Ann. Inst. H. Poincaré
Anal. Non Linéaire (Online first)
DOI 10.4171/AIHPC/105

Published by EMS Press

Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows

Antonin Chodron de Courcel, Matthew Rosenzweig, and Sylvia Serfaty

Abstract. We consider conservative and gradient flows for N-particle Riesz energies with mean-field scaling on the torus $\mathbb{T}^d$, for $d \geq 1$, and with thermal noise of McKean–Vlasov type. We prove global well-posedness and relaxation to equilibrium rates for the limiting PDE. Combining these relaxation rates with the modulated free energy of Bresch et al. (2019, 2019–2020, 2020) and recent sharp functional inequalities of the last two named authors for variations of Riesz modulated energies along a transport, we prove uniform-in-time mean-field convergence in the gradient case with a rate which is sharp for the modulated energy pseudo-distance. For gradient dynamics, this completes in the periodic case the range $d-2 \leq s < d$ not addressed by the previous work Rosenzweig and Serfaty (2023). We also combine our relaxation estimates with the relative entropy approach of Jabin and Wang (2018) for so-called $\dot{W}^{-1,\infty}$ kernels, giving a proof of uniform-in-time propagation of chaos alternative to Guillin et al. (2021).

1. Introduction

1.1. The problem

We are interested in proving mean-field convergence, i.e., the large N limiting behavior of dynamics for stochastic singular interacting particle systems of the form

$$\begin{cases} dx_i^t = \dfrac{1}{N} \displaystyle\sum_{1 \leq j \leq N : j \neq i} \mathbb{M} \nabla \mathsf{g}(x_i^t - x_j^t)\, dt + \sqrt{2\sigma}\, dW_i^t, \\ x_i^t|_{t=0} = x_i^0, \end{cases} \qquad i \in \{1, \ldots, N\}. \tag{1.1}$$

Above, $x_i^0 \in \mathbb{T}^d$, the flat torus in dimension $d \geq 1$ which we identify with $[-\frac{1}{2}, \frac{1}{2}]^d$ under periodic boundary conditions, are the pairwise distinct initial positions; $\{W_i\}_{i=1}^N$ are independent standard Brownian motions in $\mathbb{T}^d$, so that the noise in (1.1) is of so-called additive type; the coefficient σ, which may be interpreted as temperature, is nonnegative; and $\mathbb{M}$ is a $d \times d$ matrix with constant real entries. We will either choose $\mathbb{M} = -\mathbb{I}$, corresponding to *gradient/dissipative* dynamics or choose $\mathbb{M}$ to be antisymmetric, corresponding to

2020 Mathematics Subject Classification. Primary 35Q35; Secondary 35Q70, 35Q82, 35Q84, 37A60, 60K35, 82C22.
Keywords. Mean-field limit, propagation of chaos, Coulomb/Riesz gas, McKean–Vlasov, uniform in time.

Hamiltonian/conservative dynamics. Mixed flows are also allowable, but here our main results will concern gradient flows. The motivation for considering $\mathbb{T}^d$, as opposed to Euclidean space $\mathbb{R}^d$, will be explained below.

The interaction potential g that we will study is a periodic Riesz potential (indexed by a parameter $-1 \le s < d$), that is the zero average solution to

$$|\nabla|^{d-s}\mathsf{g} = \mathsf{c}_{d,s}(\delta_0 - 1), \quad \mathsf{c}_{d,s} := \begin{cases} \dfrac{4^{\frac{d-s}{2}}\Gamma((d-s)/2)\pi^{d/2}}{\Gamma(s/2)}, & -1 \le s < d, \\ \dfrac{\Gamma(d/2)(4\pi)^{d/2}}{2}, & s = 0. \end{cases} \tag{1.2}$$

The notation $|\nabla|^{d-s}$ denotes the Fourier multiplier with symbol $(2\pi|\xi|)^{d-s}$. As explained in Section 3, the potential g behaves like $|x|^{-s}$, if $-1 \le s < d$, or $-\log|x|$, if $s = 0$, near the origin. This is a model choice for studying systems, such as (1.1), with interactions between particles that become singular as the inter-particle distance tends to zero. The family of potentials defined by (1.2) includes the physically important *Coulomb* case $s = d-2$, as well as the *sub-Coulomb* range $s < d-2$ and *super-Coulomb* range $d-2 < s < d$. We are primarily interested in the super-Coulomb case, as we will explain momentarily. Unlike in the setting of $\mathbb{R}^d$, where $\mathsf{g}(x) = |x|^{-s}$ for $s \neq 0$ and $\mathsf{g}(x) = -\log|x|$ for $s = 0$, the potential g on $\mathbb{T}^d$ does not have a simple form (see [61, 62] for various representations of periodic Riesz potentials). However, one can show that in a neighborhood of the origin, g equals its Euclidean analogue plus a smooth correction (see (3.1) below). We limit ourselves to the *potential* case $s < d$, in which $\mathsf{g} \in L^1(\mathbb{T}^d)$. The *hypersingular* case $s \ge d$ is also interesting, but is a fundamentally different regime and will not be considered in this article. We refer to [19, Chapter 8], [59, 60], and references therein for more on this case.

Applications of systems of the form (1.1) are numerous. Since this topic has been discussed at length elsewhere, we will not repeat this discussion. Instead, we refer the reader to the introduction of [91], the survey [67], and the recent lecture notes [38, 52].

One can show by a truncation and stopping time argument that there is a unique, local strong solution to system (1.1). When $s \le d-2$, one can then use the energy of the system (which has nonincreasing expectation) to show that the solution is global. In particular, with probability 1, the particles never collide. This has been shown in [91, Section 4] for the case $s < d-2$ in Euclidean space, but the argument is adaptable to the periodic setting without issue and, with a little more work, to the Coulomb case $s = d-2$ as well. When $d-2 < s < d$, the aforementioned global existence argument fails, in short because $\Delta\mathsf{g} \to +\infty$ as $|x| \to \infty$, as opposed to $-\infty$ when $s < d-2$. Consequently, it is unclear how to make sense of the system of SDEs (1.1), except on very short timescales which a priori vanish as $N \to \infty$.[1] Accordingly, rather than work with (1.1) directly, we work

[1] Elias Hess-Childs has recently informed us that it is possible to show well-posedness of the SDEs in the gradient flow case for $\max(0, d-2) < s < d$.

with the *Liouville/forward Kolmogorov equation*

$$\begin{cases} \partial_t f_N = -\sum_{i=1}^{N} \mathrm{div}_{x_i}\left(f_N \frac{1}{N} \sum_{1 \le j \le N : j \ne i} \mathbb{M}\nabla \mathrm{g}(x_i - x_j)\right) + \sigma \sum_{i=1}^{N} \Delta_{x_i} f_N, \\ f_N|_{t=0} = f_N^0, \end{cases} \tag{1.3}$$

which is obtainable from (1.1) through Itô's formula. Here, the initial positions of the particles are thought of as random vectors in $\mathbb{T}^d$ distributed according to a probability density f_N^0, and f_N^t is the law of the solution $\underline{x}_N^t := (x_1^t, \ldots, x_N^t)$ to (1.1).

Establishing the mean-field limit refers to showing the weak convergence (in the sense of probability measures) as $N \to \infty$ of the *empirical measure*

$$\mu_N^t := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^t}$$

associated to a solution $\underline{x}_N^t := (x_1^t, \ldots, x_N^t)$ of system (1.1). For fixed t, we note that the empirical measure is a random Borel probability measure on $\mathbb{T}^d$. If the points x_i^0, which themselves depend on N, are such that μ_N^0 converges to some sufficiently regular measure μ^0, then a formal application of Itô's lemma leads to the expectation that for $t > 0$, μ_N^t converges as $N \to \infty$ to the solution of the Cauchy problem

$$\begin{cases} \partial_t \mu = -\mathrm{div}(\mu \mathbb{M}\nabla \mathrm{g} * \mu) + \sigma \Delta \mu, \\ \mu|_{t=0} = \mu^0, \end{cases} \qquad (t,x) \in \mathbb{R}_+ \times \mathbb{T}^d. \tag{1.4}$$

While the underlying N-body dynamics are stochastic, we stress that equation (1.4) is completely deterministic, and the noise has been averaged out to become diffusion, which is consistent with the independence of the Brownian motions and the mean-field limit being a law-of-large-numbers-type result. Proving the convergence of the empirical measure is closely related to proving *propagation of molecular chaos*: if $f_N^0(x_1, \ldots, x_N)$ is the initial law of the distribution of the N particles in $\mathbb{R}^d$ and if f_N^0 converges to some factorized law $(\mu^0)^{\otimes N}$, then the k-point marginals $f_{N;k}^t$ converge for all time to $(\mu^t)^{\otimes k}$. It is known that mean-field convergence and propagation of chaos are qualitatively equivalent (e.g., see [64]), though quantitative results for one form of convergence do not a priori carry over to the other.

The topic of mean-field limits for singular interactions has seen tremendous progress in recent years. In particular, we mention the works of Jabin and Wang [68] which allowed so-called $\dot{W}^{-1,\infty}$ interactions to be treated via a relative entropy method; the introduction of the *modulated energy* by Duerinckx [46] to noiseless systems of the form (1.1), following earlier usage in a different context by the third author in [98]; the generalization of the modulated energy method to all super-Coulomb interactions in [99], which allowed cases without noise (conservative, dissipative, or mixed) to be treated; and the *modulated free energy method* of Bresch et al. [22–24], which combines both the relative entropy and modulated energy approaches in a physical way to treat the case (of gradient flows only) with noise. Subsequent work by the last two authors with Nguyen [79] generalized the modulated energy method to sub-Coulomb interactions (cf. [31, 63]), and to singular interactions that are not exactly of Riesz type (e.g., repulsive Lennard-Jones potentials in the case of gradient flows). Further extensions and improvements of the modulated energy method concerning regularity of solutions to (1.4) [88, 89] and incorporation of multiplicative noise [86] have been achieved by the second author. Much progress has also been made by Lacker [69], who introduced a novel usage of the relative entropy in conjunction with the BBGKY hierarchy to obtain the sharp $O(k^2/N^2)$ rate for the asymptotic factorization of the k-point marginals measured by the relative entropy, but only for less singular cases, such as bounded interactions. The recent work of Bresch et al. [21] also introduces a novel usage of the BBGKY hierarchy to prove uniform-in-N weighted L^p estimates for the k-point marginals, which allows second-order systems with degenerate noise and singular interactions (e.g., Coulomb in dimension 2) of Vlasov–Fokker–Planck type to be treated,[2] as well as first-order systems with interactions more singular than in [68]. We emphasize that these last two works strongly rely on the dissipative effect of the noise, i.e., they require $\sigma > 0$.

The aforementioned work [22–24] of Bresch et al. focuses on treating as general as possible *repulsive* singular interactions with a mildly *attractive* part (e.g., logarithmic). In particular, the latter work proves the mean-field limit for the Patlak–Keller–Segel (PKS) equation on $\mathbb{T}^2$, which corresponds to (1.2) with $s = 0$ and g replaced by $-\mathsf{g}$, up to, but not including, the critical temperature.[3] However, when considering only repulsive Riesz interactions of Coulomb or super-Coulomb type, their modulated free energy method, which leverages algebraic cancellations specific to the gradient flow structure, can lead to a much quicker proof of convergence, as outlined in the introduction of [99]. In addition, their work, which was restricted to the torus, left as an assumption the existence of a sufficiently regular limiting solution. The essential content of this restriction to the torus is the need for compactness of the underlying domain in order to prove lower bounds for the density that are uniform in space, which are needed in order to show certain norms involving $\log \mu^t$ are finite. Such pointwise bounds are seemingly incompatible with the setting of $\mathbb{R}^d$, without some form of confinement in the form of an external potential,[4] as μ^t vanishes as $|x| \to \infty$ and the L^∞ norm of μ^t vanishes as $t \to \infty$.

---

[2]The work of Lacker [69] (see Remarks 2.11 and 4.5 in that work) is also capable of proving (sharp) propagation of chaos for second-order kinetic models with degenerate noise, but again only for less singular interactions.

[3]In the literature on PKS dynamics, the critical parameter for the global existence vs. finite-time blowup is typically formulated in terms of a critical mass with fixed unit temperature (i.e., diffusion coefficient). Since the mass in our setting is normalized to 1, this critical mass can be equivalently expressed as a critical temperature.

[4]In unpublished work by the last two authors with Huang [65], we show how to extend the modulated free energy to the case of $\mathbb{R}^d$ when a confining term $-\nabla V_{\text{ext}}$ is added to the dynamics in (1.1) and provided one starts from initial data which are small perturbations of the equilibrium for equation (1.4) with the additional confining term.

Our goal in this paper is to present a streamlined version of mean-field convergence for periodic Riesz interactions (1.2) in the case $d-2 \leq s < d$, along with a complete analysis of the limiting equation (1.4). Specifically, we prove (1.4) is globally well posed (either in the dissipative or conservative case), and solutions and their derivatives satisfy exponentially fast relaxation estimates (see Sections 4 and 5 below). By combining these relaxation estimates with the modulated free energy method and new sharp functional inequalities for the variations of Coulomb/Riesz modulated energies obtained by the last two named authors [93], we manage to show the first instance of *uniform-in-time* convergence for singular dissipative flows, with a rate which is sharp in N.

There have been a number of results obtained by probabilistic arguments over the years on uniform-in-time mean-field convergence/propagation of chaos for McKean–Vlasov-type systems. We mention the sample of works [5, 36, 45, 48, 76, 95], which are related to the long-time dynamics of nonlinear Fokker–Planck equations, e.g., [8–10, 16, 17, 34–36, 76]. Importantly, these results are restricted to regular potentials. We also note that uniform-in-time propagation of chaos may fail for certain potentials [12, 13]. It is only very recently that uniform-in-time results for the much more difficult case of singular potentials have been obtained.

The uniform-in-time convergence (without sharp rate) was previously shown by the last two authors in [91] only for $d \geq 3$ and $0 \leq s < d-2$ on Euclidean space, though the proof is adaptable to the torus. The idea of our proof for the relaxation to equilibrium of the limiting solutions is inspired by the method of [91], which itself builds on earlier ideas of Carlen and Loss [29]. We note that uniform-in-time convergence (without sharp rate) was established by Guillin et al. [55] for $\dot{W}^{-1,\infty}$ kernels through a refinement of the argument in [68], in particular the exploitation of the Fisher information through a uniform-in-time log-Sobolev inequality. We also mention that recent work of Lacker and Le Flem [70] builds on [69] to obtain uniform-in-time propagation of chaos with a sharp rate. The results of [70] have been subsequently extended to slightly more singular interactions – though not covering the case $s = 0$ of (1.2) – in [57], subject to a number of conditions. We mention recent work of Guillin et al. [56], which proves uniform-in-time propagation of chaos for one-dimensional log and Riesz gases, exploiting the convexity of the log/Riesz interaction in dimension one – and only dimension one (cf. [11]). Finally, we mention recent work [92] (subsequent to the first version of this paper) by the last two named authors on the stronger notion of *generation of chaos*, conditional on certain logarithmic Sobolev inequalities, which asserts that systems become chaotic as both $N \to \infty$ and $t \to \infty$ regardless of the initial condition.

Lest the reader think otherwise, uniform-in-time convergence is not merely aesthetically pleasing. It is important for both theory and practice, such as when using a particle system to approximate the limiting equation or its equilibrium states and for quantifying stochastic gradient methods, such as those used in machine learning for general interaction kernels.

1.2. The modulated free energy method

In order to present our results, let us introduce the modulated free energy from [22–24], which is a combination of two quantities: the modulated energy from [46,99] and the relative entropy from [66,68]. See also earlier incarnations – in other contexts – of modulated energy/relative entropy methods in [20, 44, 103].

The modulated energy is a Coulomb/Riesz-based "metric" that can be understood as a renormalization of the negative-order homogeneous Sobolev norm corresponding to the energy space of equation (1.4). More precisely, it is defined to be

$$F_N(\underline{x}_N, \mu) := \frac{1}{2} \int_{(\mathbb{T}^d)^2 \setminus \triangle} \mathsf{g}(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y),$$

where we remove the infinite self-interaction of each particle by excising the diagonal $\triangle := \{(x,x) \in (\mathbb{T}^d)^2\}$. Since we work in the statistical setting of the Liouville equation (1.3), we need to average this quantity with respect to the joint law f_N of the positions $\underline{x}_N$. We then define the (normalized) relative entropy with respect to the N-fold tensor product of a probability density μ (denoted $\mu^{\otimes N}$, which is the distribution of N iid random points in $\mathbb{T}^d$ with law μ) as

$$H_N(f_N|\mu^{\otimes N}) := \frac{1}{N} \int_{(\mathbb{T}^d)^N} \log\left(\frac{f_N}{\mu^{\otimes N}}\right) df_N.$$

With the modulated energy and relative entropy, we now define the *modulated free energy* following [22–24]:

$$E_N(f_N, \mu) := \sigma H_N(f_N|\mu^{\otimes N}) + \int_{(\mathbb{T}^d)^N} F_N(\underline{x}_N, \mu)\, df_N(\underline{x}_N). \tag{1.5}$$

As explained above, the consideration of $\mathbb{T}^d$, as opposed to $\mathbb{R}^d$, stems from the need for a confined domain.

When there is no noise (i.e., $\sigma = 0$), the relative entropy is unnecessary and a pure modulated energy approach suffices, which is consistent with the weighting of the relative entropy by σ in (1.5). This has been shown in [46, 79, 99], the last of which treats the full range $0 \le s < d$ for (1.2). Initially, it was unclear whether a pure modulated energy approach could also handle noise of the form in (1.1).[5] Such an extension was finally shown in [91], but only for $s < d - 2$. This limitation stems from treating the nontrivial quadratic variation contribution to the evolution of the expectation of the modulated energy as a term which is nonpositive up to negligible error. Such nonpositivity is no longer expected to hold if $d - 2 \le s < d$, and we are skeptical a pure modulated energy approach is feasible for $d - 2 \le s < d$. Note that this work also makes sense of and works

[5] A pure modulated energy approach is known to be well suited to a completely different kind of noise, of multiplicative type, thanks to [79, 86].

directly with the SDE (1.1), and not the Liouville equation (1.3), so that the modulated energy is a stochastic process.

The modulated free energy method comes at the cost of only treating the gradient flows case. This method consists of computing the evolution of the quantity $E_N(f_N^t, \mu^t)$, given solutions f_N^t and μ^t of equations (1.3) and (1.4), respectively, and establishing an inequality in caricature of the form

$$\frac{d}{dt} E_N(f_N^t, \mu^t) \leq C(E_N(f_N^t, \mu^t) + N^{-\beta}),$$

where C is some constant depending on norms of μ^t, and $\beta > 0$ is some exponent determined by d, s. One then concludes by the Grönwall–Bellman lemma. Not only is it a physically well-motivated quantity, but mathematically, $E_N(f_N^t, \mu^t)$ is a good quantity for showing propagation of chaos because it metrizes both convergence of the k-point marginals (thanks to the relative entropy) and convergence of the empirical measure (thanks to the modulated energy). See Remark 6.5 below for further elaboration. The beautiful observation of [22–24] is that when computing $\frac{d}{dt} E_N(f_N^t, \mu^t)$, the contribution of the noise to the evolution of the modulated energy cancels exactly with terms coming from the relative entropy, but only for the gradient flow case. Then one is left with having to control the average with respect to the measure $df_N(\underline{x}_N)$ of an expression of the form

$$I := \int_{(\mathbb{T}^d)^2 \setminus \triangle} (v(x) - v(y)) \cdot \nabla \mathsf{g}(x-y) \, d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y), \tag{1.6}$$

where v is a vector field, by $E_N(f_N, \mu)$ and some negligible error. Note that the expression (1.6) arises naturally by pushing forward the measure $\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu$ under the map $\mathbb{I} + tv$ in the modulated energy $F_N(\underline{x}_N, \mu)$ and computing the first derivative at $t = 0$, or in other words, computing the first variation of the modulated energy along the transport v.

Such a functional inequality was previously shown by the third author [99, Proposition 1.1] in the form[6]

$$\begin{aligned} |I| \leq C\|\nabla v\|_{L^\infty}\Big(F_N(\underline{x}_N, \mu) + \frac{(\log N)}{2dN}\mathbf{1}_{s=0} + C(1 + \|\mu\|_{L^\infty})N^{-\frac{d-s}{d(d+1)}}\Big) \\ + (\text{other terms}), \end{aligned} \tag{1.7}$$

for a constant C depending only d, s. The (other terms) are not so important for our discussion, and we choose not to make them explicit. These functional inequalities have proven to be extremely powerful for mean-field limit and related problems, and we mention a sample of recent applications [30, 53, 58, 78, 82, 87, 90, 100]. The original functional

[6] Strictly speaking, this cited work considers $\mathbb{R}^d$, not $\mathbb{T}^d$, but the argument is adaptable to the torus, as for instance shown in [90, Proposition 3.9]. See also Section 6.2 below for explanation.

inequality (1.7) has since been improved in the Coulomb case $s = d - 2$ in [100, Corollary 4.3], [90, Proposition 3.9], where the exponent $-\frac{d-s}{d(d+1)}$ is improved to $-1 + \frac{s}{d}$. This is sharp in the sense that the modulated energy scales as $N \to \infty$ like $N^{-1+\frac{s}{d}}$: the minimal value of the modulated energy among all point configurations $\underline{x}_N$, for a fixed background density μ, scales like $N^{-1+\frac{s}{d}}$. See [62] specifically for the periodic case and [81, 94, 96, 97] for the Euclidean case with a confining potential. Recent work by the last two authors [93] goes further, in particular covering the full range $d - 2 \le s < d$, and replaces the right-hand side in (1.7) by

$$C\|\nabla v\|_{L^\infty}\Big(F_N(\underline{x}_N, \mu) + \frac{\log(\|\mu\|_{L^\infty})}{2dN}\mathbf{1}_{s=0} + \|\mu\|_{L^\infty}^{\frac{s}{d}} N^{-1+\frac{s}{d}}\Big). \tag{1.8}$$

This estimate is sharp in its rate $N^{-1+\frac{s}{d}}$. Moreover, it is improved in its dependence on $\|\mu\|_{L^\infty}$ and, as first observed in [91], this improved dependence can be used in conjunction with decay estimates for $\|\mu^t\|_{L^\infty}$ to obtain uniform-in-time bounds. The question of functional inequalities with sharp dependence on N in the sub-Coulomb range $s < d - 2$ remains an interesting open problem.

Returning to the modulated free energy method, one wants to apply the above described functional inequalities with $v = u^t := \sigma\nabla \log \mu^t + \nabla \mathrm{g} * \mu^t$, where μ^t solves (1.4). A point we stress to the reader is that when there is no noise, there is no $\nabla \log \mu^t$ term, and thus the vector field is not the same in the modulated energy and modulated free energy methods. We have now arrived at a PDE question, which is control on the Lipschitz seminorm of u^t. Assuming that μ^t is bounded from below, this translates to $W^{2,\infty}$ control on μ^t. Another point we stress to the reader is that control of $\|\nabla^{\otimes 2} \log \mu^t\|_{L^\infty}$ is delicate on the Euclidean space due to the decay of μ^t to 0 at infinity. This issue, of course, disappears on the torus (likely more generally a bounded domain with appropriate boundary conditions), since the solution to (1.4) remains bounded from below, provided the initial data is (see Lemma 4.6 below). This is our main reason for considering the periodic setting. It is possible, however, to implement the modulated free energy method on $\mathbb{R}^d$ with a confining potential V_{ext} added to the dynamics (1.1), (1.4) and for solutions of (1.4) which start near equilibrium (which is no longer uniform) [65].

We now come to the main concern of the present article. In light of the work [91] on uniform-in-time convergence for sub-Coulomb Riesz interactions, it is natural to ask whether such a uniform-in-time result is also possible for the modulated free energy, which would then yield uniform-in-time convergence for the full Riesz range $-1 \le s < d$, at least in the periodic setting. Such a result also necessitates having a satisfactory solution theory for the limiting equation (1.4), in particular global solutions in $W^{2,\infty}$. The well-posedness of (1.4), even locally in time, is taken for granted in [23], which sketches the use of the modulated free energy for local-in-time convergence for general Riesz interactions. Additionally, one seeks estimates for the modulated free energy which are sharp in their dependence on N. Such estimates are obtained in the forthcoming work [93] for the Coulomb/super-Coulomb case without noise, but to our knowledge no work to date has achieved the sharp rate of convergence for Riesz interactions with noise.

1.3. Main theorem

We are now ready to state our main results, which establish in complete generality the global existence and asymptotic behavior of solutions to (1.4) in both the conservative and dissipative cases for $d-2 \le s < d$ and show the first uniform-in-time propagation of chaos result for both Coulomb and super-Coulomb gradient flows on the torus. Moreover, the convergence is at the sharp rate $N^{-1+\frac{s}{d}}$. The function space notation in the statements of Theorems 1.1 and 1.2 below is standard, but we recall it anyway for the reader's benefit in Section 1.4.

Theorem 1.1. *Let $d \ge 1$, $d-2 \le s < d$, and $\sigma > 0$. Define the space $X := L^\infty(\mathbb{T}^d) \cap \dot{W}^{\alpha,p}(\mathbb{T}^d)$, where*

$$\begin{cases} \alpha \ge 0,\ 1 \le p \le \infty, & \text{if } d-2 \le s \le d-1 \\ \alpha > \max(d-s+1, d-s+\frac{d}{p}),\ 1 \le p \le \infty, & \text{if } d-1 < s < d. \end{cases}$$

Assume further that the initial datum $\mu^0 \ge 0$ if $\mathbb{M} = -\mathbb{I}$. Then equation (1.4) is globally well posed in the space $C([0,\infty), X)$, smooth on $(0,\infty) \times \mathbb{T}^d$, $\inf_{\mathbb{T}^d} \mu^t \ge \inf_{\mathbb{T}^d} \mu^0$, and for any $n \ge 0$ and $1 \le p \le \infty$ we have, for all $t > 0$,[7]

$$\begin{aligned} \|\nabla^{\otimes n}(\mu^t - 1)\|_{L^p} \le \mathbf{W}(&\|\mu^0\|_{L^\infty}, \|\mu^0 - 1\|_{L^1}, \mathcal{F}_\sigma(\mu^0), \\ &\|\mu^0\|_{\dot{H}^{s+d-1}} \mathbf{1}_{d-1<s<d}, \sigma^{-1})(\sigma t)^{-m} e^{-Ct}, \end{aligned} \tag{1.9}$$

where $m > 0$ depends on n, s, d, p; C depends on $n, s, d, \mathbb{M}$; and $\mathbf{W}\colon [0,\infty)^5 \to [0,\infty)$ is continuous, nondecreasing in its arguments, and depends on parameters similar to C. Here, $\mathcal{F}_\sigma(\mu^0)$ is the free energy associated to equation (1.4) in the case $M = -\mathbb{I}$ (see (4.20) below), and $\mathbf{W}$ is independent of $\mathcal{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric.

Theorem 1.2. *Let $d \ge 1$, $d-2 \le s < d$, and $\sigma > 0$. Let f_N be an entropy solution to (1.3), in the sense of Definition 6.1, and let $\mu^0 \in \mathcal{P}(\mathbb{T}^d) \cap W^{2,\infty}(\mathbb{T}^d)$ with associated solution $\mu \in C([0,\infty), \mathcal{P}(\mathbb{T}^d) \cap W^{2,\infty}(\mathbb{T}^d))$. Assume further that $\inf_{\mathbb{T}^d} \mu^0 > 0$. Suppose now that $\mathbb{M} = -\mathbb{I}$, so that we consider gradient flows. Define the quantity*

$$\mathcal{E}_N^t := E_N(f_N^t, \mu^t) + \frac{\log(N\|\mu^t\|_{L^\infty})}{2Nd} \mathbf{1}_{s=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1},$$

where $\mathsf{C} > 0$ is a certain constant to ensure that $\mathcal{E}_N^t \ge 0$ (see (6.10) below). There exists a function $\mathcal{A}\colon [0,\infty) \to [0,\infty)$, depending on s, d, σ, $\inf_{\mathbb{T}^d} \mu^0$, $\|\mu^0\|_{W^{2,\infty}}$, $\|\mu^0\|_{\dot{H}^{1+s-d}}$, $\|\mu^0 - 1\|_{L^1}$, $\mathcal{F}_\sigma(\mu^0)$, such that $\mathcal{A}^0 = 1$, $\sup_{t \ge 0} \mathcal{A}^t < \infty$, and

$$\forall t \ge 0, \quad \mathcal{E}_N(f_N^t, \mu^t) \le \mathcal{A}^t \mathcal{E}_N(f_N^0, \mu^0). \tag{1.10}$$

[7] The notation $\mathbf{1}_{d-1<s<d}$ denotes the indicator function for the condition $d-1 < s < d$.

Using ideas inspired by the proof of uniform-in-time propagation chaos in Theorem 1.2, we are also able to give a proof of uniform-in-time propagation of chaos for systems like (1.1) but with $\mathbb{M}\nabla \mathsf{g}$ replaced by a kernel k which belongs to the space $\dot{W}^{-1,\infty}$ (i.e., it is the divergence of an L^∞ matrix field). A precise statement of the result is given in Section 2.4 with Theorem 2.14. This improves the result of Jabin–Wang [68], which had a growing factor e^{Ct} in their relative entropy estimate. Our result should be understood as a refinement of their original proof, as the main novelty is the incorporation of decay estimates for the derivatives of μ^t to obtain a uniform-in-time result. As mentioned in Section 1.1, Guillin et al. previously obtained a uniform-in-time version of the Jabin–Wang result, also through a refinement of the original proof of [68], but not relying on decay estimates. See Section 2.4 for comparison between the two proofs.

The modulated free energy method was originally developed [22, 24] to treat propagation of chaos for the gradient dynamics of the d-dimensional *attractive log gas*, which coincides with the aforementioned Patlak–Keller–Segel model if $d = 2$. The cited works show a nonsharp rate for propagation of chaos, which deteriorates exponentially fast in time, leaving as a question whether a uniform-in-time rate is possible. In forthcoming work [39], we answer this question for sufficiently high temperatures using the modulated free energy and relaxation estimates for the limiting equation. The attractive case is substantially more difficult than the repulsive case considered here due to the existence of phase transitions: at a certain critical temperature, the long-time dynamics of the system completely change and one encounters issues of nonuniqueness and instability of stationary states. In fact, we show that a uniform-in-time estimate for the modulated free energy may fail if the temperature σ is too low.

We close this subsection with the following remarks concerning Theorems 1.1 and 1.2. We defer a discussion of the proofs of these results until Section 2.

Remark 1.3. The relaxation/decay estimate (1.9) is not the most general possible statement. One also has estimates which hold for fractional derivatives $|\nabla|^\alpha$, $\alpha > 0$. We refer the reader to Section 5 for further details. The fact that we have an exponential decay as $t \to \infty$, as opposed to an algebraic decay, as for instance in [91], is a special feature of the confined setting of the torus vs. Euclidean space. The reader may easily convince themselves of this by ignoring the nonlinearity and considering the asymptotic behavior of solutions μ^t to the linear heat equation, for which $\|\mu^t\|_{L^\infty}$ decays at the optimal rate $O(t^{-\frac{d}{2}})$ on $\mathbb{R}^d$.

Remark 1.4. Concerning regularity assumptions, Bresch et al. [23] assume – but do not prove – the existence of a local solution $\mu \in C([0,T], W^{2,\infty}(\mathbb{T}^d))$, for some $T > 0$, to equation (1.4), which remains bounded from below on $[0,T]$. For such a solution and for $t \in [0,T]$, they can prove an estimate with the same structure as (1.10), but with nonsharp exponents.

Remark 1.5. An explicit form of the right-hand side in the bound (1.10) is given in Section 6.3 (see inequality (6.22)). We have not presented the explicit form above, so as to keep the introduction accessible.

Remark 1.6. As is by now well known, the vanishing of the relative entropy or modulated energy, and therefore the modulated free energy, as $N \to \infty$, implies propagation of chaos. We refer to Remark 6.5 below for further explanation.

Remark 1.7. To the best of our knowledge, equation (1.4) has not been studied in the complete generality presented here. Some special cases are treated (on $\mathbb{R}^d$) in [14, 32, 37, 40, 42, 43, 50]. However, the decay estimates seem generally to be new. If $d = 2$ and $s = 0$ and $\mathbb{M}$ is a 90° rotation, then this is the well-known Navier–Stokes in vorticity form (e.g., see [50]). Staying in dimension two, but letting $0 < s < 2$, this is the generalized SQG equation with subcritical dissipation (e.g., see [43]). If $d = 2$ and $s = 0$, but now $\mathbb{M} = -\mathbb{I}$, then this is a repulsive analogue of the famous Patlak–Keller–Segel equation (e.g., see [15]). Usually, these equations are studied on $\mathbb{R}^d$, but many of the results are expected to carry over to $\mathbb{T}^d$ mutatis mutandis. We also mention that the case without temperature (and generally on $\mathbb{R}^d$, not $\mathbb{T}^d$) has been studied in several works, e.g., [3, 14, 26–28, 32, 33, 37, 40, 41, 43, 47, 74, 75, 77, 101, 104, 106].

Remark 1.8. The results of Theorems 1.1 and 1.2 are valid for any $\sigma > 0$ (i.e., positive temperature), but essentially all our estimates blow up as $\sigma \to 0$. Naturally, one asks if it is possible to have a uniform-in-time mean-field convergence/propagation of chaos result when $\sigma = 0$ (i.e., zero temperature). Interestingly, the answer is yes. For instance, for two-dimensional Coulomb gradient dynamics, we can show that any L^∞ solution of (1.4) which is bounded from below converges exponentially fast to the uniform distribution as $t \to \infty$, and this relaxation can be combined with the refinement of the modulated energy developed by the second author in [88] to obtain uniform-in-time mean-field convergence. These findings and others will be reported elsewhere.

1.4. Notation

We close the introduction with the basic notation used throughout the article without further comment. We mostly follow the conventions of [79, 91].

Given nonnegative quantities A and B, we write $A \lesssim B$ if there exists a constant $C > 0$, independent of A and B, such that $A \le CB$. The dependence of the implicit constant on a parameter p is denoted by $\lesssim_p$. If $A \lesssim B$ and $B \lesssim A$, we write $A \sim B$. Throughout this paper, C will be used to denote a generic constant which may change from line to line. Also, $(\cdot)_+$ denotes the positive part of a number.

The natural numbers are denoted by $\mathbb{N}$ excluding zero, and $\mathbb{N}_0$ including zero. Similarly, $\mathbb{R}_+$ denotes the positive reals. Given $N \in \mathbb{N}$ and points $x_{1,N}, \dots, x_{N,N}$ in some set X, $\underline{x}_N = (x_{1,N}, \dots, x_{N,N}) \in X^N$. Given $x \in \mathbb{R}^d$ and $r > 0$, $B(x,r)$ and $\partial B(x,r)$ respectively denote the ball and sphere centered at x of radius r. Given a function f, we denote the support of f by supp f. The notation $\nabla^{\otimes k} f$ denotes the k-tensor field with components $(\partial^k_{i_1\cdots i_k} f)_{1\le i_1,\ldots,i_k\le d}$.

The space of Borel probability measures on $\mathbb{T}^d$ is denoted by $\mathcal{P}(\mathbb{T}^d)$. If μ is absolutely continuous with respect to Lebesgue measure, we will abuse notation by writing μ for both the measure and its density function. The Banach space of continuous, bounded functions on $\mathbb{R}^d$ is denoted by $C(\mathbb{T}^d)$, and is equipped with the uniform norm $\|\cdot\|_\infty$. The Banach space of k-times continuously differentiable functions with bounded derivatives up to order k is denoted by $C^k(\mathbb{T}^d)$ and is equipped with the natural norm; also, $C^\infty := \bigcap_{k=1}^\infty C^k$. The subspace of smooth functions with compact support is denoted with a subscript c.

The Fourier multiplier, with symbol $2\pi|\xi|$, is denoted by $|\nabla| = (-\Delta)^{-\frac{1}{2}}$. Functions of $|\nabla|$ can be defined through the spectral calculus (i.e., by using the Fourier transform). For integers $n \in \mathbb{N}_0$ and exponents $1 \le p \le \infty$, $W^{n,p}$ denotes the usual Sobolev space. For general $\alpha \in \mathbb{R}$ and $1 < p < \infty$, $W^{\alpha,p}$ denotes the Bessel potential space defined by

$$\left\{\mu \in \mathcal{D}'(\mathbb{T}^d) \colon \|(I-\Delta)^{\alpha/2}\mu\|_{L^p} < \infty\right\},$$

in other words, the space of distributions μ such that $(I-\Delta)^{\alpha/2}\mu$ is an L^p function. When α is a positive integer, then $W^{\alpha,p}$ coincides with the classical Sobolev space above. For $p \in \{1,\infty\}$, these fractional Sobolev spaces are awkward to consider and will be generally avoided in this paper. When $p = 2$, we instead use the customary notation H^α. As is convention in the literature, a $\dot{}$ superscript indicates the corresponding homogeneous space.

2. Roadmap for the paper

We give here a roadmap for the paper, in particular the various results contained in it and their relations to one another. We also take this opportunity to comment on the general strategy behind the proofs of Theorems 1.1 and 1.2.

In Section 3 we review some basic facts about periodizations of Riesz potentials and estimates for the heat kernel. This section may be skipped upon first reading and consulted as necessary.

2.1. Well-posedness and L^p control

In Section 4, we take up the first part of Theorem 1.1 by showing the global well-posedness of the limiting equation (1.4), its basic properties, and the relaxation to the uniform distribution in L^p norm. Section 4.1 considers the well-posedness. The local well-posedness (Proposition 2.1), which allows for $\mathbb{M}$ to be either conservative or dissipative, proceeds through a fixed point argument for the mild formulation of (1.4). This technique is classical, but some care is needed in the case $d-1 < s < d$, as the vector field $\mathbb{M}\nabla \mathsf{g} * \mu$ loses derivatives compared to μ.

Proposition 2.1. *Let $d \geq 1$, $s < d$, and $\sigma > 0$.*

- $(s \leq d-1)$ *If $\mu^0 \in L^\infty(\mathbb{T}^d)$, then there exists a time*

$$T \geq \Big(\frac{\sigma^{1/2}}{C\|\mu^0\|_{L^\infty}}\Big)^2 \mathbf{1}_{s<d-1} + \Big(\frac{\sigma^{\frac{d}{2p}+\frac{1}{2}}}{C_p\|\mu^0\|_{L^\infty}}\Big)^{\frac{2p}{p-d}} \mathbf{1}_{s=d-1},$$

where $d < p < \infty$ is arbitrary, such that equation (4.2) has a unique solution $\mu \in C([0,T], L^\infty)$. Moreover, if μ_1, μ_2 are two solutions to (4.2) on $[0,T]$, then

$$\|\mu_1 - \mu_2\|_{C([0,T],L^\infty)} \leq 2\|\mu_1^0 - \mu_2^0\|_{L^\infty}. \tag{2.1}$$

- $(s > d-1)$ *Let $1 \leq p < \infty$ and $\alpha \geq s+1-d$ satisfy $p > d$ or $\alpha > s-d+\frac{d}{p}$. If $\mu^0 \in L^\infty(\mathbb{T}^d) \cap \dot{W}^{\alpha,p}(\mathbb{T}^d)$,[8] then for arbitrary $\delta \in (s+d-1, 1)$, there exists a time*

$$T \geq \Big(\frac{\sigma^{\frac{1+\delta}{2}}}{C_\delta\|\mu^0\|_{L^\infty\cap\dot{W}^{\alpha,p}}}\Big)^{\frac{2}{1-\delta}},$$

such that equation (4.2) has a unique solution $\mu \in C([0,T], L^\infty \cap \dot{W}^{\alpha,p})$. Moreover, if μ_1, μ_2 are two solutions to (4.2) on $[0,T]$, then

$$\|\mu_1 - \mu_2\|_{C([0,T],L^\infty\cap\dot{W}^{\alpha,p})} \leq 2\|\mu_1^0 - \mu_2^0\|_{L^\infty\cap\dot{W}^{\alpha,p}}. \tag{2.2}$$

The constant C above depends only on d, s, $\mathbb{M}$ if $d-2 \leq s \leq d-1$ and C_p additionally on p if $s = d-1$; C_δ depends additionally on α, δ, p if $d-1 < s < d$.

- (Blowup) *Let $X = L^\infty \cap \dot{W}^{\alpha,p}$, where α, p are as above if $d-1 < s < d$, equipped with its natural norm. Let $\mu \in C([0,T_{\max}), X)$ be the maximal lifespan solution obtained by iterating the local existence argument. If $T_{\max} < \infty$, then*

$$\limsup_{T\to T_{\max}^-} \|\mu\|_{C([0,T],X)} = \infty.$$

After proving Proposition 2.1, we establish important properties of solutions in Lemma 4.6, such as conservation of mass and the minimum/maximum principle, that will be useful in the sequel. The subsection concludes with Proposition 2.2, showing solutions are global. Here we crucially use the repulsive assumption for the dissipative case. The case $s \leq d-1$ is an immediate consequence of the maximum principle, while the case $s > d-1$ follows from a nonlinear Grönwall inequality. In the latter case, the controlling norm, which depends on a fractional derivative of the initial data, a priori may grow in time, but this will be ruled out later by Proposition 2.6 discussed below.

[8] We exclude the case $p = \infty$ because expressions of the form $\| |\nabla|^\alpha \mu\|_{L^\infty}$ are awkward from the point of view of harmonic analysis. If $\alpha = n$ is a positive integer, then there is no issue in adapting our proof to the usual Sobolev spaces $\dot{W}^{n,\infty}$.

Proposition 2.2 (Global well-posedness). *Under the same assumptions as in the statement of Proposition 2.1, there exists a unique global solution μ to (4.2) in*

$$\begin{cases} C([0,\infty), L^\infty(\mathbb{T}^d)), & d-2 \le s \le d-1, \\ C([0,\infty), L^\infty \cap \dot{W}^{\alpha,p}(\mathbb{T}^d)), & d-1 < s < d. \end{cases}$$

Next, Section 4.2 shows (Lemma 4.9) that all L^p norms of solutions are nonincreasing and, in fact, are decreasing exponentially in time for conservative dynamics when restricted to zero-mass solutions. The latter establishes a cheap form of convergence to the uniform distribution for the conservative case. Section 4.3 reviews the gradient flow structure for the dissipative case and uses the dissipation of free energy,

$$\mathcal{F}_\sigma(\mu) := \sigma \int_{\mathbb{T}^d} \log(\mu)\, d\mu + \frac{1}{2} \int_{(\mathbb{T}^d)^2} \mathrm{g}(x-y)\, d\mu^{\otimes 2}(x,y),$$

with the logarithmic Sobolev inequality for the uniform measure (Lemma 4.11) to obtain exponential-in-time decay of the free energy. Interpolating with the L^p control provided by Lemma 4.9 yields exponential-in-time decay of all norms $\|\mu^t - 1\|_{L^p}$ for finite p.

Lemma 2.3. *Let μ be a probability density solution of equation (1.4). If $\mathbb{M} = -\mathbb{I}$, then*

$$\forall t \ge 0, \quad \mathcal{F}_\sigma(\mu^t) \le \mathcal{F}_\sigma(\mu^0) e^{-8\pi^2 \sigma t}, \tag{2.3}$$

and for $1 \le p < r \le \infty$,

$$\forall t \ge 0, \quad \|\mu^t - 1\|_{L^p} \le (1 + \|\mu^0\|_{L^\infty})^{1 - \frac{\frac{1}{p} - \frac{1}{r}}{1 - \frac{1}{r}}} \left(e^{-4\pi^2 \sigma t} \sqrt{2\mathcal{F}_\sigma(\mu^0)/\sigma}\right)^{\frac{\frac{1}{p} - \frac{1}{r}}{1 - \frac{1}{r}}}. \tag{2.4}$$

Similarly, if $\mathbb{M}$ is antisymmetric, then

$$\forall t \ge 0, \quad \mathrm{Ent}(\mu^t) \le e^{-8\pi^2 \sigma t}\, \mathrm{Ent}(\mu^0),$$

and for $1 \le p < r \le \infty$,

$$\forall t \ge 0, \quad \|\mu^t - 1\|_{L^p} \le (1 + \|\mu^0\|_{L^r})^{1 - \frac{\frac{1}{p} - \frac{1}{r}}{1 - \frac{1}{r}}} \left(e^{-4\pi^2 \sigma t} \sqrt{2\,\mathrm{Ent}(\mu^0)}\right)^{\frac{\frac{1}{p} - \frac{1}{r}}{1 - \frac{1}{r}}}.$$

Lastly, Section 4.4 establishes a smoothing property for solutions (hypercontractivity), asserting a higher L^q norm $\|\mu^t\|_{L^q}$ is controlled by a lower L^p norm $\|\mu^t\|_{L^p}$ at the cost of a factor of $e^{C\sigma t}/(\sigma t)^{n_{p,q}}$, for some positive integer $n_{p,q}$. In the conservative case, one may replace μ^t by $\mu^t - 1$ and the result still holds. The proof adapts an argument from [91] – which in turn was an extension of a work by Carlen and Loss [29] – to the periodic setting. This hypercontractive estimate is a priori only useful for short times, due to the exponential factor. But by combining Lemmas 4.13 and 4.9 with a time translation trick, we improve the factor to $\min(\sigma t, 1)^{-n_{p,q}}$. In other words, the smoothing costs nothing for large times. This result also gives exponential-in-time decay of $\|\mu^t - 1\|_{L^p}$, for any $1 \le p \le \infty$, in the conservative case.

Corollary 2.4. *Let μ be a solution to equation (1.4). Then, for any $1 \le p \le q \le \infty$, there exists a constant $C_1 > 0$ depending on $d,\ p,\ q$ such that*

$$\forall t > 0, \quad \|\mu^t\|_{L^q} \le C_1(\min(\sigma t, 1))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^0\|_{L^p}. \tag{2.5}$$

Additionally, if $\mathbb{M}$ is antisymmetric, and $1 \le p < \infty$, then there is a constant C_2 depending on $d,\ p$ such that

$$\forall t > 0, \quad \|\mu^t - 1\|_{L^q} \le C_1(\min(\sigma t, 1))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{-C_2\sigma t}\|\mu^0 - 1\|_{L^p}. \tag{2.6}$$

Corollary 2.4 does not yield a rate of decay for $\|\mu^t - 1\|_{L^p}$ in the gradient flow case $\mathbb{M} = -\mathbb{I}$ and Lemma 2.3 by itself does not give a rate of decay when $p = \infty$. However, by combining Lemma 2.3 with Lemma 4.13, we can obtain such a rate of decay, but only under the restriction $d - 2 \le s \le d - 1$. The reason for this restriction is that $\mathbb{M}\nabla \mathsf{g} * \mu^t$ loses derivatives compared to μ^t if $d - 1 < s < d$.

Corollary 2.5. *Suppose that $d \ge 2$, $d - 2 \le s \le d - 1$ and that $\mathbb{M} = -\mathbb{I}$. Then there exist constants $C_1, C_2 > 0$ depending only on the dimension d, such that*

$$\forall t > 0, \quad \|\mu^t - 1\|_{L^\infty} \le C_1\sigma^{-d-\frac{3}{2}}(\sigma t)^{-\frac{(d^2+d-1)}{2}}e^{-C_2\sigma t}\sqrt{\mathcal{F}_\sigma(\mu^0)}.$$

2.2. Derivative decay

In Section 5 we take up the second part of Theorem 1.1 by showing the decay estimates for derivatives of μ^t of arbitrary degree, which contribute a major portion of the technical effort in this paper.

Proposition 2.6. *Let μ^t be a solution to (1.4) with $\int_{\mathbb{T}^d}\mu^0 = 1$. If $\mathbb{M} = -\mathbb{I}$, further assume that $\mu^0 \ge 0$. Let $\alpha > 0$, $\varepsilon > 0$, and $n \in \mathbb{N}$.*

- *When* $\max(0, d-2) \le s \le d - 1$, *there exist constants $C, C_\varepsilon > 0$, for any $\varepsilon > 0$, and functions $\mathbf{W}_{n,q}, \mathbf{W}_{\alpha,q}\colon [0,\infty)^4 \to [0,\infty)$, continuous, nondecreasing, and polynomial in their arguments, such that for every $t > 0$,*[9]

$$\begin{aligned}\| |\nabla|^\alpha \mu^t\|_{L^q} \le \mathbf{W}_{\alpha,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^q}, \mathcal{F}_\sigma(\mu^0)) \\ \times (\sigma t)^{-\frac{\alpha}{2}}(1 + C_\varepsilon(\sigma t)^{-\varepsilon}\mathbf{1}_{s=d-1\wedge q=1})e^{-C\sigma t}\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}\|\nabla^{\otimes n}\mu^t\|_{L^q} \le \mathbf{W}_{n,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^q}, \mathcal{F}_\sigma(\mu^0)) \\ \times (\sigma t)^{-\frac{n}{2}}\|\mu^0\|_{L^q}(1 + C_\varepsilon(\sigma t)^{-\varepsilon}\mathbf{1}_{s=d-1\wedge q=1})e^{-C\sigma t}.\end{aligned} \tag{2.8}$$

[9] The notation $\mathbf{1}_{a\wedge b}$ denotes the indicator function which is 1 if both conditions a and b hold, and 0 otherwise.

- *When $d-1<s<d$, there exist a constant $C>0$ and functions $\mathbf{W}_{\alpha,q}, \mathbf{W}_{n,q} \colon [0,\infty)^5 \to [0,\infty)$, which are continuous, nondecreasing, and polynomial in their arguments, such that for every $t>0$,*

$$\begin{aligned} \| |\nabla|^\alpha \mu^t \|_{L^q} \leq \mathbf{W}_{\alpha,q}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0)) \\ \times (\sigma t)^{-\frac{\alpha}{2}}(1+(\sigma t)^{-\varepsilon}\mathbf{1}_{q=\infty})e^{-C\sigma t} \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} \|\nabla^{\otimes n} \mu^t \|_{L^q} \leq \mathbf{W}_{n,q}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0)) \\ \times (\sigma t)^{-\frac{n}{2}}(1+(\sigma t)^{-\varepsilon}\mathbf{1}_{q=\infty})e^{-C\sigma t}, \end{aligned} \tag{2.10}$$

where $\lambda_2 := 1+s-d$.

We divide into two cases based on the value of s: $d-2 \leq s \leq d-1$ in Section 5.1 and $d-1<s<d$ in Section 5.2. Beginning with the former, the proof proceeds through two steps. First, we show that $\|\nabla^{\otimes n}\mu^t\|_{L^p}$, for any $1 \leq p \leq \infty$ and $n \geq 1$, is bounded by a function of L^q norms of $\mu^0 - 1$ (and in the gradient case, also the free energy of μ^0) and some polynomial of $(\sigma t)^{-1}$ for $\sigma t \leq 1$, or in other words, a short-time smoothing effect. The proof of this lemma is an induction argument, through the mild formulation of the equation, asserting that control of $\|\nabla^{\otimes m}\mu^t\|_{L^p}$ for orders $m \leq n$, as a function of L^q norms of μ^0-1 (and the free energy of μ^0) and $(\sigma t)^{-1}$, for σt small, implies short-time control of $\|\nabla^{\otimes m}\mu^t\|_{L^p}$ as a similar function of the initial data and $(\sigma t)^{-1}$ for $m \leq n+1$.

Lemma 2.7. *Let $d \geq 2$ and $d-2 \leq s \leq d-1$. For each $n \in \mathbb{N}$ and $1 \leq p \leq q \leq \infty$, there exists a function $\mathbf{W}_{n,p,q} \colon [0,\infty)^4 \to [0,\infty)$, continuous, nondecreasing in its arguments, and vanishing if any of its arguments is zero such that the following holds. If μ is a smooth solution to equation (1.4) on $[0,T]$, then for all $t \in (0,\sigma^{-1}]$,*

$$\begin{aligned} \|\nabla^{\otimes n}\mu^t\|_{L^q} \leq \mathbf{W}_{n,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0-1\|_{L^p}, \mathcal{F}_\sigma(\mu^0)) \\ \times (1+C_\varepsilon(\sigma t)^{-\varepsilon}\mathbf{1}_{s=d-1 \wedge q=1}) \\ \times \Big(\|\mu^0-1\|_{L^p} + \Big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}}(\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}}\mathbf{1}_{p<\infty} \\ + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon}(\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}}\mathbf{1}_{p=\infty}\Big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\ \times (\sigma t)^{-\frac{n}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}, \end{aligned}$$

where $C>0$ depends on n, d, p, q, $C_\epsilon>0$ depends on $\epsilon \in (0,d^{-1})$, $C_\varepsilon>0$ depends on $\varepsilon>0$, which can be made arbitrarily small. The function $\mathbf{W}_{n,p,q}$ additionally depends on d, s, $\mathbb{M}$. Moreover, it is independent of $\mathcal{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric and is independent of $\|\mu^0-1\|_{L^p}$ if $\mathbb{M}=-\mathbb{I}$.

Next we combine Lemma 2.7 with Lemmas 4.9 and 2.3 and a time translation trick to obtain exponential decay of $\|\nabla^{\otimes m}\mu^t\|_{L^p}$ for any order m and $1 \leq p \leq \infty$.

Lemma 2.8. *Let $d \geq 2$ and $d-2 \leq s \leq d-1$. For each $n \in \mathbb{N}$ and $1 \leq p \leq q \leq \infty$, there exists a function $\mathbf{W}_{n,p,q}\colon [0,\infty)^4 \to [0,\infty)$, continuous, nondecreasing, and polynomial in its arguments, such that the following holds. If μ is a solution to equation (1.4), then*

$$\begin{aligned}\forall t > 0, \quad \|\nabla^{\otimes n}\mu^t\|_{L^q} \leq{}& \mathbf{W}_{n,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathscr{F}_\sigma(\mu^0))e^{-C\sigma t} \\ &\times \min(\sigma t, 1)^{-\frac{n}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}(1 + C_\varepsilon \min(\sigma t, 1)^{-\varepsilon}\mathbf{1}_{s=d-1\wedge q=1}),\end{aligned}$$

where $C > 0$ depends on n, d, p, q and $C_\varepsilon > 0$ depends on $\varepsilon > 0$, which can be made arbitrarily small. The function $\mathbf{W}_{n,p,q}$ additionally depends on d, s, $\mathbb{M}$. Moreover, it is independent of $\mathscr{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric.

A scheme similar to that described above is followed in Section 5.2 for the case $d-1 < s < d$, but due to the loss of regularity in the velocity field, we need additional bounds on the Sobolev norms of the solution (see Lemma 2.9), which are established through an energy-method-type argument, in order to implement the induction.

Lemma 2.9. *Let $d \geq 1$ and $d-1 < s < d$. Let μ be a solution to (1.4) with unit mass and such that $\mu^0 \geq 0$ if $\mathbb{M} = -\mathbb{I}$. For $\alpha > 0$, there exists a $C > 0$ depending only on d, s, α, σ, $\mathbb{M}$, such that, for all $t > 0$,*

$$\|\mu^t\|^2_{\dot{H}^\alpha} \leq \|\mu^0\|^2_{\dot{H}^\alpha} + \sigma^{-1}\widetilde{\mathbf{W}}_\alpha(\|\mu^0\|_{L^\infty}, \|\mu^0 - 1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}, \mathscr{F}_\sigma(\mu^0)/\sigma), \tag{2.11}$$

where $\widetilde{\mathbf{W}}_\alpha\colon [0,\infty)^3 \to [0,\infty)$ is a continuous, nondecreasing, polynomial function of its arguments, which does not depend on its third argument if $\mathbb{M}$ is antisymmetric. Additionally, there exists a $T_* > 0$ such that $\|\mu^t\|^2_{\dot{H}^\alpha}$ is decreasing on $[T_*, \infty)$.

Lemma 2.10. *Let $d \geq 1$ and $d-1 < s < d$. For every $\alpha > 0$, and $1 \leq q \leq \infty$, there exists a function $\mathbf{W}_{\alpha,q}\colon [0,\infty)^5 \to [0,\infty)$, which is continuous, nondecreasing, polynomial in its arguments, such that for any solution μ to (1.4), it holds that, for all $t \in (0, \sigma^{-1}]$,*

$$\begin{aligned}\||\nabla|^\alpha\mu^t\|_{L^q} \leq{}& (\sigma t)^{-\frac{\alpha}{2}}(1 + C_\varepsilon(\sigma t)^{-\varepsilon}\mathbf{1}_{q=\infty}) \\ &\times \mathbf{W}_{\alpha,q}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \mathscr{F}_\sigma(\mu^0)),\end{aligned} \tag{2.12}$$

where $C > 0$ depends on d, α, s, q, $\mathbb{M}$, $\varepsilon > 0$ is arbitrary, and $C_\varepsilon > 0$ depends only on d, s, ε. The function $\mathbf{W}_{\alpha,q}$ additionally depends on d, s, $\mathbb{M}$ and is independent of its fifth argument if $\mathbb{M}$ is antisymmetric.

Lemma 2.11. *Let $d \geq 1$ and $d-1 < s < d$. For every $\alpha > 0$, and $1 \leq q \leq \infty$, there exists a function $\mathbf{W}_{\alpha,q}\colon [0,\infty)^5 \to [0,\infty)$, which is continuous, nondecreasing, polynomial in its arguments, such that for a solution μ to (1.4), it holds that, for all $t > 0$,*

$$\begin{aligned}\||\nabla|^\alpha\mu^t\|_{L^q} \leq{}& \min(\sigma t, 1)^{-\frac{\alpha}{2}}(1 + C_\varepsilon \min(\sigma t, 1)^{-\varepsilon}\mathbf{1}_{q=\infty})e^{-C\sigma t} \\ &\times \mathbf{W}_{\alpha,q}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \mathscr{F}_\sigma(\mu^0)),\end{aligned} \tag{2.13}$$

where $C > 0$ *depends on* d, α, s, q, $\mathbb{M}$, $\varepsilon > 0$ *is arbitrary, and* $C_\varepsilon > 0$ *depends only on* d, s, ε. *The function* $\mathbf{W}_{\alpha,q}$ *additionally depends on* d, s, $\mathbb{M}$ *and is independent of its fifth argument if* $\mathbb{M}$ *is antisymmetric.*

Let us mention that ideas related to our method of proof in this section have been used, for instance, for two-dimensional Navier–Stokes (see [51, Section 2.4]) and perhaps in other contexts as well; but to our knowledge there has not been a treatment at the level of generality and for such singular vector fields as in our equation (1.4).

2.3. Modulated free energy

In Section 6 we combine our decay estimates with the modulated free energy to prove uniform-in-time propagation of chaos for system (1.1). This then completes the proof of Theorem 1.2.

Section 6.1 reviews the notion of entropy solutions to the forward Kolmogorov equation. The existence of entropy solutions is presented in Appendix A. Entropy solutions are a suitable notion of a weak solution that allows one to establish the key dissipation inequality behind the modulated free energy method, as stated in Proposition 2.12 below. We refer to [24, Proposition 2.3] for a proof. The very interesting phenomenon is that, compared to a pure modulated energy approach such as in [91], the modulated free energy yields crucial cancellations in the dissipative case at positive temperature.

Proposition 2.12. *Assume that* f_N *is an entropy solution to the Liouville equation* (1.3) *and that* $\mu \in C([0,\infty), W^{2,\infty}(\mathbb{T}^d))$ *solves equation* (1.4). *Then the modulated free energy defined by* (1.5) *satisfies that*

$$\begin{aligned} E_N(f_N^t, \mu^t) \le E_N(f_N^0, \mu^0) \\ - \frac{1}{2}\int_0^t \int_{(\mathbb{T}^d)^N} \int_{(\mathbb{T}^d)^2 \setminus \triangle} (u^\tau(x) - u^\tau(y)) \\ \cdot \nabla \mathrm{g}(x-y)\, d(\mu_N^\tau - \mu^\tau)^{\otimes 2}(x,y)\, df_N^\tau, \end{aligned} \tag{2.14}$$

where $u^t := \sigma \nabla \log \mu^t + \nabla \mathrm{g} * \mu^t$.

We now want to control the right-hand side of (2.14) by the modulated free energy itself and conclude by application of the Grönwall–Bellman lemma. The control of the right-hand side is done in several ways in the literature. In [79], the authors use commutator estimates together with a renormalization procedure implemented through a smearing of the Dirac masses, which allows for Riesz-like potentials that are not exactly our potential g. In particular, this method works for full range $0 \le s < d$. Instead of smearing of the Dirac masses, Bresch et al. [23] employed a regularization of the kernel $\nabla \mathrm{g}$ and a functional inequality (of the same kind as in [99]) for this regularized kernel, which may be understood as a commutator estimate, though this connection is not made in [23].

When considering the exact Riesz potential in the Coulomb/super-Coulomb case $d-2 \le s < d$, one can prove functional inequalities (1.7) using integration by parts. More precisely, one uses the Caffarelli–Silvestre extension procedure to replace g by the kernel of a local operator[10] in the extended space $\mathbb{R}^{d+k}$ and then exploits a *stress–energy tensor* structure and integration parts. Combining this with the smearing procedure mentioned in the preceding paragraph, this method allows for *sharp* estimates. This is an advantage over the approaches of [23, 79], in particular the latter work which encounters an inefficiency in the kernel regularization. Of course, the cost to the stress-tensor approach is the rigidity of the interaction – it must be exactly Riesz.

In the forthcoming article [93], a proof of the sharp version of (1.7) with right-hand side (1.8) is given in Euclidean space using the above-described stress-tensor approach. As there is a version of the Caffarelli–Silvestre extension procedure for the torus (reviewed in Section 6.2), the proof can be straightforwardly adapted to the setting of this paper to give Proposition 2.13 stated below. We sketch the proof of this proposition in Section 6.2. We also review the truncation/smearing procedure in the periodic setting used to express the modulated energy as a renormalized energy, as this is used in the proof.

Proposition 2.13. *Assume $\mu \in L^\infty(\mathbb{T}^d)$ is such that $\int_{\mathbb{T}^d} \mu = 1$. For any pairwise distinct configuration $\underline{x}_N \in (\mathbb{T}^d)^N$ and any Lipschitz map $v\colon \mathbb{T}^d \to \mathbb{R}^d$, we have*

$$\left| \int_{(\mathbb{T}^d)^2 \setminus \triangle} (v(x) - v(y)) \cdot \nabla \mathrm{g}(x-y)\, d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\Big)^{\otimes 2}(x,y) \right|$$
$$\leq C\|\nabla v\|_{L^\infty}\Big(F_N(\underline{x}_N,\mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2dN}\mathbf{1}_{s=0} + C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{-1+\frac{s}{d}}\Big), \quad (2.15)$$

where C depends only on s, d.

Finally, in Section 6.3 we implement the Grönwall argument underlying the proof of Theorem 1.2. The idea is to combine the use of functional inequalities in the modulated free energy method, as described in Section 1.2, with the relaxation estimates of Section 5. Such a combination was first observed in [91] (for just the modulated energy). An important difference, though, with the cited work is that the density no longer decays to zero as $t \to \infty$, but rather to 1. Only derivatives of the density decay. There is an additional ingredient concerning the sharpness of the functional inequalities. As advertised in the title of our work, the factor $N^{-1+\frac{s}{d}}$ is of the same order as the modulated energy as $N \to \infty$ and is therefore sharp. To obtain the exponent $-1+\frac{s}{d}$, we crucially rely on the aforementioned new functional inequalities.

More precisely, we consider the quantity

$$\mathcal{E}_N^t := E_N(f_N^t, \mu^t) + \frac{\log(N\|\mu^t\|_{L^\infty})}{2Nd}\mathbf{1}_{s=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1},$$

where $\mathsf{C} > 0$ is the constant in (6.10) and the reader will recall the definition of the modulated free energy from (1.5). The inclusion of the last two terms is to obtain a nonnegative

[10] A degenerate elliptic operator with an A_2 weight, for which there is a good theory [49].

quantity. Using (2.14) and (2.15), averaging with respect to f_N^t, one obtains the inequality

$$\mathcal{E}_N^t \leq \mathcal{E}_N^0 + C \int_0^t \|\nabla u^\tau\|_{L^\infty} \Big(F_N(\underline{x}_N^\tau, \mu^\tau) + \frac{\log(N\|\mu^\tau\|_{L^\infty})}{2Nd} \mathbf{1}_{s=0} + C\|\mu^\tau\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1} \Big) \, d\tau, \tag{2.16}$$

where the vector field u^τ is given by[11]

$$u^\tau := \sigma \nabla \log(\mu^\tau) + \nabla \mathsf{g} * \mu^\tau.$$

In arriving at (2.16), we have implicitly used that $\|\mu^t\|_{L^\infty}$ is nonincreasing. The constant C depends only on d, s and taking C above larger if necessary, we may assume that $\mathsf{C} \geq C$. Since the relative entropy is nonnegative, it may be added to the expression inside the parentheses and the inequality still holds. Applying the Grönwall–Bellman lemma, we obtain

$$\mathcal{E}_N(f_N^t, \mu^t) \leq \mathcal{E}_N(f_N^0, \mu^0) \exp\Big(C \int_0^t \|\nabla u^\tau\|_{L^\infty} \, d\tau \Big), \tag{2.17}$$

A uniform-in-time bound for $\mathcal{E}_N(f_N^t, \mu^t)$ now follows from Proposition 2.6. The details are given in Section 6.3.

2.4. Application to $\dot{W}^{-1,\infty}$ kernels

Finally, in Section 7 we show how our decay estimates approach may be combined with the relative entropy approach of [68] in a straightforward manner to give a proof of uniform-in-time propagation of chaos for mean-field McKean–Vlasov systems

$$\begin{cases} dx_i^t = \dfrac{1}{N} \displaystyle\sum_{1 \leq j \leq N : j \neq i} \mathsf{k}(x_i^t - x_j^t) \, dt + \sqrt{2\sigma} \, dW_i^t, \\ x_i^t|_{t=0} = x_i^0, \end{cases} \quad i \in \{1, \ldots, N\}. \tag{2.18}$$

The vector field/kernel $\mathsf{k} \colon \mathbb{T}^d \setminus \{0\} \to \mathbb{T}^d$ is assumed to satisfy the following conditions:

(i) $\mathsf{k} \in L^1(\mathbb{T}^d)$,

(ii) $\operatorname{div} \mathsf{k} = 0$ in the sense of distributions,[12]

(iii) $\mathsf{k}^\alpha = \partial_\beta V^{\alpha\beta}$ for an $L^\infty(\mathbb{T}^d)$ matrix-valued field $(V^{\alpha\beta})_{\alpha,\beta=1}^d$.

Note that k is no longer assumed to be potential (i.e., $\mathsf{k} = \nabla \mathsf{g}$ for some g). We remark that the last condition (iii) amounts to requiring that k belong to the negative-order Sobolev space $\dot{W}^{-1,\infty}(\mathbb{T}^d)$. A sufficient condition is that $\mathsf{k} \in L^{d,\infty}(\mathbb{T}^d)$ (the weak L^d space).

[11] Note that this is *not* the same vector field as in the pure modulated energy approach due the presence of the first term coming from the relative entropy.

[12] This condition is not strictly necessary; it would suffice to have $\operatorname{div} \mathsf{k} \in \dot{W}^{-1,\infty}$. But we impose it to simplify the presentation.

A model example is the two-dimensional periodic Biot–Savart kernel (i.e., $\mathsf{k} = \mathbb{M}\nabla \mathsf{g}$ for a $90°$ rotation matrix $\mathbb{M}$ and g the periodic Coulomb potential), as explained in [68, p. 531].

Analogously to (1.4), the mean-field limit of system (2.18) is determined by the solution of the Cauchy problem

$$\begin{cases} \partial_t \mu = -\operatorname{div}(\mu \mathsf{k} * \mu) + \sigma \Delta \mu, \\ \mu|_{t=0} = \mu^0, \end{cases} \qquad (t,x) \in [0,\infty) \times \mathbb{T}^d. \tag{2.19}$$

By assumption (ii), the vector field $u := \mathsf{k} * \mu$ is divergence-free. Hence, equation (2.19) is a transport–diffusion equation with a divergence-free vector field. Such an equation fits into the framework of Section 4 for the range $d-2 \le s < d-1$. Indeed, an examination of the proofs of Proposition 2.1 and Lemma 2.7 reveals that we only used that $\nabla \mathsf{g} \in L^1$, no other specific structure on g. Thus, one may repeat the aforementioned proofs with $\nabla \mathsf{g}$ replaced by $\mathsf{k} \in L^1$ and no other change. Since k is divergence-free, the equation conserves average/mass and the lower and upper bounds for the initial data as in Lemma 4.6, respectively. In particular, given $\mu^0 \in L^\infty(\mathbb{T}^d) \cap \dot{W}^{\alpha,p}(\mathbb{T}^d)$ for any $\alpha \ge 0$ and $1 \le p < \infty$,[13] there is a unique global solution $\mu \in C([0,\infty), L^\infty \cap \dot{W}^{\alpha,p})$. If μ is a zero-mean solution, then all the L^p norms are nonincreasing, and for any $1 \le q \le \infty$, we have the decay estimates

$$\begin{aligned} \forall t > 0,\ n \ge 0, \quad \|\nabla^{\otimes n} \mu^t\|_{L^q} \le \mathbf{W}_{n,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}) \|\mu^0\|_{L^1} e^{-C\sigma t} \\ \times (\sigma t)^{-\frac{n}{2} - \frac{d}{2}(1-\frac{1}{q})}. \end{aligned} \tag{2.20}$$

Theorem 2.14. *Let $d \ge 1$, k be a kernel satisfying assumptions (i), (ii), (iii). Let $f_N \in L^\infty([0,\infty), L^1(\mathbb{T}^d))$ be an entropy solution to the Liouville equation (1.3),[14] and let $\mu \in L^\infty([0,\infty), \mathcal{P}(\mathbb{T}^d) \cap W^{2,\infty}(\mathbb{T}^d))$ be a solution to equation (2.19) with $\inf_{\mathbb{T}^d} \mu^0 > 0$. Then*

$$\forall t \ge 0, \quad H_N(f_N^t | (\mu^t)^{\otimes N}) \le \Big(H_N(f_N^0 | (\mu^0)^{\otimes N}) + \frac{\mathcal{C}^t}{N} \Big) e^{\mathcal{C}^t}, \tag{2.21}$$

where $\mathcal{C}\colon [0,\infty) \to [0,\infty)$ is a continuous, nondecreasing function such that $\mathcal{C}^0 = 0$ and $\sup_{t\ge 0} \mathcal{C}^t < \infty$. In addition, $\mathcal{C}$ depends on d, σ, $\|\mathsf{k}\|_{\dot{W}^{-1,\infty}}$, $\|\mu^0\|_{W^{2,\infty}}$, $\inf_{\mathbb{T}^d} \mu^0$.

Remark 2.15. By Remark 6.5, such a bound implies propagation of chaos with an explicit rate, though one not expected to be optimal.

Remark 2.16. By exploiting the Fisher information as in [55] (see the next paragraph), one can obtain a factor of the form $e^{-C_2 t}$ for large t in front of the term $H_N(f_N^0 | (\mu^0)^{\otimes N})$ in (2.21).

---

[13] If α is integral, then $p = \infty$ is allowed.

[14] The existence of such a solution is sketched in [68, Section 1.5].

As commented in Section 1.1, Guillin et al. [55] have previously shown a comparable result to Theorem 2.14. Their proof again is a refinement of [68] and uses techniques more common in the probability community (e.g., uniform-in-time log-Sobolev inequalities to exploit the Fisher information in the entropy dissipation), as opposed to the "PDE approach" of the present paper. They also need uniform-in-time bounds on the $W^{2,\infty}$ norm of μ^t, which they obtain through standard energy methods, akin to the proof of Lemma 2.9, and Sobolev embedding. Relaxation estimates neither enter into their proof nor are established. Therefore, our proof presented below should be taken as an alternative to their work. In particular, our proof demonstrates there is no need to use the Fisher information.

3. Periodic Riesz potentials and heat kernel estimates

We recall from the introduction that g is the unique distributional solution to the equation

$$|\nabla|^{d-s}\mathsf{g} = \mathsf{c}_{d,s}(\delta_0 - 1), \quad x \in \mathbb{T}^d,$$

subject to the constraint that $\int_{\mathbb{T}^d} \mathsf{g} = 0$. Equivalently, g is the distribution with Fourier coefficients $\hat{\mathsf{g}}(\xi) = \mathsf{c}_{d,s}(2\pi|\xi|)^{s-d}\mathbf{1}_{\xi\neq 0}$ for $\xi \in \mathbb{Z}^d$. One can show that $\mathsf{g} \in C^\infty(\mathbb{T}^d \setminus \{0\})$. Moreover, if we let

$$\mathsf{g}_E(x) := -\log|x|\mathbf{1}_{s=0} + |x|^{-s}\mathbf{1}_{0<s<d} \quad \forall x \in \mathbb{R}^d$$

denote the Euclidean log/Riesz potential, then

$$\mathsf{g} - \mathsf{g}_E \in C^\infty\Big(B\Big(0, \frac{1}{4}\Big)\Big). \tag{3.1}$$

For proofs of these facts, we refer the reader to [62]. In particular, these facts imply that $\mathsf{g} \in L^{\frac{d}{s},\infty}(\mathbb{T}^d)$ (the weak $L^{\frac{d}{s}}$ space), a fortiori in $L^1(\mathbb{T}^d)$, and

$$\forall n \geq 0,\ x \in \mathbb{T}^d \setminus \{0\}, \quad |\nabla^{\otimes n}\mathsf{g}(x)| \lesssim_n |x|^{-s-n} + 1.$$

We let $e^{t\Delta}$ denote the Fourier multiplier on $\mathbb{T}^d$ with coefficients $(e^{-4\pi^2 t|\xi|^2})_{\xi\in\mathbb{Z}^d}$, and we let K_t denote the convolution kernel of $e^{t\Delta}$. It is easy to check from the Fourier representation that $\mathsf{K}_t \in C^\infty(\mathbb{T}^d)$ and $\int_{\mathbb{T}^d} \mathsf{K}_t = 1$, for every $t > 0$. One can explicitly write K_t as the periodization of the Euclidean heat kernel,

$$\mathsf{K}_t(x) = (4\pi t)^{-d/2} \sum_{n\in\mathbb{Z}^d} e^{-\frac{|x-n|^2}{4t}}.$$

For instance, see [19, Section 10.3]. Since $\widehat{\mathsf{K}_t}(\xi) = e^{-4\pi^2 t|\xi|^2}$, it follows that for any $m > \frac{d}{2}$,

$$\|\mathsf{K}_t - 1\|_{L^\infty} \leq \sum_{\xi\in\mathbb{Z}^d:\xi\neq 0} e^{-4\pi^2 t|\xi|^2} \lesssim \sum_{\xi\in\mathbb{Z}^d:\xi\neq 0} (4\pi^2 t|\xi|^2)^{-m} \lesssim t^{-m}. \tag{3.2}$$

The decay as $t \to \infty$ may, in fact, be improved to exponential by applying $\frac{d}{dt}$ to the second expression in (3.2) and using Grönwall's lemma:

$$\forall t \geq t_0, \quad \|\mathsf{K}_t - 1\|_{L^\infty} \leq e^{-4\pi^2(t-t_0)} \sum_{\xi \in \mathbb{Z}^d : \xi \neq 0} e^{-4\pi^2 t_0 |\xi|^2}.$$

Additionally, by Riemann sum approximation, we have

$$\|\mathsf{K}_t - 1\|_{L^\infty} \leq \sum_{\xi \in \mathbb{Z}^d : \xi \neq 0} e^{-4\pi^2 t |\xi|^2} \approx t^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-4\pi^2 |\xi|^2} d\xi,$$

which shows that $\|\mathsf{K}_t\|_{L^\infty} = O(t^{-\frac{d}{2}})$ as $t \to 0$. Hence,

$$\|\mathsf{K}_t - 1\|_{L^\infty} \lesssim_d \min(t,1)^{-\frac{d}{2}} e^{-4\pi^2 \max(t,1)}.$$

One can repeat the same analysis for derivatives and use interpolation to show

$$\forall n \in \mathbb{N}_0,\ t > 0, \quad \|\nabla^{\otimes n}(\mathsf{K}_t - 1)\|_{L^p} \lesssim_{n,d,p} \min(t,1)^{-\frac{d}{2}(1-\frac{1}{p}) - \frac{n}{2}} e^{-C_{n,p} \max(t,1)} \tag{3.3}$$

for any $1 \leq p \leq \infty$. In fact, one can show that if $m(D)$ is a Fourier multiplier with symbol $m(\xi)$ which is homogeneous of degree κ and has symbol $m \in C^\infty(\mathbb{R}^d \setminus \{0\})$, then

$$\|m(D)(\mathsf{K}_t - 1)\|_{L^p} \lesssim_{n,d,p,m(D)} \min(t,1)^{-\frac{d}{2}(1-\frac{1}{p}) - \frac{\kappa}{2}} e^{-C_{p,m(D)} \max(t,1)}.$$

From these properties, we deduce that if $\int_{\mathbb{T}^d} \mu = 0$, then

$$\forall t > 0, \quad \|m(D)e^{t\Delta}\mu\|_{L^p} \lesssim_{n,d,p,q,m(D)} \|\mu\|_{L^q} \min(t,1)^{-\frac{d}{2}(\frac{1}{q}-\frac{1}{p}) - \frac{\kappa}{2}} \times e^{-C_{p,m(D)} \max(t,1)} \tag{3.4}$$

for any $1 \leq q \leq p \leq \infty$. We will use this property, sometimes referred to as *hypercontractivity*, in the remaining body of the paper without further comment.

4. Well-posedness and L^p control for the mean-field equation

In this section we prove well-posedness for the limiting PDE (1.4) in all cases $d \geq 1$ and $d-2 \leq s < d$. The case where $0 \leq s < d-2$ and on $\mathbb{R}^d$ was previously treated in [91], and the proof here is an adaptation of that argument; however, the super-Coulomb case, in particular when $d-1 < s < d$, is more complicated, as the reader will see, due to the loss of regularity in the velocity field $\mathbb{M}\nabla \mathsf{g} * \mu$. While to the best of our knowledge, equation (1.4) has not been previously considered in the literature in its full generality, some special cases of this well-posedness result are known in certain function spaces, as commented in the introduction.

We record here two observations about the solution class to (1.4) that will be important in the sequel.

Remark 4.1. By rescaling time, we may always normalize the mass to be unital up to a change of temperature. More precisely, suppose that μ is a solution to (1.4). Letting $\bar{\mu} = \int_{\mathbb{T}^d} \mu^0$, set $\nu^t := \bar{\mu}^{-1}\mu^{t/\bar{\mu}}$. Then, using the chain rule,

$$\partial_t \nu^t = -\bar{\mu}^{-2}\operatorname{div}(\mu^{t/\bar{\mu}}\nabla \mathsf{g} * \mu^{t/\bar{\mu}}) + \sigma\bar{\mu}^{-2}\Delta\mu^{t/\bar{\mu}} = -\operatorname{div}(\nu^t\nabla \mathsf{g} * \nu^t) + \tilde{\sigma}\Delta\nu^t,$$

where $\tilde{\sigma} := \sigma/\bar{\mu}$.

Remark 4.2. If $c \in \mathbb{R}$, then setting $\nu := \mu - c$, one computes

$$\partial_t \nu = -\operatorname{div}(\nu\mathbb{M}\nabla \mathsf{g} * \nu) - c\operatorname{div}(\mathbb{M}\nabla \mathsf{g} * \nu) + \sigma\Delta\nu.$$

If $\mathbb{M}$ is antisymmetric, then $\operatorname{div}(\mathbb{M}\nabla \mathsf{g} * \mu) = 0$, so $\mu - c$ is a solution to equation (1.4). Indeed, using that the (i,j) entry $\mathbb{M}^{ij}$ of $\mathbb{M}$ is a constant by assumption, the commutativity of differentiation, and $\mathbb{M}^{ij} = -\mathbb{M}^{ji}$,

$$\operatorname{div}(\mathbb{M}\nabla \mathsf{g} * \mu) = \mathbb{M}^{ij}\partial_i\partial_j(\mathsf{g} * \mu) = -\mathbb{M}^{ji}\partial_j\partial_i(\mathsf{g} * \mu) = -\operatorname{div}(\mathbb{M}\nabla \mathsf{g} * \mu), \tag{4.1}$$

since the sum over i, j is symmetric under swapping $i \leftrightarrow j$. In particular, one can always take $c = \int_{\mathbb{T}^d} \mu$ and reduce to considering zero-mean solutions in the conservative case.

4.1. Local well-posedness and basic properties

We prove Proposition 2.1 on local well-posedness of (1.4). We rewrite the equation in the mild form as

$$\mu^t = e^{t\sigma\Delta}\mu^0 - \int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau. \tag{4.2}$$

Let us remark here that the regularity assumptions on μ^0 in Proposition 2.1 are not optimal, but we have imposed them to simplify the proof and because we will need such regularity for the modulated free energy method in Section 6. Indeed, an examination of the proof of Proposition 2.1 below reveals that it would suffice to have $\mu \in L^p$ for sufficiently high p.

Proof of Proposition 2.1. We first consider the case $d-2 \le s \le d-1$. Note that in this case $\mu \mapsto \mathbb{M}\nabla \mathsf{g} * \mu$ is an order $s+1-d$ operator, which is either smoothing ($s < d-1$) or of the same order ($s = d-1$) compared to the regularity of μ. Consider the Banach space[15]

$$X := C([0,T], L^\infty(\mathbb{T}^d)),$$

for some $T > 0$ to be determined. For fixed $\mu^0 \in L^\infty$, the right-hand side of (4.2) defines a map $\mathcal{T}$,

$$\mu \mapsto e^{t\sigma\Delta}\mu^0 - \int_0^{(\cdot)} e^{\sigma((\cdot)-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau.$$

[15] If we were to adapt the proof to $\mathbb{R}^d$, the definition of X would need to be modified to $X := C([0,T], L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$.

We aim to show that $\mathcal{T}$ is a contraction on the ball $B_R \subset X$, for $R, T > 0$ appropriately chosen. Once we have shown this, we can appeal to the Banach fixed point theorem to obtain a unique solution to (4.2) in the class X.

By the triangle inequality, for any $t \geq 0$,

$$\begin{aligned}\|\mu^t\|_{L^\infty} &\leq \|e^{t\sigma\Delta}\mu^0\|_{L^\infty} + \left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau\right\|_{L^\infty} \\ &\leq \|\mu^0\|_{L^\infty} + \left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau\right\|_{L^\infty},\end{aligned}$$

where the second line follows from the heat kernel being in L^1. If $s < d-1$, then we may use Minkowski's inequality together with $\|e^{(t-\tau)\sigma\Delta}\operatorname{div}\|_{L^1} \lesssim (\sigma(t-\tau))^{-1/2}$ to obtain

$$\begin{aligned}\left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau\right\|_{L^\infty} &\leq \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}}\|\mu^\tau\|_{L^\infty}\|\nabla \mathsf{g} * \mu^\tau\|_{L^\infty} \\ &\lesssim \|\mu\|_X^2 \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}}\,d\tau \\ &\lesssim \|\mu\|_X^2 (t/\sigma)^{\frac{1}{2}}, \end{aligned} \tag{4.3}$$

where we use that $\nabla \mathsf{g} \in L^1$.[16] If $s = d-1$, then $\mathbb{M}\nabla \mathsf{g} * \mu^\tau$ is a matrix transformation of the Riesz transform vector of μ^τ, which is not bounded on L^∞. Instead, we use the smoothing property (3.4) to obtain

$$\|e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau\|_{L^\infty} \lesssim (\sigma(t-\tau))^{-\frac{d}{2p}-\frac{1}{2}}\|\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^p}$$

for any $p < \infty$ such that $\frac{d}{2p} + \frac{1}{2} < 1$ (i.e., $p > d$). By Hölder's inequality and the boundedness of the Riesz transform on L^p,

$$\|\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^p} \leq \|\mu^\tau\|_{L^\infty}\|\mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^p} \lesssim \|\mu^\tau\|_{L^\infty}\|\mu^\tau\|_{L^p}.$$

Since $\|\cdot\|_{L^p} \leq \|\cdot\|_{L^\infty}$,

$$\left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\,d\tau\right\|_{L^\infty} \lesssim \sigma^{-\frac{d}{2p}-\frac{1}{2}} t^{\frac{1}{2}-\frac{d}{2p}}\|\mu\|_X^2.$$

Repeating the preceding argument, we also have for any $\mu_1, \mu_2 \in X$,

$$\begin{aligned}&\left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu_1^\tau \mathbb{M}\nabla \mathsf{g} * \mu_1^\tau)\,d\tau - \int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu_2^\tau \mathbb{M}\nabla \mathsf{g} * \mu_2^\tau)\,d\tau\right\|_{L^\infty} \\ &\quad\lesssim \|\mu_1 - \mu_2\|_X(\|\mu_1\|_X + \|\mu_2\|_X)\big((t/\sigma)^{\frac{1}{2}}\mathbf{1}_{s<d-1} + \sigma^{-\frac{d}{2p}-\frac{1}{2}} t^{\frac{1}{2}-\frac{d}{2p}}\mathbf{1}_{s=d-1}\big).\end{aligned}$$

[16] Note this is obviously false for $\mathbb{R}^d$, and this step needs to be modified with the usual L^∞ Riesz potential interpolation estimate.

Suppose that $\mu_1, \mu_2 \in B_R \subset X$. Then we have shown

$$\|\mathcal{T}(\mu_1) - \mathcal{T}(\mu_2)\|_X \le CR\|\mu_1 - \mu_2\|_X((T/\sigma)^{\frac{1}{2}}\mathbf{1}_{s<d-1} + \sigma^{-\frac{d}{2p}-\frac{1}{2}}T^{\frac{1}{2}-\frac{d}{2p}}\mathbf{1}_{s=d-1})$$

for a constant $C > 0$ depending on d, s, $\mathbb{M}$. Fix $R > 2\|\mu^0\|_{L^\infty}$, where μ^0 is the initial datum for the Cauchy problem. Evidently, for $\mu \in B_R$,

$$\|\mathcal{T}\mu\|_X < \frac{R}{2} + 2CR^2((T/\sigma)^{\frac{1}{2}}\mathbf{1}_{s<d-1} + \sigma^{-\frac{d}{2p}-\frac{1}{2}}T^{\frac{1}{2}-\frac{d}{2p}}\mathbf{1}_{s=d-1}).$$

Choosing $T > 0$ such that

$$2CR((T/\sigma)^{\frac{1}{2}}\mathbf{1}_{s<d-1} + \sigma^{-\frac{d}{2p}-\frac{1}{2}}T^{\frac{1}{2}-\frac{d}{2p}}\mathbf{1}_{s=d-1}) = \frac{1}{2},$$

we have that $\mathcal{T}(B_R) \subset B_R$. Now for any $\mu_1, \mu_2 \in B_R$, we additionally have

$$\|\mathcal{T}(\mu_1) - \mathcal{T}(\mu_2)\|_X < \frac{1}{2}\|\mu_1 - \mu_2\|_X,$$

which shows that $\mathcal{T}$ is a contraction on B_R. We can also extract from our analysis the local Lipschitz dependence on the initial data,

$$\forall \mu_1, \mu_2 \in B_R, \quad \|\mu_1 - \mu_2\|_X \le 2\|\mu_1^0 - \mu_2^0\|_{L^\infty}.$$

We now consider the case $d - 1 < s < d$. The velocity field $\mathbb{M}\nabla \mathsf{g} * \mu$ now has less regularity than μ, and we need to exploit additional smoothing to avoid this loss of regularity. Let α, p be as in the statement of the proposition. Recycling notation, we will show that the map $\mathcal{T}$ above is a contraction on the ball B_R of the Banach space

$$X := C([0,T], L^\infty(\mathbb{T}^d) \cap \dot{W}^{\alpha,p}(\mathbb{T}^d)),$$

for some $T, R > 0$ to be determined.

By arguing similarly to (4.3), we observe that for any $q > d$,

$$\begin{aligned}
&\left\|\int_0^t e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\, d\tau\right\|_{L^\infty} \\
&\qquad \le \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2q}}\|\mu^\tau\|_{L^\infty}\|\nabla \mathsf{g} * \mu^\tau\|_{L^q}\, d\tau.
\end{aligned} \tag{4.4}$$

If $\infty > p > d$, then we may choose $q = p$ and estimate

$$\|\nabla \mathsf{g} * \mu^\tau\|_{L^q} \lesssim \| \, |\nabla|^{1+s-d}\mu^\tau\|_{L^p} \lesssim \| \, |\nabla|^\alpha \mu^\tau\|_{L^p}. \tag{4.5}$$

If $p \le d$, then our assumption $\alpha > s - d + \frac{d}{p}$ implies by Sobolev embedding that we may find a $q > d$ such that

$$\|\nabla \mathsf{g} * \mu^\tau\|_{L^q} \lesssim \| \, |\nabla|^\alpha \mu^\tau\|_{L^p}. \tag{4.6}$$

We now conclude that

$$\left\| \int_0^t e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M} \nabla g * \mu^\tau)\, d\tau \right\|_{L^\infty} \lesssim \|\mu\|_X^2 \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2q}}\, d\tau \\ \lesssim \|\mu\|_X^2 \sigma^{-\frac{1}{2}-\frac{d}{2q}} t^{\frac{1}{2}-\frac{d}{2q}}. \tag{4.7}$$

Next we let $\delta \in (s+1-d, 1)$. Then, again using the smoothing of the heat kernel, followed by the fractional Leibniz rule (e.g., see [54, Theorem 7.6.1] or [72, Theorem 1.5]),[17] we find

$$\| |\nabla|^\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M} \nabla g * \mu^\tau)\|_{L^p} \\ \lesssim (\sigma(t-\tau))^{-\frac{1}{2}-\frac{\delta}{2}} (\| |\nabla|^{\alpha-\delta} \mu^\tau \|_{L^{p_1}} \|\nabla g * \mu^\tau\|_{L^{p_2}} \\ + \|\mu^\tau\|_{L^{\tilde{p}_1}} \| |\nabla|^{\alpha-\delta} \nabla g * \mu^\tau \|_{L^{\tilde{p}_2}}), \tag{4.8}$$

where $p \le p_1, \tilde{p}_1, p_2, \tilde{p}_2 \le \infty$ satisfy $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{\tilde{p}_1} + \frac{1}{\tilde{p}_2} = \frac{1}{p}$. We choose $\tilde{p}_1 = \infty$ and $\tilde{p}_2 = p$, so that

$$\|\mu^\tau\|_{L^{\tilde{p}_1}} \| |\nabla|^{\alpha-\delta} \nabla g * \mu^\tau \|_{L^{\tilde{p}_2}} \lesssim \|\mu^\tau\|_{L^\infty} \| |\nabla|^\alpha \mu^\tau \|_{L^p}, \tag{4.9}$$

since $\alpha - \delta + s + 1 - d < \alpha$ by choice of δ. We choose $p_1 = \frac{\alpha p}{\alpha - \delta}$ and $p_2 = \frac{\alpha p}{\delta}$, which are evidently Hölder conjugate to p. Then, by the Gagliardo–Nirenberg interpolation inequality (e.g., see [6, Theorem 2.44]),[18]

$$\| |\nabla|^{\alpha-\delta} \mu^\tau \|_{L^{p_1}} \lesssim \|\mu^\tau\|_{L^\infty}^{\frac{\delta}{\alpha}} \| |\nabla|^\alpha \mu^\tau \|_{L^p}^{1-\frac{\delta}{\alpha}} \tag{4.10}$$

and

$$\|\nabla g * \mu^\tau\|_{L^{p_2}} \lesssim \| |\nabla|^\delta \mu^\tau \|_{L^{p_2}} \lesssim \|\mu^\tau\|_{L^\infty}^{1-\frac{\delta}{\alpha}} \| |\nabla|^\alpha \mu^\tau \|_{L^p}^{\frac{\delta}{\alpha}}. \tag{4.11}$$

Combining estimates, we conclude

$$\left\| |\nabla|^\alpha \int_0^t e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M} \nabla g * \mu^\tau)\, d\tau \right\|_{L^p} \lesssim \|\mu\|_X^2 \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{\delta}{2}}\, d\tau \\ \lesssim \sigma^{-\frac{1}{2}-\frac{\delta}{2}} t^{\frac{1}{2}-\frac{\delta}{2}} \|\mu\|_X^2. \tag{4.12}$$

Putting together the estimates (4.7), (4.12) and using the properties of the heat kernel for the linear part of the map $\mathcal{T}$, we arrive at

$$\|\mathcal{T}(\mu)\|_X \le \|\mu^0\|_{L^\infty \cap \dot{W}^{\alpha,p}} + C\|\mu\|_X^2 (\sigma^{-\frac{1}{2}-\frac{d}{2q}} T^{\frac{1}{2}-\frac{d}{2q}} + \sigma^{-\frac{1}{2}-\frac{\delta}{2}} T^{\frac{1}{2}-\frac{\delta}{2}}), \tag{4.13}$$

[17] The estimates are stated for $\mathbb{R}^d$, but they carry over to $\mathbb{T}^d$ as well when restricted to zero-mean functions.

[18] Again, this result is stated on $\mathbb{R}^d$ but is adaptable to $\mathbb{T}^d$ for zero-mean functions.

for some constant $C > 0$ depending on $d, s, p, \alpha, \delta, \mathbb{M}$. Completely analogous analysis also shows that

$$\|\mathcal{T}(\mu_1) - \mathcal{T}(\mu_2)\|_X \le C(\|\mu_1\|_X + \|\mu_2\|_X)\|\mu_1 - \mu_2\|_X \times (\sigma^{-\frac{1}{2}-\frac{d}{2q}} T^{\frac{1}{2}-\frac{d}{2q}} + \sigma^{-\frac{1}{2}-\frac{\delta}{2}} T^{\frac{1}{2}-\frac{\delta}{2}}). \tag{4.14}$$

Without loss of generality, we may choose δ sufficiently close to 1 and q sufficiently close to d so that $\delta = \frac{d}{q}$. Letting $R > 2\|\mu^0\|_{L^\infty \cap \dot{W}^{\alpha,p}}$, we see that if we choose T so that

$$2CR\sigma^{-\frac{1+\delta}{2}} T^{\frac{1-\delta}{2}} = \frac{1}{2},$$

then $\mathcal{T}$ is a contraction on B_R. So by the Banach fixed point theorem, we obtain a solution to (4.2) in X. Additionally, we can extract from the reasoning used to obtain (4.13), (4.14) that if μ_1, μ_2 are two solutions to (4.2) in B_R with initial data μ_1^0, μ_2^0, respectively, then

$$\|\mu_1 - \mu_2\|_X \le 2\|\mu_1^0 - \mu_2^0\|_{L^\infty \cap \dot{W}^{\alpha,p}},$$

which shows local Lipschitz continuity of the solution map on the initial data.

We conclude with the blowup criterion for the solution. Suppose that $T_{\max} < \infty$, but

$$\limsup_{T \to T_{\max}^-} \|\mu\|_{C([0,T],X)} < \infty.$$

Then there exists an $M > 0$, such that for every $T < T_{\max}$, $\|\mu\|_{C([0,T],X)} \le M$. We choose T sufficiently close to $T_{\max}$ so that $T_{\max} - T$ is less than the lower bound for the time of existence given for a solution by Proposition 2.1. Using Proposition 2.1 with initial datum μ^T, we can then increase the lifespan of the solution beyond $T_{\max}$, which is a contradiction. This completes the proof of Proposition 2.1. ∎

Before proving some further properties of solutions to equation (1.4), we record a series of remarks about Proposition 2.1.

Remark 4.3. The proof of local well-posedness makes no assumptions on the *sign* of the interaction or the symmetry properties of the matrix $\mathbb{M}$. In particular, it is valid for both conservative and gradient flows, as well as repulsive and attractive interactions.

Remark 4.4. By using the fractional Leibniz rule as in (4.8), but skipping the Gagliardo–Nirenberg interpolation, one can also show in the case $d-2 \le s \le d-1$ that given $\mu^0 \in L^\infty \cap \dot{W}^{\alpha,p}$, for any $\alpha \ge 0$ and $1 \le p < \infty$, there exists a unique solution to (4.2) in $C([0,T], L^\infty \cap \dot{W}^{\alpha,p})$ for some $T > 0$.

Remark 4.5. Using the dependence on initial data estimates (2.1) and (2.2), we see that we can always approximate solutions to (4.2) by C^∞ solutions, in particular classical solutions.

The solutions we have constructed necessarily conserve mass and have nonincreasing L^1 norm (conserved if $\mathbb{M}$ is antisymmetric); note that we did not limit ourselves to nonnegative solutions in Proposition 2.1, so the mass and L^1 norm a priori do not coincide. Moreover, solutions preserve the upper and lower bounds of the initial data. One may interpret this as a form of maximum/minimum principle.

Lemma 4.6. *Let μ be a solution of equation (1.4) as in Proposition 2.1. The following hold:*

- (Mass conservation) *For $t \geq 0$, $\int_{\mathbb{T}^d} \mu^t \, dx = \int_{\mathbb{T}^d} \mu^0 \, dx$.*
- (Nonincreasing L^1) *For $t \geq 0$, $\|\mu^t\|_{L^1} \leq \|\mu^0\|_{L^1}$.*
- (Maximum/minimum principle) *If $c_1 \leq \mu^0 \leq c_2$ a.e., then for $t \geq 0$, $c_1 \leq \mu^t \leq c_2$ a.e.*[19]

Proof. We only sketch the proof of the first two assertions. It suffices by approximation to consider a smooth solution. Then

$$\frac{d}{dt} \int_{\mathbb{T}^d} \mu^t \, dx = \int_{\mathbb{T}^d} \operatorname{div}((\mu^t \mathbb{M} \nabla \mathsf{g} * \mu^t) + \sigma \nabla \mu^t) \, dx = 0$$

by the fundamental theorem of calculus and periodicity, which gives mass conservation. For nonincrease of the L^1 norm, we regularize the function $|\cdot|$ by $\sqrt{\varepsilon^2 + |\cdot|^2}$, which is C^∞ for fixed $\varepsilon > 0$. By the chain rule,

$$\frac{d}{dt} \int_{\mathbb{T}^d} \sqrt{\varepsilon^2 + |\mu^t|^2} \, dx = \int_{\mathbb{T}^d} \frac{\mu^t (\operatorname{div}(\mu^t \mathbb{M} \nabla \mathsf{g} * \mu^t) + \sigma \Delta \mu^t)}{\sqrt{\varepsilon^2 + |\mu^t|^2}}$$

and the result is obtained by integration by parts and letting $\varepsilon \to 0$.

For the third assertion, consider first the lower bound with $c_1 = 0$. If $\mu^0 \geq 0$ a.e., then the solution $\mu^t \geq 0$ a.e. on its lifespan. Indeed, let μ^t_+, μ^t_- denote the positive and negative parts of μ^t respectively. Then

$$\frac{d}{dt} \int_{\mathbb{T}^d} \mu^t_\pm \, dx = \frac{1}{2} \frac{d}{dt} \int_{\mathbb{T}^d} (|\mu^t| \pm \mu^t) \, dx \leq 0,$$

where the final inequality follows from the first two assertions. In particular, if $\int_{\mathbb{T}^d} \mu^0_- \, dx = 0$, then $0 \leq \int_{\mathbb{T}^d} \mu^t_- \, dx \leq \int_{\mathbb{T}^d} \mu^0_- \, dx = 0$, which implies that $\mu^t_- = 0$ a.e. The same reasoning establishes the upper bound with $c_2 = 0$: if $\mu^0 \leq 0$ a.e., then $\mu^t \leq 0$ a.e. for $t \geq 0$.

Using this result, we can also show in the conservative case that if $c_1 \leq \mu^0 \leq c_2$ a.e. (i.e., the initial data is bounded from above and below), then $c_1 \leq \mu^t \leq c_2$ a.e. for every $t > 0$. Simply consider $\mu^t - c_1$ and $\mu^t - c_2$ which are solutions to (1.4) with initial data

[19] As we show later in Section 5 that μ^t is smooth, a fortiori continuous, for $t > 0$, the inequalities hold pointwise everywhere.

$\mu^0 - c_1 \geq 0$ and $\mu^0 - c_2 \leq 0$ a.e., respectively. For the gradient flow case, let $0 \leq c \leq \bar{\mu}$ and use integration by parts to compute

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{T}^d}(\mu^t - c)_- \, dx &= -\int_{\mu^t \leq c}(\operatorname{div}(\mu^t \nabla \mathsf{g} * \mu^t) + \sigma\Delta\mu^t)\, dx \\ &= -\int_{\mu^t \leq c}\nabla\mu^t \cdot \nabla\mathsf{g} * \mu^t \, dx + \mathsf{c}_{d,s}\int_{\mu^t \leq c}\mu^t |\nabla|^{2+s-d}(\mu^t - \bar{\mu})\, dx \\ &\quad - \sigma\int_{\mu^t = c}\nabla\mu^t \cdot \frac{\nabla\mu^t}{|\nabla\mu^t|}\, dx. \end{aligned} \tag{4.15}$$

The last term is nonpositive and may be discarded. For the first term, observe that $\nabla\mu^t \mathbf{1}_{\mu^t \leq c} = -\nabla(\mu^t - c)_-$ a.e. Hence, integrating by parts,

$$-\int_{\mu^t \leq c}\nabla\mu^t \cdot \nabla\mathsf{g} * \mu^t = \mathsf{c}_{d,s}\int_{\mathbb{T}^d}(\mu^t - c)_- |\nabla|^{2+s-d}(\mu^t)\, dx.$$

Similarly, $\mu^t \mathbf{1}_{\mu^t \leq c} = -(\mu^t - c)_- + c\mathbf{1}_{\mu^t \leq c}$, which implies that the right-hand side of (4.15) is

$$\leq \mathsf{c}_{d,s} c \int_{\mu^t \leq c}|\nabla|^{2+s-d}(\mu^t)\, dx. \tag{4.16}$$

In particular, in the Coulomb case $s = d - 2$ and assuming $\mu^t \geq 0$, the right-hand side is $\leq \mathsf{c}_{d,s} c |\{\mu^t \leq c\}|(c - \bar{\mu})$. For general $d - 2 < s < d$, we use the definition of the fractional Laplacian on $\mathbb{T}^d$ (e.g., see [43, Proposition 2.2]) to write

$$\int_{\mu^t \leq c}|\nabla|^{2+s-d}(\mu^t)\, dx = C_{s,d}\sum_{k \in \mathbb{Z}^d}\int_{\mu^t \leq c}\int_{\mathbb{T}^d}\frac{\mu^t(x) - \mu^t(y)}{|x - y - k|^{d+(2+s-d)}}\, dy\, dx.$$

Evidently,

$$\int_{\mu^t \leq c}\int_{\mu^t > c}\frac{\mu^t(x) - \mu^t(y)}{|x - y - k|^{d+(2+s-d)}}\, dy\, dx \leq 0.$$

Since by swapping $x \leftrightarrow y$ and making the change of variable $-k \mapsto k$,

$$\begin{aligned}&\sum_k \int_{\mu^t \leq c}\int_{\mu^t \leq c}\frac{\mu^t(x) - \mu^t(y)}{|x - y - k|^{d+(2+s-d)}}\, dy\, dx \\ &\qquad = \sum_k \int_{\mu^t \leq c}\int_{\mu^t \leq c}\frac{\mu^t(y) - \mu^t(x)}{|x - y - k|^{d+(2+s-d)}}\, dy\, dx = 0,\end{aligned}$$

we conclude that the right-hand side of (4.16) is ≤ 0. Hence, $\int_{\mathbb{T}^d}(\mu^t - c)_- \, dx$ is nonincreasing, and so if c is chosen such that $\inf \mu^0 \geq c$, then $\int_{\mathbb{T}^d}(\mu^t - c)_- \, dx = 0$ for every t in the lifespan of μ. This implies that $\inf \mu^t \geq \inf \mu^0$. An analogous argument shows that if $c \geq \bar{\mu}$, then $\int_{\mathbb{T}^d}(\mu^t - c)_+ \, dx$ is nonincreasing. In particular, if $c \geq \sup \mu^0$, then $\int_{\mathbb{T}^d}(\mu^t - c)_+ \, dx = 0$ on the lifespan of μ, implying $\sup \mu^t \leq \sup \mu^0$. ■

We conclude this subsection with the proof of Proposition 2.2, showing that the solutions given by Proposition 2.1 are, in fact, global (i.e., $T_{\max} = \infty$). The case $d-2 \leq s \leq d-1$ is a triviality since we know from Lemma 4.6 that the L^∞ norm is nonincreasing. For the case $d-1 < s < d$, we are not able to show yet – but we will in the next section – that $\|\mu^t\|_{\dot{W}^{\alpha,p}}$ is controlled by $\|\mu^0\|_{\dot{W}^{\alpha,p}}$. But this is unnecessary: it suffices to show that $\|\mu^t\|_{\dot{W}^{\alpha,p}}$ cannot blow up in finite time.

Proof of Proposition 2.2. For the case $d-2 \leq s \leq d-1$, we have already explained the proof. For the case $d-1 < s < d$, we simply need to revisit the proof of Proposition 2.1. Recalling estimates (4.4), (4.5), (4.6), we have

$$\begin{aligned}&\left\| \int_0^t e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\, d\tau \right\|_{L^\infty} \\ &\qquad \lesssim \|\mu^0\|_{L^\infty} \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2q}} \|\, |\nabla|^\alpha \mu^\tau\|_{L^p}\, d\tau,\end{aligned}$$

for some $q > d$ depending on p. Recalling estimates (4.8), (4.9), (4.10), (4.11),

$$\begin{aligned}&\left\| |\nabla|^\alpha \int_0^t e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\, d\tau \right\|_{L^p} \\ &\qquad \lesssim \|\mu^0\|_{L^\infty} \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{\delta}{2}} \|\, |\nabla|^\alpha \mu^\tau\|_{L^p}\, d\tau,\end{aligned}$$

for $\delta \in (s+1-d, 1)$. Note here that in the gradient flow case, we are implicitly using that the interaction is repulsive to control $\|\mu^t\|_{L^\infty} \leq \|\mu^0\|_{L^\infty}$ (recall Lemma 4.6). Hence,

$$\begin{aligned}\|\mu^t\|_{L^\infty \cap \dot{W}^{\alpha,p}} \lesssim \|\mu^0\|_{L^\infty \cap \dot{W}^{\alpha,p}} + \|\mu^0\|_{L^\infty} \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2q}} \|\, |\nabla|^\alpha \mu^\tau\|_{L^p}\, d\tau \\ + \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{\delta}{2}} \|\, |\nabla|^\alpha \mu^\tau\|_{L^p}\, d\tau.\end{aligned}$$

By adjusting q or δ if necessary, we may assume without loss of generality that $\frac{\delta}{2} = \frac{d}{2q}$. The singular Grönwall lemma [80, Chapter 5, Lemma 6.7] now implies that for any $T > 0$,

$$\sup_{0 \leq t \leq T} \|\mu^t\|_{L^\infty \cap \dot{W}^{\alpha,p}} < \infty,$$

completing the proof of the proposition. ∎

4.2. Control of L^p norms

We saw in Lemma 4.6 that the L^1 and L^∞ norms of a solution are nonincreasing. By interpolation, this implies that for any $1 \leq p \leq \infty$,

$$\forall t \geq 0, \quad \|\mu^t\|_{L^p} \leq \|\mu^0\|_{L^1}^{\frac{1}{p}} \|\mu^0\|_{L^\infty}^{1-\frac{1}{p}}.$$

As we show in Lemma 4.9 below, it is possible to control the L^p norm, for any $1 < p < \infty$, of the solution in terms of the L^p norm of the initial datum.

Before stating and proving the lemma, we recall some technical preliminaries. The first is a type of Poincaré inequality adapted from [1, Lemma 3.2]. Note that when $p = 2$, the inequality is the usual Poincaré inequality, which is a trivial consequence of Plancherel's theorem. We include a sketch of the proof for the reader's convenience.

Lemma 4.7. *Let $d \ge 1$ and $1 \le p < \infty$. There exists a constant $C > 0$ depending only on d, p, such that for any f with $\int_{\mathbb{T}^d} f = 0$ and $|f|^{\frac{p}{2}} \operatorname{sgn}(f) \in W^{1,2}(\mathbb{T}^d)$, it holds that*

$$\int_{\mathbb{T}^d} |f|^p \, dx \le C \int_{\mathbb{T}^d} |\nabla |f|^{\frac{p}{2}}|^2 \, dx.$$

Proof. The proof is by contradiction. Suppose there is a sequence of f_n with $\int_{\mathbb{T}^d} f_n = 0$, $\|f_n\|_{L^p} = 1$, and $\|\nabla(|f_n|^{\frac{p}{2}} \operatorname{sgn}(f_n))\|_{L^2} \to 0$ as $n \to \infty$. By approximation, we may assume without loss of generality that f_n is C^∞. Set $g_n := |f_n|^{\frac{p}{2}} \operatorname{sgn}(f_n)$. Since g_n is bounded in $W^{1,2}(\mathbb{T}^d)$, the Rellich–Kondrachov embedding implies that (up to a subsequence) there is a $g \in W^{1,2}(\mathbb{T}^d)$ such that $\|g_n - g\|_{L^2} \to 0$ as $n \to \infty$. Since $1 = \|f_n\|_{L^p}^p = \|g_n\|_{L^2}^2$, it follows that $1 = \|g\|_{L^2}$. Since $\|\nabla g_n\|_{L^2} \to 0$, it follows from lower semicontinuity that $\nabla g = 0$, which implies that $g = c$ for some constant c. Since $\|g\|_{L^2} = 1$, in fact, $c = 1$, which implies that $f := |g|^{\frac{2}{p}} \operatorname{sgn}(g) = 1$. Since $f_n = |g_n|^{\frac{2}{p}} \operatorname{sgn}(g_n) \to f$ in L^p by the fact that $g_n \to g$ in L^2, we also have that $f_n \to f$ in L^1. This implies that $\int_{\mathbb{T}^d} f \, dx = 0$, which contradicts that $c = 1$ above. ∎

The second ingredient is a positivity lemma for the fractional Laplacian, adapted from [43, Lemma 2.5], which one may view as a "cheap" version of the Stroock–Varopoulos inequality (e.g., see [14, Proposition 3.1]).

Lemma 4.8 (Positivity lemma). *For $0 \le \alpha \le 2$, $f \in C^\infty(\mathbb{T}^d)$,*[20] *and $1 \le q < \infty$, it holds that*

$$\int_{\mathbb{T}^d} |f|^{q-1} \operatorname{sgn}(f) |\nabla|^\alpha f \, dx \ge 0. \tag{4.17}$$

We now come to the main lemma of this subsection, which uses the assumption that g is repulsive in the gradient flow case and $d - 2 \le s < d$, in contrast to Proposition 2.1.

Lemma 4.9. *Let $d - 2 \le s < d$ and $1 \le p < \infty$. If μ^t is a solution to equation (1.4), with the condition that $\mu^t \ge 0$ if $\mathbb{M} = -\mathbb{I}$, then*

$$\forall t \ge 0, \quad \|\mu^t\|_{L^p} \le \|\mu^0\|_{L^p}.$$

Furthermore, in the case when $\mathbb{M}$ is antisymmetric, there is a constant $C > 0$ depending only on d, p, such that

$$\forall t \ge 0, \quad \|\mu^t - \bar{\mu}\|_{L^p} \le e^{-C\sigma t} \|\mu^0 - \bar{\mu}\|_{L^p}. \tag{4.18}$$

[20] The C^∞ assumption is qualitative. One just needs enough regularity for the integral in (4.17) to make sense.

Proof. Without loss of generality, suppose that μ is a classical solution to equation (1.4). Using the chain rule, we compute

$$\begin{aligned}\frac{d}{dt}\|\mu^t\|_{L^p}^p &= p\int_{\mathbb{T}^d} |\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\partial_t\mu^t\,dx\\ &= p\sigma\int_{\mathbb{T}^d} |\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\Delta\mu^t\,dx\\ &\quad - p\int_{\mathbb{T}^d} |\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx.\end{aligned}$$

By the product rule,

$$\begin{aligned}&- p\int_{\mathbb{T}^d} |\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx\\ &\quad= -p\int_{\mathbb{T}^d}|\mu^t|^p\operatorname{div}(\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx - p\int_{\mathbb{T}^d}|\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\nabla\mu^t\cdot(\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx.\end{aligned}$$

Since

$$p|\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\nabla\mu^t = \nabla(|\mu^t|^p),$$

an integration by parts reveals that

$$-p\int_{\mathbb{T}^d} |\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx = -(p-1)\int_{\mathbb{T}^d}|\mu^t|^p\operatorname{div}(\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx.$$

Finally, write

$$\operatorname{sgn}(\mu^t) = \lim_{\epsilon\to0^+}\frac{\mu^t}{\sqrt{\epsilon^2+|\mu^t|^2}}.$$

Integrating by parts,

$$\begin{aligned}&p\int_{\mathbb{T}^d}|\mu^t|^{p-1}\operatorname{sgn}(\mu^t)\Delta\mu^t\,dx\\ &\quad= \lim_{\epsilon\to0^+}\Bigg(-p(p-1)\int_{\mathbb{T}^d}|\mu^t|^{p-2}\frac{|\mu^t|}{\sqrt{\epsilon^2+|\mu^t|^2}}|\nabla\mu^t|^2\,dx\\ &\qquad\qquad - p\int_{\mathbb{T}^d}|\mu^t|^{p-1}\frac{|\nabla\mu^t|^2}{\sqrt{\epsilon^2+|\mu^t|^2}}\,dx + p\int_{\mathbb{T}^d}|\mu^t|^{p-1}\frac{|\mu^t|^2|\nabla\mu^t|^2}{(\epsilon^2+|\mu^t|^2)^{3/2}}\,dx\Bigg)\\ &\quad= -p(p-1)\int_{\mathbb{T}^d}|\mu^t|^{p-2}|\nabla\mu^t|^2\,dx\\ &\quad= -\frac{4(p-1)}{p}\int_{\mathbb{T}^d}|\nabla|\mu^t|^{p/2}|^2\,dx.\end{aligned}$$

After a little bookkeeping, we arrive at

$$\begin{aligned}\frac{d}{dt}\|\mu^t\|_{L^p}^p &= -(p-1)\int_{\mathbb{T}^d}|\mu^t|^p\operatorname{div}(\mathbb{M}\nabla\mathsf{g}*\mu^t)\,dx\\ &\quad - 4\sigma\frac{(p-1)}{p}\int_{\mathbb{T}^d}|\nabla|\mu^t|^{\frac{p}{2}}|^2\,dx.\end{aligned} \tag{4.19}$$

Suppose that $\mathbb{M}$ is antisymmetric. Then, recalling (4.1), identity (4.19) becomes

$$\frac{d}{dt}\|\mu^t\|_{L^p}^p = -4\sigma\frac{(p-1)}{p}\|\nabla|\mu^t|^{\frac{p}{2}}\|_{L^2}^2,$$

which implies that $\|\mu^t\|_{L^p}^p$ is nonincreasing and, in fact, strictly decreasing unless $|\mu^t|$ is constant. Additionally, since $\mu^t - \bar{\mu}$ is also a solution to equation (1.4), replacing μ^t above with $\mu^t - \bar{\mu}$, we find that

$$\frac{d}{dt}\|\mu^t - \bar{\mu}\|_{L^p}^p = -4\sigma\frac{(p-1)}{p}\|\nabla|\mu^t - \bar{\mu}|^{\frac{p}{2}}\|_{L^2}^2 \le -\frac{4C_{d,p}\sigma(p-1)}{p}\|\mu^t - \bar{\mu}\|_{L^p}^p,$$

where the final inequality follows from application of Lemma 4.7. From Grönwall's lemma, we then obtain

$$\|\mu^t - \bar{\mu}\|_{L^p}^p \le e^{-C_{d,p}\sigma t}\|\mu^0 - \bar{\mu}\|_{L^p}^p,$$

where we have redefined $C_{d,p}$ compared to above.

Suppose now that $\mathbb{M} = -\mathbb{I}$. Then

$$\operatorname{div}(\nabla \mathsf{g} * \mu^t) = -\mathsf{c}_{d,s}|\nabla|^{s-d+2}(\mu^t - \bar{\mu}),$$

so that (4.19) becomes

$$\begin{aligned}\frac{d}{dt}\|\mu^t\|_{L^p}^p &= -(p-1)\mathsf{c}_{d,s}\int_{\mathbb{T}^d}|\mu^t|^p|\nabla|^{s-d+2}(\mu^t - \bar{\mu})\,dx \\ &\quad -\frac{4\sigma(p-1)}{p}\int_{\mathbb{T}^d}|\nabla|\mu^t|^{\frac{p}{2}}|^2\,dx.\end{aligned}$$

If $s = d-2$, then

$$\int_{\mathbb{T}^d}|\mu^t|^p|\nabla|^{s-d+2}(\mu^t - \bar{\mu})\,dx = \|\mu^t\|_{L^{p+1}}^{p+1} - \bar{\mu}\|\mu^t\|_{L^p}^p \ge 0,$$

where we have used that $\mu^t \ge 0$ by assumption, so that $\bar{\mu} = \|\mu^t\|_{L^1} \le \|\mu^t\|_{L^p}$, and that $\|\cdot\|_{L^{p+1}} \ge \|\cdot\|_{L^p}$. If $d-2 < s < d$, then we may apply Lemma 4.8 with $f = \mu^t$ (again using that $\mu^t \ge 0$ by assumption) to obtain

$$\int_{\mathbb{T}^d}|\mu^t|^p|\nabla|^{s-d+2}(\mu^t - \bar{\mu})\,dx \ge 0.$$

In all cases, we conclude that

$$\frac{d}{dt}\|\mu^t\|_{L^p}^p \le -\frac{4\sigma(p-1)}{p}\|\nabla|\mu^t|^{\frac{p}{2}}\|_{L^2}^2,$$

which completes the proof of the lemma. ■

Remark 4.10. In the Coulomb gradient flow case $\mathbb{M} = -\mathbb{I}$ and $s = d-2$, we can actually obtain an exponential rate of decay for $\|\mu^t - \bar{\mu}\|_{L^p}$, for $1 \le p \le \infty$, through a modification of the proof of Lemma 4.9. Since we do not use such a result in this paper, we do not report on the details.

4.3. Free energy and entropy control

In this subsection we show how the free energy and entropy provide a priori estimates for solutions of equation (1.4), as well as rates of convergence as $t \to \infty$ to the unique equilibrium given by the uniform measure on $\mathbb{T}^d$. We assume through this section that μ is a probability density solution. Recall from Lemma 4.6 that mass and sign are preserved and given a nonnegative solution, we may always rescale time (up to a change of temperature) to normalize the mass to be 1, as explained in Remark 4.1.

If $\mathbb{M} = -\mathbb{I}$, then the *free energy* associated to equation (1.4) is defined by

$$\begin{aligned}\mathcal{F}_\sigma(\mu) &:= \sigma \int_{\mathbb{T}^d} \log(\mu)\, d\mu + \frac{1}{2}\int_{(\mathbb{T}^d)^2} \mathsf{g}(x-y)\, d\mu^{\otimes 2}(x,y) \\ &=: \sigma\, \mathrm{Ent}(\mu) + \mathrm{Eng}(\mu),\end{aligned} \tag{4.20}$$

consisting of the entropy and the energy of μ. Evidently, both terms in the definition of $\mathcal{F}_\sigma(\mu)$ are nonnegative and, in fact, are equal to zero if and only if $\mu \equiv 1$. Equation (1.4) with $\mathbb{M} = -\mathbb{I}$ is the gradient flow of $\mathcal{F}_\sigma$, in the sense that (1.4) may be rewritten as

$$\partial_t \mu = -\operatorname{div}\Big(\mu \nabla \frac{\delta \mathcal{F}_\sigma}{\delta \mu}\Big),$$

where $\frac{\delta \mathcal{F}_\sigma}{\delta \mu}$ is the variational derivative of $\mathcal{F}_\sigma$. Consequently,

$$\frac{d}{dt}\mathcal{F}_\sigma(\mu^t) = -\mathcal{D}_\sigma(\mu^t), \tag{4.21}$$

where $\mathcal{D}_\sigma$ is the *free energy dissipation functional* given by

$$\mathcal{D}_\sigma(\mu^t) := \int_{\mathbb{T}^d} |\sigma \nabla \log(\mu^t) + \nabla(\mathsf{g} * \mu^t)|^2\, d\mu^t.$$

There is a deeper significance behind the relationship of the free energy to equation (1.4) in terms of gradient flows on the manifold of probability densities on $\mathbb{T}^d$ (e.g., see [2]). But we will not make use of this structure and therefore make no further comments on it.

Expanding the square,

$$\begin{aligned}\mathcal{D}_\sigma(\mu^t) &= \int_{\mathbb{T}^d} |\sigma \nabla \log(\mu^t) + \nabla \mathsf{g} * \mu^t|^2\, d\mu^t \\ &= \sigma^2 \int_{\mathbb{T}^d} |\nabla \log(\mu^t)|^2\, d\mu^t + 2\sigma \int_{\mathbb{T}^d} (\nabla \log(\mu^t) \cdot \nabla \mathsf{g} * \mu^t)\, d\mu^t \\ &\quad + \int_{\mathbb{T}^d} |\nabla \mathsf{g} * \mu^t|^2\, d\mu^t \\ &\geq \sigma^2 \int_{\mathbb{T}^d} |\nabla \log(\mu^t)|^2\, d\mu^t + 2\sigma \mathsf{c}_{d,s} \int_{\mathbb{T}^d} \big||\nabla|^{\frac{2+s-d}{2}}(\mu^t)\big|^2\, dx \\ &\quad + \int_{\mathbb{T}^d} |\nabla \mathsf{g} * \mu^t|^2\, d\mu^t,\end{aligned} \tag{4.22}$$

where the last inequality follows from integration by parts in the second integral and the definition of g. Thus we see that $\mathcal{D}_\sigma(\mu^t) = 0$ if and only if $\mu^t \equiv 1$. Since the uniform measure is a stationary solution to equation (1.4), uniqueness of solutions implies that if $\mu^t \equiv 1$ for some $t = t_0$, then $\mu^t \equiv 1$ for all $t \geq t_0$. Thus, the free energy is strictly decreasing unless $\mu^t \equiv 1$, at which point it is then constant for all future time.

For conservative flows, the free energy is no longer the right quantity to consider. Instead, the entropy alone suffices. Given a classical solution of equation (1.4) with anti-symmetric matrix $\mathbb{M}$, we compute

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{T}^d} \log(\mu^t)\, d\mu^t &= \int_{\mathbb{T}^d} \partial_t \mu^t (1 + \log(\mu^t))\, dx \\ &= \int_{\mathbb{T}^d} (\sigma\Delta\mu^t - \operatorname{div}(\mu^t \mathbb{M}\nabla \mathrm{g} * \mu^t))\, dx \\ &\quad - \int_{\mathbb{T}^d} \operatorname{div}(\mu^t \mathbb{M}\nabla \mathrm{g} * \mu^t) \log(\mu^t)\, dx + \sigma \int_{\mathbb{T}^d} \log(\mu^t)\Delta\mu^t\, dx.\end{aligned}$$

By the fundamental theorem of calculus, the first term in the right-hand side of the second equality is obviously zero. The second term is also zero. To see this, we integrate by parts using that $\operatorname{div}(\mathbb{M}\nabla \mathrm{g} * \mu^t) = 0$,

$$\begin{aligned}-\int_{\mathbb{T}^d} \operatorname{div}(\mu^t \mathbb{M}\nabla \mathrm{g} * \mu^t) \log(\mu^t)\, dx &= \int_{\mathbb{T}^d} \nabla\mu^t \cdot \mathbb{M}\nabla \mathrm{g} * \mu^t\, dx \\ &= -\int_{\mathbb{T}^d} \operatorname{div}(\mathbb{M}\nabla \mathrm{g} * \mu^t)\, d\mu^t = 0.\end{aligned}$$

For the third term above, we also integrate by parts to obtain

$$\sigma \int_{\mathbb{T}^d} \log(\mu^t)\Delta\mu^t\, dx = -\sigma \int_{\mathbb{T}^d} \nabla \log(\mu^t) \cdot \nabla\mu^t\, dx = -\sigma \int_{\mathbb{T}^d} |\nabla \log(\mu^t)|^2\, d\mu^t.$$

The right-hand side is the *Fisher information* of μ^t. Putting everything together, we find

$$\frac{d}{dt}\int_{\mathbb{T}^d} \log(\mu^t)\, d\mu^t = -\sigma \int_{\mathbb{T}^d} |\nabla \log(\mu^t)|^2\, d\mu^t. \tag{4.23}$$

Similarly to the free energy dissipation functional, we note that the right-hand side is zero if and only if $\mu^t \equiv 1$. If $\mu^t \equiv 1$ for some $t = t_0$, then the uniqueness of solutions and the fact that 1 is a stationary solution imply $\mu^t \equiv 1$ for all $t \geq t_0$. Thus, the entropy is strictly decreasing unless at some time t_0, $\mu^{t_0} \equiv 1$, at which point the solution remains identically 1 and the entropy is 0 for all subsequent time $t \geq t_0$.

The dissipation of free energy/entropy can be combined with the log-Sobolev inequality for the uniform measure on $\mathbb{T}^d$ to obtain rates of convergence to equilibrium as $t \to \infty$. We reproduce this log-Sobolev inequality from [55, Lemma 3].

Lemma 4.11. *For any probability density f on $\mathbb{T}^d$,*

$$\int_{\mathbb{T}^d} \log(f)\, df \leq \frac{1}{8\pi^2}\int_{\mathbb{T}^d} |\nabla \log(f)|^2\, df.$$

Using Lemma 4.11, we may obtain the following exponential rate of decay for the free energy/entropy. Pinsker's inequality (e.g., see [102, Remark 22.12]) and interpolation imply an exponential rate of convergence to the uniform distribution as $t \to \infty$ in any L^p norm, for finite p.

Proof of Lemma 2.3. By approximation, we may assume without loss of generality that μ is a classical solution. Consider first the gradient flow case. From (4.22), we find

$$\frac{d}{dt}\mathcal{F}_\sigma(\mu^t) \le -\sigma^2 \int_{\mathbb{T}^d} |\nabla \log(\mu^t)|^2 \, d\mu^t - 2\sigma \mathsf{c}_{d,s} \int_{\mathbb{T}^d} ||\nabla|^{\frac{2+s-d}{2}}(\mu^t)|^2 \, dx.$$

Observe from Plancherel's theorem that

$$\begin{aligned}\int_{\mathbb{T}^d} ||\nabla|^{\frac{2+s-d}{2}}(\mu^t)|^2 \, dx &\ge (2\pi)^2 \int_{\mathbb{T}^d} ||\nabla|^{\frac{s-d}{2}}(\mu^t)|^2 \, dx \\ &= \frac{(2\pi)^2}{\mathsf{c}_{d,s}} \int_{(\mathbb{T}^d)^2} \mathsf{g}(x-y) \, d(\mu^t)^{\otimes 2}(x,y).\end{aligned}$$

Applying Lemma 4.11, we then find

$$\begin{aligned}\frac{d}{dt}\mathcal{F}_\sigma(\mu^t) &\le -8\pi^2\sigma^2 \int_{\mathbb{T}^d} \log(\mu^t) \, d\mu^t - 2\sigma(2\pi)^2 \int_{(\mathbb{T}^d)^2} \mathsf{g}(x-y) \, d(\mu^t)^{\otimes 2}(x,y) \\ &\le -8\pi^2\sigma \mathcal{F}_\sigma(\mu^t).\end{aligned}$$

By Grönwall's lemma, we conclude that

$$\mathcal{F}_\sigma(\mu^t) \le \mathcal{F}_\sigma(\mu^0) e^{-8\pi^2\sigma t}. \tag{4.24}$$

For the conservative case, the argument is essentially the same, except we now use the entropy dissipation (4.23) instead of the free energy dissipation (4.21). We ultimately obtain

$$\mathrm{Ent}(\mu^t) \le e^{-8\pi^2\sigma t} \, \mathrm{Ent}(\mu^0). \tag{4.25}$$

By Pinsker's inequality, (4.24) and (4.25) respectively imply

$$\frac{\sigma}{2}\|\mu^t - 1\|_{L^1}^2 + \frac{1}{2}\|\mu^t - 1\|_{\dot{H}^{\frac{s-d}{2}}}^2 \le \mathcal{F}_\sigma(\mu^0) e^{-8\pi^2\sigma t} \qquad \text{if } \mathbb{M} = -\mathbb{I}, \tag{4.26}$$

$$\frac{1}{2}\|\mu^t - 1\|_{L^1}^2 \le e^{-8\pi^2\sigma t} \, \mathrm{Ent}(\mu^0) \qquad \text{if } \mathbb{M} \text{ is antisymmetric.} \tag{4.27}$$

By Lemma 4.6, we know that $\|\mu^t\|_{L^r} \le \|\mu^0\|_{L^r}$. Combining this fact with (4.26), (4.27), and interpolation, we conclude that for any $1 \le p < \infty$, if $\mathbb{M} = -\mathbb{I}$, then

$$\begin{aligned}\|\mu^t - 1\|_{L^p} &\le \|\mu^t - 1\|_{L^1}^{\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}} \|\mu^t - 1\|_{L^r}^{1-\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}} \\ &\le (1 + \|\mu^0\|_{L^r})^{1-\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}} \left(e^{-4\pi^2\sigma t}\sqrt{2\mathcal{F}_\sigma(\mu^0)/\sigma}\right)^{\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}},\end{aligned}$$

and if $\mathbb{M}$ is antisymmetric, then similarly

$$\|\mu^t - 1\|_{L^p} \le (1 + \|\mu^0\|_{L^r})^{1-\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}} \left(e^{-4\pi^2\sigma t}\sqrt{2\,\mathrm{Ent}(\mu^0)}\right)^{\frac{\frac{1}{p}-\frac{1}{r}}{1-\frac{1}{r}}}.$$

This completes the proof of the lemma. ∎

4.4. L^p–L^q smoothing

In the previous subsections, the decay estimate for $\|\mu^t - 1\|_{L^p}$ required at least control on $\|\mu^0\|_{L^{p+}}$ when using the free energy/entropy. Moreover, we have not yet obtained a decay estimate for $\|\mu^t - 1\|_{L^\infty}$. In this subsection we show that it is possible to obtain decay estimates in terms of much weaker control on the initial data.

To do this, we need a log-Sobolev inequality for the uniform measure on $\mathbb{T}^d$. We could not find the exact form that we need in the literature, so we include a proof below (without any outright claims of originality).

Lemma 4.12. *There exist constants $C_{\mathrm{LS},1}, C_{\mathrm{LS},2} > 0$, which depend only on d, such that for any $f \in C^\infty(\mathbb{T}^d)$, with $\bar{f} := \int_{\mathbb{T}^d} f\,dx$, and any $a > 0$, we have*

$$\int_{\mathbb{T}^d} f^2 \log\left(\frac{f^2}{\int_{\mathbb{T}^d} f^2}\right) dx + \frac{d}{2}\log(a)\|f\|_{L^2}^2 \le a\frac{d}{2}(C_{\mathrm{LS},1}\|\nabla f\|_{L^2}^2 + C_{\mathrm{LS},2}|\bar{f}|^2).$$

Proof. The argument is classical and proceeds through Jensen's inequality and Sobolev embedding. Suppose first that $\int_{\mathbb{T}^d} f^2\,dx = 1$, so that f^2 is a probability density on $\mathbb{T}^d$. Then, for $p > 2$, write

$$\int_{\mathbb{T}^d} f^2 \log f\,dx = \frac{1}{p-2}\int_{\mathbb{T}^d} f^2 \log(f^{p-2})\,dx \le \frac{p}{2(p-2)}\log(\|f\|_{L^p}^2).$$

From the inequality $\log t \le at - \log a$, for any $t, a > 0$, it follows that

$$\frac{p}{2(p-2)}\log(\|f\|_{L^p}^2) \le \frac{p}{2(p-2)}(a\|f\|_{L^p}^2 - \log a).$$

Observe from the triangle inequality that $\|f\|_{L^p} \le \|f - \bar{f}\|_{L^p} + |\bar{f}|$. If $d \ge 3$, then we choose $p = \frac{2d}{d-2}$ and use Sobolev embedding to obtain that the right-hand side is

$$\le \frac{d}{4}(2aC_{\mathrm{Sob},\frac{2d}{d-2}}^2\|\nabla f\|_{L^2}^2 + 2a|\bar{f}|^2 - \log a).$$

If $d \le 2$, then we choose any $2 < p < \infty$ and use Sobolev embedding plus interpolation to instead obtain

$$\begin{aligned}&\frac{p}{2(p-2)}\log(\|f\|_{L^p}^2)\\&\quad= \frac{p}{2(p-2)}\cdot\left(\frac{2p}{d(p-2)}\right)^{-1}\log\left(\|f\|_{L^p}^{\frac{4p}{d(p-2)}}\right)\end{aligned}$$

$$
\begin{aligned}
&\le \frac{d}{4}\big(a(\|f-\bar{f}\|_{L^p}+|\bar{f}|)^{\frac{4p}{d(p-2)}}-\log a\big)\\
&\le \frac{d}{4}\Big(a\Big(C_{\mathrm{Sob},p}\|f-\bar{f}\|_{L^2}^{1-d(\frac{1}{2}-\frac{1}{p})}\|f-\bar{f}\|_{\dot{H}^1}^{d(\frac{1}{2}-\frac{1}{p})}+|\bar{f}|\Big)^{\frac{4p}{d(p-2)}}-\log a\Big)\\
&\le \frac{d}{4}\Big(2^{\frac{4p}{d(p-2)}-1}a\Big(|\bar{f}|^{\frac{4p}{d(p-2)}}+C_{\mathrm{Sob},p}^{\frac{4p}{d(p-2)}}2^{\frac{4p}{d(p-2)}(1-d(\frac{1}{2}-\frac{1}{p}))}\|\nabla f\|_{L^2}^2\Big)-\log a\Big)\\
&\le \frac{d}{4}\Big(2^{\frac{4p}{d(p-2)}-1}a\Big(|\bar{f}|^{2}+C_{\mathrm{Sob},p}^{\frac{4p}{d(p-2)}}2^{\frac{4p}{d(p-2)}(1-d(\frac{1}{2}-\frac{1}{p}))}\|\nabla f\|_{L^2}^2\Big)-\log a\Big),
\end{aligned}
$$

where we have implicitly used above that $|\bar{f}|\le\|f\|_{L^2}=1$ and the convexity of $|\cdot|^{\frac{4p}{d(p-2)}}$. To remove the assumption $\int_{\mathbb{T}^d}f^2=1$, we apply the preceding argument to $g:=f/\|f\|_{L^2}$, which satisfies $\int_{\mathbb{T}^d}g^2\,dx=1$. This then gives the inequality in the statement of the lemma. ∎

Next we show that for any time $t>0$, the L^p norm of μ^t controls the L^q norm of μ^t, for $1\le p\le q\le\infty$, at the cost of a factor blowing up like $(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}$ as $t\to 0$ and like $e^{C\sigma t}$ as $t\to\infty$. In other words, this gain of integrability, sometimes called *hypercontractivity*, is only useful for short positive times. This is in contrast to the setting of $\mathbb{R}^d$, where one has this L^p–L^q control with only a factor of $(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}$, which yields the optimal decay of solutions as $t\to\infty$ (cf. [91, Proposition 3.8]).

Lemma 4.13. *Let $d\ge 1$ and $d-2\le s<d$. If μ is a solution to equation (1.4), then for $1\le p\le q\le\infty$,*

$$
\forall t>0,\quad \|\mu^t\|_{L^q}\le C_{p,q,d}(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{C_{p,q,d}\sigma t}\|\mu^0\|_{L^p},
$$

where $C_{p,q,d}>0$ depends only on p, q, d.

Proof. We have already seen that we may assume without loss of generality that μ is spatially smooth on its lifespan and μ is C^∞ in time. Therefore, there are no issues of regularity or decay in justifying the computations to follow. The proof is based on an adaptation of an argument, originally due to Carlen and Loss [29] and extended in [91].

For given p,q as above, let $r\colon[0,T]\to[p,q]$ be a C^1 increasing function to be specified momentarily. Replacing the absolute value $|\cdot|$ with $(\epsilon^2+|\cdot|^2)^{1/2}$, differentiating, then sending $\epsilon\to 0^+$, we find that

$$
\begin{aligned}
r(t)^2\|\mu^t\|_{L^{r(t)}}^{r(t)-1}\frac{d}{dt}\|\mu^t\|_{L^{r(t)}} &= \dot{r}(t)\int_{\mathbb{T}^d}|\mu^t|^{r(t)}\log\Big(\frac{|\mu^t|^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}}\Big)\,dx\\
&\quad+\sigma r(t)^2\int_{\mathbb{T}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\Delta\mu^t\,dx\\
&\quad-r(t)^2\int_{\mathbb{T}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathrm{g}*\mu^t)\,dx.
\end{aligned}
$$

Above, we have used the calculus identity

$$
\frac{d}{dt}x(t)^{y(t)}=\dot{y}(t)x(t)^{y(t)}\log x(t)+y(t)\dot{x}(t)x(t)^{y(t)-1}
$$

for C^1 functions $x(t) > 0$ and $y(t)$. As shown in the proof of Lemma 4.9,

$$\begin{aligned}\sigma r(t)^2 \int_{\mathbb{T}^d} |\mu^t|^{r(t)-1} \operatorname{sgn}(\mu^t)\Delta\mu^t \, dx \\ - r(t)^2 \int_{\mathbb{T}^d} |\mu^t|^{r(t)-1} \operatorname{sgn}(\mu^t) \operatorname{div}(\mu^t \mathbb{M}\nabla \mathrm{g} * \mu^t)\, dx \\ \le -4\sigma(r(t)-1) \int_{\mathbb{T}^d} |\nabla|\mu^t|^{r(t)/2}|^2 \, dx.\end{aligned}$$

Hence,

$$\begin{aligned} r(t)^2 \|\mu^t\|_{L^{r(t)}}^{r(t)-1} \frac{d}{dt}\|\mu^t\|_{L^{r(t)}} \le \dot r(t) \int_{\mathbb{T}^d} |\mu^t|^{r(t)} \log\Big(\frac{|\mu^t|^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}}\Big)\, dx \\ - 4\sigma(r(t)-1) \int_{\mathbb{T}^d} |\nabla|\mu^t|^{r(t)/2}|^2 \, dx. \end{aligned} \tag{4.28}$$

We apply Lemma 4.12 to the right-hand side of inequality (4.28) with choice $a = \frac{8\sigma(r(t)-1)}{\dot r(t) d C_{\mathrm{LS},1}}$ and $f = |\mu^t|^{r(t)/2}$ to obtain that

$$\begin{aligned} r(t)^2 \|\mu^t\|_{L^{r(t)}}^{r(t)-1} \frac{d}{dt}\|\mu^t\|_{L^{r(t)}} \le -\dot r(t) \frac{d}{2} \log\Big(\frac{8\sigma(r(t)-1)}{\dot r(t) d C_{\mathrm{LS},1}}\Big) \|\mu^t\|_{L^{r(t)}}^{r(t)} \\ + \frac{4\sigma C_{\mathrm{LS},2}(r(t)-1)}{C_{\mathrm{LS},1}} \|\mu^t\|_{L^{r(t)/2}}^{r(t)}, \end{aligned} \tag{4.29}$$

with the implicit understanding that $\dot r(t) > 0$ (i.e., r is strictly increasing). Define

$$G(t) := \log(\|\mu^t\|_{L^{r(t)}}).$$

Then it follows from (4.29) that

$$\begin{aligned} \frac{d}{dt} G(t) = \frac{1}{\|\mu^t\|_{L^{r(t)}}} \frac{d}{dt}\|\mu^t\|_{L^{r(t)}} \le -\frac{\dot r(t)}{r(t)^2} \frac{d}{2} \log\Big(\frac{8\sigma(r(t)-1)}{\dot r(t) d C_{\mathrm{LS},1}}\Big) \\ + \frac{4\sigma C_{\mathrm{LS},2}(r(t)-1)}{C_{\mathrm{LS},1} r(t)^2} \frac{\|\mu^t\|_{L^{r(t)/2}}^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}}. \end{aligned} \tag{4.30}$$

By Hölder's inequality, $\frac{\|\mu^t\|_{L^{r(t)/2}}}{\|\mu^t\|_{L^{r(t)}}} \le 1$. Setting $s(t) := 1/r(t)$ and writing $\frac{r-1}{\dot r} = -\frac{s(1-s)}{\dot s}$, we find from (4.30) that

$$\begin{aligned} \frac{d}{dt} G(t) \le \dot s(t) \Big(\frac{d}{2} \log\Big(\frac{8\sigma}{d C_{\mathrm{LS},1}} s(t)(1-s(t))\Big)\Big) + \frac{d}{2}(-\dot s(t)) \log(-\dot s(t)) \\ + \frac{4\sigma C_{\mathrm{LS},2}}{C_{\mathrm{LS},1}} (s(t) - s(t)^2). \end{aligned}$$

So by the fundamental theorem of calculus,

$$\begin{aligned} G(T) - G(0) \le \int_0^T \dot s(t) \frac{d}{2} \log\Big(\frac{8\sigma}{d C_{\mathrm{LS},1}} s(t)(1-s(t))\Big)\, dt \\ - \frac{d}{2} \int_0^T \dot s(t) \log(-\dot s(t))\, dt + \frac{4\sigma C_{\mathrm{LS},2}}{C_{\mathrm{LS},1}} \int_0^T (s(t) - s(t)^2)\, dt. \end{aligned} \tag{4.31}$$

We require that $s(0) = 1/p$ and $s(T) = 1/q$, so by the fundamental theorem of calculus,

$$\begin{aligned}\int_0^T &\dot{s}(t)\frac{d}{2}\log\Big(\frac{8\sigma}{dC_{\mathrm{LS},1}}s(t)(1-s(t))\Big)\,dt\\ &= \frac{d}{2}\Big(\log\Big(\frac{8\sigma}{dC_{\mathrm{LS},1}}\Big)s + \log(s^s(1-s)^{-(1-s)}) - 2s\Big)\Big|_{s=1/p}^{s=1/q}.\end{aligned}$$

Using the convexity of $a \mapsto a\log a$, we minimize the second integral in the right-hand side of (4.31) by choosing $s(t)$ to linearly interpolate between $s(0) = 1/p$ and $s(T) = 1/q$, i.e.,

$$\dot{s}(t) = \frac{1}{T}\Big(\frac{1}{q} - \frac{1}{p}\Big), \quad 0 \le t \le T.$$

Thus, by the fundamental theorem of calculus,

$$-\frac{d}{2}\int_0^T \dot{s}(t)\log(-\dot{s}(t))\,dt = -\frac{d}{2}\Big(\frac{1}{p} - \frac{1}{q}\Big)\log\Big(\frac{T}{1/p - 1/q}\Big)$$

and

$$\begin{aligned}\frac{4\sigma C_{\mathrm{LS},2}}{C_{\mathrm{LS},1}}\int_0^T (s(t) - s(t)^2)\,dt &= \Big(\frac{1}{T}\Big(\frac{1}{q} - \frac{1}{p}\Big)\Big)^{-1}\frac{4\sigma C_{\mathrm{LS},2}}{C_{\mathrm{LS},1}}\int_0^T \dot{s}(t)(s(t) - s(t)^2)\,dt\\ &= \Big(\frac{1}{T}\Big(\frac{1}{q} - \frac{1}{p}\Big)\Big)^{-1}\frac{4\sigma C_{\mathrm{LS},2}}{C_{\mathrm{LS},1}}\Big(\frac{s^2}{2} - \frac{s^3}{3}\Big)\Big|_{s=1/p}^{s=1/q}.\end{aligned}$$

The desired conclusion now follows from a little bookkeeping and exponentiating both sides of the inequality. ■

Combining Lemma 4.13 with Lemma 4.9, we obtain Corollary 2.4.

Proof of Corollary 2.4. If $\sigma t \le 1$, then by Lemma 4.13,

$$\begin{aligned}\|\mu^t\|_{L^q} &\le C_{p,q,d}(\sigma(t/2))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{C_{p,q,d}\sigma(t/2)}\|\mu^{t/2}\|_{L^p}\\ &\le C_{p,q,d}(\sigma(t/2))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{C_{p,q,d}\sigma(t/2)}\|\mu^0\|_{L^p},\end{aligned}$$

where the final inequality is by Lemma 4.9. If $\sigma t > 1$, then by time-translation invariance of solutions and Lemma 4.13 again,

$$\|\mu^t\|_{L^q} \le C_{p,q,d}(\sigma(1/2\sigma))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{C_{p,q,d}\sigma(1/2\sigma)}\|\mu^{t-\frac{1}{2\sigma}}\|_{L^p} \le C'_{p,q,d}\|\mu^0\|_{L^p}.$$

In the conservative case, we may also obtain from combining Lemmas 4.9 and 4.13,

$$\begin{aligned}\|\mu^t - 1\|_{L^q} &\le C_{p,q,d}\Big(\sigma\min\Big(\frac{t}{2}, \frac{1}{2\sigma}\Big)\Big)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{C_{p,q,d}\sigma\min(\frac{t}{2},\frac{1}{2\sigma})}\|\mu^{\max(\frac{t}{2},t-\frac{1}{2\sigma})} - 1\|_{L^p}\\ &\le C'_{p,q,d}(\min(\sigma t, 1))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}e^{-C_{p,d}\sigma\max(\frac{t}{2},t-\frac{1}{2\sigma})}\|\mu^0 - 1\|_{L^p},\end{aligned}$$

provided that $p < \infty$. This then completes the proof of the corollary. ■

Lastly, we prove Corollary 2.5.

Proof of Corollary 2.5. The argument exploits the mild formulation of equation (4.2), together with the smoothing properties of the heat kernel. Arguing similarly to the proof of Proposition 2.1, for any $1 \le q \le \infty$, we have

$$\begin{aligned}\|\mu^t - 1\|_{L^q} &\le \|e^{t\sigma\Delta}(\mu^0 - 1)\|_{L^q} + \int_0^t \|e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^q}\, d\tau \\ &\le C_{d,p,q}(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^0 - 1\|_{L^p} \\ &\quad + C_{d,p,q}\int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^\tau\|_{L^{p_1}}\|\mathbb{M}\nabla\mathsf{g} * (\mu^\tau - 1)\|_{L^{p_2}}\, d\tau,\end{aligned}$$

where $\frac{d}{2}(\frac{1}{p} - \frac{1}{q}) < \frac{1}{2}$ and $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$. Using Corollary 2.4 on $\|\mu^\tau\|_{L^{p_1}}$ and Young's inequality/boundedness of Riesz transforms (assuming $p_2 < \infty$) on $\|\mathbb{M}\nabla\mathsf{g} * (\mu^\tau - 1)\|_{L^{p_2}}$, we find

$$\begin{aligned}&\int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^\tau\|_{L^{p_1}}\|\mathbb{M}\nabla\mathsf{g} * (\mu^\tau - 1)\|_{L^{p_2}}\, d\tau \\ &\quad \lesssim \int_0^t (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})} \min(\sigma\tau, 1)^{-\frac{d}{2}(\frac{1}{r_1}-\frac{1}{p_1})}\|\mu^0\|_{L^{r_1}} \sup_{0<\tau\le t}\|\mu^\tau - 1\|_{L^{p_2}}\, d\tau,\end{aligned}$$

where r_1 is chosen so that $\frac{d}{2}(\frac{1}{r_1} - \frac{1}{p}) \le \frac{1}{2}$. Choose $p_1 = \infty$ and $p_2 = p$. If $\sigma t \le 1$, then by rescaling time, the integral in the last line becomes

$$\sigma^{-1}(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^0\|_{L^d} \sup_{0<\tau\le t}\|\mu^\tau - 1\|_{L^p}\int_0^1 (1-\tau)^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\tau^{-\frac{1}{2}}\, d\tau,$$

which implies that

$$\|\mu^t - 1\|_{L^q} \le C_{p,q,d}\,\sigma^{-1}(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})}\|\mu^0\|_{L^d}\sup_{0<\tau\le t}\|\mu^\tau - 1\|_{L^p}. \tag{4.32}$$

Using time translation and an iteration argument, the L^p norm in the right-hand side can be reduced to an L^1 norm, to which we can applying Lemma 2.3.

More precisely, suppose that $q = \infty$. Fix a time t_0, and fix a step size $\sigma\kappa \le \min(\frac{1}{2(d+1)}, \frac{\sigma t_0}{2(d+1)})$. We choose a sequence of exponents

$$\infty = p_0 > p_1 > \cdots > p_d > p_{d+1} = 1,$$

such that $\frac{d}{2}(\frac{1}{p_{i+1}} - \frac{1}{p_i}) < \frac{1}{2}$. By translation of the initial time, we observe from (4.32) that for any $t_0 - i\kappa \le t \le t_0$,

$$\|\mu^t - 1\|_{L^{p_i}} \le C_d\,\sigma^{-1}(\sigma\kappa)^{-\frac{d}{2}(\frac{1}{p_{i+1}}-\frac{1}{p_i})}\|\mu^{t-(d+1)\kappa}\|_{L^d}\sup_{t-(i+1)\kappa<\tau\le t}\|\mu^\tau - 1\|_{L^{p_{i+1}}}.$$

Implicitly, we have used that $\|\mu^\tau\|_{L^d}$ is nonincreasing. This implies

$$\|\mu^t - 1\|_{L^\infty} \le C_d \sigma^{-d-1}(\sigma\kappa)^{-\frac{d}{2}} \|\mu^{t-(d+1)\kappa}\|_{L^d}^{d+1} \sup_{t-(d+1)\kappa<\tau\le t} \|\mu^\tau - 1\|_{L^1}.$$

Applying Lemma 4.13 to $\|\mu^{t-(d+1)\kappa}\|_{L^d}$ to go from L^d to L^1 and applying Lemma 2.3 to $\|\mu^\tau - 1\|_{L^1}$, the preceding right-hand side is

$$\le C_d \sigma^{-d-\frac{3}{2}} (\min(\sigma t, 1))^{-\frac{(d^2+d-1)}{2}} e^{-c_d \sigma t} \sqrt{\mathcal{F}_\sigma(\mu^0)}$$

for constants $C_d, c_d > 0$ depending on d. Note we have implicitly used that $\|\mu^0\|_{L^1} = 1$. This now completes the proof. ∎

5. Derivative decay estimates for the mean-field equation

In this section we prove Proposition 2.6 on the exponential rate of decay as $t \to \infty$ for the L^q norms of derivatives (of arbitrarily large order) of solutions to equation (1.4), which we know are global by Proposition 2.2. In particular, we show that solutions are smooth for $t > 0$. We assume throughout this section that $\int_{\mathbb{T}^d} \mu^0 = 1$ and that $\mu^0 \ge 0$ if $\mathbb{M} = -\mathbb{I}$.

Before starting the proof, let us record some remarks about the statement of Proposition 2.6

Remark 5.1. The constants and functions in the statement of the proposition additionally may depend on d, s, $\mathbb{M}$. One may extract a more explicit dependence of $\mathbf{W}_{\alpha,q}$, $\mathbf{W}_{n,q}$ on their arguments from the proof of Proposition 2.6; but we do not find it enlightening and so do not present it. We only remark that $\mathbf{W}_{\alpha,q}$, $\mathbf{W}_{n,q}$ do not depend on the argument $\mathcal{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric (conservative case).

Remark 5.2. The exponent λ_2 is chosen so that $\|\nabla \mathsf{g} * \mu\|_{L^2} \lesssim \|\mu\|_{\dot{H}^{\lambda_2}}$. See Remark 5.5 below for further motivation.

Remark 5.3. By using the time-translation trick, one can apply Proposition 2.6, going from $\tau = t$ to $\tau = t_0 := \min(\frac{t}{2}, t - \frac{1}{2\sigma})$, leading to the argument $\|\mu^{t_0}\|_{L^\infty}$ in $\mathbf{W}_{\alpha,q}$, $\mathbf{W}_{n,q}$. One can then use Corollary 2.4 to eliminate the norm $\|\mu^{t_0}\|_{L^\infty}$ at the cost of additional factors of $(\sigma t)^{-1}$. Since this does not help us – and we know that $\mu^t \in L^\infty$ for any $t > 0$ automatically – we have chosen to keep the L^∞ dependence for simplicity.

So as to make the presentation easier to digest, we break the proof of Proposition 2.6 into a series of lemmas, which are proved in the upcoming two subsections. In Section 5.1 we treat the range $d-2 \le s \le d-1$, showing (2.7), (2.8). Then in Section 5.2 we treat the harder, remaining range $d-1 < s < d$, showing (2.9), (2.10). This then completes the proof of Proposition 2.6 and, together with the results of Sections 4.3 and 4.4, establishes assertion (1.9) from Theorem 1.1.

5.1. The case $d-2 \leq s \leq d-1$

We begin with the temporal decay estimates for the L^p norms of the derivatives of μ^t (note $\nabla^{\otimes n}\mu^t$ has zero average for $n \geq 1$ and similarly for $|\nabla|^\alpha \mu^t$ if $\alpha > 0$) in the easier case $d-2 \leq s \leq d-1$. The first step is to prove Lemma 2.7 on the short-time gain of regularity.

Proof of Lemma 2.7. Our starting point is the following identity, which follows from commutativity of Fourier multipliers:

$$\partial_\alpha \mu^t = e^{t\sigma\Delta}\partial_\alpha \mu^0 - \int_0^t e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\alpha (\mu^\tau \mathbb{M}\nabla \mathrm{g} * \mu^\tau)\, d\tau, \tag{5.1}$$

where $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ is a multi-index of order $|\alpha| = 1$. The general case $|\alpha| \geq 1$ will be handled by induction. As the heat kernel is singular at $\tau = t$, we divide the integration over $[0,t]$ into $[0,(1-\varepsilon)t]$ and $[(1-\varepsilon)t, t]$, for some $\varepsilon \in (0,1)$ to be determined. Applying the triangle and Minkowski inequalities to the right-hand side of (5.1) leads us to

$$\begin{aligned}\|\partial_\alpha \mu^t\|_{L^q} \leq \|e^{t\sigma\Delta}\partial_\alpha\mu^0\|_{L^q} &+ \int_0^{t(1-\varepsilon)} \|e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\alpha(\mu^\tau \mathbb{M}\nabla \mathrm{g} * \mu^\tau)\|_{L^q}\, d\tau \\ &+ \int_{t(1-\varepsilon)}^t \|e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\alpha(\mu^\tau \mathbb{M}\nabla \mathrm{g} * \mu^\tau)\|_{L^q}\, d\tau.\end{aligned} \tag{5.2}$$

We respectively denote by $J_1(t)$, $J_2(t)$, $J_3(t)$ the three terms in the right-hand side of the previous inequality and proceed to estimate each of them individually.

First, $J_1(t)$ is a consequence of the heat kernel estimate (3.3) and Young's inequality:

$$J_1(t) \lesssim (\min(\sigma t, 1))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1}{2}} e^{-C\sigma t}\|\mu^0 - 1\|_{L^p}, \tag{5.3}$$

for any $1 \leq p \leq q \leq \infty$.

Now consider $J_2(t)$. By (3.4) and Hölder's inequality, we have for any $1 \leq p \leq q \leq \infty$,

$$\begin{aligned}&\|\partial_\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathrm{g} * \mu^\tau)\|_{L^q} \\ &\quad\lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau), 1)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} \|\mu^\tau \mathbb{M}\nabla \mathrm{g} * \mu^\tau\|_{L^p} \\ &\quad\lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau), 1)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} \|\mu^\tau\|_{L^\infty} \|\mathbb{M}\nabla \mathrm{g} * \mu^\tau\|_{L^p} \\ &\quad\lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau), 1)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} \\ &\qquad \times (\min(\sigma\tau, 1))^{-\frac{1}{2}} \|\mu^0\|_{L^d} \|\mathbb{M}\nabla \mathrm{g} * \mu^\tau\|_{L^p},\end{aligned} \tag{5.4}$$

where we applied Corollary 2.4 to $\|\mu^\tau\|_{L^\infty}$ to obtain the last line. If $p = \infty$ (and so $q = \infty$ as well) and $s = d-1$, the Riesz transform $\nabla \mathrm{g}*$ is not bounded on L^p and so we will be out of luck in trying to estimate $\|\mathbb{M}\nabla \mathrm{g} * \mu^t\|_{L^p}$ in terms of μ^t. Instead, we modify the preceding argument to obtain that for any $d < r < \infty$,

$$\begin{aligned}
&\|\partial_\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^\infty} \\
&\quad\lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau),1)^{-\frac{d}{2r}-1} \\
&\qquad \times (\min(\sigma\tau,1))^{-\frac{1}{2}+\frac{d}{2r}} \|\mu^0\|_{L^{\frac{rd}{r-d}}} \|\mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^r}.
\end{aligned} \tag{5.5}$$

Combining estimates (5.4) and (5.5), we have shown that for any $1 \le p \le q \le \infty$ and $d < r < \infty$,

$$\begin{aligned}
&\|\partial_\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^q} \\
&\quad\lesssim \min(\sigma(t-\tau),1)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} e^{-C\sigma(t-\tau)} (\min(\sigma\tau,1))^{-\frac{1}{2}} \\
&\qquad \times \Big(\|\mu^0\|_{L^d} \|\mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^p} \mathbf{1}_{\substack{s<d-1 \\ s=d-1\wedge p<\infty}} \\
&\qquad\quad + \min(\sigma(t-\tau),1)^{-\frac{d}{2r}} (\min(\sigma\tau,1))^{\frac{d}{2r}} \|\mu^0\|_{L^{\frac{rd}{r-d}}} \|\mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^r} \mathbf{1}_{s=d-1\wedge p=\infty} \Big).
\end{aligned}$$

We then use (2.6) from Corollary 2.4 if $\mathbb{M}$ is antisymmetric, or (2.4) from Lemma 2.3 if $\mathbb{M} = -\mathbb{I}$, to bound for $1 \le r < \infty$ (same for r replaced by p),

$$\begin{aligned}
\|\mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^r} &\lesssim \|\mu^\tau - 1\|_{L^r} \\
&\lesssim e^{-C'\sigma\tau} \|\mu^0 - 1\|_{L^r} \mathbf{1}_{\mathbb{M}\text{ a.s.}} \\
&\quad + (1 + \|\mu^0\|_{L^\infty})^{1-\frac{1}{r}} \Big(e^{-C''\sigma\tau} \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma} \Big)^{\frac{1}{r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}}.
\end{aligned}$$

Since $\tau \le t$ and $\sigma t \le 1$ by assumption, we can drop the $\min(\cdot)$ and exponential factors above. It now follows that

$$\begin{aligned}
&\|\partial_\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^q} \\
&\quad\lesssim (\sigma(t-\tau))^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} (\sigma\tau)^{-\frac{1}{2}} \\
&\qquad \times \bigg(\|\mu^0\|_{L^d} \Big(\|\mu^0 - 1\|_{L^p} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + (1 + \|\mu^0\|_{L^\infty})^{1-\frac{1}{p}} \Big(\sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma} \Big)^{\frac{1}{p}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \Big) \mathbf{1}_{p<\infty} \\
&\qquad\quad + \|\mu^0\|_{L^{d+}} \Big(\|\mu^0 - 1\|_{L^r} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + (1 + \|\mu^0\|_{L^\infty})^{1-\frac{1}{r}} \Big(\sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma} \Big)^{\frac{1}{r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \Big) \\
&\qquad\quad \times \left(\frac{\tau}{(t-\tau)} \right)^{\frac{d}{2r}} \mathbf{1}_{p=\infty} \bigg).
\end{aligned}$$

Adjusting C, C', $\frac{C''}{r}$, $\frac{C''}{p}$ if necessary, we may assume without loss of generality that $C = C' = \frac{C''}{r} = \frac{C''}{p}$. Recalling the definition of $J_2(t)$ from (5.2) and using the dilation invariance of Lebesgue measure, we arrive at

$$\begin{aligned}
J_2(t) &\lesssim \frac{A_\varepsilon}{\sigma} \|\mu^0\|_{L^{d*}} (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1}{2}} \\
&\quad \times \bigg(\Big(\|\mu^0 - 1\|_{L^p} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \Big) \mathbf{1}_{p<\infty} \\
&\qquad + \Big(\|\mu^t - 1\|_{L^r} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{r}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \Big) \mathbf{1}_{p=\infty} \bigg),
\end{aligned} \tag{5.6}$$

where $d^* := d^+\mathbf{1}_{s=d-1\wedge p=\infty} + d(1-\mathbf{1}_{s=d-1\wedge p=\infty})$, $d < r < \infty$ is arbitrary, and

$$A_\varepsilon := \int_0^{1-\varepsilon} (1-\tau)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1}\tau^{-\frac{1}{2}}\Big(\mathbf{1}_{p<\infty} + \Big(\frac{\tau}{(1-\tau)}\Big)^{\frac{d}{2r}}\mathbf{1}_{p=\infty}\Big)\,d\tau. \tag{5.7}$$

Finally, for $J_3(t)$, we have by the product rule and triangle inequality,

$$\begin{aligned}&\|\partial_\alpha e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau)\|_{L^q}\\&\quad\lesssim e^{-C\sigma(t-\tau)}\min(\sigma(t-\tau),1)^{-\frac{1}{2}}(\|\partial_\alpha\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau\|_{L^q}+\|\mu^\tau\mathbb{M}\nabla\mathsf{g}*\partial_\alpha\mu^\tau\|_{L^q}).\end{aligned}$$

If $s < d-1$, then since $\nabla\mathsf{g} \in L^1$, we obtain from application of estimate (2.5) of Corollary 2.4 to $\|\mu^\tau\|_{L^\infty}$,

$$\begin{aligned}&\|\partial_\alpha e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau)\|_{L^q}\\&\quad\lesssim e^{-C\sigma(t-\tau)}\min(\sigma(t-\tau),1)^{-\frac{1}{2}}\|\mu^\tau\|_{L^\infty}\|\partial_\alpha\mu^\tau\|_{L^q}\\&\quad\lesssim e^{-C\sigma(t-\tau)}\min(\sigma(t-\tau),1)^{-\frac{1}{2}}\min(\sigma\tau,1)^{-\frac{1}{2}}\|\mu^0\|_{L^d}\|\partial_\alpha\mu^\tau\|_{L^q}.\end{aligned} \tag{5.8}$$

If $s = d-1$ and $q > 1$, we modify the argument (to account for $\nabla\mathsf{g} \notin L^1$) to obtain, for any $1 < r < q \le \infty$,

$$\begin{aligned}\|\partial_\alpha e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau)\|_{L^q} &\lesssim e^{-C\sigma(t-\tau)}\min(\sigma(t-\tau),1)^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{r}-\frac{1}{q})}\\&\quad\times(\|\mu^\tau\mathbb{M}\nabla\mathsf{g}*\partial_\alpha\mu^\tau\|_{L^r}+\|\partial_\alpha\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau\|_{L^r}).\end{aligned}$$

We have by Hölder's and Young's inequalities,

$$\begin{aligned}\|\mu^\tau\mathbb{M}\nabla\mathsf{g}*\partial_\alpha\mu^\tau\|_{L^r} &\lesssim \|\mathbb{M}\nabla\mathsf{g}*\partial_\alpha\mu^\tau\|_{L^r}\|\mu^\tau\|_{L^\infty}\\&\lesssim \|\partial_\alpha\mu^\tau\|_{L^r}\|\mu^\tau\|_{L^\infty}\\&\le \|\partial_\alpha\mu^\tau\|_{L^q}\|\mu^\tau\|_{L^\infty}.\end{aligned} \tag{5.9}$$

Again, by Hölder's inequality plus the boundedness of the Riesz transform,

$$\begin{aligned}\|\partial_\alpha\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau\|_{L^r} &\le \|\partial_\alpha\mu^\tau\|_{L^q}\|\mathbb{M}\nabla\mathsf{g}*\mu^\tau\|_{L^{\frac{rq}{q-r}}}\\&\lesssim \|\partial_\alpha\mu^\tau\|_{L^q}\|\mu^\tau-1\|_{L^{\frac{rq}{q-r}}}.\end{aligned} \tag{5.10}$$

Taking r close enough to q so that $\frac{d}{r}-\frac{d}{q} < 1$, it follows from (5.9), (5.10), and application of (2.5) to $\|\mu^\tau\|_{L^\infty}$,

$$\begin{aligned}&\|\partial_\alpha e^{\sigma(t-\tau)\Delta}\operatorname{div}(\mu^\tau\mathbb{M}\nabla\mathsf{g}*\mu^\tau)\|_{L^q}\\&\quad\lesssim e^{-C\sigma(t-\tau)}\min(\sigma(t-\tau),1)^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{r}-\frac{1}{q})}\\&\qquad\times\min(\sigma\tau,1)^{-\frac{1}{2}+\frac{d}{2}(\frac{1}{r}-\frac{1}{q})}\|\mu^0\|_{L^{\frac{d}{1-d(\frac{1}{r}-\frac{1}{q})}}}\|\partial_\alpha\mu^\tau\|_{L^q}.\end{aligned} \tag{5.11}$$

Combining estimates (5.8) and (5.11) and dropping the $\min(\cdot)$ and exponential factors using the assumption $\sigma t \leq 1$, we arrive at

$$J_3(t) \lesssim \|\mu^0\|_{L^{d*}} \int_{t(1-\varepsilon)}^t (\sigma(t-\tau))^{-\frac{1}{2}} (\sigma\tau)^{-\frac{1}{2}} \|\partial_\alpha \mu^\tau\|_{L^q} \times \left(\mathbf{1}_{s<d-1} + \left(\frac{\tau}{(t-\tau)}\right)^{0+} \mathbf{1}_{s=d-1,q>1}\right) d\tau. \tag{5.12}$$

Combining estimates (5.3), (5.6), (5.12), we obtain

$$\begin{aligned}
\|\partial_\alpha \mu^t\|_{L^q} &\lesssim \|\mu^0 - 1\|_{L^p} (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1}{2}} \\
&+ \frac{A_\varepsilon}{\sigma} \|\mu^0\|_{L^{d*}} (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1}{2}} \\
&\times \Big(\big(\|\mu^0 - 1\|_{L^p} \mathbf{1}_{\mathbb{M} \text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \big) \mathbf{1}_{p<\infty} \\
&\quad + \big(\|\mu^0 - 1\|_{L^r} \mathbf{1}_{\mathbb{M} \text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{r}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \big) \mathbf{1}_{p=\infty} \Big) \\
&+ \|\mu^0\|_{L^{d*}} \int_{t(1-\varepsilon)}^t (\sigma(t-\tau))^{-\frac{1}{2}} (\sigma\tau)^{-\frac{1}{2}} \|\partial_\alpha \mu^\tau\|_{L^q} \\
&\times \left(\mathbf{1}_{s<d-1} + \left(\frac{\tau}{(t-\tau)}\right)^{0+} \mathbf{1}_{s=d-1,q>1}\right) d\tau.
\end{aligned} \tag{5.13}$$

To close the estimate for $\|\partial_\alpha \mu^t\|_{L^q}$, we define the function (for $0 < t \leq \sigma^{-1}$)

$$\phi(t) := \sup_{0<\tau\leq t} (\sigma\tau)^{\frac{d}{2}(\frac{1}{p}-\frac{1}{q})+\frac{1}{2}} \|\partial_\alpha \mu^\tau\|_{L^q}.$$

Using this notation, we rearrange (5.13), using the dilation invariance of Lebesgue measure, to obtain the inequality

$$\begin{aligned}
\phi(t) &\leq C\|\mu^0 - 1\|_{L^p} \\
&+ \frac{CA_\varepsilon \|\mu^0\|_{L^{d*}}}{\sigma} \Big(\big(\|\mu^0 - 1\|_{L^p} \mathbf{1}_{\mathbb{M} \text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \big) \mathbf{1}_{p<\infty} \\
&\quad + \big(\|\mu^0 - 1\|_{L^r} \mathbf{1}_{\mathbb{M} \text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{r}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}} \big) \mathbf{1}_{p=\infty} \Big) \\
&+ \frac{CB_\varepsilon \|\mu^0\|_{L^{d*}}}{\sigma} \phi(t),
\end{aligned} \tag{5.14}$$

where $C > 0$ depends only on d, s, p, q, and

$$B_\varepsilon := \int_{1-\varepsilon}^1 (1-\tau)^{-\frac{1}{2}} \tau^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-1} \left(\mathbf{1}_{s<d-1} + \left(\frac{\tau}{1-\tau}\right)^{0+} \mathbf{1}_{s=d-1,q>1}\right) d\tau.$$

The fact that we did not pick up any factors of t in (5.14) precisely explains our choice of the exponents in the factors $(t-\tau)$ and τ above. Since the integral in the definition of B_ε

decreases monotonically to zero as $\varepsilon \to 1^-$, we may choose ε sufficiently small so that $C\|\mu^0\|_{L^{d*}} B_\varepsilon \le \frac{\sigma}{2}$. Thus,

$$\begin{aligned}\|\partial_\alpha \mu^t\|_{L^q} \le \mathbf{W}_{1,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathscr{F}_\sigma(\mu^0))(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1}{2}} \\ \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\ + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big), \quad (5.15)\end{aligned}$$

for any $\epsilon \in (0, d^{-1})$, where $\mathbf{W}_{1,p,q}$ is a continuous, nondecreasing, polynomial function of its arguments. Furthermore, $\mathbf{W}_{1,p,q}$ does not depend on $\mathscr{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric.

Let us now bootstrap from the case $|\alpha| = 1$ to the general case $n = |\alpha| \ge 1$. As our induction hypothesis, assume that, for all $|\beta| \le n-1$, $t \in (0, \sigma^{-1}]$, $1 \le p \le q \le \infty$,

$$\begin{aligned}\|\partial_\beta \mu^t\|_{L^q} \le \mathbf{W}_{|\beta|,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathscr{F}_\sigma(\mu^0))(\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{|\beta|}{2}} \\ \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\ + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big), \quad (5.16)\end{aligned}$$

where $\varepsilon \in (0, d^{-1})$ and $\mathbf{W}_{|\beta|,p,q}$ is a continuous, nondecreasing, polynomial function of its arguments. Analogously to (5.2), we have

$$\begin{aligned}\|\partial_\alpha \mu^t\|_{L^q} \le \|e^{\sigma t\Delta}\partial_\alpha \mu^0\|_{L^q} + \int_0^{t(1-\varepsilon)} \|e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\alpha(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^q}\, d\tau \\ + \int_{t(1-\varepsilon)}^t \|e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\alpha(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^q}\, d\tau.\end{aligned}$$

Repeating the arguments for $J_1(t)$ and $J_2(t)$ above, we have

$$\begin{aligned}J_1(t) &\lesssim (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{n}{2}} \|\mu^0 - 1\|_{L^p}, \\ J_2(t) &\lesssim \frac{A_\varepsilon \|\mu^0\|_{L^{d*}}}{\sigma} (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{n}{2}} \\ &\quad \times \Big(\big(\|\mu^0 - 1\|_{L^p} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}}\big)\mathbf{1}_{p<\infty} \\ &\qquad + \big(\|\mu^0 - 1\|_{L^r} \mathbf{1}_{\mathbb{M}\text{ a.s.}} + \|\mu^0\|_{L^\infty}^{1-\frac{1}{r}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2r}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}}\big)\mathbf{1}_{p=\infty}\Big),\end{aligned}$$

where now

$$A_\varepsilon := \int_0^{1-\varepsilon} (1-\tau)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{|\alpha|+1}{2}} \tau^{-\frac{1}{2}} \Big(\mathbf{1}_{p<\infty} + \Big(\frac{\tau}{1-\tau}\Big)^{0+} \mathbf{1}_{p=\infty}\Big)\, d\tau.$$

For J_3, we apply the Leibniz rule,

$$\partial_\alpha(\mu \mathbb{M}\nabla \mathsf{g} * \mu) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial_\beta \mu \mathbb{M}\nabla \mathsf{g} * \partial_{\alpha-\beta}\mu,$$

and note that estimates (5.8), (5.9), (5.10) also hold for the $\beta = 0$, $\beta = \alpha$ terms. For the terms with $\beta \notin \{0, \alpha\}$, we use the induction hypothesis (5.16). If $s < d-1$ or $s = d-1$ and $q \notin \{1, \infty\}$,

$$
\begin{aligned}
&\int_{t(1-\varepsilon)}^{t} \|e^{\sigma(t-\tau)\Delta} \operatorname{div} \partial_\beta \mu^\tau \mathbb{M}\nabla \mathsf{g} * \partial_{\alpha-\beta}\mu^\tau\|_{L^q}\, d\tau \\
&\quad\lesssim \int_{t(1-\varepsilon)}^{t} (\sigma(t-\tau))^{-\frac{1}{2}} \|\partial_\beta \mu^\tau\|_{L^\infty} \|\partial_{\alpha-\beta}\mu^\tau\|_{L^q}\, d\tau \\
&\quad\lesssim \mathbf{W}_{|\beta|,d,\infty}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^d}, \mathcal{F}_\sigma(\mu^0)) \\
&\qquad\times \mathbf{W}_{n-|\beta|,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathcal{F}_\sigma(\mu^0)) \\
&\qquad\times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\qquad\times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_{\epsilon'} \|\mu^0\|_{L^\infty}^{1-\epsilon'} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon'}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\qquad\times \int_{t(1-\varepsilon)}^{t} (\sigma(t-\tau))^{-\frac{1}{2}} (\sigma\tau)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q}) - \frac{|\alpha|+1}{2}}\, d\tau \\
&\quad\lesssim \sigma^{-1}\mathbf{W}_{|\beta|,d,\infty}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^d}, \mathcal{F}_\sigma(\mu^0)) \\
&\qquad\times \mathbf{W}_{n-|\beta|,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathcal{F}_\sigma(\mu^0)) \\
&\qquad\times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\qquad\times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_{\epsilon'} \|\mu^0\|_{L^\infty}^{1-\epsilon'} (\mathcal{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon'}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\qquad\times (\sigma t)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q}) - \frac{(n+1)}{2}} B_\varepsilon,
\end{aligned}
$$

where $\varepsilon, \varepsilon' \in (0, d^{-1})$ and

$$
B_\varepsilon := \int_{1-\varepsilon}^{1} (1-\tau)^{-\frac{1}{2}} \tau^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q}) - \frac{(n+1)}{2}}\, d\tau.
$$

If $s = d-1$ and $q = \infty$, then similarly to before, we argue

$$
\begin{aligned}
&\|e^{\sigma(t-\tau)\Delta} \operatorname{div}(\partial_\beta \mu^\tau \mathbb{M}\nabla \mathsf{g} * \partial_{\alpha-\beta}\mu^\tau)\|_{L^q} \\
&\quad\lesssim (\sigma(t-\tau))^{-\frac{1}{2}-\frac{d}{2}(\frac{1}{r}-\frac{1}{q})} \|\partial_\beta \mu^\tau\|_{L^q} \|\partial_{\alpha-\beta}\mu^\tau\|_{L^\infty}
\end{aligned}
$$

$$
\begin{aligned}
&\lesssim \mathbf{W}_{|\beta|,p,q}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^p}, \mathscr{F}_\sigma(\mu^0)) \\
&\quad \times \mathbf{W}_{n-|\beta|,d+,\infty}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^{d+}}, \mathscr{F}_\sigma(\mu^0)) \\
&\quad \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\quad \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_{\epsilon'} \|\mu^0\|_{L^\infty}^{1-\epsilon'} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon'}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\quad \times (\sigma(t-\tau))^{-\frac{1}{2}} (\sigma\tau)^{-\frac{|\alpha|+1}{2} - \frac{d}{2}(\frac{1}{p} - \frac{1}{q})} \Big(\frac{\tau}{t-\tau}\Big)^{0+}.
\end{aligned}
$$

Above, we have neglected the case $s = d - 1$ and $q = 1$, as our arguments do not work. However, we have by Hölder's inequality that, for any $r > 1$,

$$
\begin{aligned}
\|\partial_\alpha \mu^t\|_{L^1} &\leq \|\partial_\alpha \mu^t\|_{L^r} \\
&\lesssim \mathbf{W}_{|\alpha|,p,r}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^r}, \mathscr{F}_\sigma(\mu^0))(\sigma t)^{-\frac{d}{2}(1-\frac{1}{r}) - \frac{|\alpha|}{2}} \\
&\quad \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\lesssim \mathbf{W}_{|\alpha|,p,r}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^1}^{\frac{1}{r}} (1 + \|\mu^0\|_{L^\infty})^{1-\frac{1}{r}}, \mathscr{F}_\sigma(\mu^0)) \\
&\quad \times (\sigma t)^{-\frac{d}{2}(1-\frac{1}{r}) - \frac{|\alpha|}{2}} \\
&\quad \times \Big(\|\mu^0 - 1\|_{L^p} + \big(\|\mu^0\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^0\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^0)/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big).
\end{aligned}
$$

Combining the estimates above and following the same reasoning used to obtain (5.15) in the case $|\alpha| = 1$, one completes the proof of the induction step. Therefore, the proof of Lemma 2.7 is complete. ∎

We now combine Lemma 2.7 with Lemmas 4.9 and 2.3 to prove Lemma 2.8 on the long-time exponential decay of $\|\nabla^{\otimes n}\mu^t\|_{L^q}$. This then establishes estimate (2.8) of Proposition 2.6.

Proof of Lemma 2.8. Fix $t > 0$ and assume that $\sigma t > 1$ (otherwise, the desired result is covered by Lemma 2.7). Let $\sigma t_0 = \sigma t - \frac{1}{2}$. Translating time, we may apply Lemma 2.7 to obtain

$$
\begin{aligned}
\|\nabla^{\otimes n}\mu^t\|_{L^q} &\leq \mathbf{W}_{n,p,q}(\|\mu^{t_0}\|_{L^\infty}, \sigma^{-1}, \|\mu^{t_0} - 1\|_{L^p}, \mathscr{F}_\sigma(\mu^{t_0}))(\sigma(t-t_0))^{-\frac{n}{2} - \frac{d}{2}(\frac{1}{p} - \frac{1}{q})} \\
&\quad \times \Big(\|\mu^{t_0} - 1\|_{L^p} + \big(\|\mu^{t_0}\|_{L^\infty}^{1-\frac{1}{p}} (\mathscr{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{1}{2p}} \mathbf{1}_{p<\infty} \\
&\qquad\qquad + C_\epsilon \|\mu^{t_0}\|_{L^\infty}^{1-\epsilon} (\mathscr{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{\epsilon}{2}} \mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\quad \times (1 + C_\varepsilon(\sigma(t-t_0))^{-\varepsilon} \mathbf{1}_{s=d-1 \wedge q=1})
\end{aligned}
$$

$$
\begin{aligned}
&\le C\mathbf{W}_{n,p,q}(\|\mu^{t_0}\|_{L^\infty},\sigma^{-1},\|\mu^{t_0}-1\|_{L^p},\mathcal{F}_\sigma(\mu^{t_0}))(1+C'_\varepsilon\mathbf{1}_{s=d-1\wedge q=1}) \\
&\quad\times\Big(\|\mu^{t_0}-1\|_{L^p}+\big(\|\mu^{t_0}\|_{L^\infty}^{1-\frac{1}{p}}(\mathcal{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{1}{2p}}\mathbf{1}_{p<\infty} \\
&\qquad\qquad+C_\epsilon\|\mu^{t_0}\|_{L^\infty}^{1-\epsilon}(\mathcal{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{\epsilon}{2}}\mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big),
\end{aligned}
$$

where the second inequality follows from $\sigma(t-t_0)>\frac{1}{2}$. We have $\|\mu^{t_0}\|_{L^\infty}\le\|\mu^0\|_{L^\infty}$. Applying (4.18) from Lemma 4.9 to $\|\mu^{t_0}-1\|_{L^p}$ and (2.3) from Lemma 2.3 to $\mathcal{F}_\sigma(\mu^0)$, then using that $\mathbf{W}_{n,p,q}$ is nondecreasing in its arguments, we find

$$
\begin{aligned}
&\mathbf{W}_{n,p,q}(\|\mu^{t_0}\|_{L^\infty},\sigma^{-1},\|\mu^{t_0}-1\|_{L^p},\mathcal{F}_\sigma(\mu^{t_0})) \\
&\quad\times\Big(\|\mu^{t_0}-1\|_{L^p}+\big(\|\mu^{t_0}\|_{L^\infty}^{1-\frac{1}{p}}(\mathcal{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{1}{2p}}\mathbf{1}_{p<\infty} \\
&\qquad\qquad+C_\epsilon\|\mu^{t_0}\|_{L^\infty}^{1-\epsilon}(\mathcal{F}_\sigma(\mu^{t_0})/\sigma)^{\frac{\epsilon}{2}}\mathbf{1}_{p=\infty}\big)\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\le\widetilde{\mathbf{W}}_{n,p,q}(\|\mu^0\|_{L^\infty},\sigma^{-1},\|\mu^0-1\|_{L^p},\mathcal{F}_\sigma(\mu^0))e^{-C'\sigma t},
\end{aligned}
$$

where $C'>0$ and $\widetilde{\mathbf{W}}_{n,p,q}\colon[0,\infty)^4\to[0,\infty)$ is a continuous, nondecreasing, polynomial function of its arguments, which is independent of $\mathcal{F}_\sigma(\mu^0)$ if $\mathbb{M}$ is antisymmetric. This completes the proof. ∎

Remark 5.4. Using the fractional Leibniz rule in place of the ordinary Leibniz rule, one can adapt the proof of Lemma 2.7, then use the same argument as in the proof of Lemma 2.8, to also obtain, for any $\alpha>0$ and $\varepsilon>0$, for all $t>0$,

$$
\begin{aligned}
\||\nabla|^\alpha\mu^t\|_{L^q}&\le\mathbf{W}_{\alpha,p,q}(\|\mu^0\|_\infty,\sigma^{-1},\|\mu^0-1\|_{L^p},\mathcal{F}_\sigma(\mu^0))e^{-C\sigma t}\min(\sigma t,1)^{-\frac{\alpha}{2}-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})} \\
&\quad\times(1+C_\varepsilon\min(\sigma t,1)^{-\varepsilon}\mathbf{1}_{s=d-1\wedge q=1}),
\end{aligned}
\tag{5.17}
$$

where $\mathbf{W}_{\alpha,p,q}\colon[0,\infty)^4\to[0,\infty)$ is a function with the same properties as $\mathbf{W}_{n,p,q}$ above. This establishes estimate (2.7) of Proposition 2.6. Alternatively, one can obtain the decay estimate (5.17) following the proof of Lemma 2.10 presented in the next subsection.

5.2. The case $d-1<s<d$

Next we establish the analogue of Lemma 2.7 in the more difficult case $d-1<s<d$. Recall from above that the difficulty stems from the loss of regularity in the vector field $\mathbb{M}\nabla\mathsf{g}*\mu$, an issue we saw in the proof of Proposition 2.1 for local well-posedness.

As an intermediate step, we first prove Lemma 2.9 on the uniform-in-time bound for L^2 norms of (fractional) derivatives of μ^t, which by Sobolev embedding, will yield uniform-in-time control of the quantity $\||\nabla|^{s-d+1}\mu^\tau\|_{L^p}$, for any $1\le p\le\infty$, arising in estimation of the vector field $\mathbb{M}\nabla\mathsf{g}*\mu^\tau$.

Proof of Lemma 2.9. We may assume that $\mu^0\ne1$; otherwise, the left-hand side of (2.11) is identically zero and there is nothing to prove. By Remark 4.5, we may assume without loss of generality that μ is a classical solution. We have for $\alpha > 0$,

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\||\nabla|^\alpha\mu^t\|_{L^2}^2 &= \int_{\mathbb{T}^d}|\nabla|^\alpha\mu^t(\sigma\Delta|\nabla|^\alpha\mu^t - \operatorname{div}|\nabla|^\alpha(\mu^t\mathbb{M}\nabla\mathrm{g}*\mu^t))\,dx\\ &= -\sigma\int_{\mathbb{T}^d}|\nabla|\nabla|^\alpha\mu^t|^2\,dx\\ &\quad+\int_{\mathbb{T}^d}\nabla|\nabla|^\alpha\mu^t\cdot|\nabla|^\alpha(\mu^t\mathbb{M}\nabla\mathrm{g}*\mu^t)\,dx, \end{aligned}\tag{5.18}$$

where the ultimate equality follows from integration by parts.

Consider the second term in (5.18). By Cauchy–Schwarz and the fractional Leibniz rule (e.g., see [54, Theorem 7.6.1]), we have, for any exponent $2 \le p \le \infty$,

$$\begin{aligned}&\left|\int_{\mathbb{T}^d}\nabla|\nabla|^\alpha\mu^t\cdot|\nabla|^\alpha(\mu^t\mathbb{M}\nabla\mathrm{g}*\mu^t)\,dx\right|\\ &\quad\le\|\nabla|\nabla|^\alpha\mu^t\|_{L^2}\||\nabla|^\alpha(\mu^t\mathbb{M}\nabla\mathrm{g}*\mu^t)\|_{L^2}\\ &\quad\lesssim\|\nabla|\nabla|^\alpha\mu^t\|_{L^2}\big(\||\nabla|^\alpha\mu^t\|_{L^p}\|\mathbb{M}\nabla\mathrm{g}*\mu^t\|_{L^{\frac{2p}{p-2}}}\\ &\qquad\qquad+\|\mu^t\|_{L^\infty}\|\mathbb{M}\nabla\mathrm{g}*|\nabla|^\alpha\mu^t\|_{L^2}\big).\end{aligned}\tag{5.19}$$

We choose $p = \frac{2(1+\alpha)}{\alpha}$. Then, by the fractional Gagliardo–Nirenberg interpolation inequalities (e.g., see [6, Theorem 2.44]),

$$\begin{aligned}\||\nabla|^\alpha\mu^t\|_{L^p}&\lesssim\|\mu^t\|_{\dot H^{1+\alpha}}^{\frac{\alpha}{1+\alpha}}\|\mu^t-1\|_{L^\infty}^{\frac{1}{1+\alpha}},\\ \|\mathbb{M}\nabla\mathrm{g}*\mu^t\|_{L^{\frac{2p}{p-2}}}&\lesssim\|\mu^t\|_{\dot H^{1+\alpha}}^{\frac{s+1-d}{1+\alpha}}\|\mu^t-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{\alpha-s+d}{1+\alpha}},\end{aligned}$$

which allows one to handle the first product inside the parentheses in (5.19). For the second product, we trivially estimate

$$\|\mathbb{M}\nabla\mathrm{g}*|\nabla|^\alpha\mu^t\|_{L^2}\lesssim\|\mu^t\|_{\dot H^{1+\alpha+s-d}}\le\|\mu^t-1\|_{L^2}^{\frac{d-s}{1+\alpha}}\|\mu^t\|_{\dot H^{1+\alpha}}^{\frac{1+\alpha+s-d}{1+\alpha}}.$$

Combining the above estimates, we obtain

$$\begin{aligned}\frac{d}{dt}\|\mu^t\|_{\dot H^\alpha}^2&\le-\sigma\|\mu^t\|_{\dot H^{1+\alpha}}^2+C\|\mu^t\|_{\dot H^{1+\alpha}}^{2+\frac{s-d}{1+\alpha}}\|\mu^t-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{\alpha-s+d}{1+\alpha}}\|\mu^t-1\|_{L^\infty}^{\frac{1}{1+\alpha}}\\ &\quad+C\|\mu^t\|_{\dot H^{1+\alpha}}^{2+\frac{s-d}{1+\alpha}}\|\mu^t\|_{L^\infty}\|\mu^t-1\|_{L^2}^{\frac{d-s}{1+\alpha}},\end{aligned}\tag{5.20}$$

for some constant $C > 0$ depending only on $d, s, \alpha, \mathbb{M}$. By Plancherel's theorem,

$$\|\mu^t\|_{\dot H^{1+\alpha}}^2\ge\|\mu^t\|_{\dot H^{1+\alpha}}^{2+\frac{s-d}{1+\alpha}}(\|\mu^t\|_{L^2}^2-1)^{\frac{d-s}{2(1+\alpha)}}\ge\|\mu^t\|_{\dot H^{1+\alpha}}^{2+\frac{s-d}{1+\alpha}}(\|\mu^0\|_{L^2}^2-1)^{\frac{d-s}{2(1+\alpha)}},$$

where the final inequality follows from $\|\mu^t\|_{L^2} \ge \|\mu^0\|_{L^2} > 1$, since the L^2 norm is nonincreasing. Thus, using that $\|\mu^t\|_{L^\infty}$ is nonincreasing and the triangle inequality, it follows from (5.20) that

$$\frac{d}{dt}\|\mu^t\|_{\dot{H}^\alpha}^2 \le \|\mu^t\|_{\dot{H}^{1+\alpha}}^{2+\frac{s-d}{1+\alpha}}\Big(-\sigma(\|\mu^0\|_{L^2}^2-1)^{\frac{d-s}{2(1+\alpha)}} + C\|\mu^t-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{\alpha-s+d}{1+\alpha}}(1+\|\mu^0\|_{L^\infty})^{\frac{1}{1+\alpha}} + C\|\mu^0\|_{L^\infty}\|\mu^t-1\|_{L^2}^{\frac{d-s}{1+\alpha}}\Big). \tag{5.21}$$

Applying the exponential decay of $\|\mu^t-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}$, $\|\mu^t-1\|_{L^2}$, given by estimate (4.18) of Lemma 4.9 in the conservative case, or estimate (2.4) of Lemma 2.3 in the dissipative case, we see that there is a $T_* > 0$, a lower bound which is explicitly computable, such that the right-hand side of (5.21) is < 0 for all $t > T_*$. Hence, $\|\mu^t\|_{\dot{H}^\alpha}^2$ is strictly decreasing on (T_*, ∞).

Using Young's product inequality, we see that for any $\varepsilon > 0$, the right-hand side of (5.20) is

$$\le \Big(-\sigma + \Big(2+\frac{s-d}{1+\alpha}\Big)\varepsilon\Big)\|\mu^t\|_{\dot{H}^{1+\alpha}}^2 + \frac{(d-s)}{2(1+\alpha)}\Big(C\varepsilon^{-\frac{2+\frac{s-d}{1+\alpha}}{2}}\|\mu^t-1\|_{L^\infty}^{\frac{1}{1+\alpha}}\|\mu^t-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{\alpha+d-s}{1+\alpha}}\Big)^{\frac{2(1+\alpha)}{d-s}} + \frac{(d-s)}{2(1+\alpha)}\Big(C\varepsilon^{-\frac{2+\frac{s-d}{1+\alpha}}{2}}\|\mu^t\|_{L^\infty}\|\mu^t-1\|_{L^2}^{\frac{d-s}{1+\alpha}}\Big)^{\frac{2(1+\alpha)}{d-s}}.$$

Choosing ε sufficiently small depending on d, s, α, σ, we see that the first term is nonpositive. Using that $\|\mu^t\|_{L^\infty}$ is nonincreasing, we now conclude from the fundamental theorem of calculus that

$$\|\mu^t\|_{\dot{H}^\alpha}^2 \le \|\mu^0\|_{\dot{H}^\alpha}^2 + C_\varepsilon\int_0^t\Big((1+\|\mu^0\|_{L^\infty})^{\frac{2}{d-s}}\|\mu^\tau-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{2(\alpha+d-s)}{d-s}} + \|\mu^0\|_{L^\infty}^{\frac{2(1+\alpha)}{d-s}}\|\mu^\tau-1\|_{L^2}^2\Big)d\tau.$$

Using estimate (4.18) from Lemma 4.9 in the conservative case and (2.4) from Lemma 2.3 in the dissipative case, the preceding right-hand side is controlled by

$$\begin{aligned}
&\|\mu^0\|_{\dot{H}^\alpha}^2 + C_\varepsilon\int_0^t e^{-C\sigma\tau}\Big((1+\|\mu^0\|_{L^\infty})^{\frac{2}{d-s}} \\
&\qquad\times\Big(\|\mu^0-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}^{\frac{2(\alpha+d-s)}{d-s}}\mathbf{1}_{\mathbb{M}\text{ a.s.}} + (1+\|\mu^0\|_{L^\infty})^{\frac{2(\alpha+d-s)}{d-s}-1}\sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma}\,\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big) \\
&\qquad+\|\mu^0\|_{L^\infty}^{\frac{2(1+\alpha)}{d-s}}\Big(\|\mu^0-1\|_{L^2}^2\mathbf{1}_{\mathbb{M}\text{ a.s.}} + (1+\|\mu^0\|_{L^\infty})\sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma}\,\mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big)\Big)d\tau \\
&\le \|\mu^0\|_{\dot{H}^\alpha}^2 + \frac{C_\varepsilon C}{\sigma}\widetilde{\mathbf{W}}_\alpha(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}, \mathcal{F}_\sigma(\mu^0)/\sigma),
\end{aligned}$$

where $\widetilde{\mathbf{W}}_\alpha$ is a continuous, nondecreasing function of its arguments, vanishing if any of its arguments is zero. Also, $\widetilde{\mathbf{W}}_\alpha$ does not depend on its third argument if $\mathbb{M}$ is antisymmetric and does not depend on its second argument if $\mathbb{M} = -\mathbb{I}$. Implicitly, we have used above that $\|\cdot\|_{L^2} \leq \|\cdot\|_{L^{\frac{2(\alpha+d-s)}{d-s}}}$ in arriving at the final inequality. With this final estimate, the proof of the lemma is complete. ∎

Remark 5.5. If $1 \leq p \leq 2$, then by Hölder's inequality and Lemma 2.9 with $\alpha = s - d + 1$,

$$\begin{aligned}\| |\nabla|^{s-d+1}\mu^t\|_{L^p} &\leq \|\mu^t\|_{\dot{H}^{s-d+1}} \\ &\leq 2\Bigg(\|\mu^0\|_{\dot{H}^{s-d+1}} \\ &\qquad + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{s-d+1}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathscr{F}_\sigma(\mu^0)/\sigma})}\Bigg).\end{aligned}$$

If $2 < p < \infty$, then by Sobolev embedding and Lemma 2.9 with $\alpha = 1 + s - d + d(\frac{1}{2} - \frac{1}{p})$,

$$\begin{aligned}&\| |\nabla|^{s-d+1}\mu^t\|_{L^p} \\ &\quad\lesssim \|\mu^t\|_{\dot{H}^{1+s-d+d(\frac{1}{2}-\frac{1}{p})}} \\ &\quad\leq \Bigg(\|\mu^0\|_{\dot{H}^{s+1-d(\frac{1}{2}+\frac{1}{p})}} \\ &\qquad + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{1+d(\frac{1}{2}-\frac{1}{p})+s-d}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2+2d(\frac{1}{2}-\frac{1}{p})}{d-s}}}, \sqrt{\mathscr{F}_\sigma(\mu^0)/\sigma})}\Bigg).\end{aligned}$$

If $p = \infty$, due to the failure of endpoint Sobolev embedding, we instead have the preceding bound with an arbitrarily small ϵ added to $1 + s - d + d(\frac{1}{2} - \frac{1}{p})$. In all cases, there exist $\lambda_p > 0$ defined by

$$\lambda_p := \begin{cases} 1+s-d, & 1 \leq p \leq 2, \\ 1+s-d+d\Big(\dfrac{1}{2}-\dfrac{1}{p}\Big), & 2 < p < \infty, \\ \Big(1+s-\dfrac{d}{2}\Big)+, & p = \infty, \end{cases}$$

such that for any $1 \leq p \leq \infty$, for all $t \geq 0$,

$$\|\nabla \mathrm{g} * \mu^t\|_{L^p} \leq C\Big(\|\mu^0\|_{\dot{H}^{\lambda_p}} + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{\lambda_p}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2(\lambda_p+d-s)}{d-s}}}, \mathscr{F}_\sigma(\mu^0)/\sigma)}\Big).$$

With Lemma 2.9 in hand, we are now ready to prove Lemma 2.10, which is the analogue of Lemma 2.7 in the case $d - 1 < s < d$.

Proof of Lemma 2.10. We first prove assertion (2.12) in the case $q = 2$. We will then treat general L^q norms by Hölder's inequality ($q \le 2$) and Gagliardo–Nirenberg interpolation ($q > 2$). The reason for this approach is that we need an a priori uniform-in-time bound on $\| |\nabla|^{s+1-d} \mu^\tau \|_{L^q}$ if we try to directly start with general q, and the only way we know how to obtain such a bound is through an intermediate L^2 estimate (i.e., Remark 5.5) and Sobolev embedding as commented above.

Starting from (5.2) with ∂_α replaced by $|\nabla|^\alpha$ and recycling notation, we define

$$
\begin{aligned}
J_1(t) &:= \|e^{\sigma t \Delta} |\nabla|^\alpha \mu^0\|_{L^2}, \\
J_2(t) &:= \int_0^{t(1-\varepsilon)} \|e^{\sigma(t-\tau)\Delta} \operatorname{div} |\nabla|^\alpha (\mu^\tau \mathbb{M} \nabla \mathsf{g} * \mu^\tau)\|_{L^2}\, d\tau, \\
J_3(t) &:= \int_{t(1-\varepsilon)}^t \|e^{\sigma(t-\tau)\Delta} \operatorname{div} |\nabla|^\alpha (\mu^\tau \mathbb{M} \nabla \mathsf{g} * \mu^\tau)\|_{L^2}\, d\tau,
\end{aligned}
$$

for $\varepsilon \in (0,1)$ to be determined. Analogously to (5.3), heat kernel estimates give

$$
J_1(t) \lesssim e^{-C\sigma t} \min(\sigma t, 1)^{-\frac{\alpha}{2}} \|\mu^0 - 1\|_{L^2}.
$$

For $J_2(t)$, we also have

$$
\begin{aligned}
&\| |\nabla|^\alpha e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M} \nabla \mathsf{g} * \mu^\tau)\|_{L^2} \\
&\quad \lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau), 1)^{-\frac{1+\alpha}{2}} \min(\sigma\tau, 1)^{-\frac{1}{2}} \|\mu^0\|_{L^d} \|\mathbb{M} \nabla \mathsf{g} * \mu^\tau\|_{L^2} \\
&\quad \lesssim e^{-C\sigma(t-\tau)} \min(\sigma(t-\tau), 1)^{-\frac{1+\alpha}{2}} \min(\sigma\tau, 1)^{-\frac{1}{2}} \|\mu^0\|_{L^d} \\
&\qquad \times \left(\|\mu^0\|_{\dot{H}^{\lambda_2}} + \sqrt{\sigma^{-1} \widetilde{\mathbf{W}}_{\lambda_2}\left(\|\mu^0\|_{L^\infty}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma}\right)} \right),
\end{aligned}
$$

where we have used (2.5) from Corollary 2.4 in the first inequality and Lemma 2.9 in the second. Thus,

$$
\begin{aligned}
J_2(t) &\lesssim \frac{A_{\varepsilon,\alpha}}{\sigma} (\sigma t)^{-\frac{\alpha}{2}} \|\mu^0\|_{L^d} \\
&\qquad \times \left(\|\mu^0\|_{\dot{H}^{\lambda_2}} + \sqrt{\sigma^{-1} \widetilde{\mathbf{W}}_{\lambda_2}\left(\|\mu^0\|_{L^\infty}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma}\right)} \right),
\end{aligned}
$$

where, similarly to (5.7),

$$
A_{\varepsilon,\alpha} := \int_0^{1-\varepsilon} (1-\tau)^{-\frac{d}{2}(\frac{1}{p}-\frac{1}{q})-\frac{1+\alpha}{2}} \tau^{-\frac{1}{2}}\, d\tau.
$$

For $J_3(t)$, we choose $\delta' \in (1+s-d, 1)$ so that

$$
\alpha + 1 + s - d - \delta' < \alpha. \tag{5.22}
$$

Using the fractional Leibniz rule (see [54, Theorem 7.6.1]), we find that

$$\begin{aligned}
&\| |\nabla|^{\alpha} e^{\sigma(t-\tau)\Delta} \operatorname{div}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^2} \\
&\quad\lesssim \min(\sigma(t-\tau),1)^{-\frac{1+\delta'}{2}} \| |\nabla|^{\alpha-\delta'}(\mu^\tau \mathbb{M}\nabla \mathsf{g} * \mu^\tau)\|_{L^2} \\
&\quad\lesssim \min(\sigma(t-\tau),1)^{-\frac{1+\delta'}{2}} \big(\| |\nabla|^{\alpha-\delta'}\mu^\tau\|_{L^{p_1}} \|\mathbb{M}\nabla\mathsf{g} * \mu^\tau\|_{L^{p_2}} \\
&\qquad\qquad + \|\mu^\tau\|_{L^{\tilde{p}_1}} \| |\nabla|^{\alpha-\delta'} \mathbb{M}\nabla \mathsf{g} * \mu^\tau\|_{L^{\tilde{p}_2}}\big),
\end{aligned} \tag{5.23}$$

where $\frac{1}{p_1}+\frac{1}{p_2}=\frac{1}{\tilde{p}_1}+\frac{1}{\tilde{p}_2}=\frac{1}{2}$. Choose $(\tilde{p}_1,\tilde{p}_2)=(\infty,2)$, so that by condition (5.22),

$$\|\mu^\tau\|_{L^{\tilde{p}_1}} \| |\nabla|^{\alpha-\delta'}\mathbb{M}\nabla\mathsf{g} * \mu^\tau\|_{L^{\tilde{p}_2}} \lesssim \|\mu^\tau\|_{L^\infty} \| |\nabla|^{\alpha}\mu^\tau\|_{L^2}.$$

Note that $\alpha-\delta' < \alpha-(s+1-d)$, by choice of δ'. So, using Gagliardo–Nirenberg interpolation, we have for the choice $(p_1,p_2)=(\frac{2\alpha}{\alpha-(s+1-d)},\frac{2\alpha}{s+1-d})$ (which the reader may check is Hölder conjugate to 2),

$$\begin{aligned}
\| |\nabla|^{\alpha-\delta'}\mu^\tau\|_{L^{p_1}} &\lesssim \| |\nabla|^{\alpha-(s+1-d)}\mu^\tau\|_{L^{p_1}} \lesssim \|\mu^\tau-1\|_{L^\infty}^{\frac{s+1-d}{\alpha}} \|\mu^\tau\|_{\dot{H}^\alpha}^{1-\frac{s+1-d}{\alpha}}, \\
\|\mathbb{M}\nabla\mathsf{g} * \mu^\tau\|_{L^{p_2}} &\lesssim \| |\nabla|^{s+1-d}\mu^\tau\|_{L^{p_2}} \lesssim \|\mu^\tau-1\|_{L^\infty}^{1-\frac{s+1-d}{\alpha}} \|\mu^\tau\|_{\dot{H}^\alpha}^{\frac{s+1-d}{\alpha}}.
\end{aligned}$$

Evidently, the preceding implies

$$\| |\nabla|^{\alpha-\delta'}\mu^\tau\|_{L^{p_1}} \|\mathbb{M}\nabla\mathsf{g} * \mu^\tau\|_{L^{p_2}} \lesssim \|\mu^\tau-1\|_{L^\infty}\|\mu^\tau\|_{\dot{H}^\alpha}$$

and in turn that the right-hand side of (5.23) is

$$\begin{aligned}
&\lesssim \min(\sigma(t-\tau),1)^{-\frac{1+\delta'}{2}} \|\mu^\tau\|_{L^\infty} \| |\nabla|^\alpha \mu^\tau\|_{L^2} \\
&\lesssim \min(\sigma(t-\tau),1)^{-\frac{1+\delta'}{2}} \min(\sigma\tau,1)^{-\frac{1-\delta'}{2}} \|\mu^0\|_{L^{\frac{d}{1-\delta'}}} \| |\nabla|^\alpha\mu^\tau\|_{L^2},
\end{aligned}$$

where the second line is by (2.5) from Corollary 2.4 applied to $\|\mu^\tau\|_{L^\infty}$. Hence, defining $\phi(t) := \sup_{t\geq\tau>0}(\sigma\tau)^{\frac{\alpha}{2}} \| |\nabla|^\alpha\mu^\tau\|_{L^2}$ and using dilation invariance of Lebesgue measure, we obtain the estimate

$$J_3(t) \leq \frac{CB_{\varepsilon,\alpha}\|\mu^0\|_{L^{\frac{d}{1-\delta'}}}}{\sigma(\sigma t)^{\frac{\alpha}{2}}}\phi(t),$$

where

$$B_{\varepsilon,\alpha} := \int_{1-\varepsilon}^{1} (1-\tau)^{-\frac{1+\delta'}{2}} \tau^{-\frac{\alpha+(1-\delta')}{2}}\, d\tau.$$

Note that δ' may be chosen independently of α, hence we have omitted the dependence on it from our notation. Choosing ε sufficiently close to 1 so that $CB_{\varepsilon,\alpha}\|\mu^0\|_{L^{\frac{d}{1-\delta'}}} < \frac{\sigma}{2}$, we arrive at

$$\begin{aligned}
\phi(t) &\leq C_{\alpha,2}\|\mu^0\|_{L^2} \\
&\quad + \frac{A_{\varepsilon,\alpha}\|\mu^0\|_{L^d}}{\sigma}\Bigg(\|\mu^0\|_{\dot{H}^{\lambda_2}} \\
&\qquad + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{\lambda_2}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma})}\Bigg).
\end{aligned} \tag{5.24}$$

From these L^2 estimates, we now obtain general L^q estimates. If $1 \le q \le 2$, then Hölder's inequality implies that $\sup_{0<t\le\sigma^{-1}}(\sigma t)^{\frac{\alpha}{2}}\| |\nabla|^{\alpha}\mu^t\|_{L^q}$ is controlled by the right-hand side of (5.24). If $2 < q < \infty$, then choosing $\beta = \frac{q\alpha}{2}$, Gagliardo–Nirenberg interpolation gives

$$\begin{aligned}
\| |\nabla|^{\alpha}\mu^t\|_{L^q} &\lesssim \|\mu^t - 1\|_{L^\infty}^{1-\frac{2}{q}}\|\mu^t\|_{\dot H^\beta}^{\frac{2}{q}} \\
&\le (\sigma t)^{-\frac{\alpha}{2}}\|\mu^0\|_{L^\infty}^{1-\frac{2}{q}} \\
&\quad\times\Bigg(C_{\beta,2}\|\mu^0-1\|_{L^2} + \frac{A_{\varepsilon,\beta}\|\mu^0\|_{L^d}}{\sigma} \\
&\qquad\times\Big(\|\mu^0\|_{\dot H^{\lambda_2}} + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{\lambda_2}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma})}\Big)\Bigg)^{\frac{2}{q}}. \quad (5.25)
\end{aligned}$$

If $q=\infty$, then for $\epsilon > 0$ and $1 < r < \infty$, we have by Sobolev embedding and (5.25),

$$\begin{aligned}
\| |\nabla|^{\alpha}\mu^t\|_{L^\infty} &\lesssim \| |\nabla|^{\alpha+\frac{d}{r}+\epsilon}\mu^t\|_{L^r} \\
&\le (\sigma t)^{-\frac{\alpha+\frac{d}{r}+\epsilon}{2}}\|\mu^0\|_{L^\infty}^{1-\frac{2}{r}} \\
&\quad\times\Bigg(C_{\beta,2}\|\mu^0-1\|_{L^2} + \frac{A_{\varepsilon,\beta}\|\mu^0\|_{L^d}}{\sigma} \\
&\qquad\times\Big(\|\mu^0\|_{\dot H^{\lambda_2}} + \sqrt{\sigma^{-1}\widetilde{\mathbf{W}}_{\lambda_2}(\|\mu^0\|_{L^\infty}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \sqrt{\mathcal{F}_\sigma(\mu^0)/\sigma})}\Big)\Bigg)^{\frac{2}{r}}, \quad (5.26)
\end{aligned}$$

where $\beta := \alpha + \frac{d}{r} + \epsilon$. Choosing r arbitrarily large, this completes the proof of the lemma. ■

Remark 5.6. A posteriori, one can infer from Lemma 2.10 that for all $n \in \mathbb{N}$ and $1 \le q \le \infty$, there exists a function $\mathbf{W}_{n,q}$ with the same properties as $\mathbf{W}_{\alpha,q}$, such that, for all $t \in (0, \sigma^{-1}]$,

$$\begin{aligned}
\|\nabla^{\otimes n}\mu^t\|_{L^q} &\le (\sigma t)^{-\frac{\alpha}{2}}(1 + C_\varepsilon(\sigma t)^{-\varepsilon}\mathbf{1}_{q=\infty}) \\
&\qquad\times \mathbf{W}_{\alpha,q}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot H^{\lambda_2}}, \sigma^{-1}, \|\mu^0-1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0)).
\end{aligned}$$

Indeed, the case $q < \infty$ follows from (2.12) using the L^q boundedness of the Fourier multiplier $\frac{\nabla}{|\nabla|}$. The case $q = \infty$ follows from $\frac{\nabla}{|\nabla|^{1+\varepsilon}}$ being bounded on L^∞, for $\varepsilon > 0$.

Similarly to Lemma 2.8, we now combine Lemma 2.10 with Lemmas 4.9 and 2.3 to show Lemma 2.11, giving estimate (2.9) of Proposition 2.6. Estimate (2.10) then follows from the preceding remark.

Proof of Lemma 2.11. Fix $t > 0$ and assume that $\sigma t > 1$ (otherwise, there is nothing to prove). Let $\sigma t_0 = \sigma t - \frac{1}{2} > \frac{\sigma t}{2}$. Then, translating time and applying Lemma 2.10, we obtain for any $\beta > 0$,

$$\begin{aligned}
\| |\nabla|^{\beta} \mu^t \|_{L^2} &\leq (\sigma(t-t_0))^{-\frac{\beta}{2}} (1 + C_\varepsilon (\sigma(t-t_0))^{-\varepsilon} \mathbf{1}_{q=\infty}) \\
&\qquad \times \mathbf{W}_{\beta,2}(\|\mu^{t_0}\|_{L^\infty}, \|\mu^{t_0}\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^{t_0} - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^{t_0})) \\
&\leq C(1 + C'_\varepsilon \mathbf{1}_{q=\infty}) \\
&\qquad \times \mathbf{W}_{\beta,2}(\|\mu^{t_0}\|_{L^\infty}, \|\mu^{t_0}\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^{t_0} - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^{t_0})). \qquad (5.27)
\end{aligned}$$

By Lemmas 2.9, 4.9, and 2.3, and the nondecreasing property of $\mathbf{W}_{\beta,2}$,

$$\begin{aligned}
&\mathbf{W}_{\beta,2}(\|\mu^{t_0}\|_{L^\infty}, \|\mu^{t_0}\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^{t_0} - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^{t_0})) \\
&\quad \leq \mathbf{W}_{\beta,2}\Big(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}^2 + \sigma^{-1}\widetilde{\mathbf{W}}_{\lambda_2}(\|\mu^0\|_{L^\infty}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0)/\sigma), \sigma^{-1}, \\
&\qquad\qquad e^{-\frac{C\sigma t}{2}} \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}} \mathbf{1}_{\mathbb{M} \text{ a.s.}} \\
&\qquad\qquad\quad + (1 + \|\mu^0\|_{L^\infty})^{1-\frac{d-s}{2}} \left(e^{-2\pi^2\sigma t} \sqrt{2\mathcal{F}_\sigma(\mu^0)/\sigma}\right)^{\frac{d-s}{2}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}}, \\
&\qquad\qquad e^{-2\pi^2\sigma t} \mathcal{F}_\sigma(\mu^0)\Big). \qquad (5.28)
\end{aligned}$$

Let us denote the right-hand side by $\widetilde{\mathbf{W}}_{\beta,2}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \sigma^{-1}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0))$. Now for $\alpha > 0$, by interpolation, then combining (5.27) and (5.28) for $\beta = 2\alpha$,

$$\begin{aligned}
\| |\nabla|^{\alpha} \mu^t \|_{L^2} &\leq \|\mu^t - 1\|_{L^2}^{\frac{1}{2}} \| |\nabla|^{2\alpha} \mu^t \|_{L^2}^{\frac{1}{2}} \\
&\leq \Big(e^{-C\sigma t} \|\mu^0 - 1\|_{L^2} \mathbf{1}_{\mathbb{M} \text{ a.s.}} \\
&\qquad + (1 + \|\mu^0\|_{L^\infty})^{\frac{1}{2}} \left(e^{-4\pi^2\sigma t} \sqrt{2\mathcal{F}_\sigma(\mu^0)/\sigma}\right)^{\frac{1}{2}} \mathbf{1}_{\mathbb{M}=-\mathbb{I}}\Big)^{\frac{1}{2}} \\
&\qquad \times \widetilde{\mathbf{W}}_{\beta,2}(\|\mu^0\|_{L^\infty}, \|\mu^0\|_{\dot{H}^{\lambda_2}}, \|\mu^0 - 1\|_{L^{\frac{2}{d-s}}}, \mathcal{F}_\sigma(\mu^0))^{\frac{1}{2}},
\end{aligned}$$

where in the last line we have also used Lemmas 4.9 and 2.3. Upon relabeling, this yields (2.13) for $q = 2$. For $1 \leq q < 2$, we may simply appeal to Hölder's inequality. For $2 < q \leq \infty$, we appeal to Gagliardo–Nirenberg interpolation similarly to (5.25), (5.26). ■

With the proof of Lemma 2.10 complete, the reader will see, after a little bookkeeping, that the proof of Proposition 2.6 is also complete.

6. The modulated free energy approach

In this section we explain how to prove uniform-in-time propagation of chaos for system (1.1) (in the gradient flow case) using the modulated free energy approach. This will then complete the proof of our main result Theorem 1.2.

6.1. Entropy solutions

First, we must clarify what we mean by a solution to the Liouville equation (1.3), since the kernel $\nabla \mathsf{g}$ is singular. We recall from [24, 68] (e.g., see [24, Definition 2.1]) the definition of an entropy solution to the Liouville equation (1.3). The proof of existence of an entropy solution to (1.3) is sketched in [24, Section 4.2] for the (attractive) case $s = 0$. Following a similar argument, we sketch a proof of existence for the Riesz case (1.2) in Appendix A. In principle, entropy solutions need not be unique, though this is immaterial for our purposes.

Definition 6.1. Let $T > 0$. We say that $f_N \in L^\infty([0,T], L^1((\mathbb{T}^d)^N))$, with $f_N^t \geq 0$ and $\int_{(\mathbb{T}^d)^N} df_N^t = 1$, is an entropy solution to equation (1.3) on the interval $[0,T]$ if it solves (1.3) in the sense of distributions and for $0 \leq t \leq T$,

$$\begin{aligned}\int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^t}{G_N}\Big)\, df_N^t + \sigma \sum_{i=1}^N \int_0^t \int_{(\mathbb{T}^d)^N} \Big|\nabla_{x_i} \log\Big(\frac{f_N^\tau}{G_N}\Big)\Big|^2 df_N^\tau \\ \leq \int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^0}{G_N}\Big)\, df_N^0,\end{aligned}$$

where $G_N := \exp(-\frac{1}{2N\sigma}\sum_{1\leq i\neq j\leq N} \mathsf{g}(x_i - x_j))$. We say that the entropy solution is global if the above holds on $[0,\infty)$.

Lemma 6.2. *If f_N^0 is a probability density on $(\mathbb{T}^d)^N$ such that*

$$\int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^0}{G_N}\Big)\, df_N^0 < \infty,$$

then there exists a global entropy solution to (1.3) *with initial datum f_N^0.*

Remark 6.3. Given that the only entropy solutions we show exist are limits of sequences of smooth solutions to a regularized problem, there seems no harm in taking as part of Definition 6.1 that f_N can be expressed as such a limit.

6.2. The modulated energy and functional inequalities on the torus

As a first step to establishing the functional inequality of Proposition 2.13, we need to discuss properties of the modulated energy, in particular the electric formulation as a renormalized energy following [81, 90, 93, 99].[21]

The distribution $\frac{1}{\mathsf{c}_{d,s}}\mathsf{g}_E$ is the kernel of the nonlocal operator $|\nabla|^{d-s}$ in $\mathbb{R}^d$. However, as popularized by Caffarelli and Silvestre [25], g_E is the restriction to $\mathbb{R}^d \times \{0\}$ of the kernel

$$\mathsf{G}_E(X) := |X|^{-s} \quad \forall X = (x,z) \in \mathbb{R}^d \times \mathbb{R}^k,$$

[21] Strictly speaking, only the first and third cited works consider the periodic setting; but the arguments are adaptations of the Euclidean case anyway.

which satisfies (in the sense of distributions)

$$-\frac{1}{\bar{\mathsf{c}}_{d,s}}\operatorname{div}(|z|^\gamma \nabla \mathsf{G}_E) = \delta_0, \quad \mathbb{R}^d \times \mathbb{R}^k,$$

for $\gamma = s+1-d$ and $k=0$ if $s=d-2$ and $k=1$ if $d-2<s<d$.[22] We generally use capital letters (e.g., X) to denote points of the extended space $\mathbb{R}^d \times \mathbb{R}^k$. Such a representation also holds on $\mathbb{T}^d$, as shown in [84,85]. Namely, let G denote the unique solution of

$$-\frac{1}{\bar{\mathsf{c}}_{d,s}}\operatorname{div}(|z|^\gamma \nabla \mathsf{G}) = \delta_0 - \delta_{\mathbb{T}^d\times\{0\}}, \quad \mathbb{T}^d \times \mathbb{R}^k, \tag{6.1}$$

with $\int_{\mathbb{T}^d\times\mathbb{R}^k} \mathsf{G}\, d\delta_{\mathbb{T}^d\times\{0\}} = 0$. Here, $\delta_{\mathbb{T}^d\times\{0\}}$ denotes the restriction to $\mathbb{T}^d$ viewed as a subspace of $\mathbb{T}^d\times\mathbb{R}^k$.

Following [81], we will also use in Section 6 the following truncation of the extended potential G. For $0<\eta<\frac{1}{4}$, we let

$$\mathsf{G}_\eta := \min(\mathsf{G}_E(\eta), \mathsf{G}_E) + \mathsf{G} - \mathsf{G}_E - \mathsf{C}_\eta, \quad \mathbb{T}^d\times\mathbb{R}^k, \tag{6.2}$$

where

$$\mathsf{C}_\eta := \int_{\mathbb{T}^d\times\mathbb{R}^k} \big(\min(\mathsf{G}_E(\eta), \mathsf{G}_E) + \mathsf{G} - \mathsf{G}_E\big)\, d\delta_{\mathbb{T}^d\times\{0\}}(X).$$

The constant C_η is to enforce that G_η has zero average on $\mathbb{T}^d\times\{0\}$. The reader may check that

$$-\frac{1}{\bar{\mathsf{c}}_{d,s}}\operatorname{div}(|z|^\gamma \nabla \mathsf{G}_\eta) = \delta_0^{(\eta)} - \delta_{\mathbb{T}^d\times\{0\}}, \quad \mathbb{T}^d\times\mathbb{R}^k, \tag{6.3}$$

where $\delta_0^{(\eta)}$ is the positive measure supported on the sphere $\partial B(0,\eta) \subset \mathbb{T}^d\times\mathbb{R}^k$ defined by

$$\int_{\mathbb{T}^d\times\mathbb{R}^k} \varphi\, d\delta_0^{(\eta)} = -\frac{1}{\bar{\mathsf{c}}_{d,s}} \int_{\partial B(0,\eta)} \varphi(X)|z|^\gamma \mathsf{g}_E'(\eta) \quad \forall \varphi \in C(\mathbb{T}^d\times\mathbb{R}^k), \tag{6.4}$$

where g_E is viewed as a function on $\mathbb{R}$ (through radial symmetry) with an abuse of notation. Given $X \in \mathbb{T}^d\times\mathbb{R}^k$, we let $\delta_X^{(\eta)} := \delta_0^{(\eta)}(\cdot - X)$ denote the translate by X.

We introduce the notation

$$H_N := \mathsf{G} * \Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i} - \tilde{\mu}\Big), \tag{6.5}$$

$$H_{N,\vec{\eta}} := \frac{1}{N}\sum_{i=1}^N \mathsf{G}_{\eta_i}(\cdot - X_i) - \mathsf{G} * \tilde{\mu},$$

[22]The constant $\bar{\mathsf{c}}_{d,s}$ should not be confused with the constant $\mathsf{c}_{d,s}$ in (1.2).

where $\tilde{\mu} := \mu\delta_{\mathbb{T}^d\times\{0\}}$ is the identification of μ as a probability measure on $\mathbb{T}^d \times \mathbb{R}^k$ and $\vec{\eta} = (\eta_1, \ldots, \eta_N)$ is an N-tuple of smearing length scales. We use the notation $X_i = (x_i, 0)$ to denote points x_i embedded in the extended space $\mathbb{T}^d \times \mathbb{R}^k$. We also let $H_N^i(X) := H_N(X) - \frac{1}{N}\mathsf{G}(X - X_i)$. Observe from (6.1), (6.3) that

$$-\frac{1}{\bar{\mathsf{c}}_{d,s}}\operatorname{div}(|z|^\gamma \nabla H_{N,\vec{\eta}}) = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta_i)} - \tilde{\mu}. \tag{6.6}$$

Consider the quantity

$$\begin{aligned}\mathcal{F}^{\vec{\eta}} := \frac{1}{2\bar{\mathsf{c}}_{d,s}}\Bigg(\int_{\mathbb{T}^d\times\mathbb{R}^k} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2\, dX - \frac{\bar{\mathsf{c}}_{d,s}}{N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} \mathsf{G}_{\eta_i}\, d\delta_0^{(\eta_i)} \\ - \frac{2\bar{\mathsf{c}}_{d,s}}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} \mathsf{F}_{\eta_i}(x - x_i)\, d\tilde{\mu}(x)\Bigg),\end{aligned}$$

where $\mathsf{F}_{\eta_i} := \mathsf{G} - \mathsf{G}_{\eta_i}$. Using identity (6.6) and integration by parts, it is straightforward that $\mathcal{F}^{\vec{\eta}}$ converges to $F_N(\underline{x}_N, \mu)$ as $\max_i \eta_i \to 0$. One can say more: the expression $\mathcal{F}^{\vec{\eta}}$ is monotonically decreasing with respect to the parameters η_i and becomes equal to the modulated energy $F_N(\underline{x}_N, \mu)$ when the η_i are sufficiently small so that the balls $B(X_i, \eta_i)$ are disjoint.

Proposition 6.4. *Assume $d \geq 1$ and $d - 2 \leq s < d$. Let $\eta_i, \alpha_i \in (0, \frac{1}{4})$ such that $\eta_i \geq \alpha_i$. Given a pairwise distinct configuration $\underline{x}_N \in (\mathbb{T}^d)^N$ and a density $\mu \in L^\infty(\mathbb{T}^d)$ with $\int_{\mathbb{T}^d} \mu = 1$, then*

$$\mathcal{F}^{\vec{\eta}} \leq \mathcal{F}^{\vec{\alpha}}.$$

Defining the nearest-neighbor-type length scale[23]

$$\mathsf{r}_i := \frac{1}{4}\min\Big(\min_{1\leq j\leq N: j\neq i} |x_i - x_j|, (N\|\mu\|_{L^\infty})^{-\frac{1}{d}}\Big) \quad \forall 1 \leq i \leq N,$$

then

$$F_N(\underline{x}_N, \mu) = \mathcal{F}^{\vec{\eta}} \quad \text{if } \eta_i \leq \mathsf{r}_i \text{ for every } 1 \leq i \leq N.$$

From this relation, it follows that there is a constant $C > 0$ depending only d, s such that

$$\begin{aligned}&\frac{1}{2\bar{\mathsf{c}}_{d,s}}\int_{\mathbb{T}^d\times\mathbb{R}^k} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2\, dX \\ &\qquad \leq C\Bigg(F_N(\underline{x}_N, \mu) + \frac{1}{2N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} \mathsf{G}_{\eta_i}\, d\delta_0^{(\eta_i)} + C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}\Bigg),\end{aligned} \tag{6.7}$$

$$\frac{1}{N^2}\sum_{i=1}^N \mathsf{g}_E(\mathsf{r}_i) \leq C\Bigg(F_N(\underline{x}_N, \mu) + \frac{1}{2N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} \mathsf{G}_{\eta_i}\, d\delta_0^{(\eta_i)} + C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}\Bigg). \tag{6.8}$$

[23] The idea to have a length scale which depends on each point originates in [71, 99]. The recognition of the importance, in particular for proving uniform-in-time convergence results, of weighting the typical inter-particle distance $N^{-1/d}$ by the maximum density of the points is due to [91].

Proof. We sketch the proof, which originates in [81] and has been adapted to [99, Lemma 3.2], [4, Lemma B.1]. Given $\alpha_i \leq \eta_i$, write

$$H_{N,\vec{\eta}} = H_{N,\vec{\alpha}} + \frac{1}{N}\sum_{i=1}^{N}(\mathsf{G}_{\eta_i} - \mathsf{G}_{\alpha_i})(X - X_i)$$

and expand the first term in $\mathcal{F}^{\vec{\eta}}$. Using integration by parts and the identities (6.3), (6.6), we obtain

$$\mathcal{F}^{\vec{\alpha}} - \mathcal{F}^{\vec{\eta}} = \frac{1}{2N^2}\sum_{1\leq i\neq j\leq N}\int_{\mathbb{T}^d\times\mathbb{R}^k}(\mathsf{G}_{\alpha_i} - \mathsf{G}_{\eta_i})(X - X_i)\, d(\delta_{X_j}^{(\alpha_j)} + \delta_{X_j}^{(\eta_j)})(X). \tag{6.9}$$

From definition (6.2) of G_η, we see that $\mathsf{G}_{\alpha_i} - \mathsf{G}_{\eta_i} \geq 0$ with support in the closed ball $\overline{B(0,\eta_i)}$. Since $\delta_{X_j}^{(\alpha_j)}$, $\delta_{X_j}^{(\eta_j)}$ are positive measures, it follows that the integral in (6.9) is nonnegative and vanishes if the balls $B(X_i,\eta_i)$, $B(X_j,\eta_j)$ are disjoint, for $i \neq j$. Letting $\max_i \alpha_i \to 0$ now yields $\mathcal{F}^{\vec{\eta}} = F_N(\underline{x}_N,\mu)$.

Relation (6.7) is an immediate consequence of the end result of the preceding paragraph and Hölder's inequality. Relation (6.8) follows by the same reasoning as the proof of [4, Lemma B.1] and using (3.1). ∎

Remark 6.5. Since one may directly estimate the self-interaction term, relation (6.7) implies that the modulated energy is nonnegative up to a term vanishing as $N \to \infty$:

$$F_N(\underline{x}_N,\mu) \geq -\frac{\log(N\|\mu\|_{L^\infty})}{2dN}\mathbf{1}_{s=0} - \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{-1+\frac{s}{d}}, \tag{6.10}$$

where $\mathsf{C} > 0$ depends only on d,s. We use a special font for the constant C to distinguish it in later computations. As previously commented, the order of the term $N^{-1+\frac{s}{d}}$ is sharp. Furthermore, it is known (e.g., see [99, Proposition 3.6]) that the modulated energy is coercive in the sense that it controls a negative-order Sobolev distance between the empirical measure $\mu_N := \frac{1}{N}\sum_i \delta_{x_i}$ and the density μ: for any $\zeta > \frac{d}{2} + d - s$,

$$\begin{aligned}\|\mu_N - \mu\|_{\dot{H}^{-\zeta}} &\leq C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}\\ &\quad + C\Big(F_N(\underline{x}_N,\mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2dN}\mathbf{1}_{s=0} + C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}\Big)^{1/2},\end{aligned}$$

where $C > 0$ depends only on d,s. From this relation, one can deduce that if the N-point configuration $\underline{x}_N$ is regarded as a random vector in $(\mathbb{T}^d)^N$ with law f_N, so that μ_N is a random element in $\mathcal{P}(\mathbb{T}^d)$, then

$$\begin{aligned}\mathbb{E}_{f_N}(\|\mu_N - \mu\|_{\dot{H}^{-\zeta}}^2) &\leq C\mathbb{E}_{f_N}\Big(F_N(\underline{x}_N,\mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2dN}\mathbf{1}_{s=0}\Big)\\ &\quad + C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}(1 + \|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}).\end{aligned}$$

This yields a bound for the difference $f_{N;k} - (\mu)^{\otimes k}$ in a negative-order Sobolev space (see [91, Remark 1.5]).

The relative entropy is obviously nonnegative by Jensen's inequality. Moreover, by sub-additivity, the total N-particle relative entropy controls the relative entropy of the k-point marginals. Using Pinsker's inequality, it follows that

$$\|f_{N;k} - \mu^{\otimes k}\|_{L^1} \leq \sqrt{2kH_k(f_{N;k}|\mu^{\otimes k})} \leq \sqrt{2kH_N(f_N|\mu^{\otimes N})}.$$

The implied rate $O(k/N)$ for the relative entropy between $f_{N;k}$ and $\mu^{\otimes k}$ is in general *not sharp*, as recently demonstrated by Lacker [69], who shows that $O(k^2/N^2)$ is the sharp rate. We note, however, that this cited work is limited to interactions less singular than Riesz (e.g., bounded).

In any case, we conclude that the modulated free energy metrizes both propagation of chaos in the sense of convergence of marginals in the L^1 norm and convergence of the empirical measure in expected Sobolev distance. It is therefore a good quantity for quantitatively proving mean-field convergence.

We now prove Proposition 2.13.

Proof of Proposition 2.13. We only sketch the proof. For more details, we refer to the upcoming work [93].

First, the reader may check using (6.3), (6.1) that if $W = (w, 0)$, $Y = (y, 0) \in \mathbb{T}^d \times \mathbb{R}^k$, and $\eta \in (0, \frac{1}{4})$, then

$$\int_{\mathbb{T}^d \times \mathbb{R}^k} \mathsf{G}(X - W)\, d\delta_Y^{(\eta)}(X) = \mathsf{G}_\eta(W - Y). \tag{6.11}$$

We identify the vector field v as a vector field on $\mathbb{T}^d \times \mathbb{R}^k$ by defining $v(X) := (v(x), 0)$. Desymmetrizing and breaking up the measure,

$$\begin{aligned}
&\int_{(\mathbb{T}^d \times \mathbb{R}^k)^2 \setminus \triangle} (v(X) - v(Y)) \cdot \nabla \mathsf{G}(X - Y)\, d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i} - \tilde{\mu}\Big)^{\otimes 2}(X, Y) \\
&\quad = \sum_{i=1}^N \frac{2}{N} \int_{\mathbb{T}^d \times \mathbb{R}^k} v(X_i) \cdot \nabla \mathsf{G}(X_i - Y)\, d\Big(\frac{1}{N}\sum_{j \neq i} \delta_{X_j} - \tilde{\mu}\Big)(Y) \\
&\qquad - 2\int_{(\mathbb{T}^d \times \mathbb{R}^k)^2 \setminus \triangle} v(X) \cdot \nabla \mathsf{G}(X - Y)\, d\tilde{\mu}(X)\, d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i} - \tilde{\mu}\Big)(Y) \\
&\quad = \frac{2}{N}\sum_{i=1}^N v(X_i) \cdot \nabla H_N^i(X_i) - 2\int_{\mathbb{T}^d} v \cdot \nabla H_N\, d\mu,
\end{aligned} \tag{6.12}$$

where the reader will recall the definitions of H_N, H_N^i from (6.5). Using the identities

$$\begin{aligned}
H_N^i(X) &= H_{N,\vec{\eta}}(X) - \frac{1}{N}\mathsf{G}_{\eta_i}(X - X_i) \quad \text{in } B(X_i, \eta_i), \\
H_N(X) &= H_{N,\vec{\eta}}(X) + \frac{1}{N}\sum_{i=1}^N (\mathsf{G} - \mathsf{G}_{\eta_i})(X - X_i),
\end{aligned}$$

we rewrite expression (6.12) as the sum $\mathrm{Term}_1 + \mathrm{Term}_2 + \mathrm{Term}_3$, where

$$\begin{aligned}
\mathrm{Term}_1 &= 2\int_{\mathbb{T}^d\times\mathbb{R}^k} v\cdot\nabla H_{N,\vec{\eta}}\, d\Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta_i)} - \mu\Big),\\
\mathrm{Term}_2 &= \frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} (v(X_i)-v(X))\cdot\nabla H_N^i(X)\, d\delta_{X_i}^{(\eta_i)}(X)\\
&\quad -\frac{2}{N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} (v(X)-v(X_i))\cdot\nabla \mathsf{G}_{\eta_i}(X-X_i)\, d\delta_{X_i}^{(\eta_i)}(X)\\
&\quad +\frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} (v(X)-v(X_i))\cdot\nabla(\mathsf{G}_{\eta_i}-\mathsf{G})(X-X_i)\, d\tilde{\mu}(X), \qquad (6.13)\\
\mathrm{Term}_3 &= \frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla H_N^i\, d(\delta_{X_i}-\delta_{X_i}^{(\eta_i)})\\
&\quad -\frac{2}{N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla \mathsf{G}_{\eta_i}(X-X_i)\, d\delta_{X_i}^{(\eta_i)}(X)\\
&\quad +\frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla(\mathsf{G}_{\eta_i}-\mathsf{G})(X-X_i)\, d\tilde{\mu}(X).
\end{aligned}$$

First, we claim $\mathrm{Term}_3 = 0$. Indeed, unpacking the definition of H_N^i,

$$\begin{aligned}
\mathrm{Term}_3 &= \frac{2}{N^2}\sum_{1\le i\neq j\le N} \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla\mathsf{G}(X-X_j)\, d(\delta_{X_i}-\delta_{X_i}^{(\eta_i)})(X)\\
&\quad -\frac{2}{N}\sum_{i=1}^N \int_{(\mathbb{T}^d\times\mathbb{R}^k)^2} v(X_i)\cdot\nabla\mathsf{G}(X-Y)\, d\tilde{\mu}(Y)\, d(\delta_{X_i}-\delta_{X_i}^{(\eta_i)})(X)\\
&\quad -\frac{2}{N^2}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla \mathsf{G}_{\eta_i}(X-X_i)\, d\delta_{X_i}^{(\eta_i)}(X)\\
&\quad +\frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla(\mathsf{G}_{\eta_i}-\mathsf{G})(X-X_i)\, d\tilde{\mu}(X). \qquad (6.14)
\end{aligned}$$

Thanks to (6.11), we have

$$\int_{\mathbb{T}^d\times\mathbb{R}^k} \nabla\mathsf{G}(X-X_j)\, d(\delta_{X_i}-\delta_{X_i}^{(\eta_i)})(X) = \nabla\mathsf{G}(X_i-X_j) - \nabla\mathsf{G}_{\eta_i}(X_i-X_j),$$

which vanishes since $\eta_i \le \mathsf{r}_i$ by assumption and $\mathsf{G}_{\eta_i} = \mathsf{G}$ outside $B(0,\eta_i)\subset\mathbb{T}^d\times\mathbb{R}^k$. Thus, the first line of (6.14) vanishes. By the same reasoning, the second line of (6.14) equals

$$-\frac{2}{N}\sum_{i=1}^N \int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla(\mathsf{G}-\mathsf{G}_{\eta_i})(X_i-Y)\, d\tilde{\mu}(Y),$$

and therefore the second line cancels with the second term on the last line of (6.14). It remains to show that

$$\int_{\mathbb{T}^d\times\mathbb{R}^k} v(X_i)\cdot\nabla\mathsf{G}_{\eta_i}(X-X_i)\,d\delta_{X_i}^{(\eta_i)}(X)=0.$$

This is a consequence of the fundamental theorem of calculus, the observation

$$\begin{aligned}\nabla\mathsf{G}_{\eta_i}(X-X_i)(\delta_{X_i}^{(\eta_i)}-\delta_{X_i})(X)&=-\frac{1}{\mathsf{c}_{d,s}}\nabla\mathsf{G}_{\eta_i}(X-X_i)\operatorname{div}(|z|^\gamma\nabla\mathsf{G}_{\eta_i}(X-X_i))\\&=-\frac{1}{\mathsf{c}_{d,s}}\operatorname{div}[\mathsf{G}_{\eta_i}(X-X_i),\mathsf{G}_{\eta_i}(X-X_i)],\end{aligned}$$

and that the last k components of v vanish and the trace of $\nabla_x\mathsf{G}_{\eta_i}$ to $\mathbb{T}^d\times\{0\}$ has zero average. Above, $[\cdot,\cdot]$ denotes the stress–energy tensor, which is the $(d+k)\times(d+k)$ tensor defined by

$$[\varphi,\psi]^{ij}:=|z|^\gamma(\partial_i\varphi\partial_j\psi+\partial_j\varphi\partial_i\psi)-|z|^\gamma\nabla\varphi\cdot\nabla\psi\delta_{ij},\quad 1\le i,j\le d+k,$$

for test functions φ,ψ on $\mathbb{T}^d\times\mathbb{R}^k$.

We write Term_1 in terms of the divergence of the stress-energy tensor as in [99] and integrate by parts to obtain

$$|\mathrm{Term}_1|\le C\|\nabla v\|_{L^\infty}\int_{\mathbb{T}^d\times\mathbb{R}^k}|z|^\gamma|\nabla H_{N,\vec{\eta}}|^2\,dX. \tag{6.15}$$

Finally, consider Term_2. Since $\eta_i\le\mathsf{r}_i$, $\operatorname{supp}(\nabla\mathsf{F}_{\eta_i})\subset\overline{B(0,\eta_i)}$ implies that the second and third lines simplify to

$$\begin{aligned}-\frac{2}{N}\sum_{i=1}^N\frac{1}{N}\Bigg(\frac{1}{N}\int_{\mathbb{T}^d\times\mathbb{R}^k}&\nabla\mathsf{F}_{\eta_i}(X-X_i)\cdot(v(X)-v(X_i))\,d\delta_{X_i}^{(\eta_i)}(X)\\&-\int_{\mathbb{T}^d}\nabla\mathsf{f}_{\eta_i}(x-x_i)\cdot(v(x)-v(x_i))\,d\mu(x)\Bigg),\end{aligned}$$

where f_{η_i} is the trace of F_{η_i} to $\mathbb{T}^d\times\{0\}$. Using $|\nabla\mathsf{F}_\eta|=\eta^{-s-1}$ on the support of $\delta_0^{(\eta)}$, we may bound the first term inside the parentheses by $\frac{1}{N}\eta_i^{-s}\|\nabla v\|_{L^\infty}$, and using $|\nabla\mathsf{f}_\eta|\le|\nabla\mathsf{g}|$, we may bound the second term by $C\eta_i^{d-s}\|\mu\|_{L^\infty}\|\nabla v\|_{L^\infty}$. Using the explicit form (6.4) of $\delta_{X_i}^{(\eta_i)}$ and the mean value theorem, we bound each summand in the first line of (6.13) by

$$C\eta_i\|\nabla v\|_{L^\infty}\int_{\partial B(X_i,\eta_i)}|\nabla H_N^i|\frac{|z|^\gamma}{\eta_i^{s+1}}\,d\mathcal{H}^{d+k-1}(X), \tag{6.16}$$

where $\mathcal{H}^{d+k-1}$ denotes the $(d+k-1)$-dimensional Hausdorff measure in $\mathbb{T}^d\times\mathbb{R}^k$ (equivalent to surface measure). We set $\eta_i=t\mathsf{r}_i$ and average (with respect to Lebesgue measure) over $t \in [\frac{1}{2}, 1]$. After using Cauchy–Schwarz, it follows that the average of (6.16) is

$$\leq \frac{C}{N}\|\nabla v\|_{L^\infty}\Big(\frac{1}{N}\sum_{i=1}^{N} \mathsf{r}_i^{-s} + \int_{\mathbb{T}^d\times\mathbb{R}^k} |z|^\gamma |\nabla H_{N,\vec{\mathsf{r}}}|^2\, dX\Big). \tag{6.17}$$

After a little bookkeeping and using relations (6.7), (6.8) from Proposition 6.4 to bound the right-hand sides of (6.15) and (6.17), we arrive at the statement of the proposition. ∎

6.3. Conclusion of the Grönwall argument

We now have all the ingredients necessary to show a uniform-in-time bound for the modulated free energy. This then completes the proof of Theorem 1.2.

Recalling inequality (2.17), we only need to exhibit decay of the Lipschitz seminorm $\|\nabla u^\tau\|_{L^\infty}$. By the triangle inequality,

$$\begin{aligned}\|\nabla u^\tau\|_{L^\infty} &\leq \sigma\|\nabla^{\otimes 2}\log(\mu^\tau)\|_{L^\infty} + \|\nabla^{\otimes 2}\mathsf{g} * \mu^\tau\|_{L^\infty}\\ &= \sigma\|\nabla^{\otimes 2}\log(\mu^\tau)\|_{L^\infty} + \|\mathsf{g} * \nabla^{\otimes 2}\mu^\tau\|_{L^\infty}.\end{aligned}$$

Assume μ^0 is bounded from below, i.e., $\kappa := \inf_{\mathbb{T}^d}\mu^0 > 0$. Note that $\int_{\mathbb{T}^d}\mu^0 = 1$ implies $\kappa \leq 1$, since $\mathbb{T}^d$ has unit volume. Then $\mu^\tau \geq \kappa$ uniformly in τ by Lemma 4.6, and we have by the chain rule that

$$\begin{aligned}\|\nabla^{\otimes 2}\log(\mu^\tau)\|_{L^\infty} &\leq \Big\|\frac{\nabla^{\otimes 2}\mu^\tau}{\mu^\tau}\Big\|_{L^\infty} + \Big\|\frac{(\nabla\mu^\tau)^{\otimes 2}}{(\mu^\tau)^2}\Big\|_{L^\infty}\\ &\lesssim \frac{\|\nabla^{\otimes 2}\mu^\tau\|_{L^\infty}}{\kappa} + \frac{\|\nabla\mu^\tau\|_{L^\infty}^2}{\kappa^2}.\end{aligned} \tag{6.18}$$

Since g is in L^1,

$$\|\mathsf{g} * \nabla^{\otimes 2}\mu^t\|_{L^\infty} \lesssim \|\nabla^{\otimes 2}\mu^t\|_{L^\infty}. \tag{6.19}$$

Combining (6.18), (6.19), we obtain

$$\|\nabla u^t\|_{L^\infty} \lesssim \Big(1 + \frac{\sigma}{\kappa}\Big)\|\nabla^{\otimes 2}\mu^\tau\|_{L^\infty} + \frac{\sigma\|\nabla\mu^\tau\|_{L^\infty}^2}{\kappa^2}.$$

For $1 \leq n \leq 2$, we may use estimate (2.8) from Proposition 2.6 to find

$$\begin{aligned}&\|\nabla u^t\|_{L^\infty}\\ &\quad\leq \Big(1 + \frac{\sigma}{\kappa}\Big)\mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^\infty}, \mathscr{F}_\sigma(\mu^0))e^{-C\sigma t}\min(\sigma t, 1)^{-1}\\ &\qquad + \frac{\sigma}{\kappa^2}\big(\mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty}, \sigma^{-1}, \|\mu^0 - 1\|_{L^\infty}, \mathscr{F}_\sigma(\mu^0))e^{-C\sigma t}\min(\sigma t, 1)^{-\frac{1}{2}}\big)^2\end{aligned} \tag{6.20}$$

if $d-2 \le s \le d-1$, and estimate (2.10) from Proposition 2.6 to find

$$\begin{aligned}
\|\nabla u^t\|_{L^\infty} \le \Big(1+\frac{\sigma}{\kappa}\Big)e^{-C\sigma t}\min(\sigma t,1)^{-1}(1+C_\varepsilon\min(\sigma t,1)^{-\varepsilon})\\
\times \mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty},\|\mu^0\|_{\dot H^{\lambda_2}},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))\\
+\frac{\sigma}{\kappa^2}\big(e^{-C\sigma t}\min(\sigma t,1)^{-\frac12}(1+C_\varepsilon\min(\sigma t,1)^{-\varepsilon})\\
\times \mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty},\|\mu^0\|_{\dot H^{\lambda_2}},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))\big)^2 \quad (6.21)
\end{aligned}$$

if $d-1<s<d$ and where $\varepsilon>0$ is arbitrary. By Proposition 2.1, there is a time $T_0>0$, comparable to

$$\begin{cases}
\dfrac{\sigma}{\|\mu^0\|_{L^\infty}^2}, & s<d-1,\\
\Big(\dfrac{\sigma^{\frac{d}{2p}+\frac12}}{C_p\|\mu^0\|_{L^\infty}}\Big)^{\frac{2p}{p-d}}, & s=d-1,\\
\Big(\dfrac{\sigma^{\frac{1+\delta}{2}}}{C_\delta\|\mu^0\|_{W^{2,\infty}}}\Big)^{\frac{2}{1-\delta}}, & d-1<s<d,
\end{cases}$$

for some $p\in(d,\infty)$ and $\delta\in(s+d-1,1)$, such that $\|\mu\|_{C([0,T_0],W^{2,\infty})}\le 2\|\mu^0\|_{W^{2,\infty}}$. We then divide the integration over the subintervals $[0,T_0]$ and $[T_0,t]$, assuming that $t\ge T_0$ without loss of generality. On $[0,T_0]$, we use the trivial estimate given by the choice of T_0. On $[T_0,t]$, we use the decay estimates (6.20), (6.21) and we obtain

$$\begin{aligned}
&\int_0^\infty \|\nabla u^\tau\|_{L^\infty}\,d\tau\\
&\le 2C_1T_0\Big(\Big(1+\frac{\sigma}{\kappa}\Big)\|\mu^0\|_{W^{2,\infty}}+\frac{\sigma\|\mu^0\|_{W^{2,\infty}}^2}{\kappa^2}\Big)\\
&\quad + C_1\int_{T_0}^\infty\Bigg(\bigg(\Big(\Big(1+\frac{\sigma}{\kappa}\Big)\mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))\\
&\qquad\times e^{-C\sigma\tau}\min(\sigma\tau,1)^{-1}\\
&\qquad+\frac{\sigma}{\kappa^2}\big(\mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))e^{-C\sigma t}\\
&\qquad\times\min(\sigma\tau,1)^{-\frac12}\big)^2\bigg)\mathbf{1}_{d-2\le s\le d-1}\\
&\qquad+\bigg(\Big(1+\frac{\sigma}{\kappa}\Big)e^{-C\sigma\tau}\min(\sigma\tau,1)^{-1}(1+C_\varepsilon\min(\sigma\tau,1)^{-\varepsilon})\\
&\qquad\times\mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty},\|\mu^0\|_{\dot H^{\lambda_2}},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))\\
&\qquad+\frac{\sigma}{\kappa^2}\big(e^{-C\sigma\tau}\min(\sigma\tau,1)^{-\frac12}(1+C_\varepsilon\min(\sigma\tau,1)^{-\varepsilon})\\
&\qquad\times\mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty},\|\mu^0\|_{\dot H^{\lambda_2}},\sigma^{-1},\|\mu^0-1\|_{L^\infty},\mathscr{F}_\sigma(\mu^0))\big)^2\bigg)\\
&\qquad\times\mathbf{1}_{d-1<s<d}\Bigg)\,d\tau. \qquad (6.22)
\end{aligned}$$

Evidently, the integral over $[T_0,\infty)$ is finite. Applying this bound, we obtain the uniform-in-time estimate (written in compact form)

$$\begin{aligned}\sup_{t\ge 0}\mathcal{E}_N(f_N^t,\mu^t) &\le \mathcal{E}_N(f_N^0,\mu^0)\\ &\qquad\times\big(\mathbf{W}_1(\|\mu^0\|_{W^{2,\infty}},\sigma^{-1},\kappa^{-1},\|\mu^0-1\|_{L^\infty},\mathcal{F}_\sigma(\mu^0))\mathbf{1}_{d-2\le s\le d-1}\\ &\qquad\quad+\mathbf{W}_2(\|\mu^0\|_{W^{2,\infty}},\sigma^{-1},\kappa^{-1},\|\mu^0-1\|_{L^\infty},\mathcal{F}_\sigma(\mu^0))\mathbf{1}_{d-1<s<d}\big),\end{aligned}$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are continuous, nondecreasing in their arguments. This completes the proof of Theorem 1.2.

7. Application to $\dot{W}^{-1,\infty}$ kernels

In this final section of the paper, we sketch the proof of Theorem 2.14, in particular focusing on the main steps and how decay estimates allow one to obtain a uniform-in-time result. For justification of the differential identities below – especially, the consideration needed given that f_N is only a weak solution – we refer to the original article of Jabin–Wang [68, Section 2].

One may verify that if μ is a solution to (2.19), then $\mu^{\otimes N}$ is a solution to the Cauchy problem

$$\begin{cases}\partial_t \bar{f}_N+\sum_{i=1}^N(\mathsf{k}*\mu)(x_i)\cdot\nabla_{x_i}\bar{f}_N=\sigma\sum_{i=1}^N\Delta_{x_i}\bar{f}_N, & (t,\underline{x}_N)\in[0,\infty)\times(\mathbb{T}^d)^N.\\ \bar{f}_N|_{t=0}=(\mu^0)^{\otimes N}, \end{cases}$$

Using this equation, one can show that (see [68, Lemma 2])

$$\begin{aligned}&\frac{d}{dt}H_N(f_N^t|(\mu^t)^{\otimes N})\\ &\quad\le-\frac{1}{N}\sum_{i=1}^N\int_{(\mathbb{T}^d)^N}(\mathsf{k}*(\mu_{\underline{x}_N}-\mu))(x_i)\cdot\nabla_{x_i}\log((\mu^t)^{\otimes N})\,df_N^t(\underline{x}_N)\\ &\qquad-\frac{\sigma}{N}\int_{(\mathbb{T}^d)^N}\Big|\nabla_{x_i}\log\Big(\frac{f_N^t}{(\mu^t)^{\otimes N}}\Big)\Big|^2\,df_N^t,\end{aligned}$$

where $\mu_{\underline{x}_N}:=\frac{1}{N}\sum_{j=1}^N\delta_{x_j}$ and we set $\mathsf{k}(0):=0$, which is harmless given we are modifying on a measure zero set. By assumption (iii), there exists an L^∞ matrix field $(V^{\alpha\beta})_{\alpha,\beta=1}^d$ such that $\mathsf{k}^\alpha=\partial_\beta V^{\alpha\beta}$. Integrating by parts in the variable x_i^β and performing some manipulation, one arrives at the inequality

$$\begin{aligned}\frac{d}{dt}H_N(f_N^t|(\mu^t)^{\otimes N})&\le\frac{1}{N\sigma}\sum_{i=1}^N\int_{(\mathbb{T}^d)^N}|(V*(\mu_{\underline{x}_N}-\mu^t))(x_i)|^2|\nabla\log(\mu^t)(x_i)|^2\,df_N^t\\ &\quad+\frac{1}{N}\sum_{i=1}^N\int_{(\mathbb{T}^d)^N}(V*(\mu_{\underline{x}_N}-\mu))(x_i):\frac{\nabla^{\otimes 2}\mu^t(x_i)}{\mu^t(x_i)}\,df_N^t.\end{aligned}\tag{7.1}$$

To control the right-hand side, we recall the following convexity inequality (see [68, Lemma 1]), sometimes called the Donsker–Varadhan lemma, which allows one to change the N-particle law with respect to which we compute expectations. This is useful because we do not have much information about the N-particle law f_N, as opposed to the tensorized mean-field law $(\mu)^{\otimes N}$.

Lemma 7.1. *Let $\mu_N, \nu_N \in \mathcal{P}((\mathbb{T}^d)^N)$, and let $\Phi \in L^\infty((\mathbb{T}^d)^N)$. Then, for all $\eta > 0$,*

$$\int_{(\mathbb{T}^d)^N} \Phi \, d\mu_N \le \frac{1}{\eta}\bigg(H_N(\mu_N|\nu_N) + \frac{1}{N}\log\bigg(\int_{(\mathbb{T}^d)^N} e^{\eta N \Phi}\, d\nu_N\bigg)\bigg).$$

Applying Lemma 7.1 to each of the two terms in the right-hand side of (7.1), we obtain

$$\begin{aligned} &\frac{d}{dt} H_N(f_N^t|(\mu^t)^{\otimes N}) \\ &\quad \le \Big(\frac{\|\nabla \log \mu^t\|_{L^\infty}^2}{\sigma\eta} + \frac{1}{\eta'}\Big) H_N(f_N^t|(\mu^t)^{\otimes N}) \\ &\qquad + \frac{\|\nabla \log \mu^t\|_{L^\infty}^2}{N^2\sigma\eta} \sum_{\alpha,\beta=1}^{d}\sum_{i=1}^{N} \log\Big(\int_{(\mathbb{T}^d)^N} \exp\big(\eta N |(V^{\alpha\beta} * (\mu_{\underline{x}_N} - \mu^t))(x_i)|^2\big) \\ &\qquad\qquad \times d(\mu^t)^{\otimes N}(\underline{x}_N)\Big) \\ &\qquad + \frac{1}{N\eta'}\log\Big(\int_{(\mathbb{T}^d)^N} \exp\Big(\eta' \sum_{i=1}^{N} (V * (\mu_{\underline{x}_N} - \mu^t))(x_i) : \frac{\nabla^{\otimes 2}\mu^t(x_i)}{\mu^t(x_i)}\Big) \\ &\qquad\qquad \times d(\mu^t)^{\otimes N}(\underline{x}_N)\Big), \end{aligned} \tag{7.2}$$

where the value of the parameters $\eta, \eta' > 0$ will be specified momentarily. Note that by symmetry, the first integral in the right-hand side is independent of the index i.

To close the estimate for the relative entropy, we now recall two functional inequalities. The first is a law of large numbers at exponential scale. For a proof, see [68, Section 4]; but note the result is a consequence of classical exponential inequalities for sums of random vectors [83, 105].

Proposition 7.2. *There exist constants $C_1, C_2 > 0$ such that for any $\phi \in L^\infty(\mathbb{T}^d)$ with $\|\phi\|_{L^\infty} \le 1$ and any probability measure μ on $\mathbb{T}^d$,*

$$\int_{(\mathbb{T}^d)^N} \exp\bigg(\frac{N}{C_1}\bigg|\int_{\mathbb{T}^d} \phi(x) d(\mu_{\underline{x}_N} - \mu)(x)\bigg|^2\bigg) d\mu^{\otimes N}(\underline{x}_N) \le C_2.$$

The next inequality is a large deviation estimate, which is proved in [68, Theorem 4].[24] A much simpler probabilistic proof of this estimate has been given in [73, Section 5].

[24] As noted by Jabin–Wang, this estimate and more would follow from classical large deviations work [7], which in turn builds on [18], if ϕ were continuous.

Proposition 7.3. *Let μ be a probability density on $\mathbb{T}^d$. Suppose that $\phi \in L^\infty((\mathbb{T}^d)^2)$ satisfies*

$$\forall x \in \mathbb{T}^d, \quad \int_{\mathbb{T}^d} \phi(x,z)\, d\mu(z) = 0 \quad \textit{and} \quad \forall z \in \mathbb{T}^d, \quad \int_{\mathbb{T}^d} \phi(x,z)\, d\mu(x) = 0. \tag{7.3}$$

Then there is a universal constant $C_3 > 0$ such that if $\sqrt{C_3}\|\phi\|_{L^\infty} < 1$, then

$$\int_{(\mathbb{T}^d)^N} \exp\left(N \int_{(\mathbb{T}^d)^2} \phi(x,z)\, d\mu_{\underline{x}_N}^{\otimes 2}(x,z) \right) d\mu^{\otimes N}(\underline{x}_N) \le \frac{2}{1 - C_3\|\phi\|_{L^\infty}^2}.$$

Let us now see how to use Propositions 7.2 and 7.3 to complete the proof of the estimate for the evolution of the relative entropy. Let C_1, C_2 be the constants in the statement of Proposition 7.2. For $1 \le \alpha, \beta \le d$, we set

$$\phi(x) := \sqrt{\eta C_1} V^{\alpha\beta}(x_1 - x) \quad \forall x \in \mathbb{T}^d,$$

so that

$$\exp\big(\eta N |(V^{\alpha\beta} * (\mu_{\underline{x}_N} - \mu^t))(x_i)|^2\big) = \exp\left(\frac{N}{C_1} \left| \int_{\mathbb{T}^d} \phi(x)\, d(\mu_{\underline{x}_N} - \mu^t)(x) \right|^2 \right).$$

We choose $\eta > 0$ sufficiently small so that $\sqrt{\eta C_1} \max_{\alpha,\beta} \|V^{\alpha\beta}\|_{L^\infty} = 1$, which ensures that $\|\phi\|_{L^\infty} \le 1$. Applying Proposition 7.2 pointwise in t, we obtain

$$\begin{aligned} &\frac{\|\nabla \log \mu^t\|_{L^\infty}^2}{N^2 \sigma \eta} \sum_{\alpha,\beta=1}^{d} \sum_{i=1}^{N} \log\left(\int_{(\mathbb{T}^d)^N} \exp\big(\eta N |(V^{\alpha\beta} * (\mu_{\underline{x}_N} - \mu^t))(x_i)|^2\big)\, d(\mu^t)^{\otimes N} \right) \\ &\quad \le \frac{d^2 \|\nabla \log \mu^t\|_{L^\infty}^2 C_1 \|V\|_{L^\infty}^2 (\log C_2)}{N\sigma}. \end{aligned} \tag{7.4}$$

Next, set

$$\phi(x,z) := \eta'(V(x-z) - V * \mu^t(x)) : \frac{\nabla^{\otimes 2}\mu^t(x)}{\mu^t(x)} \quad \forall x, z \in \mathbb{T}^d.$$

Using that $\mathsf{k}^\alpha = \partial_\beta V^{\alpha\beta}$ is divergence-free, one checks that ϕ satisfies condition (7.3). Now

$$\eta' \sum_{i=1}^{N} (V * (\mu_{\underline{x}_N} - \mu^t))(x_i) : \frac{\nabla^{\otimes 2}\mu^t(x_i)}{\mu^t(x_i)} = N \int_{(\mathbb{T}^d)^2} \phi(x,z)\, d(\mu_{\underline{x}_N})^{\otimes 2}(x,z).$$

We choose $\eta' > 0$ to satisfy

$$\sqrt{C_3}\|\phi\|_{L^\infty} \le 2\sqrt{C_3}\eta' \|V\|_{L^\infty} \frac{\|\nabla^{\otimes 2}\mu^t\|_{L^\infty}}{\inf \mu^t} = \frac{1}{2},$$

so that by applying Proposition 7.3 pointwise in t,

$$\frac{1}{N\eta'}\log\left(\int_{(\mathbb{T}^d)^N}\exp\left(\eta'\sum_{i=1}^N (V*(\mu_{\underline{x}_N}-\mu^t))(x_i):\frac{\nabla^{\otimes 2}\mu^t(x_i)}{\mu^t(x_i)}\right)d(\mu^t)^{\otimes N}(\underline{x}_N)\right)$$
$$\leq \frac{4\sqrt{C_3}\|V\|_{L^\infty}\|\nabla^{\otimes 2}\mu^t\|_{L^\infty}(\log 4)}{N\inf\mu^t}. \tag{7.5}$$

Applying the estimates (7.4), (7.5) to the right-hand side of (7.2) and substituting in our choices for η, η', we find

$$\begin{aligned}\frac{d}{dt}&H_N(f_N^t|(\mu^t)^{\otimes N})\\ &\leq \left(\frac{\|\nabla\log\mu^t\|_{L^\infty}^2 C_1\|V\|_{L^\infty}^2}{\sigma}+\frac{4\sqrt{C_3}\|V\|_{L^\infty}\|\nabla^{\otimes 2}\mu^t\|_{L^\infty}}{\inf\mu^t}\right)H_N(f_N^t|(\mu^t)^{\otimes N})\\ &\quad+\frac{d^2\|\nabla\log\mu^t\|_{L^\infty}^2 C_1\|V\|_{L^\infty}^2(\log C_2)}{N\sigma}\\ &\quad+\frac{4\sqrt{C_3}\|V\|_{L^\infty}\|\nabla^{\otimes 2}\mu^t\|_{L^\infty}(\log 4)}{N\inf\mu^t}.\end{aligned} \tag{7.6}$$

By the local well-posedness theory for (2.19), there exists a time $T_0>0$ comparable to $\frac{\sigma}{\|\mu^0\|_{W^{2,\infty}}^2}$, such that $\|\mu\|_{C([0,T_0],W^{2,\infty})}\leq 2\|\mu^0\|_{W^{2,\infty}}$. Since $\inf\mu^t\geq\inf\mu^0$, we have

$$\begin{aligned}&\forall t\in[0,T_0], && \|\nabla\log\mu^t\|_{L^\infty}\leq\frac{2\|\mu^0\|_{W^{2,\infty}}}{\inf\mu^0},\\ &\forall t\geq T_0, && \|\nabla\log\mu^t\|_{L^\infty}\leq\frac{\mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1})}{\inf\mu^0}e^{-C_4\sigma t}(\sigma t)^{-\frac{d+1}{2}},\\ &\forall t\geq T_0, && \|\nabla^{\otimes 2}\mu^t\|_{L^\infty}\leq\mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1})e^{-C_4\sigma t}(\sigma t)^{-\frac{d+2}{2}},\end{aligned}$$

where the second and third assertions follow from use of (2.20). Integrating both sides of the differential inequality (7.6), then applying the Grönwall–Bellman lemma, we ultimately find that

$$H_N(f_N^t|(\mu^t)^{\otimes N})\leq e^{\mathfrak{C}^t}\left(H_N(f_N^0|(\mu^0)^{\otimes N})+\frac{\mathfrak{C}^t}{N}\right),$$

where

$$\begin{aligned}\mathfrak{C}^t:=&\frac{C_1\min(t,T_0)}{\sigma}\left(\frac{\|V\|_{L^\infty}\|\mu^0\|_{W^{2,\infty}}}{\inf\mu^0}\right)^2+\frac{\|V\|_{L^\infty}\|\mu^0\|_{W^{2,\infty}}\min(t,T_0)}{\inf\mu^0}\\ &+\left(\frac{C_1}{\sigma^2}\left(\frac{\|V\|_{L^\infty}}{\inf\mu^0}\right)^2\left(\frac{\mathbf{W}_{1,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1})e^{-C_2\sigma T_0}}{(\sigma T_0)^{\frac{d+1}{2}}}\right)^2\right.\\ &\left.+\frac{C_1\|V\|_{L^\infty}\mathbf{W}_{2,\infty}(\|\mu^0\|_{L^\infty},\sigma^{-1})e^{-C_2\sigma T_0}}{\sigma(\sigma T_0)^{\frac{d+2}{2}}\inf\mu^0}\right)\mathbf{1}_{t\geq T_0}.\end{aligned}$$

By inspection, one checks that $\mathfrak{C}^0=0$ and $\sup_{t\geq 0}\mathfrak{C}^t<\infty$. This completes the proof of Theorem 2.14.

A. Proof of Lemma 6.2

We give here the proof of Lemma 6.2 on the existence of entropy solutions to the Liouville equation (1.3).

Let $\chi \in C_c^\infty$ be a bump function, with values between 0 and 1, which is identically 1 on $B(0, \frac{1}{16})$ and zero outside $B(0, \frac{1}{8})$. Given $\varepsilon > 0$, set $\chi_\varepsilon(x) := \varepsilon^{-d}\chi(x/\varepsilon)$ and define the truncated potential $\mathrm{g}_{(\varepsilon)} := \mathrm{g}_E(1-\chi_\varepsilon) + (\mathrm{g} - \mathrm{g}_E)$.[25] Evidently, $\mathrm{g}_{(\varepsilon)} \in C^\infty(\mathbb{T}^d)$ coincides with g if $|x| \geq \frac{\varepsilon}{8}$.

Consider the Cauchy problem for the regularized Liouville equation,

$$\begin{cases} \partial_t f_{N,\varepsilon} = -\sum_{i=1}^N \operatorname{div}_{x_i}\Big(f_{N,\varepsilon} \frac{1}{N} \sum_{1 \leq j \leq N : j \neq i} \mathbb{M}\nabla \mathrm{g}_{(\varepsilon)}(x_i - x_j)\Big) + \sigma \sum_{i=1}^N \Delta_{x_i} f_{N,\varepsilon}, \\ f_{N,\varepsilon}|_{t=0} = f_N^0. \end{cases} \tag{A.1}$$

By standard well-posedness theory for transport–diffusion equations (e.g., see [6, Section 3.4]), (A.1) has a solution $f_{N,\varepsilon} \in L^\infty([0,\infty), \mathcal{P}((\mathbb{T}^d)^N))$, which is C^∞ for positive times. Letting $G_{N,\varepsilon}$ denote the analogue of G_N from Definition 6.1 with g replaced by $\mathrm{g}_{(\varepsilon)}$, the reader may verify the entropy bound

$$\begin{aligned} \forall t \geq 0, \quad & \int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t + \sigma \sum_{i=1}^N \int_0^t \int_{(\mathbb{T}^d)^N} \Big|\nabla_{x_i} \log\Big(\frac{f_{N,\varepsilon}^\tau}{G_{N,\varepsilon}}\Big)\Big|^2 df_{N,\varepsilon}^\tau \\ & \leq \int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^0}{G_{N,\varepsilon}}\Big)\, df_N^0. \end{aligned} \tag{A.2}$$

Since g_E is decreasing and by the properties of χ, the preceding right-hand is

$$\leq \int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^0}{G_N}\Big)\, df_N^0.$$

From relation (3.1),

$$\begin{aligned} -\log(G_{N,\varepsilon}) &= \frac{1}{2N\sigma} \sum_{1 \leq i \neq j \leq N} \mathrm{g}_{(\varepsilon)}(x_i - x_j) \\ &\geq -\frac{N}{\sigma}\Big(\sup_{|x| \geq \frac{1}{4}} |\mathrm{g}(x)| + \sup_{|x| \leq \frac{1}{4}} |\mathrm{g}_E(x) - \mathrm{g}(x)|\Big) > -\infty, \end{aligned}$$

it follows from (A.2) that

$$\begin{aligned} \sup_{\varepsilon > 0} \sup_{t \geq 0} \int_{(\mathbb{T}^d)^N} \log f_{N,\varepsilon}^t \, df_{N,\varepsilon}^t < & \int_{(\mathbb{T}^d)^N} \log\Big(\frac{f_N^0}{G_N}\Big)\, df_N^0 \\ & + \frac{N}{\sigma}\Big(\sup_{|x| \geq \frac{1}{4}} |\mathrm{g}(x)| + \sup_{|x| \leq \frac{1}{4}} |\mathrm{g}_E(x) - \mathrm{g}(x)|\Big). \end{aligned} \tag{A.3}$$

[25] Note that $\mathrm{g}_{(\varepsilon)}$ should not be confused with the truncated potential g_ε from Section 3.

Hence, by the Dunford–Pettis theorem, after passing to a subsequence, $f_{N,\varepsilon}$ converges weakly in $L^1([0,\infty), L^1((\mathbb{T}^d)^N))$ to an element $f_N \in L^1([0,\infty), L^1((\mathbb{T}^d)^N))$. It is easy to check that $f_N^t \geq 0$ for a.e. (t,x).[26] We in fact have $f \in L^\infty([0,\infty), L\log L((\mathbb{T}^d)^N))$. Indeed, let $\rho \in L^1([0,\infty))$ be a temporal function. Then, for any $\varphi \in C((\mathbb{T}^d)^N)$, we have

$$\lim_{\varepsilon\to 0}\int_0^\infty \rho(t)\bigg(\int_{(\mathbb{T}^d)^N}\varphi\, df_{N,\varepsilon}^t - \int_{(\mathbb{T}^d)^N} e^\varphi\bigg)\,dt = \int_0^\infty \rho(t)\bigg(\int_{(\mathbb{T}^d)^N}\varphi\, df_N^t - \int_{(\mathbb{T}^d)^N} e^\varphi\bigg)\,dt.$$

Now by the variational formulation of entropy, if $\varphi \geq 0$,

$$\begin{aligned}\int_0^\infty \rho\bigg(\int_{(\mathbb{T}^d)^N}\varphi\, df_{N,\varepsilon}^t - \int_{(\mathbb{T}^d)^N} e^\varphi\bigg)\,dt &\leq \int_0^\infty \rho\bigg(-1 + \int_{(\mathbb{T}^d)^N}\log f_{N,\varepsilon}^t\, df_{N,\varepsilon}^t\bigg)\,dt \\ &\leq \bigg(-1 + \sup_{\substack{\varepsilon'>0\\ \tau\geq 0}}\int_{(\mathbb{T}^d)^N}\log f_{N,\varepsilon'}^\tau\, df_{N,\varepsilon}^\tau\bigg)\|\rho\|_{L^1}.\end{aligned}$$

Since ρ was arbitrary, the preceding implies that

$$\text{a.e. } t\geq 0,\quad \bigg(\int_{(\mathbb{T}^d)^N}\varphi\, df_N^t - \int_{(\mathbb{T}^d)^N} e^\varphi\bigg) \leq \bigg(-1 + \sup_{\varepsilon'>0}\sup_{\tau\geq 0}\int_{(\mathbb{T}^d)^N}\log f_{N,\varepsilon'}^\tau\, df_{N,\varepsilon}^\tau\bigg).$$

Taking the sup over $\varphi \in C((\mathbb{T}^d)^N)$ in the left-hand side and again using the variational formulation of entropy, we see that

$$\text{a.e. } t\geq 0,\quad \int_{(\mathbb{T}^d)^N}\log f_N^t\, df_N^t \leq \sup_{\varepsilon'>0}\sup_{\tau\geq 0}\int_{(\mathbb{T}^d)^N}\log f_{N,\varepsilon'}^\tau\, df_{N,\varepsilon}^\tau, \tag{A.4}$$

the right-hand side of which is bounded by the right-hand side of (A.3).

Additionally, for any temporal test function $\rho \in C^\infty([0,\infty))$, we have by the weak convergence with spacetime test function $(x,t)\mapsto \rho(t)$,

$$\int_0^\infty \rho(t)\,dt = \lim_{\varepsilon\to 0}\int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\,df_{N,\varepsilon}^t\,dt = \int_0^\infty \rho(t)\int_{(\mathbb{T}^d)^N} df_N^t\,dt.$$

Since ρ was arbitrary, this implies that $\int_{(\mathbb{T}^d)^N} df_N^t = 1$ for a.e. t.

Now for any fixed $\varepsilon_0 \in (0,1]$, we have

$$\forall \varepsilon \leq \frac{\varepsilon_0}{2},\quad -\int_{(\mathbb{T}^d)^N}\log G_{N,\varepsilon_0}\, df_{N,\varepsilon}^t \leq -\int_{(\mathbb{T}^d)^N}\log G_{N,\varepsilon}\, df_{N,\varepsilon}^t.$$

[26] By redefinition of f_N on a set of measure zero, we may assume without loss of generality that $f_N \geq 0$ everywhere.

Note that $\log G_{N,\varepsilon_0}$ is C^∞ and so may be taken as a test function. Let $\rho \in C^\infty([0,\infty))$ be an arbitrary test function of time; the weak convergence of $f_{N,\varepsilon}$ to f_N implies

$$\lim_{\varepsilon\to 0}\int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\log G_{N,\varepsilon_0}\,df_{N,\varepsilon}^t\,dt = \int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\log G_{N,\varepsilon_0}\,df_N^t\,dt.$$

So by monotone convergence,

$$\begin{aligned}\int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\log G_N\,df_N^t\,dt &= \lim_{\varepsilon_0\to 0} -\int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\log G_{N,\varepsilon_0}\,df_N^t\,dt\\ &\le \liminf_{\varepsilon\to 0} -\int_0^\infty\int_{(\mathbb{T}^d)^N}\rho(t)\log G_{N,\varepsilon}\,df_{N,\varepsilon}^t\,dt.\end{aligned}$$

Since ρ was arbitrary, the preceding inequality together with (A.4) implies that

$$\text{a.e. } t>0,\quad \int_{(\mathbb{T}^d)^N}\log\Big(\frac{f_N^t}{G_N}\Big)\,df_N^t \le \liminf_{\varepsilon\to 0}\int_{(\mathbb{T}^d)^N}\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\,df_{N,\varepsilon}^t.$$

Next, observe that $\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}$ converges in the sense of spacetime distributions to $\frac{f_N}{G_N}$. Note also from the monotonicity of g_E and the properties of χ that

$$\forall \varepsilon < \frac{\varepsilon_0}{2},\quad \int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}|^2}{\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}}\,dG_{N,\varepsilon_0} \le \int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}|^2}{\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}}\,dG_{N,\varepsilon}.$$

By weak lower semicontinuity, we have

$$\int_0^T\int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_N}{G_N}|^2}{\frac{f_N}{G_N}}\,dG_{N,\varepsilon_0}\,dt \le \liminf_{\varepsilon\to 0}\int_0^T\int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}|^2}{\frac{f_{N,\varepsilon}}{G_{N,\varepsilon}}}\,dG_{N,\varepsilon_0}\,dt.$$

By another application of the monotone convergence theorem, it follows that

$$\begin{aligned}\sigma\int_0^T\int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_N}{G_N}|^2}{\frac{f_N}{G_N}}\,dG_N\,dt &= \sigma\lim_{\varepsilon_0\to 0}\int_0^T\int_{(\mathbb{T}^d)^N}\frac{|\nabla\frac{f_N}{G_N}|^2}{\frac{f_N}{G_N}}\,dG_{N,\varepsilon_0}\,dt\\ &\le \int_{(\mathbb{T}^d)^N}\log\Big(\frac{f_N^0}{G_N}\Big)\,df_N^0 - \int_{(\mathbb{T}^d)^N}\log\Big(\frac{f_N^T}{G_N}\Big)\,df_N^T.\end{aligned}$$

Finally, we check that the limit f_N satisfies the original Liouville equation (1.3) in the distributional sense on $[0,\infty)\times(\mathbb{T}^d)^N$. Observe that for any test function $\varphi \in C^\infty((\mathbb{T}^d)^N)$ and a.e. t,

$$\int_{(\mathbb{T}^d)^N}\varphi\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\,df_N^t$$

is absolutely convergent. Indeed, this follows since for a.e. t, $\nabla\log(\frac{f_N^t}{G_N})\sqrt{f_N^t} \in L^2$ and therefore by Cauchy–Schwarz,

$$\begin{aligned}\int_{(\mathbb{T}^d)^N} \Big|\varphi\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\Big|\, df_N^t &\le \Big(\int_{(\mathbb{T}^d)^N} |\varphi|^2\, df_N^t\Big)^{1/2}\Big(\int_{(\mathbb{T}^d)^N} \Big|\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\Big|^2 df_N^t\Big)^{1/2} \\ &\le \|\varphi\|_{L^\infty}\Big(\int_{(\mathbb{T}^d)^N} \Big|\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\Big|^2 df_N^t\Big)^{1/2}, \end{aligned} \tag{A.5}$$

where we use that f_N^t is a probability density to obtain the final line. We now want to show that for any spacetime test function ψ,

$$\lim_{\varepsilon\to 0}\int_0^\infty\int_{(\mathbb{T}^d)^N} \psi^t\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t\, dt = \int_0^\infty\int_{(\mathbb{T}^d)^N} \psi^t\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\, df_N^t\, dt.$$

Let $M \gg 1$ and decompose

$$\begin{aligned}\int_{(\mathbb{T}^d)^N} \psi^t\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t &= \int_{(\mathbb{T}^d)^N} \psi^t\mathbf{1}_{G_N^{-1}\ge M}\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t \\ &\quad + \int_{(\mathbb{T}^d)^N} \psi^t\mathbf{1}_{G_N^{-1}< M}\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t.\end{aligned}$$

Observe that

$$r := \inf_{i\ne j}\big\{|x_i - x_j| \colon G_N^{-1}(\underline{x}_N) \le M\big\} > 0.$$

Hence, for all $\varepsilon \le r$, $\nabla\log(G_{N,\varepsilon}^{-1})\mathbf{1}_{G_N^{-1}\le M} = \nabla\log(G_N^{-1})\mathbf{1}_{G_N^{-1}\le M}$. So, by weak convergence,

$$\begin{aligned}\lim_{\varepsilon\to 0}\int_0^\infty\int_{(\mathbb{T}^d)^N} &\psi^t\mathbf{1}_{G_N^{-1}<M}\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t \\ &= \int_0^\infty\int_{(\mathbb{T}^d)^N} \psi^t\mathbf{1}_{G_N^{-1}<M}\nabla\log\Big(\frac{f_N^t}{G_N}\Big)\, df_N^t.\end{aligned}$$

Finally, arguing similarly to (A.5),

$$\begin{aligned}&\Big|\int_{(\mathbb{T}^d)^N} \psi^t\mathbf{1}_{G_N^{-1}\ge M}\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t\Big| \\ &\quad\le \|\psi^t\|_{L^\infty}\Big(\int_{(\mathbb{T}^d)^N}\Big|\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\Big|^2 df_{N,\varepsilon}^t\Big)^{1/2}\Big(\int_{(\mathbb{T}^d)^N}\mathbf{1}_{G_N^{-1}\ge M}\, df_{N,\varepsilon}^t\Big)^{1/2} \\ &\quad\le \frac{\|\psi^t\|_{L^\infty}}{(\log M)^{1/2}}\Big(\int_{(\mathbb{T}^d)^N}\Big|\nabla\log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\Big|^2 df_{N,\varepsilon}^t\Big)^{1/2}\sup_{\varepsilon,t>0}\Big(\int_{(\mathbb{T}^d)^N}\log(G_N^{-1})\, df_{N,\varepsilon}^t\Big)^{1/2}.\end{aligned}$$

Hence by Cauchy–Schwarz,

$$\begin{aligned}\int_0^\infty \Big| \int_{(\mathbb{T}^d)^N} \psi^t \mathbf{1}_{G_N^{-1} \geq M} \nabla \log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\, df_{N,\varepsilon}^t \Big| \\ \leq \sup_{\varepsilon,t>0} \Big(\int_{(\mathbb{T}^d)^N} \log(G_N^{-1})\, df_{N,\varepsilon}^t\Big)^{1/2} (\log M)^{-1/2} \Big(\int_0^\infty \|\psi^t\|_{L^\infty}^2\, dt\Big)^{1/2} \\ \times \sup_{\varepsilon>0} \Big(\int_0^\infty \int_{(\mathbb{T}^d)^N} \Big|\nabla \log\Big(\frac{f_{N,\varepsilon}^t}{G_{N,\varepsilon}}\Big)\Big|^2 df_{N,\varepsilon}^t\, dt\Big)^{1/2},\end{aligned}$$

which tends to zero as $M \to \infty$. This last step completes the proof of the lemma.

Acknowledgments. This work was completed as a result of the first author's internship visit to the Courant Institute of Mathematical Sciences, NYU, which he thanks for its hospitality. The second author thanks Toan T. Nguyen for helpful discussion and feedback about the results of this paper.

Funding. The second author was supported by the Simons Foundation through the Simons Collaboration on Wave Turbulence and by NSF grants DMS-2052651, DMS-2206085. The third author was supported by the Simons Foundation through a Simons Investigator award and by NSF grant DMS-2000205.

References

[1] N. D. Alikakos and R. Rostamian, Large time behavior of solutions of Neumann boundary value problem for the porous medium equation. *Indiana Univ. Math. J.* **30** (1981), no. 5, 749–785 Zbl 0598.76100 MR 625601

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. 2nd edn., Lectures Math. ETH Zürich, Birkhäuser, Basel, 2008 Zbl 1090.35002 MR 2401600

[3] L. Ambrosio and S. Serfaty, A gradient flow approach to an evolution problem arising in superconductivity. *Comm. Pure Appl. Math.* **61** (2008), no. 11, 1495–1539 Zbl 1171.35005 MR 2444374

[4] S. Armstrong and S. Serfaty, Local laws and rigidity for Coulomb gases at any temperature. *Ann. Probab.* **49** (2021), no. 1, 46–121 Zbl 1457.82004 MR 4203333

[5] M. Arnaudon and P. Del Moral, A second order analysis of McKean–Vlasov semigroups. *Ann. Appl. Probab.* **30** (2020), no. 6, 2613–2664 Zbl 1504.60136 MR 4187123

[6] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*. Grundlehren Math. Wiss. 343, Springer, Heidelberg, 2011 Zbl 1227.35004 MR 2768550

[7] G. Ben Arous and M. Brunaud, Méthode de Laplace: Étude variationnelle des fluctuations de diffusions de type "champ moyen". *Stochastics Stochastics Rep.* **31** (1990), no. 1-4, 79–144 Zbl 0705.60046 MR 1080535

[8] S. Benachour, B. Roynette, D. Talay, and P. Vallois, Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* **75** (1998), no. 2, 173–201 Zbl 0932.60063 MR 1632193

[9] S. Benachour, B. Roynette, and P. Vallois, Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl.* **75** (1998), no. 2, 203–224 Zbl 0932.60064 MR 1632197

[10] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti, A non-Maxwellian steady distribution for one-dimensional granular media. *J. Statist. Phys.* **91** (1998), no. 5-6, 979–990 Zbl 0921.60057 MR 1637274

[11] R. J. Berman and M. Önnheim, Propagation of chaos for a class of first order models with singular mean field interactions. *SIAM J. Math. Anal.* **51** (2019), no. 1, 159–196 Zbl 1418.60054 MR 3904425

[12] L. Bertini, G. Giacomin, and K. Pakdaman, Dynamical aspects of mean field plane rotators and the Kuramoto model. *J. Stat. Phys.* **138** (2010), no. 1-3, 270–290 Zbl 1187.82067 MR 2594897

[13] L. Bertini, G. Giacomin, and C. Poquet, Synchronization and random long time dynamics for mean-field plane rotators. *Probab. Theory Related Fields* **160** (2014), no. 3-4, 593–653 Zbl 1309.60093 MR 3278917

[14] P. Biler, C. Imbert, and G. Karch, The nonlocal porous medium equation: Barenblatt profiles and other weak solutions. *Arch. Ration. Mech. Anal.* **215** (2015), no. 2, 497–529 Zbl 1308.35197 MR 3294409

[15] A. Blanchet, On the parabolic–elliptic Patlak–Keller–Segel system in dimension 2 and higher. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2011–2012*, Exp. No. VIII, Sémin. Équ. Dériv. Partielles, École Polytech., Palaiseau, 2013 Zbl 1326.35400 MR 3379847

[16] F. Bolley, I. Gentil, and A. Guillin, Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.* **208** (2013), no. 2, 429–445 Zbl 1264.35040 MR 3035983

[17] F. Bolley, A. Guillin, and F. Malrieu, Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *M2AN Math. Model. Numer. Anal.* **44** (2010), no. 5, 867–884 Zbl 1201.82029 MR 2731396

[18] E. Bolthausen, Laplace approximations for sums of independent random vectors. *Probab. Theory Relat. Fields* **72** (1986), no. 2, 305–318 Zbl 0572.60007 MR 836280

[19] S. V. Borodachov, D. P. Hardin, and E. B. Saff, *Discrete energy on rectifiable sets*. Springer Monogr. Math., Springer, New York, 2019 Zbl 1437.41002 MR 3970999

[20] Y. Brenier, Convergence of the Vlasov–Poisson system to the incompressible Euler equations. *Comm. Partial Differential Equations* **25** (2000), no. 3-4, 737–754 Zbl 0970.35110 MR 1748352

[21] D. Bresch, P.-E. Jabin, and J. Soler, A new approach to the mean-field limit of Vlasov–Fokker–Planck equations. 2022, arXiv:2203.15747

[22] D. Bresch, P.-E. Jabin, and Z. Wang, On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model. *C. R. Math. Acad. Sci. Paris* **357** (2019), no. 9, 708–720 Zbl 1428.35617 MR 4018082

[23] D. Bresch, P.-E. Jabin, and Z. Wang, Modulated free energy and mean field limit. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Anné 2019–2020*, Exp. No. II, Inst. Hautes Études Sci., Bures-sur-Yvette, 2019–2020 Zbl 1506.35247 MR 4632269

[24] D. Bresch, P.-E. Jabin, and Z. Wang, Mean-field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.* **172** (2023), no. 13, 2591–2641 MR 4658923

[25] L. Caffarelli and L. Silvestre, An extension problem related to the fractional Laplacian. *Comm. Partial Differential Equations* **32** (2007), no. 7-9, 1245–1260 Zbl 1143.26002 MR 2354493

[26] L. Caffarelli, F. Soria, and J. L. Vázquez, Regularity of solutions of the fractional porous medium flow. *J. Eur. Math. Soc. (JEMS)* **15** (2013), no. 5, 1701–1746 Zbl 1292.35312 MR 3082241

[27] L. Caffarelli and J. L. Vazquez, Nonlinear porous medium flow with fractional potential pressure. *Arch. Ration. Mech. Anal.* **202** (2011), no. 2, 537–565 Zbl 1264.76105 MR 2847534

[28] L. A. Caffarelli and J. L. Vázquez, Asymptotic behaviour of a porous medium equation with fractional diffusion. *Discrete Contin. Dyn. Syst.* **29** (2011), no. 4, 1393–1404 Zbl 1211.35043 MR 2773189

[29] E. A. Carlen and M. Loss, Optimal smoothing and decay estimates for viscously damped conservation laws, with applications to the 2-D Navier–Stokes equation. *Duke Math. J.* **81** (1995), no. 1, 135–157 Zbl 0859.35011 MR 1381974

[30] J. A. Carrillo and Y.-P. Choi, Mean-field limits: From particle descriptions to macroscopic equations. *Arch. Ration. Mech. Anal.* **241** (2021), no. 3, 1529–1573 Zbl 1478.35170 MR 4284530

[31] J. A. Carrillo, Y.-P. Choi, and M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, pp. 1–46, CISM Courses and Lect. 553, Springer, Vienna, 2014 MR 3331178

[32] J. A. Carrillo, L. C. F. Ferreira, and J. C. Precioso, A mass-transportation approach to a one dimensional fluid mechanics model with nonlocal velocity. *Adv. Math.* **231** (2012), no. 1, 306–327 Zbl 1252.35224 MR 2935390

[33] J. A. Carrillo, Y. Huang, M. C. Santos, and J. L. Vázquez, Exponential convergence towards stationary states for the 1D porous medium equation with fractional pressure. *J. Differential Equations* **258** (2015), no. 3, 736–763 Zbl 1307.35311 MR 3279352

[34] J. A. Carrillo, R. J. McCann, and C. Villani, Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana* **19** (2003), no. 3, 971–1018 Zbl 1073.35127 MR 2053570

[35] J. A. Carrillo, R. J. McCann, and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.* **179** (2006), no. 2, 217–263 Zbl 1082.76105 MR 2209130

[36] P. Cattiaux, A. Guillin, and F. Malrieu, Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** (2008), no. 1-2, 19–40 Zbl 1169.35031 MR 2357669

[37] D. Chae, P. Constantin, D. Córdoba, F. Gancedo, and J. Wu, Generalized surface quasi-geostrophic equations with singular velocities. *Comm. Pure Appl. Math.* **65** (2012), no. 8, 1037–1066 Zbl 1244.35108 MR 2928091

[38] L.-P. Chaintron and A. Diez, Propagation of chaos: A review of models, methods and applications. II. Applications. *Kinet. Relat. Models* **15** (2022), no. 6, 1017–1173 Zbl 1496.82017 MR 4489769

[39] A. Chodron de Courcel, M. Rosenzweig, and S. Serfaty, The attractive log gas: Stability, uniqueness, and propagation of chaos. 2023, arXiv:2311.14560

[40] Y.-P. Choi and I.-J. Jeong, Classical solutions for fractional porous medium flow. *Nonlinear Anal.* **210** (2021), article no. 112393 Zbl 1466.35356 MR 4253949

[41] P. Constantin, A. J. Majda, and E. Tabak, Formation of strong fronts in the 2-D quasi-geostrophic thermal active scalar. *Nonlinearity* **7** (1994), no. 6, 1495–1533 Zbl 0809.35057 MR 1304437

[42] P. Constantin and J. Wu, Behavior of solutions of 2D quasi-geostrophic equations. *SIAM J. Math. Anal.* **30** (1999), no. 5, 937–948 Zbl 0957.76093 MR 1709781

[43] A. Córdoba and D. Córdoba, A maximum principle applied to quasi-geostrophic equations. *Comm. Math. Phys.* **249** (2004), no. 3, 511–528 Zbl 1309.76026 MR 2084005

[44] C. M. Dafermos, The second law of thermodynamics and stability. *Arch. Rational Mech. Anal.* **70** (1979), no. 2, 167–179 Zbl 0448.73004 MR 546634

[45] F. Delarue and A. Tse, Uniform in time weak propagation of chaos on the torus. 2021, arXiv:2104.14973

[46] M. Duerinckx, Mean-field limits for some Riesz interaction gradient flows. *SIAM J. Math. Anal.* **48** (2016), no. 3, 2269–2300 Zbl 1348.82050 MR 3514713

[47] M. Duerinckx, Well-posedness for mean-field evolutions arising in superconductivity. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35** (2018), no. 5, 1267–1319 Zbl 1393.35230 MR 3813965

[48] A. Durmus, A. Eberle, A. Guillin, and R. Zimmer, An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** (2020), no. 12, 5387–5398 Zbl 1471.60123 MR 4163850

[49] E. B. Fabes, C. E. Kenig, and R. P. Serapioni, The local regularity of solutions of degenerate elliptic equations. *Comm. Partial Differential Equations* **7** (1982), no. 1, 77–116 Zbl 0498.35042 MR 643158

[50] T. Gallay and C. E. Wayne, Global stability of vortex solutions of the two-dimensional Navier–Stokes equation. *Comm. Math. Phys.* **255** (2005), no. 1, 97–129 Zbl 1139.35084 MR 2123378

[51] M.-H. Giga, Y. Giga, and J. Saal, *Nonlinear partial differential equations*. Progr. Nonlinear Differential Equations Appl. 79, Birkhäuser Boston, Boston, MA, 2010 Zbl 1215.35001 MR 2656972

[52] F. Golse, Mean-field limits in statistical dynamics. 2022, arXiv:2201.02005

[53] F. Golse and T. Paul, Mean-field and classical limit for the N-body quantum dynamics with Coulomb interaction. *Comm. Pure Appl. Math.* **75** (2022), no. 6, 1332–1376 Zbl 1502.81072 MR 4415778

[54] L. Grafakos, *Modern Fourier analysis*. 3rd edn., Grad. Texts in Math. 250, Springer, New York, 2014 Zbl 1304.42002 MR 3243741

[55] A. Guillin, P. L. Bris, and P. Monmarché, Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems. 2021, arXiv:2108.08675

[56] A. Guillin, P. Le Bris, and P. Monmarché, On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas. *J. Éc. polytech. Math.* **10** (2023), 867–916 Zbl 07690441 MR 4585552

[57] Y. Han, Entropic propagation of chaos for mean field diffusion with L^p interactions via hierarchy, linear growth and fractional noise. 2022, arXiv:2205.02772

[58] D. Han-Kwan and M. Iacobelli, From Newton's second law to Euler's equations of perfect fluids. *Proc. Amer. Math. Soc.* **149** (2021), no. 7, 3045–3061 Zbl 1477.35165 MR 4257814

[59] D. Hardin, E. B. Saff, R. Shu, and E. Tadmor, Dynamics of particles on a curve with pairwise hyper-singular repulsion. *Discrete Contin. Dyn. Syst.* **41** (2021), no. 12, 5509–5536 Zbl 1483.31022 MR 4342127

[60] D. P. Hardin, T. Leblé, E. B. Saff, and S. Serfaty, Large deviation principles for hypersingular Riesz gases. *Constr. Approx.* **48** (2018), no. 1, 61–100 Zbl 1398.82044 MR 3825947

[61] D. P. Hardin, E. B. Saff, and B. Simanek, Periodic discrete energy for long-range potentials. *J. Math. Phys.* **55** (2014), no. 12, article no. 123509 Zbl 1310.74022 MR 3390564

[62] D. P. Hardin, E. B. Saff, B. Z. Simanek, and Y. Su, Next order energy asymptotics for Riesz potentials on flat tori. *Int. Math. Res. Not. IMRN* (2017), no. 12, 3529–3556 Zbl 1405.31008 MR 3693658

[63] M. Hauray, Wasserstein distances for vortices approximation of Euler-type equations. *Math. Models Methods Appl. Sci.* **19** (2009), no. 8, 1357–1384 Zbl 1188.35126 MR 2555474

[64] M. Hauray and S. Mischler, On Kac's chaos and related problems. *J. Funct. Anal.* **266** (2014), no. 10, 6055–6157 Zbl 1396.60102 MR 3188710

[65] J. Huang, M. Rosenzweig, and S. Serfaty, The modulated free energy method on $\mathbb{R}^d$. Unpublished

[66] P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.* **271** (2016), no. 12, 3588–3627 Zbl 1388.60163 MR 3558251

[67] P.-E. Jabin and Z. Wang, Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, pp. 379–402, Model. Simul. Sci. Eng. Technol., Birkhäuser/Springer, Cham, 2017 MR 3644596

[68] P.-E. Jabin and Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.* **214** (2018), no. 1, 523–591 Zbl 1402.35208 MR 3858403

[69] D. Lacker, Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.* **4** (2023), no. 2, 377–432 Zbl 1515.82109 MR 4595391

[70] D. Lacker and L. Le Flem, Sharp uniform-in-time propagation of chaos. *Probab. Theory Related Fields* **187** (2023), no. 1-2, 443–480 Zbl 07735855 MR 4634344

[71] T. Leblé and S. Serfaty, Fluctuations of two dimensional Coulomb gases. *Geom. Funct. Anal.* **28** (2018), no. 2, 443–508 Zbl 1423.60045 MR 3788208

[72] D. Li, On Kato–Ponce and fractional Leibniz. *Rev. Mat. Iberoam.* **35** (2019), no. 1, 23–100 Zbl 1412.35261 MR 3914540

[73] T. S. Lim, Y. Lu, and J. H. Nolen, Quantitative propagation of chaos in a bimolecular chemical reaction–diffusion model. *SIAM J. Math. Anal.* **52** (2020), no. 2, 2098–2133 Zbl 1439.60094 MR 4091878

[74] F. Lin and P. Zhang, On the hydrodynamic limit of Ginzburg–Landau vortices. *Discrete Contin. Dynam. Systems* **6** (2000), no. 1, 121–142 Zbl 1034.35128 MR 1739596

[75] S. Lisini, E. Mainini, and A. Segatti, A gradient flow approach to the porous medium equation with fractional pressure. *Arch. Ration. Mech. Anal.* **227** (2018), no. 2, 567–606 Zbl 1384.35095 MR 3740382

[76] F. Malrieu, Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13** (2003), no. 2, 540–560 Zbl 1031.60085 MR 1970276

[77] N. Masmoudi and P. Zhang, Global solutions to vortex density equations arising from sup-conductivity. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **22** (2005), no. 4, 441–458 Zbl 1070.35036 MR 2145721

[78] M. Ménard, Mean-field limit derivation of a monokinetic spray model with gyroscopic effects. 2022, arXiv:2204.02145

[79] Q.-H. Nguyen, M. Rosenzweig, and S. Serfaty, Mean-field limits of Riesz-type singular flows. *Ars Inven. Anal.* (2022), article no. 4 MR 4462479

[80] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*. Appl. Math. Sci. 44, Springer, New York, 1983 Zbl 0516.47023 MR 710486

[81] M. Petrache and S. Serfaty, Next order asymptotics and renormalized energy for Riesz interactions. *J. Inst. Math. Jussieu* **16** (2017), no. 3, 501–569 Zbl 1373.82013 MR 3646281

[82] I. B. Porat, Derivation of Euler's equations of perfect fluids from von Neumann's equation with magnetic field. *J. Stat. Phys.* **190** (2023), no. 7, article no. 121 Zbl 07723869 MR 4614118

[83] Y. V. Prohorov, The extension of S. N. Bernšteĭn's inequalities to a multi-dimensional case. *Teor. Verojatnost. i Primenen.* **13** (1968), 266–274. Zbl 0167.46903 MR 230353

[84] L. Roncal and P. R. Stinga, Transference of fractional Laplacian regularity. In *Special functions, partial differential equations, and harmonic analysis*, pp. 203–212, Springer Proc. Math. Stat. 108, Springer, Cham, 2014 Zbl 1323.35201 MR 3297661

[85] L. Roncal and P. R. Stinga, Fractional Laplacian on the torus. *Commun. Contemp. Math.* **18** (2016), no. 3, article no. 1550033 Zbl 1383.35246 MR 3477397

[86] M. Rosenzweig, The mean-field limit of stochastic point vortex systems with multiplicative noise. 2020, arXiv:2011.12180

[87] M. Rosenzweig, From quantum many-body systems to ideal fluids. 2021, arXiv:2110.04195

[88] M. Rosenzweig, Mean-field convergence of point vortices to the incompressible Euler equation with vorticity in L^∞. *Arch. Ration. Mech. Anal.* **243** (2022), no. 3, 1361–1431 Zbl 1509.35200 MR 4381143

[89] M. Rosenzweig, The mean-field approximation for higher-dimensional Coulomb flows in the scaling-critical L^∞ space. *Nonlinearity* **35** (2022), no. 6, 2722–2766 Zbl 1496.35324 MR 4443920

[90] M. Rosenzweig, On the rigorous derivation of the incompressible Euler equation from Newton's second law. *Lett. Math. Phys.* **113** (2023), no. 1, article no. 13 Zbl 1509.35201 MR 4540782

[91] M. Rosenzweig and S. Serfaty, Global-in-time mean-field convergence for singular Riesz-type diffusive flows. *Ann. Appl. Probab.* **33** (2023), no. 2, 754–798 Zbl 1516.35414 MR 4564418

[92] M. Rosenzweig and S. Serfaty, Modulated logarithmic Sobolev inequalities and generation of chaos. 2023, arXiv:2307.07587

[93] M. Rosenzweig and S. Serfaty, Sharp estimates for the variations of Coulomb and Riesz modulated energies, applications to supercritical mean-field limits. In preparation

[94] N. Rougerie and S. Serfaty, Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.* **69** (2016), no. 3, 519–605 Zbl 1338.82043 MR 3455593

[95] S. Salem, A gradient flow approach of uniform in time propagation of chaos for particles in double a well confinement. 2018, arXiv:1810.08946

[96] E. Sandier and S. Serfaty, 1D log gases and the renormalized energy: Crystallization at vanishing temperature. *Probab. Theory Related Fields* **162** (2015), no. 3-4, 795–846 Zbl 1327.82005 MR 3383343

[97] E. Sandier and S. Serfaty, 2D Coulomb gases and the renormalized energy. *Ann. Probab.* **43** (2015), no. 4, 2026–2083 Zbl 1328.82006 MR 3353821

[98] S. Serfaty, Mean field limits of the Gross–Pitaevskii and parabolic Ginzburg–Landau equations. *J. Amer. Math. Soc.* **30** (2017), no. 3, 713–768 Zbl 1406.35377 MR 3630086

[99] S. Serfaty, Mean field limit for Coulomb-type flows. *Duke Math. J.* **169** (2020), no. 15, 2887–2935 Zbl 1475.35341 MR 4158670

[100] S. Serfaty, Gaussian fluctuations and free energy expansion for Coulomb gases at any temperature. *Ann. Inst. Henri Poincaré Probab. Stat.* **59** (2023), no. 2, 1074–1142 Zbl 1517.82008 MR 4575026

[101] S. Serfaty and J. L. Vázquez, A mean field equation as limit of nonlinear diffusions with fractional Laplacian operators. *Calc. Var. Partial Differential Equations* **49** (2014), no. 3-4, 1091–1120 Zbl 1290.35316 MR 3168624

[102] C. Villani, *Optimal transport*. Grundlehren Math. Wiss. 338, Springer, Berlin, 2009 Zbl 1156.53003 MR 2459454

[103] H.-T. Yau, Relative entropy and hydrodynamics of Ginzburg–Landau models. *Lett. Math. Phys.* **22** (1991), no. 1, 63–80 Zbl 0725.60120 MR 1121850

[104] V. I. Yudovič, Non-stationary flows of an ideal incompressible fluid. *Ž. Vyčisl. Mat i Mat. Fiz.* **3** (1963), 1032–1066 Zbl 0147.44303 MR 158189

[105] V. V. Yurinskiĭ, Exponential inequalities for sums of random vectors. *J. Multivariate Anal.* **6** (1976), no. 4, 473–499 Zbl 0346.60001 MR 428401

[106] X. Zhou and W. Xiao, Well-posedness of a porous medium flow with fractional pressure in Sobolev spaces. *Electron. J. Differential Equations* (2017), article no. 238 Zbl 1375.35417 MR 3711191

Received 19 April 2023; revised 18 August 2023; accepted 1 September 2023.

Antonin Chodron de Courcel
Laboratoire Alexandre Grothendieck, Institut des Hautes Études Scientifiques, Université Paris-Saclay, CNRS, 35 Routes de Chartres, 91440 Bures-sur-Yvette, France; antonin.chodron-de-courcel@polytechnique.edu

Matthew Rosenzweig
Department of Mathematical Sciences, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA; mrosenz2@andrew.cmu.edu

Sylvia Serfaty
Department of Mathematics, Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, USA; serfaty@cims.nyu.edu