Ars Inveniendi Analytica (202X), Paper No. X, XX pp.
DOI [journal inserts paper DOI]
ISSN: 2769-8505

Mean-field limits of Riesz-type singular flows

Quoc-Hung Nguyen
Academy of Mathematics and Systems Science, Chinese Academy of Sciences

Matthew Rosenzweig
Massachusetts Institute of Technology

Sylvia Serfaty
Courant Institute of Mathematical Sciences, New York University

Communicated by Francesco Maggi

Abstract. *We provide a proof of mean-field convergence of first-order dissipative or conservative dynamics of particles with Riesz-type singular interaction (the model interaction is an inverse power s of the distance for any $0 < s < d$) when assuming a certain regularity of the solutions to the limiting evolution equations. It relies on a modulated-energy approach, as introduced in previous works where it was restricted to the Coulomb and super-Coulombic cases. The method is also capable of incorporating multiplicative noise of transport type into the dynamics. It relies in extending functional inequalities of* [Ser20b, Ros20, Ser20a] *to more general interactions, via a new, robust proof that exploits a certain commutator structure.*

Keywords. mean-field limits, propagation of chaos, Coulomb and Riesz interactions, modulated energy, renormalized commutator.

1. Introduction

1.1. **Main goal and method.** In this article, we consider the first-order mean-field dynamics of interacting particle systems of the form

$$\begin{cases} \dot{x}_i = \dfrac{1}{N} \displaystyle\sum_{1\leqslant j\leqslant N: j\neq i} \mathbb{M}\nabla \mathrm{g}(x_i - x_j) \\ x_i|_{t=0} = x_i^0, \end{cases} \qquad i \in \{1,\dots,N\}. \tag{1.1}$$

Here, $N \in \mathbb{N}$ is the number of particles, $x_i^0 \in \mathbb{R}^d$ are the pairwise distinct initial configurations and $\mathbb{M}$ is a matrix satisfying

$$\forall \xi \in \mathbb{R}^d, \quad \langle \mathbb{M}\xi, \xi\rangle \leqslant 0, \tag{1.2}$$

where $\langle\cdot,\cdot\rangle$ is the standard Euclidean scalar product. Taking $\mathbb{M} = -\mathbb{I}$ yields *gradient-flow/ dissipative* dynamics, while taking $\mathbb{M}$ to be an antisymmetric matrix yields *Hamiltonian* dynamics. General $\mathbb{M}$ allow for possible mixed flows. We assume that the potential g is repulsive, so that particles never collide and there is a unique global solution to the system of ODEs (1.1). The model case we have in mind is g a Riesz potential depending on a parameter $0 \leqslant s < d$ according to the convention

$$\mathrm{g}(x) = \begin{cases} -\log|x|, & s = 0 \\ |x|^{-s}, & s > 0. \end{cases} \tag{1.3}$$

We will in fact consider a slightly more general class of potentials with singularity of the type (1.3), as explained in the next subsection. Based on the value of the parameter s and its relation to the dimension d, we classify the potential g as *sub-Coulombic* if $0 \leqslant s < d-2$, *Coulombic* if $s = d-2$, and *super-Coulombic* if $d-2 < s < d$.

Motivations for studying systems of the form (1.1) are numerous: interacting particles in physics, particle approximation to PDEs, finding equilibrium states of interaction energies, biological and sociological models, large neural networks (see for instance the introduction in [Ser20b]). Riesz interactions are particularly interesting for physics and approximation theory, see for instance [DRAW02, BHS19].

By mean-field limit, we mean proving the convergence as $N \to \infty$ of the *empirical measure*

$$\mu_N^t \coloneqq \frac{1}{N}\sum_{i=1}^N \delta_{x_i^t} \tag{1.4}$$

associated to a solution $\underline{x}_N^t \coloneqq (x_1^t,\dots,x_N^t)$ of the system (1.1). 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 derivation leads to the expectation that for $t > 0$, μ_N^t converges to the solution of the Cauchy problem with initial data μ^0 for the limiting evolution equation

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

with $*$ denoting the usual convolution. Proving the convergence of the empirical measure is closely related to proving *propagation of molecular chaos* (see [Gol16, HM14, Jab14] and references therein) which means showing that if $f_N^0(x_1,\dots,x_N)$ is the initial probability density of seeing particles at positions $(x_1,\dots,x_N)$ and if f_N^0 converges to some factorized state $(\mu^0)^{\otimes N}$, then the k-point marginals $f_{N,k}^t$ converge for all time to $(\mu^t)^{\otimes k}$. Our result implies a convergence of this type as well (see Theorem 1.6 below).

We note that we restrict to the case $s<d$ because for $s\geqslant d$, called the *hypersingular case*, the potential g is not integrable near the origin and the convolution $\mathrm{g}*\mu$ no longer makes sense for a nonnegative measure μ. In fact, the expected evolution equation is no longer of the form (1.5), see for instance [HSST21]. This is the reason for restricting ourselves to the *potential case* $s<d$.

Mean-field limits for systems of particles with regular (typically, Lipschitz) interactions have been understood for a long time [Dob79, Szn91], with a trajectorial approach consisting in comparing the positions of the particles to the characteristics of the limiting PDE. In contrast, systems with singular interaction are far more challenging and only in recent years have breakthroughs been made to cover the full potential case $s<d$. Various approaches have been put forward to study them. They usually involve finding a good metric to measure the distance of the empirical measure to its evolving expected limit and showing a Gronwall relation on the time evolution of that metric. The use of the ∞-Wasserstein metric has been successful in the sub-Coulombic case $s<d-2$, [Hau09, CCH14]. The modulated-energy approach of [Due16, Ser20b], itself inspired from [Ser17], has, in contrast, been successful in the Coulomb and super-Coulombic case $d-2\leqslant s<d$ (in low dimension and the dissipative case in [Due16], and generally in [Ser20b]). There, the metric used, called the modulated energy, is a Coulomb/Riesz-based metric, based on the interaction itself, and which can be understood as a renormalized negative-order homogeneous Sobolev norm. To be more precise, the modulated energy or (squared) "distance" between $\frac{1}{N}\sum_{i=1}^N\delta_{x_i}$ and μ is defined to be

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

where we excise the diagonal $\triangle$ in order to remove the infinite self-interaction of each particle. Again, a Gronwall-type inequality can be proved for this modulated energy.

In parallel, work of Jabin and Wang [JW16, JW18] put forward a relative-entropy based method which can handle moderately singular interactions, and is particularly tailored for systems with added noise/diffusion. The two approaches (modulated energy and relative entropy) were recently combined in [BJW19b, BJW19a, BJW20] into a *modulated free energy* method capable of treating *dissipative* singular flows with additive noise. In the course of that combination, these works also relaxed the assumptions on the interaction (their setting is the torus, which changes the long-range effects though), in particular allowing even for attractive potentials, such as those of Patlak-Keller-Segel-type. Note that the one-dimensional case can be handled via a Wasserstein-gradient-flow approach

[CFP12, BO19] using the fact that in (and only in) dimension 1, the Riesz interaction is convex.

Our present concern is showing that the modulated-energy method functions beyond the Coulombic or super-Coulombic Riesz case. Proving an exact Gronwall relation on the modulated energy requires showing that the derivative of the modulated energy along particle displacement is bounded by the modulated energy itself. This is in essence a functional inequality, which can be proven independently of any dynamics. Plugging into (1.6) the solution $\underline{x}_N^t$ of (1.1) and the solution μ^t of (1.5), direct computation (see [Ser20b, Lemma 2.1]) yields the inequality

$$\frac{d}{dt}F_N(\underline{x}_N^t,\mu^t) \leqslant \int_{(\mathbb{R}^d)^2\setminus\triangle} \big(u^t(x)-u^t(y)\big)\cdot\nabla \mathsf{g}(x-y)\, d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i^t}-\mu^t\right)^{\otimes 2}(x,y), \tag{1.7}$$

where $u^t := \mathbb{M}\nabla(\mathsf{g}*\mu^t)$. Thus, we need to estimate the right-hand side, which is done by proving the following functional inequality: for any configuration $\underline{x}_N := (x_1,\dots,x_N)$, any suitably decaying, bounded probability density μ, and any suitably regular vector field $v:\mathbb{R}^d\to\mathbb{R}^d$,

$$\int_{(\mathbb{R}^d)^2\setminus\triangle} (v(x)-v(y))\cdot\nabla \mathsf{g}(x-y)\, d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}-\mu\right)^{\otimes 2}(x,y) \;\leqslant\; CF_N(\underline{x}_N,\mu)+CN^{-\alpha} \tag{1.8}$$

for some $C>0$ and $\alpha>0$ depending only on norms of μ and v. The precise statement of the functional inequality (1.8) is given in Theorem 4.1 below.

In the prior works [Due16, Ser20b] that used the modulated-energy method, the exact Coulomb nature of g and the fact that the potential $\mathsf{h} := \mathsf{g}*\mu$ solves a *local equation* in terms of μ was used crucially as it allowed to rewrite the left-hand side of (1.8) in terms of the *stress-energy tensor* associated to h. Thanks to a dimension extension procedure popularized by Caffarelli and Silvestre [CS07], the super-Coulombic case $d-2<s<d$ could also be handled this way by going through a local operator and a stress-energy tensor. This approach was rigid in the exact form of the interaction. Later in [BJW19a], a Fourier-based approach to proving a relation similar to (1.8) was employed and allows, in the case of the torus, to give more flexibility on the interaction, which still needs to have Riesz-type behavior in physical and Fourier space. In fact, the assumptions placed there are similar to the general assumptions imposed in this article, a point on which we elaborate more in Section 1.3 below. In [Ser20b], in addition to the stress-tensor structure, a truncation or smearing procedure with a *truncation radius depending on the points* and a *monotonicity* with respect to the smearing radii were crucially used. This allows for a "renormalization" of the stress-energy tensor and the energy in order to account for the removal of the diagonal in (1.6) and (1.8).

Our main achievement here is to show that the stress-energy tensor structure is not needed to prove (1.8), nor is a Fourier-based approach. Rather, some underlying *commutator structure* to the left-hand side of (1.8) can be exploited for all Riesz interactions (with $s<d$). We write here of commutator structure because one may view the left-hand side of

(1.8) as a renormalization of the (dualized) commutator

$$\int_{\mathbb{R}^d} \big(v \cdot \nabla(\mathrm{g} * f) - \mathrm{g} * \big(\nabla \cdot (vf)\big)\big) dx, \tag{1.9}$$

or equivalently

$$\left\langle f, \left[v, \frac{\nabla}{(-\Delta)^{\frac{d-s}{2}}} \right] f \right\rangle_{L^2}, \tag{1.10}$$

where

$$f := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i} - \mu.$$

In fact, in the exact Riesz cases, we could also prove (1.8) via off-the-shelf commutator estimates from the harmonic analysis literature on paraproducts, specifically refinements by Li [Li19] of the classical fractional Leibnitz rules due to Kato-Ponce [KP88] and Kenig-Ponce-Vega [KPV93]. Instead, we provide a fairly direct and essentially physical-space proof based on integration by parts. The Coulomb case $s = d-2$ is the most delicate. In that case–and only to obtain the sharp estimate–we appeal to a deep result of Christ and Journé [CJ87] (see also [SSS19, Lai20] for extensions) for commutators of Calderón-Zygmund operators. Following the terminology of Christ and Journé, such expressions are known as *Calderón d-commutators*, generalizations of Calderón's classic one-dimensional commutator [Cal80]. With a little work, the structure in the left-hand side of (1.8) allows for a reduction to estimating such a Calderón d-commutator. We mention that such commutator expressions often appear in studying fluid problems. See for example [Mar08, CCC$^+$12, HSSS18].

The truncation or smearing procedure of [Ser20b] is still needed and is done in a slightly different way depending on whether one is in the sub-Coulombic/Coulombic or the super-Coulombic case. In the sub-Coulombic/Coulombic case, the nature of the interaction makes g superharmonic, which makes the smearing monotone in the right way. The way we handle super-Coulombic interactions is then by noticing that a subharmonic Riesz potential becomes superharmonic when viewed in a larger dimension. We thus appeal again to a dimension-extension procedure (different from that of [Ser20b], however) to perform the smearing of the Dirac masses without other modification of the method. This thus provides a unified treatment of all Riesz interactions. Since the stress-energy tensor structure is no longer used, we do not need the interaction to be the kernel of a local operator and to be exactly an inverse power. Consequently, we can relax the assumptions in a way similar to [BJW19a].

Our approach is, in principle, robust to the inclusion of multiplicative noise of transport type. Treating the case with such noise was done in [Ros20] for the two-dimensional Coulomb case, with the proof generalizing to higher dimensions. Due to the nonzero quadratic variation of the noise, it involves proving another functional inequality (see Theorem 6.1 below), which is exactly a second-order version of (1.8). In [Ros20], the Coulomb nature of the interaction was strongly used. But with the commutator-based approach and the new insights of this article, it is then not too difficult to prove the desired functional

inequality in all the Riesz cases we are considering. With this second-order functional inequality and assuming the existence of suitably regular solutions to the limiting stochastic PDEs (SPDEs) to justify the computations, one can follow the road map of [Ros20] to bound the expected magnitude of the modulated energy. Such a bound, of course, implies convergence in law of the empirical measure to the solution of the limiting SPDE.

Added in proof: While our manuscript was under review, a new approach to mean-field limits of first- and second-order systems with additive noise was introduced by Bresch *et al.* [BJS22]. This approach is based on estimates for the BBGKY hierarchy which take advantage of the diffusion to bound weighted L^p norms of the marginals, and has similarities with the hierarchy approach of Lacker [Lac21]. In particular, the work [BJS22] manages to prove the mean-field limit for the repulsive Vlasov-Fokker-Planck equation with Coulomb potential in dimension 2 and improves upon the prior work of Jabin-Wang [JW18].

1.2. **Assumptions and statement of main results.** We now state the precise assumptions for the class of interaction potentials we consider. In the statements below and throughout this article, the notation $\mathbf{1}_{(\cdot)}$ denotes the indicator function for the condition $(\cdot)$.

For $d \geqslant 2$ and $0 \leqslant s \leqslant d-2$, we assume the following:

$$\mathrm{g}(x) = \mathrm{g}(-x) \quad \forall x \in \mathbb{R}^d \setminus \{0\} \tag{1.11}$$

$$\lim_{x\to 0} \mathrm{g}(x) = +\infty \tag{1.12}$$

$$\exists r_0 > 0 \text{ such that } \quad \Delta \mathrm{g} \leqslant 0 \quad \text{in } B(0,r_0)^1 \tag{1.13}$$

$$\forall k \geqslant 0, \qquad |\nabla^{\otimes k} \mathrm{g}(x)| \leqslant C\left(\frac{1}{|x|^{s+k}} + |\log|x|| \mathbf{1}_{s=k=0}\right) \quad \forall x \in \mathbb{R}^d \setminus \{0\} \tag{1.14}$$

$$|x||\nabla \mathrm{g}(x)| + |x|^2 |\nabla^{\otimes 2} \mathrm{g}(x)| \leqslant C\mathrm{g}(x) \quad \forall x \in B(0,r_0)\setminus\{0\} \tag{1.15}$$

$$\frac{C_1}{|\xi|^{d-s}} \leqslant \hat{\mathrm{g}}(\xi) \leqslant \frac{C_2}{|\xi|^{d-s}} \quad \forall \xi \in \mathbb{R}^d \setminus \{0\} \tag{1.16}$$

and

$$\exists c_s < 1 \text{ such that } \quad \mathrm{g}(y) < c_s \mathrm{g}(x) \quad \forall x, y \in B(0,r_0) \setminus \{0\} \text{ with } |y| \geqslant 2|x|. \tag{1.17}$$

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

$$\mathrm{k}(x-y) := (x-y) \otimes \nabla^{\otimes(2k+1)} \mathrm{g}(x-y) \tag{1.18}$$

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

is associated to a Calderón-Zygmund operator in $\mathbb{R}^d$.[2] For $s=0$, we replace the assumption (1.17) with

$$\exists c_0>0 \text{ such that } \quad \mathrm{g}(x)-\mathrm{g}(y)\geqslant c_0 \quad \forall x,y\in B(0,r_0)\setminus\{0\} \text{ with } |y|\geqslant 2|x|. \tag{1.19}$$

If $\max\{d-2,0\}<s<d$ or $d=1$ and $s=0$, then we expect the local superharmonicity assumption (1.13) to fail. The idea, though, is that g becomes superharmonic when viewing it as an interaction potential in an extended space of dimension $d+m$. Indeed, $\Delta|X|^{-s}=-s(d+m-s-2)|X|^{-s-2}$ in $\mathbb{R}^{d+m}$, which becomes $\leqslant 0$ if m is large enough. We will denote points in $\mathbb{R}^{d+m}$ by $X=(x,z)$ with $x\in\mathbb{R}^d$ and $z\in\mathbb{R}^m$. Our assumption for the case $\max\{d-2,0\}<s<d$ or $d=1$ and $s=0$ is thus that there exists an integer $m\geqslant 0$, $\mathsf{G}:\mathbb{R}^{d+m}\to\mathbb{R}$, and constants $C_1,C_2,C>0$ depending only on d,s,m, such that

$$\mathsf{G}(x,0)=\mathrm{g}(x) \quad \forall (x,0)\in\mathbb{R}^{d+m} \tag{1.20}$$

$$\mathsf{G}(X)=\mathsf{G}(-X) \quad \forall X\in\mathbb{R}^{d+m}\setminus\{0\} \tag{1.21}$$

$$\lim_{X\to 0}\mathsf{G}(X)=+\infty \tag{1.22}$$

$$\exists r_0>0 \text{ such that } \quad \Delta\mathsf{G}\leqslant 0 \quad \text{in } B(0,r_0)\subset\mathbb{R}^{d+m} \tag{1.23}$$

$$\forall k\geqslant 0, \quad |\nabla^{\otimes k}\mathsf{G}(X)|\leqslant C\left(\frac{1}{|X|^{s+k}}+|\log|X||\mathbf{1}_{s=k=0\wedge d=1}\right) \quad \forall X\in\mathbb{R}^{d+m}\setminus\{0\} \tag{1.24}$$

$$|X||\nabla\mathsf{G}(X)|+|X|^2|\nabla^{\otimes 2}\mathsf{G}(X)|\leqslant C\mathsf{G}(X) \quad \forall X\in B(0,r_0)\setminus\{0\} \tag{1.25}$$

$$\frac{C_1}{|\Xi|^{d+m-s}}\leqslant\hat{\mathsf{G}}(\Xi)\leqslant\frac{C_2}{|\Xi|^{d+m-s}} \quad \forall\Xi\in\mathbb{R}^{d+m}\setminus\{0\}. \tag{1.26}$$

We also add

$$\begin{cases}\exists c_s<1 \text{ such that } \ \mathsf{G}(Y)<c_s\mathsf{G}(X) \quad \forall X,Y\in B(0,r_0)\setminus\{0\} \text{ with } |Y|\geqslant 2|X|, & s>0\\ \exists c_0>0 \text{ such that } \ \mathsf{G}(X)-\mathsf{G}(Y)\geqslant c_0 \quad \forall X,Y\in B(0,r_0)\setminus\{0\} \text{ with } |Y|\geqslant 2|X|, & s=0.\end{cases} \tag{1.27}$$

If $s=d-1$ and $m=1$, then we also assume that the $(\mathbb{R}^d)^{\otimes 4}$-valued kernel

$$\mathsf{K}(X-Y):=(X-Y)\otimes\nabla^{\otimes 3}\mathsf{G}(X-Y), \qquad \forall X\neq Y \tag{1.28}$$

is associated to a Calderón-Zygmund operator in $\mathbb{R}^{d+m}$. We can simultaneously treat the cases $0\leqslant s\leqslant d-2$ and $\max\{d-2,0\}<s<d$ or $d=1$ and $s=0$ by setting $m=0$ if $0\leqslant s\leqslant d-2$. If g satisfies the assumptions (1.11) – (1.19) in the case $0\leqslant s\leqslant d-2$ or (1.20) – (1.28) in the case $d-2<s<d$, then we call g an *admissible potential*.

The main result of this article is the following functional inequality for the modulated energy (cf. [Ser20b, Theorem 1], [BJW19a, Theorem 3.2]).

[2] Sufficient and necessary conditions for this Calderón-Zygmund property are explained in [Gra14a, Section 5.4]. The reader may check that this condition is satisfied if g is the genuine Coulomb potential and more generally if g satisfies $|\nabla^{\otimes k}\hat{\mathrm{g}}(\xi)|\leqslant\frac{C}{|\xi|^{d+k-s}}$ $\forall\xi\in\mathbb{R}^d\setminus\{0\}$.

Theorem 1.1. *Let* g *be an admissible potential. Assume the equation* (1.5) *admits a solution* $\mu \in L^\infty([0,T],\mathscr{P}(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d))$*, for some* $T>0$*, such that*

$$\begin{aligned}(1.29)\quad &\||\nabla|^{s-d}\nabla\mu\|_{L^\infty([0,T],L^\infty)}\mathbf{1}_{s\geqslant d-1}+\|\nabla^{\otimes 2}\mathrm{g}*\mu\|_{L^\infty([0,T],L^\infty)}\mathbf{1}_{s=d-2}\\ &+\left(\|\nabla^{\otimes 2}\mathrm{g}*\mu\|_{L^\infty([0,T],C^{0,\alpha})}+\|\nabla\mathrm{g}*\mu\|_{L^\infty([0,T],W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}})}\right)\mathbf{1}_{s>d-2}<\infty.\end{aligned}$$

If $s=0$*, then also assume that* $\int_{\mathbb{R}^d}\log(1+|x|)d\mu^t(x)<\infty$ *for all* $t\in[0,T]$*. Here,* $|\nabla|=(-\Delta)^{1/2}$*,* m *is the dimension of the extension coordinate space* $\mathbb{R}^m$ *for* g*, and* $0<\alpha\leqslant 1$*.*

Let $\underline{x}_N$ *solve* (1.1)*. Then there exist positive constants* C_3, C_4*, depending only on* $|\mathbb{M}|$*,* $\|\mu^0\|_{L^\infty}$ *and the norms of* μ *controlled by* (1.29) *and on the potential only through the constants in assumptions* (1.13)–(1.28)*, and an exponent* $\beta>0$*, depending only on* d,s*, such that*

$$(1.30)\qquad F_N(\underline{x}_N^t,\mu^t)\leqslant\left(F_N(\underline{x}_N^0,\mu^0)+C_3N^{-\beta}\right)e^{C_4t}\qquad\forall\, t\in[0,T].$$

In particular, using the notation (1.4)*, if* $\mu_N^0\rightharpoonup\mu^0$ *in the weak-* topology for measures and*

$$(1.31)\qquad \lim_{N\to\infty}F_N(\underline{x}_N^0,\mu^0)=0,$$

then

$$(1.32)\qquad \mu_N^t\rightharpoonup\mu^t\qquad\forall\, t\in[0,T].$$

Remark 1.2. *The reader may infer from Section 5 the precise dependence of the constants* C_3, C_4 *above on the norms of the solution* μ *as well as the size of the exponent* β*. We have omitted this precise dependence from the statement of Theorem 1.1 so as to present the result in the most accessible terms.*

We have not attempted to optimize the regularity/integrability assumptions on the solution μ*. But we expect that one can proceed in the direction of the works* [Ros22b, Ros22a] *to only require that* μ *belongs to a function space which is critical or borderline with respect to the scaling of the equation. For more discussion on the existence of (necessarily unique) solutions satisfying the condition* (1.29)*, we refer the reader to* [Ser20b, Subsection 1.3].

Remark 1.3. *In the case* $s=0$*, the condition* $\int_{\mathbb{R}^d}\log(1+|x|)d\mu(x)<\infty$ *is to ensure that the convolution* $\mathrm{g}*\mu$ *is a well-defined function. This assumption is purely qualitative, in the sense that none of our estimates will depend on it. One may check through a Gronwall argument that if the initial datum* μ^0 *satisfies this condition, then* μ^t *also satisfies this condition for all* $t\in[0,T]$*.*

Remark 1.4. *Sufficient conditions for the initial modulated energy* $F_N(\underline{x}_N^0,\mu^0)$ *to tend vanish as* $N\to\infty$ *are explained in* [Due16, Remark 1.2(c)]*. In particular, convergence of the energy for* (1.1) *to the energy of* (1.5) *together with weak-* convergence of the initial empirical measure to the initial datum* μ^0 *suffice.*

Remark 1.5. *The weak-* convergence of* μ_N^t *to* μ^t *is quantitative, as shown in Theorem 2.4. In fact, one can also express the convergence in terms of a suitably negative-order Sobolev norm using Theorem 2.4.*

Remark 1.6. *As is well-known, Theorem 1.1 implies propagation of chaos for the marginals of the system* (1.1) *with initial data* $\underline{x}_N^0$ *randomly chosen according to the law* $(\mu^0)^{\otimes N}$. *See* [Ser20b, Remark 3.7] *for details.*

1.3. **Allowable potentials and further extensions.** Assumptions (1.20) - (1.28) allow for a large class of interaction potentials beyond the standard Riesz or log cases. Indeed, the simplest, nontrivial extension is the Lenard-Jones-type potential

$$\mathrm{g}(x) = |x|^{-s} + \varphi_1(x)|x|^{-s_1} + \cdots + \varphi_n(x)|x|^{-s_n}, \tag{1.33}$$

where $0 < s < d$, $0 \leqslant s_n < \cdots < s_1 < s$, and $\varphi_1, \ldots, \varphi_n$ are smooth, compactly supported functions–not necessarily nonnegative. It is straightforward to check that assumptions (1.20) - (1.28) are satisfied.

In the gradient-flow case, we can even consider genuine Lenard-Jones potentials, where $\varphi_1, \ldots, \varphi_n$ are identically equal to real-valued constants $a_1, \ldots, a_n$. Indeed, let χ be a smooth, radial bump function adapted to the region $|\xi| \geqslant 2$. Introducing a small parameter $\kappa > 0$, we split

$$\mathrm{g} = \mathrm{g}_{good} + \mathrm{g}_{bad}, \tag{1.34}$$

where

$$\begin{aligned} \hat{\mathrm{g}}_{bad}(\xi) &:= c|\xi|^{s-d} + \chi(\kappa\xi)\left(\sum_{j=1}^{n} \tilde{c}_j |\xi|^{s_j - d}\right), \\ \hat{\mathrm{g}}_{good}(\xi) &:= \big(1 - \chi(\kappa\xi)\big)\left(\sum_{j=1}^{n} \tilde{c}_j |\xi|^{s_j - d}\right). \end{aligned} \tag{1.35}$$

Here, the constants $c, \tilde{c}_1, \ldots, \tilde{c}_n$ are normalization constants from taking the Fourier transform. If $d - 2 < s < d$, then we set $\mathsf{G}_{bad}(X) := \mathrm{g}_{bad}(|X|)$ and $\mathsf{G}_{good}(X) := \mathrm{g}_{good}(|X|)$ for $X \in \mathbb{R}^{d+m}$. Since the second term in the definition of $\hat{\mathrm{g}}_{bad}$ is smooth, its inverse Fourier transform and its derivatives are rapidly decaying. It is straightforward to check that

$$\mathrm{g}_{bad}(x) \geqslant \frac{1}{|x|^s} - C\sum_{j=1}^{n} \frac{1}{|x|^{s_j}} - C, \qquad \forall\, x \neq 0, \tag{1.36}$$

where $C > 0$ is some constant. In particular, there exists $0 < r_0 \ll 1$, such that for all $|x| \leqslant r_0$,

$$\mathrm{g}_{bad}(x) \geqslant \frac{1}{2|x|^s}. \tag{1.37}$$

Furthermore,

$$-\Delta \mathsf{G}_{bad}(X) \geqslant s(d + m - s - 2)|X|^{-2-s} - C\sum_{j=1}^{n} |X|^{-s_j - 2} - C > 0, \tag{1.38}$$

for $|X| \leqslant r_0$ (possibly choosing r_0 smaller) and $d + m - s - 2 > 0$. Lastly, it follows from the radial symmetry of G_{bad} that

$$C_1|\Xi|^{-d-m+s} \leqslant |\hat{\mathsf{G}}_{bad}(\Xi)| \leqslant C_2|\Xi|^{-d-m+s}, \qquad \forall\, \Xi \neq 0. \tag{1.39}$$

Since $\hat{\mathrm{g}}_{good}$ has compact support and is locally integrable, it belongs to C^∞. Moreover, direct computation shows that $\nabla \mathrm{g}_{good} \in \dot{H}^{\frac{d-s}{2}}$ if $2(s_n+1) > s$. Thus, we can rewrite the system (1.1) as

$$\begin{cases} \dot{x}_i = -\dfrac{1}{N} \displaystyle\sum_{\substack{1 \le j \le N \\ j \neq i}} \left(\nabla \mathrm{g}_{bad}(x_i - x_j) + \mathsf{F}(x_i - x_j) \right) \\ x_i(0) = x_i^0, \end{cases} \tag{1.40}$$

where $\mathsf{F} := \nabla \mathrm{g}_{good}$ is now a smooth added force. The potential g_{bad} satisfies assumptions (1.20) – (1.28), and the contribution of F can be handled using [Ser20b, Lemma 2.3]. One may modify the preceding argument to allow for Lenard-Jones potentials with log interactions as well.

Given that Bresch *et al.* [BJW19b, BJW19a, BJW20] also consider a general class of singular interaction potentials, including those of Riesz type, let us compare and contrast our Theorem 1.1 to their main result. Bresch *et al.* consider only the gradient-flow case of the system (1.1) in the periodic setting with random initial data. So rather than working with particle trajectories, they work with the associated Liouville equation. They also allow for additive noise in the dynamics, but we shall ignore this aspect for the purposes of this discussion. Through a Gronwall-type argument, they prove a functional inequality for the modulated free energy. As previously remarked, this quantity combines the relative entropy of earlier work [JW16, JW18] with the modulated energy of [Due16, Ser20b]. By standard arguments, their functional inequality then implies propagation of chaos. Instead of the smearing and extension procedure used in our article, Bresch *et al.* introduce a clever construction of a regularized potential [BJW19a, Lemma 4.1], which tames the singularity of the Dirac masses. To replace the stress-energy tensor argument from [Ser20b] based on the Caffarelli-Silvestre extension, they substitute the regularized potential g_{reg} into the right-hand side of inequality (1.7), use some elementary Fourier analysis [BJW19a, Lemma 5.2] to bound the resulting expression, and finally estimate the error from this substitution. Although it is never mentioned in their work, their Fourier-analytic argument is, in fact, an elementary commutator estimate, a perspective that we develop in detail in this paper.

As in our work, the potentials they consider are even, singular at the origin, and have nonnegative Fourier transforms. Given that the torus is compact, they only need to assume that the potential is in $L^p(\mathbb{T}^d)$, for some $p > 1$. Similar to assumptions (1.14), (1.24) and (1.15), (1.25), Bresch *et al.* need pointwise control on the potential and its derivatives to sufficiently high order. They also need a doubling condition analogous to (1.17), (1.19), (1.27). A key difference, though, is that no pointwise control on $\hat{\mathrm{g}}(\xi)$ is needed, aside from nonnegativity. Bresch *et al.* only need some control on the the first derivative of $\hat{\mathrm{g}}$. Also, superharmonicity plays no role in their work, in sharp contrast to our work

It is worth mentioning that the relative entropy portion of the functional used by Bresch *et al.* is unnecessary and in fact one can use their regularized potential construction with their elementary commutator estimate in order to treat both the conservative and gradient flow cases of the particle system (1.1). Moreover, one can work directly with the particle

dynamics, as opposed to with the Liouville equation. In some sense, one of the new contributions of our paper is the recognition that this is possible, albeit using different methods. In forthcoming work by the second and third authors [RS21], we will also show that this is possible when one includes so-called additive noise in the dynamics, distinct from the multiplicative noise considered in this article (see Section 6.1 for a comparison); and, in fact, one can obtain bounds on the modulated energy which are *global* in time.

Compared to the work of Bresch *et al.*, the advantages of our approach are that it is not limited to the periodic setting; it is pathwise in the sense that no randomization of the initial data is needed, avoiding the use of modulated free energy in favor of the modulated energy; and, perhaps most importantly, the rate of convergence is explicit. In [BJW19a], for instance, an explicit rate is never presented. If one does attempt to optimize the rate in their work, then one sees that the use of the regularized potential construction [BJW19a, Lemma 4.1] is a bottleneck, and the rate one does obtain is suboptimal compared to our work. Additionally, the regularity requirements for the solution μ to (1.5) of [BJW19a] are more severe than our work. The cost of these advantages is the need for stronger pointwise control on the potential and its Fourier transform, in particular we have to worry about decay at infinity in physical space and blow-up near the origin in Fourier space.

Finally, as commented in the previous subsection, our method of proof also works when multiplicative noise is added to the right-hand side of the N-body problem (1.1). The expected limiting equation is now a stochastic PDE, so that the modulated energy is now a stochastic process. By combining Itô's lemma for the modulated energy with pathwise analysis in the form of first- and second-order commutator estimates, we can perform a similar Gronwall argument for the expectation of the magnitude of the modulated energy. In fact, one can bound moments of arbitrarily large degree. We defer a precise statement of the model under consideration and the new functional inequality until Section 6.

1.4. **Organization of article.** Before introducing the basic notation of the article, let us briefly comment on the organization of the remaining body of the article. In Section 2, we introduce the smearing procedure and establish properties of the modulated energy following the outline of [Ser20b, Section 3], but for the more general class of potentials satisfying assumptions (1.20) - (1.27). In Section 3, we present our new commutator estimate Theorem 3.1, and in Section 4, we use the smearing procedure from Section 2 to renormalize our commuator estimate, obtaining Theorem 4.1. Then in Section 5, we use the analysis of Section 4, specifically Theorem 4.1, together with a delicate optimization of the floating small parameters in order to close the Gronwall argument for the modulated energy. This then completes the proof of Theorem 1.1. Lastly, in Section 6, we close the paper by showing how to extend our modulated-energy approach to variants of the system (1.1) with multiplicative noise added to the dynamics using second-order commutator estimates.

1.5. **Notation.** We close the introduction by specifying the basic notation used in the body of 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 $A \leqslant CB$. 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 $A \sim_p B$.

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 $x_{1,N}, \dots, x_{N,N}$ in some set X, we will write $\underline{x}_N$ to denote the N-tuple $(x_{1,N}, \dots, x_{N,N})$. Given $x \in \mathbb{R}^d$ 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 $\operatorname{supp} f$.

We make frequent use of tensor notation. $A \otimes B$ denotes the usual tensor product between $A = (A^{j_1 \cdots j_n}_{i_1 \cdots i_m})$ and $B = (B^{j'_1 \cdots j'_{n'}}_{i'_1 \cdots i'_{m'}})$ with components

$$(A \otimes B)^{j_1 \cdots j_n j'_1 \cdots j'_{n'}}_{i_1 \cdots i_m i'_1 \cdots i'_{m'}} = A^{j_1 \cdots j_n}_{i_1 \cdots i_m} B^{j'_1 \cdots j'_{n'}}_{i'_1 \cdots i'_{m'}}. \tag{1.41}$$

We identify 2-tensors as matrices and use the notation : to denote the Frobenius inner product. Given a function f, $\nabla^{\otimes k} f$ denotes the $(\mathbb{R}^d)^{\otimes k}$-valued field whose components are given by the partial derivatives $\partial_{i_1 \cdots i_k} f$, for any $1 \leqslant i_1, \dots, i_k \leqslant d$.

We denote the space of probability measures on $\mathbb{R}^d$ by $\mathscr{P}(\mathbb{R}^d)$. When μ is in fact absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$, we shall abuse notation by writing μ for both the measure and its density function. We denote the Banach space of complex-valued continuous functions on $\mathbb{R}^d$ by $C(\mathbb{R}^d)$ 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 $C^k(\mathbb{R}^d)$ equipped with the natural norm, and we define $C^\infty := \bigcap_{k=1}^\infty C^k$. Functions f which belong to $C^{k-1}(\mathbb{R}^d)$ with $\nabla^{\otimes k} f$ α-Hölder continuous are denoted by $C^{k,\alpha}(\mathbb{R}^d)$, which we equip with the usual inhomogeneous norm. We use the subscript c to denote the subspace with compact support. Also, we denote the Schwartz space by $\mathscr{S}(\mathbb{R}^d)$ and the Bessel potential space by $W^{s,p}(\mathbb{R}^d)$.

1.6. **Acknowledgments.** Q.H.N. is supported by the Shanghai Tech University startup fund, the National Natural Science Foundation of China (12050410257), and NSFC Fund's staff support. M.R. is supported by the Simons Foundation through the Simons Collaboration on Wave Turbulence. S.S. is supported by NSF grant DMS-2000205 and by the Simons Foundation through the Simons Investigator program.

2. The modulated energy

In this section, we perform a functional-analytic study of the modulated energy F_N for an arbitrary pairwise distinct configuration $\underline{x}_N \in (\mathbb{R}^d)^N$ and probability density $\mu \in L^\infty$ which has a logarithmic-type decay at infinity in the case $s = 0$. This last assumption is purely qualitative (our estimates will not depend on it): it serves to ensure that the convolution $g * \mu$ is well-defined.

We start by considering the case where (1.13) holds, which is satisfied by all the sub-Coulombic Riesz cases $0 < s < d-2$ as well as the log case $s = 0$ if $d \geqslant 2$.

2.1. **Smearing.** Since (1.13) holds, g is superharmonic in $B(0,r_0)$. For any sufficiently integrable function f, we have

$$\frac{d}{dr}\fint_{\partial B(x,r)} f\,d\mathscr{H}^{d-1} = \frac{1}{d|B(0,1)|r^{d-1}}\int_{B(x,r)} \Delta f\,dy, \tag{2.1}$$

where $\mathscr{H}^{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure on $\mathbb{R}^d$. This immediately implies that g satisfies a mean-value inequality:

$$\forall B(x,r)\subset B(0,r_0), \qquad \mathrm{g}(x)\geqslant \fint_{\partial B(x,r)} \mathrm{g}\,dy. \tag{2.2}$$

We note that this inequality holds despite the singularity of g at the origin. Indeed, $\Delta\mathrm{g}$ is locally integrable if $0\leqslant s<d-2$ by assumption (1.14). If $s=d-2$, then $\Delta\mathrm{g}$ is locally integrable away from the origin, in which case (2.1) holds. Using an approximation argument, (2.2) also follows.

Next, for any $0<\eta<\min\{\frac{1}{2},\frac{r_0}{2}\}$ and $x\in\mathbb{R}^d$, we let $\delta_x^{(\eta)}$ be the uniform probability measure on $\partial B(x,\eta)$ and set

$$\mathrm{g}_\eta := \mathrm{g}*\delta_0^{(\eta)}. \tag{2.3}$$

It follows directly from (2.2) that

$$\mathrm{g}_\eta(x)\leqslant \mathrm{g}(x) \quad \forall x\in B(0,r_0-\eta)\setminus\{0\} \tag{2.4}$$

and from (2.1) and (1.14) that

$$|\mathrm{g}(x)-\mathrm{g}_\eta(x)|\leqslant \frac{C\eta^2}{|x|^{s+2}} \quad \forall |x|\geqslant 2\eta, \tag{2.5}$$

where the constant $C>0$ depends on r_0.

Observe that in view of (2.4) and (1.14), the self-interaction of the smeared point mass $\delta_{x_0}^{(\eta)}$ satisfies the relation

$$\int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)\,d\delta_{x_0}^{(\eta)}(x)\,d\delta_{x_0}^{(\eta)}(y) = \int_{\mathbb{R}^d} \mathrm{g}_\eta(x)\,d\delta_0^{(\eta)}(x) \leqslant \int_{\mathbb{R}^d} \mathrm{g}(x)\,d\delta_0^{(\eta)}(x) = \mathrm{g}_\eta(0)\leqslant C\eta^{-s}. \tag{2.6}$$

If $\mathrm{g}(x)=|x|^{-s}$ then a direct computation with a rescaling is that

$$\int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)\,d\delta_{x_0}^{(\eta)}(x)\,d\delta_{x_0}^{(\eta)}(y) = \eta^{-s}\int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)\,d\delta_0^{(1)}(x)\,d\delta_0^{(1)}(y) = \eta^{-s}\mathrm{g}_1(1), \tag{2.7}$$

where we abuse notation by letting 1 denote any choice of unit vector. Similarly, if $\mathrm{g}(x)=-\log|x|$, then

$$\begin{aligned}\int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)\,d\delta_{x_0}^{(\eta)}(x)\,d\delta_{x_0}^{(\eta)}(y) &= -\log\eta + \int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)\,d\delta_0^{(1)}(x)\,d\delta_0^{(1)}(y)\\ &= -\log\eta + \mathrm{g}_1(1).\end{aligned} \tag{2.8}$$

2.2. **Monotonicity, coerciveness and local interaction control.** By definition (1.6) of $F_N(\underline{x}_N,\mu)$ and the relation (2.6), we may write

$$F_N(\underline{x}_N,\mu)=\lim_{\alpha_i\to 0}\left(\int_{(\mathbb{R}^d)^2}\mathrm{g}(x-y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\alpha_i)}-\mu\right)^{\otimes 2}(x,y)-\frac{1}{N^2}\sum_{i=1}^N\int_{\mathbb{R}^d}\mathrm{g}_{\alpha_i}d\delta_0^{\alpha_i}\right). \tag{2.9}$$

This is a way to express F_N as a renormalized version of the undefined expression

$$\int \mathrm{g}(x-y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu\right)^{\otimes 2}, \tag{2.10}$$

where the renormalization occurs via the smearing of the Dirac masses.

One may then examine the change in the expression in the limit when increasing α_i to fixed smearing radii η_i. As first noted in [PS17, LSZ17], the expression is almost decreasing in the radii, with errors that can be explicitly controlled. In addition, thanks to the Fourier transform bounds (1.26), once the diagonal has been reinserted in this way we have

$$\left\|\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\eta_i)}-\mu\right\|^2_{\dot H^{\frac{s-d}{2}}(\mathbb{R}^d)}\sim\int_{(\mathbb{R}^d)^2}\mathrm{g}(x-y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\eta_i)}-\mu\right)^{\otimes 2}(x,y), \tag{2.11}$$

thus allowing us to control a true Sobolev distance between $\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\eta_i)}$ and μ.

The next proposition expresses this crucial monotonicity property (cf. [Ser20b, Proposition 3.3]) and shows that the modulated energy is bounded from below and coercive, and controls the small scale interactions. It covers potentials with singularity up to and including Coulomb.

Proposition 2.1. *Let $d\geqslant 2$ and $0\leqslant s\leqslant d-2$. Suppose that $\underline{x}_N\in(\mathbb{R}^d)^N$ is pairwise distinct and $\mu\in\mathscr{P}(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. In the case $s=0$, also suppose that $\int_{\mathbb{R}^d}\log(1+|x|)d\mu(x)<\infty$. There exists a constant C depending only on the potential g through assumptions* (1.13) - (1.19)*, such that for every set of truncation parameters $0<\eta_1,\dots,\eta_N<r_0/2$, we have*

$$\begin{aligned}\frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|\leqslant\frac12 r_0}}\big(\mathrm{g}(x_j-x_i)-\mathrm{g}_{\eta_i}(x_j-x_i)\big)_+ + C^{-1}\left\|\frac{1}{N}\sum_{i=1}^N\delta_{x_i}^{(\eta_i)}-\mu\right\|^2_{\dot H^{\frac{s-d}{2}}}&\\ \leqslant F_N(\underline{x}_N,\mu)+\frac{1}{N^2}\sum_{i=1}^N\mathrm{g}_{\eta_i}(0)+\frac{C}{N}\sum_{i=1}^N\eta_i^2&\\ +\frac{C}{N}\sum_{i=1}^N\|\mu\|_{L^\infty}\Big(\eta_i^{d-s}+(\eta_i^d|\log\eta_i|)\mathbf{1}_{s=0}+(\eta_i^2|\log\eta_i|)\mathbf{1}_{s=d-2}\Big),&\end{aligned} \tag{2.12}$$

where r_0 is the constant in (1.13).

Proof. Adding and subtracting $\delta_{x_i}^{(\eta_i)}$ and regrouping terms, we find that

$$\begin{aligned} F_N(\underline{x}_N,\mu) = \int_{(\mathbb{R}^d)^2} \mathrm{g}(x-y)d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - \mu\right)^{\otimes 2}(x,y) - \frac{1}{N^2}\sum_{i=1}^N \int_{\mathbb{R}^d} \mathrm{g}_{\eta_i} d\delta_0^{(\eta_i)} \\ - \frac{2}{N}\sum_{i=1}^N \int_{\mathbb{R}^d} \big(\mathrm{g}(y-x_i) - \mathrm{g}_{\eta_i}(y-x_i)\big)\, d\mu(y) \\ + \frac{1}{N^2}\sum_{1\leqslant i\neq j\leqslant N} \int_{\mathbb{R}^d} \big(\mathrm{g}(y-x_i) - \mathrm{g}_{\eta_i}(y-x_i)\big)\, d(\delta_{x_j}+\delta_{x_j}^{(\eta)})(y). \end{aligned} \tag{2.13}$$

From inequalities (2.4) and (2.5), we see that

$$\begin{aligned} \frac{1}{N^2}\sum_{1\leqslant i\neq j\leqslant N} \int_{\mathbb{R}^d} \big(\mathrm{g}(y-x_i) - \mathrm{g}_{\eta_i}(y-x_i)\big)\, d(\delta_{x_j}+\delta_{x_j}^{(\eta_j)})(y) \\ > \frac{1}{N^2} \sum_{\substack{1\leqslant i\neq j\leqslant N \\ |x_i-x_j|<r_0-\eta}} \big(\mathrm{g}(x_j-x_i) - \mathrm{g}_{\eta_i}(x_j-x_i)\big)_+ - \frac{C}{N}\sum_{i=1}^N \eta_i^2. \end{aligned} \tag{2.14}$$

Next, using (2.4) and (2.5) again, we find that

$$\begin{aligned} \int_{B(0,r_0-\eta)} |\mathrm{g}(x) - \mathrm{g}_\eta(x)|\,dx &\leqslant 2\int_{B(0,2\eta)} \mathrm{g}(x)\,dx + C\eta^2 \int_{B(0,r_0)\setminus B(0,2\eta)} \frac{dx}{|x|^{s+2}} \\ &\leqslant C\Big(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0} + (\eta^2|\log\eta|)\mathbf{1}_{s=d-2}\Big). \end{aligned} \tag{2.15}$$

Since μ is a probability density, we deduce from Hölder's inequality that

$$\begin{aligned} -\frac{2}{N}\sum_{i=1}^N \int_{\mathbb{R}^d} \big(\mathrm{g}(y-x_i) - \mathrm{g}_{\eta_i}(y-x_i)\big) d\mu(y) \geqslant -\frac{C}{N}\sum_{i=1}^N \eta_i^2 \\ - \frac{C}{N}\|\mu\|_{L^\infty} \sum_{i=1}^N \Big(\eta_i^{d-s} + (\eta_i^d|\log\eta_i|)\mathbf{1}_{s=0} + (\eta_i^2|\log\eta_i|)\mathbf{1}_{s=d-2}\Big). \end{aligned} \tag{2.16}$$

Inserting (2.14) and (2.16) into (2.13) and using (2.11), the result follows. □

Since $\mathrm{g}_\eta(0) \leqslant C(\eta^{-s} + |\log\eta|\mathbf{1}_{s=0})$, the right-hand side of (2.12) is easily optimized by choosing $\eta_i = N^{-\frac{1}{2+s}} \leqslant N^{-1/d}$. Despite not having a sign, F_N is bounded below:

$$F_N(\underline{x}_N,\mu) \geqslant -C(1+\|\mu\|_{L^\infty})\Big(N^{-\frac{2}{2+s}} + (N^{-1}\log N)\mathbf{1}_{s=0} + (N^{-\frac{2}{d}}\log N)\mathbf{1}_{s=d-2}\Big), \tag{2.17}$$

which, by triangle inequality, implies that

$$\begin{aligned} |F_N(\underline{x}_N,\mu)| \leqslant F_N(\underline{x}_N,\mu) + 2C(1+\|\mu\|_{L^\infty})\Big(N^{-\frac{2}{2+s}} + (N^{-1}\log N)\mathbf{1}_{s=0} \\ + (N^{-\frac{2}{d}}\log N)\mathbf{1}_{s=d-2}\Big). \end{aligned} \tag{2.18}$$

2.3. **Renormalization in extended space.** Similar to before, we define $\delta_X^{(\eta)}$ to be the uniform probability measure on $\partial B(X,\eta) \subset \mathbb{R}^{d+m}$ and define the mollified potential

$$\mathsf{G}_\eta := \mathsf{G} * \delta_0^{(\eta)}. \tag{2.19}$$

With the observation that G is superharmonic in $B(0,r_0)\subset\mathbb{R}^{d+m}$, we have the analogue of (2.4) and (2.5):

$$\mathsf{G}_\eta(X)\leqslant \mathsf{G}(X)\qquad \forall X\in B(0,r_0-\eta)\setminus\{0\} \tag{2.20}$$

and, from (2.1) and (1.14),

$$|\mathsf{G}(X)-\mathsf{G}_\eta(X)|\leqslant \frac{C\eta^2}{|X|^{s+2}}\qquad \forall |X|\geqslant 2\eta, \tag{2.21}$$

where $C>0$ depends on r_0. From (2.20), we obtain the analogue of (2.6):

$$\int_{(\mathbb{R}^{d+m})^2}\mathsf{G}(x-y)d\delta_0^{(\eta)}(x)d\delta_0^{(\eta)}(y)\leqslant \mathsf{G}_\eta(0)\leqslant C\big(\eta^{-s}+|\log\eta|\mathbf{1}_{d=1\wedge s=0}\big). \tag{2.22}$$

We may now repeat all of the analysis of Section 2.2 in the extended space $\mathbb{R}^{d+m}$. The modulated energy is still defined by (1.6). The only change is in the mollification procedure which happens in extended space. Analogous to the lower bound of Theorem 2.1, we have a lower bound for the modulated energy in the super-Coulombic case via this mollification procedure.

Proposition 2.2. *Let $d\geqslant 1$ and $\max\{d-2,0\}<s<d$ or $d=1$ and $s=0$. Suppose that $\underline{x}_N\in(\mathbb{R}^d)^N$ is a pairwise distinct configuration and $\mu\in\mathscr{P}(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. If $s=0$, also assume that $\int_{\mathbb{R}^d}\log(1+|x|)d\mu(x)<\infty$. There exists a constant C depending on s,d and on g only through assumptions* (1.23) – (1.27)*, such that for every $\eta_i<\min\{\frac12,\frac12 r_0\}$, we have*

$$\begin{aligned}\frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|\leqslant\frac12 r_0}}\big(\mathrm{g}(x_j-x_i)-\mathsf{G}_{\eta_i}(x_j-x_i,0)\big)_+ + C^{-1}\left\|\frac1N\sum_{i=1}^N\delta_{X_j}^{(\eta_i)}-\tilde\mu\right\|^2_{\dot H^{\frac{s-d-m}{2}}(\mathbb{R}^{d+m})}\\ \leqslant F_N(\underline{x}_N,\mu)+\frac{1}{N^2}\sum_{i=1}^N\mathsf{G}_{\eta_i}(0)+\frac{C}{N}\sum_{i=1}^N\Big(\|\mu\|_{L^\infty}\Big(\eta_i^{d-s}+(\eta_i^d|\log\eta_i|)\mathbf{1}_{s=0}\Big)+\eta_i^2\Big),\end{aligned} \tag{2.23}$$

where r_0 is the constant in (1.23)*, $X_i:=(x_i,0)$, and*

$$\tilde\mu(X):=\mu(x)\delta_{\mathbb{R}^d\times\{0\}}(X). \tag{2.24}$$

Proof. Following the proof of Theorem 2.1, we write

$$\begin{aligned}F_N(\underline{x}_N,\mu)=&\int_{(\mathbb{R}^{d+m})^2}\mathsf{G}(X-Y)d\left(\frac1N\sum_{i=1}^N\delta_{X_i}^{(\eta_i)}-\tilde\mu\right)^{\otimes2}(x,y)\\ &-\frac{1}{N^2}\sum_{i=1}^N\int_{(\mathbb{R}^{d+m})^2}\mathsf{G}(X-Y)d(\delta_0^{(\eta_i)})^{\otimes2}(X,Y)\\ &-\frac2N\sum_{i=1}^N\int_{\mathbb{R}^{d+m}}\big(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i)\big)d\tilde\mu(Y)\\ &+\frac{1}{N^2}\sum_{1\leqslant i\neq j\leqslant N}\int_{\mathbb{R}^d}\big(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i)\big)d(\delta_{X_j}+\delta_{X_j}^{(\eta_j)})(Y).\end{aligned} \tag{2.25}$$

The third line reduces to an integral over $\mathbb{R}^d$, identical to those in (2.16). Analogously to (2.15), in view of (2.20), (1.24) and (2.21) we have

$$\begin{aligned}\int_{B(0,r_0-\eta)\subset\mathbb{R}^d} |\mathsf{g}(x)-\mathsf{G}_\eta(x,0)|dx &\leqslant 2\int_{B(0,2\eta)} \mathsf{g}(x) + C\eta^2 \int_{B(0,r_0)\setminus B(0,2\eta)} \frac{dx}{|x|^{2+s}} \\ &\leqslant C\left(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0}\right). \end{aligned} \tag{2.26}$$

The third line in (2.25) is thus bounded below by

$$-\frac{C}{N}\sum_{i=1}^N \left(\eta_i^2 + \|\mu\|_{L^\infty}\left(\eta_i^{d-s} + (\eta_i^d|\log\eta_i|)\mathbf{1}_{s=0}\right)\right) \tag{2.27}$$

as in (2.16). For the fourth line, we use (2.21) and that $\delta_{X_j}^{(\eta_j)}$ is a positive measure on $\mathbb{R}^{d+m}$ to obtain

$$\begin{aligned}&\frac{1}{N^2}\sum_{1\leqslant i\neq j\leqslant N}\int_{\mathbb{R}^{d+m}} \left(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i)\right)d(\delta_{X_j}+\delta_{X_j}^{(\eta_j)})(Y) \\ &\geqslant \frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|\leqslant r_0-\eta}}\int_{\mathbb{R}^{d+m}} \left(\mathsf{G}(Y-X_i)-\mathsf{G}_{\eta_i}(Y-X_i)\right)d\delta_{X_j}(Y) - \frac{C}{N}\sum_{i=1}^N \eta_i^2 \\ &\geqslant \frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|\leqslant r_0-\eta}} \left(\mathsf{g}(x_j-x_i)-\mathsf{G}_{\eta_j}(X_j-X_i)\right)_+ - \frac{C}{N}\sum_{i=1}^N \eta_i^2. \end{aligned} \tag{2.28}$$

Finally, noting that we have by virtue of (1.26) that

$$\int_{(\mathbb{R}^{d+m})^2} \mathsf{G}(X-Y)d\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta_i)}-\tilde{\mu}\right)^{\otimes 2}(X,Y) \geqslant C_1\left\|\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta_i)}-\tilde{\mu}\right\|^2_{\dot{H}^{\frac{s-d-m}{2}}(\mathbb{R}^{d+m})}, \tag{2.29}$$

we arrive at the inequality in the statement of the proposition. □

2.4. **Control of microscale interactions.** As recognized for the first time by the third author in [Ser20b], an interesting choice of parameters η_i is to take each η_i equal to a fixed fraction of the distance from x_i to its nearest neighbor(s). This ensures that the balls $B(x_i,\eta_i)$ are pairwise disjoint, while also keeping the η_i large enough that the right-hand side error in (2.12) and (2.23) is kept small. However, in this paper, we are able to dispense with the requirement that the balls be disjoint; so, we can choose the η_i to all be equal to some parameter η or ε, to be optimized at the end.

Inequalities (2.12) and (2.23) allow us to control the small-scale interactions. Indeed, applying them to $\eta_i = 3\varepsilon$ we arrive at the following corollary.

Corollary 2.3. *Let $0 \leq s < d$. For any pairwise distinct $\underline{x}_N \in (\mathbb{R}^d)^N$ and $\mu \in \mathscr{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ (with $\int_{\mathbb{R}^d} \log(1+|x|)\, d\mu(x) < \infty$ if $s = 0$), we have for all $0 < \varepsilon < \min\{\frac{1}{8}, \frac{r_0}{8}\}$,*

$$\frac{1}{N^2} \sum_{\substack{1 \leq i \neq j \leq N \\ |x_i - x_j| \leq \varepsilon}} \big(\mathrm{g}(x_i - x_j)\mathbf{1}_{s>0} + \mathbf{1}_{s=0}\big) \leq C\bigg(F_N(\underline{x}_N, \mu) + C\Big(\frac{\varepsilon^{-s} + |\log\varepsilon|\mathbf{1}_{s=0}}{N} + \varepsilon^2 + \|\mu\|_{L^\infty}(\varepsilon^{d-s} + (\varepsilon^d|\log\varepsilon|)\mathbf{1}_{s=0} + (\varepsilon^2|\log\varepsilon|)\mathbf{1}_{s=d-2})\Big)\bigg), \tag{2.30}$$

where C depends on s, d and on g only through the assumptions (1.13) – (1.27).

Proof. We use assumption (1.27), which ensures that for $|x_i - x_j| \leq \varepsilon$, we have

$$\mathrm{g}(x_i - x_j) - \mathrm{g}_{3\varepsilon}(x_i - x_j) \geqslant (1 - c_s)\mathrm{g}(x_i - x_j)\mathbf{1}_{s>0} + c_0\mathbf{1}_{s=0}. \tag{2.31}$$

□

2.5. **Coerciveness.** From Theorem 2.1 or Theorem 2.2, we easily deduce that the modulated energy F_N controls $\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu$ in the weak-* topology (cf. [Ser20b, Proposition 3.6]).

Proposition 2.4. *Let $\phi \in \mathscr{S}(\mathbb{R}^d)$ be a test function. For any $\eta > 0$, any $\underline{x}_N$ and μ as in Theorem 2.1 or Theorem 2.2, we have*

$$\left|\int_{\mathbb{R}^d} \phi\, d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)\right| \leqslant \|\phi\|_{C^{0,\alpha}}\eta^\alpha + C\|\phi\|_{\dot{H}^{\frac{d-s}{2}}}\bigg(F_N(\underline{x}_N, \mu) + C\frac{\eta^{-s} + |\log\eta|\mathbf{1}_{s=0}}{N} + C\|\mu\|_{L^\infty}(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0} + (\eta^2|\log\eta|)\mathbf{1}_{s=d-2}) + \eta^2\bigg)^{1/2}, \tag{2.32}$$

where C depends on s, d, m and the potential g through assumptions (1.13) - (1.27).

Proof. Fix $\phi \in \mathscr{S}(\mathbb{R}^d)$, and let $\tilde{\phi} \in \mathscr{S}(\mathbb{R}^{d+m})$ be an extension of ϕ such that

$$\begin{aligned} \|\tilde{\phi}\|_{\dot{H}^{\frac{d+m-s}{2}}(\mathbb{R}^{d+m})} &\lesssim \|\phi\|_{\dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d)}, \\ \|\tilde{\phi}\|_{C^{0,\alpha}(\mathbb{R}^{d+m})} &\lesssim \|\phi\|_{C^{0,\alpha}(\mathbb{R}^d)}. \end{aligned} \tag{2.33}$$

For instance, take $\tilde{\phi}(x,z) := (e^{|z|\Delta_{\mathbb{R}^d}}\phi)(x)$. We may now write

$$\begin{aligned} \left|\int_{\mathbb{R}^d} \phi d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)\right| &\leqslant \left|\int_{\mathbb{R}^{d+m}} \tilde{\phi} d\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta)} - \tilde{\mu}\right)\right| + \left|\int_{\mathbb{R}^{d+m}} \tilde{\phi}\frac{1}{N}\sum_{i=1}^N d(\delta_{X_i} - \delta_{X_i}^{(\eta)})\right| \\ &\leqslant \|\tilde{\phi}\|_{\dot{H}^{\frac{m+d-s}{2}}(\mathbb{R}^{d+m})}\left\|\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta)} - \tilde{\mu}\right\|_{\dot{H}^{\frac{s-d-m}{2}}(\mathbb{R}^{d+m})} + \|\phi\|_{C^{0,\alpha}(\mathbb{R}^d)}\eta^\alpha. \end{aligned} \tag{2.34}$$

We conclude after applying the estimate (2.12) or (2.23). □

By duality, Theorem 2.4 implies that the modulated energy controls inhomogeneous Sobolev norms: for $\gamma > \frac{d+2}{2}$,

$$\left\| \frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu \right\|_{H^{-\gamma}}^2 \leqslant CF_N(\underline{x}_N, \mu) + C(1+\|\mu\|_{L^\infty})\Big(N^{\frac{-2}{2+s}} + (N^{-1}\log N)\mathbf{1}_{s=0} + (N^{-\frac{2}{d}}\log N)\mathbf{1}_{s=d-2}\Big). \tag{2.35}$$

Indeed, Sobolev embedding implies that $H^\gamma(\mathbb{R}^d) \subset C^1(\mathbb{R}^d)$, from which the preceding bound follows. Further refinements of such norm control, such as in terms of Besov spaces, are possible. For more details, we refer to [Ros22b, Proposition 3.10] and [Ros22a, Proposition 3.8].

3. Commutator estimates

In this section, we present commutator estimates without renormalization, which is the most technical component of the paper. This is a purely space-based functional inequality (cf. [BJW19a, Lemma 5.1]) which dispenses with the stress-energy tensor structure of [Due16, Ser20b].

Proposition 3.1. *Let $0 \leqslant s < d$. Let $\mathrm{g} \in C^\infty(\mathbb{R}^d \setminus \{0\})$ such that $\mathrm{g}(x) = \mathrm{g}(-x)$ in $B(0, r_0)$, and*[3]

$$\forall k \geqslant 1, \qquad |\nabla^{\otimes k}\mathrm{g}(x)| \lesssim |x|^{-k-s} \quad \forall x \neq 0. \tag{3.1}$$

If $s \leqslant d-2$, then also assume that

$$|\hat{\mathrm{g}}(\xi)| \lesssim |\xi|^{s-d} \qquad \forall \xi \neq 0. \tag{3.2}$$

If $s = d-2k$, for some $k \in \mathbb{N}$, also assume that g is such that the the kernel

$$\mathrm{k}(x-y) := (x-y)\otimes \nabla^{\otimes(1+2k)}\mathrm{g}(x-y) \qquad \forall x \neq y \tag{3.3}$$

is associated to a Calderón-Zygmund operator. Let v be a Lipschitz continuous vector field from $\mathbb{R}^d$ to $\mathbb{R}^d$. There exists a constant $C > 0$, depending only s, d and the potential g through (3.1) – (3.3)*, such that for any $f, g \in \mathscr{S}(\mathbb{R}^d)$,*[4] *we have*

$$\begin{aligned} &\left| \int_{(\mathbb{R}^d)^2} (v(x)-v(y))\cdot\nabla \mathrm{g}(x-y) f(x) g(y)\, dx\, dy \right| \\ &\leqslant C\Big(\|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2}\Big) \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \end{aligned} \tag{3.4}$$

Consequently, the integral in the left-hand side of (3.4) *extends to a bounded bilinear form $B_v(\cdot,\cdot)$ on $\dot{H}^{\frac{s-d}{2}}(\mathbb{R}^d)$ satisfying the bound* (3.4).

Proof. We divide the proof into the super-Coulombic, Coulombic, and sub-Coulombic cases, beginning with the super-Coulombic. We will not explicitly track the dependence

[3] The hypothesis (3.1) is only necessary for all k up to some integer depending only on d, s, which may be determined from examination of the proof of the proposition.

[4] If $s = 0$, then implicit is the requirement that the Fourier transforms of f, g vanish sufficiently rapidly at the origin so that the $\dot{H}^{-d/2}$ norm converges. This may be ensured by assuming a log-weighted L^1 condition on f, g.

of the implicit constants; but the reader should keep in mind that they depend on the parameters s, d as well as the implicit constants in (3.1), (3.2) as well as the size and smoothness condition constants for the kernel (3.3) and the L^2 operator norm of the associated Calderón-Zygmund operator.

The super-Coulombic case. Define

$$\mathsf{k}_v(x,y) := (v(x)-v(y))\cdot\nabla \mathsf{g}(x-y) \qquad \forall x \neq y. \tag{3.5}$$

By approximation, we may assume that $v \in C^\infty$. Furthermore, by density in $\dot{H}^{\frac{s-d}{2}}$, we may assume without loss of generality that $f, g \in \mathscr{S}(\mathbb{R}^d)$ with Fourier support away from the origin. Additionally, by dominated convergence, we may assume that g has compact support. Write

$$f = \operatorname{div}\nabla\Delta^{-1} f =: \operatorname{div} f_1 \tag{3.6}$$

and similarly for g. Note that $f_1 = (f_1^i), g_1 = (g_1^i)$ are $\mathbb{R}^d$-valued vector fields. Then

$$\int_{(\mathbb{R}^d)^2} \mathsf{k}_v(x,y) f(x) g(y) dxdy = \int_{(\mathbb{R}^d)^2} \mathsf{k}_v(x,y) \operatorname{div} f_1(x) \operatorname{div} g_1(y) dxdy. \tag{3.7}$$

We want to integrate by parts to move the derivatives onto the kernel k_v. But since k_v is singular along the diagonal $x = y$, we need to exercise some care.

To this end, we write the preceding right-hand side as

$$\lim_{\varepsilon\to 0^+} \int_{|x-y|\geqslant\varepsilon} \mathsf{k}_v(x,y) \operatorname{div} f_1(x) \operatorname{div} g_1(y) dxdy. \tag{3.8}$$

Integrating by parts in y, we find that the preceding integral equals

$$\begin{aligned} &\int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) g_1(y) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y) \\ &\qquad - \int_{|x-y|\geqslant\varepsilon} \nabla_y \mathsf{k}_v(x,y) f(x) \cdot g_1(y) dxdy, \end{aligned} \tag{3.9}$$

where $\mathscr{H}^{2d-1}$ denotes the $(2d-1)$-dimensional Hausdorff measure on $\mathbb{R}^{2d}$. If $s < d-1$, then we can use the crude kernel estimate based on (1.24)

$$|\mathsf{k}_v(x,y)| \lesssim \min\left(\frac{\|\nabla v\|_{L^\infty}}{|x-y|^s}, \frac{\|v\|_{L^\infty}}{|x-y|^{s+1}} \right) \tag{3.10}$$

to directly estimate the boundary term by

$$\left| \int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) g_1(y) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y) \right| \lesssim \|\nabla v\|_{L^\infty} \|f\|_{L^1} \|g_1\|_{L^\infty} \varepsilon^{d-1-s}, \tag{3.11}$$

which tends to zero as $\varepsilon \to 0$. Otherwise, we subtract and add $g_1(x)$ to write

$$\begin{aligned}
&\int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) g_1(y) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y) \\
&= \int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) (g_1(y) - g_1(x)) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y) \\
&\quad + \int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) g_1(x) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y).
\end{aligned} \tag{3.12}$$

In the sequel, it will be convenient to sometimes write in coordinates and employ the convention of Einstein summation. Writing $z = x - y$, so that by Taylor's theorem,

$$\mathsf{k}_v(x,y) = -z^{i_1} \partial_{i_1} v(x) \cdot \nabla \mathsf{g}(z) + O(\|\nabla^{\otimes 2} v\|_{L^\infty} |z|^{1-s}), \tag{3.13}$$

we see that

$$\begin{aligned}
\int_{|x-y|=\varepsilon} \mathsf{k}_v(x,y) f(x) g_1(x) \cdot \frac{(x-y)}{|x-y|} d\mathscr{H}^{2d-1}(x,y) &= O(\|\nabla^{\otimes 2} v\|_{L^\infty} \|f g_1\|_{L^1} \varepsilon^{d-s}) \\
&\quad + \int_{\mathbb{R}^d} \int_{\partial B(0,\varepsilon)} (z^{i_1} \partial_{i_1} v(x) \cdot \nabla \mathsf{g}(z)) f(x) \left(g_1(x) \cdot \frac{z}{|z|} \right) d\mathscr{H}^{d-1}(z) dx.
\end{aligned} \tag{3.14}$$

Note that the first term in the right-hand side vanishes as $\varepsilon \to 0$ since $s < d$ and the second term vanishes by assumption (1.21) and the invariance of the measure under $z \mapsto -z$.

For the non-boundary term, we integrate by parts now in x to obtain

$$\begin{aligned}
&-\int_{|x-y|\geqslant\varepsilon} \nabla_y \mathsf{k}_v(x,y) \operatorname{div} f_1(x) \cdot g_1(y) dx dy \\
&\qquad = \int_{|x-y|\geqslant\varepsilon} \nabla_x \nabla_y \mathsf{k}_v(x,y) : (f_1(x) \otimes g_1(y)) dx dy \\
&\qquad\quad + \int_{|x-y|=\varepsilon} \left(\frac{(x-y)}{|x-y|} \otimes \nabla_y \mathsf{k}_v(x,y) \right) : \big(f_1(x) \otimes g_1(y) \big) d\mathscr{H}^{2d-1}(x,y).
\end{aligned} \tag{3.15}$$

We subtract and add $f_1(y)$ to write

$$\begin{aligned}
&\int_{|x-y|=\varepsilon} \left(\frac{(x-y)}{|x-y|} \otimes \nabla_y \mathsf{k}_v(x,y) \right) : \big(f_1(x) \otimes g_1(y) \big) d\mathscr{H}^{2d-1}(x,y) \\
&= \int_{|x-y|=\varepsilon} \left(\frac{(x-y)}{|x-y|} \otimes \nabla_y \mathsf{k}_v(x,y) \right) : \big((f_1(x) - f_1(y)) \otimes g_1(y) \big) d\mathscr{H}^{2d-1}(x,y) \\
&\quad + \int_{|x-y|=\varepsilon} \left(\frac{(x-y)}{|x-y|} \otimes \nabla_y \mathsf{k}_v(x,y) \right) : \big(f_1(y) \otimes g_1(y) \big) d\mathscr{H}^{2d-1}(x,y).
\end{aligned} \tag{3.16}$$

By Taylor's theorem,

$$\begin{aligned}
&\int_{|x-y|=\varepsilon}\left(\frac{(x-y)}{|x-y|}\otimes\nabla_y \mathsf{k}_v(x,y)\right):\big((f_1(x)-f_1(y))\otimes g_1(y)\big)d\mathscr{H}^{2d-1}(x,y)\\
&=\int_{|x-y|=\varepsilon}\left(\frac{(x-y)}{|x-y|}\otimes\nabla_y \mathsf{k}_v(x,y)\right):\Big((x-y)^{i_1}\partial_{i_1}f_1(y)\otimes g_1(y)\Big)d\mathscr{H}^{2d-1}(x,y)\\
&\quad+O(\|v\|_{W^{1,\infty}}\|\nabla^{\otimes 2}f_1\|_{L^\infty}\|g_1\|_{L^1}\varepsilon^{d-s}).
\end{aligned}\tag{3.17}$$

Proceeding similarly as above, we have that

$$\begin{aligned}
&\left|\int_{|x-y|=\varepsilon}\left(\frac{(x-y)}{|x-y|}\otimes\nabla_y \mathsf{k}_v(x,y)\right):\Big((x-y)^{i_1}\partial_{i_1}f_1(y)\otimes g_1(y)\Big)d\mathscr{H}^{2d-1}(x,y)\right|\\
&\lesssim\|v\|_{W^{2,\infty}}\|\nabla f_1 g_1\|_{L^1}\varepsilon^{d-s}.
\end{aligned}\tag{3.18}$$

Using that

$$\nabla_x\nabla_y \mathsf{k}_v(x,y)=\nabla_x\nabla_y \mathsf{k}_v(y,x)\tag{3.19}$$

by assumption (1.21), we can swap x and y to obtain

$$\begin{aligned}
&\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):\big(f_1(x)\otimes g_1(y)\big)dxdy\\
&=-\frac{1}{2}\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):\big((f_1(x)-f_1(y))\otimes(g_1(x)-g_1(y))\big)dxdy\\
&\quad+\frac{1}{2}\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):\big(f_1(x)\otimes g_1(x)\big)dxdy\\
&\quad+\frac{1}{2}\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):\big(f_1(y)\otimes g_1(y)\big)dxdy.
\end{aligned}\tag{3.20}$$

Note that the last two terms are equal by symmetry and by Stokes theorem,

$$\begin{aligned}
&\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):(f_1(y)\otimes g_1(y))dxdy\\
&=-\int_{|x-y|=\varepsilon}\left(\frac{(x-y)}{|x-y|}\otimes\nabla_y \mathsf{k}_v(x,y)\right):\big(f_1(y)\otimes g_1(y)\big)d\mathscr{H}^{2d-1}(x,y).
\end{aligned}\tag{3.21}$$

After a little bookkeeping, we realize we have shown that

$$\begin{aligned}
\int_{|x-y|\geqslant\varepsilon}\mathsf{k}_v(x,y)f(x)g(y)dxdy&=O(\|\nabla^{\otimes 2}v\|_{L^\infty}\|fg_1\|_{L^1}\varepsilon^{d-s})\\
&\quad+O(\|v\|_{W^{2,\infty}}\|\nabla f_1 g_1\|_{L^1}\varepsilon^{d-s})+O(\|v\|_{W^{1,\infty}}\|\nabla^{\otimes 2}f_1\|_{L^\infty}\|g_1\|_{L^1}\varepsilon^{d-s})\\
&\quad-\frac{1}{2}\int_{|x-y|\geqslant\varepsilon}\nabla_x\nabla_y \mathsf{k}_v(x,y):\big((f_1(x)-f_1(y))\otimes(g_1(x)-g_1(y))\big)dxdy.
\end{aligned}\tag{3.22}$$

The first three terms of the right-hand side vanish as $\varepsilon \to 0$ since $s < d$. For the last term, direct computation reveals

$$\nabla_x \nabla_y \mathsf{k}_v(x,y) = -(v(x)-v(y)) \cdot \nabla^{\otimes 3} \mathrm{g}(x-y) - \big(\nabla v(y) + \nabla v(x)\big) : \nabla^{\otimes 2} \mathrm{g}(x-y). \tag{3.23}$$

Therefore, using the triangle inequality and mean-value theorem, we find from assumption (3.1) that

$$|\nabla_x \nabla_y \mathsf{k}_v(x,y)| \lesssim \frac{\|\nabla v\|_{L^\infty}}{|x-y|^{s+2}}. \tag{3.24}$$

This bound together with Cauchy-Schwarz implies that

$$\begin{aligned} &\int_{|x-y|\geqslant \varepsilon} |\nabla_x \nabla_y \mathsf{k}_v(x,y)||f_1(x)-f_1(y)||g_1(x)-g_1(y)|dxdy \\ &\lesssim \|\nabla v\|_{L^\infty} \bigg(\int_{(\mathbb{R}^d)^2} \frac{|f_1(x)-f_1(y)|^2}{|x-y|^{s+2}} dxdy \bigg)^{1/2} \bigg(\int_{(\mathbb{R}^d)^2} \frac{|g_1(x)-g_1(y)|^2}{|x-y|^{s+2}} dxdy \bigg)^{1/2}. \end{aligned} \tag{3.25}$$

Provided that $d-2 < s < d$, the preceding right-hand side is equivalent to (for instance, see [DNPV12, Proposition 3.4])

$$\|\nabla v\|_{L^\infty} \|f_1\|_{\dot{H}^{1+\frac{s-d}{2}}} \|g_1\|_{\dot{H}^{1+\frac{s-d}{2}}} \lesssim \|\nabla v\|_{L^\infty} \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}, \tag{3.26}$$

where the ultimate inequality follows from the definition of f_1, g_1. In conclusion,

$$\lim_{\varepsilon \to 0} \bigg| \int_{|x-y|\geqslant \varepsilon} \mathsf{k}_v(x,y) f(x) g(y) dxdy \bigg| \lesssim \|\nabla v\|_{L^\infty} \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \tag{3.27}$$

The Coulombic case. We can modify the preceding argument to also work in the Coulombic case $s = d-2$. To see this, we observe from above that

$$\begin{aligned} &\int_{(\mathbb{R}^d)^2} \mathsf{k}_v(x,y) f(x) g(y) dxdy \\ &\quad = \lim_{\varepsilon \to 0} \bigg(\int_{|x-y|=\varepsilon} \frac{(x-y)_{i_1}}{|x-y|} \partial_{y_{i_2}} \mathsf{k}_v(x,y) f_1^{i_1}(y) g_1^{i_2}(y) d\mathscr{H}^{2d-1}(x,y) \\ &\qquad + \int_{|x-y|=\varepsilon} \frac{(x-y)_{i_1}}{|x-y|} \partial_{y_{i_2}} \mathsf{k}_v(x,y) (f_1(x)-f_1(y))^{i_1} g_1^{i_2}(y) d\mathscr{H}^{2d-1}(x,y) \\ &\qquad\qquad + \int_{|x-y|\geqslant \varepsilon} \partial_{x_{i_1}} \partial_{y_{i_2}} \mathsf{k}_v(x,y) f_1^{i_1}(x) g_1^{i_2}(y) dxdy \bigg). \end{aligned} \tag{3.28}$$

The second boundary term vanishes as $\varepsilon \to 0$ by (3.17) and (3.18). We rewrite the first boundary term as

$$\int_{\mathbb{R}^d} f_1^{i_1}(y) g_1^{i_2}(y) \bigg(\int_{\partial B(0,\varepsilon)} \bigg(\frac{z_{i_1}}{|z|} \partial_{y_{i_2}} \mathsf{k}_v(y+z,y) \bigg) d\mathscr{H}^{d-1}(z) \bigg) dy. \tag{3.29}$$

It is straightforward to check that as $\varepsilon \to 0$, this expression tends to

$$\begin{aligned} &-\int_{\mathbb{R}^d} f_1^{i_1}(y) g_1^{i_2}(y) \Big(\varepsilon^d \int_{\partial B(0,1)} \frac{z_{i_1}}{|z|} \partial_{i_3} v^{i_4}(y) z^{i_3} \partial_{i_2 i_4} \mathsf{g}(\varepsilon z)) d\mathscr{H}^{d-1}(z) \Big) dy \\ &-\int_{\mathbb{R}^d} f_1^{i_1}(y) g_1^{i_2}(y) \Big(\varepsilon^{d-1} \int_{\partial B(0,1)} \frac{z_{i_1}}{|z|} (\partial_{i_2} v(y) \cdot \nabla \mathsf{g}(\varepsilon z)) d\mathscr{H}^{d-1}(z) \Big) dy. \end{aligned} \tag{3.30}$$

Using Cauchy-Schwarz and assumption (3.1), we can bound this term directly (up to a constant) by

$$\|\nabla v\|_{L^\infty} \|f_1\|_{L^2} \|g_1\|_{L^2} \lesssim \|\nabla v\|_{L^\infty} \|f\|_{\dot{H}^{-1}} \|g\|_{\dot{H}^{-1}}. \tag{3.31}$$

For the non-boundary term, we write out $\nabla_x \nabla_y \mathsf{k}_v(x,y)$ to obtain

$$\begin{aligned} &-\int_{|x-y| \geqslant \varepsilon} \big(\partial_{i_1} \nabla \mathsf{g}(x-y) \cdot \partial_{i_2} v(y) + \partial_{i_2} \nabla \mathsf{g}(x-y) \cdot \partial_{i_1} v(x) \big) f_1^{i_1}(x) g_1^{i_2}(y)) dx dy \\ &-\int_{|x-y| \geqslant \varepsilon} (v(x) - v(y)) \cdot \nabla \partial_{i_1 i_2} \mathsf{g}(x-y) f_1^{i_1}(x) g_1^{i_2}(y)) dx dy. \end{aligned} \tag{3.32}$$

Observe from assumption (3.2) that the kernel $\nabla^{\otimes 2} \mathsf{g}(x-y)$ defines a $(\mathbb{R}^d)^{\otimes 2}$-valued Fourier multiplier with L^∞ symbol, and therefore by Cauchy-Schwarz,

$$\begin{aligned} \sup_{\varepsilon > 0} \left| \int_{|x-y| \geqslant \varepsilon} \big(\partial_{i_1} \nabla \mathsf{g}(x-y) \cdot \partial_{i_2} v(y) + \partial_{i_2} \nabla \mathsf{g}(x-y) \cdot \partial_{i_1} v(x) \big) f_1^{i_1}(x) g_1^{i_2}(y)) dx dy \right| \\ \lesssim \|\nabla v\|_{L^\infty} \|f_1\|_{L^2} \|g_1\|_{L^2} \lesssim \|\nabla v\|_{L^\infty} \|f\|_{\dot{H}^{-1}} \|g\|_{\dot{H}^{-1}}. \end{aligned} \tag{3.33}$$

For the remaining term, we use Taylor's theorem to write

$$\begin{aligned} (v(x) - v(y)) \cdot \nabla \partial_{i_1 i_2} \mathsf{g}(x-y) &= \left(\int_0^1 \nabla v(y + t(x-y)) \cdot (x-y) dt \right) \cdot \nabla \partial_{i_1 i_2} \mathsf{g}(x-y) \\ &= \left(\int_0^1 \nabla v(y + t(x-y)) dt \right) \mathsf{k}_{i_1 i_2}(x-y), \end{aligned} \tag{3.34}$$

where the $(\mathbb{R}^d)^{\otimes 2}$-valued kernel $\mathsf{k}_{i_1 i_2}$ is defined by

$$\mathsf{k}_{i_1 i_2}(x-y) := (x-y) \otimes \nabla \partial_{i_1 i_2} \mathsf{g}(x-y). \tag{3.35}$$

By assumption (3.3), $\mathsf{k}_{i_1 i_2}$ is a standard Calderón-Zygmund kernel which is associated to a Calderón-Zygmund operator (see [Gra14a, Theorem 5.4.1]). Consequently, the second term in (3.32) is a Calderón d-commutator, which by the Christ-Journé theorem [CJ87], is $\lesssim$

$$\|\nabla v\|_{L^\infty} \|f_1\|_{L^2} \|g_1\|_{L^2} \lesssim \|\nabla v\|_{L^\infty} \|f\|_{\dot{H}^{-1}} \|g\|_{\dot{H}^{-1}}. \tag{3.36}$$

The sub-Coulombic case. Let us now show that the above integration-by-parts argument works for the case $0 \leqslant s < d-2$. First, suppose that $d-4 \leqslant s < d-2$. An examination of the above argument shows that

$$\int_{(\mathbb{R}^d)^2} \mathsf{k}_v(x,y) f(x) g(y) dx dy = \lim_{\varepsilon \to 0} \int_{|x-y| \geqslant \varepsilon} \nabla_x \nabla_y \mathsf{k}_v(x,y) : f_1(x) \otimes g_1(y) dx dy. \tag{3.37}$$

Substituting in identity (3.23) yields

$$\lim_{\varepsilon\to 0}\Big(-\int_{|x-y|\geqslant\varepsilon}(v(x)-v(y))\cdot\nabla\partial_{i_1i_2}\mathrm{g}(x-y)f_1^{i_1}(x)g_1^{i_2}(y)dxdy \tag{3.38}$$
$$-\int_{|x-y|\geqslant\varepsilon}\big(\partial_{i_2}v(y)\cdot\nabla\partial_{i_1}\mathrm{g}(x-y)+\partial_{i_1}v(x)\cdot\nabla\partial_{i_2}\mathrm{g}(x-y)\big)f_1^{i_1}(x)g_1^{i_2}(y)dxdy\Big).$$

For the first term, we can repeat the proof above for the super-Coulombic ($d-4<s<d-2$) and Coulombic ($s=d-4$) cases with g replaced by $\partial_{i_1i_2}\mathrm{g}$, and writing f_1 as the divergence of a 2-tensor,

$$f_1^{i_1}=\mathrm{div}\nabla\Delta^{-1}f_1^{i_1}=:\mathrm{div}\,f_2^{i_1}, \tag{3.39}$$

and similarly for g_1. For the second term in (3.38), we can pass to the limit $\varepsilon\to 0$ (note that $\nabla^{\otimes 2}\mathrm{g}$ is locally integrable), so that by Fubini-Tonelli, we now consider

$$\int_{\mathbb{R}^d}(\nabla\partial_{i_1}\mathrm{g}*f_1^{i_1})(y)\cdot\partial_{i_2}v(y)g_1^{i_2}(y)dy+\int_{\mathbb{R}^d}\partial_{i_1}v(x)f_1^{i_1}(x)\cdot(\nabla\partial_{i_2}\mathrm{g}*g_1^{i_2})(x)dx. \tag{3.40}$$

Since f_1,g_1 satisfy the same bounds, we can use the $x\leftrightarrow y$ swapping symmetry of the kernel to only consider the first term in the preceding expression. Writing $\mathrm{I}=|\nabla|^{\frac{d-s}{2}-1}|\nabla|^{\frac{s-d}{2}+1}$, the Cauchy-Schwarz inequality implies

$$\Big|\int_{\mathbb{R}^d}(\nabla\partial_{i_1}\mathrm{g}*f_1^{i_1})(x)\cdot\partial_{i_2}v(x)g_1^{i_2}(x)dx\Big| \tag{3.41}$$
$$\lesssim\||\nabla|^{\frac{d-s}{2}-1}(\nabla\partial_{i_1}\mathrm{g}*f_1^{i_1})\|_{L^2}\||\nabla|^{\frac{s-d}{2}+1}(\partial_{i_2}vg_1^{i_2})\|_{L^2}.$$

By Plancherel and assumption (3.2), it then follows that

$$\||\nabla|^{\frac{d-s}{2}-1}(\nabla\partial_{i_1}\mathrm{g}*f_1^{i_1})\|_{L^2}\lesssim\|f\|_{\dot{H}^{\frac{s-d}{2}}}, \tag{3.42}$$

and by duality,

$$\begin{aligned}\||\nabla|^{\frac{s-d}{2}+1}(\partial_{i_2}vg_1^{i_2})\|_{L^2}&\leqslant\sup_{h\in\mathscr{S}(\mathbb{R}^d):\|h\|_{L^2}\leqslant 2}\Big|\int_{\mathbb{R}^d}|\nabla|^{\frac{s-d}{2}+1}h(x)\partial_{i_2}v(x)g_1^{i_2}(x)dx\Big|\\&\leqslant\sup_{h\in\mathscr{S}(\mathbb{R}^d):\|h\|_{L^2}\leqslant 2}\||\nabla|^{\frac{d-s}{2}-1}(|\nabla|^{\frac{s-d}{2}+1}h\partial_{i_2}v)\|_{L^2}\||\nabla|^{\frac{s-d}{2}+1}g_1^{i_2}\|_{L^2}\\&\lesssim\|g\|_{\dot{H}^{\frac{s-d}{2}}}\sup_{h\in\mathscr{S}(\mathbb{R}^d):\|h\|_{L^2}\leqslant 2}\||\nabla|^{\frac{d-s}{2}-1}(|\nabla|^{\frac{s-d}{2}+1}h\partial_{i_2}v)\|_{L^2}.\end{aligned} \tag{3.43}$$

By fractional Leibnitz rule (for instance, see [Gra14b, Theorem 7.6.1])

$$\||\nabla|^{\frac{d-s}{2}-1}(|\nabla|^{\frac{s-d}{2}+1}h\partial_{i_2}v)\|_{L^2}\lesssim\|h\|_{L^{p_1}}\|\nabla v\|_{L^{q_1}}+\||\nabla|^{\frac{s-d}{2}+1}h\|_{L^{p_2}}\||\nabla|^{\frac{d-s}{2}-1}\nabla v\|_{L^{q_2}} \tag{3.44}$$

for any $1<p_1,q_1,p_2,q_2\leqslant\infty$ with $p_i^{-1}+q_i^{-1}=2^{-1}$. We choose $p_1=2$ and $p_2=\frac{2d}{s+2}$. So by Hardy-Littlewood-Sobolev and the L^p boundedness of Riesz transforms, the right-hand side is $\lesssim$

$$\|h\|_{L^2}\Big(\|\nabla v\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}}\Big). \tag{3.45}$$

After a little bookkeeping, we arrive at the estimate

$$\left|\int_{\mathbb{R}^d}(\nabla\partial_{i_1}\mathrm{g}*f_1^{i_1})(x)\cdot\partial_{i_2}v(x)g_1^{i_2}(x)dx\right| \lesssim \|f\|_{\dot{H}^{\frac{s-d}{2}}}\|g\|_{\dot{H}^{\frac{s-d}{2}}}\left(\|\nabla v\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}}\right). \tag{3.46}$$

For the general case $0\leqslant s<d-2$, we let $k\in\mathbb{N}$ be such that $d-2(k+1)\leqslant s<d-2k$, and proceed by induction on k, using the argument just given above for the induction step. The existence of the form extension B_v follows by the usual density argument. □

4. Renormalization of the commutator estimate

In order to apply Theorem 3.1 to f,g which fail to be in $\dot{H}^{\frac{s-d}{2}}$, such as the Dirac mass, we now have to deal with the nondiagonal renormalization which is carried out through the smearing procedure. The main result of this section is the following proposition, which we view as a *renormalized commutator estimate* and which is the workhorse of this article.

Theorem 4.1 stated below should be compared to [Ser20b, Proposition 1.1], which, as previously discussed, treated the exact Coulomb/super-Coulombic Riesz case, as well as the log case in dimension 1. That result only required that the vector field v be in $W^{1,\infty}$. In contrast, Theorem 4.1 requires in the super-Coulombic case the stronger assumption that $v\in C^{1,\alpha}$, for any positive $\alpha\in(0,1]$, and belong to the Bessel potential space $W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}}$. The reason for this difference is precisely the combination of the commutator estimate with the extension/smearing procedure used in the present article. Namely, we want to apply Theorem 3.1 by using the smearing procedure to renormalize. But this forces us to work in the extended space $\mathbb{R}^{d+m}$ when we apply Theorem 3.1 and then to find a way to control the right-hand side of the proposition in terms of the modulated energy and norms of v and μ, which live in $\mathbb{R}^d$. Consequently, we choose an extension $\tilde{v}:\mathbb{R}^{d+m}\to\mathbb{R}^{d+m}$ of v and apply Theorem 3.1 with $\tilde{v}$. It is easy to construct an extension satisfying $\|\tilde{v}\|_{W^{1,\infty}}\lesssim\|v\|_{W^{1,\infty}}$ simply by multiplying v by a bump function in the new variable. But such an extension does not satisfy the auxiliary Sobolev condition, which is present (unless $s=d-1$) since we always choose m so that $d+m-s-2\geqslant 0$. Finding an extension which satisfies both conditions essentially amounts to showing that the trace operator

$$\mathrm{Tr}: W^{1,\infty}(\mathbb{R}^{d+m})\cap W^{\frac{d+m-s}{2},\frac{2(d+m)}{d+m-2-s}}(\mathbb{R}^{d+m})\to W^{1,\infty}(\mathbb{R}^d)\cap W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}}(\mathbb{R}^d) \tag{4.1}$$

is surjective with a bounded right inverse. If one could show such a result, which we do not know how to do, then one could drop the $C^{1,\alpha}$ assumption.

Proposition 4.1. *Let $d\geqslant 1$, $0\leqslant s<d$. Let $\underline{x}_N\in(\mathbb{R}^d)^N$ be pairwise distinct, and $\mu\in\mathscr{P}(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. If $s=0$, assume that $\int_{\mathbb{R}^d}\log(1+|x|)d\mu(x)<\infty$ and if $s\geqslant d-1$, assume that $\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}<\infty$. Let v be a continuous vector field on $\mathbb{R}^d$ such that*

$$\begin{cases}\|\nabla v\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}}\mathbf{1}_{s<d-2}<\infty, & 0\leqslant s\leqslant d-2\\ \|v\|_{C^{1,\alpha}}+\|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}}}<\infty, & d-2<s<d\end{cases} \tag{4.2}$$

for some $0 < \alpha \leqslant 1$ *and where* m *is the dimension of the extension space. There exists a constant* C *depending only* d, s, α, m *and on the potential* g *through assumptions* (1.13) - (1.28)*, such that*

$$
\begin{aligned}
(4.3)\quad & \left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} (v(x)-v(y)) \cdot \nabla \mathrm{g}(x-y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) \right| \\
& \leqslant C \Bigg(\|\nabla v\|_{L^\infty} + \| |\nabla|^{\frac{d-s}{2}} v \|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2} + \left(\|v\|_{C^{1,\alpha}} + \|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)}, \frac{2(d+m)}{d+m-2-s}}} \right) \mathbf{1}_{s>d-2} \Bigg) \\
& \times \Bigg(F_N(\underline{x}_N, \mu) + C(1+\|\mu\|_{L^\infty}) \Big(N^{-\frac{\min\{d-s,2\}}{\min\{d,s+2\}(s+1)}} + (N^{-1} \log N) \mathbf{1}_{s=0} + (N^{-\frac{2}{d(d-1)}} \log N) \mathbf{1}_{s=d-2} \Big) \\
& \qquad + C N^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}} \Big(\| |\nabla|^{s+1-d} \mu \|_{L^\infty} \mathbf{1}_{s<d-1} + \| |\nabla|^{s-d} \nabla \mu \|_{L^\infty} \mathbf{1}_{s \geqslant d-1} \Big) \Bigg).
\end{aligned}
$$

Proof. We need to introduce an extension of k_v to $\mathbb{R}^{d+m}$. By the surjectivity of the trace theorem for Besov and fractional Sobolev spaces (see [Tri10, Section 2.7.2]),[5] there exists a vector field $\tilde{v}$ on $\mathbb{R}^{d+m}$ whose $(m+1)$-th through $(m+d)$-th components vanish identically, $\tilde{v}(x,0) = (v(x),0)$, and

$$
\begin{aligned}
\|\tilde{v}\|_{C^{1,\alpha}(\mathbb{R}^{d+m})} &\lesssim \|v\|_{C^{1,\alpha}(\mathbb{R}^d)}, \\
\|\tilde{v}\|_{W^{\frac{d+m-s}{2}, \frac{2(d+m)}{d+m-2-s}}(\mathbb{R}^{d+m})} &\lesssim \|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)}, \frac{2(d+m)}{d+m-2-s}}(\mathbb{R}^d)}.
\end{aligned} \tag{4.4}
$$

With this notation, we define

$$
\mathsf{K}_v(X,Y) := \nabla \mathsf{G}(X-Y) \cdot (\tilde{v}(X) - \tilde{v}(Y)) \qquad \forall X \neq Y \in \mathbb{R}^{d+m}. \tag{4.5}
$$

We note that from assumption (1.24) that

$$
|\mathsf{K}_v(X,Y)| \leqslant C \|\nabla \tilde{v}\|_{L^\infty} \left(|X-Y|^{-s} \mathbf{1}_{s>0} + \mathbf{1}_{s=0} \right) \tag{4.6}
$$

and

$$
|\nabla_X \mathsf{K}_v(X,Y)| \leqslant C \|\nabla \tilde{v}\|_{L^\infty} |X-Y|^{-s-1}. \tag{4.7}
$$

For simplicity of notation, we again let X_i denote $(x_i, 0)$.

Adding and subtracting the smeared point mass $\delta_{X_i}^{(\eta)}$, for some η to be chosen momentarily, we obtain the decomposition

$$
\int_{(\mathbb{R}^d)^2 \setminus \triangle} \mathsf{k}_v(x,y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) = \mathrm{Term}_1 + \mathrm{Term}_2 + \mathrm{Term}_3, \tag{4.8}
$$

[5] This is where the $C^{1,\alpha}$ assumption is used, even though our estimates in the proof will only depend on the Lipschitz seminorm of the extended vector field $\tilde{v}$. It is possible to relax $C^{1,\alpha}$ to the Besov space $B^1_{\infty,1}$, but this is still stronger than $W^{1,\infty}$.

where

$$\mathrm{Term}_1 := \int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)\, d\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta)} - \tilde{\mu}\right)^{\otimes 2}(X,Y), \tag{4.9}$$

$$\mathrm{Term}_2 := -\frac{2}{N}\sum_{j=1}^N \int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)\, d\tilde{\mu}(X)\, d(\delta_{X_j} - \delta_{X_j}^{(\eta)})(Y), \tag{4.10}$$

$$\mathrm{Term}_3 := \frac{1}{N^2}\sum_{1\leqslant i,j\leqslant N} \int_{(\mathbb{R}^{d+m})^2\setminus\triangle} \mathsf{K}_\nu(X,Y)\, d(\delta_{X_i} + \delta_{X_i}^{(\eta)})(X)\, d(\delta_{X_j} - \delta_{X_j}^{(\eta)})(Y). \tag{4.11}$$

As before, $\tilde{\mu} := \mu\delta_{\mathbb{R}^d\times\{0\}}$. We proceed to estimate each of the Term_j individually.

Estimate for Term_1. This term is exactly of the form that can be handled by Theorem 3.1 with

$$f = g = \frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta)} - \tilde{\mu} \in \dot{H}^{\frac{s-d-m}{2}}(\mathbb{R}^{d+m}). \tag{4.12}$$

We deduce from that proposition that

$$|\mathrm{Term}_1| \leqslant C\left(\|\nabla\tilde{v}\|_{L^\infty} + \||\nabla|^{\frac{d+m-s}{2}}\tilde{v}\|_{L^{\frac{2(d+m)}{d+m-2-s}}}\mathbf{1}_{s<d+m-2}\right)\left\|\frac{1}{N}\sum_{i=1}^N \delta_{X_i}^{(\eta)} - \tilde{\mu}\right\|^2_{\dot{H}^{\frac{s-d-m}{2}}}. \tag{4.13}$$

Applying either Theorem 2.1 or Theorem 2.2, we can then bound

$$\begin{aligned}|\mathrm{Term}_1| &\leqslant C\left(\|\nabla\tilde{v}\|_{L^\infty} + \||\nabla|^{\frac{d+m-s}{2}}\tilde{v}\|_{L^{\frac{2(d+m)}{d+m-2-s}}}\mathbf{1}_{s<d+m-2}\right)\\ &\times\left(F_N(\underline{x}_N,\mu) + \frac{\mathsf{G}_\eta(0)}{N} + C\left(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0} + (\eta^2|\log\eta|)\mathbf{1}_{s=d-2}\right)\|\mu\|_{L^\infty} + C\eta^2\right).\end{aligned} \tag{4.14}$$

Estimate for Term_2. Suppose first that $0 \leqslant s < d-1$. We break Term_2 into a sum over the index i of terms

$$\begin{aligned}&\int_{|X-X_i|<2\eta}\int_{\mathbb{R}^{d+m}} \mathsf{K}_\nu(X,Y)\, d\left(\delta_{X_i} - \delta_{X_i}^{(\eta)}\right)(Y)\, d\tilde{\mu}(X)\\ &\qquad + \int_{|X-X_i|\geqslant 2\eta}\int_{\mathbb{R}^{d+m}} \mathsf{K}_\nu(X,Y)\, d\left(\delta_{X_i} - \delta_{X_i}^{(\eta)}\right)(Y)\, d\tilde{\mu}(X).\end{aligned} \tag{4.15}$$

For the first type of integral, we use that $\delta_{X_i}^{(\eta)}$ is a probability measure to rewrite

$$\int_{\mathbb{R}^{d+m}} \mathsf{K}_\nu(X,Y)\, d\left(\delta_{X_i} - \delta_{X_i}^{(\eta)}\right)(Y) = \int_{\mathbb{R}^{d+m}} \left(\mathsf{K}_\nu(X,X_i) - \mathsf{K}_\nu(X,Y)\right) d\delta_{X_i}^{(\eta)}(Y), \tag{4.16}$$

and unpacking the definition (4.5) of K_ν, we have

$$\begin{aligned}&\mathsf{K}_\nu(X,X_i) - \mathsf{K}_\nu(X,Y)\\ &\qquad = (\tilde{v}(Y) - \tilde{v}(X_i))\cdot\nabla\mathsf{G}(X-X_i) + (\tilde{v}(X) - \tilde{v}(Y))\cdot(\nabla\mathsf{G}(X-X_i) - \nabla\mathsf{G}(X-Y)).\end{aligned} \tag{4.17}$$

Using (1.24), the mean-value theorem, and triangle inequality

$$\int_{\mathbb{R}^{d+m}} |\mathsf{K}_\nu(X,X_i)-\mathsf{K}_\nu(X,Y)|d\delta_{X_i}^{(\eta)}(Y) \leqslant C\eta\|\nabla\tilde v\|_{L^\infty}\left(|X-X_i|^{-s-1}+\int_{\mathbb{R}^{d+m}}\big(|X-Y|^{-s-1}\big)d\delta_{X_i}^{(\eta)}(Y)\right) \tag{4.18}$$

for all $|X-X_i|\leqslant 2\eta$. Integrating the left-hand side with respect $d\tilde\mu(X)$ and using Fubini-Tonelli, we find that the first term in (4.15) is controlled by

$$C\eta\|\nabla\tilde v\|_{L^\infty}\||\nabla|^{s+1-d}\mu\|_{L^\infty}, \tag{4.19}$$

where we also use that $|\cdot|^{-s-1}$ is (up to a constant) the convolution kernel of the Fourier multiplier $|\nabla|^{s+1-d}$ in $\mathbb{R}^d$.

For the second type of contributions we use the mean-value inequality and (4.7) to bound

$$\begin{aligned}&\left|\int_{|X-X_i|\geqslant 2\eta}\int_{\mathbb{R}^{d+m}}\mathsf{K}_\nu(X,Y)d\left(\delta_{X_i}-\delta_{X_i}^{(\eta)}\right)(Y)d\tilde\mu(X)\right|\\ &\leqslant C\eta\|\nabla\tilde v\|_{L^\infty}\int_{|X-X_i|\geqslant 2\eta}|X-X_i|^{-s-1}d\tilde\mu(X)\\ &\leqslant C\eta\|\nabla\tilde v\|_{L^\infty}\||\nabla|^{s+1-d}\mu\|_{L^\infty}.\end{aligned} \tag{4.20}$$

Now if $d-1\leqslant s<d$, we integrate by parts in X to write

$$\begin{aligned}&\int_{(\mathbb{R}^{d+m})^2}\mathsf{K}_\nu(X,Y)d\Big(\delta_{X_i}-\delta_{X_i}^{(\eta)}\Big)(Y)d\tilde\mu(X)\\ &=-\int_{(\mathbb{R}^{d+m})^2}\tilde{\mathsf{K}}_\nu(X,Y)\cdot d\Big(\delta_{X_i}-\delta_{X_i}^{(\eta)}\Big)(Y)d\widetilde{\nabla\mu}(X)\\ &\quad-\int_{\mathbb{R}^d}\operatorname{div}\tilde v(x,0)\big(\mathsf{G}(x-x_i,0)-\mathsf{G}_\eta(x-x_i,0)\big)d\mu(x),\end{aligned} \tag{4.21}$$

where $\widetilde{\nabla\mu}(X):=\nabla\mu(x)\delta_{\mathbb{R}^d\times\{0\}}(X)$ and we have defined the vector-valued kernel

$$\tilde{\mathsf{K}}_\nu(X,Y):=(\tilde v(X)-\tilde v(Y))\mathsf{G}(X-Y). \tag{4.22}$$

Proceeding as in the case $0\leqslant s<d-1$ with μ replaced by $\nabla\mu$, the first term in the preceding right-hand side is bounded by

$$C\eta\|\nabla\tilde v\|_{L^\infty}\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}. \tag{4.23}$$

Using that

$$\int_{|x-x_i|\leqslant 2\eta}\big|\mathsf{G}(x-x_i,0)-\mathsf{G}_\eta(x-x_i,0)\big|\,dx\leqslant C\eta^{d-s} \tag{4.24}$$

by assumption (1.24) and also using the bound (2.21), we see from Hölder's inequality that the second term is bounded by

$$C\|\nabla\tilde v\|_{L^\infty}\|\mu\|_{L^\infty}\eta^{d-s}. \tag{4.25}$$

After a little bookkeeping, we conclude that

$$|\mathrm{Term}_2| \leqslant C\|\nabla\tilde{v}\|_{L^\infty}\Big(\eta\||\nabla|^{s+1-d}\mu\|_{L^\infty}\mathbf{1}_{s<d-1} + \Big(\eta\||\nabla|^{s-d}\nabla\mu\|_{L^\infty} + \eta^{d-s}\|\mu\|_{L^\infty}\Big)\mathbf{1}_{s\geqslant d-1}\Big). \tag{4.26}$$

Estimate for Term_3. We first remove the self-interaction by observing that

$$\begin{aligned}&\sum_{i=1}^N \int_{(\mathbb{R}^{d+m})^2\setminus\triangle} \mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_i}-\delta^{(\eta)}_{X_i})(Y)\\ &= -\sum_{i=1}^N \int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)d(\delta^{(\eta)}_{X_i})^{\otimes 2}(X,Y),\end{aligned} \tag{4.27}$$

so that

$$\begin{aligned}\mathrm{Term}_3 = &-\frac{1}{N^2}\sum_{i=1}^N \int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)d(\delta^{(\eta)}_{X_i})^{\otimes 2}(X,Y)\\ &+\frac{1}{N^2}\sum_{1\leqslant i\neq j\leqslant N}\int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_j}-\delta^{(\eta)}_{X_j})(Y).\end{aligned} \tag{4.28}$$

For the first sum, we use assumption (1.25) with the estimates (4.6) and (2.6) to bound it as

$$\begin{aligned}\frac{1}{N^2}\left|\sum_{i=1}^N \int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)d(\delta^{(\eta)}_{X_i})^{\otimes 2}(X,Y)\right| &\leqslant \frac{C}{N}\|\nabla\tilde{v}\|_{L^\infty}\big(\mathsf{G}_\eta(0)\mathbf{1}_{s>0}+\mathbf{1}_{s=0}\big)\\ &\leqslant \frac{C}{N}\|\nabla\tilde{v}\|_{L^\infty}\big(\eta^{-s}\mathbf{1}_{s>0}+\mathbf{1}_{s=0}\big).\end{aligned} \tag{4.29}$$

Next, we split the second sum over $1\leqslant i\neq j\leqslant N$ into a sum over "close pairs" for which $|x_i-x_j|<\varepsilon$ and a sum over "far pairs" for which $|x_i-x_j|\geqslant\varepsilon$, where $\varepsilon>2\eta$ is a parameter, to be chosen later. For the sum over close pairs, we again use assumption (1.25) with the bound (4.6) to obtain

$$\begin{aligned}&\frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|<\varepsilon}}\left|\int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_j}-\delta^{(\eta)}_{X_j})(Y)\right|\\ &\leqslant \frac{C\|\nabla\tilde{v}\|_{L^\infty}}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|<\varepsilon}}\left(\left(\mathsf{g}(x_i-x_j)+\int \mathsf{G}_\eta(X_i-Y)d\delta^{(\eta)}_{X_j}(Y)\right)\mathbf{1}_{s>0}+\mathbf{1}_{s=0}\right).\end{aligned} \tag{4.30}$$

Using (2.20), we then obtain (provided $\varepsilon < \frac{r_0}{2}$)

$$\begin{aligned}
&\frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|<\varepsilon}}\left|\int_{(\mathbb{R}^{d+m})^2}\mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_j}-\delta^{(\eta)}_{X_j})(Y)\right|\\
&\leqslant C\frac{\|\nabla\tilde v\|_{L^\infty}}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|<\varepsilon}}\big(\mathrm{g}(x_i-x_j)\mathbf{1}_{s>0}+\mathbf{1}_{s=0}\big)\\
&\leqslant C\|\nabla\tilde v\|_{L^\infty}\bigg(F_N(\underline{x}_N,\mu)+C\Big(\frac{\varepsilon^{-s}+|\log\varepsilon|\mathbf{1}_{s=0}}{N}+\varepsilon^2\\
&\qquad+\|\mu\|_{L^\infty}(\varepsilon^{d-s}+(\varepsilon^d|\log\varepsilon|)\mathbf{1}_{s=0}+(\varepsilon^2|\log\varepsilon|)\mathbf{1}_{s=d-2})\Big)\bigg),
\end{aligned}\tag{4.31}$$

where the ultimate inequality follows from application of Theorem 2.3. For the sum over far pairs, we first note that by symmetry under swapping $i\leftrightarrow j$,

$$\begin{aligned}
&\sum_{\substack{1\leqslant i,j\leqslant N\\ |x_i-x_j|\geqslant\varepsilon}}\int_{(\mathbb{R}^{d+m})^2}\mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_j}-\delta^{(\eta)}_{X_j})(Y)\\
&=\sum_{\substack{1\leqslant i,j\leqslant N\\ |x_i-x_j|\geqslant\varepsilon}}\int_{(\mathbb{R}^{d+m})^2}\big(\mathsf{K}_\nu(X_i,X_j)-\mathsf{K}_\nu(X,Y)\big)d\delta^{(\eta)}_{X_i}(X)d\delta^{(\eta)}_{X_j}(Y).
\end{aligned}\tag{4.32}$$

Since $\varepsilon>2\eta$, using the mean-value inequality and (4.7), we find

$$\frac{1}{N^2}\sum_{\substack{1\leqslant i\neq j\leqslant N\\ |x_i-x_j|\geqslant\varepsilon}}\left|\int_{(\mathbb{R}^{d+m})^2}\mathsf{K}_\nu(X,Y)d(\delta_{X_i}+\delta^{(\eta)}_{X_i})(X)d(\delta_{X_j}-\delta^{(\eta)}_{X_j})(Y)\right|\leqslant\frac{C\eta\|\nabla\tilde v\|_{L^\infty}}{\varepsilon^{s+1}}.\tag{4.33}$$

Putting together the estimates (4.29), (4.31), and (4.33), we obtain

$$\begin{aligned}
|\mathrm{Term}_3|\leqslant C\|\nabla\tilde v\|_{L^\infty}\bigg(&F_N(\underline{x}_N,\mu)+C\Big(\frac{\eta^{-s}+|\log\eta|\mathbf{1}_{s=0}}{N}+\frac{\varepsilon^{-s}+|\log\varepsilon|\mathbf{1}_{s=0}}{N}+\varepsilon^2\\
&+\|\mu\|_{L^\infty}(\varepsilon^{d-s}+(\varepsilon^d|\log\varepsilon|)\mathbf{1}_{s=0}+(\varepsilon^2|\log\varepsilon|)\mathbf{1}_{s=d-2})+\eta\varepsilon^{-s-1}\Big)\bigg).
\end{aligned}\tag{4.34}$$

Conclusion. Combining the estimates (4.14), (4.26), (4.34) obtained above, we have shown that

$$\begin{aligned}
&\left|\int_{(\mathbb{R}^d)^2\setminus\Delta}\mathsf{k}_\nu(x,y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu\right)^{\otimes2}(x,y)\right|\\
&\leqslant C\Big(\|\nabla\tilde v\|_{L^\infty}+\||\nabla|^{\frac{d+m-s}{2}}\tilde v\|_{L^{\frac{2(d+m)}{d+m-2-s}}}\mathbf{1}_{s<d+m-2}\Big)\bigg(F_N(\underline{x}_N,\mu)+C\Big(\frac{\eta^{-s}+|\log\eta|\mathbf{1}_{s=0}}{N}+\eta^2\\
&\qquad+\|\mu\|_{L^\infty}(\eta^{d-s}+(\eta^d|\log\eta|)\mathbf{1}_{s=0}+(\eta^2|\log\eta|)\mathbf{1}_{s=d-2})\Big)\bigg)
\end{aligned}\tag{4.35}$$

$$+C\|\nabla\tilde{v}\|_{L^\infty}\Big((\eta\||\nabla|^{s+1-d}\mu\|_{L^\infty})\mathbf{1}_{s<d-1}+(\eta\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}+\eta^{d-s}\|\mu\|_{L^\infty})\mathbf{1}_{s\geqslant d-1}$$
$$+\frac{\eta^{-s}+|\log\eta|\mathbf{1}_{s=0}}{N}+\frac{\varepsilon^{-s}+|\log\varepsilon|\mathbf{1}_{s=0}}{N}+\varepsilon^2$$
$$+\|\mu\|_{L^\infty}(\varepsilon^{d-s}+(\varepsilon^d|\log\varepsilon|)\mathbf{1}_{s=0}+(\varepsilon^2|\log\varepsilon|)\mathbf{1}_{s=d-2})+\eta\varepsilon^{-s-1}\Big).$$

We may now optimize over η and ε by taking

$$\begin{cases}\eta=\varepsilon^{s+2}\text{ and }\varepsilon=N^{-1}, & s=0\\ \eta=\varepsilon^{s+3}\text{ and }\varepsilon=N^{-\frac{1}{(s+2)(s+1)}}, & 0<s\leqslant d-2\\ \eta=\varepsilon^{d+1}\text{ and }\varepsilon=N^{-\frac{1}{d(s+1)}}, & d-2<s<d.\end{cases}\tag{4.36}$$

Using (4.4) to control the norms of $\tilde{v}$ in terms of norms of v, we then find

$$\left|\int_{(\mathbb{R}^d)^2\setminus\Delta}\mathsf{k}_v(x,y)d\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu\right)^{\otimes 2}(x,y)\right|\leqslant C\Big(\|\nabla v\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}}\mathbf{1}_{s<d-2}+\tag{4.37}$$
$$+\left(\|v\|_{C^{1,\alpha}}+\|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}}}\right)\mathbf{1}_{s>d-2}\Big)\Big(F_N(\underline{x}_N,\mu)$$
$$+C(1+\|\mu\|_{L^\infty})\Big(N^{-\frac{\min\{d-s,2\}}{\min\{d,s+2\}(s+1)}}+(N^{-1}\log N)\mathbf{1}_{s=0}+(N^{-\frac{2}{d(d-1)}}\log N)\mathbf{1}_{s=d-2}\Big)$$
$$+CN^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}}\Big(\||\nabla|^{s+1-d}\mu\|_{L^\infty}\mathbf{1}_{s<d-1}+\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}\mathbf{1}_{s\geqslant d-1}\Big)\Big).$$

Thus, the proof of the proposition is complete. □

5. Gronwall argument

We now have all the ingredients for our Gronwall argument on the modulated energy. Set $u:=\mathbb{M}\nabla\mathsf{g}*\mu$. Applying the bound of Theorem 4.1 pointwise in time with $v=u^t$ to the right-hand side of inequality (1.7) and then integrating with respect to time, we find that for $0\leqslant t\leqslant T$,

$$|F_N(\underline{x}_N^t,\mu^t)|\leqslant|F_N(\underline{x}_N^0,\mu^0)|+C\int_0^t\Big(\|\nabla u^\tau\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}u^\tau\|_{L^{\frac{2d}{d-2-s}}}\mathbf{1}_{s<d-2}\tag{5.1}$$
$$+\left(\|u^\tau\|_{C^{1,\alpha}}+\|u^\tau\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)},\frac{2(d+m)}{d+m-2-s}}}\right)\mathbf{1}_{s>d-2}\Big)\Big(F_N(\underline{x}_N^\tau,\mu^\tau)$$
$$+C(1+\|\mu^\tau\|_{L^\infty})\Big(N^{-\frac{\min\{d-s,2\}}{\min\{d,s+2\}(s+1)}}+(N^{-1}\log N)\mathbf{1}_{s=0}+(N^{-\frac{2}{d(d-1)}}\log N)\mathbf{1}_{s=d-2}\Big)$$
$$+CN^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}}\Big(\||\nabla|^{s+1-d}\mu^\tau\|_{L^\infty}\mathbf{1}_{s<d-1}+\||\nabla|^{s-d}\nabla\mu^\tau\|_{L^\infty}\mathbf{1}_{s\geqslant d-1}\Big)\Big)d\tau.$$

If $0\leqslant s<d-2$, then by Sobolev embedding, assumption (1.16), and Hölder's inequality,

$$\||\nabla|^{\frac{d-s}{2}}u^\tau\|_{L^{\frac{2d}{d-2-s}}}+\|\nabla u^\tau\|_{L^\infty}\lesssim 1+\|\mu^\tau\|_{L^\infty}.\tag{5.2}$$

If $0 \leqslant s < d-1$, then it also follows from Hölder's inequality that

$$\| |\nabla|^{s+1-d} \mu^\tau \|_{L^\infty} \lesssim 1 + \|\mu^\tau\|_{L^\infty}. \tag{5.3}$$

Applying the Gronwall-Bellman lemma, we conclude that for all $0 \leqslant t \leqslant T$,

$$\begin{aligned} |F_N(\underline{x}_N^t, \mu^t)| \leqslant A_N(t) \exp\Bigg(C \int_0^t \Big(&\|\nabla u^\tau\|_{L^\infty} + \| |\nabla|^{\frac{d-s}{2}} u^\tau \|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2} \\ &+ \Big(\|u^\tau\|_{C^{1,\alpha}} + \|u^\tau\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)}, \frac{2(d+m)}{d+m-2-s}}} \Big) \mathbf{1}_{s>d-2} \Big) d\tau \Bigg), \end{aligned} \tag{5.4}$$

where the time-dependent prefactor A_N is defined by

$$\begin{aligned} A_N(t) \coloneqq C \int_0^t \Bigg(&(1 + \|\mu^\tau\|_{L^\infty}) \Big(N^{-\frac{\min\{d-s,2\}}{\min\{d,s+2\}(s+1)}} + (N^{-1} \log N) \mathbf{1}_{s=0} + (N^{-\frac{2}{d(d-1)}} \log N) \mathbf{1}_{s=d-2} \Big) \\ &+ C N^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}} \Big(\| |\nabla|^{s+1-d} \mu^\tau \|_{L^\infty} \mathbf{1}_{s<d-1} + \| |\nabla|^{s-d} \nabla \mu^\tau \|_{L^\infty} \mathbf{1}_{s \geqslant d-1} \Big) \Bigg) d\tau + |F_N(\underline{x}_N^0, \mu^0)|. \end{aligned} \tag{5.5}$$

Thus, the proof of our bound for the evolution of the modulated energy and therefore of Theorem 1.1 is complete.

6. Multiplicative Noise

6.1. **Overview.** In this last section, we show how our method of proof can in principle be extended to treat the mean-field limit of first-order systems, such as (1.1), with multiplicative noise added to the dynamics:

$$\begin{cases} \dot{x}_i = \dfrac{1}{N} \displaystyle\sum_{1 \leqslant j \leqslant N : j \neq i} \mathbb{M} \nabla \mathrm{g}(x_i - x_j) + \sum_{k=1}^{\infty} \sigma_k(x_i) \circ \dot{W}_k \\ x_i|_{t=0} = x_i^0, \end{cases} \qquad i \in \{1, \dots, N\}. \tag{6.1}$$

Here, the σ_k are C^∞ vector fields on $\mathbb{R}^d$, the W_k are independent real Brownian motions, and $\circ$ denotes that product should be interpreted in the Stratonovich sense.[6] The correspond mean-field equation is no longer deterministic, but instead gains a stochastic term:

$$\begin{cases} \partial_t \mu = -\operatorname{div}((\mathbb{M} \nabla \mathrm{g} * \mu)\mu) - \displaystyle\sum_{k=1}^{\infty} \operatorname{div}(\sigma_k \mu) \circ \dot{W}_k \\ \mu(0) = \mu^0 \end{cases} \qquad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d. \tag{6.2}$$

The model example of a system of the form (6.1) is the *stochastic point vortex model* of Flandoli, Gubinelli, and Priola [FGP11], for which g is the 2D Coulomb potential. The limiting SPDE is now the 2D incompressible Euler vorticity equation with an additional stochastic transport term, the well-posedness of which has been studied by Brzeźniak, Flandoli, and Maurelli [BFM16]. We emphasize that the noise in (6.1) is very different than

[6] When the $\circ$ is not present, the product should be interpreted in the Itô sense throughout this paper.

the so-called additive noise model,

$$\dot{x}_i = \frac{1}{N} \sum_{1 \leqslant j \leqslant N : j \neq i} \mathbb{M} \nabla \mathrm{g}(x_i - x_j) + \sqrt{2\sigma} \underline{\dot{W}}_i, \tag{6.3}$$

where now $\sigma > 0$ is a constant and the $\underline{W}_i$ are independent Brownian motions in $\mathbb{R}^d$. Indeed, in equation (6.3), the noise is independent for each particle, whereas in equation (6.1), the spatially dependent noise acting on each particle is the same. This difference clearly manifests itself at the limiting level. As has been rigorously shown by numerous authors [Osa87b, Osa87a, FHM14, LY16, JW18, LLY19, BJW19b, BJW19a, BJW20], the empirical measure of the system (6.3) converges in law to the solution of the deterministic diffusive/viscous equation

$$\partial_t \mu = -\operatorname{div}((\mathbb{M} \nabla \mathrm{g} * \mu)\mu) + \sigma \Delta \mu. \tag{6.4}$$

As mentioned in the introduction, the second author [Ros20] was the first to study the convergence in law of the empirical measure for (6.1) to the solution (if one exists) of the SPDE (6.2).[7] By developing a stochastic extension of the modulated-energy approach of [Ser20b] and introducing the commutator perspective adopted in this paper, he proved a quantitative estimate for the convergence in the 2D Coulomb case corresponding to the aforementioned stochastic point vortex model. The proof can be extended to dimensions $d \geqslant 3$, but, importantly, it very strongly depends on the Coulomb nature of the interaction. At the time of the present article, this is the only such result for singular interactions of which we are aware (see [CF16] for the case of regular interactions).

The idea of [Ros20] is to consider the modulated energy $F_N(\underline{x}_N^t, \mu^t)$ as before, noting that it is now a real-valued stochastic process which is almost surely finite. A formal application of Itô's lemma (in Stratonovich form) to $F_N(\underline{x}_N^t, \mu^t)$ yields the stochastic differential inequality

$$\begin{aligned} \frac{d}{dt} F_N(\underline{x}_N^t, \mu^t) \leqslant & \int_{(\mathbb{R}^d)^2 \setminus \triangle} \big(u^t(x) - u^t(y)\big) \cdot \nabla \mathrm{g}(x-y) d \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i^t} - \mu^t \right)^{\otimes 2} (x,y) \\ & + \sum_{k=1}^{\infty} \int_{(\mathbb{R}^d)^2 \setminus \triangle} \nabla \mathrm{g}(x-y) \cdot \big(\sigma_k(x) - \sigma_k(y)\big) d \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i^t} - \mu^t \right)^{\otimes 2} (x,y) \circ \dot{W}_k, \end{aligned} \tag{6.5}$$

[7] An earlier work of Coghi and Maurelli [CM20] considered this problem but with an N-dependent asymptotically vanishing truncation of the potential in (6.1) to distances much larger than the typical interparticle distance $N^{-1/d}$.

where again $u^t := \mathbb{M}\nabla \mathsf{g} * \mu^t$. Converting from Stratonovich to Itô, the second line becomes

$$\begin{aligned}(6.6)\quad &\sum_{k=1}^{\infty}\int_{(\mathbb{R}^d)^2\setminus\triangle}\nabla \mathsf{g}(x-y)\cdot\big(\sigma_k(x)-\sigma_k(y)\big)d\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i^t}-\mu^t\right)^{\otimes 2}(x,y)\dot{W}_k\\ &+\frac{1}{2}\sum_{k=1}^{\infty}\int_{(\mathbb{R}^d)^2\setminus\triangle}\nabla \mathsf{g}(x-y)\cdot(\sigma_k\cdot\nabla\sigma_k(x)-\sigma_k\cdot\nabla\sigma_k(y))d\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i^t}-\mu^t\right)^{\otimes 2}(x,y)\\ &\qquad+\frac{1}{2}\sum_{k=1}^{\infty}\int_{(\mathbb{R}^d)^2\setminus\triangle}\nabla^{\otimes 2}\mathsf{g}(x-y):\big(\sigma_k(x)-\sigma_k(y)\big)^{\otimes 2}d\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i^t}-\mu^t\right)^{\otimes 2}(x,y).\end{aligned}$$

The first term in (6.6), which is now an Itô integral, should have zero expectation, therefore we can ignore it. The second term has the same commutator structure as in the functional inequality (1.8) with $\nu = \sum_{k=1}^{\infty}\sigma_k\cdot\nabla\sigma_k$. Consequently, we can treat this expression as a renormalized commutator using Theorem 4.1. The third term is more complicated. It appears similar to the first two terms, except now there are two derivatives on g and two symmetrized expressions $\sigma_k(x)-\sigma_k(y)$ which, in principle, should cancel out the singularities induced by the derivatives. However, Theorem 4.1 does not cover such expressions. Accordingly, the goal of this section is to prove its second-order analogue, which is Theorem 6.2 stated below. Again, the idea will be to prove a second-order version of the commutator estimate of Theorem 3.1 using integration by parts and then to obtain a renormalization of this estimate using the smearing procedure.

The preceding computation to arrive at (6.5) and (6.6) is completely formal. We cannot directly invoke Itô's lemma because the potential g is singular. Also, we have not specified whether there is a solution to the system (6.1) or in what sense the equation (6.2) holds. The N-body problem is relatively straightforward to make sense of. Using a truncation of the potential near the origin and a stopping time argument in the spirit of [FGP11, Section 3], one can show there is a unique strong solution. In particular, with probability one, the particles never collide. The SPDE (6.2) is more difficult to make sense of. In fact, we are unaware of any well-posedness results for this equation aside from the 2D Euler case in [BFM16], a gap in the literature which should be filled. Accordingly, we will not attempt to rigorously justify equation (6.5) in this article. Instead, we will only show in this section how to use our methods to estimate the last term in (6.6) *pathwise* in the noise. Since the sub-Coulombic case is strictly easier than the Coulomb case, we expect that the results of [BFM16] can be generalized to this case without issue. When combined with the results of the present paper and the reasoning in [Ros20], all computations should be justifiable in the sub-Coulombic case in a straightforward manner.

The assumptions (1.11) – (1.19) in the case $0 \leqslant s \leqslant d-2$ and (1.20) – (1.28) in the case $d-2 < s < d$ carry over. But we also need to supplement (1.18) and (1.28), respectively, with the assumptions that the $(\mathbb{R}^d)^{\otimes(4+2k)}$-valued kernel

$$(6.7)\qquad \mathsf{k}(x-y) := (x-y)^{\otimes 2}\otimes\nabla^{\otimes(2+2k)}\mathsf{g}(x-y)$$

and the $(\mathbb{R}^d)^{\otimes 6}$-valued kernel

$$\mathsf{K}(X-Y) := (X-Y)^{\otimes 2} \otimes \nabla^{\otimes 4} \mathsf{G}(X-Y) \tag{6.8}$$

are associated to Calderón-Zygmund operators on $\mathbb{R}^d$ and $\mathbb{R}^{d+m}$.

6.2. **Second-order commutator estimate.** Following the strategy of Section 3, we prove a second-order version of the commutator estimate Theorem 3.1. With more work, one can extend the method of proof to arbitrary k-th order, but having no need for such generality, we will not do so in this work.

Proposition 6.1. *Let $0 \leqslant s < d$. Let $\mathrm{g} \in C^\infty(\mathbb{R}^d \setminus \{0\})$ such that $\mathrm{g}(x) = \mathrm{g}(-x)$ in $B(0, r_0)$, and*

$$\forall k \geqslant 1, \qquad |\nabla^{\otimes k} \mathrm{g}(x)| \lesssim |x|^{-k-s} \qquad \forall x \neq 0. \tag{6.9}$$

If $s \leqslant d-2$, then also assume that

$$|\hat{\mathrm{g}}(\xi)| \lesssim |\xi|^{s-d} \qquad \forall \xi \neq 0. \tag{6.10}$$

If $s = d-2k$, for some $k \in \mathbb{N}$, also assume that the $(\mathbb{R}^d)^{\otimes(4+2k)}$-valued kernel

$$\mathsf{k}(x,y) := (x-y)^{\otimes 2} \otimes \nabla^{\otimes(2+2k)} \mathrm{g}(x-y) \tag{6.11}$$

is associated to a Calderón-Zygmund operator. Let v be a Lipschitz continuous vector field on $\mathbb{R}^d$. There exists a constant $C > 0$, depending only on s, d, and the potential g through (3.1) – (3.3)*, such that for any $f, g \in \mathscr{S}(\mathbb{R}^d)$,[8] we have*

$$\begin{aligned} &\left| \int_{(\mathbb{R}^d)^2} (v(x)-v(y))^{\otimes 2} : \nabla^{\otimes 2} \mathrm{g}(x-y) f(x) g(y) dx dy \right| \\ &\qquad \leqslant C \left(\|\nabla v\|_{L^\infty} + \| |\nabla|^{\frac{d-s}{2}} v \|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2} \right)^2 \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \end{aligned} \tag{6.12}$$

Consequently, the integral in the left-hand side of (6.12) *extends to a bounded bilinear form $B_v(\cdot,\cdot)$ on $\dot{H}^{\frac{s-d}{2}}(\mathbb{R}^d)$ satisfying the bound* (6.12)*.*

Proof. We follow the outline of the proof of Theorem 3.1. Since the arguments are very similar, we only sketch the proof, focusing on what is different. Again, we will not explicitly track the dependence of the implicit constants.

The super-Coulombic case. Recycling the kernel notation k_v, define

$$\mathsf{k}_v(x,y) := (v(x)-v(y))^{\otimes 2} : \nabla^{\otimes 2} \mathrm{g}(x-y) \qquad \forall x \neq y. \tag{6.13}$$

Note that by assumption (1.24) and the mean-value theorem, we have the kernel estimate

$$|\mathsf{k}_v(x,y)| \leqslant C \min\left(\frac{\|\nabla v\|_{L^\infty}^2}{|x-y|^s}, \frac{\|v\|_{L^\infty}^2}{|x-y|^{s+2}} \right). \tag{6.14}$$

By approximation, we assume that $v \in C^\infty$; $f, g \in \mathscr{S}(\mathbb{R}^d)$ with Fourier support away from the origin; and g has compact support. Writing

$$f = \operatorname{div} \nabla \Delta^{-1} f =: \operatorname{div} f_1 \tag{6.15}$$

[8] If $s = 0$, then it is again implicit that the Fourier transforms of f, g vanish sufficiently rapidly at the origin so that the $\dot{H}^{-d/2}$ norm converges.

and similarly for g, we find

$$\int_{(\mathbb{R}^d)^2} \mathsf{k}_v(x,y) f(x) g(y) dxdy = \lim_{\varepsilon \to 0^+} \int_{|x-y|\geqslant \varepsilon} \mathsf{k}_v(x,y) \operatorname{div} f_1(x) \operatorname{div} g_1(y) dxdy. \tag{6.16}$$

Integrating by parts and using assumption (6.9) to estimate the boundary terms as in the proof of Theorem 3.1, we find that

$$\begin{aligned} \int_{|x-y|\geqslant \varepsilon} \mathsf{k}_v(x,y) f(x) g(y) dxdy &= O(\|\nabla^{\otimes 2} v\|_{L^\infty}^2 \|f g_1\|_{L^1} \varepsilon^{d-s}) \\ &+ O(\|v\|_{W^{2,\infty}}^2 \|\nabla f_1 g_1\|_{L^1} \varepsilon^{d-s}) + O(\|v\|_{W^{1,\infty}}^2 \|\nabla^{\otimes 2} f_1\|_{L^\infty} \|g_1\|_{L^1} \varepsilon^{d-s}) \\ &- \frac{1}{2} \int_{|x-y|\geqslant \varepsilon} \nabla_x \nabla_y \mathsf{k}_v(x,y) : \big((f_1(x) - f_1(y)) \otimes (g_1(x) - g_1(y)) \big) dxdy. \end{aligned} \tag{6.17}$$

The first three terms of the right-hand side vanish as $\varepsilon \to 0$ since $s < d$. For the last term, direct computation reveals

$$\begin{aligned} \nabla_x \nabla_y \mathsf{k}_v(x,y) &= -\nabla^{\otimes 4} \mathsf{g}(x-y) : (v(x) - v(y))^{\otimes 2} \\ &- 2\nabla^{\otimes 3} \mathsf{g}(x-y) : \big((\nabla v(x) + \nabla v(y)) \otimes (v(x) - v(y)) \big) \\ &- 2\nabla^{\otimes 2} \mathsf{g}(x-y) : (\nabla v(x) \otimes \nabla v(y)). \end{aligned} \tag{6.18}$$

So using the triangle inequality and mean-value theorem, we find that

$$\big| \nabla_x \nabla_y \mathsf{k}_v(x,y) \big| \lesssim \frac{\|\nabla v\|_{L^\infty}^2}{|x-y|^{s+2}}, \tag{6.19}$$

which implies by Cauchy-Schwarz that

$$\begin{aligned} &\int_{|x-y|\geqslant \varepsilon} |\nabla_x \nabla_y \mathsf{k}_v(x,y)| |f_1(x) - f_1(y)| |g_1(x) - g_1(y)| dxdy \\ &\lesssim \|\nabla v\|_{L^\infty}^2 \left(\int_{(\mathbb{R}^d)^2} \frac{|f_1(x) - f_1(y)|^2}{|x-y|^{s+2}} dxdy \right)^{1/2} \left(\int_{(\mathbb{R}^d)^2} \frac{|g_1(x) - g_1(y)|^2}{|x-y|^{s+2}} dxdy \right)^{1/2} \\ &\lesssim \|\nabla v\|_{L^\infty}^2 \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \end{aligned} \tag{6.20}$$

In conclusion,

$$\lim_{\varepsilon \to 0} \left| \int_{|x-y|\geqslant \varepsilon} \mathsf{k}_v(x,y) f(x) g(y) dxdy \right| \lesssim \|\nabla v\|_{L^\infty}^2 \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \tag{6.21}$$

The Coulombic case. To modify the proof of Theorem 3.1 in the Coulombic case $s = d-2$, we first note that if $|x-y| = \varepsilon$, then writing $x = y + z$,

$$\begin{aligned} \frac{(x-y)^{i_1}}{|x-y|} \partial_{y_{i_2}} \mathsf{k}_v(x,y) &= \frac{z^{i_1}}{|z|} \partial_{y_{i_2}} \mathsf{k}_v(y+z,y) \\ &= -\frac{z^{i_1}}{|z|} \Big((z^{i_3} \partial_{i_3} v)^{\otimes 2} : \nabla^{\otimes 2} \partial_{i_2} \mathsf{g}(z) + 2(\partial_{i_2} v(y) \otimes (z^{i_3} \partial_{i_3} v(y))) : \nabla^{\otimes 2} \mathsf{g}(z) \Big) \\ &+ O(\|v\|_{W^{2,\infty}}^2 \varepsilon^d), \end{aligned} \tag{6.22}$$

as $\varepsilon \to 0$, where to obtain the error bound we use assumption (6.9). So by Cauchy-Schwarz and assumption (6.9),

$$\begin{aligned}\left|\int_{\mathbb{R}^d}(f_1(y)\otimes g_1(y)) : \int_{\partial B(0,\varepsilon)} \frac{z}{|z|}\otimes \nabla_y \mathsf{k}_v(y+z,y)\, d\mathscr{H}^{d-1}(z)dy\right| &\lesssim \|\nabla v\|_{L^\infty}^2 \|f_1\|_{L^2}\|g_1\|_{L^2}\\ &\lesssim \|\nabla v\|_{L^\infty}^2 \|f\|_{\dot H^{-1}}\|g_1\|_{\dot H^{-1}}.\end{aligned} \tag{6.23}$$

This takes care of the first boundary term in the identity (3.28). The second boundary term vanishes as $\varepsilon \to 0$ by modifying (3.17), (3.18) to account for the new definition of k_v.

For the non-boundary term in (3.28), we see, using the identity (6.18), that it is equal to

$$\begin{aligned}&-\int_{|x-y|\geqslant\varepsilon} \nabla^{\otimes 2}\partial_{i_1 i_2}\mathrm{g}(x-y) : (v(x)-v(y))^{\otimes 2} f_1^{i_1}(x) g_1^{i_2}(y)\,dxdy\\ &-2\int_{|x-y|\geqslant\varepsilon} \nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y) : \partial_{i_1}v(x)\otimes (v(x)-v(y)) f_1^{i_1}(x) g_1^{i_2}(y)\,dxdy\\ &-2\int_{|x-y|\geqslant\varepsilon} \nabla^{\otimes 2}\partial_{i_1}\mathrm{g}(x-y) : (v(x)-v(y))\otimes \partial_{i_2}v(y) f_1^{i_1}(x) g_1^{i_2}(y)\,dxdy\\ &-2\int_{|x-y|\geqslant\varepsilon} \nabla^{\otimes 2}\mathrm{g}(x-y) : (\partial_{i_1}v(x)\otimes \partial_{i_2}v(y)) f_1^{i_1}(x) g_1^{i_2}(y)\,dxdy.\end{aligned} \tag{6.24}$$

By Taylor's theorem,

$$\begin{aligned}&\partial_{i_1}v(x)\otimes (v(x)-v(y)) : \nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y)\\ &= \partial_{i_1}v(x)\otimes\left(\int_0^1 \nabla v(y+t(x-y))\cdot(x-y)\,dt\right) : \nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y)\\ &=: \partial_{i_1}v(x)\otimes\left(\int_0^1 \partial_{i_3}v(y+t(x-y))\,dt\right) : \mathsf{k}_{i_2}^{i_3}(x-y),\end{aligned} \tag{6.25}$$

where for fixed i_2, i_3, the $(\mathbb{R}^d)^{\otimes 2}$-valued kernel $\mathsf{k}_{i_2}^{i_3}(x-y)$ is defined by

$$\mathsf{k}_{i_2}^{i_3}(x-y) := (x-y)^{i_3}\nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y), \qquad \forall x\neq y. \tag{6.26}$$

By assumption (6.11), $\mathsf{k}_{i_2}^{i_3}$ is associated to a Calderón-Zygmund operator. Hence, the expression

$$\left(\int_0^1 \partial_{i_3}v(y+t(x-y))\,dt\right)\cdot \mathsf{k}_{i_2}^{i_3}(x-y) \tag{6.27}$$

is the kernel of an $(\mathbb{R}^d)^{\otimes 2}$-valued Calderón d-commutator, which by the Christ-Journé theorem [CJ87] is bounded on $L^2(\mathbb{R}^d)$. So by Cauchy-Schwarz,

$$\begin{aligned}
&\left|\int_{|x-y|\geqslant\varepsilon} \nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y) : \partial_{i_1} v(x)\otimes(v(x)-v(y)) f_1^{i_1}(x) g_1^{i_2}(y)\,dx\,dy\right| \\
&\lesssim \|\nabla v\|_{L^\infty}\|\partial_{i_1} v\otimes f_1^{i_1}\|_{L^2}\|g_1\|_{L^2} \\
&\lesssim \|\nabla v\|_{L^\infty}^2\|f_1\|_{L^2}\|g_1\|_{L^2} \\
&\lesssim \|\nabla v\|_{L^\infty}^2\|f\|_{\dot H^{-1}}\|g\|_{\dot H^{-1}}. \qquad (6.28)
\end{aligned}$$

By symmetry under swapping the $i_1 \leftrightarrow i_2$, this takes care of the second and third lines of (6.24). For the fourth line, similarly using Taylor's theorem yields

$$\begin{aligned}
&\big(v(x)-v(y)\big)^{\otimes 2} : \nabla^{\otimes 2}\partial_{i_1 i_2}\mathrm{g}(x-y) \\
&= \left(\int_0^1 \nabla v(y+t(x-y))\cdot(x-y)\,dt\right)^{\otimes 2} : \nabla^{\otimes 2}\partial_{i_1 i_2}\mathrm{g}(x-y) \\
&=: \left(\int_0^1 \partial_{i_3} v(y+t(x-y))\,dt\right)\otimes\left(\int_0^1 \partial_{i_4} v(y+t(x-y))\,dt\right) : \mathsf{k}_{i_1 i_2}^{i_3 i_4}(x-y), \qquad (6.29)
\end{aligned}$$

where the $(\mathbb{R}^d)^{\otimes 2}$-valued kernel $\mathsf{k}_{i_1 i_2}^{i_3 i_4}$ is defined by

$$\mathsf{k}_{i_1 i_2}^{i_3 i_4}(x-y) := (x-y)^{\otimes 2, i_3 i_4}\nabla^{\otimes 2}\partial_{i_1 i_2}\mathrm{g}(x-y) \qquad \forall x\neq y. \qquad (6.30)$$

By assumption (6.11), $\mathsf{k}_{i_1 i_2}^{i_3 i_4}$ is associated to a Calderón-Zygmund operator. Hence, the expression (6.29) is the kernel of a Calderón d-commutator, and we can apply the Christ-Journé theorem [CJ87] like before to conclude that the modulus of the fourth line of (6.24) is bounded by

$$\|\nabla v\|_{L^\infty}^2\|f\|_{\dot H^{-1}}\|g\|_{\dot H^{-1}}. \qquad (6.31)$$

The sub-Coulombic case. Finally, in the sub-Coulombic case $0\leqslant s<d-2$, the induction argument proceeds exactly as in the proof of Theorem 3.1. Instead of identity (3.38), we now have

$$\begin{aligned}
&\int_{(\mathbb{R}^d)^2}\mathsf{k}_v(x,y)f(x)g(y)\,dx\,dy \\
&= \lim_{\varepsilon\to 0}\bigg(-\int_{|x-y|\geqslant\varepsilon}\nabla^{\otimes 2}\partial_{i_1 i_2}\mathrm{g}(x-y):\big(v(x)-v(y)\big)^{\otimes 2} f_1^{i_1}(x)g_1^{i_2}(y)\,dx\,dy \\
&\quad -2\int_{|x-y|\geqslant\varepsilon}\nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y):\partial_{i_1} v(x)\otimes\big(v(x)-v(y)\big) f_1^{i_1}(x)g_1^{i_2}(y)\,dx\,dy \qquad (6.32) \\
&\quad -2\int_{|x-y|\geqslant\varepsilon}\nabla^{\otimes 2}\partial_{i_1}\mathrm{g}(x-y):\big(v(x)-v(y)\big)\otimes\partial_{i_2} v(y) f_1^{i_1}(x)g_1^{i_2}(y)\,dx\,dy \\
&\quad -\int_{|x-y|\geqslant\varepsilon}\nabla^{\otimes 2}\mathrm{g}(x-y):\big(\partial_{i_1} v(x)\otimes\partial_{i_2} v(y)+\partial_{i_2} v(y)\otimes\partial_{i_1} v(x)\big) f_1^{i_1}(x)g_1^{i_2}(y)\,dx\,dy\bigg).
\end{aligned}$$

Fix i_1, i_2. For the first term, we can repeat repeat the proof above for the super-Coulombic ($d-4<s<d-2$) and Coulombic ($s=d-4$) cases with g replaced by $\partial_{i_1 i_2}\mathrm{g}$.

The second and third terms are symmetric, so it suffices to consider the second term. Writing

$$\begin{aligned}\nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y):\partial_{i_1}v(x)\otimes\big(v(x)-v(y)\big)&=\partial_{j_2}\partial_{j_1 i_2}\mathrm{g}(x-y)\partial_{i_1}v^{j_1}(x)\big(v(x)-v(y)\big)^{j_2}\\&=\nabla\partial_{j_1 i_2}\mathrm{g}(x-y)\cdot\big(v(x)-v(y)\big)\partial_{i_1}v^{j_1}(x)\end{aligned}\tag{6.33}$$

and therefore

$$\begin{aligned}&\lim_{\varepsilon\to 0}\int_{|x-y|\geqslant\varepsilon}\nabla^{\otimes 2}\partial_{i_2}\mathrm{g}(x-y):\partial_{i_1}v(x)\otimes\big(v(x)-v(y)\big)f_1^{i_1}(x)g_1^{i_2}(y)\,dxdy\\&=\lim_{\varepsilon\to 0}\int_{|x-y|\geqslant\varepsilon}\nabla\partial_{j_1 i_2}\mathrm{g}(x-y)\cdot\big(v(x)-v(y)\big)\partial_{i_1}v^{j_1}(x)f_1^{i_1}(x)g_1^{i_2}(y)\,dxdy,\end{aligned}\tag{6.34}$$

we see that we can apply Theorem 3.1 to this expression with g replaced by $\partial_{j_1 i_2}\mathrm{g}$, f replaced by $\partial_{i_1}v^{j_1}f_1^{i_1}$ and g replaced by $g_1^{i_2}$. Thus, its modulus is bounded by

$$\|\nabla v\|_{L^\infty}\|\partial_{i_1}v^{j_1}f_1^{i_1}\|_{L^2}\|g_1^{i_2}\|_{L^2}\lesssim\|\nabla v\|_{L^\infty}^2\|f\|_{\dot H^{-1}}\|g\|_{\dot H^{-1}}.\tag{6.35}$$

Finally, for the fourth term, $\nabla^{\otimes 2}\mathrm{g}$ is locally integrable, so we can use dominated convergence to pass to the limit $\varepsilon\to 0$. By symmetry under swapping $x\leftrightarrow y$, it suffices to consider the expression

$$\int_{(\mathbb{R}^d)^2}\nabla^{\otimes 2}\mathrm{g}(x-y):\big(\partial_{i_1}v(x)\otimes\partial_{i_2}v(y)\big)f_1^{i_1}(x)g_1^{i_2}(y)\,dxdy.\tag{6.36}$$

Writing $\mathrm{I}=|\nabla|^{\frac{d-s}{2}-1}|\nabla|^{\frac{s-d}{2}+1}$ and using Cauchy-Schwarz, the modulus of the preceding expression is bounded by

$$\||\nabla|^{\frac{d-s}{2}-1}(\nabla^{\otimes 2}\mathrm{g}*(\partial_{i_2}vg_1^{i_2}))\|_{L^2}\||\nabla|^{\frac{s-d}{2}+1}(\partial_{i_1}vf_1^{i_1})\|_{L^2}.\tag{6.37}$$

By Plancherel's theorem and assumption (6.10),

$$\||\nabla|^{\frac{d-s}{2}-1}(\nabla^{\otimes 2}\mathrm{g}*(\partial_{i_2}vg_1^{i_2}))\|_{L^2}\lesssim\||\nabla|^{\frac{s-d}{2}+1}(\partial_{i_2}vg_1^{i_2})\|_{L^2}.\tag{6.38}$$

Using (3.43) and (3.45), we conclude that

$$\begin{aligned}&\left|\int_{(\mathbb{R}^d)^2}\nabla^{\otimes 2}\mathrm{g}(x-y):\big(\partial_{i_1}v(x)\otimes\partial_{i_2}v(y)\big)f_1^{i_1}(x)g_1^{i_2}(y)\,dxdy\right|\\&\lesssim\left(\|\nabla v\|_{L^\infty}+\||\nabla|^{\frac{d-s}{2}}v\|_{L^{\frac{2d}{d-2-s}}}\right)^2\|f\|_{\dot H^{\frac{s-d}{2}}}\|g\|_{\dot H^{\frac{s-d}{2}}}.\end{aligned}\tag{6.39}$$

Just as before, the general case $d-2(k+1)\leqslant s<d-2k$ follows by induction on k, and the existence of the form extension B_v by density of $\mathscr{S}$ in $\dot H^{\frac{s-d}{2}}$. □

6.3. **Renormalization of second-order commutator estimate.** We now use the smearing procedure to obtain a renormalization of Theorem 6.1, which is the second-order analogue of Theorem 4.1.

Proposition 6.2. *Let $d \geqslant 1$, $0 \leqslant s < d$. Let $\underline{x}_N \in (\mathbb{R}^d)^N$ be pairwise distinct, and $\mu \in \mathscr{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. If $s = 0$, assume that $\int_{\mathbb{R}^d} \log(1+|x|) d\mu(x) < \infty$ and if $s \geqslant d-1$, assume that $\||\nabla|^{s-d}\nabla\mu\|_{L^\infty} < \infty$. Let be v a continuous vector field on $\mathbb{R}^d$ such that*

$$\begin{cases} \|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2} < \infty, & 0 \leqslant s \leqslant d-2 \\ \|v\|_{C^{1,\alpha}} + \|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)}, \frac{2(d+m)}{d+m-2-s}}} < \infty, & d-2 < s < d, \end{cases} \tag{6.40}$$

for some $0 < \alpha \leqslant 1$ and where m is the dimension of the extension space. There exists a constant C depending only d, s, α, m and on the potential g through assumptions (1.13) - (1.28)*, such that*

$$\begin{aligned} &\left| \int_{(\mathbb{R}^d)^2 \setminus \triangle} (v(x)-v(y))^{\otimes 2} \cdot \nabla^{\otimes 2} \mathsf{g}(x-y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) \right| \\ &\leqslant C \Bigg(\|\nabla v\|_{L^\infty} + \||\nabla|^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-2-s}}} \mathbf{1}_{s<d-2} + \left(\|v\|_{C^{1,\alpha}} + \|v\|_{W^{\frac{d(d+m-s)+2m}{2(d+m)}, \frac{2(d+m)}{d+m-2-s}}} \right) \mathbf{1}_{s>d-2} \Bigg)^2 \\ &\quad \times \Bigg(F_N(\underline{x}_N, \mu) + C(1+\|\mu\|_{L^\infty}) \Big(N^{-\frac{\min\{d-s,2\}}{\min\{d,s+2\}(s+1)}} + (N^{-1}\log N)\mathbf{1}_{s=0} \\ &\qquad + (N^{-\frac{2}{d(d-1)}} \log N)\mathbf{1}_{s=d-2} \Big) \Bigg) \\ &\quad + \left(N^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}} \|v\|_{C^{1,\alpha}}^2 \Big(\||\nabla|^{s+1-d}\mu\|_{L^\infty} \mathbf{1}_{s<d-1} + \||\nabla|^{s-d}\nabla\mu\|_{L^\infty} \mathbf{1}_{s\geqslant d-1} \Big) \right. \\ &\left. + \|v\|_{C^{1,\alpha}} \|\operatorname{div} v\|_{L^\infty} \left(\|\mu\|_{L^\infty} N^{-\frac{(d-s)\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}} + N^{-\frac{\min\{d+1,s+3\}}{\min\{d,s+2\}(s+1)}} \Big(\|\mu\|_{L^\infty} \log N + 1 \Big) \mathbf{1}_{s=d-1} \right) \right). \end{aligned} \tag{6.41}$$

Proof. We follow the outline of the proof of Theorem 4.1, focusing on what is different in the second-order setting. Let $\tilde{v}$ be the same extension of the vector field v as before. Recycling the kernel notation K_v, we define

$$\mathsf{K}_v(X,Y) := \nabla^{\otimes 2}\mathsf{G}(X-Y) : (\tilde{v}(X) - \tilde{v}(Y))^{\otimes 2} \qquad \forall X \neq Y \in \mathbb{R}^{d+m}. \tag{6.42}$$

We note that from assumption (1.24) that

$$|\mathsf{K}_v(X,Y)| \leqslant C\|\nabla\tilde{v}\|_{L^\infty}^2 \left(|X-Y|^{-s}\mathbf{1}_{s>0} + \mathbf{1}_{s=0} \right) \tag{6.43}$$

and

$$|\nabla_X \mathsf{K}_v(X,Y)| \leqslant C\|\nabla\tilde{v}\|_{L^\infty}^2 |X-Y|^{-s-1}. \tag{6.44}$$

We again let X_i denote $(x_i, 0)$.

Adding and subtracting the smeared point mass $\delta_{X_i}^{(\eta)}$, we obtain the decomposition

$$\int_{(\mathbb{R}^d)^2 \setminus \triangle} \mathsf{k}_v(x,y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) = \mathrm{Term}_1 + \mathrm{Term}_2 + \mathrm{Term}_3, \tag{6.45}$$

where $\mathrm{Term}_1, \ldots, \mathrm{Term}_3$ are defined as before. We proceed to estimate each of the Term_j individually.

Estimate for Term_1. We apply Theorem 6.1 with

$$f = g = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_i}^{(\eta)} - \tilde{\mu} \in \dot{H}^{\frac{s-d-m}{2}}(\mathbb{R}^{d+m}) \tag{6.46}$$

followed by Theorem 2.1 or Theorem 2.2 to obtain

$$\begin{aligned} |\mathrm{Term}_1| &\leq C\Big(\|\nabla\tilde{v}\|_{L^\infty} + \||\nabla|^{\frac{d+m-s}{2}}\tilde{v}\|_{L^{\frac{2(d+m)}{d+m-2-s}}}\mathbf{1}_{s<d+m-2}\Big)^2 \\ &\times\left(F_N(\underline{x}_N,\mu) + \frac{\mathsf{G}_\eta(0)}{N} + C\|\mu\|_{L^\infty}\Big(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0} + (\eta^2|\log\eta|)\mathbf{1}_{s=d-2}\Big) + C\eta^2\right). \end{aligned} \tag{6.47}$$

Estimate for Term_2. Suppose first that $0 \leq s < d-1$. We break Term_2 into a sum over the index i of terms

$$\int_{\mathbb{R}^{d+m}}\int_{\mathbb{R}^{d+m}} (\mathsf{K}_\nu(X,X_i) - \mathsf{K}_\nu(X,Y))\, d\delta_{X_i}^{(\eta)}(Y)\, d\tilde{\mu}(X). \tag{6.48}$$

Unpacking the definition (6.42) of K_ν, we have

$$\begin{aligned} \mathsf{K}_\nu(X,X_i) - \mathsf{K}_\nu(X,Y) &= \big(\nabla^{\otimes 2}\mathsf{G}(X-X_i) - \nabla^{\otimes 2}\mathsf{G}(X-Y)\big) : ((\tilde{v}(X)-\tilde{v}(Y))\otimes(\tilde{v}(X)-\tilde{v}(X_i))) \\ &+ \nabla^{\otimes 2}\mathsf{G}(X-Y) : ((\tilde{v}(X)-\tilde{v}(Y))\otimes(\tilde{v}(Y)-\tilde{v}(X_i))) \\ &+ \nabla^{\otimes 2}\mathsf{G}(X-X_i) : ((\tilde{v}(Y)-\tilde{v}(X_i))\otimes(\tilde{v}(X)-\tilde{v}(X_i))). \end{aligned} \tag{6.49}$$

By considering the cases $|X-X_i| \leq 2\eta$ and $|X-X_i| > 2\eta$, it follows from using (1.24), the mean-value theorem, and triangle inequality that

$$|\mathsf{K}_\nu(X,X_i) - \mathsf{K}_\nu(X,Y)| \leq C\eta\|\nabla\tilde{v}\|_{L^\infty}^2\big(|X-X_i|^{-s-1} + |Y-X_i|^{-s-1}\big) \tag{6.50}$$

for all $Y \in \mathrm{supp}(\delta_{X_i}^{(\eta)})$. So, we find that

$$|\mathrm{Term}_2|\mathbf{1}_{s<d-1} \leq C\eta\|\nabla\tilde{v}\|_{L^\infty}^2\||\nabla|^{s+1-d}\mu\|_{L^\infty}. \tag{6.51}$$

Now if $d-1 \leq s < d$, we integrate by parts in X to write

$$\begin{aligned} &\int_{(\mathbb{R}^{d+m})^2} \mathsf{K}_\nu(X,Y)\, d\Big(\delta_{X_i} - \delta_{X_i}^{(\eta)}\Big)(Y)\, d\tilde{\mu}(X) \\ &\qquad = -\int_{(\mathbb{R}^{d+m})^2} \tilde{\mathsf{K}}_\nu(X,Y)\cdot d\Big(\delta_{X_i} - \delta_{X_i}^{(\eta)}\Big)(Y)\, d\widetilde{\nabla\mu}(X) \\ &\qquad -2\int_{(\mathbb{R}^{d+m})^2} \mathrm{div}\,\tilde{v}(X)\nabla\mathsf{G}(X-Y)\cdot(\tilde{v}(X)-\tilde{v}(Y))\, d\Big(\delta_{X_i} - \delta_{X_i}^{(\eta)}\Big)(Y)\, d\tilde{\mu}(X), \end{aligned} \tag{6.52}$$

where we have defined the vector-valued kernel

$$\tilde{\mathsf{K}}_\nu(X,Y) := \nabla\mathsf{G}(X-Y)\cdot(\tilde{v}(X)-\tilde{v}(Y))^{\otimes 2}. \tag{6.53}$$

Proceeding as in the case $0 \leq s < d-1$, the first term in the right-hand side of (6.52) is bounded by

$$C\eta\|\nabla\tilde{v}\|_{L^\infty}^2\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}. \tag{6.54}$$

For the second term, we observe that

$$\begin{aligned}
&\int_{\mathbb{R}^{d+m}} (\tilde v(X)-\tilde v(Y))\cdot\nabla G(X-Y)\,d(\delta_{X_i}-\delta^{(\eta)}_{X_i})(Y)\\
&=\int_{\mathbb{R}^{d+m}} ((\tilde v(X)-\tilde v(X_i))\cdot\nabla \mathsf{G}(X-X_i)-(\tilde v(X)-\tilde v(Y))\cdot\nabla G(X-Y))\,d\delta^{(\eta)}_{X_i}(Y)\\
&=(\tilde v(X)-\tilde v(X_i))\cdot\big(\nabla\mathsf{G}(X-X_i)-\nabla\mathsf{G}_\eta(X-X_i)\big)\\
&\quad-\int_{\mathbb{R}^{d+m}} (\tilde v(X_i)-\tilde v(Y))\cdot\nabla\mathsf{G}(X-Y)\,d\delta^{(\eta)}_{X_i}(Y). \qquad (6.55)
\end{aligned}$$

If $|X-X_i|\leqslant 2\eta$, then it follows from the second line using the mean-value theorem and assumption (1.24) that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^{d+m}} (\tilde v(X)-\tilde v(Y))\cdot\nabla G(X-Y)\,d(\delta_{X_i}-\delta^{(\eta)}_{X_i})(Y)\right| \qquad (6.56)\\
&\qquad\leqslant C\|\nabla\tilde v\|_{L^\infty}\left(|X-X_i|^{-s}+\int_{\mathbb{R}^{d+m}}|X-Y|^{-s}\,d\delta^{(\eta)}_{X_i}(Y)\right).
\end{aligned}$$

If $|X-X_i|>2\eta$, then it follows from the third line using the mean-value theorem, assumption (1.24), and repeating the proof of (2.21) with G replaced by $\nabla\mathsf{G}$ that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^{d+m}} (\tilde v(X)-\tilde v(Y))\cdot\nabla G(X-Y)\,d(\delta_{X_i}-\delta^{(\eta)}_{X_i})(Y)\right| \qquad (6.57)\\
&\qquad\leqslant C\|\nabla\tilde v\|_{L^\infty}\big(\eta^2|X-X_i|^{-s-2}+\eta|X-X_i|^{-s-1}\big),
\end{aligned}$$

where we also use that the reverse triangle inequality and that $\delta^{(\eta)}_{X_i}$ is a probability measure. Combining (6.56) and (6.57), we obtain

$$\begin{aligned}
&\left|\int_{(\mathbb{R}^{d+m})^2} \operatorname{div}\tilde v(X)\nabla\mathsf{G}(X-Y)\cdot(\tilde v(X)-\tilde v(Y))\,d\Big(\delta_{X_i}-\delta^{(\eta)}_{X_i}\Big)(Y)\,d\tilde\mu(X)\right| \qquad (6.58)\\
&\leqslant C\|\nabla\tilde v\|_{L^\infty}\|\operatorname{div}\tilde v\|_{L^\infty}\bigg(\int_{|X-X_i|\leqslant 2\eta}|X-X_i|^{-s}\,d\tilde\mu(X)+\int_{|X-X_i|\leqslant 2\eta}\int_{\mathbb{R}^{d+m}}|X-Y|^{-s}\,d\delta^{(\eta)}_{X_i}(Y)\,d\tilde\mu(X)\\
&\qquad+\int_{|X-X_i|>2\eta}\big(\eta^2|X-X_i|^{-s-2}+\eta|X-X_i|^{-s-1}\big)\,d\tilde\mu(X)\bigg).
\end{aligned}$$

Evidently,

$$\int_{|X-X_i|\leqslant 2\eta}|X-X_i|^{-s}\,d\tilde\mu(X)=\int_{|x-x_i|\leqslant 2\eta}|x-x_i|^{-s}\,d\mu(x)\leqslant C\eta^{d-s}\|\mu\|_{L^\infty}. \qquad (6.59)$$

Similarly, if $|X-X_i|\leqslant 2\eta$ and $Y\in\operatorname{supp}(\delta^{(\eta)}_{X_i})$, then $|X-Y|\leqslant 3\eta$. So by direct majorization and Fubini-Tonelli,

$$\begin{aligned}
\int_{|X-X_i|\leqslant 2\eta}\int_{\mathbb{R}^{d+m}}|X-Y|^{-s}\,d\delta^{(\eta)}_{X_i}(Y)\,d\tilde\mu(X)&\leqslant\int_{\mathbb{R}^{d+m}}\int_{|X-Y|\leqslant 3\eta}|X-Y|^{-s}\,d\tilde\mu(X)\,d\delta^{(\eta)}_{X_i}(Y)\\
&\leqslant C\eta^{d-s}\|\mu\|_{L^\infty}. \qquad (6.60)
\end{aligned}$$

Lastly, if $s > d-1$, then by dilation invariance of Lebesgue measure,

$$\int_{|X-X_i|>2\eta} \left(\eta^2 |X-X_i|^{-s-2} + \eta |X-X_i|^{-s-1}\right) d\tilde{\mu}(X) \leqslant C\eta^{d-s}\|\mu\|_{L^\infty}. \tag{6.61}$$

If $s = d-1$, then we make the modification

$$\begin{aligned}\int_{|x-x_i|>2\eta} |x-x_i|^{-d} d\mu(x) &= \left(\int_{1\geqslant|x-x_i|>2\eta} |x-x_i|^{-d} d\mu(x) + \int_{|x-x_i|>1} |x-x_i|^{-d} d\mu(x)\right) \\ &\leqslant C\|\mu\|_{L^\infty} |\log\eta| + 1, \end{aligned} \tag{6.62}$$

where we use that μ is a probability density. Thus, we shown that

$$\begin{aligned}&\left|\int_{(\mathbb{R}^{d+m})^2} \operatorname{div}\tilde{v}(X)\nabla \mathsf{G}(X-Y)\cdot(\tilde{v}(X)-\tilde{v}(Y)) d\left(\delta_{X_i} - \delta_{X_i}^{(\eta)}\right)(Y) d\tilde{\mu}(X)\right| \\ &\qquad \leqslant C\|\nabla\tilde{v}\|_{L^\infty}\|\operatorname{div}\tilde{v}\|_{L^\infty}\left(\|\mu\|_{L^\infty}\eta^{d-s} + \eta\left(\|\mu\|_{L^\infty}|\log\eta|+1\right)\mathbf{1}_{s=d-1}\right).\end{aligned} \tag{6.63}$$

Putting together the estimates (6.54) and (6.63) with (6.51), we conclude that

$$\begin{aligned}|\mathrm{Term}_2| \leqslant C\Big(&\eta\|\nabla\tilde{v}\|_{L^\infty}^2 \||\nabla|^{s+1-d}\mu\|_{L^\infty}\mathbf{1}_{s<d-1} + \eta\|\nabla\tilde{v}\|_{L^\infty}^2\||\nabla|^{s-d}\nabla\mu\|_{L^\infty}\mathbf{1}_{s\geqslant d-1} \\ &+ \|\nabla\tilde{v}\|_{L^\infty}\|\operatorname{div}\tilde{v}\|_{L^\infty}\left(\|\mu\|_{L^\infty}\eta^{d-s} + \eta\left(\|\mu\|_{L^\infty}|\log\eta|+1\right)\mathbf{1}_{s=d-1}\right)\Big).\end{aligned} \tag{6.64}$$

Estimate for Term_3. This step proceeds as for Term_3 in the proof of Theorem 4.1, except now using the kernel bounds (6.43) and (6.44). Ultimately, we obtain

$$\begin{aligned}|\mathrm{Term}_3| \leqslant C\|\nabla\tilde{v}\|_{L^\infty}^2\Big(&F_N(\underline{x}_N,\mu) + C\Big(\frac{\eta^{-s}+|\log\eta|\mathbf{1}_{s=0}}{N} + \frac{\varepsilon^{-s}+|\log\varepsilon|\mathbf{1}_{s=0}}{N} + \varepsilon^2 \\ &+ \|\mu\|_{L^\infty}(\varepsilon^{d-s} + (\varepsilon^d|\log\varepsilon|)\mathbf{1}_{s=0} + (\varepsilon^2|\log\varepsilon|)\mathbf{1}_{s=d-2}) + \eta\varepsilon^{-s-1}\Big)\Big).\end{aligned} \tag{6.65}$$

Conclusion. Combining the estimates (4.14), (4.26), (4.34) obtained above, we have shown that

$$\begin{aligned}&\left|\int_{(\mathbb{R}^d)^2\setminus\Delta} \mathsf{k}_\nu(x,y) d\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x,y)\right| \\ &\leqslant C\left(\|\nabla\tilde{v}\|_{L^\infty} + \||\nabla|^{\frac{d+m-s}{2}}\tilde{v}\|_{L^{\frac{2(d+m)}{d+m-2-s}}}\mathbf{1}_{s<d+m-2}\right)^2\Big(F_N(\underline{x}_N,\mu) + \frac{\eta^{-s}+|\log\eta|\mathbf{1}_{s=0}}{N} \\ &\qquad + C\|\mu\|_{L^\infty}\left(\eta^{d-s} + (\eta^d|\log\eta|)\mathbf{1}_{s=0} + (\eta^2|\log\eta|)\mathbf{1}_{s=d-2}\right) + C\eta^2\Big) \\ &\quad + C\Big((\eta\|\nabla\tilde{v}\|_{L^\infty}^2\||\nabla|^{s+1-d}\mu\|_{L^\infty})\mathbf{1}_{s<d-1} + (\eta\|\nabla\tilde{v}\|_{L^\infty}^2\||\nabla|^{s-d}\nabla\mu\|_{L^\infty})\mathbf{1}_{s\geqslant d-1} \\ &\qquad + \|\nabla\tilde{v}\|_{L^\infty}\|\operatorname{div}\tilde{v}\|_{L^\infty}\left(\|\mu\|_{L^\infty}\eta^{d-s} + \eta\left(\|\mu\|_{L^\infty}|\log\eta|+1\right)\mathbf{1}_{s=d-1}\right)\Big)\end{aligned} \tag{6.66}$$

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

We may now optimize over η and ε exactly as in (4.36). Using the extension bounds (4.4) and noting that $\operatorname{div}\tilde{v}(x,0)=\operatorname{div}v(x)$, since $\tilde{v}$ vanishes in the $(m+1)$-th through $(m+d)$-th components, we arrive at the stated inequality. □

REFERENCES

[BFM16] Zdzisław Brzeźniak, Franco Flandoli, and Mario Maurelli. Existence and uniqueness for stochastic 2D Euler flows with bounded vorticity. *Arch. Ration. Mech. Anal.*, 221(1):107–142, 2016.

[BHS19] Sergiy V. Borodachov, Douglas P. Hardin, and Edward B. Saff. *Discrete energy on rectifiable sets*. Springer Monographs in Mathematics. Springer, New York, 2019.

[BJS22] Didier Bresch, Pierre-Emmanuel Jabin, and Juan Soler. A new approach to the mean-field limit of Vlasov-Fokker-Planck equations. *arXiv preprint arXiv:2203.15747*, 2022.

[BJW19a] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Modulated free energy and mean field limit. *Séminaire Laurent Schwartz–EDP et applications*, pages 1–22, 2019.

[BJW19b] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak-Keller-Segel model. *C. R. Math. Acad. Sci. Paris*, 357(9):708–720, 2019.

[BJW20] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Mean-field limit and quantitative estimates with singular attractive kernels. *arXiv preprint arXiv:2011.08022*, 2020.

[BO19] Robert J. Berman and Magnus Önnheim. Propagation of chaos for a class of first order models with singular mean field interactions. *SIAM J. Math. Anal.*, 51(1):159–196, 2019.

[Cal80] A.-P. Calderón. Commutators, singular integrals on Lipschitz curves and applications. In *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, pages 85–96. Acad. Sci. Fennica, Helsinki, 1980.

[CCC+12] Dongho Chae, Peter Constantin, Diego Córdoba, Francisco Gancedo, and Jiahong Wu. Generalized surface quasi-geostrophic equations with singular velocities. *Comm. Pure Appl. Math.*, 65(8):1037–1066, 2012.

[CCH14] José Antonio Carrillo, Young-Pil Choi, and Maxime Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, pages 1–46. Springer, Vienna, 2014.

[CF16] Michele Coghi and Franco Flandoli. Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.*, 26(3):1407–1442, 2016.

[CFP12] José A. Carrillo, Lucas C. F. Ferreira, and Juliana C. Precioso. A mass-transportation approach to a one dimensional fluid mechanics model with nonlocal velocity. *Adv. Math.*, 231(1):306–327, 2012.

[CJ87] Michael Christ and Jean-Lin Journé. Polynomial growth estimates for multilinear singular integral operators. *Acta Math.*, 159(1-2):51–80, 1987.

[CM20] Michele Coghi and Mario Maurelli. Regularized vortex approximation for 2D Euler equations with transport noise. *Stoch. Dyn.*, 20(6):2040002, 27, 2020.

[CS07] Luis Caffarelli and Luis Silvestre. An extension problem related to the fractional Laplacian. *Comm. Partial Differential Equations*, 32(7-9):1245–1260, 2007.

[DNPV12] Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[Dob79] R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979.

[DRAW02] Thierry Dauxois, Stefano Ruffo, Ennio Arimondo, and Martin Wilkens. Dynamics and thermodynamics of systems with long-range interactions: an introduction. In *Dynamics and thermodynamics of systems with long-range interactions (Les Houches, 2002)*, volume 602 of *Lecture Notes in Phys.*, pages 1–19. Springer, Berlin, 2002.

[Due16] Mitia Duerinckx. Mean-field limits for some Riesz interaction gradient flows. *SIAM Journal on Mathematical Analysis*, 48(3):2269–2300, 2016.

[FGP11] Franco. Flandoli, Massimiliano Gubinelli, and Enrico Priola. Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations. *Stochastic Process. Appl.*, 121(7):1445–1463, 2011.

[FHM14] Nicolas Fournier, Maxime Hauray, and Stéphane Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014.

[Gol16] François Golse. 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*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, 2016.

[Gra14a] Loukas Grafakos. *Classical Fourier Analysis*. Number 249 in Graduate Texts in Mathematics. Springer, third edition, 2014.

[Gra14b] Loukas Grafakos. *Modern Fourier Analysis*. Number 250 in Graduate Texts in Mathematics. Springer, third edition, 2014.

[Hau09] Maxime Hauray. Wasserstein distances for vortices approximation of Euler-type equations. *Math. Models Methods Appl. Sci.*, 19(8):1357–1384, 2009.

[HM14] Maxime Hauray and Stéphane Mischler. On Kac's chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014.

[HSSS18] Mahir Hadžić, Andreas Seeger, Charles K. Smart, and Brian Street. Singular integrals and a problem on mixing flows. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(4):921–943, 2018.

[HSST21] Douglas Hardin, Edward B. Saff, Ruiwen Shu, and Eitan Tadmor. Dynamics of particles on a curve with pairwise hyper-singular repulsion. *Discrete Contin. Dyn. Syst.*, 41(12):5509–5536, 2021.

[Jab14] Pierre-Emmanuel Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014.

[JW16] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.*, 271(12):3588–3627, 2016.

[JW18] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.

[KP88] Tosio Kato and Gustavo Ponce. Commutator estimates and the Euler and Navier-Stokes equations. *Comm. Pure Appl. Math.*, 41(7):891–907, 1988.

[KPV93] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle. *Comm. Pure Appl. Math.*, 46(4):527–620, 1993.

[Lac21] Daniel Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *arXiv preprint arXiv:2105.02983*, 2021.

[Lai20] Xudong Lai. Multilinear estimates for Calderón commutators. *Int. Math. Res. Not. IMRN*, (20):7097–7138, 2020.

[Li19] Dong Li. On Kato-Ponce and fractional Leibniz. *Rev. Mat. Iberoam.*, 35(1):23–100, 2019.

[LLY19] Lei Li, Jian-Guo Liu, and Pu Yu. On the mean field limit for Brownian particles with Coulomb interaction in 3D. *J. Math. Phys.*, 60(11):111501, 34, 2019.

[LSZ17] Thomas Leblé, Sylvia Serfaty, and Ofer Zeitouni. Large deviations for the two-dimensional two-component plasma. *Comm. Math. Phys.*, 350(1):301–360, 2017.

[LY16] Jian-Guo Liu and Rong Yang. Propagation of chaos for large Brownian particle system with Coulomb interaction. *Res. Math. Sci.*, 3:Paper No. 40, 33, 2016.

[Mar08] Fabien Marchand. Existence and regularity of weak solutions to the quasi-geostrophic equations in the spaces L^p or $\dot{H}^{1/2}$. *Comm. Math. Phys.*, 277(1):45–67, Jan 2008.

[Osa87a] Hirofumi Osada. Limit points of empirical distributions of vortices with small viscosity. In *Hydrodynamic behavior and interacting particle systems (Minneapolis, Minn., 1986)*, volume 9 of *IMA Vol. Math. Appl.*, pages 117–126. Springer, New York, 1987.

[Osa87b] Hirofumi Osada. Propagation of chaos for the two-dimensional Navier-Stokes equation. In *Probabilistic methods in mathematical physics (Katata/Kyoto, 1985)*, pages 303–334. Academic Press, Boston, MA, 1987.

[PS17] Mircea Petrache and Sylvia Serfaty. Next order asymptotics and renormalized energy for Riesz interactions. *J. Inst. Math. Jussieu*, 16(3):501–569, 2017.

[Ros20] Matthew Rosenzweig. The mean-field limit of stochastic point vortex systems with multiplicative noise. *arXiv preprint arXiv:2011.12180*, 2020.

[Ros22a] Matthew Rosenzweig. The mean-field approximation for higher-dimensional Coulomb flows in the scaling-critical L^∞ space. *Nonlinearity*, 35(6):2722–2766, may 2022.

[Ros22b] Matthew Rosenzweig. Mean-field convergence of point vortices to the incompressible Euler equation with vorticity in L^∞. *Arch. Ration. Mech. Anal.*, 243(3):1361–1431, 2022.

[RS21] Matthew Rosenzweig and Sylvia Serfaty. Global-in-time mean-field convergence for singular Riesz-type diffusive flows. *arXiv preprint arXiv:2108.09878*, 2021. To appear in Ann. Appl. Prob.

[Ser17] Sylvia Serfaty. Mean field limits of the Gross-Pitaevskii and parabolic Ginzburg-Landau equations. *J. Amer. Math. Soc.*, 30(3):713–768, 2017.

[Ser20a] Sylvia Serfaty. Gaussian fluctuations and free energy expansion for 2d and 3d Coulomb gases at any temperature. *arXiv preprint arXiv:2003.11704*, 2020. To appear in Annales IHP, Probabilités et Statistiques.

[Ser20b] Sylvia Serfaty. Mean field limit for Coulomb-type flows. *Duke Math. J.*, 169(15):2887–2935, 10 2020. Appendix with Mitia Duerinckx.

[SSS19] Andreas Seeger, Charles K. Smart, and Brian Street. Multilinear singular integral forms of Christ-Journé type. *Mem. Amer. Math. Soc.*, 257(1231):v+134, 2019.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[Tri10] Hans Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010. Reprint of 1983 edition, Also published in 1983 by Birkhäuser Verlag.

QUOC-HUNG NGUYEN, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES (CAS), BEIJING, CHINA

Email address: qhnguyen@amss.ac.cn

MATTHEW ROSENZWEIG, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, DEPARTMENT OF MATHEMATICS, CAMBRIDGE, MA

Email address: mrosenzw@mit.edu

SYLVIA SERFATY, COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK CITY, NY

Email address: serfaty@cims.nyu.edu