View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Hal-Diderot

Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation

Benjamin Jourdain, Julien Reygner

To cite this version:

Benjamin Jourdain, Julien Reygner. Propagation of chaos for rank-based interacting diffusions and long time behaviour of a scalar quasilinear parabolic equation. Stochastic partial differential equations: analysis and computations, 2013, 1 (3), pp.455-506. <10.1007/s40072-013-0014-2>. <hal-00755269v2>

HAL Id: hal-00755269

https://hal-enpc.archives-ouvertes.fr/hal-00755269v2

Submitted on 24 Jun 2013

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.

PROPAGATION OF CHAOS FOR RANK-BASED INTERACTING DIFFUSIONS AND LONG TIME BEHAVIOUR OF A SCALAR QUASILINEAR PARABOLIC EQUATION

BENJAMIN JOURDAIN AND JULIEN REYGNER

Abstract. We study a quasilinear parabolic Cauchy problem with a cumulative distribution function on the real line as an initial condition. We call 'probabilistic solution' a weak solution which remains a cumulative distribution function at all times. We prove the uniqueness of such a solution and we deduce the existence from a propagation of chaos result on a system of scalar diffusion processes, the interactions of which only depend on their ranking. We then investigate the long time behaviour of the solution. Using a probabilistic argument and under weak assumptions, we show that the flow of the Wasserstein distance between two solutions is contractive. Under more stringent conditions ensuring the regularity of the probabilistic solutions, we finally derive an explicit formula for the time derivative of the flow and we deduce the convergence of solutions to equilibrium.

Introduction

Let $a, b : [0,1] \to \mathbb{R}$ be continuous functions, with $a \geq 0$. For all $u \in [0,1]$, let us define $A(u) = \int_0^u a(v)\mathrm{d}v$ and $B(u) = \int_0^u b(v)\mathrm{d}v$. Let m be a probability distribution on $\mathbb{R}$. We are interested in the nonlinear Cauchy problem on $[0,+\infty) \times \mathbb{R}$:

$$\begin{cases} \partial_t F_t(x) = \dfrac{1}{2}\partial_x^2\big(A(F_t(x))\big) - \partial_x\big(B(F_t(x))\big), \\ F_0(x) = H * m(x), \end{cases} \tag{1}$$

where $H * \cdot$ refers to the spatial convolution with the Heaviside function.

The partial differential equation in (1) is called a scalar *quasilinear parabolic* equation. It is a model for several usual nonlinear evolution equations, such as the porous medium equation, for which $B(u) = 0$ and the diffusion term has the particular form $A(u) = u^q$, $q > 1$; or conservation laws, in which the diffusion term is linear, i.e. $A(u) = \sigma^2 u$ with $\sigma^2 \geq 0$. A conservation law is said to be *viscous* if $\sigma^2 > 0$ and *inviscid* if $\sigma^2 = 0$. A particular case of a conservation law is the Burgers equation, for which $B(u) = u^2$.

In this article, we introduce a probabilistic approximation of the Cauchy problem by means of a system of scalar diffusion processes, interacting through their ranking. We then use this probabilistic representation to study the long time behaviour of the solution.

A *weak solution* to the Cauchy problem (1) is a continuous mapping $F : t \in [0,+\infty) \mapsto F_t \in L^1_{\mathrm{loc}}(\mathbb{R})$ such that for all $t \geq 0$, F_t takes its values in $[0,1]$ and for all $g \in C_{\mathrm{c}}^\infty([0,+\infty) \times \mathbb{R})$,

$$\begin{aligned} &\int_{\mathbb{R}} g(t,x)F_t(x)\mathrm{d}x - \int_{\mathbb{R}} g(0,x)H * m(x)\mathrm{d}x \\ &\quad = \int_{\mathbb{R}}\int_0^t \left\{\frac{1}{2}A(F_s(x))\partial_x^2 g(s,x) + B(F_s(x))\partial_x g(s,x) + F_s(x)\partial_s g(s,x)\right\}\mathrm{d}s\mathrm{d}x, \end{aligned} \tag{2}$$

where $C_{\mathrm{c}}^\infty([0,+\infty) \times \mathbb{R})$ refers to the space of real-valued C^∞ functions with compact support in $[0,+\infty) \times \mathbb{R}$.

2000 *Mathematics Subject Classification.* 65C35, 60K35, 35K59.

Key words and phrases. Nonlinear evolution equation, particle system, propagation of chaos, Wasserstein distance, long time behaviour.

This work is supported by the French National Research Agency under the grant ANR-12-BLAN-Stab. The final publication is available at http://link.springer.com/article/10.1007/s40072-013-0014-2.

When A is C^2 on $[0,1]$, we say that F has the *classical regularity* if F is $C^{1,2}$ on $(0,+\infty)\times\mathbb{R}$ and solves (1) in the classical sense. The space derivative p of a solution F with classical regularity satisfies the nonlinear Fokker-Planck equation

$$\partial_t p_t(x) = \frac{1}{2}\partial_x^2\big(a(H*p_t(x))p_t(x)\big) - \partial_x\big(b(H*p_t(x))p_t(x)\big), \tag{3}$$

therefore it is natural to consider the associated *nonlinear* stochastic differential equation

$$\begin{cases} X_t = X_0 + \displaystyle\int_0^t b(H*P_s(X_s))\mathrm{d}s + \int_0^t \sigma(H*P_s(X_s))\mathrm{d}W_s, \\ P_t \text{ is the distribution of } X_t, \end{cases} \tag{4}$$

where $\sigma(u) := a(u)^{1/2}$, X_0 has distribution m and is independent of the Brownian motion W. Due to the discontinuity of the Heaviside function, a direct study of this equation by classical techniques, such as the use of fixed-point theorems [38], seems out of reach (except when the diffusion coefficient σ is constant, see [24]); therefore we introduce a linearized approximation of (4).

In Section 1, we call *particle system* a solution $X^n \in C([0,+\infty),\mathbb{R}^n)$ to the stochastic differential equation in $\mathbb{R}^n$

$$\begin{cases} X_t^{i,n} = X_0^i + \displaystyle\int_0^t b(H*\mu_s^n(X_s^{i,n}))\mathrm{d}s + \int_0^t \big(c_n + \sigma(H*\mu_s^n(X_s^{i,n}))\big)\,\mathrm{d}W_s^i, \\ \mu_t^n = \dfrac{1}{n}\displaystyle\sum_{i=1}^n \delta_{X_t^{i,n}}, \end{cases} \tag{5}$$

where $(X_0^i)_{i\geq 1}$ is a sequence of i.i.d. random variables with marginal distribution m, independent of the $\mathbb{R}^n$-valued Brownian motion $(W^1,\dots,W^n)$. Here, the term $c_n > 0$ has been added in the diffusion in order to ensure the well definition of solutions, and it is only required to vanish when $n\to+\infty$. In Proposition 1.1, we prove that as soon as the function A is increasing and m has a finite first order moment, the flow of time marginals $t\mapsto\mu_t^n$ of the empirical distribution of the particle system converges in probability to the unique mapping $t\mapsto P(t)$ such that the function $F:(t,x)\mapsto(H*P(t))(x)$ is a weak solution to the Cauchy problem (1). This function will be referred to as the *probabilistic solution* of the Cauchy problem. Our analysis is essentially based on results obtained for the particular case of the porous medium equation [26]. A crucial argument in the extension of this work is the uniqueness of weak solutions to the Cauchy problem stated in Proposition 1.2, the proof of which is adaptated from works by Wu, Zhao, Yin and Lin [42] and Liu and Wang [32] (see Appendix A). Proposition 1.1 has strong connections with recent results by Shkolnikov [36], see Remark 1.6.

We then give two representations of the mapping $t\mapsto P(t)$ as the flow of time marginals of a probability distribution on the space of sample-paths $C([0,+\infty),\mathbb{R})$. More precisely, in Subsection 1.2, we give necessary and sufficient conditions on the coefficients of the Cauchy problem for the empirical distribution $\mu^n = (1/n)\sum_{i=1}^n \delta_{X^{i,n}}$ of the particle system in $C([0,+\infty),\mathbb{R})$ to converge in probability to the law P of a weak solution to (4). In Subsection 1.3, we define the *reordered particle system* as the reflected diffusion process obtained by increasingly reordering the positions $(X_t^{1,n},\dots,X_t^{n,n})$ of the particles. We prove in Proposition 1.15 that the associated empirical distribution $\tilde{\mu}^n$ converges in probability to a probability distribution $\tilde{P}$ on $C([0,+\infty),\mathbb{R})$ with time marginals $P(t)$.

Our motivation for introducing the particle system (5) is the study of nonlinear evolution problems, as it has been done for particular equations by Jourdain [25, 26, 27]. However, such systems of so-called *rank-based interacting particles* also arise in several contexts (see the introduction of [22] for references), and have received much attention lately. In particular, motivated by the study of the Atlas model of equity markets introduced by Fernholz [16] (see also Banner, Fernholz and

Karatzas [4]), much work has been done about the rank-based stochastic differential equation

$$\mathrm{d}X_t^i = \sum_{j=1}^n \mathbf{1}_{\{X_t^i = Y_t^j\}} b_j \mathrm{d}t + \sum_{j=1}^n \mathbf{1}_{\{X_t^i = Y_t^j\}} \sigma_j \mathrm{d}W_t^i, \tag{6}$$

where $(Y_t^1, \dots, Y_t^n)$ refers to the increasing reordering of $(X_t^1, \dots, X_t^n)$. The case $n = 2$ is exhaustively studied by Fernholz, Ichiba, Karatzas and Prokaj [15]. For $n \geq 3$, Ichiba, Karatzas and Shkolnikov [21] show that strong solutions can be defined as long as there is no triple collision. Triple collisions are studied in [20]. Concentration of measure bounds for the local time at collisions and statistics related to this system are given by Pal and Shkolnikov in [33].

As far as the long time behaviour of solutions to (6) is concerned, Ichiba, Papathanakos, Banner, Karatzas and Fernholz [23] prove that under some convexity assumption on the sequence of drift coefficients (b_j), the process of spacings $(Y_t^2 - Y_t^1, \dots, Y_t^n - Y_t^{n-1})$ converges in total variation to its unique stationary distribution when $t \to +\infty$. When the sequence of diffusion coefficients is such that $\sigma_2^2 - \sigma_1^2 = \cdots = \sigma_n^2 - \sigma_{n-1}^2$, this stationary distribution is the product of exponential distributions. These results extend the work by Pal and Pitman [34], in which $\sigma_j = 1$ for all j. In this case, the particle system solution to (6) does not have any equilibrium, as the process of its center of mass is a drifted Brownian motion. However, the convergence to equilibrium in total variation of its projection on the hyperplane $\{x_1 + \cdots + x_n = 0\}$ can be deduced from the long time behaviour of the process of spacings [34]. Convergence rates are provided by Ichiba, Pal and Shkolnikov [22] using Lyapounov functionals. Based on the Poincaré inequality satisfied by the stationary distribution, Jourdain and Malrieu [29] prove the convergence to equilibrium in χ^2 distance with an exponential rate, which is uniform in n. However, due to the lack of scaling property in the dimension n for the χ^2 distance, one cannot deduce from their result the convergence to equilibrium of the probabilistic solution F_t to the Cauchy problem (1), which is the purpose of Sections 2 and 3 of this article.

In many cases, transport metrics, and the Wasserstein distance in particular, are contractive for the flow of solutions to parabolic equations: see von Renesse and Sturm [41] for the linear Fokker-Planck equation, Carrillo, McCann and Villani [11], Cattiaux, Guillin and Malrieu [13] and the recent work by Bolley, Gentil and Guillin [8] for the granular media equation and Bolley, Guillin and Malrieu [7] for the kinetic Vlasov-Fokker-Planck equation. We will prove such a contractivity property by a probabilistic argument and without further regularity assumption, and then take advantage of it to state the convergence to equilibrium of the solutions.

Let us first recall some useful properties of the one-dimensional Wasserstein distance (see Villani [40] for a complete introduction). Let $p \geq 1$. For all probability distributions μ and ν on $\mathbb{R}$, we define

$$W_p(\mu, \nu) := \inf_{(X,Y) \in \Pi(\mu,\nu)} \mathbb{E}\left(|X - Y|^p\right)^{1/p},$$

where $\Pi(\mu, \nu)$ refers to the set of random couples (X, Y) such that X has marginal distribution μ and Y has marginal distribution ν. As soon as both μ and ν have a finite moment of order p, then $W_p(\mu, \nu) < +\infty$.

Given a right-continuous nondecreasing function F, we define its pseudo-inverse as $F^{-1}(u) := \inf\{x \in \mathbb{R} : F(x) > u\}$. Then it is a remarkable feature of the one-dimensional case that the Wasserstein distance $W_p(\mu, \nu)$ can be expressed in terms of the pseudo-inverses of the cumulative distribution functions $F_\mu := H * \mu$ and $F_\nu := H * \nu$ as

$$W_p^p(\mu, \nu) = \int_0^1 |F_\mu^{-1}(u) - F_\nu^{-1}(u)|^p \mathrm{d}u. \tag{7}$$

This leads to the following useful expressions: let $(x_1, \dots, x_n) \in \mathbb{R}^n$, we denote by $(y_1, \dots, y_n)$ its increasing reordering and by μ^n its empirical distribution. Then for all probability distribution μ,

$$W_p^p(\mu^n, \mu) = \sum_{i=1}^n \int_{(i-1)/n}^{i/n} |y_i - F_\mu^{-1}(u)|^p \mathrm{d}u \leq \sum_{i=1}^n \int_{(i-1)/n}^{i/n} |x_i - F_\mu^{-1}(u)|^p \mathrm{d}u. \tag{8}$$

In particular when μ is the empirical distribution μ'^n of some vector $(x'_1, \dots, x'_n)$ with increasing reordering $(y'_1, \dots, y'_n)$,

$$W_p^p(\mu^n, \mu'^n) = \frac{1}{n}\sum_{i=1}^n |y_i - y'_i|^p \leq \frac{1}{n}\sum_{i=1}^n |x_i - x'_i|^p. \tag{9}$$

We finally point out the fact that we will indifferently refer to the Wasserstein distance between μ and ν as $W_p(\mu, \nu)$ or $W_p(F_\mu, F_\nu)$.

In Section 2 we study the evolution of the Wasserstein distance between two probabilistic solutions F and G of the Cauchy problem (1) with different initial conditions F_0 and G_0. We use the contractivity of the reordered particle system in Proposition 2.1 to prove that the flow $t \mapsto W_p(F_t, G_t)$ is nonincreasing. Then we provide an explicit expression of the time derivative of the flow in Proposition 2.3. Our work is related to results exposed in the review papers by Carrillo and Toscani [10] and Carrillo, Di Francesco and Lattanzio [9]; a further review is given in Remark 2.5.

Section 3 is dedicated to the convergence to equilibrium of the solutions. We call *stationary solution* the cumulative function F_∞ of a probability distribution m_∞ with a finite first order moment, such that if F is the probabilistic solution of the Cauchy problem (1) with $F_0 = F_\infty$, then for all $t \geq 0$, $F_t = F_\infty$. It is clear from (2) that F_∞ is a stationary solution if and only if it solves the *stationary equation* $(1/2)\partial^2(A(F_\infty)) - \partial(B(F_\infty)) = 0$, in the sense of distributions.

We solve the stationary equation in Proposition 3.1, extending the results of Jourdain and Malrieu [29] who deal with the viscous conservation law. Using the results of Section 2 as well as the probabilistic approximation built in Section 1, we then prove in Theorem 3.6 that the probabilistic solutions converge to the stationary solutions in Wasserstein distance.

In [29], the solutions of the viscous conservation law are proven to converge exponentially fast to equilibrium in χ^2 distance, under the condition that the initial measure m be close enough to the stationary solution m_∞. Theorem 3.6 does not involve such a condition, but the proof does not provide any indication on the rate of convergence of F_t to the equilibrium. As we remark in Subsection 3.3, one can recover an exponential rate of convergence in (quadratic) Wasserstein distance in the setting of [29].

Notations. Given a separable metric space S, we denote by $\mathcal{P}(S)$ the set of Borel probability distributions on S, equipped with the topology of weak convergence. The space $C([0, +\infty), S)$ of continuous functions from $[0, +\infty)$ to S is provided with the topology of the uniform convergence on the compact sets of $[0, +\infty)$. Besides, for all probability distribution $\mu \in \mathcal{P}(C([0, +\infty), S))$, the marginal distribution at time $t \geq 0$ is denoted by $\mu_t \in \mathcal{P}(S)$. The *canonical application* $\mathcal{P}(C([0, +\infty), S)) \to C([0, +\infty), \mathcal{P}(S))$ associates the distribution μ with the flow of its time marginals $t \mapsto \mu_t$. Finally, if f is a real-valued bounded function then $||f||_\infty$ refers to the supremum of the function $|f|$.

Assumptions. Our results are valid under various assumptions on the degeneracy of the parabolic equation. Let us introduce the following conditions:

(D1) The function A is increasing.
(D2) For all $u \in (0, 1]$, $a(u) > 0$.
(D3) There exists $\underline{a} > 0$ such that, for all $u \in [0, 1]$, $a(u) \geq \underline{a}$.

Obviously, (D1) is weaker than (D2), which is weaker than (D3).

We also introduce the two following conditions on the regularity of the coefficients:

(R1) The function a is C^1 on $[0, 1]$.
(R2) The function a is C^2 on $[0, 1]$, the function b is C^1 on $[0, 1]$ and there exists $\beta > 0$ such that the functions a'' and b' are β-Hölder continuous.

The condition (R1) is a natural necessary condition for the Cauchy problem (1) to admit classical solutions. The stronger condition (R2) will be used in Lemma 1.7 to ensure the existence of classical solutions to the Fokker-Planck equation (3).

Finally, the existence and integrability of stationary solutions will depend on the two following equilibrium conditions:

(E1) For all $u \in (0,1)$, $B(u) > 0$, $B(1) = 0$ and the function $a/2B$ is locally integrable on $(0,1)$.
(E2) The function $a/2B$ is such that

$$\int_0^{1/2} \frac{a(u)u}{2|B(u)|}\mathrm{d}u + \int_{1/2}^1 \frac{a(u)(1-u)}{2|B(u)|}\mathrm{d}u < +\infty.$$

1. Probabilistic approximation of the solution

1.1. **Existence and uniqueness of the probabilistic solution.** Following [5], for all $n \geq 1$ there exists a unique weak solution $X^n = (X_t^{1,n}, \dots, X_t^{n,n})_{t\geq 0}$ to the stochastic differential equation (5). We call it the *particle system* and denote by μ^n the random variable in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$ defined by $\mu^n := (1/n)\sum_{i=1}^n \delta_{X^{i,n}}$.

Let $T > 0$, possibly $T = +\infty$. We denote by $\mathcal{P}_1(T)$ the set of continuous mappings $t \in [0,T) \mapsto P(t) \in \mathcal{P}(\mathbb{R})$ such that for all $t \in [0,T)$, the probability distribution $P(t)$ has a finite first order moment and the function $t \mapsto \int_{\mathbb{R}} |x| P(t)(\mathrm{d}x)$ is locally integrable on $[0,T)$. Let $\mathcal{F}(T) := \{F : (t,x) \mapsto (H * P(t))(x); P \in \mathcal{P}_1(T)\}$; note that $\mathcal{F}(T) \subset C([0,T), L^1_{\mathrm{loc}}(\mathbb{R}))$. The particular sets $\mathcal{P}_1(+\infty)$ and $\mathcal{F}(+\infty)$ are simply denoted by $\mathcal{P}_1$ and $\mathcal{F}$.

Throughout this article, we will call *probabilistic solution* the solution to the Cauchy problem (1) given by the following proposition.

Proposition 1.1. *Under the nondegeneracy condition (D1) and the assumption that m has a finite first order moment, there exists a unique weak solution F to the Cauchy problem (1) in $\mathcal{F}$, and it writes $F : (t,x) \mapsto (H * P(t))(x)$ where the mapping $t \mapsto P(t)$ is the limit in probability, in $C([0,+\infty), \mathcal{P}(\mathbb{R}))$, of the sequence of mappings $t \mapsto \mu_t^n$.*

The proof of Proposition 1.1 relies on Proposition 1.2 and Lemmas 1.3 and 1.4. For all $n \geq 1$, let π'^n denote the distribution of μ^n in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$. In Lemma 1.3, we prove that the sequence $(\pi'^n)_{n\geq 1}$ is tight. Since the canonical application $\mathcal{P}(C([0,+\infty),\mathbb{R})) \to C([0,+\infty), \mathcal{P}(\mathbb{R}))$ is continuous, then the sequence $(\pi^n)_{n\geq 1}$ of the distributions of the random mappings $t \mapsto \mu_t^n$ in $C([0,+\infty), \mathcal{P}(\mathbb{R}))$ is tight. Let π^∞ be the limit of a converging subsequence, that we still index by n for convenience. Lemma 1.4 combined with Proposition 1.2 proves that π^∞ concentrates on a single point $P \in \mathcal{P}_1$, which is such that the function $(t,x) \mapsto (H * P(t))(x)$ is a weak solution of (1).

Proposition 1.2. *Assume that the nondegeneracy condition (D1) holds.*

1. *Let $T > 0$, possibly $T = +\infty$. Let F^1 and $F^2 \in \mathcal{F}(T)$, such that for all $g \in C_c^\infty([0,T) \times \mathbb{R})$, F^1 and F^2 satisfy (2). Then $F^1 = F^2$ in $\mathcal{F}(T)$.*
2. *There is at most one weak solution to the Cauchy problem (1) in $\mathcal{F}$.*

Proof. The second point of the proposition clearly follows from the first point, and the first point is proved in Appendix A. □

Lemma 1.3. *The sequence $(\pi'^n)_{n\geq 1}$ is tight.*

Proof. Since the distribution of $(X^{1,n}, \dots, X^{n,n})$ in $C([0,+\infty), \mathbb{R}^n)$ is symmetric, according to Sznitman [38, Proposition 2.2, p. 177], $(\pi'^n)_{n\geq 1}$ is tight if and only if the sequence of the distributions of the variables $X^{1,n} \in C([0,+\infty),\mathbb{R})$ is tight. This latter fact classically follows from the fact that for all $n \geq 1$, $X_0^{1,n} = X_0^1$ has distribution m on the one hand, and from the Kolmogorov criterion as well as the boundedness of the coefficients a and b and the sequence $(c_n)_{n\geq 1}$ on the other hand. □

Lemma 1.4. *Under the assumption that m has a finite first order moment, the distribution π^∞ is concentrated on the set of mappings $P \in \mathcal{P}_1$ such that the function $(t,x) \mapsto (H * P(t))(x)$ is a weak solution to the Cauchy problem (1).*

Proof. We first prove that π^∞ concentrates on $\mathcal{P}_1$. Let μ^∞ be a variable in $C([0,+\infty),\mathcal{P}(\mathbb{R}))$ with distribution π^∞. We will prove that for all $t\geq 0$,

$$\sup_{s\in[0,t]}\int_{\mathbb{R}}|x|\mu^\infty(s)(\mathrm{d}x)<+\infty \qquad \text{a.s.},$$

so that taking t in a countable unbounded subset of $[0,+\infty)$ yields $\mu^\infty\in\mathcal{P}_1$ almost surely.

Let $t\geq 0$. For all $M\geq 0$, the function $f_M:\mu\mapsto\sup_{s\in[0,t]}\int_{\mathbb{R}}(|x|\wedge M)\mu(s)(\mathrm{d}x)$ is continuous and bounded on $C([0,+\infty),\mathcal{P}(\mathbb{R}))$. For fixed n,

$$\begin{aligned}\mathbb{E}(f_M(\mu^n))&\leq\frac{1}{n}\sum_{i=1}^n\mathbb{E}\left(\sup_{s\in[0,t]}|X_s^{i,n}|\right)\\&\leq\int_{\mathbb{R}}|x|m(\mathrm{d}x)+t||b||_\infty+\frac{1}{n}\sum_{i=1}^n\left[\mathbb{E}\left(\sup_{s\in[0,t]}\left|\int_0^s\left(c_n+\sigma(H*\mu_r^n(X_r^{i,n}))\right)\mathrm{d}W_r^i\right|^2\right)\right]^{1/2}\\&\leq C,\end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the second line and the Doob inequality as well as the fact that m has a finite first order moment in the third line. The constant C depends neither on M nor on n. As a consequence, $\liminf_{M\to+\infty}\mathbb{E}(f_M(\mu^\infty))\leq C$ and by Fatou's lemma,

$$C\geq\mathbb{E}\left(\liminf_{M\to+\infty}f_M(\mu^\infty)\right)\geq\mathbb{E}\left(\sup_{s\in[0,t]}\liminf_{M\to+\infty}\int_{\mathbb{R}}(|x|\wedge M)\mu^\infty(s)(\mathrm{d}x)\right).$$

By the monotone convergence theorem,

$$\liminf_{M\to+\infty}\int_{\mathbb{R}}(|x|\wedge M)\mu^\infty(s)(\mathrm{d}x)=\lim_{M\to+\infty}\int_{\mathbb{R}}(|x|\wedge M)\mu^\infty(s)(\mathrm{d}x)=\int_{\mathbb{R}}|x|\mu^\infty(s)(\mathrm{d}x),$$

so that $\mathbb{E}\left(\sup_{s\in[0,t]}\int_{\mathbb{R}}|x|\mu^\infty(s)(\mathrm{d}x)\right)\leq C$, which yields the expected result.

It now remains to prove that $(t,x)\mapsto(H*\mu^\infty(t))(x)$ is almost surely a weak solution of the Cauchy problem (1). The computation is made for the porous medium equation in [26, Lemma 1.5] and can be straightforwardly extended; it relies on the uniform continuity of a and b on $[0,1]$ and the fact that $c_n\to 0$. □

Remark 1.5. Without assuming neither (D1) nor the existence of a first order moment for m, one can still prove that the sequence $(\pi^n)_{n\geq1}$ is tight and that the limit of any converging subsequence concentrates on weak solutions to the Cauchy problem (1), that of course not necessarily belong to $\mathcal{F}$. Thus, the existence of weak solutions holds under very weak assumptions.

Remark 1.6. The law of large numbers for the sequence of mappings $t\mapsto\mu_t^n$ stated in Proposition 1.1 has recently been addressed under more restrictive conditions on the initial condition m and the coefficients a and b. In [36], Shkolnikov studies the particle system (5) with the specific condition that the process of spacings between two particles with consecutive positions in $\mathbb{R}$ be stationary (the description of the stationary distribution is given in [23]). Then in the case where a is affine and b is C^1, with b' uniformly negative, the sequence of mappings $t\mapsto\mu_t^n$ is proven to converge in probability, in $C([0,+\infty),\mathcal{P}(\mathbb{R}))$, to the unique mapping $t\mapsto P(t)$ such that $(t,x)\mapsto(H*P(t))(x)$ is a weak solution to the Cauchy problem (1) for a specified initial condition F_0. In place of our Proposition 1.2, the author uses Gilding's theorem for uniqueness [19, Theorem 4] and therefore needs to assume that any weak solution to the Cauchy problem is continuous on $[0,+\infty)\times\mathbb{R}$ when m does not weight points.

In the more recent article by Dembo, Shkolnikov, Varadhan and Zeitouni [14], the stationarity assumption is removed and the continuity of F is obtained as a consequence of mild regularity and nondegeneracy assumptions on the coefficients of the Cauchy problem (1). More precisely, the authors establish a large deviation principle for the sequence $(\pi^n)_{n\geq1}$, with a rate function that is infinite on the set of mappings $t\mapsto P(t)$ such that the function $(t,x)\mapsto(H*P(t))(x)$ is discontinuous. They also prove that a zero of the rate function is a mapping $t\mapsto P(t)$ such that the function $(t,x)\mapsto(H*P(t))(x)$ is a continuous weak solution to (1), and deduce the law of large numbers as a consequence of Gilding's uniqueness theorem.

Both approaches heavily rely on the continuity of the solution F to (1), as it is a crucial condition to use Gilding's uniqueness theorem. While we address the regularity of F in Lemma 1.7 below, we insist on the fact that our proof of Proposition 1.1 does not require that F be continuous, which allows us to relax the regularity and nondegeneracy assumptions on m, a and b with respect to [36, 14]. However, the regularity of F plays a more important role in establishing the law of large numbers for the sequence of empirical distributions $\mu^n \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$ in Subsection 1.2. Therefore, in the proof of Lemma 1.10, we prove that, under the nondegeneracy condition (D2), the function F_t is continuous on $\mathbb{R}$, $\mathrm{d}t$-almost everywhere.

We conclude this subsection by discussing the regularity of the probabilistic solution F. For all finite $T > 0$, we denote by $C^{1,2}_{\mathrm{b}}([0,T]\times\mathbb{R})$ the set of $C^{1,2}$ functions on $[0,T]\times\mathbb{R}$ that are bounded together with their derivatives. For all $l > 0$, the Hölder spaces $H^l(\mathbb{R})$ and $H^{l/2,l}([0,T]\times\mathbb{R})$ are defined as in [31, p. 7].

Lemma 1.7. *Assume that the uniform ellipticity condition (D3) and the regularity condition (R2) hold, that m has a finite first order moment and that $H * m$ is in the Hölder space $H^l(\mathbb{R})$, with $l = 3+\beta$. Then for all finite $T > 0$, the probabilistic solution F to (1) is in $C^{1,2}_{\mathrm{b}}([0,T]\times\mathbb{R})$. In particular, it is a classical solution to (1).*

Proof. Fix a finite $T > 0$. Then owing to the assumptions (D3), (R2) and on the regularity of $H * m$, the classical result of Ladyženskaja, Solonnikov and Ural'ceva [31, Theorem 8.1, p. 495] ensures that the Cauchy problem in divergence form

$$\begin{cases} \partial_t \tilde{F}_t(x) = \partial_x \left(\frac{1}{2} a(\tilde{F}_t(x)) \partial_x \tilde{F}_t(x) - B(\tilde{F}_t(x)) \right), \\ \tilde{F}_0(x) = H * m(x), \end{cases} \tag{10}$$

admits a classical bounded solution $\tilde{F}$, which belongs to the Hölder space $H^{l/2,l}([0,T]\times\mathbb{R})$, with $l = 3+\beta$. Certainly, $\tilde{F}$ satisfies (2) for all $g \in C^\infty_{\mathrm{c}}([0,T)\times\mathbb{R})$. Let us now prove that $\tilde{F} \in \mathcal{F}(T)$. On the one hand, by the maximum principle [31, Theorem 2.5, p. 18], for all $t \in [0,T]$, $||\tilde{F}_t||_\infty \leq 1$. On the other hand, the space derivative $\tilde{p} := \partial_x \tilde{F}$ is $C^{1,2}$ on $[0,T]\times\mathbb{R}$ and satisfies the linear parabolic equation

$$\partial_t \tilde{p}_t(x) = \tilde{a}(t,x)\partial_x^2 \tilde{p}_t(x) + \tilde{b}(t,x)\partial_x \tilde{p}_t(x) + \tilde{c}(t,x)\tilde{p}_t(x),$$

where

$$\begin{aligned} \tilde{a}(t,x) &:= \frac{1}{2} a(\tilde{F}_t(x)), \\ \tilde{b}(t,x) &:= \frac{3}{2} a'(\tilde{F}_t(x))\partial_x \tilde{F}_t(x) - b(\tilde{F}_t(x)), \\ \tilde{c}(t,x) &:= \frac{1}{2} a''(\tilde{F}_t(x))(\partial_x \tilde{F}_t(x))^2 - b'(\tilde{F}_t(x))\partial_x \tilde{F}_t(x). \end{aligned}$$

The coefficients $\tilde{a}$, $\tilde{b}$ and $\tilde{c}$ are continuous and bounded in $[0,T]\times\mathbb{R}$ and, due to the condition (D3), the operator is parabolic. By the maximum principle [18, Theorem 9, p. 43], and since $\tilde{p}_0 \geq 0$, then $\tilde{p}_t(x) \geq 0$ for all $(t,x) \in [0,T]\times\mathbb{R}$.

As a consequence, for all $t \in [0,T]$, $\tilde{p}_t$ is the density of a nonnegative bounded measure on $\mathbb{R}$, with total mass lower than 1. Let us now prove that the mapping $t \mapsto \tilde{p}_t(x)\mathrm{d}x$ is continuous for the topology of weak convergence. Since $\tilde{p}$ is continuous on $[0,T]\times\mathbb{R}$, the mapping $t \mapsto \tilde{p}_t(x)\mathrm{d}x$ is continuous for the topology of vague convergence. Besides, as, for all $s,t \in [0,T]$, $\sup_{x\in\mathbb{R}} |F_t(x) - F_s(x)| \leq ||\partial_t F||_\infty |t-s|$, the total mass $t \mapsto \int_{\mathbb{R}} \tilde{p}_t(x)\mathrm{d}x$ is continuous.

Hence, the continuous mapping $t \mapsto \tilde{p}_t(x)\mathrm{d}x$ is a measure-valued solution to the linear Fokker-Planck equation

$$\partial_t \mu_t = \frac{1}{2}\partial_x^2\big(a(\tilde{F}_t(x))\mu_t\big) - \partial_x\big(b(\tilde{F}_t(x))\mu_t\big),$$

the coefficients of which are measurable and bounded functions on $[0,T]\times\mathbb{R}$. Therefore by Figalli [17, Theorem 2.6], and since $\tilde{p}_0$ is a probability density, then for all $t \in [0,T]$, $\tilde{p}_t$ is the density

of the distribution of $\tilde{X}_t$, where $(\tilde{X}_t)_{t\in[0,T]}$ is a weak solution to the stochastic differential equation

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t b(\tilde{F}_s(\tilde{X}_s))\mathrm{d}s + \int_0^t \sigma(\tilde{F}_s(\tilde{X}_s))\mathrm{d}\tilde{W}_s,$$

where $\tilde{X}_0$ has distribution m and is independent of the Brownian motion $\tilde{W}$. Now one easily deduces from the assumption that m has a finite first order moment and from the boundedness of σ and b that $\tilde{F} \in \mathcal{F}(T)$. Therefore, by the first part of Proposition 1.2, $\tilde{F}$ is the restriction to $[0,T]$ of the probabilistic solution F to (1) given by Proposition 1.1. Hence, $F \in C^{1,2}_{\mathrm{b}}([0,T]\times\mathbb{R})$ and the fact that F is a classical solution to (1) now follows from the fact that T is arbitrarily large. □

Remark 1.8. The regularity assumption on the initial condition $H*m$ is far from being necessary for the probabilistic solution F to have the classical regularity. For instance, it is known for the case of the viscous conservation law that F has the classical regularity even for a discontinuous initial condition $H*m$ (see [29, Corollary 1.2]).

1.2. **The nonlinear martingale problem.** The propagation of chaos result of Proposition 1.1 only deals with the flow of time marginals of the empirical distribution μ^n. A natural further question is the convergence in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$ towards the solution to a proper *nonlinear martingale problem*.

Recall that the distribution of the random variable μ^n in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$ is denoted by π'^n, and by Lemma 1.3, it is tight. Let X refer to the canonical process on the probability space $C([0,+\infty),\mathbb{R})$, namely $X_t(\omega) := \omega_t$ for all $\omega \in C([0,+\infty),\mathbb{R})$. We shall also denote by $C^2_{\mathrm{b}}(\mathbb{R})$ the space of C^2 functions $\phi : \mathbb{R} \to \mathbb{R}$ such that ϕ, ϕ' and ϕ'' are bounded.

Definition 1.9. *A probability distribution $P \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$ is called a solution to the nonlinear martingale problem if:*

- $P_0 = m$;
- *for all $\phi \in C^2_{\mathrm{b}}(\mathbb{R})$, the process M^ϕ defined by $M^\phi_t := \phi(X_t) - \phi(X_0) - \int_0^t L(P_s)\phi(X_s)\mathrm{d}s$ is a P-martingale, where, for all $\mu \in \mathcal{P}(\mathbb{R})$, $L(\mu)\phi(x) := b(H*\mu(x))\phi'(x) + (1/2)a(H*\mu(x))\phi''(x)$;*
- $\mathrm{d}t$*-almost everywhere, P_t does not weight points.*

Following [26, Lemma 1.2], if P is a solution to the nonlinear martingale problem, then the function $(t,x) \mapsto H*P_t(x)$ is a weak solution to the Cauchy problem (1). Besides, since the coefficients a and b are bounded, it is easily seen that if m has a finite first order moment, then this solution belongs to $\mathcal{F}$, therefore it coincides with the probabilistic solution F given by Proposition 1.1.

Owing to Lévy's characterization of the Brownian motion, a probability distribution solving the nonlinear martingale problem is the distribution of a weak solution to the nonlinear stochastic differential equation (4). Reciprocally, the distribution P of a weak solution to (4) is a solution to the nonlinear martingale problem if and only if, $\mathrm{d}t$-almost everywhere, P_t does not weight points. When there exists a unique solution to the nonlinear martingale problem, we will refer to the associated weak solution X of (4) as the *nonlinear diffusion process*.

Let us first investigate the existence of a solution to the nonlinear martingale problem.

Lemma 1.10. *Under the nondegeneracy condition (D2) and the assumption that m has a finite first order moment, the limit of any converging subsequence of $(\pi'^n)_{n\geq 1}$ concentrates on the set of solutions to the nonlinear martingale problem.*

Proof. By Lemma 1.3, the sequence $(\pi'^n)_{n\geq 1}$ is tight. Let π'^∞ denote the limit of a converging subsequence, that we still index by n for convenience. Let Q refer to the canonical variable in the probability space $\mathcal{P}(C([0,+\infty),\mathbb{R}))$. Since the variables X^i_0 are i.i.d. with marginal distribution m, then π'^∞-a.s., $Q_0 = m$. Let us now prove that π'^∞-a.s., $\mathrm{d}t$-almost everywhere, Q_t does not weight points. By Proposition 1.1, π'^∞-a.s., for all $t \geq 0$ one has $H*Q_t = F_t$ where F is the probabilistic solution to the Cauchy problem (1). Therefore it is enough to prove that, $\mathrm{d}t$-almost everywhere, the function F_t is continuous on $\mathbb{R}$. In this purpose, we first remark that the mapping $t \mapsto P(t)$ solves a linear Fokker-Planck equation. Indeed, since A and B are C^1 on $[0,1]$, the

functions $(t,x) \mapsto A(F_t(x))$, $(t,x) \mapsto B(F_t(x))$ are of finite variation and the associated Stieltjes measures write $\mathrm{d}(A(F_t(x))) = \bar{a}(t,x)P(t)(\mathrm{d}x)$, $\mathrm{d}(B(F_t(x))) = \bar{b}(t,x)P(t)(\mathrm{d}x)$, where

$$\bar{a}(t,x) := \begin{cases} a(F_t(x)) & \text{if } F_t \text{ is continuous in } x, \\ \dfrac{A(F_t(x)) - A(F_t(x^-))}{F_t(x) - F_t(x^-)} & \text{otherwise,} \end{cases}$$

and $\bar{b}(t,x)$ is similarly defined. Remark that the functions $\bar{a}$ and $\bar{b}$ are bounded, and $||\bar{a}||_\infty \leq ||a||_\infty$, $||\bar{b}||_\infty \leq ||b||_\infty$. As a consequence, the continuous mapping $t \mapsto P(t)$ is a measure-valued solution on $[0,+\infty)$ to the Fokker-Planck equation with measurable and bounded coefficients

$$\partial_t P(t) = \frac{1}{2}\partial_x^2(\bar{a}(t,x)P(t)) - \partial_x(\bar{b}(t,x)P(t)).$$

Fix a finite $T > 0$. Then by Figalli [17, Theorem 2.6], there exists a probability distribution $\bar{P} \in \mathcal{P}(C([0,T],\mathbb{R}))$ such that, for all $t \in [0,T]$, $\bar{P}_t = P(t)$ and $\bar{P}$ is the distribution of a weak solution $(\bar{X}_t)_{t\in[0,T]}$ on $[0,T]$ to the stochastic differential equation

$$\bar{X}_t = \bar{X}_0 + \int_0^t \bar{b}(s,\bar{X}_s)\mathrm{d}s + \int_0^t \bar{\sigma}(s,\bar{X}_s)\mathrm{d}\bar{W}_s,$$

where $\bar{\sigma}(t,x) := \bar{a}(t,x)^{1/2}$. The process $(\bar{X}_t)_{t\in[0,T]}$ satisfies the condition of Bogachev, Krylov and Röckner [6, Remark 2.2.3, p. 63]. Hence the positive measure $\bar{\sigma}(t,x)\bar{P}_t(\mathrm{d}x)\mathrm{d}t$ admits a density $\rho(t,x) \in L^2_{\mathrm{loc}}([0,T]\times\mathbb{R})$. Now for all $t \in [0,T]$, F_t is the cumulative distribution function of $\bar{X}_t$ so that $F_t(x) > 0$, $\bar{P}_t$-almost everywhere. By (D2) and the definition of $\bar{\sigma}$, we deduce that $\bar{P}_t$-almost everywhere, $\bar{\sigma}(t,x) > 0$, therefore

$$\bar{P}_t(\mathrm{d}x)\mathrm{d}t = \mathbf{1}_{\{\bar{\sigma}(t,x)>0\}}\frac{\rho(t,x)}{\bar{\sigma}(t,x)}\mathrm{d}x\mathrm{d}t$$

and consequently, $\bar{P}_t(\mathrm{d}x)$ admits a density $\mathrm{d}t$-almost everywhere in $[0,T]$. We conclude by taking T arbitrarily large.

We finally prove that π'^∞-a.s., for all $\phi \in C^2_{\mathrm{b}}(\mathbb{R})$, the process M^ϕ defined by $M^\phi_t := \phi(X_t) - \phi(X_0) - \int_0^t L(Q_s)\phi(X_s)\mathrm{d}s$ is a Q-martingale. We will proceed as in the proof of [26, Lemma 1.6]. Let $\phi \in C^2_{\mathrm{b}}(\mathbb{R})$, $k \geq 1$, $0 \leq s_1 \leq \cdots \leq s_k \leq s \leq t$ and $g : \mathbb{R}^k \to \mathbb{R}$ continuous and bounded. For all $Q \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$, we define

$$\mathcal{G}(Q) := \left\langle Q, g(X_{s_1},\ldots,X_{s_k})\left(\phi(X_t) - \phi(X_s) - \int_s^t L(Q_r)\phi(X_r)\mathrm{d}r\right)\right\rangle.$$

By Itô's formula, for all $n \geq 1$,

$$\mathcal{G}(\mu^n) = \frac{1}{n}\sum_{i=1}^n g(X_{s_1},\ldots,X_{s_k})\Bigg(\int_s^t \phi'(X^{i,n}_r)\left(c_n + \sigma(H*\mu^n_t(X^{i,n}_r))\right)\mathrm{d}W^i_r \\ + \int_s^t \phi''(X^{i,n}_r)\left(c_n\sigma(H*\mu^n_t(X^{i,n}_r)) + \frac{c_n^2}{2}\right)\mathrm{d}r\Bigg),$$

so that, since $(c_n)_{n\geq 1}$, σ, g, ϕ' and ϕ'' are bounded, $\lim_{n\to+\infty}\mathbb{E}(\mathcal{G}(\mu^n)^2) = 0$. We now check that the functional $\mathcal{G}$ is continuous at all $P \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$ such that, $\mathrm{d}r$-almost everywhere, P_r does not weight points. Let $(P^q)_{q\geq 1}$ be a sequence of probability distributions on $C([0,+\infty),\mathbb{R})$ weakly converging to $P \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$ such that, $\mathrm{d}r$-almost everywhere, P_r does not weight points. Then, for all $q \geq 1$,

$$\mathcal{G}(P^q) = \left\langle P^q, g(X_{s_1},\ldots,X_{s_k})\int_s^t (L(P^q_r) - L(P_r))\phi(X_r)\mathrm{d}r\right\rangle \\ + \left\langle P^q, g(X_{s_1},\ldots,X_{s_k})\left(\phi(X_t) - \phi(X_s) - \int_s^t L(P_r)\phi(X_r)\mathrm{d}r\right)\right\rangle. \tag{11}$$

On the one hand, as g and the derivatives of ϕ are bounded, there exists $C > 0$ independent of q such that, for all $q \geq 1$,

$$\left|\left\langle P^q, g(X_{s_1}, \ldots, X_{s_k}) \int_s^t (L(P_r^q) - L(P_r))\phi(X_r)\mathrm{d}r \right\rangle\right|$$
$$\leq C \int_s^t \left(\sup_{x\in\mathbb{R}} |b(H * P_r^q(x)) - b(H * P_r(x))| + \sup_{x\in\mathbb{R}} |a(H * P_r^q(x)) - a(H * P_r(x))| \right) \mathrm{d}r.$$

By Dini's theorem, $\mathrm{d}r$-almost everywhere in $[s,t]$, $H * P_r^q$ converges uniformly to $H * P_r$ on $\mathbb{R}$ when $q \to +\infty$. As the functions b and a are bounded and uniformly continuous on $[0,1]$, by Lebesgue's theorem, the right-hand side above goes to 0 when $q \to +\infty$.

On the other hand, $\mathrm{d}r$-almost everywhere in $[s,t]$, the function $x \mapsto L(P_r)\phi(x)$ is continuous on $\mathbb{R}$ and uniformly bounded in r, therefore by Lebesgue's theorem again, the function $(x_r)_{r\geq 0} \mapsto g(x_{s_1}, \ldots, x_{s_k})(\phi(x_t) - \phi(x_s) - \int_s^t L(P_r)\phi(x_r)\mathrm{d}r)$ is continuous on $C([0,+\infty),\mathbb{R})$. As a consequence, the second term in the right-hand side of (11) converges to $\mathcal{G}(P)$ and we conclude that $\lim_{q\to+\infty} \mathcal{G}(P^q) = \mathcal{G}(P)$.

Since we have proved that π'^{∞}-a.s., $\mathrm{d}r$-almost everywhere, Q_r does not weight points, then

$$\lim_{n\to+\infty} \mathbb{E}^{\pi'^n}(\mathcal{G}(Q)^2) = \mathbb{E}^{\pi'^{\infty}}(\mathcal{G}(Q)^2),$$

which rewrites

$$\mathbb{E}^{\pi'^{\infty}}(\mathcal{G}(Q)^2) = \lim_{n\to+\infty} \mathbb{E}(\mathcal{G}(\mu^n)^2) = 0.$$

As a consequence, taking ϕ, $(s_1, \ldots, s_k, s, t)$, g in countable subsets leads to the conclusion that π'^{∞}-a.s., Q solves the nonlinear martingale problem. □

We now address the uniqueness of solutions to the nonlinear martingale problem. The following criterion is due to Stroock and Varadhan [37].

Lemma 1.11. *Under the assumptions of Proposition 1.1, if the function $(t,x) \mapsto a(F_t(x))$ is uniformly positive on the compact sets of $[0,+\infty) \times \mathbb{R}$, then there is at most one solution to the nonlinear martingale problem.*

Proof. Let P and Q denote two solutions to the nonlinear martingale problem. Then they both solve the following linear martingale problem in $R \in \mathcal{P}(C([0,+\infty),\mathbb{R}))$:

- $R_0 = m$;
- for all $\phi \in C_{\mathrm{b}}^2(\mathbb{R})$, the process $\tilde{M}^{\phi}$ defined by
$$\tilde{M}_t^{\phi} := \phi(X_t) - \phi(X_0) - \int_0^t \left\{ \phi'(X_s) b(F_s(X_s)) + \frac{1}{2}\phi''(X_s) a(F_s(X_s)) \right\} \mathrm{d}s$$
is a R-martingale.

The functions $a(F_s(x))$ and $b(F_s(x))$ are measurable and bounded, and $a(F_s(x))$ is uniformly positive on the compact sets of $[0,+\infty) \times \mathbb{R}$. By [37, Exercise 7.3.3, p. 192], $P = Q$. □

Lemma 1.12. *Assume that m has a finite first order moment, and either the uniform ellipticity condition (D3) holds, or the nondegeneracy condition (D2) holds and $F_0(x) > 0$ for all $x \in \mathbb{R}$. Then the function $(t,x) \mapsto a(F_t(x))$ is uniformly positive on the compact sets of $[0,+\infty) \times \mathbb{R}$.*

Proof. If (D3) holds, the result is obvious. Now if $F_0(x) > 0$, for all compact subset $K \in [0,+\infty) \times \mathbb{R}$, by the first part of Lemma B.1 in Appendix B, there exists $u_0 > 0$ such that for all $(t,x) \in K$, $F_t(x) \geq u_0$. If (D2) holds in addition, for all $(t,x) \in K$, $a(F_t(x)) \geq \inf_{u\geq u_0} a(u) > 0$. □

We conclude this subsection by stating a propagation of chaos result for the empirical distribution μ^n of the particle system in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$.

Corollary 1.13. *Under the assumptions of Lemma 1.12, there exists a unique solution P to the nonlinear martingale problem, and it is the limit in probability, in $\mathcal{P}(C([0,+\infty),\mathbb{R}))$, of the sequence of empirical distributions μ^n.*

1.3. **The reordered particle system.** For all $t \geq 0$, let $Y_t^n := (Y_t^{1,n}, \dots, Y_t^{n,n})$ denote the increasing reordering of the vector $(X_t^{1,n}, \dots, X_t^{n,n})$. Then the sample-paths of the process Y^n are in $C([0,+\infty), D_n)$, where D_n refers to the polyhedron $\{(y_1, \dots, y_n) \in \mathbb{R}^n : y_1 \leq \dots \leq y_n\}$.

It is well known that Y^n is a normally reflected Brownian motion on ∂D_n, with constant drift vector and constant diagonal diffusion matrix. More precisely, according to [26], the process $(\beta^1, \dots, \beta^n)$ defined by $\beta_t^i = \sum_{j=1}^n \int_0^t \mathbf{1}_{\{X_s^{j,n} = X_s^{i,n}\}} \mathrm{d}W_s^j$ is a Brownian motion. By the Itô-Tanaka formula,

$$Y_t^{i,n} = Y_0^{i,n} + b(i/n)t + (c_n + \sigma(i/n))\beta_t^i + V_t^i, \tag{12}$$

where V is a $\mathbb{R}^n$-valued continuous process with finite variation $|V|$ which writes $V_t^i = \int_0^t (\gamma_s^i - \gamma_s^{i+1}) \mathrm{d}|V|_s$ with $\mathrm{d}|V|_t$-a.e., $\gamma_t^1 = \gamma_t^{n+1} = 0$, $\gamma_t^i \geq 0$ and $\gamma_t^i (Y_t^{i,n} - Y_t^{i-1,n}) = 0$. We shall now refer to the process Y^n as the *reordered particle system* and denote by $\tilde{\mu}^n \in \mathcal{P}(C([0,+\infty), \mathbb{R}))$ its empirical distribution.

Lemma 1.14 (Tanaka [39])**.** *For a given random variable $Y_0^n \in D_n$ and an independent $\mathbb{R}^n$-valued Brownian motion $(\beta^1, \dots, \beta^n)$, there exists a unique process $(Y^n, V) \in C([0,+\infty), D_n \times \mathbb{R}^n)$ satisfying all the above conditions.*

For $Q \in \mathcal{P}(C([0,+\infty), \mathbb{R}))$ and $t_1, \dots, t_k \geq 0$, let us denote by $Q_{t_1,\dots,t_k} \in \mathcal{P}(\mathbb{R}^k)$ the finite-dimensional marginal distribution of Q. Let us define $\mathcal{A}$ as the set of probability distributions $Q \in \mathcal{P}(C([0,+\infty), \mathbb{R}))$ such that, for all $0 \leq t_1 < \dots < t_k$, $Q_{t_1,\dots,t_k}$ is the distribution of $(H * Q_{t_1})^{-1}(U), \dots, (H * Q_{t_k})^{-1}(U))$ where U is a uniform random variable on $[0,1]$. Remark that any $Q \in \mathcal{A}$ is exactly determined by the flow of its one-dimensional marginals $t \mapsto Q_t$.

Proposition 1.15. *Under the assumptions of Proposition 1.1, the empirical distribution $\tilde{\mu}^n$ of the reordered particle system converges in probability, in $\mathcal{P}(C([0,+\infty), \mathbb{R}))$, to the unique $\tilde{P} \in \mathcal{A}$ such that for all $t \geq 0$, $\tilde{P}_t = P(t)$, where the mapping $t \mapsto P(t)$ is given by Proposition 1.1. In particular, for all $t \geq 0$, $H * \tilde{P}_t = F_t$ where F is the probabilistic solution to the Cauchy problem (1).*

Proof. Let $\tilde{\pi}^n$ refer to the distribution of $\tilde{\mu}^n$ in $\mathcal{P}(C([0,+\infty), \mathbb{R}))$. According to Sznitman [38], the tightness of $(\tilde{\pi}^n)_{n \geq 1}$ is equivalent to the tightness of the sequence of the distributions of the variables $Y^{\theta_n,n} \in C([0,+\infty), \mathbb{R})$ where θ_n is a uniform random variable in the set $\{1, \dots, n\}$, independent of Y^n. For all $n \geq 1$, $Y_0^{\theta_n,n}$ has distribution m. Besides, for $s \leq t$ and $p \geq 1$,

$$\mathbb{E}\left(|Y_s^{\theta_n,n} - Y_t^{\theta_n,n}|^p\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left(|Y_s^{i,n} - Y_t^{i,n}|^p\right) = \mathbb{E}\left(W_p^p(\tilde{\mu}_s^n, \tilde{\mu}_t^n)\right).$$

Now, by (9) and exchangeability of $(X_t^{1,n})_{t \geq 0}, \dots, (X_t^{n,n})_{t \geq 0}$,

$$\mathbb{E}\left(W_p^p(\tilde{\mu}_s^n, \tilde{\mu}_t^n)\right) \leq \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left(|X_s^{i,n} - X_t^{i,n}|^p\right) = \mathbb{E}\left(|X_s^{1,n} - X_t^{1,n}|^p\right),$$

so that, by the proof of Lemma 1.3 and the Kolmogorov criterion, the sequence $(\tilde{\pi}^n)_{n \geq 1}$ is tight.

It is clear from the definition of the reordered particle system that for all $t \geq 0$, $\mu_t^n = \tilde{\mu}_t^n$, therefore it is already known from Proposition 1.1 that $\tilde{\mu}_t^n$ converges in distribution to $P(t)$. Consequently, the proof of Proposition 1.15 requires nothing but a uniqueness result for the support of any limit point of $(\tilde{\pi}^n)_{n \geq 1}$.

The latter is a consequence of the following remark. Certainly, $\tilde{\mu}^n \in \mathcal{A}$, and the set $\mathcal{A}$ is closed in $\mathcal{P}(C([0,+\infty), \mathbb{R}))$ by [27, Lemma 3.5]. Hence, any limit point of $(\tilde{\pi}^n)_{n \geq 1}$ is concentrated on the unique probability distribution $\tilde{P} \in \mathcal{A}$ such that for all $t \geq 0$, $\tilde{P}_t = P(t)$. □

1.4. **Propagation of chaos in Wasserstein distance.** The original particle system defined by (5) is exchangeable, therefore the propagation of chaos result stated in Proposition 1.1 implies that the distribution $P_t^{1,n}$ of $X_t^{1,n}$ converges weakly to $P(t)$ in $\mathcal{P}(\mathbb{R})$. This convergence result can be strengthened in Wasserstein distance.

Corollary 1.16. *Under the nondegeneracy condition (D1) and the assumption that m has a finite moment of order $p \geq 1$, then $P_t^{1,n}$ and $P(t)$ have a finite moment of order p, and*

$$\lim_{n\to+\infty} W_p(P_t^{1,n}, P(t)) = 0, \qquad \lim_{n\to+\infty} \mathbb{E}[W_p^p(\mu_t^n, P(t))] \to 0.$$

Proof. Let $t \geq 0$. As just seen before the corollary, $P_t^{1,n}$ converges weakly to $P(t)$. To prove that this convergence holds in the Wasserstein distance of order p, it is sufficient to prove that the sequence $(|X_t^{1,n}|^p)_{n\geq 1}$ is uniformly integrable (see Villani [40, Theorem 6.9, p. 108]). For all $q \geq 1$,

$$\mathbb{E}\left(|X_t^{1,n} - X_0^1|^q\right) \leq 2^{q-1}\left((t||b||_\infty)^q + \mathbb{E}\left(\left|\int_0^t \left(c_n + \sigma(H * \mu_s^n(X_s^{1,n}))\right) \mathrm{d}W_s^1\right|^q\right)\right) \leq C, \tag{13}$$

where C does not depend on n. Thus, the sequence $(|X_t^{1,n} - X_0^1|^p)_{n\geq 1}$ is uniformly integrable, and since $|X_t^{1,n}|^p \leq 2^{p-1}(|X_t^{1,n} - X_0^1|^p + |X_0^1|^p)$ then the sequence $(|X_t^{1,n}|^p)_{n\geq 1}$ is uniformly integrable. Therefore $P_t^{1,n}$ and $P(t)$ have a finite moment of order p and $W_p(P_t^{1,n}, P(t)) \to 0$.

Let $M \geq 0$. Then, by (8),

$$\begin{aligned} W_p^p(\mu_t^n, P(t)) &= \sum_{i=1}^n \int_{(i-1)/n}^{i/n} (|Y_t^{i,n} - F_t^{-1}(u)|^p - M)^+ \mathrm{d}u + \sum_{i=1}^n \int_{(i-1)/n}^{i/n} (|Y_t^{i,n} - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u \\ &\leq \sum_{i=1}^n \int_{(i-1)/n}^{i/n} |Y_t^{i,n} - F_t^{-1}(u)|^p \mathbf{1}_{\{|Y_t^{i,n} - F_t^{-1}(u)|^p \geq M\}} \mathrm{d}u + \sum_{i=1}^n \int_{(i-1)/n}^{i/n} (|Y_t^{i,n} - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u. \end{aligned}$$

On the one hand,

$$\sum_{i=1}^n \int_{(i-1)/n}^{i/n} (|Y_t^{i,n} - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u = \int_0^1 (|(H * \mu_t^n)^{-1}(u) - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u,$$

and the function $\mu \in \mathcal{P}(\mathbb{R}) \mapsto \int_0^1 (|(H * \mu)^{-1}(u) - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u$ is continuous and bounded. Therefore, by Proposition 1.1,

$$\lim_{n\to+\infty} \mathbb{E}\left(\sum_{i=1}^n \int_{(i-1)/n}^{i/n} (|Y_t^{i,n} - F_t^{-1}(u)|^p \wedge M) \mathrm{d}u\right) = 0.$$

On the other hand, remarking that for all $x, y \in \mathbb{R}$,

$$\begin{aligned} |x-y|^p \mathbf{1}_{\{|x-y|^p \geq M\}} &\leq |x-y|^p \mathbf{1}_{\{|x| \geq |y| \vee M^{1/p}/2\}} + |x-y|^p \mathbf{1}_{\{|y| \geq |x| \vee M^{1/p}/2\}} \\ &\leq 2^p |x|^p \mathbf{1}_{\{|x|^p \geq M/2^p\}} + 2^p |y|^p \mathbf{1}_{\{|y|^p \geq M/2^p\}}, \end{aligned} \tag{14}$$

we write

$$\begin{aligned} &\sum_{i=1}^n \int_{(i-1)/n}^{i/n} |Y_t^{i,n} - F_t^{-1}(u)|^p \mathbf{1}_{\{|Y_t^{i,n} - F_t^{-1}(u)|^p \geq M\}} \mathrm{d}u \\ &\quad \leq \frac{2^p}{n} \sum_{i=1}^n |Y_t^{i,n}|^p \mathbf{1}_{\{|Y_t^{i,n}|^p \geq M/2^p\}} + 2^p \int_0^1 |F_t^{-1}(u)|^p \mathbf{1}_{\{|F_t^{-1}(u)|^p \geq M/2^p\}} \mathrm{d}u \\ &\quad = \frac{2^p}{n} \sum_{i=1}^n |X_t^{i,n}|^p \mathbf{1}_{\{|X_t^{i,n}|^p \geq M/2^p\}} + 2^p \int_0^1 |F_t^{-1}(u)|^p \mathbf{1}_{\{|F_t^{-1}(u)|^p \geq M/2^p\}} \mathrm{d}u. \end{aligned}$$

We deduce from the exchangeability of the variables $X_t^{1,n}, \ldots, X_t^{n,n}$, the uniform integrability of $(|X_t^{1,n}|^p)_{n\geq 1}$ and the finiteness of $\int_0^1 |F_t^{-1}(u)|^p \mathrm{d}u = \int_{\mathbb{R}} |x|^p P(t)(\mathrm{d}x)$ that

$$\lim_{M\to+\infty} \sup_{n\geq 1} \mathbb{E}\left(\sum_{i=1}^n \int_{(i-1)/n}^{i/n} |Y_t^{i,n} - F_t^{-1}(u)|^p \mathbf{1}_{\{|Y_t^{i,n} - F_t^{-1}(u)|^p \geq M\}} \mathrm{d}u\right) = 0,$$

so that $\mathbb{E}[W_p^p(\mu_t^n, P(t))] \to 0$. □

2. Contraction of the Wasserstein distance between two solutions

Let F_0 and G_0 be the cumulative functions of two probability distributions with a finite first order moment. Under the condition (D1), by Proposition 1.1 there exist a unique probabilistic solution F to the Cauchy problem (1) with initial condition F_0, and a unique probabilistic solution G to the Cauchy problem (1) with initial condition G_0. This section addresses the behaviour of the flow $t \mapsto W_p(F_t, G_t)$. In Proposition 2.1 we prove that it is nonincreasing if $W_p(F_0, G_0) < +\infty$, using only the contractivity of the reordered particle system. Then, assuming the classical regularity of F and G, we provide an explicit expression of the time derivative of the flow $t \mapsto W_p^p(F_t, G_t)$.

We point out the fact that we will sometimes call *expectation* or *moment* of a cumulative distribution function the expectation or the moment of the derivated probability distribution.

2.1. Monotonicity of the flow.

We first deduce from a natural coupling between two versions of the reordered particle system that the flow $t \mapsto W_p(F_t, G_t)$ is nonincreasing.

Proposition 2.1. *Assume that the nondegeneracy condition (D1) holds and that F_0 and G_0 have a finite first order moment. Then, for all $p \geq 1$,*

- *if $W_p(F_0, G_0) < +\infty$, then the flow $t \mapsto W_p(F_t, G_t)$ in nonincreasing;*
- *if $W_p(F_0, G_0) = +\infty$, then for all $t \geq 0$, $W_p(F_t, G_t) = +\infty$.*

Proof. We deduce the monotonicity property from the contractive behaviour of the reordered particle system. Let $(\beta^1, \dots, \beta^n)$ be a $\mathbb{R}^n$-valued Brownian motion and let $U^1, \dots, U^n$ be independent uniform variables on $[0, 1]$. Let us denote by $(U^{(1)}, \dots, U^{(n)})$ the increasing reordering of $(U^1, \dots, U^n)$. For all $1 \leq i \leq n$, let $Y_0^{F,i} := F_0^{-1}(U^{(i)})$ and $Y_0^{G,i} := G_0^{-1}(U^{(i)})$. By Lemma 1.14, there exists a unique strong solution $(Y^F, V^F) \in C([0, +\infty), D_n \times \mathbb{R}^n)$ to the reflected stochastic differential equation

$$Y_t^{F,i,n} = Y_0^{F,i} + b(i/n)t + (c_n + \sigma(i/n))\beta_t^i + V_t^{F,i},$$

and similarly, we denote by (Y^G, V^G) the unique strong solution in $C([0, +\infty), D_n \times \mathbb{R}^n)$ to the reflected stochastic differential equation

$$Y_t^{G,i,n} = Y_0^{G,i} + b(i/n)t + (c_n + \sigma(i/n))\beta_t^i + V_t^{G,i}.$$

By the beginning of Subsection 1.3 and the Yamada-Watanabe theorem, the process $Y^{F,n}$ (resp. $Y^{G,n}$) has the same distribution as the increasing reordering of the particle system $X^{F,n}$ (resp. $X^{G,n}$) solution to (5) with initial conditions i.i.d. according to the cumulative distribution function F_0 (resp. G_0). In particular, the propagation of chaos result of Proposition 1.15 applies to the empirical distributions $\tilde{\mu}^{F,n} := (1/n)\sum_{i=1}^n \delta_{Y^{F,i,n}}$ and $\tilde{\mu}^{G,n} := (1/n)\sum_{i=1}^n \delta_{Y^{G,i,n}}$.

Now, for all $t \geq 0$, (9) yields $W_p^p(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n}) = (1/n)\sum_{i=1}^n |Y_t^{F,i,n} - Y_t^{G,i,n}|^p$. The following inequality, the proof of which is postponed below, is crucial:

$$\forall 0 \leq s \leq t, \qquad W_p^p(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n}) \leq W_p^p(\tilde{\mu}_s^{F,n}, \tilde{\mu}_s^{G,n}). \tag{15}$$

Case $W_p(F_0, G_0) < +\infty$. If both F_0 and G_0 have a finite moment of order p, then owing to Corollary 1.16, one can extract a subsequence along which $W_p^p(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n})$ goes to $W_p^p(F_t, G_t)$ and $W_p^p(\tilde{\mu}_s^{F,n}, \tilde{\mu}_s^{G,n})$ goes to $W_p^p(F_s, G_s)$ almost surely, then conclude by using (15).

Assuming only $W_p(F_0, G_0) < +\infty$, we shall now proceed as in the proof of Corollary 1.16 to show that for all $t \geq 0$,

$$\lim_{n \to +\infty} \mathbb{E}(W_p^p(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n})) = W_p^p(F_t, G_t),$$

which results in the claimed assertion thanks to (15).

For all $M \geq 0$, by (9) we write

$$\begin{aligned} W_p^p(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n}) &= \frac{1}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_t^{G,i,n}|^p \\ &= \frac{1}{n}\sum_{i=1}^n (|Y_t^{F,i,n} - Y_t^{G,i,n}|^p \wedge M) + \left(|Y_t^{F,i,n} - Y_t^{G,i,n}|^p - M\right)^+. \end{aligned}$$

By Proposition 1.1, $\tilde{\mu}_t^{F,n}$ converges in probability to the probability distribution $\mathrm{d}F_t$ with cumulative distribution function F_t, and similarly $\tilde{\mu}_t^{G,n}$ converges in probability to $\mathrm{d}G_t$. Therefore, the couple $(\tilde{\mu}_t^{F,n}, \tilde{\mu}_t^{G,n})$ converges in probability to $(\mathrm{d}F_t, \mathrm{d}G_t)$ and

$$\lim_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n (|Y_t^{F,i,n} - Y_t^{G,i,n}|^p \wedge M)\right) = \int_0^1 (|F_t^{-1}(u) - G_t^{-1}(u)|^p \wedge M)\mathrm{d}u.$$

By the monotone convergence theorem and (7),

$$\lim_{M\to+\infty} \lim_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n (|Y_t^{F,i,n} - Y_t^{G,i,n}|^p \wedge M)\right) = W_p^p(F_t, G_t) \in [0, +\infty].$$

It now remains to check that

$$\lim_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n (|Y_t^{F,i,n} - Y_t^{G,i,n}|^p - M)^+\right) = 0. \tag{16}$$

Using (14) twice results in

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \left(|Y_t^{F,i,n} - Y_t^{G,i,n}|^p - M\right)^+ &\leq \frac{1}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_t^{G,i,n}|^p \mathbf{1}_{\{|Y_t^{F,i,n} - Y_t^{G,i,n}|^p \geq M\}} \\
&\leq \frac{4^p}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_0^{F,i}|^p \mathbf{1}_{\{|Y_t^{F,i,n} - Y_0^{F,i}|^p \geq M/4^p\}} \\
&\quad + \frac{4^p}{n}\sum_{i=1}^n |Y_t^{G,i,n} - Y_0^{G,i}|^p \mathbf{1}_{\{|Y_t^{G,i,n} - Y_0^{G,i}|^p \geq M/4^p\}} \\
&\quad + \frac{2^p}{n}\sum_{i=1}^n |Y_0^{F,i} - Y_0^{G,i}|^p \mathbf{1}_{\{|Y_0^{F,i} - Y_0^{G,i}|^p \geq M/2^p\}}.
\end{aligned}$$

On the one hand, by the construction of $Y_0^{F,i}$ and $Y_0^{G,i}$,

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |Y_0^{F,i} - Y_0^{G,i}|^p \mathbf{1}_{\{|Y_0^{F,i} - Y_0^{G,i}|^p \geq M/2^p\}}\right) \\
&\quad = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |F_0^{-1}(U^{(i)}) - G_0^{-1}(U^{(i)})|^p \mathbf{1}_{\{|F_0^{-1}(U^{(i)}) - G_0^{-1}(U^{(i)})|^p \geq M/2^p\}}\right) \\
&\quad = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |F_0^{-1}(U^{i}) - G_0^{-1}(U^{i})|^p \mathbf{1}_{\{|F_0^{-1}(U^{i}) - G_0^{-1}(U^{i})|^p \geq M/2^p\}}\right) \\
&\quad = \int_0^1 |F_0^{-1}(u) - G_0^{-1}(u)|^p \mathbf{1}_{\{|F_0^{-1}(u) - G_0^{-1}(u)|^p \geq M/2^p\}}\mathrm{d}u,
\end{aligned}$$

and the right-hand side does not depend on n. Since $W_p(F_0, G_0) < +\infty$, it goes to 0 when $M \to +\infty$. On the other hand,

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_0^{F,i}|^p \mathbf{1}_{\{|Y_t^{F,i,n} - Y_0^{F,i}|^p \geq M/4^p\}}\right) &\leq \frac{4^p}{M}\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_0^{F,i}|^{p+1}\right) \\
&\leq \frac{4^p}{M}\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |X_t^{F,i,n} - X_0^{F,i}|^{p+1}\right),
\end{aligned}$$

where we have used the inequality (9) in the last line. By (13), there exists $C > 0$ independent of n such that $\mathbb{E}(|X_t^{F,i,n} - X_0^{F,i}|^{p+1}) \leq C$. Then

$$\lim_{M\to+\infty} \limsup_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n |Y_t^{F,i,n} - Y_0^{F,i}|^p \mathbf{1}_{\{|Y_t^{F,i,n} - Y_0^{F,i}|^p \geq M/4^p\}}\right) = 0;$$

and likewise,

$$\lim_{M\to+\infty}\limsup_{n\to+\infty}\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}|Y_t^{G,i,n}-Y_0^{G,i}|^p\mathbf{1}_{\{|Y_t^{G,i,n}-Y_0^{G,i}|^p\geq M/4^p\}}\right)=0,$$

which completes the proof of (16).

Case $W_p(F_0,G_0)=+\infty$. By the triangle inequality,

$$W_p(F_0,G_0)\leq W_p(F_0,F_t)+W_p(G_0,G_t)+W_p(F_t,G_t).$$

According to Proposition 1.15, there exists a subsequence (that we still index by n for convenience) along which $\tilde{\mu}_0^{F,n}$ converges to the distribution with cumulative function F_0 almost surely in $\mathcal{P}(\mathbb{R})$, and $\tilde{\mu}_t^{F,n}$ converges to the distribution with cumulative function F_t almost surely in $\mathcal{P}(\mathbb{R})$. Recalling that the Wasserstein distance is lower semicontinuous on $\mathcal{P}(\mathbb{R})$ (see [40, Remark 6.12]), we get $W_p(F_0,F_t)\leq\liminf_{n\to+\infty}W_p(\tilde{\mu}_0^{F,n},\tilde{\mu}_t^{F,n})$ so that by Fatou's lemma, (9) and (13),

$$W_p^p(F_0,F_t)\leq\liminf_{n\to+\infty}\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}|Y_t^{F,i,n}-Y_0^{F,i}|^p\right)<+\infty,$$

and similarly $W_p(G_0,G_t)<+\infty$. As a consequence, if $W_p(F_0,G_0)=+\infty$ then $W_p(F_t,G_t)=+\infty$.

Proof of (15). Recall that

$$\begin{aligned}Y_t^{F,i,n}&=Y_0^{F,i}+b(i/n)t+(c_n+\sigma(i/n))\beta_t^i+V_t^{F,i},\\ Y_t^{G,i,n}&=Y_0^{G,i}+b(i/n)t+(c_n+\sigma(i/n))\beta_t^i+V_t^{G,i},\end{aligned}$$

with $\mathrm{d}V_t^{F,i}=(\gamma_t^{F,i}-\gamma_t^{F,i+1})\mathrm{d}|V^F|_t$, $\mathrm{d}V_t^{G,i}=(\gamma_t^{G,i}-\gamma_t^{G,i+1})\mathrm{d}|V^G|_t$. Thus, for all $1\leq i\leq n$, the process $Y^{F,i,n}-Y^{G,i,n}$ is of finite variation, hence

$$\begin{aligned}\mathrm{d}\sum_{i=1}^{n}|Y_t^{F,i,n}-Y_t^{G,i,n}|^p&=p\sum_{i=1}^{n}|Y_t^{F,i,n}-Y_t^{G,i,n}|^{p-2}(Y_t^{F,i,n}-Y_t^{G,i,n})\mathrm{d}(Y_t^{F,i,n}-Y_t^{G,i,n})\\ &=p\sum_{i=1}^{n}|Y_t^{F,i,n}-Y_t^{G,i,n}|^{p-2}(Y_t^{F,i,n}-Y_t^{G,i,n})(\gamma_t^{F,i}-\gamma_t^{F,i+1})\mathrm{d}|V^F|_t\\ &\quad+p\sum_{i=1}^{n}|Y_t^{G,i,n}-Y_t^{F,i,n}|^{p-2}(Y_t^{G,i,n}-Y_t^{F,i,n})(\gamma_t^{G,i}-\gamma_t^{G,i+1})\mathrm{d}|V^G|_t,\end{aligned}$$

where, for all $p\geq 1$, we take the convention that $|z|^{p-2}z=0$ when $z=0$. The two terms of the last member above are symmetric; we only deal with the first one, which we rewrite $S_t\mathrm{d}|V^F|_t$. By the Abel transform, $S_t=p\sum_{i=2}^{n}\gamma_t^{F,i}u(Y_t^{F,i-1,n},Y_t^{G,i-1,n},Y_t^{F,i,n},Y_t^{G,i,n})$, where $u(x_1,y_1,x_2,y_2):=|x_2-y_2|^{p-2}(x_2-y_2)-|x_1-y_1|^{p-2}(x_1-y_1)$. Recall that $\mathrm{d}|V^F|_t$-a.e., $\gamma_t^{F,i}\geq 0$ and $\gamma_t^{F,i}(Y_t^{F,i,n}-Y_t^{F,i-1,n})=0$, and remark that for fixed $y_1\leq y_2$, the expression $u(x,y_1,x,y_2)$ remains nonpositive when $x\in\mathbb{R}$. Then $\mathrm{d}|V^F|_t$-a.e., $S_t\leq 0$ and the proof of (15) is completed. □

2.2. **Derivative of the flow.** We will now compute the time derivative of the flow $t\mapsto W_p^p(F_t,G_t)$ when F and G have the classical regularity. We first derive a nonlinear evolution equation for the pseudo-inverse function F_t^{-1}. Of course, the same holds for G_t^{-1}.

Lemma 2.2. *Assume that the uniform ellipticity condition (D3) and the regularity condition (R1) hold, that F_0 has a finite first order moment, that the probabilistic solution F to the Cauchy problem (1) with initial conditon F_0 has the classical regularity and that for all $0<t_1<t_2$, the function $(t,x)\mapsto\partial_xF_t(x)$ is bounded on $[t_1,t_2]\times\mathbb{R}$. Then the pseudo-inverse function $(t,u)\mapsto F_t^{-1}(u)$ is $C^{1,2}$ on $(0,+\infty)\times(0,1)$ and satisfies*

$$\partial_tF_t^{-1}(u)=b(u)-\frac{1}{2}\partial_u\left(\frac{a(u)}{\partial_uF_t^{-1}(u)}\right).\tag{17}$$

Proof. On the one hand, since F is a classical solution to (1), then

$$\partial_t F_t(x) = \frac{1}{2}\left(a'(F_t(x))(\partial_x F_t(x))^2 + a(F_t(x))\partial_x^2 F_t(x)\right) - b(F_t(x))\partial_x F_t(x). \tag{18}$$

On the other hand, the lower bound in the Aronson estimate (47) allows to apply the implicit functions theorem to $(t,x,u) \mapsto F_t(x) - u$ and deduce that $(t,u) \mapsto F_t^{-1}(u)$ is $C^{1,2}$ on $(0,+\infty)\times\mathbb{R}$. Besides,

$$\partial_u F_t^{-1}(u) = \frac{1}{\partial_x F_t(F_t^{-1}(u))} \qquad \text{and} \qquad \partial_u^2 F_t^{-1}(u) = -\frac{\partial_x^2 F_t(F_t^{-1}(u))}{[\partial_x F_t(F_t^{-1}(u))]^3}. \tag{19}$$

For all $t > 0$, since F_t is a continuous function then $F_t(F_t^{-1}(u)) = u$, so that derivating with respect to t yields $0 = \partial_t F_t(F_t^{-1}(u)) + \partial_x F_t(F_t^{-1}(u))\partial_t F_t^{-1}(u)$. Using (18) and (19), we write

$$\partial_t F_t^{-1}(u) = b(u) - \frac{1}{2}\left(\frac{a'(u)}{\partial_u F_t^{-1}(u)} - \frac{a(u)\partial_u^2 F_t^{-1}(u)}{[\partial_u F_t^{-1}(u)]^2}\right),$$

from which we deduce (17). □

For all function $f : \mathbb{R} \to [0,1]$, we define the function $\mathrm{tl}(f,\cdot) : \mathbb{R} \to [0,1]$ by

$$\mathrm{tl}(f,x) := \mathbf{1}_{\{x\geq 0\}}(1 - f(x)) + \mathbf{1}_{\{x\leq 0\}} f(x).$$

Then the main result of this subsection is the following.

Proposition 2.3. *Assume that the conditions of Lemma 2.2 are satisfied with both F and G. Let $p \geq 2$ such that $W_p(F_0,G_0) < +\infty$ and $|x|^{p-1}(\mathrm{tl}(F_0,x) + \mathrm{tl}(G_0,x)) \to 0$ when $x \to \pm\infty$. Then for all $0 < t_1 < t_2$,*

$$\begin{aligned} &W_p^p(F_{t_2},G_{t_2}) - W_p^p(F_{t_1},G_{t_1}) \\ &\quad = -\frac{p(p-1)}{2}\int_{t_1}^{t_2}\int_0^1 a(u)|F_t^{-1}(u) - G_t^{-1}(u)|^{p-2}\frac{\left(\partial_u F_t^{-1}(u) - \partial_u G_t^{-1}(u)\right)^2}{\partial_u F_t^{-1}(u)\partial_u G_t^{-1}(u)}\mathrm{d}u\mathrm{d}t. \end{aligned}$$

Proof. See Appendix B. □

Remark 2.4. A straightforward strategy to prove Proposition 2.3 is the following: for $\epsilon_1,\epsilon_2 \in (0,1/2)$, let

$$W_{p,\epsilon_1,\epsilon_2}^p(F_t,G_t) := \int_{\epsilon_1}^{1-\epsilon_2} |F_t^{-1}(u) - G_t^{-1}(u)|^p \mathrm{d}u,$$

then certainly $W_{p,\epsilon_1,\epsilon_2}^p(F_t,G_t) \to W_p^p(F_t,G_t)$ when $\epsilon_1,\epsilon_2 \to 0$. Now,

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}W_{p,\epsilon_1,\epsilon_2}^p(F_t,G_t) \\ &\quad = p\int_{\epsilon_1}^{1-\epsilon_2}(F_t^{-1}(u) - G_t^{-1}(u))|F_t^{-1}(u) - G_t^{-1}(u)|^{p-2}\left(\partial_t F_t^{-1}(u) - \partial_t G_t^{-1}(u)\right)\mathrm{d}u \\ &\quad = \frac{p}{2}\int_{\epsilon_1}^{1-\epsilon_2}(F_t^{-1}(u) - G_t^{-1}(u))|F_t^{-1}(u) - G_t^{-1}(u)|^{p-2}\partial_u\left(\frac{a(u)}{\partial_u G_t^{-1}(u)} - \frac{a(u)}{\partial_u F_t^{-1}(u)}\right)\mathrm{d}u \\ &\quad = \frac{p}{2}\left[a(u)(F_t^{-1}(u) - G_t^{-1}(u))|F_t^{-1}(u) - G_t^{-1}(u)|^{p-2}\left(\frac{1}{\partial_u G_t^{-1}(u)} - \frac{1}{\partial_u F_t^{-1}(u)}\right)\right]_{\epsilon_1}^{1-\epsilon_2} \\ &\qquad - \frac{p(p-1)}{2}\int_{\epsilon_1}^{1-\epsilon_2} a(u)|F_t^{-1}(u) - G_t^{-1}(u)|^{p-2}\frac{\left(\partial_u F_t^{-1}(u) - \partial_u G_t^{-1}(u)\right)^2}{\partial_u F_t^{-1}(u)\partial_u G_t^{-1}(u)}\mathrm{d}u, \end{aligned}$$

where we have used (17) at the second line and integrated by parts in the last line. Hence, Proposition 2.3 holds as soon as the boundary terms vanish, i.e.

$$\liminf_{\epsilon_1,\epsilon_2\to 0}\int_{t_1}^{t_2}\left[a(F_t^{-1} - G_t^{-1})|F_t^{-1} - G_t^{-1}|^{p-2}\left(\frac{1}{\partial_u G_t^{-1}} - \frac{1}{\partial_u F_t^{-1}}\right)\right]_{\epsilon_1}^{1-\epsilon_2}\mathrm{d}t = 0.$$

However, we were not able to provide a rigorous account of this statement. In Appendix B, we use a different expression of $W_p^p(F_t,G_t)$ in terms of F_t, G_t to compute the time derivative of the flow.

Remark 2.5. Although the setting is different, the result of Proposition 2.1 obtained by a probabilistic approximation is comparable to the result of Carrillo, Di Francesco and Lattanzio [9, Theorem 5.1], the proof of which relies on the deterministic operator splitting method. In Lemma 2.2, the nonlinear evolution equation for the pseudo-inverse of the solution to the Cauchy problem generalizes in a rigorous way the proposed extensions of the work by Carrillo and Toscani [10, Section 3]. In Carrillo, Gualdani and Toscani [12], the time derivative of the flow of the Wasserstein distance between two solutions is computed by the method described in Remark 2.4 for the case of compactly supported solutions at all times, therefore the boundary terms necessarily vanish in the integration by parts.

The same method is also applied by Alfonsi, Jourdain and Kohatsu-Higa [1], where the authors get rid of the boundary terms using Gaussian estimates on the density. As Lemma B.2 shows, under the uniform ellipticity condition (D3), such estimates still hold in our case as soon as the tails of F_0 or G_0 are not heavier than Gaussian. Since we are willing to use Proposition 2.3 to compare F_t with the stationary solution F_∞, we would therefore need the tails of F_∞ not to be heavier than Gaussian. But according to Remark 3.5 below, under the condition (D3), the tails of F_∞ cannot be lighter than exponential.

3. Convergence to equilibrium

This section is divided into three parts. In Subsection 3.1, we solve the stationary equation. In Subsection 3.2, we prove the convergence of solutions to stationary solutions. In Subsection 3.3, we discuss the (lack of) rate of convergence to equilibrium.

3.1. The stationary equation.

We recall that the *stationary equation* is the following:

$$\frac{1}{2}\partial_x^2\big(A(F_\infty(x))\big) - \partial_x\big(B(F_\infty(x))\big) = 0 \tag{20}$$

As mentionned in the introduction, the stationary solutions for the Cauchy problem are the cumulative distribution functions, with a finite first order moment, solving (20) in the sense of distributions. In Proposition 3.1, we solve the stationary equation, and we give a criterion for integrability in Corollary 3.4.

Proposition 3.1. *Under the nondegeneracy condition (D1), a necessary and sufficient condition for the existence of cumulative distribution functions solving the stationary equation is $B(1) = 0$, $B(u) \geq 0$ and the local integrability of the function $a/2B$ on $(0,1)$. Then all the solutions are continuous.*

If in addition $B(u) > 0$ for all $u \in (0,1)$, which corresponds to the equilibrium condition (E1), then F_∞ is a solution if and only if there exists $\bar{x} \in \mathbb{R}$ such that for all $x \in \mathbb{R}$, $F_\infty(x) = \Psi^{-1}(x+\bar{x})$, where the function Ψ is defined by

$$\forall u \in (0,1), \qquad \Psi(u) := \int_{1/2}^{u} \frac{a(v)}{2B(v)}\mathrm{d}v. \tag{21}$$

In this case, $\bar{x} = \Psi(F_\infty(0))$.

Proof of Proposition 3.1. We first prove the necessary condition. Let F be a cumulative distribution function (we shall write F instead of F_∞ in the proof), solving (20) in the sense of distributions. Then there exists $c \in \mathbb{R}$ such that the function $x \mapsto (1/2)A(F(x))$ is absolutely continuous with respect to the Lebesgue measure, with density $B(F(x)) + c$. Since by the condition (D1), A is increasing, then F is continuous. Hence, $B(F(x)) + c$ is a nonnegative, continuous and integrable function, so that taking the limit $x \to -\infty$ yields $B(0) + c = c = 0$. We deduce $B(u) \geq 0$ and $B(1) = 0$ by taking the limit $x \to +\infty$.

It remains to prove that $a(u)/2B(u)$ is locally integrable in $(0,1)$. We use the convention that $a(u)/2B(u) = 0$ when $a(u) = B(u) = 0$, and we define $U := \{u \in (0,1) : B(u) = 0 \text{ and } a(u) \neq 0\}$. Let $0 < \alpha < \beta < 1$ and $x_- := F^{-1}(\alpha)$, $x_+ := F^{-1}(\beta)$. We denote by $\mathrm{d}F(x)$ the Stieltjes measure associated with the continuous function F of finite variation. Then by the chain rule formula [35, (4.6) p. 6], the Radon measure $(1/2)a(F(x))\mathrm{d}F(x)$ has the density $B(F(x))$ with respect to the

Lebesgue measure. Therefore $\mathrm{d}F(x)$-a.e., $F(x) \notin U$. By the change of variable formula [35, (4.9) p. 8],

$$\int_{x_-}^{x_+} \frac{a(F(x))}{2B(F(x))} \mathbf{1}_{\{F(x)\notin U\}} \mathrm{d}F(x) = \int_{x_-}^{x_+} \frac{a(F(x))}{2B(F(x))} \mathrm{d}F(x) = \int_{\alpha}^{\beta} \frac{a(u)}{2B(u)} \mathrm{d}u,$$

and the left-hand side is bounded by $x_+ - x_- < +\infty$.

We now prove that the condition is sufficient. For $u \in (0,1)$, let us define $\Psi(u)$ as in (21). Then Ψ is absolutely continuous and since A is increasing, so is Ψ. Thus Ψ^{-1} is continuous, with finite variation. In case Ψ has a finite limit in 1 (resp. 0), we write $\Psi^{-1}(x) = 1$ (resp. 0) for $x \in [\Psi(1), +\infty)$ (resp. $(-\infty, \Psi(0)]$) so that Ψ^{-1} is a cumulative distribution function on $\mathbb{R}$. We now check that Ψ^{-1} is a solution, in the sense of distributions, of the stationary equation. Let $\phi \in C_{\mathrm{c}}^{\infty}(\mathbb{R})$ be a test function. Then by the integration by parts formula [35, (4.5) p. 6] and the chain rule formula applied to $A(\Psi^{-1}(x))$,

$$-\frac{1}{2}\int_{\mathbb{R}} \phi'(x) A(\Psi^{-1}(x)) \mathrm{d}x = \frac{1}{2}\int_{\mathbb{R}} \phi(x) a(\Psi^{-1}(x)) \mathrm{d}\Psi^{-1}(x).$$

By the change of variable formula and the definition of Ψ,

$$\frac{1}{2}\int_{\mathbb{R}} \phi(x) a(\Psi^{-1}(x)) \mathrm{d}\Psi^{-1}(x) = \frac{1}{2}\int_0^1 \phi(\Psi(u)) a(u) \mathrm{d}u = \int_0^1 \phi(\Psi(u)) B(u) \Psi'(u) \mathrm{d}u,$$

and performing a new change of variables in the last member above yields

$$-\frac{1}{2}\int_{\mathbb{R}} \phi'(x) A(\Psi^{-1}(x)) \mathrm{d}x = \int_{\mathbb{R}} \phi(x) B(\Psi^{-1}(x)) \mathrm{d}x,$$

i.e. Ψ^{-1} is a solution in the sense of distributions of the stationary equation.

We finally assume that $B(u) > 0$ for all $u \in (0,1)$ and prove that F is a solution if and only if it is a translation of Ψ^{-1}. If F is a solution, since $B(u) > 0$ then Ψ is C^1 on $(0,1)$ so that the chain rule formula gives

$$\mathrm{d}(\Psi(F(x))) = \Psi'(F(x)) \mathrm{d}F(x) = \frac{a(F(x))}{2B(F(x))} \mathrm{d}F(x) = \mathrm{d}x,$$

where the last equality holds due to $B(u) > 0$. Then $\Psi(F(x)) - \Psi(F(0)) = x$ and $F(x) = \Psi^{-1}(x + \bar{x})$ with $\bar{x} = \Psi(F(0))$. Reciprocally it is immediate that all the translations of Ψ^{-1} solve the stationary equation. □

Remark 3.2. When the condition that $B(u) > 0$ on $(0,1)$ is not fulfilled, then one can exhibit solutions that are not translations of each other. For instance, let $a(u) = u(1-u)|u - 1/2|^{3/2}$ and $B(u) = u(1-u)(u-1/2)^2$. Then a is continuous on $[0,1]$, its antiderivative satisfies (D1), B is C^1 on $[0,1]$, $B(0) = B(1) = 0$ and $B(u) \geq 0$. Besides, $\Psi(u) = \mathrm{sgn}(u - 1/2)|u - 1/2|^{1/2}$. For all $h \geq 0$, let us define

$$F_{\infty,h}(x) := \begin{cases} 0 & x < -1/\sqrt{2}, \\ 1/2 - x^2 & -1/\sqrt{2} \leq x < 0, \\ 1/2 & 0 \leq x < h, \\ 1/2 + (x-h)^2 & h \leq x < h + 1/\sqrt{2}, \\ 1 & x \geq h + 1/\sqrt{2}. \end{cases}$$

Then $F_{\infty,0} = \Psi^{-1}$ and for all $h > 0$, $F_{\infty,h}$ solves the stationary equation although it is not a translation of Ψ^{-1}.

In order to apply the results of Section 2, we need criteria ensuring the existence of a first order moment as well as the classical regularity for a stationary solution F_∞. They come as corollaries to Proposition 3.1.

Corollary 3.3. *Under the nondegeneracy condition (D2), the regularity condition (R1) and the equilibrium condition (E1), all the stationary solutions F_∞ are C^2 on $\mathbb{R}$.*

Proof. By Proposition 3.1 and the condition (E1), it is enough to prove that Ψ^{-1} is C^2 on $\mathbb{R}$, which follows from the inverse function theorem, since Ψ is C^2 on $(0,1)$ by (R1) and (E1), and $\Psi'(u) > 0$ for all $u \in (0,1)$ by (D2). □

Corollary 3.4. *Under the nondegeneracy condition (D1) and the equilibrium condition (E1), the solutions to the stationary equation have a finite first order moment if and only if (E2) holds.*

Proof. According to Proposition 3.1, it is enough to prove the statement for the solution Ψ^{-1}, the first order moment of which is given by

$$\int_{1/2}^{1} \Psi(u)\mathrm{d}u - \int_{0}^{1/2} \Psi(u)\mathrm{d}u = \int_{0}^{1/2} \frac{a(u)u}{2B(u)}\mathrm{d}u + \int_{1/2}^{1} \frac{a(u)(1-u)}{2B(u)}\mathrm{d}u$$

by the Fubini-Tonelli theorem. The finiteness of the right-hand side is the condition (E2). □

Remark 3.5. In the case of the viscous conservation law and under the condition (E1), it is sufficient that $b(0) > 0$ and $b(1) < 0$ for (E2) to hold, and in this case the probability distributions derivated from the stationary solutions have exponential tails and satisfy a Poincaré inequality (see [29, Lemma 2.1]). In the general case of the stationary equation (20), and under the equilibrium condition (E1), these results extend as follows.

- Under the uniform ellipticity condition (D3), if $b(0) > 0$ and $b(1) < 0$ then it is clear from the expression of Ψ that the stationary solutions still have exponential tails, and they consequently satisfy a Poincaré inequality. If $b(0) = 0$ (resp. $b(1) = 0$), then the left (resp. the right) tail is heavier than exponential.
- Under the nondegeneracy condition (D1), as soon as the cumulative distribution function Ψ^{-1} admits a positive density p, then it satisfies a Poincaré inequality if and only if it satisfies the Hardy criterion (see [2, Theorem 6.2.2, p. 99]), namely

$$\sup_{x\geq 0} \int_{x}^{+\infty} p(y)\mathrm{d}y \int_{0}^{x} \frac{\mathrm{d}y}{p(y)} < +\infty, \quad \sup_{x\leq 0} \int_{-\infty}^{x} p(y)\mathrm{d}y \int_{x}^{0} \frac{\mathrm{d}y}{p(y)} < +\infty.$$

  Letting $y = \Psi(v)$, we rewrite

$$\int_{0}^{x} \frac{\mathrm{d}y}{p(y)} = \int_{1/2}^{\Psi^{-1}(x)} (\Psi'(v))^2 \mathrm{d}v,$$

  so that that the stationary solutions satisfy a Poincaré inequality if and only if

$$\sup_{u\geq 1/2} (1-u) \int_{1/2}^{u} \left(\frac{a(v)}{2B(v)}\right)^2 \mathrm{d}v < +\infty, \quad \sup_{u\leq 1/2} u \int_{u}^{1/2} \left(\frac{a(v)}{2B(v)}\right)^2 \mathrm{d}v < +\infty.$$

3.2. **Convergence in Wasserstein distance.** We now state the main result of the article, namely the convergence to equilibrium of the probabilistic solutions in Wasserstein distance.

Theorem 3.6. *Let us assume that:*

- *the coefficients of the Cauchy problem (1) satisfy the uniform ellipticity condition (D3), the regularity condition (R2) and the equlibrium conditions (E1) and (E2);*
- *the probability distribution m has a finite first order moment;*
- $W_2(H * m, \Psi^{-1}) < +\infty$.

Let F be the probabilistic solution the Cauchy problem (1) with initial condition $H * m$. Then there exists a unique stationary solution F_∞ such that F_0 and F_∞ have the same expectation, and for all $p \geq 2$ such that $W_p(H * m, \Psi^{-1}) < +\infty$,

$$\forall 1 \leq q < p, \qquad \lim_{t\to+\infty} W_q(F_t, F_\infty) = 0.$$

Proof. The proof is in 6 steps.

Step 1. We first prove the existence and uniqueness of a stationary solution F_∞ such that F_0 and F_∞ have the same expectation. By the condition (E2), the integral $\int_0^1 \Psi(u)\mathrm{d}u$ is defined. Owing to the condition (E1) and according to Proposition 3.1, the stationary solutions are the

functions of the form $F_\infty(x) = \Psi^{-1}(x+\bar{x})$, $\bar{x} \in \mathbb{R}$. By Corollary 3.4, the expectation of such a cumulative distribution function exists and it is given by

$$\int_0^1 F_\infty^{-1}(u)\mathrm{d}u = \int_0^1 \Psi(u)\mathrm{d}u - \bar{x}.$$

Thus, the unique stationary solution with the same expectation as F_0 is given by $F_\infty(x) = \Psi^{-1}(x+\bar{x})$, with $\bar{x} = \int_0^1(\Psi(u) - F_0^{-1}(u))\mathrm{d}u$.

Step 2. We now introduce a smooth approximation of the initial condition F_0 in order to use Lemma 1.7. Let ζ be a C^∞ probability density on $\mathbb{R}$ with compact support and such that $\int_{\mathbb{R}} x\zeta(x)\mathrm{d}x = 0$. For all $\alpha > 0$ and $x \in \mathbb{R}$, we define $\zeta_\alpha(x) := \alpha^{-1}\zeta(\alpha^{-1}x)$ and $F_0^\alpha(x) := F_0 * \zeta_\alpha(x)$. Then F_0^α is C^∞ on $\mathbb{R}$ and it is the cumulative distribution function of $X_0 + \alpha Z$, where X_0 has distribution m, Z has density ζ and X_0 and Z are independent. As a consequence, F_0^α has a finite first order moment and it has the same expectation as F_0 and F_∞ due to Step 1. By Lemma 1.7, the probabilistic solution F^α to the Cauchy problem (1) with initial condition F_0^α has the classical regularity and for all finite $T > 0$, it belongs to $C_{\mathrm{b}}^{1,2}([0,T]\times\mathbb{R})$.

Using the obvious coupling $(X_0, X_0+\alpha Z)$ of F_0 and F_0^α, we note that for all $q \geq 1$, $W_q(F_0, F_0^\alpha) \leq \alpha\mathbb{E}(|Z|^q)^{1/q} < +\infty$. This leads to the following remarks:

- As F_∞ is a translation of Ψ^{-1}, by the triangle inequality, $W_q(H*m, \Psi^{-1})$ and $W_q(F_0^\alpha, F_\infty)$ are simultaneously finite or infinite.
- Using the triangle inequality again and Proposition 2.1, we get

$$W_q(F_t, F_\infty) \leq W_q(F_t, F_t^\alpha) + W_q(F_t^\alpha, F_\infty) \leq \alpha\mathbb{E}(|Z|^q)^{1/q} + W_q(F_t^\alpha, F_\infty).$$

  Hence, as soon as, for all $\alpha > 0$, $\limsup_{t\to+\infty} W_q(F_t^\alpha, F_\infty) = 0$, then taking α arbitrarily small yields $\lim_{t\to+\infty} W_q(F_t, F_\infty) = 0$.

We now fix $\alpha > 0$. The remaining steps are dedicated to the proof of the fact that, for all $p \geq 2$ such that $W_p(H*m, \Psi^{-1}) < +\infty$, for all $1 \leq q < p$, $\lim_{t\to+\infty} W_q(F_t^\alpha, F_\infty) = 0$.

Step 3. We prove that for all $t \geq 0$, the expectation of F_t^α remains constant. By Corollary 1.13, for all $t \geq 0$, F_t^α is the marginal cumulative distribution function of the nonlinear diffusion process X^α solution to (4) with initial condition having cumulative distribution function F_0^α. Since σ and b are bounded,

$$\forall t \geq 0, \qquad \mathbb{E}(X_t^\alpha) = \mathbb{E}(X_0^\alpha) + \int_0^t \mathbb{E}[b(F_s^\alpha(X_s^\alpha))]\mathrm{d}s.$$

But for all $s > 0$, F_s^α is continuous so that $\mathbb{E}[b(F_s^\alpha(X_s^\alpha))] = \int_0^1 b(u)\mathrm{d}u = B(1)$. By (E1), we conclude that $\mathbb{E}(X_t^\alpha) = \mathbb{E}(X_0^\alpha)$.

Step 4. We now describe the evolution of the Wasserstein distance $W_2(F_t^\alpha, F_\infty)$. We are willing to use Proposition 2.3, therefore we need to check that F^α and F_∞ satisfy the assumptions of Lemma 2.2. It is the case for F^α thanks to Lemma 1.7. The stationary solution F_∞ has a finite first order moment owing to the condition (E2) and Corollary 3.4, it is C^2 on $\mathbb{R}$ by the condition (R2) and Corollary 3.3, and from the definition of Ψ^{-1} and condition (D3) it follows that the derivative of F_∞ is bounded by $2||B||_\infty/\underline{a}$.

Moreover, by the assumption that $W_2(H*m, \Psi^{-1}) < +\infty$ and Step 2, $W_2(F_0^\alpha, F_\infty) < +\infty$; and since both F_0^α and F_∞ have a finite first order moment, $|x|(\mathrm{tl}(F_0^\alpha, x) + \mathrm{tl}(F_\infty, x))$ vanishes when $x \to \pm\infty$. Thus, Proposition 2.3 applies to F^α and F_∞ with $p = 2$ and yields, for all $0 < t_1 < t_2$,

$$W_2^2(F_{t_2}^\alpha, F_\infty) - W_2^2(F_{t_1}^\alpha, F_\infty) = -\int_{t_1}^{t_2}\int_0^1 a(u)\frac{\left(\partial_u(F_t^\alpha)^{-1}(u) - \partial_u F_\infty^{-1}(u)\right)^2}{\partial_u(F_t^\alpha)^{-1}(u)\partial_u F_\infty^{-1}(u)}\mathrm{d}u\mathrm{d}t \leq 0.$$

Using the uniform ellipticity condition (D3), we can then assert that

$$\liminf_{t\to+\infty}\int_0^1 \frac{\left(\partial_u(F_t^\alpha)^{-1}(u) - \partial_u F_\infty^{-1}(u)\right)^2}{\partial_u(F_t^\alpha)^{-1}(u)\partial_u F_\infty^{-1}(u)}\mathrm{d}u = 0,$$

and extract a sequence $(t_n)_{n\geq 1}$ growing to $+\infty$ such that the integral above goes to 0 along $(t_n)_{n\geq 1}$. Let us prove that for all $u \in (0,1)$,

$$\lim_{n\to+\infty} \int_{1/2}^{u} |\partial_u (F_{t_n}^\alpha)^{-1}(v) - \partial_u F_\infty^{-1}(v)| \mathrm{d}v = 0. \tag{22}$$

Let $0 < \epsilon < 1/2$. By the Cauchy-Schwarz inequality and the condition (E1),

(23)

$$\begin{aligned}\int_\epsilon^{1-\epsilon} |\partial_u (F_{t_n}^\alpha)^{-1} - \partial_u F_\infty^{-1}| \mathrm{d}v &\leq \left(\int_\epsilon^{1-\epsilon} \partial_u (F_{t_n}^\alpha)^{-1} \partial_u F_\infty^{-1} \mathrm{d}v \int_\epsilon^{1-\epsilon} \frac{\left(\partial_u (F_{t_n}^\alpha)^{-1} - \partial_u F_\infty^{-1}\right)^2}{\partial_u (F_{t_n}^\alpha)^{-1} \partial_u F_\infty^{-1}} \mathrm{d}v \right)^{1/2} \\ &\leq \left(\sup_{v\in[\epsilon,1-\epsilon]} \frac{a(v)}{2B(v)} \int_\epsilon^{1-\epsilon} \partial_u (F_{t_n}^\alpha)^{-1} \mathrm{d}v \int_\epsilon^{1-\epsilon} \frac{\left(\partial_u (F_{t_n}^\alpha)^{-1} - \partial_u F_\infty^{-1}\right)^2}{\partial_u (F_{t_n}^\alpha)^{-1} \partial_u F_\infty^{-1}} \mathrm{d}v \right)^{1/2}.\end{aligned}$$

The first integral can be bounded uniformly in n as follows:

$$\begin{aligned}\int_\epsilon^{1-\epsilon} \partial_u (F_{t_n}^\alpha)^{-1}(v) \mathrm{d}v &= (F_{t_n}^\alpha)^{-1}(1-\epsilon) - (F_{t_n}^\alpha)^{-1}(\epsilon) \\ &\leq \frac{2}{\epsilon} \int_0^{\epsilon/2} \left((F_{t_n}^\alpha)^{-1}(1-\epsilon+v) - (F_{t_n}^\alpha)^{-1}(\epsilon - v) \right) \mathrm{d}v \\ &\leq \frac{2}{\epsilon} \left(\int_0^{\epsilon/2} \left(F_\infty^{-1}(1-\epsilon+v) - F_\infty^{-1}(\epsilon-v) \right) \mathrm{d}v + \int_0^1 |(F_{t_n}^\alpha)^{-1}(v) - F_\infty^{-1}(v)| \mathrm{d}v \right) \\ &\leq \frac{2}{\epsilon} \left(\int_0^{\epsilon/2} \left(F_\infty^{-1}(1-\epsilon+v) - F_\infty^{-1}(\epsilon-v) \right) \mathrm{d}v + \int_0^1 |(F_0^\alpha)^{-1}(v) - F_\infty^{-1}(v)| \mathrm{d}v \right),\end{aligned}$$

where the last inequality is due to Proposition 2.1. We deduce that the right-hand side of (23) goes to 0, so that taking $\epsilon \leq u \wedge (1-u)$ yields (22).

Step 5. We extract a subsequence of $(t_n)_{n\geq 1}$, that we still index by n for convenience, such that $\lim_{n\to+\infty} (F_{t_n}^\alpha)^{-1}(1/2) - F_\infty^{-1}(1/2) = \ell \in [-\infty, +\infty]$. Then using Step 4, for all $u \in (0,1)$ one has $(F_{t_n}^\alpha)^{-1}(u) - F_\infty^{-1}(u) \to \ell$. Besides, since by Proposition 2.1,

$$\sup_{t\geq 0} \int_0^1 |(F_t^\alpha)^{-1}(u) - F_\infty^{-1}(u)|^2 \mathrm{d}u = W_2^2(F_0^\alpha, F_\infty) < +\infty,$$

then the functions $(u \mapsto (F_{t_n}^\alpha)^{-1}(u) - F_\infty^{-1}(u))_{n\geq 1}$ are uniformly integrable. We deduce using Step 3 that

$$\ell = \lim_{n\to+\infty} \int_0^1 \left((F_{t_n}^\alpha)^{-1}(u) - F_\infty^{-1}(u) \right) \mathrm{d}u = 0.$$

Step 6. Let $p \geq 2$ such that $W_p(F_0^\alpha, \Psi^{-1}) < +\infty$. Then by Step 2, $W_p(F_0^\alpha, F_\infty) < +\infty$; therefore, for all $1 \leq q < p$, the functions $(u \mapsto |(F_{t_n}^\alpha)^{-1}(u) - F_\infty^{-1}(u)|^q)_{n\geq 1}$ are uniformly integrable, and using Step 5 we have

$$\lim_{n\to+\infty} \int_0^1 |(F_{t_n}^\alpha)^{-1}(u) - F_\infty^{-1}(u)|^q \mathrm{d}u = 0.$$

But according to Proposition 2.1, the flow $t \mapsto W_q(F_t^\alpha, F_\infty)$ is nonincreasing. As a consequence $\lim_{t\to+\infty} W_q(F_t^\alpha, F_\infty) = 0$ and the proof is completed by virtue of Step 2. □

3.3. **Rate of convergence.** We first recall the result of convergence to equilibrium stated in [29], where $A(u) = \sigma^2 u$ with $\sigma^2 > 0$. Then it is easily checked that the conditions (R1), (E1) and (E2) are satisfied if B is C^2 on $[0,1]$, with $B(1) = 0$, $b(0) > 0$, $b(1) < 0$ and $B(u) > 0$ on $(0,1)$. Then according to Remark 3.5, all the stationary solutions F_∞ admit a positive density p_∞ and satisfy a Poincaré inequality. Under these assumptions, we have the following convergence result.

Lemma 3.7. [29, Lemma 2.8] *There exist $\eta > 0$ and $c > 0$ depending on A and B such that for all cumulative distribution function F_0 with a finite first order moment, calling F_∞ the sationary solution with the same expectation as F_0, as soon as $\int (F_0 - F_\infty)^2/p_\infty \mathrm{d}x \leq \eta$ then*

$$\forall t \geq 0, \qquad \int_{\mathbb{R}} \frac{(F_t(x) - F_\infty(x))^2}{p_\infty(x)} \mathrm{d}x \leq \frac{\exp(-ct)}{c} \int_{\mathbb{R}} \frac{(F_0(x) - F_\infty(x))^2}{p_\infty(x)} \mathrm{d}x.$$

According to [28, Proposition 1.4], the quadratic Wasserstein distance between two probability distributions μ and ν on $\mathbb{R}$ such that μ admits a positive density p satisfies the inequality

$$W_2^2(\mu, \nu) \leq 4 \int_{\mathbb{R}} \frac{(H * \mu(x) - H * \nu(x))^2}{p(x)} \, \mathrm{d}x.$$

Hence, the convergence result of Lemma 3.7 can be translated in terms of the Wasserstein distance.

Corollary 3.8. *Under the assumptions of Lemma 3.7, as soon as $\int (F_0 - F_\infty)^2/p_\infty \mathrm{d}x$ is small enough, then $W_2(F_t, F_\infty)$ converges to 0 exponentially fast.*

Appendix A. Proof of Proposition 1.2

This appendix is dedicated to the proof of the first point of Proposition 1.2, which states that there is at most one weak solution to the Cauchy problem (1) in the set $\mathcal{F}(T)$, for all $T > 0$, possibly $T = +\infty$. The proof is adaptated from Wu, Zhao, Yin and Lin [42, Section 3.2] as well as Liu and Wang [32], who provide uniqueness of bounded weak solutions to the *initial-boundary value* problem, namely the Cauchy problem (1) in the strip $[0, T) \times (0, 1)$ with boundary conditions at $x = 0$ and $x = 1$. At an intuitive level, one can see our restriction to the set $\mathcal{F}(T)$ as some boundary conditions at $x = -\infty$ and $x = +\infty$.

We shall follow the so-called Holmgren's approach which consists in turning the proof of uniqueness for (1) into a proof of existence for an *adjoint problem*. Recall that we make the following nondegeneracy assumption:

(D1) The function A is increasing.

Let $T > 0$, possibly $T = +\infty$, and let F^1, $F^2 \in \mathcal{F}(T)$ such that for all $g \in C_c^\infty([0, T) \times \mathbb{R})$, both F^1 and F^2 satisfy (2). Then, for all $t \in [0, T)$, the function $F_t^2 - F_t^1$ is integrable on $\mathbb{R}$ and the function $(s, x) \mapsto F_s^2(x) - F_s^1(x)$ is integrable on $Q_t := (0, t) \times \mathbb{R}$. Therefore, for all $t \in [0, T)$ and for all $g \in C_c^\infty([0, T) \times \mathbb{R})$, (2) yields

$$\int_{Q_t} (F_s^2 - F_s^1) \left\{ \frac{1}{2} \tilde{A} \partial_x^2 g + \tilde{B} \partial_x g + \partial_s g \right\} \mathrm{d}s \mathrm{d}x = \int_{\mathbb{R}} (F_t^2(x) - F_t^1(x)) g(t, x) \mathrm{d}x; \tag{24}$$

where

$$\tilde{A}(s, x) = \int_0^1 a\left((1 - \theta) F_s^1(x) + \theta F_s^2(x)\right) \mathrm{d}\theta,$$

and

$$\tilde{B}(s, x) = \int_0^1 b\left((1 - \theta) F_s^1(x) + \theta F_s^2(x)\right) \mathrm{d}\theta.$$

Remark A.1. For all $t \in [0, T)$, by a classical regularization argument the integral equality (24) holds true for all function g in the space $C_b^{1,2}([0, t] \times \mathbb{R})$ of real-valued $C^{1,2}$ functions bounded together with their derivatives.

Let $f \in C_c^\infty([0, T) \times \mathbb{R})$. Then there exists $t \in [0, T)$ such that $\operatorname{Supp} f \subset [0, t) \times \mathbb{R}$. Let us introduce the *adjoint problem* to (1) as

$$\begin{cases} \frac{1}{2} \tilde{A} \partial_x^2 g + \tilde{B} \partial_x g + \partial_s g = f & (s, x) \in [0, t) \times \mathbb{R}, \\ g(t, x) = 0 & x \in \mathbb{R}. \end{cases} \tag{25}$$

The coefficients $\tilde{A}$ and $\tilde{B}$ may not be smooth enough to allow the adjoint problem to admit classical solutions. Therefore we introduce a suitable approximation of (25). For small $\delta, \eta > 0$, let

$$\begin{aligned} G_\delta &:= \{(s, x) \in [0, t] \times \mathbb{R} : |F_s^1(x) - F_s^2(x)| < \delta\}, \\ F_\delta &:= \{(s, x) \in [0, t] \times \mathbb{R} : |F_s^1(x) - F_s^2(x)| \geq \delta\}, \end{aligned} \tag{26}$$

and let us define

$$\lambda_\eta^\delta(s,x) = \begin{cases} 0 & \text{on } G_\delta, \\ \left[\frac{1}{2}(\eta + \tilde{A}(s,x))\right]^{-1/2} \tilde{B}(s,x) & \text{on } F_\delta. \end{cases}$$

Since A is increasing and F^1, F^2 are bounded, there exist $L(\delta) > 0$ and $K(\delta) > 0$ independent of η such that

$$\begin{aligned} &\tilde{A}(s,x) \geq L(\delta) && (s,x) \in F_\delta, \\ &|\lambda_\eta^\delta(s,x)| \leq K(\delta) && (s,x) \in [0,t] \times \mathbb{R}. \end{aligned}$$

Let ξ be a C^∞ probability density on $\mathbb{R}^2$ such that $\operatorname{Supp} \xi \subset [-1,1] \times [-1,1]$. For all $\epsilon > 0$, let $\xi_\epsilon := \epsilon^{-2}\xi(\epsilon^{-1}s, \epsilon^{-1}x)$ and define $\tilde{A}_\epsilon = \tilde{A} * \xi_\epsilon$ and $\lambda_{\eta,\epsilon}^\delta = \lambda_\eta^\delta * \xi_\epsilon$. Then $\tilde{A}_\epsilon$ and $\lambda_{\eta,\epsilon}^\delta$ are C^∞ functions and all their derivatives are bounded on $[0,t] \times \mathbb{R}$. Besides,

$$\begin{aligned} &\lim_{\epsilon \to 0} \tilde{A}_\epsilon(s,x) = \tilde{A}(s,x) && \text{a.e. in } [0,t] \times \mathbb{R}, \\ &\lim_{\epsilon \to 0} \lambda_{\eta,\epsilon}^\delta(s,x) = \lambda_\eta^\delta(s,x) && \text{a.e. in } [0,t] \times \mathbb{R}, \\ &\tilde{A}_\epsilon(s,x) \leq C && (s,x) \in [0,t] \times \mathbb{R}, \\ &|\lambda_{\eta,\epsilon}^\delta(s,x)| \leq K(\delta) && (s,x) \in [0,t] \times \mathbb{R}, \end{aligned} \tag{27}$$

where C refers to a positive constant independent of ϵ, δ and η, and $K(\delta)$ refers to a positive constant depending only on δ. In the sequel, the values of C and $K(\delta)$ can change from one line to another.

We finally define

$$\tilde{B}_{\eta,\epsilon}^\delta(s,x) = \lambda_{\eta,\epsilon}^\delta(s,x) \left[\frac{1}{2}(\eta + \tilde{A}_\epsilon(s,x))\right]^{1/2},$$

and emphasize the fact that

$$||\tilde{B}_{\eta,\epsilon}^\delta||_\infty \leq K(\delta). \tag{28}$$

We are now able to introduce the *approximate adjoint problem*

$$\begin{cases} \frac{1}{2}(\eta + \tilde{A}_\epsilon)\partial_x^2 g + \tilde{B}_{\eta,\epsilon}^\delta \partial_x g + \partial_s g = f & (s,x) \in [0,t) \times \mathbb{R}, \\ g(t,x) = 0 & x \in \mathbb{R}. \end{cases} \tag{29}$$

The coefficients of the equation are bounded, globally Lipschitz continuous, the operator is uniformly parabolic, and the right-hand side f is continuous and bounded. Therefore the Cauchy problem (29) admits a unique classical bounded solution $g_{\eta,\epsilon}^\delta$ (see [30, p. 369]). Since the coefficients of the equation and f are C^∞ on $[0,t] \times \mathbb{R}$, then so is $g_{\eta,\epsilon}^\delta$ (see [18, p. 263]). Owing to the Feynman-Kac formula, $g_{\eta,\epsilon}^\delta$ has the following probabilistic representation:

$$\forall (s,x) \in [0,t) \times \mathbb{R}, \qquad g_{\eta,\epsilon}^\delta(s,x) = -\mathbb{E}\left[\int_s^t f(r, Z_r^{s,x}) \mathrm{d}r\right] \tag{30}$$

where, for a given standard Brownian motion W, $(Z_r^{s,x})_{r \in [0,t]}$ is the unique strong solution of the stochastic differential equation

$$Z_r^{s,x} = x + \int_s^r \tilde{B}_{\eta,\epsilon}^\delta(u, Z_u^{s,x}) \mathrm{d}u + \int_s^r (\eta + \tilde{A}_\epsilon(u, Z_u^{s,x}))^{1/2} \mathrm{d}W_u. \tag{31}$$

Lemma A.2. *The functions $g_{\eta,\epsilon}^\delta$, $\partial_x g_{\eta,\epsilon}^\delta$ and $\partial_x^2 g_{\eta,\epsilon}^\delta$ are such that:*

$$\sup_{[0,t] \times \mathbb{R}} |g_{\eta,\epsilon}^\delta(s,x)| \leq C, \tag{32}$$

$$\sup_{s \in [0,t]} \int_{\mathbb{R}} |g_{\eta,\epsilon}^\delta(s,x)| \mathrm{d}x \leq K(\delta), \tag{33}$$

$$\sup_{s \in [0,t]} |\partial_x g_{\eta,\epsilon}^\delta(s,x)| \leq \kappa(\epsilon,\delta,\eta) \exp(-x^2/\kappa(\epsilon,\delta,\eta)), \tag{34}$$

$$\sup_{s \in [0,t]} |\partial_x^2 g_{\eta,\epsilon}^\delta(s,x)| \leq \kappa(\epsilon,\delta,\eta) \exp(-x^2/\kappa(\epsilon,\delta,\eta)), \tag{35}$$

where the value of $\kappa(\epsilon,\delta,\eta)$ can change from one line to another.

Proof. The inequality (32) directly follows from the Feynman-Kac formula (30). Besides, since f has a compact support in $[0,t]\times\mathbb{R}$, say $\mathrm{Supp}\, f\subset[0,t]\times[x_-,x_+]$, one has

$$|g^\delta_{\eta,\epsilon}(s,x)|\leq ||f||_\infty\int_s^t\mathbb{P}(Z^{s,x}_r\in[x_-,x_+])\mathrm{d}r,$$

and for $x>x_+$,

$$\mathbb{P}(Z^{s,x}_r\in[x_-,x_+])\leq\mathbb{P}(x-Z^{s,x}_r\geq x-x_+)\leq\frac{\mathbb{E}[(x-Z^{s,x}_r)^2]}{(x-x_+)^2}.$$

Owing to (31) and (28),

$$\mathbb{E}[(x-Z^{s,x}_r)^2]\leq 2\left((r-s)^2||\tilde{B}^\delta_{\eta,\epsilon}||^2_\infty+(r-s)||\eta+\tilde{A}_\epsilon||_\infty\right)\leq K(\delta),$$

and similar arguments for $x<x_-$ yield (33).

In order to prove (34) and (35), let us take the derivative with respect to x of the problem (29). Then the function $\partial_x g^\delta_{\eta,\epsilon}$ is the unique classical solution of the Cauchy problem

$$\begin{cases}\frac{1}{2}(\eta+\tilde{A}_\epsilon)\partial^2_x g^1+\left(\frac{1}{2}\partial_x\tilde{A}_\epsilon+\tilde{B}^\delta_{\eta,\epsilon}\right)\partial_x g^1+\partial_x\tilde{B}^\delta_{\eta,\epsilon}g^1+\partial_s g^1=\partial_x f,\\ g^1(t,x)=0,\end{cases}\tag{36}$$

and the Feynman-Kac formula now writes

$$\begin{aligned}\partial_x g^\delta_{\eta,\epsilon}(s,x)&=-\mathbb{E}\left[\int_s^t\partial_x f(r,Z^{1,s,x}_r)\exp\left(\int_s^r\partial_x\tilde{B}^\delta_{\eta,\epsilon}(u,Z^{1,s,x}_u)\mathrm{d}u\right)\mathrm{d}r\right]\\&=-\int_s^t\int_{\mathbb{R}}\partial_x f(r,z)G^1(s,x;r,z)\mathrm{d}z\mathrm{d}r\end{aligned}\tag{37}$$

where $(Z^{1,s,x}_r)_{r\in[0,t]}$ is the associated diffusion process and $G^1(s,x;r,z)$ is the *fundamental solution* of (36). Following Friedman [18, p. 24], there exists some constant $\kappa>0$ depending on the coefficients of (36) (therefore, on ϵ, δ and η) such that, for all $s<r$,

$$|G^1(s,x;r,z)|\leq\frac{\kappa}{(r-s)^{1/2}}\exp\left(-\frac{(z-x)^2}{\kappa(r-s)}\right),\tag{38}$$

$$|\partial_x G^1(s,x;r,z)|\leq\frac{\kappa}{r-s}\exp\left(-\frac{(z-x)^2}{\kappa(r-s)}\right).\tag{39}$$

For $x>x_+$, (37) combined with (38) yields

$$\begin{aligned}|\partial_x g^\delta_{\eta,\epsilon}(s,x)|&\leq\int_s^t\int_{z=x_-}^{x_+}||\partial_x f||_\infty\frac{\kappa}{(r-s)^{1/2}}\exp\left(-\frac{(z-x_+)^2}{\kappa(r-s)}\right)\mathrm{d}z\mathrm{d}r\\&\leq\kappa||\partial_x f||_\infty(x_+-x_-)\exp\left(-\frac{(x-x_+)^2}{\kappa(t-s)}\right)\int_s^t\frac{\mathrm{d}r}{(r-s)^{1/2}}\\&\leq 2(t-s)^{1/2}\kappa||\partial_x f||_\infty(x_+-x_-)\exp\left(-\frac{(x-x_+)^2}{\kappa(t-s)}\right),\end{aligned}$$

and similar arguments for $x<x_-$ lead to (34). Likewise, for $x>x_+$, using (39) one gets

$$\begin{aligned}|\partial^2_x g^\delta_{\eta,\epsilon}(s,x)|&\leq\int_s^t\int_{z=x_-}^{x_+}||\partial_x f||_\infty\frac{\kappa}{r-s}\exp\left(-\frac{(z-x_+)^2}{\kappa(r-s)}\right)\mathrm{d}z\mathrm{d}r\\&\leq\kappa||\partial_x f||_\infty(x_+-x_-)\int_s^t\frac{1}{r-s}\exp\left(-\frac{(x-x_+)^2}{\kappa(r-s)}\right)\mathrm{d}r.\end{aligned}$$

Writing, thanks to the change of variable $v=(x-x_+)/(\kappa(r-s))^{1/2}$

$$\int_s^t\frac{1}{r-s}\exp\left(-\frac{(x-x_+)^2}{\kappa(r-s)}\right)\mathrm{d}r=\int_{(x-x_+)/(\kappa(t-s))^{1/2}}^{+\infty}\frac{2}{v}\exp(-v^2)\mathrm{d}v,$$

and using the fact that, as soon as $x \geq x_+ + (t\kappa)^{1/2}$,

$$\forall v \geq \frac{x - x_+}{(\kappa(t-s))^{1/2}}, \qquad \frac{1}{v} \leq \frac{\kappa(t-s)}{(x-x_+)^2} v,$$

we deduce that for $x \geq x_+ + (t\kappa)^{1/2}$,

$$\int_{(x-x_+)/(\kappa(t-s))^{1/2}}^{+\infty} \frac{2}{v} \exp(-v^2) \mathrm{d}v \leq \frac{\kappa(t-s)}{(x-x_+)^2} \exp\left(-\frac{(x-x_+)^2}{\kappa(t-s)}\right) \leq \exp\left(-\frac{(x-x_+)^2}{\kappa(t-s)}\right).$$

By similar arguments for $x < x_-$, one finally concludes to (35). □

By the definition of $g^\delta_{\eta,\epsilon}$,

$$\begin{aligned} \int_{[0,+\infty)\times\mathbb{R}} (F^2_s(x) - F^1_s(x)) f(s,x) \mathrm{d}s\mathrm{d}x &= \int_{Q_t} (F^2_s(x) - F^1_s(x)) f(s,x) \mathrm{d}s\mathrm{d}x \\ &= \int_{Q_t} (F^2_s - F^1_s) \left\{ \frac{1}{2}(\eta + \tilde{A}_\epsilon)\partial^2_x g^\delta_{\eta,\epsilon} + \tilde{B}^\delta_{\eta,\epsilon}\partial_x g^\delta_{\eta,\epsilon} + \partial_s g^\delta_{\eta,\epsilon} \right\} \mathrm{d}s\mathrm{d}x. \end{aligned} \tag{40}$$

It follows from the boundedness of $\tilde{A}_\epsilon$ and $\tilde{B}^\delta_{\eta,\epsilon}$ and from Lemma A.2 that

$$\sup_{s\in[0,t]} |\partial_s g^\delta_{\eta,\epsilon}(s,x)| \leq \kappa(\epsilon,\delta,\eta) \exp(-\kappa(\epsilon,\delta,\eta)x^2).$$

Consequently, $g^\delta_{\eta,\epsilon} \in C^{1,2}_{\mathrm{b}}([0,t]\times\mathbb{R})$, therefore due to remark A.1,

$$\int_{Q_t} (F^2_s - F^1_s) \left\{ \frac{1}{2}\tilde{A}\partial^2_x g^\delta_{\eta,\epsilon} + \tilde{B}\partial_x g^\delta_{\eta,\epsilon} + \partial_s g^\delta_{\eta,\epsilon} \right\} \mathrm{d}s\mathrm{d}x = 0. \tag{41}$$

As a conclusion, subtracting (41) to (40),

$$\begin{aligned} &\int_{[0,+\infty)\times\mathbb{R}} (F^2_s(x) - F^1_s(x)) f(s,x) \mathrm{d}s\mathrm{d}x \\ &\quad = \int_{Q_t} (F^2_s - F^1_s) \left\{ \frac{1}{2}(\eta + \tilde{A}_\epsilon - \tilde{A})\partial^2_x g^\delta_{\eta,\epsilon} + (\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon} \right\} \mathrm{d}s\mathrm{d}x. \end{aligned} \tag{42}$$

We now have to prove that the right-hand side of (42) goes to 0 as $\epsilon, \delta, \eta \to 0$. In this purpose, we closely follow the line of [32]. In particular, the proofs of our Lemmas A.3 and A.4 are nothing but transcriptions of the proofs of Lemmas 1 and 3 in [32] to the framework of an unbounded domain Q_t and weak solutions in $\mathcal{F}(T)$. Then the estimates (32)–(35) ensure that the computations still make sense.

Recall that C refers to a positive constant that does not depend on ϵ, η or δ.

Lemma A.3. [32, Lemma 1] *The functions $\partial_x g^\delta_{\eta,\epsilon}$ and $\partial^2_x g^\delta_{\eta,\epsilon}$ are such that:*

$$\int_{Q_t} \frac{1}{2}(\eta + \tilde{A}_\epsilon)(\partial^2_x g^\delta_{\eta,\epsilon})^2 \mathrm{d}s\mathrm{d}x \leq \frac{K(\delta)}{\eta} + C, \tag{43}$$

$$\int_{Q_t} (\partial_x g^\delta_{\eta,\epsilon})^2 \mathrm{d}s\mathrm{d}x \leq \frac{K(\delta)}{\eta} + C. \tag{44}$$

Lemma A.4. [32, Lemma 3] *The function $\partial_x g^\delta_{\eta,\epsilon}$ is such that:*

$$\sup_{s\in[0,t]} \int_{\mathbb{R}} |\partial_x g^\delta_{\eta,\epsilon}(s,x)| \mathrm{d}x \leq C. \tag{45}$$

The estimates of Lemmas A.2, A.3 and A.4 give sufficient uniformity over the derivatives of $g^\delta_{\eta,\epsilon}$ to conclude.

Proposition A.5. *The right-hand side of (42) is arbitrarily small when $\epsilon, \eta, \delta \to 0$.*

Proof. For lighter notations, let us denote $\bar{F}(s,x) = F^2_s(x) - F^1_s(x)$, and

$$I := \int_{Q_t} \bar{F}\frac{1}{2}(\tilde{A}_\epsilon - \tilde{A})\partial^2_x g^\delta_{\eta,\epsilon}\mathrm{d}s\mathrm{d}x + \int_{Q_t} \bar{F}\frac{\eta}{2}\partial^2_x g^\delta_{\eta,\epsilon}\mathrm{d}s\mathrm{d}x + \int_{Q_t} \bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon}\mathrm{d}s\mathrm{d}x$$
$$=: I_1 + I_2 + I_3.$$

Recall that since $F^1, F^2 \in \mathcal{F}(T)$, then $\bar{F} \in (L^1 \cap L^\infty)(Q_t)$.

Owing to the Cauchy-Schwarz inequality,

$$|I_1| \leq \left(\int_{Q_t} |\bar{F}|\frac{(\tilde{A}_\epsilon - \tilde{A})^2}{2(\eta + \tilde{A}_\epsilon)}\mathrm{d}s\mathrm{d}x\right)^{1/2} \left(\int_{Q_t} |\bar{F}|\frac{1}{2}(\eta + \tilde{A}_\epsilon)(\partial^2_x g^\delta_{\eta,\epsilon})^2\mathrm{d}s\mathrm{d}x\right)^{1/2}. \tag{46}$$

Using (27), by dominated convergence the first integral in the right-hand side of (46) goes to 0 as $\epsilon \to 0$ for fixed η and δ. According to Lemma A.3, the second integral in (46) is bounded by $C + K(\delta)/\eta$. Therefore, for fixed η and δ, $\lim_{\epsilon\to 0} I_1 = 0$.

Let $\alpha > 0$. Recalling the definition (26) of F_α and G_α, let us write

$$|I_2| \leq \frac{\eta}{2}\int_{F_\alpha} |\bar{F}\partial^2_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x + \frac{\eta}{2}\int_{G_\alpha} |\bar{F}\partial^2_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x.$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}\int_{F_\alpha} |\bar{F}\partial^2_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x &\leq \left(\int_{F_\alpha} |\bar{F}|\frac{\mathrm{d}s\mathrm{d}x}{(1/2)(\eta + \tilde{A}_\epsilon)}\right)^{1/2} \left(\int_{F_\alpha} |\bar{F}|\frac{1}{2}(\eta + \tilde{A}_\epsilon)(\partial^2_x g^\delta_{\eta,\epsilon})^2\mathrm{d}s\mathrm{d}x\right)^{1/2}\\ &\leq C\left(\sup_{F_\alpha}\frac{1}{(1/2)(\eta + \tilde{A}_\epsilon)}\right)^{1/2}\left(C + \frac{K(\delta)}{\eta}\right)^{1/2}\\ &\leq \frac{C}{L(\alpha)}\left(C + \frac{K(\delta)}{\eta}\right)^{1/2},\end{aligned}$$

where $L(\alpha)$ only depends on α. Likewise,

$$\begin{aligned}\int_{G_\alpha} |\bar{F}\partial^2_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x &\leq \left(\int_{G_\alpha} |\bar{F}|\frac{\mathrm{d}s\mathrm{d}x}{(1/2)(\eta + \tilde{A}_\epsilon)}\right)^{1/2} \left(\int_{G_\alpha} |\bar{F}|\frac{1}{2}(\eta + \tilde{A}_\epsilon)(\partial^2_x g^\delta_{\eta,\epsilon})^2\mathrm{d}s\mathrm{d}x\right)^{1/2}\\ &\leq \frac{C}{\eta^{1/2}}\left(\alpha\left(C + \frac{K(\delta)}{\eta}\right)\right)^{1/2};\end{aligned}$$

so that for fixed η, α,

$$\limsup_{\epsilon\to 0}|I_2| \leq C\eta\left(C + \frac{K(\delta)}{\eta}\right)^{1/2}\left(\frac{1}{L(\alpha)} + \left(\frac{\alpha}{\eta}\right)^{1/2}\right).$$

Finally, let us write

$$|I_3| \leq \int_{F_\delta} |\bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x + \int_{G_\delta} |\bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x.$$

We first deal with the integral on G_δ. On account of Lemma A.3, for given δ and η the family $(|\partial_x g^\delta_{\eta,\epsilon}|)_{\epsilon>0}$ is bounded in $L^2(Q_t)$. Therefore there exists a sequence $(\epsilon_k)_{k\geq 1}$ decreasing to 0, such that $|\partial_x g^\delta_{\eta,\epsilon_k}|$ converges weakly in $L^2(Q_t)$ to a function $h \geq 0$ when $k \to +\infty$. From now on, the convergence $\epsilon \to 0$ will always be understood along the sequence $(\epsilon_k)_{k\geq 1}$. According to Lemma A.4, for all compact subset $D \subset Q_t$,

$$\int_D h\mathrm{d}s\mathrm{d}x = \lim_{\epsilon\to 0}\int_D |\partial_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x \leq C$$

so that

$$\int_{Q_t} h\mathrm{d}s\mathrm{d}x \leq C.$$

Furthermore, on G_δ one has $\tilde{B}^\delta_{\eta,\epsilon} \to 0$ a.e. when $\epsilon \to 0$, and $||\tilde{B}^\delta_{\eta,\epsilon}||_\infty \leq K(\delta)$. Since $\bar{F} \in (L^1 \cap L^\infty)(Q_t) \subset L^2(Q_t)$, by dominated convergence one deduces that $\mathbf{1}_{G_\delta}|\bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})|$ converges strongly in $L^2(Q_t)$ to $\mathbf{1}_{G_\delta}|\bar{F}\tilde{B}|$. Finally,

$$\lim_{\epsilon\to 0}\int_{G_\delta} |\bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x = \int_{G_\delta} |\bar{F}\tilde{B}|h\mathrm{d}s\mathrm{d}x \leq C\delta.$$

We now turn to the integral on F_δ. By the Cauchy-Schwarz inequality,

$$\int_{F_\delta} |\bar{F}(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})\partial_x g^\delta_{\eta,\epsilon}|\mathrm{d}s\mathrm{d}x \leq \left(\int_{F_\delta} |\bar{F}|(\tilde{B}^\delta_{\eta,\epsilon} - \tilde{B})^2\mathrm{d}s\mathrm{d}x\right)^{1/2}\left(\int_{F_\delta} |\bar{F}|(\partial_x g^\delta_{\eta,\epsilon})^2\mathrm{d}s\mathrm{d}x\right)^{1/2}.$$

Owing to Lemma A.3 and the boundedness of $\bar{F}$,

$$\int_{F_\delta} |\bar{F}|(\partial_x g^\delta_{\eta,\epsilon})^2\mathrm{d}s\mathrm{d}x \leq C + \frac{K(\delta)}{\eta}$$

and by construction, $\tilde{B}^\delta_{\eta,\epsilon} \to \tilde{B}$ a.e. in F_δ when $\epsilon \to 0$, while $||\tilde{B}^\delta_{\eta,\epsilon}||_\infty \leq K(\delta)$. By dominated convergence, on concludes that for fixed η and δ, $\limsup_{\epsilon\to 0}|I_3| \leq C\delta$.

Combining the previous estimates, let us now write

$$\limsup_{\epsilon\to 0}|I| \leq C\eta\left(C + \frac{K(\delta)}{\eta}\right)^{1/2}\left(\frac{1}{L(\alpha)} + \left(\frac{\alpha}{\eta}\right)^{1/2}\right) + C\delta$$

and conclude by taking consecutively $\eta \to 0$, $\alpha \to 0$ and $\delta \to 0$. □

It follows from Proposition A.5 and (42) that

$$\int_{[0,T)\times\mathbb{R}} (F^2_s(x) - F^1_s(x))f(s,x)\mathrm{d}s\mathrm{d}x = 0.$$

Since f is arbitrary, $F^1_s(x) = F^2_s(x)$ a.e. in Q_t, and this holds for all $t \in [0,T)$. As a consequence, $F^1 = F^2$ in $\mathcal{F}(T)$.

Appendix B. Proof of Proposition 2.3

This appendix is dedicated to the computation of the time derivative of the flow $t \mapsto W_p^p(F_t, G_t)$ of the Wasserstein distance between two solutions F and G of the Cauchy problem (1) with respective initial conditions F_0 and G_0 and classical regularity. In Subsection B.1, we gather some tail estimates on the solutions F_t and G_t as well as their space derivatives. In Subsection B.2, we use a new expression of $W_p^p(F_t, G_t)$ in terms of F_t and G_t to prove Proposition 2.3.

B.1. Tail estimates.

In this subsection we are concerned with the asymptotic behaviour of $F_t(x)$ and $\partial_x F_t(x)$ when $|x|$ is large.

Lemma B.1. *Under the assumptions of Proposition 1.1, for all $t > 0$, there exists a finite constant $C(t) > 0$ such that the function $t \mapsto C(t)$ is nondecreasing and:*

$$\forall x \leq -C(t), \qquad \frac{1}{2}F_0(2x - C(t)) \leq F_t(x) \leq F_0\left(\frac{x}{2} + C(t)\right) + \exp\left(-\frac{x^2}{C(t)}\right),$$

$$\forall x \geq C(t), \qquad \frac{1}{2}[1 - F_0(2x + C(t))] \leq 1 - F_t(x) \leq 1 - F_0\left(\frac{x}{2} - C(t)\right) + \exp\left(-\frac{x^2}{C(t)}\right).$$

Proof. Fix a finite $T > 0$. We will use the process $(\bar{X}_t)_{t\in[0,T]}$ introduced in the proof of Lemma 1.10. Let $t \in [0,T]$ and define $C_1(t) := t||\bar{b}||_\infty$. If $x \leq -C_1(t)$, then

$$\begin{aligned}
F_t(x) &= \mathbb{P}\left(\bar{X}_0 + \int_0^t \bar{b}(r,\bar{X}_r)\mathrm{d}r + \int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x\right)\\
&\geq \mathbb{P}\left(\bar{X}_0 + \int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x - t||\bar{b}||_\infty\right)\\
&\geq \mathbb{P}\left(\bar{X}_0 \leq 2(x - t||\bar{b}||_\infty), \quad \left|\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r\right| \leq t||\bar{b}||_\infty - x\right)\\
&= \mathbb{P}\left(\left|\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r\right| \leq t||\bar{b}||_\infty - x \quad \Big| \quad \bar{X}_0 \leq 2x - C_2(t)\right)\mathbb{P}(\bar{X}_0 \leq 2x - C_2(t)),
\end{aligned}$$

where $C_2(t) := 2C_1(t) = 2t||\bar{b}||_\infty$. By Chebyshev's inequality,

$$\begin{aligned}
&\mathbb{P}\left(\left|\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r\right| \leq t||\bar{b}||_\infty - x \quad \Big| \quad \bar{X}_0 \leq 2x - C_2(t)\right)\\
&\qquad \geq 1 - \frac{\mathbb{E}\left(\left|\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r\right|^2 \quad \Big| \quad \bar{X}_0 \leq 2x - C_2(t)\right)}{(t||\bar{b}||_\infty - x)^2}\\
&\qquad \geq 1 - \frac{t||\bar{\sigma}^2||_\infty}{(t||\bar{b}||_\infty - x)^2},
\end{aligned}$$

and the right-hand side is larger than $1/2$ as soon as $x \leq -C_3(t) := -(2t||\bar{\sigma}^2||_\infty)^{1/2}$.

As far as the upper bound is concerned, for $x \leq -C_1(t)$,

$$\begin{aligned}
F_t(x) &= \mathbb{P}\left(\bar{X}_0 + \int_0^t \bar{b}(r,\bar{X}_r)\mathrm{d}r + \int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x\right)\\
&\leq \mathbb{P}\left(\bar{X}_0 + \int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x + t||\bar{b}||_\infty\right)\\
&= \int_{\mathbb{R}} \mathbb{P}\left(\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x - y + t||\bar{b}||_\infty \quad \Big| \quad \bar{X}_0 = y\right) m(\mathrm{d}y).
\end{aligned}$$

Let us fix $x \leq -C_1(t)$. For $x_0 \in \mathbb{R}$, let us split the integral in the right-hand side above in two parts, integrating respectively on $(-\infty, x_0]$ and $(x_0, +\infty)$. Then the first part can be bounded by $F_0(x_0)$, whereas for the second part the exponential Markov inequality yields, for all $\lambda > 0$,

$$\mathbb{P}\left(\int_0^t \bar{\sigma}(r,\bar{X}_r)\mathrm{d}\bar{W}_r \leq x - y + t||\bar{b}||_\infty \quad \Big| \quad \bar{X}_0 = y\right) \leq \exp\left(\lambda(x - y + t||\bar{b}||_\infty) + t\lambda^2\frac{||\bar{\sigma}^2||_\infty}{2}\right);$$

and finally, $F_t(x) \leq F_0(x_0) + \exp\left(\lambda(x - x_0 + t||\bar{b}||_\infty) + t\lambda^2||\bar{\sigma}^2||_\infty/2\right)$. As soon as $x_0 > x + t||\bar{b}||_\infty$, optimizing this expression in $\lambda > 0$ yields

$$F_t(x) \leq F_0(x_0) + \exp\left(-\frac{(x - x_0 + t||\bar{b}||_\infty)^2}{2t||\bar{\sigma}^2||_\infty}\right).$$

We now choose $x_0 = (x + t||\bar{b}||_\infty)/2$, then $x_0 < 0$ and $F_0(x_0) = F_0((x + t||\bar{b}||_\infty)/2) = F_0(x/2 + C_4(t))$ with $C_4(t) := t||\bar{b}||_\infty/2$. Moreover, for $x \leq -C_2(t) = -2t||\bar{b}||_\infty$,

$$\exp\left(-\frac{(x - x_0 + t||\bar{b}||_\infty)^2}{2t||\bar{\sigma}^2||_\infty}\right) = \exp\left(-x^2\frac{(1 + t||\bar{b}||_\infty/x)^2}{8t||\bar{\sigma}^2||_\infty}\right) \leq \exp\left(-\frac{x^2}{C_5(t)}\right),$$

where $C_5(t) := 32t||\bar{\sigma}^2||_\infty$. We get the first part of the lemma by taking $C(t)$ as the maximum of $C_1(t), \ldots, C_5(t)$, and the second part follows similarly. □

Assuming the classical regularity of F, we now derive some estimates on the probability density $\partial_x F_t$ from the celebrated Aronson inequalities on the fundamental solution of a parabolic equation with divergence form operator [3]. Due to the possible dispersion of the initial condition, our upper bound contains an extra tail term in addition to the classical Gaussian term.

Lemma B.2. *Under the assumptions of Lemma 2.2, for all $0 < t_1 < t_2$, there exists a positive constant $K > 0$, depending on t_1 and t_2, such that for all $(t, x) \in (t_1, t_2] \times \mathbb{R}$,*

$$\frac{1}{K(t-t_1)^{1/2}} \exp\left(-\frac{Kx^2}{t-t_1}\right) \leq \partial_x F_t(x) \leq \frac{K}{(t-t_1)^{1/2}} \left(\exp\left(-\frac{x^2}{K(t-t_1)}\right) + \mathrm{tl}\left(F_{t_1}, \frac{x}{2}\right)\right). \tag{47}$$

Proof. The assumptions of Lemma 2.2 together with Corollary 1.13 ensure that the nonlinear martingale problem of Subsection 1.2 has a unique weak solution P. We denote by X the associated nonlinear diffusion process. Let $\Gamma(s, y; t, \mathrm{d}x) = \mathbb{P}(X_t \in \mathrm{d}x | X_s = y)$ be the transition probability of X. The generator

$$Lf := \frac{1}{2} a(F_t(x)) \partial_x^2 f + b(F_t(x)) \partial_x f$$

is uniformly elliptic, and by the regularity assumptions on A, B and F, it rewrites in the divergence form

$$Lf = \frac{1}{2} \partial_x \big(a(F_t(x)) \partial_x f\big) - \left(\frac{1}{2} a'(F_t(x)) \partial_x F_t(x) - b(F_t(x))\right) \partial_x f.$$

Let $0 < t_1 < t_2$. The assumption of boundedness of $\partial_x F_t(x)$ on $[t_1, t_2] \times \mathbb{R}$ ensures that the coefficients of the latter form are bounded. Thus, owing to [3], $\Gamma(s, y; t, \mathrm{d}x)$ admits a density $g(s, y; t, x)$ and there exist some positive constants γ_i, κ_i, $i \in \{1, 2\}$, depending on t_1 and t_2, such that for all $t \in (t_1, t_2]$,

$$\frac{\kappa_1}{(t-t_1)^{1/2}} \exp\left(-\frac{(x-y)^2}{\gamma_1(t-t_1)}\right) \leq g(t_1, y; t, x) \leq \frac{\kappa_2}{(t-t_1)^{1/2}} \exp\left(-\frac{(x-y)^2}{\gamma_2(t-t_1)}\right).$$

Hence, $p_t^1(x) \leq \partial_x F_t(x) \leq p_t^2(x)$, where

$$p_t^i(x) := \frac{\kappa_i}{(t-t_1)^{1/2}} \int_{\mathbb{R}} \exp\left(-\frac{(x-y)^2}{\gamma_i(t-t_1)}\right) P_{t_1}(\mathrm{d}y).$$

For all $x \geq 0$,

$$\int_{-\infty}^{x/2} \exp\left(-\frac{(x-y)^2}{\gamma_2(t-t_1)}\right) P_{t_1}(\mathrm{d}y) \leq \exp\left(-\frac{x^2}{4\gamma_2(t-t_1)}\right)$$

while

$$\int_{x/2}^{+\infty} \exp\left(-\frac{(x-y)^2}{\gamma_2(t-t_1)}\right) P_{t_1}(\mathrm{d}y) \leq 1 - F_{t_1}(x/2).$$

Likewise, for $x \leq 0$,

$$\int_{x/2}^{+\infty} \exp\left(-\frac{(x-y)^2}{\gamma_2(t-t_1)}\right) P_{t_1}(\mathrm{d}y) \leq \exp\left(-\frac{x^2}{4\gamma_2(t-t_1)}\right)$$

while

$$\int_{-\infty}^{x/2} \exp\left(-\frac{(x-y)^2}{\gamma_2(t-t_1)}\right) P_{t_1}(\mathrm{d}y) \leq F_{t_1}(x/2),$$

so that the upper bound of (47) holds for any $K \geq \kappa_2 \vee (4\gamma_2)$.

As far as the lower bound is concerned, there exist $x_- < 0 < x_+$ such that $F_{t_1}(x_+) - F_{t_1}(x_-) \geq 1/2$. Then for all $x \in \mathbb{R}$,

$$\begin{aligned} p_t^1(x) &\geq \frac{\kappa_1}{(t-t_1)^{1/2}} \int_{x_-}^{x_+} \exp\left(-\frac{(x-y)^2}{\gamma_1(t-t_1)}\right) P_{t_1}(\mathrm{d}y) \\ &\geq \frac{\kappa_1}{2(t-t_1)^{1/2}} \exp\left(-\frac{((x_+ \vee -x_-) + |x|)^2}{\gamma_1(t-t_1)}\right). \end{aligned}$$

Now there exists $K > 0$ large enough, depending on κ_1, γ_1 and $x_+ \vee -x_-$, such that for all $x \in \mathbb{R}$,

$$\frac{\kappa_1}{2} \exp\left(-\frac{((x_+ \vee -x_-) + |x|)^2}{\gamma_1(t-t_1)}\right) \geq \frac{1}{K} \exp\left(-\frac{Kx^2}{t-t_1}\right),$$

which results in the lower bound of (47). □

B.2. **Proof of Proposition 2.3.** We first give a general formula for the Wasserstein distance $W_p(F,G)$ between two cumulative distribution functions F and G.

Lemma B.3. *Let F and G be two cumulative distribution functions on $\mathbb{R}$. Then, for all $p>1$,*

$$W_p^p(F,G) = p(p-1)\int_{\mathbb{R}^2} \mathbf{1}_{\{x<y\}}\left([G(x)-F(y)]^+ + [F(x)-G(y)]^+\right)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y. \tag{48}$$

Proof. Let us split the right-hand side of (48) into two symmetric integrals in F and G. Thanks to the Fubini-Tonelli theorem, the first integral writes:

$$\begin{aligned}
&\int_{\mathbb{R}^2} \mathbf{1}_{\{x<y\}}[G(x)-F(y)]^+ p(p-1)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y \\
&\quad= \int_{\mathbb{R}^2} \mathbf{1}_{\{x<y;G(x)\geq F(y)\}}\left(\int_0^1 \mathbf{1}_{\{F(y)<u\leq G(x)\}}\mathrm{d}u\right)p(p-1)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y \\
&\quad= \int_0^1\int_{\mathbb{R}^2} \mathbf{1}_{\{x<y;F(y)<u\leq G(x)\}}p(p-1)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y\mathrm{d}u.
\end{aligned}$$

By the definition of the pseudo-inverse functions F^{-1} and G^{-1}, note that for all $x,y\in\mathbb{R}$ and $u\in(0,1)$, $F(y)<u$ if and only if $y<F^{-1}(u)$ and $G(x)\geq u$ if and only if $x\geq G^{-1}(u)$. Thus, the right-hand side above rewrites

$$\begin{aligned}
&\int_0^1\int_{\mathbb{R}^2} \mathbf{1}_{\{G^{-1}(u)\leq x<y<F^{-1}(u)\}}p(p-1)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y\mathrm{d}u \\
&\quad= \int_0^1 \mathbf{1}_{\{G^{-1}(u)<F^{-1}(u)\}}\int_{x=G^{-1}(u)}^{F^{-1}(u)}\int_{y=x}^{F^{-1}(u)} p(p-1)(y-x)^{p-2}\mathrm{d}y\mathrm{d}x\mathrm{d}u \\
&\quad= \int_0^1 \mathbf{1}_{\{G^{-1}(u)<F^{-1}(u)\}}(F^{-1}(u)-G^{-1}(u))^p\mathrm{d}u,
\end{aligned}$$

and we conclude using the symmetry in F and G of the two integrals in the right-hand side of (48). □

We are now ready to complete the proof of Proposition 2.3.

Proof of Proposition 2.3. For all $t>0$, (48) yields

$$W_p^p(F_t,G_t) = p(p-1)(I(F_t,G_t)+I(G_t,F_t)), \tag{49}$$

where we define $I(F_t,G_t) := \int_{\mathbb{R}^2}\mathbf{1}_{\{x<y\}}[G_t(x)-F_t(y)]^+(y-x)^{p-2}\mathrm{d}x\mathrm{d}y$. The assumption that $W_p(F_0,G_0)<+\infty$ combined with Proposition 2.1 ensures that both $I(F_t,G_t)$ and $I(G_t,F_t)$ are finite.

For all $M\geq 0$, let us denote

$$I_M(F_t,G_t) := \int_{\mathbb{R}^2}\mathbf{1}_{\{-M\leq x<y\leq M\}}[G_t(x)-F_t(y)]^+(y-x)^{p-2}\mathrm{d}x\mathrm{d}y,$$

then by the monotone convergence theorem, $\lim_{M\to+\infty} I_M(F_t,G_t) = I(F_t,G_t)<+\infty$. Owing to the assumption of classical regularity on F and G, the function $t\mapsto I_M(F_t,G_t)$ is C^1 on $(0,+\infty)$ and for all $t>0$,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}I_M(F_t,G_t) &= \int_{\mathbb{R}^2}\mathbf{1}_{\{-M\leq x<y\leq M\}}\mathbf{1}_{\{G_t(x)\geq F_t(y)\}}(\partial_t G_t(x)-\partial_t F_t(y))(y-x)^{p-2}\mathrm{d}x\mathrm{d}y \\
&= \int_{\mathbb{R}^2}\mathbf{1}_{\{-M\leq x<y\leq M;G_t(x)\geq F_t(y)\}}\partial_x\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x)-B(G_t(x))\right)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y \\
&\quad-\int_{\mathbb{R}^2}\mathbf{1}_{\{-M\leq x<y\leq M;G_t(x)\geq F_t(y)\}}\partial_x\left(\frac{1}{2}a(F_t(y))\partial_x F_t(y)-B(F_t(y))\right)(y-x)^{p-2}\mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{50}$$

Let us define $\varphi_M^+(x) := M\wedge F_t^{-1}(G_t(x))$ and $\varphi_M^-(x) := (-M)\vee G_t^{-1}(F_t(x))$. Then the first integral in the right-hand side of (50) rewrites

$$\int_{y=-M}^{M}\mathbf{1}_{\{F_t(y)\leq G_t(y)\}}\int_{x=\varphi_M^-(y)}^{y}(y-x)^{p-2}\partial_x\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x)-B(G_t(x))\right)\mathrm{d}x\mathrm{d}y$$

and integrating by parts, we get

$$\begin{aligned}
&\int_{x=\varphi_M^-(y)}^{y} (y-x)^{p-2}\partial_x\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\right)\mathrm{d}x \\
&\quad = -(y-\varphi_M^-(y))^{p-2}\left(\frac{1}{2}a(G_t(-M)\vee F_t(y))\partial_x G_t(\varphi_M^-(y)) - B(G_t(-M)\vee F_t(y))\right) \\
&\qquad + \int_{x=\varphi_M^-(y)}^{y} (p-2)(y-x)^{p-3}\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\right)\mathrm{d}x.
\end{aligned}$$

Now

$$\begin{aligned}
&\int_{y=-M}^{M} \mathbf{1}_{\{F_t(y)\le G_t(y)\}}\int_{x=\varphi_M^-(y)}^{y} (p-2)(y-x)^{p-3}\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\right)\mathrm{d}x\mathrm{d}y \\
&\quad = \int_{x=-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\right)\int_{y=x}^{\varphi_M^+(x)} (p-2)(y-x)^{p-3}\mathrm{d}y\mathrm{d}x \\
&\quad = \int_{x=-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\left(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\right)(\varphi_M^+(x)-x)^{p-2}\mathrm{d}y\mathrm{d}x,
\end{aligned}$$

so that the first integral in the right-hand side of (50) finally writes

$$\begin{aligned}
&\int_{-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\Big\{(\varphi_M^+(x)-x)^{p-2}\Big(\frac{1}{2}a(G_t(x))\partial_x G_t(x) - B(G_t(x))\Big) \\
&- (x-\varphi_M^-(x))^{p-2}\Big(\frac{1}{2}a(G_t(-M)\vee F_t(x))\partial_x G_t(\varphi_M^-(x)) - B(G_t(-M)\vee F_t(x))\Big)\Big\}\mathrm{d}x
\end{aligned}$$

whereas the second integral similarly writes

$$\begin{aligned}
&\int_{-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\Big\{(\varphi_M^+(x)-x)^{p-2}\Big(\frac{1}{2}a(F_t(M)\wedge G_t(x))\partial_x F_t(\varphi_M^+(x)) - B(F_t(M)\wedge G_t(x))\Big) \\
&- (x-\varphi_M^-(x))^{p-2}\Big(\frac{1}{2}a(F_t(x))\partial_x F_t(x) - B(F_t(x))\Big)\Big\}\mathrm{d}x.
\end{aligned}$$

Hence, we deduce that for all $0 < t_1 \le t_2$,

$$I_M(F_{t_2}, G_{t_2}) - I_M(F_{t_1}, G_{t_1}) = \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} I_M(F_t, G_t)\mathrm{d}t = J_M^1 + J_M^2 + J_M^3,$$

where

$$\begin{aligned}
J_M^1 &:= \int_{t_1}^{t_2}\int_{-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\{(\varphi_M^+(x)-x)^{p-2}[B(F_t(M)\wedge G_t(x)) - B(G_t(x))] \\
&\qquad + (x-\varphi_M^-(x))^{p-2}[B(G_t(-M)\vee F_t(x)) - B(F_t(x))]\}\mathrm{d}x\mathrm{d}t; \\
J_M^2 &:= \frac{1}{2}\int_{t_1}^{t_2}\int_{-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\{(\varphi_M^+(x)-x)^{p-2}a(G_t(x))\partial_x G_t(x) \\
&\qquad + (x-\varphi_M^-(x))^{p-2}a(F_t(x))\partial_x F_t(x)\}\mathrm{d}x\mathrm{d}t; \\
J_M^3 &:= -\frac{1}{2}\int_{t_1}^{t_2}\int_{-M}^{M} \mathbf{1}_{\{F_t(x)\le G_t(x)\}}\{(\varphi_M^+(x)-x)^{p-2}a(F_t(M)\wedge G_t(x))\partial_x F_t(\varphi_M^+(x)) \\
&\qquad + (x-\varphi_M^-(x))^{p-2}a(G_t(-M)\vee F_t(x))\partial_x G_t(\varphi_M^-(x))\}\mathrm{d}x\mathrm{d}t.
\end{aligned}$$

Integral term J^1_M. Since B is C^1 on $[0,1]$,

$$\begin{aligned}|J^1_M| &\leq \int_{t_1}^{t_2}\int_{-M}^{M}\mathbf{1}_{\{F_t(x)\leq G_t(x)\}}||b||_\infty\{(\varphi^+_M(x)-x)^{p-2}|F_t(M)\wedge G_t(x)-G_t(x)| \\ &\quad + (x-\varphi^-_M(x))^{p-2}|G_t(-M)\vee F_t(x)-F_t(x)|\}\mathrm{d}x\mathrm{d}t \\ &\leq \int_{t_1}^{t_2}\int_{-M}^{M}||b||_\infty(2M)^{p-2}\{[G_t(x)-F_t(M)]^+ + [G_t(-M)-F_t(x)]^+\}\mathrm{d}x\mathrm{d}t \\ &\leq \int_{t_1}^{t_2}||b||_\infty(2M)^{p-1}\{\mathrm{tl}(F_t,M)+\mathrm{tl}(G_t,-M)\}\mathrm{d}t.\end{aligned}$$

By Lemma B.1, for all $M\geq 2C(t_2)$, for all $t\in[t_1,t_2]$,

$$\begin{aligned}&\mathrm{tl}(F_t,M)\leq \mathrm{tl}(F_0,M/2-C(t_2))+\exp(-M^2/C(t_2)),\\ &\mathrm{tl}(G_t,-M)\leq \mathrm{tl}(G_0,-M/2+C(t_2))+\exp(-M^2/C(t_2)),\end{aligned}$$

so that $|J^1_M|\to 0$ when $M\to+\infty$ due to the tail assumption on F_0 and G_0.

Integral term J^2_M. By the monotone convergence theorem,

$$\begin{aligned}\lim_{M\to+\infty}J^2_M = \frac{1}{2}\int_{t_1}^{t_2}\int_{\mathbb{R}}\mathbf{1}_{\{F_t(x)\leq G_t(x)\}}\{(F_t^{-1}(G_t(x))-x)^{p-2}a(G_t(x))\partial_x G_t(x) \\ + (x-G_t^{-1}(F_t(x)))^{p-2}a(F_t(x))\partial_x F_t(x)\}\mathrm{d}x\mathrm{d}t,\end{aligned}$$

and the limit is finite as

$$\begin{aligned}&\int_{t_1}^{t_2}\int_{\mathbb{R}}\mathbf{1}_{\{F_t(x)\leq G_t(x)\}}(F_t^{-1}(G_t(x))-x)^{p-2}a(G_t(x))\partial_x G_t(x)\mathrm{d}x\mathrm{d}t \\ &\quad\leq ||a||_\infty\int_{t_1}^{t_2}\int_{\mathbb{R}}|F_t^{-1}(G_t(x))-x|^{p-2}\partial_x G_t(x)\mathrm{d}x\mathrm{d}t \\ &\quad= ||a||_\infty\int_{t_1}^{t_2}W^{p-2}_{p-2}(F_t,G_t)\mathrm{d}t \\ &\quad\leq ||a||_\infty(t_2-t_1)W^{p-2}_{p-2}(F_0,G_0)<+\infty,\end{aligned}$$

due to Proposition 2.1 (we take the convention that $W^0_0(F_t,G_t)=1$).

Integral term J^3_M. Note that

$$\begin{aligned}&\int_{t_1}^{t_2}\int_{-M}^{M}\mathbf{1}_{\{F_t(x)\leq G_t(x)\}}(\varphi^+_M(x)-x)^{p-2}a(F_t(M)\wedge G_t(x))\partial_x F_t(\varphi^+_M(x))\mathrm{d}x\mathrm{d}t \\ &= \int_{t_1}^{t_2}\int_{-M}^{M}\mathbf{1}_{\{F_t(x)\leq G_t(x);F_t(M)\leq G_t(x)\}}(M-x)^{p-2}a(F_t(M))\partial_x F_t(M)\mathrm{d}x\mathrm{d}t+ \\ &\int_{t_1}^{t_2}\int_{-M}^{M}\mathbf{1}_{\{F_t(x)\leq G_t(x);F_t(M)>G_t(x)\}}(F_t^{-1}(G_t(x))-x)^{p-2}a(G_t(x))\partial_x F_t(F_t^{-1}(G_t(x)))\mathrm{d}x\mathrm{d}t.\end{aligned}\tag{51}$$

According to Lemmas B.1 and B.2, letting $C:=C(t_1/2)$, for $M\geq 4C$, the first integral in the right-hand side of (51) is bounded by

$$\begin{aligned}&||a||_\infty\int_{t_1}^{t_2}(2M)^{p-1}\partial_x F_t(M)\mathrm{d}t \\ &\quad\leq ||a||_\infty(2M)^{p-1}\int_{t_1}^{t_2}\frac{K}{(t-t_1/2)^{1/2}}\left(\exp\left(-\frac{M^2}{K(t-t_1/2)}\right)+\mathrm{tl}\left(F_{t_1/2},\frac{M}{2}\right)\right)\mathrm{d}t \\ &\quad\leq K||a||_\infty(2M)^{p-1}\frac{(t_2-t_1)}{(t_1/2)^{1/2}}\left(\exp\left(-\frac{M^2}{K(t_2-t_1/2)}\right)+\mathrm{tl}\left(F_0,\frac{M}{4}-C\right)+\exp\left(-\frac{M^2}{C}\right)\right),\end{aligned}$$

and the right-hand side of the last inequality vanishes when $M \to +\infty$, whereas the second integral in the right-hand side of (51) converges monotonically to

$$\int_{t_1}^{t_2} \int_{\mathbb{R}} \mathbf{1}_{\{F_t(x) \le G_t(x)\}} (F_t^{-1}(G_t(x)) - x)^{p-2} a(G_t(x)) \partial_x F_t(F_t^{-1}(G_t(x))) \mathrm{d}x\mathrm{d}t.$$

The second term in J_M^3 is similar.

Conclusion. Taking the limit $M \to +\infty$ in the equality $I_M(F_{t_2}, G_{t_2}) - I_M(F_{t_1}, G_{t_1}) - (J_M^1 + J_M^2) = J_M^3$ now yields

$$\begin{aligned}
&I(F_{t_2}, G_{t_2}) - I(F_{t_1}, G_{t_1}) - \frac{1}{2} \int_{t_1}^{t_2} \int_{\mathbb{R}} \mathbf{1}_{\{F_t(x) \le G_t(x)\}} \{(F_t^{-1}(G_t(x)) - x)^{p-2} a(G_t(x)) \partial_x G_t(x) \\
&+ (x - G_t^{-1}(F_t(x)))^{p-2} a(F_t(x)) \partial_x F_t(x)\} \mathrm{d}x\mathrm{d}t \\
&\quad = -\frac{1}{2} \left(\int_{t_1}^{t_2} \int_{\mathbb{R}} \mathbf{1}_{\{F_t(x) \le G_t(x)\}} (F_t^{-1}(G_t(x)) - x)^{p-2} a(G_t(x)) \partial_x F_t(F_t^{-1}(G_t(x))) \mathrm{d}x\mathrm{d}t \right. \\
&\qquad \left. + \int_{t_1}^{t_2} \int_{\mathbb{R}} \mathbf{1}_{\{F_t(x) \le G_t(x)\}} (x - G_t^{-1}(F_t(x)))^{p-2} a(F_t(x)) \partial_x G_t(G_t^{-1}(F_t(x))) \mathrm{d}x\mathrm{d}t \right).
\end{aligned}$$

The left-hand side of the equality above is finite and the integrands in both integrals of the right-hand side are nonnegative. Hence, all the integrals involved are absolutely convergent, and we deduce

$$\begin{aligned}
&I(F_{t_2}, G_{t_2}) - I(F_{t_1}, G_{t_1}) \\
&\quad = \frac{1}{2} \int_{t_1}^{t_2} \int_{\mathbb{R}} \mathbf{1}_{\{F_t(x) \le G_t(x)\}} \{(F_t^{-1}(G_t(x)) - x)^{p-2} \left(a(G_t(x)) (\partial_x G_t(x) - \partial_x F_t(F_t^{-1}(G_t(x)))) \right) \\
&\qquad + (x - G_t^{-1}(F_t(x)))^{p-2} \left(a(F_t(x)) (\partial_x F_t(x) - \partial_x G_t(G_t^{-1}(F_t(x)))) \right) \} \mathrm{d}x\mathrm{d}t.
\end{aligned}$$

By symmetry and using (49), we now conclude

$$\begin{aligned}
&W_p^p(F_{t_2}, G_{t_2}) - W_p^p(F_{t_1}, G_{t_1}) \\
&\quad = \frac{p(p-1)}{2} \int_{t_1}^{t_2} \int_{\mathbb{R}} \{|F_t^{-1}(G_t(x)) - x|^{p-2} \left(a(G_t(x)) (\partial_x G_t(x) - \partial_x F_t(F_t^{-1}(G_t(x)))) \right) \\
&\qquad + |x - G_t^{-1}(F_t(x))|^{p-2} \left(a(F_t(x)) (\partial_x F_t(x) - \partial_x G_t(G_t^{-1}(F_t(x)))) \right) \} \mathrm{d}x\mathrm{d}t \\
&\quad = \frac{p(p-1)}{2} \int_{t_1}^{t_2} \left\{ \int_0^1 |F_t^{-1}(u) - G_t^{-1}(u)|^{p-2} a(u) \left(\frac{1}{\partial_u G_t^{-1}(u)} - \frac{1}{\partial_u F_t^{-1}(u)} \right) \partial_u G_t^{-1}(u) \mathrm{d}u \right. \\
&\qquad \left. + \int_0^1 |F_t^{-1}(u) - G_t^{-1}(u)|^{p-2} a(u) \left(\frac{1}{\partial_u F_t^{-1}(u)} - \frac{1}{\partial_u G_t^{-1}(u)} \right) \partial_u F_t^{-1}(u) \mathrm{d}u \right\} \mathrm{d}t \\
&\quad = -\frac{p(p-1)}{2} \int_{t_1}^{t_2} \int_0^1 a(u) |F_t^{-1}(u) - G_t^{-1}(u)|^{p-2} \frac{(\partial_u F_t^{-1}(u) - \partial_u G_t^{-1}(u))^2}{\partial_u F_t^{-1}(u) \partial_u G_t^{-1}(u)} \mathrm{d}u\mathrm{d}t;
\end{aligned}$$

which completes the proof. □

References

[1] A. Alfonsi, B. Jourdain, and A. Kohatsu-Higa. Pathwise optimal transport bounds between a one-dimensional diffusion and its Euler scheme. Preprint available at http://arxiv.org/abs/1209.0576, 2012.

[2] Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[3] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.

[4] Adrian D. Banner, Robert Fernholz, and Ioannis Karatzas. Atlas models of equity markets. *Ann. Appl. Probab.*, 15(4):2296–2330, 2005.

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

[6] V. I. Bogachev, N. V. Krylov, and M. Rëkner. Elliptic and parabolic equations for measures. *Uspekhi Mat. Nauk*, 64(6(390)):5–116, 2009.

[7] François Bolley, Arnaud Guillin, and Florent Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5):867–884, 2010.
[8] François Bolley, Ivan Gentil, and Arnaud Guillin. Uniform convergence to equilibrium for granular media. *Archive for Rational Mechanics and Analysis*, 208(2):429–445, 2013.
[9] J. A. Carrillo, M. Di Francesco, and C. Lattanzio. Contractivity and asymptotics in Wasserstein metrics for viscous nonlinear scalar conservation laws. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)*, 10(2):277–292, 2007.
[10] J. A. Carrillo and G. Toscani. Wasserstein metric and large-time asymptotics of nonlinear diffusion equations. In *New trends in mathematical physics*, pages 234–244. World Sci. Publ., Hackensack, NJ, 2004.
[11] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.
[12] José Antonio Carrillo, Maria Pia Gualdani, and Giuseppe Toscani. Finite speed of propagation in porous media by mass transportation methods. *C. R. Math. Acad. Sci. Paris*, 338(10):815–818, 2004.
[13] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.
[14] A. Dembo, M. Shkolnikov, S.R.S. Varadhan, and O. Zeitouni. Large deviations for diffusions interacting through their ranks. Preprint available at http://arxiv.org/abs/1211.5223, 2012.
[15] E. Fernholz, Tomoyuki Ichiba, Ioannis Karatzas, and Vilmos Prokaj. Planar diffusions with rank-based characteristics and perturbed tanaka equations. *Probability Theory and Related Fields*, pages 1–32, 2012. To appear.
[16] E. Robert Fernholz. *Stochastic portfolio theory*, volume 48 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2002. Stochastic Modelling and Applied Probability.
[17] Alessio Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
[18] Avner Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
[19] B. H. Gilding. Improved theory for a nonlinear degenerate parabolic equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 16(2):165–224, 1989.
[20] Tomoyuki Ichiba and Ioannis Karatzas. On collisions of Brownian particles. *Ann. Appl. Probab.*, 20(3):951–977, 2010.
[21] Tomoyuki Ichiba, Ioannis Karatzas, and Mykhaylo Shkolnikov. Strong solutions of stochastic equations with rank-based coefficients. *Probability Theory and Related Fields*, pages 1–20, 2012.
[22] Tomoyuki Ichiba, Soumik Pal, and Mykhaylo Shkolnikov. Convergence rates for rank-based models with applications to portfolio theory. To appear in Probability Theory and Related Fields, 2012.
[23] Tomoyuki Ichiba, Vassilios Papathanakos, Adrian Banner, Ioannis Karatzas, and Robert Fernholz. Hybrid atlas models. *Ann. Appl. Probab.*, 21(2):609–644, 2011.
[24] B. Jourdain. Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized Burgers' equations. *ESAIM Probab. Statist.*, 1:339–355 (electronic), 1995/97.
[25] B. Jourdain. Diffusion processes associated with nonlinear evolution equations for signed measures. *Methodol. Comput. Appl. Probab.*, 2(1):69–91, 2000.
[26] B. Jourdain. Probabilistic approximation for a porous medium equation. *Stochastic Process. Appl.*, 89(1):81–99, 2000.
[27] B. Jourdain. Probabilistic characteristics method for a one-dimensional inviscid scalar conservation law. *Ann. Appl. Probab.*, 12(1):334–360, 2002.
[28] Benjamin Jourdain. Equivalence of the Poincaré inequality with a transport-chi-square inequality in dimension one. *Electronic Communications in Probability*, 17(43):1–12, 2012.
[29] Benjamin Jourdain and Florent Malrieu. Propagation of chaos and Poincaré inequalities for a system of particles interacting through their CDF. *Ann. Appl. Probab.*, 18(5):1706–1736, 2008.
[30] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
[31] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural′ceva. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1967.
[32] Qiang Liu and Chunpeng Wang. Uniqueness of the bounded solution to a strongly degenerate parabolic problem. *Nonlinear Anal.*, 67(11):2993–3002, 2007.
[33] S. Pal and M. Shkolnikov. Concentration of measure for systems of brownian particles interacting through their ranks. Preprint available at http://arxiv.org/abs/1011.2443, 2010.
[34] Soumik Pal and Jim Pitman. One-dimensional Brownian particle systems with rank-dependent drifts. *Ann. Appl. Probab.*, 18(6):2179–2207, 2008.
[35] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
[36] Mykhaylo Shkolnikov. Large systems of diffusions interacting through their ranks. *Stochastic Processes and their Applications*, 122(4):1730 – 1747, 2012.
[37] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

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

[39] Hiroshi Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.

[40] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[41] Max-K. von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[42] Zhuoqun Wu, Junning Zhao, Jingxue Yin, and Huilai Li. *Nonlinear diffusion equations*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Translated from the 1996 Chinese original and revised by the authors.

Université Paris-Est, CERMICS, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2

E-mail address: jourdain@cermics.enpc.fr

Laboratoire de Probabilités et Modèles Aléatoires (CNRS UMR. 7599), Université Paris 6 — Pierre et Marie Curie, U.F.R. Mathématiques, Case 188, 4 place Jussieu, 75252 Paris Cedex 05

Université Paris-Est, CERMICS, 6 et 8 avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée Cedex 2

E-mail address: julien.reygner@upmc.fr