



**HAL**  
open science

# A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The two-dimensional case

Milica Tomasevic

► **To cite this version:**

Milica Tomasevic. A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The two-dimensional case. 2019. hal-02044262v1

**HAL Id: hal-02044262**

**<https://inria.hal.science/hal-02044262v1>**

Preprint submitted on 21 Feb 2019 (v1), last revised 1 Aug 2019 (v2)

**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.

# A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The two-dimensional case.

Milica Tomasevic\*

**Abstract:** In Talay and Tomasevic [20] we proposed a new stochastic interpretation of the parabolic-parabolic Keller-Segel system without cut-off. It involved an original type of McKean-Vlasov interaction kernel which involved all the past time marginals of its probability distribution in a singular way. In the present paper, we study this McKean-Vlasov representation in the two-dimensional case. In this setting there exists a possibility of a blow-up in finite time for the Keller-Segel system if some parameters of the model are large. Indeed, we prove the well-posedness of the McKean-Vlasov equation under some constraints involving a parameter of the model and the initial datum. Under these constraints, we also prove the global existence for the Keller-Segel model in the plane. To obtain this result, we combine PDE analysis and stochastic analysis techniques.

**Key words:** Keller-Segel system; Singular McKean-Vlasov non-linear stochastic differential equation.

**Classification:** 60H30, 60H10.

## 1 Introduction

The standard  $d$ -dimensional parabolic-parabolic Keller-Segel model for chemotaxis reads:

$$\begin{cases} \partial_t \rho(t, x) = \nabla \cdot (\frac{1}{2} \nabla \rho - \chi \rho \nabla c)(t, x), & t > 0, x \in \mathbb{R}^d, \\ \alpha \partial_t c(t, x) = \frac{1}{2} \Delta c(t, x) - \lambda c(t, x) + \rho(t, x), & t > 0, x \in \mathbb{R}^d. \\ \rho(0, x) = \rho_0(x), \quad c(0, x) = c_0(x). \end{cases} \quad (1)$$

It describes the time evolution of the density  $\rho_t$  of a cell population and of the concentration  $c_t$  of a chemical attractant. The parameter  $\chi$  is called the chemotactic sensitivity and, together with the total mass  $M := \int \rho_0(x) dx$ , plays an important role in the well-posedness theory for (1). For a very thorough review on the analytic results on the Keller-Segel model and its variations see the reviews of Horstmann [9, 10].

When  $\alpha = 0$ , we have the parabolic-elliptic version of this system, while when  $\alpha = 1$  (more generally  $\alpha \geq 0$ ) the system is known as the parabolic-parabolic Keller-Segel model.

In Talay and Tomasevic [20], we propose the following stochastic representation of the parabolic-parabolic version of the above system: The family  $(\rho_t)_{t \geq 0}$  is seen as a family of one dimensional time marginal distributions of the law of a McKean-Vlasov stochastic process and the family  $(c_t)_{t \geq 0}$  as its transformation. Namely, we consider the following stochastic differential equation:

$$\begin{cases} dX_t = b_0^{(d)}(t, X_t) dt + \left\{ \int_0^t (K_{t-s}^{(d)} * \rho_s)(X_t) ds \right\} dt + dW_t, & t > 0, \\ \rho_s(y) dy := \mathcal{L}(X_s), \quad X_0 \sim \rho_0(x) dx, \end{cases} \quad (2)$$

where

$$K_t^{(d)}(x) := \chi e^{-\lambda t} \nabla(g_t^{(d)}(x)) \quad \text{and} \quad b_0^{(d)}(t, x) := \chi e^{-\lambda t} \nabla(c_0 * g_t^{(d)}(x)).$$

Here,  $g_t^{(d)}(x) := \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}}$ ,  $(W_t)_{t \geq 0}$  is a  $d$ -dimensional Brownian motion on a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$  and  $X_0$  is an  $\mathbb{R}^d$ -valued  $\mathcal{F}_0$ -measurable random variable. Notice that the formulation requires that the one dimensional time marginals of the law of the solution are absolutely continuous with respect to Lebesgue's measure. In addition, the process interacts with all the past time marginals of its probability distribution through a functional involving a singular kernel. Nevertheless, imagine we are able to construct a

---

\*CMAP - Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France; milica.tomasevic@polytechnique.edu.

solution to (2) and extract the family of one-dimensional marginals  $(\rho_t)_{t \geq 0}$ . Then, define the function  $c(t, x)$  as

$$c(t, x) := e^{-\lambda t} (g_t^{(d)} * c_0)(x) + \int_0^t \rho(t-s, \cdot) * e^{-\lambda s} g_s^{(d)}(x) ds.$$

The couple  $(\rho, c)$  is our stochastic interpretation of the system in (1).

The first difficulty when dealing with (2) is the singular nature of the kernel  $K^{(d)}$ . Namely, once the derivative of the heat kernel is computed, a singularity of order  $\sim t^{-(\frac{d}{2}+1)}$  emerges and needs to be integrated in time. The second difficulty is the interaction with all the past time one dimensional marginals of the law of the process. This is particularly inconvenient once the system of interacting particles related to (2) is written. Namely, the particle system will be of non-Markovian nature where each particle interacts in the current time with all the past of all the other particles through a singular functional involving the kernel  $K^{(d)}$ .

In [20], the authors overcome these difficulties and validate the above stochastic interpretation in the framework of  $d = 1$  with no constraints on the parameters. That is to say, under the assumptions  $\rho_0 \in L^1(\mathbb{R})$  and  $c_0 \in \mathcal{C}_b^1(\mathbb{R})$  they prove the global (in time) well-posedness of the Mc-Kean Vlasov SDE (2) in  $d = 1$  and as consequence the same holds for the Keller-Segel PDEs. This result generalizes the results in Osaki and Yagi [16] and Hillen and Potapov [8]. One of the key points there is that the one-dimensional kernel ( $K^{(1)}$ ) belongs to the space  $L^1((0, T); L^1(\mathbb{R}))$ . Then, by the help of precise  $L^\infty(\mathbb{R})$ -norm density estimates the singularity is tamed. Namely, Picard's iteration process is used to exhibit a weak solution to (2). In each step the  $L^\infty([0, T] \times \mathbb{R})$ -norm of the drift and  $L^1((0, T); L^\infty(\mathbb{R}))$ -norm of the marginal densities were controlled simultaneously. These controls were obtained thanks to a probabilistic method which exhibits sharp density estimates for a process whose drift is uniformly bounded in space and time (see Qian and Zheng [18] for the general idea and [20] for a generalization on time inhomogeneous processes).

Furthermore, in  $d = 1$ , Jabir *et. al* [11] prove the well-posedness of the interacting particle system associated to (2) in  $d = 1$  and its propagation of chaos. The non-Markovian nature of the system is treated with techniques based on Girsanov theorem. The calculation was based on the fact that the kernel  $K^{(1)}$  is in  $L^1((0, T); L^2(\mathbb{R}))$ . An interesting point there was that in order to get the tightness in number of particles and the propagation of chaos, the authors needed to use the so called "partial" Girsanov transformations removing a finite number of particles from the system. This was necessary in order to exhibit estimates on the moments on the corresponding exponential martingales that behave satisfactory when the number of particles goes to infinity.

Contrary to the one-dimensional case, a blow up may occur for the Keller-Segel system in the two-dimensional setting if the parameter  $\chi$  is large. In the parabolic-elliptic version of the system, its behaviour has been completely understood. There, the system exhibits the "threshold" behaviour: if  $M\chi < 8\pi$  the solutions are global in time, if  $M\chi > 8\pi$  every solution blows-up in finite time (see e.g. Blanchet *et. al* [2] and Nagai and Ogawa [15]). The fully parabolic model does not exhibit the same behaviour. It has been proved that when  $M\chi < 8\pi$  one has global existence (see Calvez and Corrias [5], Mizogouchi [14]). However, in Biller *et. al* [1] the authors find an initial configuration of the system in which a global solution in some sense exists with  $M\chi > 8\pi$ . Then, Herrero and Velázquez [7] construct a radially symmetric solution on a disk that blows-up and develops  $\delta$ -function type singularities. Finally, unique solution with any positive mass exists when the parameter  $\alpha$  is large enough (Corrias *et. al* [6]). Thus, in the case of parabolic-parabolic model, the value  $8\pi$  can still be understood as a threshold, but in a different sense: under it there is global existence, over it there exists a solution that blows up.

The goal of this paper is to prove well-posedness of the equation (2) in the case  $d = 2$  and then get a new well-posedness result for the parabolic-parabolic Keller-Segel equation in  $d = 2$  by means of the above stochastic representation. All this under some constraints on the size of  $\chi$  and with an appropriate (minimal) conditions on initial datum.

In order to achieve this, we will perform a regularization of the singular interaction kernel and combine probabilistic and PDE techniques in order to prove that the regularized process converges (weakly) to the non-regularized one. The increase in the space dimension has a significant impact on the techniques used in [20] to prove the well-posedness of the McKean-Vlasov SDE. A generalisation to the multidimensional case of the results in [18] can be found in Qian *et al.* [17] in the case of time homogeneous drifts. There, the authors show that the estimate of the transition density of a  $d$ -dimensional stochastic process is a product of one-dimensional estimates provided that the Euclidean norm of the drift vector is uniformly bounded. With the arguments we used in  $d = 1$ , one can easily extend the results in [17] to time inhomogeneous drifts. However, the estimates obtained are too crude and the choice of  $L^\infty$ -norm in space seems no longer to be

adapted to the problem. That is why, a new functional space needs to be introduced and the next obvious choice are the  $L^q$ -spaces. Analyzing a priori the mild equation for the one-dimensional time marginals  $\rho_t$ , one can notice that if  $\rho_0 \in L^1(\mathbb{R}^2)$ , the term depending on the initial condition gives an  $L^q$ -norm in space of order  $\sim t^{-(1-1/q)}$ . After that, if we were to impose  $\sup_{t \leq T} t^{1-\frac{1}{q}} \|\rho_t\| < \infty$  for a  $q > 2$ , then one would obtain that the non-linear drift is of order  $\sim t^{-\frac{1}{2}}$ . The goal will be to obtain these estimates for the regularized versions of the drift and density uniformly in the regularization parameter. The price to pay will be a restriction on the parameter  $\chi$ . This will enable us to conclude tightness of the probability laws of the regularized processes and it will remain to solve the martingale problem corresponding to (2). We emphasize here that with the above choice of density spaces, the drift of the stochastic process will not belong to the framework of Krylov and Röckner [13]. Thus, the techniques we used in the one-dimensional case on the associated particle system will not work. A consequence of the existence result for (2) is the global existence for (1) in  $d = 2$ . This generalizes the result in [6] by removing the assumption on the smallness of the initial datum (for more details see the next section).

The plan of the paper is the following: In Section 2 we state our main results and compare them with the above mentioned Keller-Segel literature. In Section 3 we present our regularization procedure and we obtain density estimates of the regularized processes independent of the regularization parameter. This estimates enable us to prove tightness and then the well-posedness of the corresponding martingale problem in Section 4. Then, in Section 5 we prove the global existence for the Keller-Segel system in  $d = 2$ .

In all the paper  $C$  will denote a generic constant,  $C_p$  will denote a constant depending on a parameter  $p$  and  $C(p_1, p_2, \dots)$  will denote a constant depending on parameters  $p_1, p_2, \dots$ . In addition, from now on we will drop the dimension index in the definitions of the interaction kernel ( $K^{(d)}$ ) and linear drift ( $b_0^{(d)}$ ).

## 2 Main results

On a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$ , equipped with a 2-dimensional Brownian motion  $(W_t)_{t \geq 0}$ , we consider the following non-linear stochastic process:

$$\begin{cases} dX_t = dW_t + b_0(t, X_t)dt + \chi \left\{ \int_0^t e^{-\lambda(t-s)} \int K_{t-s}(X_t - y) p_s(y) dy ds \right\} dt, & t \leq T \\ p_s(y) dy := \mathcal{L}(X_s), & X_0 \sim \rho_0, \end{cases} \quad (3)$$

where  $X_0$  is an  $\mathbb{R}^2$ -valued  $\mathcal{F}_0$ -measurable random variable,  $g_t$  denotes the probability density of  $W_t$  and for  $(t, x) \in (0, T] \times \mathbb{R}^2$

$$b_0(t, x) := \chi e^{-\lambda t} (\nabla c_0 * g_t)(x) \text{ and } K_t(x) := \nabla g_t(x) = -\frac{x}{2\pi t^2} e^{-\frac{|x|^2}{2t}}.$$

Here  $|x|$  denotes the Euclidean norm. Notice that  $K_t$  is a two dimensional vector. We denote its coordinates by  $K_t^i$  with  $i = 1, 2$  and

$$b^i(t, x; p) = b_0^i(t, X_t) + \chi \int_0^t e^{-\lambda(t-s)} \int K_{t-s}^i(X_t - y) p_s(y) dy ds dt.$$

Let us define the notion of a weak solution to (3).

**Definition 2.1.** *The family  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t), X, W)$  is said to be a weak solution to the equation (3) up to time  $T > 0$  if:*

1.  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$  is a filtered probability space.
2. The process  $X := (X_t)_{t \in [0, T]}$  is  $\mathbb{R}^2$ -valued, continuous, and  $(\mathcal{F}_t)$ -adapted. In addition, the probability distribution of  $X_0$  has density  $\rho_0$ .
3. The process  $W := (W_t)_{t \in [0, T]}$  is a two-dimensional  $(\mathcal{F}_t)$ -Brownian motion.
4. The probability distribution  $\mathbb{P} \circ X^{-1}$  has time marginal densities  $(p_t, t \in (0, T])$  with respect to Lebesgue measure which satisfy for any

$$\forall 1 < q < \infty \exists C_q > 0, \quad \sup_{t \leq T} t^{1-\frac{1}{q}} \|p_t\|_{L^q(\mathbb{R}^2)} \leq C_q. \quad (4)$$

5. For any  $t \in [0, T]$  and  $x \in \mathbb{R}^2$ , one has  $\int_0^t |b_0(s, x)| ds < \infty$ .

6.  $\mathbb{P}$ -a.s. the pair  $(X, W)$  satisfies (3).

**Remark 2.2.** Notice that under the condition  $c_0 \in H^1(\mathbb{R}^2)$  one gets applying Hölder's inequality and (11),

$$\int_0^t |b_0(s, x)| ds \leq C \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \int_0^t \frac{1}{\sqrt{s}} ds.$$

Similarly, (4) implies

$$\int_0^t \left| \int_0^s K_{s-u} * p_u(x) du \right| ds \leq C \int_0^t \frac{1}{\sqrt{s}} ds.$$

Moreover, if  $c_0 \in H^1(\mathbb{R}^2)$  and if the marginals of  $X_t$  satisfy (4), then

$$\forall 2 \leq r \leq \infty \exists C_r > 0, \quad \sup_{t \leq T} t^{\frac{1}{2} - \frac{1}{r}} \|b(t, \cdot; p)\|_{L^r(\mathbb{R}^2)} \leq C_r. \quad (5)$$

Now, we present our first main result.

**Theorem 2.3.** Let  $T > 0$  and suppose that  $X_0$  has a probability density function  $\rho_0$ . Furthermore, assume that  $c_0 \in H^1(\mathbb{R}^2)$ . Then, Equation (3) admits a weak solution under the following condition

$$A\chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} + B\sqrt{\chi} < 1, \quad (6)$$

where  $A$  and  $B$  are defined as in Proposition 3.11 below.

To prove the above result, we do not apply Picard's iteration procedure since in each iteration step we will need a well-posedness result for a linear SDE whose drift satisfies (5). In view of Krylov and Röckner [13], the well-posedness follows from a finite  $L^p((0, T); L^r(\mathbb{R}^2))$ -norm of the drift with  $\frac{1}{p} + \frac{1}{r} < \frac{1}{2}$ . Unfortunately, the property in (5) will imply the opposite condition  $\frac{1}{p} + \frac{1}{r} > \frac{1}{2}$  for the same norm to be finite. That is why to prove Theorem 2.3 we will use a regularization method. The goal is to prove that the time marginals of the regularized version of (3) satisfy the property (4) with uniform constants with respect to the regularization parameter. Then, the tightness will follow thanks to (5) for  $r = \infty$ . It will remain, then, to solve the non-linear martingale problem related to (3). The well-posedness of the regularized equation is an adaptation of the results in Sznitman [19] presented in a general way in the Appendix. In addition, the incompatibility of (5) and the condition in [13] makes us doubt that Girsanov transform techniques would work and that the law of (3) is absolutely continuous w.r.t. Wiener's measure even under (6).

The next objective is to use Theorem 2.3 to get a well-posedness of the Keller-Segel model in  $d = 2$ . The system reads

$$\begin{cases} \frac{\partial \rho}{\partial t}(t, x) = \nabla \cdot \left( \frac{1}{2} \nabla \rho(t, x) - \chi \rho(t, x) \nabla c(t, x) \right), & t > 0, \quad x \in \mathbb{R}^2, & (7a) \\ \frac{\partial c}{\partial t}(t, x) = \frac{1}{2} \Delta c(t, x) - \lambda c(t, x) + \rho(t, x), & t > 0, \quad x \in \mathbb{R}^2, & (7b) \\ \rho(0, x) = \rho_0(x), \quad c(0, x) = c_0, & & \end{cases}$$

where  $\chi > 0$  and  $\lambda \geq 0$ . Notice that the two diffusion coefficient are deliberately chosen to be equal to  $\frac{1}{2}$  in order to have unit diffusion coefficient and standard Gaussian kernel in the formulation of (3).

The new functions  $\tilde{\rho}(t, x) := \frac{\rho(t, x)}{M}$  and  $\tilde{c}(t, x) := \frac{c(t, x)}{M}$  satisfy the system (7) with the new parameter  $\tilde{\chi} := \chi M$ . Therefore, w.l.o.g. we may and do thereafter assume that  $M = 1$ . We consider the following notion of solution to (7):

**Definition 2.4.** Given the functions  $\rho_0$  and  $c_0$ , and the constants  $\chi > 0$ ,  $\lambda \geq 0$ ,  $T > 0$ , the pair  $(\rho, c)$  is said to be a solution to (7) if  $\rho(t, \cdot)$  is a probability density function for every  $0 \leq t \leq T$ , one has

$$\forall 1 < q < \infty \exists C_q > 0 : \quad \sup_{t \leq T} t^{1 - \frac{1}{q}} \|\rho(t, \cdot)\|_{L^q(\mathbb{R}^2)} \leq C_q,$$

and the following equality

$$\rho(t, x) = g_t * \rho_0(x) - \chi \sum_{i=1}^2 \int_0^t \nabla_i g_{t-s} * \nabla_i c(s, \cdot) \rho(s, \cdot)(x) ds \quad (8)$$

is satisfied in the sense of the distribution with

$$c(t, x) = e^{-\lambda t}(g(t, \cdot) * c_0)(x) + \int_0^t e^{-\lambda s}(g_s * \rho(t - s, \cdot))(x) ds. \quad (9)$$

Notice that the function  $c(t, x)$  defined by (9) is a mild solution to (7b). These solutions are known as integral solutions and they have already been studied in PDE literature for the two-dimensional Keller-Segel model (see [6] and references therein).

A consequence of Theorem 2.3 is the well-posedness of (7).

**Corollary 2.5.** *Let  $\rho_0$  a probability density function and  $c_0 \in H^1(\mathbb{R}^2)$ . Under the condition (6) the system (7) admits a global solution in the sense of Definition 2.4.*

Let us compare the above result with the literature mentioned in the introduction. In [5] the authors obtain the global existence in sub-critical case assuming:

- i)  $\rho_0 \in L^1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2, \log(1 + |x|^2)dx)$  and  $\rho_0 \log \rho_0 \in L^1(\mathbb{R}^2)$ ;
- ii)  $c_0 \in H^1(\mathbb{R}^2)$  if  $\lambda > 0$  or  $c_0 \in L^1(\mathbb{R}^2)$  and  $|\nabla c_0| \in L^2(\mathbb{R}^2)$  if  $\lambda = 0$ ;
- iii)  $\rho_0 c_0 \in L^1(\mathbb{R}^2)$ .

We should emphasize that their sub-critical condition translates into  $4\chi < 8\pi$  for (7) due to the additional diffusion coefficients in it and the assumption  $M = 1$ . In the same sub-critical case, the global existence result is obtained in [14] assuming  $\rho_0 \in L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$  and  $c_0 \in H^1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ . Our result does not assume any additional conditions other than that  $\rho_0$  is a probability density function and  $c_0 \in H^1(\mathbb{R}^2)$ . The price to pay is the smallness condition (6) that not just involves the parameter  $\chi$ , but the initial datum as well.

Corollary 2.5 is more appropriate to be compared to the result in [6, Thm. 2.1]. Indeed, the assumptions on initial conditions are the same and as well the notion of solution. However, the setting and the objectives are different. In [6], the parameter  $\alpha$  (see (1)) is not fixed to be equal to 1 and plays an important role. The goal is to prove the global existence for any positive mass  $M$  and  $\chi = 1$ . This is achieved under the following conditions:

C1: There exists  $\delta = \delta(M, \alpha)$  such that  $\|\nabla c_0\|_{L^2(\mathbb{R}^2)} < \delta$ ,

C2: There exists a constant  $C = C(\alpha)$  such that  $M < C(\alpha)$ .

The condition C2 is similar to the Condition (6) for  $\chi$ . We cannot totally compare them as the constants are not explicitly written in [6]. What is important is that  $C(\alpha)$  grows with  $\alpha$ , so one can have  $M$  as large as one likes in C2 as soon as  $\alpha$  is large enough as well. In this paper the objective is to get results for the classical Keller-Segel model ( $\alpha = 1$ ) with respect to the chemo-attractant sensitivity. When we assume the same ( $\alpha = 1$ ,  $M = 1$  and  $\chi > 0$ ) in the framework of [6], we see that we have removed the assumption on the smallness of the initial datum (C1). The reason lies in our method: in [6] the Banach's fixed point is used to construct solutions locally in time (where C1 emerges) and then such solution is globalized (where C2 emerges). In our case only a condition of C2 type appears as, thanks to our regularization procedure, we directly construct a global solution. The well-posedness of the regularized equation comes from the Appendix. In addition, as our condition is explicitly written in Section 3, one can analyze the constants in order to find the optimal condition on  $\chi$ .

## 3 Regularization

### 3.1 Preliminaries

The following technical lemmas will be used throughout the paper.

**Lemma 3.1.** *Let  $t > 0$  and  $i \in \{1, 2\}$ . Then, for any  $1 \leq q < \infty$  one has*

$$\|K_t^i\|_{L^q(\mathbb{R}^2)} = \|\nabla_i g_t\|_{L^q(\mathbb{R}^2)} = \frac{C_1(q)}{t^{\frac{3}{2} - \frac{1}{q}}}, \quad (10)$$

where

$$C_1(q) = \frac{2^{\frac{1}{q}-\frac{1}{2}}}{\pi^{1-\frac{1}{2q}} q^{\frac{1}{q}+\frac{1}{2}}} \left( \Gamma\left(\frac{q+1}{2}\right) \right)^{\frac{1}{q}}.$$

Here  $\Gamma(x)$  denotes the Gamma function:  $\Gamma(x) = \int_0^\infty z^{x-1} e^{-z} dz$ .

Notice that for  $q \geq 2$ , the  $L^1((0, T); L^q(\mathbb{R}^2))$ -norm of  $K^i$  explodes. This was not the case in  $d = 1$  and it is in this sense that the two-dimensional kernel is more singular.

*Proof.* Let  $1 \leq q < \infty$ . One has

$$\begin{aligned} \|K_t^i\|_{L^q(\mathbb{R}^2)} &= \|\nabla_i g_t\|_{L^q(\mathbb{R}^2)} = \frac{1}{2\pi t^2} \left( \int_{\mathbb{R}^2} |x_i|^q e^{-\frac{q|x|^2}{2t}} dx \right)^{\frac{1}{q}} = \frac{1}{2\pi t^2} \left( \int_{\mathbb{R}} e^{-q\frac{x^2}{2t}} dx \int_{\mathbb{R}} |x|^q e^{-q\frac{x^2}{2t}} dx \right)^{\frac{1}{q}} \\ &= \frac{1}{2\pi t^2} \left( \frac{\sqrt{2\pi t}}{\sqrt{q}} 2 \int_0^\infty x^q e^{-q\frac{x^2}{2t}} dx \right)^{\frac{1}{q}}. \end{aligned}$$

Apply the change of variables  $\frac{qx^2}{2t} = y$ . It comes

$$\begin{aligned} \|K_t^i\|_{L^q(\mathbb{R}^2)} &= \|\nabla_i g_t\|_{L^q(\mathbb{R}^2)} = \frac{1}{2\pi t^2} \left( \frac{\sqrt{2\pi t}}{\sqrt{q}} 2 \left(\frac{2t}{q}\right)^{\frac{q-1}{2}} \int_0^\infty y^{\frac{q-1}{2}} e^{-y} \frac{t}{q} dy \right)^{\frac{1}{q}} \\ &= \frac{1}{2\pi t^2} \left(\frac{2t}{q}\right)^{\frac{1}{q}+\frac{1}{2}} \pi^{\frac{1}{2q}} \left( \Gamma\left(\frac{q+1}{2}\right) \right)^{\frac{1}{q}}. \end{aligned}$$

This ends the proof. □

The change of variables  $\frac{x}{\sqrt{t}} = z$  leads to

**Lemma 3.2.** *Let  $t > 0$ . Then, for any  $1 \leq q < \infty$  one has*

$$\|g_t\|_{L^q(\mathbb{R}^2)} = \frac{C_2(q)}{t^{1-\frac{1}{q}}}, \quad (11)$$

where

$$C_2(q) = \frac{1}{(2\pi)^{1-\frac{1}{q}} q^{\frac{1}{q}}}.$$

The functions  $C_1(q)$  and  $C_2(q)$  will be used only when we need the explicit constants in a computation. In all other cases we will use notation  $C_q$  that may change from line to line.

Now, for  $0 < a, b < 1$ , we denote

$$\beta(a, b) := \int_0^1 \frac{1}{u^a (1-u)^b} du. \quad (12)$$

**Lemma 3.3.** *Let  $t > 0$  and  $0 < a, b < 1$ . Then,*

$$\int_0^t \frac{1}{s^a (t-s)^b} ds = t^{1-(a+b)} \beta(a, b).$$

*Proof.* Observe that

$$\int_0^t \frac{1}{s^a (t-s)^b} ds = \frac{1}{t^{(a+b)}} \int_0^t \frac{1}{\left(\frac{s}{t}\right)^a \left(1-\frac{s}{t}\right)^b} ds.$$

The change of variables  $\frac{s}{t} = u$  implies the result. □

Finally, we state here the two standard convolution inequalities in their general form. The following is proven in Brezis [3, Thm. 4.15]:



**Lemma 3.4** (The convolution inequality). *Let  $f \in L^p(\mathbb{R}^d)$  and  $g \in L^1(\mathbb{R}^d)$  with  $l \leq p \leq \infty$  and  $\frac{1}{l} + \frac{1}{q} - 1 \geq 0$ . Then,  $f * g \in L^p(\mathbb{R}^d)$  and*

$$\|f * g\|_{L^p(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}. \quad (13)$$

The following is an extension of Lemma 13 and it is proven in [3, Thm. 4.33]:

**Lemma 3.5** (The convolution inequality). *Let  $f \in L^p(\mathbb{R}^d)$  and  $g \in L^q(\mathbb{R}^d)$  with  $l \leq p, q \leq \infty$ . Then,  $f * g \in L^r(\mathbb{R}^d)$  and*

$$\|f * g\|_{L^r(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)}. \quad (14)$$

## 3.2 Regularization

We define the regularized version of the interaction kernel  $K$  and the linear part of the drift as follows. For  $\varepsilon > 0$  and  $(t, x) \in (0, T) \times \mathbb{R}^2$  define

$$K_t^\varepsilon := -\frac{x}{2\pi(t+\varepsilon)^2} e^{-\frac{|x|^2}{2t}}, \quad g_t^\varepsilon(x) := \frac{1}{2\pi(t+\varepsilon)} e^{-\frac{|x|^2}{2t}} \quad \text{and} \quad b_0^\varepsilon(t, x) := \chi e^{-\lambda t} (\nabla c_0 * g_t^\varepsilon)(x).$$

The regularized Mc-Kean-Vlasov equation reads

$$\begin{cases} dX_t^\varepsilon = dW_t + b_0^\varepsilon(t, X_t^\varepsilon) dt + \chi \left\{ \int_0^t e^{-\lambda(t-s)} \int K_{t-s}^\varepsilon(X_t^\varepsilon - y) \mu_s^\varepsilon(dy) ds \right\} dt, t \leq T, \\ \mu_s^\varepsilon := \mathcal{L}(X_s^\varepsilon), \quad X_0^\varepsilon \sim \rho_0, \end{cases} \quad (15)$$

Set

$$b^\varepsilon(t, x; \mu^\varepsilon) := b_0^\varepsilon(t, x) + \chi \int_0^t e^{-\lambda(t-s)} \int K_{t-s}^\varepsilon(x - y) \mu_s^\varepsilon(dy) ds.$$

It is clear that there exists  $C_\varepsilon > 0$  such that for any  $t \in (0, T)$ , one has

$$\forall x, y \in \mathbb{R}^2, \quad |b_0^\varepsilon(t, x) - b_0^\varepsilon(t, y)| + |K_t^\varepsilon(x) - K_t^\varepsilon(y)| \leq C_\varepsilon |x - y| \quad \text{and} \quad |b_0^\varepsilon(t, x)| + |K_t^\varepsilon(x)| \leq C_\varepsilon. \quad (16)$$

Notice that  $C_\varepsilon \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ . Similar computations as the ones to get (10) and (11) lead to the following estimates. For  $t \in (0, T]$  and  $1 \leq q < \infty$ , one has

$$\|K_t^{\varepsilon, i}\|_{L^q(\mathbb{R}^2)} \leq \frac{C_1(q)}{(t+\varepsilon)^{\frac{3}{2}-\frac{1}{q}}} \quad \text{and} \quad \|g_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq \frac{C_2(q)}{(t+\varepsilon)^{1-\frac{1}{q}}}. \quad (17)$$

**Proposition 3.6.** *Let  $T > 0, \chi > 0, \nabla c_0 \in L^2(\mathbb{R}^2)$  and  $\rho_0$  a probability density function on  $\mathbb{R}^2$ . Then, for any  $\varepsilon > 0$ , Equation (15) admits a unique strong solution. Moreover, the one dimensional time marginals of the law of the solution admit probability density functions,  $(p_t^\varepsilon)_{t \leq T}$ . In addition, for  $t \in (0, T)$ ,  $p_t^\varepsilon$  satisfies the following mild equation in the sense of the distributions:*

$$p_t^\varepsilon = g_t * \rho_0 - \sum_{i=1}^2 \int_0^t \nabla_i g_{t-s} * (p_s^\varepsilon b^{\varepsilon, i}(s, \cdot; p_s^\varepsilon)) ds. \quad (18)$$

*Proof.* In view of (16) and Theorem 6.1 from the Appendix, the strong solution to Equation (15) is uniquely well defined. In addition, as the drift term is bounded, we can apply Girsanov's transformation and conclude that the one dimensional time marginals of the law of the solution admit probability density functions. By classical arguments (see e.g. [21]), one can prove that for  $t \in (0, T)$ ,  $p_t^\varepsilon$  satisfies (18) in sense of the distributions.  $\square$

In the sequel, for  $1 < q < \infty$ , uniform in  $\varepsilon$  estimates on  $\sup_{t \leq T} t^{1-\frac{1}{q}} \|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)}$  will be crucial. They will imply uniform in  $\varepsilon$  estimates on  $\sup_{t \leq T} t^{\frac{1}{2}-\frac{1}{r}} \|b^{\varepsilon, i}(t, \cdot)\|_{L^r(\mathbb{R}^2)}$  for  $2 \leq r \leq \infty$ . In particular, for  $i = 1, 2$  and  $t \in (0, T]$ , one will have

$$\|b^{\varepsilon, i}(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \frac{C}{\sqrt{t}}.$$

The latter will enable us to prove tightness of the probability laws of  $(X^\varepsilon)$ .



### 3.3 Density estimates

**Lemma 3.7.** *Let  $t > 0$ . Then, the function  $b_0^i(t, \cdot)$  is continuous on  $\mathbb{R}^2$  and for  $r \in [2, \infty]$ , one has*

$$\|b_0^i(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \frac{C_2(\frac{2r}{r+2})}{t^{\frac{1}{2}-\frac{1}{r}}}.$$

*Proof.* As  $\nabla_i c_0$  is only in  $L^2(\mathbb{R}^2)$  we can not apply the classical results of convolution with a continuous function. The continuity of  $b_0^i(t, \cdot) = \chi \nabla_i c_0 * g_t$  is a direct consequence of [3, Ex. 4.30-3.] as for a  $t > 0$  both  $g_t$  and  $\nabla_i c_0$  belong to  $L^2(\mathbb{R}^2)$ . However, in this case, one can use the particular form of the functions involved in the convolution to prove the continuity. Let  $x_n \rightarrow x$  in  $\mathbb{R}^2$  as  $n \rightarrow \infty$ . To prove  $g_t * \nabla_i c_0(x_n) \rightarrow g_t * \nabla_i c_0(x)$  we need to bound  $|g_t(x_n - y) \nabla_i c_0(y)|$  with an  $h(y) \in L^1(\mathbb{R}^2)$ . As  $g_t$  is continuous we would then apply the dominated convergence theorem. Let  $R > 0$ . Then there exists  $n_0 \geq 1$  such that for  $n \geq n_0$  one has that  $|x_n - x| \leq R$ . Then, by reverse triangular inequality one has that

$$e^{-\frac{|x_n - y|^2}{2t}} \leq e^{-\frac{(|x_n - x| - |y - x|)^2}{2t}} \leq e^{\frac{R^2}{2t}} e^{-\frac{(|y - x| - R)^2}{2t}}.$$

Thus, we define  $h(y) = \nabla_i c_0(y) \frac{1}{2\pi t} e^{\frac{R^2}{2t}} e^{-\frac{(|y - x| - R)^2}{2t}}$  and conclude  $h \in L^1(\mathbb{R}^2)$  by Cauchy-Schwarz inequality.

Let  $q \geq 1$  be such that  $\frac{1}{q} + \frac{1}{2} = 1 + \frac{1}{r}$ . By the convolution inequality (14), one has

$$\|b_0^i(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \chi \|\nabla_i c_0\|_{L^2(\mathbb{R}^2)} \|g_t\|_{L^q(\mathbb{R}^2)}.$$

In view of estimates on  $\|g_t\|_{L^q(\mathbb{R}^2)}$  and the relation above between  $r$  and  $q$ , one has

$$\|b_0^i(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \|g_t\|_{L^{\frac{2r}{r+2}}} \leq \chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \frac{C_2(\frac{2r}{r+2})}{t^{1-(\frac{1}{r}+\frac{1}{2})}}.$$

□

Repeating the arguments as in the preceding proof, one gets

**Lemma 3.8.** *For  $t > 0$  and  $r \in [2, \infty]$  one has*

$$\|b_0^{\varepsilon, i}(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \frac{C_2(\frac{2r}{r+2})}{(t + \varepsilon)^{\frac{1}{2}-\frac{1}{r}}}.$$

**Lemma 3.9.** *Let  $p_0$  a probability density function on  $\mathbb{R}^2$  and  $1 < q < \infty$ . One has*

$$\limsup_{t \rightarrow 0} t^{1-\frac{1}{q}} \|g_t * p_0\|_{L^q(\mathbb{R}^2)} = 0.$$

*Proof.* The proof is a special case of Lemma 8 in [4]. Let  $f \in C_K(\mathbb{R}^2)$ . Using the standard convolution inequality (13), one has

$$\begin{aligned} t^{1-\frac{1}{q}} \|g_t * p_0\|_{L^q(\mathbb{R}^2)} &\leq t^{1-\frac{1}{q}} \|g_t * f\|_{L^q(\mathbb{R}^2)} + t^{1-\frac{1}{q}} \|g_t * (p_0 - f)\|_{L^q(\mathbb{R}^2)} \\ &\leq t^{1-\frac{1}{q}} \|g_t\|_{L^1(\mathbb{R}^2)} \|f\|_{L^q(\mathbb{R}^2)} + t^{1-\frac{1}{q}} \|g_t\|_{L^q(\mathbb{R}^2)} \|p_0 - f\|_{L^1(\mathbb{R}^2)} \\ &\leq \|f\|_{L^q(\mathbb{R}^2)} t^{1-\frac{1}{q}} + C \|p_0 - f\|_{L^1(\mathbb{R}^2)}. \end{aligned}$$

Thus,

$$\limsup_{t \rightarrow 0} t^{1-\frac{1}{q}} \|g_t * p_0\|_{L^q(\mathbb{R}^2)} \leq C \|p_0 - f\|_{L^1(\mathbb{R}^2)}.$$

Since  $f$  is arbitrary, the r.h.s. can be arbitrarily small ( see e.g. [3, Theorem 4.3] ).

□

Now, let us define

$$\mathcal{N}_q^\varepsilon(t) := \sup_{s \in (0, t)} s^{1-\frac{1}{q}} \|p_s^\varepsilon\|_{L^q(\mathbb{R}^2)}. \quad (19)$$

The following lemma provides a first estimate for  $\mathcal{N}_q^\varepsilon(t)$  for a fixed  $\varepsilon > 0$ . This estimate is not the optimal one in  $\varepsilon$ , but it is necessary in order to be sure that all the quantities we work with are well defined. Also, it will be used in order to obtain the limit behaviour of  $\mathcal{N}_q^\varepsilon(t)$  as  $t \rightarrow 0$ .

**Lemma 3.10.** *Let  $0 < t \leq T$  and  $\varepsilon > 0$  fixed. For any  $1 < q < \infty$ , there exists  $C_\varepsilon(T, \chi) > 0$  such that*

$$\mathcal{N}_q^\varepsilon(t) \leq C_\varepsilon(T, \chi). \quad (20)$$

Moreover, one has

$$\lim_{t \rightarrow 0} \mathcal{N}_q^\varepsilon(t) = 0. \quad (21)$$

As  $K^\varepsilon$  is smooth, we can propose a simplified version of the arguments in [4, p. 285-286] for the proof of (21).

*Proof.* The drift of the regularized stochastic equation is bounded. Indeed,  $|K_t^\varepsilon| \leq \frac{C}{\varepsilon^{3/2}}$  and Lemma 3.8 imply

$$\|b^{\varepsilon, i}(t, \cdot; p^\varepsilon)\|_{L^\infty(\mathbb{R}^2)} \leq \frac{C}{\sqrt{\varepsilon}} + \frac{Ct}{\varepsilon^{3/2}} =: C_\varepsilon(1+t).$$

For  $1 < q < \infty$  and  $q'$  such that  $\frac{1}{q} + \frac{1}{q'} = 1$  integrate (18) w.r.t. a test function  $f \in L^{q'}(\mathbb{R}^2)$  and apply Hölder's inequality. It comes

$$\left| \int p_t^\varepsilon(x) f(x) dx \right| \leq \|f\|_{L^{q'}(\mathbb{R}^2)} \left( \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + \sum_{i=1}^2 \int_0^t \|\nabla_i g_{t-s} * (p_s^\varepsilon b_s^{\varepsilon, i})\|_{L^q(\mathbb{R}^2)} ds \right). \quad (22)$$

a) Assume  $1 < q < 2$ . The above drift bound and the convolution inequality (13) applied in (22), lead to

$$\|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + C_\varepsilon(1+t) \sum_{i=1}^2 \int_0^t \|\nabla_i g_{t-s}\|_{L^q(\mathbb{R}^2)} \|p_s^\varepsilon\|_{L^1(\mathbb{R}^2)} ds.$$

In view of (10), we deduce that

$$\int_0^t \|\nabla_i g_{t-s}\|_{L^q(\mathbb{R}^2)} \|p_s^\varepsilon\|_{L^1(\mathbb{R}^2)} ds \leq C_q \int_0^t \frac{1}{(t-s)^{\frac{3}{2}-\frac{1}{q}}} = C_q t^{\frac{1}{q}-\frac{1}{2}}.$$

Thus,

$$t^{1-\frac{1}{q}} \|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq t^{1-1/q} \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + t^{1-\frac{1}{q}+\frac{1}{q}-1/2} C_\varepsilon(1+t). \quad (23)$$

To get (20), in (23) use the convolution inequality (13) and that  $\|g_t\|_{L^q(\mathbb{R}^2)} = \frac{C}{t^{1-\frac{1}{q}}}$ . To get (21) use Lemma 3.9 for the first term of the r.h.s. of (23) and the fact that the second term tends to zero as  $t \rightarrow 0$ .

b) Let  $q \geq 2$  and  $\frac{1}{p_1} = \frac{1}{p_2} = \frac{1}{2} + \frac{1}{2q}$ . Then,  $1 < p_1, p_2 < 2$  and  $1 + \frac{1}{q} = \frac{1}{p_1} + \frac{1}{p_2}$ . The convolution inequality (14) and the drift estimate applied in (22), lead to

$$\|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + C_\varepsilon(1+T) \sum_{i=1}^2 \int_0^t \|\nabla_i g_{t-s}\|_{L^{p_1}(\mathbb{R}^2)} \|p_s^\varepsilon\|_{L^{p_2}(\mathbb{R}^2)} ds.$$

In view of (10) and the result in a), one has

$$t^{1-\frac{1}{q}} \|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq t^{1-\frac{1}{q}} \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + C_\varepsilon(1+T) t^{1-\frac{1}{q}} \int_0^t \frac{C_q}{(t-s)^{\frac{3}{2}-\frac{1}{p_1}}} \frac{C(\varepsilon, T)}{s^{1-\frac{1}{p_2}}} ds.$$

Apply Lemma 3.3 and use the relation between the exponents. It comes:

$$t^{1-\frac{1}{q}} \|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq t^{1-\frac{1}{q}} \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + t^{1-\frac{1}{q}} \frac{C(\varepsilon, T)}{t^{\frac{1}{2}-\frac{1}{q}}} \beta \left(1 - \frac{1}{p_2}, \frac{3}{2} - \frac{1}{p_1}\right).$$

Repeating the last steps as in a), one can obtain the desired result.  $\square$

The following proposition enables one to control  $\mathcal{N}_q^\varepsilon(t)$  for a fixed  $q$  and uniformly on small  $\varepsilon$ .

**Proposition 3.11.** *Let  $T > 0$  and fix a  $q \in (2, 4)$ . Then, there exists  $C > 0$  such that for any  $t \in (0, T]$ ,  $\mathcal{N}_q^\varepsilon(t)$  defined in (19) satisfies*

$$\forall 0 < \varepsilon < 1: \quad \mathcal{N}_q^\varepsilon(t) \leq C,$$

provided that

$$A\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + B\sqrt{\chi} < 1 \quad (24)$$

where,  $C_1$ ,  $C_2$  and  $\beta(\cdot, \cdot)$  being defined as in Lemmas 3.1, 3.2 and Eq. 12 respectively,

$$A = C_1(q')C_2\left(\frac{2q}{q+2}\right)\beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right) \quad \text{and} \quad B = 2\sqrt{C_2(q)C_1(q')C_1(1)\beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right)\beta\left(1 - \frac{1}{q}, \frac{1}{2}\right)}.$$

*Proof.* Let  $q' > 1$  be such that  $\frac{1}{q} + \frac{1}{q'} = 1$ . Integrating (18) w.r.t. a test function  $f \in L^{q'}(\mathbb{R}^2)$ , one again starts from

$$\left| \int p_t^\varepsilon(x)f(x)dx \right| \leq \|f\|_{L^{q'}(\mathbb{R}^2)} \left( \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} + \sum_{i=1}^2 \int_0^t \|\nabla_i g_{t-s} * (p_s^\varepsilon b_s^{\varepsilon,i})\|_{L^q(\mathbb{R}^2)} ds \right). \quad (25)$$

Let us fix  $i \in \{1, 2\}$ ,  $s < t$  and denote  $A_s^i := \|\nabla_i g_{t-s} * (p_s^\varepsilon b_s^{\varepsilon,i})\|_{L^q(\mathbb{R}^2)}$ . Observe that  $\frac{1}{q'} + \frac{2}{q} = 1 + \frac{1}{q}$ . Apply the convolution inequality (14) and then use (10). It comes

$$A_s^i \leq \|\nabla_i g_{t-s}\|_{L^{q'}(\mathbb{R}^2)} \|p_s^\varepsilon b_s^{\varepsilon,i}\|_{L^{\frac{q}{2}}(\mathbb{R}^2)} \leq \frac{C_1(q') \|b_s^{\varepsilon,i}\|_{L^q(\mathbb{R}^2)} s^{1-\frac{1}{q}} \|p_s^\varepsilon\|_{L^q(\mathbb{R}^2)}}{(t-s)^{\frac{3}{2}-\frac{1}{q'}} s^{1-\frac{1}{q}}} \leq C_1(q') \mathcal{N}_q^\varepsilon(t) \frac{\|b_s^{\varepsilon,i}\|_{L^q(\mathbb{R}^2)}}{(t-s)^{\frac{3}{2}-\frac{1}{q'}} s^{1-\frac{1}{q}}}.$$

In view of Lemma 3.8, (17) and Lemma 3.3, we get

$$\begin{aligned} \|b_s^{\varepsilon,i}\|_{L^q(\mathbb{R}^2)} &\leq \frac{C_2\left(\frac{2q}{q+2}\right)\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)}}{(s+\varepsilon)^{\frac{1}{2}-\frac{1}{q}}} + \chi \int_0^s \|K_{s-u}^{i,\varepsilon}\|_{L^1(\mathbb{R}^2)} \|p_u^\varepsilon\|_{L^q(\mathbb{R}^2)} du \\ &\leq \frac{C_2\left(\frac{2q}{q+2}\right)\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)}}{s^{\frac{1}{2}-\frac{1}{q}}} + \chi C_1(1) \mathcal{N}_q^\varepsilon(t) \int_0^s \frac{1}{\sqrt{s-u} u^{1-\frac{1}{q}}} ds \\ &\leq \frac{C_2\left(\frac{2q}{q+2}\right)\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + \chi C_1(1) \mathcal{N}_q^\varepsilon(t) \beta\left(1 - \frac{1}{q}, \frac{1}{2}\right)}{s^{\frac{1}{2}-\frac{1}{q}}}. \end{aligned}$$

It comes

$$A_s^i \leq C_1(q') \chi \mathcal{N}_q^\varepsilon(t) \frac{C_2\left(\frac{2q}{q+2}\right)\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + C_1(1) \mathcal{N}_q^\varepsilon(t) \beta\left(1 - \frac{1}{q}, \frac{1}{2}\right)}{(t-s)^{\frac{3}{2}-\frac{1}{q'}} s^{\frac{3}{2}-\frac{2}{q}}}.$$

Plug this into (22). The condition  $q \in (2, 4)$  ensures that  $\frac{3}{2} - \frac{2}{q} < 1$  and  $\frac{3}{2} - \frac{1}{q'} < 1$ . Thus, Lemma 3.3 leads to

$$\begin{aligned} \left| \int p_t^\varepsilon(x)f(x)dx \right| &\leq \|f\|_{L^{q'}(\mathbb{R}^2)} \left( \|g_t * \rho_0\|_{L^q(\mathbb{R}^2)} \right. \\ &\quad \left. + 2C_1(q') \chi \mathcal{N}_q^\varepsilon(t) \frac{C_2\left(\frac{2q}{q+2}\right)\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + C_1(1) \mathcal{N}_q^\varepsilon(t) \beta\left(1 - \frac{1}{q}, \frac{1}{2}\right)}{t^{1-\frac{1}{q}}} \beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right) \right). \end{aligned}$$

Take  $\sup_{\|f\|_{L^{q'}=1}}$  in the preceding inequality. It follows from the convolution inequality (13) and (11) that

$$\|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)} \leq \frac{C_2(q)}{t^{1-\frac{1}{q}}} + 2C_1(q') \chi \beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right) \mathcal{N}_q^\varepsilon(t) \frac{C_2\left(\frac{2q}{q+2}\right)\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + C_1(1) \mathcal{N}_q^\varepsilon(t) \beta\left(1 - \frac{1}{q}, \frac{1}{2}\right)}{t^{1-\frac{1}{q}}}.$$

Let us denote

$$K_1 := 2C_1(q')C_1(1)\beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right)\beta\left(1 - \frac{1}{q}, \frac{1}{2}\right) \quad \text{and} \quad K_2 := 2C_1(q')C_2\left(\frac{2q}{q+2}\right)\beta\left(\frac{3}{2} - \frac{2}{q}, \frac{3}{2} - \frac{1}{q'}\right).$$

After rearranging the terms,

$$0 \leq K_1 \chi (\mathcal{N}_q^\varepsilon(t))^2 + (K_2 \chi \|\nabla c_0\|_{L^2(\mathbb{R}^2)} - 1) \mathcal{N}_q^\varepsilon(t) + C_2(q). \quad (26)$$

Under the assumptions

$$K_2\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} - 1 < 0 \text{ and } (K_2\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} - 1)^2 - 4K_1C_2(q)\chi > 0,$$

the polynomial function

$$P(z) = K_1\chi z^2 + (K_2\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} - 1)z + C_2(q)$$

admits two positive roots. In view of Lemma 3.10 and (26), one has that  $\lim_{t \rightarrow 0} \mathcal{N}_q^\varepsilon(t) = 0$  and  $P(\mathcal{N}_q^\varepsilon(t)) > 0$  for any  $t \in [0, T]$ . Necessarily, for any  $t \in [0, T]$   $\mathcal{N}_q^\varepsilon(t)$  is bounded from above by the smaller root of the polynomial function  $P(z)$ . As the constants do not depend on  $T$  and  $\varepsilon$ , this estimate is uniform in time and does not depend on the regularization parameter.

Notice that the above condition is equivalent to

$$K_2\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} + 2\sqrt{K_1C_2(q)\chi} < 1.$$

Denote  $A := K_2$  and  $B := 2\sqrt{C_2(q)K_1}$  to finish the proof.  $\square$

**Remark 3.12.** *The upper bound  $C$  of  $\mathcal{N}_q^\varepsilon(t)$  is given by*

$$C = \frac{1 - A\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)} - \sqrt{(1 - A\chi\|\nabla c_0\|_{L^2(\mathbb{R}^2)})^2 - B^2\chi}}{2K_1\chi}.$$

Now, we will continue analyzing  $\mathcal{N}_r^\varepsilon(t)$ , for different values of  $r$  w.r.t. the  $q \in (2, 4)$  fixed in Proposition 3.11. We will see that different arguments are used when  $r < q$  and  $r > q$ . The result obtained for  $r < q$  will be used to control  $\|b_t^\varepsilon\|_{L^r(\mathbb{R}^2)}$ , for  $r \geq 2$ .

**Corollary 3.13.** *Same assumptions as in Proposition 3.11. Then, for  $1 < r < q$ , it holds*

$$\forall 0 < \varepsilon < 1, \quad \mathcal{N}_r^\varepsilon(T) \leq C_r.$$

*Proof.* Let  $1 < r < q$ . Define  $\theta := \frac{1-\frac{1}{r}}{1-\frac{1}{q}}$ . Then,  $\frac{1}{r} = 1 - \theta + \frac{\theta}{q}$ . As  $p_t^\varepsilon \in L^1(\mathbb{R}^2)$ , "interpolation inequalities" (see [3, p. 93]) lead to

$$\|p_t^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq \|p_t^\varepsilon\|_{L^1(\mathbb{R}^2)}^{1-\theta} \|p_t^\varepsilon\|_{L^q(\mathbb{R}^2)}^\theta \leq \frac{C^\theta}{t^{\theta(1-\frac{1}{q})}} =: \frac{C_r}{t^{1-\frac{1}{r}}}.$$

$\square$

**Corollary 3.14.** *Same assumptions as in Proposition 3.11. Then, for  $2 \leq r \leq \infty$ ,*

$$\forall 0 < \varepsilon < 1, \quad \|b_t^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq \frac{C_r(\chi, \|\nabla c_0\|_{L^2(\mathbb{R}^2)})}{t^{\frac{1}{2}-\frac{1}{r}}}$$

*Proof.* In view of Lemma 3.8, one has for  $i \in \{1, 2\}$

$$\|b_t^{i,\varepsilon}\|_{L^r(\mathbb{R}^2)} \leq \frac{C(\chi, \|\nabla c_0\|_{L^2(\mathbb{R}^2)})}{t^{\frac{1}{2}-\frac{1}{r}}} + \chi \int_0^t \|K_{t-s}^{\varepsilon,i} * p_s^\varepsilon\|_{L^r(\mathbb{R}^2)} ds. \quad (27)$$

Let  $q \in (2, 4)$  fixed in Proposition 3.11.

a) For  $r \in [2, q)$ , Corollary 3.13 immediately implies

$$\|K_{t-s}^{\varepsilon,i} * p_s^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq \frac{C}{\sqrt{t-s} s^{1-\frac{1}{r}}}.$$

b) For  $q \leq r \leq \infty$ , choose  $p_1$  such that  $\frac{1}{p_1} := 1 + \frac{1}{r} - \frac{1}{q}$ . Notice that, as  $2 < q \leq r$ , it follows that  $\frac{1}{2} < \frac{1}{p_1} \leq 1$ . Applying the convolution inequality (14) and Corollary 3.13, one has

$$\|K_{t-s}^{\varepsilon,i} * p_s^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq \frac{C}{(t-s)^{\frac{3}{2}-\frac{1}{p_1}} s^{1-\frac{1}{q}}}.$$

To finish the proof, in both cases, one plugs the obtained estimates in (27) and applies Lemma 3.3.  $\square$

**Corollary 3.15.** *Same assumptions as in Proposition 3.11. Then, for  $q < r < \infty$ , one has*

$$\forall 0 < \varepsilon < 1, \quad \mathcal{N}_r^\varepsilon(T) \leq C_r.$$

*Proof.* Let  $1 < q_1, q_2 < 2$  such that  $\frac{1}{q_1} = \frac{1}{q_2} = \frac{1}{2} + \frac{1}{2r}$ . Then,  $1 + \frac{1}{r} = \frac{1}{q_1} + \frac{1}{q_2}$ . Convolution inequality (14) leads to

$$\|p_t^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq \|g_t * \rho_0\|_{L^r(\mathbb{R}^2)} + \sum_{i=1}^2 \int_0^t \|\nabla_i g_{t-s}\|_{L^{q_1}(\mathbb{R}^2)} \|p_s^\varepsilon b_s^{\varepsilon,i}\|_{L^{q_2}(\mathbb{R}^2)} ds.$$

Let us apply Hölder's inequality for  $\frac{1}{\lambda_1} + \frac{1}{\lambda_2} = 1$  such that  $\lambda_1 = \frac{q}{2}$ ,

$$\|p_s^\varepsilon b_s^{\varepsilon,i}\|_{L^{q_2}(\mathbb{R}^2)} \leq \|p_s^\varepsilon\|_{L^{\lambda_1 q_2}(\mathbb{R}^2)} \|b_s^{\varepsilon,i}\|_{L^{\lambda_2 q_2}(\mathbb{R}^2)}.$$

Notice that  $1 < \lambda_1 < 2$  since  $2 < q < 4$  by hypothesis. Then,  $\lambda_2 > 2$ , thus  $\lambda_2 q_2 > 2$ . In addition,  $\lambda_1 q_2 = \frac{q}{2} q_2 < q$ . In view of Corollaries 3.13 and 3.14, one has

$$\|p_s^\varepsilon b_s^{\varepsilon,i}\|_{L^{q_2}(\mathbb{R}^2)} \leq \frac{C}{s^{1 - \frac{1}{\lambda_1 q_2} + \frac{1}{2} - \frac{1}{\lambda_2 q_2}}} = \frac{C}{s^{\frac{3}{2} - \frac{1}{q_2}}}.$$

Therefore,

$$t^{1 - \frac{1}{q}} \|p_t^\varepsilon\|_{L^r(\mathbb{R}^2)} \leq C + t^{1 - \frac{1}{q}} \int_0^t \frac{C}{(t-s)^{\frac{3}{2} - \frac{1}{q_1}} s^{\frac{3}{2} - \frac{1}{q_2}}} ds.$$

Apply Lemma 3.3 to finish the proof.  $\square$

Notice that the choice of the constants  $A$  and  $B$  depends only on the initially chosen  $q \in (2, 4)$ . One may analyze the constants in Condition (24) in function of  $q$  to get an optimal condition on  $\chi$ .

## 4 Proof of Theorem 2.3

### 4.1 Tightness

**Proposition 4.1.** *Let  $T > 0$ . Denote  $\varepsilon_k = \frac{1}{k}$ , for  $k \in \mathbb{N}$ .  $\mathbb{P}^k$  denotes the law of the solutions to (15) regularized with  $\varepsilon_k$ . If the initial law  $\rho_0$  is a probability density,  $\nabla c_0 \in L^2(\mathbb{R}^2)$  and  $\chi > 0$  are such that Condition (24) is satisfied, then the probability laws  $(\mathbb{P}^k)_{k \geq 1}$  are tight in  $C([0, T]; \mathbb{R}^2)$  w.r.t.  $k \in \mathbb{N}$ .*

*Proof.* For  $m > 2$  and  $0 < s < t \leq T$ , notice that

$$\mathbb{E}|X_t^\varepsilon - X_s^\varepsilon|^m \leq \mathbb{E} \left( \left( \int_s^t b^{\varepsilon,1}(u, X_u^\varepsilon) du \right)^2 + \left( \int_s^t b^{\varepsilon,2}(u, X_u^\varepsilon) du \right)^2 \right)^{\frac{m}{2}} + \mathbb{E}|W_t - W_s|^m.$$

In view of the drift estimate for  $r = \infty$  in Corollary 3.14, one has

$$\mathbb{E}|X_t^\varepsilon - X_s^\varepsilon|^m \leq \left( 2 \int_s^t \frac{C(\chi, \|\nabla c_0\|_{L^2(\mathbb{R}^2)})}{\sqrt{u}} du \right)^m + C(t-s)^{\frac{m}{2}} \leq C(\chi, \|\nabla c_0\|_{L^2(\mathbb{R}^2)})(t-s)^{\frac{m}{2}}.$$

Kolmogorov's criterion implies tightness.  $\square$

### 4.2 Existence

In order to prove the existence of a weak solution, we will prove that the following non-linear martingale problem related to (3) admits a solution under the hypothesis of Theorem 2.3.

**Definition 4.2.** *A probability measure  $\mathbb{Q}$  on the canonical space  $\mathcal{C}([0, T]; \mathbb{R}^2)$  equipped with its canonical filtration and a canonical process  $(w_t)$  is a solution to the non-linear martingale problem (MP) if:*

(i)  $\mathbb{Q}_0 = \rho_0$ .

(ii) *For any  $t \in (0, T]$ , the one dimensional time marginals of  $\mathbb{Q}$ , denoted by  $\mathbb{Q}_t$ , have densities  $q_t$  w.r.t. Lebesgue measure on  $\mathbb{R}$ . In addition, they satisfy*

$$\forall r \in (1, \infty) \exists C > 0 : \quad \sup_{t \in (0, T)} t^{1 - \frac{1}{r}} \|q_t\|_{L^r(\mathbb{R}^2)} \leq C.$$

(iii) For any  $f \in C_K^2(\mathbb{R}^2)$  the process  $(M_t)_{t \leq T}$ , defined as

$$M_t := f(w_t) - f(w_0) - \int_0^t \left[ \frac{1}{2} \Delta f(w_u) + \nabla f(w_u) \cdot (b_0(u, w_u) + \chi \int_0^u \int K_{u-\tau}(w_u - y) q_\tau(y) dy d\tau) \right] du$$

is a  $\mathbb{Q}$ -martingale.

In view of Proposition 4.1, there exists a weakly convergent subsequence of  $(\mathbb{P}^k)_{k \geq 1}$  that we will still denote by  $(\mathbb{P}^k)_{k \geq 1}$ . Denote its limit by  $\mathbb{P}^\infty$ . Let us prove that  $\mathbb{P}^\infty$  solves the martingale problem (MP).

- i) Each  $\mathbb{P}_0^k$  has density  $\rho_0$ , and therefore  $\mathbb{P}_0^\infty$  also has density  $\rho_0$ .
- ii) Define the functional  $\Lambda_t(\varphi)$  by

$$\Lambda_t(\varphi) := \int_{\mathbb{R}^2} \varphi(y) \mathbb{P}_t^\infty(dy), \quad \varphi \in C_K(\mathbb{R}^2).$$

By weak convergence we have

$$\Lambda_t(\varphi) = \lim_{k \rightarrow \infty} \int \varphi(y) p_t^k(y) dy,$$

and thus for any  $1 < r < \infty$  and its conjugate  $r'$ , in view of Proposition 3.11 and Corollaries 3.13 and 3.15 one has

$$|\Lambda_t(\varphi)| \leq \frac{C}{t^{1-\frac{1}{r}}} \|\varphi\|_{L^{r'}(\mathbb{R}^2)}.$$

Therefore, for each  $0 < t \leq T$ ,  $\Lambda_t$  is a bounded linear functional on a dense subset of  $L^{r'}(\mathbb{R}^2)$ . Thus,  $\Lambda_t$  can be extended to a linear functional on  $L^{r'}(\mathbb{R}^2)$ . By Riesz-representation theorem (e.g. [3, Thm. 4.11 and 4.14]), there exists a unique  $p_t^\infty \in L^r(\mathbb{R}^2)$  such that  $\|p_t^\infty\|_{L^r(\mathbb{R}^2)} \leq \frac{C}{t^{1-\frac{1}{r}}}$  and  $p_t^\infty$  is the probability density of  $\mathbb{P}_t^\infty(dy)$ .

iii) Set

$$M_t^\infty := f(w_t) - f(w_0) - \int_0^t \left[ \Delta f(w_u) + \nabla f(w_u) \cdot (b_0(u, w_u) + \frac{\chi}{8\pi} \int_0^u e^{-\lambda(u-\tau)} \int K_{u-\tau}(w_u - y) p_\tau^\infty(y) dy d\tau) \right] du.$$

In order to prove that  $(M_t^\infty)_{t \leq T}$  is a  $\mathbb{P}^\infty$  martingale, we will check that for any  $N \geq 1$ ,  $0 \leq t_1 < \dots < t_N < s \leq t \leq T$  and any  $\phi \in C_b((\mathbb{R}^2)^N)$ , one has

$$\mathbb{E}_{\mathbb{P}^\infty} [(M_t^\infty - M_s^\infty) \phi(w_{t_1}, \dots, w_{t_N})] = 0. \quad (28)$$

As  $\mathbb{P}^k$  solves the non-linear martingale problem related to (15) with  $\varepsilon_k = \frac{1}{k}$ , one has

$$M_t^k := f(w_t) - f(x(0)) - \chi \int_0^t \left[ \Delta f(w_u) + \nabla f(w_u) \cdot (b_0^{\varepsilon_k}(u, w_u) + \frac{\chi}{8\pi} \int_0^u e^{-\lambda(u-\tau)} (K_{u-\tau}^{\varepsilon_k} * p_\tau^k)(w_u) d\tau) \right] du$$

is a martingale under  $\mathbb{P}^k$ . Thus,

$$\begin{aligned} 0 &= \mathbb{E}_{\mathbb{P}^k} [(M_t^k - M_s^k) \phi(w_{t_1}, \dots, w_{t_N})] = \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) (f(w_t) - f(w_s))] \\ &+ \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \Delta f(w_u) du] + \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0^{\varepsilon_k}(u, w_u) du] \\ &+ \frac{\chi}{8\pi} \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u e^{-\lambda(u-\tau)} (K_{u-\tau}^{\varepsilon_k} * p_\tau^k)(w_u) d\tau du]. \end{aligned}$$

Since  $(\mathbb{P}^k)$  weakly converges to  $\mathbb{P}^\infty$ , the first two terms on the r.h.s. converge to their analogues in (28). It remains to check the convergence of the last two terms. We will analyze separately the parts coming from the linear and non-linear drifts.

**Linear part** Notice that for  $t > 0$  and  $x \in \mathbb{R}^2$

$$|b_0^{\varepsilon_k}(t, x) - b_0(t, x)| \leq C \frac{\chi e^{-\lambda t} \varepsilon_k}{t(t + \varepsilon_k)} \left| \int_{\mathbb{R}^2} \nabla c_0(x - y) e^{-\frac{|y|^2}{2t}} dy \right| \leq \frac{\varepsilon_k \|\nabla c_0\|_{L^2(\mathbb{R}^2)}}{\sqrt{t(t + \varepsilon_k)}}.$$

Thus,  $\|b_0^{\varepsilon_k}(t, \cdot) - b_0(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \rightarrow 0, k \rightarrow \infty$  and from Lemmas 3.7 and 3.8 we have

$$\|b_0^{\varepsilon_k}(t, \cdot) - b_0(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \frac{C}{\sqrt{t}}.$$

Similarly, for  $t > 0$  and  $r > 2$ ,

$$\|b_0^{\varepsilon_k}(t, \cdot) - b_0(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \|\nabla c_0\|_{L^2(\mathbb{R}^2)} \frac{\varepsilon_k}{t(t + \varepsilon_k)} C t^{\frac{1}{r} + \frac{1}{2}}.$$

Therefore,

$$\|b_0^{\varepsilon_k}(t, \cdot) - b_0(t, \cdot)\|_{L^r(\mathbb{R}^2)} \rightarrow 0, k \rightarrow \infty \quad (29)$$

and from Lemmas 3.7 and 3.8 we have

$$\|b_0^{\varepsilon_k}(t, \cdot) - b_0(t, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \frac{C}{t^{\frac{1}{2} - \frac{1}{r}}}. \quad (30)$$

Now, observe that

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}^k}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0^{\varepsilon_k}(u, w_u) du] - \mathbb{E}_{\mathbb{P}^\infty}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0(u, w_u) du] \\ &= (\mathbb{E}_{\mathbb{P}^k}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0^{\varepsilon_k}(u, w_u) du] - \mathbb{E}_{\mathbb{P}^k}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0(u, w_u) du]) \\ &+ (\mathbb{E}_{\mathbb{P}^k}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0(u, w_u) du] - \mathbb{E}_{\mathbb{P}^\infty}[\phi(\dots) \int_s^t \nabla f(w_u) \cdot b_0(u, w_u) du]) =: I_k + II_k. \end{aligned}$$

We start from  $II_k$ . Define for  $x \in C([0, T]; \mathbb{R}^2)$  the functional

$$F(x) := \phi(x_{t_1}, \dots, x_{t_N}) \int_s^t \nabla f(x_u) \cdot b_0(u, x_u) du.$$

In view of Lemma 3.7, for  $u > 0$  and  $i = 1, 2$ , the function  $b_0^i(u, \cdot)$  is bounded and continuous on  $\mathbb{R}^2$  and one has  $\|b_0^i(t, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \frac{C}{\sqrt{t}}$ . By dominated convergence one gets that  $F(\cdot)$  is continuous. In addition,  $F(\cdot)$  is bounded on  $C([0, T]; \mathbb{R}^2)$ . Thus, by weak convergence,  $II_k \rightarrow 0$ , as  $k \rightarrow \infty$ .

We turn to  $I_k$ :

$$|I_k| \leq \|\phi\|_\infty \int_s^t \sum_{i=1}^2 \int_{\mathbb{R}^2} |\nabla_i f(z)(b_0^{\varepsilon_k, i}(u, z) - b_0^i(u, z))| p_u^k(z) dz ds.$$

Apply the Hölder's inequality for  $\frac{1}{\lambda} + \frac{1}{\lambda'} = 1$  such that  $1 < \lambda < 2$ . In view of Corollary 3.13, one has

$$|I_k| \leq \|\phi\|_\infty \|\nabla f\|_\infty \int_s^t \frac{C}{u^{1-\frac{1}{\lambda}}} \sum_{i=1}^2 \|b_0^{\varepsilon_k, i}(u, \cdot) - b_0^i(u, \cdot)\|_{L^{\lambda'}(\mathbb{R}^2)} du.$$

In view of (29),  $\|b_0^{\varepsilon_k, i}(u, \cdot) - b_0^i(u, \cdot)\|_{L^{\lambda'}(\mathbb{R}^2)} \rightarrow 0$  as  $k \rightarrow \infty$ . In addition (30) leads to

$$\frac{C}{u^{1-\frac{1}{\lambda}}} \sum_{i=1}^2 \|b_0^{\varepsilon_k, i}(u, \cdot) - b_0^i(u, \cdot)\|_{L^{\lambda'}(\mathbb{R}^2)} \leq \frac{C}{u^{\frac{1}{\lambda'} + \frac{1}{2} - \frac{1}{\lambda'}}}.$$

By dominated convergence,  $I_k \rightarrow 0$ , as  $k \rightarrow \infty$ .



**Non-linear part** Let us first analyze the difference of the two drifts. Fix  $0 < s < t$ ,  $x \in \mathbb{R}^2$  and  $i \in 1, 2$ . Notice that

$$|K_{t-s}^{\varepsilon_k, i}(x) - K_{t-s}^i(x)| \leq \frac{(t-s)\varepsilon_k + \varepsilon_k^2}{(t-s)^2(t-s+\varepsilon_k)^2} |x^i| e^{\frac{|x|^2}{2(t-s)}}.$$

Thus for any  $x \in \mathbb{R}^2$  and any  $0 < s < t$ , we have that  $|K_{t-s}^{\varepsilon_k, i}(x) - K_{t-s}^i(x)| \rightarrow 0, k \rightarrow \infty$ . After integration, for any  $1 < r < 2$  one has

$$\|K_{t-s}^{\varepsilon_k, i} - K_{t-s}^i\|_{L^r(\mathbb{R}^2)} \leq C_r \frac{(t-s)\varepsilon_k + \varepsilon_k^2}{(t-s)^2(t-s+\varepsilon_k)^2} (t-s)^{\frac{1}{2} + \frac{1}{r}}.$$

Therefore, for any  $0 < s < t$ , one gets

$$\|K_{t-s}^{\varepsilon_k, i} - K_{t-s}^i\|_{L^r(\mathbb{R}^2)} \rightarrow 0, k \rightarrow \infty. \quad (31)$$

In addition, (10) and (17) lead to

$$\|K_{t-s}^{\varepsilon_k, i} - K_{t-s}^i\|_{L^r(\mathbb{R}^2)} \leq \|K_{t-s}^{\varepsilon_k, i}\|_{L^r(\mathbb{R}^2)} + \|K_{t-s}^i\|_{L^r(\mathbb{R}^2)} \leq \frac{C_r}{(t-s)^{\frac{3}{2} - \frac{1}{r}}}. \quad (32)$$

For  $t > 0$ ,  $x \in \mathbb{R}^2$  and  $i \in 1, 2$ , one has

$$\begin{aligned} & \left| \int_0^t (K_{t-s}^{\varepsilon_k, i} * p_s^k)(x) ds - \int_0^t (K_{t-s}^i * p_s^\infty)(x) ds \right| \leq \left| \int_0^t (K_{t-s}^{\varepsilon_k, i} * p_s^k)(x) ds - \int_0^t (K_{t-s}^i * p_s^k)(x) ds \right| \\ & + \left| \int_0^t (K_{t-s}^i * p_s^k)(x) ds - \int_0^t (K_{t-s}^i * p_s^\infty)(x) ds \right| =: A_k + B_k. \end{aligned}$$

We start from  $B_k$ . For  $s < t$  and  $i = 1, 2$ , the kernel  $K_{t-s}^i(\cdot)$  is a continuous and bounded function on  $\mathbb{R}^2$ . Thus, by weak convergence we have that  $\lim_{k \rightarrow \infty} (K_{t-s}^i * p_s^k)(x) = (K_{t-s}^i * p_s^\infty)(x)$ . In addition, for  $r > 2$  Hölder's inequality, part *ii*) and Proposition 3.11 lead to

$$|(K_{t-s}^i * p_s^k)(x) - (K_{t-s}^i * p_s^\infty)(x)| \leq \frac{C_r}{(t-s)^{\frac{3}{2} - \frac{1}{r}} s^{1 - \frac{1}{r}}}.$$

As the bound is integrable in  $(0, t)$ , the dominated convergence theorem implies that  $B_k \rightarrow 0$ , as  $k \rightarrow \infty$ . In  $A_k$  we apply the Hölder's inequality with  $1 < r < 2$  and the density bounds from Corollary 3.13. It comes

$$|A_k| \leq \int_0^t \|K_{t-s}^{\varepsilon_k, i} - K_{t-s}^i\|_{L^{r'}(\mathbb{R}^2)} \frac{C_r}{s^{1 - \frac{1}{r}}} ds.$$

In view of (31) and (32), one can apply the dominated convergence. Thus,  $A_k \rightarrow 0$ , as  $k \rightarrow \infty$ . Finally, we obtain

$$\lim_{k \rightarrow \infty} \left| \int_0^t (K_{t-s}^{\varepsilon_k, i} * p_s^k)(x) ds - \int_0^t (K_{t-s}^i * p_s^\infty)(x) ds \right| = 0. \quad (33)$$

As in the linear part, we decompose

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau}^{\varepsilon_k} * p_\tau^k)(w_u) d\tau du] \\ & - \mathbb{E}_{\mathbb{P}^\infty} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau} * p_\tau^\infty)(w_u) d\tau du] \\ & \leq \left( \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau}^{\varepsilon_k} * p_\tau^k)(w_u) d\tau du] \right. \\ & \left. - \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau} * p_\tau^\infty)(w_u) d\tau du] \right) \\ & + \left( \mathbb{E}_{\mathbb{P}^k} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau} * p_\tau^\infty)(w_u) d\tau du] \right. \\ & \left. - \mathbb{E}_{\mathbb{P}^\infty} [\phi(\dots) \int_s^t \nabla f(w_u) \cdot \int_0^u (K_{u-\tau} * p_\tau^\infty)(w_u) d\tau du] \right) \\ & =: C_k + D_k. \end{aligned}$$

Start from  $D_k$ . Similarly to the linear part, we need the boundness and continuity of the functional

$$H(x) := \phi(x(t_1), \dots, x(t_N)) \int_s^t \nabla f(x(u)) \cdot \int_0^u (K_{u-\tau} * p_\tau^\infty)(x(u)) d\tau du, \quad x \in \mathcal{C}([0, T]; \mathbb{R}^2).$$

The continuity comes from the fact that the kernel is a continuous function on  $\mathbb{R}^2$  whenever  $\tau < u$ . Namely, if  $x_n \in \mathcal{C}([0, T]; \mathbb{R}^2)$  converges to  $x \in \mathcal{C}([0, T]; \mathbb{R}^2)$ , then  $K_{u-\tau}^i(x_n(u) - y) \rightarrow K_{u-\tau}^i(x(u) - y)$ . In addition  $|K_{u-\tau}^i(x_n(u) - y)p_\tau^\infty(y)| \leq \frac{C}{(u-\tau)^{3/2}} p_\tau^\infty(y)$ , for  $i \in \{1, 2\}$ , as  $n \rightarrow \infty$ . Thus, by dominated convergence, for  $\tau < u$  one has

$$K_{u-\tau}^i * p_\tau^\infty(x_n(u)) \rightarrow K_{u-\tau}^i * p_\tau^\infty(x(u)), \quad n \rightarrow \infty.$$

For  $\frac{1}{r} + \frac{1}{r'} = 1$  such that  $r > 2$  apply Hölder's inequality and after the estimate in *ii*). It comes

$$|K_{u-\tau}^i * p_\tau^\infty(x_n(u))| \leq \frac{C_r}{(u-\tau)^{\frac{3}{2}-\frac{1}{r'}} \tau^{1-\frac{1}{r}}}.$$

By dominated convergence,

$$\int_0^u (K_{u-\tau}^i * p_\tau^\infty)(x_n(u)) d\tau \rightarrow \int_0^u (K_{u-\tau}^i * p_\tau^\infty)(x(u)) d\tau, \quad n \rightarrow \infty.$$

Moreover, in view of Lemma 3.3, one has

$$\left| \nabla f(x_n(u)) \cdot \int_0^u K_{u-\tau} * p_\tau^\infty(x_n(u)) d\tau \right| \leq C \|\nabla f\|_\infty \frac{\beta(1 - \frac{1}{r}, \frac{3}{2} - \frac{1}{r'})}{\sqrt{u}}.$$

Finally, after one more application of dominated convergence the continuity of the functional  $H$  follows. This procedure obviously implies  $H$  is a bounded functional on  $\mathcal{C}([0, T]; \mathbb{R}^2)$ . Thus, by weak convergence,  $D_k$  converges to zero.

We turn to  $C_k$ . Let us just for this part denote by  $b^i(u, z) := \int_0^u K_{u-\tau}^i * p_\tau^\infty(z) d\tau$  and  $b^{k,i}(u, z) := \int_0^u K_{u-\tau}^{\varepsilon_k, i} * p_\tau^\infty(z) d\tau$ . Notice that

$$|C_k| \leq \|\phi\|_\infty \int_s^t \sum_{i=1}^2 \int_{\mathbb{R}^2} |\nabla_i f(z) (b^{k,i}(u, z) - b^i(u, z))| p_u^k(z) dz.$$

After Hölder inequality for  $\frac{1}{r} + \frac{1}{r'} = 1$  such that  $r > 2$ , one has

$$|C_k| \leq \|\phi\|_\infty \int_s^t \frac{C}{u^{1-\frac{1}{r'}}} \sum_{i=1}^2 \left( \int |\nabla_i f(z)|^r |b^{k,i}(u, z) - b^i(u, z)|^r dz \right)^{1/r} du.$$

Let  $u > 0$ . In view of (33),  $|b^{k,i}(u, z) - b^i(u, z)|^r \rightarrow 0$  as  $k \rightarrow \infty$ . Now, we do not omit  $|\nabla_i f(z)|^q$  as in the linear part. Instead, we use it in order to integrate in space with respect to drift bounds. Namely, for  $u > 0$  and  $i = 1, 2$ , we have seen that  $|b^{k,i}(u, \cdot)| + |b^i(u, \cdot)| \leq \frac{C}{\sqrt{u}}$ . Thus,

$$|\nabla_i f(z)|^r |b^{k,i}(u, z) - b^i(u, z)|^r \leq \frac{C}{u^{\frac{r}{2}}} |\nabla_i f(z)|^r.$$

By dominated convergence,

$$\|\nabla_i f(\cdot) (b^{k,i}(u, \cdot) - b^i(u, \cdot))\|_{L^r(\mathbb{R}^2)} \rightarrow 0, \quad k \rightarrow \infty.$$

Using that  $\|b^{k,i}(u, \cdot)\|_{L^r(\mathbb{R}^2)} + \|b^i(u, \cdot)\|_{L^r(\mathbb{R}^2)} \leq \frac{C}{u^{\frac{1}{2}-\frac{1}{r}}}$ , one gets

$$\frac{1}{u^{1-\frac{1}{r'}}} \|\nabla_i f(\cdot) (b^{k,i}(u, \cdot) - b^i(u, \cdot))\|_{L^r(\mathbb{R}^2)} \leq \|\nabla_i f(\cdot)\|_\infty \frac{C}{u^{\frac{1}{2}-\frac{1}{r}+1-\frac{1}{r'}}$$

Thus, by dominated convergence, we get that  $C_k \rightarrow 0$ , as  $k \rightarrow \infty$ .

As all the terms converge, we get that (28) holds true. Thus, the process  $(M_t^\infty)_{t \leq T}$  is a  $\mathbb{P}^\infty$  martingale.

## 5 Application to the two-dimensional Keller-Segel model

In this section we prove Corollary 2.5. The parameter  $\lambda$  does not play any role in the above results. Therefore, we will assume here  $\lambda = 0$ . It is easy to extend the following arguments for  $\lambda > 0$ .

Denote by  $\rho(t, \cdot) \equiv p_t(x)$  the time marginals of the probability distribution constructed in Theorem 2.3. As such,  $\rho$  satisfies for any  $1 \leq q < \infty$ ,

$$\sup_{t \leq T} t^{1-\frac{1}{q}} \|\rho(t, \cdot)\|_{L^q(\mathbb{R}^2)} \leq C_q.$$

The corresponding drift function satisfies for any  $1 \leq r \leq \infty$ ,

$$t^{\frac{1}{2}-\frac{1}{r}} \|b(t, \cdot; \rho)\|_{L^r(\mathbb{R}^2)} \leq C_r.$$

Following the arguments in Proposition [20] one may derive the mild equation for  $\rho(t, \cdot)$ . The above estimates ensure that everything is well defined. Thus, one arrives to the following: for any  $f \in C_K^\infty(\mathbb{R}^2)$  and any  $t \in (0, T]$ ,

$$\begin{aligned} \int f(y) \rho(t, y) dy &= \int f(y) (g_t * \rho_0)(y) dy \\ &- \sum_{i=1}^2 \int f(y) \int_0^t [\nabla_i g_{t-s} * (b^i(s, \cdot; \rho) \rho(s, \cdot))](y) ds dy. \end{aligned}$$

Thus  $\rho$  satisfies in the sense of the distributions

$$\rho(t, \cdot) = g_t * \rho_0 - \sum_{i=1}^2 \int_0^t \nabla_i g_{t-s} * (b^i(s, \cdot; \rho) \rho(s, \cdot)) ds. \quad (34)$$

Now, define the function  $c(t, x)$  as

$$c(t, x) := (g(t, \cdot) * c_0)(x) + \int_0^t \rho(t-s, \cdot) * g(s, \cdot)(x) ds.$$

Thanks to the density estimates  $c(t, x)$  is well defined for all  $x \in \mathbb{R}^2$  as soon as  $t > 0$ . Indeed,

$$|c(t, x)| \leq \frac{\|c_0\|_{L^2(\mathbb{R}^2)}}{\sqrt{t}} + C \int_0^t \|\rho(t-s, \cdot)\|_{L^2(\mathbb{R}^2)} \|g_s\|_{L^2(\mathbb{R}^2)} ds \leq \frac{\|c_0\|_{L^2(\mathbb{R}^2)}}{\sqrt{t}} + C\beta\left(\frac{1}{2}, \frac{1}{2}\right).$$

It is obvious that  $c(t, \cdot) \in L^2(\mathbb{R}^2)$ . Thanks to the density estimates and the fact that  $g_t$  is strongly derivable as soon as  $t > 0$ ,  $c(t, x)$  is derivable in any point  $x$  and

$$\frac{\partial}{\partial x_i} c(t, x) = \nabla_i (g(t, \cdot) * c_0)(x) + \int_0^t (\rho_{t-s} * \nabla_i g(s, \cdot))(x) ds.$$

The fact that  $c_0 \in H^1(\mathbb{R}^2)$  enables us to write  $\nabla_i (g(t, \cdot) * c_0) = (g(t, \cdot) * \nabla_i c_0)$ . Now, remark that  $\chi \frac{\partial}{\partial x_i} c(t, x)$  is exactly the drift in (34). Thus, the couple  $(\rho, c)$  satisfies Definition 2.4.

Assume there exists another couple  $(\rho^1, c^1)$  satisfying Definition 2.4 with the above initial conditions  $(\rho_0, c_0)$ . As such, they satisfy

$$\forall 1 \leq q < \infty \exists C > 0 : \quad \sup_{t \leq T} t^{1-\frac{1}{q}} \|\rho_t^1\|_{L^q(\mathbb{R}^2)} \leq C$$

and

$$\forall 2 \leq r \leq \infty \exists C > 0 : \quad \sup_{t \leq T} t^{\frac{1}{2}-\frac{1}{r}} \|\nabla c_t^1\|_{L^r(\mathbb{R}^2)} \leq C.$$

## 6 Appendix

Let  $T > 0$ . On a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$  equipped with a  $d$ -dimensional Brownian motion  $(W)$  and an  $\mathcal{F}_0$ -measurable random variable  $X_0$ , we study the stochastic equation

$$\begin{cases} dX_t = dW_t + \left\{ \int_0^t \int_{\mathbb{R}^d} L(t-s, X_t - y) \mathbb{Q}_s(dy) ds \right\} dt, & t \leq T, \\ \mathbb{Q}_s := \mathcal{L}(X_s), & X_0 \sim q_0, \end{cases} \quad (35)$$

where  $L$  maps  $[0, T] \times \mathbb{R}^d$  to  $\mathbb{R}^d$ . In this section we show how to adapt the Proof of Theorem 1.1 in [19] in the framework of the additional time interaction in (35).

Firstly, the assumption about the regularity of the interaction kernel in [19] needs to be replaced by the following hypothesis on the interaction  $L$ :

**Hypothesis (H0).** *The function  $L : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies*

$$\begin{aligned} \forall (t, x) \in (0, T) \times \mathbb{R}^d, & \quad |L(t, x)| \leq h_1(t), \\ \forall (t, x, y) \in (0, T) \times \mathbb{R}^d \times \mathbb{R}^d, & \quad |L(t, x) - L(t, y)| \leq h_2(t)|x - y|, \end{aligned}$$

where  $h_i : (0, T) \rightarrow \mathbb{R}^+$  is such that there exists  $D_T > 0$  such that for any  $t \leq T$ , one has  $\int_0^t h_i(s) ds \leq D_T$ .

Notice that the time interaction induces a slight change in (H0) with respect to what is assumed on the interaction kernel in [19]. We still assume the kernel is bounded and Lipschitz in space, but in order to treat the additional integral in time, we introduce the functions  $h_1$  and  $h_2$ .

Let  $C((0, T); \mathbb{R}^d)$  be a set of continuous  $\mathbb{R}^d$ -valued functions defined on  $(0, T)$  and  $\mathcal{P}_T$  be the set of probability measures on  $C((0, T); \mathbb{R}^d)$ . For a  $\mathbb{Q} \in \mathcal{P}_T$  and  $(t, x) \in (0, T) \times \mathbb{R}^d$  denote by

$$b(t, x; (\mathbb{Q}_s)_{s \leq t}) := \int_0^t \int_{\mathbb{R}^d} L(t-s, x-y) \mathbb{Q}_s(dy) ds.$$

In view of Hypothesis (H0), for a given  $\mathbb{Q} \in \mathcal{P}_T$  one has that

$$\forall (t, x, y) \in (0, T) \times \mathbb{R}^d \times \mathbb{R}^d : \begin{cases} |b(t, x; (\mathbb{Q}_s)_{s \leq t})| \leq D_T, \\ |b(t, x; (\mathbb{Q}_s)_{s \leq t}) - b(t, y; (\mathbb{Q}_s)_{s \leq t})| \leq D_T|x - y|. \end{cases} \quad (36)$$

**Theorem 6.1.** *Under the hypothesis (H0), Equation (35) admits a unique strong solution.*

Let us show how the calculations in [19] change in this setting. We adopt the definition of the Wasserstein distance given there.

*Proof of Theorem 6.1.* To prove Theorem 6.1, one should search for a fixed point of the map  $\Phi : \mathcal{P}_T \rightarrow \mathcal{P}_T$  that to a given  $m \in \mathcal{P}_T$  associates the law of the solution to the following SDE:

$$\begin{cases} dX_t = dW_t + b(t, X_t; (m_s)_{s \leq t}) dt, \\ X_0 \sim p_0. \end{cases}$$

Notice that this equation is well-defined in strong sense thanks to (36) (see e.g. [12, Thm. 5.2.9]). To exhibit the fixed point, the following contraction inequality should be demonstrated for  $m_1, m_2 \in \mathcal{P}_T$ :

$$D_{1,t}(\Phi(m_1), \Phi(m_2)) \leq C_T \int_0^t D_{1,u}(m_1, m_2) du.$$

To prove the latter, follow the steps in [19]. Always use (H0) when dealing with the time interaction. Let  $m_1, m_2 \in \mathcal{P}_T$ . For  $i = 1, 2$ , associate to  $m_i$  the law of the solution of

$$X_t^i = X_0 + W_t + \int_0^t \int_0^s b(u, X_u^i; (m_{i,r})_{r \leq u}) du ds.$$

Then,

$$|X_s^1 - X_s^2| \leq \left| \int_0^s \int_0^u \left[ \int_{C((0,T);\mathbb{R}^d)} L(u - \alpha, X_u^1 - w_\alpha) dm_1(w) - \int_{C((0,T);\mathbb{R}^d)} L(u - \alpha, X_u^2 - w_\alpha) dm_2(w) \right] d\alpha du \right| =: F(s).$$

Taking  $\sup_{s \leq t}$  of the previous expression and an expectation on both sides, one has

$$\mathbb{E} [\sup_{s \leq t} |X_s^1 - X_s^2|] \leq \mathbb{E} [\sup_{s \leq t} F(s)]. \quad (37)$$

Taking  $\pi$  to be any coupling of  $m_1$  and  $m_2$ , it comes

$$F(s) = \left| \int_0^s \int_0^u \int_{C((0,T);\mathbb{R}^d) \times C((0,T);\mathbb{R}^d)} [L(u - \alpha, X_u^1 - w_\alpha^1) - L(u - \alpha, X_u^2 - w_\alpha^2)] d\pi(w^1, w^2) d\alpha du \right|.$$

In view of (H0), one has

$$|L(u - \alpha, X_u^1 - w_\alpha^1) - L(u - \alpha, X_u^2 - w_\alpha^2)| \leq (2h_1(u - \alpha) + h_2(u - \alpha)) |X_u^1 - w_\alpha^1 - X_u^2 + w_\alpha^2| \wedge 1.$$

Therefore,

$$\begin{aligned} \mathbb{E} [\sup_{s \leq t} F(s)] &\leq C \mathbb{E} \int_0^t \int_0^u (h_1(u - \alpha) + h_2(u - \alpha)) \\ &\quad \int_{C((0,T);\mathbb{R}^d) \times C((0,T);\mathbb{R}^d)} [|X_u^1 - X_u^2| \wedge 1 + |w_\alpha^1 - w_\alpha^2| \wedge 1] d\pi(w^1, w^2) d\alpha du. \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{E} [\sup_{s \leq t} F(s)] &\leq 2C_T \mathbb{E} \int_0^t |X_u^1 - X_u^2| \wedge 1 du \\ &\quad + \int_0^t \int_0^u (h_1(u - \alpha) + h_2(u - \alpha)) \int_{C((0,T);\mathbb{R}^d) \times C((0,T);\mathbb{R}^d)} |w_\alpha^1 - w_\alpha^2| \wedge 1 d\pi(w^1, w^2) d\alpha du. \end{aligned}$$

Using that  $|X_u^1 - X_u^2| \leq \sup_{r \leq u} |X_r^1 - X_r^2|$  and  $|w_\alpha^1 - w_\alpha^2| \leq \sup_{r \leq u} |w_r^1 - w_r^2|$  and applying Fubini's theorem, one obtains

$$\mathbb{E} [\sup_{s \leq t} F(s)] \leq C_T \left[ \int_0^t \mathbb{E} [\sup_{r \leq u} |X_r^1 - X_r^2| \wedge 1] du + \int_0^t \int_{C((0,T);\mathbb{R}^d) \times C((0,T);\mathbb{R}^d)} \sup_{r \leq u} |w_r^1 - w_r^2| \wedge 1 d\pi(w^1, w^2) du \right].$$

Coming back to (37), one gets

$$\mathbb{E} [\sup_{s \leq t} |X_s^1 - X_s^2|] \leq C_T \left[ \int_0^t \mathbb{E} [\sup_{r \leq u} |X_r^1 - X_r^2| \wedge 1] du + \int_0^t \int_{C((0,T);\mathbb{R}^d) \times C((0,T);\mathbb{R}^d)} \sup_{r \leq u} |w_r^1 - w_r^2| \wedge 1 d\pi(w^1, w^2) du \right].$$

Taking an infimum over all couplings  $\pi$  of  $m_1$  and  $m_2$  of the above expression, leads to

$$\mathbb{E} [\sup_{s \leq t} |X_s^1 - X_s^2|] \leq C_T \int_0^t \mathbb{E} [\sup_{r \leq u} |X_r^1 - X_r^2| \wedge 1] du + C_T \int_0^t D_{1,u}(m_1, m_2) du.$$

Gronwall's lemma implies

$$\mathbb{E} [\sup_{s \leq t} |X_s^1 - X_s^2| \wedge 1] \leq C_T \int_0^t D_{1,u}(m_1, m_2) du. \quad (38)$$

As  $X^1$  and  $X^2$  have laws  $\Phi(m_1)$  and  $\Phi(m_2)$ , respectively, a standard property of the Wasserstein distance together with (38), lead to the contraction inequality

$$D_{1,t}(\Phi(m_1), \Phi(m_2)) \leq C_T \int_0^t D_{1,u}(m_1, m_2) du.$$

Once the contraction inequality is obtained one repeats the arguments in [19] to finish the proof.  $\square$

## References

- [1] BILER, P., CORRIAS, L., AND DOLBEAULT, J. Large mass self-similar solutions of the parabolic-parabolic Keller-Segel model of chemotaxis. *J. Math. Biol.* 63, 1 (2011), 1–32.
- [2] BLANCHET, A., DOLBEAULT, J., AND PERTHAME, B. T. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations* (2006), No. 44, 32.
- [3] BREZIS, H. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.
- [4] BREZIS, H. M., AND CAZENAVE, T. A nonlinear heat equation with singular initial data. *J. Anal. Math.* 68 (1996), 277–304.
- [5] CALVEZ, V., AND CORRIAS, L. The parabolic-parabolic Keller-Segel model in  $\mathbb{R}^2$ . *Commun. Math. Sci.* 6, 2 (2008), 417–447.
- [6] CORRIAS, L., ESCOBEDO, M., AND MATOS, J. Existence, uniqueness and asymptotic behavior of the solutions to the fully parabolic Keller-Segel system in the plane. *J. Differential Equations* 257, 6 (2014), 1840–1878.
- [7] HERRERO, M. A., AND VELÁZQUEZ, J. J. L. A blow-up mechanism for a chemotaxis model. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 24, 4 (1997), 633–683 (1998).
- [8] HILLEN, T., AND POTAPOV, A. The one-dimensional chemotaxis model: global existence and asymptotic profile. *Math. Methods Appl. Sci.* 27, 15 (2004), 1783–1801.
- [9] HORSTMANN, D. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.* 105, 3 (2003), 103–165.
- [10] HORSTMANN, D. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II. *Jahresber. Deutsch. Math.-Verein.* 106, 2 (2004), 51–69.
- [11] JABIR, J.-F., TALAY, D., AND TOMAŠEVIĆ, M. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller-Segel model without smoothing. *arXiv:1712.07490v3* (2017).
- [12] KARATZAS, I., AND SHREVE, S. E. *Brownian motion and stochastic calculus*, second ed., vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.
- [13] KRYLOV, N. V., AND RÖCKNER, M. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* 131, 2 (2005), 154–196.
- [14] MIZOGUCHI, N. Global existence for the Cauchy problem of the parabolic-parabolic Keller-Segel system on the plane. *Calc. Var. Partial Differential Equations* 48, 3-4 (2013), 491–505.
- [15] NAGAI, T., AND OGAWA, T. Global existence of solutions to a parabolic-elliptic system of drift-diffusion type in  $\mathbf{R}^2$ . *Funkcial. Ekvac.* 59, 1 (2016), 67–112.
- [16] OSAKI, K., AND YAGI, A. Finite dimensional attractor for one-dimensional Keller-Segel equations. *Funkcial. Ekvac.* 44, 3 (2001), 441–469.
- [17] QIAN, Z., RUSSO, F., AND ZHENG, W. Comparison theorem and estimates for transition probability densities of diffusion processes. *Probab. Theory Related Fields* 127, 3 (2003), 388–406.
- [18] QIAN, Z., AND ZHENG, W. Sharp bounds for transition probability densities of a class of diffusions. *C. R. Math. Acad. Sci. Paris* 335, 11 (2002), 953–957.
- [19] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math*. Springer, Berlin, 1991, pp. 165–251.
- [20] TALAY, D., AND TOMAŠEVIĆ, M. A new stochastic interpretation of Keller-Segel equations: the 1-D case. *arXiv:1712.10254v3*.
- [21] TOMAŠEVIĆ, M. Ph.d. thesis.