GLOBAL-IN-TIME MEAN-FIELD CONVERGENCE FOR SINGULAR RIESZ-TYPE DIFFUSIVE FLOWS

BY MATTHEW ROSENZWEIG[1,a], SYLVIA SERFATY[2,b]

[1]*Mathematics, Massachusetts Institute of Technology,* [a]*mrosenzw@mit.edu*

[2]*Mathematics, Courant Institute of Mathematical Sciences, New York University,* [b]*serfaty@cims.nyu.edu*

We consider the mean-field limit of systems of particles with singular interactions of the type $-\log|x|$ or $|x|^{-s}$, with $0 < s < d - 2$, and with an additive noise in dimensions $d \geq 3$. We use a modulated-energy approach to prove a quantitative convergence rate to the solution of the corresponding limiting PDE. When $s > 0$, the convergence is global in time, and it is the first such result valid for both conservative and gradient flows in a singular setting on $\mathbb{R}^d$. The proof relies on an adaptation of an argument of Carlen-Loss [9] to show a decay rate of the solution to the limiting equation, and on an improvement of the modulated-energy method developed in [17, 63, 50], making it so that all prefactors in the time derivative of the modulated energy are controlled by a decaying bound on the limiting solution.

1. Introduction.

1.1. *The problem.* We consider the first-order mean-field dynamics of stochastic interacting particle systems of the form

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

Above, $x_i^0 \in \mathbb{R}^d$ are the pairwise distinct initial positions, $\mathbb{M}$ is a $d \times d$ matrix such that

$$
(1.2) \qquad \mathbb{M}\xi \cdot \xi \leq 0 \qquad \forall \xi \in \mathbb{R}^d,
$$

and $W_1, \ldots, W_N$ are independent standard Brownian motions in $\mathbb{R}^d$, so that the noise in (1.1) is of so-called additive type. There are several choices for $\mathbb{M}$. For instance, choosing $\mathbb{M} = -\mathbb{I}$ yields *gradient-flow/dissipative* dynamics, while choosing $\mathbb{M}$ to be antisymmetric yields *Hamiltonian/conservative* dynamics. Mixed flows are also permitted. The potential g is assumed to be repulsive, which, as we shall later show in Section 4, ensures that there is a unique, global strong solution to the system (1.1). In particular, with probability one, the particles never collide. The model case for g is either a logarithmic or Riesz potential indexed by a parameter $0 \leq s < d - 2$, according to

$$
(1.3) \qquad \mathsf{g}(x) =
\begin{cases}
-\log|x|, & s = 0 \\
|x|^{-s}, & 0 < s < d - 2.
\end{cases}
$$

The above restriction on s means that we are considering potentials that are *sub-Coulombic*: their singularity is below that of the Coulomb potential, which corresponds to $s = d - 2$. As explained precisely in the next subsection, we can consider a general class of potentials g which have sub-Coulombic-type behavior.

Systems of the form (1.1) have numerous applications in the physical and life sciences as well as economics. Examples include vortices in viscous fluids [51, 14, 53, 43], models of the collective motion of microscopic organisms [56, 68, 57, 22, 40, 20], aggregation phenomena [3, 12, 42], and opinion dynamics [30, 37, 48, 70]. For more discussion on applications, we refer the reader to the survey of Jabin and Wang [34] and references therein.

Establishing the mean-field limit consists of showing the convergence in a suitable topology as $N \to \infty$ of the *empirical measure*

$$(1.4) \qquad \mu_N^t := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^t}$$

associated to a solution $\underline{x}_N^t := (x_1^t, \ldots, x_N^t)$ of the system (1.1). We remark that for fixed t, the empirical measure is a random Borel probability measure on $\mathbb{R}^d$. If the points x_i^0, which themselves depend on N, are such that μ_N^0 converges to some regular measure μ^0, then a formal calculation using Itô's lemma leads to the expectation that for $t > 0$, μ_N^t converges to the solution of the Cauchy problem with initial datum μ^0 of the limiting evolution equation

$$(1.5) \qquad \begin{cases} \partial_t \mu = -\operatorname{div}(\mu \mathbb{M} \nabla \mathsf{g} * \mu) + \sigma \Delta \mu \\ \mu|_{t=0} = \mu^0 \end{cases} \qquad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$$

as the number of particles $N \to \infty$. While the underlying N-body dynamics are stochastic, we emphasize that the equation (1.5) is completely deterministic, and the noise has been averaged out to become diffusion. Proving the convergence of the empirical measure is closely related to proving *propagation of molecular chaos* (see [23, 29, 32] and references therein): 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 $(\mu^t)^{\otimes k}$.

The mean-field problem for the system (1.1) with $\sigma > 0$ and interactions which are regular (e.g. globally Lipschitz) has been understood for many years now [44, 67, 46, 42] (see also [3, 2, 47, 38, 16] for more recent developments still in the regular case). The classical approach consists in comparing the trajectories of the original system (1.1) to those of a cooked-up symmetric particle system coupled to (1.1). Subsequent work has focused on treating the more challenging singular interactions — initially by compactness-type arguments that yield qualitative convergence [52, 54, 53, 19, 22, 40, 20, 39] and later by more quantitative methods that yield an explicit rate for propagation of chaos [31, 35, 4, 6]. To our knowledge, the best results in the literature can quantitatively prove propagation of chaos for singular interactions up to and including the Coulomb case for $d = 2$ conservative dynamics [35][1] and arbitrary $0 \le s < d$ in for repulsive, dissipative dynamics [4, 6]. Unlike the previous aforementioned works which utilize the noise in an essential way, the methods of [35, 4] allow for taking vanishing diffusion: $\sigma = \sigma_N \ge 0$, where $\sigma_N \to 0$ as $N \to \infty$. We also mention that the recent preprint [69] has gone beyond the mean-field limit and shown the convergence of the fluctuations of the empirical measure to (1.1) to a generalized Ornstein-Uhlenbeck process for singular potentials including the two-dimensional Coulomb case. These state-of-the-art works are limited to the periodic setting. In particular, up to the time of the present article, the conservative case $s > 0$ in dimensions $d \ge 2$ has been open.

When there is no noise in the system (1.1) (i.e. $\sigma = 0$), much more is known mathematically about the mean-field limit thanks to recent advances that are capable of treating the full potential case $s < d$. Approaches vary, but they all typically involve finding a good metric to

---

[1]More generally, g such that $\nabla \mathsf{g} \in \dot{W}^{-1,\infty}$, which allows for the $s = 0$ case if $d \ge 3$.

measure the distance between the empirical measure and its expected limit and then proving a Gronwall relation for the evolution of this metric. The ∞-Wasserstein metric allowed to treat the sub-Coulombic case $s < d - 2$ [28, 10]. A Wassertein-gradient-flow approach [11, 1] can also treat the one-dimensional case using the convexity of the Riesz interaction in (and only in) dimension one. The modulated-energy approach of [17, 63], inspired by the prior work [62], managed to treat the more difficult Coulomb and super-Coulombic case $d - 2 \leq s < d$ for the model potential (1.3). In very recent work by the authors together with Nguyen [50], this modulated-energy approach has been redeveloped to allow to treat the full range $s < d$ and under fairly general assumptions for the potential g. In these works, 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 the equation (1.5). More precisely, it is defined to be

$$(1.6) \qquad F_N(\underline{x}_N, \mu) := \int_{(\mathbb{R}^d)^2 \setminus \triangle} \mathsf{g}(x - y) d\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} - \mu \right)^{\otimes 2} (x, y),$$

where we remove the infinite self-interaction of each particle by excising the diagonal $\triangle$.

Contemporaneous to the development of the modulated-energy approach, Jabin and Wang [33, 35] introduced a relative-entropy method capable of treating the mean-field limit of (1.1) when the interaction is moderately singular and which works well with or without noise. The relative-entropy and modulated-energy approaches were recently combined into a *modulated free energy* method [5, 4, 6] that allows for treating the mean-field limit of (1.1) in the dissipative case, but not the conservative case, set on the torus. More precisely, the authors show in [4] how to use the modulated free energy to prove mean-field convergence of repulsive gradient flows under fairly general assumptions on the interaction, including the full (periodic) Riesz case $0 \leq s < d$, with fixed or vanishing viscosity (i.e. $\sigma > 0$ or $\sigma_N \to 0$ as $N \to \infty$). In [5, 6], they prove mean-field convergence for viscous repulsive-attractive gradient flows with repulsive part of the form in [4] and log-type attractive part. In particular, this latter work allows to treat the mean-field limit of the Patlak-Keller-Segel equation in the subcritical regime.

In this article, we show for the first time that the modulated-energy approach of [17, 63, 50] can be extended to treat the mean-field limit of (1.1) in the sub-Coulombic case $0 \leq s < d - 2$.[2] No incorporation of the entropy, as in the modulated-free-energy approach of [5, 4, 6] is needed. Moreover, the modulated-energy approach is well-suited for exploiting the dissipation of the limiting equation (1.5) to obtain rates of convergence in N which are *uniform* over the entire interval $[0, \infty)$. In other words, mean-field convergence holds globally in time. At the time of completion of this manuscript, this is, to the best of our knowledge, the first instance of such a result for singular potentials. Obtaining a uniform-in-time convergence is important in both theory and practice — for instance, 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 other interaction kernels (for instance, see [13, 45, 60]).

Lastly, we mention that previous uses of the modulated energy in the stochastic setting [58, 50] were limited to the case of multiplicative noise, which behaves very differently in the limit as $N \to \infty$. Most notably, the limiting evolution equation is stochastic.

---

[2]Our ability to use the modulated-energy approach in the sub-Coulombic case crucially relies on our recent work [50] with Nguyen, since the previous works [17, 63] could only treat this way the Coulomb/super-Coulombic case.

4

1.2. *Formal idea.* Let us sketch our main proof in the model case (1.3). For simplicity of exposition, let us also restrict ourselves to the simpler range $d - 4 < s < d - 2$. We note that the modulated energy (1.6) is a real-valued continuous stochastic process. Formally by Itô's lemma (see Section 6 for the rigorous computation), it satisfies the stochastic differential inequality (cf. [63, Lemma 2.1])

$$(1.7) \quad \frac{d}{dt} F_N(\underline{x}_N^t, \mu^t) \leq \int_{(\mathbb{R}^d)^2 \setminus \triangle} \nabla \mathsf{g}(x - y) \cdot \left(u^t(x) - u^t(y) \right) d(\mu_N^t - \mu^t)^{\otimes 2}(x, y)$$

$$+ \sigma \int_{(\mathbb{R}^d)^2 \setminus \triangle} \Delta \mathsf{g}(x - y) d(\mu_N^t - \mu^t)^{\otimes 2}(x, y)$$

$$+ \frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^d \setminus \{x_i^t\}} \nabla \mathsf{g}(x_i^t - y) d(\mu_N^t - \mu^t)(y) \cdot \dot{W}_i^t,$$

where we have set $u^t := \mathbb{M} \nabla \mathsf{g} * \mu^t$. The third term in the right-hand side (formally) has zero expectation and may be ignored for the purposes of this discussion. The first term is the contribution of the drift and also appears in the deterministic case, but the second term is new and due to the nonzero quadratic variation of Brownian motion. Observe that

$$(1.8) \quad \Delta \mathsf{g}(x - y) = -(d - s - 2)|x - y|^{-s-2}.$$

Since $0 \leq s < d - 2$ by assumption, we see that $\Delta \mathsf{g}$ is superharmonic and equals a constant multiple of $-\tilde{\mathsf{g}}$, where $\tilde{\mathsf{g}}$ is the kernel of the Riesz potential operator $(-\Delta)^{\frac{s+2-d}{2}}$. We would like to conclude that the second term in the right-hand side of (1.7) is nonpositive by Plancherel's theorem and therefore may be discarded, but the excision of the diagonal $\triangle$ obstructs this reasoning. Fortunately, prior work of the second author [63, Proposition 3.3] gives the lower bound

$$(1.9) \quad \int_{(\mathbb{R}^d)^2 \setminus \triangle} \tilde{\mathsf{g}}(x - y) d\left(\mu_N^t - \mu^t\right)^{\otimes 2}(x, y) \geq -\frac{1}{N^2} \sum_{i=1}^{N} \tilde{\mathsf{g}}(\eta_i) - \frac{C\|\mu^t\|_{L^\infty}}{N} \sum_{i=1}^{N} \eta_i^{d-s-2}$$

for all choices of parameters $\eta_i > 0$. Here, C is a constant that just depends on s, d. We emphasize that (1.9) is a functional inequality which holds independently of any underlying dynamics. The choice of η_i that balances the decay in N between the two terms in the right-hand side of the inequality (1.9) is the typical interparticle distance $N^{-1/d}$. Since the L^∞ norm of μ^t is a source of decay and we wish to distribute it between terms, we instead choose

$$(1.10) \quad \eta_i = (\|\mu^t\|_{L^\infty} N)^{-1/d} \qquad \forall 1 \leq i \leq N,$$

so that the right-hand side of (1.9) is bounded from below by

$$(1.11) \quad -C\sigma \|\mu^t\|_{L^\infty}^{\frac{s+2}{d}} N^{-\frac{d-s}{d}},$$

providing a bound from above for the corresponding term in (1.7). Note that since μ^t is time-dependent, our choice for η_i above depends on time, a trick previously used by the first author [59, 58] to study the mean-field limit for point vortices with possible multiplicative noise when μ^t belongs to a function space which is invariant or critical under the scaling of the equation.

It remains to consider the first term in the right-hand side of (1.7), which, as previously mentioned, also appears in the deterministic case. This expression has the structure of a commutator which has been *renormalized* through the exclusion of diagonal in order to accommodate the singularity of the Dirac masses. As shown in [59, 50], one can make this commutator

intuition rigorous (see Propositions 5.7 and 5.15 below) and, revisiting the estimates there together with some elementary potential analysis, we are able to optimize the dependence in $\|\mu\|_{L^\infty}$ of the estimate and show the pathwise and pointwise-in-time bound

$$(1.12) \quad \left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} \nabla \mathbf{g}(x-y) \cdot \left(u^t(x) - u^t(y) \right) d(\mu_N^t - \mu^t)^{\otimes 2}(x,y) \right|$$

$$\leq C\|\mu^t\|_{L^\infty}^{\frac{s+2}{d}} \left(|F_N(\underline{x}_N^t, \mu^t)| + (1 + \|\mu^t\|_{L^\infty})N^{-\beta} \right),$$

where again $C, \beta > 0$ are constants depending only s, d.

Now taking expectations of both sides of (1.7), integrating with respect to time, and using Remark 5.2 below to control $|F_N(\underline{x}_N^t, \mu^t)|$ by $F_N(\underline{x}_N^t, \mu^t)$, we find that

$$(1.13)$$

$$\mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq |F_N(\underline{x}_N^0, \mu^0)| + C\|\mu^t\|_{L^\infty}^{\frac{s}{d}} N^{-\beta} + C \int_0^t \|\mu^\tau\|_{L^\infty}^{\frac{s+2}{d}} \mathbb{E}(|F_N(\underline{x}_N^\tau, \mu^\tau)|) d\tau$$

$$+ C\sigma N^{-\beta} \int_0^t \|\mu^\tau\|_{L^\infty}^{\frac{s+2}{d}} d\tau.$$

The structure of the right-hand side of (1.13) allows us to leverage the decay rate of the solution to (1.5). In Proposition 3.8, we show the decay rate

$$\|\mu^t\|_{L^\infty} \leq \min\{C(\sigma t)^{-\frac{d}{2}}, \|\mu^0\|_{L^\infty}\}.$$

This is done by revisiting work of Carlen and Loss [9] on the optimal decay of nonlinear viscously damped conservation laws, which was essentially restricted to divergence-free vector fields, and adapting it to treat the case of (1.5).

Once this is done, an application of the Gronwall-Bellman lemma to (1.13) yields a uniform-in-time bound if $s > 0$, while if $s = 0$, long-range effects only allow us to obtain an $O(t^{\frac{C}{\sigma}})$ growth estimate.

1.3. *Assumptions on the potential and main results.*   We now state the precise assumptions for the class of interaction potentials we consider. This class corresponds to the sub-Coulombic sub-class of the larger class of potentials considered by the authors in collaboration with Nguyen in [50]. We remark that Bresch-Jabin-Wang [4] were the first to prove mean-field limits by placing assumptions of suitable pointwise bounds for the interaction kernel and its Fourier transform, as opposed to assuming Riesz or log potentials. In the statement below and throughout this article, the notation $\mathbf{1}_{(\cdot)}$ denotes the indicator function for the condition $(\cdot)$.

For $d \geq 3$ and $0 \leq s < d - 2$, we assume the following:

(i)

$$\mathbf{g}(x) = \mathbf{g}(-x)$$

(ii)

$$\lim_{x \to 0} \mathbf{g}(x) = \infty$$

(iii)

$$\exists r_0 > 0 \text{ such that } \quad \Delta \mathbf{g} \leq 0 \quad \text{in } B(0, r_0)^3$$

---

[3]Here, we mean $\mathbf{g}$ is superharmonic in $B(0, r_0)$ in the sense of distributions, which implies that $\Delta \mathbf{g}(x) \leq 0$ for almost every $x \in B(0, r_0)$.

6

(iv)

$$\forall k \geq 0, \qquad |\nabla^{\otimes k}\mathsf{g}(x)| \leq C\left(\frac{1}{|x|^{s+k}} + |\log|x||\mathbf{1}_{s=k=0}\right) \quad \forall x \in \mathbb{R}^d \setminus \{0\}$$

(v)

$$|x||\nabla\mathsf{g}(x)| + |x|^2|\nabla^{\otimes 2}\mathsf{g}(x)| \leq C\mathsf{g}(x) \quad \forall x \in B(0,r_0)\setminus\{0\}$$

(vi)

$$\frac{C_1}{|\xi|^{d-s}} \leq \hat{\mathsf{g}}(\xi) \leq \frac{C_2}{|\xi|^{d-s}} \quad \forall \xi \in \mathbb{R}^d \setminus \{0\}$$

where $\hat{\cdot}$ denotes the Fourier transform.

(vii)

$$\begin{cases} \exists c_s < 1 \text{ such that } & \mathsf{g}(x) < c_s\mathsf{g}(y) \quad \forall x,y \in B(0,r_0) \text{ with } |y| \geq 2|x|, \qquad s > 0 \\ \exists c_0 > 0 \text{ such that } & \mathsf{g}(x) - \mathsf{g}(y) \geq c_0 \quad \forall x,y \in B(0,r_0) \text{ with } |y| \geq 2|x|, \quad s = 0 \end{cases}$$

(viii) If $d - 4 < s < d - 2$, then we also assume that there is an $m \in \mathbb{N}$ and $\mathsf{G} : \mathbb{R}^{d+m} \to \mathbb{R}$ such that

$$-\Delta\mathsf{g}(x) = \mathsf{G}(x,0) \quad \forall (x,0) \in \mathbb{R}^{d+m}$$

(1.14) $$\mathsf{G}(X) = \mathsf{G}(-X)$$

(1.15) $$\exists r_0 > 0 \text{ such that } \quad \Delta\mathsf{G}(X) \leq 0 \quad \text{in } B(0,r_0) \subset \mathbb{R}^{d+m}$$

(1.16) $$\forall k \geq 0, \quad |\nabla^{\otimes k}\mathsf{G}(X)| \leq \frac{C}{|X|^{s+2+k}} \quad \forall X \in B(0,r_0)$$

(1.17) $$\hat{\mathsf{G}}(\Xi) \geq 0 \quad \forall \Xi \in \mathbb{R}^{d+m} \setminus \{0\}.$$

(ix) In the cases $s = d - 2k \geq 0$, for some positive integer k, we also assume that the $(\mathbb{R}^d)^{\otimes(2k+2)}$-valued kernel

$$\mathsf{k}(x - y) := (x - y) \otimes \nabla^{\otimes(2k+1)}\mathsf{g}(x - y)$$

is associated to a Calderón-Zygmund operator.[4]

(x) In all cases,

$$\mathbb{M} : \nabla^{\otimes 2}\mathsf{g}(x) \geq 0 \quad \forall x \in \mathbb{R}^d \setminus \{0\},$$

where : denotes the Frobenius inner product.

We shall say that any potential g satisfying assumptions (i) – (x) is an *admissible potential*. We refer to [50, Subsection 1.3] for a discussion of the types of potentials permitted under these assumptions. Compared to that work, only assumptions (viii) and (x) are new. The former is to ensure that our modulated-energy method can be applied to g̃ and thus to the diffusion term in (1.7), while the latter is to ensure that the solutions of (1.5) satisfy the temporal decay bounds of the heat equation. Note that (x) is automatically satisfied if $\mathbb{M}$ is antisymmetric. Additionally, if $\mathbb{M} = -\mathbb{I}$ so that we consider gradient-flow dynamics, then

---

[4]Sufficient and necessary conditions for this Calderón-Zygmund property are explained in [24, Section 5.4]. The reader may check that this condition is satisfied if g is the Riesz potential $|x|^{2k-d}$.

(x) amounts to requiring that g is globally superharmonic (i.e. $r_0 = \infty$ in (iii)). In general, though, we do not require global superharmonicity except where explicitly stated.

We assume that we are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, \mathbb{P})$ on which a countable collection of independent standard d-dimensional Brownian motions $(W_n)_{n=1}^{\infty}$ are defined. Moreover, $(\mathcal{F}_t)_{t\geq0}$ is the complete filtration generated by the Brownian motions. All stochastic processes considered in this article are defined on this probability space.

Let $\underline{x}_N^0 \in (\mathbb{R}^d)^N$ be an N-tuple of distinct points in $\mathbb{R}^d$. As shown in Proposition 4.6 (more generally Section 4), there exists a unique, global strong solution $\underline{x}_N$ to the Cauchy problem for (1.1). Moreover, with probability one, the particles x_i^t and x_j^t never collide on the interval $[0, \infty)$. Let $\mu^0 \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty$ be a probability measure with L^∞ density with respect to Lebesgue measure. We abuse notation here and throughout the article by using the same symbol to denote both the measure and its density.

As shown in Proposition 3.8 (more generally Section 3), there is a unique, global solution to the Cauchy problem for (1.5) in the class $C([0, \infty); \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$. In the case of logarithmic interactions (i.e. $s = 0$), we also assume that μ^0 satisfies the logarithmic growth condition $\int_{\mathbb{R}^d} \log(1 + |x|)d\mu^0(x)$, which is propagated locally uniformly by the evolution (see Remark 3.6).

Since $\underline{x}_N$ is stochastic, $\{F_N(\underline{x}_N^t, \mu^t)\}_{t\geq0}$ is a real-valued stochastic process. It is straightforward to check from the non-collision of the particles and Hölder's inequality that $F_N(\underline{x}_N^t, \mu^t)$ is almost surely finite on the interval $[0, \infty)$ and a continuous process. Our main theorem is a quantitative estimate for the expected magnitude of $F_N(\underline{x}_N^t, \mu^t)$.

The first result of this article is the following functional inequality for the expected magnitude of the modulated energy. In the case $0 < s < d - 2$, we get a linear growth estimate, while in the case $s = 0$, we have superlinear growth of size $O(t^{\frac{\sigma+C}{\sigma}})$ as $t \to \infty$.

THEOREM 1.1. *Let $d \geq 3$, $0 \leq s < d - 2$, and $\sigma > 0$. Let $\underline{x}_N$ be a solution to the system (1.1), and let $\mu \in C([0, \infty); \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$ be a solution to the PDE (1.5). If $s = 0$, also assume that $\int_{\mathbb{R}^d} \log(1 + |x|)d\mu^0(x) < \infty$. There exists a constant $C > 0$ depending only $s, d, \sigma, \|\mu\|_{L^\infty}$, and the potential g through assumptions (i) − (ix) and an exponent $\beta > 0$ depending only s, d, such that following holds. For all $t \geq 0$ and all N sufficiently large depending on $\|\mu^0\|_{L^\infty}$, we have that*

$$(1.18) \qquad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq C\Big(1 + t + t^{\frac{\sigma+C}{\sigma}}\mathbf{1}_{s=0}\Big)\Big(|F_N(\underline{x}_N^0, \mu^0)| + N^{-\beta}\Big).$$

Assume now that $r_0 = \infty$ in assumption (iii). In other words, the potential g is globally superharmonic, as opposed to just in a neighborhood of the origin. The second result of this article is a functional inequality for the expected magnitude of the modulated energy which yields a global bound in the case $0 < s < d - 2$. In the case $s = 0$, we have an almost global bound, in the sense that the growth is $O(t^{\frac{1}{\sigma}^+})$ as $t \to \infty$, which can be arbitrarily small by choosing the diffusion strength σ arbitrarily large.

THEOREM 1.2. *Impose the same assumptions as Theorem 1.1 with the additional condition that $r_0 = \infty$. Choose any exponent $\frac{d}{s+2} < p \leq \infty$. Then there exist constants $C, C_p > 0$ depending on $s, d, \sigma, \|\mu\|_{L^\infty}$, and the potential g through assumptions (i) − (ix) and exponents β_p depending on s, d, such that the following holds. For all $t \geq 0$ and all N sufficiently large depending on $\|\mu^0\|_{L^\infty}$, we have that*

$$(1.19) \qquad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq C\Big(1 + \Big(t^{\frac{1}{\sigma}}\log(1 + t)\Big)\mathbf{1}_{s=0}\Big)\Big(|F_N(\underline{x}_N^0, \mu^0)| + C_p N^{-\beta_p}\Big).$$

We close this subsection with some remarks on the statements of Theorems 1.1 and 1.2, further extensions, and interesting questions for future work.

REMARK 1.3. An examination of Sections 7.1 and 7.2 will reveal to the interested reader the precise dependence of the constants C, β (C_p, β_p) on the norms of μ, s, d (on p), and other underlying parameters. We have omitted the explicit dependence and simplified the statements of our final bounds (see (7.10) and (7.15)) in order to make the results more accessible to the reader.

Additionally, we have not attempted to optimize the regularity/integrability assumptions for μ. One can show that in the case $s > 0$, the linear and global bounds of Theorems 1.1 and 1.2, respectively, still hold if we replace the L^∞ assumption with $\mu \in L^p$ for finite p sufficiently large depending on s, d. This, though, comes at the cost of slower decay in N.

REMARK 1.4. Sufficient conditions for $F_N(\underline{x}_N^0, \mu^0)$ to vanish as $N \to \infty$ are that the energy of (1.1) converges to the energy of (1.5) and that $\mu_N^0 \overset{*}{\rightharpoonup} \mu^0$ in the weak-* topology for $\mathcal{P}(\mathbb{R}^d)$. See [17, Remark 1.2(c)] for more details.

REMARK 1.5. It is well-known [50, Proposition 2.4] that the modulated energy $F_N(\underline{x}_N, \mu)$ controls convergence in negative-order Sobolev spaces. Note that since we are restricted to the sub-Coulombic setting, the extension implicit in the cited proposition can be ignored. Consequently, Theorems 1.1 and 1.2 yield a quantitative bound for the expected squared H^s norm of $\mu_N - \mu$, for $s < -\frac{d+2}{2}$, of the form

$$(1.20) \qquad \mathbb{E}\left(\left\| \mu_N^t - \mu^t \right\|_{H^s}^2 \right) \leq C \rho(t) \left(|F_N(\underline{x}_N^0, \mu^0)| + N^{-\beta} \right),$$

where $\rho(t)$ is the time factor from either Theorem 1.1 or Theorem 1.2. From this Sobolev convergence and standard arguments (see [29, Section 1]), one deduces convergence in law of the empirical measure μ_N to μ.

We can also deduce an explicit rate for propagation of chaos for the system (1.1). Indeed, suppose that $\underline{x}_N^0$ are initially distributed in $(\mathbb{R}^d)^N$ according to some probability density f_N^0. Let f_N^t denote the law of $\underline{x}_N^t$, and let $f_{N,k}^t$ denote the k-particle marginal of f_N^t. Then using for instance [61, (7.21)], we see that for any symmetric test function $\varphi \in C_c^\infty((\mathbb{R}^d)^k)$,

$$(1.21) \qquad \left| \int_{(\mathbb{R}^d)^k} \varphi d\left(f_{N,k}^t - (\mu^t)^{\otimes k} \right) \right|$$

$$\leq Ck \sup_{\underline{x}_{k-1} \in (\mathbb{R}^d)^{k-1}} \| \varphi(\underline{x}_{k-1}, \cdot) \|_{H^{-s}(\mathbb{R}^d)} \int_{(\mathbb{R}^d)^N} \| \mu_N^t - \mu^t \|_{H^s(\mathbb{R}^d)} df_N^0,$$

for any $s < -\frac{d+2}{2}$. Combining (1.21) with (1.20) and using duality now yields an explicit rate for propagation of chaos in H^s norm.

REMARK 1.6. It is an interesting problem to obtain analogues of Theorems 1.1 and 1.2 when a confining potential V is added to the right-hand sides of (1.1), (1.5). In this case, one has a nontrivial equilibrium for the equation (1.5) as $t \to \infty$, which clearly breaks our proof, in particular the Carlen-Loss argument used to prove Proposition 3.8 below. But we still might expect that the difference $\mu_N^t - \mu^t$ decays to zero in a suitable topology as $t \to \infty$ and that our argument may be salvaged by incorporating the long-time equilibrium for μ^t. We plan to investigate this problem in future work. Note that by using so-called similarity variables (see [21, Section 1]), one can obtain a local-in-time bound for the modulated energy

from Theorems 1.1 and 1.2 in the model interaction case (1.3) and with a quadratic potential. In fact, using our prior work [50] and by appropriately modifying the similarity coordinates (see [64, Section 8]), one can also obtain a local-in-time bound for the log or Riesz case $0 \leq s < d$ in any dimension $d \geq 1$ with a quadratic potential and without noise.

1.4. *Comparison with prior results.* At the time of completion of this manuscript, our Theorem 1.2 is, to the best of our knowledge, the first time that a quantitative rate of convergence for the mean-field limit of (1.1) has been shown to hold uniformly on the interval $[0, \infty)$. Additionally, in the case $s = 0$, the rate of convergence as $N \to \infty$ is (up to a $\log N$ factor) optimal. Existing results in the literature for singular potentials (e.g. [35, 4]) have at least an exponential growth in time due to a reliance on a Gronwall-type argument without exploiting the dissipation of the limiting equation. Those works (e.g. [42, 12, 18, 16]) that do have a uniform convergence result are restricted to regular potentials and require confinement.

Additionally, our theorem is at the level of the empirical measure for the original SDE dynamics for (1.1), as opposed to their associated *Liouville/forward Kolmogorov equations* for the joint law of the process $\underline{x}_N^t = (x_1^t, \ldots, x_N^t)$,

$$(1.22) \qquad \partial_t f_N = -\sum_{i=1}^N \text{div}_{x_i} \left(f_N \frac{1}{N} \sum_{1 \leq i \neq j \leq N} \mathbb{M} \nabla \mathsf{g}(x_i - x_j) \right) + \sigma \sum_{i=1}^N \Delta_{x_i} f_N.$$

Namely, no randomization of the initial data is needed, although as discussed in Remark 1.5, our result implies convergence of this form as well. This stands in contrast to previous work [35, 4] whose starting point is the Liouville equation (1.22).

The costs of the strong bounds we obtain with Theorems 1.1 and 1.2 are two-fold. First, we need somewhat stronger assumptions on the potential g than in [35, 4]–especially the latter work. The reader may find a detailed comparison in [50, Subsection 1.3]. Second, and more importantly, our results are limited to the sub-Coulombic range $0 \leq s < d - 2$ and dimensions $d \geq 3$. The Coulomb case is barely just out of the reach. Indeed, the diligent reader will note that if g is the Coulomb potential, then $\Delta \mathsf{g} = -\delta$, which is obviously no longer a function. Thus, the argument we described in the previous subsection to bound the second term in the right-hand side of (1.7) no longer applies. The situation is even worse when $s > d - 2$ as $\Delta \mathsf{g} > 0$, meaning what was previously a dissipation term should now cause the modulated energy to grow in time. Having resolved the case $0 \leq s < d - 2$ for dimensions $d \geq 3$, it is now only an open problem to prove the mean-field limit of (1.1) in the conservative case when $d - 2 \leq s < d$ and $d \geq 2$, with the exception of $s = 0$ and $d = 2$ in the periodic setting.

During the final proofreading of the manuscript of this article, Guillin, Le Bris, and Monmarché posted to the arXiv their preprint [27] showing uniform-in-time propagation of chaos in L^1 norm in the periodic setting $\mathbb{T}^d$ for singular interactions in the range $0 \leq s \leq d - 2$ and $d \geq 2$ with $N^{-\frac{1}{2}}$ rate. In particular, they can treat the viscous vortex model corresponding to the $d = 2$ Coulomb case. Their method is very different from ours, as it is based on the relative-entropy method of Jabin-Wang [35]. Moreover, they assume random initial data and work at the level of the Liouville equation, and their interaction kernel is assumed to be integrable, have zero distributional divergence, and equal to the divergence of an L^∞ matrix field. Due to the reliance on the method of [35], the work of Guillin *et al.* does not seem capable of treating potentials g with $|\nabla \mathsf{g}|$ having singularity much worse than $|x|^{-1}$ (e.g. g has a $\log$ singularity). Additionally, their result is limited to the periodic setting, due to a need for compactness of domain, and to conservative flows. We also note that they impose much stronger regularity assumptions on their solutions to (1.5) than we do.

1.5. *Organization of article.* In Section 2, we review some estimates for Riesz potential operators and the interaction potential g that are used frequently in the paper.

Section 3 is devoted to the study of the limiting equation (1.5), showing that it is globally well-posed and moreover the L^p norms of solutions satisfy the optimal decay bounds (see Propositions 3.1 and 3.8).

In Section 4, we show that the N-particle system (1.1) has well-defined dynamics (see Proposition 4.6) in the sense that there exists a unique strong solution and with probability one, the particles never collide. We also introduce in this section a truncation and stopping time procedure that will be used again later in Section 6.

In Section 5, we review properties of the modulated energy and renormalized commutator estimates from the perspective of our recent joint work [50]. We also prove refinements (see Section 5.2) of the results from that work in the case where g is globally superharmonic.

Finally, in Sections 6 and 7, we prove our main results, Theorems 1.1 and 1.2. Section 6 gives the rigorous computation of the Itô equation (cf. (1.7)) satisfied by the process $F_N(\underline{x}_N^t, \mu^t)$. The main result is Proposition 6.3, which establishes an integral inequality for $\mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|)$. Using this inequality together with the decay bound of Proposition 3.8 and the results of Section 5, we close our Gronwall argument in Section 7, completing the proofs of Theorems 1.1 and 1.2.

1.6. *Notation.* We close the introduction with the basic notation used throughout the article without further comment.

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 $A \leq CB$. If $A \lesssim B$ and $B \lesssim A$, we write $A \sim B$. To emphasize the dependence of the constant C on some parameter p, we sometimes write $A \lesssim_p B$ or $A \sim_p B$. Also $(\cdot)_+$ denotes the positive part of a number.

We denote the natural numbers excluding zero by $\mathbb{N}$ and including zero by $\mathbb{N}_0$. Similarly, we denote the positive real numbers by $\mathbb{R}_+$. Given $N \in \mathbb{N}$ and points $x_{1,N}, \ldots, x_{N,N}$ in some set X, we will write $\underline{x}_N$ to denote the N-tuple $(x_{1,N}, \ldots, x_{N,N})$. Given $x \in \mathbb{R}^n$ and $r > 0$, we denote the ball and sphere centered at x of radius r by $B(x,r)$ and $\partial B(x,r)$, respectively. Given a function f, we denote the support of f by $\mathrm{supp}\, f$. We use the notation $\nabla^{\otimes k} f$ to denote the tensor with components $\partial_{i_1 \cdots i_k}^k f$.

We denote the space of Borel probability measures on $\mathbb{R}^n$ by $\mathcal{P}(\mathbb{R}^n)$. When μ is in fact absolutely continuous with respect to Lebesgue measure, we shall abuse notation by writing μ for both the measure and its density function. We denote the Banach space of complex-valued continuous, bounded functions on $\mathbb{R}^n$ by $C(\mathbb{R}^n)$ equipped with the uniform norm $\|\cdot\|_\infty$. More generally, we denote the Banach space of k-times continuously differentiable functions with bounded derivatives up to order k by $C^k(\mathbb{R}^n)$ equipped with the natural norm, and we define $C^\infty := \bigcap_{k=1}^\infty C^k$. We denote the subspace of smooth functions with compact support by $C_c^\infty(\mathbb{R}^n)$. We denote the Schwartz space of functions by $\mathcal{S}(\mathbb{R}^n)$ and the space of tempered distributions by $\mathcal{S}'(\mathbb{R}^n)$.

2. Potential estimates. We review some facts about Riesz potential estimates. For a more thorough discussion, we refer to [65, 66, 24, 25].

Let $d \geq 1$ For $s > -d$, we define the Fourier multiplier $|\nabla|^s = (-\Delta)^{s/2}$ by

$$(2.1) \qquad ((-\Delta)^{s/2} f) := (|\cdot|^s \widehat{f}(\cdot))^\vee, \qquad f \in \mathcal{S}(\mathbb{R}^d).$$

Since, for $s \in (-d, 0)$, the inverse Fourier transform of $|\xi|^s$ is the tempered distribution

$$(2.2) \qquad \frac{2^s \Gamma(\frac{d+s}{2})}{\pi^{\frac{d}{2}} \Gamma(-\frac{s}{2})} |x|^{-s-d},$$

it follows that

$$(2.3) \qquad ((-\Delta)^{s/2}f)(x) = \frac{2^s\Gamma(\frac{d+s}{2})}{\pi^{\frac{d}{2}}\Gamma(-\frac{s}{2})}\int_{\mathbb{R}^d}\frac{f(y)}{|x-y|^{s+d}}dy \qquad \forall x \in \mathbb{R}^d.$$

For $s \in (0,d)$, we define the *Riesz potential operator* of order s by $\mathcal{I}_s := (-\Delta)^{-s/2}$ on $\mathcal{S}(\mathbb{R}^d)$.

REMARK 2.1.   If $0 < s < d$, we see that the model potential (1.3) corresponds to g is a constant times the kernel of $\mathcal{I}_{d-s}$. If $s = 0$, then g is a multiple of the inverse Fourier transform of the tempered distribution P.V.$|\xi|^{-d} - c\delta_0(\xi)$, for some constant c. The subtraction of the Dirac mass is to cure the singularity of $|\xi|^{-d}$ near the origin.

$\mathcal{I}_s$ extends to a well-defined operator on any L^p space, the extension also denoted by $\mathcal{I}_s$ with an abuse of notation, as a consequence of the *Hardy-Littlewood-Sobolev (HLS) lemma*.

PROPOSITION 2.2.   *Let $d \geq 1$, $s \in (0,d)$, and $1 < p < q < \infty$ satisfy the relation*

$$(2.4) \qquad \frac{1}{p} - \frac{1}{q} = \frac{s}{d}.$$

Then for all $f \in \mathcal{S}(\mathbb{R}^d)$,

$$(2.5) \qquad \|\mathcal{I}_s(f)\|_{L^q} \lesssim \|f\|_{L^p},$$

$$(2.6) \qquad \|\mathcal{I}_s(f)\|_{L^{\frac{d}{d-s},\infty}} \lesssim \|f\|_{L^1},$$

where $L^{r,\infty}$ denotes the weak-L^r space. Consequently, $\mathcal{I}_s$ has a unique extension to L^p, for all $1 \leq p < \infty$.

The next lemma allows us to control the L^∞ norm of $\mathcal{I}_s(f)$ in terms of the L^1 norm and L^p norm, for some p depending on s, d. We omit the proof as it is a straightforward application of Hölder's inequality.

LEMMA 2.3.   *Let $d \geq 1$, $s \in (0,d)$, and $\frac{d}{s} < p \leq \infty$. Then for all $f \in L^1(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$, it holds that $\mathcal{I}_s(f) \in C(\mathbb{R}^d)$ and*

$$(2.7) \qquad \|\mathcal{I}_s(f)\|_{L^\infty} \lesssim \|f\|_{L^1}^{1-\frac{d-s}{d(1-\frac{1}{p})}}\|f\|_{L^p}^{\frac{d-s}{d(1-\frac{1}{p})}}.$$

When the convolution with a Riesz potential is replaced by convolution with the log potential, we have an analogue of Lemma 2.3.

LEMMA 2.4.   *Let $d \geq 1$ and $1 < p \leq \infty$. For all $f \in L^p(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d}\log(1+|x|)|f|dx < \infty$, it holds that $\log|\cdot| * f \in C_{loc}(\mathbb{R}^d)$ and*

$$(2.8) \quad |(\log|\cdot| * f)(x)| \lesssim (1+|x|)^{\frac{d(p-1)}{p}}\log(1+|x|)$$

$$+ \int_{\mathbb{R}^d}\log(1+|y|)|f(y)|dy \qquad \forall x \in \mathbb{R}^d.$$

REMARK 2.5.   If $0 \leq s < d$, then for any integer $1 \leq k < d - s$, assumption (iv) implies that $|\nabla^{\otimes k}\mathbf{g}|$ is bounded from above by a constant multiple of the kernel of $\mathcal{I}_{d-s-k}$. Lemma 2.3 implies that

$$(2.9) \qquad \|\nabla^{\otimes k}\mathbf{g} * f\|_{L^\infty} \lesssim \|f\|_{L^1}^{1-\frac{s+k}{d}}\|f\|_{L^\infty}^{\frac{s+k}{d}}, \qquad f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d).$$

We shall use this estimate frequently in the sequel.

3. The mean-field equation. We start by discussing the well-posedness of the Cauchy problem for and asymptotic decay of solutions to the mean-field PDE (1.5). The latter property is strictly a consequence of the diffusion and is the crucial ingredient to obtain a rate of convergence for the mean-field limit beyond the standard exponential bound given by the Gronwall-Bellman lemma. The results of this section are perhaps mathematical folklore. We present them not for claim for originality but since we could not find them conveniently stated in the literature.

3.1. *Local well-posedness.* We start by proving local well-posedness for the equation (1.5) for initial data μ^0 in the Banach space $X := L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. That is, we show existence, uniqueness, and continuous dependence on the initial data. The proof proceeds by a contraction mapping argument for the mild formulation of (1.5). In the next subsection, we will upgrade this local well-posedness to global well-posedness through estimates for the temporal decay of the L^p norms of the solution.

Let us introduce the mild formulation of equations (1.5), on which we base our notion of solution. With $e^{t\Delta}$ denoting the heat flow, we write

$$(3.1) \qquad \mu^t = e^{t\sigma\Delta}\mu^0 - \int_0^t e^{(t-\kappa)\sigma\Delta} \operatorname{div}(\mu^\kappa \mathbb{M}\nabla g * \mu^\kappa) d\kappa.$$

PROPOSITION 3.1. *Suppose $d \geq 3$ and $0 \leq s < d-2$. Let $\mu^0 \in X := L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, and let $R > 0$ be such that $\|\mu^0\|_X \leq R$. Then there exists a unique solution $\mu \in C([0,T];X)$ to equation (3.1) such that $T \sim \sigma R^{-2}$ and*

$$(3.2) \qquad \|\mu\|_{C([0,T];X)} \leq 2R.$$

Moreover, if $\|\mu_1^0\|_X, \|\mu_2^0\|_X \leq R$, then there exists $T' \sim \sigma R^{-2}$ such that their associated solutions μ_1, μ_2 satisfy the bound

$$(3.3) \qquad \|\mu_1 - \mu_2\|_{C([0,T'];X)} \leq 2\|\mu_1^0 - \mu_2^0\|_X.$$

PROOF. Let $R > 0$, let $\mu^0 \in X$ with $\|\mu^0\|_X \leq R$, and let $\mathcal{T}$ denote the map

$$(3.4) \qquad \mu \mapsto e^{(\cdot)\sigma\Delta}\mu^0 - \int_0^{(\cdot)} e^{(\cdot-\kappa)\sigma\Delta} \operatorname{div}(\mu^\kappa \mathbb{M}\nabla g * \mu^\kappa) d\kappa.$$

We claim that for any appropriate choice of T, this map is a contraction on the closed ball of radius $2R$ in $C([0,T];X)$. Indeed,

$$\|\mathcal{T}\mu\|_{C([0,T];X)} \leq \|\mu^0\|_X + \int_0^T \|e^{(\cdot-\kappa)\sigma\Delta} \operatorname{div}(\mu^\kappa \mathbb{M}\nabla g * \mu^\kappa)\|_{C([0,T];X)} d\kappa$$

$$\leq R + C \int_0^T (\sigma\kappa)^{-1/2} \|\mu \mathbb{M}\nabla g * \mu\|_{C([0,T];X)} d\kappa$$

$$(3.5) \qquad \leq R + C(T/\sigma)^{1/2} \|\mu \mathbb{M}\nabla g * \mu\|_{C([0,T];X)}.$$

By Hölder's inequality and Remark 2.5,

$$(3.6) \qquad \|\mu^t \mathbb{M}\nabla g * \mu^t\|_{L^p} \leq \|\mu^t\|_{L^p} \|\nabla g * \mu^t\|_{L^\infty} \lesssim \|\mu^t\|_{L^p} \|\mu^t\|_{L^1}^{1-\frac{s+1}{d}} \|\mu^t\|_{L^\infty}^{\frac{s+1}{d}}.$$

for any exponent $1 \leq p \leq \infty$. Consequently, if $\|\mu\|_{C([0,T];X)} \leq 2R$, then

$$(3.7) \qquad \|\mu \mathbb{M}\nabla g * \mu\|_{C([0,T];X)} \lesssim R^2,$$

which implies that

$$\|\mathcal{T}\mu\|_{C([0,T];X)} \leq R + C(T/\sigma)^{1/2}R^2, \tag{3.8}$$

for some constant C depending only on s, d and the potential g through assumption (iv). Similarly, for $\mu_1, \mu_2 \in B_{2R} \subset C([0,T];X)$,

$$\|\mathcal{T}\mu_1 - \mathcal{T}\mu_2\|_{C([0,T];X)} \lesssim (T/\sigma)^{1/2}\Big(\|(\mu_1 - \mu_2)\nabla\mathbb{M}g * \mu_1\|_{C([0,T];X)} \tag{3.9}$$
$$+ \|\mu_2\mathbb{M}\nabla g * (\mu_1 - \mu_2)\|_{C([0,T];X)}\Big).$$

Using inequality (3.6), the preceding right-hand side is $\lesssim$

$$(T/\sigma)^{1/2}R\|\mu_1 - \mu_2\|_{C([0,T];X)}. \tag{3.10}$$

After a little bookkeeping, we see that there is a constant $C > 0$ such that if

$$C(T/\sigma)^{1/2}R < 1, \tag{3.11}$$

then $\mathcal{T}$ is indeed a contraction on the closed ball B_{2R}. Consequently, the contraction mapping theorem implies there is a unique solution to equation (3.1) in $C([0,T];X)$.

We can also prove Lipschitz-continuous dependence on the initial data. Indeed, let $\|\mu_i^0\|_X \leq R$ for $i = 1, 2$. Then the preceding result tells us there exist unique solutions μ_i in $C([0,T];X)$ for some $T \sim \sigma R^{-2}$ and that $\|\mu_i\|_{C([0,T];X)} \lesssim R$. Using inequality (3.9), we find that

$$\|\mu_1 - \mu_2\|_{C([0,T];X)} \leq \|\mu_1^0 - \mu_2^0\|_X + C(T/\sigma)^{1/2}R\|\mu_1 - \mu_2\|_{C([0,T];X)}. \tag{3.12}$$

Provided that $C(T/\sigma)^{1/2}R < 1$, we have the bound

$$\|\mu_1 - \mu_2\|_{C([0,T];X)} \leq (1 - C(T/\sigma)^{1/2}R)^{-1}\|\mu_1^0 - \mu_2^0\|_X. \tag{3.13}$$

$\square$

REMARK 3.2.   By a Gronwall argument for the energy

$$\sum_{k=0}^{n}\int_{\mathbb{R}^d}(1 + |x|^2)^m|\nabla^{\otimes k}\mu(x)|^2 dx \tag{3.14}$$

for arbitrarily large integers $m, n \in \mathbb{N}$, it is easy to see that if the initial datum $\mu^0 \in \mathcal{S}(\mathbb{R}^d)$, then it remains spatially Schwartz on its lifespan. This property combined with the dependence bound (3.13) allows to approximate solutions in $C([0,T];X)$ by Schwartz-class solutions.

REMARK 3.3.   For a Schwartz-class solution μ, equation (1.5) and the divergence theorem yield

$$\frac{d}{dt}\int_{\mathbb{R}^d}\mu^t dx = \int_{\mathbb{R}^d}\big(-\operatorname{div}(\mu^t\mathbb{M}\nabla g * \mu^t) + \sigma\Delta\mu^t\big)dx = 0. \tag{3.15}$$

So by approximation, solutions $\mu \in C([0,T];X)$ obey conservation of mass.

REMARK 3.4.   If $\mu \in C([0,T];X)$, then for any $1 \leq p \leq \infty$, it holds that $\|\mu^t\|_{L^p} \leq \|\mu^{t'}\|_{L^p}$ for all $0 \leq t' \leq t \leq T$. Indeed, suppose that μ is Schwartz-class and $p \geq 1$ is finite. Then using equation (1.5), we see that

$$\frac{d}{dt}\|\mu^t\|_{L^p}^p = -p\int_{\mathbb{R}^d}|\mu^t|^{p-2}\mu^t\operatorname{div}\big(\mu^t\mathbb{M}\nabla g * \mu^t\big)dx + p\sigma\int_{\mathbb{R}^d}|\mu^t|^{p-2}\mu^t\Delta\mu^t dx. \tag{3.16}$$

14

It follows from integration by parts and the product rule that
(3.17)
$$\int_{\mathbb{R}^d} |\mu^t|^{p-2}\mu^t \Delta\mu^t dx = -\int_{\mathbb{R}^d} \left((p-2)(|\mu^t|^{p-4}\mu^t\nabla\mu^t)\mu^t + |\mu^t|^{p-2}\nabla\mu^t\right) \cdot \nabla\mu^t dx \leq 0.$$

Similarly,

$$(3.18) \quad -\int_{\mathbb{R}^d} |\mu^t|^{p-2}\mu^t \operatorname{div}\left(\mu^t \mathbb{M}\nabla\mathbf{g} * \mu^t\right) dx$$

$$= \int_{\mathbb{R}^d} \left((p-2)(|\mu^t|^{p-4}\mu^t\nabla\mu^t)\mu^t + |\mu^t|^{p-2}\nabla\mu^t\right) \cdot \mu^t \mathbb{M}\nabla\mathbf{g} * \mu^t dx.$$

Writing $(|\mu^t|^{p-4}\mu^t\nabla\mu^t)(\mu^t)^2 = p^{-1}\nabla(|\mu^t|^p)$ and $|\mu^t|^{p-2}\mu^t\nabla\mu^t = p^{-1}\nabla(|\mu^t|^p)$, it follows from integrating by parts that

$$(3.19) \quad -p\int_{\mathbb{R}^d} |\mu^t|^{p-2}\mu^t \operatorname{div}\left(\mu^t\mathbb{M}\nabla\mathbf{g} * \mu^t\right)dx = -(p-1)\int_{\mathbb{R}^d} |\mu^t|^p \operatorname{div}(\mathbb{M}\nabla\mathbf{g} * \mu^t)dx \leq 0$$

where the final inequality follows from assumption (x). This takes care of the case $p < \infty$. For $p = \infty$, we take the limit $p \to \infty^-$.

REMARK 3.5.   Since Remark 3.4 implies the L^1 and L^∞ norms of solutions are nonincreasing and the time of existence in Proposition 3.1 is proportional to $\|\mu^0\|_X^{-2}$, it follows from iterating Proposition 3.1 that solutions exist globally in $C([0,\infty); X)$.

REMARK 3.6.   Let μ be a nonnegative Schwartz-class solution to (1.5). Then using equation (1.5), integrating by parts, and using the chain rule,

$$\frac{d}{dt}\int_{\mathbb{R}^d} \log(1+|x|)\mu^t(x)dx = -\int_{\mathbb{R}^d} \log(1+|x|)\operatorname{div}\left(\mu^t(\mathbb{M}\nabla\mathbf{g} * \mu^t)\right)(x)dx$$

$$+ \sigma\int_{\mathbb{R}^d} \log(1+|x|)\Delta\mu^t(x)dx$$

$$= \int_{\mathbb{R}^d} \frac{x}{|x|(1+|x|)} \cdot \mathbb{M}\nabla\mathbf{g} * \mu^t(x)\mu^t(x)dx$$

$$(3.20) \qquad\qquad - \sigma\int_{\mathbb{R}^d} (1+|x|)^{-2}\mu^t(x)dx.$$

Hence, for any $T > 0$,

$$\int_{\mathbb{R}^d} \log(1+|x|)\mu^T(x)dx \leq \int_{\mathbb{R}^d} \log(1+|x|)\mu^0(x)dx + T\sup_{0\leq t\leq T} \|\nabla\mathbf{g} * \mu^t\|_{L^\infty}\|\mu^t\|_{L^1}$$

$$\lesssim \int_{\mathbb{R}^d} \log(1+|x|)\mu^0(x)dx + T\sup_{0\leq t\leq T} \|\mu^t\|_{L^1}^{2-\frac{s+1}{d}}\|\mu^t\|_{L^\infty}^{\frac{s+1}{d}}$$

$$(3.21) \qquad\qquad \leq \int_{\mathbb{R}^d} \log(1+|x|)\mu^0(x)dx + T\|\mu^0\|_{L^1}^{2-\frac{s+1}{d}}\|\mu^0\|_{L^\infty}^{\frac{s+1}{d}},$$

where the penultimate line follows from Remark 2.5 and the ultimate line from the nonincreasing property of L^p norms. By approximation and continuous dependence, it follows that if $\mu^0 \in X$ satisfies $\int_{\mathbb{R}^d} \log(1+|x|)\mu^0(x)dx < \infty$, then μ^t does as well for all $t > 0$.

REMARK 3.7.   Using assumption (x), it is not hard to also show that the minimum value of μ^t is nondecreasing in time. Consequently, if $\mu^0 \geq 0$, then $\mu \geq 0$ on its lifespan.

3.2. *Asymptotic decay.*   We now show the L^p norms of the solutions obtained in previous subsection satisfy the same temporal decay estimates as the linear heat equation. This follows the method of [9] and extends it to non divergence-free vector fields.

PROPOSITION 3.8.   *Suppose that $\mu \in C([0,\infty); X)$ is a solution to equation (1.5). If $\mathbb{M} : \nabla^{\otimes 2} \mathsf{g} \neq 0$, then assume that $\mu \geq 0$. Let $1 \leq p \leq q \leq \infty$. Then for all $t > 0$,*

$$(3.22) \qquad \|\mu^t\|_{L^q} \leq \left(\frac{K(q)}{K(p)} \right)^{\frac{d}{2}} \left(\frac{4\pi \sigma t}{1/p - 1/q} \right)^{-\frac{d}{2}\left(\frac{1}{p} - \frac{1}{q} \right)} \|\mu^0\|_{L^p},$$

where

$$(3.23) \qquad K(r) := \frac{r'^{1/r'}}{r^{1/r}}, \qquad 1 \leq r \leq \infty.$$

In the conservative case, the velocity field $\mathbb{M}\nabla\mathsf{g} * \mu$ is divergence-free and therefore Proposition 3.8 follows from the seminal work [9] of Carlen and Loss. In the general case where we only know that the matrix $\mathbb{M}$ is negative semidefinite (i.e. condition (1.2) holds), we make a small modification of their proof. As it is a crucial ingredient, we first recall the sharp form of Gross's log-Sobolev inequality [26, 8, 7]. Proposition 3.9 below is reproduced from [9] (see equation (1.17) in that work).

PROPOSITION 3.9.   *Let $a > 0$. Then for all $f \in H^1(\mathbb{R}^d)$,*

$$(3.24) \quad \int_{\mathbb{R}^d} |f(x)|^2 \log\left(\frac{|f(x)|^2}{\|f\|_{L^2}^2} \right) dx + \left(d + \frac{d\log a}{2} \right) \int_{\mathbb{R}^d} |f(x)|^2 dx \leq \frac{a}{\pi} \int_{\mathbb{R}^d} |\nabla f(x)|^2 dx.$$

Moreover, equality holds if and only if f is a scalar multiple and translate of $f_a(x) := a^{-d/4} e^{-\pi|x|^2/2a}$.

PROOF OF PROPOSITION 3.8.   Using Remark 3.2 and continuous dependence on the initial data, we may assume without loss of generality that μ is spatially Schwartz on its lifespan and μ is C^∞ in time. Therefore, there are no issues of regularity or decay in justifying the computations to follow. Additionally, let us rescale time by defining $\mu_\sigma(t, x) := \mu(t/\sigma, x)$, which now satisfies the equation

$$(3.25) \qquad \partial_t \mu_\sigma = -\sigma^{-1} \operatorname{div}(\mu_\sigma \mathbb{M}\nabla\mathsf{g} * \mu_\sigma) + \Delta\mu_\sigma.$$

It suffices to show

$$(3.26) \qquad \|\mu_\sigma^t\|_{L^q} \leq \left(\frac{K(q)}{K(p)} \right)^{\frac{d}{2}} \left(\frac{4\pi t}{1/p - 1/q} \right)^{-\frac{d}{2}\left(\frac{1}{p} - \frac{1}{q} \right)} \|\mu_\sigma^0\|_{L^p},$$

since replacing t with σt yields the desired result. To simplify the notation, we drop the σ subscript in what follows and assume that μ solves equation (3.25).

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

$$(3.27) \quad \begin{aligned} r(t)^2 \|\mu^t\|_{L^{r(t)}}^{r(t)-1} \frac{d}{dt} \|\mu^t\|_{L^{r(t)}} &= \dot{r}(t) \int_{\mathbb{R}^d} |\mu^t|^{r(t)} \log\left(\frac{|\mu^t|^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}} \right) dx \\ &\quad + r(t)^2 \int_{\mathbb{R}^d} |\mu^t|^{r(t)-1} \operatorname{sgn}(\mu^t) \partial_t \mu^t dx. \end{aligned}$$

16

Above, we have used the calculus identity

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

for C^1 functions $x(t) > 0$ and $y(t)$. Substituting in equation (1.5), the right-hand side of (3.27) equals

(3.29)
$$\dot{r}(t)\int_{\mathbb{R}^d}|\mu^t|^{r(t)}\log\left(\frac{|\mu^t|^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}}\right)dx + r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\Delta\mu^t dx$$
$$- \frac{r(t)^2}{\sigma}\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathbf{g}*\mu^t)dx.$$

By the product rule,

(3.30) $$- r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathbf{g}*\mu^t)dx$$

$$= -r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)}\operatorname{div}(\mathbb{M}\nabla\mathbf{g}*\mu^t)dx - r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\nabla\mu^t\cdot(\mathbb{M}\nabla\mathbf{g}*\mu^t)dx.$$

We recognize

(3.31) $$r(t)|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\nabla\mu^t = \nabla(|\mu^t|^{r(t)}).$$

Therefore integrating by parts, the second term in the right-hand side of (3.30) equals

(3.32) $$r(t)\int_{\mathbb{R}^d}|\mu^t|^{r(t)}\operatorname{div}(\mathbb{M}\nabla\mathbf{g}*\mu^t)dx.$$

Thus, equality (3.30) and assumption (x) (and that $\mu \geq 0$ by assumption if $\mathbb{M}:\nabla^{\otimes 2}\mathbf{g}$ does not vanish on $\mathbb{R}^d\setminus\{0\}$) now give

$$- \frac{r(t)^2}{\sigma}\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\operatorname{div}(\mu^t\mathbb{M}\nabla\mathbf{g}*\mu^t)dx$$
$$= -\frac{r(t)(r(t)-1)}{\sigma}\int_{\mathbb{R}^d}|\mu^t|^{r(t)}(\mathbb{M}:\nabla^{\otimes 2}\mathbf{g}*\mu^t)dx$$

(3.33) $$\leq 0.$$

Finally, write

(3.34) $$\operatorname{sgn}(\mu^t) = \lim_{\varepsilon\to 0^+}\frac{\mu^t}{\sqrt{\varepsilon^2 + |\mu^t|^2}}.$$

Integrating by parts,

$$r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\operatorname{sgn}(\mu^t)\Delta\mu^t dx$$

$$= \lim_{\varepsilon\to 0^+}\left(- r(t)^2(r(t)-1)\int_{\mathbb{R}^d}|\mu^t|^{r(t)-2}\frac{|\mu^t|}{\sqrt{\varepsilon^2+|\mu^t|^2}}|\nabla\mu^t|^2 dx\right.$$

$$\left. - r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\frac{|\nabla\mu^t|^2}{\sqrt{\varepsilon^2+|\mu^t|^2}}dx + r(t)^2\int_{\mathbb{R}^d}|\mu^t|^{r(t)-1}\frac{|\mu^t|^2|\nabla\mu^t|^2}{(\varepsilon^2+|\mu^t|^2)^{3/2}}dx\right)$$

$$= -r(t)^2(r(t)-1)\int_{\mathbb{R}^d}|\mu^t|^{r(t)-2}|\nabla\mu^t|^2 dx$$

$$(3.35) \quad = -4(r(t) - 1) \int_{\mathbb{R}^d} |\nabla|\mu^t|^{r(t)/2}|^2 dx.$$

After a little bookkeeping, we realize that we have shown

$$(3.36) \quad r(t)^2 \|\mu^t\|_{L^{r(t)}}^{r(t)-1} \frac{d}{dt} \|\mu^t\|_{L^{r(t)}} \leq \dot{r}(t) \int_{\mathbb{R}^d} |\mu^t|^{r(t)} \log\left(\frac{|\mu^t|^{r(t)}}{\|\mu^t\|_{L^{r(t)}}^{r(t)}}\right) dx$$
$$- 4(r(t) - 1) \int_{\mathbb{R}^d} |\nabla|\mu^t|^{r(t)/2}|^2 dx.$$

The remainder of the proof follows that of Carlen and Loss. We include the details for the sake of completeness.

We apply Proposition 3.9 pointwise in time with choice $a = \frac{4\pi(r(t)-1)}{\dot{r}(t)}$ and $f = |\mu^t|^{r(t)/2}$ to obtain that

$$(3.37) \quad r(t)^2 \|\mu^t\|_{L^{r(t)}}^{r(t)-1} \frac{d}{dt} \|\mu^t\|_{L^{r(t)}} \leq -\dot{r}(t)\left(d + \frac{d}{2}\log\left(\frac{4\pi(r(t)-1)}{\dot{r}(t)}\right)\right)\|\mu^t\|_{L^{r(t)}}^{r(t)}.$$

Implicit here is the requirement that $\dot{r}(t) > 0$. Define the function

$$(3.38) \quad G(t) := \log\left(\|\mu^t\|_{L^{r(t)}}\right).$$

Then

$$(3.39) \quad \frac{d}{dt} G(t) = \frac{1}{\|\mu^t\|_{L^{r(t)}}} \frac{d}{dt} \|\mu^t\|_{L^{r(t)}} \leq -\frac{\dot{r}(t)}{r(t)^2}\left(d + \frac{d}{2}\log\left(\frac{4\pi(r(t)-1)}{\dot{r}(t)}\right)\right).$$

Set $s(t) := 1/r(t)$, so that the preceding inequality becomes, after writing $\frac{r-1}{\dot{r}} = -\frac{s(1-s)}{\dot{s}}$,

$$(3.40) \quad \frac{d}{dt} G(t) \leq \dot{s}(t)\left(d + \frac{d}{2}\log(4\pi s(t)(1-s(t)))\right) + \frac{d}{2}(-\dot{s}(t))\log(-\dot{s}(t)).$$

So by the fundamental theorem of calculus,

$$(3.41) \quad G(t) - G(0) \leq \int_0^T \dot{s}(t)\left(d + \frac{d}{2}\log(4\pi s(t)(1-s(t)))\right) dt$$
$$- \frac{d}{2}\int_0^T \dot{s}(t)\log(-\dot{s}(t))dt.$$

We require that $s(0) = 1/p$ and $s(T) = 1/q$, so by the fundamental theorem of calculus,
$$(3.42)$$
$$\int_0^T \dot{s}(t)\left(d + \frac{d}{2}\log(4\pi s(t)(1-s(t)))\right) dt = \frac{d}{2}\left(\log(4\pi)s + \log\left(s^s(1-s)^{-(1-s)}\right)\right)\big|_{s=1/p}^{s=1/q}.$$

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

$$(3.43) \quad \dot{s}(t) = \frac{1}{T}\left(\frac{1}{q} - \frac{1}{p}\right), \qquad 0 \leq t \leq T.$$

Thus,

$$(3.44) \quad -\frac{d}{2}\int_0^T \dot{s}(t)\log(-\dot{s}(t))dt = -\frac{d}{2}\left(\frac{1}{p} - \frac{1}{q}\right)\log\left(\frac{T}{1/p - 1/q}\right).$$

The desired conclusion now follows from a little bookkeeping and exponentiating both sides of the inequality (3.41). $\qquad \square$

REMARK 3.10. Our extension of the Carlen-Loss [9] method to non-divergence-free vector fields is not limited to proving optimal decay estimates. In fact, it seems we have come across a more general property for which certain parabolic theory is valid. For example, suppose one considers linear equations of the form

(3.45)
$$\begin{cases} \partial_t \mu = \Delta \mu + \operatorname{div}(b\mu) + c\mu \\ \mu^t|_{t=s} = \mu^s \end{cases} \qquad (t,x) \in (s,\infty) \times \mathbb{R}^d,$$

where b is a vector field and c is a scalar, for simplicity both assumed to be C^∞. If $\operatorname{div} b \leq 0$, then using the same reasoning as in the proof of Proposition 3.8, one can show that the solution μ to (3.45) satisfies the decay estimates

(3.46)

$$\|\mu^t\|_{L^q} \leq \left(\frac{K(q)}{K(p)}\right)^{\frac{d}{2}} \left(\frac{4\pi(t-s)}{1/p-1/q}\right)^{-\frac{d}{2}\left(\frac{1}{p}-\frac{1}{q}\right)} e^{\int_s^t \|c^\kappa\|_{L^\infty} d\kappa} \|\mu^s\|_{L^p} \qquad \forall 0 < s \leq t < \infty.$$

Now one can easily show that equation (3.45) has a (smooth) fundamental solution, and a well-known problem in parabolic theory is to obtain Gaussian upper and lower bounds for such fundamental solutions, since such bounds imply Hölder continuity of the fundamental solution by an argument of Nash [49]. One can adapt the Carlen-Loss argument, which in turn is an adaptation of an earlier argument of Davies [15], to obtain a Gaussian upper bound from (3.46). In certain cases (e.g. [55, 41]), this Gaussian upper bound then implies a corresponding lower bound, and it would seem that these results also hold under the assumption that $\operatorname{div} b \leq 0$. We hope to investigate this line of inquiry more in future work.

4. N-particle dynamics. In this section, we discuss the well-posedness of the SDE system (1.1) for fixed $N \in \mathbb{N}$, in particular that with probability one, the particles never collide. We also discuss stability of the system under regularization of the potential. These regularizations will be important in the sequel when we attempt to apply Itô's lemma to functions which are singular at the origin.

To prove the well-posedness of the system (1.1), we first consider the well-posedness of the corresponding system where the potential g has been smoothly truncated at a short distance $\varepsilon > 0$ from the origin but otherwise is the same. If the particles remain more than distance ε from one another, then they do not see the truncation and therefore the truncation plays no role. This observation will guide us throughout this subsection.

Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be a nonnegative bump function satisfying

(4.1)
$$\chi(x) = \begin{cases} 1, & |x| \leq 1/2 \\ 0, & |x| \geq 1. \end{cases}$$

Given $\varepsilon > 0$, define

(4.2)
$$\mathrm{g}_{(\varepsilon)}(x) := \mathrm{g}(x)(1 - \chi(x/\varepsilon)).$$

The notation $\mathrm{g}_{(\varepsilon)}$ should not be confused with the notation g_ε in (5.2) used later in Section 5. By assumption (iv), $\mathrm{g}_{(\varepsilon)} \in C^\infty$ with

(4.3)
$$\|\mathrm{g}_{(\varepsilon)}\|_{L^\infty} \lesssim \begin{cases} -\log \varepsilon, & s = 0 \\ \varepsilon^{-s}, & 0 < s < d - 2 \end{cases}$$

and $\|\nabla^{\otimes k} \mathrm{g}_{(\varepsilon)}\|_{L^\infty} \lesssim \varepsilon^{-(s+k)}$ for $k \geq 1$,

(4.4)
$$\mathrm{g}_{(\varepsilon)}(x) = \begin{cases} \mathrm{g}(x), & |x| \geq \varepsilon \\ 0, & |x| \leq \varepsilon/2. \end{cases}$$

Define now the truncated version of the system (1.1) by

(4.5)
$$\begin{cases} dx_{i,\varepsilon} = \dfrac{1}{N} \sum_{1 \le j \le N : j \ne i} \mathbb{M}\nabla\mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})dt + \sqrt{2\sigma}dW_i \\ x_{i,\varepsilon}|_{t=0} = x_i^0. \end{cases}$$

Since the vector field $\mathbb{M}\nabla\mathbf{g}_{(\varepsilon)}$ is smooth with bounded derivatives of all order, global well-posedness of (4.5) is trivial. The equality (4.4) implies that if

(4.6)
$$\inf_{0 \le t \le T} \min_{1 \le i \ne j \le N} |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \ge \varepsilon,$$

then $x_{i,\varepsilon} = x_i$ on $[0,T]$ for every $1 \le i \le N$. In other words, the truncated dynamics coincide with the untruncated dynamics, just as remarked at the beginning of the subsection. Accordingly, we can define the stopping time

(4.7)
$$\tau_\varepsilon := \inf\{0 \le t \le T : \min_{1 \le i \ne j \le N} |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \ge 2\varepsilon\},$$

so that on the random time interval $[0, \tau_\varepsilon(\omega)]$, $\underline{x}_{N,\varepsilon}(\omega) \equiv \underline{x}_N(\omega)$, where $\omega \in \Omega$ is a sample in the underlying probability space.

REMARK 4.1.   For later use, we observe (for instance, see [36, Section 3.2.C]) that the quadratic variation of $x_{i,\varepsilon}$ is the $d \times d$ matrix with components

(4.8)
$$[x_{i,\varepsilon}]^{t,\alpha\beta} = 2\sigma t \delta^{\alpha\beta}, \qquad \alpha, \beta \in \{1, \dots, d\},$$

where $\delta^{\alpha\beta}$ is the Kronecker δ-function. Similarly, for $i \ne j$, the quadratic variation of $x_{i,\varepsilon} - x_{j,\varepsilon}$ is given by

(4.9)
$$[x_{i,\varepsilon} - x_{j,\varepsilon}]^{t,\alpha\beta} = [\sqrt{2\sigma}(W_i - W_j)^\alpha, \sqrt{2\sigma}(W_i - W_j)^\beta]^t = 4\sigma\delta^{\alpha\beta}t.$$

We first show that with probability one, the particles cannot escape to infinity by controlling the expectation of the *moment of inertia*

(4.10)
$$I_{N,\varepsilon} := \sum_{i=1}^N |x_{i,\varepsilon}|^2.$$

LEMMA 4.2.   *Suppose $s = 0$. There exists a constant $C > 0$ depending only on the dimension d, such that for all $T > 0$,*

(4.11)
$$\mathbb{E}\left(\sup_{0 \le t \le T} I_{N,\varepsilon}^t\right) \le C\left(I_{N,\varepsilon}^0 + N^2 T + \sigma(N+T)\right)e^{C\sigma T}.$$

PROOF.   We sketch the proof. Applying Itô's lemma with $f(x) = |x|^2$, we find that with probability one,

(4.12)
$$\forall t \ge 0, \qquad |x_{i,\varepsilon}^t|^2 - |x_i^0|^2 = 2\int_0^t x_{i,\varepsilon}^\kappa \cdot \sum_{\substack{1 \le j \le N \\ j \ne i}} \mathbb{M}\nabla\mathbf{g}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa)d\kappa$$
$$+ 2\sqrt{2\sigma}\int_0^t x_{i,\varepsilon}^\kappa \cdot dW_i^\kappa + 2d\sigma t.$$

20

Since $\nabla\mathsf{g}$ is odd by assumption (i), it follows from the triangle inequality and (iv) that

$$2\left|\sum_{i=1}^{N} x_{i,\varepsilon}\cdot\sum_{\substack{1\leq j\leq N\\ j\neq i}}\mathbb{M}\nabla\mathsf{g}(x_{i,\varepsilon}-x_{j,\varepsilon})\right|=\left|\sum_{1\leq i\neq j\leq N}(x_{i,\varepsilon}-x_{j,\varepsilon})\cdot\mathbb{M}\nabla\mathsf{g}(x_{i,\varepsilon}-x_{j,\varepsilon})\right|$$

(4.13)
$$\leq CN^2.$$

By the Burkholder-Davis-Gundy inequality, denoting again $[\cdot]$ for the quadratic variation, we have

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\left|\int_0^t x_{i,\varepsilon}^\kappa\cdot dW_i^\kappa\right|\right)\lesssim\mathbb{E}\left(\sqrt{\left[\int_0^{(\cdot)}x_{i,\varepsilon}^\kappa\cdot dW_i^\kappa\right]^T}\right)$$

(4.14)
$$\lesssim\mathbb{E}\left(\sqrt{\int_0^T|x_{i,\varepsilon}^\kappa|^2 d\kappa}\right).$$

We find after a little bookkeeping that

$$\mathbb{E}\left(\sup_{0\leq t\leq T}I_{N,\varepsilon}^t\right)\lesssim I_{N,\varepsilon}^0+\sigma\left(T+N^{1/2}\mathbb{E}\left(\left|\int_0^T I_{N,\varepsilon}^\kappa d\kappa\right|^{1/2}\right)\right)$$

$$\leq I_{N,\varepsilon}^0+CN^2T+\sigma\left(T+N^{1/2}\mathbb{E}\left(\int_0^T I_{N,\varepsilon}^\kappa d\kappa\right)^{1/2}\right)$$

(4.15)
$$\leq I_{N,\varepsilon}^0+CN^2T+\sigma\left(T+N+\int_0^T\mathbb{E}\left(I_{N,\varepsilon}^\kappa\right)d\kappa\right),$$

where the second line follows from Jensen's inequality and the third line from the elementary inequality $ab\leq\frac{a^2+b^2}{2}$ together with Fubini-Tonelli to interchange the expectation with the temporal integration. The desired conclusion now follows from the Gronwall-Bellman lemma. $\qquad\square$

REMARK 4.3. By Chebyshev's lemma, Lemma 4.2 implies that with probability one,

(4.16)
$$\lim_{R\to\infty}\inf\{t\geq 0:\max_{1\leq i\neq j\leq N}|x_{i,\varepsilon}^t-x_{j,\varepsilon}^t|\geq R\}=\infty.$$

Next, define the function

(4.17)
$$H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}):=\sum_{1\leq i\neq j\leq N}\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}-x_{j,\varepsilon}),$$

which has the interpretation of the energy of the system (4.5).

LEMMA 4.4. *There exists a constant $C>0$ depending only on s,d, and the potential g through assumptions (iv), (v), and (vii), such that for all $0<\varepsilon<\min\{\frac{1}{2},\frac{r_0}{2}\}$, where r_0 is as in (iii), and $T>0$, it holds that*

(4.18)
$$\left(\min_{|x|\leq 2\varepsilon}\mathsf{g}(x)\right)\mathbb{P}(\tau_\varepsilon<T)$$

$$\lesssim\begin{cases}\mathbb{E}\left(\frac{CN^2\left(N+e^{C\sigma T}(\sigma(N+T)+N^2T+I_{N,\varepsilon}^0)\right)}{2}+H_{N,\varepsilon}(\underline{x}_N^0)\right), & s=0\\ \mathbb{E}(H_{N,\varepsilon}(\underline{x}_N^0)), & 0<s<d-2.\end{cases}$$

In particular, by assumption (ii), $\mathbb{P}(\tau_\varepsilon < T) \to 0$ as $\varepsilon \to 0^+$.

REMARK 4.5.   Lemma 4.4 actually permits us to choose $\varepsilon = \varepsilon_N$ to depend on N in such a way that $\varepsilon_N \to 0$ as $N \to \infty$ and $\mathbb{P}(\tau_{\varepsilon_N} < T) \to 0$ with an explicit rate. For example, supposing that g is of the form (1.3), we have $\min_{|x| \le 2\varepsilon} \mathsf{g}(x) \ge C^{-1}(\varepsilon^{-s} + |\log \varepsilon| \mathbf{1}_{s=0})$. Assuming that $x_1^0, \ldots, x_N^0$ are IID with law μ^0, then $\mathbb{E}(H_{N,\varepsilon}^0) = O(N^2)$ uniformly in ε. If $s > 0$, then we can choose ε_N such that $\varepsilon_N N^{\frac{2}{s}} \to 0$ as $N \to \infty$. If $s = 0$, then we also have that $\mathbb{E}(I_{N,\varepsilon}^0) = O(N)$ uniformly in ε. Hence, we can choose ε_N such that $N^4/|\log \varepsilon_N| \to 0$ as $N \to \infty$.

PROOF OF LEMMA 4.4.   By Itô's lemma applied to $\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})$, it holds with probability one that

$$(4.19) \quad \forall t \ge 0, \qquad H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t) = H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^0)$$

$$+ 2 \sum_{1 \le i \ne j \le N} \sum_{\substack{1 \le k \le N \\ k \ne i}} \int_0^t \mathbb{M} \nabla \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{k,\varepsilon}^\kappa) \cdot \nabla \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) d\kappa$$

$$+ \sqrt{2\sigma} \sum_{1 \le i \ne j \le N} \int_0^t \nabla \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) \cdot d(W_i - W_j)^\kappa$$

$$+ \sigma \sum_{1 \le i \ne j \le N} \int_0^t \nabla^{\otimes 2} \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) : \mathbb{I} \, d\kappa.$$

The second line is nonnegative by condition (1.2) and therefore may be discarded. For the third line, we note that the Itô integral is a martingale with zero initial expectation. So by Doob's optional sampling theorem,

$$(4.20) \qquad \mathbb{E}\left(\int_0^{\tau_\varepsilon} \nabla \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) \cdot d(W_i - W_j)^\kappa \right) = 0.$$

For the fourth line, we note that

$$(4.21) \qquad \nabla^{\otimes 2} \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) : \mathbb{I} = \Delta \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa).$$

Since $\Delta \mathsf{g} \le 0$ by assumption (iii) and $\mathsf{g} = \mathsf{g}_{(\varepsilon)}$ outside the ball $B(x, \varepsilon)$, it follows that

$$(4.22) \qquad \Delta \mathsf{g}_{(\varepsilon)}(x) \le 0 \qquad \forall |x| \ge \varepsilon.$$

Consequently, it holds with probability one that

$$(4.23) \qquad \int_0^{\tau_\varepsilon} \nabla^{\otimes 2} \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) : \mathbb{I} \, d\kappa \le 0.$$

Taking expectations of both sides of equation (4.19), we find that

$$(4.24) \qquad \mathbb{E}\left(H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^{\tau_\varepsilon}) \right) \le \mathbb{E}\left(H_{N,\varepsilon}(\underline{x}_N^0) \right).$$

We now want to use this bound to control the minimal distance between particles. To this end, we separately consider the cases $s = 0$ and $0 < s < d - 2$. If $s = 0$, we use the moment of inertia to control the possible large negative values of g. Using assumptions (iv), (v), and

(vii) (provided that $2\varepsilon \leq r_0$), we see that for any $t \geq 0$,

$$H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t) \geq -C \sum_{\substack{1 \leq i \neq j \leq N \\ |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \geq 1}} \left| \log |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \right| + \sum_{\substack{1 \leq i \neq j \leq N \\ |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \leq 2\varepsilon}} \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^t - x_{j,\varepsilon}^t)$$

$$\geq C^{-1} \left(\min_{|x| \leq 2\varepsilon} \mathsf{g}(x) \right) |\{(i,j) \in \{1,\ldots,N\}^2 : i \neq j \text{ and } \varepsilon \leq |x_{i,\varepsilon}^t - x_{j,\varepsilon}^t| \leq 2\varepsilon\}|$$

$$(4.25) \qquad - CN^2 \max\left(\log 2, \log\left(\sum_{i=1}^N |x_{i,\varepsilon}^t| \right) \right).$$

Note that if $\sum_{i=1}^N |x_{i,\varepsilon}^t| \geq 1$, then by Cauchy-Schwarz and since $\log|x| \leq |x|$,

$$(4.26) \qquad \log\left(\sum_{i=1}^N |x_{i,\varepsilon}^t| \right) \leq N^{1/2} |I_{N,\varepsilon}^t|^{1/2} \leq \frac{N + I_{N,\varepsilon}^t}{2},$$

where we recall that $I_{N,\varepsilon}$ is the moment of inertia (4.10). Modulo a null set, this lower bound implies that

$$\{\tau_\varepsilon < T\} \subset \{\exists i \neq j \in \{1,\ldots,N\}^2 \text{ such that } \varepsilon \leq |x_{i,\varepsilon}^{\tau_\varepsilon} - x_{j,\varepsilon}^{\tau_\varepsilon}| \leq 2\varepsilon\}$$

$$(4.27) \qquad \subset \{H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^{\tau_\varepsilon}) \geq -\frac{CN^2(N + I_{N,\varepsilon}^{\tau_\varepsilon})}{2} + C^{-1} \min_{|x| \leq 2\varepsilon} \mathsf{g}(x)\}.$$

So by Chebyshev's inequality and inequality (4.24),

$$(4.28) \qquad \mathbb{E}\left(\frac{CN^2(N + I_{N,\varepsilon}^{\tau_\varepsilon})}{2} \right) + \mathbb{E}\left(H_{N,\varepsilon}(\underline{x}_N^0) \right) \geq C^{-1} \left(\min_{|x| \leq 2\varepsilon} \mathsf{g}(x) \right) \mathbb{P}(\tau_\varepsilon < T).$$

which in view of Lemma 4.2 concludes the proof if $s = 0$.

If $0 < s < d - 2$, then it follows from assumptions (iv), (v), and (vii) (provided that $2\varepsilon \leq r_0$)that

$$H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}) \geq \sum_{\substack{1 \leq i \neq j \leq N \\ |x_{i,\varepsilon} - x_{j,\varepsilon}| \leq 2\varepsilon}} \mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})$$

$$(4.29) \qquad \geq -CN^2 + \frac{\left(\min_{|x| \leq 2\varepsilon} \mathsf{g}(x) \right)}{C} |\{(i,j) : i \neq j \text{ and } \varepsilon \leq |x_{i,\varepsilon} - x_{j,\varepsilon}| \leq 2\varepsilon\}|,$$

which implies that

$$(4.30) \qquad C^{-1} \left(\min_{|x| \leq 2\varepsilon} \mathsf{g}(x) \right) \mathbb{P}(\tau_\varepsilon < T) \leq \mathbb{E}(H_{N,\varepsilon}(\underline{x}_N^0)) + CN^2,$$

completing the proof of the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The next proposition, which is the main result of this section, shows that there is a unique solution to the Cauchy problem for (1.1) in the strong sense.

PROPOSITION 4.6. *With probability one,*

$$(4.31) \qquad \underline{x}_N^t := \lim_{\varepsilon \to 0^+} \underline{x}_{N,\varepsilon}^t \text{ exists } \forall t \geq 0,$$

and we can unambiguously define $\underline{x}_N$ as the unique strong solution to (1.1). *Moreover,*

$$(4.32) \qquad \mathbb{P}\left(\forall t \geq 0, \quad \min_{1 \leq i \neq j \leq N} |x_i^t - x_j^t| > 0 \right) = 1.$$

PROOF. Note that if

$$(4.33) \qquad 2\varepsilon \leq \min_{1 \leq i \neq j \leq N} |x_i^0 - x_j^0|,$$

then $H_{N,\varepsilon}(\underline{x}_N^0) = H_N(\underline{x}_N^0)$. Choose a sequence $\varepsilon_k > 0$ such that

$$(4.34) \qquad \sum_{k=1}^\infty \left(\min_{|x| \leq 2\varepsilon_k} \mathsf{g}(x) \right)^{-1} < \infty.$$

Then by Lemma 4.4,

$$(4.35) \qquad \sum_{k=1}^\infty \mathbb{P}(\tau_{\varepsilon_k} < T) < \infty,$$

so by Borel-Cantelli,

$$(4.36) \qquad \mathbb{P}\left(\limsup_{k \to \infty} \{\tau_{\varepsilon_k} < T\} \right) = 0.$$

Consequently, for almost every sample $\omega \in \Omega$, there exists $\varepsilon(\omega) > 0$ such that for all $0 < \varepsilon \leq \varepsilon(\omega)$,

$$(4.37)$$
$$\underline{x}_{N,\varepsilon}^t(\omega) = \underline{x}_{N,\varepsilon(\omega)}^t(\omega) \text{ on } [0,T] \quad \text{and} \quad \inf_{0 \leq t \leq T} \min_{1 \leq i \neq j \leq N} |x_{i,\varepsilon}^t(\omega) - x_{j,\varepsilon}^t(\omega)| \geq 2\varepsilon(\omega).$$

Since $T > 0$ was arbitrary, we note that the preceding a.s. statement in fact holds globally in time. □

5. Modulated energy and renormalized commutator estimates. In this section, we review the properties of the modulated energy $F_N(\underline{x}_N, \mu)$ established in the authors' joint work with Nguyen [50] along with the associated renormalized commutator estimate proven in that work. These previous results will suffice to prove Theorem 1.1. In the case of potentials which are globally superharmonic (i.e. $r_0 = \infty$ in assumption (iii)), we also prove sharper versions (in terms of their $\|\mu\|_{L^\infty}$ dependence) of the results of [50] that are crucial to obtain global bounds of Theorem 1.2.

Throughout this section, we always assume that μ is a probability measure with density in $L^\infty(\mathbb{R}^d)$. If $0 < s < d$, then it is immediate from Lemma 2.3 that $\mathsf{g} * \mu$ is a bounded, continuous function (it is actually $C^{k,\alpha}$ for some $k \in \mathbb{N}_0$ and $\alpha > 0$ depending on the value of s) and therefore the modulated energy is well-defined. If $s = 0$, then we need to impose a suitable decay assumption on μ to compensate for the logarithmic growth of g at infinity. For example,

$$(5.1) \qquad \int_{\mathbb{R}^d} \log(1 + |x|) d\mu(x) < \infty.$$

5.1. *Review of results from [50].* We start by reviewing the results of [50, Section 2] on the properties of the modulated energy under the general assumptions on the potential contained in (i) – (ix). The statements presented below are specialized to the sub-Coulombic setting (i.e. $0 \leq s \leq d - 2$), and the relevant proofs, as well as further comments, may be found in [50, Sections 2, 4].

With r_0 as in (iii) and $0 < \eta < \min\{\frac{1}{2}, \frac{r_0}{2}\}$, we let $\delta_x^{(\eta)}$ denote the uniform probability measure on the sphere $\partial B(x, \eta)$ and set

$$(5.2) \qquad \mathsf{g}_\eta := \mathsf{g} * \delta_0^{(\eta)}.$$

Since g is superharmonic in $B(0, r_0)$ by assumption (iii), it follows from the formula ($\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure in $\mathbb{R}^d$)

$$(5.3) \qquad \frac{d}{dr}\fint_{\partial B(x,r)} f d\mathcal{H}^{d-1} = \frac{1}{d|B(0,1)|r^{d-1}} \int_{B(x,r)} \Delta f \, dy,$$

which holds for any sufficiently integrable f, and an approximation argument that

$$(5.4) \qquad \mathsf{g}_\eta(x) \le \mathsf{g}(x) \quad \forall x \in B(0, r_0 - \eta) \setminus \{0\}$$

and (using assumption (iv))

$$(5.5) \qquad |\mathsf{g}(x) - \mathsf{g}_\eta(x)| \le \frac{C\eta^2}{|x|^{s+2}} \qquad \forall |x| \ge 2\eta,$$

where the constant C depends on r_0. By virtue of the mean value inequality (5.4) and assumption (iv), the self-interaction of the smeared point mass $\delta_{x_0}^{(\eta)}$ satisfies the relation

$$\int_{(\mathbb{R}^d)^2} \mathsf{g}(x-y) d\delta_{x_0}^{(\eta)}(x) d\delta_{x_0}^{(\eta)}(y) = \int_{\mathbb{R}^d} \mathsf{g}_\eta d\delta_0^{(\eta)} \le \int_{\mathbb{R}^d} \mathsf{g} d\delta_0^{(\eta)} = \mathsf{g}_\eta(0)$$

$$(5.6) \qquad\qquad\qquad\qquad\qquad \le C\big(\eta^{-s} + |\log\eta| \mathbf{1}_{s=0}\big).$$

The next result we recall [50, Proposition 2.1] expresses the crucial monotonicity property of the modulated energy with respect to the smearing radii when expressed as the limit

$$(5.7) \quad F_N(\underline{x}_N, \mu)$$

$$= \lim_{\alpha_i \to 0} \left(\int_{(\mathbb{R}^d)^2} \mathsf{g}(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}^{(\alpha_i)} - \mu\right)^{\otimes 2}(x,y) - \frac{1}{N^2}\sum_{i=1}^N \int_{\mathbb{R}^d} \mathsf{g}_{\alpha_i} d\delta_0^{(\alpha_i)} \right).$$

It also shows that the modulated energy is bounded from below, coercive, and controls the microscale interactions [50, Corollary 2.3].

PROPOSITION 5.1. *Let $d \ge 3$ and $0 \le s \le d-2$. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is pairwise distinct and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. If $s = 0$, also suppose that $\int_{\mathbb{R}^d} \log(1+|x|) d\mu(x) < \infty$. There exists a constant C depending only on s, d and the potential g through assumptions* (iii), (iv), (vi), *such that for every choice of $0 < \eta_1, \ldots, \eta_N < \min\{\frac{1}{2}, \frac{r_0}{2}\}$,*

$$(5.8) \quad \frac{1}{N^2} \sum_{\substack{1 \le i \ne j \le N \\ |x_i - x_j| \le \frac{r_0}{2}}} (\mathsf{g}(x_j - x_i) - \mathsf{g}_{\eta_i}(x_j - x_i))_+ + C^{-1} \left\| \frac{1}{N}\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - \mu \right\|_{\dot{H}^{\frac{s-d}{2}}}^2$$

$$\le F_N(\underline{x}_N, \mu) + \frac{C}{N} \sum_{i=1}^N \left(\left(\eta_i^2 + \frac{\eta_i^{-s}(1 + |\log\eta_i| \mathbf{1}_{s=0})}{N} \right) \right.$$

$$\left. + C\|\mu\|_{L^\infty} \eta_i^{d-s}(1 + |\log\eta_i|(\mathbf{1}_{s=0} + \mathbf{1}_{s=d-2})) \right).$$

REMARK 5.2. Since $0 \le s \le d-2$ by assumption, we can balance the error terms in (5.8) by setting

$$(5.9) \qquad \eta_i^2 = \frac{\eta_i^{-s}}{N} \iff \eta_i = N^{-\frac{1}{s+2}},$$

which, in particular, yields the lower bound

$$(5.10) \qquad F_N(\underline{x}_N, \mu) \geq -C(1 + \|\mu\|_{L^\infty})N^{-\frac{2}{s+2}}(1 + (\log N)(\mathbf{1}_{s=0} + \mathbf{1}_{s=d-2})).$$

REMARK 5.3.   If instead of (vi), we only assume that $\hat{\mathsf{g}} \geq 0$ on $\mathbb{R}^d \setminus \{0\}$, then the proof of [50, Proposition 2.1] yields the bound

$$(5.11) \quad \frac{1}{N^2} \sum_{\substack{1 \leq i \neq j \leq N \\ |x_i - x_j| \leq \frac{r_0}{2}}} (\mathsf{g}(x_j - x_i) - \mathsf{g}_{\eta_i}(x_j - x_i))_+ \leq F_N(\underline{x}_N, \mu)$$

$$+ \frac{C}{N} \sum_{i=1}^N \left(\left(\eta_i^2 + \frac{\eta_i^{-s}(1 + |\log \eta_i| \mathbf{1}_{s=0})}{N} \right) \right.$$

$$\left. + C\|\mu\|_{L^\infty} \eta_i^{d-s}(1 + |\log \eta_i|(\mathbf{1}_{s=0} + \mathbf{1}_{s=d-2})) \right).$$

The next result we recall [50, Proposition 2.2] concerns the analogue of Proposition 5.1 in the case $d - 2 < s < d$. One of the key new insights from [50] is that although superharmonicity may fail in the space $\mathbb{R}^d$, as it does for the Riesz potential $|x|^{-s}$, superharmonicity may be restored by considering the potential as the restriction of a potential G (i.e. $\mathsf{g}(x) = \mathsf{G}(x, 0)$) in an extended space $\mathbb{R}^{d+m}$, where the size of m depends on the value of s so as to make G superharmonic in a neighborhood of the origin. Namely, suppose that $\mathsf{g} : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ is such that there exists $\mathsf{G} : \mathbb{R}^{d+m} \setminus \{0\} \to \mathbb{R}$ with $\mathsf{g}(x) = \mathsf{G}(x, 0)$ and satisfying conditions (1.14) – (1.17). With the notation $X = (x, z) \in \mathbb{R}^{d+m}$ and $X_i = (x_i, 0)$, we let $\delta_X^{(\eta)}$ denote the uniform probability measure on the sphere $\partial B(X, r) \subset \mathbb{R}^{d+m}$ and set

$$(5.12) \qquad \mathsf{G}_\eta := \mathsf{G} * \delta_0^{(\eta)}$$

for $0 < \eta < \min\{\frac{1}{2}, \frac{r_0}{2}\}$. Analogously to (5.4), (5.5), (5.6), we have

$$(5.13) \qquad \mathsf{G}_\eta(X) \leq \mathsf{G}(X) \qquad \forall X \in B(0, r_0 - \eta) \setminus \{0\},$$

$$(5.14) \qquad |\mathsf{G}(X) - \mathsf{G}_\eta(X)| \leq \frac{C\eta^2}{|X|^{s+2}} \qquad \forall |X| \geq 2\eta,$$

and

$$(5.15) \qquad \int_{(\mathbb{R}^{d+m})^2} \mathsf{G}(x - y) d\delta_0^{(\eta)}(x) d\delta_0^{(\eta)}(y) \leq \mathsf{G}_\eta(0) \leq C\big(\eta^{-s} + |\log \eta| \mathbf{1}_{d=1 \wedge s=0}\big).$$

Again, the constant C in (5.14) depends on r_0.

PROPOSITION 5.4.   Let $d \geq 3$ and $d - 2 < s < d$. Let g, G be as above. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is a pairwise distinct configuration and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. There exists a constant $C > 0$ only depending on s, d and on g, G, such that for every $0 < \eta_1, \ldots, \eta_N < \min\{\frac{1}{2}, \frac{r_0}{2}\}$, we have

$$(5.16)$$

$$\frac{1}{N^2} \sum_{\substack{1 \leq i \neq j \leq N \\ |x_i - x_j| \leq \frac{r_0}{2}}} (\mathsf{g}(x_j - x_i) - \mathsf{G}_{\eta_i}(x_j - x_i, 0))_+ \leq F_N(\underline{x}_N, \mu) + \sum_{i=1}^N \frac{\eta_i^{-s}(1 + |\log \eta_i| \mathbf{1}_{s=0})}{N^2}$$

$$+ \frac{C}{N} \sum_{i=1}^N \left(\|\mu\|_{L^\infty} \eta_i^{d-s} + \eta_i^2 \right).$$

REMARK 5.5.   Strictly speaking, the inequality (5.16) differs from [50, (2.23), Proposition 2.2] by the omission of a positive term (a suitable squared norm of $\mu_N^t - \mu^t$). The reason we have omitted this term is because we no longer assume that $\hat{\mathsf{G}}(\Xi) \sim |\Xi|^{s-d-m}$. Instead, condition (1.17) only tells us that

$$(5.17) \qquad \int_{(\mathbb{R}^{d+m})^2} \mathsf{G}(X - Y) d\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}^{(\eta_i)} - \tilde{\mu} \right)(X, Y) \geq 0,$$

which is good enough for the purposes of this article.

As an application of Proposition 5.1 if $0 \leq s \leq d - 4$ and Proposition 5.4 if $d - 4 < s < d - 2$ using assumption (viii), expressions like the second term appearing in the right-hand side of (1.7), which is due to the nonzero quadratic variation of the Brownian motion when we calculate the Itô equation for the modulated energy $F_N(\underline{x}_N^t, \mu^t)$, are nonpositive up to a controllable error.

COROLLARY 5.6.   *Let $d \geq 3$ and $0 \leq s < d - 2$. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is pairwise distinct and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Let g be a potential satisfying assumptions (i), (iii), (iv), (vi), (viii). There exists a constant $C > 0$ depending only on s, d and g, such that*

$$(5.18)$$
$$\int_{(\mathbb{R}^d)^2 \setminus \triangle} (-\Delta \mathsf{g})(x - y) d\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} - \mu \right)^{\otimes 2}(x, y) \geq -C(1 + \|\mu\|_{L^\infty}) N^{-\frac{\min\{2, d-s-2\}}{\min\{s+4, d\}}}.$$

PROOF.   If $0 \leq s \leq d - 4$, we use (5.11) applied with potential $-\Delta \mathsf{g}$ and with each $\eta_i = N^{-\frac{1}{s+4}}$. If $d - 4 < s < d - 2$, we use (5.16) applied with extended potential G for $-\Delta \mathsf{g}$ given by assumption (viii) and with each $\eta_i = N^{-\frac{1}{d}}$.   □

Finally, we close this section by recalling [50, Proposition 4.1] the renormalized commutator estimate from that work. As commented in the introduction, such estimates are the main workhorse to close Gronwall arguments based on the modulated energy.

PROPOSITION 5.7.   *Let $d \geq 3$ and $0 \leq s \leq d - 2$. Let $\underline{x}_N \in (\mathbb{R}^d)^N$ be a pairwise distinct configuration, and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. If $s = 0$, assume that $\int_{\mathbb{R}^d} \log(1 + |x|) d\mu(x) < \infty$. Let v be a continuous vector field on $\mathbb{R}^d$. There exists a constant $\hat{C}$ depending only d, s and on the potential g through assumptions (i) – (vii), (ix), such that*

$$(5.19) \qquad \left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} (v(x) - v(y)) \cdot \nabla \mathsf{g}(x - y) d\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} - \mu \right)^{\otimes 2}(x, y) \right|$$

$$\leq C \Bigg(\|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s < d-2} \Bigg) \Bigg(F_N(\underline{x}_N, \mu) + C N^{-\frac{s+3}{(s+2)(s+1)}} \||\nabla|^{s+1-d} \mu\|_{L^\infty}$$

$$+ C(1 + \|\mu\|_{L^\infty}) N^{-\frac{2}{(s+2)(s+1)}} \Big(1 + (\log N)(\mathbf{1}_{s=0} + \mathbf{1}_{s=d-2}) \Big) \Bigg).$$

5.2. *New estimates for globally superharmonic potentials.*   We now assume that $r_0 = \infty$ in the condition (iii), i.e. g is globally superharmonic. Under this more restrictive assumption, which holds in the model potential case (1.3), we can obtain versions of Proposition 5.1 and Proposition 5.4 that have a better balance of factors of $\|\mu\|_{L^\infty}$ between terms. In particular, there is no η^2 error term like there is in the right-hand side of inequality (5.8). This is important because if μ^t is time-dependent and satisfies the decay bound (3.22), this term will contribute linear growth in time when integrated. As we shall see in Section 7.2, this better balance will be crucial to show that the error terms (i.e. those which are not $F_N(\underline{x}_N^t, \mu^t)$) that result when estimating the right-hand side of (1.7) are integrable in time over the interval $[0, \infty)$.

PROPOSITION 5.8.   *Let $d \geq 3$ and $0 \leq s \leq d - 2$. Assume that g is a potential satisfying conditions (i), (iii), (iv), (vi) with $r_0 = \infty$. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is pairwise distinct and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. In the case $s = 0$, also suppose that $\int_{\mathbb{R}^d} \log(1 + |x|) d\mu(x) < \infty$. For any $\infty \geq p > \frac{d}{s+2}$, there exist constants $C, C_p > 0$ depending only on s, d and the potential g through the assumed conditions, such that for every choice of $0 < \eta_1, \ldots, \eta_N < 2^{-\frac{dp-d+2p}{d(p-1)}} \|\mu\|_{L^\infty}^{-\frac{1}{d}}$,*

(5.20)

$$\frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} (\mathrm{g}(x_j - x_i) - \mathrm{g}_{\eta_i}(x_j - x_i))_+ + C^{-1} \left\| \frac{1}{N} \sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - \mu \right\|_{\dot{H}^{\frac{s-d}{2}}}^2 \leq F_N(\underline{x}_N, \mu)$$

$$+ \frac{C_p \|\mu\|_{L^\infty}^{\gamma_{s,p}}}{N} \sum_{i=1}^N \eta_i^{\lambda_{s,p}} (1 + (|\log \eta_i| + |\log \|\mu\|_{L^\infty}|) \mathbf{1}_{s=0}) + \sum_{i=1}^N \frac{C \eta_i^{-s}(1 + |\log \eta_i| \mathbf{1}_{s=0})}{N^2},$$

where the exponents $\gamma_{s,p}, \lambda_{s,p}$ are defined by

(5.21) $$\gamma_{s,p} := \frac{2p + sp - s}{dp + 2p - d}, \qquad \lambda_{s,p} := \frac{2p(d - s)}{dp + 2p - d}.$$

PROOF.   We modify the proof of [50, Proposition 2.1]. Adding and subtracting $\delta_{x_i}^{(\eta_i)}$ and regrouping terms yields the decomposition

$$F_N(\underline{x}_N, \mu) = \int_{(\mathbb{R}^d)^2} \mathrm{g}(x - y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - \mu \right)^{\otimes 2} (x, y) - \frac{1}{N^2} \sum_{i=1}^N \int_{\mathbb{R}^d} \mathrm{g}_{\eta_i} d\delta_0^{(\eta_i)}$$

(5.22)

$$- \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} (\mathrm{g}(y - x_i) - \mathrm{g}_{\eta_i}(y - x_i)) d\mu(y)$$

$$+ \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} \int_{\mathbb{R}^d} (\mathrm{g}(y - x_i) - \mathrm{g}_{\eta_i}(y - x_i)) d(\delta_{x_j} + \delta_{x_j}^{(\eta)})(y).$$

Since $\Delta \mathrm{g} \leq 0$ on $\mathbb{R}^d$ by assumption (iii) and $\delta_{x_j}^{(\eta_j)}$ is a positive measure, we have from (5.4) that

(5.23)

$$\frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} \int_{\mathbb{R}^d} (\mathrm{g}(y - x_i) - \mathrm{g}_{\eta_i}(y - x_i)) d(\delta_{x_j} + \delta_{x_j}^{(\eta_j)})(y)$$

$$\geq \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} (\mathrm{g}(x_j - x_i) - \mathrm{g}_{\eta_i}(x_j - x_i))_+.$$

Let $R \geq 2\eta$ be a parameter to be specified shortly. Using assumption (iv) and the estimate (5.5), we find that

$$\left| \int_{\mathbb{R}^d} (\mathbf{g}(y - x_i) - \mathbf{g}_\eta(y - x_i)) d\mu(y) \right|$$

$$\leq \|\mu\|_{L^\infty} \int_{|y - x_i| \leq R} (|\mathbf{g}(y - x_i)| + |\mathbf{g}_\eta(y - x_i)|) dy$$

$$+ \int_{|y - x_i| > R} |\mathbf{g}(y - x_i) - \mathbf{g}_\eta(y - x_i)| \, d\mu(y)$$

$$\tag{5.24} \leq C\|\mu\|_{L^\infty} \left(R^{d-s}(1 + |\log R|\mathbf{1}_{s=0}) \right) + C\eta^2 \int_{|y - x_i| > R} |y - x_i|^{-s-2} d\mu(y)$$

By Hölder's inequality,

$$\int_{|y - x_i| > R} |y - x_i|^{-s-2} d\mu(y) \leq C(p(s+2) - d)^{-\frac{1}{p}} R^{\frac{d - p(s+2)}{p}} \|\mu\|_{L^{p'}}$$

$$\tag{5.25} \leq C(p(s+2) - d)^{-\frac{1}{p}} R^{\frac{d - p(s+2)}{p}} \|\mu\|_{L^\infty}^{\frac{1}{p}}$$

for any Hölder conjugate p, p' with $\frac{d}{s+2} < p \leq \infty$. Implicitly, we have used that μ is a probability density, and the constant C is independent of p. Setting

$$\tag{5.26} \|\mu\|_{L^\infty} R^{d-s} = \eta^2 R^{\frac{d - p(s+2)}{p}} \|\mu\|_{L^\infty}^{\frac{1}{p}},$$

we get

$$\tag{5.27} R = \left(\eta^2 \|\mu\|_{L^\infty}^{\frac{1-p}{p}} \right)^{\frac{p}{dp - d + 2p}}.$$

In order for $R \geq 2\eta$, we need

$$\tag{5.28} \left(\eta^2 \|\mu\|_{L^\infty}^{\frac{1-p}{p}} \right)^{\frac{p}{dp - d + 2p}} \geq 2\eta \iff \eta \leq 2^{-\frac{dp - d + 2p}{d(p-1)}} \|\mu\|_{L^\infty}^{-\frac{1}{d}}.$$

Substituting in the above choice of R, we obtain that

$$\tag{5.29}$$

$$\frac{1}{N} \sum_{i=1}^{N} \left| \int_{\mathbb{R}^d} (\mathbf{g}(y - x_i) - \mathbf{g}_{\eta_i}(y - x_i)) d\mu(y) \right| \leq \frac{C}{N} \|\mu\|_{L^\infty}^{\frac{2p - s + sp}{dp - d + 2p}} (p(s+2) - d)^{-\frac{1}{dp + 2p - d}}$$

$$\times \sum_{i=1}^{N} \eta_i^{\frac{2p(d-s)}{dp + 2p - d}} \left(1 + \left(\frac{p}{dp + 2p - d} \left| \log \eta_i^2 \|\mu\|_{L^\infty}^{\frac{1-p}{p}} \right| \right) \mathbf{1}_{s=0} \right).$$

Next, we use the relation (5.6) to bound

$$\tag{5.30} \frac{1}{N^2} \sum_{i=1}^{N} \left| \int_{\mathbb{R}^d} \mathbf{g}_{\eta_i} d\delta_0^{(\eta_i)} \right| \leq \sum_{i=1}^{N} \frac{C\eta_i^{-s}(1 + |\log \eta_i|\mathbf{1}_{s=0})}{N^2}.$$

Collecting (5.23), (5.29), (5.30) and using the assumption (vi) with Plancherel's theorem for the remaining term in (5.22), we arrive at the inequality in the statement of the proposition. $\qquad \square$

REMARK 5.9. Evidently, the constant C_p in Proposition 5.8 blows up as $p \to \frac{d}{s+2}^+$.

REMARK 5.10.   Dropping the s, p subscripts in $\gamma_{s,p}, \lambda_{s,p}$, we balance the error terms by setting

$$(5.31) \qquad C_p \|\mu\|_{L^\infty}^\gamma \eta_i^\lambda = \frac{\eta_i^{-s}}{N} \iff \eta_i = C_p^{-\frac{1}{\lambda+s}} \|\mu\|_{L^\infty}^{-\frac{\gamma}{\lambda+s}} N^{-\frac{1}{\lambda+s}},$$

which, for possibly larger constant $C_p > 0$, implies the lower bound

$$(5.32) \qquad F_N(\underline{x}_N, \mu) \geq -C_p \|\mu\|_{L^\infty}^{\frac{s}{d}} N^{-\frac{\lambda}{\lambda+s}} (1 + (|\log \|\mu\|_{L^\infty}| + \log N) \mathbf{1}_{s=0}).$$

REMARK 5.11.   Just as in Remark 5.3, if instead of (vi), we only assume that $\hat{g} \geq 0$ on $\mathbb{R}^d \setminus \{0\}$, then we have the bound

$$(5.33)$$

$$\frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} (\mathsf{g}(x_j - x_i) - \mathsf{g}_{\eta_i}(x_j - x_i))_+ \leq F_N(\underline{x}_N, \mu) + \sum_{i=1}^N \frac{C \eta_i^{-s} (1 + |\log \eta_i| \mathbf{1}_{s=0})}{N^2}$$

$$+ \frac{C_p \|\mu\|_{L^\infty}^\gamma}{N} \sum_{i=1}^N \eta_i^\lambda (1 + (|\log \eta_i| + |\log \|\mu\|_{L^\infty}|) \mathbf{1}_{s=0}).$$

Under the global superharmonicity assumption, we can also obtain a version of Proposition 5.4 without an η^2 term and where every error term that is increasing in η has a factor of $\|\mu\|_{L^\infty}$.

PROPOSITION 5.12.   *Let $d \geq 3$ and $d - 2 < s < d$. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is a pairwise distinct configuration and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Let g be a potential satisfying (i), (iii), (iv) with $r_0 = \infty$. Let $\mathsf{G} : \mathbb{R}^{d+m} \setminus \{0\} \to \mathbb{R}$ be an extension $\mathsf{G}(x, 0) = \mathsf{g}(x)$ such that G satisfies conditions (1.14), (1.15), (1.16), (1.17) with $r_0 = \infty$. There exists a constant $C > 0$ only depending on s, d and on g, G, such that for every $\eta_1, \ldots, \eta_N > 0$, we have*
$$(5.34)$$
$$\frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} (\mathsf{g}(x_j - x_i) - \mathsf{G}_{\eta_i}(x_j - x_i, 0))_+ \leq F_N(\underline{x}_N, \mu) + \frac{C}{N} \sum_{i=1}^N \left(\frac{\eta_i^{-s}}{N} + \|\mu\|_{L^\infty} \eta_i^{d-s} \right).$$

PROOF.   We modify the proof of [50, Proposition 2.2] in the same spirit as we did for Proposition 5.8. Adding and subtracting $\delta_{X_i}^{(\eta_i)}$ and regrouping terms, we find that

$$
\begin{aligned}
(5.35) \qquad F_N(\underline{x}_N, \mu) &= \int_{(\mathbb{R}^{d+m})^2} \mathsf{G}(X - Y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{X_i}^{(\eta_i)} - \tilde{\mu} \right)^{\otimes 2} (x, y) \\
&\quad - \frac{1}{N^2} \sum_{i=1}^N \int_{(\mathbb{R}^{d+m})^2} \mathsf{G}(X - Y) d(\delta_0^{(\eta_i)})^{\otimes 2}(X, Y) \\
&\quad - \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^{d+m}} (\mathsf{G}(Y - X_i) - \mathsf{G}_{\eta_i}(Y - X_i)) d\tilde{\mu}(Y) \\
&\quad + \frac{1}{N^2} \sum_{1 \leq i \neq j \leq N} \int_{\mathbb{R}^d} (\mathsf{G}(Y - X_i) - \mathsf{G}_{\eta_i}(Y - X_i)) d(\delta_{X_j} + \delta_{X_j}^{(\eta_j)})(Y).
\end{aligned}
$$

By the inequality (5.13) and since $\delta_{X_i}^{(\eta_i)}$ is a positive measure in $\mathbb{R}^{d+m}$, we have the lower bound

(5.36)
$$\frac{1}{N^2}\sum_{1\leq i\neq j\leq N}\int_{\mathbb{R}^d}(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i))d(\delta_{X_j}+\delta_{X_j}^{(\eta_j)})(Y)$$
$$\geq\frac{1}{N^2}\sum_{1\leq i\neq j\leq N}(\mathsf{g}(x_j-x_i)-\mathsf{G}_\eta(x_j-x_i,0))_+.$$

Letting $R\geq 2\eta$ be a parameter to be determined, we have that

(5.37) $\displaystyle\int_{\mathbb{R}^{d+m}}(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i))d\tilde{\mu}(Y)=\int_{\mathbb{R}^d}(\mathsf{g}(y-x_i)-\mathsf{G}_{\eta_i}(y-x_i,0))d\mu(y)$

by definition of $\tilde{\mu}$. Since $s>d-2$ by assumption, we can use (5.14) and (1.16) to obtain the bound

(5.38) $\displaystyle\int_{|y-x_i|\geq R}|\mathsf{g}(y-x_i)-\mathsf{G}_{\eta_i}((y-x_i,0))|d\mu(y)\leq C\eta_i^2 R^{d-s-2}\|\mu\|_{L^\infty}.$

Using 1.16 to estimate directly the integral over $|y-x_i|<R$ as in (5.24), it follows that

(5.39) $\displaystyle\left|\int_{\mathbb{R}^{d+m}}(\mathsf{g}(y-x_i)-\mathsf{G}_{\eta_i}(y-x_i,0))d\mu(y)\right|\leq C\|\mu\|_{L^\infty}\left(R^{d-s}+\eta_i^2 R^{d-s-2}\right).$

We can them optimize the choice of R by setting $R=2\eta_i$. After a little bookkeeping, we have shown that

(5.40) $\displaystyle\frac{1}{N}\sum_{i=1}^N\left|\int_{\mathbb{R}^{d+m}}(\mathsf{g}(y-x_i)-\mathsf{G}_{\eta_i}(y-x_i,0))d\mu(y)\right|\leq C\eta_i^{d-s}\|\mu\|_{L^\infty}.$

Finally, using the relation (5.15), we have the self-interaction bound

(5.41) $\displaystyle\frac{1}{N^2}\sum_{i=1}^N\left|\int_{(\mathbb{R}^{d+m})^2}\mathsf{G}(X-Y)d(\delta_0^{(\eta_i)})^{\otimes 2}(X,Y)\right|\leq\frac{C\eta^{-s}}{N}$

and using assumption 1.17 with Plancherel's theorem, we have the lower bound

(5.42) $\displaystyle\int_{(\mathbb{R}^{d+m})^2}\mathsf{G}(X-Y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\eta_i)}-\tilde{\mu}\right)(X,Y)\geq 0.$

Combining these observations with (5.36), (5.40), we arrive at the inequality in the statement of the proposition. $\qquad\square$

REMARK 5.13. Similar to Remark 5.10, we can balance the error terms in (5.34) by choosing $\eta_i=(\|\mu\|_{L^\infty}N)^{-\frac{1}{d}}$, which implies the lower bound

(5.43) $$F_N(\underline{x}_N,\mu)\geq -C\|\mu\|_{L^\infty}^{\frac{s}{d}}N^{-\frac{d-s}{d}}.$$

Analogous to Corollary 5.6, we can use assumption (viii) for admissible potentials g together with Proposition 5.8 (if $0\leq s\leq d-4$) and Proposition 5.12 (if $d-4<s<d-2$) to obtain the following result.

COROLLARY 5.14.   *Let $d \geq 3$ and $0 \leq s < d-2$. Suppose that $\underline{x}_N \in (\mathbb{R}^d)^N$ is pairwise distinct and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Let g be a potential satisfying assumptions (i), (iii), (iv), (vi), (viii) with $r_0 = \infty$. There exists a constant $C > 0$ depending only on s, d and g, G, such that the following holds. If $0 \leq s \leq d-4$, then for any $\infty \geq p > \frac{d}{s+4}$, C also depends on p and*

$$(5.44) \quad \int_{(\mathbb{R}^d)^2 \setminus \triangle} (-\Delta \mathsf{g})(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y) \geq -C_p \|\mu\|_{L^\infty}^{\frac{s+2}{d}} N^{-\frac{\lambda_{s+2,p}}{\lambda_{s+2,p}+s+2}}.$$

where $\lambda_{s+2,p}$ is as defined in (5.21). If $d-4 < s < d-2$, then

$$(5.45) \quad \int_{(\mathbb{R}^d)^2 \setminus \triangle} (-\Delta \mathsf{g})(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y) \geq -C\|\mu\|_{L^\infty}^{\frac{s+2}{d}} N^{-\frac{d-s-2}{d}}.$$

PROOF.   If $0 \leq s \leq d-4$, then we use (5.33) with s replaced by $s+2$ and choosing $\eta_i = \|\mu\|_{L^\infty}^{-\frac{\gamma_{s+2,p}}{\lambda_{s+2,p}+s+2}} N^{-\frac{1}{\lambda_{s+2,p}+s+2}}$. If $d-4 < s < d-2$, then we use (5.34) with s replaced by $s+2$ and choosing $\eta_i = (\|\mu\|_{L^\infty} N)^{-\frac{1}{d}}$. $\qquad\square$

Repeating the proof of [50, Proposition 4.1], except now using Proposition 5.8 instead of Proposition 5.1, we can obtain a renormalized commutator estimate (cf. Proposition 5.7) with better distribution of norms of μ.

PROPOSITION 5.15.   *Let $d \geq 3$ and $0 \leq s \leq d-2$. Let $\underline{x}_N \in (\mathbb{R}^d)^N$ be pairwise distinct, and $\mu \in \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. If $s = 0$, assume that $\int_{\mathbb{R}^d} \log(1+|x|) d\mu(x) < \infty$. Let v be a continuous vector field on $\mathbb{R}^d$. For every $\infty \geq p > \frac{d}{s+2}$, there exists a constant C_p depending only d, s and on the potential g through assumptions (i) – (vii), (ix) such that for all $N > \left(2^{\frac{dp-d+p}{(p-1)}}\|\mu\|_{L^\infty}\right)^{\frac{(s+1)(s+\lambda_{s,p})}{d}}$,*

$$(5.46) \quad \left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} (v(x)-v(y)) \cdot \nabla \mathsf{g}(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y) \right|$$

$$\leq C\|\nabla v\|_{L^\infty} \||\nabla|^{s+1-d}\mu\|_{L^\infty} N^{-\frac{s+1+\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}}$$

$$+ C\left(\|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2}\right)\left(F_N(\underline{x}_N, \mu)\right.$$

$$\left. + C_p\left(1+\|\mu\|_{L^\infty}^{\gamma_{s,p}}\right) N^{-\frac{\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}}\left(1 + (\log N + |\log \|\mu\|_{L^\infty}|)\mathbf{1}_{s=0}\right)\right),$$

where $\gamma_{s,p}, \lambda_{s,p}$ are as in (5.21).

PROOF.   Since s, p are fixed, we drop the subscripts in $\gamma_{s,p}, \lambda_{s,p}$. Repeating the steps in the proof of [50, Proposition 4.1], we find that the left-hand side of (5.46) is controlled by

$$(5.47) \quad C\left(\|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2}\right)\left(F_N(\underline{x}_N, \mu)\right.$$

$$\left. + C_p\|\mu\|_{L^\infty}^{\gamma} \eta^\lambda (1 + (|\log \eta| + |\log\|\mu\|_{L^\infty}|)\mathbf{1}_{s=0}) + \frac{C\eta^{-s}(1+|\log\eta|\mathbf{1}_{s=0})}{N}\right)$$

$$+ C\|\nabla v\|_{L^\infty} \left(\frac{\varepsilon^{-s}(1 + |\log\varepsilon|\mathbf{1}_{s=0})}{N} + \eta\||\nabla|^{s+1-d}\mu\|_{L^\infty} + \frac{\eta}{\varepsilon^{s+1}} \right.$$

$$\left. + C_p\|\mu\|_{L^\infty}^\gamma \varepsilon^\lambda(1 + (|\log\varepsilon| + |\log\|\mu\|_{L^\infty}|)\mathbf{1}_{s=0}) \right),$$

where $0 < 2\eta \le \varepsilon < 2^{-\frac{dp-d+2p}{d(p-1)}}\|\mu\|_{L^\infty}^{-\frac{1}{d}}$ and $\infty \ge p > \frac{d}{s+2}$. Since $0 \le s \le d-2$ by assumption, we balance error terms (i.e. those terms which are not $F_N(\underline{x}_N, \mu)$) by setting

$$(5.48) \qquad \frac{\eta}{\varepsilon^{s+1}} = \frac{\eta^{-s}}{N} = \varepsilon^\lambda,$$

which yields $\eta = \varepsilon^{\lambda+s+1}$ and $\varepsilon = N^{-\frac{1}{(1+s)(s+\lambda)}}$. To ensure that $\varepsilon < 2^{-\frac{dp-d+2p}{d(p-1)}}\|\mu\|_{L^\infty}^{-\frac{1}{d}}$, we require that

$$(5.49) \qquad N^{-\frac{1}{(1+s)(s+\lambda)}} < 2^{-\frac{dp-d+2p}{d(p-1)}}\|\mu\|_{L^\infty}^{-\frac{1}{d}} \iff 2^{\frac{(dp-d+2p)(1+s)(s+\lambda)}{d(p-1)}}\|\mu\|_{L^\infty}^{\frac{(1+s)(s+\lambda)}{d}} < N.$$

Substituting these choices back into (5.47), we arrive at the inequality in the statement of the proposition. □

6. Evolution of the modulated energy.

Our next task is to rigorously compute the time-derivative of the modulated energy, which we recall is a real-valued stochastic process. Since the potential g is not C^2 due to its singularity at the origin, we cannot directly apply Itô's lemma to $F_N(\underline{x}_N^t, \mu^t)$, as we formally did in the introduction to obtain (1.7). Instead, we proceed by a truncation and stopping time argument, similar to that used Section 4 to prove the well-posedness of the N-body dynamics.

We define the *truncated modulated energy*

$$(6.1) \qquad F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t, \mu^t) := \int_{(\mathbb{R}^d)^2\setminus\triangle} \mathsf{g}_{(\varepsilon)}(x-y)d(\mu_{N,\varepsilon}^t - \mu^t)^{\otimes 2}(x,y),$$

where $\mathsf{g}_{(\varepsilon)}$ is as defined in (4.2), $\underline{x}_{N,\varepsilon}$ is the solution to the truncated system (4.5), and $\mu_{N,\varepsilon}$ denotes the empirical measure induced by $\underline{x}_{N,\varepsilon}$. Comparing this expression to the definition of $F_N(\underline{x}_N, \mu)$ above, we have just replaced the potential g with the potential $\mathsf{g}_{(\varepsilon)}$ and replaced $\underline{x}_N$ with $\underline{x}_{N,\varepsilon}$. Thanks to the regularity of the truncated potential $\mathsf{g}_{(\varepsilon)}$, we can *rigorously* apply Itô's lemma to $F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t, \mu^t)$.

LEMMA 6.1. *Let $\underline{x}_{N,\varepsilon}$ be a solution to the system* (4.5), *and let $\mu \in C([0,\infty); L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$ be a solution to equation* (1.5). *Then for every choice of ε satisfying*

$$(6.2) \qquad 0 < \varepsilon \le \frac{1}{2}\min_{1\le i\ne j\le N}|x_i^0 - x_j^0|,$$

it holds with probability one that for all $t \ge 0$,

$$(6.3) \quad F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t, \mu^t) = F_{N,\varepsilon}(\underline{x}_N^0, \mu^0)$$

$$+ \frac{2}{N^3}\sum_{\substack{1\le i,j,k\le N \\ k,j\ne i}}\int_0^t \nabla\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa)\cdot\mathbb{M}\nabla\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{k,\varepsilon}^\kappa)d\kappa$$

$$+ \frac{2}{N}\sum_{i=1}^N \int_0^t (\mathsf{g}_{(\varepsilon)} * \mathrm{div}(u^\kappa\mu^\kappa))(x_{i,\varepsilon}^\kappa)d\kappa$$

$$+ \frac{2}{N^2} \sum_{1 \le i \ne j \le N} \int_0^t (\nabla \mathbf{g}_{(\varepsilon)} * \mu^\kappa)(x_{i,\varepsilon}^\kappa) \cdot \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) d\kappa$$

$$- 2 \int_0^t \langle \mathbf{g}_{(\varepsilon)} * \mathrm{div}(u^\kappa \mu^\kappa), \mu^\kappa \rangle_{L^2} d\kappa + 2\sigma \int_0^t \int_{(\mathbb{R}^2)^2 \setminus \triangle} \Delta \mathbf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa$$

$$+ \frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d \setminus \{x_{i,\varepsilon}^\kappa\}} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)(y) \cdot dW_i^\kappa,$$

where $\mu_{N,\varepsilon} := \frac{1}{N} \sum_{i=1}^N \delta_{x_{i,\epsilon}}$ *and* $u := \mathbb{M} \nabla \mathbf{g} * \mu$.

PROOF. By approximation, we may assume without loss of generality that μ is smooth and rapidly decaying at infinity. We split the modulated energy into a sum of three terms, defined by

$$(6.4) \qquad \mathrm{Term}_1 := \frac{1}{N^2} \sum_{1 \le i \ne j \le N} \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}),$$

$$(6.5) \qquad \mathrm{Term}_2 := -\frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^2} \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - y) d\mu(y) = -\frac{2}{N} \sum_{i=1}^N \mathbf{g}_{(\varepsilon)} * \mu(x_{i,\varepsilon}),$$

$$(6.6) \qquad \mathrm{Term}_3 := \int_{(\mathbb{R}^2)^2} \mathbf{g}_{(\varepsilon)}(x-y) d\mu^{\otimes 2}(x,y) = \langle \mathbf{g}_{(\varepsilon)} * \mu, \mu \rangle_{L^2},$$

and compute the stochastic/deterministic differential equation satisfied by each of these terms. Of course, $\mathrm{Term}_1, \ldots, \mathrm{Term}_3$ depend on ε, but since ε is fixed, we omit this dependence.

Term_1 By Itô's lemma and Remark 4.1, we have, for $i \ne j$, that $\mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})$ satisfies the SDE

$$dg_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})$$

$$= \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) \cdot d(x_{i,\varepsilon} - x_{j,\varepsilon}) + \frac{1}{2} \nabla^{\otimes 2} \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) : d[x_{i,\varepsilon} - x_{j,\varepsilon}]$$

$$= \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) \cdot \left(\frac{1}{N} \sum_{\substack{1 \le k \le N \\ k \ne i}} \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{k,\varepsilon}) \right.$$

$$\left. - \frac{1}{N} \sum_{\substack{1 \le k \le N \\ k \ne j}} \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{j,\varepsilon} - x_{k,\varepsilon}) \right) dt + \sqrt{2\sigma} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) \cdot d(W_i - W_j)$$

$$(6.7) \qquad + 2\sigma (\nabla^{\otimes 2} \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) : \mathbb{I}) dt.$$

Evidently,

$$(6.8) \qquad 2\sigma \nabla^{\otimes 2} \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon}) : \mathbb{I} = 2\sigma \Delta \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{j,\varepsilon})$$

Thus by symmetry under swapping $i \leftrightarrow j$, and after integrating in time, we obtain

(6.9)

$$\mathrm{Term}_1(t) = \frac{1}{N^2} \sum_{1 \le i \ne j \le N} \mathbf{g}(x_i^0 - x_j^0) + \frac{2\sqrt{2\sigma}}{N^2} \sum_{1 \le i \ne j \le N} \int_0^t \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) \cdot dW_i^\kappa$$

$$+ \frac{2}{N^3} \sum_{1 \leq i \neq j \leq N} \sum_{\substack{1 \leq k \leq N \\ k \neq i}} \int_0^t \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) \cdot \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{k,\varepsilon}^\kappa) d\kappa$$

$$+ 2\sigma \sum_{1 \leq i \neq j \leq N} \int_0^t \Delta \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) d\kappa,$$

provided that $\varepsilon \leq \frac{1}{2} \min_{i \neq j} |x_{i,0} - x_{j,0}|$.

Term$_2$ Defining $f(t,x) := (\mathbf{g}_{(\varepsilon)} * \mu^t)(x)$, we first observe from equation (1.5) that

$$(6.10) \qquad \partial_t f = \mathbf{g}_{(\varepsilon)} * (-\operatorname{div}(\mu u) + \sigma \Delta \mu)),$$

Applying Itô's lemma with the time-dependent function f, we find that

$$df(t, x_{i,\varepsilon}) = \partial_t f(t, x_{i,\varepsilon})dt + \nabla f(t, x_{i,\varepsilon}) \cdot dx_{i,\varepsilon} + \frac{1}{2} \nabla^{\otimes 2} f(t, x_{i,\varepsilon}) : d[x_{i,\varepsilon}]$$

$$= -\mathbf{g}_{(\varepsilon)} * \operatorname{div}(u^t \mu^t)(x_{i,\varepsilon})dt + \sigma(\mathbf{g}_{(\varepsilon)} * \Delta \mu^t)(x_{i,\varepsilon})dt$$

$$+ \nabla(\mathbf{g}_{(\varepsilon)} * \mu^t)(x_{i,\varepsilon}) \cdot \frac{1}{N} \sum_{\substack{1 \leq k \leq N \\ k \neq i}} \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon} - x_{k,\varepsilon})dt$$

$$(6.11) \qquad + \sqrt{2\sigma} \nabla(\mathbf{g}_{(\varepsilon)} * \mu^t)(x_{i,\varepsilon}) \cdot dW_i + \sigma(\nabla^{\otimes 2}(\mathbf{g}_{(\varepsilon)} * \mu^t)(x_{i,\varepsilon}) : \mathbb{I})dt,$$

where we also use Remark 4.1 to obtain the ultimate line. Noting that

$$(6.12) \qquad \sigma \nabla^{\otimes 2}(\mathbf{g}_{(\varepsilon)} * \mu^t)(x_{i,\varepsilon}) : \mathbb{I} = \sigma \Delta(\mathbf{g}_{(\varepsilon)} * \mu^t)(x_{i,\varepsilon}),$$

we conclude that

$$(6.13) \quad \text{Term}_2(t) = -\frac{2}{N} \mathbf{g}_{(\varepsilon)} * \mu^0(x_i^0) + \frac{2}{N} \sum_{i=1}^N \int_0^t \mathbf{g}_{(\varepsilon)} * \operatorname{div}(u^\kappa \mu^\kappa)(x_{i,\varepsilon}^\kappa) d\kappa$$

$$- \frac{2}{N^2} \sum_{1 \leq i \neq k \leq N} \int_0^t \nabla(\mathbf{g}_{(\varepsilon)} * \mu^\kappa)(x_{i,\varepsilon}^\kappa) \cdot \mathbb{M} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{k,\varepsilon}^\kappa) d\kappa$$

$$- \frac{2\sigma}{N} \sum_{i=1}^N \int_0^t (\mathbf{g}_{(\varepsilon)} * \Delta \mu^\kappa)(x_{i,\varepsilon}^\kappa) d\kappa - \frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^N \int_0^t \nabla(\mathbf{g}_{(\varepsilon)} * \mu^\kappa)(x_{i,\varepsilon}^\kappa) \cdot dW_i^\kappa.$$

Term$_3$ Using equation (6.10) and symmetry under swapping $x \leftrightarrow y$, we find that

$$(6.14) \qquad \text{Term}_3(t) = \langle \mathbf{g}_\varepsilon * \mu^0, \mu^0 \rangle + 2 \int_0^t \langle \mathbf{g}_{(\varepsilon)} * (-\operatorname{div}(u^\kappa \mu^\kappa) + \sigma \Delta \mu^\kappa), \mu^\kappa \rangle_{L^2} d\kappa.$$

Combining the identities (6.9), (6.13) and (6.14) completes the proof of the lemma. $\qquad \square$

We now proceed to group terms following the proof of [63, Lemma 2.1]. We leave filling in the details to the reader, as it requires nothing new from the aforementioned work.

LEMMA 6.2. *Let $\underline{x}_{N,\varepsilon}$ and μ be as in Lemma 6.1. Then for every ε satisfying (6.2), it holds with probability 1 that for all $t \geq 0$,*

$$(6.15) \quad F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^t, \mu^t) - F_{N,\varepsilon}(\underline{x}_N^0, \mu^0)$$

$$\leq \int_0^t \int_{(\mathbb{R}^d)^2 \setminus \triangle} (u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y)) \cdot \nabla \mathbf{g}_{(\varepsilon)}(x - y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x, y)$$

$$+ 2\sigma \int_0^t \int_{(\mathbb{R}^d)^2 \backslash \triangle} \Delta \mathbf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa$$

$$+ \frac{2}{N} \sum_{i=1}^N \int_0^t (\mathbf{g}_{(\varepsilon)} * \mathrm{div}((u^\kappa - u_\varepsilon^\kappa)\mu^\kappa))(x_{i,\varepsilon}^\kappa) d\kappa$$

$$- 2 \int_0^t \langle \mathbf{g}_{(\varepsilon)} * \mathrm{div}((u^\kappa - u_\varepsilon^\kappa)\mu^\kappa), \mu^\kappa \rangle_{L^2} d\kappa$$

$$+ \frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^N \int_0^t \mathrm{P.V.} \int_{\mathbb{R}^d \backslash \{x_{i,\varepsilon}^\kappa\}} \nabla \mathbf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)(y) \cdot dW_i^\kappa,$$

where $u_\varepsilon := \mathbb{M}\nabla \mathbf{g}_{(\varepsilon)} * \mu$.

We are now prepared to remove the truncation by passing to the limit $\varepsilon \to 0^+$. To this end, we recall from Section 4 the stopping time τ_ε defined in (4.7). From Proposition 4.6, we know that $\lim_{\varepsilon \to 0} \tau_\varepsilon = \infty$ a.s. The next proposition, the culmination of our work so far, is the main result of this subsection. It gives a functional inequality for the expected magnitude of the modulated energy, which serves as the first step in our Gronwall argument, and should be interpreted as the "rigorous version" of the inequality (1.7) from the introduction.

PROPOSITION 6.3.   *For all $t \geq 0$, we have the inequality*

(6.16)

$$\mathbb{E}\big(F_N(\underline{x}_N^t, \mu^t) - F_N(\underline{x}_N^0, \mu^0)\big) \leq 2\sigma \mathbb{E}\left(\int_0^t \int_{(\mathbb{R}^d)^2 \backslash \triangle} \Delta \mathbf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa\right)$$

$$+ \mathbb{E}\left(\int_0^t \left| \int_{(\mathbb{R}^d)^2 \backslash \triangle} (u^\kappa(x) - u^\kappa(y)) \cdot \nabla \mathbf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) \right| d\kappa\right).$$

REMARK 6.4.   By Proposition 5.1, Proposition 6.3 also implies that there is a constant $C > 0$ such that

(6.17) $\mathbb{E}\big(\big|F_N(\underline{x}_N^t, \mu^t)\big| - \big|F_N(\underline{x}_N^0, \mu^0)\big|\big)$

$$\leq C\left(\frac{\eta^{-s}(1 + |\log \eta|\mathbf{1}_{s=0})}{N} + \eta^2 + \|\mu\|_{L^\infty}\eta^{d-s}(1 + |\log \eta|\mathbf{1}_{s=0})\right)$$

$$+ 2\sigma \mathbb{E}\left(\int_0^t \int_{(\mathbb{R}^d)^2 \backslash \triangle} \Delta \mathbf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa\right)$$

$$+ \mathbb{E}\left(\int_0^t \left| \int_{(\mathbb{R}^d)^2 \backslash \triangle} (u^\kappa(x) - u^\kappa(y)) \cdot \nabla \mathbf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) \right| d\kappa\right)$$

for any choice of $0 < \eta < \min\{\frac{1}{2}, \frac{r_0}{2}\}$. Similarly, using Proposition 5.8 if g is globally superharmonic, Proposition 6.3 also implies that for any $\infty \geq p > \frac{d}{s+2}$, there is a $C_p > 0$ such that

(6.18)

$$\mathbb{E}\big(\big|F_N(\underline{x}_N^t, \mu^t)\big| - \big|F_N(\underline{x}_N^0, \mu^0)\big|\big) \leq C_p \|\mu\|_{L^\infty}^{\gamma_{s,p}} \eta^{\lambda_{s,p}}(1 + (|\log \eta| + |\log \|\mu\|_{L^\infty}|)\mathbf{1}_{s=0})$$

$$+ \frac{C\eta^{-s}(1+|\log\eta|\mathbf{1}_{s=0})}{N} + 2\sigma\mathbb{E}\left(\int_0^t \int_{(\mathbb{R}^d)^2\backslash\triangle} \Delta\mathsf{g}(x-y)d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y)d\kappa\right)$$

$$+ \mathbb{E}\left(\int_0^t \left|\int_{(\mathbb{R}^d)^2\backslash\triangle} (u^\kappa(x) - u^\kappa(y)) \cdot \nabla\mathsf{g}(x-y)d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y)\right| d\kappa\right)$$

for any $0 < \eta < 2^{-\frac{dp-d+2p}{d(p-1)}}\|\mu\|_{L^\infty}^{-\frac{1}{d}}$.

PROOF OF PROPOSITION 6.3. Fix $t > 0$. By mollifying the initial datum and using the continuous dependence in Proposition 3.1, we assume without loss of generality that μ is C^∞. Since

(6.19) $$\frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^N \int_0^{(\cdot)} \text{P.V.} \int_{\mathbb{R}^d\backslash\{x_{i,\varepsilon}^\kappa\}} \nabla\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - y)d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)(y) \cdot dW_i^\kappa$$

is a sum of square-integrable martingales with zero initial expectation, Doob's optional sampling theorem implies that for every $\varepsilon > 0$,

(6.20) $$\mathbb{E}\left(\frac{2\sqrt{2\sigma}}{N} \sum_{i=1}^N \int_0^{\tau_\varepsilon\wedge t} \text{P.V.} \int_{\mathbb{R}^d\backslash\{x_{i,\varepsilon}^\kappa\}} \nabla\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - y)d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)(y) \cdot dW_i^\kappa\right) = 0.$$

Next, consider the expression

(6.21) $$\int_{(\mathbb{R}^d)^2\backslash\triangle} \Delta(\mathsf{g}_{(\varepsilon)} - \mathsf{g})(x-y)d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y).$$

Observe that by definition (4.2) of $\mathsf{g}_{(\varepsilon)}$,

(6.22) $$(\mathsf{g}_{(\varepsilon)} - \mathsf{g}) * \Delta\mu^\kappa(x) = -\int_{\mathbb{R}^d} \mathsf{g}(x-y)\chi_\varepsilon(x-y)\Delta\mu^\kappa(y)dy.$$

Integrating by parts twice to move the derivatives off μ^κ and then applying Cauchy-Schwarz and using (iv) for g, we find

(6.23) $$\left|\int_{\mathbb{R}^d} \mathsf{g}(x-y)\chi_\varepsilon(x-y)\Delta\mu^\kappa(y)dy\right| \lesssim \frac{\|\mu^\kappa\|_{L^\infty}}{\varepsilon^{2+s}}\left(\int_{|x-y|\le 2\varepsilon} dy\right) \lesssim \|\mu^0\|_{L^\infty}\varepsilon^{d-2-s},$$

where in the ultimate inequality we use the nonincreasing property of L^p norms. Thus,

(6.24) $$\left|\int_{\mathbb{R}^d} (\mathsf{g}_{(\varepsilon)} - \mathsf{g}) * \Delta\mu^\kappa(x)d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)(x)\right| \lesssim \|\mu^0\|_{L^\infty}\varepsilon^{d-2-s}.$$

Since for every $0 \le \kappa \le \tau_\varepsilon$,

(6.25) $$\sum_{1\le i\neq j\le N} \Delta\mathsf{g}_{(\varepsilon)}(x_{i,\varepsilon}^\kappa - x_{j,\varepsilon}^\kappa) = \sum_{1\le i\neq j\le N} \Delta\mathsf{g}(x_i^\kappa - x_j^\kappa),$$

using $\lim_{\varepsilon\to 0}\tau_\varepsilon = \infty$ a.s., it follows that with probability one, for all $0 \le \kappa \le t$,

(6.26) $$\lim_{\varepsilon\to 0^+} \left|\int_{(\mathbb{R}^d)^2\backslash\triangle} \left(\Delta\mathsf{g}_{(\varepsilon)}(x-y)d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y)\right.\right.$$

$$\left.\left. - \Delta\mathsf{g}(x-y)d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y)\right)\right| = 0.$$

So by dominated convergence,

$$
\begin{aligned}
(6.27) \quad \lim_{\varepsilon \to 0^+} \mathbb{E}&\left(\int_0^{t \wedge \tau_\varepsilon} \int_{(\mathbb{R}^d)^2 \setminus \triangle} \Delta \mathsf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa \right) \\
&= \mathbb{E}\left(\int_0^t \int_{(\mathbb{R}^d)^2 \setminus \triangle} \Delta \mathsf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa \right).
\end{aligned}
$$

Next, consider the expression

$$
(6.28) \qquad \int_0^t \int_{(\mathbb{R}^d)^2 \setminus \triangle} (u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y)) \cdot \nabla \mathsf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y) d\kappa.
$$

We want to show that

$$
\begin{aligned}
(6.29) \quad \lim_{\varepsilon \to 0} \mathbb{E}&\left(\int_0^{t \wedge \tau_\varepsilon} \left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} (u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y)) \cdot \nabla \mathsf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa - \mu^\kappa)^{\otimes 2}(x,y) \right. \right. \\
&\left. \left. - \int_{(\mathbb{R}^d)^2 \setminus \triangle} (u^\kappa(x) - u^\kappa(y)) \cdot \nabla \mathsf{g}(x-y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x,y) \right| d\kappa \right) = 0.
\end{aligned}
$$

We break up the demonstration of (6.29) into three parts.

• Almost surely, we have that for all $0 \le \kappa \le \tau_\varepsilon$

$$
\begin{aligned}
(6.30) \quad \int_{(\mathbb{R}^d)^2 \setminus \triangle}& (u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y)) \cdot \nabla \mathsf{g}_{(\varepsilon)}(x-y) d(\mu_{N,\varepsilon}^\kappa)^{\otimes 2}(x,y) \\
&= \frac{1}{N^2} \sum_{1 \le i \ne j \le N} \left(u_\varepsilon^\kappa(x_i^\kappa) - u_\varepsilon^\kappa(x_j^\kappa) \right) \cdot \nabla \mathsf{g}(x_i^\kappa - x_j^\kappa).
\end{aligned}
$$

Write

$$
(6.31) \qquad u - u_\varepsilon = \mathbb{M} \nabla (\mathsf{g} - \mathsf{g}_{(\varepsilon)}) * \mu.
$$

We see from the same reasoning as in the estimate (6.23) that

$$
(6.32) \qquad \|\nabla^{\otimes 2}(\mathsf{g} - \mathsf{g}_{(\varepsilon)}) * \mu^\kappa\|_{L^\infty} \lesssim \|\mu^\kappa\|_{L^\infty} \varepsilon^{d-s-2} \le \|\mu^0\|_{L^\infty} \varepsilon^{d-s-2}.
$$

Hence by the mean-value theorem and using assumption (v) for g,

$$
\begin{aligned}
(6.33) \quad &\left| \left((\nabla(\mathsf{g}_{(\varepsilon)} - \mathsf{g}) * \mu^\kappa)(x_i^\kappa) - (\nabla(\mathsf{g}_{(\varepsilon)} - \mathsf{g}) * \mu^\kappa)(x_j^\kappa) \right) \cdot \nabla \mathsf{g}(x_i^\kappa - x_j^\kappa) \right| \\
&\quad \lesssim \|\mu^0\|_{L^\infty} \varepsilon^{d-s-2} |x_i^\kappa - x_j^\kappa| |\nabla \mathsf{g}(x_i^\kappa - x_j^\kappa)| \\
&\quad \lesssim \begin{cases} \|\mu^0\|_{L^\infty} \varepsilon^{d-2}, & s = 0 \\ \|\mu^0\|_{L^\infty} \varepsilon^{d-2-s} \mathsf{g}(x_i^\kappa - x_j^\kappa), & 0 < s < d-2, \end{cases}
\end{aligned}
$$

where to obtain the ultimate line in the case $0 < s < d-2$ we assume $|x_i^\kappa - x_j^\kappa| < r_0$. For the case $s = 0$, the preceding estimate suffices. For the case $0 < s < d-2$, we need to deal with the factor $\mathsf{g}(x_i - x_j)$. To this end, we set $H_N(\underline{x}_N) \coloneqq \sum_{1 \le i \ne j \le N} \mathsf{g}(x_i - x_j)$. Since g is positive inside the ball $B(0, r_0)$ and $|\mathsf{g}| \le C$ outside $B(0, r_0)$ by assumptions (iv) and (v), we find that

$$
\mathbb{E}\left(\int_0^{t \wedge \tau_\varepsilon} H_N(\underline{x}_N^\kappa) d\kappa \right) = \mathbb{E}\left(\int_0^{t \wedge \tau_\varepsilon} H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^\kappa) d\kappa \right)
$$

$$\leq \mathbb{E}\left(\int_0^t \left(H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^\kappa) + CN^2\right)d\kappa\right)$$

$$= \int_0^t \mathbb{E}\left(H_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^\kappa) + CN^2\right)d\kappa$$

(6.34)
$$\leq t\left(\mathbb{E}(H_N(\underline{x}_N^0)) + CN^2\right),$$

where the ultimate line follows from the proof of (4.24), which is valid for any t, and the fact that we may always assume $\varepsilon < \min_{i\neq j}|x_i^0 - x_j^0|$. It now follows that

(6.35)
$$\mathbb{E}\left(\int_0^{t\wedge\tau_\varepsilon}\left|\int_{(\mathbb{R}^d)^2\setminus\triangle}((u_\varepsilon^\kappa - u^\kappa)(x) - (u_\varepsilon^\kappa - u^\kappa)(y))\cdot\nabla\mathbf{g}_{(\varepsilon)}(x-y)d(\mu_{N,\varepsilon}^\kappa)^{\otimes 2}(x,y)\right|d\kappa\right)$$

vanishes as $\varepsilon \to 0^+$.

- Almost surely, for $0 \leq \kappa \leq \tau_\varepsilon$,

(6.36)
$$\int_{(\mathbb{R}^d)^2\setminus\triangle}(u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y))\cdot\nabla\mathbf{g}_{(\varepsilon)}(x-y)d\mu_{N,\varepsilon}^\kappa(x)d\mu^\kappa(y)$$

$$= \frac{1}{N}\sum_{i=1}^N(\mathbb{M}\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa)(x_i^\kappa)\cdot(\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa)(x_i^\kappa) - \frac{1}{N}\sum_{i=1}^N\left(\mathbf{g}_{(\varepsilon)}*(\mathrm{div}(\mu^\kappa u_\varepsilon^\kappa))\right)(x_i^\kappa).$$

By Remark 2.5, the same reasoning as (6.23) and the nonincreasing property of L^p norms, we have

$$\left|(\mathbb{M}\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa)\cdot(\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa) - (\mathbb{M}\nabla\mathbf{g}*\mu^\kappa)\cdot(\nabla\mathbf{g}*\mu^\kappa)\right|$$

$$\leq \|\nabla(\mathbf{g}_{(\varepsilon)} - \mathbf{g})*\mu^\kappa\|_{L^\infty}(\|\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa\|_{L^\infty} + \|\nabla\mathbf{g}*\mu^\kappa\|_{L^\infty})$$

(6.37)
$$\lesssim \varepsilon^{d-s-1}\|\mu^0\|_{L^\infty}^{\frac{d+s+1}{d}}.$$

Similarly,

$$\left|\left(\mathbf{g}_{(\varepsilon)}*\mathrm{div}(\mu^\kappa u_\varepsilon^\kappa)\right)(x_i^\kappa) - \mathbf{g}*(\mathrm{div}(\mu^\kappa u^\kappa))(x_i^\kappa)\right|$$

$$\lesssim \varepsilon^{d-s-1}\|\mu^\kappa u_\varepsilon^\kappa\|_{L^\infty} + \|\mu^\kappa(u_\varepsilon^\kappa - u^\kappa)\|_{L^1}^{1-\frac{s+1}{d}}\|\mu^\kappa(u_\varepsilon^\kappa - u^\kappa)\|_{L^\infty}^{\frac{s+1}{d}}$$

(6.38)
$$\lesssim \varepsilon^{d-s-1}\|\mu^0\|_{L^\infty}^{\frac{d+s+1}{d}}.$$

After a little bookkeeping, we find that

(6.39)
$$\mathbb{E}\left(\int_0^{t\wedge\tau_\varepsilon}\left|\int_{(\mathbb{R}^d)^2\setminus\triangle}((u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y))\cdot\nabla\mathbf{g}_{(\varepsilon)}(x-y)d\mu_{N,\varepsilon}^\kappa(x)d\mu^\kappa(y)\right.\right.$$

$$\left.\left. - \int_{(\mathbb{R}^d)^2\setminus\triangle}(u^\kappa(x) - u^\kappa(y))\cdot\nabla\mathbf{g}(x-y)d\mu_N^\kappa(x)d\mu^\kappa(y)\right|d\kappa\right) \lesssim t\varepsilon^{d-s-1}\|\mu^0\|_{L^\infty}^{\frac{d+s+1}{d}},$$

which evidently vanishes as $\varepsilon \to 0^+$.

- Observe that

(6.40)
$$\int_{(\mathbb{R}^d)^2\setminus\triangle}(u_\varepsilon^\kappa(x) - u_\varepsilon^\kappa(y))\cdot\nabla\mathbf{g}_{(\varepsilon)}(x-y)d(\mu^\kappa)^{\otimes 2}(x,y)$$

$$= 2\int_{\mathbb{R}^d}u_\varepsilon^\kappa(x)\cdot(\nabla\mathbf{g}_{(\varepsilon)}*\mu^\kappa)(x)d\mu^\kappa(x).$$

By Lemma 2.3 and the nonincreasing property of L^p norms, arguing similarly as above, we have

$$\int_{\mathbb{R}^d} \left| u_\varepsilon^\kappa(x) \cdot (\nabla \mathsf{g}_{(\varepsilon)} * \mu^\kappa)(x) - u^\kappa(x) \cdot (\nabla \mathsf{g} * \mu^\kappa)(x) \right| d\mu^\kappa(x)$$

$$\leq \|\nabla(\mathsf{g} - \mathsf{g}_{(\varepsilon)}) * \mu^\kappa\|_{L^\infty} (\|\nabla \mathsf{g}_{(\varepsilon)} * \mu^\kappa\|_{L^\infty} + \|\nabla \mathsf{g} * \mu^\kappa\|_{L^\infty})$$

(6.41)
$$\lesssim \varepsilon^{d-s-1} \|\mu^0\|_{L^\infty}^{\frac{d+s+1}{d}}.$$

Therefore,

(6.42) $\quad \mathbb{E}\left(\int_0^{t \wedge \tau_\varepsilon} \int_{\mathbb{R}^d} \left| u_\varepsilon^\kappa(x) \cdot (\nabla \mathsf{g}_{(\varepsilon)} * \mu^\kappa)(x) - u^\kappa(x) \cdot (\nabla \mathsf{g} * \mu^\kappa)(x) \right| d\mu^\kappa(x) d\kappa \right)$

$$\lesssim t \varepsilon^{d-s-1} \|\mu^0\|_{L^\infty}^{\frac{d+s+1}{d}},$$

which evidently tends to zero as $\varepsilon \to 0^+$. With this last bit, the desired result (6.29) now follows.

Combining the above results, we see that we have shown the inequality

(6.43) $\quad \lim_{\varepsilon \to 0^+} \mathbb{E}\left(F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}^{\tau_\varepsilon \wedge t}, \mu^{\tau_\varepsilon \wedge t}) - F_N(\underline{x}_N^0, \mu^0) \right)$

$$\leq 2\sigma \mathbb{E}\left(\int_0^t \int_{\mathbb{R}^d} \Delta \mathsf{g}(x - y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x, y) d\kappa \right)$$

$$+ \mathbb{E}\left(\int_0^t \left| \int_{(\mathbb{R}^2)^2 \setminus \triangle} ((u^\kappa(x) - u^\kappa(y)) \cdot \nabla \mathsf{g}(x - y) d(\mu_N^\kappa - \mu^\kappa)^{\otimes 2}(x, y) \right| d\kappa \right).$$

Since $F_{N,\varepsilon}(\underline{x}_{N,\varepsilon}, \mu)$ is bounded from below uniformly in the noise and time by virtue of Proposition 5.1, we can conclude the proof by applying Fatou's lemma. $\qquad \square$

7. Gronwall argument. We now have all the ingredients necessary to conclude our Gronwall argument for the modulated energy, thereby proving Theorems 1.1 and 1.2. We divide this section into two subsections. In the first subsection, we consider the case where the admissible potential g is only superharmonic in a neighborhood of the origin (i.e. $r_0 < \infty$ in assumption (iii)). For such potentials, we obtain decay bounds for the modulated $F_N(\underline{x}_N^t, \mu^t)$ as $N \to \infty$ which grow linearly in time. The conclusion is the proof of Theorem 1.1. In the second subsection, we consider admissible potentials which are superharmonic on $\mathbb{R}^d$ (i.e. $r_0 = \infty$ in assumption (iii)). Under this stronger assumption, we can prove decay bounds for $F_N(\underline{x}_N^t, \mu^t)$ which are uniform on the interval $[0, \infty)$. The conclusion is the proof of Theorem 1.2.

7.1. *Linear-in-time estimates.* Applying Corollary 5.6 and Proposition 5.7 pointwise in time to the first and second terms, respectively, of the right-hand side of inequality (6.16) and using Remark 5.2 to control $|F_N(\underline{x}_N^t, \mu^t)|$ in terms of $F_N(\underline{x}_N^t, \mu^t)$, we find that

(7.1) $\quad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq |F_N(\underline{x}_N^0, \mu^0)| + C(1 + \|\mu^t\|_{L^\infty}) N^{-\frac{2}{2+s}} (1 + (\log N) \mathbf{1}_{s=0})$

$$+ C\sigma \int_0^t (1 + \|\mu^\kappa\|_{L^\infty}) N^{-\frac{\min\{2, d-s-2\}}{\min\{s+4, d\}}} d\kappa$$

$$+ C \int_0^t \left(\|\nabla u^\kappa\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} u^\kappa\|_{L^{\frac{2d}{d-2-s}}} \right) \left(F_N(\underline{x}_N^\kappa, \mu^\kappa) + C N^{-\frac{s+3}{(s+2)(s+1)}} \||\nabla|^{s+1-d} \mu^\kappa\|_{L^\infty} \right.$$

$$\left. + C(1 + \|\mu^\kappa\|_{L^\infty}) N^{-\frac{2}{(s+2)(s+1)}} \left(1 + (\log N) \mathbf{1}_{s=0} \right) \right) d\kappa,$$

where we have defined $u := \mathbb{M} \nabla g * \mu$. We remind the reader that the constant C depends on r_0 from assumption (iii).

Using Remark 2.5 and the fact that $\|\mu^\kappa\|_{L^1} = 1$, we see that

$$(7.2) \qquad \qquad \|\nabla u^\kappa\|_{L^\infty} \leq C \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}}.$$

Similarly, using the commutativity of Fourier multipliers together with the Hardy-Littlewood-Sobolev lemma, followed by Hölder's inequality

$$\||\nabla|^{\frac{d-s}{2}} u^\kappa\|_{L^{\frac{2d}{d-2-s}}} = \|\mathbb{M} \nabla g * (|\nabla|^{\frac{d-s}{2}} \mu^\kappa)\|_{L^{\frac{2d}{d-2-s}}} \leq C \|\mathcal{I}_{\frac{d-s-2}{2}}(\mu^\kappa)\|_{L^{\frac{2d}{d-2-s}}}$$

$$\leq C \|\mu^\kappa\|_{L^{\frac{d}{d-2-s}}}$$

$$(7.3) \qquad \qquad \leq C \|\mu^\kappa\|_{L^\infty}^{\frac{2+s}{d}}.$$

By another application of Lemma 2.3,

$$(7.4) \qquad \qquad \||\nabla|^{s+1-d} \mu^\kappa\|_{L^\infty} \leq \|\mu^\kappa\|_{L^\infty}^{\frac{s+1}{d}}.$$

Applying the bounds (7.2), (7.3), (7.4) to the right-hand side of (7.1), then applying the Gronwall-Bellman lemma, we find that

$$(7.5) \qquad \qquad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq A_N^t \exp\left(C \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} d\kappa \right),$$

where the time-dependent prefactor A_N^t is defined by

$$(7.6) \quad A_N^t := |F_N(\underline{x}_N^0, \mu^0)| + C(1 + \|\mu^t\|_{L^\infty}) N^{-\frac{2}{2+s}} (1 + (\log N) \mathbf{1}_{s=0})$$

$$+ C\sigma \int_0^t (1 + \|\mu^\kappa\|_{L^\infty}) N^{-\frac{\min\{2, d-s-2\}}{\min\{s+4, d\}}} d\kappa + C \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} \left(\|\mu^\kappa\|_{L^\infty}^{\frac{s+1}{d}} N^{-\frac{s+3}{(s+2)(s+1)}} \right.$$

$$\left. + C(1 + \|\mu^\kappa\|_{L^\infty}) N^{-\frac{2}{(s+2)(s+1)}} \left(1 + (\log N) \mathbf{1}_{s=0} \right) \right) d\kappa.$$

Without loss of generality, we may assume that $t \geq 1$. Split the interval $[0, t]$ into $[0, 1]$ and $[1, t]$. On $[0, 1]$, we use the trivial bound $\|\mu^\kappa\|_{L^\infty} \leq \|\mu^0\|_{L^\infty}$; and on $[1, t]$, we use the bound $\|\mu^\kappa\|_{L^\infty} \leq C(\sigma\kappa)^{-\frac{d}{2}}$, which comes from Proposition 3.8. It then follows that

$$\int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} d\kappa \leq \|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + C \int_1^t (\sigma\kappa)^{-\frac{s+2}{2}} d\kappa$$

$$(7.7) \qquad \qquad \leq \|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + C\left(\frac{2}{s\sigma^{\frac{s+2}{2}}} \mathbf{1}_{s>0} + \frac{(\log t)}{\sigma} \mathbf{1}_{s=0} \right),$$

$$(7.8) \qquad \qquad \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{2s+3}{d}} d\kappa \leq \|\mu^0\|_{L^\infty}^{\frac{2s+3}{d}} + \frac{C}{\sigma^{\frac{2s+3}{2}}},$$

and

$$(7.9) \quad A_N^t \leq |F_N(\underline{x}_N^0, \mu^0)| + C(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{2}{2+s}} (1 + (\log N)\mathbf{1}_{s=0})$$

$$+ C\sigma t(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{\min\{2, d-s-2\}}{\min\{s+4, d\}}} + C\Big(\|\mu^0\|_{L^\infty}^{\frac{2s+3}{d}} + \sigma^{-\frac{2s+3}{2}}\Big) N^{-\frac{s+3}{(s+2)(s+1)}}$$

$$+ Ct(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{2}{(s+2)(s+1)}} (1 + (\log N)\mathbf{1}_{s=0}).$$

Applying the preceding bounds and inserting into (7.5), we conclude

$$(7.10) \quad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq \exp\Bigg(C\Big(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0}\Big)\Bigg)$$

$$\times \Bigg(|F_N(\underline{x}_N^0, \mu^0)| + C(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{2}{2+s}} (1 + (\log N)\mathbf{1}_{s=0})$$

$$+ C\sigma t(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{\min\{2, d-s-2\}}{\min\{s+4, d\}}} + C\Big(\|\mu^0\|_{L^\infty}^{\frac{2s+3}{d}} + \sigma^{-\frac{2s+3}{2}}\Big) N^{-\frac{s+3}{(s+2)(s+1)}}$$

$$+ Ct(1 + \|\mu^0\|_{L^\infty}) N^{-\frac{2}{(s+2)(s+1)}} (1 + (\log N)\mathbf{1}_{s=0})\Bigg).$$

Comparing (7.10) to (1.18), we see that we have proved Theorem 1.1.

7.2. *Global-in-time estimates.*

We now assume that the potential g is globally superharmonic and show, using the results of Section 5.2, global-in-time bounds for the modulated energy $F_N(\underline{x}_N^t, \mu^t)$ for the range $0 < s < d - 2$ and almost-global-in-time bounds if $s = 0$. This proves Theorem 1.2.

Applying Corollary 5.14 and Proposition 5.15 pointwise in time to the first and second terms, respectively, of the right-hand side of inequality (6.16) and using Remark 5.10 to control $|F_N(\underline{x}_N^t, \mu^t)|$ in terms of $F_N(\underline{x}_N^t, \mu^t)$, we find that

$$(7.11) \quad \mathbb{E}\big(|F_N(\underline{x}_N^t, \mu^t)|\big) \leq |F_N(\underline{x}_N^0, \mu^0)|$$

$$+ C_p\|\mu^t\|_{L^\infty}^{\frac{s}{d}} N^{-\frac{\lambda_{s,p}}{\lambda_{s,p}+s}} (1 + (|\log\|\mu\|_{L^\infty}| + \log N)\mathbf{1}_{s=0})$$

$$+ C\sigma \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} \Big(C_q N^{-\frac{\lambda_{s+2,q}}{\lambda_{s+2,q}+s+2}}\mathbf{1}_{0 \leq s \leq d-4} + N^{-\frac{d-s-2}{d}}\mathbf{1}_{s>d-4}\Big) d\kappa$$

$$+ C \int_0^t \|\nabla u^\kappa\|_{L^\infty} \||\nabla|^{s+1-d}\mu^\kappa\|_{L^\infty} N^{-\frac{s+1+\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}} d\kappa$$

$$+ C \int_0^t \Big(\|\nabla u^\kappa\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} u^\kappa\|_{L^{\frac{2d}{d-2-s}}}\Big)\Big(F_N(\underline{x}_N^\kappa, \mu^\kappa)$$

$$+ C_p\big(1 + \|\mu^\kappa\|_{L^\infty}^{\gamma_{s,p}}\big) N^{-\frac{\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}} (1 + (\log N + |\log\|\mu^\kappa\|_{L^\infty}|)\mathbf{1}_{s=0})\Big) d\kappa$$

for any choices $\infty \geq p > \frac{d}{s+2}$ and $\infty \geq q > \frac{d}{s+4}$. The reader will recall the exponents $\gamma_{s,p}, \lambda_{s,p}$ from (5.21). Implicit here is the assumption that $N > \big(2^{\frac{dp-d+p}{(p-1)}}\|\mu^\kappa\|_{L^\infty}\big)^{\frac{(s+1)(s+\lambda_{s,p})}{d}}$ for every $\kappa \in [0, t]$, as required by Proposition 5.15. We can satisfy this constraint by assuming that $N > \big(2^{\frac{dp-d+p}{(p-1)}}\|\mu^0\|_{L^\infty}\big)^{\frac{(s+1)(s+\lambda_{s,p})}{d}}$, since $\|\mu^\kappa\|_{L^\infty}$ is nonincreasing.

Applying the bounds (7.2), (7.3), (7.4) to the right-hand side of (7.11), then applying the Gronwall-Bellman lemma, we find that

$$(7.12) \qquad \mathbb{E}(|F_N(\underline{x}_N^t, \mu^t)|) \leq B_N^t \exp\left(C \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} \, d\kappa \right),$$

where the time-dependent prefactor B_N^t is given by

$$(7.13) \quad B_N^t \coloneqq |F_N(\underline{x}_N^0, \mu^0)| + C_p \|\mu^t\|_{L^\infty}^{\frac{s}{d}} N^{-\frac{\lambda_{s,p}}{\lambda_{s,p}+s}} \left(1 + \left(|\log\|\mu^t\|_{L^\infty}| + \log N \right)\mathbf{1}_{s=0} \right)$$

$$+ C \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{3+2s}{d}} N^{-\frac{s+1+\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}} \, d\kappa$$

$$+ C\sigma \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} \left(C_q N^{-\frac{\lambda_{s+2,q}}{\lambda_{s+2,q}+s+2}} \mathbf{1}_{0\leq s\leq d-4} + N^{-\frac{d-s-2}{d}} \mathbf{1}_{s>d-4} \right) d\kappa$$

$$+ C_p \int_0^t \|\mu^\kappa\|_{L^\infty}^{\frac{s+2}{d}} \left(1 + \|\mu^\kappa\|_{L^\infty}^{\gamma_{s,p}} \right) N^{-\frac{\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}} \left(1 + (\log N + |\log\|\mu^\kappa\|_{L^\infty}|)\mathbf{1}_{s=0} \right) d\kappa.$$

Assuming $t \geq 1$ and splitting the interval $[0,t]$ into $[0,1],[1,t]$ exactly as in the last subsection, we find that

$$(7.14) \quad B_N^t \leq |F_N(\underline{x}_N^0, \mu^0)| + C\left(\|\mu^0\|_{L^\infty}^{\frac{3+2s}{d}} + \sigma^{-\frac{3+2s}{2}} \right) N^{-\frac{s+1+\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}}$$

$$+ C_p \min\{\|\mu^0\|_{L^\infty}^{\frac{s}{d}}, (\sigma t)^{-\frac{s}{2}}\} N^{-\frac{\lambda_{s,p}}{\lambda_{s,p}+s}} \left(1 + \left(\max\{|\log\|\mu^0\|_{L^\infty}|, |\log(\sigma t)|\} + \log N \right)\mathbf{1}_{s=0} \right)$$

$$+ C\sigma \left(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0} \right) \left(C_q N^{-\frac{\lambda_{s+2,q}}{\lambda_{s+2,q}+s+2}} \mathbf{1}_{0\leq s\leq d-4} \right.$$

$$\left. + N^{-\frac{d-s-2}{d}} \mathbf{1}_{s>d-4} \right) + C_p\left(1 + \|\mu^0\|_{L^\infty}^{\gamma_{s,p}} \right) \left(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0} \right).$$

Applying this bound to the right-hand side of (7.12) and using (7.7) for the exponential factor, we conclude that

$$(7.15) \quad \mathbb{E}\left(|F_N(\underline{x}_N^t, \mu^t)| \right) \leq \exp\left(C\left(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0} \right) \right)$$

$$\times \left(|F_N(\underline{x}_N^0, \mu^0)| + C_p \min\{\|\mu^0\|_{L^\infty}^{\frac{s}{d}}, (\sigma t)^{-\frac{s}{2}}\} N^{-\frac{\lambda_{s,p}}{\lambda_{s,p}+s}} \left(1 \right.\right.$$

$$+ \left(\max\{|\log\|\mu^0\|_{L^\infty}|, |\log(\sigma t)|\} + \log N \right)\mathbf{1}_{s=0} \Big) + C\sigma\Big(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}}$$

$$+ \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0} \Big) \left(C_q N^{-\frac{\lambda_{s+2,q}}{\lambda_{s+2,q}+s+2}} \mathbf{1}_{0\leq s\leq d-4} + N^{-\frac{d-s-2}{d}} \mathbf{1}_{s>d-4} \right)$$

$$+ C_p\left(1 + \|\mu^0\|_{L^\infty}^{\gamma_{s,p}} \right) \left(\|\mu^0\|_{L^\infty}^{\frac{s+2}{d}} + \frac{2}{s\sigma^{\frac{s+2}{2}}}\mathbf{1}_{s>0} + (\log t^{\frac{1}{\sigma}})\mathbf{1}_{s=0} \right)$$

$$+ C\left(\|\mu^0\|_{L^\infty}^{\frac{3+2s}{d}} + \sigma^{-\frac{3+2s}{2}} \right) N^{-\frac{s+1+\lambda_{s,p}}{(s+\lambda_{s,p})(1+s)}} \bigg).$$

REFERENCES

[1] BERMAN, R. J. and ÖNNHEIM, M. (2019). Propagation of chaos for a class of first order models with singular mean field interactions. *SIAM J. Math. Anal.* **51** 159–196. https://doi.org/10.1137/18M1196662 MR3904425

[2] BOLLEY, F., CAÑIZO, J. A. and CARRILLO, J. A. (2011). Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.* **21** 2179–2210. https://doi.org/10.1142/S0218202511005702 MR2860672

[3] BOLLEY, F., GUILLIN, A. and MALRIEU, F. (2010). Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.* **44** 867–884. https://doi.org/10.1051/m2an/2010045 MR2731396

[4] BRESCH, D., JABIN, P.-E. and WANG, Z. (2019). Modulated free energy and mean field limit. *Séminaire Laurent Schwartz–EDP et applications* 1–22.

[5] BRESCH, D., JABIN, P.-E. and WANG, Z. (2019). 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** 708–720. https://doi.org/10.1016/j.crma.2019.09.007 MR4018082

[6] BRESCH, D., JABIN, P.-E. and WANG, Z. (2020). Mean-field limit and quantitative estimates with singular attractive kernels. *arXiv preprint arXiv:2011.08022*.

[7] CARLEN, E. A. (1991). Some integral identities and inequalities for entire functions and their application to the coherent state transform. *J. Funct. Anal.* **97** 231–249. https://doi.org/10.1016/0022-1236(91)90022-W MR1105661

[8] CARLEN, E. A. and LOSS, M. (1990). Extremals of functionals with competing symmetries. *J. Funct. Anal.* **88** 437–456. https://doi.org/10.1016/0022-1236(90)90114-Z MR1038450

[9] CARLEN, E. A. and LOSS, M. (1995). Optimal smoothing and decay estimates for viscously damped conservation laws, with applications to the 2-D Navier-Stokes equation. *Duke Math. J.* **81** 135 – 157. https://doi.org/10.1215/S0012-7094-95-08110-1

[10] CARRILLO, J. A., CHOI, Y.-P. and HAURAY, M. (2014). The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds. CISM Courses and Lect.* **553** 1–46. Springer, Vienna. https://doi.org/10.1007/978-3-7091-1785-9_1 MR3331178

[11] CARRILLO, J. A., FERREIRA, L. C. F. and PRECIOSO, J. C. (2012). A mass-transportation approach to a one dimensional fluid mechanics model with nonlocal velocity. *Adv. Math.* **231** 306–327. https://doi.org/10.1016/j.aim.2012.03.036 MR2935390

[12] CATTIAUX, P., GUILLIN, A. and MALRIEU, F. (2008). Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** 19–40. https://doi.org/10.1007/s00440-007-0056-3 MR2357669

[13] CHIZAT, L. and BACH, F. (2018). On the Global Convergence of Gradient Descent for Over-Parameterized Models Using Optimal Transport. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems. NIPS18* 3040-3050. Curran Associates Inc., Red Hook, NY, USA.

[14] CHORIN, A. J. (1973). Numerical study of slightly viscous flow. *J. Fluid Mech.* **57** 785–796. https://doi.org/10.1017/S0022112073002016 MR395483

[15] DAVIES, E. B. (1987). Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.* **109** 319–333. https://doi.org/10.2307/2374577 MR882426

[16] DELARUE, F. and TSE, A. (2021). Uniform in time weak propagation of chaos on the torus. *arXiv preprint arXiv:2104.14973*.

[17] DUERINCKX, M. (2016). Mean-Field Limits for Some Riesz Interaction Gradient Flows. *SIAM Journal on Mathematical Analysis* **48** 2269-2300. https://doi.org/10.1137/15M1042620

[18] DURMUS, A., EBERLE, A., GUILLIN, A. and ZIMMER, R. (2020). An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** 5387–5398. https://doi.org/10.1090/proc/14612 MR4163850

[19] FOURNIER, N., HAURAY, M. and MISCHLER, S. (2014). Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)* **16** 1423–1466. https://doi.org/10.4171/JEMS/465 MR3254330

[20] FOURNIER, N. and JOURDAIN, B. (2017). Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.* **27** 2807–2861. https://doi.org/10.1214/16-AAP1267 MR3719947

[21] GALLAY, T. and WAYNE, C. E. (2005). Global stability of vortex solutions of the two-dimensional Navier-Stokes equation. *Comm. Math. Phys.* **255** 97–129. https://doi.org/10.1007/s00220-004-1254-9 MR2123378

[22] GODINHO, D. and QUIÑINAO, C. (2015). Propagation of chaos for a subcritical Keller-Segel model. *Annales de l'I.H.P. Probabilités et statistiques* **51** 965–992. https://doi.org/10.1214/14-AIHP606 MR3365970

[23] GOLSE, F. (2016). On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*. *Lect. Notes Appl. Math. Mech.* **3** 1–144. Springer. https://doi.org/10.1007/978-3-319-26883-5_1 MR3468297

[24] GRAFAKOS, L. (2014). *Classical Fourier Analysis*, third ed. *Graduate Texts in Mathematics* **249**. Springer.

[25] GRAFAKOS, L. (2014). *Modern Fourier Analysis*, third ed. *Graduate Texts in Mathematics* **250**. Springer.

[26] GROSS, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. https://doi.org/10.2307/2373688 MR420249

[27] GUILLIN, A., BRIS, P. L. and MONMARCHÉ, P. (2021). Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems. *arXiv preprint arXiv:2108.08675*.

[28] HAURAY, M. (2009). Wasserstein distances for vortices approximation of Euler-type equations. *Math. Models Methods Appl. Sci.* **19** 1357–1384. https://doi.org/10.1142/S0218202509003814 MR2555474

[29] HAURAY, M. and MISCHLER, S. (2014). On Kac's chaos and related problems. *J. Funct. Anal.* **266** 6055–6157. https://doi.org/10.1016/j.jfa.2014.02.030 MR3188710

[30] HEGSELMANN, R. and KRAUSE, U. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. *Journal of Artificial Societies and Social Simulation* **5** 1-2.

[31] HOLDING, T. (2016). Propagation of chaos for Hölder continuous interaction kernels via Glivenko-Cantelli. *arXiv preprint arXiv:1608.02877*.

[32] JABIN, P.-E. (2014). A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models* **7** 661–711. https://doi.org/10.3934/krm.2014.7.661 MR3317577

[33] JABIN, P.-E. and WANG, Z. (2016). Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.* **271** 3588–3627. https://doi.org/10.1016/j.jfa.2016.09.014 MR3558251

[34] JABIN, P.-E. and WANG, Z. (2017). Mean field limit for stochastic particle systems. In *Act. Part. Vol. 1. Advances theory, Model. Appl.. Model. Simul. Sci. Eng. Technol.* 379–402. Birkhäuser/Springer, Cham.

[35] JABIN, P.-E. and WANG, Z. (2018). Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.* **214** 523–591. https://doi.org/10.1007/s00222-018-0808-y MR3858403

[36] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian motion and stochastic calculus*, second ed. *Graduate Texts in Mathematics* **113**. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4612-0949-2 MR1121940

[37] KRAUSE, U. (2000). A discrete nonlinear and non-autonomous model of consensus formation. *Communications in Difference Equations*.

[38] LACKER, D. (2021). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *arXiv preprint arXiv:2105.02983*.

[39] LI, L., LIU, J.-G. and YU, P. (2019). On the mean field limit for Brownian particles with Coulomb interaction in 3D. *J. Math. Phys.* **60** 111501, 34. https://doi.org/10.1063/1.5114854 MR4026330

[40] LIU, J.-G. and YANG, R. (2016). Propagation of chaos for large Brownian particle system with Coulomb interaction. *Res. Math. Sci.* **3** Paper No. 40, 33. https://doi.org/10.1186/s40687-016-0086-5 MR3572548

[41] MAEKAWA, Y. (2008). A lower bound for fundamental solutions of the heat convection equations. *Arch. Ration. Mech. Anal.* **189** 45–58. https://doi.org/10.1007/s00205-007-0088-y MR2403599

[42] MALRIEU, F. (2003). Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13** 540–560. https://doi.org/10.1214/aoap/1050689593 MR1970276

[43] MARCHIORO, C. and PULVIRENTI, M. (2012). *Mathematical theory of incompressible nonviscous fluids* **96**. Springer Science & Business Media.

[44] MCKEAN, H. P. JR. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)* 41–57. Air Force Office Sci. Res., Arlington, Va. MR0233437

[45] MEI, S., MONTANARI, A. and NGUYEN, P.-M. (2018). A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA* **115** E7665–E7671. https://doi.org/10.1073/pnas.1806579115 MR3845070

[46] MÉLÉARD, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995). Lecture Notes in Math.* **1627** 42–95. Springer, Berlin. https://doi.org/10.1007/BFb0093177 MR1431299

[47] MISCHLER, S., MOUHOT, C. and WENNBERG, B. (2015). A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probab. Theory Related Fields* **161** 1–59. https://doi.org/10.1007/s00440-013-0542-8 MR3304746

[48] MOTSCH, S. and TADMOR, E. (2011). A new model for self-organized dynamics and its flocking behavior. *J. Stat. Phys.* **144** 923–947. https://doi.org/10.1007/s10955-011-0285-9 MR2836613

[49] NASH, J. (1958). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* **80** 931–954. https://doi.org/10.2307/2372841 MR100158

[50] NGUYEN, Q. H., ROSENZWEIG, M. and SERFATY, S. (2021). Mean-field limits of Riesz-type singular flows with possible multiplicative transport noise. *arXiv preprint arXiv:2107.02592*.

[51] ONSAGER, L. (1949). Statistical hydrodynamics. *Nuovo Cimento (9)* **6** 279–287. MR36116

[52] OSADA, H. (1986). Propagation of chaos for the two-dimensional Navier-Stokes equation. *Proc. Japan Acad. Ser. A Math. Sci.* **62** 8–11. MR839794

[53] OSADA, H. (1987). Propagation of chaos for the two-dimensional Navier-Stokes equation. In *Probabilistic methods in mathematical physics (Katata/Kyoto, 1985)* 303–334. Academic Press, Boston, MA. MR933829

[54] OSADA, H. (1987). Limit points of empirical distributions of vortices with small viscosity. In *Hydrodynamic behavior and interacting particle systems (Minneapolis, Minn., 1986). IMA Vol. Math. Appl.* **9** 117–126. Springer, New York. https://doi.org/10.1007/978-1-4684-6347-7_10 MR914990

[55] OSADA, H. (1987). Diffusion processes with generators of generalized divergence form. *J. Math. Kyoto Univ.* **27** 597–619. https://doi.org/10.1215/kjm/1250520601 MR916761

[56] OTHMER, H. G. and STEVENS, A. (1997). Aggregation, blowup, and collapse: the ABCs of taxis in reinforced random walks. *SIAM J. Appl. Math.* **57** 1044–1081. https://doi.org/10.1137/S0036139995288976 MR1462051

[57] PERTHAME, B. (2007). *Transport equations in biology. Frontiers in Mathematics*. Birkhäuser Verlag, Basel. MR2270822

[58] ROSENZWEIG, M. (2020). The Mean-Field Limit of Stochastic Point Vortex Systems with Multiplicative Noise. *arXiv preprint arXiv:2011.12180*.

[59] ROSENZWEIG, M. (2022). Mean-Field Convergence of Point Vortices to the Incompressible Euler Equation with Vorticity in L^∞. *Arch. Ration. Mech. Anal.* **243** 1361–1431. https://doi.org/10.1007/s00205-021-01735-3 MR4381143

[60] ROTSKOFF, G. M. and VANDEN-EIJNDEN, E. (2018). Trainability and accuracy of neural networks: An interacting particle system approach. *arXiv preprint arXiv:1805.00915*.

[61] ROUGERIE, N. and SERFATY, S. (2016). Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.* **69** 519–605. https://doi.org/10.1002/cpa.21570 MR3455593

[62] SERFATY, S. (2017). Mean field limits of the Gross-Pitaevskii and parabolic Ginzburg-Landau equations. *J. Amer. Math. Soc.* **30** 713–768. https://doi.org/10.1090/jams/872 MR3630086

[63] SERFATY, S. (2020). Mean field limit for Coulomb-type flows. *Duke Math. J.* **169** 2887–2935. Appendix with Mitia Duerinckx. https://doi.org/10.1215/00127094-2020-0019

[64] SERFATY, S. and VÁZQUEZ, J. L. (2014). A mean field equation as limit of nonlinear diffusions with fractional Laplacian operators. *Calc. Var. Partial Differential Equations* **49** 1091–1120. https://doi.org/10.1007/s00526-013-0613-9 MR3168624

[65] STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions* **2**. Princeton University Press.

[66] STEIN, E. M. and MURPHY, T. S. (1993). *Harmonic Analysis (PMS-43): Real-Variable Methods, Orthogonality, and Oscillatory Integrals. (PMS-43)*. Princeton University Press.

[67] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. https://doi.org/10.1007/BFb0085169 MR1108185

[68] TOPAZ, C. M., BERTOZZI, A. L. and LEWIS, M. A. (2006). A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.* **68** 1601–1623. https://doi.org/10.1007/s11538-006-9088-6 MR2257718

[69] WANG, Z., ZHAO, X. and ZHU, R. (2021). Gaussian fluctuations for interacting particle systems with singular kernels. *arXiv preprint arXiv:2105.13201*.

[70] XIA, H., WANG, H. and XUAN, Z. (2011). Opinion Dynamics: A Multidisciplinary Review and Perspective on Future Research. *Int. J. Knowl. Syst. Sci.* **2** 72-91. https://doi.org/10.4018/jkss.2011100106