



**HAL**  
open science

# Vlasov limit for a chain of oscillators with Kac potentials

Alejandro Fernandez Montero

► **To cite this version:**

Alejandro Fernandez Montero. Vlasov limit for a chain of oscillators with Kac potentials. 2020.  
hal-02957947

**HAL Id: hal-02957947**

**<https://hal.science/hal-02957947v1>**

Preprint submitted on 5 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Vlasov limit for a chain of oscillators with Kac potentials

Alejandro Fernandez Montero \*

## Abstract

We consider a chain of anharmonic oscillators with local mean field interaction and long-range stochastic exchanges of velocity. Even if the particles are not exchangeable, we prove the convergence of the empirical measure associated with this chain to a solution of a Vlasov-type equation. We then use this convergence to prove energy diffusion for a restricted class of anharmonic potentials.

## 1 Model and results

### 1.1 Introduction

The study of chains of interacting oscillators has drawn a lot of attention over the past few years. Deriving Fourier's law from an anharmonic chain is a major open problem in statistical mechanics [4]. Mathematically, the model consists in a system of  $N$  oscillators, whose displacement and momentum are denoted by  $X^i \in \mathbb{R}^d$  and  $V^i \in \mathbb{R}^d$  respectively for  $1 \leq i \leq N$ . Particles interact via a Hamiltonian dynamics, with Hamiltonian given by

$$\mathcal{H} = \sum_{i=1}^N \left( \frac{1}{2} |V^i|^2 + \frac{1}{2} \sum_k \phi_k W(X^i - X^{i+k}) + U(X^i) \right),$$

where generally  $(\phi_k)_{k \in \mathbb{Z}}$  is such that  $\phi_k = 0$  when  $|k| > K$  for some integer value  $K$ , *i.e.* the interaction is only between oscillators with close lattice index.  $W$  is a pair potential modelling the interaction between particles and  $U$  is a pinning potential. It is known since [20] that for nearest neighbor harmonic interaction (*i.e.*  $K = 1$ ,  $W(x) = x^2$ ,  $U(x) = x^2$ ) the transport of energy is ballistic and therefore Fourier's law is not valid. The study of the anharmonic chain seems nevertheless out of reach for the moment. However, the model has drawn attention over the past few years with on the one hand the proof of the convergence to the unique invariant measure for anharmonic chains coupled to two heat baths with different temperatures (see *e.g.* [9] and [5]), and on the other hand the study of energy transport in harmonic chains with additional conservative stochastic collisions that enable to derive hydrodynamic limits (see [1] for a review). In fact, stochastic collisions give enough ergodicity to derive such limits, but for long time scales, calculations rely heavily on the harmonic structure of the interactions.

In this paper, we consider a chain with so-called Kac potentials (see [19] for a detailed introduction), *i.e.* we define the coefficients  $\phi_k$  by

$$\phi_k = \frac{1}{\ell N} \phi \left( \frac{k}{\ell N} \right), \quad (1)$$

for  $|k| \leq \ell N$ , where  $\ell$  is a small parameter and  $\phi$  is a smooth even function, with support included in  $[-1/2, 1/2]$  and normalized so that  $\int_{-1/2}^{1/2} \phi(r) dr = 1$ . Therefore, the model has a

---

\*CMAP, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex-France; E-mail: [alejandromontero@polytechnique.edu](mailto:alejandromontero@polytechnique.edu)



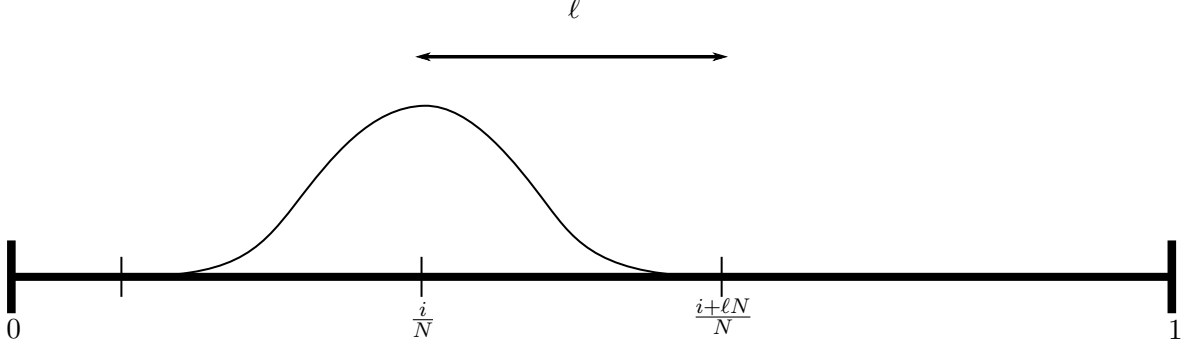


Figure 1 – Different scales of the model and shape of the functions  $\Phi_\ell$  and  $\Gamma_\ell$

To properly define the dynamics for the whole system, we should specify the correlations between the point processes  $\mathcal{N}^{\mu^N, r^i}$ , which is essentially that the jumps between two different pairs are independent. The detailed definition of those measures is left to Section 2. The equation (3) is thus a mean field type equation and we expect that the empirical measure  $\mu_t^N$  converges to a solution of the following Vlasov-type equation

$$\begin{aligned} \partial_t f_t + v \cdot \nabla_x f_t - \left( \int_{\mathbb{T} \times E} \Phi_\ell(r - r') \nabla W(x - x') f_t(r', z') dr' dz' + \nabla U(x) \right) \cdot \nabla_v f_t \\ = \bar{\gamma} \int_{\mathbb{T} \times E} \Gamma_\ell(r - r') (f_t(r, x, v') f_t(r', x', v) - f_t(r, z) f_t(r', z')) dr' dz'. \end{aligned} \quad (6)$$

We will actually prove this statement by introducing the following nonlinear martingale problem. For any measure  $\nu \in \mathcal{M}^1(\mathbb{T} \times E)$ , the space of probability measures on  $\mathbb{T} \times E$ , define the operator  $\mathcal{L}[\nu]$  by

$$\mathcal{L}[\nu]\psi = \mathcal{A}[\nu]\psi + \bar{\gamma}\mathcal{S}[\nu]\psi, \quad (7)$$

for all  $\psi \in C_b^1(\mathbb{T} \times E)$ , the space of bounded continuously differentiable real functions of  $\mathbb{T} \times E$ .  $\mathcal{A}[\nu]$  is a drift operator given by

$$\mathcal{A}[\nu]\psi(r, x, v) = v \cdot \nabla_x \psi(r, x, v) - \left( \int_{\mathbb{T} \times E} \Phi_\ell(r - r') \nabla W(x - x') d\nu(r', z') + \nabla U(x) \right) \cdot \nabla_v \psi(r, x, v),$$

and

$$\mathcal{S}[\nu]\psi(r, x, v) = \int_{\mathbb{T} \times E} (\psi(r, x, v') - \psi(r, x, v)) \Gamma_\ell(r - r') d\nu(r', z')$$

is a pure jump operator that exchanges velocities. Denote by  $\mathcal{D} = D(\mathbb{R}_+, \mathbb{T} \times E)$  the set of right continuous functions with left limits on  $\mathbb{R}_+$  with values in  $\mathbb{T} \times E$ , by  $\mathcal{F}$  the product  $\sigma$ -field on  $\mathcal{D}$ , and let  $(\mathcal{F}_t)_{t \geq 0}$  be the filtration generated by the canonical process  $Y = (r, Z) = (r, X, V)$  on  $(\mathcal{D}, \mathcal{F}, \mu)$ . Then the probability measure  $\mu \in M^1(\mathcal{D})$  is said to solve the *nonlinear martingale problem* starting at  $\nu_0 \in \mathcal{M}^1(\mathbb{T} \times E)$  if  $\mu_0 = \nu_0$  and, for any  $\psi \in C_b^1(\mathbb{T} \times E)$ ,

$$M_t^\psi = \psi(Y_t) - \psi(Y_0) - \int_0^t \mathcal{L}[\mu_s]\psi(Y_s) ds \quad (8)$$

is a martingale under  $\mu$ ,  $\mu_s$  denoting the time marginal of  $\mu$  at time  $s$ . It is then straightforward to check that a solution to the martingale problem is a weak solution to the Vlasov-type equation (6). In particular,  $r_t = r_0$  for any  $t > 0$ , and we will indifferently use either the notation  $r_t$ ,  $r_0$  or  $r$  to refer to the spatial coordinate of  $(Y_t)_{t \geq 0}$  in  $\mathbb{T}$ . The martingale problem (8) is nonlinear in the sense that it is defined using the marginals  $(\mu_t)_{t \geq 0}$ , which are themselves only defined as a

byproduct of the solutions to this problem. Taking expectations in (8), it is then straightforward to check that the flow of time-marginals  $(\mu_t)_{t \geq 0}$  associated with a solution  $\mu$  to the martingale problem is a weak solution to the Vlasov-type equation (6) in the sense of distributions.

Several existence and uniqueness results exist for the solutions of SDEs and nonlinear martingale problems with similar Poisson-type exchanges in the case of the Kac model (see for instance [7, 13], and also [15] for a model with additional mean field interaction). Due to the non-exchangeability of the particles' coordinates in our model those results cannot be applied straightforwardly in our setting and we have to study it from scratch.

### 1.3 Results

We will always assume that the following hypotheses hold

**(H1)**  $W, U \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ , the set of twice differentiable functions of  $\mathbb{R}^d$  taking values in  $\mathbb{R}$ . Moreover  $\nabla W$  and  $\nabla U$  are uniformly Lipschitz and satisfy  $\nabla W(0) = \nabla U(0) = 0$ . There exists a positive constant  $c$  such that for all  $x \in \mathbb{R}^d$ ,

$$|x|^2 \leq cU(x), \quad |\nabla W(x)|^2 \leq cW(x).$$

**(H2)**  $\phi, \gamma \in \mathcal{C}_c^\infty(\mathbb{R})$ , the set of infinitely differentiable functions of  $\mathbb{R}$  with compact support, and in particular, their support is included in  $[-1/2, 1/2]$ . Moreover  $\phi$  and  $\gamma$  are non-negative, even functions that are non-increasing on  $[0, 1/2]$  and such that  $\int_{-1/2}^{1/2} \phi(r) dr = \int_{-1/2}^{1/2} \gamma(r) dr = 1$ .

**(H3)**  $\int_{\mathbb{T} \times E} \left( \frac{1}{2}|v|^2 + \frac{1}{2} \int_{\mathbb{T} \times E} \Phi_\ell(r - r') W(x - x') d\mu_0(r', z') + U(x) \right) d\mu_0(r, z) < \infty$ .

**(H4)**  $\mu_0 \in \mathcal{M}^1(\mathbb{T} \times E)$  has a density  $f_0$  with respect to the Lebesgue measure and its  $r$ -marginal is the uniform measure on  $\mathbb{T}$ . In particular, for any  $r \in \mathbb{T}$ ,  $f_0(r, \cdot)$  is a probability density on  $E$ .

Moreover, there exist a probability density  $h$  on  $E$  with finite first moment  $\int_E |z| h(z) dz < \infty$ , and a constant  $C > 0$  such that for any  $r, r' \in \mathbb{T}$  and  $z \in E$ ,

$$|f_0(r, z) - f_0(r', z)| \leq C|r - r'|h(z).$$

The Lipschitz assumption in **(H1)** is classical in mean field theory. The two inequalities on  $U$  and  $W$  are technical assumptions that hold for harmonic potentials. Hypothesis **(H3)** is a moment assumption which, by conservation of energy, is crucial. **(H4)** is the minimal regularity hypothesis on the  $r$  variable at time 0 that we will need to prove the mean field limit. It holds for local Gibbs measures, for which the temperature is a regular function of the  $r$  variable for instance. Under these hypotheses, we then have the following proposition.

**Proposition 1.** *There is a unique solution to the nonlinear martingale problem (8) starting at  $\mu_0$ .*

In particular, Proposition 1 implies existence of weak solutions to (6). Denote by  $\mathcal{W}_1$  the Wasserstein distance associated with the Euclidean norm  $|\cdot|$  on  $\mathbb{T} \times E$

$$\mathcal{W}_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{T} \times E} |y - y'| d\pi(y, y'), \quad (9)$$

where  $\Pi(\mu, \nu)$  is the set of couplings of the probability measures  $\mu, \nu \in \mathcal{M}^1(\mathbb{T} \times E)$ , and we used the shortened notation  $y := (r, z) = (r, x, v)$ . We are now able to state the mean field convergence result.

**Theorem 1.** *Let  $\mu$  be the solution of the nonlinear martingale problem (8) starting at  $\mu_0$ . Assume that the initial coordinates  $(Z_0^i)_{i \leq N}$ , with  $Z_0^i = (X_0^i, V_0^i)$ , are independent and with respective density distribution  $f_0(i/N, z)$  for all  $1 \leq i \leq N$ . Then there exist two positive constants  $K_1$  and  $K_2$  such that for any  $1/N < \epsilon_N < \ell$ ,*

$$\mathbb{E} \left[ \mathcal{W}_1(\mu_t^N, \mu_t) \right] \leq K_1 \left( (N\epsilon_N)^{-\frac{1}{4(d+1)}} + \frac{\epsilon_N}{\ell} + \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) e^{K_2(1+\bar{\gamma})t}.$$

The constants  $K_1$  and  $K_2$  in the theorem do not depend on the three parameters  $N$ ,  $\ell$  and  $\bar{\gamma}$ , but only on the potentials  $W$  and  $U$ , on the functions  $\phi$  and  $\gamma$ , and the initial measure  $\mu_0$ .  $\epsilon_N$  is a coarse-graining parameter that naturally appears in the proof of Theorem 1 and is precisely defined in Section 2. In particular, for fixed  $\bar{\gamma}$ , choosing  $\epsilon_N = \ell^{\frac{2d+2}{2d+3}} N^{-\frac{1}{2d+3}}$ , we deduce that there exist positive constants  $K$  and  $K'$  such that for any  $\ell > 1/N$ ,

$$\mathbb{E} \left[ \mathcal{W}_1(\mu_t^N, \mu_t) \right] \leq K(\ell N)^{-\frac{1}{2(2d+3)}} e^{K't}.$$

One of the main features of the particle system we consider is that, contrary to classical mean field theory for gases, the sequence  $(Z_t^i)_{i \leq N}$  is not exchangeable. The behavior of  $Z_t^i$  is intrinsically tied with the position  $r^i$  in the chain. In particular, one cannot prove the mean field limit by comparing the law of one typical oscillator  $Z_t^i$  at time  $t$  to  $\mu_t$ , as usually done in the mean field theory (see [8, 14, 23]). Instead, the whole system  $(r^i, Z_t^i)_{i \leq N}$  has to be compared to  $\mu_t$  in its entirety.

Another difficulty comes from the fact that, even if  $\nabla W$  is uniformly Lipschitz over  $\mathbb{R}^d$ , the map  $(r, x) \mapsto \Phi_\ell(r) \nabla W(x)$  is not uniformly Lipschitz over  $\mathbb{T} \times \mathbb{R}^d$  in general. Consequently, classical mean field limit proofs do not readily work in this situation, and even the proof of Proposition 1 is not straightforward. We bypass this difficulty for Proposition 1 by proving a contraction estimate for a well-suited distance, the sliced Wasserstein distance (see Definition 1). Its proof is postponed to Section 3.

Theorem 1 is proved in Section 2. The proof is based on a coupling of the particle system  $(Y^i)_{i \leq N} := (r^i, Z^i)_{i \leq N}$  to a new system  $(\tilde{Y}^i)_{i \leq N}$ , whose law is based on the solution  $\mu$  to the nonlinear martingale problem (8), and which are driven by the same Poisson measures as the original particle system. We use an original coupling over mesoscopic boxes in  $\mathbb{T}$ , and control directly the averages  $1/N \sum \mathbb{E}[|Y^i - \tilde{Y}^i|]$  to circumvent the aforementioned difficulties arising in this mean field limit. Moreover, contrary to classical McKean-Vlasov theory (see [23] for instance), since the stochastic terms contribute via conservative exchanges, the new system  $(\tilde{r}^i, \tilde{Z}^i)_{i \leq N}$  is heavily correlated. We will therefore introduce a third system of independent processes, based on the techniques developed by Cortez and Fontbona in [6] for the Kac model (see also [12] for an example with continuous exchanges).

Finally, in Section 4, we study transport of energy in appropriate scales for the system by using this mean field limit. We will consider the following additional moment and symmetry hypothesis on the initial distribution  $f_0$ :

**(H5)** for any  $(r, z) \in \mathbb{T} \times E$ ,  $f_0(r, z) = f_0(r, -z)$ .

**(H6)**  $\int_{\mathbb{T} \times E} (|v|^{2+2b} + |x|^{2+2b}) d\mu_0(r, x, v) < \infty$  for some  $b > 0$ .

We prove in Lemma 7 that the symmetry **(H5)** is preserved at any later time  $t$  for the solution of the nonlinear martingale problem. Therefore, for a harmonic interaction potential  $W(x) = |x|^2/2$ ,  $\mu$  is a weak solution to the simpler equation

$$\begin{aligned} \partial_t f_t + v \cdot \nabla_x f_t - \left( x + \nabla U(x) \right) \cdot \nabla_v f_t \\ = \bar{\gamma} \int_{\mathbb{T} \times E} \Gamma_\ell(r - r') (f_t(r, x, v') f_t(r', x', v) - f_t(r, z) f_t(r', z')) dr' dz', \quad (10) \end{aligned}$$

and the term coming from the interaction potential is therefore reduced to an additional pinning term (associated with a harmonic potential). Therefore, energy is only transmitted by the noise in the mean field limit. Define the energy of particle  $i$  at time  $t$  by

$$\mathcal{E}_t^i := \frac{1}{2}|V_t^i|^2 + U(X_t^i) + \frac{1}{4} \sum_{k=-\ell N}^{\ell N} \phi_k |X_t^i - X_t^{i+k}|^2, \quad (11)$$

we prove that  $\mathcal{E}_t^i$  evolves diffusively for a class of anharmonic pinning potentials  $U$ .

**Proposition 2.** *Suppose  $W$  is harmonic and  $U(x) = |x|^2 \psi(x/|x|)$ , where  $\psi \in C^2(\mathbb{S}^{d-1}, \mathbb{R}_+^*)$ .  $\frac{1}{N} \sum_{i=1}^N \mathcal{E}_{t\ell^{-2}}^i \delta_{i/N}$  converges to the solution of*

$$\begin{cases} \partial_t e_t = \bar{\gamma} \frac{c_\gamma}{2} \partial_{xx}^2 e_t \\ e_0(r) = \int_E \left( \frac{1}{2}|v|^2 + U(x) + \frac{1}{2}|x|^2 \right) d\mu_0(r, x, v), \end{cases}$$

where  $c_\gamma = \frac{1}{2} \int_{-1/2}^{1/2} u^2 \gamma(u) du$ , in the sense that for any  $g \in C^4(\mathbb{T})$  and any  $t > 0$ ,

$$\lim_{\ell \rightarrow 0} \lim_{N \rightarrow \infty} \sup_{t \leq T} \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N \mathcal{E}_{t\ell^{-2}}^i g\left(\frac{i}{N}\right) - \int_{\mathbb{T}} e_t(r) g(r) dr \right| \right] = 0.$$

The potentials  $U$  in Proposition 2 are exactly  $C^2$  homogeneous functions of degree 2, which satisfy **(H1)** and for which one can prove an equipartition theorem to close the diffusion equation for the energy. In particular, we will prove that one can construct a function  $\ell = \ell(N) = c(\log N)^{-1/2}$  for some constant  $c > 0$ , such that

$$\lim_{N \rightarrow \infty} \sup_{t \leq T} \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N \mathcal{E}_{t\ell(N)^{-2}}^i g\left(\frac{i}{N}\right) - \int_{\mathbb{T}} e_t(r) g(r) dr \right| \right] = 0. \quad (12)$$

## 2 Mean field limit

To define fully define the system (3) with the correlation structure of the different Poisson random measures, we use a construction similar to [6] for the Kac model. Instead of considering a collection of Poisson random measure on  $\mathbb{R}_+ \times \mathbb{R}^d$  that select velocities as in (3), we first define a global Poisson random measure  $\mathcal{N}$  on  $\mathbb{R}_+ \times \mathbb{T}^2$  that selects the pairs of particles which exchange velocities.  $\mathcal{N}$  selects points in the periodic square  $\mathbb{T}^2$  with intensity

$$\bar{\gamma} N^2 \sum_{i=1}^N \left( \gamma_0 \mathbb{1}_{(r,r') \in \Lambda^i \times \Lambda^i} + \frac{1}{2} \sum_{\substack{k=-\ell N \\ k \neq 0}}^{\ell N} \gamma_k \mathbb{1}_{(r,r') \in \Lambda^i \times \Lambda^{i+k}} \right) dt dr dr', \quad (13)$$

where  $\Lambda^i = [(i-1/2)/N, (i+1/2)/N]$  for  $-N < i \leq N$ . Note that we identify  $\Lambda^i$  with  $\Lambda^{i-N}$  for any  $1 \leq i \leq N$  as we work in the torus. The velocity exchanges between two different particles with indices  $i$  and  $i+k$  only happen when the Poisson random measure hits either  $\Lambda^i \times \Lambda^{i+k}$  or  $\Lambda^{i+k} \times \Lambda^i$ , which happens at rate  $\bar{\gamma} \gamma_k$ .

We associate to the particle indexed by  $i$  a Poisson random measure  $\mathcal{N}^i$  on  $\mathbb{R}_+ \times \mathbb{T}$  that only selects the velocity exchanges between this particle and one of its neighbours:

$$\mathcal{N}^i(dt, dr) = \mathcal{N}(dt, \Lambda^i, dr) + \mathcal{N}(dt, dr, \Lambda^i) - \mathcal{N}(dt, \Lambda^i, \Lambda^i). \quad (14)$$

$\mathcal{N}^i$  has thus intensity

$$\bar{\gamma} N \sum_{k=-\ell N}^{\ell N} \gamma_k \mathbb{1}_{r \in \Lambda^{i+k}} dt dr. \quad (15)$$





both systems. In particular, we will take the initial conditions  $\bar{Y}_0^i$  and  $\tilde{Y}_0^i$  to be equal almost surely for any  $i \leq N$ .

Secondly, to prove a law of large numbers type result for  $\bar{\mu}_t^N$ , we need the variables  $\bar{Y}_t^i$  to have a similar enough distribution. But, even if we have some regularity on the  $r$ -parameter for the initial distribution  $\mu_0$  by **(H3)**, we do not know any regularity property at time  $t$ . In particular, we do not know to what extent two nonlinear processes  $\bar{Y}^i$  and  $\bar{Y}^j$  with close spatial parameters  $\bar{r}^i$  and  $\bar{r}^j$  have a similar distribution. Therefore, we will deal with truly identically distributed random variables, by dividing  $\mathbb{T}$  into mesoscopic boxes and requiring that nonlinear processes in the same mesoscopic box have the same law.

Let  $\epsilon_N$  be a parameter representing the macroscopic size of the aforementioned boxes such that  $1/N < \epsilon_N < 1$ , and assume  $\epsilon_N$  goes to 0 with  $N$  at a rate to be defined later on. Assume for simplicity that the parameter  $\epsilon_N$  is such that  $\epsilon_N^{-1}$  is an integer. We subdivide  $\mathbb{T}$  in exactly  $\epsilon_N^{-1}$  boxes  $(B_j)_{1 \leq j \leq \epsilon_N^{-1}}$ , where  $B_j = ]j\epsilon_N, (j+1)\epsilon_N[$ , for  $1 \leq j \leq \epsilon_N^{-1}$ . We add the extra technical assumption that the integer  $\epsilon_N^{-1}$  divides  $N$ , so that each box contains exactly  $N\epsilon_N$  particles of the original system. Denote by  $NB_j$  the set of indices of the particles in the box  $B_j$ :

$$NB_j = \{jN\epsilon_N + 1, jN\epsilon_N + 2, \dots, (j+1)N\epsilon_N\}.$$

Denote by  $\mu_t^j$  the distribution  $\mu_t$  conditioned on the fact that the coordinate  $r$  belongs to  $B_j$ :

$$\mu_t^j(r, z) = \epsilon_N^{-1} \mathbb{1}_{r \in B_j} \mu_t(r, z). \quad (19)$$

For any  $1 \leq j \leq \epsilon_N^{-1}$ , the variables  $(\tilde{Y}_0^i)_{i \in NB_j}$  (and therefore  $(\bar{Y}_0^i)_{i \in NB_j}$  as well) are then taken independent and with distribution  $\mu_0^j$ . In particular,  $\tilde{r}_0^i$  is a uniformly distributed over  $B_j$  for any  $i \in NB_j$ .

Notice that, in a given box  $B_j$ , the spatial parameters  $(\tilde{r}^i)_{i \in NB_j}$  are not ordered as their respective indices  $i \in NB_j$  in general. Yet, with this construction, we can always localize  $\tilde{r}^i$  and get the uniform estimate

$$|i/N - \tilde{r}^i| \leq \epsilon_N. \quad (20)$$

This estimate, together with energy bounds involving  $\nabla W$ , will help compensating the fact we do not have the uniform Lipschitz property for  $\Phi_\ell(r)\nabla W(x)$ . The idea of doing couplings over mesoscopic boxes  $(B_j)_{1 \leq j \leq \epsilon_N^{-1}}$  will appear several times during the proof due to the local mean field structure of our problem.

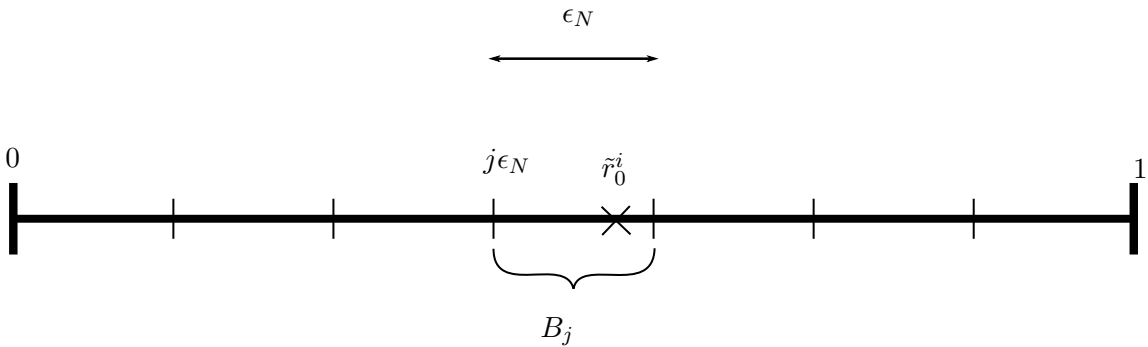


Figure 3 – Boxes  $B_j$  of size  $\epsilon_N$ . For  $i \in NB_j$ ,  $\tilde{r}_0^i$  is a uniform random variable in  $B_j$

Before turning to the proof of Theorem 1, let us state a series of lemmas that we will use and whose proofs are postponed to the end of the section. In the whole proof and in the lemmas,  $K$ ,

$K'$  are constants that change from one line to another that depend on  $U, W, \phi, \gamma$ , but not on  $N, \ell, \bar{\gamma}$  since we want to track the dependence of our convergence rates on these terms. Define for any  $1 \leq i \leq N$  the probability density on  $\mathbb{T}$

$$\sigma^i(r) = N \sum_{k=-\ell N}^{\ell N} \gamma_k \mathbb{1}_{r \in \Lambda^{i+k}}, \quad (21)$$

from which  $\mathcal{N}^i$  selects a neighbor of particle  $i$  to exchange velocities with, see (15). Define also the weighted distribution

$$w^{\mu_t, \tilde{r}_0^i}(v') = \int_{\mathbb{T} \times \mathbb{R}^d} \Gamma_\ell(\tilde{r}_0^i - r') d\mu_t(r', x', v'), \quad (22)$$

which is the velocity marginal of  $\mu_t$ , modulated by the function  $\Gamma_\ell$  around the axial parameter  $\tilde{r}_0^i$ . This is the distribution we expect for the updates of velocity of the nonlinear processes with axial parameter  $\tilde{r}_0^i$ . In the following lemma, adapted from [6], we construct the map  $\Pi^i$  introduced in (17).

**Lemma 1.** *Let  $1 \leq j \leq \epsilon_N^{-1}$  and  $i \in NB_j$ . Then for any  $N$ , there exists a measurable mapping*

$$\Pi^i : \mathbb{R}_+ \times \mathbb{T} \times B_j \times (\mathbb{R}^d)^{2\ell N+1} \rightarrow \mathbb{R}^d,$$

such that, for any  $\tilde{r} \in B_j$  and any  $\mathbf{v} \in (\mathbb{R}^d)^{2\ell N+1}$ , if  $r$  is drawn from (21), then  $\Pi_t^i(r, \tilde{r}, \mathbf{v})$  has distribution  $w^{\mu_t, \tilde{r}}$  (see (22)). Moreover, for any coupling  $\pi_t$  of  $\mu_t^N$  and  $\mu_t$ , the following bounds hold

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T}} |V_t^{[Nr]} - \Pi_t^i(r, \tilde{r}_0^i, \mathbf{V}_t^i)| \sigma^i(r) dr \right] &\leq K \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v' - v''| d\pi_t(y', y'') \right] \\ &+ \frac{K}{\ell^2} \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v''| \mathbb{1}_{\{|\tilde{r}_0^i - r''| < \ell/2\}} |r' - r''| d\pi_t(y', y'') \right] + \frac{K\epsilon_N}{\ell}, \end{aligned} \quad (23)$$

and also

$$\frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_{B_j} |\Pi_t^i(r, \tilde{r}_0^i, \mathbf{V}_t^i)| \sigma^i(r) dr \right] \leq K \frac{\epsilon_N^{1/2}}{\ell^{1/2}}. \quad (24)$$

From (23), we can control the difference of the velocities' updates between the original system and the nonlinear processes. (24) is a useful bound that we will need. The next lemma shows the marginal distribution at time  $t$  of the nonlinear process  $\tilde{Y}^i$  for  $i \in NB_j$  defined in (17) is indeed  $\mu_t^j$ , where  $\mu_t^j$  is defined in (19).

**Lemma 2.** *Let  $1 \leq j \leq \epsilon_N^{-1}$  and  $i \in NB_j$ . The process  $(\tilde{Y}_t^i)_{t \geq 0}$  satisfying (17) and such that  $\tilde{Y}_0^i$  has distribution  $\mu_0^j$  is well defined and has distribution  $\mu_t^j$ .*

The next lemma shows that we can choose a good coupling of the initial conditions of the original system  $Y_0^i$  and of the nonlinear processes  $\tilde{Y}_0^i$ .

**Lemma 3.** *There exists a coupling of  $(Y_0^i)_{i \leq N}$  and  $(\tilde{Y}_0^i)_{i \leq N}$  such that, for any  $1 \leq i \leq N$ ,*

$$\mathbb{E} \left[ |Y_0^i - \tilde{Y}_0^i| \right] \leq K\epsilon_N,$$

for some constant  $K$ .

This last lemma gives uniform in time moment bounds by conservation type arguments.

**Lemma 4.** *The solution  $\mu$  of the nonlinear martingale problem (8) satisfies*

$$\sup_{t \geq 0} \int_{\mathbb{T} \times E} |z|^2 d\mu_t(r, z) < \infty.$$

*For the particle system, the following bound holds*

$$\sup_{N \geq 1} \sup_{t \geq 0} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |V_t^i|^2 \right] < \infty.$$

## 2.1 Proof of Theorem 1

*Proof.* Recall that in (18), we got

$$\mathbb{E} \left[ \mathcal{W}_1(\mu_t^N, \mu_t) \right] \leq \mathbb{E} \left[ \mathcal{W}_1(\mu_t^N, \tilde{\mu}_t^N) \right] + \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^N, \mu_t) \right]$$

by triangular inequality. It suffices to prove the convergence of the two terms on the right-hand side to conclude. Following the terminology in [6], we will call the first step that consists in bounding  $\mathbb{E}[\mathcal{W}_1(\mu_t^N, \tilde{\mu}_t^N)]$  the *coupling* and the second step that consists in bounding  $\mathbb{E}[\mathcal{W}_1(\tilde{\mu}_t^N, \mu_t)]$  the *decoupling*.

*Step 1: Coupling*

First, as the two measures  $\mu_t^N$  and  $\tilde{\mu}_t^N$  are atomic, it is easy to bound

$$\mathcal{W}_1(\mu_t^N, \tilde{\mu}_t^N) \leq \frac{1}{N} \sum_{i=1}^N |Y_t^i - \tilde{Y}_t^i|,$$

and we are led to bound the expectation of the term on the right-hand side. By the evolution equations (16) for  $(Y^i)_{1 \leq i \leq N}$  and (17) for  $(\tilde{Y}^i)_{1 \leq i \leq N}$  and the Lipschitz property for  $\nabla U$ , we can bound:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |Y_t^i - \tilde{Y}_t^i| \right] &\leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |Y_0^i - \tilde{Y}_0^i| \right] + \int_0^t \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |Y_s^i - \tilde{Y}_s^i| \right] ds \\ &\quad + \int_0^t ds \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} \Phi_\ell \left( \frac{i}{N} - r' \right) \nabla W(X_s^i - x') d\mu_s^N(r', z') \right. \right. \\ &\quad \quad \left. \left. - \int_{\mathbb{T} \times E} \Phi_\ell(\tilde{r}_0^i - r'') \nabla W(\tilde{X}_s^i - x'') d\mu_s(r'', z'') \right| \right] \\ &\quad + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_0^t \int_{\mathbb{T}} |V^{[Nr]} - \Pi_s^i(r) - (V_{s-}^i - \tilde{V}_{s-}^i)| d\mathcal{N}^i(s, r) \right]. \end{aligned} \quad (25)$$

Denoting by  $\pi_s$  a coupling between  $\mu_s^N$  and  $\mu_s$ , the force term involving  $\nabla W$  in (25), can be

bounded by:

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \left| \Phi_\ell \left( \frac{i}{N} - r' \right) \nabla W(X_s^i - x') - \Phi_\ell(\tilde{r}_0^i - r'') \nabla W(\tilde{X}_s^i - x'') \right| d\pi_s(y', y'') \right] \\
& \leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T} \times E} \Phi_\ell \left( \frac{i}{N} - r' \right) \left| \nabla W(X_s^i - x') - \nabla W(\tilde{X}_s^i - x') \right| d\mu_s^N(y') \right] \\
& \quad + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \Phi_\ell \left( \frac{i}{N} - r' \right) \left| \nabla W(\tilde{X}_s^i - x') - \nabla W(\tilde{X}_s^i - x'') \right| d\pi_s(y', y'') \right] \\
& \quad + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \left| \Phi_\ell \left( \frac{i}{N} - r' \right) - \Phi_\ell \left( \frac{i}{N} - r'' \right) \right| \left| \nabla W(\tilde{X}_s^i - x'') \right| d\pi_s(y', y'') \right] \\
& \quad + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \left| \Phi_\ell \left( \frac{i}{N} - r'' \right) - \Phi_\ell(\tilde{r}_0^i - r'') \right| \left| \nabla W(\tilde{X}_s^i - x'') \right| d\mu_s(y'') \right] \\
& \leq A_1 + A_2 + A_3 + A_4,
\end{aligned}$$

by introducing the four cross terms in the second line and denoting by  $A_k$  the term numbered  $1 \leq k \leq 4$  at the last line. By the Lipschitz property for  $\nabla W$ , the first term is bounded by

$$A_1 \leq K \frac{1}{N} \sum_{i=1}^N \sum_{k=-\ell N}^{\ell N} \Phi_\ell \left( \frac{i-k}{N} \right) \mathbb{E} \left[ \left| X_s^i - \tilde{X}_s^i \right| \right] \leq K' \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| X_s^i - \tilde{X}_s^i \right| \right], \quad (26)$$

since  $\sum_k \Phi_\ell(k/N)$  is bounded. Similarly, the second term is bounded by

$$A_2 \leq K' \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |x' - x''| d\pi_s(y', y'') \right]. \quad (27)$$

It remains to choose a good coupling to bound this term. The fourth term is also relatively easy to bound by the Lipschitz property for  $\Phi_\ell$ , which has Lipschitz constant equal to  $Lip(\phi)\ell^{-2}$ :

$$\begin{aligned}
A_4 & \leq \frac{K}{\ell^2} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \left| \frac{i}{N} - \tilde{r}_0^i \right| \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} \left| \nabla W(\tilde{X}_s^i - x'') \right| d\mu_s(y'') \right] \\
& \leq \frac{K\epsilon_N}{\ell^2} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} \left| \nabla W(\tilde{X}_s^i - x'') \right| d\mu_s(y'') \right] \\
& \leq \frac{K'\epsilon_N}{\ell}.
\end{aligned} \quad (28)$$

We used at the first line that  $\Phi_\ell$  has support included in  $[-\ell/2, \ell/2]$  to introduce the indicator function, and we used at second line the uniform bound (20). For the last line, we used the following lemma.

**Lemma 5.** *Under hypothesis (H1) on  $W$ , the following uniform in time bound holds*

$$\sup_{0 < \ell < 1} \sup_{N \geq 1} \sup_{t \geq 0} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T} \times E} \frac{1}{\ell} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} \left| \nabla W(\tilde{X}_t^i - x'') \right| d\mu_t(y'') \right] < \infty.$$

The proof of Lemma 5 is postponed to the end of the section. Using the same arguments, the third term is bounded by

$$A_3 \leq \frac{K}{\ell^2} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |r' - r''| \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} \left| \nabla W(\tilde{X}_s^i - x'') \right| d\pi_s(y', y'') \right]. \quad (29)$$

It just remains to choose an appropriate coupling  $\pi_s$  of  $\mu_s^N$  and  $\mu_s$  to bound (27) and (29), keeping in mind that we will need to apply Gronwall's lemma at some point. A first option would be to take  $\pi_s$  to be the optimal coupling, which would give the best bound for (27), but this gives no information on the term (29) since  $\nabla W$  is unbounded. The local mean field structure of the system we consider actually leads us to introduce again couplings over mesoscopic boxes. More precisely, recall from (19) the notation  $\mu_s^j$  for the measure  $\mu_s$  conditioned on the fact that  $r \in B_j$ . We similarly define the empirical probability measure on the box  $B_j$ :

$$\mu_s^{N,j} := \frac{1}{N\epsilon_N} \sum_{i \in NB_j} \delta_{Y_s^i}.$$

Then, denoting by  $\pi_s^j$  the optimal coupling between  $\mu_s^{N,j}$  and  $\mu_s^j$  for  $\mathcal{W}_1$ , it is easy to see that

$$\pi_s := \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \pi_s^j$$

is a coupling of  $\mu_s^N$  and  $\mu_s$ . In words, this coupling consists in choosing a box  $B_j$  uniformly at random over  $(B_j)_{1 \leq j \leq \epsilon_N^{-1}}$  and then, given that the coordinates  $r', r'' \in B_j$ , select the optimal coupling  $\pi^j$ . With this approach, the term (27) can then be bounded by

$$A_2 \leq K' \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |x' - x''| d\pi_s(y', y'') \right] \leq K' \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\mu_s^{N,j}, \mu_s^j) \right],$$

and the term (29) is bounded by

$$\begin{aligned} A_3 &\leq K \frac{\epsilon_N}{\ell^2} \frac{1}{N} \sum_{i=1}^N \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} |\nabla W(\tilde{X}_s^i - x'')| d\mu_s^j(y'') \right] \\ &\leq K \frac{\epsilon_N}{\ell^2} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} |\nabla W(\tilde{X}_s^i - x'')| d\mu_s(y'') \right] \\ &\leq K' \frac{\epsilon_N}{\ell}, \end{aligned}$$

by Lemma 5. All in all, assembling the last two inequalities and the bounds (26) and (28), the force terms involving  $\nabla W$  and  $\nabla U$  in (25) are bounded by

$$K \int_0^t \left( \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |X_s^i - \tilde{X}_s^i| \right] + \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\mu_s^{N,j}, \mu_s^j) \right] + \frac{\epsilon_N}{\ell} \right) ds. \quad (30)$$

Finally, it remains to control the terms coming from the exchanges of velocities. Recall that  $\mathcal{N}^i$  has intensity  $\bar{\gamma} \sigma^i(r) dr dt$ , with  $\sigma^i$  defined in (21). Since  $(t, r) \mapsto \Pi_t^i(r)$  is measurable, one can bound

$$\begin{aligned} B &:= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_0^t |V^{[Nr]} - \Pi_s^i(r) - (V_{s-}^i - \tilde{V}_{s-}^i)| d\mathcal{N}^i(s, r) \right] \\ &\leq \bar{\gamma} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_0^t ds \int_{\mathbb{T}} |V^{[Nr]} - \Pi_s^i(r)| \sigma^i(r) dr \right] + \bar{\gamma} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_0^t ds \int_{\mathbb{T}} |V_s^i - \tilde{V}_s^i| d\sigma^i(r) dr \right] \\ &\leq \bar{\gamma} \int_0^t \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T}} |V^{[Nr]} - \Pi_s^i(r)| \sigma^i(r) dr \right] ds + \bar{\gamma} \int_0^t \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ |V_s^i - \tilde{V}_s^i| \right] ds. \end{aligned}$$

By (23) in Lemma 1, the first term is bounded by:

$$K\bar{\gamma} \int_0^t ds \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v' - v''| d\pi_s(y', y'') \right] \\ + K \int_0^t \left( \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v''| \frac{1}{\ell^2} \mathbb{1}_{\{|\tilde{r}_0^i - r''| < \ell/2\}} |r' - r''| d\pi_s(y', y'') \right] + \frac{\epsilon_N}{\ell} \right) ds,$$

for any coupling  $\pi_s$  of  $\mu_s^N$  and  $\mu_s$ . Once again we choose the coupling  $\pi_s = \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \pi_s^j$  over the mesoscopic boxes  $(B_j)_{1 \leq j \leq \epsilon_N^{-1}}$  and we get the bound:

$$K\bar{\gamma} \epsilon_N \int_0^t \left( \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} [\mathcal{W}_1(\mu_s^{N,j}, \mu_s^j)] + \frac{1}{\ell} \int_{\mathbb{T} \times E} |v''| d\mu_s + \frac{1}{\ell} \right) ds.$$

Using Lemma 4 to bound  $\int_{\mathbb{T} \times E} |v''| d\mu_s$  by a constant, we finally can bound the terms coming from the exchanges of velocity in (25) by

$$B \leq K\bar{\gamma} \int_0^t \left( \frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |V_s^i - \tilde{V}_s^i| ] + \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} [\mathcal{W}_1(\mu_s^{N,j}, \mu_s^j)] + \frac{\epsilon_N}{\ell} \right) ds. \quad (31)$$

Combining the bound on the force terms (30) and the bound on the exchanges of velocities (31), we obtain

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_t^i - \tilde{Y}_t^i| ] \leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_0^i - \tilde{Y}_0^i| ] \\ + K(1 + \bar{\gamma}) \int_0^t \left( \frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_s^i - \tilde{Y}_s^i| ] + \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} [\mathcal{W}_1(\mu_s^{N,j}, \mu_s^j)] + \frac{\epsilon_N}{\ell} \right) ds. \quad (32)$$

As our goal is to ultimately combine this inequality with (18), we can introduce the empirical measure associated with the nonlinear processes  $(\tilde{Y}_s^i)_{i \in NB_j}$  to replace the term  $\mathcal{W}_1(\mu_s^{N,j}, \mu_s^j)$ . Define

$$\tilde{\mu}_s^{N,j} = \frac{1}{N\epsilon_N} \sum_{i \in NB_j} \delta_{\tilde{Y}_s^i},$$

the obvious analogue of  $\mu_s^{N,j}$  for the system of nonlinear processes  $(\tilde{Y}_s^i)_{i \in NB_j}$ . We then apply the triangular inequality and a straightforward choice of coupling between  $\mu_s^{N,j}$  and  $\tilde{\mu}_s^{N,j}$  to bound

$$\mathcal{W}_1(\mu_s^{N,j}, \mu_s^j) \leq \mathcal{W}_1(\mu_s^{N,j}, \tilde{\mu}_s^{N,j}) + \mathcal{W}_1(\tilde{\mu}_s^{N,j}, \mu_s^j) \leq \frac{1}{N\epsilon_N} \sum_{i \in NB_j} |Y_s^i - \tilde{Y}_s^i| + \mathcal{W}_1(\tilde{\mu}_s^{N,j}, \mu_s^j).$$

Summing this last inequality over  $j$ , we obtain the mean value  $1/N \sum_{i=1}^N |Y_s^i - \tilde{Y}_s^i|$  for the first term on the right-hand side. Inserting it in (32), we finally get

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_t^i - \tilde{Y}_t^i| ] \leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_0^i - \tilde{Y}_0^i| ] \\ + K(1 + \bar{\gamma}) \int_0^t \left( \frac{1}{N} \sum_{i=1}^N \mathbb{E} [ |Y_s^i - \tilde{Y}_s^i| ] + \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} [\mathcal{W}_1(\tilde{\mu}_s^{N,j}, \mu_s^j)] + \frac{\epsilon_N}{\ell} \right) ds. \quad (33)$$

Step 2: Decoupling

Observing (18) and (33), and since

$$\mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^N, \mu_t) \right] \leq \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \mu_t^j) \right],$$

we can see that it suffices to control the right-hand side term

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}-1} \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \mu_t^j) \right]$$

to conclude the proof. As noted earlier, the sequence of variables  $(\tilde{Y}_t^i)_{i \leq N}$  is correlated since particles with close indices  $i, i'$  are driven by two Poisson  $\mathcal{N}^i$  and  $\mathcal{N}^{i'}$  that share some atoms. Following [6], we introduce a new system  $(\bar{Y}_t^i)_{i \leq N}$  of nonlinear processes by replacing those shared atoms. In particular, we take advantage of our box structure  $(B_j)_{1 \leq j \leq \epsilon_N^{-1}}$  to only require independence of  $(\bar{Y}_t^i)_{i \in NB_j}$  on each box.

The construction of the nonlinear processes  $(\bar{Y}_t^i)_{1 \leq i \leq N}$  is the following. Recall that  $\mathcal{N}^i$  was defined from a macroscopic Poisson random measure  $\mathcal{N}$  by (14). We define a Poisson random measure  $\mathcal{M}$  is on  $\mathbb{R}_+ \times \mathbb{T}^2$ , independent of  $\mathcal{N}$ , with same intensity given by (13). For  $1 \leq j \leq \epsilon_N^{-1}$  and  $i \in NB_j$ , we define  $\mathcal{M}^i$  by

$$\mathcal{M}^i(dt, dr) = \mathcal{N}(dt, \Lambda^i, dr) + \mathcal{N}(dt, dr, \Lambda^i) \mathbb{1}_{r \notin B_j} + \mathcal{M}(dt, dr, \Lambda^i) \mathbb{1}_{r \in B_j \setminus \Lambda^i}.$$

By their definitions, the Poisson random measures  $\mathcal{M}^i$  and  $\mathcal{M}^{i'}$  do not share any atom if  $i, i'$  are in the same set  $NB_j$  for some  $j$ . The processes  $(\bar{Y}_t^i)_{1 \leq i \leq N}$  are now defined by the same equations (17) as  $(\tilde{Y}_t^i)_{1 \leq i \leq N}$ , but replacing  $\mathcal{N}^i$  by  $\mathcal{M}^i$ , and with initial values  $\bar{Y}_0^i = \tilde{Y}_0^i$ .

**Lemma 6.** For any  $1 \leq j \leq \epsilon_N^{-1}$ , the processes  $(\bar{Y}_t^i)_{i \in NB_j}$ , defined by

$$\begin{cases} d\bar{X}_t^i = \bar{V}_t^i dt \\ d\bar{V}_t^i = - \left( \int_{\mathbb{T} \times E} \Phi_\ell(\bar{r}_0^i - r') \nabla W(\tilde{X}_t^i - x') d\mu_t(y') + \nabla U(\bar{X}_t^i) \right) dt + \int_{\mathbb{T}} (\Pi_t^i(r') - \bar{V}_t^i) d\mathcal{M}^i(t, r') \\ \bar{Y}_0^i = \tilde{Y}_0^i, \end{cases} \quad (34)$$

are independent.

Denoting by  $\bar{\mu}_t^{N,j}$  the empirical measure associated with  $(\bar{Y}_t^i)_{i \in NB_j}$ , the triangular inequality gives

$$\mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \mu_t^j) \right] \leq \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \bar{\mu}_t^{N,j}) \right] + \mathbb{E} \left[ \mathcal{W}_1(\bar{\mu}_t^{N,j}, \mu_t^j) \right], \quad (35)$$

and it suffices to control  $\mathbb{E}[\mathcal{W}_1(\tilde{\mu}_t^{N,j}, \bar{\mu}_t^{N,j})]$  to conclude, the other term being controlled by the law of large numbers result in Appendix A. The bound

$$\mathcal{W}_1(\tilde{\mu}_t^j, \bar{\mu}_t^j) \leq \frac{1}{N\epsilon_N} \sum_{i \in NB_j} |\tilde{Y}_t^i - \bar{Y}_t^i| \quad (36)$$

is easily obtained for all  $1 \leq j \leq \epsilon_N^{-1}$ . As we are ultimately interested in founding a bound for

the expectation and mean over  $j$  of  $\mathcal{W}_1(\tilde{\mu}_t^j, \bar{\mu}_t^j)$ , we are finally led to bound

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{Y}_t^i - \bar{Y}_t^i|] &\leq K \int_0^t ds \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{Y}_s^i - \bar{Y}_s^i|] \\ &+ \int_0^t ds \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T} \times E} \Phi_\ell(\tilde{r}_0^i - r') \left| \nabla W(\tilde{X}_s^i - x') - \nabla W(\bar{X}_s^i - x') \right| d\mu_s(r', z') \right] \\ &+ \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| \int_0^t \int_{\mathbb{T}} (\Pi_s^i(r) - \tilde{V}_{s-}^i) d\mathcal{N}^i(s, r) - \int_0^t \int_{\mathbb{T}} (\Pi_s^i(r) - \bar{V}_{s-}^i) d\mathcal{M}^i(s, r) \right| \right], \end{aligned} \quad (37)$$

by the Lipschitz property for  $\nabla U$ . Notice that we used that  $\tilde{r}_0^i = \bar{r}_0^i$  for the term at the second line. Therefore, by the Lipschitz property for  $\nabla W$ , we obtain the bound

$$K' \int_0^t ds \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{Y}_s^i - \bar{Y}_s^i|] \quad (38)$$

for the first two terms in (37). The last term in (37) can be bounded by splitting it in the three following terms:

$$\begin{aligned} C_1 + C_2 + C_3 &= \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t \int_{\mathbb{T}} |\tilde{V}_{s-}^i - \bar{V}_{s-}^i| \left( d\mathcal{N}(s, \Lambda^i, r) + d\mathcal{N}(s, r, \Lambda^i) \mathbb{1}_{r \notin B_j} \right) \right] \\ &+ \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t \int_{\mathbb{T}} |\Pi_s^i(r) - \tilde{V}_{s-}^i| d\mathcal{N}(s, r, \Lambda^i) \mathbb{1}_{r \in B_j \setminus \Lambda^i} \right] \\ &+ \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t \int_{\mathbb{T}} |\Pi_s^i(r) - \bar{V}_{s-}^i| d\mathcal{M}(s, r, \Lambda^i) \mathbb{1}_{r \in B_j \setminus \Lambda^i} \right], \end{aligned}$$

where the first term  $C_1$  corresponds to the simultaneous jumps of  $\tilde{V}^i$  and  $\bar{V}^i$ , *i.e.* the shared atoms of  $\mathcal{N}^i$  and  $\mathcal{M}^i$  for  $1 \leq i \leq N$ , and the two other terms  $C_2$  and  $C_3$  respectively correspond to the jumps of  $\tilde{V}^i$  alone and to the jumps of  $\bar{V}^i$  alone. Computing the expectation with respect to the Poisson integral in  $C_1$ , we get:

$$C_1 = \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \frac{\bar{\gamma}}{2} \left( 1 + \sum_{k: i+k \notin NB_j} \gamma_k \right) \mathbb{E} \left[ \int_0^t |\tilde{V}_s^i - \bar{V}_s^i| ds \right] \leq \bar{\gamma} \int_0^t ds \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{V}_s^i - \bar{V}_s^i|]. \quad (39)$$

As for the term  $C_2$  coming from the jumps of  $\mathcal{N}^i$  alone, one immediately gets:

$$\begin{aligned} C_2 &\leq \bar{\gamma} \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t ds \int_{B_j} \left( |\Pi_s^i(r)| + |\tilde{V}_s^i| \right) \sigma^i(r) dr \right] \\ &\leq \bar{\gamma} \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t ds \int_{B_j} |\Pi_s^i(r)| \sigma^i(r) dr \right] + \bar{\gamma} \sum_{k=-N\epsilon_N}^{N\epsilon_N} \gamma_k \int_0^t \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{V}_s^i|] ds \\ &\leq \bar{\gamma} \frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_0^t ds \int_{B_j} |\Pi_s^i(r)| \sigma^i(r) dr \right] + K\bar{\gamma} \frac{\epsilon_N}{\ell} \int_0^t \frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\tilde{V}_s^i|] ds, \end{aligned} \quad (40)$$

where we used the symmetry and monotonic properties **(H2)** of  $\gamma$  to bound the integral  $\int_{B_j} \sigma^i(r) dr$  by the sum  $\sum_{k=-N\epsilon_N}^{N\epsilon_N} \gamma_k$  at the second line, and then bounded this sum by  $K\epsilon_N/\ell$



by the definition (2) of  $\gamma_k$ . The sum in the second term of (40) is equal to

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| \tilde{V}_s^i \right| \right] = \sum_{j=1}^{\epsilon_N^{-1}} \int_{B_j \times E} |v| d\mu_s = \int_{\mathbb{T} \times E} |v| d\mu_s < \infty,$$

which is finite by Lemma 4. By the bound (24) of Lemma 1, we can bound the first term in (40) and finally get the following bound for  $C_2$ :

$$C_2 \leq Kt\bar{\gamma} \left( \frac{\epsilon_N^{1/2}}{\ell^{1/2}} + \frac{\epsilon_N}{\ell} \right) \leq K't\bar{\gamma} \frac{\epsilon_N^{1/2}}{\ell^{1/2}}, \quad (41)$$

for  $N$  large enough. The same bound can be obtained for the term  $C_3$  coming from the jumps of  $\mathcal{M}^i$  alone. Combining the bound on the force terms (38), the bound on the simultaneous jumps (39) and the bounds on the asynchronous jumps (41) in (37), we get

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| \tilde{Y}_t^i - \bar{Y}_t^i \right| \right] \leq K \int_0^t \left( (1 + \bar{\gamma}) \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| \tilde{Y}_s^i - \bar{Y}_s^i \right| \right] + \bar{\gamma} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) ds.$$

By Gronwall's lemma and (36), we finally get

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \bar{\mu}_t^{N,j}) \right] \leq \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \left( e^{K(1+\bar{\gamma})t} - 1 \right). \quad (42)$$

Since by Lemma 4, the measure  $\mu_t$  admits a moment of order 2 for all  $t \geq 0$ , we can apply the results of Appendix A and get

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\bar{\mu}_t^{N,j}, \mu_t^j) \right] \leq K'(N\epsilon_N)^{-\frac{1}{4(d+1)}} + K'\epsilon_N.$$

Combining this result with (42) in (35) finally gives

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\tilde{\mu}_t^{N,j}, \mu_t^j) \right] \leq \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} e^{K(1+\bar{\gamma})t} + K'(N\epsilon_N)^{-\frac{1}{4(d+1)}} + K'\epsilon_N. \quad (43)$$

### Conclusion of the proof of Theorem 1

Combining (33) and (43), and applying Gronwall's lemma, we get the bound:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| Y_t^i - \tilde{Y}_t^i \right| \right] &\leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| Y_0^i - \tilde{Y}_0^i \right| \right] e^{K(1+\bar{\gamma})t} \\ &\quad + K' \left( \frac{\epsilon_N}{\ell} + (N\epsilon_N)^{-\frac{1}{4(d+1)}} + \epsilon_N + \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) e^{K(1+\bar{\gamma})t}. \end{aligned}$$

By Lemma 3, we can bound  $\mathbb{E} \left[ \left| Y_0^i - \tilde{Y}_0^i \right| \right]$  by  $\epsilon_N$  for all  $i$ . Since  $\epsilon_N$  is smaller than  $\epsilon_N/\ell$ , we finally get

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \left| Y_t^i - \tilde{Y}_t^i \right| \right] \leq K' \left( \frac{\epsilon_N}{\ell} + (N\epsilon_N)^{-\frac{1}{4(d+1)}} + \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) e^{K(1+\bar{\gamma})t}.$$

Finally, combining this bound together with (43) in (18), gives

$$\mathbb{E} \left[ \mathcal{W}_1(\mu_t^N, \mu_t) \right] \leq K' \left( \frac{\epsilon_N}{\ell} + (N\epsilon_N)^{-\frac{1}{4(d+1)}} + \frac{\bar{\gamma}}{1 + \bar{\gamma}} \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) e^{K(1+\bar{\gamma})t}.$$

□

## 2.2 Proof of the intermediate lemmas

*Proof of Lemma 1.* 1. *Construction of  $\Pi^i$*

The construction of  $\Pi^i$  is based on Lemma 3 in [6]. Let  $1 \leq j \leq \epsilon_N^{-1}$  and  $i \in NB_j$ . Fix  $\tilde{r} \in B_j$  and  $\mathbf{v}^i = (v^{i+k})_{-\ell N \leq k \leq \ell N} \in (\mathbb{R}^d)^{2\ell N+1}$ . Recall from (22) the notation for the weighted measure  $w^{\mu_t, \tilde{r}}$  around  $\tilde{r}$ , from which we select a new velocity for particle  $\tilde{Y}^i$ . Denote also by  $w^{\mathbf{v}^i}$  the weighted measure on  $\mathbb{R}^d$  corresponding to the update of velocity among  $\mathbf{v}^i$ :

$$w^{\mathbf{v}^i} = \sum_{k=-\ell N}^{\ell N} \gamma_k \delta_{v^{i+k}}.$$

From a random variable  $r$  with distribution  $\sigma^i$  (see (21)), we want to construct a couple of variables  $(v^{[Nr]}, \Pi_t^i(r, \tilde{r}, \mathbf{v}^i))$  with respective distributions  $w^{\mathbf{v}^i}$  and  $w^{\mu_t, \tilde{r}}$  for all  $(t, \tilde{r}, \mathbf{v}^i)$ . In addition, we want the distance between those variables to be small in the sense that

$$\int_{\mathbb{T}} |v^{[Nr]} - \Pi_t^i(r, \tilde{r}, \mathbf{v}^i)| \sigma^i(r) dr = \mathcal{W}_1(w^{\mathbf{v}^i}, w^{\mu_t, \tilde{r}}). \quad (44)$$

It is straightforward to see that  $v^{[Nr]}$  has indeed distribution  $w^{\mathbf{v}^i}$  in that case. The construction of a function  $\Pi^i(r)$  with distribution  $w^{\mu_t, \tilde{r}}$  such that (44) holds principally amounts to the construction of a random variable with given law from a uniform random variable.

In fact, let  $p_{t, \tilde{r}, \mathbf{v}^i}^i(v, \tilde{v})$  be the optimal coupling of  $w^{\mathbf{v}^i}$  and  $w^{\mu_t, \tilde{r}}$ . For a couple of random variables  $(V, \tilde{V})$  with distribution  $p_{t, \tilde{r}, \mathbf{v}^i}^i$ , define  $P_{t, \tilde{r}, \mathbf{v}^i}^{i,k}$  to be the conditional law of  $\tilde{V}$ , given that  $V = V^{i+k}$ , i.e.:

$$P_{t, \tilde{r}, \mathbf{v}^i}^{i,k}(B) = p_{t, \tilde{r}, \mathbf{v}^i}^i(\{v^{i+k}\} \times B | \{v^{i+k}\} \times \mathbb{R}^d) = \frac{1}{\gamma_k} p_{t, \tilde{r}, \mathbf{v}^i}^i(\{v^{i+k}\} \times B), \quad (45)$$

for any Borel subset  $B$  of  $\mathbb{R}^d$ . The second equality comes from the fact that  $p_{t, \tilde{r}, \mathbf{v}^i}^i(\{v^{i+k}\} \times \mathbb{R}^d) = w^{\mathbf{v}^i}(\{v^{i+k}\}) = \gamma_k$ . Then there exists a function

$$q_{t, \tilde{r}, \mathbf{v}^i}^{i,k} : \Lambda^{i+k} \rightarrow \mathbb{R}^d, \quad (46)$$

such that, if  $r$  is uniformly distributed on  $\Lambda^{i+k}$ ,  $q_{t, \tilde{r}, \mathbf{v}^i}^{i,k}(r)$  has distribution  $P_{t, \tilde{r}, \mathbf{v}^i}^{i,k}$ . Defining  $\Pi^i$  by

$$\Pi_t^i(r, \tilde{r}, \mathbf{v}^i) = \sum_{k=-\ell N}^{\ell N} q_{t, \tilde{r}, \mathbf{v}^i}^{i,k}(r) \mathbb{1}_{r \in \Lambda^{i+k}}, \quad (47)$$

the pair  $(v^{[Nr]}, \Pi_t^i(r, \tilde{r}, \mathbf{v}^i))$  has indeed distribution  $p_{t, \tilde{r}, \mathbf{v}^i}^i$  if  $r$  has distribution  $\sigma^i(r) dr$ . Indeed,

$$\begin{aligned} \mathbb{P}(v^{[Nr]} = V_t^{i+k}, \Pi_t^i \in B) &= \mathbb{P}(r \in \Lambda^{i+k}, \Pi_t^i \in B) \\ &= \mathbb{P}(r \in \Lambda^{i+k}) \mathbb{P}(\Pi_t^i \in B | r \in \Lambda^{i+k}) \\ &= \gamma_k \mathbb{P}(q_{t, \tilde{r}, \mathbf{v}^i}^{i,k}(r) \in B) \\ &= p_{t, \tilde{r}, \mathbf{v}^i}^i(\{V_t^{i+k}\} \times B). \end{aligned}$$

(44) is now obvious from the fact  $(V_t^{[Nr]}, \Pi_t(r))$  has distribution  $p_{t, \tilde{r}, \mathbf{v}^i}^i$  when  $r$  is distributed by  $\sigma^i(r) dr$ . To ensure that the construction of  $\Pi^i$  detailed above can be done with the desired measurability properties, we follow [6]. Since the mapping  $(t, \tilde{r}, \mathbf{v}^i) \mapsto (w^{\mathbf{v}^i}, w^{\mu_t, \tilde{r}})$  is measurable, then we know from Corollary 5.22 in [26] that there exists a measurable mapping

$(t, \tilde{r}, \mathbf{v}^i) \mapsto p_{t, \tilde{r}, \mathbf{v}^i}^i$  such that  $p_{t, \tilde{r}, \mathbf{v}^i}^i(v, \tilde{v})$  is an optimal coupling between those two measures.  $P_{t, \tilde{r}, \mathbf{v}^i}^{i, k}$  defined in (45) is a kernel from  $\mathbb{R}_+ \times B_j \times (\mathbb{R}^d)^{2\ell N+1}$  to  $\mathbb{R}^d$ , so by Lemma 2.22 in [17],  $q_{t, \tilde{r}, \mathbf{v}^i}^{i, k}$  in (46) can be defined as a measurable function of  $(t, r, \tilde{r}, \mathbf{v}^i)$ . Finally the definition (47) ensures that  $\Pi^i$  is a measurable function of  $(t, r, \tilde{r}, \mathbf{v}^i)$ .

## 2. Proof of (23)

By (44), we can directly bound  $\mathcal{W}_1(w^{\mathbf{V}_t^i}, w^{\mu_t, \tilde{r}_0^i})$  to obtain a bound on (23). By Kantorovich-Rubinstein duality formula (see [26]), this term can be rewritten:

$$\mathcal{W}_1(w^{\mathbf{V}_t^i}, w^{\mu_t, \tilde{r}_0^i}) = \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \left| \sum_{k=-\ell N}^{\ell N} \gamma_k \varphi(V_t^{i+k}) - \int_{\mathbb{T} \times E} \Gamma_\ell(\tilde{r}_0^i - r'') \varphi(v'') d\mu_t(r'', z'') \right|. \quad (48)$$

By the definition (2) of  $\gamma_k$ , we have

$$\left| \gamma_k - \frac{1}{N} \Gamma_\ell \left( \frac{k}{N} \right) \right| = \left| \int_{\Lambda^k} \left( \Gamma_\ell(r) - \Gamma_\ell \left( \frac{k}{N} \right) \right) dr \right| \leq \frac{K}{\ell^2} \int_{\Lambda^0} |r| dr \leq \frac{K'}{N^2 \ell^2},$$

using the Lipschitz property for  $\Gamma_\ell$  for the first inequality. As a consequence, we can introduce the empirical measure  $\mu_t^N$  in (48) by replacing  $\gamma_k$  and get the bound:

$$\begin{aligned} & \mathcal{W}_1(w^{\mathbf{V}_t^i}, w^{\mu_t, \tilde{r}_0^i}) \\ & \leq \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \left| \int_{\mathbb{T} \times E} \varphi(v') \Gamma_\ell \left( \frac{i}{N} - r' \right) d\mu_t^N(r', z') - \int_{\mathbb{T} \times E} \varphi(v'') \Gamma_\ell(\tilde{r}_0^i - r'') d\mu_t(r'', z'') \right| \\ & \quad + \frac{K}{N \ell^2} \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{\mathbb{T} \times E} |\varphi(v')| \mathbb{1}_{\{|i/N - r'| < \ell/2\}} d\mu_t^N(r', z'). \end{aligned}$$

Therefore, bounding  $\varphi(v') \leq |v'|$  in the last term on the one hand, and taking expectations and the mean over  $1 \leq i \leq N$  on the other hand, we get:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \mathcal{W}_1(w^{\mathbf{V}_t^i}, w^{\mu_t, \tilde{r}_0^i}) \right] & \leq \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \left| \int_{\mathbb{T} \times E} \varphi(v') \Gamma_\ell \left( \frac{i}{N} - r' \right) d\mu_t^N(r', z') \right. \right. \\ & \quad \left. \left. - \int_{\mathbb{T} \times E} \varphi(v'') \Gamma_\ell(\tilde{r}_0^i - r'') d\mu_t(r'', z'') \right| \right] \\ & \quad + \frac{K}{N \ell^2} \frac{1}{N} \sum_{i=1}^N \frac{1}{N} \sum_{k=-\ell N}^{\ell N} \mathbb{E} \left[ |V_t^{i+k}| \right]. \end{aligned} \quad (49)$$

The sum  $1/N \sum_i \mathbb{E}[|V_t^i|]$  is bounded by Lemma 4, therefore the last term in (49) is bounded by

$$\frac{1}{N \ell^2} \frac{1}{N} \sum_{i=1}^N \frac{1}{N} \sum_{k=-\ell N}^{\ell N} \mathbb{E} \left[ |V_t^{i+k}| \right] \leq \frac{K}{N \ell}. \quad (50)$$

We can now focus on the first term of (49). Let  $\pi_t$  be a coupling between  $\mu_t^N$  and  $\mu_t$ . Introducing

cross terms, we bound the supremum term in the expectation:

$$\begin{aligned}
& \sup_{\substack{Lip(\varphi) \leq 1 \\ \varphi(0)=0}} \left| \int_{\mathbb{T} \times E} \varphi(v') \Gamma_\ell \left( \frac{i}{N} - r' \right) d\mu_t^N(r', z') - \int_{\mathbb{T} \times E} \varphi(v'') \Gamma_\ell(\tilde{r}_0^i - r'') d\mu_t(r'', z'') \right| \\
& \leq \sup_{\substack{Lip(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{(\mathbb{T} \times E)^2} |\varphi(v') - \varphi(v'')| \Gamma_\ell \left( \frac{i}{N} - r' \right) d\pi_t(y', y'') \\
& \quad + \sup_{\substack{Lip(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{(\mathbb{T} \times E)^2} |\varphi(v'')| \left| \Gamma_\ell \left( \frac{i}{N} - r' \right) - \Gamma_\ell \left( \frac{i}{N} - r'' \right) \right| d\pi_t(y', y'') \\
& \quad + \sup_{\substack{Lip(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{(\mathbb{T} \times E)^2} |\varphi(v'')| \left| \Gamma_\ell \left( \frac{i}{N} - r'' \right) - \Gamma_\ell(\tilde{r}_0^i - r'') \right| d\mu_t(r'', z'') \\
& \leq \int_{(\mathbb{T} \times E)^2} |v' - v''| \Gamma_\ell \left( \frac{i}{N} - r' \right) d\pi_t(y', y'') + \frac{K}{\ell^2} \int_{(\mathbb{T} \times E)^2} |v''| \mathbb{1}_{\{|\tilde{r}_0^i - r''| < \ell/2\}} |r' - r''| d\pi_t(y', y'') \\
& \quad + \frac{K\epsilon_N}{\ell^2} \int_{(\mathbb{T} \times E)^2} \mathbb{1}_{\{|\tilde{r}_0^i - r''| < \ell/2\}} |v''| d\mu_t(r'', z''), \tag{51}
\end{aligned}$$

where we used that  $|\phi(v'')| \leq |v''|$  and the Lipschitz property for  $\Gamma_\ell$  at the last line, as well as the fact that  $i/N, \tilde{r}_0^i \in B_j$  for some  $j$  and therefore their difference is smaller than  $\epsilon_N$  for the last term (see (20)). Using Lemma 4, we can bound  $\int_{\mathbb{T} \times E} |v''| d\mu_t$  by a constant. Combining (50) and (51) in (49), and taking into account the expectation and the mean over  $i$ , we finally get the bound

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \mathcal{W}_1(w^{\mathbf{V}_t^i}, w^{\mu_t, \tilde{r}_0^i}) \right] \leq K \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v' - v''| d\pi_t(y', y'') \right] \\
& \quad + \frac{K}{\ell^2} \mathbb{E} \left[ \int_{(\mathbb{T} \times E)^2} |v''| \mathbb{1}_{\{|\tilde{r}_0^i - r''| < \ell/2\}} |r' - r''| d\pi_t(y', y'') \right] + \frac{K\epsilon_N}{\ell} + \frac{K}{N\ell},
\end{aligned}$$

for all coupling  $\pi_t$  of  $\mu_t^N$  and  $\mu_t$ . As  $\epsilon_N > 1/N$ , (23) follows.

### 3. Proof of (24)

Using the definition (47) of  $\Pi^i$  and the definition of  $q^{i,k}$  we get

$$\int_{B_j} \left| \Pi_t^i(r, \tilde{r}_0^i, \mathbf{V}_t^i) \right| \sigma^i(r) dr = \int_{\mathbb{R}^d \times \mathbb{R}^d} |v''| \mathbb{1}_{\{v' \in \{V_t^k, k \in NB_j\}\}} dp_{t, \tilde{r}_0^i, \mathbf{V}_t^i}^i(v', v'').$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}
\int_{B_j} \left| \Pi_t^i(r, \tilde{r}_0^i, \mathbf{V}_t^i) \right| \sigma^i(r) dr & \leq \left( \int_{\mathbb{R}^d} \mathbb{1}_{\{v' \in \{V_t^k, k \in NB_j\}\}} dw^{\mathbf{V}_t^i}(v') \right)^{1/2} \left( \int_{\mathbb{R}^d} |v''|^2 dw^{\mu_t, \tilde{r}_0^i}(v'') \right)^{1/2} \\
& \leq \left( \sum_{\substack{k=-\ell N \\ i+k \in NB_j}}^{\ell N} \gamma_k \right)^{1/2} \left( \int_{\mathbb{T} \times E} \Gamma_\ell(\tilde{r}_0^i - r'') |v''|^2 d\mu_t(r'', z'') \right)^{1/2} \\
& \leq K \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \left( 1 + \frac{1}{2} \int_{\mathbb{T} \times E} \Gamma_\ell(\tilde{r}_0^i - r) |v|^2 d\mu_t(r'', z'') \right),
\end{aligned}$$

using at the last line that the coefficients  $\gamma_k$  are bounded by  $K/\ell$  and using the standard inequality  $(1+x)^{1/2} \leq 1+x/2$  to bound the square root. Summing this last inequality over

$i, j$ , taking expectations and using that  $\tilde{r}_0^i$  is uniform random variable on  $B_j$  we get:

$$\frac{1}{N} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{i \in NB_j} \mathbb{E} \left[ \int_{B_j} |\Pi_t^i(r)| \sigma^i(r) dr \right] \leq K \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \left( 1 + \frac{1}{2} \int_{\mathbb{T} \times E} |v''|^2 d\mu_t(r'', z'') \right).$$

Applying Lemma 4 gives finally (24). □

*Proof of Lemma 2.* Fix  $1 \leq j \leq \epsilon_N^{-1}$  and  $i \in NB_j$ . First, by the construction of  $\Pi^i$  in Lemma 1,  $(t, \omega, r) \mapsto \Pi_t^i(r, \tilde{r}_0^i(\omega), \mathbf{V}_{t-}^i(\omega))$  is measurable with respect to the product of the predictable sigma field in  $(t, \omega)$  and the Borel sigma field of  $\mathbb{T}$ . Therefore the integral with respect to the Poisson random measure in (17) is well-defined. Using the terminology of the next section, we are first going to check that the law of the process  $\tilde{Y}^i$  solves the *linear* martingale problem (56) associated with  $\mu$  and starting at  $\mu_0^j$  and finish with a uniqueness argument to prove that its law is in fact  $\mu^j$ . For any  $\psi \in C_b^1(\mathbb{T} \times E)$ , by a direct computation from (17), we have

$$\begin{aligned} \psi(\tilde{Y}_t^i) &= \psi(\tilde{Y}_0^i) + \int_0^t \tilde{V}_s^i \cdot \nabla_x \psi(\tilde{Y}_s^i) ds \\ &\quad - \int_0^t ds \left( \int_{\mathbb{T} \times E} \Phi_\ell(\tilde{r}_s^i - r') \nabla W(\tilde{X}_s^i - x') d\mu_s(r', z') + \nabla U(\tilde{X}_s^i) \right) \cdot \nabla_v \psi(\tilde{Y}_s^i) \\ &\quad + \bar{\gamma} \int_0^t ds \int_{\mathbb{T}} \left( \psi(\tilde{r}_s^i, \tilde{X}_s^i, \Pi_s^i(r')) - \psi(\tilde{r}_s^i, \tilde{X}_s^i, \tilde{V}_s^i) \right) \sigma^i(r') dr' + M_t, \end{aligned} \quad (52)$$

where  $M_t$  is the compensated martingale associated with the Poisson integral. By the definition of  $\Pi^i$  in Lemma 1, the last integral can be written

$$\begin{aligned} &\int_{\mathbb{T}} \left( \psi(\tilde{r}_s^i, \tilde{X}_s^i, \Pi_s^i(r')) - \psi(\tilde{r}_s^i, \tilde{X}_s^i, \tilde{V}_s^i) \right) \sigma^i(r') dr' \\ &= \int_{\mathbb{R}^d} \left( \psi(\tilde{r}_s^i, \tilde{X}_s^i, v') - \psi(\tilde{Y}_s^i) \right) d\mu^{\mu_s, \tilde{r}_0^i}(v') \\ &= \int_{\mathbb{T} \times E} \Gamma_\ell(\tilde{r}_0^i - r') \left( \psi(\tilde{r}_s^i, \tilde{X}_s^i, v') - \psi(\tilde{Y}_s^i) \right) d\mu_s(r', x', v'), \end{aligned}$$

and we can therefore rewrite (52) as

$$\psi(\tilde{Y}_t^i) = \psi(\tilde{Y}_0^i) + \int_0^t \mathcal{L}[\mu_s] \psi(\tilde{Y}_s) ds + M_t.$$

Hence the law of  $\tilde{Y}^i$  solves the linear martingale problem associated with  $\mu$  and starting at  $\mu_0^j$ . We are now left with the task of proving that this property is sufficient to characterize the law  $\mu_t^j(r, z) = \epsilon_N^{-1} \mathbb{1}_{r \in B_j} \mu_t(r, z)$ .

To do so, we construct  $\epsilon_N^{-1}$  processes  $(\hat{Y}^j)_{1 \leq j \leq \epsilon_N^{-1}}$ , whose respective laws solve the linear martingale problem associated with  $\mu$  and starting at  $\mu_0^j$ . Let  $U$  be a uniform  $\mathcal{F}_0$ -measurable random variable on  $\mathbb{T}$ , independent of any  $\hat{Y}^j$ . Now define the process  $Y$  on  $\mathbb{T} \times E$  by

$$Y = \sum_{j=1}^{\epsilon_N^{-1}} \hat{Y}^j \mathbb{1}_{U \in B_j}.$$

We can check that the law of  $Y$  solves the linear martingale problem associated with  $\mu$  and starting at  $\mu_0$ . This suffices to deduce that  $Y$  has law  $\mu$  by uniqueness of the solutions to the

linear martingale problem. In fact, it is easy to see that  $Y_0$  has distribution  $\mu_0$  and, for any  $\psi \in C_b^1(\mathbb{T} \times E)$ , one has

$$\begin{aligned}\psi(Y_t) &= \sum_{j=1}^{\epsilon_N^{-1}} \psi(\hat{Y}_t^j) \mathbb{1}_{U \in B_j} = \sum_{j=1}^{\epsilon_N^{-1}} \left( \psi(\hat{Y}_0^j) + \int_0^t \mathcal{L}[\mu_s] \psi(\hat{Y}_s^j) ds + M_t^{\psi, j} \right) \mathbb{1}_{U \in B_j} \\ &= \psi(Y_0) + \int_0^t \mathcal{L}[\mu_s] \psi(Y_s) ds + M_t^\psi,\end{aligned}$$

where  $M^{\psi, j}$  are martingales and  $M^\psi = \sum M^{\psi, j} \mathbb{1}_{U \in B_j}$  is therefore a martingale.

It now remains to check that  $\hat{Y}^j$  has distribution  $\mu^j$ . For any measurable function  $\varphi$  and for any  $1 \leq j \leq \epsilon_N^{-1}$ , by independence of  $U$  and  $\hat{Y}^j$ ,

$$\mathbb{E} \left[ \varphi(Y_t) \mathbb{1}_{r_t \in B_j} \right] = \mathbb{E} \left[ \varphi(\hat{Y}_t^j) \mathbb{1}_{U \in B_j} \right] = \epsilon_N \mathbb{E} \left[ \varphi(\hat{Y}_t^j) \right].$$

Therefore  $\hat{Y}^j$  has distribution  $\mu^j$  as expected, since

$$\mathbb{E} \left[ \varphi(\hat{Y}_t^j) \right] = \epsilon_N^{-1} \mathbb{E} \left[ \varphi(Y_t) \mathbb{1}_{r_t \in B_j} \right] = \int_{\mathbb{T} \times E} \varphi d\mu_t^j,$$

and this concludes the proof.  $\square$

*Proof of Lemma 3.* Recall that  $r_0^i$  is uniformly distributed over  $B_j$ , where  $j$  is such that  $i \in NB_j$ . We therefore already have the uniform bound

$$\left| \frac{i}{N} - \tilde{r}_0^i \right| \leq \epsilon_N.$$

Recall that  $Z_0^i$  has distribution  $f_0(i/N, z) dz$  and  $\tilde{Z}_0^i$  has distribution  $\epsilon_N^{-1} \int_{B_j} f(r, z) dr dz$ , where the integral is only over  $r$ , *i.e.* the distribution of  $\tilde{Z}_0^i$  is the marginal of  $\mu_0^j$  on the  $z$  coordinates. Therefore, we can choose  $(Z_0^i, \tilde{Z}_0^i)$  to be distributed as the optimal coupling between those two measures and get, by Kantorovich-Rubinstein duality:

$$\begin{aligned}\mathbb{E} \left[ |Z_0^i - \tilde{Z}_0^i| \right] &= \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_E \varphi(z) f_0 \left( \frac{i}{N}, z \right) dz - \epsilon_N^{-1} \int_{B_j \times E} \varphi(z) f_0(r, z) dr dz \\ &= \epsilon_N^{-1} \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(z) \left( f_0 \left( \frac{i}{N}, z \right) - f_0(r, z) \right) dr dz \\ &\leq \epsilon_N^{-1} \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} |\varphi(z)| \left| \frac{i}{N} - r \right| h(z) dr dz \\ &\leq \epsilon_N^{-1} \int_{B_j} \left| \frac{i}{N} - r \right| dr \int_E |z| h(z) dz \\ &\leq K \epsilon_N\end{aligned}$$

using hypothesis **(H4)** at the third line and at the last line. This concludes the proof.  $\square$

*Proof of Lemma 4.* Let  $Y_t$  be the canonical process associated with the solution to the nonlinear martingale problem (8) starting at  $\mu_0$ . Defining the energy

$$\mathcal{E}[\mu_t](Y_t) := \frac{1}{2} |V_t|^2 + \frac{1}{2} \int_{\mathbb{T} \times E} \Phi_\ell(r_0 - r') W(X_t - x') d\mu_t(y') + U(X_t), \quad (53)$$

we can see that its expectation is constant, *i.e.*

$$\mathbb{E}[\mathcal{E}[\mu_t](Y_t)] = \int_{\mathbb{T} \times E} \mathcal{E}[\mu_t](y) d\mu_t(y) = \int_{\mathbb{T} \times E} \mathcal{E}[\mu_0](y) d\mu_0(y),$$

which is finite by assumption **(H3)**. In particular, by assumption **(H1)** on  $U$ , we get the moment bound we expected.

A similar argument holds for the particle system defined by (16), which clearly conserves energy in the sense that:

$$\begin{aligned} \int_{\mathbb{T} \times E} \mathcal{E}[\mu_t^N](y) d\mu_t^N(y) &= \sum_{i=1}^N \left( \frac{1}{2} |V_t^i|^2 + \frac{1}{2} \sum_{k=-\ell N}^{\ell N} \phi_k W(X_t^i - X_t^{i+k}) + U(X_t^i) \right) \\ &= \int_{\mathbb{T} \times E} \mathcal{E}[\mu_0^N](y) d\mu_0^N(y) \end{aligned}$$

almost surely for all  $t > 0$ . □

*Proof of Lemma 5.* The proof of this lemma is similar to that of Lemma 4 and is again based on conservation of energy. Since the law of  $\tilde{Y}_t^i$  is given by  $\mu_t^j$  if  $i \in NB_j$ , one has:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T} \times E} \frac{1}{\ell} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} |\nabla W(\tilde{X}_t^i - x'')| d\mu_t(r'', z'') \right] \\ = \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \int_{(\mathbb{T} \times E)^2} \frac{1}{\ell} \mathbb{1}_{|r' - r''| < \ell/2} |\nabla W(x' - x'')| d\mu_t^j(r', z') d\mu_t(r'', z'') \\ = \int_{(\mathbb{T} \times E)^2} \frac{1}{\ell} \mathbb{1}_{|r' - r''| < \ell/2} |\nabla W(x' - x'')| d\mu_t(r', z') d\mu_t(r'', z''). \end{aligned}$$

By Jensen's inequality and assumption **(H1)** on  $W$ , we get

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_{\mathbb{T} \times E} \frac{1}{\ell} \mathbb{1}_{|\tilde{r}_0^i - r''| < \ell/2} |\nabla W(\tilde{X}_t^i - x'')| d\mu_t(r'', z'') \right] \\ \leq \left( c \int_{(\mathbb{T} \times E)^2} \frac{1}{\ell} \mathbb{1}_{|r' - r''| < \ell/2} W(x' - x'') d\mu_t(r', z') d\mu_t(r'', z'') \right)^{1/2}. \end{aligned}$$

It just remains to show that this last term is bounded by a constant that does not depend on  $\ell$ . This fact is true if we replace  $\ell^{-1} \mathbb{1}_{|r' - r''| < \ell/2}$  by  $\Phi_\ell(r' - r'')$ , by conservation of energy, following the proof of Lemma 4. However,  $\Phi_\ell(r)$  is not necessarily bigger than  $\ell^{-1} \mathbb{1}_{|r| < \ell/2}$ , in particular for  $r$  close to  $\ell/2$  or  $-\ell/2$ . Instead, we are going to take advantage of the integration over  $r'$  and  $r''$  and of the translation invariance of the Lebesgue measure to properly achieve this bound.

Let  $a$  be the only real number in  $[0, 1/2]$  such that  $\phi(a) = \phi(-a) = \|\phi\|_\infty/2$ . Adding several times shifted version of the function  $\phi$ , we can upper bound the indicator function. More precisely, there exists a finite integer  $m$  such that

$$\frac{\|\phi\|_\infty}{2} \mathbb{1}_{|r| \leq 1/2} \leq \sum_{k=-m}^m \phi(r + ka).$$

From this, we deduce that

$$\frac{1}{\ell} \mathbb{1}_{|r| \leq \ell/2} \leq \frac{2}{\|\phi\|_\infty} \sum_{k=-m}^m \Phi_\ell(r + ka).$$





given  $r_0$ .  $w^{\mu_t, r_0}$  is the velocity marginal of  $\mu_t$ , weighted by the function  $\Gamma_\ell$  around  $r_0$ , as defined in (22) (replacing obviously  $\tilde{r}_0^i$  by  $r_0$  in this formula). Moreover, we set  $r_t = r_0$  for all time  $t \geq 0$ . Formally, (54) is given by taking the limit  $\mu$  of  $\mu^N$  in (3).

To deal with the nonlinear problems, we will need to work with solutions of the associated linear problems. Let  $Q \in \mathcal{M}^1(\mathcal{D})$  be a fixed probability measure. We say that  $\mu$  is a solution to the *linear martingale problem* associated with  $Q$  and starting at  $\nu_0 \in \mathcal{M}^1(\mathbb{T} \times E)$  if  $\mu_0 = \nu_0$  and

$$M_t^\psi = \psi(Y_t) - \psi(Y_0) - \int_0^t \mathcal{L}[Q_s] \psi(Y_s) ds \quad (56)$$

is a martingale under  $\mu$  for any  $\psi \in C_b^1(\mathbb{T} \times E)$ , where  $\mathcal{L}$  is the operator defined by (7). Similarly, the *linear mean field SDE* associated with  $Q$  will refer to the SDE (54) where  $\mu$  is replaced by  $Q$ .

**Proposition 3.** *Let  $\nu_0 \in \mathcal{M}^1(\mathbb{T} \times E)$  be such that its  $r$ -marginal is the Lebesgue measure on  $\mathbb{T}$  and  $\int_{\mathbb{T} \times E} |z| d\nu_0$ . Let  $Y_0$  be random variable with law  $\nu_0$  and let  $Q \in \mathcal{D}$  such that for all  $T \geq 0$ ,*

$$\sup_{t \leq T} \int_{\mathbb{T} \times E} |z| dQ_t(r, z) < \infty. \quad (57)$$

*There is a unique solution to the linear stochastic differential equation, both pathwise and in law. Its law  $\mu$  is the unique solution to the linear martingale problem (56) associated with  $Q$  starting at  $\nu_0$ . Moreover, for any  $T > 0$ ,  $\mu$  satisfies*

$$\sup_{t \leq T} \int_{\mathbb{T} \times E} |z| d\mu_t(r, z) < \infty, \quad (58)$$

*and for any  $t > 0$ , and any measurable  $\varphi : \mathbb{T} \rightarrow \mathbb{R}$*

$$\int_{\mathbb{T} \times E} \varphi(r) d\mu_t(r, z) = \int_{\mathbb{T} \times E} \varphi(r) d\nu_0(r, z). \quad (59)$$

The proof of Proposition 3 is classical and left to the reader. We only give the main steps for sake of completeness. It suffices to prove first a contraction result for the solutions of the linear SDE, which is straightforward in the linear case. Then the law of the solutions of the linear SDE solve the linear martingale problem (56). To prove that it is the only one, we can follow [7, 13] and adapt the classical representation arguments from [10], [18] or [25] by proving that the canonical process associated with a solution to the martingale problem can be represented as a solution to the linear SDE on an enlarged probability space.

With this result, we are now able to prove Proposition 1. The proof is inspired by the proof of Graham [15] and relies on a contraction estimate in the space  $\mathcal{M}^1(\mathcal{D}_T)$  for some  $T > 0$ , where  $\mathcal{D}_T := D([0, T], \mathbb{T} \times E)$  is the set of right-continuous functions with left-limits defined on the time interval  $[0, T]$  with values in  $\mathbb{T} \times E$ . To take into account the particular role of the parameter  $r$ , we will use a new appropriate Wasserstein-type distance. Let us first introduce the more classical distance that we will use throughout the proof. From the uniform distance

$$\|y^1 - y^2\|_{\infty, T} := \sup_{t \leq T} |y_t^1 - y_t^2| \quad (60)$$

between  $y^1, y^2 \in \mathcal{D}_T$  and from the Skorokhod distance  $\|\cdot\|_{0, T}$  (see [3]), we can consider the associated Wasserstein distances between  $Q^1, Q^2 \in \mathcal{M}^1(\mathcal{D}_T)$ :

$$\mathcal{W}_{\infty, T}(Q^1, Q^2) = \inf_{\pi \in \Pi(Q^1, Q^2)} \int \|y - y'\|_{\infty, T} d\pi(y, y'),$$

and similarly for  $\mathcal{W}_{0, T}$ . In particular

$$\mathcal{W}_{0, T}(Q^1, Q^2) \leq \mathcal{W}_{\infty, T}(Q^1, Q^2), \quad (61)$$

holds for any measures  $Q^1$  and  $Q^2$  by comparison of the two distances  $\|\cdot\|_{\infty,T}$  and  $\|\cdot\|_{0,T}$ . We will also use the notation  $\mathcal{W}_{\infty,T}$  for Wasserstein distances between probability measures on the space of trajectories on  $\mathbb{T} \times E$ ,  $E$  and  $\mathbb{R}^d$ . We will only deal with measures  $Q^1, Q^2 \in \mathcal{M}^1(\mathcal{D})$  satisfying the following properties

**(h1)**  $\sup_{t \leq T} \int_{\mathbb{T} \times E} |z| dQ_t^i < \infty$  for  $i = 1, 2$  and for all  $T > 0$ .

**(h2)**  $Q^1$  and  $Q^2$  only charge trajectories  $(r_t, x_t, v_t)_{t \geq 0}$  with constant  $r_t = r_0$ . Moreover, the  $r_0$ -marginal of  $Q_0^1$  and  $Q_0^2$  is the Lebesgue measure on  $\mathbb{T}$ .

Under **(h2)**, define the probability measures  $Q^{1,r}$  and  $Q^{2,r}$  on  $D(\mathbb{R}_+, E)$  to be the conditional distributions of  $Q^1$  and  $Q^2$  on the  $z$ -trajectories given  $r$ .

**Definition 1** (Sliced Wasserstein distance). *The sliced Wasserstein distance between two measures  $Q^1, Q^2 \in \mathcal{M}^1(\mathcal{D}_T)$  satisfying **(h2)** is the distance*

$$\mathcal{S}\mathcal{W}_{\infty,T}(Q^1, Q^2) := \int_{\mathbb{T}} dr \mathcal{W}_{\infty,T}(Q^{1,r}, Q^{2,r}). \quad (62)$$

Such a terminology has been used in a different context (see Chapter 5.5.4 in [21]), although the ideas behind both definitions are close. Therefore we will stick to this terminology. We will actually prove the contraction estimate for sliced Wasserstein distance by constructing an explicit coupling of the solutions of the linear martingale problem associated with  $Q^1$  and  $Q^2$ . Moreover, the bound

$$\mathcal{W}_{\infty,T}(Q^1, Q^2) \leq \mathcal{S}\mathcal{W}_{\infty,T}(Q^1, Q^2).$$

follows directly by considering the coupling  $(r, Z, r, Z') \in (\mathcal{D}_T)^2$  of  $Q^1$  and  $Q^2$ , where  $r$  is a uniform random variable on  $\mathbb{T}$  and  $(Z, Z')$  has distribution given by the optimal coupling of  $Q^{1,r}$  and  $Q^{2,r}$ . Combined with (61), we get

$$\mathcal{W}_{0,T}(Q^1, Q^2) \leq \mathcal{S}\mathcal{W}_{\infty,T}(Q^1, Q^2). \quad (63)$$

*Proof of Proposition 1.* In the whole proof,  $K$  and  $K'$  are positive constants that change from one line to another and depend on the parameters of the problem  $W, U, \phi, \gamma$  but also  $\ell$  and  $\bar{\gamma}$ , since we do not need to keep track of these elements anymore.

### 1. Contraction estimate for the sliced Wasserstein distance on a short time interval

Let  $Q^1$  and  $Q^2$  be two elements of  $\mathcal{M}^1(\mathcal{D})$  satisfying **(h1)** and **(h2)**. We construct two coupled solutions  $Y^1 = (r^1, Z^1)$  and  $Y^2 = (r^2, Z^2)$  of the linear SDE associated with  $Q^1$  and  $Q^2$  respectively and with initial distribution  $\nu_0$ . We first couple their initial positions by setting  $r_0^1 = r_0^2 = r_0$ , where  $r_0$  is a uniform random variable on  $\mathbb{T}$ , and we set  $Z_0^1 = Z_0^2$ , with  $Z_0^1$  distributed with the probability measure  $\nu_0^{r_0}$ . Denoting by  $\mu^i$  the law of  $Y^i$ ,  $i = 1, 2$ , we immediately get

$$\mathcal{S}\mathcal{W}_{\infty,T}(\mu^1, \mu^2) \leq \mathbb{E} \left[ \|Z^1 - Z^2\|_{\infty,T} \right]. \quad (64)$$

We then construct the Poisson random measures  $\mathcal{N}^{Q^1, r_0}, \mathcal{N}^{Q^2, r_0}$  that drive the two processes. By **(h2)**, the time intensities of the Poisson random measures  $\mathcal{N}^{Q^1, r_0}$  and  $\mathcal{N}^{Q^2, r_0}$  are equal (to  $\bar{\gamma}$ ), thus we can take the same jump times  $(T_n)_{n \geq 0}$  for both. We will take  $V_{T_n}^1$  and  $V_{T_n}^2$  to be distributed as the optimal coupling of  $w^{Q_{T_n}^1, r_0}$  and  $w^{Q_{T_n}^2, r_0}$ , that is to say:

$$\mathbb{E} \left[ |V_{T_n}^1 - V_{T_n}^2| | r_0, T_n \right] = \mathcal{W}_1(w^{Q_{T_n}^1, r_0}, w^{Q_{T_n}^2, r_0}). \quad (65)$$

We have a good control on (65) with the sliced distance by the following argument. Let  $\pi^r \in \Pi(Q^{1,r}, Q^{2,r})$  be the optimal coupling between  $Q^{1,r}$  and  $Q^{2,r}$  for  $\mathcal{W}_{\infty,T}$ , for  $r \in \mathbb{T}$ . Using

Kantorovich-Rubinstein duality formula and hypothesis **(h2)**, we can bound  $\mathcal{W}_1(w^{Q_t^1, r_0}, w^{Q_t^2, r_0})$  for all  $t \geq 0$ :

$$\begin{aligned} \mathcal{W}_1(w^{Q_t^1, r_0}, w^{Q_t^2, r_0}) &= \sup_{Lip(\varphi) \leq 1} \left| \int_{\mathbb{T}} dr \Gamma_\ell(r_0 - r) \int_{E^2} (\varphi(v') - \varphi(v'')) d\pi_t^r(z', z'') \right| \\ &\leq K \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2), \end{aligned} \quad (66)$$

bounding  $\Gamma_\ell$  by a constant and the difference of the Lipschitz functions by  $|v' - v''|$ . As a consequence we control what happens at jump times. We will only be concerned by the jump times before  $T$ , so set  $S_n = T \wedge T_n$ . The quantity  $d_n := \|Z_t^1 - Z_t^2\|_{\infty, S_n}$  is clearly bounded by

$$d_n \leq \sup_{t < S_n} |Z_t^1 - Z_t^2| + |Z_{S_n}^1 - Z_{S_n}^2|.$$

The last term can be bounded in expectation using that  $X$  is time-continuous and  $V$  is only discontinuous at jump times, in which case (66) controls the difference of the two velocities:

$$\mathbb{E} \left[ |Z_{S_n}^1 - Z_{S_n}^2| \right] \leq \mathbb{E} \left[ \sup_{t < S_n} |Z_t^1 - Z_t^2| \right] + K \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2).$$

Combining this with the previous inequality gives:

$$\mathbb{E}[d_n] \leq K' \mathbb{E}[d_{n-1}] + K' \mathbb{E} \left[ \sup_{S_{n-1} < t < S_n} |Z_t^1 - Z_t^2| \right] + K' \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2). \quad (67)$$

It just remains to control properly the second term of (67) to get a recursive inequality on  $\mathbb{E}[d_n]$ . This second term corresponds to a deterministic evolution between two jump times, so we need in particular to control the difference between the force terms at time  $t$ . By the Lipschitz property for  $\nabla W$  and **(h2)**, we have

$$\begin{aligned} &\left| \int_{\mathbb{T} \times E} \Phi_\ell(r_0 - r') \nabla W(X_t^1 - x') dQ_t^1(r', z') - \int_{\mathbb{T} \times E} \Phi_\ell(r_0 - r'') \nabla W(X_t^2 - x'') dQ_t^2(r'', z'') \right| \\ &\leq \int_{\mathbb{T}} dr \Phi_\ell(r_0 - r) \int_{E^2} |\nabla W(X_t^1 - x') - \nabla W(X_t^2 - x'')| d\pi_t^r(z', z'') \\ &\leq K |X_t^1 - X_t^2| + K \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2), \end{aligned}$$

introducing cross terms at the last line. The contribution of the difference between the pinning terms  $\nabla U$  gives an extra  $K |X_t^1 - X_t^2|$  term. Hence, we get for  $S_{n-1} < t < S_n$

$$\begin{aligned} |Z_t^1 - Z_t^2| &\leq |Z_{S_{n-1}}^1 - Z_{S_{n-1}}^2| + K \int_{S_{n-1}}^t |Z_s^1 - Z_s^2| ds + K T \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2) \\ &\leq d_{n-1} + K T \sup_{S_{n-1} < s < S_n} |Z_s^1 - Z_s^2| + K T \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2). \end{aligned}$$

Taking the supremum for  $S_{n-1} < t < S_n$  on the left-hand side and choosing  $T$  sufficiently small so that  $K T < 1$ , we get the following bound:

$$\sup_{S_{n-1} < t < S_n} |Z_t^1 - Z_t^2| \leq \frac{1}{1 - K T} d_{n-1} + \frac{K T}{1 - K T} \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2). \quad (68)$$

We can now combine this bound (68) for the deterministic dynamic evolution to (67) to get that:

$$\begin{aligned} \mathbb{E}[d_n] &\leq K \frac{2 - K T}{1 - K T} \mathbb{E}[d_{n-1}] + \frac{K}{1 - K T} \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2) \\ &\leq a_T \mathbb{E}[d_{n-1}] + b_T \mathcal{S} \mathcal{W}_{\infty, T}(Q^1, Q^2), \end{aligned}$$

where we denoted for simplicity

$$a_T = K \frac{2 - KT}{1 - KT}, \quad b_T = \frac{K}{1 - KT}.$$

By recursion and using the initial condition  $d_0 = 0$ , we get

$$\mathbb{E}[d_n] \leq \frac{a_T^n - 1}{a_T - 1} b_T \mathcal{SW}_{\infty, T}(Q^1, Q^2). \quad (69)$$

Now that we control the trajectories until time  $S_n$ , we can extend this control up to time  $T$ . Let  $N_T$  be the number of jumps during the interval of time  $[0, T]$ . Then

$$\mathbb{E} \left[ \|Z^1 - Z^2\|_{\infty, T} \right] \leq \mathbb{E}[d_{N_T}] + \mathbb{E} \left[ \sup_{T_{N_T} < t \leq T} |Z_t^1 - Z_t^2| \right]. \quad (70)$$

As there is no jump between times  $T_{N_T}$  and  $T$ , we can apply the same estimates that lead to (68) to the second term on the right-hand side of (70):

$$\begin{aligned} \mathbb{E} \left[ \sup_{T_{N_T} < t \leq T} |Z_t^1 - Z_t^2| \right] &\leq \frac{1}{1 - KT} \mathbb{E}[d_{N_T}] + \frac{KT}{1 - KT} \mathcal{SW}_{\infty, T}(Q^1, Q^2) \\ &\leq \left( \frac{a_T}{K} - 1 \right) \mathbb{E}[d_{N_T}] + \left( \frac{b_T}{K} - 1 \right) \mathcal{SW}_{\infty, T}(Q^1, Q^2), \end{aligned}$$

in terms of  $a_T$  and  $b_T$ . All in all, combining this last inequality with (69) and (70) leads to

$$\begin{aligned} \mathbb{E} \left[ \|Z_t^1 - Z_t^2\|_{\infty, T} \right] &\leq \frac{a_T}{K} \mathbb{E}[d_{N_T}] + \left( \frac{b_T}{K} - 1 \right) \mathcal{SW}_{\infty, T}(Q^1, Q^2) \\ &\leq \left( \frac{a_T b_T}{K(a_T - 1)} \left( \mathbb{E}[a_T^{N_T}] - 1 \right) + \frac{b_T}{K} - 1 \right) \mathcal{SW}_{\infty, T}(Q^1, Q^2). \end{aligned} \quad (71)$$

As  $N_T$  has a Poisson distribution with parameter  $\bar{T}$  given  $r_0$ , the coefficient multiplying  $\mathcal{SW}_{\infty, T}(Q^1, Q^2)$  can be computed explicitly:

$$\begin{aligned} c_T &:= \frac{a_T b_T}{K(a_T - 1)} \left( \mathbb{E}[a_T^{N_T}] - 1 \right) + \frac{b_T}{K} - 1 \\ &= \frac{a_T b_T}{K(a_T - 1)} \left( \exp(\bar{\gamma} T (a_T - 1)) - 1 \right) + \frac{b_T}{K} - 1. \end{aligned}$$

Since  $b_0 = K$ , it is easy to see that  $c_0 = 0$ . As  $c_T$  is a continuous function of  $T$ , we can choose  $T$  small enough so that  $c_T < 1$ . Combining (71) with (64) finally gives the following contraction estimate for a short time  $T > 0$ :

$$\mathcal{SW}_{\infty, T}(\mu^1, \mu^2) \leq c_T \mathcal{SW}_{\infty, T}(Q^1, Q^2). \quad (72)$$

## 2. Existence and uniqueness of solutions to the martingale problem

Uniqueness in  $\mathcal{M}^1(\mathcal{D}_T)$  follows from the contraction estimate (72). Existence in  $\mathcal{M}^1(\mathcal{D}_T)$  follows from a classical iteration procedure. By (58) and (59) in Proposition 3, the solutions of the linear martingale problem are well defined iteratively and satisfy hypotheses **(h1)** and **(h2)**. By the contraction estimate (72), we conclude that the sequence is Cauchy for  $\mathcal{SW}_{\infty, T}$  and therefore for  $\mathcal{W}_{0, T}$  by (63). From the completeness of  $(\mathcal{D}_T, \|\cdot\|_{0, T})$ ,  $\mathcal{W}_{0, T}$  makes  $\mathcal{M}^1(\mathcal{D}_T)$  a complete metric space (see [26]). Consequently, the sequence converges in  $\mathcal{M}^1(\mathcal{D}_T)$ .

Finally, we extend existence and uniqueness to  $\mathcal{M}^1(\mathcal{D})$  by the following argument. Let  $\mu$  be the solution to the nonlinear martingale problem starting at  $\nu_0$  on  $\mathcal{M}^1(\mathcal{D}_T)$  and let  $Q^1$  and  $Q^2$  be two measures on  $\mathcal{M}^1(\mathcal{D})$  such that  $Q_t^1 = Q_t^2 = \mu_t$  for  $t \leq T$ . Define  $\mu^1$  and  $\mu^2$  to be the solutions of the linear martingale problem associated with  $Q^1$  and  $Q^2$  on  $\mathcal{M}^1(\mathcal{D}_{2T})$ . Thus  $\mu_T^1 = \mu_T^2$  by uniqueness of the linear martingale problem on  $\mathcal{M}^1(\mathcal{D}_T)$  and the same contraction estimate as (72) hold for the interval  $[T, 2T]$  between  $\mu^1, \mu^2$  and  $Q^1, Q^2$ . Thus existence and uniqueness extends to  $\mathcal{M}^1(\mathcal{D}_{2T})$  and by immediate recurrence to  $\mathcal{M}^1(\mathcal{D})$ .  $\square$

The same Picard iteration used in the previous proof enables to prove the following Proposition.

**Proposition 4.** *There is existence and uniqueness in law of solutions to the SDE (54).*

## 4 Energy transport

We finish by studying the transport of energy in our model for a class of pinning potentials  $U$  and proving Proposition 2. The proof is in two parts. In the first one, we prove a diffusion equation for the energy associated with the limit measure  $\mu_t$  at the timescale  $t\ell^{-2}$ . In the second part, we prove that the energy of the particle system converges to the energy associated with  $\mu_t$  in that timescale.

### 4.1 Diffusion equation for the limit measure

First, let us start by proving that the symmetry (H5) is preserved at any time  $t > 0$  for symmetric potentials.

**Lemma 7.** *Under (H5), if  $U$  and  $W$  are symmetric, then the symmetry  $\mu_t(r, z) = \mu_t(r, -z)$  for any  $(r, z) \in \mathbb{T} \times E$  is preserved at any later time  $t > 0$ .*

*Proof.* Let  $(Y_t)_{t \geq 0} = (r_t, X_t, V_t)_{t \geq 0}$  be the canonical process associated with the solution  $\mu$  to the nonlinear martingale problem (8). Define  $\mu^*(r, z) = \mu(r, -z)$ ,  $\psi^*(r, z) = \psi(r, -z)$  and  $Y_t^* = (r_t, -Z_t)$  for simplicity. Then, we can rewrite

$$\begin{aligned} & \mathcal{A}[\mu_s]\psi(Y_s) \\ &= V_s \cdot \nabla_v \psi(Y_s) - \left( \nabla U(X_s) + \int_{\mathbb{T} \times E} \Phi_\ell(r_s - r') \nabla W(X_s - x') d\mu_s(r', z') \right) \cdot \nabla_x \psi(Y_s) \\ &= -V_s \cdot \nabla_v \psi^*(Y_s^*) + \left( \nabla U(X_s) + \int_{\mathbb{T} \times E} \Phi_\ell(r_s - r') \nabla W(X_s + x') d\mu_s^*(r', z') \right) \cdot \nabla_x \psi^*(Y_s^*) \\ &= \mathcal{A}[\mu_s^*]\psi^*(Y_s^*), \end{aligned}$$

using antisymmetry of  $\nabla U$  and  $\nabla W$  at the last line. Similarly,

$$\mathcal{S}[\mu_s]\psi(Y_s) = \int_{\mathbb{T} \times E} \Gamma_\ell(r_s - r') (\psi^*(r_s, -X_s, v') - \psi^*(Y_s^*)) d\mu_s^*(r', z') = \mathcal{S}[\mu_s^*]\psi^*(Y_s^*).$$

As a consequence, for any  $\psi \in C_b^1(\mathbb{T} \times E)$ ,

$$\begin{aligned} \psi^*(Y_t^*) - \psi^*(Y_0^*) + \int_0^t \mathcal{L}[\mu_s^*]\psi^*(Y_s^*) ds &= \psi(Y_t) - \psi(Y_0) + \int_0^t \mathcal{L}[\mu_s]\psi(Y_s) ds \\ &= M_t^\psi, \end{aligned}$$

where  $M_t^\psi$  is a martingale. Since  $Y^*$  has distribution  $\mu^*$ ,  $\mu^*$  is also solution to the nonlinear martingale problem (8) starting at  $\mu_0^* = \mu_0$ . By uniqueness,  $\mu^* = \mu$ , which implies the property we were looking for.  $\square$

As a consequence of this symmetry property, the integral  $\int_{\mathbb{R}^d} x' d\mu_t(r', z')$  is null. In particular, if the interaction potential is harmonic  $W(x) = |x|^2/2$ , then the Vlasov equation (6) is reduced to the simplified version (10). We will always consider harmonic interaction potentials from now on. Let  $Y$  be a solution to the nonlinear SDE (54) with spatial parameter  $r \in \mathbb{T}$  and recall the notation from (53):

$$\begin{aligned}\mathcal{E}[\mu_t](Y_t) &= \frac{1}{2}|V_t|^2 + \frac{1}{2} \int_{\mathbb{T} \times E} \Phi_\ell(r - r') W(X_t - x') d\mu_t(r', z') + U(X_t) \\ &= \frac{1}{2}|V_t|^2 + \frac{1}{4}|X_t|^2 + \frac{1}{4} \int_{\mathbb{T} \times E} \Phi_\ell(r - r') |x'|^2 d\mu_t(r', z') + U(X_t),\end{aligned}\quad (73)$$

since the term  $\frac{1}{2} X_t \cdot \int_{\mathbb{T} \times E} \Phi_\ell(r - r') x' d\mu_t(r', z') = 0$  by Lemma 7. Define also the local energy associated with the distribution  $\mu$  of  $Y$ , given  $r_0 = r$ , by

$$\begin{aligned}\mathcal{E}_t(r) &= \mathbb{E}[\mathcal{E}[\mu_t](Y_t) | r_0 = r] \\ &= \int_E \left( \frac{1}{2}|v|^2 + \frac{1}{2} \int_{\mathbb{T} \times E} \Phi_\ell(r - r') W(x - x') d\mu_t(r', z') + U(x) \right) d\mu_t(r, x, v) \\ &= \int_E \left( \frac{1}{2}|v|^2 + \frac{1}{4} \int_{\mathbb{T} \times E} \Phi_\ell(r - r') |x'|^2 d\mu_t(r', z') + \frac{1}{4}|x|^2 + U(x) \right) d\mu_t(r, x, v).\end{aligned}\quad (74)$$

When  $W$  is harmonic, the contribution of the neighboring parts near the axial position  $r$  in (74) is only given by the integral  $\int_{\mathbb{T} \times E} \Phi_\ell(r - r') |x'|^2 d\mu_t(r', z')$ , which does not involve the variable  $x$ . Let us investigate in more detail the evolution of  $\mathcal{E}_t(r)$ . By a straightforward calculation from (54), we get that for any  $g \in \mathcal{C}^2(\mathbb{T})$ ,

$$\begin{aligned}\int_{\mathbb{T}} \mathcal{E}_{t\ell^{-2}}(r) g(r) dr - \int_{\mathbb{T}} \mathcal{E}_0(r) g(r) dr \\ &= \mathbb{E}[\mathcal{E}[\mu_{t\ell^{-2}}](Y_{t\ell^{-2}}) g(r)] - \mathbb{E}[\mathcal{E}[\mu_0](Y_0) g(r)] \\ &= - \int_0^{t\ell^{-2}} ds \int_{\mathbb{T}^2} g(r) \Phi_\ell(u) j_s^{r, r+u; a} dr du - \bar{\gamma} \int_0^{t\ell^{-2}} ds \int_{\mathbb{T}^2} g(r) \Gamma_\ell(u) j_s^{r, r+u; s} dr du,\end{aligned}\quad (75)$$

where

$$\begin{aligned}j_s^{r, r+u; a} &= \frac{1}{2} \int_{E^2} (v + v') \cdot \nabla W(x - x') d\mu_s(r + u, z') d\mu_s(r, z) \\ &= \frac{1}{2} \int_E v \cdot x d\mu_s(r, z) - \frac{1}{2} \int_E v' \cdot x' d\mu_s(r + u, z')\end{aligned}$$

is the mechanical contribution to the energy current and

$$j_s^{r, r+u; s} = \int_E \frac{1}{2} |v|^2 d\mu_s(r, z) - \int_E \frac{1}{2} |v|^2 d\mu_s(r + u, z').$$

is the stochastic contribution to the energy current. Rearranging terms by an integration by parts in (75), since  $r \in \mathbb{T}$  and  $\Phi_\ell$  and  $\Gamma_\ell$  are symmetric, we get

$$\begin{aligned}\int_{\mathbb{T}} \mathcal{E}_{t\ell^{-2}}(r) g(r) dr - \int_{\mathbb{T}} \mathcal{E}_0(r) g(r) dr \\ &= \int_0^{t\ell^{-2}} ds \int_{\mathbb{T} \times E} \frac{1}{2} v \cdot x d\mu_s(r, z) \int_{\mathbb{T}} \Phi_\ell(u) (g(r + u) - g(r)) du \\ &\quad + \bar{\gamma} \int_0^{t\ell^{-2}} ds \int_{\mathbb{T} \times E} \frac{1}{2} |v|^2 d\mu_s(r, z) \int_{\mathbb{T}} \Gamma_\ell(u) (g(r + u) - g(r)) du \\ &= c_\phi \int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} v \cdot x g''(r) d\mu_{s\ell^{-2}}(r, z) + \bar{\gamma} c_\gamma \int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} |v|^2 g''(r) d\mu_{s\ell^{-2}}(r, z) + O(t\ell),\end{aligned}\quad (76)$$

where the error term comes from a Taylor expansion of  $g$  at the last line and we used the uniform in time moment bound of Lemma 4. The constant  $c_\phi$  is equal to

$$c_\phi = \frac{1}{2} \int_{-1/2}^{1/2} u^2 \phi(u) du,$$

and similarly for  $c_\gamma$ . Notice that, even if we have the symmetry property given by Lemma 7, this is not sufficient to prove that the hamiltonian contribution  $\int x \cdot v d\mu_t(r, z)$  in (76) vanishes. However, it is easy to prove from (54) that

$$\begin{aligned} \int_{\mathbb{T} \times E} |x|^2 g''(r) d\mu_{t\ell^{-2}}(r, z) - \int_{\mathbb{T} \times E} |x|^2 g''(r) d\mu_0(r, z) \\ = \ell^{-2} \int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} x \cdot v g''(r) d\mu_{s\ell^{-2}}(r, z), \end{aligned} \quad (77)$$

and consequently, applying uniform in time moment bound of Lemma 4 to the left-hand side of (77), we get an estimate of the right-hand side of (77):

$$\int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} x \cdot v g''(r) d\mu_{s\ell^{-2}}(r, z) = O(t\ell^2). \quad (78)$$

To prove that  $\mathcal{E}_{t\ell^{-2}}(r)$  evolves diffusively, it remains to close the equation (76) by replacing the kinetic energy integral  $\int_E \frac{1}{2} |v|^2 d\mu_{s\ell^{-2}}(r, z)$ , by the energy  $\mathcal{E}_{t\ell^{-2}}(r)$ . To do so, we prove a result of equipartition of energy in Lemma 8. This result states that long time integrals of kinetic energy can be actually replaced by a fraction of time integrals of the total energy. We prove it only for specific pinning potentials for which the identity

$$x \cdot \nabla U(x) = 2U(x)$$

holds. For a positive continuously differentiable function  $U$ , it is well known that satisfying this identity is equivalent to be homogeneous of degree 2. In particular,  $U$  is such that  $U(x) = |x|^2 \psi(x/|x|)$ , where  $\psi \in C^2(\mathbb{S}^{d-1}, \mathbb{R}_+^*)$ . It is straightforward to check that  $U$  satisfies (H1) in this setting. We furthermore require  $\psi$  to be symmetric for Lemma 7 to hold.

**Lemma 8.** *Let  $U(x) = |x|^2 \psi(x/|x|)$ , where  $\psi \in C^2(\mathbb{S}^{d-1}, \mathbb{R}_+^*)$  is a symmetric function. For any  $G \in C^2(\mathbb{T})$ , we have the following time equipartition of energy.*

$$\int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} |v|^2 G(r) d\mu_{s\ell^{-2}}(r, z) = \frac{1}{2} \int_0^t ds \int_{\mathbb{T}} \mathcal{E}_{s\ell^{-2}}(r) G(r) dr + O((1 + \bar{\gamma})t\ell^2). \quad (79)$$

*Proof.* Let  $(Y_s)_{s \geq 0} = (r_0, X_s, V_s)_{t \geq 0}$  be a solution to the nonlinear SDE (54). We compute for any  $T > 0$ :

$$\begin{aligned} X_T \cdot V_T - X_0 \cdot V_0 \\ = \int_0^T |V_s|^2 ds - \int_0^T X_s \cdot (\nabla U(X_s) + X_s) ds + \int_0^T \int_{\mathbb{R}^d} X_s \cdot (v' - V_{s-}) d\mathcal{N}^{\mu, r_0}(s, v'). \end{aligned}$$

Using that  $x \cdot \nabla U(x) = 2U(x)$ , we can introduce the potential  $U$  at the last line. Then, using (73), we can introduce  $\mathcal{E}[\mu_s](Y_s)$  and get

$$\begin{aligned} X_T \cdot V_T - X_0 \cdot V_0 \\ = 4 \int_0^T \frac{1}{2} |V_s|^2 ds - 2 \int_0^T \mathcal{E}[\mu_s](Y_s) ds + \int_0^T ds \left( \frac{1}{2} \int_{\mathbb{T} \times E} \Phi_\ell(r_0 - r') |x'|^2 d\mu_s(r', z') - \frac{1}{2} |X_s|^2 \right) \\ + \int_0^T \int_{\mathbb{R}^d} X_s \cdot (v' - V_{s-}) d\mathcal{N}^{\mu, r_0}(s, v'). \end{aligned}$$

Multiplying by  $G(r_0)$  on both sides for  $G \in \mathcal{C}^2(\mathbb{T})$  and taking expectations gives:

$$\begin{aligned} & \int_{\mathbb{T} \times E} x \cdot vG(r) d\mu_T(r, z) - \int_{\mathbb{T} \times E} x \cdot vG(r) d\mu_0(r, z) \\ &= 4 \int_0^T ds \int_{\mathbb{T} \times E} \frac{1}{2} |v|^2 G(r) d\mu_s(r, z) - 2 \int_0^T ds \int_{\mathbb{T}} \mathcal{E}_s(r) G(r) dr \\ & \quad + \ell^2 c_\phi \int_0^T ds \int_{\mathbb{T} \times E} \frac{1}{2} |x|^2 G''(r) d\mu_s(r, z) + \ell^2 \bar{\gamma} c_\gamma \int_0^T ds \int_{\mathbb{T} \times E} x \cdot vG''(r) d\mu_s(r, z) + o(T\ell^2), \end{aligned}$$

where we used a Taylor expansion for the two terms at the last line. Applying this result for  $T = t\ell^{-2}$ , and rearranging terms, we get:

$$\begin{aligned} & \int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} |v|^2 G(r) d\mu_{s\ell^{-2}}(r, z) - \frac{1}{2} \int_0^t ds \int_{\mathbb{T}} \mathcal{E}_{s\ell^{-2}}(r) G(r) dr \\ &= \ell^2 \frac{1}{4} \left( \int_{\mathbb{T} \times E} x \cdot vG(r) d\mu_{t\ell^{-2}}(r, z) - \int_{\mathbb{T} \times E} x \cdot vG(r) d\mu_0(r, z) \right) \\ & \quad + \ell^2 \frac{c_\phi}{4} \int_0^t ds \int_{\mathbb{T} \times E} \frac{1}{2} |x|^2 G''(r) d\mu_{s\ell^{-2}}(r, z) + \ell^2 \bar{\gamma} \frac{c_\gamma}{4} \int_0^t ds \int_{\mathbb{T} \times E} x \cdot vG''(r) d\mu_{s\ell^{-2}}(r, z) + o(t\ell^2). \end{aligned}$$

Applying the uniform moment bound in Lemma 4, we deduce that the right-hand side is a  $O((1 + \bar{\gamma})\ell^2)$  and this concludes the proof.  $\square$

Combining the equipartition result (79) and the control on the Hamiltonian current (78) in (76), we finally get that  $\mathcal{E}_{t\ell^{-2}}(r)$  evolves diffusively:

$$\int_{\mathbb{T}} \mathcal{E}_{t\ell^{-2}}(r) g(r) dr - \int_{\mathbb{T}} \mathcal{E}_0(r) g(r) dr = \bar{\gamma} \frac{c_\gamma}{2} \int_0^t ds \int_{\mathbb{T}} \mathcal{E}_{s\ell^{-2}}(r) g''(r) dr + O(t\ell + t\bar{\gamma}\ell^2), \quad (80)$$

for any  $g \in \mathcal{C}^4(\mathbb{T})$ .

## 4.2 Convergence of the particle system's energy

We deduce Proposition 2 from the convergence of the microscopic energy to  $\mathcal{E}_t(r)$  proven in the next lemma. Let us fix  $\bar{\gamma}$  and let  $c(N, \ell, t)$  denote the constant appearing in Theorem 1, *i.e.*

$$c(N, \ell, t) = K_1 \left( (N\epsilon_N)^{-\frac{1}{4(d+1)}} + \frac{\epsilon_N^{1/2}}{\ell^{1/2}} \right) e^{K_2 t}.$$

Recall from (11) the definition of the microscopic energy

$$\mathcal{E}_t^i = \mathcal{E}[\mu_t^N](Y_t^i) = \frac{1}{2} |V_t^i|^2 + \frac{1}{4} \int_{\mathbb{T} \times E} \Phi_\ell \left( \frac{i}{N} - r' \right) |X^i - x'|^2 d\mu_t^N(r', z') + U(X_t^i).$$

**Lemma 9.** *Let  $G \in \mathcal{C}^1(\mathbb{T})$ . Under hypothesis (H6), there exist positive constants  $K, K'$  such that for any time  $T > 0$  and a constant  $M > 0$  large enough*

$$\mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^N \mathcal{E}_T^i G \left( \frac{i}{N} \right) - \int_{\mathbb{T}} \mathcal{E}_T(r) G(r) dr \right| \right] \leq M^2 c(N, \ell, T) + \frac{K}{M^b} e^{K'T} + O(\ell^2).$$

In this result,  $b$  is the constant appearing in the moment hypothesis (H6). Taking  $T = t\ell^{-2}$ ,  $M = c(N, \ell, T)^{-1/3}$  and  $\epsilon_N = \ell^{\frac{2d+2}{2d+3}} N^{-\frac{1}{2d+3}}$  gives the result stated in Proposition 2. In particular, one can find a constant  $c > 0$  and choose  $\ell = \ell(N) = c(\log N)^{-1/2}$  so that (12) holds.



*Proof.* First, let us simplify the problem. In the definition (74) of the energy  $\mathcal{E}_T(r)$ , we expect that the approximation

$$\frac{1}{4} \int_{\mathbb{T}} \Phi_\ell(r-r') |x'|^2 d\mu_T(r', z') \approx \frac{1}{4} |x|^2$$

is true when  $\ell$  is small. We can actually make this approximation rigorous when we integrate  $\mathcal{E}_T(r)G(r)$ :

$$\begin{aligned} \int_{\mathbb{T}} \mathcal{E}_T(r)G(r)dr &= \int_{\mathbb{T} \times E} \left( \left( \frac{1}{2}|v|^2 + \frac{1}{4}|x|^2 + U(x) \right) G(r) + \frac{1}{4}|x|^2 \int_{\mathbb{T}} \Phi_\ell(u)G(r+u)du \right) d\mu_T(r, z) \\ &= \int_{\mathbb{T} \times E} \left( \frac{1}{2}|v|^2 + \frac{1}{2}|x|^2 + U(x) \right) G(r) d\mu_T(r, z) + O(\ell^2), \end{aligned}$$

by a Taylor expansion of  $G$ . By the same procedure for the microscopic system, we have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathcal{E}_T^i G\left(\frac{i}{N}\right) &= \int_{\mathbb{T} \times E} \left( \frac{1}{2}|v|^2 + \frac{1}{2}|x|^2 + U(x) \right) G(r) d\mu_T^N(r, z) \\ &\quad - \frac{1}{2} \int_{(\mathbb{T} \times E)^2} x \cdot x' G(r) d\mu_T^N(r, z) d\mu_T^N(r', z') + \delta_\ell, \end{aligned}$$

where  $\delta_\ell$  is such that  $\mathbb{E}[|\delta_\ell|] = O(\ell^2)$  by conservation of energy. Notice that we have a non-vanishing extra term for the microscopic version. We introduce

$$H(x, v) = \frac{1}{2}|v|^2 + \frac{1}{2}|x|^2 + U(x)$$

for notational convenience. The proof of the lemma now boils down to bound the following two terms

$$\mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} H(x, v)G(r) d\mu_T^N(r, z) - \int_{\mathbb{T} \times E} H(x, v)G(r) d\mu_T(r, z) \right| \right] \quad (81)$$

and

$$\mathbb{E} \left[ \left| \int_{(\mathbb{T} \times E)^2} x \cdot x' G(r) d\mu_T^N(r, z) d\mu_T^N(r', z') - \int_{(\mathbb{T} \times E)^2} x \cdot x' G(r) d\mu_T^N(r, z) d\mu_T(r', z') \right| \right], \quad (82)$$

since the second integral in (82) is null by Lemma 7.

For that, we are going to use Theorem 1. However, since the Wasserstein  $\mathcal{W}_1$  distance only enables to control differences of integrals with respect to Lipschitz functions with Lipschitz constant less than 1, we have to cut the large values of  $H$  and  $x \cdot x'$ . To control (81), we therefore introduce a sequence of functions  $H^M$  depending on a parameter  $M$  such that  $H^M$  approximates  $H$ :

$$H^M(z) := H(z) \mathbb{1}_{H(z) < M} + M \left( 2 - \exp\left(1 - \frac{H(z)}{M}\right) \right) \mathbb{1}_{H(z) \geq M}.$$

With this choice, it is easy to check that  $H^M/M^2$  is a Lipschitz continuous function such that

$$\left\| \frac{1}{M^2} H^M \right\|_\infty \leq \frac{2}{M}, \quad \text{Lip}\left(\frac{1}{M^2} H^M\right) \leq \frac{1}{M^2}.$$

In particular,

$$\text{Lip}\left(\frac{1}{M^2} H^M G\right) \leq \text{Lip}\left(\frac{1}{M^2} H^M\right) \|G\|_\infty + \left\| \frac{1}{M^2} H^M \right\|_\infty \text{Lip}(G) \leq \frac{1}{M^2} \|G\|_\infty + \frac{2}{M} \text{Lip}(G),$$

which is less than one for  $M$  large enough. Introducing cross terms in (81), we now bound it by

$$\begin{aligned} & \mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} H(z) G(r) d\mu_T^N(r, z) - \int_{\mathbb{T} \times E} H(z) G(r) d\mu_T(r, z) \right| \right] \\ & \leq \|G\|_\infty \mathbb{E} \left[ \int_{\mathbb{T} \times E} |H(z) - H^M(z)| d\mu_T^N(r, z) \right] + \|G\|_\infty \int_{\mathbb{T} \times E} |H(z) - H^M(z)| d\mu_T(r, z) \\ & \quad + M^2 \mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} \frac{1}{M^2} H^M(z) G(r) d\mu_T^N(r, z) - \int_{\mathbb{T} \times E} \frac{1}{M^2} H^M(z) G(r) d\mu_T(r, z) \right| \right]. \end{aligned} \quad (83)$$

We can bound the last term by  $M^2 \mathbb{E}[\mathcal{W}(\mu_t^N, \mu_t)]$  and get

$$\mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} H^M(z) G(r) d\mu_T^N(r, z) - \int_{\mathbb{T} \times E} H^M(z) G(r) d\mu_T(r, z) \right| \right] \leq M^2 c(N, \ell, T). \quad (84)$$

It just remains to control the first two terms in (83) which correspond to the cut parts. The first term in (83) can be bounded by

$$\begin{aligned} \mathbb{E} \left[ \int_{\mathbb{T} \times E} |H(z) - H^M(z)| d\mu_T^N(r, z) \right] & \leq \mathbb{E} \left[ \int_{\mathbb{T} \times E} H(z) \mathbb{1}_{H(z) \geq M} d\mu_T^N(r, z) \right] \\ & \leq \frac{1}{M^b} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N H(X_T^i, V_T^i)^{1+b} \right], \end{aligned} \quad (85)$$

by a Markov inequality, with the view of using the moment hypothesis (H6). We now derive a moment type bound at time  $T$ . Using the dynamics (16), and the fact that  $\nabla_v H = v$  and  $\nabla_x H = x + \nabla U(x)$ , we compute the last term:

$$\begin{aligned} & \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N H(X_T^i, V_T^i)^{1+b} \right] - \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N H(X_0^i, V_0^i)^{1+b} \right] \\ & = (1+b) \int_0^T \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \sum_{k=-\ell N}^{\ell N} \phi_k V^i \cdot X^{i+k} H(X_s^i, V_s^i)^b \right] \\ & \quad + \bar{\gamma} \int_0^T \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \sum_{k=-\ell N}^{\ell N} \gamma_k \left( H(X_s^i, V_s^{i+k})^{1+b} - H(X_s^i, V_s^i)^{1+b} \right) \right] ds, \end{aligned} \quad (86)$$

the first term being the hamiltonian contribution and the second one the stochastic contribution. We then bound the product

$$\begin{aligned} V^i \cdot X^{i+k} H(X_s^i, V_s^i)^b & \leq \left( \frac{1}{2} |V^i|^2 + \frac{1}{2} |X^{i+k}|^2 \right) H(X_s^i, V_s^i)^b \\ & \leq H(X_s^i, V_s^i)^{1+b} + H(X^{i+k}, V^{i+k}) H(X_s^i, V_s^i)^b \\ & \leq K H(X_s^i, V_s^i)^{1+b} + K' H(X_s^{i+k}, V_s^{i+k})^{1+b}, \end{aligned}$$

where  $K$  and  $K'$  are two constants. We bounded directly the two terms by  $H$  at the second line, and then used Young's inequality at the third line. Summing this inequality over  $i$  and  $k$ , we obtain that the first term in the right-hand side of (86) is thus bounded by

$$\int_0^T \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \sum_{k=-\ell N}^{\ell N} \phi_k V^i \cdot X^{i+k} H(X_s^i, V_s^i)^b \right] \leq K \int_0^T \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N H(X_s^i, V_s^i)^{1+b} \right], \quad (87)$$

for some constant  $K$ . The same bound can be obtained for the second term in the right-hand side of (86). All in all, By Gronwall's inequality and (H6), we get:

$$\mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N H(X_T^i, V_T^i)^{1+\epsilon} \right] \leq K e^{K'T}.$$

Eventually, the term (85) corresponding to the large energies is bounded by

$$\mathbb{E} \left[ \int_{\mathbb{T} \times E} |H(z) - H^M(z)| d\mu_T^N(r, z) \right] \leq \frac{K}{M^b} e^{K'T}.$$

The same bound can be obtained for the second term in (83). Combining both bounds together with (84) in (83) gives the bound we wanted for the convergence of  $H$  in (81):

$$\mathbb{E} \left[ \left| \int_{\mathbb{T} \times E} H(z) G(r) d\mu_T^N(r, z) - \int_{\mathbb{T} \times E} H(z) G(r) d\mu_T(r, z) \right| \right] \leq M^2 c(N, \ell, T) + \frac{K}{M^b} e^{K'T}.$$

(82) can be controlled in the same spirit, and this concludes the proof.  $\square$

## A Convergence of some empirical measure for the Wasserstein distance

For any  $1 \leq j \leq \epsilon_N^{-1}$ , let  $(\bar{Y}^i)_{i \in NB_j}$  be a family of independent random variables on  $B_j \times \mathbb{R}^d \times \mathbb{R}^d$  with law  $\mu^j(r, z) = \epsilon_N^{-1} \mu(r, z) \mathbb{1}_{r \in B_j}$ , and  $\bar{Y}^i = (\bar{r}^i, \bar{X}^i, \bar{V}^i)$ . In this section, we prove that the empirical measure  $\bar{\mu}^{N,j} = 1/(N\epsilon_N) \sum \delta_{\bar{Y}^i}$  satisfies a law of large numbers in some sense. More precisely, we prove that

**Proposition 5.** *If  $\mu$  has a finite moment of order 2 in the sense that  $\int_{\mathbb{T} \times E} |z|^2 d\mu(r, z) < \infty$ . Then there exists some positive constant  $K$  such that*

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\bar{\mu}^{N,j}, \mu^j) \right] \leq K(N\epsilon_N)^{-\frac{1}{4(d+1)}} + K\epsilon_N.$$

Notice that more refined versions of law of large numbers for empirical measures in the Wasserstein distance exist (see [11] for instance). But the proposition we prove here has an easy proof and is well suited for our problem.

*Proof.* Let  $M > 0$ . For any  $1 \leq j \leq \epsilon_N^{-1}$ , using Kantorovich-Rubinstein duality formula, we can bound:

$$\begin{aligned} \mathbb{E} \left[ \mathcal{W}_1(\bar{\mu}^{N,j}, \mu^j) \right] &\leq \mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| \leq M\}} d(\bar{\mu}^{N,j} - \mu^j) \right] \\ &\quad + \mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| > M\}} d(\bar{\mu}^{N,j} - \mu^j) \right]. \end{aligned} \quad (88)$$

The second term in (88) can be bounded using Markov inequality type arguments. First, we bound it by

$$\begin{aligned} &\mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| > M\}} d(\bar{\mu}^{N,j} - \mu^j) \right] \\ &\leq \mathbb{E} \left[ \int_{B_j \times E} (r + |z|) \mathbb{1}_{\{|z| > M\}} d(\bar{\mu}^{N,j} + \mu^j) \right] \\ &\leq 2\epsilon_N^{-1} \int_{B_j \times E} \mathbb{1}_{\{|z| > M\}} d\mu(r, z) + 2\epsilon_N^{-1} \int_{B_j \times E} |z| \mathbb{1}_{\{|z| > M\}} d\mu(r, z). \end{aligned}$$

Taking the mean over  $j$  and then applying Markov inequality in the last expression gives:

$$\begin{aligned}
\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \sup_{\substack{Lip(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{\mathbb{T} \times E} \varphi(r, z) \mathbb{1}_{\{|z| > M\}} d(\bar{\mu}^{N,j} - \mu^j) \right] \\
\leq 2 \int_{\mathbb{T} \times E} \mathbb{1}_{\{|z| > M\}} d\mu(r, z) + 2 \int_{\mathbb{T} \times E} |z| \mathbb{1}_{\{|z| > M\}} d\mu(r, z) \\
\leq \frac{K}{M}, \tag{89}
\end{aligned}$$

for some constant  $K$ . To bound the first term in (88), let us subdivide  $[-M, M]^{2d}$  in  $n^{2d}$  disjoint cubes  $(C_k)_{1 \leq k \leq n^{2d}}$ , for some integer  $n \geq 1$ . More precisely, each cube  $C_k$  is of the form  $[i_1 M/n, (i_1 + 1)M/n] \times \dots \times [i_{2d} M/n, (i_{2d} + 1)M/n]$  for  $-n \leq i_1, \dots, i_{2d} \leq n - 1$ . Let  $z_k \in C_k$  be a point in the cube  $C_k$ , let us say the center of  $C_k$  to fix ideas. Then for any  $\varphi$  such that  $Lip(\varphi) \leq 1$  and  $\varphi(0) = 0$ , one has

$$\begin{aligned}
& \left| \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| \leq M\}} d(\bar{\mu}^{N,j} - \mu^j) \right| \\
&= \left| \sum_{k=1}^{n^{2d}} \int_{B_j \times C_k} \varphi(r, z) d(\bar{\mu}^{N,j} - \mu^j) \right| \\
&\leq \sum_{k=1}^{n^{2d}} \left( \left| \int_{B_j \times C_k} \varphi(j \epsilon_N, z_k) d(\bar{\mu}^{N,j} - \mu^j) \right| + \int_{B_j \times C_k} |\varphi(r, z) - \varphi(j \epsilon_N, z_k)| d(\bar{\mu}^{N,j} + \mu^j) \right) \\
&\leq \sum_{k=1}^{n^{2d}} \left( \int_{B_j \times C_k} (1 + |z_k|) d|\bar{\mu}^{N,j} - \mu^j| + \left( \epsilon_N + \frac{M}{n} \right) (\bar{\mu}^{N,j}(B_j \times C_k) + \mu^j(B_j \times C_k)) \right) \\
&\leq (1 + M) \sum_{k=1}^{n^{2d}} \left| \bar{\mu}^{N,j}(B_j \times C_k) - \mu^j(B_j \times C_k) \right| + 2 \frac{M}{n} + 2\epsilon_N, \tag{90}
\end{aligned}$$

where we used at the third line that  $\varphi(j \epsilon_N, z_k) \leq r + |z_k|$  for the first term, and the Lipschitz property for the second term. Then we used that  $|z_k| \leq M$  at the last line. For any  $k \leq n^{2d}$  and any  $1 \leq j \leq \epsilon_N^{-1}$ , by independence of the  $(\bar{Y}^i)_{i \in NB_j}$  we get

$$\begin{aligned}
\mathbb{E} \left[ \left( \bar{\mu}^{N,j}(B_j \times C_k) - \mu^j(B_j \times C_k) \right)^2 \right] &= \frac{1}{(N\epsilon_N)^2} \sum_{i, i' \in NB_j} \mathbb{E} \left[ \mathbb{1}_{\bar{Z}^i \in C_k} \mathbb{1}_{\bar{Z}^{i'} \in C_k} \right] \\
&\quad - \frac{2}{N\epsilon_N} \sum_{i \in NB_j} \mathbb{P}(\bar{Z}^i \in C_k) \mu^j(B_j \times C_k) + \mu^j(B_j \times C_k)^2 \\
&= \frac{1}{N\epsilon_N} \mu^j(B_j \times C_k) \left( 1 - \mu^j(B_j \times C_k) \right) \\
&\leq \frac{1}{N\epsilon_N} \mu^j(B_j \times C_k). \tag{91}
\end{aligned}$$

Taking the mean over  $j$  in (90) and using (91) gives, by Cauchy-Schwarz inequality

$$\begin{aligned} \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| \leq M\}} d(\bar{\mu}^{N,j} - \mu^j) \right] \\ \leq (1 + M) \epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \sum_{k=1}^{n^{2d}} \mathbb{E} \left[ \left| \bar{\mu}^{N,j}(B_j \times C_k) - \mu^j(B_j \times C_k) \right| \right] + 2 \frac{M}{n} + 2\epsilon_N \\ \leq \frac{2M\epsilon_N}{(N\epsilon_N)^{1/2}} \sum_{j=1}^{\epsilon_N^{-1}} \sum_{k=1}^{n^{2d}} \sqrt{\mu^j(B_j \times C_k)} + 2 \frac{M}{n} + 2\epsilon_N, \end{aligned}$$

for  $M > 1$ . Applying Cauchy-Schwarz inequality to the two sums at the last line and using that for any  $1 \leq j \leq \epsilon_N^{-1}$ , we have  $\sum_{k=1}^{n^{2d}} \mu^j(B_j \times C_k) \leq 1$ , we get

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \sup_{\substack{\text{Lip}(\varphi) \leq 1 \\ \varphi(0)=0}} \int_{B_j \times E} \varphi(r, z) \mathbb{1}_{\{|z| \leq M\}} d(\bar{\mu}^{N,j} - \mu^j) \right] \leq \frac{2Mn^d}{(N\epsilon_N)^{1/2}} + 2 \frac{M}{n} + 2\epsilon_N. \quad (92)$$

Combining (89) and (92) into (88) finally gives

$$\epsilon_N \sum_{j=1}^{\epsilon_N^{-1}} \mathbb{E} \left[ \mathcal{W}_1(\bar{\mu}^{N,j}, \mu^j) \right] \leq K \left( \frac{Mn^d}{(N\epsilon_N)^{1/2}} + \frac{M}{n} + \epsilon_N + \frac{1}{M} \right),$$

which choosing  $M = (N\epsilon_N)^{\frac{1}{4(d+1)}}$  and  $n = \lfloor M^2 \rfloor$  yields the desired result.  $\square$

## Acknowledgements

The author would like to thank Thierry Bodineau for suggesting this subject, and for the useful discussions and suggestions. The author would also like to thank Joaquín Fontbona and Nicolas Fournier for their relevant remarks.

## References

- [1] G. Basile, C. Bernardin, M. Jara, T. Komorowski, and S. Olla. Thermal conductivity in harmonic lattices with random collisions. In Stefano Lepri, editor, *Thermal transport in low dimensions: from statistical physics to nanoscale heat transfer*, volume 921 of *Springer Lecture Notes in Physics*, page 215. Springer, 2016.
- [2] G. Basile, C. Bernardin, and S. Olla. Thermal conductivity for a momentum conservative model. *Communications in Mathematical Physics*, 287(1):67–98, 2009.
- [3] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., second edition, 1999.
- [4] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. Fourier’s law: A challenge to theorists. In *in Mathematical Physics 2000*, pages 128–150. College Press, 2000.
- [5] P. Carmona. Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths. *Stochastic Processes and their Applications*, 117(8):1076 – 1092, 2007.

- [6] R. Cortez and J. Fontbona. Quantitative propagation of chaos for generalized Kac particle systems. *Ann. Appl. Probab.*, 26(2):892–916, 04 2016.
- [7] L. Desvillettes, C. Graham, and S. Méléard. Probabilistic interpretation and numerical approximation of a Kac equation without cutoff. *Stochastic Processes and their Applications*, 84(1):115–135, November 1999.
- [8] R. L. Dobrushin. Vlasov equations. *Functional Analysis and Its Applications*, 13(2):115–123, 1979.
- [9] J.-P. Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Communications in Mathematical Physics*, 212(1):105–164, 2000.
- [10] N. El Karoui and J.-P. Lepeltier. Représentation des processus ponctuels multivariés à l’aide d’un processus de poisson. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 39(2):111–133, 1977.
- [11] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707, August 2015.
- [12] N. Fournier and A. Guillin. From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules. *Ann. Sci. Ec. Norm. Supér.*, 50:157–199, 2017.
- [13] N. Fournier and S. Méléard. A Markov process associated with a Boltzmann equation without cutoff and for non-Maxwell molecules. *Journal of Statistical Physics*, 104(1):359–385, 2001.
- [14] F. Golse. On the dynamics of large particle systems in the mean field limit. Lecture notes, 2013.
- [15] C. Graham. Nonlinear diffusion with jumps. *Annales de l’I.H.P. Probabilités et statistiques*, 28(3):393–402, 1992.
- [16] M. Jara, T. Komorowski, and S. Olla. Superdiffusion of energy in a chain of harmonic oscillators with noise. *Communications in Mathematical Physics*, 339:46, 2015.
- [17] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [18] J.-P. Lepeltier and B. Marchal. Problème des martingales et équations différentielles stochastiques associées à un opérateur intégro-différentiel. *Annales de l’I.H.P. Probabilités et statistiques*, 12(1):43–103, 1976.
- [19] E. Presutti. *Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics*. Theoretical and Mathematical Physics. Springer Berlin Heidelberg, 2008.
- [20] Z. Rieder, J. L. Lebowitz, and E. Lieb. Properties of a harmonic crystal in a stationary nonequilibrium state. *Journal of Mathematical Physics*, 8(5):1073–1078, 1967.
- [21] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications. Springer International Publishing, 2015.
- [22] H. Suda. A family of fractional diffusion equations derived from stochastic harmonic chains with long-range interactions. preprint, 2019.
- [23] A.-S. Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d’Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.
- [24] S. Tamaki and K. Saito. Energy current correlation in solvable long-range interacting systems. *Phys. Rev. E*, 101:042118, Apr 2020.
- [25] H. Tanaka. On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. *Proc. Intern. Symp. SDE, Kyoto*, pages 409–425, 1976.

- [26] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.