ECOLE POLYTECHNIQUE

CENTRE DE MATHÉMATIQUES APPLIQUÉES
UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 41 50. Fax: 01 69 33 30 11

http://www.cmap.polytechnique.fr/

Measurability of optimal transportation and convergence rate for Landau type interacting particle systems

Joaquin Fontbona, Hélène Guérin, Sylvie Méléard

R.I. N^0 613 *March 2007*

Measurability of optimal transportation and convergence rate for Landau type interacting particle systems

Joaquin Fontbona*, Hélène Guérin †, Sylvie Méléard‡

Abstract

In this paper, we consider nonlinear diffusion processes driven by space-time white noises, which have an interpretation in terms of partial differential equations. For a specific choice of coefficients, they correspond to the Landau equation arising in kinetic theory. A particular feature is that the diffusion matrix of this process is a linear function the law of the process, and not a quadratic one, as in the McKean-Vlasov model. The main goal of the paper is to construct an easily simulable diffusive interacting particle system, converging towards this nonlinear process and to obtain an explicit pathwise rate. This requires to find a significant coupling between finitely many Brownian motions and the infinite dimensional white noise process. The key idea will be to construct the right Brownian motions by pushing forward the white noise processes, through the Brenier map realizing the optimal transport between the law of the nonlinear process, and the empirical measure of independent copies of it. A striking problem then is to establish the joint measurability of this optimal transport map with respect to the space variable and the parameters (time and randomness) making the marginals vary. We shall prove a general measurability result for the mass transportation problem in terms of the support of the transfert plans, in the sense of set-valued mappings. This will allow us to construct the coupling and to obtain explicit convergence rates.

Key words and phrases: Landau type interacting particle systems, nonlinear white noise driven SDE, pathwise coupling, measurability of optimal transport, predictable transport process.
MSC: 60K35, 49Q20, 82C40, 82C80, 60G07.

1 Introduction and main statements

Consider the nonlinear diffusion processes in $\mathbb{R}^d$ of the following type:

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^d} \sigma(X_s - y) W_P(dy, ds) + \int_0^t \int_{\mathbb{R}^d} b(X_s - y) P_s(dy) ds \tag{1}$$

*DIM-CMM, Universidad de Chile, Casilla 170-3, Correo 3, Santiago-Chile, e-mail:fontbona@dim.uchile.cl. Supported by Fondecyt Proyect 1040689, ECOS-Conicyt C05E02, Millennium Nucleus Information and Randomness ICM P04-069-F and FONDAP Applied Mathematics

†IRMAR, Université Rennes 1, Campus de Beaulieu, 35042 Rennes-France, e-mail:helene.guerin@univ-rennes1.fr. Supported by ECOS-Conicyt C05E02 and Millennium Nucleus Information and Randomness ICM P04-069-F

‡CMAP, Ecole Polytechnique, CNRS, route de Saclay, 91128 Palaiseau Cedex-France e-mail: sylvie.meleard@polytechnique.edu. Supported by ECOS-Conicyt C05E02 and Millennium Nucleus Information and Randomness ICM P04-069-F

where P_t is the law of X_t, and W_P is a $\mathbb{R}^d$ valued space-time white noise on $[0,T] \times \mathbb{R}^d$ with independent coordinates, each of which having covariance measure $P_t(dy) \otimes dt$.

The nonlinear process (1) was introduced by Funaki [3], who obtained existence and uniqueness results for Lipschitz coefficients $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\otimes d}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$, see also Guerin [7] for a different approach. It has an important interpretation in terms of partial differential equations issued from kinetic theory. More precisely, for a specific choice of coefficients σ and b, the laws $(P_t)_t$ are a weak solution of the spatially homogeneous Landau (also called Fokker-Planck-Landau) equations for Maxwell potential:

$$\frac{\partial f}{\partial t}(t,v) = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial}{\partial v_i}\left\{\int_{\mathbb{R}^d} a_{ij}(v-v_*)\left[f(t,v_*)\frac{\partial f}{\partial v_j}(t,v) - f(t,v)\frac{\partial f}{\partial v_{*j}}(t,v_*)\right]dv_*\right\}, \tag{2}$$

with $a_{ij}(v) := (\sigma\sigma^*)_{ij}(v) = |v|^2\delta_{ij} - v_i v_j$ and $b_i(v) = \nabla \cdot a_{i\cdot}(v)$. The equations (2) model collisions of particles in a plasma and can be obtained as limit of the Boltzmann equations when collisions become grazing, see Funaki [4], Goudon [5], Villani [17] [18] and Guérin-Méléard [8].

In this work, we shall prove the convergence in law of an easily simulable mean field interacting particle system towards the nonlinear process (1) at an explicit pathwise rate. This problem is of great interest in order to construct a tractable simulation algorithm for the law P_t and thus, in particular, for solutions f of equation (2). To our knowledge, there is no result on convergence rates of the deterministic numerical methods used at present for the Landau equation, which are reviewed in [2]. The interest of our approach is that it is based on the diffusive nature of the equation, and that it addresses a large class of nonlinear processes. The fact that we want to deal with simulable systems will necessitate a coupling between finite dimensional and infinite dimensional stochastic processes. We shall introduce a coupling argument based on new results on measurability of the optimal mass transportation problem.

We consider a particle system which is naturally related to the nonlinear process. Indeed, notice that the diffusion matrix associated with (1) is defined on $\mathbb{R}^d$ by

$$a(x,P_t) := \int_{\mathbb{R}^d} \sigma(x-y)\sigma^*(x-y)P_t(dy) = [(\sigma\sigma^*) * P_t](x). \tag{3}$$

Thus, if in order to approximate the white noise driven stochastic differential equation (1), we heuristically replace P_t in (3) by an empirical measure of $n \in \mathbb{N}^*$ particles in $\mathbb{R}^d$, we are led to consider the following system driven by n^2 independent Brownian motions (B^{ik}):

$$X_t^{i,n} = X_0^i + \frac{1}{\sqrt{n}}\int_0^t \sum_{k=1}^n \sigma(X_s^{i,n} - X_s^{k,n})dB_s^{ik} + \frac{1}{n}\int_0^t \sum_{k=1}^n b(X_s^{i,n} - X_s^{k,n})ds,\ i = 1,\ldots,n. \tag{4}$$

To be more precise, if $\mu_t^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_t^{i,n}}$ is the empirical measure of the system, the mappings

$$f(t,\omega,x) \mapsto \frac{1}{\sqrt{n}}\int_0^t \sum_{k=1}^n f(s,\omega,X_s^{k,n})dB_s^{ik}, i = 1,\ldots,n, \tag{5}$$

define (for suitably measurable functions f) orthogonal martingale measures in the sense of Walsh [20], with covariance measure $\mu_t^n \otimes dt$.

By adapting techniques of Méléard-Roelly [11] based on martingale problems, one can show propagation of chaos for system (4) with as limit the process (1). This says in particular that the covariance measure of (5) converges in law to $P_t \otimes dt$ when n goes to infinity. But in turn, the arguments of [11] do not give any information about speed of convergence.
To estimate the distance between the law of the particles and the law of the nonlinear process, we need to construct a significant coupling between finitely many Brownian motions and the white noises processes. This problem is much more subtle than in the McKean-Vlasov model (cf. Sznitman [16] or Méléard [12]), where each particle is coupled with a limiting process through a single Brownian motion that drives them both. The well known $\frac{1}{\sqrt{n}}-$ convergence rate in that model is consequence of the standard L^2-law of large numbers in $\mathbb{R}^d$ and of the fact that the diffusion and drift coefficients of the nonlinear process depend linearly on the limiting law through expectations with respect to it. In the present Landau model, we have to deal with the space-time random fields (5), which have fluctuations of constant order in n. This is also reflected in the fact that it is the *squared* diffusion matrix of (1), that depends linearly on P_t (see (3)). It is hence not clear where a convergence rate can be deduced from.

Let X^i,$i = 1, \dots, n$ be n independent copies of the nonlinear process in some probability space, and ν_t^n their empirical measure at time t (observe that it samples P_t). We shall construct particles (4) on the same probability space, in such way that they will converge pathwise in L^2 on finite time intervals, at the same rate at which the Wasserstein distance W_2 between P_t and ν_t^n goes to 0. Let us state our main result on the process (1):

Theorem 1.1. *Let $n \in \mathbb{N}$ and assume usual Lipschitz hypothesis on σ and b, and that the law P_0 of X_0^i has finite second order moment. Assume moreover that P_t has a density with respect to Lebesgue measure for each $t > 0$.*
Then, in the same probability space as $(X^1, \dots, X^n)$ there exist independent standard Brownian motions $(B^{ik})_{1 \le i,k \le n}$ such that the particle system $(X^{i,n})_{i=1}^n$ defined in (4) satisfies

$$E\left(\sup_{t \in [0,T]} |X_t^{i,n} - X_t^i|^2 \right) \le C \exp(C'T) \int_0^T E(W_2^2(\nu_s^n, P_s))ds$$

for constants C, C' that do not depend on n.

Thanks to available convergence results for empirical measures of i.i.d samples (see e.g. [14]), Theorem 1.1 will allow us to obtain, under some additional moment assumptions on P_0, the speed of convergence $n^{\frac{-2}{d+4}}$ for the pathwise law of the system (see Corollary 6.2). We remark that the absolute continuity condition of Theorem 1.1 can be obtained under non-degeneracy of the matrix $\sigma\sigma^*$ by using for instance Malliavin calculus [13]; it is also true for the specific coefficients of the Landau equation (2) despite their degeneracy, and for some generalizations (see Guérin [6]).

The proof of Theorem 1.1 relies on new results on the optimal mass transportation problem. For general background on the theory of mass transportation, we refer to Villani [19]. Recall that if μ and ν are probability measures in $\mathbb{R}^d$ with finite second moment, the first of them having a density, then the optimal mass transportation problem with quadratic cost between μ and ν has a unique solution, which is a probability measure on $\mathbb{R}^{2d}$ of the form $\pi(dx, dy) = \mu(dx)\delta_{T(x)}(dy)$. The so-called Brenier or optimal transport map $T(x)$ is (μ a.s. equal to) the gradient of some convex function in $\mathbb{R}^d$, and pushes forward μ to ν.

Let now W_P^i be the white noise process driving the i-th nonlinear process X^i. The key idea in Theorem 1.1 will be to construct Brownian motions $(B^{ik})_{k=1\ldots n}$ in an "optimal" pathwise way from W_P^i. Heuristically, this will consist in pushing forward the martingale measure W_P^i through the Brenier maps $T^{t,\omega,n}(x)$ realizing the optimal transport between P_t and $\nu_t^n(\omega)$ (this is the reason for the absolute continuity assumption on P_t). But to give such a construction a rigorous sense, we must make sure that we can compute stochastic integrals of $T^{t,\omega,n}(x)$ with respect to $W_P^i(dx, dt)$. ¿From the basic definition of stochastic integration with respect to space-time white noise (cf. [20]), this requires the existence of a measurable version of $(t, \omega, x) \mapsto T^{t,\omega,n}(x)$ being moreover predictable in (t, ω). A striking problem then is that no available result in the mass transportation theory can provide any information about joint measurability properties of the optimal transport map, with respect to the space variable and some parameter making the marginals vary. Nevertheless, we will show that a suitable "predictable transportation process" exists:

Theorem 1.2. *There exists a measurable process $(t, \omega, x) \mapsto T^n(t, \omega, x)$ that is predictable in (t, ω) with respect to the filtration associated to $(W_P^1, \ldots, W_P^n)$ and $(X_0^1, \ldots, X_0^n)$, and such that for $dt \otimes \mathbb{P}(d\omega)$ almost every (t, ω),*

$$T^n(t, \omega, x) = T^{t,\omega,n}(x) \quad P_t(dx)\text{-almost surely.}$$

This statement is consequence of a general abstract result about "measurability" of the mass transportation problem. To be more explicit, recall that the optimality of a transfert plan π is determined by its support (it is equivalent to the support being cyclically monotone, see McCann [10] or Villani [19]). On the other hand, without assumptions (besides moments) on the marginals μ and ν, the solution π of the mass transportation problem may not be unique. A basic question then is how to formulate, in a general setting, the adequate property of "measurability" of the solution(s) π with respect to the data (μ, ν). As we shall see, the natural formulation requires to introduce notions and techniques from set-valued analysis. Then, we shall prove the following

Theorem 1.3. *Let $\mathcal{P}_2(\mathbb{R}^d)$ be the space of Borel probability measures in $\mathbb{R}^2$ with finite second order moment, endowed with the Wasserstein distance and its Borel $\sigma-$field. Denote by $\Pi^*(\mu, \nu)$ the set of solutions of the mass transportation problem with quadratic cost associated with $(\mu, \nu) \in (\mathcal{P}_2(\mathbb{R}^d))^2$. The function assigning to (μ, ν) the* **set** *of $\mathbb{R}^{2d}$:*

$$\bigcup_{\pi \in \Pi^*(\mu,\nu)} supp(\pi),$$

is measurable in the sense of set-valued mappings.

In particular, this ensures that if μ_λ and ν_λ vary in a measurable way with respect to some parameter λ, so that in each of the associated optimal transportation problems uniqueness holds, then the support of the solution π_λ also "varies" in a measurable way. This will be the key to our results.

The rest of this work is organized as follows. In Section 2 we review the Wasserstein distance and the mass transportation problem with quadratic cost in $\mathbb{R}^d$ (in particular the characterization of its minimizers). In Section 3 we prove Theorem 1.3 and a consequence needed to prove Theorem 1.1. In Section 4, we state some properties about process (1) and we heuristically describe our coupling between space-time white noises and Brownian motions. In Section 5 we construct the "predictable transportation process" of Theorem 1.2 needed to rigorously define the coupling. Section 6 is devoted to complete the proof of Theorem 1.1 and to obtain explicit convergence rates.

2 The mass transportation problem with quadratic cost in $\mathbb{R}^d$ and the Wasserstein distance

We denote the space of Borel probability measures in $\mathbb{R}^d$ by $\mathcal{P}(\mathbb{R}^d)$, and by $\mathcal{P}_2(\mathbb{R}^d)$ the subspace of probability measures having finite second order moment.
Given $\pi \in \mathcal{P}_2(\mathbb{R}^{2d})$, we respectively denote by π_1 and π_2 its first and second marginals on $\mathbb{R}^d$. On the other hand, for any two probability measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\pi \in \mathcal{P}_2(\mathbb{R}^{2d})$, we write

$$\pi <^{\mu}_{\nu}$$

if $\pi_1 = \mu$ and $\pi_2 = \nu$. Such π is refereed to as a "transfert plan" between μ and ν.

Definition 2.1. *The Wassertein distance W_2 on $\mathcal{P}_2(\mathbb{R}^d)$ is defined by*

$$W_2^2(\mu,\nu) := \inf_{\pi <^{\mu}_{\nu}} \int_{\mathbb{R}^2} |x-y|^2 \pi(dx,dy).$$

Then, $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a Polish space, see e.g. Rachev and Rüschendorf [14]. The topology is stronger that the usual weak topology. More precisely, one has the following result (see for instance Villani, [19] Theorem 7.12)

Theorem 2.2. *Let $\mu^n, \mu \in \mathcal{P}(\mathbb{R}^d)$. The following are then equivalent:*

i) $W_2(\mu^n, \mu) \to 0$ *when* $n \to \infty$.

ii) μ^n *converges weakly to* μ *and*

$$\int_{\mathbb{R}^d} |x|^2 \mu^n(dx) \to \int_{\mathbb{R}^d} |x|^2 \mu(dx).$$

iii) We have

$$\int_{\mathbb{R}^d} \varphi(x)\mu^n(dx) \to \int_{\mathbb{R}^d} \varphi(x)\mu(dx)$$

for all continuous function $\varphi : \mathbb{R}^d \to \mathbb{R}$ *such that* $|\varphi(x)| \leq C(1+|x|^2)$ *for some* $C \in \mathbb{R}$.

We shall denote by L the mapping $L : \mathcal{P}_2(\mathbb{R}^{2d}) \to \mathbb{R}$ defined by

$$L(\pi) = \int_{\mathbb{R}^2} |x-y|^2 \pi(dx,dy).$$

Remark 2.3. *It is not hard to check that L is lower semi continuous (l.s.c) for the weak topology. Moreover, L is continuous for the Wasserstein topology in $\mathcal{P}_2(\mathbb{R}^{2d})$ by part iii) of Theorem 2.2.*

Fix now $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, and denote by $\Pi^*(\mu,\nu)$ the subset of $\mathcal{P}_2(\mathbb{R}^{2d})$ of minimizers of the *Monge-Kantorovich transportation problem* with quadratic cost for the pair of marginals (μ,ν) . That is,

$$\Pi^*(\mu,\nu) := argmin_{\pi <^{\mu}_{\nu}} L(\pi).$$

It is well known that $\Pi^*(\mu,\nu)$ is non-empty. Indeed, it is not hard to see that for the weak topology, $\{\pi \in \mathcal{P}_2(\mathbb{R}^{2d}) : \pi <^{\mu}_{\nu}\}$ is a compact set, and the lower semi-continuity of L implies the existence of minimizers (see e.g. [19] Chapter 1 for details).

We shall next recall the characterization of minimizers of the transportation problem with quadratic cost. We need the notion of sub-differential of a convex function:

Definition 2.4. *Let $\varphi : A \subset \mathbb{R}^d \to]-\infty, \infty]$ be a proper (i.e. $\varphi \not\equiv +\infty$) lower semi-continuous (l.s.c) convex function. The sub-differential of φ at x is*

$$\partial\varphi(x) = \{y \in \mathbb{R}^d : \varphi(z) \geq \varphi(x) + \langle y, z - x\rangle, \forall z \in \mathbb{R}^d\}.$$

Elements of $\partial\varphi(x)$ are called sub-gradients of φ at point x. The graph of $\partial\varphi$ is

$$Gr(\partial\varphi) = \{(x, y) \in \mathbb{R}^{2d} : y \in \partial\varphi(x)\}$$

and it is a closed set.

Recall that φ is differentiable at x if and only if $\partial\varphi(x)$ is a singleton (in which case $\partial\varphi(x) = \{\nabla\varphi(x)\}$). Also, the set $\{x \in \mathbb{R}^d : \varphi \text{ is differentiable at } x\}$ is borelian, see e.g. McCann [10].

We next summarize results in pioneer works in this domain, Knott-Smith [9], Brenier [1] and McCann [10], Rachev and Rüschendorf [14]. See also Villani [19] for a complete discussion on these questions, proofs and background.

Theorem 2.5. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ and $\pi <^{\mu}_{\nu}$ be a transfert plan. We have*

a) $\pi \in \Pi^*(\mu, \nu)$ if and only if there exists a proper l.s.c. convex function φ such that

$$supp(\pi) \subset Gr(\partial\varphi)$$

or, equivalently

$$\pi(\{(x, y) \in \mathbb{R}^2 : y \in \partial\varphi(x)\}) = 1.$$

b) Assume that μ does not charge sets of Hausdorff dimension less or equal than $d - 1$ and that $\pi \in \Pi^*(\mu, \nu)$. Then,

i) the set $\{x \in \mathbb{R}^d : \varphi \text{ is not differentiable at } x\}$ has null μ-measure.

ii) We have

$$\pi(dx, dy) = \mu(dx) \otimes \delta_{\nabla\varphi(x)}(dy).$$

ii) If T is a measurable mapping such that $\pi(dx, dy) = \mu(dx) \otimes \delta_{T(x)}(dy)$, then $T(x) = \nabla\varphi(x)$, $\mu(dx) - a.s.$.

iii) $\pi \in \Pi^*(\mu, \nu)$ is unique.

This result will be useful later in the particular case when the measure μ is absolutely continuous with respect to Lebesgue measure.

3 Measurability of the mass transportation problem

We now introduce the basic notions on "multi-applications" or "set-valued mappings" that we need to prove Theorem 1.3. For general background, we refer the reader to Appendix A in Rockafellar and Wets [15].

Definition 3.1. *Let X, Y be two sets.*

i) A function S on X taking values in the set of subsets of Y is called a set-valued mapping or multi-application. We write $S : X \rightrightarrows Y$.

ii) For any $A \subset Y$, the inverse image of A through S is the set

$$S^{-1}(A) := \{x \in X : S(x) \cap A \neq \emptyset\}.$$

iii) If $(X, \mathcal{A})$ is a measurable space and (Y, Θ) a topological space, we say that $S : X \rightrightarrows Y$ is measurable if for all $\theta \in \Theta$,

$$S^{-1}(\theta) \in \mathcal{A}.$$

(Of course, if $S(x) = \{s(x)\}$ is singleton for all x, measurability of S is equivalent to that of s.)

Consider $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the Wasserstein distance and the Borel $\sigma-$field. We define a set-valued mapping

$$\Psi : (\mathcal{P}_2(\mathbb{R}^d))^2 \rightrightarrows \mathbb{R}^{2d}$$

by

$$\Psi(\mu, \nu) := \{(x, y) : \exists \pi \in \Pi^*(\mu, \nu) \ s.t. \ (x, y) \in supp(\pi)\}.$$

Our goal is to prove that Ψ is measurable. We shall need some further notions on set-valued mappings.

Definition 3.2. *Let X be a set, and (Y, Ξ) and (Z, Θ) be topological spaces.*

i) A set-valued mapping $S : X \rightrightarrows Y$ is closed-valued if for all $x \in X$, $S(x)$ is a closed set of (Y, Ξ).

ii) A set-valued mapping $U : Y \rightrightarrows Z$ is inner semicontinous (i.s.c) if for all $\theta \in \Theta$,

$$S^{-1}(\theta) \in \Xi$$

The following results can be found in Appendix A of [15], in the case of set-valued mappings in $\mathbb{R}^d$. For completeness we provide proofs in a more general context.

Lemma 3.3. *Let $(X, \mathcal{A})$ be a measurable space and (Y, Ξ) a topological space.*

i) $S : X \rightrightarrows Y$ is measurable if and only if the closed-valued mapping $x \rightrightarrows Cl(S(x))$ is measurable, where $Cl(S(x))$ is the topological closure of the set $S(x)$.

ii) Assume that (Y, d) is a separable metric space and that $S : X \rightrightarrows Y$ is closed-valued. Then, S is measurable if and only if for all closed set F of Y,

$$S^{-1}(F) \in \mathcal{A}.$$

iii) Let (Y, Ξ) and (Z, Θ) be topological spaces, $S : X \rightrightarrows Y$ be measurable and $U : Y \rightrightarrows Z$ be i.s.c. Then, the multi-application $U \circ S : X \rightrightarrows Z$, defined by

$$U \circ S(x) := \bigcup_{y \in S(x)} U(y)$$

is measurable.

Proof *i)* For any open set $\theta \in \Xi$, $S(x) \cap \theta \neq \emptyset$ if and only if $Cl(S(x)) \cap \theta \neq \emptyset$.
ii) "Only if" part: since Y is a metric space, we use that every closed set F is the intersection of some countable collection of open sets (θ_n). Therefore,

$$\{x \in X : S(x) \cap F \neq \emptyset\} = \bigcap_{n \in \mathbb{N}} \{x \in X : S(x) \cap \theta_n \neq \emptyset\} \in \mathcal{A}.$$

"If" part: (Y, d) being separable, we can express every open set θ as the union of some countable collection (B_n) of closed balls. We then have that

$$\{x \in X : S(x) \cap \theta \neq \emptyset\} = \bigcup_{n \in \mathbb{N}} \{x \in X : S(x) \cap B_n \neq \emptyset\} \in \mathcal{A}.$$

iii) Straightforward:

$$\begin{aligned}(U \circ S)^{-1}(\theta) &= \{x \in X : \left(\cup_{y \in S(x)} U(y)\right) \cap \theta \neq \emptyset\} = \{x \in X : \ \exists y \in S(x) \ s.t. \ U(y) \cap \theta \neq \emptyset\} \\ &= \{x \in X : S(x) \cap (U^{-1}(\theta)) \neq \emptyset\}.\end{aligned}$$

The function U being i.s.c., $U^{-1}(\theta)$ belongs to Ξ, which allows us to conclude.

□

Now we can proceed to the

Proof of Theorem 1.3

We observe first that $\Psi(\mu, \nu) = U \circ S(\mu, \nu)$, where S and U are the set valued mappings respectively defined by

$$(\mu, \nu) \rightrightarrows S(\mu, \nu) := \Pi^*(\mu, \nu)$$

and $U : \mathcal{P}_2(\mathbb{R}^{2d}) \rightrightarrows \mathbb{R}^d$ by

$$U(\pi) := supp(\pi)$$

We will therefore split the proof in several parts:
a) S is a closed valued mapping
First notice that $\pi \mapsto \pi_i$ is continuous for the Wasserstein topology. Indeed, $W_2(\pi^n, \pi) \to 0$ implies that π^n converges weakly to π, and then π_i^n converges weakly to π_i for $i = 1, 2$. Moreover, we have $\int_{\mathbb{R}^d} |x|^2 \pi_1^n(dx) = \int_{\mathbb{R}^{2d}} |x|^2 \pi^n(dx, dy) \to \int_{\mathbb{R}^{2d}} |x|^2 \pi(dx, dy) = \int_{\mathbb{R}^d} |x|^2 \pi_1(dx)$ by Theorem 2.2, and then the asserted continuity follows.
Consequently, $\pi \mapsto W_2(\pi_1, \pi_2)$ too is continuous. Therefore,

$$\Pi^*(\mu, \nu) = \{\pi : \pi <_{\nu}^{\mu}\} \cap \{\pi : L(\pi) - W_2(\pi_1, \pi_2) = 0\}$$

is the intersection of two closed sets $\mathcal{P}_2(\mathbb{R}^{2d})$.

b) Inverse images through S of closed sets are closed sets
Let $F \subset \mathcal{P}_2(\mathbb{R}^d)$ be a closed set and $(\mu^n, \nu^n) \in S^{-1}(F)$, $n \in \mathbb{N}$, be a sequence converging to (μ, ν) in $(\mathcal{P}_2(\mathbb{R}^d))^2$. Then, $\mu^n \to \mu$ and $\nu^n \to \nu$ weakly, and (μ^n) and (ν^n) are tight.
But since $(\mu^n, \nu^n) \in S^{-1}(F)$ for each n, there exists π_n s.t. $\pi^n <_{\nu^n}^{\mu^n}$, and then (π_n) too is tight (by considering products of compact sets).
Let (π^{n_k}) be a weakly convergent subsequence with limit π. Then, clearly $\pi <_{\nu}^{\mu}$. We will prove that $L(\pi) = W_2(\mu, \nu)$ and that $\pi \in F$, which will mean that $(\mu, \nu) \in S^{-1}(F)$ and finish the proof.

We have

$$\int_{\mathbb{R}^{2d}} \left(|x|^2 + |y|^2\right) \pi^{n_k}(dx, dy) = \int_{\mathbb{R}^d} |x|^2 \mu^{n_k}(dx) + \int_{\mathbb{R}^d} |y|^2 \nu^{n_k}(dy) \to$$
$$\int_{\mathbb{R}^d} |x|^2 \mu(dx) + \int_{\mathbb{R}^d} |y|^2 \nu(dy) = \int_{\mathbb{R}^{2d}} \left(|x|^2 + |y|^2\right) \pi(dx, dy),$$

which implies that $W_2(\pi^n, \pi) \to 0$ and $\pi \in F$. Finally, by the continuity of $\pi \mapsto L(\pi) - W_2(\pi_1, \pi_2)$ we get that

$$0 = L(\pi^{n_k}) - W_2(\pi_1^{n_k}, \pi_2^{n_k}) = L(\pi) - W_2(\mu, \nu).$$

c) The mapping U is i.s.c.
Let θ be an open set of $\mathbb{R}^{2d}$. We must check that

$$\{\pi \in \mathcal{P}_2(\mathbb{R}^{2d}) : supp(\pi) \cap \theta \neq \emptyset\} = \{\pi \in \mathcal{P}_2(\mathbb{R}^{2d}) : \pi(\theta) > 0\}$$

is open, or equivalently, that

$$\{\pi \in \mathcal{P}_2(\mathbb{R}^{2d}) : \pi(\theta) = 0\}$$

is closed in $\mathcal{P}_2(\mathbb{R}^{2d})$. Assume that $\pi, \pi^n \in \mathcal{P}_2(\mathbb{R}^{2d})$, with π^n such that $\pi^n(\theta) = 0$ for all $n \in \mathbb{N}$, and moreover that $W_2(\pi^n, \pi) \to 0$. Then π^n converges weakly to π, and so by the Portemanteau theorem, we have

$$0 = \liminf_n \pi^n(\theta) \geq \pi(\theta).$$

d) Conclusion
By parts *a)* and *b)* and Lemma 3.3 *ii)* we get that S is measurable. By *c)* and Lemma 3.3 *iii)* $U \circ S$ is measurable and the proof is finished.

□

The following corollary will be useful in the specific setting needed to prove Theorem 1.1:

Corollary 3.4. *Let (E, Σ) be a measurable space, and $\lambda \in E \mapsto (\mu_\lambda, \nu_\lambda) \in (\mathcal{P}_2(\mathbb{R}^d))^2$ and $\xi : E \to \mathbb{R}^d$ be measurable functions. Then, the set*

$$\{(\lambda, x) : (x, \xi(\lambda)) \in Cl(\Psi)(\mu_\lambda, \nu_\lambda)\}$$

belongs to $\Sigma \otimes \mathcal{B}(\mathbb{R}^d)$

Proof By Lemma 3.3 *i)* and Theorem 1.3 we get that $Cl(\Psi)$ is measurable. Moreover, it is not hard to check that the mapping

$$(\lambda, x) \rightrightarrows Cl(\Psi)(\mu_\lambda, \nu_\lambda) - (x, \xi(\lambda))$$

is measurable and closed-valued. Then, we just have to notice that

$$(x, \xi(\lambda)) \in Cl(\Psi)(\mu_\lambda, \nu_\lambda) \text{ if and only if } [Cl(\Psi)(\mu_\lambda, \nu_\lambda) - (x, \xi(\lambda))] \cap C \neq \emptyset$$

for the closed set $C = \{0\}$.

□

4 A coupling between space-time white noise and Brownian motions via optimal transport

In all the sequel, we refer the reader to Walsh [20] for background on space-time white noise processes and stochastic integration with respect to martingale measures.
Assume that $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\otimes d}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz continuous and with linear growth. Then, by results of [3] or [7] we can construct in some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a sequence $(X^i)_{i\in\mathbb{N}}$ of independent copies of the nonlinear processes,

$$X_t^i = X_0^i + \int_0^t \int_{\mathbb{R}^d} \sigma(X_s^i - y) W_P^i(dy, ds) + \int_0^t \int_{\mathbb{R}^d} b(X_s^i - y) P_s(dy) ds, \tag{6}$$

where the W_P^i are independent space-time $\mathbb{R}^d$-valued white noises defined on $[0,\infty) \times \mathbb{R}^d$. Each of the d (independent) coordinates of W_P^i has covariance measure $P_t(dy) \otimes dt$, where P_t is the law of X_t. The initial conditions $(X_0^1, \dots, X_0^n, \dots)$ are independent and identically distributed with law P_0, and independent of the white noises. The pathwise law of X^i is denoted by P, and it is uniquely determined.
Denote by $\mathcal{F}_t^n$ the complete right continuous σ-field generated by

$$\{(W_P^1([0,s] \times A^1), \dots, W_P^n([0,s] \times A^n)) : 0 \le s \le t, A^i \in \mathcal{B}(\mathbb{R}^d)\}$$

and $(X_0^1, \dots, X_0^n)$. We also denote by

$$\mathcal{P}red^n$$

the predictable field generated by continuous $(\mathcal{F}_t^n)$-adapted processes.

In what follows, we fix a finite time horizon $T > 0$. Under usual Lipschitz assumptions on the coefficients, there is propagation of the moments of the law P_0, as proved in Guérin [7].

Lemma 4.1. *If $E(|X_0|^k) < \infty$ for some $k \ge 2$, then*

$$E\left(\sup_{t\in[0,T]} |X_t|^k\right) < \infty.$$

The continuity of X and the previous uniform bound imply that $t \mapsto \int_{\mathbb{R}^d} |x|^k P_t(dx)$ is continuous.

Throughout the sequel, the assumptions of Theorem 1.1 on P_0 and P_t are enforced, in particular, the condition $E(\sup_{t\in[0,T]} |X_t|^2) < \infty$ will hold by the previous lemma.

We shall now present the main idea of the coupling we introduce to prove Theorem 1.1. Basically, this consists in constructing for each n, n^2 Brownian motions in a pathwise way, from the realizations of the n white noises $(W_P^1, \dots, W_P^n)$. The key for that will be to use the optimal transport maps between the marginal P_t of the nonlinear process and the empirical measures of samples of that law. More precisely, write

$$\nu_t^n := \frac{1}{n} \sum_{i=1}^n \delta_{X_t^i}$$

and notice that for each $\omega \in \Omega$, $(\nu_t^n, 0 \le t \le T)$ is an element of $C([0,T], \mathcal{P}_2(\mathbb{R}^d))$. Thus, for each $t \in [0,T]$, $n \in \mathbb{N}$ and ω, and we can consider the optimal coupling problem with quadratic cost between $\nu_t^n(\omega)$ and P_t,

$$\inf_{\pi <^{P_t}_{\nu_t^n(\omega)}} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \pi(dx,dy) \right\}.$$

By the assumption on P_t and Theorem 2.5, the following properties hold **for each fixed pair** $(t,\omega) \in]0,T] \times \Omega$:

Lemma 4.2. *a) There exists a unique $\pi^{t,\omega,n}$, such that*

$$W_2^2(P_t, \nu_t^n(\omega)) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \pi^{t,\omega,n}(dxdy).$$

b) There is a $P_t(dx) - a.e.$ unique measurable function $T^{t,\omega,n} : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\pi^{t,\omega,n}(dx,dy) = \delta_{T^{t,\omega,n}(x)}(dy) P_t(dx).$$

In particular, under $P_t(dx)$ the law of $T^{t,\omega,n}(x)$ is $\nu_t^n(\omega)$.

c) We have

$$W_2^2(P_t, \nu_t^n(\omega)) = \int_{\mathbb{R}^2} |x - T^{t,\omega,n}(x)|^2 P_t(dx).$$

We would like to construct n^2 independent Brownian motions by "transporting" the n independent white noises $(W_P^1, \ldots, W_P^n)$ through the transport mappings $T^{s,\omega,n}(x)$. As pointed out in the introduction, to do so we must at least be able to define stochastic integrals of functions of the form $(t,\omega,x) \mapsto f(T^{t,\omega,n}(x))$, with respect to the white noise processes. The existence of a version $T^n(t,\omega,x)$ of $T^{t,\omega,n}(x)$ having good enough properties, will be established in next section, when we shall prove Theorem 1.2.

Before doing so, we observe that if Theorem 1.2 holds, then the following processes $B_t^{ik} = B_t^{ik,n}$ will be well defined from (6).

Proposition 4.3. *For each $n \in \mathbb{N}^*$, define*

$$B_t^{ik,n}(\omega) := \sqrt{n} \int_0^t \int_{\mathbb{R}^d} \mathbf{1}_{\{T^n(s,\omega,x) = X_s^k(\omega)\}} W_P^i(dx,ds), \quad i,k = 1 \ldots n \tag{7}$$

Then, $(B^{ik,n})_{1 \le i,k \le n}$ are n^2 independent standard Brownian motions in $\mathbb{R}^d$.

These are the right Brownian motions we need to construct (4). The proof of Proposition 4.3 will given in Section 6.

5 Construction of the predictable "transport process"

Our goal now in this section is to show that for each $n \in \mathbb{N}^*$, there exists a process $(t,\omega,x) \mapsto T^n(t,\omega,x)$ defined $\mathbb{P}(d\omega) \otimes dt \otimes P_t(dx)$-almost everywhere, which is measurable with respect to $\mathcal{P}red^n \otimes \mathcal{B}(\mathbb{R}^d)$, and such that:

for $dt \otimes \mathbb{P}(d\omega)$ almost every (t,ω),

$$T^n(t,\omega,x) = T^{t,\omega,n}(x) \quad P_t(dx)\text{-almost surely .}$$

Since $(X^1,\dots,X^n)$ are independent copies of the nonlinear process and $P_t = law(X^i_t)$ has a density with respect to Lebesgue measure, for each $t \in [0,T]$ we have that

$$\mathbb{P}(\exists i \neq j : X^i_t = X^j_t) = 0.$$

Notice also that for fixed (i,j) with $i \neq j$ the following set

$$\{(t,\omega) : X^i_t(\omega) = X^j_t(\omega)\}$$

belongs to $\mathcal{P}red^n$ since $(t,\omega) \mapsto |X^i_t(\omega) - X^j_t(\omega)|$ is adapted and continuous in t.
By Fubini's theorem we then see that

$$\int_{[0,T]\times\Omega} \mathbf{1}_{\{X^i_t = X^j_t\}}(t,\omega)\mathbb{P}(d\omega) \otimes dt = 0$$

Remark 5.1. *Consequently, there is a predictable set of* $[0,T]\times\Omega$*,*

$$\Omega'_T \in \mathcal{P}red^n$$

of full $\mathbb{P}(d\omega)\otimes dt$*-measure and such that*

$$\textit{for all } (t,\omega) \in \Omega'_T, \quad X^i_t(\omega) \neq X^j_t(\omega) \textit{ for all } i,j \in \{1,\dots,n\}.$$

Let us denote by $(\mathcal{P}red^n)'$ the $\sigma-$field $\mathcal{P}red^n$ restricted to Ω'_T.
Recall that for each (t,ω), the set of solutions $\Pi^*\left(P_t, \frac{1}{n}\sum_{i=1}^n \delta_{X^i_t}(\omega)\right)$ of the optimal transport problem between P_t and $\frac{1}{n}\sum_{i=1}^n \delta_{X^i_t}(\omega)$ is a singleton that we have denoted by $\pi^{t,\omega,n}$.
Let us define now the sets

$$A^{i,n} := \left\{(t,\omega,x) \in \Omega'_T \times \mathbb{R}^d : (x, X^i_t(\omega)) \in supp(\pi^{t,\omega,n})\right\}, \quad i = 1,\dots,n.$$

The sets $A^{i,n}$ are predictable, as proved in the following lemma.

Lemma 5.2. *We have* $A^{i,n} \in (\mathcal{P}red^n)' \otimes \mathcal{B}(\mathbb{R}^d)$*.*

Proof Observe that the deterministic process $(t,\omega) \mapsto P_t \in \mathcal{P}_2(\mathbb{R}^d)$ is $\mathcal{P}red^n$-measurable. Indeed, if $(f_n)_{n\in\mathbb{N}\setminus\{0,1\}}$ is a countable dense subset of the space of continuous functions in $\mathbb{R}^d$ with compact support, and $f_0(x) = 1, f_1(x) = |x|^2$, then the topology of $\mathcal{P}_2(\mathbb{R}^d)$ is generated by the real mappings $m \mapsto \int f_n(x)m(dx)$. It is therefore enough that $(t,\omega) \to \int f_n(x)P_t(dx)$ be $\mathcal{P}red^n$-measurable, which is clear since $t \mapsto P_t$ is continuous.
Next we will apply Corollary 3.4 to the measurable space

$$(E,\Sigma) = (\Omega'_T, (\mathcal{P}red^n)'),$$

$\lambda = (t,\omega)$, and the $(\mathcal{P}red^n)'$-measurable functions given by

$$(t,\omega) \to \left(P_t, \frac{1}{n}\sum_{j=1}^n \delta_{X^j_t}(\omega)\right) \in \mathcal{P}_2(\mathbb{R}^{2d}) \ \text{ and } \ (t,\omega) \mapsto \xi^i(t,\omega) = X^i_t(\omega) \in \mathbb{R}^d.$$

For each $(t,\omega) \in \Omega'_T$, with Ψ denoting the multi-application defined in Theorem 1.3, we simply have in the current setting that

$$\Psi\left(P_t, \frac{1}{n}\sum_{j=1}^{n}\delta_{X_t^j}(\omega)\right) = supp(\pi^{t,\omega,n}).$$

Corollary 3.4 implies the result.

□

Recall from basic measure theory that if E_1 and E_2 are measurable spaces and $A \subseteq E_1 \times E_2$ is an element of their product $\sigma-$field, then, for each $\lambda_1 \in E_1$, the *fiber* of A at λ_1 is the set

$$[A]_{\lambda_1} := \{\lambda \in E_2 : (\lambda_1, \lambda) \in A\},$$

and it is always measurable in E_2.
We can now proceed to the

Proof of Theorem 1.2:

We split the proof in several parts.

a) The sets $A^{i,n}, i = 1 \dots n$ form a partition of $\Omega'_T \times \mathbb{R}^d$ up to $\mathbb{P}(d\omega) \otimes dt \otimes P_t(dx)$-null sets.
For $i \neq j$ write

$$\begin{aligned} A^{ij,n} &:= \{(t,\omega,x) \in (\Omega'_T \times \mathbb{R}^d) : (x, X_t^i(\omega)) \in supp(\pi^{t,\omega,n}) \text{ and } (x, X_t^j(\omega)) \in supp(\pi^{t,\omega,n})\} \\ &= A^{i,n} \cap A^{j,n}, \end{aligned}$$

and denote by $[A^{ij,n}]_{(t,\omega)} := \{x \in \mathbb{R}^d : (t,\omega,x) \in A^{ij,n}\} \in \mathcal{B}(\mathbb{R}^d)$ the fiber of $A^{ij,n}$ at $(t,\omega) \in \Omega'_T$. Then, we have

$$\begin{aligned} P_t([A^{ij,n}]_{(t,\omega)}) &= P_t(\{x \in \mathbb{R}^d : (x, X_t^i(\omega)), (x, X_t^j(\omega)) \in supp(\pi^{t,\omega,n})\}) \\ &\leq P_t(\{x \in \mathbb{R}^d : X_t^i(\omega), X_t^j(\omega) \in \partial\varphi^{t,\omega,n}(x)\}), \end{aligned}$$

where $\varphi^{t,\omega,n}$ is a proper l.s.c. convex function given by Theorem 2.5 *a)*. But since $(t,\omega) \in \Omega'_T$, we have $X_t^i(\omega) \neq X_t^j(\omega)$, and so

$$X_t^i(\omega), X_t^j(\omega) \in \partial\varphi^{t,\omega,n}(x) \Longrightarrow \varphi \text{ is not differentiable in } x.$$

We obtain by Theorem 2.5 *b)* that $P_t([A^{ij,n}]_{(t,\omega)}) = 0$, and then

$$E\left(\int_{[0,T]\times\Omega\times\mathbb{R}^d} \mathbf{1}_{A^{i,n}\cap A^{j,n}}(t,\omega,x) P_t(dx)dt\right) = 0.$$

On the other hand, since $T^{t,\omega,n}(x) \in \{X_t^1(\omega), \dots, X_t^n(\omega)\}$ $P_t(dx)$ a.s., we have for all (t,ω) that

$$\begin{aligned} P_t\left(\left[\left(\bigcup_{i=1}^{n} A^{i,n}\right)^c\right]_{(t,\omega)}\right) &= P_t(\{x \in \mathbb{R}^d : \text{ for all } i = 1,\dots,n, (x, X_t^i(\omega)) \notin supp(\pi^{t,\omega,n})\}) \\ &\leq P_t(\{x \in \mathbb{R}^d : (x, T^{t,\omega,n}(x)) \notin supp(\pi^{t,\omega,n})\}) \\ &= \pi^{t,\omega,n}(supp(\pi^{t,\omega,n})^c) \\ &= 0 \end{aligned}$$

Defining the set

$$\tilde{\Omega}_T := (\Omega'_T \times \mathbb{R}^d) \bigcap \left(\bigcup_{i=1}^{n} A^{i,n} \backslash \left(\bigcup_{k \neq j} A^{kj,n} \right) \right) \in \mathcal{P}red^n \otimes \mathcal{B}(\mathbb{R}^d)$$

we deduce that $P_t\left([\tilde{\Omega}^c_T]_{(t,\omega)}\right) = 0$ for all $(t,\omega) \in \Omega'_T$. Therefore,

$$\mathbb{E}\left(\int_0^T \int_{\mathbb{R}^d} \mathbf{1}_{\tilde{\Omega}^c_T}(t,\omega,x) P_t(dx) dt\right) = \int_{\Omega'_T} P_t\left([\tilde{\Omega}^c_T]_{(t,\omega)}\right) \, dt \otimes \mathbb{P}(d\omega) = 0,$$

and so $\tilde{\Omega}_T$ has full $\mathbb{P}(d\omega) \otimes dt \otimes P_t(dx)-$measure. This proves assertion *a)*.
We can now define a $\mathcal{P}red^n \otimes \mathcal{B}(\mathbb{R}^d)-$measurable function by

$$T^n(t,\omega,x) := \sum_{i=1}^{n} \mathbf{1}_{A^{i,n} \cap \tilde{\Omega}_T}(t,\omega,x) X^i_t(\omega). \tag{8}$$

b) For $\mathbb{P}(d\omega) \otimes dt$ almost every (t,ω), $T^n(t,\omega,x) = T^{t,\omega,n}(x)$ holds $P_t(dx)$ almost surely.
By Theorem 2.5, *b)*, this is equivalent to prove that

$$\pi^{t,\omega,n}(dx,dy) = P_t(dx) \otimes \delta_{T^n(t,\omega,x)}(dy) \quad \mathbb{P}(d\omega) \otimes dt - a.e.$$

We fix now $(t,\omega) \in \Omega'_T$ and $C, D \in \mathcal{B}(\mathbb{R}^d)$.
We have by definition of $T^{t,\omega,n}$ that

$$\begin{aligned} \pi^{t,\omega,n}(C \times D) &= \int_{\mathbb{R}^d} \mathbf{1}_C(x) \mathbf{1}_D(T^{t,\omega,n}(x)) P_t(dx) \\ &= \int_{\mathbb{R}^d} \mathbf{1}_{C \cap [\tilde{\Omega}_T]_{(t,\omega)}}(x) \mathbf{1}_D(T^{t,\omega,n}(x)) P_t(dx), \end{aligned}$$

the latter because $P_t\left([\tilde{\Omega}^c]_{(t,\omega)}\right) = 0$. Notice that on the other hand, by definition of $A^{i,n}, \tilde{\Omega}_T$ and T^n, for all $(t,\omega,x) \in A^{i,n} \cap \tilde{\Omega}_T$ we have that

$$\{y : (x,y) \in supp(\pi^{t,\omega,n})\} = \{X^i_t(\omega)\} = \{T^n(t,\omega,x)\}.$$

This implies that $\tilde{\Omega}_T \subset \{(t,\omega,x) \in \Omega'_T \times \mathbb{R}^d : \{y : (x,y) \in supp(\pi^{t,\omega,n})\}$ is a singleton$\}$.
Now, let $F^{t,\omega} \in \mathcal{B}(\mathbb{R}^d)$ be a measurable set with $P_t(F^{t,\omega}) = 1$ and such that $T^{t,\omega,n}(x) = \nabla\varphi^{t,\omega,n}(x)$ is defined for all $x \in F^{t,\omega}$. Then, on $F^{t,\omega} \cap [\tilde{\Omega}_T]_{(t,\omega)}$ it must hold that

$$T^n(t,\omega,x) = T^{t,\omega,n}(x) = \nabla\varphi^{t,\omega,n}(x),$$

and we conclude that for all $(t,\omega) \in \Omega'_T$,

$$\begin{aligned} \pi^{t,\omega,n}(C \times D) &= \int_{\mathbb{R}^d} \mathbf{1}_{C \cap F^{t,\omega} \cap [\tilde{\Omega}_T]_{(t,\omega)}}(x) \mathbf{1}_D(T^{t,\omega,n}(x)) P_t(dx), \\ &= \int_{\mathbb{R}^d} \mathbf{1}_{C \cap F^{t,\omega} \cap [\tilde{\Omega}_T]_{(t,\omega)}}(x) \mathbf{1}_D(T^n(t,\omega,x)) P_t(dx) \\ &= \int_{\mathbb{R}^d} \mathbf{1}_C(x) \mathbf{1}_D(T^n(t,\omega,x)) P_t(dx) \end{aligned}$$

□

We point out that Theorem 1.2 implies

Corollary 5.3. *$T^n(t,\omega,x) = T^{t,\omega,n}(x)$ holds $\mathbb{P}(d\omega) \otimes dt \otimes P_t(dx)$-almost surely . Consequently, $T^{t,\omega,n}(x)$ is measurable with respect to the completed $\sigma-$field of $\mathcal{P}red^n \otimes \mathcal{B}(\mathbb{R}^d)$ with respect to $\mathbb{P}(d\omega) \otimes dt \otimes P_t(dx)$.*

6 Pathwise convergence and rates for stochastic particle systems to Landau process

Proof of Proposition 4.3
¿From the proof of Theorem 1.2, it is clear that integrals with respect to the measures $\mathbf{1}_{A^{k,n}\cap\tilde{\Omega}_T}P_t(dx)\otimes dt$ and $\mathbf{1}_{A^{k,n}}P_t(dx)\otimes dt$ are indistinguishable. By considering quadratic variations, the same is seen to hold for the stochastic integrals with respect to $\mathbf{1}_{A^{k,n}\cap\tilde{\Omega}_T}W^i_P(dx,dt)$ and $\mathbf{1}_{A^{k,n}}W^i_P(dx,dt)$. Write

$$B^{ik,n,m}_t$$

for the $m-$th coordinate of the process $B^{ik,n}_t$ in (7), which is a real valued continuous local martingale with respect to $\mathcal{F}^n_t$ (see [20]). Then, we have that

$$\begin{aligned}\langle B^{ik,n,m},B^{i'k',n,m'}\rangle_t(\omega)=&n\delta_{(i,m),(i',m')}\int_0^t\int_{\mathbb{R}^d}\mathbf{1}_{A^{k,n}\cap A^{k',n}\cap\tilde{\Omega}_T}(s,\omega,x)P_s(dx)ds\\ =&\,n\delta_{(i,k,m),(i',k',m')}\int_0^t\int_{\mathbb{R}^d}\mathbf{1}_{A^{k,n}\cap\tilde{\Omega}_T}(s,\omega,x)P_s(dx)ds,\end{aligned}$$

by step (a) in the proof of Theorem 1.2. Now, for $(s,\omega)\in\Omega'_T$ the points $X^1_s(\omega),\dots,X^n_s(\omega)$ are all different, and consequently we have that

$$\begin{aligned}\int_{\mathbb{R}^d}\mathbf{1}_{A^{k,n}\cap\tilde{\Omega}_T}(s,\omega,x)P_s(dx)=&P_s(\{x:T^n(s,\omega,x)=X^k_s(\omega)\})\\ =&P_s(\{x:T^{s,\omega,n}(x)=X^k_s(\omega)\})\\ =&\pi^{s,\omega,n}(\{(x,y):y=X^k_s(\omega)\})\\ =&\nu^n_s(X^k_s(\omega))\\ =&\frac{1}{n}\end{aligned}$$

Thus, we have $\langle B^{ik,n,m},B^{i'k',n,m'}\rangle_t=t\delta_{(i,k,m),(i',k',m')}$, and the result follows.

∎

We now are ready to prove Theorem 1.1.

Proof of Theorem 1.1, *a)* Let us fix $n\in\mathbb{N}^*$, and define for $i=1,\dots n$,

$$X^{i,n}_t=X^i_0+\frac{1}{\sqrt{n}}\int_0^t\sum_{k=1}^n\sigma(X^{i,n}_s-X^{k,n}_s)dB^{ik,n}_s+\frac{1}{n}\int_0^t\sum_{k=1}^n b(X^{i,n}_s-X^{k,n}_s)ds$$

or equivalently, in an indistinguishable way,

$$\begin{aligned}X^{i,n}_t=&X^i_0+\int_0^t\int_{\mathbb{R}^d}\sum_{k=1}^n\sigma(X^{i,n}_s-X^{k,n}_s)\mathbf{1}_{A^{k,n}}(s,y)W^i_P(dy,ds)\\ &+\int_0^t\int_{\mathbb{R}^d}\sum_{k=1}^n b(X^{i,n}_s-X^{k,n}_s)\mathbf{1}_{A^{k,n}}(s,y)P_s(dy)ds\end{aligned}$$

By standard arguments and the fact that the sets $A^{k,n}$ are disjoint (step (a) of the proof of Theorem 1.2), we have

$$E\left(|X_t^{i,n}-X_t^i|^2\right) \le \int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left[\sigma(X_s^{i,n}-X_s^{k,n})-\sigma(X_s^i-y)\right]^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds$$
$$+\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left[b(X_s^{i,n}-X_s^{k,n})-b(X_s^i-y)\right]^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \tag{9}$$

The first term in the right hand side of (9) is bounded by

$$\begin{aligned}
& C\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left[\sigma(X_s^{i,n}-X_s^{k,n})-\sigma(X_s^i-X_s^{k,n})\right]^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left[\sigma(X_s^i-X_s^{k,n})-\sigma(X_s^i-T^n(s,y))\right]^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left[\sigma(X_s^i-T^n(s,y))-\sigma(X_s^i-y)\right]^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \\
\le\ & C\int_0^t E\left(\sum_{k=1}^n \left(\left|X_s^{i,n}-X_s^i\right|^2 \int_{\mathbb{R}^d} \mathbf{1}_{A^{k,n}}(s,y)P_s(dy)\right)\right) ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(\left|X_s^{k,n}-T^n(s,y))\right|^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(\sum_{k=1}^n \left(|T^n(s,\omega,y)-y|^2 \mathbf{1}_{A^{k,n}}(s,y)\right)\right) P_s(dy)ds \\
=\ & C\int_0^t E\left(\left|X_s^{i,n}-X_s^i\right|^2\right) ds \\
& +C\int_0^t E\left(\sum_{k=1}^n \left(\left|X_s^{k,n}-X_s^k\right|^2 \int_{\mathbb{R}^d} \mathbf{1}_{A^{k,n}}(s,y)P_s(dy)\right)\right) ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(|T^n(s,\omega,y)-y|^2\right) P_s(dy)ds \\
=\ & C\int_0^t E\left(\left|X_s^{i,n}-X_s^i\right|^2\right) ds + C\int_0^t \frac{1}{n} E\left(\sum_{k=1}^n \left|X_s^{k,n}-X_s^k\right|^2\right) ds \\
& +C\int_0^t \int_{\mathbb{R}^d} E\left(|T^n(s,\omega,y)-y|^2\right) P_s(dy)ds \\
& = 2C\int_0^t E\left(\left|X_s^{i,n}-X_s^i\right|^2\right) ds + C\int_0^t E\left(W_2^2(\nu_s^n,P_s)\right) ds
\end{aligned}$$

by exchangeability of $((X^{1,n},X^1),\ldots,(X^{n,n},X^n))$. A similar bound is obtained for the second term in (9). We deduce by Gronwall's lemma that

$$E\left(|X_t^{i,n}-X_t^i|^2\right) \le C\exp(C'T)\int_0^t E(W_2^2\left(\nu_s^n,P_s\right))\,ds$$

By a little finer argument using a Burkholder-Davis-Gundy inequality, we can obtain as usual an estimate of the form

$$E\left(\sup_{t\in[0,T]}|X_t^{i,n}-X_t^i|^2\right)\leq C\exp(C'T)\int_0^T E\left(W_2^2(\nu_s^n,P_s)\right)ds$$

□

We recall a result proved in Rachev and Rüschendorf [14] giving L^2-rates of convergence of empirical measures in the Wasserstein metric.

Theorem 6.1. *([14] Theorem 10.2.1) Let* μ *a probability on* $\mathbb{R}^d$ *and let* $Y^1, Y^2, \ldots, Y^n$ *be independent identically distributed random variables with law* μ*. Let* μ_n *be the empirical measure of these variables. Then, if* μ *has high enough finite absolute moments:* $c := \int_{\mathbb{R}^d}|y|^{d+5}\mu(dy) < \infty$*, there is a constant* C *depending only on* c *and on the dimension* d*, such that*

$$E\left(W_2^2(\mu_n,\mu)\right)\leq Cn^{\frac{-2}{d+4}}.$$

Denote by $\mathcal{W}_2$ the Wasserstein distance between probability measures Q on the path space $\mathcal{C}_T := C([0,T],\mathbb{R}^d)$, such that $\int_{\mathcal{C}_T}\sup_{0\leq t\leq T}|x(t)|^2Q(dx) < \infty$.
¿From the previous result and Lemma 4.1, it is simple to deduce the following

Corollary 6.2. *Let* P *be the pathwise law of the nonlinear process* (1)*. Under the assumptions of Theorem 1.1 and moreover that* $\int_{\mathbb{R}^d}|y|^{d+5}P_0(dy) < \infty$*, we have that*

$$\mathcal{W}_2^2(law(X^{1,n}),P)\leq C_{T,d}n^{\frac{-2}{d+4}}.$$

□

The previous results are the first convergence rates obtained so far for stochastic particle systems of the "Landau type" (4), and they are not specific to the particular coefficients of the Landau equation (2). They justify the interest of the particle systems introduced in (4) and are the first step in the construction and the numerical study of a simulation algorithm for $(P_t)_t$. We notice that since we deal with space-time random fields, the dependence of the results on the dimension d is somewhat expectable, as opposite to the situation in the McKean-Vlasov model. The techniques we have introduced provide some insight about that dependence.

Acknowledgements The authors are very grateful to Roberto Cominetti for helpful suggestions about the theory of set-valued mappings.

References

[1] Brenier, Y. Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* **44**, (1991), 375–417.

[2] Cordier, S. and Mancini, S. A brief review on numerical methods for the collisions operators, *ENUMATH Proceedings 2001*, Springer (2003).

[3] Funaki, T. A certain class of diffusion processes associated with non linear parabolic equations, *Z. Wahrsch. Verw. Gebiete* **67**, (1984), 331–348.

[4] Funaki, T. The diffusion approximation of the spatially homogeneous Boltzmann equation, *Duke Math. J.* **52**, (1985), 1–23.

[5] Goudon, T. Sur l'équation de Boltzmann homogène et sa relation avec l'équation de Landau-Fokker-Planck: influence des collisions rasantes, *CRAS Paris Sér. I Math.* **324** (1997), 265–270.

[6] Guérin, H. Existence and regularity of a weak function-solution for some Landau equation with a stochastic approach, *Stochastic Process. Appl.* **101** (2002), 303–25.

[7] Guérin, H. Solving Landau equation for some soft potentials through a probabilistic approach, *Ann. Appl. Probab.* **13** (2003), 515–539.

[8] Guérin, H. and Méléard, S. Convergence from Boltzmann to Landau processes with soft potential and particle approximations, *J. Statist. Physics* **111** (2003), 931–66.

[9] Knott, M; Smith, C.S. On the optimal mapping of distributions, *J. Optim. Theory Appl.* **43** (1984), 39–49.

[10] McCann, R.J. Existence and uniqueness of monotone measure-preserving maps, *Duke Math. J.* **80**, (1995), no 2, 309–323.

[11] Méléard, S.; Roelly-Coppoletta, S. Système de particules et mesures-martingales: un théorème de propagation du chaos, Séminaire de Probabilités XXII, Lect. Notes in Maths 1321, Springer (1988), 438–448.

[12] Méléard S. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), 42–95, Lecture Notes in Math., 1627, Springer, Berlin, 1996.

[13] Nualart, D. The Malliavin calculus and related topics. Second edition, Springer-Verlag (2006).

[14] Rachev, S.T.; Ruschendorf, L. Mass Transportation Problems, Volumes I and II, Springer (1998).

[15] Rockafeller, R.T.; Wets, R. J-B.: Variational Analysis, Springer (1998).

[16] Sznitman A.S. Topics in propagation of chaos. *Ecole d'été de probabilités de Saint-Flour XIX-1989, Lect. Notes in Math.* **1464**, Springer (1991).

[17] Villani, C. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations, *ARMA* **143** (1998), 273–307.

[18] Villani, C. On the spatially homogeneous Landau equation for Maxwellian molecules, *Math. Mod. Meth. Appl. Sci.*

[19] Villani, C. Topics in Optimal Transportation, Graduate Studies in Mathematics Vol. 58, AMS (2003). **8** (1998), 957-983.

[20] Walsh, J.B. An introduction to stochastic partial differential equations, *École d'été de Probabilités de Saint-Flour XIV, Lect. Notes in Math.* **1180** (1984), 265-437.