

Journées

ÉQUATIONS AUX DÉRIVÉES PARTIELLES

Aussois, 2–6 juin 2025

Matthew Rosenzweig

Commutators, mean-field, and supercritical mean-field limits for Coulomb/Riesz gases

J. É. D. P. (2025), Exposé n° VII, 32 p.

<<https://doi.org/10.5802/jedp.698>>



*Article mis en ligne dans le cadre du
Centre Mersenne pour l'édition scientifique ouverte
<http://www.centre-mersenne.org/>*

RÉSEAU THÉMATIQUE AEDP DU CNRS

Commutators, mean-field, and supercritical mean-field limits for Coulomb/Riesz gases

Matthew Rosenzweig

*Commutateurs, limite de champ moyen et limite de champ moyen supercritique pour
les gaz de Coulomb/Riesz*

Résumé

Cet article est un complément à l'exposé de l'auteur intitulé « Commutator estimates and mean-field limits for Coulomb/Riesz gases », présenté aux *Journées Équations aux Dérivées Partielles* 2025 à Aussois. L'objectif est de fournir une présentation concise et accessible des estimations fines de commutateurs récemment obtenues pour les énergies modulées associées aux interactions de type Coulomb/Riesz, ainsi que d'expliquer comment ces estimations conduisent à des résultats optimaux pour les limites de champ moyen et de champ moyen supercritique de la dynamique des gaz de Coulomb/Riesz, via la méthode de l'énergie modulée. L'exposé s'appuie principalement sur les travaux [129, 131] avec Serfaty et [84] avec Hess-Childs et Serfaty.

Abstract

This paper is intended as a companion to the author's talk "Commutator estimates and mean-field limits for Coulomb/Riesz gases" at the 2025 *Journées équations aux dérivées partielles* in Aussois. The goal is to provide a concise, accessible account of sharp commutator estimates recently obtained for modulated energies associated to Coulomb/Riesz interactions and how these estimates lead to optimal results for mean-field and supercritical mean-field limits of Coulomb/Riesz gas dynamics via the modulated-energy method. The exposition centers on the works [129, 131] with Serfaty and [84] with Hess-Childs and Serfaty.

1. Introduction

When studying systems of N distinct points $X_N = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ with interaction energy

$$\sum_{1 \leq i \neq j \leq N} g(x_i, x_j), \quad (1.1)$$

one is led to comparing the sequence of empirical measures $\mu_N := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ to an average, or *mean-field*, density μ . This comparison is conveniently performed by considering a *modulated energy*, or (squared) "distance" between μ_N and μ , defined by

$$F_N(X_N, \mu) := \frac{1}{2} \int_{(\mathbb{R}^d)^2 \setminus \Delta} g(x, y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)(x) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)(y), \quad (1.2)$$

M.R. is supported by NSF grants DMS-2441170, DMS-2345533, DMS-2342349.

2020 *Mathematics Subject Classification*: 35Q70, 35Q35, 35Q83, 82C22, 82C70, 76B47.

Keywords: Commutator estimates, modulated energy, mean-field limit, Riesz interactions, supercritical mean-field limit.

where $\Delta \subset (\mathbb{R}^d)^2$ denotes the diagonal. Physically, the quantity corresponds to the total energy of a system of N discrete charges located at the positions x_i neutralized by a background charge distribution μ with the infinite self-interaction of each charge removed, the latter corresponding to the excision of the diagonal.

The quantity (1.2) originated in the study of the statistical mechanics of Coulomb/Riesz gases [116, 133, 135, 136] and was later used in the derivation of mean-field dynamics [39, 112, 138] and following works. The notion of modulated energy to prove a stability estimate is classical, with the term first appearing to this author’s knowledge in the influential work of Brenier [13]. In the nonsingular case, where the diagonal may be reinserted, the modulated energy coincides with the square of what is known in the statistics literature as *maximum mean discrepancy (MMD)* [58, 59, 60], a type of integral probability metric [110]. MMDs for varying choices of kernels are widely used in statistics and machine learning contexts as distances between probability measures (see, e.g., [81, 93, 108]).

An essential point in this comparison is to control derivatives of F_N along a transport $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$, i.e. the quantities

$$\begin{aligned} \frac{d^n}{dt^n} \Big|_{t=0} F_N((\mathbb{I} + tv)^{\oplus N}(X_N), (\mathbb{I} + tv)\#\mu) \\ = \frac{1}{2} \int_{(\mathbb{R}^d)^2 \setminus \Delta} \nabla_x^{\otimes n} \mathbf{g}(x, y) : (v(x) - v(y))^{\otimes n} d(\mu_N - \mu)^{\otimes 2}, \end{aligned} \quad (1.3)$$

where $\mathbb{I} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the identity, $(\mathbb{I} + tv)^{\oplus N}(X_N) := (x_1 + tv(x_1), \dots, x_N + tv(x_N))$, and $:$ denotes the inner product between the tensors. Here, $\nabla^{\otimes n}$ denotes the n -tensor with components $(\partial_{\bar{\alpha}})_{\bar{\alpha} \in [d]^n}$.

For applications to the mean-field limit, v is the velocity field of the limiting dynamics as $N \rightarrow \infty$. For applications to central limit theorems (CLTs) for the fluctuations (the next-order description), v is the gradient of a test function evolved along the adjoint linearized mean-field flow [27, 86], and these inequalities are at the core of the “transport” approach to fluctuations of canonical Gibbs ensembles [5, 98, 115, 139] (cf. loop/Dyson–Schwinger equations in random matrix theory, e.g. [4, 9, 10]).

The desired control is a functional inequality of the form

$$|(1.3)| \leq C_1(F_N(X_N, \mu) + C_2 N^{-\alpha}), \quad (1.4)$$

where $\alpha \in (0, \infty]$ depends on $\mathbf{d}, \mathbf{g}, C_1 > 0$ is a constant depending on $\mathbf{d}, \mathbf{g}, n, v$, and $C_2 > 0$ depends on $\mathbf{d}, n, \mathbf{g}, \mu$. As observed in [121], the expressions (1.3) may be viewed as the quadratic form of a *commutator*, akin to the famous Calderón commutator [24, 30, 31, 137]. Indeed, ignoring the exclusion of the diagonal,

$$\int_{(\mathbb{R}^d)^2} (v(x) - v(y)) \cdot \nabla_x \mathbf{g}(x, y) f(x) g(y) dx dy = \left\langle f, \left[v \cdot, \nabla \mathbf{g}^* \right] g \right\rangle_{L^2}, \quad (1.5)$$

where $\langle \cdot, \cdot \rangle_{L^2}$ denotes the L^2 inner product. For this reason, estimates of the form (1.4) are called *commutator estimates*. The expressions (1.3) also have a *stress-energy tensor* structure, a point on which we elaborate in Section 2.1.

The discussion so far has concerned a general (conditionally) positive definite kernel \mathbf{g} . To have any hope of showing a functional inequality of the form (1.4), one needs to impose some assumptions on \mathbf{g} . Of particular interest and relevance for applications are singular interactions, model examples of which are the *log/Riesz potentials*¹ given by

$$\mathbf{g}(x, y) = \mathbf{g}(x - y) := \begin{cases} \frac{1}{s} |x - y|^{-s}, & s \neq 0 \\ -\log |x - y|, & s = 0. \end{cases} \quad (1.6)$$

We consider any dimension $\mathbf{d} \geq 1$ and assume that $-2 < s < \mathbf{d}$. Up to a normalizing constant $c_{\mathbf{d}, s}$, the potential \mathbf{g} is the fundamental solution of the fractional Laplacian $(-\Delta)^{\frac{\mathbf{d}-s}{2}} = |\nabla|^{\mathbf{d}-s}$, i.e. $|\nabla|^{\mathbf{d}-s} \mathbf{g} = c_{\mathbf{d}, s} \delta_0$. The cases $0 \leq s < \mathbf{d}$ and $-2 < s < 0$ are respectively referred to as the *singular* and *nonsingular* regimes because of the behavior of \mathbf{g} at the origin. The restriction $s > -2$ is natural

¹The log case being the $s \rightarrow 0$ limit of the Riesz case, the term “Riesz” will always include the log case.

because the Riesz potential ceases to be conditionally positive definite at $s = -2$.² The restriction to the *potential* regime $s < d$, in which g is locally integrable, is to exclude the *hypersingular* case, which is not of the mean-field type considered in this paper (e.g. see [72, 73]). The case $s = d - 2$ corresponds to the classical *Coulomb* potential from electrostatics/gravitation. Based on this threshold, it is convenient to call the interaction *sub-Coulomb* if $s < d - 2$ and *super-Coulomb* if $s > d - 2$. More generally, long-range interactions of the form (1.6) play an important role in physics [36], approximation theory [8], and machine learning [1, 65, 80, 81], just to list a few areas. Motivations for considering systems of the form (3.1) are extensively reviewed in the lecture notes [140].

Specializing to Riesz potentials and focusing on the first-order $n = 1$ case, the control (1.4) was first proved by Leblé and Serfaty [98] in the 2D Coulomb case, then generalized by Serfaty to the super-Coulomb case [138]. The proof relied on the identification of the aforementioned stress-energy tensor structure in (1.3) and using integration by parts. Subsequently, the author [121] used the commutator structure in (1.3) to give a new proof in the Coulomb case. This commutator perspective was then used together with Q.H. Nguyen and Serfaty [112] to show the desired functional inequality for the full singular regime as well as a broader class of Riesz-type interactions including Lennard–Jones-type potentials. The argument of [112] has also been extended to the nonsingular regime in the joint work [132] with Slepčev and Wang. As recognized in [131], the stress-energy and commutator perspectives are in fact two sides of the same coin. These functional inequalities are crucial not only for proving CLTs for fluctuations of Riesz gases [98, 115, 139], but also for deriving mean-field limits [33, 112, 118, 123, 124, 131, 138] and large deviation principles [82]. They have also been used to show the joint classical and mean-field limits for quantum systems of particles [52], and the supercritical mean-field limits of classical [69, 106, 125, 129] and quantum systems of particles [117, 122]. The reader will see two applications, mean-field and supercritical mean-field limits, below.

The main problem for these functional inequalities has been to determine the optimal size of the additive error, i.e. the exponent α in (1.4), which a priori depends on d and s . In formulating optimality, the appropriate comparison is with the minimal size of $F_N(X_N, \mu)$. When $0 \leq s < d$, it is known that $F(X_N, \mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2dN} \mathbf{1}_{s=0} \geq -C\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}$ for $C = C(d, s) > 0$ and this lower bound is sharp [84, 131]. The log term in the $s = 0$ case is *not* an additive error like $\|\mu\|_{L^\infty}^{\frac{s}{d}} N^{\frac{s}{d}-1}$, but rather the right quantity to consider is $F(X_N, \mu) + \frac{\log(N\|\mu\|_{L^\infty})}{2dN} \mathbf{1}_{s=0}$. In controlling the magnitude of (1.3), one needs a right-hand side which is also nonnegative. Thus, the best error one may hope for is of size $N^{\frac{s}{d}-1}$. When $-2 < s < 0$, the modulated energy is nonnegative (it is a squared MMD as mentioned above), and its minimal value is $\propto N^{\frac{s}{d}-1}$ [83]. Moreover, the functional inequality holds without additive error because the interaction is nonsingular [132].

Only very recently has the program to show the sharp $N^{\frac{s}{d}-1}$ error rate been completed by the author and collaborators: first, the Coulomb case [98, 125, 139], then the entire (super-)Coulomb case [131], and finally the entire singular Riesz case [84]. The completion of this program has yielded the sharp rates of convergence for the mean-field limit in the modulated energy distance, as well as rates for supercritical mean-field limits under optimal scaling assumptions. It has also been crucial for studying fluctuations around the mean-field limit in and out of thermal equilibrium [86, 98, 115, 120, 126, 139].

In the remainder of this paper, we describe the results and applications advertised in the preceding discussion. The body is organized as follows. In Section 2, we present our sharp commutator estimates in the form of Theorem 2.1, discuss their proof (Sections 2.1 and 2.2), followed by localized and higher-order variants Theorems 2.5 and 2.6 (Section 2.3), and conclude with a discussion of transport regularity and defective estimates (Section 2.4). In Section 3, we turn to mean-field limits, recalling the background and main result Theorem 3.1 (Sections 3.1 and 3.2), and then explaining the role of commutator estimates and modulated energy in the proof of Theorem 3.1 (Section 3.3), including the regularity assumptions (Section 3.4), and end with some comments on the positive-temperature generalization known as modulated free energy (Section 3.5). Finally, Section 4 is devoted to supercritical mean-field limits: we review the combined mean-field/quasineutral

²We say that a kernel $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is *conditionally positive definite (CPD)* iff for any signed Borel measure ρ with zero mass, it holds that $\int_{(\mathbb{R}^d)^2} k(x, y) d\rho(x) d\rho(y) \geq 0$.

regime (Section 4.1); discuss scaling assumptions and prior work (Section 4.3); present an informal version of the main result Theorem 4.1 (Section 4.5); outline the method of proof; and close with comments on optimality (Section 4.6) and extensions to the sub-Coulomb (Section 4.7) and non-monokinetic regimes (Section 4.8).

Acknowledgments

The author thanks the organizers of the 2025 *Journées équations aux dérivées partielles* for the opportunity to present his research in the stimulating and scenic environment of Aussois, and he thanks the staff of the Centre Paul Langevin for their gracious hospitality. He also thanks the Institute for Computational and Experimental Research in Mathematics (ICERM) for its hospitality, where part of the research for this project was carried out during the Fall 2021 semester program “Hamiltonian Methods in Dispersive and Wave Evolution Equations,” as well as the Courant Institute of Mathematical Sciences at NYU for their hospitality during his visits in November 2021 and April 2024, where part of the reported research was carried out. Finally, and most importantly, the author thanks his collaborators Sylvia Serfaty and Elias Hess-Childs, without whom the works described here would not have come to fruition.

2. Commutator estimates

We elaborate on the commutator estimates for Riesz potentials highlighted in the introduction. To compactify the notation, given a point configuration X_N , reference measure μ , vector field v , and $n \in \mathbb{N}$, let us write

$$A_n[X_N, \mu, v] := \int_{(\mathbb{R}^d)^2 \setminus \Delta} \nabla \mathbf{g}(x-y) \cdot (v(x) - v(y)) d \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2}(x, y). \quad (2.1)$$

Summarizing the state of the art for global, first-order commutator estimates is the following theorem, combining the results of [84, 112, 127, 131, 132]. In particular, taking $p = \infty$ in (2.4), (2.5) below yields the announced sharp $N^{\frac{s}{d}-1}$ additive error. As commented above, there is no additive error in the nonsingular case (2.7). Comparing the sub-Coulomb estimates (2.5) and (2.6), the transport regularity in the latter is optimal even though its additive error is not, while the additive error in the former is optimal even though the transport regularity is not. We discuss this point more in Section 2.4. Later, in Section 2.3, we discuss generalizations of Theorem 2.1 to localized and higher-order commutators.

Theorem 2.1. *Let $-2 < s < d$. Let $\mu \in L^1 \cap L^p$ for $\frac{d}{d-s} < p \leq \infty$ with $\int_{\mathbb{R}^d} d\mu = 1$, and associated to μ , define the length scales*

$$\lambda := (N \|\mu\|_{L^p})^{-\frac{1}{d}}, \quad (2.2)$$

$$\kappa := (N^{\frac{1}{s+1}} \|\mu\|_{L^p})^{-\frac{1}{d}}, \quad (2.3)$$

When $-2 < s \leq 0$, assume that $\int_{(\mathbb{R}^d)^2} |\mathbf{g}|(x-y) d\mu^{\otimes 2} < \infty$ to ensure that the modulated energy is finite. Let v be a Lipschitz vector field. Then for any pairwise distinct point configuration $X_N \in (\mathbb{R}^d)^N$, the following hold.

[131] *If $s \geq \max(0, d-2)$, then*

$$|A_1[X_N, \mu, v]| \leq C \|\nabla v\|_{L^\infty} \left(F_N(X_N, \mu) - \frac{\log \lambda}{2N} \mathbf{1}_{s=0} + C_p \|\mu\|_{L^p} \lambda^{\frac{d(p-1)}{p}-s} \right). \quad (2.4)$$

[84] *If $0 \leq s < d-2$, then for any $a \in (d, d+2)$,*

$$|A_1[X_N, \mu, v]| \leq C (\|\nabla v\|_{L^\infty} + \|\nabla v\|_{L^{\frac{2d}{d-2s}}}^{\frac{3}{2}}) \left(F_N(X_N, \mu) - \frac{\log \lambda}{2N} \mathbf{1}_{s=0} + C_p \|\mu\|_{L^p} \lambda^{\frac{d(p-1)}{p}-s} \right), \quad (2.5)$$

[112, 127] and for any $\frac{d}{d-s-1} < p \leq \infty$,

$$|A_1[X_N, \mu, v]| \leq C \left(\|\nabla v\|_{L^\infty} + \|\nabla v\|_{L^{\frac{d-s}{2}}}^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-s-2}}} \mathbf{1}_{s < d-2} \right) \times \left(F_N(X_N, \mu) + C_p \|\mu\|_{L^p} \kappa^{\frac{d(p-1)}{p}-s} (1 - (\log \kappa) \mathbf{1}_{s=0}) \right). \quad (2.6)$$

[132] If $-2 < s < 0$, then

$$|A_1[X_N, \mu, v]| \leq C \left(\|\nabla v\|_{L^\infty} + \|\nabla v\|_{L^{\frac{d-s}{2}}}^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-s-2}}} \mathbf{1}_{s < d-2} \right) F_N(X_N, \mu) \quad (2.7)$$

In all cases, the constant $C > 0$ depends only on d, s and the constants $C_a, C_p > 0$ additionally depend on a, p , respectively.

Remark 2.2. Here, optimality is understood in the “worst case” sense, i.e. that there exists a point configuration X_N that achieves the bound. This should be contrasted with the “average” sense, where $X_N \sim f_N \in \mathcal{P}((\mathbb{R}^d)^N)$ and one estimates $\mathbb{E}_{X_N \sim f_N} [(1.3)]$. It is well-known that averaging may produce smaller errors. For instance, if $f_N \sim \mu^{\otimes N}$ (i.e. μ -iid), then $\mathbb{E}_{X_N \sim f_N} [(1.3)] \approx N^{-1}$. More generally, one can prove *high-temperature/entropic* commutator estimates for bounded interactions [90, 102] and even some Riesz cases [37].

Remark 2.3. Theorem 2.1 (and its proof) hold mutatis mutandis on the flat torus \mathbb{T}^d where \mathbf{g} is now taken to be the solution of $|\nabla|^{d-s} \mathbf{g} = c_{d,s}(\delta_0 - 1)$.

Given a measure ν with density in $\mathcal{S}(\mathbb{R}^d)$ and $\int_{\mathbb{R}^d} d\nu = 1$, letting $f := \nu - \mu$, we can find a sequence of pairwise distinct point configurations X_N such that, as $N \rightarrow \infty$,

$$F_N(X_N, \mu) \rightarrow \frac{1}{2} \int_{(\mathbb{R}^d)^2} \mathbf{g}(x-y) f(x) f(y) dx dy = \frac{C_{d,s}}{2} \|f\|_{\dot{H}^{\frac{s-d}{2}}}^2 \quad (2.8)$$

and $A_1[X_N, \mu, v] \rightarrow$

$$\frac{1}{2} \int_{(\mathbb{R}^d)^2} \nabla \mathbf{g}(x-y) \cdot (v(x) - v(y)) f(x) f(y) dx dy = \frac{1}{2} \left\langle f, [v \cdot, \nabla |\nabla|^{s-d}] f \right\rangle_{L^2}. \quad (2.9)$$

Thus, a commutator estimate with an additive error involving the singular centered measure $\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu$ implies a commutator estimate without an additive error for the regular centered measure f . More precisely, Theorem 2.1 implies the following standard or *unrenormalized* commutator estimate (see [112, Proposition 3.1]).

Proposition 2.4. Let $d \geq 1$ and $-2 < s < d$. Then there exists $C > 0$, depending only on d and s , such that for all $f, g \in \mathcal{S}(\mathbb{R}^d)$,³ we have

$$\left| \int_{(\mathbb{R}^d)^2} (v(x) - v(y)) \cdot \nabla \mathbf{g}(x-y) f(x) g(y) dx dy \right| \leq C \left(\|\nabla v\|_{L^\infty} + \|\nabla v\|_{L^{\frac{d-s}{2}}}^{\frac{d-s}{2}} v\|_{L^{\frac{2d}{d-s-2}}} \mathbf{1}_{s < d-2} \right) \|f\|_{\dot{H}^{\frac{s-d}{2}}} \|g\|_{\dot{H}^{\frac{s-d}{2}}}. \quad (2.10)$$

In fact, as we describe below, the philosophy of [112, 121, 131], and to a more implicit degree [39, 98, 138, 139], proceeds in the reverse order: first show the unrenormalized commutator estimate of Proposition 2.4, then combine it with a renormalization procedure to allow for the irregular test measure $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$. It is this renormalization step, which is the most technically involved part of the proof, that leads to the additive N -dependent errors in the singular case $0 \leq s < d$. For this reason, the estimates (2.4), (2.5), (2.6) are sometimes referred to as *renormalized* commutator estimates to distinguish them from the unrenormalized estimate (2.10). In the nonsingular case $-2 < s < 0$, no renormalization is necessary, and one can apply a commutator estimate directly, leading to the absence of additive errors in (2.7).

We will only discuss the proofs of the (super-)Coulomb estimate (2.4) (Section 2.1) and the sub-Coulomb estimate (2.5) (Section 2.2), as these are two sharp, singular cases. The reader interested in the proof of the nonsharp sub-Coulomb estimate (2.6) may consult [112]. The estimate (2.7), shown in [132], follows from a straightforward extension of the method of [112].

³If $s = 0$, then we implicitly assume f, g are zero-mean to avoid the low-frequency divergence in $\dot{H}^{-\frac{d}{2}}$.

2.1. Coulomb/super-Coulomb case

Following the chronology of the literature, we start with the (super-)Coulomb case $s \geq \max(0, d-2)$ as considered in [131].

As in prior works [39, 98, 138], the starting point is the *electric reformulation* of the modulated energy (1.2) as a (renormalized) version of the energy

$$\int_{\mathbb{R}^d} |\nabla h_N|^2, \quad h_N := \mathbf{g} * (\mu_N - \mu), \quad (2.11)$$

where h_N denotes the electric potential and as before, μ_N is the empirical measure. This reformulation is valid for the Coulomb case. In the super-Coulomb case, we may instead use the Caffarelli–Silvestre extension procedure [22]. It allows to view \mathbf{g} as the kernel of a local operator $-\nabla \cdot (|z|^\gamma \nabla \cdot)$ and replace (2.11) by the same quantity with weight $|z|^\gamma$, once working in the extended space \mathbb{R}^{d+1} , with z being the $(d+1)$ -th variable. The fact that such a procedure is not available in the same way if $s < d-2$ (one has to consider Dirichlet-type energies with higher derivatives; see, e.g., [32, 148]) is what restricts us so far to the range $s \geq d-2$.

The key to the proof of inequalities such as (2.4) was the observation of a *stress-energy tensor structure* in the left-hand side. More precisely, letting

$$(T_h)_{ij} := 2\partial_i h \partial_j h - |\nabla h|^2 \delta_{ij}. \quad (2.12)$$

denote the stress-energy tensor (from classical mechanics or calculus of variations) associated to h , it is well-known that if h is smooth enough, then

$$\nabla \cdot T_h = 2\nabla h \Delta h, \quad (2.13)$$

where the divergence may be taken with respect either rows or columns, as the tensor is symmetric. Ignoring the question of the diagonal in (1.2), the main point is that in the Coulomb case, using (2.13), one may rewrite the left-hand side of (2.4) after desymmetrization as

$$2 \int_{(\mathbb{R}^d)^2} v(x) \cdot \nabla \mathbf{g}(x-y) d(\mu_N - \mu)^{\otimes 2} = -\frac{2}{c_{d,s}} \int_{\mathbb{R}^d} v \cdot \nabla h_N \Delta h_N = -\frac{1}{c_{d,s}} \int_{\mathbb{R}^d} v \cdot \nabla \cdot T_{h_N}. \quad (2.14)$$

An integration by parts then allows to formally conclude, since $\int_{\mathbb{R}^d} |T_{h_N}| \leq \int_{\mathbb{R}^d} |\nabla h_N|^2$, which is the energy. The preceding reasoning is purely formal because $|\nabla h_N|^2$ diverges near each point of the configurations (this is related to the diagonal excision), and this computation needs to be properly renormalized, which is the main technical roadblock in the proof. Things work the same in the Riesz case, after extending the space and adding the appropriate weight. We refer to [131, Section 3] for the details.

Considering a pure stress-tensor approach to the proof of higher-order estimates (2.38) for $n \geq 2$, it is unclear that an algebraic manipulation like (2.14) can be found, leading to delicate proofs in [98, 139] which do not seem extendable to $n \geq 3$. In [131], we are able to exhibit a suitable, albeit more complicated, stress-tensor structure in the higher-order variations (1.3). This structure requires us to consider not just the electric potential h_N , but also *iterated commutators* of it. Let us describe this more precisely at first order in the Coulomb case. We will discuss the higher commutators later in Section 2.3.

Given a measure f in \mathbb{R}^d of integral 0 (think of $f = \mu_N - \mu$), $h^f = \mathbf{g} * f$ its Coulomb potential, and a vector-field v , we define the *first commutator* of h^f as

$$\kappa^f(x) := \int_{\mathbb{R}^d} \nabla \mathbf{g}(x-y) \cdot (v(x) - v(y)) df(y). \quad (2.15)$$

It is a commutator because it can be rewritten as

$$\kappa^f = -\mathbf{g} * (\nabla \cdot (vf)) + v \cdot \nabla \mathbf{g} * f = -h^{\nabla \cdot (vf)} + v \cdot \nabla h^f. \quad (2.16)$$

In fact, the commutator κ^f is intimately tied to the stress-tensor itself, via the relation (2.14), which can be rewritten in general as

$$\begin{aligned} 2c_{d,s} \int_{\mathbb{R}^d} \kappa^f dw &= c_{d,s} \int_{\mathbb{R}^d} \kappa^f dw + c_{d,s} \int_{\mathbb{R}^d} \kappa^w df = - \int_{\mathbb{R}^d} v \nabla \cdot [\nabla h^f, \nabla h^w] \\ &= \int \nabla v : [\nabla h^f, \nabla h^w], \end{aligned} \quad (2.17)$$

where

$$[\nabla h^f, \nabla h^w]_{ij} := \partial_i h^f \partial_j h^w + \partial_j h^f \partial_i h^w - \nabla h^f \cdot \nabla h^w \delta_{ij} \quad (2.18)$$

is the bilinear generalization of the stress-energy tensor T_h from (2.12), which we note is symmetric.

Considered separately, each term in the definition of the function κ^f is one derivative less regular than h^f ; however thanks to its commutator structure, some compensation happens. By choosing $w = -\frac{1}{c_{d,s}} \Delta \kappa^f$, integrating by parts on the left-hand side, and using Cauchy–Schwarz on the right-hand side, we can show the \dot{H}^1 control

$$\int_{\mathbb{R}^d} |\nabla \kappa^f|^2 \leq C \|\nabla v\|_{L^\infty} \int_{\text{supp } \nabla v} |\nabla h^f|^2. \quad (2.19)$$

Compared to classical proofs of commutator estimates via paraproducts, the proof of (2.19) is remarkably simple, thanks to the use of the stress tensor. Moreover, the right-hand side is localized to the support of the transport, something not evidently obtainable with classical proofs.

2.2. sub-Coulomb case

We now discuss the sub-Coulomb case $s < d - 2$ following the later work [84] joint with Hess-Childs and Serfaty. In fact, the cited work gives a new proof for the full Riesz case $0 \leq s < d$, but the estimates are worse in the (super-)Coulomb case in terms of their transport regularity assumptions. We comment more on this point in Section 2.4.

There are two main ingredients in the proof: (1) a new potential truncation scheme based on an integral representation of the Riesz potential and (2) commutator estimates of Kato–Ponce type.

The starting point for the first ingredient is the identity

$$\forall x \neq 0, \quad \mathbf{g}(x) = c_{\phi,d,s} \int_0^\infty t^{d-s} \phi_t(x) \frac{dt}{t}, \quad (2.20)$$

where ϕ is any sufficiently nice radial function with nonnegative Fourier transform $\widehat{\phi} \geq 0$, $\phi_t(x) := t^{-d} \phi(x/t)$, and $c_{\phi,d,s}$ depends on ϕ, d, s . In the log case $s = 0$, the integral has to be renormalized at large scales to obtain a convergent expression due to the fact that $\log|x|$ does not decay at infinity. The reader may consult [84, Section 2] for precise assumptions ensuring the validity of the identity. The formula, which is a consequence of scaling, expresses that the Riesz potential is a weighted average of the family of approximate identities $\{\phi_t\}_{t>0}$. In [134, Chapter 6], Rubin refers to this as a “wavelet type representation.” The identity (2.20) may also be understood in terms of the scaling/inversion properties of Mellin transforms [49, Appendix B].

The utility of the representation (2.20) is that it conveniently allows us to truncate the singularity of the Riesz potential simply by cutting off the integral for small t . More precisely, given a scale $\eta > 0$, we define the *truncated potential*

$$\mathbf{g}_\eta(x) := c_{\phi,d,s} \int_\eta^\infty t^{d-s} \phi_t(x) \frac{dt}{t}. \quad (2.21)$$

The truncated potential has several important properties that are crucial to the analysis (see [84, Lemma 2.1] for all of the properties of \mathbf{g}_η):

- (1) While \mathbf{g} is singular at the origin, \mathbf{g}_η is finite and bounded by $\mathbf{g}(\eta)$.
- (2) \mathbf{g}_η is a positive definite kernel (i.e. repulsive, or with nonnegative Fourier transform).
- (3) $\mathbf{g}_\eta \leq \mathbf{g}$ and the difference $\mathbf{g} - \mathbf{g}_\eta$ decays rapidly at scale η outside the ball $B(0, \eta)$.

Without using the representation (2.20), it is not easy to construct a truncation \mathbf{g}_η satisfying all of the three requirements at once, and a standard mollification procedure fails to do so. Moreover,

the choice of ϕ will require care. Thanks to these properties, splitting $F_N(X_N, \mu)$ into

$$F_N(X_N, \mu) = \frac{1}{2} \int_{(\mathbb{R}^d)^2} \mathbf{g}_\eta(x-y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2} + \frac{1}{2} \int_{(\mathbb{R}^d)^2 \setminus \Delta} (\mathbf{g} - \mathbf{g}_\eta)(x-y) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2} - \frac{\mathbf{g}_\eta(0)}{2N}, \quad (2.22)$$

where the diagonal can be reinserted in the first integral up to the well-controlled last term on the second line, we easily obtain two sets of controls by F_N . The first is on the first term on the right-hand side, which is nonnegative (thanks to the repulsive nature of \mathbf{g}_η) and is the square of a genuine MMD. The second is a control on the second term on the right-hand side which, up to well-controlled additive error terms, is also nonnegative and, thanks to properties of \mathbf{g}_η , controls $\frac{1}{N^2} \sum_{i \neq j} \mathbf{g}(x_i - x_j)$, i.e., the interaction energy coming from particle interactions at small scales. See [84, Proposition 2.11] for the precise statements of these controls. The former control allows to show in [84, Proposition 2.15] that the modulated energy controls the squared inhomogeneous Sobolev norm $\|\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\|_{H^{-\frac{d}{2}-\varepsilon}}^2$, for any $\varepsilon > 0$, up to $O(\lambda^{d-s})$ error (i.e. coercivity). Choosing $\eta \propto N^{-1/d}$ (the microscale), the latter control allows to bound in terms of the modulated energy the interaction energy due to nearest neighbors at the microscale, a fortiori bounding the number of points with nearest-neighbor distance below the microscale, see [84, Corollary 2.14]. Let us mention that the potential truncation introduced here plays a key role in our forthcoming work [83] on optimal quantization via MMDs with kernels having power law-type spectra (e.g. Riesz).

In contrast to [112, 116, 131, 138], there is no need to consider a smearing/regularizing of the charges δ_{x_i} , nor to use smearing radii that are point-dependent and configuration-dependent; we only need to consider the truncated potential \mathbf{g}_η . In these previous works, the strategy behind an estimate of the form in Theorem 2.1 is to first prove an estimate valid when the diagonal is reinserted and $\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu$ is replaced by a more regular distribution f . One then reduces to this regular setting by replacing the point charges δ_{x_i} by smeared charges $\delta_{x_i}^{(\eta)}$ and estimating the error from this replacement, the so-called renormalization step. Important to implementing the renormalization is to show that the modulated energy controls both the potential energy of the difference $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}^{(\eta)} - \mu$ and the small-scale interactions.

The strategy in [84] differs crucially in two regards. First, rather than regularize $\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu$ through smearing, we regularize \mathbf{g} through the truncation \mathbf{g}_η .⁴ Second, to prove an estimate for \mathbf{g}_η , we use the definition (2.21) to reduce to proving estimates for

$$\int_{(\mathbb{R}^d)^2} (v(x) - v(y)) \cdot \nabla \phi\left(\frac{x-y}{t}\right) d\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu\right)^{\otimes 2}(x, y), \quad t \in (0, \infty), \quad (2.23)$$

which we integrate over (η, ∞) with respect to the measure $t^{d-s} \frac{dt}{t}$. Provided the dependence on v in our estimate is scale-invariant, this reduces to proving an estimate for $t = 1$.

We now come to the second main ingredient of our proof. The identity (2.20) is valid for a large class of scaling functions ϕ , but not all choices of ϕ will yield the desired estimate. Indeed, it is a elementary Fourier computation that no estimate of the form

$$\int_{(\mathbb{R}^d)^2} (v(x) - v(y)) \cdot \nabla \phi(x-y) f(x) f(y) \leq C_v \int_{(\mathbb{R}^d)^2} \phi(x-y) f(x) f(y) \quad (2.24)$$

can hold when $\hat{\phi}$ decays super-polynomially (e.g. Gaussian). It turns out that a good choice for ϕ is a Bessel potential, i.e. $\hat{\phi}(\xi) = \langle 2\pi\xi \rangle^{-a}$, which is the kernel of the inhomogeneous Fourier multiplier $\langle \nabla \rangle^a$, where we use the Japanese bracket notation $\langle x \rangle^a := (1 + |x|^2)^{a/2}$ for $a \in \mathbb{R}$ and $x \in \mathbb{R}^d$. The Bessel potential is a screening of the Riesz potential which preserves its local behavior at the origin but avoids the issues at low frequency that lead to the slow spatial decay of the Riesz potential. Setting $h := \phi * f$ and using that $\langle \nabla \rangle^a h = f$, the choice of Bessel potential reduces to proving a

⁴The work [18] also directly regularized the interaction \mathbf{g} at small length scales. However, the procedure in that work is completely different than in [84], limited to the spatially periodic setting, and does not yield optimal error estimates.

product rule for $\langle \nabla \rangle^{a/2}$. Such product rules are known as Kato–Ponce estimates in the harmonic analysis literature on account of their origins [91]. This is the content of [84, Proposition 3.1], which is the main technical lemma. Here, we rely on commutator estimates of Li [101], as well as a local representation of fractional powers $\langle \nabla \rangle^{a/2}$ via dimension extension (see [84, Section 3]), analogous to the representation for the fractional Laplacian popularized by Caffarelli and Silvestre [22].

The parameter a for the Bessel potential is subject to constraints. We need $a > d$, so that ϕ is continuous, bounded. On the other hand, if a is too large, then we will not be able to obtain an estimate whose dependence on v scales properly. This leads to the constraint $a < d + 2$. Even with the constraints on a , it is somewhat remarkable that an intermediate inhomogeneous estimate allows for a homogeneous estimate in the end.

The identity (2.20) itself suggests a class of potentials beyond the exact Riesz case by replacing the function t^{d-s} with a general function of t that is pointwise controlled above and below by t^{d-s} . We call these *Riesz-type* potentials. The reader may find their proper definition in [84, Section 2.2] and should compare this class with the related, but different, class of Riesz-type potentials previously introduced in [112]. In fact, the estimate Theorem 2.1(2.5) is obtained as a special case of a more general functional inequality for Riesz-type potentials (see [84, Theorem 4.1]).

2.3. Localized and higher-order estimates

Having discussed first-order estimates, let us now see how to obtain the second- and higher-order, possibly localized, estimates using stress-tensor and commutator structures, following [131]. The results discussed in this subsection are limited to the (super-)Coulomb case. It is unclear how to obtain (sharp) higher-order estimates in the sub-Coulomb case with the method of [84]. For more discussion on this problem, see [84, Section 1.5] or [84, Sections 5.2, 5.3].

Ignoring the question of renormalization, and sticking to the Coulomb case, we have seen in (2.14) (combining with (1.3)) that the first variation of the modulated energy along the transport map $\mathbb{I} + tv$ is

$$-\frac{1}{2c_{d,s}} \int_{\mathbb{R}^d} v \cdot \nabla \cdot T_{h_N} dx = \frac{1}{2c_{d,s}} \int_{\mathbb{R}^d} \nabla v : T_{h_N} dx. \quad (2.25)$$

To compute the second variation of the modulated energy, we thus need to compute the first variation of $\operatorname{div} T_{h_N}$ when again μ_N and μ are pushed forward by $\mathbb{I} + tv$. In view of the expression for T_{h_N} in (2.12), it suffices to compute the derivative of h_N^t (with obvious notation) at $t = 0$, and since $h_N = \mathbf{g} * (\mu_N - \mu)$, the definition of the push-forward yields that

$$\left. \frac{d}{dt} \right|_{t=0} h_N^t = -\mathbf{g} * (\operatorname{div}(v(\mu_N - \mu))) = -h^{\operatorname{div}(v(\mu_N - \mu))} = \kappa^{\mu_N - \mu} - v \cdot \nabla h_N, \quad (2.26)$$

after using (2.16). The L^2 norm of the gradient of this expression is one derivative more singular than the energy $\int_{\mathbb{R}^d} |\nabla h_N|^2$. Still with $f = \mu_N - \mu$, inserting (2.26) into the variation of (2.25), we can decompose the second-order variation as

$$\frac{1}{c_{d,s}} \int_{\mathbb{R}^d} \partial_i v^j : ([\nabla h_N, \nabla \kappa^f] - [\nabla h_N, \nabla(v \cdot \nabla h_N)]). \quad (2.27)$$

Thanks to the \dot{H}^1 estimate (2.19) for the commutator κ^f , the first term on the right-hand side can be directly controlled by $C_v \int |\nabla h_N|^2$ as desired, while the second one can be transformed into similarly controllable terms (albeit C_v is now quadratic in v and depends on $v, \nabla^{\otimes 2} v$) by means of integration by parts of the advection operator $v \cdot \nabla$.

This argument can then be iterated at next order by introducing the family of n -th order commutators $\kappa^{(0),f} = h^f$ and

$$\kappa^{(n),f} := \int_{\mathbb{R}^d} \nabla^{\otimes n} \mathbf{g}(x - y) : (v(x) - v(y))^{\otimes n} df(y), \quad (2.28)$$

together with “transported” commutators

$$\kappa_t^{(n),f} := \int_{\mathbb{R}^d} \nabla^{\otimes n} \mathbf{g}(x - y - tv(y)) : (v(x) - v(y))^{\otimes n} df(y), \quad (2.29)$$

and observing the elegant recursive formula

$$\partial_t \kappa_t^{(n),f} = \kappa_t^{(n+1),f} - v \cdot \nabla \kappa_t^{(n),f}, \quad (2.30)$$

where t is the same parameter as in the transport map $\mathbb{I} + tv$. In other words, the high-order “time-dependent” commutators $\kappa_t^{(n),f}$ satisfy a hierarchy of transport equations with a source coupling the $(n+1)$ -th order commutator to the n -th order commutator. An iteration of the same argument as for (2.19) using this recursion allows to prove the estimate

$$\int_{\mathbb{R}^d} |\nabla \kappa^{(n),f}|^2 \leq C_v \int_{\text{supp } \nabla v} |\nabla h^f|^2, \quad (2.31)$$

valid for general functions f such that the right-hand side is finite, with the constant C_v now being n -linear in v and involving derivatives up to order n of v . One can see the relation (2.31) as an *L^2 -based regularity theory for arbitrary-order commutators*.

The discussion so far has been limited to the Coulomb case, the complete details of which are presented in [131, Appendix A]. In generalizing to the Riesz case $\mathfrak{s} \neq \mathfrak{d} - 1, \mathfrak{d} - 2$, we run into issues. The vector field v may be trivially extended to $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$, with $\mathfrak{k} = 1$ if $\mathfrak{s} > \mathfrak{d} - 2$, by fixing the last component to zero, \mathfrak{g} may be trivially extended through radial symmetry, and all distributions f on $\mathbb{R}^{\mathfrak{d}}$ viewed as living on the boundary $\mathbb{R}^{\mathfrak{d}} \times \{0\}^{\mathfrak{k}}$. Consequently, $\kappa^{(n),f}$ may be viewed as a function on $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$. Going through the computations above, all integrals $\int_{\mathbb{R}^{\mathfrak{d}}}$ should be replaced $\int_{\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}}$ $|z|^\gamma$ where $\gamma = \mathfrak{s} + 2 - \mathfrak{d} - \mathfrak{k}$. The obstruction lies in obtaining the L^2 gradient bound (2.31) by duality.

Setting $L := -\text{div}(|z|^\gamma \nabla \cdot)$, we would like to take $w = \frac{1}{c_{\mathfrak{d},\mathfrak{s}}} L \kappa^{(n),f}$ in $\int_{\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}} \kappa^{(n),f} w$ and then integrate by parts to conclude an estimate for $\int_{\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}} |z|^\gamma |\nabla \kappa^{(n),f}|^2$. However, this choice for w is a priori not supported on $\mathbb{R}^{\mathfrak{d}} \times \{0\}^{\mathfrak{k}}$. If $\text{supp } \phi \not\subset \mathbb{R}^{\mathfrak{d}} \times \{0\}^{\mathfrak{k}}$, it is not necessarily true that $\mathfrak{g} * L \phi = c_{\mathfrak{d},\mathfrak{s}} \phi$, for a test function ϕ on $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$, unless $\gamma = 0$. This forces us to come up with a new approach.

The starting point is the observation $L \mathfrak{g} = c_{\mathfrak{d},\mathfrak{s}} \delta_0$ in $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$, which underlies the Caffarelli–Silvestre representation of the fractional Laplacian as a degenerate elliptic operator in $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$. From this observation, it follows that $\kappa^{(n),f}$ satisfies the equation

$$\begin{aligned} L \kappa^{(n),f} &= c_{\mathfrak{d},\mathfrak{s}} (-1)^n \sum_{\sigma \in \mathbb{S}_n} \partial_{i_{\sigma_1}} v^{i_1} \dots \partial_{i_{\sigma_n}} v^{i_n} f \delta_{\mathbb{R}^{\mathfrak{d}} \times \{0\}^{\mathfrak{k}}} - n \partial_i (|z|^\gamma \partial_i v \cdot \nu^{(n-1),f}) \\ &\quad - n |z|^\gamma \partial_i v \cdot \partial_i \nu^{(n-1),f} + n(n-1) |z|^\gamma (\partial_i v)^{\otimes 2} : \mu^{(n-2),f}, \end{aligned} \quad (2.32)$$

where \mathbb{S}_n denotes the symmetric group on $[n]$ and a repeated index denotes summation over that index. In the right-hand side, ν is a vector field on $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$ that “morally” is like $\nabla \kappa^{(n),f}$, while μ is a symmetric matrix field on $\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}$ that morally is like $\nabla^{\otimes 2} \kappa^{(n),f}$ (see [131, (4.12), (4.14)] respectively, for the precise definitions). The right-hand side of (2.32) is good because it only depends on lower-order (i.e. $n-1, n-2$) commutators. Following the standard method for the L^2 regularity of elliptic equations, we want to obtain an estimate for $\int_{\mathbb{R}^{\mathfrak{d}+\mathfrak{k}}} |z|^\gamma |\nabla \kappa^{(n),f}|^2$ by testing the equation (2.32) against $\kappa^{(n),f}$ and using an induction hypothesis that will guarantee good bounds on the lower order commutators and their derivatives. There is however a difficulty involving the nondivergence form of the right-hand side of (2.32) and L^2 estimates for $\kappa^{(n),f}$ (see the beginning of [131, Section 4.1] for elaboration), which requires the use of a Poincaré inequality when $n \geq 3$ and yields an estimate that deteriorates when ℓ , the size of the support of v , gets large.

As a result, when $\text{supp } v$ is not compact, which is not the case for the aforementioned CLT application, this approach does not work, and the right-hand side of (2.38) is infinite. However, obtaining unlocalized L^2 commutator estimates is much easier and may be done following the method of [112]: integration by parts and the difference quotient characterization of the Sobolev seminorm $\dot{H}^{\frac{\mathfrak{s}-\mathfrak{d}}{2}}$ when $\mathfrak{s} > \mathfrak{d} - 2$, and the L^2 boundedness of Calderón \mathfrak{d} -commutators when $\mathfrak{s} = \mathfrak{d} - 2$.

For the precise statements of the results described above, we refer the reader to [131, Theorem 4.1] (and more generally, Section 4.1), which is our omnibus L^2 regularity estimate for commutators.

This discussion has so far left aside the question of renormalization, which is that of dealing with the singularities of the Dirac masses in h_N , and which is arguably the most delicate part of the analysis. As mentioned in Section 2.1 for the first-order estimates, the renormalization is handled as in [138] and subsequent works via a charge smearing and potential truncation at a lengthscale crucially *depending on each point* x_i and proportional to its nearest-neighbor distance. The renormalization in the $n \geq 2$ case is much more delicate, occupying the largest part of the paper and requiring several innovations.

First, replacing the Diracs in the left-hand side of (2.38) by their smearings $\delta_{x_i}^{(\eta_i)}$ will not allow us to apply our unrenormalized commutator estimate unless $s = d - 2$ because $\mu \delta_{\mathbb{R}^d \times \{0\}^k} - \frac{1}{N} \sum_{i=1}^N \delta_{x_i}^{(\eta_i)}$ is not supported on $\mathbb{R}^d \times \{0\}^k$. To overcome this difficulty, we introduce a new smearing $\rho_{x_i}^{(\eta_i)}$, obtained simply by mollifying the Dirac in \mathbb{R}^d , and estimate the electric energy difference of the two regularizations in terms of the modulated energy. We can then directly apply our L^2 commutator bound to $\mu - \frac{1}{N} \sum_{i=1}^N \rho_{x_i}^{(\eta_i)}$. A novelty compared to previous work is that we will choose the mollifier to have vanishing moments up to high enough order in order to obtain optimal error rates. For the details of the described construction and estimates, see [131, Section 5.1].

Second, to handle the error from this smearing, we establish an L^∞ -based *regularity theory for commutators* (see [131, Sections 4.5–4.6]), which shows that the L^∞ norm of “horizontal” derivatives $\nabla_x^{\otimes m} \kappa^{(n),f}$, where $\nabla_x := (\partial_1, \dots, \partial_d)$, *restricted to arbitrarily small balls* can be estimated in terms of the L^2 norm of $\nabla \kappa^{(n),f}$ in a double ball. Commutator estimates do not typically allow such localization, and therefore this result may be of independent interest. The proof relies on regularity theory for second-order divergence-form elliptic operators with A_2 weights à la Fabes et al. [45] in the extended space \mathbb{R}^{d+k} , which, by the Caffarelli–Silvestre extension, may be viewed as a regularity theory for the fractional Laplacian. The proof also crucially uses the commutator structure in the form of equation (2.32) for $L\kappa^{(n),f}$, along with recursions for $\kappa^{(n),f}$, $\nu^{(n),f}$, and $\mu^{(n),f}$. When $\ell \asymp 1$, this local regularity theory may be avoided with cruder estimates, which have the benefit of requiring less regularity on the vector field v . However, when $\ell \ll 1$, these cruder estimates give errors that scale sub-optimally.

Third, we improve truncation errors estimates from [138], obtaining the optimal dependence on μ , and show new localizable *mesoscale interaction energy estimates*, which generalize those of [139] to the Riesz case. See [131, Sections 2.2–2.3].

The complete details of the renormalization step may be found in [131, Section 5.2] and more generally, Section 5 of the cited work. Having discussed the proof strategy and main innovations, we conclude this subsection with statements of the main results.

The first theorem is a global estimate, which does not require any localization of the vector field v . The additive error term is $\lambda^{d-s} \propto (N^{-1+\frac{s}{d}})$, which is the announced optimal estimate. Compared to the localized estimate (2.38) below, the regularity requirement for v is not dependent on the order n of the commutator. As the reader may check, if $s > d - 1$, we only need $\|\nabla v\|_{L^\infty}$ ($m = 0$) to achieve the sharp error λ^{d-s} . If $d - 2 < s \leq d - 1$, then we also need $\|\nabla^{\otimes 2} v\|_{L^\infty}$ ($m = 1$). If $s = d - 2$, then we additionally need $\|\nabla^{\otimes 3} v\|_{L^\infty}$ ($m = 2$).

Theorem 2.5 (Global estimates). *There exists a constant $C > 0$ depending only d, s such that the following holds. Let $\mu \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \mu = 1$. If $s \leq 0$, suppose further that $\int_{(\mathbb{R}^d)^2} |\mathbf{g}(x - y)| d|\mu|^{\otimes 2} < \infty$. Let $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a Lipschitz vector field, suppose that $\lambda < 1$, where $\lambda := (N \|\mu\|_{L^\infty(\mathbb{R}^d)})^{-\frac{1}{d}}$. For any integers $m \geq 0$ and $n \geq 1$, there exists a constant $C > 0$ depending only d, s, m, n such that the following holds. Let $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a $C^{m,1}$ vector field. For any pairwise distinct configuration $X_N \in (\mathbb{R}^d)^N$, we have*

$$\begin{aligned}
|A_n[X_N, \mu, v]| &\leq C(\|\nabla v\|_{L^\infty})^n \left(F_N(X_N, \mu) - \left(\frac{\log \lambda}{2N} \right) \mathbf{1}_{s=0} + C\|\mu\|_{L^\infty} \lambda^{d-s} \right) \\
&+ C \sum_{q=1}^{m+1} \sum_{r=0}^q D_{v,q,r} \left(\lambda^{m+1} + \lambda^{m+1-q} \left(F_N(X_N, \mu) - \left(\frac{\log \lambda}{2N} \right) \mathbf{1}_{s=0} + C\|\mu\|_{L^\infty} \lambda^{d-s} \right) \right. \\
&\quad \left. + \lambda^{m+1} \|\mu\|_{L^\infty} \begin{cases} 1, & s + q \neq d \\ \log(1/\lambda), & s + q = d \end{cases} \right. \\
&\quad \left. + \lambda^{m+1} \begin{cases} \|\mu\|_{L^\infty} \lambda^{d-s-q}, & s + q > d \\ \|\mu\|_{L^\infty} (1 - \lambda^{d-s-q}) + \|\mu\|_{L^1}, & s + q < d \\ \|\mu\|_{L^\infty} \log(1/\lambda) + \|\mu\|_{L^1}, & s + q = d \end{cases} \right), \quad (2.33)
\end{aligned}$$

where

$$D_{v,q,r} := \sum_{\substack{1 \leq c_1, \dots, c_r \\ c_1 + \dots + c_r = m+1 - (q-r)}} \|\nabla^{\otimes c_1} v\|_{L^\infty} \cdots \|\nabla^{\otimes c_r} v\|_{L^\infty} \|\nabla v\|_{L^\infty}^{n-r}. \quad (2.34)$$

In studying the fluctuations of Coulomb/Riesz gases on mesoscopic lengths $N^{-1/d} \ll \ell \ll 1$, it is crucial that commutator estimates admit a *local* form, in which the right-hand side can be expressed in terms of the modulated energy restricted to the region where v is supported (up to boundary errors). Such estimates are key analytic input to the transport method for proving CLTs [98, 115, 139] (see [140, Part 3] for a consolidated treatment).

Our second theorem is a localized generalization of Theorem 2.1, as well as a higher-order version. Now, the modulated energy on the right-hand side is localized to a region comparable to the support of v and the estimate scales with the diameter of the support. The additive error term is in $N^{-1} \lambda^{d-s} \propto (N^{-2+\frac{s}{d}})$ per point, which is optimal. However, the error now depends on higher-order derivatives of v and deteriorates when the size of the support of v gets large. We mention that Theorem 2.6 was an important component to the recent work [115] of Peilen and Serfaty on the fluctuations of super-Coulomb Riesz gases.

To formulate the result, let $\Omega \subset \mathbb{R}^d$, meant to represent $\text{supp } v$. Associated to Ω , define the microscopic length scale

$$\lambda := (N \|\mu\|_{L^\infty(\Omega)})^{-\frac{1}{d}}, \quad (2.35)$$

which can be thought of as the typical inter-particle distance, and define $\widehat{\Omega} := \{x : \text{dist}(x, \Omega) \leq \frac{1}{4}\lambda\}$. We will use ℓ to denote the typical size of Ω (or of $\text{supp } v$), which may also depend on N and whose only constraint is to remain larger than the microscopic scale λ . Finally, we let $I_\Omega := \{1 \leq i \leq N : x_i \in \Omega\}$ and use $\#$ to denote the cardinality of a finite set.

Theorem 2.6 (Localized estimates). *There exists a constant $C > 0$ depending only d, s such that the following holds. Let $\mu \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \mu = 1$. If $s \leq 0$, suppose further that $\int_{(\mathbb{R}^d)^2} |\mathbf{g}(x-y)| |\mu|^{\otimes 2} < \infty$. Let $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a Lipschitz vector field and Ω be a closed set containing a 2λ -neighborhood of $\text{supp } \nabla v$, where λ is defined as in (2.35) and $\lambda < 1$. For any pairwise distinct configuration $X_N \in (\mathbb{R}^d)^N$, it holds that*

$$|A_1[X_N, \mu, v]| \leq C \|\nabla v\|_{L^\infty} \left(\int_{\Omega \times [-\ell, \ell]^k} |z|^\gamma |\nabla h_{N, \vec{r}}|^2 + C \frac{\#I_\Omega \|\mu\|_{L^\infty(\widehat{\Omega})} \lambda^{d-s}}{N} \right). \quad (2.36)$$

Here, $h_{N, \vec{r}}$ is the truncated electric potential and $\vec{r} = (r_i)_{i=1}^N$ is the nearest-neighbor type distance defined by

$$h_{N, \vec{r}} := \frac{1}{N} \sum_{i=1}^N \mathbf{g}_{r_i}(\cdot - x_i) - \mathbf{g} * \mu, \quad \mathbf{g}_{r_i} := \min(\mathbf{g}, \mathbf{g}(r_i)), \quad r_i := \frac{1}{4} \min\left(\min_{j \neq i} |x_i - x_j|, \lambda\right), \quad (2.37)$$

where we use the radial symmetry of \mathbf{g} to abuse notation.

Suppose in addition that $v \in C^{2n-1,1}$ and that Ω' is a ball of radius ℓ containing a λ -neighborhood of $\text{supp } v$ and Ω contains a 5ℓ -neighborhood of Ω' ,⁵ where λ, ℓ satisfy $\min(\lambda, \lambda/\ell) < \frac{1}{2}$. For any $n \geq 2$, we have

$$|A_n[X_N, \mu, v]| \leq C \sum_{p=0}^n (\ell \|\nabla^{\otimes 2} v\|_{L^\infty})^p \sum_{\substack{1 \leq c_1, \dots, c_{n-p} \\ n-p \leq c_1 + \dots + c_{n-p} \leq 2n}} \lambda^{-(n-p) + \sum_{k=1}^p c_{n-k}} \\ \times \|\nabla^{\otimes c_1} v\|_{L^\infty} \cdots \|\nabla^{\otimes c_{n-p}} v\|_{L^\infty} \left(\int_{\Omega \times [-\ell, \ell]^k} |z|^\gamma |\nabla h_{N, \vec{r}}|^2 + C \frac{\#I_\Omega \|\mu\|_{L^\infty(\widehat{\Omega})} \lambda^{d-s}}{N} \right), \quad (2.38)$$

where $C > 0$ may depend on n and the summation of the c_i is understood as vacuous when $p = n$.

Note that the estimate (2.38) does not allow to take $\Omega = \mathbb{R}^d$ because that would mean $\ell = \infty$, in which case the right-hand side is infinite. This is an unfortunate limitation of the result, one we

⁵For integer $k \geq 0$ and real $\alpha \in [0, 1]$, $C^{k, \alpha}$ denotes the inhomogeneous Hölder space of k -times continuously differentiable functions whose k -th derivative is α -Hölder continuous.

believe is technical. We refer the interested reader to [131, Section 4.1] and [84, Section 5.3] for a discussion of the issue and possible ways to circumvent it.

Remark 2.7. While the conditions on Ω, Ω', λ may appear circular upon first read, it is indeed possible for them to hold. See [131, Remark 1.4] for details.

2.4. Transport regularity and defective estimates

Besides the sharpness of the additive errors in these commutator estimates, discussed above, another important question, which we recently considered in the joint work [84] with Hess-Childs and Serfaty, is the optimal regularity assumption for the transport velocity v in the aforementioned commutator estimates. Compared to the former, the latter question has received little attention in the literature.

It is natural to ask whether the Lipschitz assumption on v may be weakened while still preserving the scaling. For example, can $\|\nabla v\|_{L^\infty}$ be replaced by $\|\nabla v\|_{BMO}$? It is also unclear whether, in the sub-Coulomb regime $s < d - 2$, the additional term of the form $\| |\nabla|^{\frac{a}{2}} v \|_{L^{\frac{2d}{a-2}}}$ is necessary and how small the exponent a may be taken. By Sobolev embedding, $\| |\nabla|^{\frac{a}{2}} v \|_{L^{\frac{2d}{a-2}}} \lesssim \| |\nabla|^{\frac{a'}{2}} v \|_{L^{\frac{2d}{a'-2}}}$ for $2(d+1) > a' \geq a$. Moreover, there is a gap between our two sub-Coulomb estimates (2.5), (2.6).

This regularity question is motivated by applications. It is a general principle in PDE that if an equation has a scaling invariance, then a natural function space to study the equation is one that is invariant under the equation's scaling [92]. Such *scaling-critical* or *scaling-invariant* spaces are expected to be thresholds for the well-posedness/ill-posedness of the equation along with other phenomena, such as finite-time blowup vs. global existence.

When $\mathbf{V}^t = 0$, the reader may check that if μ is a solution to the mean-field equation (3.3) with $\beta = \infty$, then $\tilde{\mu}(t, x) := a\mu(bt, cx)$, for $a, b, c > 0$, is also a solution provided that $bc^{d-2-s} = a$. Fixing $b = 1$ so that the solution lifespan is preserved under rescaling, yields the relation $a = c^{d-2-s}$. For $1 \leq p \leq \infty$, the reader may check that

$$\|\tilde{\mu}\|_{L^p} = c^{d-s-2-\frac{d}{p}} \|\mu\|_{L^p}, \quad (2.39)$$

which implies that $p = \frac{d}{d-s-2}$ is the scaling-critical exponent. Thus, the potential can be at most Coulomb in order for there to be a critical L^p space. More generally, letting $\dot{W}^{a,p}$ denote the homogeneous Bessel potential (fractional Sobolev) space (see, e.g., [141, Section 3.3] or [53, Section 1.3]), we see that $\dot{W}^{a,p}$ is a critical space if and only if $a = \frac{d}{p} + s + 2 - d$. Other scales of function spaces, such as Hölder-Zygmund spaces (see, e.g., [141, Section 4]), are also interesting to consider as critical spaces.

When $\mu \in \dot{W}^{\frac{d}{p}+s+2-d,p}$, for $1 < p < \infty$, some basic Fourier analysis shows that the associated mean-field velocity $v = \mathbb{M}\nabla\mathbf{g} * \mu$ is in $\dot{W}^{\frac{d}{p}+1,p}$. Unfortunately, $\dot{W}^{\frac{d}{p}+1,p}$ does not embed into the homogeneous Lipschitz space $\dot{W}^{1,\infty}$. If $f \in \dot{W}^{\frac{d}{p}+1,p}$, then one only has $\|\nabla f\|_{BMO} \leq C\|f\|_{\dot{W}^{\frac{d}{p}+1,p}}$. Even when $s + 2 - d = k$ is an integer and taking $p = \infty$, having $\mu \in \dot{W}^{k,\infty}$ does not imply that $v \in \dot{W}^{1,\infty}$. This is problematic because all of the commutator estimates of Theorem 2.1 feature $\|\nabla v\|_{L^\infty}$ on the right-hand side. Thus, to show mean-field convergence when the mean-field density μ^t belongs to the critical space, it would be desirable to have a commutator estimate with a regularity assumption that is satisfied when $\mu \in \dot{W}^{\frac{d}{p}+s+2-d,p}$.

In [84], we showed that estimate (2.10) can fail if either of the vector field norms is weakened. We also showed necessary conditions for analogous commutator estimates to hold for a broader class of interaction potentials. In particular, the norm controlling v must involve enough derivatives to compensate for the high-frequency decay of $\widehat{\mathbf{g}}$. This demonstrates a more general principle of regularity trade-off between the interaction and transport for commutator estimates. All of these results, which may be found in [84, Section 3], are obtained through explicit counterexamples. Given that the renormalized commutator estimates of Theorem 2.1 imply unrenormalized commutator estimates, the problem reduces to showing counterexamples when $\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu$ is replaced by a zero-mean Schwartz function.

The second contribution of [84], motivated by applications to mean-field convergence at scaling-critical regularity, concerns substitutes for the functional inequality of Theorem 2.1 when considering the mean-field vector field $v = \mathbb{M}\nabla\mathbf{g} * \mu$ for μ in a critical space. The results of the previous

paragraph give a negative answer to the possibility of a commutator estimate holding in this case, regardless of the size of the additive error. That said, a *defective* commutator estimate still holds, by which we mean that there exists a locally integrable function $\rho : [0, \infty) \rightarrow [0, \infty)$, called the *defect*, such that for all $\epsilon > 0$,

$$|\mathbf{A}_1[X_N, \mu, v]| \leq \rho(\epsilon) C_v \left(\mathbf{F}_N(X_N, \mu) + C_\mu N^{-\alpha} \right) + \epsilon^\beta N^\gamma C'_v \left(\mathbf{F}_N(X_N, \mu) + C'_\mu N^{-\alpha} \right), \quad (2.40)$$

where $\alpha > 0$, $\beta, \gamma \geq 0$ and $C_v, C'_v, C_\mu, C'_\mu > 0$ are constants depending on \mathbf{d}, \mathbf{s} and quantitatively on norms of v, μ , respectively. As we will explain later in Section 3.4, if $\rho(\epsilon) = O(|\log \epsilon|)$ as $\epsilon \rightarrow 0^+$, then a defective commutator estimate suffices to conclude an estimate for the modulated energy that in turn implies mean-field convergence.

In [84, Theorem 2.6], we showed a new defective estimate with $\rho(\epsilon) = O(|\log \epsilon|^\theta)$ for $\theta \in (0, 1)$. Previous work [123, 124] by the author, limited to the Coulomb case, had obtained a defect corresponding to $\theta = 1$. This improvement is made possible by the celebrated *Brezis–Wainger–Hansson inequality* [21, 71]: if $1 < p < \infty$, $1 \leq q \leq \infty$, and $aq > \mathbf{d}$, then

$$\forall f \in \mathcal{S}(\mathbb{R}^{\mathbf{d}}), \quad \|f\|_{L^\infty} \leq C \|f\|_{\dot{W}^{\frac{\mathbf{d}}{p}, p}} \left(1 + \log^{\frac{p-1}{p}} \left(1 + \frac{\|f\|_{\dot{W}^{a, q}}}{\|f\|_{\dot{W}^{\frac{\mathbf{d}}{p}, p}}} \right) \right), \quad (2.41)$$

where $C = C(\mathbf{d}, p, a, q) > 0$. Remark that the exponent $\frac{p-1}{p}$ is sharp. Brezis and Wainger [21, Theorem 1] first proved the estimate (2.41), which is a generalization of an earlier $\mathbf{d} = p = 2$ inequality due to Brezis and Gallouët [20]. Independently, Hansson [71] proved an equivalent inequality expressed in potential-theoretic language, and an alternative proof of (2.41) is also given in [44]. The inequality (2.41) is a substitute for the false endpoint Sobolev embedding $\|f\|_{L^\infty} \lesssim \|f\|_{\dot{W}^{\frac{\mathbf{d}}{p}, p}}$, except when $\mathbf{d} = p = 1$. In some sense, it is also dual to the famous Trudinger–Moser inequality [109, 144] (see [21, Theorem A]).

3. Mean-field limits

We transition to applications of Theorem 2.1, beginning with sharp rates of convergence for the mean-field dynamics of Riesz gases. This section draws on the works [84, 84, 131] joint with Hess-Childs and Serfaty.

3.1. Background and motivation

In recent years, the study of mean-field limits for interacting particle systems has seen significant progress. A basic model is the first-order system

$$dx_i^t = \frac{1}{N} \sum_{1 \leq j \leq N: j \neq i} \mathbb{M} \nabla_x \mathbf{g}(x_i^t, x_j^t) dt - \mathbf{V}^t(x_i^t) dt + \sqrt{2/\beta} dW_i^t, \quad i \in [N], \quad (3.1)$$

where $\beta \in (0, \infty]$ has the interpretation of inverse temperature, W_1, \dots, W_N are iid \mathbf{d} -dimensional Wiener processes, and the stochastic differential dW_i is understood in the Itô sense. Here, \mathbf{g} is an *interaction potential* to be specified, and \mathbf{V}^t is a (possibly time-dependent) *external field*. \mathbb{M} is a $\mathbf{d} \times \mathbf{d}$ matrix with real entries that specifies the type of dynamics, with $\mathbb{M} = -\mathbb{I}$ (*gradient/dissipative*) and \mathbb{M} antisymmetric (*Hamiltonian/conservative*) being the most common. We assume that

$$\mathbb{M} \xi \cdot \xi \leq 0 \quad \forall \xi \in \mathbb{R}^{\mathbf{d}}, \quad (3.2)$$

which is a repulsivity assumption. One can check that our assumption (3.2) for \mathbb{M} ensures that the energy for (3.1) is almost surely (a.s.) nonincreasing, therefore if the particles are initially separated, they a.s. remain separated for all time, so that there is a unique, global strong solution to the system (3.1).

The mean-field limit refers to the convergence as $N \rightarrow \infty$ of the empirical measure μ_N^t associated to a solution X_N^t of the system (3.1). Assuming the initial points X_N^0 are such that μ_N^0 converges to a sufficiently regular measure μ^0 , a formal calculation leads one to expect that for $t > 0$, μ_N^t converges to the solution of the *mean-field equation*

$$\partial_t \mu = \operatorname{div}((\mathbf{V} - \mathbb{M} \nabla \mathbf{g} * \mu) \mu) + \frac{1}{\beta} \Delta \mu. \quad (3.3)$$

Convergence of the empirical measure is qualitatively equivalent to *propagation of molecular chaos* (see [50, 79] and references therein). This means that if $f_N^0(x_1, \dots, x_N) = \text{Law}(X_N^0)$ is μ^0 -chaotic (i.e. the k -point marginals $f_{N;k}^0 \rightharpoonup (\mu^0)^{\otimes k}$ as $N \rightarrow \infty$ for every fixed k), then $f_N^t = \text{Law}(X_N^t)$ is μ^t -chaotic. There is also the stronger notion of *generation of chaos*, discussed in [130], which expresses that f_N^t becomes μ^t -chaotic as $t \rightarrow \infty$.

Mean-field limits for systems of the form (3.1) have a long history, beginning with regular drifts (typically, Lipschitz) [38, 142] and then progressing to increasingly more singular interactions over the years. A proper review of this history is beyond the scope of these pages. The interested reader may consult the survey [28], the lecture notes [50, 51, 87, 89, 140], and the introductions of [112, 138]. Instead, we focus on the much more recent and challenging treatment of singular interactions, specifically log/Riesz potentials, which we consider here under the same assumptions as above, focusing on the zero temperature case $\beta = \infty$. At the end, in Section 3.5, we comment on the positive temperature case.

Mean-field convergence for the full range $-2 < s < d$ has only been resolved in the last few years: the nonsingular case $-2 < s < 0$ [132], the sub-Coulomb case $s < d-2$, [25, 75], the Coulomb/super-Coulomb case $d-2 \leq s < d$ [6, 26, 39, 138], and the full singular case $0 \leq s < d$ [112]. These advances are due in large part to the modulated-energy method, which we review in the next subsection.

Other techniques, such as based on weighted estimates for hierarchies of cumulants or marginals, can show propagation of chaos even for some Riesz interactions [15]. This approach is better suited to the positive temperature/diffusive case, in contrast to our zero-temperature setting, where stronger results are known [15, 16, 42]. Although unable to treat the full potential Riesz range, these approaches do have advantages in terms of working for both first- and second-order dynamics, are robust to the precise form of the interaction, and in some cases achieve sharp rates [85, 94, 95, 146], though for essentially at most log singularities. These results imply probabilistic statements about the mean-field convergence rates, but they do not imply the deterministic convergence rates obtainable with our methods.

3.2. Main results

The optimal rate of convergence as $N \rightarrow \infty$ of the empirical measure μ_N^t to the solution μ^t of (3.3) in the distance F_N is $N^{\frac{s}{d}-1}$. The reason is that minimizers X_N of Riesz interaction energies with a suitable confinement V in \mathbb{R}^d or on the flat torus are known to satisfy $|F_N(X_N, \mu_V)| \propto N^{\frac{s}{d}-1}$, where μ_V is the minimizer of the mean-field energy, called the equilibrium measure (see Section 4.1). This has been shown for the (super-)Coulomb case [116, 133, 135, 136], the general singular case $0 \leq s < d$ on the flat torus [74], and for the nonsingular case $-2 < s < 0$ [83]. Since minimizers are critical points of the energy, it follows that when $V^t = \nabla V$, they yield stationary solutions of (3.1) with $M = -\mathbb{I}$ and $\beta = \infty$. Similarly, μ_V is a stationary solution of (3.3). Sharp rates for the nonstationary case were only completely established recently: the Coulomb case [124, 139], (super-)Coulomb case [131], full singular case [84], and the nonsingular case [132].

Theorem 3.1 below summarizes the state of the art for mean-field convergence (at zero temperature) in the full potential Riesz case $-2 < s < d$ with the sharp rate. The $d-2 \leq s < d$ portion is taken from [131], the $0 \leq s < d-2$ portion from [84], and the $-2 < s < 0$ portion from [132]. Remark that there is no additive $N^{\frac{s}{d}-1}$ error in the nonsingular case $-2 < s \leq 0$ because for any suitable μ , it has been shown that $\min_{X_N} F_N(X_N, \mu) \propto N^{\frac{s}{d}-1}$, in particular nonnegative. As explained in the next subsection, the proof is a quick consequence of the by now standard modulated-energy method and Theorem 2.1. Following the statement, we record some remarks on the theorem's assumptions and implications.

Theorem 3.1. *Let g be of the form (1.6), and suppose that for some $a \in (d, d+2)$ and $T > 0$, V satisfies*

$$\int_0^T \left(\|\nabla V^t\|_{L^\infty} + \|\|\nabla\|^{\frac{3}{2}} \nabla V^t\|_{L^{\frac{2d}{a-2}}} \mathbf{1}_{0 \leq s < d-2} + \|\|\nabla\|^{\frac{d-s}{2}} V^t\|_{L^{\frac{2d}{d-s}}} \mathbf{1}_{-2 < s < 0} \right) dt < \infty. \quad (3.4)$$

Assume the equation (3.3) admits a solution $\mu \in L^\infty([0, T], \mathcal{P}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$ such that $\int_{(\mathbb{R}^d)^2} |\mathbf{g}|(x - y) d(\mu^t)^{\otimes 2} < \infty$ for every $t \in [0, T]$ and

$$\mathcal{N}(u^T) := \int_0^T \left(\|\nabla u^t\|_{L^\infty} + \|\nabla\|_{\frac{3}{2}} u^t\|_{L^{\frac{2d}{d-2}}} \mathbf{1}_{0 \leq s < d-2} + \|\nabla\|^{\frac{d-s}{2}} u^t\|_{L^{\frac{2d}{d-s}}} \mathbf{1}_{-2 < s < 0} \right) dt < \infty, \quad (3.5)$$

where $u^t := -\mathbb{M}\nabla\mathbf{g} * \mu^t + \mathbf{V}^t$.

There exists a constant $C > 0$, depending only on \mathbf{d}, \mathbf{s} , such that for any solution X_N of (3.1),

$$\begin{aligned} \mathbf{F}_N(X_N^t, \mu^t) + \frac{\log(N\|\mu^t\|_{L^\infty})}{2N\mathbf{d}} \mathbf{1}_{\mathbf{s}=0} + C\|\mu^t\|_{L^\infty}^{\frac{5}{d}} N^{\frac{5}{d}-1} \\ \leq C e^{C\mathcal{N}(u^t)} \left(\mathbf{F}_N(X_N^0, \mu^0) + \sup_{\tau \in [0, t]} \left(\frac{\log(N\|\mu^\tau\|_{L^\infty})}{2N\mathbf{d}} \mathbf{1}_{\mathbf{s}=0} + C\|\mu^\tau\|_{L^\infty}^{\frac{5}{d}} N^{\frac{5}{d}-1} \right) \right). \end{aligned} \quad (3.6)$$

Consequently, if $\lim_{N \rightarrow \infty} \mathbf{F}_N(X_N^0, \mu^0) = 0$, then $\mu_N^t \rightharpoonup \mu^t$ for every $t \in [0, T]$.

Remark 3.2. When $\mathbf{V}^t = 0$, the norm $\|\mu^t\|_{L^\infty}$. Consequently,

$$\sup_{\tau \in [0, t]} \left(\frac{\log(N\|\mu^\tau\|_{L^\infty})}{2N\mathbf{d}} \mathbf{1}_{\mathbf{s}=0} + C\|\mu^\tau\|_{L^\infty}^{\frac{5}{d}} N^{\frac{5}{d}-1} \right) \leq \frac{\log(N\|\mu^0\|_{L^\infty})}{2N\mathbf{d}} \mathbf{1}_{\mathbf{s}=0} + C\|\mu^0\|_{L^\infty}^{\frac{5}{d}} N^{\frac{5}{d}-1}. \quad (3.7)$$

Remark 3.3. The reader may check that the above condition for $\mathbf{V}^t = -c^t x + \Phi^t(x)$ is satisfied, where $\int_0^T |c^t| dt < \infty$, and $\int_0^T (\|\nabla\Phi^t\|_{L^\infty} + \|\nabla\|_{\frac{3}{2}} \Phi^t\|_{L^{\frac{2d}{d-2}}} \mathbf{1}_{\mathbf{a}>2}) dt < \infty$. The regularity assumption (3.5) for μ^t follows from a number of works of in the PDE literature. We refer to [138, Section 1.3] for references in the (super-)Coulomb case and [84, Section 5.2] for direct verification in the sub-Coulomb case.

Remark 3.4. By a standard argument (see, e.g., [138, Remark 3.7] or [127, Remark 1.5] for details), Theorem 3.1 implies propagation of chaos for the system (3.1). However, this ‘‘global-to-local’’ argument in general leads to a suboptimal rate [94].

3.3. Modulated energy and commutators

Since the empirical measure μ_N^t associated with the microscopic system (3.1) is a weak solution of the mean-field equation (3.3), a natural approach for showing mean-field convergence is a weak-strong stability estimate for (3.3). For Riesz interactions, such stability is conveniently formulated in terms of the modulated energy due to its coercivity, vanishing of the modulated energy metrizes the convergence of μ_N^t to μ^t .

The modulated energy approach to mean-field convergence consists of establishing a Grönwall relation for $\mathbf{F}_N(X_N^t, \mu^t)$, where X_N^t is a solution of the microscopic system (3.1) and μ^t is a solution of the mean-field equation (3.3). The key observation is that $\mathbf{F}_N(X_N^t, \mu^t)$ satisfies the dissipation inequality (see, e.g., [138, Section 3.1])

$$\frac{d}{dt} \mathbf{F}_N(X_N^t, \mu^t) \leq \int_{(\mathbb{R}^d)^2 \setminus \Delta} \nabla\mathbf{g}(x - y) \cdot (u^t(x) - u^t(y)) d(\mu_N^t - \mu^t)^{\otimes 2}, \quad (3.8)$$

where u^t is the mean-field velocity from (3.5). The right-hand side is exactly a first-order commutator expression with $v = u^t$. Importantly, the transport only depends on the mean-field density μ^t , not the empirical measure μ_N^t . The beauty of the modulated-energy method is that with this inequality in hand, one can apply the estimates of Theorem 2.1 and conclude by Grönwall’s lemma. For the exact details on closing the differential inequality and obtaining the estimate (3.6) stated in Theorem 3.1, we refer to [131, Section 6] or [84, Section 5.1].

3.4. Optimal regularity assumptions

The assumption that $\mu^t \in L^\infty$ is not so important and can be greatly relaxed at the cost of worsening the rate of convergence. More fundamental is the condition (3.5) in the statement of Theorem 3.1, needed to apply Theorem 2.1, which enforces some regularity for μ^t . Given the relationship between the scaling-critical regularity and that for which uniqueness is expected to hold for (3.3), it is natural to ask if the result holds when μ^t belongs to the critical space $\dot{W}_p^{\frac{d}{p} + s + 2 - d, p}$

for $1 < p < \infty$. Provided that $p \leq \frac{2d}{d-s-2}$, one has $\|\nabla^{\frac{d-s}{2}} u^t\|_{L^{\frac{2d}{d-s-2}}} \lesssim \|\mu^t\|_{\dot{W}^{\frac{d}{p}+s+2-d,p}}$. But in general, ∇u^t is only in BMO when $\mu^t \in \dot{W}^{\frac{d}{p}+s+2-d,p}$. As we observed in [123, 124], even in this borderline regularity case, one can still obtain a Grönwall-type⁶ estimate for the modulated energy, which is weaker than (3.6) but still suffices for proving mean-field convergence. Tying back into the discussion of Section 2.4, the key ingredient was a defective commutator estimate. These works [123, 124] considered only the Coulomb case, where the natural scale-invariant space is L^∞ , and had a short-time restriction for $d \geq 3$.

In [84, Theorem 2.8], we used our new defective commutator estimate to generalize the results of [123, 124] to the entire Riesz case without any restrictions on the interval of time for which mean-field convergence holds. In the sub-Coulomb case, we showed that if the initial mean-field density $\mu^\circ \in L^{\frac{d}{d-s-2}}$, which is scaling-critical, then there is a unique global solution to the mean-field equation (3.3) which is a (quantitative) mean-field limit of the microscopic system (3.1). For details on closing the resulting differential inequality for the modulated energy after applying the defective commutator estimate, we refer to [84, Sections 2.3, 4.4].

3.5. Positive temperature

So far, we have only discussed the zero-temperature case $\beta = \infty$. At positive temperature $\beta < \infty$, where the dynamics of (3.1) are now stochastic, the pure modulated-energy approach described above can be extended, considering now moments of the modulated energy, at least in the sub-Coulomb case $s < d - 2$. This restriction is due to the Itô correction to the evolution of the energy, which is not clear how to suitably control when $s \geq d - 2$ (barely missing the Coulomb case). We refer to [82, 84, 127] for details on this extension and the resulting convergence rates.

For gradient dynamics (i.e. $\mathbb{M} = -\mathbb{I}$) at positive temperature, Bresch et al. [17, 18, 19] beautifully observed that the modulated energy combines naturally with the previously used relative entropy [48, 64, 90, 128] in the form of the *modulated free energy*

$$\begin{aligned} \mathbb{E}_N(f_N^t, \mu^t) &:= \frac{1}{\beta N} \mathbb{H}(f_N^t \mid (\mu^t)^{\otimes N}) + \mathbb{E}_{X_N \sim f_N^t} \left[\mathbb{F}_N(X_N, \mu^t) \right] \\ &= \frac{1}{N} \mathbb{H}(f_N^t \mid \mathbb{Q}_{N,\beta}(\mu^t)) - \frac{\log \mathbb{K}_{N,\beta}(\mu^t)}{N}, \end{aligned} \quad (3.9)$$

where \mathbb{H} is the relative entropy/KL divergence. The second equality is the reformulation in terms of the modulated Gibbs measure and partition function introduced by the author and Serfaty [130],

$$d\mathbb{Q}_{N,\beta}(\mu) := \frac{e^{-\beta N \mathbb{F}_N(X_N, \mu)}}{\mathbb{K}_{N,\beta}(\mu)} d\mu^{\otimes N}, \quad \mathbb{K}_{N,\beta}(\mu) := \int_{(\mathbb{R}^d)^N} e^{-\beta N \mathbb{F}_N(X_N, \mu)} d\mu^{\otimes N}(X_N). \quad (3.10)$$

The quantity (3.9) is well-suited to proving entropic propagation of chaos [17, 18, 19, 23, 33, 128, 130] and can even handle logarithmically attractive interactions [17, 19, 34]. Commutator estimates play the same essential role in the positive temperature setting. More precisely, if $f_N^t := \text{Law}(X_N^t)$, the original BJW quantity can be written in our notation as

4. Supercritical mean-field limits

We conclude with the second application of our sharp commutator estimates to supercritical mean-field limits. Namely, these estimates combined with a modulated-energy scheme allow to derive the Lake equation from the Newtonian N -body problem in the joint mean-field and quasineutral limit for (super-)Coulomb Riesz interactions under optimal scaling assumptions. The material in this section is drawn from the joint work [129] with Serfaty.

⁶More precisely, an estimate in the form of Osgood's lemma (see, e.g., [2, Section 3.1.1]).

4.1. Background and motivation

Consider a *Newtonian system* of N particles with a pairwise interaction potential \mathbf{g} and external confining potential V :

$$\begin{cases} \dot{x}_i^t = v_i^t \\ \dot{v}_i^t = -\gamma v_i^t - \frac{1}{\varepsilon^2 N} \sum_{1 \leq j \leq N: j \neq i} \nabla \mathbf{g}(x_i^t - x_j^t) - \frac{1}{\varepsilon^2} \nabla V(x_i^t) \\ (x_i^0, v_i^0) = (x_i^\circ, v_i^\circ), \end{cases} \quad 1 \leq i \leq N. \quad (4.1)$$

The positions and velocities are assumed to belong to \mathbb{R}^d . Here, $\gamma \geq 0$ is the *friction coefficient*, and $\varepsilon > 0$ is a small parameter, possibly depending on N , which encodes physical information about the system. In contrast to the first-order dynamics in (3.1), the dynamics in (4.1) are now second-order. For concreteness and because the results will ultimately be for this class, we will limit our attention to (Coulomb/super-Coulomb) Riesz potentials. Though, much of the discussion at this stage makes sense for more general interaction potentials.

We are interested in the large N and small ε limit of (4.1), the latter of which we interpret as a *quasineutral limit* elaborated on in the next subsection. There are two (mathematically equivalent) motivations for our setup.

The first motivation is non-neutral plasmas [114] (see also [107, 147] for relevance to trapped neutral systems). The system (4.1) models the evolution of a trapped system of ions near thermodynamic equilibrium, meaning the spatial density $\mu_N^t := \frac{1}{N} \sum_{i=1}^N \delta_{x_i^t}$ is close to the *equilibrium measure* μ_V . This equilibrium measure is defined as the probability measure that minimizes the macroscopic or mean-field energy

$$\mathcal{E}(\mu) := \int_{\mathbb{R}^d} V d\mu + \frac{1}{2} \int_{(\mathbb{R}^d)^2} \mathbf{g}(x - y) d\mu(x) d\mu(y). \quad (4.2)$$

We refer to [140, Chap. 2] for details on the equilibrium measure in this Coulomb/Riesz context, as well as for its connection to the solution of the fractional obstacle problem, which will be also discussed below. One can see the equilibrium measure as a generalization of the uniform measure on a torus when considering an infinitely extended trapped system.

The second motivation comes from two-species globally neutral systems. In this setting, the empirical spatial density μ_N^t is close to a fixed density μ , representing the density of a stationary background of an oppositely charged species of particles (e.g. heavy positively charged ions). Alternatively, one may think of this as a one-component plasma with a nonuniform background. In this case, the term $-\nabla V(x_i^t)$ on the right-hand side of the second line of (4.1) should be replaced by the attractive force $+\nabla(\mathbf{g} * \mu)(x_i^t)$ due to the background.

The settings for each of these motivations are mathematically equivalent. The latter corresponds to $V = -\mathbf{g} * \mu + \frac{1}{2} \int_{(\mathbb{R}^d)^2} \mathbf{g}(x - y) d\mu^{\otimes 2}$. While the former corresponds to considering a background density such that $\nabla(\mathbf{g} * \mu + V) = 0$ in the support of μ . This is in particular achieved when (but not only when) μ is equal to the equilibrium measure μ_V .

Under suitable assumptions on the external potential V , our goal is to show that if the initial *empirical spatial density* $\mu_N^\circ := \frac{1}{N} \sum_{i=1}^N \delta_{x_i^\circ} \rightarrow \mu_V$ as $N \rightarrow \infty$ and the initial velocities $v_i^\circ \approx u^\circ(x_i^\circ)$, for a macroscopic vector field u° on \mathbb{R}^d , then the total empirical measure $\frac{1}{N} \sum_{i=1}^N \delta_{(x_i^t, v_i^t)}$ associated to a solution of (4.1) converges as $\varepsilon \rightarrow 0$ and $N \rightarrow \infty$ to the *monokinetic* measure $\mu_V(x) \delta_{u^t(x)}(v)$, where u^t satisfies the *Lake equation*

$$\begin{cases} \partial_t u + \gamma u + u \cdot \nabla u = -\nabla p \\ \operatorname{div}(\mu_V u) = 0. \end{cases} \quad (4.3)$$

In particular, when \mathbf{g} is the Coulomb potential, we give a microscopic counterpart to the proof of the *quasineutral* limit for Vlasov–Poisson with monokinetic data by Barré et al. [3].

Note that if μ_V is constant and $\gamma = 0$, then (4.3) is nothing but the *incompressible Euler equation*. The *pressure* p is a Lagrange multiplier to enforce the incompressibility constraint $\operatorname{div}(\mu_V u) = 0$. Multiplying both sides of the first equation of (4.3) by μ_V and taking the divergence, the pressure p is obtained from the velocity u by solving the divergence-form elliptic equation

$$-\operatorname{div}(\mu_V \nabla p) = \operatorname{div}^2(\mu_V u^{\otimes 2}) = \operatorname{div}(\mu_V u \cdot \nabla u). \quad (4.4)$$

Equation (4.3), which is also sometimes called the *anelastic equation*, appears in the modeling of atmospheric flows [104, 113] and superconductivity [29, 40, 43] and has been mathematically studied in [3, 40, 99, 100]. The equation also arises as a mean-field limit for Ginzburg–Landau vortices with pinning and forcing [43] have shown that the equation. Additionally, the Lake equation has been shown [105] to be a mean-field limit for a model of vortices in shallow water with varying topography [119].

4.2. The combined mean-field and quasineutral limit

To see how the equation (4.3) appears as a formal limiting dynamics for the empirical measure of (4.1), we argue as follows.

Suppose that the parameter $\varepsilon > 0$ is fixed. Analogous to the derivation of the equation (3.3) in Section 3.1, the mean-field limit in this context refers to the convergence as $N \rightarrow \infty$ of the total empirical measure $f_{N,\varepsilon}^t := \frac{1}{N} \sum_{i=1}^N \delta_{z_i^t}$, where $z_i^t := (x_i^t, v_i^t)$, to a solution f_ε^t of the *Vlasov equation with friction*

$$\begin{cases} \partial_t f_\varepsilon + v \cdot \nabla_x f_\varepsilon - \frac{1}{\varepsilon^2} \nabla(V + \mathbf{g} * \mu_\varepsilon) \cdot \nabla_v f_\varepsilon - \operatorname{div}_v(\gamma v f) = 0 \\ \mu_\varepsilon = \int_{\mathbb{R}^d} df_\varepsilon(\cdot, v) \\ f_\varepsilon|_{t=0} = f_\varepsilon^\circ. \end{cases} \quad (4.5)$$

When \mathbf{g} is Coulomb and $\gamma = 0$, the equation (4.5) is known as *Vlasov–Poisson*. More generally, for \mathbf{g} as in (1.6), the equation is called *Vlasov–Riesz*. It is a difficult problem to derive the Vlasov–Poisson/Vlasov–Riesz equation directly from (4.1). While the case of regular potentials \mathbf{g} (e.g. globally $C^{1,1}$) [11, 38, 41, 111] or even just potentials with bounded force $\nabla \mathbf{g}$ [88] is well understood, the Coulomb case in general remains out of reach except in dimension 1 [76, 143]. To our knowledge, the best results for singular potentials are limited to forces $\nabla \mathbf{g}$ which are square integrable at the origin [15, 77, 78] (which barely misses the 2D Coulomb case) or are for Coulomb potentials with short-distance vanishing cutoff [7, 46, 47, 54, 96, 97].⁷ Recently, Duerinckx and Serfaty [138, Appendix] proved the mean-field limit for the Vlasov–Poisson equation (and more generally super-Coulomb Vlasov–Riesz⁸) for monokinetic/cold initial data, for which the Vlasov–Poisson equation reduces to the *pressureless Euler–Poisson equation*. Putting aside this question of rigorously showing the mean-field limit, we seek to formally derive the Lake equation (4.3) from the Vlasov equation (4.5) in the limit as $\varepsilon \rightarrow 0$.

The regime under consideration is when the spatial density μ_ε^t converges to the equilibrium measure μ_V as $\varepsilon \rightarrow 0$ (this is an assumption). Decomposing the potential

$$V + \mathbf{g} * \mu_\varepsilon = (V + \mathbf{g} * \mu_V) + \mathbf{g} * (\mu_\varepsilon - \mu_V), \quad (4.6)$$

the fact that, by characterization of the equilibrium measure, $V + \mathbf{g} * \mu_V$ is constant on the support of μ_V implies that

$$\nabla(V + \mathbf{g} * \mu_\varepsilon) = \nabla \mathbf{g} * (\mu_\varepsilon - \mu_V), \quad x \in \operatorname{supp} \mu_V. \quad (4.7)$$

Assuming that the renormalized electric potential difference $\frac{1}{\varepsilon^2} \mathbf{g} * (\mu_\varepsilon - \mu_V)$ has a weak limit p as $\varepsilon \rightarrow 0$, we see that the weak limit $f := \lim_{\varepsilon \rightarrow 0} f_\varepsilon$ satisfies the equation

$$\begin{cases} \partial_t f + v \cdot \nabla_x f - \nabla p \cdot \nabla_v f - \operatorname{div}_v(\gamma v f) = 0 \\ \mu_V = \int_{\mathbb{R}^d} df(\cdot, v) \\ f|_{t=0} = f^\circ, \end{cases} \quad (4.8)$$

where f° is the weak limit of f_ε° .

Let us now define the *current* $J(x) := \int_{\mathbb{R}^d} v df(x, v)$ associated to (4.8). Integrating both sides of the first equation in (4.8) with respect to v , then using that the spatial density is equal to μ_V for all time, we find that $\operatorname{div} J^t$ is constant in time. After some calculus (see [129, Section 1.1] for the computation), one finds that

$$\partial_t J + \operatorname{div} \int_{\mathbb{R}^d} v^{\otimes 2} df(\cdot, v) = -\gamma J - \mu_V \nabla p. \quad (4.9)$$

⁷If one adds noise to the velocity equation in (4.1), corresponding to the *Vlasov–Fokker–Planck* mean-field limit, then the 2D Coulomb case has been shown [16]. See also [42] for some Riesz cases.

⁸Combining the total modulated energy introduced in this work with the commutator estimate Theorem 2.1(2.5), one can extend this result to the sub-Coulomb Vlasov–Riesz equation as well.

This equation is not closed in terms of (J, p) , since the second term on the left-hand side requires knowledge of the second velocity moment of f , which in turn depends on third moment and so on (this is the famous closure problem for moments of the Vlasov equation; see, e.g., [145]). But making the *monokinetic* or “*cold electrons*”⁹ ansatz $f(x, v) = \mu_V(x)\delta(v - u(x))$, it follows that $J = \mu_V u$. Since μ_V is independent of time, substituting this identity into (4.9) yields

$$\begin{cases} \mu_V \partial_t u + \operatorname{div}(\mu_V u^{\otimes 2}) = -\mu_V(\gamma u + \nabla p), \\ \operatorname{div}(\mu_V u) = 0. \end{cases} \quad (4.10)$$

Assuming that μ_V is positive on its support, we see that (4.10) is equivalent to (4.3).

In the plasma physics setting of Vlasov–Poisson, the limit $\varepsilon \rightarrow 0$ is called the *quasineutral limit*, and the equation (4.8), in the case $\mu_V \equiv 1$, is called the *kinetic incompressible Euler (KIE) equation* [12]. The inhomogeneous case of (4.8) does not seem to have previously appeared in the literature. Nowhere in the above reasoning did we assume a specific form for \mathbf{g} (e.g. Coulomb). This demonstrates a certain *universality* of the KIE for this kind of singular limit, which appears to be new. In this plasma physics setting, the distribution function f models the evolution of electrons against a stationary background of positively charged ions. After a rescaling to dimensionless variables, the parameter ε corresponds to the *Debye (screening) length* of the system, which is the scale at which charge separation in the plasma occurs. When the Debye length is much smaller than the length scale of observation, the plasma is said to be quasineutral, since it appears neutral to an observer. The rigorous justification of the quasineutral limit is a difficult problem and has been studied in [13, 14, 55, 56, 57, 61, 63, 66, 67, 68, 70, 103]. We refer to the survey [62] and references therein for further discussion.

The preceding formal calculations suggest that in the limit as $N + \varepsilon^{-1} \rightarrow \infty$, which one can physically interpret as a combined mean-field and quasineutral limit, the empirical measure $f_{N,\varepsilon}^t$ of the Newtonian system (4.1) converges to a solution f^t of the KIE (4.8), which reduces to the Lake equation (4.3) for monokinetic solutions. Thus, we expect that if the particle velocities $v_i^t \approx u^t(x_i)$, then the empirical measure $f_{N,\varepsilon}^t$ converges to the measure $\delta_{u^t(x)}(v)\mu_V(x)$ as $N + \varepsilon^{-1} \rightarrow \infty$, where u^t is a solution of (4.10). A more general interpretation of the limit as $N + \varepsilon^{-1} \rightarrow \infty$ is as a *supercritical mean-field limit* of the system (4.1). This terminology coined by Han-Kwan and Iacobelli [69] refers to the fact that the force experienced by a single particle in (4.1) formally diverges as $N \rightarrow \infty$, compared to being $O(1)$ for the usual $1/N$ mean-field scaling.

4.3. Scaling assumptions and prior work

The convergence of $f_{N,\varepsilon}^t$ to $\delta_{u^t(x)}(v)\mu_V(x)$ was previously shown by Han-Kwan and Iacobelli [69] in the spatially periodic Coulomb case (i.e. $x \in \mathbb{T}^d$) when $V = 0$ and $\mu_V \equiv 1$ assuming $\varepsilon N^{\frac{1}{d(d+1)}} \rightarrow \infty$ as $\varepsilon + N^{-1} \rightarrow 0$, where u^t is a solution of the incompressible Euler equation. In this work, they recognized (building on the aforementioned work [138, Appendix] for the mean-field limit) that the modulated-energy method may be also used for this supercritical limit provided one adds a suitable $O(\varepsilon^2)$ corrector to the background spatial density in the definition of the modulated potential energy. We elaborate more on this idea in Section 4.5 below. Their proof may be viewed as a generalization of Brenier’s [13] modulated-energy approach to proving the quasineutral limit of Vlasov–Poisson with monokinetic data to allow (via renormalization) for the solution of Vlasov–Poisson to be a sum of Diracs. As in all modulated-energy approaches, the key tool is the commutator estimate (1.4) and sharper additive errors allow to treat a large range of scaling relations for ε, N .

In the same setting, this convergence was subsequently improved by the author [125] to $\varepsilon^{-2}N^{-\frac{2}{d}} \rightarrow 0$ as $\varepsilon + N^{-1} \rightarrow 0$ using a sharp commutator estimate for the Coulomb case [98, 125, 139]. Sharper estimates in terms of the dependence on the solution u of Euler’s equation were also shown. [125] conjectured that the work’s scaling assumption should be in general optimal, in the sense that the incompressible Euler equation should not be the limiting evolution of the empirical measure when $\varepsilon^{-2} \ll N^{-\frac{2}{d}}$.

⁹This terminology from in the physics literature stems from the fact that the temperature of the distribution is zero.

These previous works are limited to the Coulomb case on the torus, which is an idealized setting since it assumes a spatially uniform density. Moreover, they left open the question of a rigorous justification of the universality of the Lake equation with respect to the interaction and confinement in the sense that it only depends on the equilibrium measure. Finally, they left open the optimality of the scaling relation between ε, N .

4.4. Informal statement of main results

In [129] joint with Serfaty, we settled the above questions for (super-)Coulomb Riesz interactions. To present the main result of that work, we introduce the *total modulated energy*

$$\mathbf{H}_N(Z_N^t, u^t) := \frac{1}{2N} \sum_{i=1}^N |v_i^t - u^t(x_i^t)|^2 + \frac{1}{\varepsilon^2} \mathbf{F}_N(X_N^t, \mu_V + \varepsilon^2 \mathfrak{U}^t) + \frac{1}{\varepsilon^2 N} \sum_{i=1}^N \zeta(x_i^t). \quad (4.11)$$

Here, Z_N^t is a solution of the N -particle system (4.1). The vector field u^t is not a solution of the Lake equation (4.3) but rather an extension from $\text{supp } \mu_V$ to all of \mathbb{R}^d , such that the regularity is preserved and $\text{div}(\mu_V u) = 0$ (see [129, Section 4]). The physical Cauchy problem for (4.3) is in the domain $\text{supp } \mu_V$ subject to a no-flux boundary condition, but the microscopic dynamics Z_N^t are not confined to $\text{supp } \mu_V$. The extension allows us to compare between the two settings.

The first term of (4.11) is the *modulated kinetic energy*. The second term is the *modulated potential energy* previously seen in (1.2). Here, \mathfrak{U}^t is an N -independent *corrector*, whose precise definition the reader may find in [129, Eq. (5.2)] (see also Remark 5.1 there). The third term is

$$\zeta := \mathbf{g} * \mu + V - c \quad (4.12)$$

for Robin constant c , which is nonnegative and vanishes on $\text{supp } \mu_V$. The quantity (4.11) depends on both ε, N ; but since we view $\varepsilon = \varepsilon(N)$, we omit the ε dependence to lighten the notation.

Because the precise, rigorous version requires stating some rather cumbersome regularity assumptions, we only present an informal version of the main result in Theorem 4.1 below. The interested reader may find the precise statement in [129, Theorem 5.2] and also a generalization of our findings to the case of *regular* interactions in Appendix A of the cited work.

Theorem 4.1 (Informal). *Suppose that the equilibrium measure μ_V is sufficiently regular and $\Sigma := \text{supp } \mu_V$ is a sufficiently smooth domain. Assume u is a sufficiently regular solution to (4.3) on $[0, T]$. Then there exist continuous functions $C_1, C_2, C_3 : [0, T] \rightarrow \mathbb{R}_+$, which depend on $\mathbf{d}, \mathbf{s}, \gamma$, and norms of u , such that for any solution $Z_N^t = (X_N^t, V_N^t)$ of (4.1), it holds that*

$$\mathbf{H}_N(Z_N^t, u^t) + \frac{\log N}{2\mathbf{d}N\varepsilon^2} \mathbf{1}_{\mathbf{s}=0} \leq e^{C_1} \left(\mathbf{H}_N(Z_N^0, u^0) + \frac{\log N}{2\mathbf{d}N\varepsilon^2} \mathbf{1}_{\mathbf{s}=0} + \frac{C_2^t N^{\frac{\mathbf{d}}{2}-1}}{\varepsilon^2} + C_3^t \varepsilon^2 \right). \quad (4.13)$$

In particular, if

$$\lim_{\varepsilon + \frac{1}{N} \rightarrow 0} \left(\mathbf{H}_N(Z_N^0, u^0) + \frac{\log N}{2\mathbf{d}N\varepsilon^2} \mathbf{1}_{\mathbf{s}=0} \right) = \lim_{\varepsilon + \frac{1}{N} \rightarrow 0} \frac{N^{\frac{\mathbf{d}}{2}-1}}{\varepsilon^2} = 0, \quad (4.14)$$

then we have the weak-* convergence

$$\frac{1}{N} \sum_{i=1}^N \delta_{z_i^t} \xrightarrow{\varepsilon + \frac{1}{N} \rightarrow 0} \delta_{u^t(x)}(v) \mu_V(x), \quad \forall t \in [0, T]. \quad (4.15)$$

The scaling assumption $N^{\frac{\mathbf{d}}{2}-1}/\varepsilon^2 \rightarrow 0$ in (4.14) is in general optimal, in the sense that there exists a sequence of solutions Z_N^t to (4.1) such that $\frac{1}{N} \sum_{i=1}^N \delta_{z_i^t} \rightharpoonup \delta_{u^t(x)}(v) \mu_V(x)$ as $\varepsilon + \frac{1}{N} \rightarrow 0$, but the total modulated energy $\mathbf{H}_N(Z_N^t, u^t)$ does not vanish. This follows from essentially the same reasoning explained in Section 3 for the sharp rate of convergence for the mean-field limit in the modulated energy distance. Namely, if X_N° is a critical point of the microscopic energy

$$\mathcal{H}_N(X_N) := \frac{1}{2N} \sum_{1 \leq i \neq j \leq N} \mathbf{g}(x_i - x_j) + \sum_{i=1}^N V(x_i), \quad (4.16)$$

then $Z_N^t := (x_i^t, v_i^t)_{i=1}^N$ with $(x_i^t, v_i^t) := (x_i^\circ, 0)$ is the unique (stationary) solution of (4.1) with $\gamma = 0$, which is moreover *independent* of ε . In particular, if X_N° is a minimizer of \mathcal{H}_N , then by

the results of [116, 133, 135, 136] (see [140, Chapters 11-12] for a consolidated treatment), it holds that $\frac{1}{N} \sum_{i=1}^N \delta_{x_i^\circ} \xrightarrow{N \rightarrow \infty} \mu_V$ and

$$\frac{1}{N} \sum_{i=1}^N \underbrace{\zeta(x_i^\circ)}_{=0} + \mathbf{F}_N(X_N^\circ, \mu_V) + \frac{\log N}{2dN} \mathbf{1}_{s=0} = C_{d,s}^V N^{\frac{s}{d}-1} + o(N^{\frac{s}{d}-1}) \quad \text{as } N \rightarrow \infty, \quad (4.17)$$

where $C_{d,s}^V$ is a computable constant depending only on d, s, V that encodes thermodynamic information at the microscale. That $\zeta(x_i^\circ) = 0$ for each i follows from the fact that minimizing point configurations lie in $\text{supp } \mu_V$, on which ζ vanishes. Hence,

$$\mathbf{H}_N(Z_N^t, 0) + \frac{\log N}{2dN\varepsilon^2} \mathbf{1}_{s=0} = \frac{1}{\varepsilon^2} \left(\mathbf{F}_N(X_N^\circ, \mu_V) + \frac{\log N}{2dN} \mathbf{1}_{s=0} \right) = C_{d,s}^V \frac{N^{\frac{s}{d}-1}}{\varepsilon^2} + o\left(\frac{N^{\frac{s}{d}-1}}{\varepsilon^2}\right), \quad (4.18)$$

and we cannot expect vanishing of the total modulated energy in the supercritical mean-field regime if $\varepsilon \leq N^{\frac{s-d}{2d}}$.

In light of the preceding observations, one asks what is the effective equation describing the system (4.1) as $\varepsilon + \frac{1}{N} \rightarrow 0$, assuming that $\varepsilon \leq N^{\frac{s-d}{2d}}$. Naively, one might expect that when $\varepsilon \ll N^{\frac{s-d}{2d}}$, this behaves like first sending $\varepsilon \rightarrow 0$ and then $N \rightarrow \infty$. However, this limit does not make sense in general. Indeed, multiplying both sides of the second equation of (4.1) by ε^2 and letting $\varepsilon \rightarrow 0$ for fixed N , we formally see that the limiting positions X_N^t should be a critical point of the microscopic energy \mathcal{H}_N . If each limiting velocity v_i^t is nonzero, we would not expect X_N^t to remain a critical point for all t . On the other hand, there is no mechanism to force the velocities to tend to zero even if the initial positions are a critical point of the energy (4.16). In Section 4.6, we demonstrate this limit fails for the exactly solvable 1D Coulomb case precisely when $\varepsilon^{-2} N^{\frac{s}{d}-1} = (\varepsilon N)^{-2}$ does not vanish, showing that there need not be any weak limit for the empirical measure, even in the simplest setting.

Remark 4.2. Given u° , one can produce statistically generic examples of initial data Z_N° such that (4.14) holds. See [129, Remark 1.2] for details.

4.5. Method of proof

We give some comments on the method of proof of Theorem 4.1. The complete details may be found in [129, Section 5].

The quantity (4.11) is a variant of the total modulated energy originally introduced by Duerinckx and Serfaty [138, Appendix] to treat the mean-field limit for Vlasov–Riesz in the monokinetic regime. The incorporation of the time-dependent corrector \mathfrak{U}^t in the modulated potential energy is inspired by [69]. The scaling by ε^2 in $\mu_V + \varepsilon^2 \mathfrak{U}^t$ reflects $O(\varepsilon^2)$ fluctuations around the macroscopic equilibrium spatial density μ_V . The last term in (4.11) is a new contribution of [129] and reflects that the microscopic system (4.1) is confined by an external potential V , as opposed to being set in a compact domain, e.g. \mathbb{T}^d as in [69, 125]. Although the ζ term appears to be only $O(1)$, it is in fact zero if the particles remain in the support of μ_V , i.e. the quasineutral assumption is propagated. In analogy to the relationship between [69] and [13], the total modulated energy (4.11) may be viewed as a renormalization of the total modulated energy from [3], so as to allow for the Vlasov solution $f_\varepsilon^t = \frac{1}{N} \sum_{i=1}^N \delta_{z_i^t}$.

As with all modulated-energy approaches, the proof is based on establishing a Grönwall relation for the total modulated energy (4.11). The time derivative of this quantity has several terms that require different consideration.

The main contribution from the modulated kinetic energy is trivially estimated using Cauchy–Schwarz. The correction $\varepsilon^2 \mathfrak{U}^t$ in the spatial density is to cancel out the contribution of the pressure in (4.3) when one differentiates the modulated kinetic energy. The main contribution from the modulated potential energy is a first-order commutator (1.4) with $v = \tilde{u}^t$, the extension of the solution u^t of the Lake equation to the whole space, and $\mu = \mu_V + \varepsilon^2 \mathfrak{U}^t$. To handle this term, we crucially rely on the sharp estimate Theorem 2.1(2.4), which allows for the scaling assumption (4.14).

A new term compared to [69, 125], coming from the contribution of the $\zeta(x_i^t)$, is

$$\frac{1}{N} \sum_{i=1}^N u^t(x_i^t) \cdot \nabla \zeta(x_i^t) = \int_{\mathbb{R}^d} u \cdot \nabla \zeta d\mu_N^t, \quad (4.19)$$

which has no commutator structure. Instead, we bound it by $\frac{C\|u^t\|_{W^{1,\infty}}}{N} \sum_{i=1}^N \zeta(x_i^t)$, which is sufficient to close the Grönwall loop. The proof of this bound uses that the no-flux condition satisfied by u^t on the boundary of $\Sigma = \text{supp } \mu_V$ (a consequence of taking the quasineutral limit) and some nontrivial results for the regularity of the free boundary for solutions of the obstacle problem for the fractional Laplacian (see [129, Appendix B]), which may be of independent interest. A term similar to (4.19) was encountered in [3] in the Coulomb case, where the μ_N^t is replaced by the spatial density μ^t of Vlasov–Poisson, but handled by ad hoc arguments using the local nature of the Laplacian.

There are also several residual terms that have the form $\int_{\mathbb{R}^d} \phi d(\mu_N - \mu_V - \varepsilon^2 \mathfrak{U})$, where ϕ is a function of u, \mathfrak{U} . These may be controlled by the modulated potential energy, thanks to its coercivity, plus errors which are $O(\varepsilon^{-2} N^{\frac{s}{d}-1})$, hence acceptable.

Combining the estimates for the various terms and appealing to the Grönwall–Bellman lemma yields the inequality (4.13). The weak convergence of the empirical measure follows from the vanishing of the total modulated energy (see, e.g., [125, Section 4.2]). This then completes the proof of Theorem 4.1.

4.6. The regime $\varepsilon^2 \lesssim N^{\frac{s}{d}-1}$ for the 1D Coulomb gas

Consider the 1D Coulomb case with quadratic confinement (i.e. $s = -1$, $\mathbf{g}(x) = -|x|$, $V(x) = |x|^2$). Under the assumption that the initial empirical spatial density converges to the equilibrium measure, there is in general no weak limit for the empirical current as $\varepsilon + N^{-1} \rightarrow 0$ if $\varepsilon^{-2} N^{\frac{s}{d}-1} = \varepsilon^{-2} N^{-2}$ does not vanish. Thus, for the supercritical mean-field limit in the (super-)Coulomb Riesz case, one cannot expect convergence to the Lake equation outside of the scaling assumption (4.14) in Theorem 4.1. This example uses the exact solvability of the one-dimensional Coulomb gas with quadratic confinement, which is folklore. We refer to [129, Section 6] for the proof of the proposition stated below.

Proposition 4.3. *There exists a sequence of initial positions X_N° with $F_N(X_N^\circ, \mu_V) \rightarrow 0$ and $\mu_N^\circ \rightarrow \mu_V$ as $N \rightarrow \infty$, but that if $v_i^\circ = 0$ for each $1 \leq i \leq N$, the associated empirical current $J_N^t := \frac{1}{N} \sum_{i=1}^N v_i^t \delta_{x_i^t}$*

- *converges to zero (which is the unique solution of the Lake equation (4.10) starting from zero initial datum) uniformly on $[0, \infty)$ as $\varepsilon + N^{-1} \rightarrow 0$ if $\varepsilon N \rightarrow \infty$,*
- *has no weak limit for any $t \in (0, \infty)$ as $\varepsilon + N^{-1} \rightarrow 0$ if $\varepsilon N \not\rightarrow \infty$.*

4.7. Sub-Coulomb case

[129] only considered the (super-)Coulomb regime for two reasons. First, at the time of writing, only the nonsharp commutator estimate of [112] was available in the sub-Coulomb case, and so (super-)Coulomb was the only case where we could show our results were sharp. Obviously, this is no longer the case, thanks to Theorem 2.1(2.5). Second, and more importantly, there does not seem to be an adequate regularity theory for the obstacle problem for higher powers of the fractional Laplacian (see [35] for some progress in this direction). If one assumes that the equilibrium measure has full support in \mathbb{R}^d , or restricts to the torus where full support is easily established under a smallness condition on the confinement, then all terms involving ζ vanish. One can then treat the sub-Coulomb case under the scaling assumption (4.14) by following the same proof in [129].

4.8. The non-monokinetic regime

The modulated-energy approach of [129] is currently limited to the monokinetic regime, as is the case for the kinetic mean-field limit. Given other approaches are capable of treating the mean-field limit for second-order systems without the monokinetic ansatz, a natural question is whether one

can derive the full KIE equation (4.8) as $\varepsilon + N^{-1} \rightarrow 0$, under some scaling relation between ε and N , drawing on these approaches instead of the modulated-energy method.

The answer to this question is not immediately clear. Per our knowledge, the only results deriving the full KIE (for the uniform density case) are due to Griffin-Pickering and Iacobelli [61, 63], where one essentially sandwiches together two results: a mean-field limit for a (regularized) Vlasov–Poisson $N \rightarrow \infty$ and the quasineutral limit of Vlasov–Poisson. This approach imposes a very restrictive condition on the relation between ε and N .

At least in the case of regular interactions (treated in [129, Appendix A]), preliminary calculations suggest that the modulated-energy approach of this paper can be extended beyond the monokinetic case, for possibly nonconstant density, through a multi-stream/multi-fluid decomposition, in the spirit of [56]. This direction is the subject of ongoing work.

References

- [1] F. ALTEKRÜGER, J. HERTRICH & G. STEIDL, “Neural Wasserstein Gradient Flows for Discrepancies with Riesz Kernels”, in *Proceedings of the 40th International Conference on Machine Learning* (A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato & J. Scarlett, eds.), Proceedings of Machine Learning Research, vol. 202, PMLR, 2023, p. 664-690.
- [2] H. BAHOURI, J.-Y. CHEMIN & R. DANCHIN, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Mathematischen Wissenschaften, vol. 343, Springer, 2011.
- [3] J. BARRÉ, D. CHIRON, T. GOUDON & N. MASMOUDI, “From Vlasov-Poisson and Vlasov-Poisson-Fokker-Planck systems to incompressible Euler equations: the case with finite charge”, *J. Éc. Polytech., Math.* **2** (2015), p. 247-296.
- [4] R. BAUERSCHMIDT, P. BOURGADE, M. NIKULA & H.-T. YAU, “The two-dimensional Coulomb plasma: quasi-free approximation and central limit theorem”, *Adv. Theor. Math. Phys.* **23** (2019), no. 4, p. 841-1002.
- [5] F. BEKERMANN, T. LEBLÉ & S. SERFATY, “CLT for fluctuations of β -ensembles with general potential”, *Electron. J. Probab.* **23** (2018), article no. 115 (31 pages).
- [6] R. J. BERMAN & M. ÖNNHEIM, “Propagation of Chaos for a Class of First Order Models with Singular Mean Field Interactions”, *SIAM J. Math. Anal.* **51** (2019), no. 1, p. 159-196.
- [7] N. BOERS & P. PICKL, “On mean field limits for dynamical systems”, *J. Stat. Phys.* **164** (2016), no. 1, p. 1-16.
- [8] S. V. BORODACHOV, D. P. HARDIN & E. B. SAFF, *Discrete energy on rectifiable sets*, Springer Monographs in Mathematics, Springer, 2019.
- [9] G. BOROT & A. GUIONNET, “Asymptotic expansion of β matrix models in the one-cut regime”, *Commun. Math. Phys.* **317** (2013), no. 2, p. 447-483.
- [10] ———, “Asymptotic expansion of β matrix models in the multi-cut regime”, *Forum Math. Sigma* **12** (2024), article no. e13 (93 pages).
- [11] W. BRAUN & K. HEPP, “The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles”, *Commun. Math. Phys.* **56** (1977), no. 2, p. 101-113.
- [12] Y. BRENIER, “Une formulation de type Vlasov-Poisson pour les equations d’Euler des fluides parfaits incompressibles”, 1989, Research Report, RR-1070, <https://hal.inria.fr/inria-00075489>.
- [13] ———, “Convergence of the Vlasov-Poisson system to the incompressible Euler equations”, *Commun. Partial Differ. Equations* **25** (2000), no. 3-4, p. 737-754.

- [14] Y. BRENIER & E. GRENIER, “Limite singulière du système de Vlasov-Poisson dans le régime de quasi neutralité: le cas indépendant du temps”, *Comptes Rendus. Mathématique* **318** (1994), no. 2, p. 121-124.
- [15] D. BRESCH, M. DUERINCKX & P.-E. JABIN, “A duality method for mean-field limits with singular interactions”, 2024, <https://arxiv.org/abs/2402.04695>.
- [16] D. BRESCH, P.-E. JABIN & J. SOLER, “A new approach to the mean-field limit of Vlasov-Fokker-Planck equations”, *Anal. PDE* **18** (2025), no. 4, p. 1037-1064.
- [17] D. BRESCH, P.-E. JABIN & Z. WANG, “On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak-Keller-Segel model”, *Comptes Rendus. Mathématique* **357** (2019), no. 9, p. 708-720.
- [18] ———, “Modulated free energy and mean field limit”, *Sémin. Laurent Schwartz, EDP Appl.* **2019-2020** (2020), article no. 2 (22 pages).
- [19] ———, “Mean field limit and quantitative estimates with singular attractive kernels”, *Duke Math. J.* **172** (2023), no. 13, p. 2591-2641.
- [20] H. BRÉZIS & T. GALLOUET, “Nonlinear Schrödinger evolution equations”, *Nonlinear Anal., Theory Methods Appl.* **4** (1980), no. 4, p. 677-681.
- [21] H. BRÉZIS & S. WAINGER, “A note on limiting cases of sobolev embeddings and convolution inequalities”, *Commun. Partial Differ. Equations* **5** (1980), no. 7, p. 773-789.
- [22] L. CAFFARELLI & L. SILVESTRE, “An extension problem related to the fractional Laplacian”, *Commun. Partial Differ. Equations* **32** (2007), no. 7-9, p. 1245-1260.
- [23] S. CAI, X. FENG, Y. GONG & Z. WANG, “Propagation of chaos for 2d log gas on the whole space”, 2024, <https://arxiv.org/abs/2411.14777>.
- [24] A. P. CALDERÓN, “Commutators, singular integrals on Lipschitz curves and applications”, in *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, Acad. Sci. Fennica, Helsinki, 1980, p. 85-96.
- [25] J. A. CARRILLO, Y.-P. CHOI & M. HAURAY, “The derivation of swarming models: mean-field limit and Wasserstein distances”, in *Collective dynamics from bacteria to crowds*, CISM International Centre for Mechanical Sciences, vol. 553, Springer, 2014, p. 1-46.
- [26] J. A. CARRILLO, L. C. F. FERREIRA & J. C. PRECIOSO, “A mass-transportation approach to a one dimensional fluid mechanics model with nonlocal velocity”, *Adv. Math.* **231** (2012), no. 1, p. 306-327.
- [27] A. CECCHIN & P. NIKOLAEV, “Convergence Rate for Fluctuations of Mean Field Interacting Diffusion and Application to 2D Viscous Vortex Model and Coulomb Potential”, 2025, <https://arxiv.org/abs/2509.01266>.
- [28] L.-P. CHAINTRON & A. DIEZ, “Propagation of chaos: a review of models, methods and applications. I. Models and methods”, *Kinet. Relat. Models* **15** (2022), no. 6, p. 895-1015.
- [29] S. J. CHAPMAN & G. RICHARDSON, “Vortex pinning by inhomogeneities in type-II superconductors”, *Phys. D: Nonlinear Phenom.* **108** (1997), no. 4, p. 397-407.
- [30] M. CHRIST & J.-L. JOURNÉ, “Polynomial growth estimates for multilinear singular integral operators”, *Acta Math.* **159** (1987), no. 1-2, p. 51-80.
- [31] R. R. COIFMAN & Y. MEYER, “Commutateurs d’intégrales singulières et opérateurs multilinéaires”, *Ann. Inst. Fourier* **28** (1978), no. 3, p. 177-202.

- [32] G. CORA & R. MUSINA, “The S -Polyharmonic Extension Problem and Higher-Order Fractional Laplacians”, *J. Funct. Anal.* **283** (2022), no. 5, article no. 109555 (33 pages).
- [33] A. CHODRON DE COURCEL, M. ROSENZWEIG & S. SERFATY, “Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows”, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **42** (2023), no. 2, p. 391-472.
- [34] ———, “The attractive log gas: Stability, uniqueness, and propagation of chaos”, *Commun. Am. Math. Soc.* **5** (2025), p. 695-773.
- [35] D. DANIELLI, A. HAJ ALI & A. PETROSYAN, “The obstacle problem for a higher order fractional Laplacian”, *Calc. Var. Partial Differ. Equ.* **62** (2023), no. 8, article no. 218 (22 pages).
- [36] T. DAUXOIS, S. RUFFO, E. ARIMONDO & M. 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)*, Lecture Notes in Physics, vol. 602, Springer, 2002, p. 1-19.
- [37] M. DELGADINO, R. GVALANI & M. ROSENZWEIG, “Entropic commutator estimates”, unpublished note.
- [38] R. L. DOBRUŠIN, “Vlasov equations”, *Funkts. Anal. Prilozh.* **13** (1979), no. 2, p. 48-58.
- [39] M. DUERINCKX, “Mean-Field Limits for Some Riesz Interaction Gradient Flows”, *SIAM J. Math. Anal.* **48** (2016), no. 3, p. 2269-2300.
- [40] ———, “Well-posedness for mean-field evolutions arising in superconductivity”, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **35** (2018), no. 5, p. 1267-1319, with an appendix jointly written with Julian Fischer.
- [41] ———, “On the size of chaos via Glauber calculus in the classical mean-field dynamics”, *Commun. Math. Phys.* **382** (2021), no. 1, p. 613-653.
- [42] M. DUERINCKX & P.-E. JABIN, “Correlation Estimates for Brownian Particles with Singular Interactions”, 2025, <https://arxiv.org/abs/2510.01507>.
- [43] M. DUERINCKX & S. SERFATY, “Mean-field dynamics for Ginzburg-Landau vortices with pinning and forcing”, *Ann. PDE* **4** (2018), no. 2, article no. 19 (172 pages).
- [44] H. ENGLER, “An Alternative Proof of the Brezis–Wainger Inequality”, *Commun. Partial Differ. Equations* **14** (1989), no. 4, p. 541-544.
- [45] E. B. FABES, C. E. KENIG & R. P. SERAPIONI, “The local regularity of solutions of degenerate elliptic equations”, *Commun. Partial Differ. Equations* **7** (1982), no. 1, p. 77-116.
- [46] M. FEISTL-HELD & P. PICKL, “On the Mean-Field Limit for the Vlasov-Poisson System”, 2025, <https://arxiv.org/abs/2504.01471>.
- [47] ———, “On the Mean-Field Limit for the Vlasov-Poisson System in Two Dimensions”, 2025, <https://arxiv.org/abs/2509.17821>.
- [48] X. FENG & Z. WANG, “Quantitative Propagation of Chaos for 2D Viscous Vortex Model with General Circulations on the Whole Space”, 2024, <https://arxiv.org/abs/2411.14266>.
- [49] P. FLAJOLET & R. SEDGEWICK, *Analytic combinatorics*, Cambridge University Press, 2009, xiv+810 pages.
- [50] F. 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*, Lecture Notes in Applied Mathematics and Mechanics, vol. 3, Springer, 2016, p. 1-144.

- [51] ———, “Mean-Field Limits in Statistical Dynamics”, 2022, <https://arxiv.org/abs/2201.02005>.
- [52] F. GOLSE & T. PAUL, “Mean-Field and Classical Limit for the N-Body Quantum Dynamics with Coulomb Interaction”, *Commun. Pure Appl. Math.* **75** (2022), no. 6, p. 1332-1376.
- [53] L. GRAFAKOS, *Modern Fourier Analysis*, third ed., Graduate Texts in Mathematics, vol. 250, Springer, 2014.
- [54] P. GRASS, “Microscopic derivation of Vlasov equations with singular potentials”, 2021, <https://arxiv.org/abs/2105.06509>.
- [55] E. GRENIER, “Defect measures of the Vlasov-Poisson system in the quasineutral regime”, *Commun. Partial Differ. Equations* **20** (1995), no. 7-8, p. 1189-1215.
- [56] ———, “Oscillations in quasineutral plasmas”, *Commun. Partial Differ. Equations* **21** (1996), no. 3-4, p. 363-394.
- [57] ———, “Limite quasineutre en dimension 1”, in *Journées “Équations aux Dérivées Partielles” (Saint-Jean-de-Monts, 1999)*, Univ. Nantes, Nantes, 1999, Exp. No. II, 8 pages.
- [58] A. GRETTON, K. M. BORGWARDT, M. RASCH, B. SCHÖLKOPF & A. SMOLA, “A Kernel Method for the Two-Sample-Problem”, in *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference* (B. Schölkopf, J. Platt & T. Hoffman, eds.), MIT Press, 2006, p. 513-520.
- [59] ———, “A kernel approach to comparing distributions”, in *AAAI’07: Proceedings of the 22nd national conference on Artificial intelligence - Volume 2*, Clay Mathematics Institute; AAAI Press; MIT Press, 2007, p. 1637-1641.
- [60] ———, “A kernel two-sample test”, *J. Mach. Learn. Res.* **13** (2012), p. 723-773.
- [61] M. GRIFFIN-PICKERING & M. IACOBELLI, “A mean field approach to the quasi-neutral limit for the Vlasov-Poisson equation”, *SIAM J. Math. Anal.* **50** (2018), no. 5, p. 5502-5536.
- [62] ———, “Recent developments on quasineutral limits for Vlasov-type equations”, 2020, <https://arxiv.org/abs/2009.14169>.
- [63] ———, “Singular limits for plasmas with thermalised electrons”, *J. Math. Pures Appl. (9)* **135** (2020), p. 199-255.
- [64] A. GUILLIN, P. L. BRIS & P. MONMARCHÉ, “Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems”, *J. Eur. Math. Soc.* **27** (2025), no. 6, p. 2359-2386.
- [65] P. HAGEMANN, J. HERTRICH, F. ALTEKRÜGER, R. BEINERT, J. CHEMSEDINE & G. STEIDL, “Posterior sampling based on gradient flows of the MMD with negative distance kernel”, 2023, <https://arxiv.org/abs/2310.03054>.
- [66] D. HAN-KWAN & M. HAURAY, “Stability issues in the quasineutral limit of the one-dimensional Vlasov-Poisson equation”, *Commun. Math. Phys.* **334** (2015), no. 2, p. 1101-1152.
- [67] D. HAN-KWAN & M. IACOBELLI, “Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension”, *J. Differ. Equations* **263** (2017), no. 1, p. 1-25.
- [68] ———, “The quasineutral limit of the Vlasov-Poisson equation in Wasserstein metric”, *Commun. Math. Sci.* **15** (2017), no. 2, p. 481-509.

- [69] ———, “From Newton’s second law to Euler’s equations of perfect fluids”, *Proc. Am. Math. Soc.* **149** (2021), no. 7, p. 3045-3061.
- [70] D. HAN-KWAN & F. ROUSSET, “Quasineutral limit for Vlasov-Poisson with Penrose stable data”, *Ann. Sci. Éc. Norm. Supér. (4)* **49** (2016), no. 6, p. 1445-1495.
- [71] K. HANSSON, “Imbedding theorems of Sobolev type in potential theory”, *Math. Scand.* **45** (1979), no. 1, p. 77-102.
- [72] D. P. HARDIN, T. LEBLÉ, E. B. SAFF & S. SERFATY, “Large deviation principles for hypersingular Riesz gases”, *Constr. Approx.* **48** (2018), no. 1, p. 61-100.
- [73] D. P. HARDIN, E. B. SAFF, R. SHU & E. TADMOR, “Dynamics of particles on a curve with pairwise hyper-singular repulsion”, *Discrete Contin. Dyn. Syst.* **41** (2021), no. 12, p. 5509-5536.
- [74] D. P. HARDIN, E. B. SAFF, B. Z. SIMANEK & Y. SU, “Next Order Energy Asymptotics for Riesz Potentials on Flat Tori”, *Int. Math. Res. Not.* **2017** (2017), no. 12, p. 3529-3556.
- [75] M. HAURAY, “Wasserstein distances for vortices approximation of Euler-type equations”, *Math. Models Methods Appl. Sci.* **19** (2009), no. 8, p. 1357-1384.
- [76] ———, “Mean field limit for the one dimensional Vlasov–Poisson equation”, *Sémin. Laurent Schwartz, EDP Appl.* **2012-2013** (2014), article no. XXI (16 pages).
- [77] M. HAURAY & P.-E. JABIN, “ N -particles approximation of the Vlasov equations with singular potential”, *Arch. Ration. Mech. Anal.* **183** (2007), no. 3, p. 489-524.
- [78] ———, “Particle approximation of Vlasov equations with singular forces: propagation of chaos”, *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 4, p. 891-940.
- [79] M. HAURAY & S. MISCHLER, “On Kac’s chaos and related problems”, *J. Funct. Anal.* **266** (2014), no. 10, p. 6055-6157.
- [80] J. HERTRICH, R. BEINERT, M. GRÄF & G. STEIDL, “Wasserstein gradient flows of the discrepancy with distance kernel on the line”, in *Scale Space and Variational Methods in Computer Vision* (L. Calatroni, M. Donatelli, S. Morigi, M. Prato & M. Santacesaria, eds.), Springer, 2023, p. 431-443.
- [81] J. HERTRICH, C. WALD, F. ALTEKRÜGER & P. HAGEMANN, “Generative sliced MMD flows with Riesz kernels”, 2023, <https://arxiv.org/abs/2305.11463>.
- [82] E. HESS-CHILDS, “Large deviation principles for singular Riesz-type diffusive flows”, 2023, <https://arxiv.org/abs/2312.02904>.
- [83] E. HESS-CHILDS, M. ROSENZWEIG & S. SERFATY, “Optimal quantization for Riesz Maximum Mean Discrepancies”, in preparation.
- [84] ———, “A sharp commutator estimate for all Riesz modulated energies”, 2025, <https://arxiv.org/abs/2511.13461>.
- [85] E. HESS-CHILDS & K. ROWAN, “Higher-order propagation of chaos in L^2 for interacting diffusions”, *Probab. Math. Phys.* **6** (2025), no. 2, p. 581-646.
- [86] J. HUANG, M. ROSENZWEIG & S. SERFATY, “Fluctuations around the mean field limit for noisy singular Riesz flows”, in preparation.
- [87] P.-E. JABIN, “A review of the mean field limits for Vlasov equations”, *Kinet. Relat. Models* **7** (2014), no. 4, p. 661-711.

- [88] P.-E. JABIN & Z. WANG, “Mean field limit and propagation of chaos for Vlasov systems with bounded forces”, *J. Funct. Anal.* **271** (2016), no. 12, p. 3588-3627.
- [89] ———, “Mean field limit for stochastic particle systems”, in *Act. Part. Vol. 1. Advances theory, Model. Appl.*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, 2017, p. 379-402.
- [90] ———, “Quantitative estimates of propagation of chaos for stochastic systems with W^{-1} -Anfty kernels”, *Invent. Math.* **214** (2018), no. 1, p. 523-591.
- [91] T. KATO & G. PONCE, “Commutator estimates and the Euler and Navier-Stokes equations”, *Commun. Pure Appl. Math.* **41** (1988), no. 7, p. 891-907.
- [92] S. KLAINERMAN, “PDE as a Unified Subject”, in *Visions in Mathematics: GAFA 2000 Special Volume, Part I* (N. Alon, J. Bourgain, A. Connes, M. Gromov & V. Milman, eds.), Birkhäuser, 2010, p. 279-315.
- [93] S. KOLOURI, K. NADJAH, S. SHAHRAMPOUR & U. SIMSEKLI, “Generalized Sliced Probability Metrics”, in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE Press, 2022, p. 4513-4517.
- [94] D. LACKER, “Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions”, *Probab. Math. Phys.* **4** (2023), no. 2, p. 377-432.
- [95] D. LACKER & L. LE FLEM, “Sharp uniform-in-time propagation of chaos”, *Probab. Theory Relat. Fields* **187** (2023), p. 433-480.
- [96] D. LAZAROVICI, “The Vlasov-Poisson dynamics as the mean field limit of extended charges”, *Commun. Math. Phys.* **347** (2016), no. 1, p. 271-289.
- [97] D. LAZAROVICI & P. PICKL, “A mean field limit for the Vlasov-Poisson system”, *Arch. Ration. Mech. Anal.* **225** (2017), no. 3, p. 1201-1231.
- [98] T. LEBLÉ & S. SERFATY, “Fluctuations of two dimensional Coulomb gases”, *Geom. Funct. Anal.* **28** (2018), no. 2, p. 443-508.
- [99] C. D. LEVERMORE, M. OLIVER & E. S. TITI, “Global well-posedness for models of shallow water in a basin with a varying bottom”, *Indiana Univ. Math. J.* **45** (1996), no. 2, p. 479-510.
- [100] ———, “Global well-posedness for the lake equations. Nonlinear Phenomena in Ocean Dynamics”, *Phys. D: Nonlinear Phenom.* **98** (1996), no. 2, p. 492-509.
- [101] D. LI, “On Kato-Ponce and fractional Leibniz”, *Rev. Mat. Iberoam.* **35** (2019), no. 1, p. 23-100.
- [102] T. S. LIM, Y. LU & J. H. NOLEN, “Quantitative propagation of chaos in a bimolecular chemical reaction-diffusion model”, *SIAM J. Math. Anal.* **52** (2020), no. 2, p. 2098-2133.
- [103] N. MASMOUDI, “From Vlasov-Poisson system to the incompressible Euler system”, *Commun. Partial Differ. Equations* **26** (2001), no. 9-10, p. 1913-1928.
- [104] ———, “Rigorous derivation of the anelastic approximation”, *J. Math. Pures Appl. (9)* **88** (2007), no. 3, p. 230-240.
- [105] M. MÉNARD, “Mean-field limit of point vortices for the lake equations”, 2023, <https://arxiv.org/abs/2309.10453>.
- [106] ———, “Mean-field limit derivation of a monokinetic spray model with gyroscopic effects”, *SIAM J. Math. Anal.* **56** (2024), no. 1, p. 1068-1113.

- [107] J. T. MENDONÇA, R. KAISER, H. TERÇAS & J. LOUREIRO, “Collective oscillations in ultracold atomic gas”, *Phys. Rev. A* **78** (2008), article no. 013408.
- [108] T. MODESTE & C. DOMBRY, “Characterization of translation invariant MMD on $\{\mathbb{Bbb R}^d\}$ and connections with Wasserstein distances”, *J. Mach. Learn. Res.* **25** (2024), article no. 237 (39 pages).
- [109] J. MOSER, “A sharp form of an inequality by N. Trudinger”, *Indiana Univ. Math. J.* **20** (1970), p. 1077-1092.
- [110] A. MÜLLER, “Integral probability metrics and their generating classes of functions”, *Adv. Appl. Probab.* **29** (1997), no. 2, p. 429-443.
- [111] H. NEUNZERT & J. WICK, “Die Approximation der Lösung von Integro-Differentialgleichungen durch endliche Punktmengeten”, in *Numerische Behandlung nicht-linearer Integrodifferential- und Differentialgleichungen (Tagung, Math. Forschungsinst., Oberwolfach, 1973)*, Lecture Notes in Mathematics, vol. 395, Springer, 1974, p. 275-290.
- [112] Q.-H. NGUYEN, M. ROSENZWEIG & S. SERFATY, “Mean-field limits of Riesz-type singular flows”, *Ars Inven. Anal.* **2022** (2022), article no. 4 (45 pages).
- [113] Y. OGURA & N. A. PHILLIPS, “Scale Analysis of Deep and Shallow Convection in the Atmosphere”, *J. Atmos. Sci.* **19** (1962), no. 2, p. 173-179.
- [114] T. M. O’NEIL & D. H. E. DUBIN, “Thermal equilibria and thermodynamics of trapped plasmas with a single sign of charge”, *Phys. Plasmas* **5** (1998), no. 6, p. 2163-2193.
- [115] L. PEILEN & S. SERFATY, “Local Laws and Fluctuations for Super-Coulombic Riesz Gases”, 2025, <https://arxiv.org/abs/2511.18623>.
- [116] M. PETRACHE & S. SERFATY, “Next order asymptotics and renormalized energy for Riesz interactions”, *J. Inst. Math. Jussieu* **16** (2017), no. 3, p. 501-569.
- [117] I. B. PORAT, “Derivation of Euler’s equations of perfect fluids from von Neumann’s equation with magnetic field”, *J. Stat. Phys.* **190** (2023), no. 7, article no. 121 (44 pages).
- [118] I. B. PORAT, J. A. CARRILLO & P.-E. JABIN, “Singular flows with time-varying weights”, 2025, <https://arxiv.org/abs/2503.02276>.
- [119] G. RICHARDSON, “Vortex motion in shallow water with varying bottom topography and zero Froude number”, *J. Fluid Mech.* **411** (2000), p. 351-374.
- [120] M. ROSENZWEIG, “Cumulants of Riesz gases”, unpublished note.
- [121] ———, “The Mean-Field Limit of Stochastic Point Vortex Systems with Multiplicative Noise”, forthcoming in *Comm. Pure Appl. Math.*, 2020, <https://arxiv.org/abs/2011.12180>.
- [122] ———, “From Quantum Many-Body Systems to Ideal Fluids”, 2021, <https://arxiv.org/abs/2110.04195>.
- [123] ———, “The Mean-Field Approximation for Higher-Dimensional Coulomb Flows in the Scaling-Critical L^∞ Space”, *Nonlinearity* **35** (2022), no. 6, p. 2722-2766.
- [124] ———, “Mean-Field Convergence of Point Vortices to the Incompressible Euler Equation with Vorticity in L^∞ ”, *Arch. Ration. Mech. Anal.* **243** (2022), no. 3, p. 1361-1431.
- [125] ———, “On the rigorous derivation of the incompressible Euler equation from Newton’s second law”, *Lett. Math. Phys.* **113** (2023), no. 1, article no. 13 (32 pages).

- [126] M. ROSENZWEIG & S. SERFATY, “Commutator estimates, Stein’s method, and the transport approach to fluctuations of Riesz gases”, in preparation.
- [127] ———, “Global-in-time mean-field convergence for singular Riesz-type diffusive flows”, *Ann. Appl. Probab.* **33** (2023), no. 2, p. 954-998.
- [128] ———, “Relative entropy and modulated free energy without confinement via self-similar transformation”, 2024, <https://arxiv.org/abs/2402.13977>.
- [129] ———, “The lake equation as a supercritical mean-field limit”, *J. Éc. Polytech., Math.* **12** (2025), p. 1019-1068.
- [130] ———, “Modulated logarithmic Sobolev inequalities and generation of chaos”, *Ann. Fac. Sci. Toulouse, Math. (6)* **34** (2025), no. 1, p. 107-134.
- [131] ———, “Sharp commutator estimates of all order for Coulomb and Riesz modulated energies”, *Commun. Pure Appl. Math.* **79** (2026), no. 2, p. 207-292.
- [132] M. ROSENZWEIG, D. SLEPČEV & L. WANG, “Wasserstein gradient flow of Maximum Mean Discrepancy with energy kernels”.
- [133] N. ROUGERIE & S. SERFATY, “Higher-dimensional Coulomb gases and renormalized energy functionals”, *Commun. Pure Appl. Math.* **69** (2016), no. 3, p. 519-605.
- [134] B. RUBIN, *Fractional integrals, potentials, and Radon transforms*, Chapman & Hall/CRC, 2024.
- [135] E. SANDIER & S. SERFATY, “1D log gases and the renormalized energy: crystallization at vanishing temperature”, *Probab. Theory Relat. Fields* **162** (2015), no. 3-4, p. 795-846.
- [136] ———, “2D Coulomb gases and the renormalized energy”, *Ann. Probab.* **43** (2015), no. 4, p. 2026-2083.
- [137] A. SEEGER, C. K. SMART & B. STREET, “Multilinear singular integral forms of Christ-Journé type”, *Mem. Am. Math. Soc.* **257** (2019), no. 1231, p. v+134.
- [138] S. SERFATY, “Mean field limit for Coulomb-type flows”, *Duke Math. J.* **169** (2020), no. 15, p. 2887-2935.
- [139] ———, “Gaussian fluctuations and free energy expansion for Coulomb gases at any temperature”, *Ann. Inst. Henri Poincaré, Probab. Stat.* **59** (2023), no. 2, p. 1074-1142.
- [140] ———, “Lectures on Coulomb and Riesz gases”, 2024, <https://arxiv.org/abs/2407.21194>.
- [141] E. M. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton Mathematical Series, vol. 30, Princeton University Press, 1970, xiv+287 pages.
- [142] A.-S. SZNITMAN, “Topics in propagation of chaos”, in *École d’Été de Probabilités de Saint-Flour XIX—1989*, Lecture Notes in Mathematics, vol. 1464, Springer, 1991, p. 165-251.
- [143] M. TROCHERIS, “On the derivation of the one-dimensional Vlasov equation”, *Transp. Theory Stat. Phys.* **15** (1986), no. 5, p. 597-628.
- [144] N. S. TRUDINGER, “On imbeddings into Orlicz spaces and some applications”, *J. Math. Mech.* **17** (1967), p. 473-483.
- [145] C. UHLEMANN, “Finding closure: approximating Vlasov-Poisson using finitely generated cumulants”, *J. Cosmol. Astropart. Phys.* **2018** (2018), no. 10, article no. 030 (23 pages).

- [146] S. WANG, “Sharp local propagation of chaos for mean field particles with $W^{-1,\infty}$ kernels”, *J. Funct. Anal.* **290** (2026), no. 3, article no. 111240 (39 pages).
- [147] D. J. WINELAND, J. J. BOLLINGER, W. M. ITANO & J. D. PRESTAGE, “Angular momentum of trapped atomic particles”, *J. Opt. Soc. Am. B* **2** (1985), no. 11, p. 1721-1730.
- [148] R. YANG, “On Higher Order Extensions for the Fractional Laplacian”, 2013, <https://arxiv.org/abs/1302.4413>.

Matthew Rosenzweig
Department of Mathematical Sciences
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA, USA
mrosenz2@andrew.cmu.edu