Quantitative approximation of the Burgers and Keller-Segel equations by moderately interacting particles

Christian Olivera ∗ Alexandre Richard† Milica Tomašević ‡

May 14, 2024

Abstract

In this work we obtain rates of convergence for two moderately interacting stochastic particle systems with singular kernels associated to the viscous Burgers and Keller-Segel equations. The main novelty of this work is to consider a non-locally integrable kernel. Namely for the viscous Burgers equation in $\mathbb{R}$, we obtain almost sure convergence of the mollified empirical measure to the solution of the PDE in some Bessel space with a rate of convergence of order $N^{-1/8}$, on any time interval. The same holds for the genuine empirical measure in Wasserstein distance.

In the case of the Keller-Segel equation on a $d$-dimensional torus, we obtain almost sure convergence of the mollified empirical measure to the solution of the PDE in some $L^q$ space with a rate of order $N^{-1/2d+1}$. The result holds up to the maximal existence time of the PDE, for any value of the chemo-attractant sensitivity $\chi$.

1 Introduction

In this work we consider the following nonlinear Fokker-Planck equation:

$$\begin{cases} 
\partial_t u(t, x) = \Delta u(t, x) - \nabla \cdot (u(t, x)(K*u)(t, x)), & t > 0, x \in \mathcal{D}, \\
u(0, x) = u_0(x), & x \in \mathcal{D},
\end{cases}$$

(1.1)

where $\mathcal{D}$ is either $\mathbb{R}$ or the $d$-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. We focus on two particular cases with singular interaction kernel $K$. Firstly, we consider the viscous Burgers equation on $\mathbb{R}$ with $K = \frac{1}{2}\delta_0$, where $\delta_0$ denotes the Dirac distribution. Secondly, we treat the parabolic-elliptic Keller-Segel equation on the torus in dimension $d \geq 2$ for which $K(x) = -\chi_d \frac{x}{|x|^2}$, with $\chi_d > 0$. Motivated by numerical simulation, our main objective is the quantitative approximation of (1.1) by means of a stochastic particle system in moderate interaction.

For both equations, we study the following moderately interacting system of stochastic particles (in the sense of Oelschläger [24] and Méléard and Roelly [23]):

$$dX_i^{t,N} = \frac{1}{N} \sum_{k=1}^{N} (K*V^N)(X_i^{k,N} - X_i^{t,N}) \ dt + \sqrt{2} \ dW_i^{t}, \quad t \leq T, \quad 1 \leq i \leq N,$$

(1.2)

where $(W^i)_{1 \leq i \leq N}$ are independent standard Brownian motions and $V^N(x) := N^{\alpha} V(N^\alpha x)$, for some $\alpha \in (0, 1]$. For this particle system, our objective is to quantify the convergence of the empirical measure $\mu^N = (\mu_i^N)_{1 \leq i \leq T}$ and the regularised empirical measure $u^N := V^N * \mu^N$ towards the solution of (1.1). In the Burgers case, mollifying the interaction seems to be the only way to...
define the particle systems. In the Keller-Segel case, while it is possible to study the non-mollified particle system for some values of \( \chi_d \) ([8, 15, 16, 33]), the mollification avoids discussions about the well-posedness of the particles and their collisions to study directly their convergence.

First, we obtain the global well-posedness for mild solutions of the viscous Burgers equation, i.e. for the PDE (1.1) with \( K = \frac{1}{2} \delta_0 \), in Bessel spaces \( H^\beta_p(\mathbb{R}) \) when \( \beta - 1/p > 0 \) (see Theorem 3.2). Then our main result is that for any \( T > 0 \),

\[
\lim_{N \to +\infty} N^{\epsilon - \xi} \sup_{t \in [0, T]} \| u^N_t - u_t \|_{\beta, p} = 0 \text{ a.s.,}
\]

where the convergence is with a rate \( \xi \) which in the best case is equal to \( \frac{1}{8} \) (see Theorem 3.3 and Remark 3.4). In addition, we prove that the same rate holds for the non-mollified empirical measure \( \mu^N \) in Wasserstein distance.

The approach to prove these results is to write the mild equation satisfied by \( u^N \) and then write the difference \( u - u^N \), for the mild solution \( u \) of (1.1). This difference involves several terms, including a nonlinear quadratic term and a stochastic convolution integral, which create two main difficulties:

- First, the nonlinearity prevents us from quantifying the convergence via a direct Grönwall argument. Instead, we propose here to use Bihari’s inequality. However, this approach works only on a small random time interval. Thus a delicate point is to achieve a global-in-time convergence, by controlling how this random time horizon behaves with \( N \). In addition, we need Hölder continuity for \( u^N \) in space, uniformly in \( N \). Obtaining estimates in \( H^\beta_p(\mathbb{R}) \) with \( \beta - 1/p > 0 \) ensures this property by Sobolev embedding.

- The second difficulty is related to the stochastic convolution integral, see (3.11). Due to the mild formulation of the problem, this stochastic integral is function-valued and we need to bound its moments in \( H^\beta_p(\mathbb{R}) \). To do so, we use Proposition A.3, which itself relies on a BDG inequality for infinite-dimensional martingales [36].

Secondly, we adapt the above methodology to the Keller-Segel equation and obtain an almost sure rate of convergence for the particle system. Namely, for the kernel \( K(x) = -\chi_d \frac{x}{|x|^d} \) on the \( d \)-dimensional torus (for any \( d \geq 1 \)), we obtain (see Theorem 4.4) that

\[
\lim_{N \to +\infty} N^{\epsilon - \xi} \sup_{t \in [0, T]} \| u^N_t - u_t \|_{L^q(\mathbb{T}^d)} = 0 \text{ a.s.,}
\]

where \( q > d, T \) is smaller than the maximal existence time for the PDE (see Definition 4.2) and \( \xi \) is in the best case \( \frac{1}{2q-1} \) (see Remark 4.6). This rate is slightly better than in [25], since we are on the torus where the tails of \( K \) play no role. Besides, no cut-off is applied to the drift in the present work.

The difference between the two functional frameworks, \( H^\beta_p(\mathbb{R}) \) for Burgers and \( L^q(\mathbb{T}^d) \) with \( q > d \) for Keller-Segel, can be roughly explained as follows. When \( K \) is the Keller-Segel kernel, it regularises enough to ensure that \( K \ast u^N \) is Hölder continuous in \( L^q \), with a Hölder coefficient that determines the rate of convergence \( \xi \). This is no longer the case when \( K \) is a Dirac. Nevertheless, assuming initial conditions in \( H^\beta_p(\mathbb{R}) \), we are able to prove that \( u^N \) is uniformly bounded in \( H^\beta_p(\mathbb{R}) \) and is therefore \( (\beta - 1/p) \)-Hölder continuous. In both cases, this uniform Hölder continuity is used in our Grönwall/Bihari argument to obtain a rate of convergence.

These two functional frameworks are well suited for numerical approximation. In particular, as \( H^\beta_p(\mathbb{R}) \) embeds into \( L^\infty \) for \( \beta - 1/p > 0 \), this could lead to pointwise numerical errors for approximation of the PDE by the empirical measure of stochastic particle systems. We leave this question for further research.

It is also worth noticing that, if one were to repeat our proofs for a Lipschitz continuous and bounded kernel, the rate of convergence would be \( N^{\frac{-1}{p} - \xi} \) for \( d = 1 \) and \( N^{\frac{-1}{p} - \xi} \) for \( d \geq 2 \) in \( L^1 \)-norm. We see that compared to the regular case, the two kernels we consider here slow down the speed, due to their singularity and the additional functional framework one needs to consider in order to tame the singularity.
A final remark is that with the results we obtained, the trajectorial propagation of chaos towards a non-linear stochastic process in the sense of McKean-Vlasov is not out of reach. Indeed, it would suffice to adapt to this framework the results of [25, Sec. 3].

**Literature.** Moderately interacting particle systems with regular coefficients and their trajectorial propagation of chaos of were studied in [23]. Furthermore, Jourdain and Méléard [20] studied the fluctuations associated with the convergence of the empirical measure of the system to the law of the solution of the limit nonlinear process when one imposes a smooth initial condition, obtaining a logarithmic rate of convergence. Based on a mild formulation of the empirical measure of a moderately interacting system and semigroup theory, recently Flandoli, Leimbach and Olivera [12] developed a technique to approximate nonlinear PDEs by smoothed empirical measures in strong functional topologies. This technique was also applied for a PDE-ODE system related to aggregation phenomena, see Flandoli and Leocata [11]; for non-local conservation laws, see Olivera and Simon [31]; for the 2d Navier-Stokes equation, see Flandoli, Olivera and Simon [13], etc. In [25] we further obtain convergence rates and propagation of chaos for singular repulsive and attractive Riesz potentials (including the Coulomb case).

In these previous works, the interaction kernel $K$ always had to be at least locally integrable, hence excluding the Burgers interactions ($K = \frac{1}{2} \delta_0$). However, particle approximation of the Burgers equation has a rich history and propagation of chaos were considered first by McKean [22], then Calderon and Pulvirenti [7], Gutkin and Kac [19], Oelschläger [24], Osada and Kotani [26], Sznitman [32] and Bossy and Talay [4, 5]. However, the problem of quantifying the distance between the particles and the equation was only considered in [4, 5]. One way to handle $K = \frac{1}{2} \delta_0$ is to take as initial condition a cumulative distribution function (c.d.f.). Indeed, Bossy and Talay [4] transformed the equation into an integrated version of (1.1), where now the kernel becomes the integral of a Dirac, i.e. a Heaviside function. In such framework, they proved quantitative propagation of chaos [4] and a rate of convergence of the Euler scheme [5] to the viscous Burgers PDE. In this paper, we combine the viewpoint of moderately interacting particles with semigroup techniques, to obtain a quantitative convergence of the empirical measure in Bessel norm for $L^4(\mathbb{R})$ initial conditions (i.e. without assuming the initial condition is a c.d.f.), which seems to be new in the literature.

In the second part of this work, we consider the Keller-Segel kernels ($K(x) = -\chi_2 \frac{x}{|x|^2}$), which are still singular, but locally integrable kernels. The Keller-Segel system of PDEs has been extensively studied for its property that a blow-up in finite time may occur in the equation. That is, the measure solution develops, in finite time, a singular part. For instance in $\mathbb{R}^2$, it is well known that if $\chi_2 < 8\pi$, the solution to (1.1) exists globally (in time), while when $\chi_2 > 8\pi$ a blow-up in finite time occurs (see the survey of Perthame [27], or more recently Biler [2] and the references therein). From the probabilistic side, the particle system (without smoothing) related to Keller-Segel equation in $\mathbb{R}^2$ was studied in [8, 15] and more recently in [16] and [33]. The authors prove well-posedness of the particle system and tightness-consistency for its empirical measure for all subcritical values of the parameter $\chi_2$ [15, 33] and analyze particle collisions in supercritical case [16]. Furthermore, Bresch, Jabin and Wang [6] study quantitative convergence, when $N \to \infty$, of the Liouville’s equations associated to $k$ fixed particles at a time $t$ towards $u_N$, where $u$ solves (1.1). Under the condition that $\chi_2 < 8\pi$ and assuming that $u \in L^\infty((0,T);W^{2,\infty}(\mathbb{T}^2))$, they proved using new techniques of relative entropy the above convergence with a rate in $L^\infty((0,T);L^4((\mathbb{T}^2)^k))$.

In [25], we obtained a rate of convergence for the empirical measure of the mollified particle system, for various kernels including the Keller-Segel one, but applying a cutoff on the drift term of each particle (i.e. we had $F(K \ast u^N)$ in the drift for some smooth cutoff $F$, instead of $K \ast u^N$ as in (1.2)). It was also possible to deduce a rate of convergence, but only in probability, for the particle system without cutoff [25, Corollary 1.4].

**Plan of the paper.** In Section 2, we list the notations used throughout the paper and in Subsection 2.2 we recall or prove some useful lemmas that hold in Bessel spaces both in the torus and in the Euclidean space. Section 3 and Section 4 follow the same organisation: in Subsection 3.1 (resp. Subsection 4.1) we define rigorously the particle system, prove technical lemmas on the
kernel and state the global well-posedness of the PDE for Burgers (resp. local well-posedness for Keller-Segel). Then in Subsection 3.2 (resp. Subsection 4.2), we state and prove the convergence theorems. Finally in the Appendix, we state and prove a bound on the stochastic convolution integral that appears in the expression of the mollified empirical measure of the Burgers particle system.

2 Notations and preliminaries

2.1 Notations and definitions

• We use \( \mathcal{D} \) to denote either \( \mathbb{T}^d \) or \( \mathbb{R}^d \).

• Whenever we consider a function on \( \mathbb{T}^d \), we associate it with its periodic extension which is a function on \( \mathbb{R}^d \).

• For \( (X, d_X) \) a Polish space, we consider the space \( C([0, T]; X) \) of continuous functions from \([0, T]\) to \( X \) endowed with the distance

\[
\tilde{d}(f, g) = 1 \wedge \sup_{t \in [0, T]} d_X(f(t), g(t)). \tag{2.1}
\]

Recall that this distance is topologically equivalent to the distance \( d(f, g) = \sup_{t \in [0, T]} d_X(f(t), g(t)) \).

Let us denote by \( N_\delta \) the Hölder seminorm of parameter \( \delta \in (0, 1] \), that is, for any function \( f \) defined over \( \mathcal{D} \):

\[
N_\delta(f) := \sup_{x \neq y \in \mathcal{D}} \frac{|f(x) - f(y)|}{|x - y|^\delta}. \tag{2.2}
\]

The space of continuous functions on \( \mathcal{D} \) which have finite \( N_\delta \) seminorm is the Hölder space \( C^\delta(\mathcal{D}) \).

• For \( (X, d_X) \) a Polish space, denote by \( \mathcal{M}_1(X) \) the set of Borel probability measures on \( X \) and by \( W_1 \) the Wasserstein distance.

Following [3, Section 8.3], let us introduce the Kantorovich-Rubinstein metric which reads, for any two probability measures \( \mu \) and \( \nu \) on \( \mathcal{D} \),

\[
\|\mu - \nu\|_0 = \sup \left\{ \int_{\mathcal{D}} \phi d(\mu - \nu) ; \phi \text{ Lipschitz with } \|\phi\|_{L^\infty(\mathcal{D})} \leq 1 \text{ and } \|\phi\|_{\text{Lip}} \leq 1 \right\}. \tag{2.3}
\]

• In this paper, \( (e^{t \Delta})_{t \geq 0} \) denotes the semigroup of the heat operator on \( \mathcal{D} \). That is, for \( f \in L^p(\mathcal{D}) \),

\[
e^{t \Delta} f(x) = (g^{(\mathcal{D})}_t \ast f)(x),
\]

where \( \ast_{\mathcal{D}} \) denotes the convolution on \( \mathcal{D} \) and for any \( t > 0 \), \( g^{(\mathcal{D})}_t \) is heat kernel given by

\[
g^{(\mathcal{D})}_t(x) = \begin{cases} 
\frac{1}{(2\pi t)^{d/2}} \sum_{k \in \mathbb{Z}^d} e^{-i k \cdot x / 2} & \text{on } \mathbb{T}^d, \\
\frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x|^2}{2t}} & \text{on } \mathbb{R}^d.
\end{cases}
\]

To make the notations easier, we will write \( g_t \) for \( g^{(\mathbb{T}^d)}_t \) and \( g^0_t \) for \( g^{(\mathbb{R}^d)}_t \).

• If \( u \) is a function or stochastic process defined on \([0, T] \times \mathcal{D} \), we will most of the time use the notation \( u_t \) to denote the mapping \( x \mapsto u(t, x) \).

• Depending on the context, the brackets \( \langle \cdot, \cdot \rangle \) will denote either the scalar product in some \( L^2 \) space or the duality bracket between a measure and a function.
Bessel spaces. We know to introduce briefly Bessel potential spaces on \( \mathbb{R}^d \) and on the torus \( \mathbb{T}^d \).

- For any \( \beta \in \mathbb{R} \) and \( p \geq 1 \), we denote by \( H^\beta_p(\mathbb{R}^d) \) the space
  \[
  H^\beta_p(\mathbb{R}^d) := \left\{ u \text{ tempered distribution}; \mathcal{F}^{-1}(1 + |\cdot|^2)^{\beta/2} \mathcal{F}u(\cdot) \in L^p(\mathbb{R}^d) \right\},
  \]
  where \( \mathcal{F}u \) denotes the Fourier transform of \( u \). This space is endowed with the norm
  \[
  \|u\|_{\beta,p} := \left\| \mathcal{F}^{-1}(1 + |\cdot|^2)^{\beta/2} \mathcal{F}u(\cdot) \right\|_{L^p(\mathbb{R}^d)}.
  \] (2.4)
  
  The space \( H^\beta_p(\mathbb{R}^d) \) is associated to the Bessel potential operator \( (I - \Delta)^{\beta/2} \) defined as (see e.g. [34, p.180] for more details on this operator):
  \[
  (I - \Delta)^{\beta/2} f := \mathcal{F}^{-1} \left( (1 + |\cdot|^2)^{\beta/2} \mathcal{F}f \right).
  \]

- Let \( D(\mathbb{T}^d) \) be the collection of all infinitely differentiable functions on \( \mathbb{T}^d \). Then \( D'(\mathbb{T}^d) \) stands for the topological dual of \( D(\mathbb{T}^d) \). We denote the Fourier coefficients of \( u \in D'(\mathbb{T}^d) \) by \( \hat{u}(k) := \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} \hat{u}(k) e^{-2i\pi \langle k, \cdot \rangle} \). As there should be no risk of confusion, we will use the same notations for the Bessel norm and Bessel operator on the torus as in the whole space. Hence for any \( \beta \in \mathbb{R} \), we define the Bessel potential operator \( (I - \Delta)^{\beta/2} \) applied to \( u \in D'(\mathbb{T}^d) \) by
  \[
  (I - \Delta)^{\beta/2} u(x) = \frac{1}{(2\pi)^{d/2}} \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} \hat{u}(k) e^{-2i\pi \langle k, x \rangle},
  \]
  and denote
  \[
  \|u\|_{\beta,p} = \left\| (I - \Delta)^{\beta/2} u \right\|_{L^p(\mathbb{T}^d)}. \] (2.5)

As in [29, p.168], \( H^\beta_p(\mathbb{T}^d) \) is defined for \( p \in (1, \infty) \) and \( \beta \in \mathbb{R} \) by
  \[
  H^\beta_p(\mathbb{T}^d) := \left\{ u \in D'(\mathbb{T}^d); \|u\|_{\beta,p} < \infty \right\}.
  \]

2.2 Preliminary lemmas

We conclude this section by heat kernel estimates. First, we have that for any \( t > 0 \),
  \[
  \|\nabla g_t\|_{L^1(\mathbb{T}^d)} \leq \|\nabla g_t^0\|_{L^1(\mathbb{R}^d)} \leq \frac{C}{\sqrt{t}}.
  \]
  By a convolution inequality, we deduce that
  \[
  \|\nabla \cdot e^{t\Delta}\|_{L^p(D) \to L^p(D)} \leq \frac{C}{\sqrt{t}}. \] (2.6)

**Lemma 2.1.** Let \( \beta \in \mathbb{R} \) and \( p \in (1, \infty) \). For any \( t > 0 \),
  \[
  \|g_t\|_{\beta,1} = \|(I - \Delta)^{\beta/2} g_t\|_{L^1(D)} \leq C \left( 1 \vee t^{-\frac{\beta}{2}} \right),
  \]
  and
  \[
  \|(I - \Delta)^{\beta/2} e^{t\Delta}\|_{L^p(D) \to L^p(D)} \leq C \left( 1 \vee t^{-\frac{\beta}{2}} \right).
  \]

**Proof.** These inequalities are standard when \( D = \mathbb{R}^d \), see e.g. Eq. (2.2) in [25]. We prove them now for \( D = \mathbb{T}^d \).
First, we notice that for \( k \in \mathbb{Z}^d \), \( \hat{g}_t(k) = \mathcal{F} \hat{g}_t^0(2 \pi k) \), where \( \mathcal{F} \) denotes the Fourier transform of \( \mathbb{R}^d \). Hence \( \hat{g}_t(k) = \hat{g}_t^0(2 \pi k \sqrt{t}) \) and
\[
\| (I - \Delta)^{\frac{\beta}{2}} \hat{g}_t^0 \|_{L^1(T^d)} = \frac{1}{(2 \pi)^{d/2}} \int_{T^d} \left| \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} \hat{g}_t^0(2 \pi k \sqrt{t}) e^{-2 \pi i \langle k, x \rangle} \right| dx
\]
\[
\leq \frac{1}{(2 \pi)^{d}} \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} e^{-2 \pi^2 |k|^2 t}.
\]
Since the mapping \( t \mapsto t^{\beta/2} \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} e^{-2 \pi^2 |k|^2 t} = \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} e^{-2 \pi^2 |k|^2 t} \) is bounded on \([0, 1]\) and the mapping \( t \mapsto \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\beta/2} e^{-2 \pi^2 |k|^2 t} \) is bounded on \((1, +\infty)\), this proves the first inequality.

Then for \( f \in L^p(T^d) \),
\[
(I - \Delta)^{\frac{\beta}{2}} e^{t \Delta} f(x) = \sum_{k \in \mathbb{Z}^d} (1 + |k|^2)^{\frac{\beta}{2}} \hat{g}_{2t}(k) \hat{f}(k) e^{-2 \pi i \langle k, x \rangle}
\]
\[
= (2 \pi)^{d} f \ast (I - \Delta)^{\frac{\beta}{2}} \hat{g}_{2t}.
\]
Hence the second inequality follows by a convolution inequality. \( \square \)

**Lemma 2.2.** Let \( \beta \in \mathbb{R} \) and \( p \in (1, \infty) \). There exists \( C > 0 \) such that for any distribution \( f \),
\[
C^{-1} \| f \|_{\beta + 1, p} \leq \| \nabla f \|_{\beta, p} \leq C \| f \|_{\beta + 1, p}.
\]

**Proof.** When \( \mathcal{D} = \mathbb{R}^d \), this is the equivalence of norms described in [35, Eq. (3) p.59] (note that the space \( F_{p,2}^p \) in [35] is the Bessel space used here).

When \( \mathcal{D} = T^d \), use the derivative in Fourier space and the equivalence of Bessel potential space and Triebel-Lizorkin space (defined in [29, Theorem. (ii), p.167]), see [29, Theorem. (v), p.168-169]. \( \square \)

### 3 The viscous Burgers equation

In this section, the domain is always \( \mathcal{D} = \mathbb{R} \).

#### 3.1 Setting and preliminaries

Let us now study the viscous Burgers equation on the whole space, which corresponds to Equation (1.1) with \( K = \frac{1}{2} \delta_0 \):
\[
\begin{aligned}
\partial_t u(t, x) &= \partial_{xx}^2 u(t, x) - \frac{1}{2} \partial_x (u(t, x)^2), \quad x \in \mathbb{R}, \quad t > 0, \\
u(0, x) &= u_0(x), \quad x \in \mathbb{R}.
\end{aligned}
\] (3.1)

Let \( N \geq 1 \). Let us introduce a mollifier that will be used both to regularise the interaction kernel in the particle system and its empirical measure. Let \( V : \mathbb{R} \to \mathbb{R}_+ \) be a compactly supported, smooth probability density function. For \( \alpha \in [0, 1] \) and any \( x \in \mathbb{R} \), define
\[
V^N(x) := N^{\alpha} V(N^{\alpha} x).
\] (3.2)

Below, \( \alpha \) will be restricted to some interval \((0, \alpha_0)\), see Assumption \((A_{\alpha, \beta, p})\). The corresponding particle system in moderate interaction reads
\[
\begin{aligned}
dX^i_{t^N} &= \frac{1}{2N} \sum_{k=1}^{N} V^N(X^i_{t^N} - X^k_{t^N}) \, dt + \sqrt{2} \, dW_i^t, \quad t \leq T, \quad 1 \leq i \leq N, \\
X^i_{0^N}, \quad 1 \leq i \leq N, \quad \text{are independent of} \{W^i, \quad 1 \leq i \leq N\},
\end{aligned}
\] (3.3)
where \( \{(W^i_t)_{t\geq 0}, \ i \in \mathbb{N}\} \) is a family of independent standard \( \mathbb{R} \)-valued Brownian motions defined on a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})\). Let us denote the empirical measure of \( N \) particles by

\[
\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{X^i,N},
\]

and the mollified empirical measure by

\[
u^N := V^N \ast \mu^N.
\]

It is well-known that when \( u_0 \) has good properties, the Cole-Hopf solution [9] given by

\[
u^N := V^N \ast \mu^N.
\]

We have the following well-posedness result for the Burgers equation. For our purpose, we will need some space regularity of the solutions. This will be achieved through the following mild notion of solution:

**Definition 3.1.** Let \( p > 1 \) and \( \beta > 1/p \). Given \( u_0 \in H_\beta^p(\mathbb{R}) \) and \( T > 0 \), a function \( u \) on \([0, T] \times \mathbb{R}\) is said to be a mild solution to (3.1) on \([0, T]\) if

(i) \( u \in C([0, T]; H_\beta^p(\mathbb{R})) \);

(ii) \( u \) satisfies the integral equation

\[
u^N := V^N \ast \mu^N.
\]

\[
u^N := V^N \ast \mu^N.
\]

We have the following well-posedness result for the Burgers equation.

**Theorem 3.2.** Let \( p > 1 \), \( \beta \in (1/p, 1) \), \( T > 0 \) and \( u_0 \) such that \( u_0 \in L^1 \cap H_\beta^p(\mathbb{R}) \). Then there exists a unique mild solution \( u \) to (3.1) in \( C([0, T]; H_\beta^p(\mathbb{R})) \). Besides, \( u \) coincides with the Cole-Hopf solution \( u^{CH} \) and is unique in the larger space \( L^\infty([0, T]; L^\infty(\mathbb{R})) \).

**Proof.** We proceed in three steps. In Step 1, we prove local well-posedness in the sense of Definition 3.1 using a fixed-point argument. In Steps 2 and 3, we prove that this solution coincides with \( u^{CH} \) and that \( \sup_{t \in [0,T]} \| u^{CH}_t \|_{\beta,p} < \infty \) for any \( T > 0 \). Hence a mild solution exists for any time horizon \( T \).

**Step 1.** Let \( T > 0 \) to be fixed later. We set \( B_t[u,v](x) = \frac{1}{2} \int_0^t \partial_x e^{(t-s)\Delta} (u_x v_x)(x) \, ds \), for \( u, v \in C([0, T]; H_\beta^p(\mathbb{R})), t \in [0, T] \) and \( x \in \mathbb{R} \). As the operator \((I - \Delta)^{\beta/2}\) commutes with \( \partial_x e^{t\Delta} \), we have

\[
\|B_t[u,v]\|_{\beta,p} \leq \frac{1}{2} \int_0^t \left\| \partial_x e^{(t-s)\Delta} (u_x v_x) \right\|_{\beta,p} \, ds \leq \frac{1}{2} \int_0^t \left\| \partial_x e^{(t-s)\Delta} (I - \Delta)^{\beta/2} (u_x v_x) \right\|_{L^p(\mathbb{R})} \, ds.
\]

Hence in view of (2.6), we have

\[
\sup_{t \in [0,T]} \| B_t[u,v] \|_{\beta,p} \leq C \sup_{t \in [0,T]} \int_0^t \| u_x v_x \|_{H_\beta^p(\mathbb{R})} \, ds.
\]

Now in view of [1, Corollary 2.86] and since \( \beta - 1/p > 0, H_\beta^p(\mathbb{R}) \) is an algebra. Hence we get

\[
\sup_{t \in [0,T]} \| B_t[u,v] \|_{\beta,p} \leq CT^{1/2} \sup_{t \in [0,T]} \| u_t \|_{\beta,p} \sup_{t \in [0,T]} \| v_t \|_{\beta,p}.
\]

Thus by a standard contraction principle, see for instance Theorem 13.2 in [21], we deduce that there exists a time \( T^+ > 0 \) and a unique mild solution \( u \) to (3.1) in \( C([0, T^+]; H_\beta^p(\mathbb{R})) \) for any \( T < T^+ \).
Step 2. We observe first that uniqueness of mild solutions holds in $L^\infty([0,T],[L^\infty(\mathbb{R})])$ by a Grönewall argument, for any $T > 0$. We know that $u^{CH}$ is a strong solution to the viscous Burgers equation, hence it is also a mild solution. In addition, it is in $L^\infty([0,T],L^\infty(\mathbb{R}))$, as we have in particular that $\|u_t^{CH}\|_{L^\infty(\mathbb{R})} \leq \|u^{CH}_0\|_{L^\infty(\mathbb{R})}$. Since there is $C(\{0,T\};H_p^\beta(\mathbb{R})) \subset L^\infty([0,T],L^\infty(\mathbb{R}))$ by a Sobolev embedding, $u$ and $u^{CH}$ both live in $L^\infty([0,T],L^\infty(\mathbb{R}))$ and solve the mild equation, so they are equal. Hence it only remains to prove that a solution exists in $C([0,T];H_p^\beta(\mathbb{R}))$, for any $T > 0$. We now check that $u^{CH}$ provides such a solution.

If $T^* = +\infty$, then the problem is already solved. So assume that we only obtained $T^* < +\infty$ in Step 1, and let $T > T^*/2$. Then

$$\sup_{t \in [T^*/2,T]} \|u_t^{CH}\|_{\beta,p} = \max \left( \sup_{t \in [0,T^*/2]} \|u_t^{CH}\|_{\beta,p}, \sup_{t \in [T^*/2,T]} \|u_t^{CH}\|_{\beta,p} \right)$$

$$= \max \left( \sup_{t \in [0,T^*/2]} \|u_t\|_{\beta,p}, \sup_{t \in [T^*/2,T]} \|u_t^{CH}\|_{\beta,p} \right),$$

since $u$ and $u^{CH}$ coincide on $[0,T^*/2]$ by the uniqueness proven in Step 1. We will show in Step 3 that for any $T > T^*/2$,

$$\sup_{t \in [T^*/2,T]} \|u_t^{CH}\|_{\beta,p} \leq \frac{C}{(T^*)^{1-1/(2p)}}, \quad (3.8)$$

with the constant $C$ that depends only on $\|u_0\|_{L^1(\mathbb{R})}$. Thus $u^{CH} \in C([0,T];H_p^\beta(\mathbb{R}))$, for any $T > 0$.

Step 3. Using the Sobolev embedding $W^{1,p}(\mathbb{R}) \hookrightarrow H_p^\beta(\mathbb{R})$ (which holds as $\beta < 1$), we have that for $T' > 0$,

$$\sup_{t \in [T^*/2,T']} \|u_t^{CH}\|_{\beta,p} \leq C \sup_{t \in [T^*/2,T']} \|u_t^{CH}\|_{W^{1,p}}.$$

Denote $U_0(x) = e^{-\frac{1}{2} \int_0^x u_0(z) \, dz}$. As $u_0 \in L^1(\mathbb{R})$ we have that

$$e^{-\frac{1}{2} \|u_0\|_{L^1(\mathbb{R})}} \leq U_0(x) \leq e^{\frac{1}{2} \|u_0\|_{L^1(\mathbb{R})}}. \quad (3.9)$$

By direct computation and (3.9), there is for $t \in [T^*/2,T']$

$$\|\partial_x u_t^{CH}\|_{L^p(\mathbb{R})} = \left\| \frac{(\partial_x g_t^0 * (u_0 U_0)) (g_t^0 * U_0)}{(g_t^0 * U_0)^2} \right\|_{L^p(\mathbb{R})}$$

$$\leq e^{\|u_0\|_{L^1(\mathbb{R})}} \|\partial_x g_t^0 * (u_0 U_0)\|_{L^p(\mathbb{R})} \leq C \|U_0\|_{L^1(\mathbb{R})} e^{\|u_0\|_{L^1(\mathbb{R})}} \partial_x g_t^0 \|_{L^p(\mathbb{R})},$$

from which (3.8) holds.

The restriction with respect to the parameter $\alpha$ and the hypothesis on the initial conditions of the system are given by the following assumption:
(A$^{\alpha,\beta,p}$): Assume that $p \geq 2$ and $\beta \in (\frac{1}{p}, 1)$.

(A$^{\alpha,\beta,p}_u$) The parameters $\alpha$, $\beta$ and $p$, satisfy

$$0 < \alpha < \frac{1}{2(1 + \beta - \frac{1}{p})}.$$ (3.1)

(A$^{\alpha,\beta,p}_i$) Assume that $u_0 \in L^1 \cap H^\beta_p(\mathbb{R})$ with $\|u_0\|_{L^1(\mathbb{R})} = 1$. For any $m \geq 1$, assume

$$\mathbb{E} \left[ \|u_0^N - u_0\|_{\beta,\mu}^m \right] \leq \frac{C}{N\xi},$$

where

$$\xi := \min \left( \alpha(\beta - \frac{1}{p}), \frac{1}{2} - \alpha(1 + \beta - \frac{1}{p}) \right).$$

3.2 Convergence

Recall that $u$ is the global solution of (3.7) given by Theorem 3.2. In the following, $\mathcal{L}(\mu^N)$ denotes the law of $u^N$ on the space $C([0,T], H^\gamma_p(\mathbb{R}))$, for some $\gamma > 1 - 1/p$. Note that $\mu^N$, as a sum of Dirac measures, is in $H^{-\gamma}_p$ (recall that in dimension 1, any Dirac measure is in $H^1_{1/p - 1/\epsilon}$).

Moreover, $W^{(\gamma,p)}_1$ denotes the 1-Wasserstein distance on the space of probability measures on $C([0,T], H^\gamma_p(\mathbb{R}))$ endowed with the distance $d(f,g) = 1 \wedge \sup_{t \in [0,T]} \|f_t - g_t\|_{\gamma,p}$.

**Theorem 3.3.** Assume (A$^{\alpha,\beta,p}$) holds true. Then for any $\epsilon > 0$,

(i) $\lim_{N \to +\infty} N^{\xi/\epsilon} \sup_{t \in [0,T]} \|u^N_t - u_t\|_{\beta,\mu} = 0$ a.s.

(ii) Let $\gamma > 1 - 1/p$. Then there exists $C > 0$ such that for all $N \in \mathbb{N}^*$,

$$W^{(\gamma,p)}_1(\mathcal{L}(\mu^N), \delta_u) \leq C(N^{-1/\epsilon}) + N^{-\alpha(1 + \beta - \frac{1}{p})}.$$ (3.1)

**Remark 3.4.** In view of the constraint (A$^{\alpha,\beta,p}_u$), this theorem gives the almost sure convergence of $u^N$ and $\mu^N$ for any $\alpha < \frac{1}{2}$ by choosing $p = 2$ and $\beta$ close to $\frac{1}{2}$, provided the initial condition has enough regularity. On the other hand, the best possible rate of convergence one can get here is $\xi = (\frac{1}{3})^+$, by choosing $p = 2$, $\beta = 1/2$ and $\alpha = (\frac{1}{3})^+$.

**Proof.** Similarly to [25, Eq. (2.3)], we obtain the following mild formulation for $x \in \mathbb{R}$:

$$u^N_t(x) = e^{t\Delta} u^N_0(x) - \frac{1}{2} \int_0^t \partial_x e^{(t-s)\Delta} \langle \mu^N_s, V^N(x - \cdot) u^N_s(\cdot) \rangle \, ds$$

$$+ \frac{1}{\sqrt{2N}} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \partial_x V^N(x - X^i_N) \, dW^i_s.$$ (3.2)

Hence

$$u^N_t(x) - u_t(x) = e^{t\Delta} (u^N_0 - u_0)(x) - \frac{1}{2} \int_0^t \partial_x e^{(t-s)\Delta} \langle (\mu^N_s, V^N(x - \cdot) u^N_s(\cdot)) - (u_s(x))^2 \rangle \, ds$$

$$- \frac{1}{\sqrt{2N}} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \partial_x V^N(x - X^i_N) \, dW^i_s$$

$$= e^{t\Delta} (u^N_0 - u_0)(x) + \frac{1}{2} \int_0^t \partial_x e^{(t-s)\Delta} ((u_s)^2 - (u^N_s)^2)(x) \, ds$$

$$+ E_t(x) + M^N_t(x),$$

where

$$E_t(x) = \int_0^t \int_{\mathbb{R}} (\mu^N_s, V^N(x - \cdot) u^N_s(\cdot)) \, dW^i_s.$$
where we have set
\[ E_t(x) := \frac{1}{2} \int_0^t \partial_s e^{(t-s)\Delta} \langle \mu^N_s, V^N(x - \cdot) \rangle \left( u^N_s(x) - u^N_s(\cdot) \right) \, ds, \]
\[ M^N_t(x) := -\frac{1}{\sqrt{2N}} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \partial_s V^N(x - X^N_{i,s}) \, dW^i_s. \] (3.11)

In view of the estimate (2.6),
\[ \|u^N_t - u_t\|_{\beta,p} \leq \|e^{\Delta(t-u^N_0 - u_0)}\|_{\beta,p} + C \int_0^t \frac{1}{\sqrt{t-s}} \|(u_s) - (u^N_s)\|^2_{\beta,p} \, ds \\
+ \|E_t\|_{\beta,p} + \|M^N_t\|_{\beta,p}. \]

Observe that for any \( a, b \in \mathbb{R} \), one can write \( a^2 - b^2 = 2a(a - b) - (a - b)^2 \). The latter combined with the fact that \( H^2_p(\mathbb{R}) \) is an algebra for \( \beta > 1/p \) ([1, Corollary 2.86]) leads to
\[ \|u^N_t - u_t\|_{\beta,p} \leq \|u^N_0 - u_0\|_{\beta,p} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s\|_{\beta,p} \|u_s - u^N_s\|_{\beta,p} \, ds \\
+ C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s - u^N_s\|^2_{\beta,p} \, ds \\
+ \|E_t\|_{\beta,p} + \|M^N_t\|_{\beta,p}. \]

- Let us focus on \( E \). Using Lemma 2.2 and the positivity of \( V^N \), we have
\[ \|E_t\|_{\beta,p} \leq C \int_0^t \| (I - \Delta)^{\beta/2} e^{(t-s)\Delta} \|_{L^p(\mathbb{R}) \rightarrow L^p(\mathbb{R})} \left( \int_{\mathbb{R}} \langle \mu^N_s, V^N(x - \cdot) \rangle |u^N_s(\cdot) - u^N_s(x)|^p \, dx \right)^{1/p} \, ds. \]

In view of Lemma 2.1, we get
\[ \|E_t\|_{\beta,p} \leq C \int_0^t \frac{1}{(t-s)^{1/2}} \left( \int_{\mathbb{R}} \langle \mu^N_s, V^N(x - \cdot) \rangle |u^N_s(\cdot) - u^N_s(x)|^p \, dx \right)^{1/p} \, ds. \]

Now, by the embedding \( H^\beta_p(\mathbb{R}) \hookrightarrow C^{\beta-1/p}(\mathbb{R}) \), we have the inequality \( |u^N_s(\cdot) - u^N_s(x)| \leq \|u^N_s\|_{\beta,p} \cdot |x|^{-\beta-1/p} \). Hence
\[ \|E_t\|_{\beta,p} \leq C \int_0^t \frac{\|u^N_s\|_{\beta,p}}{(t-s)^{1/2}} \left( \int_{\mathbb{R}} \langle \mu^N_s, V^N(x - \cdot) \rangle |x|^{-\beta-1/p} \, dx \right)^{1/p} \, ds. \]

Since \( V \) is compactly supported, we have that \( V^N(x - y) |y - x|^{-\beta-1/p} \leq CN^{1-\beta/p}V^N(x - y) \). This leads to
\[ \|E_t\|_{\beta,p} \leq \frac{C}{N^{\beta-1/p}} \int_0^t \frac{1}{(t-s)^{1/2}} \|u^N_s\|^2_{\beta,p} \, ds \\
\leq \frac{C}{N^{\beta-1/p}} \int_0^t \frac{1}{(t-s)^{1/2}} \left( \|u_s\|_{\beta,p} + \|u^N_s - u_s\|_{\beta,p} \right)^2 \, ds \]
and using again the boundedness of \( u \) we get
\[ \|E_t\|_{\beta,p} \leq \frac{C}{N^{\beta-1/p}} \left( 1 + \int_0^t \frac{1}{(t-s)^{1/2}} \|u^N_s - u_s\|^2_{\beta,p} \, ds \right). \]

- Since \( p \geq 2 \), by Sobolev embedding and Proposition A.3 we have for \( \varepsilon \) small enough that
\[ \forall N \in \mathbb{N}^*, \quad \left\| \sup_{s \in [0,t]} \| M^N_s \|_{\beta,\frac{\beta}{2}} \right\|_{L^p(\Omega)} \leq \left\| \sup_{s \in [0,t]} \| M^N_s \|_{\beta+\frac{\beta}{2}-\frac{\beta}{2}} \right\|_{L^p(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(1+2\beta+1/2)+\varepsilon)}. \] (3.12)
By Borel-Cantelli’s lemma, we deduce that there exists a random variable \( A_0 \) with finite moments such that almost surely,
\[
\sup_{s \in [0,t]} \| M_s^N \|_{\beta,p} \leq \frac{A_0}{N^{1-\varepsilon}}.
\]

Then gathering the previous bounds, using \((A^H_{\alpha,\beta,p})\) and the fact that \( \| u_s \|_{\beta,p} \) is bounded (Theorem 3.2), we have
\[
\| u_t^N - u_t \|_{\beta,p} \leq \frac{A}{N^{1-\varepsilon}} + C \int_0^t \frac{1}{\sqrt{t-s}} \| u_s - u_s^N \|_{\beta,p} ds + C \int_0^t \frac{1}{(t-s)^{(1+\beta)/2}} \| u_s - u_s^N \|_{\beta,p}^2 ds.
\]

Let \( r > \frac{2}{1-\varepsilon} \). Denote \( \mathcal{U}_t^N = \| u_t^N - u_t \|_{\beta,p}^r \) and
\[
A_N = \frac{A^r}{N^{r(1-\varepsilon)}}, \quad (3.13)
\]

By Hölder’s inequality, we have
\[
\mathcal{U}_t^N \leq A_N + CT^{1-r} \int_0^t \mathcal{U}_s^N ds + CT^{1-2r-1} \int_0^t (\mathcal{U}_s^N)^2 ds.
\]

Now apply Bihari’s inequality (see e.g. [10, Theorem 27]) and get that
\[
\mathcal{U}_t^N \leq G^{-1}(G(A_N) + CTt), \quad \text{for any } t \in [0, \tau_N(\omega)) \cap [0, T], \quad (3.14)
\]

where \( C_T = C \left( T^{1-2r-1} + T^{1-\varepsilon} \right) \) and
\[
G(x) = \int_1^x \frac{1}{y+y^2} dy = \log \left( \frac{2x}{1+x} \right),
\]
\[
G^{-1}(x) = \frac{e^x}{2 - e^x},
\]
\[
\tau_N(\omega) = \frac{1}{C_T} \log \left( \frac{1 + A_N(\omega)}{A_N(\omega)} \right).
\]

Hence, (3.14) reads
\[
\| u_t^N - u_t \|_{\beta,p} \leq e^{C_Tt} \frac{2A_N/(1 + A_N)}{2 - e^{C_Tt}2A_N/(1 + A_N)}, \quad \text{for any } t \in [0, \tau_N) \cap [0, T]. \quad (3.15)
\]

Let \( \tilde{\Omega} \) be a measurable subset of \( \Omega \) of measure 1 on which \( A \) is finite and (3.14) holds. For \( \omega \in \tilde{\Omega} \), there exists \( N_0(\omega) \) such that for any \( N \geq N_0(\omega) \), \( \tau_{N_0} > T \). Thus
\[
\sup_{N \geq N_0(\omega)} \sup_{t \in [0,T]} N^{r(\xi-2e)} \| u_t^N - u_t \|_{\beta,p} \leq \sup_{N \geq N_0(\omega)} N^{r(\xi-2e)} e^{C_Tt} \frac{2A_N/(1 + A_N)}{2 - e^{C_Tt}2A_N/(1 + A_N)} < \infty.
\]

Hence
\[
\limsup_{N \to +\infty} N^{r(\xi-2e)} \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \leq \limsup_{N \to +\infty} N^{r(\xi-2e)} e^{C_Tt} \frac{2A_N/(1 + A_N)}{2 - e^{C_Tt}2A_N/(1 + A_N)} = 0,
\]

which gives point \((i)\) of the theorem.

**Proof of \((ii)\).** Recall that \( W_1^{(\gamma,p)} \) denotes here the Wasserstein distance on the space \( C([0,T], H^{-\gamma}(\mathbb{R})) \) which is endowed with the distance \( d(f,g) = 1 \wedge \sup_{t \in [0,T]} \| f_t - g_t \|_{-\gamma,p} \), for some \( \gamma > 1 - \frac{1}{p} \). It
comes by choosing the trivial coupling \( \mathcal{L}(\mu^N) \otimes \delta_u \) that
\[
\mathcal{W}_1^{(\gamma,p)}(\mathcal{L}(\mu^N), \delta_u) \leq \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| \mu_t^N - u_t \|_{-\gamma,p} \right]
\leq \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{-\gamma,p} \right] + \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| \mu_t^N - u_t^N \|_{-\gamma,p} \right]
\leq \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \right] + \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \sup_{\| \phi \|_{\gamma,p'} \leq 1} \langle \mu_t^N - u_t^N, \phi \rangle \right].
\]
(3.16)

We treat the first term on the right-hand side and will prove that
\[
\mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \right] \leq CN^{-(\xi-\epsilon)}.
\]
(3.17)

Consider \( N_0 \equiv N_0(\omega) \) the smallest integer such that \( \tau_{N_0} > T + \frac{1}{e^{CT}} \log(2) \). Then we get
\[
\mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \right] \leq \mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \right] + \mathbb{P}(N_0 \geq N).
\]

The above choice of \( N_0 \) induces that for \( N \geq N_0 \), we have that \( \frac{2A_N}{1+A_N} e^{C_T T} = 2e^{C_T(T-\tau_N)} \leq 1 \).

Hence, in view of (3.15), which holds true for any \( t \leq T \) on the event \( \{ N \geq N_0 \} \), we obtain
\[
\mathbb{E} \left[ 1 \wedge \sup_{t \in [0,T]} \| u_t^N - u_t \|_{\beta,p} \right] \leq e^{\frac{C_T}{2} \mathbb{E} \mathbb{P}(N_0 \geq N)} \left[ \frac{2A_N/(1+A_N)}{2 - e^{C_T T} 2A_N/(1+A_N)} \right]^\frac{1}{p}
\]  
\[ \leq e^{\frac{C_T}{2} \mathbb{E} \mathbb{P}(N_0 \geq N)} \left[ \frac{2A_N}{1+A_N} \right]^\frac{1}{p}
\]  
\[ \leq C e^{\frac{C_T}{2} \mathbb{E} \mathbb{P}(N_0 \geq N)} N^{-(\xi-\epsilon)},
\]
using the definition (3.13) of \( A_N \) and the fact that \( A \) has finite moments.

Now we estimate \( \mathbb{P}(N_0 \geq N) \). By the definition of \( N_0 \), we have that \( \tau_{N_0-1} \leq T + \frac{1}{e^{CT}} \log(2) \).

Hence in view of the definition of \( \tau_{N_0-1} \) and \( A_N \), we deduce that
\[
N_0 \leq 1 + (2e^{C_T T} - 1) \tau_{N_0-1} \leq 1 + \frac{1}{e^{CT}} A \frac{1}{p}.
\]

Now we get
\[
\mathbb{P}(N_0 \geq N) \leq \mathbb{P} \left( A \geq \left( \frac{(N-1)\xi-\epsilon}{2e^{C_T T} - 1} \right)^\frac{1}{p} \right)
\]  
\[ \leq C \mathbb{E} A^p \left( \frac{(N-1)\xi-\epsilon}{2e^{C_T T} - 1} \right)^\frac{1}{p}.
\]
by the Markov inequality, for any \( p \geq 1 \). Hence (3.17) follows.

For the second term in the right-hand side of (3.16), we observe that \( H^\gamma \) embeds into the H"older space \( C^{\gamma - \frac{1}{p'}} \), where we recall that \( \gamma - \frac{1}{p'} > 1 - \frac{1}{p} - \frac{1}{p'} = 0 \). Hence
\[
|\langle \mu_t^N, \phi \rangle - \langle u_t^N, \phi \rangle| = |\langle \mu_t^N, (\phi - \phi * V^N) \rangle|
\]  
\[ \leq \left| \langle \mu_t^N, \int \phi \cdot (\gamma - \phi * V^N) dy \rangle \right|
\]  
\[ \leq \frac{C \| \phi \|_{\gamma - \frac{1}{p'}}}{N^{\alpha(\gamma - \frac{1}{p'})}}.
\]
Hence \( \mathcal{W}_1^{(\gamma,p)}(\mathcal{L}(\mu^N), \delta_u) \leq N^{-(\xi-\epsilon)} + N^{-\alpha(\gamma - \frac{1}{p'})} \).
Remark 3.5. We mention that it is also possible to get a rate for the Wasserstein distance directly on \( \mu^N \) (i.e. not on its law), but only for the marginals. Namely, denote by \( W_1 \) the 1-Wasserstein distance on probability measures on \( \mathbb{R} \), where \( \mathbb{R} \) is endowed with the equivalent (to the usual topology) distance \( |\cdot| \wedge 1 \). Then, we get
\[
\forall N \in \mathbb{N}^*, \quad \sup_{t \in [0,T]} \mathbb{E}[W_1(\mu^N_t, u_t)] \leq C N^{-(\xi - \varepsilon)}.
\]
(3.18)

To show that (3.18) holds, consider the McKean-Vlasov equations
\[
dX_t^i = \frac{1}{2} u_t(X_t^i) dt + dW_t^i, \quad \mathcal{L}(X_t^i) = u_t, \tag{3.19}
\]
with the same Brownian motions and initial condition \( X_0^i \) as the particle system. Note that \( u \in L^\infty([0,T]; L^\infty(\mathbb{R}^d)) \), thus by [37], the previous equation has a unique strong solution. Now consider for any \( N \in \mathbb{N}^* \) the empirical measure \( \overline{\mu}^N_\ell \) of the first \( N \) McKean-Vlasov particles. There is, for any \( t \in (0,T) \),
\[
W_1(\mu^N_t, u_t) \leq W_1(\mu^N_\ell, \overline{\mu}^N_\ell) + W_1(\overline{\mu}^N_\ell, u_t).
\]
For the first term, the coupling \( \pi^N_\ell = \frac{1}{N} \sum_{i=1}^N \delta_{(X^i_\ell, X^i_t)} \) gives
\[
\mathbb{E}[W_1(\mu^N_\ell, \overline{\mu}^N_\ell)] \leq \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^N |X^i_\ell - X^i_t| \right) \wedge 1 \right]
\]
\[
\leq \mathbb{E}[|X^i_\ell - X^i_t| \wedge 1]
\]
\[
= \mathbb{E} \left[ \frac{1}{2} \int_0^t u_s(X^i_\ell, X^i_t) ds \right] \wedge 1
\]
\[
\leq \mathbb{E} \left[ \frac{1}{2} \int_0^t \|u_s - u_t\|_{L^\infty(\mathbb{R})} ds \wedge 1 \right]
\]
\[
\leq C \sup_{s \leq t} \|u_s - u_t\|_{\beta,p} \wedge 1,
\]
that we bound using (3.17).

The remaining term \( \mathbb{E}[W_1(\overline{\mu}^N_\ell, u_t)] \) is bounded using Fournier-Guillin’s theorem [14, Theorem 1].

To apply this theorem, we need \( \int_\mathbb{R} |x|^q u_t(dx) = \mathbb{E}[|X^i_t|^q] < \infty \), which is satisfied for any \( q > 0 \) since the drift in (3.19) is bounded. So we get \( \mathbb{E}[W_1(\overline{\mu}^N_\ell, u)] \leq C N^{-1/2} \) for some universal constant \( C > 0 \), and (3.18) follows since \( \xi \leq 1/2 \).

4 The Keller-Segel equation

4.1 Setting and Preliminaries

In this whole section, the domain is \( D = \mathbb{T}^d \) with \( d \geq 2 \). We start with a precise definition of the particle system (1.2) in the case of the Keller-Segel model. The kernel \( K : \mathbb{T}^d \to \mathbb{R}^d \) of this model is the periodisation of the function defined on \([\frac{1}{2}\mathbb{T}^d \setminus \{0\}]\) by \( K_0(x) = -\chi_{\frac{1}{2}\mathbb{T}^d}(x) \).

First, notice that \( K \in L^p(\mathbb{T}^d) \) for any \( p \in \left[ 1, \frac{d}{d-1} \right) \). Hence, by the Hölder inequality, it holds
\[
\|K * f\|_{L^\infty(\mathbb{T}^d)} \leq C \|f\|_{L^q(\mathbb{T}^d)},
\]
(4.1)
for any \( q > d \).

The following property of the kernel will be frequently used:

Lemma 4.1. Let \( q \in (d, +\infty) \). Then we have
\[
\forall f \in L^q(\mathbb{T}^d), \quad \mathcal{N}_{1-\frac{d}{q}}(K * f) \leq C \|f\|_{L^q(\mathbb{T}^d)}.
\]
(4.2)
Proof. First, we will use the fact that $H^1_q(T^d)$ is continuously embedded into the Hölder space of parameter $1 - d/q$, denoted by $C^{1-d/q}(T^d)$. This comes from the equality between $H^1_q(T^d)$ and the Triebel-Lizorkin space $F^1_{q,2}(T^d)$ (see [29, Theorem. (v), p.168]), the equality between $C^{1-d/q}(T^d)$ and the Zygmund space $\mathcal{C}^{1-d/q}(T^d)$ (see [29, Theorem. (i)-(ii), p.168]), and the continuous embedding $F^1_{q,2}(T^d) \subset \mathcal{C}^{1-d/q}(T^d)$ (see [29, Corollary. (ii), p.170]). Hence we get

$$N_{1 - \frac{d}{q}}(K \ast f) \leq C \|K \ast f\|_{1,q}.$$

Now by Lemma 2.2, we have that $\|K \ast f\|_{1,q} \leq C \|\nabla K \ast f\|_{0,q} = C \|\nabla K \ast f\|_{L^1(T^d)}$, where the equality is the Littlewood-Paley theorem. We conclude using $\nabla K \in L^1(T^d)$ and a convolution inequality. □

Let us introduce a mollifier that will be used both to regularise the interaction kernel in the particle system and its empirical measure. Let $V : \mathbb{R}^d \to \mathbb{R}_+$ be a smooth probability density function with support in $(-\frac{1}{2}, \frac{1}{2})^d$. For any $x \in (-\frac{1}{2}, \frac{1}{2})^d$, define

$$V_0^N(x) := N^{d\alpha}V(N^\alpha x), \quad \text{for some } \alpha \in [0,1],$$

and let $V^N : T^d \to \mathbb{R}_+$ be the periodisation of $V_0^N$. Below, $\alpha$ will be restricted to some interval $(0, \alpha_0)$, see Assumption ($A_{\alpha,q}$).

Let $T > 0$. For each $N \in \mathbb{N}$, the particle system (1.2) reads more precisely:

$$
\left\{
\begin{aligned}
&dX^i_t = \frac{1}{N} \sum_{k=1}^N (K \ast V^N)(X^i_t - X^k_t) \, dt + \sqrt{2} \, dW^i_t, \quad t \leq T, \ 1 \leq i \leq N, \\
&X^i_0, \ 1 \leq i \leq N, \quad \text{are independent of } \{W^i, \ 1 \leq i \leq N\},
\end{aligned}
\right.
$$

where $\{(W^i_j)_{t \in [0,T]}, \ i \in \mathbb{N}\}$ is a family of independent standard $\mathbb{R}^d$-valued Brownian motions defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. The particles are periodised, hence we consider $X^i_t$ as a $T^d$-valued random variable. The empirical measure of $N$ particles and the mollified empirical measure are denoted by $\mu^N$ and $u^N$ respectively as in (3.4) and (3.5).

We will study the convergence of $u^N$ in $L^q(T^d)$ norm for any $q > d$. The restriction with respect to the parameter $\alpha$ and the hypothesis on the initial conditions of the system are given by the following assumption:

(A$_{\alpha,q}$): Let $q > d$ and assume that:

(A$_{\alpha,q}^{\text{u}}$) The parameters $\alpha$ and $q > d$ satisfy

$$0 < \alpha < \frac{1}{2d(1 - \frac{1}{q})}.$$

(A$_{\alpha,q}^{\text{u}}$) Assume $u_0 \in L^1 \cap L^q(T^d)$ such that $\|u_0\|_{L^1(T^d)} = 1$. For any $m \geq 1$, assume

$$
\mathbb{E} \left[ \left\| u^N_0 - u_0 \right\|_{L^m(T^d)}^m \right] \leq \frac{C}{N^\varrho},
$$

where

$$\varrho = \min \left( \alpha(1 - \frac{d}{q}), 1 - 2\alpha d(1 - \frac{1}{q}) \right). \quad (4.3)
$$

We aim to prove the convergence of the mollified empirical measure to the PDE (1.1). As (1.1) preserves the positivity and total mass $M := \int_{T^d} u_0(x) \, dx$, we will assume throughout the paper that $M = 1$.

Solutions to (1.1) will be understood in the following mild sense:
Definition 4.2. Let $q > d$. Given $u_0 \in L^q(\mathbb{T}^d)$ and $T > 0$, a function $u$ on $[0, T] \times \mathbb{T}^d$ is said to be a mild solution to (1.1) on $[0, T]$ if

(i) $u \in C([0, T]; L^q(\mathbb{T}^d))$;

(ii) $u$ satisfies the integral equation

$$u_t = e^{t \Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s) \Delta} (u_s K \ast u_s) \, ds, \quad 0 \leq t \leq T. \quad (4.4)$$

A function $u$ on $[0, \infty) \times \mathbb{T}^d$ is said to be a global mild solution to (1.1) if it is a mild solution to (1.1) on $[0, T]$ for all $T > 0$, and it is said to be local otherwise.

We have the following local well-posedness for the Keller-Segel equation.

Proposition 4.3. Let $q > d$ such that $u_0 \in L^q(\mathbb{T}^d)$. Then there exists a maximal time $T^* \in (0, +\infty]$ and a unique mild solution to the Fokker-Planck equation which is in $C\left([0, T], L^q(\mathbb{T}^d)\right)$, for any $T < T^*$.

Proof. Denote $X = C\left([0, T], L^q(\mathbb{T}^d)\right)$ and $\| \cdot \|_X$ the associated norm. Then define $B(u, v) = \int_0^t \nabla \cdot e^{(t-s) \Delta} (u_s K \ast v_s) \, ds$ for $u, v \in X$. In view of (2.6), we have

$$\|B(u, v)\|_X \leq C \int_0^T \frac{\|u_s K \ast v_s\|_{L^q(\mathbb{T}^d)}}{\sqrt{t - s}} \, ds \leq C \int_0^T \frac{\|u_s\|_{L^q(\mathbb{T}^d)} \|K \ast v_s\|_{L^\infty(\mathbb{T}^d)}}{\sqrt{t - s}} \, ds.$$ 

Now using (4.1), we get

$$\|B(u, v)\|_X \leq C T^{1/2} \|u\|_X \|v\|_X.$$ 

Thus by the same contraction principle as before, see Theorem 13.2 in [21], we deduce the result for $T$ small enough.

Unlike the Burgers equation, we do not expect to have a global solution for any Keller-Segel kernel $K$, depending on the dimension $d$ and the intensity $\chi_d$ of the interaction. Thus we keep working with local solutions in the sequel.

4.2 Convergence

Recall that $T^*$ is the maximal time of existence of (4.4) from Proposition 4.3, and let $T < T^*$. In the following, $\mathcal{L}(\mu^N)$ denotes the law of $\mu^N$ in the space $C([0, T]; M_1(\mathbb{T}^d))$, where $M_1(\mathbb{T}^d)$ is endowed with the Kantorovich-Rubinstein distance $\| \cdot \|_0$, and $W_1^{(\mathcal{L})}$ denotes the 1-Wasserstein distance on the space of probability measures on $C([0, T]; M_1(\mathbb{T}^d))$ endowed with the distance $d(\mu, \nu) = 1 \wedge \sup_{t \in [0, T]} |\mu_t - \nu_t|$.

Theorem 4.4. Assume that $(A_{\alpha, q})$ holds and recall that $q$ was defined in (4.3). Then for any $\varepsilon > 0$,

(i) $\lim_{N \to +\infty} N^{\varepsilon - \varepsilon} \sup_{t \in [0, T]} \|u_t^N - u_t\|_{L^q(\mathbb{T}^d)} = 0$ a.s.

(ii) Besides, there exists $C > 0$ such that for all $N \in \mathbb{N}^*$,

$$W_1^{(\mathcal{L})}(\mathcal{L}(\mu^N), \delta_u) \leq CN^{-(\varepsilon - \varepsilon)}.$$ 

Remark 4.5. This theorem also holds if one replaces the Keller-Segel kernel by an integrable kernel that satisfies the uniform bound (4.1) and the Hölder estimate (4.2). This is for instance the case of the periodised version of the Coulomb kernel given by $K(x) = \pm \nabla |x|^{- (d - 2)}$ and of the Biot-Savart kernel $K(x) = \frac{1}{\pi |x|^2}$ in dimension 2.
Remark 4.6. • In view of the constraint \((A^i)\), this theorem gives the almost sure convergence of \(u^N\) and \(\mu^N\) for any \(\alpha < \frac{1}{2(d+1)}\) by choosing \(q\) close to \(d\), provided the initial condition has enough regularity. On the other hand, the best possible rate of convergence one can get here is \(\rho = \frac{1}{2d+1}\), by choosing \(q = +\infty\) and \(\alpha = \frac{1}{2d+1}\).

• Compared to Theorem 1.3 in [25], where a cutoff was applied on the drift of the particles, but where the particles lived in \(\mathbb{R}^d\), we get the convergence of \(u^N\) and \(\mu^N\) for the same range of the parameter \(\alpha\). However, the best possible rate of convergence was \(\frac{1}{2(d+1)} - \varepsilon\), which is slightly less than here. This is explained by the tails of \(K\) that need to be integrated in \(\mathbb{T}^d\), while they play no role in \(\mathbb{R}^d\).

Remark 4.7. As already mentioned in the introduction, an interesting feature of Keller-Segel equations is that they exhibit a blow-up for large enough \(\chi_2\). In [30], a slightly modified 2-dimensional Keller-Segel equation on the torus was studied and a condition on \(\chi_2\) for existence of weak solutions was provided. Whether there is a blow-up or not, our approach provides an approximation of the PDE until maximal existence time.

Proof. We prove (i) first and then use it to prove (ii).

Proof of (i). Similarly to [25, Eq. (2.3)], we obtain the following mild formulation for \(x \in \mathbb{T}^d\):

\[
u_t^N(x) = e^{t\Delta} u_0^N(x) - \int_0^t \nabla \cdot e^{(t-s)\Delta} (\mu^N_s, V^N(x - \cdot) K * u^N_s(\cdot)) \, ds - \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X^i_s) \cdot dW_i.
\]

Hence

\[
u_t^N(x) - u_t(x) = e^{t\Delta} (u_0^N - u_0)(x) - \int_0^t \nabla \cdot e^{(t-s)\Delta} (\mu^N_s, V^N(x - \cdot) K * u^N_s(\cdot)) - u_s(x) K * u_s(x) \, ds
- \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X^i_s) \cdot dW_i
- e^{t\Delta} (u_0^N - u_0)(x) + \int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s K * u_s - u_s^N K * u_s^N)(x) \, ds + E_t(x) + M_t^N(x),
\]

where we have set

\[
E_t(x) := \int_0^t \nabla \cdot e^{(t-s)\Delta} (\mu^N_s, V^N(x - \cdot) (K * u^N_s(x) - K * u^N_s(\cdot))) \, ds,
\]

\[
M_t^N(x) := -\frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X^i_s) \cdot dW_i.
\]

(4.5)

In view of the estimate (2.6), one has

\[
\|u_t^N - u_t\|_{L^q(\mathbb{T}^d)} \leq \|e^{t\Delta} (u_0^N - u_0)\|_{L^q(\mathbb{T}^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s K * u_s - u_s^N K * u_s^N\|_{L^q(\mathbb{T}^d)} \, ds + \|E_t\|_{L^q(\mathbb{T}^d)} + \|M_t^N\|_{L^q(\mathbb{T}^d)}.
\]
It follows that
\[
\|u_t^N - u_t\|_{L^q(T^d)} \leq \|u_0^N - u_0\|_{L^q(T^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|K * u_s^N\|_{L^\infty(T^d)} \|u_s^N - u_s\|_{L^q(T^d)} ds
\]
\[+ C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s\|_{L^q(T^d)} \|K * (u_s - u_s^N)\|_{L^\infty(T^d)} ds
\]
\[+ \|E_t\|_{L^q(T^d)} + \|M_t^N\|_{L^q(T^d)}.
\]
Using (4.1), we get
\[
\|u_t^N - u_t\|_{L^q(T^d)} \leq \|u_0^N - u_0\|_{L^q(T^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^q(T^d)} \left(\|u_s^N - u_s\|_{L^q(T^d)} + \|u_s\|_{L^q(T^d)}\right) ds
\]
\[+ C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s\|_{L^q(T^d)} \|u_s - u_s^N\|_{L^q(T^d)} ds
\]
\[+ \|E_t\|_{L^q(T^d)} + \|M_t^N\|_{L^q(T^d)}.
\]
We know from Proposition 4.3 that \(u\) is bounded in \(C \left([0, T], L^q(T^d)\right)\). Hence
\[
\|u_t^N - u_t\|_{L^q(T^d)} \leq \|u_0^N - u_0\|_{L^q(T^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^q(T^d)}^2 ds
\]
\[+ C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s - u_s^N\|_{L^q(T^d)} ds
\]
\[+ \|E_t\|_{L^q(T^d)} + \|M_t^N\|_{L^q(T^d)}.
\]

- Let us bound \(E\). By using the positivity of \(V^N\), we get
\[
\|E_t\|_{L^q(T^d)} \leq C \int_0^t \frac{1}{\sqrt{t-s}} \left(\int_{T^d} \langle \mu_s^N, V^N(x - \cdot) | K * u_s^N(\cdot) - K * u_s^N(x) \rangle^q dx \right)^{\frac{1}{q}} ds.
\]
Using the Hölder continuity of \(K * u^N\) from Lemma 4.1, we get
\[
\|E_t\|_{L^q(T^d)} \leq C \int_0^t \frac{\|u_s^N\|_{L^q(T^d)}}{\sqrt{t-s}} \left(\int_{T^d} \langle \mu_s^N, V^N(x - \cdot) | - x \rangle^{1-\frac{q}{4}} dx \right)^{\frac{1}{q}} ds.
\]
Since \(V\) is compactly supported, we have that \(V^N(x - y) | y - x |^{1-\frac{q}{4}} \leq N^{-\alpha(1-\frac{q}{4})} V^N(x - y)\). Thus,
\[
\sup_{s \in [0, t]} \|E_s\|_{L^q(T^d)} \leq \frac{C}{N^{\alpha(1-d/q)}} \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N\|_{L^q(T^d)}^2 ds
\]
\[\leq \frac{C}{N^{\alpha(1-d/q)}} \int_0^t \frac{1}{\sqrt{t-s}} \left(\|u_s\|_{L^q(T^d)} + \|u_s^N - u_s\|_{L^q(T^d)}\right)^2 ds
\]
and using again the boundedness of \(u\) we get
\[
\sup_{s \in [0, t]} \|E_s\|_{L^q(T^d)} \leq \frac{C}{N^{\alpha(1-d/q)}} \left(1 + \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^q(T^d)}^2 ds\right).
\]

- We now focus on \(M^N\). For \(q \geq 2\), we have the embedding of \(H^{d(\frac{1}{2} - \frac{1}{q})}_2(T^d)\) into \(L^q(T^d)\). Hence by Proposition A.3, it comes that for any \(q \geq 2\) and any \(m \geq 1\),
\[
\forall N \in \mathbb{N}^*, \quad \left\|\sup_{s \in [0, t]} \|M_s^N\|_{L^q(T^d)}\right\|_{L^m(\Omega)} \leq C \left\|\sup_{s \in [0, t]} \|M_s^N\|_{H^{d(\frac{1}{2} - \frac{1}{q})}_2(T^d)}\right\|_{L^m(\Omega)}
\]
\[\leq C N^{-\frac{1}{2}(1-2m(\frac{1}{2} - \frac{1}{q}))^+}.
\] (4.6)
By Borel-Cantelli’s lemma, we deduce that there exists a random variable $A_0$ with finite moments such that almost surely,

$$\sup_{s \in [0,t]} \|M^N_s\|_{L^q(\mathbb{R}^d)} \leq \frac{A_0}{N^{3(1-2\alpha d(1-1/q)) - 2\varepsilon}}.$$ 

- Putting altogether the previous bounds and $(A_N^{4\alpha,q})$, there exists a nonnegative random variable $A$ with finite moments such that almost surely,

$$\|u_t^N - u_t\|_{L^3(\mathbb{R}^d)} \leq \frac{A}{N^{\alpha(1-d/q) + N^2(1-2\alpha d(1-1/q)) - 2\varepsilon}} + C \int_0^t \frac{1}{\sqrt{t-s}} \left( \|u_s^N - u_s\|_{L^3(\mathbb{R}^d)} + \|u_s^N - u_s\|^2_{L^2(\mathbb{R}^d)} \right) ds \leq \frac{A}{N^{\alpha(1-d/q) + N^2(1-2\alpha d(1-1/q)) - 2\varepsilon}} + C \int_0^t \frac{1}{\sqrt{t-s}} \left( \|u_s^N - u_s\|_{L^3(\mathbb{R}^d)} + \|u_s^N - u_s\|^2_{L^2(\mathbb{R}^d)} \right) ds.$$

Denote $V_t^N = \|u_t^N - u_t\|^3_{L^3(\mathbb{R}^d)}$ and

$$A_N = \frac{A^3}{N^{3(\varepsilon - \delta)}}.$$ 

Then apply Hölder’s inequality with the exponents $\frac{3}{2}$ and 3 to obtain

$$V_t^N \leq A_N + C T^\frac{3}{2} \int_0^t (V_s^N + (V_s^N)^2) ds.$$

Now we apply Bihari’s inequality (see e.g. [10, Theorem 27]) and get that

$$V_t^N \leq G^{-1}(G(A_N) + CTt), \text{ for any } t \in [0, \tau_N(\omega)) \cap [0, T], \tag{4.8}$$

where $C_T = C\sqrt{T}$ and

$$G(x) = \int_1^x \frac{1}{y + y^2} dy = \log \left( \frac{2x}{1+x} \right),$$

$$G^{-1}(x) = \frac{e^x}{2 - e^x},$$

$$\tau_N(\omega) = \frac{1}{C_T} \log \left( \frac{1 + A_N(\omega)}{A_N(\omega)} \right).$$

Hence, (4.8) reads

$$\|u_t^N - u_t\|^3_{L^3(\mathbb{R}^d)} \leq e^{C_T t} \frac{2A_N/(1 + A_N)}{2 - e^{C_T t}2A_N/(1 + A_N)}, \text{ for any } t \in [0, \tau_N(\omega)) \cap [0, T]. \tag{4.9}$$

Let $\tilde{\Omega}$ be a measurable subset of $\Omega$ of measure 1 on which $A$ is finite and (4.8) holds. For $\omega \in \tilde{\Omega}$, there exists $N_0(\omega)$ such that for any $N \geq N_0(\omega)$, $\tau_{N_0} > T$. Thus

$$\sup_{N \geq N_0(\omega)} \sup_{t \in [0,T]} \|u_t^N - u_t\|^3_{L^3(\mathbb{R}^d)} \leq \sup_{N \geq N_0(\omega)} N^{3(\varepsilon - 2\varepsilon)} e^{C_T t} \frac{2A_N/(1 + A_N)}{2 - e^{C_T t}2A_N/(1 + A_N)} < \infty.$$ 

Hence

$$\limsup_{N \to +\infty} N^{3(\varepsilon - 2\varepsilon)} \sup_{t \in [0,T]} \|u_t^N - u_t\|^3_{L^3(\mathbb{R}^d)} \leq \limsup_{N \to +\infty} N^{3(\varepsilon - 2\varepsilon)} e^{C_T t} \frac{2A_N/(1 + A_N)}{2 - e^{C_T t}2A_N/(1 + A_N)} = 0,$$

which gives point (i) of the theorem.
Proof of (ii). Recalling the definition (2.3) of the Kantorovich-Rubinstein distance, it comes by choosing the trivial coupling \( L(\mu^N) \otimes \delta_a \) that
\[
W_1(L(\mu^N), \delta_a) \leq E \left[ 1 \wedge \sup_{t \in [0,T]} \| \mu^N_t - u_t \|_0 \right]
\leq E \left[ 1 \wedge \sup_{t \in [0,T]} \| \mu^N_t - u^N_t \|_0 \right] + E \left[ 1 \wedge \sup_{t \in [0,T]} \sup_{\phi \in L^\infty} \langle u^N_t - u_t, \phi \rangle \right]
\leq E \left[ 1 \wedge \sup_{t \in [0,T]} \| \mu^N_t - u^N_t \|_0 \right] + E \left[ 1 \wedge \sup_{t \in [0,T]} \| u^N_t - u_t \|_{L^\infty(T)} \right].
\] (4.10)

We first treat the second term on the r.h.s. and will prove that
\[
E \left[ 1 \wedge \sup_{t \in [0,T]} \| u^N_t - u_t \|_{L^\infty(T)} \right] \leq C N^{-q(\varepsilon - \delta)}.
\] (4.11)

Consider \( N_0 \equiv N_0(\omega) \) the smallest integer such that \( \tau_{N_0} > T + \frac{1}{C_T} \log(2) \). Then we get
\[
E \left[ 1 \wedge \sup_{t \in [0,T]} \| u^N_t - u_t \|_{L^\infty(T)} \right] \leq E \left[ 1_{\{N \geq N_0\}} \sup_{t \in [0,T]} \| u^N_t - u_t \|_{L^\infty(T)} \right] + \mathbb{P}(N_0 \geq N).
\]
The above choice of \( N_0 \) induces that for \( N \geq N_0 \), we have that \( \frac{2A_N}{1 + A_N} e^{C_T T} = 2e^{C_T T \tau_{N_0}} \leq 1 \). Hence, in view of (4.9), which holds true for any \( t \leq T \) on the event \( \{ N \geq N_0 \} \), we obtain
\[
E \left[ 1_{\{N \geq N_0\}} \sup_{t \in [0,T]} \| u^N_t - u_t \|_{L^\infty(T)} \right] \leq e^{\frac{1}{2} C_T T} E \left[ \left( \frac{2A_N/(1 + A_N)}{2 - e^{C_T T}/2A_N/(1 + A_N)} \right)^{\frac{1}{2}} \right]
\leq e^{\frac{1}{2} C_T T} E \left[ \left( \frac{2A_N}{1 + A_N} \right)^{\frac{1}{2}} \right]
\leq e^{\frac{1}{2} C_T T} N^{-(\varepsilon - \delta)},
\]
using the definition (4.7) of \( A_N \) and the fact that \( A \) has finite moments.

Now we estimate \( \mathbb{P}(N_0 \geq N) \). By the definition of \( N_0 \), we have that \( \tau_{N_0-1} \leq T + \frac{1}{C_T} \log(2) \).
Hence in view of the definition of \( \tau_{N_0-1} \) and \( A_N \), we deduce that
\[
N_0 \leq 1 + \left( 2e^{C_T T} - 1 \right) \frac{1}{(N-1)^{\frac{1}{2} (\varepsilon - \delta)}} A^{\frac{1}{2} (\varepsilon - \delta)}.
\]
Now we get
\[
\mathbb{P}(N_0 \geq N) \leq \mathbb{P} \left( A \geq \frac{(N-1)^{\varepsilon - \delta}}{(2e^{C_T T} - 1)^{\frac{1}{2}}} \right)
\leq C \frac{E A^p}{(N-1)^{p(\varepsilon - \delta)}},
\]
by the Markov inequality, for any \( p \geq 1 \). Hence (4.11) follows.

For the first term in the right-hand side of (4.10), we observe that
\[
|\langle \mu^N_t, \phi \rangle - \langle u^N_t, \phi \rangle | = |\langle \mu^N_t, (\phi - \phi \ast V^N) \rangle | \leq \left( \mu^N_t \int_{\mathbb{T}^d} V(y) \left| \phi(\cdot) - \phi \left( \frac{y}{N^\alpha} \right) \right| dy \right)
\leq C \| \phi \|_{Lip} \frac{1}{N^\alpha}.
\]
In view of (4.10), (4.11) and the above inequality, the desired result follows. \( \square \)

Remark 4.8. As in Remark 3.5, one could be interested in a notion of convergence in Wasserstein distance that is stronger than the one given in Theorem 3.3(ii) and Theorem 4.4(ii), namely a
convergence at the level of the (random) empirical measure rather than its law. Indeed, at least formally, there is \( W_{1}^{(L)}(\mathcal{L}(\mu^{N}),\delta_u) \leq \mathbb{E}[V_{1}(\mu^{N},u)] \). In Remark 3.5, we were able to give a Wasserstein bound on the marginals. We could follow the same path here. We also notice that one can proceed as in \[ we can state the following bound on \( M \) in (2.4) on \( \mathbb{R} \) the torus, one can check that all computations go through similarly on both \( D \).

Appendix

The goal of this appendix is to prove the estimates (3.12) and (4.6) on the stochastic integrals defined respectively in (3.11) for Burgers and in (4.5) for Keller-Segel. Since we will deal here with both \( D = \mathbb{R} \) and \( D = \mathbb{T}^{d} \), we denote by \( \| \cdot \|_{H_{\beta}^{2}(D)} \) the Bessel norm in \( H_{\beta}^{2}(D) \), which was defined in (2.4) on \( \mathbb{R} \) and in (2.5) on the torus.

In Proposition B.3 of \[ we now give a version of Garsia-Rodemich-Rumsey’s Lemma \[ (\text{for } \mathbb{R}-\text{valued processes, this lemma already appears in } \text{[28, Corollary 4.4]}, \text{and the extension to Banach spaces is consistent with Garsia-Rodemich-Rumsey’s Lemma with no additional difficulty, see e.g. } \text{[17, Theorem A.1]}].)

**Proposition A.1.** Let \( \beta \in \mathbb{R}, m \geq 1 \) and \( \alpha \in [0,1] \) that defines \( V^{N} \) as in (3.2). For any \( \delta \in (0,1] \), there exists \( C > 0 \) such that

\[
\| M^{N}_{t} - M^{N}_{t} \|_{H_{\beta}^{2}(D)} \|_{L^{m}(\Omega)} \leq C (t-s)^{\frac{d}{2} \delta N^{-\frac{1}{2} (1-\alpha(d+4\delta + 2\beta))}}, \quad \forall s \leq t \in [0,T], \forall N \in \mathbb{N}^*.
\]

We now give a version of Garsia-Rodemich-Rumsey’s Lemma \[ for \mathbb{R}-\text{valued processes, this lemma already appears in } \text{[28, Corollary 4.4]}, \text{and the extension to Banach spaces is consistent with Garsia-Rodemich-Rumsey’s Lemma with no additional difficulty, see e.g. } \text{[17, Theorem A.1]}].

**Lemma A.2.** Let \( E \) be a Banach space and \( (Y^{n})_{n \geq 1} \) be a sequence of \( E \)-valued continuous processes on \( [0,T] \). Let \( m \geq 1 \) and \( \eta > 0 \) such that \( mn \geq 1 \) and assume that there exists a constant \( C_{0} > 0 \) and a sequence \( (\delta_{n})_{n \geq 1} \) of positive real numbers such that

\[
\left( \mathbb{E} \left[ \| Y^{n}_{s} - Y^{n}_{t} \|_{E}^{m} \right] \right)^{\frac{1}{m}} \leq C_{0} s-t|n \delta_{n}, \quad \forall s,t \in [0,T], \forall n \geq 1.
\]

Then for any \( m_{0} \in (0,m] \), there exists a constant \( C \), depending only on \( C_{0}, m_{0}, \eta \), and \( T \), such that \( \forall n \geq 1, \)

\[
\left( \mathbb{E} \left[ \sup_{t \in [0,T]} \| Y^{n}_{t} - Y^{n}_{0} \|_{E}^{m_{0}} \right] \right)^{\frac{1}{m_{0}}} \leq C \delta_{n}.
\]
These two results can be combined to obtain the following bound:

**Proposition A.3.** Let $\beta \in \mathbb{R}$, $m \geq 1$ and $\alpha \in [0, 1]$ that defines $V^N$ as in (3.2). Let $\varepsilon \in (0, 2\alpha)$. Then there exists $C > 0$ such that for any $t \in [0, T]$ and $N \in \mathbb{N}^*$,

$$\left\| \sup_{s \in [0, t]} \left\| M_s^N \right\|_{H_2^\beta(D)} \right\|_{L^\infty(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(d+2\beta)) + \varepsilon},$$

**Proof.** We aim to apply Lemma A.2 to $M^N$ in the Hilbert space $H_2^\beta(D)$. Let $\varepsilon > 0$ and $m_0 > 0$. With the notations of Proposition A.1, let us choose $\delta = \frac{\varepsilon}{2\alpha}$, $\eta = \frac{\varepsilon}{2}$ and $\delta_N = N^{-\rho}$ with $\rho = -\frac{1}{2} (1 - \alpha(1+ 2\beta)) + 2\alpha \delta = -\frac{1}{2} (1 - \alpha(1+ 2\beta)) + \varepsilon$. Hence, choosing $m \geq 1 \lor m_0$ large enough so that $m\eta > 1$, the inequality in Proposition A.1 shows that $M^N_s$ satisfies the conditions of Lemma A.2 and the desired result follows. \hfill \Box

**References**

[1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, 2011.

[2] P. Biler. Mathematical challenges in the theory of chemotaxis. *Ann. Math. Sil.*, 32(1):43–63, 2018.

[3] V. I. Bogachev. *Measure theory. Vol. II*. Springer-Verlag, Berlin, 2007.

[4] M. Bossy and D. Talay. Convergence rate for the approximation of the limit of weakly interacting particles: application to the Burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 1996.

[5] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.

[6] D. Bresch, P.-E. Jabin, and Z. Wang. Mean field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.*, 172(13):2591–2641, 2023.

[7] P. Cattiaux and L. Pêdèches. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA*, 13:447–463, 2016.

[8] J. D. Cole. On a quasi-linear parabolic equation occurring in aerodynamics. *Quart. Appl. Math.*, 9:225–236, 1951.

[9] F. Flandoli. *Some Gronwall type inequalities and applications*. Nova Science Publishers, Inc., Hauppauge, NY, 2003.

[10] F. Flandoli and M. Leocata. A particle system approach to aggregation phenomena. *J. Appl. Probab.*, 56(1):282–306, 2019.

[11] F. Flandoli, M. Leimbach, and C. Olivera. Uniform convergence of proliferating particles to the FKPP equation. *J. Math. Anal. Appl.*, 473(1):27–52, 2019.

[12] F. Flandoli, C. Olivera, and M. Simon. Uniform approximation of 2 dimensional Navier-Stokes equation by stochastic interacting particle systems. *SIAM J. Math. Anal.*, 52(6):5339–5362, 2020.

[13] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015.
[15] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 2017.

[16] N. Fournier and Y. Tardy. Collisions of the supercritical Keller-Segel particle system. *J. Eur. Math. Soc. (to appear), arXiv:2110.08490*, 2023.

[17] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[18] A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.*, 20:565–578, 1970/71.

[19] E. Gutkin and M. Kac. Propagation of chaos and the Burgers equation. *SIAM J. Appl. Math.*, 43(4):971–980, 1983.

[20] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.

[21] P. G. Lemarié-Rieusset. *Recent developments in the Navier-Stokes problem*, volume 431 of *Chapman & Hall/CRC Research Notes in Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2002.

[22] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, Lecture Series in Differential Equations, Session 7, pages 41–57. Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, Arlington, VA, 1967.

[23] S. Méléard and S. Roelly-Coppoletta. A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.*, 26(2):317–332, 1987.

[24] K. Oelschläger. A law of large numbers for moderately interacting diffusion processes. *Z. Wahrsch. Verw. Gebiete*, 69(2):279–322, 1985.

[25] C. Olivera, A. Richard, and M. Tomašević. Quantitative particle approximation of nonlinear Fokker-Planck equations with singular kernel. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 24(2):691–749, 2023.

[26] H. Osada and S. Kotani. Propagation of chaos for the Burgers equation. *J. Math. Soc. Japan*, 37(2):275–294, 1985.

[27] B. Perthame. PDE models for chemotactic movements: parabolic, hyperbolic and kinetic. *Appl. Math.*, 49(6):539–564, 2004.

[28] A. Richard, X. Tan, and F. Yang. Discrete-time simulation of stochastic Volterra equations. *Stochastic Process. Appl.*, 141:109–138, 2021.

[29] H.-J. Schmeisser and H. Triebel. *Topics in Fourier analysis and function spaces*. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1987.

[30] T. Senba and T. Suzuki. Weak solutions to a parabolic-elliptic system of chemotaxis. *J. Funct. Anal.*, 191(1):17–51, 2002.

[31] M. Simon and C. Olivera. Non-local conservation law from stochastic particle systems. *J. Dynam. Differential Equations*, 30(4):1661–1682, 2018.

[32] A.-S. Sznitman. A propagation of chaos result for Burgers’ equation. *Probab. Theory Relat. Fields*, 71(4):581–613, 1986.
[33] Y. Tardy. Convergence of the empirical measure for the Keller-Segel model in both subcritical and critical cases. Preprint arXiv:2205.04968, 2022.

[34] H. Triebel. Interpolation theory, function spaces, differential operators, volume 18 of North-Holland Mathematical Library. North-Holland Publishing Co., Amsterdam-New York, 1978.

[35] H. Triebel. Theory of function spaces, volume 78 of Monographs in Mathematics. Birkhäuser Verlag, Basel, 1983.

[36] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in UMD Banach spaces. Ann. Probab., 35(4):1438–1478, 2007.

[37] A. J. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. Mat. Sb. (N.S.), 111(153)(3):434–452, 480, 1980.