1}$.
Since the trajectories of $X^{1,N}$ are in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$, we use the Aldous criterion to prove tightness of $\left(\mathrm{Law}(X^{1,N})\right)_{N\geqslant 1}$ (see Theorem 3.11). We have to prove

(i) for all $T > 0$, $\lim_{a\to\infty} \limsup_{N\to\infty} \mathbb{P}\left(\sup_{t \leq T} |X_t^{1,N}| \geqslant a\right) = 0$;

(ii) for $T > 0$ and for all $a > 0$,

$$
\lim_{\delta \downarrow 0} \limsup_{N} \sup_{\substack{S, S' \text{stopping times:} \\ S \leqslant S' \leqslant S + \delta \leqslant T}} \mathbb{P}\left(|X_{S'}^{1,N} - X_S^{1,N}| \geqslant a\right) = 0.
$$

To prove Condition (i), we control $X_t^{1,N}$ with its initial value $X_0^{1,N}$ and the initial mean value $\overline{X}_0^N$. Recall that $X_t^{1,N} \geqslant 0$ for any $t \geqslant 0$.

■ **Exercise 5.** The aim of this exercise is to prove that for any $t \geqslant 0$,

$$
X_t^{i,N} \leqslant X_0^{i,N} + (4\lambda t + 4)\left(\overline{X}_0^N + Z_t^N\right), \tag{5.3}
$$

with $Z_t^N = \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{\mathbb{R}_+} \mathbb{1}_{\{z \leqslant f(2)\}} Q^i(\mathrm{d}s, \mathrm{d}z)$.

1. Prove that $x - 1 \geqslant \frac{x+1}{3} - \frac{4}{3}\mathbb{1}_{x \leqslant 2}$ for all $x \geqslant 0$.
2. From (5.1), give the equation satisfied by $\overline{X}_t^N = \frac{1}{N}\sum_{i=1}^N X_t^{i,N}$.

3. Using the two first questions, prove that

$$\frac{1}{N}\sum_{i=1}^{N}\int_0^t\int_{\mathbb{R}_+}\left(X_{s-}^{i,N}+1\right)\mathbb{1}_{\left\{z\leqslant f(X_{s-}^{i,N})\right\}}Q^i(\mathrm{d}s,\mathrm{d}z)\leqslant 3\overline{X}_0^N+4Z_t^N. \tag{5.4}$$

4. Deduce $\overline{X}_t^N\leqslant 4\overline{X}_0^N+4Z_t^N$.

5. Conclude to obtain (5.3).

From the above estimate, we have for any $T>0$,

$$\sup_{t\leqslant T}X_t^{i,N}\leqslant X_0^{i,N}+(4\lambda T+4)\left(\overline{X}_0^N+Z_T^N\right),$$

with $Z_T^N=\frac{1}{N}\sum_{i=1}^N\int_0^T\int_{\mathbb{R}_+}\mathbb{1}_{\{z\leqslant f(2)\}}Q^i(\mathrm{d}s,\mathrm{d}z)$ the empirical mean of N i.i.d. Poisson($f(2)T$) random variables. Consequently,

$$\sup_{N\geqslant 1}\mathbb{E}[\sup_{t\leqslant T}X_t^{i,N}]\leqslant\mathbb{E}[X_0]+(4\lambda T+4)(\mathbb{E}[X_0]+f(2)T)<\infty.$$

Using Markov inequality, we easily deduce Condition (i) of Aldous' criterion.

Let us now study Condition (ii). Let $T>0$ and $\delta>0$.
Remark: Note that, in this proof, C_T denotes a generic constant arising in the tightness argument, whose value may change from one line to another.

Taking the expectation in (5.4) and using the exchangeability, we have for any $N\geqslant 1$ and $t\in[0,T]$,

$$\int_0^t\mathbb{E}\big[X_s^{1,N}f\big(X_s^{1,N}\big)\big]\mathrm{d}s\leqslant 3\mathbb{E}[X_0]+4f(2)t\leqslant C_T, \tag{5.5}$$

with $C_T=3\mathbb{E}[X_0]+4f(2)T$.
We consider S,S' two $(\mathcal{F}_t)_{t\geqslant 0}$-stopping times, with $S\leqslant S'\leqslant S+\delta\leqslant T$. We have

$$\begin{aligned}\left|X_{S'}^{1,N}-X_S^{1,N}\right|&\leqslant\int_S^{S'}\int_{\mathbb{R}_+}X_u^{1,N}\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{1,N})\right\}}Q^1(\mathrm{d}u,\mathrm{d}z)+\frac{1}{N}\sum_{i=2}^N\int_S^{S'}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{i,N})\right\}}Q^i(\mathrm{d}u,\mathrm{d}z)\\&\quad+\lambda\int_S^{S'}X_u^{1,N}\mathrm{d}u+\lambda\int_S^{S'}\overline{X}_u^N\mathrm{d}u\\&=I_1+I_2+I_3+I_4.\end{aligned}$$

First note that, if I_1 is positive then there was at least one jump the $X^{1,N}$, and therefore

$$\begin{aligned}\mathbb{P}(I_1>0)&\leqslant\mathbb{P}\left(\int_S^{S'}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{1,N})\right\}}Q^i(\mathrm{d}u,\mathrm{d}z)\geqslant 1\right)\\&\leqslant\mathbb{E}\left[\int_S^{S+\delta}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{1,N})\right\}}Q^i(\mathrm{d}u,\mathrm{d}z)\right]=\mathbb{E}\left[\int_S^{S+\delta}f(X_u^{1,N})\mathrm{d}u\right],\end{aligned}$$

where the Markov inequality has been used in the second inequality. Since f is non-decreasing, it satisfies for any $m > 0$ and $x \geqslant 0$, $f(x) \leqslant f(m) + \frac{x}{m} f(x)$, and thus

$$\mathbb{P}(I_1 > 0) \leqslant \delta f(m) + \frac{1}{m} \int_0^T \mathbb{E}\big[X_u^{1,N} f(X_u^{1,N})\big] \mathrm{d}u \leqslant \delta f(m) + \frac{C_T}{m} = \sqrt{\delta} + \frac{C_T}{f^{-1}(\delta^{-1/2})},$$

by (5.5) and for $m = f^{-1}(\delta^{-1/2})$, where f^{-1} is the generalized inverse function: $f^{-1}(y) = \inf\{x \geqslant 0 : f(x) \geqslant y\}$.
Similarly, by exchangeability,

$$\mathbb{E}[I_2] \leqslant \frac{N-1}{N} \mathbb{E}\left[\int_S^{S+\delta} f(X_u^{1,N})\right] \mathrm{d}u \leqslant \sqrt{\delta} + \frac{C_T}{f^{-1}(\delta^{-1/2})}.$$

For the two next terms, we use that for $m > 0$ and $x \geqslant 0$, $x \leqslant m + x \frac{f(x)}{f(m)}$, and then

$$\mathbb{E}[I_3] \leqslant \lambda m \delta + \frac{\lambda}{f(m)} \int_0^T \mathbb{E}\big[X_u^{1,N} f\big(X_u^{1,N}\big)\big] \mathrm{d}u \leqslant \lambda m \delta + \frac{\lambda C_T}{f(m)} = \lambda\sqrt{\delta} + \frac{\lambda C_T}{f(\delta^{-1/2})},$$

by (5.5) and for $m = \delta^{-1/2}$. Similarly,

$$\mathbb{E}[I_3] \leqslant \lambda\sqrt{\delta} + \frac{\lambda C_T}{f(\delta^{-1/2})}.$$

Let $a > 0$, we have

$$\begin{aligned}\mathbb{P}\Big(\Big|X_{S'}^{1,N} - X_S^{1,N}\Big| \geqslant a\Big) &\leqslant \mathbb{P}\left(\bigcup_{i=1}^4 I_i \geqslant \frac{a}{4}\right) \leqslant \mathbb{P}(I_1 > 0) + \mathbb{P}(I_2 > a/4) + \mathbb{P}(I_3 > a/4) + \mathbb{P}(I_4 > a/4) \\ &\leqslant \sqrt{\delta} + \frac{C_T}{f^{-1}(\delta^{-1/2})} + \frac{4}{a}\left((1+2\lambda)\sqrt{\delta} + \frac{C_T}{f^{-1}(\delta^{-1/2})} + \frac{2\lambda C_T}{f(\delta^{-1/2})}\right),\end{aligned}$$

by the Markov inequality. The upper bound is uniform in N, S, S', and goes to 0 when $\delta \to 0$ because $\lim_{x\to\infty} f(x) = \infty$ by assumption. Consequently, Condition (ii) is satisfied, and thus $(X^{1,N})_{N\geqslant 1}$ is tight in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$.

5.2 Propagation of chaos

By Theorem 3.6, since there is uniqueness of the solution to nonlinear SDE (5.2) and the sequence $\bar{\mu}^N$ is tight, we only have to prove the consistency: the limiting distributions of $\bar{\mu}^N$ are solution of the martingale problem associated to (5.2) (see [JS03, Theorems II.2.42 and III.2.26]).
Let g_0 be the initial distribution. We denote by Z the canonical process on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$: $Z_t(\omega) = \omega_t$. For $t \geqslant 0$, define the projection $p_t : \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+) \to \mathbb{R}_+$ by $p_t(z) = z_t$. Assume that $\bar{\mu}^N$ converges in law to μ in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$ (up to a subsequence). We want to show that

(a) $\mu \circ p_0^{-1} = g_0$;

(b) $\forall t \geqslant 0$, $\mathbb{E}_\mu\Big[\int_0^t Z_s f(Z_s) \mathrm{d}x\Big] < \infty$, where $\mathbb{E}_\mu$ is the expectation under the distribution μ;

(c) for all $\varphi \in \mathcal{C}_b^2(\mathbb{R}_+, \mathbb{R})$, i.e. a bounded function of class $\mathcal{C}^2$ with bounded derivatives,

$$\varphi(Z_t) - \varphi(Z_0) - \int_0^t (\varphi(0) - \varphi(Z_u))f(Z_s)\mathrm{d}s - \int_0^t \varphi'(Z_s)\Big[\mathbb{E}_\mu[f(Z_s)] + \lambda(\mathbb{E}_\mu[Z_s] - Z_s)\Big]\mathrm{d}s$$

is a μ-martingale.

Note that [FL16] proved that $\mu(Z_t - Z_{t-} \neq 0) = 0$ for all $t \geqslant 0$ (see Step 2 in the proof of Theorem 5, Section 4).

First, we observe that $\mu \circ \pi_0^{-1}$ is the limit in law of $\bar{\mu}^N \circ \pi_0^{-1} = \frac{1}{N}\sum_{i=1}^N \delta_{X_0^{i,N}}$. Since $X_0^{i,N}$ are i.i.d. random variables with distribution g_0, by the Law of Large Numbers we deduce that $\mu \circ \pi_0^{-1} = g_0$. Second, for $t \geqslant 0$ and $K > 0$, by Fatou Lemma, and since $x \mapsto xf(x) \wedge K$ is a bounded function,

$$\begin{aligned}\mathbb{E}_\mu\left[\int_0^t (Z_s f(Z_s) \wedge K)\mathrm{d}s\right] &\leqslant \liminf_N \mathbb{E}_{\bar{\mu}^N}\left[\int_0^t (Z_s f(Z_s) \wedge K)\mathrm{d}s\right] \\ &= \liminf_N \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\int_0^t (X_s^{i,N} f(X_s^{i,N}) \wedge K)\mathrm{d}s\right] \leqslant C_t < \infty\end{aligned}$$

by (5.5), with C_t the constant introduced in Section 5.1 independent of K . The conclusion follows by letting $K \to \infty$ and using the monotone convergence theorem.
Finally, we prove Point (c). To this end, let $k \geqslant 1$, $0 \leqslant s_1 < \ldots < s_k < s < t$, $\varphi_1, \ldots, \varphi_k \in \mathcal{C}_b(\mathbb{R}_+)$, any $\varphi \in \mathcal{C}_b^2(\mathbb{R}_+)$, and define

$$\begin{aligned}F(\mu) := \mathbb{E}_\mu\Bigg[&\varphi_1(Z_{s_1})\ldots\varphi_1(Z_{s_k})\times \\ &\left(\varphi(Z_t) - \varphi(Z_s) - \int_s^t f(Z_u)(\varphi(0) - \varphi(Z_u))\mathrm{d}u - \int_s^t \varphi'(Z_u)\Big(\mathbb{E}_\mu[f(Z_u)] + \lambda(\mathbb{E}_\mu[Z_u] - Z_u)\Big)\mathrm{d}u\right)\Bigg].\end{aligned}$$

We want to show that $F(\mu) = \lim_{N\to\infty}\mathbb{E}[F(\bar{\mu}^N)] = 0$. [FL16] proves the continuity of the function F, ensuring that $F(\mu) = \lim_{N\to\infty}\mathbb{E}[F(\bar{\mu}^N)]$. We only detail here the proof of the convergence $\lim_{N\to\infty}\mathbb{E}[F(\bar{\mu}^N)] = 0$ to conclude $F(\mu) = 0$ and therefore μ is a solution of the martingale problem associated to (5.2).
The remainder of the proof uses Itô's formula for jump processes. If you are not familiar with this formula, you can stop reading here.

Remark: C_F denotes a generic constant arising in the consistency argument, whose value may change from one line to another.

We have

$$\begin{aligned}F(\bar{\mu}^N) = \frac{1}{N}\sum_{i=1}^N &\varphi_1(X_{s_1}^{i,N})\ldots\varphi_1(X_{s_k}^{i,N})\times \\ &\Bigg[\varphi(X_t^{i,N}) - \varphi(X_s^{i,N}) - \int_s^t f(X_u^{i,N})\big(\varphi(0) - \varphi(X_u^{i,N})\big)\mathrm{d}u - \lambda\int_s^t \varphi'(X_u^{i,N})(\overline{X}_u^N - X_u^{i,N})\mathrm{d}u \\ &\qquad - \frac{1}{N}\sum_{j=1}^N \int_s^t \varphi'(X_u^{i,N})f(X_u^{j,N})\mathrm{d}u\Bigg].\end{aligned}$$

Using Itô formula for the system (5.1) of jump processes, we have

$$
\begin{aligned}
&\varphi(X_t^{i,N})\\
&=\varphi(X_0^{i,N})+\int_0^t\int_{\mathbb{R}_+}\Big(\varphi(0)-\varphi(X_{u-}^{i,N})\Big)\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{i,N})\right\}}Q^i(\mathrm{d}u,\mathrm{d}z)+\lambda\int_0^t\varphi'(X_{u-}^{i,N})(\overline{X}_u^N-X_u^{i,N})\mathrm{d}u\\
&+\sum_{j\neq i}\int_0^t\int_{\mathbb{R}_+}\left(\varphi\left(X_{u-}^{j,N}+\frac{1}{N}\right)-\varphi(X_{u-}^{j,N})\right)\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{j,N})\right\}}Q^j(\mathrm{d}u,\mathrm{d}z).
\end{aligned}
$$

Introducing the compensator $\widetilde{Q}^i(\mathrm{d}u,\mathrm{d}z)=Q^i(\mathrm{d}u,\mathrm{d}z)-\mathrm{d}u\mathrm{d}z$, defining

$$
\begin{aligned}
M_t^{i,N}&=\int_0^t\int_{\mathbb{R}_+}\Big(\varphi(0)-\varphi(X_{u-}^{i,N})\Big)\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{i,N})\right\}}\widetilde{Q}^i(\mathrm{d}u,\mathrm{d}z)\\
\Delta_t^{i,N}&=\sum_{j\neq i}\int_0^t\int_{\mathbb{R}_+}\left(\varphi\left(X_{u-}^{j,N}+\frac{1}{N}\right)-\varphi(X_{u-}^{j,N})\right)\mathbb{1}_{\left\{z\leqslant f(X_{u-}^{j,N})\right\}}Q^j(\mathrm{d}u,\mathrm{d}z)\\
&\quad-\int_0^t\varphi'(X_u^{i,N})\frac{1}{N}\sum_{j=1}^N f(X_u^{j,N})\mathrm{d}u,
\end{aligned}
$$

we observe that

$$
F(\bar{\mu}^N)=\frac{1}{N}\sum_{i=1}^N\varphi_1(X_{s_1}^{i,N})\ldots\varphi_1(X_{s_k}^{i,N})\Big[\Big(M_t^{i,N}-M_s^{i,N}\Big)+\Big(\Delta_t^{i,N}-\Delta_s^{i,N}\Big)\Big].
$$

It is known that $(M^{i,N})$ are $(\mathcal{F}_t)_{t\geqslant 0}$-martingales. In addition, the Poisson measures $(Q^i)_{i\geqslant 1}$ being independent, the martingales $M^{i,N}$ are orthogonal. Using exchangeability, and the boundedness of φ_k, there exist a constant $C_F>0$ such that

$$
\mathbb{E}[\big|F(\bar{\mu}^N)\big|]\leqslant\frac{C_F}{\sqrt{N}}\mathbb{E}\left[\Big(M_t^{1,N}-M_t^{1,N}\Big)^2\right]^{1/2}+C_F\,\mathbb{E}\Big[\Big|\Delta_t^{1,N}\Big|+\big|\Delta_s^{1,N}\big|\Big]
$$

As φ is bounded, $f(x)\leqslant f(1)+xf(x)$, and (5.5), we also have

$$
\mathbb{E}\left[\Big(M_t^{1,N}-M_t^{1,N}\Big)^2\right]=\int_s^t\mathbb{E}\Big[\big(\varphi(0)-\varphi(X_u^{i,N})\big)^2f(X_u^{i,N})\Big]\mathrm{d}u\leqslant C_F.
$$

Using similar computations, we also obtain $\mathbb{E}\Big[\Big|\Delta_t^{1,N}\Big|\Big]\leqslant\frac{C_F}{\sqrt{N}}$ for all $t\in[0,T]$, and thus we deduce $\lim_{N\to\infty}\mathbb{E}\big[\big|F(\bar{\mu}^N)\big|\big]=0$.

Indeed, we observe that

$$\begin{aligned}\left|\Delta_t^{1,N}\right| \leqslant& \int_0^t \int_{\mathbb{R}_+} \left|\varphi\left(X_{u-}^{1,N} + \frac{1}{N}\right) - \varphi(X_{u-}^{1,N})\right| \mathbb{1}_{\left\{z \leqslant f(X_{u-}^{1,N})\right\}} Q^1(\mathrm{d}u, \mathrm{d}z) \\ &\left|\sum_{j=1}^N \int_0^t \int_{\mathbb{R}_+} \left|\varphi\left(X_{u-}^{j,N} + \frac{1}{N}\right) - \varphi(X_{u-}^{j,N})\right| \mathbb{1}_{\left\{z \leqslant f(X_{u-}^{j,N})\right\}} \widetilde{Q}^j(\mathrm{d}u, \mathrm{d}z)\right| \\ &+ \sum_{j=1}^N \int_0^t \left(\varphi\left|X_u^{j,N} + \frac{1}{N}\right| - \varphi(X_u^{j,N}) - \frac{1}{N}\varphi'(X_u^{i,N})\right) f(X_u^{j,N}) \mathrm{d}u \\ &= J_1 + J_2 + J_3.\end{aligned}$$

Since φ' is bounded, and by (5.5), there is a constant $c > 0$ such that

$$\mathbb{E}[J_i] \leqslant \frac{c}{N} \int_0^t \mathbb{E}\big[f(X_u^{1,N})\big] \mathrm{d}u \leqslant \frac{C_F}{N}.$$

Similarly, since φ'' is bounded

$$\mathbb{E}[J_3] \leqslant \frac{c}{N^2} \sum_{j=1}^N \int_0^t \mathbb{E}\big[f(X_u^{1,N})\big] \mathrm{d}u \leqslant \frac{C_F}{N}.$$

Finally using the independence between the Poisson measures $(Q_i)_{i \geqslant 1}$ and the boundness of φ', we obtain

$$\mathbb{E}\Big[(J_2)^2\Big] = \sum_{j=1}^N \int_0^t \mathbb{E}\left[\left(\varphi\left(X_u^{j,N} + \frac{1}{N}\right) - \varphi(X_u^{j,N})\right)^2 f(X_u^{j,N})\right] \leqslant \frac{C_F}{N}.$$

Combining all these bounds, we deduce that $\mathbb{E}\Big[\Big|\Delta_t^{1,N}\Big|\Big] \leqslant \frac{C_F}{\sqrt{N}}$.

References

[Ald78] David Aldous. Stopping times and tightness. *Ann. Probability*, 6(2):335–340, 1978.

[BGL01] Sergey G. Bobkov, Ivan Gentil, and Michel Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7):669–696, 2001.

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

[BP87] Richard F. Bass and Étienne Pardoux. Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields*, 76(4):557–572, 1987.

[BRTV98] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.*, 75(2):173–201, 1998.

[CD22a] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. I. models and methods. *Kinetic and Related Models*, 15(6):895, 2022.

[CD22b] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. II. applications. *Kinetic and Related Models*, 15(6):1017, 2022.

[CMV03] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[DMGLP15] Anna De Masi, Antonio Galves, Eva Löcherbach, and E. Presutti. Hydrodynamic limit for interacting neurons. *J. Stat. Phys.*, 158(4):866–902, 2015.

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

[Fer02] E. Robert Fernholz. *Stochastic portfolio theory*, volume 48 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2002. Stochastic Modelling and Applied Probability.

[FIKP13] E. Robert Fernholz, Tomoyuki Ichiba, Ioannis Karatzas, and Vilmos Prokaj. Planar diffusions with rank-based characteristics and perturbed Tanaka equations. *Probab. Theory Related Fields*, 156(1-2):343–374, 2013.

[FL16] Nicolas Fournier and Eva Löcherbach. On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1844–1876, 2016.

[FPPZN25] Raphaël Forien, Guodong Pang, Etienne Pardoux, and Arsene Brice Zotsa-Ngoufack. Stochastic epidemic models with varying infectivity and waning immunity. *The Annals of Applied Probability*, 35(3):2175–2216, 2025.

[GZN25] Hélène Guérin and Arsene Brice Zotsa-Ngoufack. A stochastic epidemic model with memory of the last infection and waning immunity, 2025. arXiv:2505.00601.

[IKS13] Tomoyuki Ichiba, Ioannis Karatzas, and Mykhaylo Shkolnikov. Strong solutions of stochastic equations with rank-based coefficients. *Probab. Theory Related Fields*, 156(1-2):229–248, 2013.

[JR13] Benjamin Jourdain and Julien Reygner. Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. *Stoch. Partial Differ. Equ. Anal. Comput.*, 1(3):455–506, 2013.

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

[KS18] Praveen Kolli and Mykhaylo Shkolnikov. SPDE limit of the global fluctuations in rank-based models. *Ann. Probab.*, 46(2):1042 – 1069, 2018.

[Mal03] Florent Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[SV06] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.