MEAN FIELD LIMIT FOR COULOMB-TYPE FLOWS

SYLVIA SERFATY, APPENDIX WITH MITIA DUERINCKX

ABSTRACT. We establish the mean-field convergence for systems of points evolving along the gradient flow of their interaction energy when the interaction is the Coulomb potential or a super-coulombic Riesz potential, for the first time in arbitrary dimension. The proof is based on a modulated energy method using a Coulomb or Riesz distance, assumes that the solutions of the limiting equation are regular enough and exploits a weak-strong stability property for them. The method can handle the addition of a regular interaction kernel, and applies also to conservative and mixed flows. In the appendix, it is also adapted to prove the mean-field convergence of the solutions to Newton's law with Coulomb or Riesz interaction in the monokinetic case to solutions of an Euler-Poisson type system.

1. INTRODUCTION

The derivation of effective equations for classical interacting many body systems is an important question in mathematical physics. Within it, one of the most important problems, fundamental for plasma physics, is the derivation of the Vlasov-Poisson equation from Newton's law for N particles i.e. a system of the form

$$(1.1) \qquad \begin{cases} \dot{x}_i = v_i \\ \dot{v}_i = \dfrac{1}{N} \displaystyle\sum_{i \neq j} K(x_i, x_j), \quad i = 1, \dots, N \\ x_i(0) = x_i^0 \quad v_i(0) = v_i^0 \end{cases}$$

where the pair interaction K derives from the Coulomb potential, and is still open in its full generality. The large N or mean-field limit of first order systems of the form

$$(1.2) \qquad \begin{cases} \dot{x}_i = \dfrac{1}{N} \displaystyle\sum_{i \neq j} K(x_i, x_j), \quad i = 1, \dots, N \\ x_i(0) = x_i^0 \end{cases}$$

is also a natural and interesting question. Such systems can model interacting particles in physics (for instance the point vortex system in two-dimensional fluids), from a numerical point of view they correspond to particle approximations of PDEs, or in the case of gradient flows can serve to approximate their equilibrium states. Motivations extend to biological and sociological sciences, including phenomena of flocking, swarming and aggregation (see for instance [CCH]) and the analysis of large neural networks in biology [BFT] and in machine learning [MMn, RVE, BaCh].

In the above problems, the cases where K is regular are well understood, while in contrast those where K is singular are the most difficult and least understood. In this paper we will be particularly focusing on the case where K corresponds to the Coulomb interaction (and some generalizations), arguably the most important case for physics. For further details on the general mathematical aspects of mean-field limits, we refer to the reviews [Go, Jab].

The rest of the introduction is structured as follows: in Section 1.1 we introduce the exact equations that we will study as well as their limiting effective evolution equations, and describe the state of the art on such questions. In Section 1.2 we state the main result, in Section 1.3 we comment on the assumptions, and in Section 1.4 we give an extended proof outline for the Coulomb case. Finally, Section 1.5 is an added section explaining how to treat the case of gradient flow evolutions with thermal noise.

1.1. **Problem and background.** In this paper we consider specifically systems with Coulomb, logarithmic or Riesz interaction with kernel deriving from

$$(1.3) \qquad \mathsf{g}(x) = |x|^{-\mathsf{s}} \quad \max(\mathsf{d}-2, 0) \le \mathsf{s} < \mathsf{d} \qquad \text{for any } \mathsf{d} \ge 1$$

or

$$(1.4) \qquad \mathsf{g}(x) = -\log|x| \qquad \text{for } \mathsf{d} = 1 \text{ or } 2,$$

where d is the dimension. In the case (1.3) with $\mathsf{s} = \mathsf{d} - 2$ and $\mathsf{d} \ge 3$, or (1.4) and $\mathsf{d} = 2$, g is exactly (a multiple of) the Coulomb kernel. In the other cases of (1.3) it is called a Riesz kernel. In some cases, we may add to the interaction force a regular part denoted F.

We will consider first order dynamics of the form

$$(1.5) \qquad \begin{cases} \dot{x}_i = -\dfrac{1}{N}\left(\nabla_{x_i}\mathcal{H}_N(x_1, \ldots, x_N) + \displaystyle\sum_{j \neq i} \mathsf{F}(x_i - x_j)\right), \quad i = 1, \ldots, N \\ x_i(0) = x_i^0 \end{cases}$$

or conservative evolutions of the form

$$(1.6) \qquad \begin{cases} \dot{x}_i = -\dfrac{1}{N}\mathbb{J}\nabla_{x_i}\mathcal{H}_N(x_1, \ldots, x_N), \quad i = 1, \ldots, N \\ x_i(0) = x_i^0 \end{cases}$$

where $\mathbb{J}$ is an antisymmetric matrix, the points x_i evolve in the whole space $\mathbb{R}^\mathsf{d}$ and their energy $\mathcal{H}_N$ is given by

$$(1.7) \qquad \mathcal{H}_N(x_1, \ldots, x_N) = \sum_{i \neq j} \mathsf{g}(x_i - x_j),$$

with g as above. The map $\mathsf{F} : \mathbb{R}^\mathsf{d} \to \mathbb{R}^\mathsf{d}$ is the additional interaction force that we can add in the dissipative case to illustrate the robustness of the method, we will require it to enjoy some Sobolev and Hölder regularity.

Mixed flows of the form

$$(1.8) \qquad \dot{x}_i = -\frac{1}{N}\left(\alpha I + \beta \mathbb{J}\right)\left(\nabla_{x_i}\mathcal{H}_N(x_1, \ldots, x_N) + \sum_{j \neq i} \mathsf{F}(x_i - x_j)\right) \qquad \alpha > 0$$

can be treated with exactly the same proof, so can the same dynamics (1.5), (1.6) or (1.8) with an additional forcing $\frac{1}{N}\sum_{i=1}^{N} \mathsf{V}(x_i)$ with V Lipschitz. These generalizations are left to the reader, see in particular Remark 2.2.

Studying the same evolutions with added noise

$$(1.9) \qquad dx_i = -\frac{1}{N}\left(\nabla_{x_i}\mathcal{H}_N(x_1, \ldots, x_N) + \sum_{j \neq i} \mathsf{F}(x_i - x_j)\right) dt + \sqrt{2\theta}dW_i,$$

or

$$(1.10) \qquad dx_i = -\frac{1}{N} \mathbb{J} \nabla_{x_i} \mathcal{H}_N(x_1, \ldots, x_N) dt + \sqrt{2\theta} dW_i,$$

with dW_i being N independent Brownian motions and $\theta \geq 0$ a temperature, is also very interesting and has motivations from Random Matrix Theory, it is done in particular in [JW1, JW2, BO, FHM, BCF, LLY] (see also references therein). We will comment further on this at the end of the introduction in Section 1.5.

In the appendix, we consider the second order system corresponding to Newton's law for the energy $\mathcal{H}_N$

$$(1.11) \qquad \begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = -\frac{1}{N} \nabla_{x_i} \mathcal{H}_N(x_1, \ldots, x_N), & i = 1, \ldots, N \\ x_i(0) = x_i^0, \quad v_i(0) = v_i^0, \end{cases}$$

in the so-called monokinetic case. In the Coulomb case it is the true physics model for plasmas.

Consider the *empirical measure*

$$(1.12) \qquad \mu_N^t := \frac{1}{N} \sum_{i=1}^N \delta_{x_i^t}$$

associated to a solution $X_N^t := (x_1^t, \ldots, x_N^t)$ of the flow (1.5) or (1.6). If the points x_i^0, which themselves depend on N, are such that μ_N^0 converges to some regular measure μ^0, then a formal derivation leads to expecting that for $t > 0$, μ_N^t converges to the solution of the Cauchy problem with initial data μ^0 for the limiting evolution equation

$$(1.13) \qquad \partial_t \mu = \operatorname{div}\left((\nabla \mathsf{g} + \mathsf{F}) * \mu)\mu\right)$$

in the dissipative case (1.5) or

$$(1.14) \qquad \partial_t \mu = \operatorname{div}\left(\mathbb{J} \nabla(\mathsf{g} * \mu)\mu\right)$$

in the conservative case (1.6), with $*$ denoting the usual convolution.

These equations should be understood in a weak sense. Equation (1.13) (with $\mathsf{F} = 0$) is sometimes called the *fractional porous medium equation*. The two-dimensional Coulomb version also arises as a model for the dynamics of vortices in superconductors. The construction of solutions, their regularity and basic properties, are addressed in [LZ, DZ, MZ, AS, SV] for the Coulomb case of (1.13), in [CSV, CV, XZ] for the case $\mathsf{d} - 2 < \mathsf{s} < \mathsf{d}$ of (1.13), and [De, Yu, Scho1] for the two-dimensional Coulomb case of (1.14).

Establishing the convergence of the empirical measures to solutions of the limiting equations is nontrivial because of the nonlinear terms in the equation and the singularity of the interaction g. In fact, because of the strength of the singularity, treating the case of Coulomb interactions in dimension $\mathsf{d} \geq 3$ (and even more so that of super-coulombic interactions) had remained an open question for a long time, see for instance the introduction of [JW2] and the review [Jab]. It was not even completely clear that the result was true without expressing it in some statistical sense (with respect to the initial data).

In [JW1, JW2], Jabin and Wang introduced a new approach for the related problem of the mean-field convergence of the solutions of Newton's second order system of ODEs to the Vlasov equation, which allowed them to treat all interactions kernels with bounded gradients, but still not Coulomb interaction. The same problem has been addressed in [BP, LP, La] with

results that still require a cutoff of the Coulomb interaction. Our method already allows to unlock the case of Coulomb interaction for monokinetic data, this is the topic of Appendix A. The non-monokinetic case is even much more challenging and remains open.

The previously known results on the problems we are addressing were the following:

- In dimension 2, choosing (1.4) and $\mathbb{J}$ the rotation by $\pi/2$ in (1.6) corresponds to the so-called *point vortex system* which is well-known in fluid mechanics (cf. for instance [MP]), and its mean-field convergence to the Euler equation in vorticity form (1.14) was already established [Scho2]. His proof can be readapted to treat (1.5) as well in that case. Results of similar nature were also obtained in [GHL].
- Hauray [Hau] (see also [CCH]) treated the case of all sub-Coulombic interactions ($s < d - 2$) for (a possibly higher-dimensional generalization of) (1.6), where particles can have positive and negative charges and thus can attract as well as repel. The proof, which relies on the stability in ∞-Wasserstein distance of the limiting solution, cannot be adapted to $s \geq d - 2$.
- In dimension 1 and in the dissipative case (1.5), Carrillo-Ferreira-Precioso and Berman-Onnheim [CFP, BO] proved the unconditional convergence for all $0 < s < 1$ using the framework of Wasserstein gradient flows but their method, based on the convexity of the interaction in dimension 1, does not extend to higher dimensions.
- Duerinckx [Du], inspired by the *modulated energy* method of [Sy] for Ginzburg-Landau equations (where vortices also interact like Coulomb particles in dimension 2), was able to prove the result in the dissipative case (1.5) for $d = 1$ and $d = 2$ with $s < 1$, conditional to the regularity of the limiting solution, as we have here.

In this paper, we prove the mean field convergence for (1.5) and (1.6) in all the cases (1.3) and (1.4) in every dimension. This extends Duerinckx's result, which involves overcoming serious difficulties, as described further below, and we add the possible additional interaction force F in the dissipative case. We are limited to $s < d$ and this is natural since for $s \geq d$ the interaction kernel g is no longer integrable and the limiting equation is expected to be a different one.

As in [Du], our proof is a modulated energy argument inspired from [Sy], which is a way of exploiting a weak-strong uniqueness principle for the limiting equation. As mentioned above, looking for a stability principle in some Wasserstein distance fails for the Coulomb singularity. Instead we use a distance which is built as a Coulomb (or Riesz) metric, associated to the norm

$$(1.15) \qquad \|\mu\|^2 = \iint \mathsf{g}(x - y) d\mu(x) d\mu(y).$$

We are able to show by a Gronwall argument on this metric that the equations (1.13) and (1.14) satisfy a weak-strong uniqueness principle, and this can be translated into a proof of stability and convergence to 0 of the norm of $\mu_N^t - \mu^t$ (if it is initially small, it remains small for all further times).

The proof is self-contained and quantitative. It does not require understanding any qualitative property of the trajectories of the particles, such as for instance their minimal distances along the flow.

After this work was completed, Bresch, Jabin and Wang [BJW1] were able to beautifully incorporate our modulated energy into the relative entropy method of [JW2], turning it into a *modulated free energy*, which is a physically very natural quantity. It allowed them to extend

the result of [JW2] to more singular interactions, including Coulomb. Seen from our point of view, it allows to treat the cases with added noise (1.9) (but unfortunately not (1.10)). This can be explained very succintly, we do it in Section 1.5 for the reader's convenience.

1.2. **Main result.** Let X_N denote $(x_1, \ldots, x_N)$ and let us define for any probability measure μ,

$$(1.16) \qquad F_N(X_N, \mu) = \iint_{\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}} \setminus \triangle} \mathsf{g}(x - y) d\Big(\sum_{i=1}^{N} \delta_{x_i} - N\mu\Big)(x) d\Big(\sum_{i=1}^{N} \delta_{x_i} - N\mu\Big)(y)$$

where $\triangle$ denotes the diagonal in $\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}}$. We choose for "modulated energy"

$$F_N(X_N^t, \mu^t)$$

where $X_N^t = (x_1^t, \ldots, x_N^t)$ are the solutions to (1.5) or (1.6), and μ^t solves the expected limiting PDE. The function F_N is not positive, however it is bounded below (by $-CN^{1+\frac{\mathsf{s}}{\mathsf{d}}}$ in the case (1.3), respectively $-\left(\frac{N}{\mathsf{d}} \log N\right) - CN$ in the case (1.4), see Corollary 3.5). It turns out that also metrizes at least weak convergence, as described in Proposition 3.6. One may thus think of it as a good notion of distance from μ_N^t to μ^t, more precisely $\frac{1}{N^2} F_N(X_N^t, \mu^t)$ is a good distance.

Our main result is a Gronwall inequality on the time-derivative of $F_N(X_N^t, \mu^t)$, which implies a quantitative rate of convergence of μ_N^t to μ^t in that metric.

Throughout the paper, $(\cdot)_+$ denotes the positive part of a number. The parameter s refers to the exponent in (1.3) while in the logarithmic case (1.4) it is set to be 0. The space $\dot{H}^m(\mathbb{R}^{\mathsf{d}})$ is the homogeneous Sobolev space of functions u whose Fourier transform $\hat{u}$ satisfies $|\xi|^m \hat{u}(\xi) \in L^2(\mathbb{R}^{\mathsf{d}})$.

Theorem 1. *Assume that g is of the form (1.3) or (1.4). Assume that $\mathsf{F} \in \dot{H}^{\frac{\mathsf{d}-\mathsf{s}}{2}}(\mathbb{R}^{\mathsf{d}}) \cap C^{0,\alpha}(\mathbb{R}^{\mathsf{d}})$ for some $\alpha > 0$ and $\nabla \mathsf{F} \in L^q(\mathbb{R}^{\mathsf{d}})$ for some $1 \leq q \leq \infty$. Assume (1.13), resp. (1.14), admits a solution μ^t such that, for some $T > 0$,*
(1.17)
$$\begin{cases} \mu^t \in L^\infty([0,T], L^\infty(\mathbb{R}^{\mathsf{d}})), \text{ and } \nabla^2 \mathsf{g} * \mu^t \in L^\infty([0,T], L^\infty(\mathbb{R}^{\mathsf{d}})) \quad \text{if } \mathsf{s} < \mathsf{d} - 1 \\ \mu^t \in L^\infty([0,T], C^\sigma(\mathbb{R}^{\mathsf{d}})) \text{ with } \sigma > \mathsf{s} - \mathsf{d} + 1, \text{ and } \nabla^2 \mathsf{g} * \mu^t \in L^\infty([0,T], L^\infty(\mathbb{R}^{\mathsf{d}})) \text{ if } \mathsf{s} \geq \mathsf{d} - 1. \end{cases}$$

Let X_N^t solve (1.5), respectively (1.6). Then there exist positive constants C_1, C_2 depending only on the norms of μ^t controlled by (1.17) and those of F, and an exponent $\beta < 2$ depending only on $\mathsf{d}, \mathsf{s}, \alpha, \sigma$, such that for every $t \in [0,T]$ we have

$$(1.18) \qquad F_N(X_N^t, \mu^t) \leq \Big(F_N(X_N^0, \mu^0) + C_1 N^\beta\Big) e^{C_2 t}.$$

In particular, using the notation (1.12), if $\mu_N^0 \rightharpoonup \mu^0$ and is such that

$$\lim_{N \to \infty} \frac{1}{N^2} F_N(X_N^0, \mu^0) = 0,$$

then the same is true for every $t \in [0,T]$ and

$$(1.19) \qquad \mu_N^t \rightharpoonup \mu^t$$

in the weak sense.

Establishing the convergence of the empirical measures is essentially equivalent to proving *propagation of molecular chaos* (see [Go, HM, Jab] and references therein) which means showing that if $f_N^0(x_1, \ldots, x_N)$ is the initial probability density of seeing particles at $(x_1, \ldots, x_N)$ and if f_N^0 converges to some factorized state $\mu^0 \otimes \cdots \otimes \mu^0$, then the k-point marginals $f_{N,k}$ converge for all time to $(\mu^t)^{\otimes k}$. With Remark 3.7, our result implies a convergence of this type as well.

Let us point out that using a Fourier-based point of view on (1.16) Bresch-Jabin-Wang were able (see [BJW1, BJW2]) to subsequently relax the assumptions on the interaction g: it does not need to be Coulomb or Riesz (a bit like with the added regular force in (1.5)) but may contain a mildly singular attractive part, as long as a sufficiently strong repulsive part is still present.

1.3. **Comments on the assumptions.** Let us now comment on the regularity assumption made in (1.17). First of all, one can check (see Lemma 3.1) that the assumption (1.17) is implied by

$$(1.20) \qquad \mu \in L^\infty([0,T], C^\theta(\mathbb{R}^\mathsf{d})) \quad \text{for some } \theta > \mathsf{s} - \mathsf{d} + 2,$$

which coincides with the assumption made in [Du] and is a bit stronger. This weakening of the assumption allows to include for instance the case of measures which are (a regular function times) the characteristic function of some regular set, such as in the situation of vortex patches for the Euler equation in vorticity form, corresponding to (1.6) in the two-dimensional logarithmic case. These vortex patches were first studied in [Ch2, BerCon, Si] where it was shown that if the patch initially has a $C^{1,\alpha}$ boundary this remains the case over time, and our second assumption that the velocity $\nabla \mathsf{g} * \mu^t$ be Lipschitz holds as well (see also [BK]). It is not too difficult to check that in all dimensions this second condition holds any time μ is C^σ with $\sigma > 0$ away from a finite number of $C^{1,\alpha}$ hypersurfaces. More generally, such situations with patches can be expected to naturally arise in all the Coulomb cases. For instance, in the dissipative Coulomb case (1.5) with $\mathsf{F} = 0$, in any dimension a self-similar solution in the form of (a constant multiple of) the characteristic function of an expanding ball was exhibited in [SV] and shown to be an attractor of the dynamics. For the non-Coulomb dissipative cases, the corresponding self-similar solutions, called *Barenblatt solutions*, are of the form

$$t^{-\frac{\mathsf{d}}{2+\mathsf{s}}}(a - bx^2 t^{-\frac{2}{2+\mathsf{s}}})_+^{\frac{\mathsf{s}-\mathsf{d}+2}{2}}$$

as shown in [BIK, CV] (and this formula retrieves the solution of [SV] for $\mathsf{s} = \mathsf{d} - 2$).

If the initial μ^0 is sufficiently regular, the stronger assumption (1.20) is known to hold with $T = \infty$ for the Coulomb case (see [LZ] where the proof works as well in higher dimensions), and it is known up to some $T > 0$ in the case $(\mathsf{d} - 2)_+ < \mathsf{s} \leq \mathsf{d} - 1$ [XZ]. As for (1.14), to our knowledge the desired regularity is only known in dimension 2 for the Euler equation in vorticity form (see [Wo, Ch2]), although the arguments of [XZ] written for the dissipative case seem to also apply to the conservative one. Our convergence result thus holds in these cases, under the assumption that the limit μ^0 of μ_N^0 is sufficiently regular and that $F_N(X_N^0, \mu^0) = o(N^2)$. Note that, as shown in [Du], the latter is implied by the convergence of the initial energy

$$\lim_{N \to \infty} \frac{1}{N^2} \sum_{i \neq j} \mathsf{g}(x_i^0 - x_j^0) = \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d}} \mathsf{g}(x - y) d\mu^0(x) d\mu^0(y)$$

which can be viewed as a well-preparedness condition. For any regular enough μ^0, one may for instance build initial conditions satisfying this assumption (and something even stronger) by the construction in [PS].

For $\mathsf{d} - 1 < \mathsf{s} < \mathsf{d}$, even the local in time propagation of regularity of solutions of (1.13) remains an open problem. Note that the uniqueness of regular enough solutions is always implied by the weak-strong stability argument we use, reproduced in Section 1.4 below.

Requiring some regularity of the solutions to the limiting equation for establishing convergence with relative entropy / modulated entropy / modulated energy methods is fairly common: the same situation appears for instance in [JW1, JW2] or in the derivation of the Euler equations from the Boltzmann equation via the modulated entropy method, see [SR] and references therein.

1.4. **The method.** As mentioned, our method exploits a weak-strong uniqueness principle for the solutions of (1.13), resp. (1.14), which is exactly the same as [Du, Lemma 2.1, Lemma 2.2] (and can be easily readapted to the conservative case) and states that if μ_1^t and μ_2^t are two L^∞ solutions to (1.13) such that $\nabla^2(\mathsf{g} * \mu_2) \in L^1([0, T], L^\infty)$, we have

$$
(1.21) \quad \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d}} \mathsf{g}(x - y) d(\mu_1^t - \mu_2^t)(x) d(\mu_1^t - \mu_2^t)(y)
$$
$$
\leq e^{C \int_0^t \|\nabla^2(\mathsf{g} * \mu_2^s)\| ds} \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d}} \mathsf{g}(x - y) d(\mu_1^0 - \mu_2^0)(x) d(\mu_1^0 - \mu_2^0)(y).
$$

But the Coulomb or Riesz energy (1.15) is nothing else than the fractional Sobolev $H^{-\alpha}$ norm of μ with $\alpha = \frac{\mathsf{d} - \mathsf{s}}{2}$, hence this is a good metric of convergence and implies the weak-strong uniqueness property.

A crucial ingredient is the use of the stress-energy (or energy-momentum) tensor which naturally appears when taking the *inner variations* of the energy (1.15), i.e. computing $\frac{d}{dt}|_{t=0}\|\mu \circ (I + t\psi)\|^2$ (this is standard in the calculus of variations, see for instance [He, Sec. 1.3.2]). To explain further, let us restrict for now to the Coulomb case, and set $h^\mu = \mathsf{g} * \mu$, the Coulomb potential generated by μ. In that case, we have

$$
(1.22) \quad -\Delta h^\mu = \mathsf{c}_\mathsf{d} \mu
$$

for some constant c_d depending only on d. The first key is to reexpress the Coulomb energy (1.15) as a single integral in h^μ, more precisely we easily find via an integration by parts that if μ is a measure with $\int d\mu = 0$,

$$
(1.23) \quad \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d}} \mathsf{g}(x - y) d\mu(x) d\mu(y) = \int_{\mathbb{R}^\mathsf{d}} h^\mu d\mu = -\frac{1}{\mathsf{c}_\mathsf{d}} \int_{\mathbb{R}^\mathsf{d}} h^\mu \Delta h^\mu = \frac{1}{\mathsf{c}_\mathsf{d}} \int_{\mathbb{R}^\mathsf{d}} |\nabla h^\mu|^2.
$$

The stress-energy tensor is then defined as the $\mathsf{d} \times \mathsf{d}$ tensor with coefficients

$$
(1.24) \quad [h^\mu, h^\mu]_{ij} = 2\partial_i h^\mu \partial_j h^\mu - |\nabla h^\mu|^2 \delta_{ij},
$$

with δ_{ij} the Kronecker symbol. We may compute that if μ is regular enough

$$
(1.25) \quad \text{div}\,[h^\mu, h^\mu] = 2\Delta h^\mu \nabla h^\mu = -2\mathsf{c}_\mathsf{d} \mu \nabla h^\mu.
$$

(Here the divergence is a vector with entries equal to the divergence of each row of $[h^\mu, h^\mu]$.) We thus see how this stress-energy tensor allows to give a weak meaning to the product $\mu \nabla h^\mu = \mu \nabla \mathsf{g} * \mu$, with $[h^\mu, h^\mu]$ well-defined in energy space and pointwise controlled by $|\nabla h^\mu|^2$, which can by the way serve to give a notion of weak solutions of the equation in the energy space (as in [De, LZ]). Note that in dimension 2, it is known since [De] that even

though $[h^\mu, h^\mu]$ is nonlinear, it is stable under weak limits in energy space provided μ has a sign, but this fact *does not extend* to higher dimension.

Let us now present the short proof of (1.21) as it will be a model for the main proof. We focus on the dissipative case (the conservative one is an easy adaptation) and still the Coulomb case for simplicity with no additional interaction F (when present, the additional terms can be absorbed thanks to the dissipation). Let μ_1 and μ_2 be two solutions to (1.13) and $h_i = \mathsf{g} * \mu_i$ the associated potentials, which solve (1.22). Let us compute

$$(1.26) \quad \partial_t \int_{\mathbb{R}^\mathsf{d}} |\nabla(h_1 - h_2)|^2 = 2\mathsf{c_d} \int_{\mathbb{R}^\mathsf{d}} (h_1 - h_2)\partial_t(\mu_1 - \mu_2)$$

$$= 2\mathsf{c_d} \int_{\mathbb{R}^\mathsf{d}} (h_1 - h_2)\mathrm{div}\,(\mu_1 \nabla h_1 - \mu_2 \nabla h_2)$$

$$= -2\mathsf{c_d} \int_{\mathbb{R}^\mathsf{d}} (\nabla h_1 - \nabla h_2) \cdot (\mu_1 \nabla h_1 - \mu_2 \nabla h_2)$$

$$= -2\mathsf{c_d} \int_{\mathbb{R}^\mathsf{d}} |\nabla(h_1 - h_2)|^2 \mu_1 - 2\mathsf{c_d} \int_{\mathbb{R}^\mathsf{d}} \nabla h_2 \cdot \nabla(h_1 - h_2)(\mu_1 - \mu_2)$$

In the right-hand side, we recognize from (1.25) the divergence of the stress-energy tensor $[h_1 - h_2, h_1 - h_2]$, hence

$$\partial_t \int_{\mathbb{R}^\mathsf{d}} |\nabla(h_1 - h_2)|^2 \leq \int_{\mathbb{R}^\mathsf{d}} \nabla h_2 \cdot \mathrm{div}\,[h_1 - h_2, h_1 - h_2]$$

so if $\nabla^2 h_2$ is bounded, we may integrate by parts the right-hand side and bound it pointwise by

$$\|\nabla^2 h_2\|_{L^\infty} \int_{\mathbb{R}^\mathsf{d}} |[h_1 - h_2, h_1 - h_2]| \leq 2\|\nabla^2 h_2\|_{L^\infty} \int_{\mathbb{R}^\mathsf{d}} |\nabla(h_1 - h_2)|^2,$$

and the claimed result follows by Gronwall's lemma.

In the Riesz case, the Riesz potential $h^\mu = \mathsf{g} * \mu$ is no longer the solution to a local equation, and to find a replacement to (1.22)–(1.25), we use an extension procedure as popularized by [CaffSi] in order to obtain a local integral in h^μ in the extended space $\mathbb{R}^{\mathsf{d}+1}$.

In the discrete case of the original ODE system, all the above integrals are singular and this constitutes the main difficulty overcome in this paper. In place of the second term in the right-hand side of (1.26), we then have to control a term which by symmetry can be written in the form

$$(1.27) \qquad \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d} \setminus \triangle} (\nabla h^{\mu_1}(x) - \nabla h^{\mu_2}(y)) \cdot \nabla \mathsf{g}(x - y) d(\mu_1 - \mu_2)(x) d(\mu_1 - \mu_2)(y)$$

where $\triangle$ denotes the diagonal, μ_1 is the limiting measure μ^t and μ_2 is the discrete empirical measure μ_N^t. Such terms are well known (see for instance [Scho2]), and create the main difficulty due to the singularity of g. When removing the diagonal, the positivity manifested by (1.23) is in effect lost. We are however able to prove the following crucial functional inequality.

Proposition 1.1. *Assume that μ is a probability density, with $\mu \in C^\sigma(\mathbb{R}^\mathsf{d})$ with $\sigma > \mathsf{s} - \mathsf{d} + 1$ if $\mathsf{s} \geq \mathsf{d} - 1$; respectively $\mu \in L^\infty(\mathbb{R}^\mathsf{d})$ or $\mu \in C^\sigma(\mathbb{R}^\mathsf{d})$ with $\sigma > 0$ if $\mathsf{s} < \mathsf{d} - 1$. For any*

$X_N \in (\mathbb{R}^{\mathsf{d}})^N$ *and any Lipschitz map* $\psi : \mathbb{R}^{\mathsf{d}} \to \mathbb{R}^{\mathsf{d}}$, *we have*

$$(1.28) \quad \left| \iint_{\triangle^c} (\psi(x) - \psi(y)) \cdot \nabla \mathsf{g}(x-y) d(\sum_{i=1}^{N} \delta_{x_i} - N\mu)(x) d(\sum_{i=1}^{N} \delta_{x_i} - N\mu)(y) \right|$$

$$\leq C \|\nabla \psi\|_{L^\infty} \left(F_N(X_N, \mu) + (1 + \|\mu\|_{L^\infty}) N^{2 - \frac{\mathsf{d} - \mathsf{s}}{\mathsf{d}(\mathsf{d}+1)}} + 2 \left(\frac{N}{\mathsf{d}} \log N \right) \mathbf{1}_{(1.4)} \right)$$

$$+ C \min \left(\|\psi\|_{L^\infty} \|\mu\|_{L^\infty} N^{1 + \frac{\mathsf{s}+1}{\mathsf{d}}} + \|\nabla \psi\|_{L^\infty} \|\mu\|_{L^\infty} N^{1 + \frac{\mathsf{s}}{\mathsf{d}}}, \|\psi\|_{W^{1,\infty}} \|\mu\|_{C^\sigma} N^{1 + \frac{\mathsf{s}+1-\sigma}{\mathsf{d}}} \right)$$

$$+ C \begin{cases} \|\nabla \psi\|_{L^\infty} (1 + \|\mu\|_{C^\sigma}) N^{2 - \frac{1}{\mathsf{d}}} & \textit{if } \mathsf{s} \geq \mathsf{d} - 1 \\ \|\nabla \psi\|_{L^\infty} (1 + \|\mu\|_{L^\infty}) N^{2 - \frac{1}{\mathsf{d}}} & \textit{if } \mathsf{s} < \mathsf{d} - 1, \end{cases}$$

where C *depends only on* s, d.

The right-hand side should be read as $C_{\psi,\mu,s,\mathsf{d}}(F_N(X_N, \mu) + N^\beta)$ for some $\beta < 2$. This inequality is the main novelty of the paper. Even though the term $(\psi(x) - \psi(y)) \cdot \nabla \mathsf{g}(x-y)$ has a singularity of same order as $\mathsf{g}(x-y)$ near the diagonal, the inequality is not at all obvious due to the lack of positivity of the integrand and its proof is rendered difficult by the handling of the removed diagonal terms. Note that in [GP] Golse and Paul were able to treat the mean field limit for the *quantum* Coulomb dynamics, relying on this inequality.

To give more insight into the proof of this proposition, we need to discuss the electric representation of the modulated energy F_N, again restricting to the Coulomb case. For that we introduce the electric potential

$$H_N^\mu[X_N] = \mathsf{g} * \left(\sum_{i=1}^{N} \delta_{x_i} - N\mu \right).$$

Arguing as in (1.22)–(1.23) we would like to rewrite $F_N(X_N)$ as $\int_{\mathbb{R}^{\mathsf{d}}} |\nabla H_N^\mu[X_N]|^2$. This is not quite correct due to the divergence of H_N^μ at the points x_i. This can however be corrected by using mollified Dirac masses and setting for any $\vec{\eta} = (\eta_1, \dots, \eta_N) \in \mathbb{R}^N$,

$$H_{N,\vec{\eta}}^\mu = \mathsf{g} * \left(\sum_{i=1}^{N} \delta_{x_i}^{(\eta_i)} - N\mu \right)$$

where we dropped the $[X_N]$ in the notation and let $\delta_x^{(\eta)}$ denote the uniform measure of mass 1 on $\partial B(x, \eta)$. This effectively truncates H_N^μ at scale η_i around each x_i, i.e. a scale which depends on each point. Reinserting the diagonal terms, it is then not too difficult to check that

$$F_N(X_N, \mu) = \lim_{\eta_i \to 0} \iint \mathsf{g}(x-y) d \left(\sum_{i=1}^{N} \delta_{x_i}^{(\eta_i)} - N\mu \right)(x) d \left(\sum_{i=1}^{N} \delta_{x_i}^{(\eta_i)} - N\mu \right)(y)$$

$$- \sum_{i=1}^{N} \iint \mathsf{g}(x-y) d\delta_{x_i}^{(\eta_i)}(x) d\delta_{x_i}^{(\eta_i)}(y)$$

$$= \frac{1}{\mathsf{c}_{\mathsf{d}}} \lim_{\eta_i \to 0} \left(\int_{\mathbb{R}^{\mathsf{d}}} |\nabla H_{N,\vec{\eta}}^\mu|^2 - \mathsf{c}_{\mathsf{d}} \sum_{i=1}^{N} \mathsf{g}(\eta_i) \right).$$

This effectively gives a renormalized meaning to the identity (1.23) in this setting.

The idea of expressing the interaction energy as a local integral in H_N^μ and its renormalization procedure were previously used in the study of Coulomb and Riesz energies

in [RS, PS, LS1, LS2], but it was not clear how to adapt these ideas to control (1.27): in fact in [Du] this was dealt with by a ball-construction procedure inspired from the analysis of Ginzburg-Landau vortices, which led to nonoptimal estimates and to the restriction to the dissipative case only and to $s < 1$ and $d \leq 2$.

In fact, we can say better, and this is where we depart significantly from previous works, by exploiting the fact that the expression $\int_{\mathbb{R}^d} |\nabla H^\mu_{N,\vec{\eta}}|^2 - c_d \sum_{i=1}^N g(\eta_i)$ is essentially *decreasing* with respect to η_i and constant for η_i small enough that the $B(x_i, \eta_i)$'s are disjoint, see Proposition 3.3. More precisely, we may set r_i to be $1/4$ of the minimal distance from x_i to all other points, and for $\eta_i \leq r_i$, we have the *equality* (without limits)

$$(1.29) \qquad F_N(X_N, \mu) = \frac{1}{c_d}\left(\int_{\mathbb{R}^d} |\nabla H^\mu_{N,\vec{\eta}}|^2 - c_d \sum_{i=1}^N g(\eta_i)\right) + \text{explicit negligible terms.}$$

In addition, an observation made in this paper for the first time is that when choosing precisely $\eta_i = r_i$, the potentially large terms $\int_{\mathbb{R}^d} |\nabla H^\mu_{N,\vec{r}}|^2$ and $c_d \sum_{i=1}^N g(r_i)$ are *separately* controlled by CF_N (plus good terms) and conversely, see Corollary 3.4. It now suffices to control the left-hand side of (1.28) by $\int_{\mathbb{R}^d} |\nabla H^\mu_{N,\vec{r}}|^2$. Let us emphasize that this choice of truncation r_i that depends on the point is (up to constants) the only one which is at the same time large enough so that $\int_{\mathbb{R}^d} |\nabla H^\mu_{N,\vec{\eta}}|^2$ is directly controlled by F_N and small enough that the equality (1.29) holds. In other words, since we do not have any bound from below on the distance between the points, a point-dependent truncation radius is crucial. As a side note, the idea of using truncations for proving mean-field limits is common, however what is usually done is to truncate the interaction g itself (at lengthscales possibly depending on N), and try to take limits in the resulting flow. What we do here is very different: we do not modify the interaction but desingularize the charges themselves as an intermediate step to control the singular terms.

In order to bound the left-hand side of (1.28), the key is then to interpret it as well as a single integral in terms of ∇H^μ_N in a suitable "renormalized" way, more precisely in terms of the stress tensor associated to $\nabla H^\mu_{N,\vec{\eta}}$, for $\eta_i \to 0$. The quantity then obtained is this time not monotone in η_i, however, by carefully studying its variations in η_i (this is the hardest part of the proof), we are able to control it by the (integral of the) stress tensor associated to $\nabla H_{N,\vec{r}}$, which can in turn be bounded by $\int |\nabla H^\mu_{N,\vec{r}}|^2$, and we conclude thanks to the fact that this is controlled by CF_N.

1.5. **The modulated free energy for the case with noise.** In this subsection, we explain the modulated free energy method of [BJW1], which is posterior to the first version of this work and allows to treat the case of (1.9). In this case, the mean-field limit inherits an added Laplacian:

$$(1.30) \qquad \partial_t \mu = \text{div}\left((\nabla g + F) * \mu)\mu\right) + \theta \Delta \mu.$$

Consider $f_N(x_1, \ldots, x_N)$ a symmetric probability density on configurations in $(\mathbb{R}^d)^N$, and let us again abbreviate $(x_1, \ldots, x_N)$ by X_N. Let us introduce the relative entropy [1]

$$H_N(f_N | \mu^{\otimes N}) := N \int_{\mathbb{R}^{dN}} f_N \log \frac{f_N}{\mu^{\otimes N}} dX_N$$

---

[1] Note that we take N^2 times the usual relative entropy, because all our quantities are N^2 times standard ones.

It is of course a way of measuring how close the distribution f_N is to $\mu^{\otimes N}$. Then consider

$$K_N(f_N, \mu) = \int_{\mathbb{R}^{dN}} f_N(X_N) F_N(X_N, \mu) dX_N$$

the expectation of our modulated energy F_N with respect to f_N. Bresch-Jabin-Wang introduce the modulated free energy

$$(1.31) \qquad \mathcal{F}_\theta(f_N, \mu) = \theta H_N(f_N | \mu^{\otimes N}) + K_N(f_N, \mu).$$

It has exactly the structure of a free energy in statistical physics, i.e. of the form energy plus temperature times entropy, and when temperature vanishes and f_N concentrates on one configuration, it coincides with the regular modulated energy.

Consider now f_N^t corresponding to the probability density of particles following the flow (1.9), then by Ito calculus f_N^t solves a Liouville or Kolmogorov equation

$$(1.32) \qquad \partial_t f_N^t = \sum_{i=1}^N \operatorname{div}_i \left(\frac{1}{N} \sum_{i \neq j} \nabla_i \mathcal{H}_N(x_i - x_j) f_N^t(X_N) \right) + \theta \sum_{i=1}^N \Delta_i f_N^t.$$

Their crucial observation is that when combining the relative entropy and the modulated energy in exactly the way of (1.31) and differentiating in time $\mathcal{F}_\theta(f_N^t, \mu^t)$ for μ^t a solution to the mean-field limit (1.30) and f_N^t a solution to the Liouville equation (1.32), the new and problematic terms arising in $\partial_t H_N(f_N^t, \mu^t)$ from the presence of the noise (which were an obstacle to treat the Coulomb case in [JW2]) exactly cancel with the new terms arising in $\partial_t K_N(f_N^t, \mu^t)$ (this does not happen in the conservative case of (1.10) though). This allows them to obtain the following crucial identity:

$$(1.33) \quad \frac{d}{dt} \mathcal{F}_\theta(f_N^t, \mu^t)$$
$$\leq -\int \left(\iint_{\triangle^c} \nabla \mathsf{g}(x - y) \cdot (\psi^t(x) - \psi^t(y)) d(\sum_{i=1}^N \delta_{x_i} - N\mu^t)(x) d(\sum_{i=1}^N \delta_{x_i} - N\mu^t)(y) \right) df_N^t(X_N)$$

with this time

$$\psi^t = \nabla h^{\mu^t} + \theta \frac{\nabla \mu^t}{\mu^t}.$$

Once this identity is observed, Proposition 1.1 directly applies (if μ^t is assumed regular enough) and yields

$$\frac{d}{dt} \mathcal{F}_\theta(f_N^t, \mu^t) \leq C K_N(f_N^t, \mu^t) + o(N^2) \leq C \mathcal{F}_\theta(f_N^t, \mu^t) + o(N^2)$$

which allows to directly conclude via Gronwall's lemma the mean-field convergence in the case with noise.

The rest of the paper is organized as follows: we start by deriving the main result assuming the result of Proposition 1.1. In Section we present the details of the electric formulation in the general Riesz case and prove the main properties of the modulated energy (monotonicity, bound from below, coerciveness). We conclude in Section 4 with the discussion of the stress-energy tensor and the proof of Proposition 1.1. The paper finishes with Appendix A on the derivation of the Vlasov-Poisson system in the monokinetic case.

2. Main proof

In all the paper, we will use the notation $\mathbf{1}_{(1.4)}$ to indicate a term which is only present in the logarithmic cases (1.4) and $\mathbf{1}_{\mathsf{s}<\mathsf{d}-1}$ for a term which is present only if $\mathsf{s} < \mathsf{d} - 1$. Below, the principal value (which may be omitted for $\mathsf{s} < \mathsf{d} - 1$) is defined by $P.V. \int_{\mathbb{R}^{\mathsf{d}} \setminus \{x\}} :=$ $\lim_{r\to 0} \int_{\mathbb{R}^{\mathsf{d}} \setminus B(x,r)}$.

Differentiating from formula (1.16), we have

Lemma 2.1. *If X_N^t is a solution of* (1.5), *then*

$$(2.1) \quad \partial_t F_N(X_N^t, \mu^t) = -2N^2 \int_{\mathbb{R}^{\mathsf{d}}} \left| P.V. \int_{\mathbb{R}^{\mathsf{d}} \setminus \{x\}} \nabla \mathsf{g}(x - y) d(\mu_N^t - \mu^t)(y) \right|^2 d\mu_N^t(x)$$

$$+ 2N^2 \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{F} * (\mu^t - \mu_N^t) \cdot \nabla(h^{\mu_N^t} - h^{\mu^t}) d\mu_N^t$$

$$-N^2 \iint_{\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}} \setminus \triangle} \left((\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(x) - (\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(y) \right) \cdot \nabla \mathsf{g}(x-y) d(\mu_N^t - \mu^t)(x) d(\mu_N^t - \mu^t)(y).$$

If X_N^t is a solution of (1.6), *then*

$$(2.2)$$
$$\partial_t F_N(X_N^t, \mu^t) = -N^2 \iint_{\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}} \setminus \triangle} \mathbb{J} \left(\nabla h^{\mu^t}(x) - \nabla h^{\mu^t}(y) \right) \cdot \nabla \mathsf{g}(x-y)(\mu_N^t - \mu^t)(x)(\mu_N^t - \mu^t)(y).$$

Remark 2.2. *In the case of an added term $\frac{1}{N} \sum_{i=1}^{N} \mathsf{V}(x_i)$ in the evolutions, one obtains an additional term*

$$-2N^2 \iint_{\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}} \setminus \triangle} (\mathsf{V}(x) - \mathsf{V}(y)) \cdot \nabla \mathsf{g}(x - y) d(\mu_N^t - \mu^t)(x) d(\mu_N^t - \mu^t)(y)$$

which can be handled like the others using Proposition 1.1 if V is globally Lipschitz.

Proof. We note that if $\mathsf{s} \geq \mathsf{d} - 1$, $\nabla \mathsf{g}$ is not integrable near 0, so $\nabla \mathsf{g} * \mu$ should be understood in a distributional sense and $\mu \nabla(\mathsf{g} * \mu) = \mu \mathsf{g} * \nabla \mu$ as well, assuming that μ is regular enough. We may also check that this distributional definition is equivalent to defining

$$\nabla h^\mu(x) = P.V. \int_{\mathbb{R}^{\mathsf{d}} \setminus \{x\}} \nabla \mathsf{g}(x - y) d\mu(y).$$

In the case (1.5), we have

$$
\begin{aligned}
\partial_t F_N(X_N^t, \mu^t) &= N^2 \partial_t \iint_{\mathbb{R}^d \times \mathbb{R}^d} \mathsf{g}(x-y) d\mu^t(x) d\mu^t(y) + \partial_t \sum_{i \neq j} \mathsf{g}(x_i^t - x_j^t) \\
&\quad -2N \partial_t \sum_{i=1}^N \int_{\mathbb{R}^d} \mathsf{g}(x_i^t - y) d\mu^t(y) \\
&= -2N^2 \int_{\mathbb{R}^d} |\nabla h^{\mu^t}|^2(x) d\mu^t(x) - 2N^2 \int_{\mathbb{R}^d} \nabla h^{\mu^t}(x) \cdot \mathsf{F} * \mu^t(x) d\mu^t(x) \\
&\quad -2 \sum_{i=1}^N \left| \sum_{j \neq i} \nabla \mathsf{g}(x_i^t - x_j^t) \right|^2 - 2 \sum_{i=1}^N \left(\sum_{j \neq i} \nabla \mathsf{g}(x_i^t - x_j^t) \cdot \sum_{j \neq i} \mathsf{F}(x_i^t - x_j^t) \right) \\
&\quad +2N \sum_{i \neq j} \nabla h^{\mu^t}(x_i^t) \cdot \left(\nabla \mathsf{g}(x_i^t - x_j^t) + \mathsf{F}(x_i^t - x_j^t) \right) \\
&\quad +2N \sum_{i=1}^N P.V. \int_{\mathbb{R}^d \setminus \{x_i^t\}} (\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(x) \cdot \nabla \mathsf{g}(x - x_i^t) d\mu^t(x).
\end{aligned}
$$

We then recombine the terms to obtain

$$
\begin{aligned}
\partial_t F_N(X_N^t, \mu^t) &= -2N^2 \int_{\mathbb{R}^d} \left| P.V. \int_{\mathbb{R}^d \setminus \{x\}} \nabla \mathsf{g}(x-y) d(\mu_N^t - \mu^t)(y) \right|^2 d\mu_N^t(x) \\
&\quad -2N^2 \int_{\mathbb{R}^d} |\nabla h^{\mu^t}|^2 d\mu^t - 2N^2 \int_{\mathbb{R}^d} \nabla h^{\mu^t}(x) \cdot \mathsf{F} * \mu^t(x) d\mu^t(x) \\
&\quad -2N \int_{\mathbb{R}^d} \left(\nabla h^{\mu_N^t}(x) \cdot \int_{\mathbb{R}^d \setminus \{x\}} \mathsf{F}(x-y) d\mu_N^t(y) \right) d\mu_N^t(x) + 2N^2 \int_{\mathbb{R}^d} |\nabla h^{\mu^t}|^2 d\mu_N^t \\
&\quad +2N^2 \int_{\mathbb{R}^d} \nabla h^{\mu^t}(x) \cdot \int_{\mathbb{R}^d \setminus \{x\}} (\nabla \mathsf{g} + \mathsf{F})(x-y) d\mu_N^t(y) d\mu_N^t(x) \\
&\quad +2N^2 \int_{\mathbb{R}^d} P.V. \int_{\mathbb{R}^d \setminus \{y\}} (\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(x) \cdot \nabla \mathsf{g}(x-y) d\mu^t(x) d\mu_N^t(y).
\end{aligned}
$$

We next recognize that all the terms except the first can be recombined and symmetrized into

$$
\begin{aligned}
-N^2 \iint_{\triangle^c} & \left((\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(x) - (\nabla h^{\mu^t} + \mathsf{F} * \mu^t)(y) \right) \cdot \nabla \mathsf{g}(x-y) d(\mu_N^t - \mu^t)(x) d(\mu_N^t - \mu^t)(y) \\
& + 2N^2 \int_{\mathbb{R}^d} \mathsf{F} * (\mu^t - \mu_N^t) \cdot \nabla (h^{\mu_N^t} - h^{\mu^t}) d\mu_N^t
\end{aligned}
$$

which gives the desired formula.

In the case (1.6) we have

$$
\partial_t F_N(X_N^t, \mu^t) = N^2 \partial_t \iint \mathsf{g}(x-y) d\mu^t(x) d\mu^t(y) + \partial_t \sum_{i \neq j} \mathsf{g}(x_i^t - x_j^t) - 2N \partial_t \sum_{i=1}^N \int_{\mathbb{R}^d} \mathsf{g}(x_i^t - y) d\mu^t(y)
$$

$$
= 2N \sum_{i \neq j} \nabla h^{\mu^t}(x_i^t) \cdot \mathbb{J} \nabla \mathsf{g}(x_i^t - x_j^t) + 2N \sum_{i=1}^N P.V. \int_{\mathbb{R}^d \setminus \{x_i^t\}} \mathbb{J} \nabla h^{\mu^t}(x) \cdot \nabla \mathsf{g}(x - x_i^t) d\mu^t(x)
$$

We then rewrite this as

$$\partial_t F_N(X_N^t, \mu^t) = 2N^2 \int_{\mathbb{R}^d} \nabla h^{\mu^t}(x) \cdot \int_{\mathbb{R}^d \setminus \{x\}} \mathbb{J} \nabla \mathsf{g}(x-y) d\mu_N^t(y) d\mu_N^t(x)$$

$$+ 2N^2 \int_{\mathbb{R}^d} P.V. \int_{\mathbb{R}^d \setminus \{y\}} \mathbb{J} \nabla h^{\mu^t}(x) \cdot \nabla \mathsf{g}(x-y) d\mu^t(x) d\mu_N^t(y).$$

By antisymmetry of $\mathbb{J}$, we recognize that the right-hand side can be symmetrized into

$$-N^2 \iint_{\triangle^c} \mathbb{J}\left(\nabla h^{\mu^t}(x) - \nabla h^{\mu^t}(y)\right) \cdot \nabla \mathsf{g}(x-y) d(\mu_N^t - \mu^t)(x) d(\mu_N^t - \mu^t)(y).$$

$\square$

The main point is thus to control the last term in the right-hand side of (2.1) or (2.2) which is done via Proposition 1.1.

For the dissipative case with the added force, we will also need

Lemma 2.3. *Assume* $\mathsf{F} \in \dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d) \cap C^{0,\alpha}(\mathbb{R}^d)$ *for some* $\alpha > 0$. *Then there exists* $\lambda > 0$ *and* $C > 0$ *depending only on* $\alpha, \mathsf{s}, \mathsf{d}$ *such that for every* t,

$$N^2 \int_{\mathbb{R}^d} \mathsf{F} * (\mu^t - \mu_N^t) \cdot \nabla(h^{\mu_N^t} - h^{\mu^t}) d\mu_N^t \le N^2 \int_{\mathbb{R}^d} \left| P.V. \int_{\mathbb{R}^d \setminus \{x\}} \nabla \mathsf{g}(x-y) d(\mu_N^t - \mu^t)(y) \right|^2 d\mu_N^t(x)$$

$$+ C\|\mathsf{F}\|_{\dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d)}^2 \left(F_N(X_N^t, \mu^t) + (1 + \|\mu^t\|_{L^\infty})N^{1+\frac{s}{d}} + 2\left(\frac{N}{d}\log N\right)\mathbf{1}_{(1.4)}\right) + C\|\mathsf{F}\|_{C^{0,\alpha}(\mathbb{R}^d)}^2 N^{2-\frac{2\lambda}{d}}.$$

Noting that by assumption $\mu^t \in \cap_{p=1}^\infty L^p(\mathbb{R}^d)$ and taking p to be the conjugate exponent to q,

$$\|\nabla \mathsf{F} * \mu^t\|_{L^\infty} \le \|\nabla \mathsf{F}\|_{L^q}\|\mu^t\|_{L^p},$$

we then immediately deduce from Lemma 2.1, Lemma 2.3 and Proposition 1.1 that

$$\partial_t F_N(X_N^t, \mu^t) \le C\left(\|\nabla^2 h^{\mu^t}\|_{L^\infty(\mathbb{R}^d)} + \|\nabla \mathsf{F}\|_{L^q(\mathbb{R}^d)} + \|\mathsf{F}\|_{\dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d)}^2\right)$$

$$\times \left[\left(F_N(X_N^t, \mu^t) + (1 + \|\mu^t\|_{L^\infty})N^{2-\frac{d-s}{d(d+1)}} + N^{\frac{3}{2}} + 2\left(\frac{N}{d}\log N\right)\mathbf{1}_{(1.4)}\right) + C\|\mathsf{F}\|_{C^{0,\alpha}(\mathbb{R}^d)}^2 N^{2-\frac{2\lambda}{d}}\right.$$

$$\left. + C\begin{cases}(1 + \|\mu^t\|_{C^\sigma})N^{2-\frac{1}{d}} + \left(\|\nabla h^{\mu^t}\|_{L^\infty} + \|\nabla^2 h^{\mu^t}\|_{L^\infty}\right)\|\mu\|_{C^\sigma}N^{1+\frac{s+1-\sigma}{d}}\right] \text{ if } \mathsf{s} \ge \mathsf{d}-1 \\ (1 + \|\mu^t\|_{L^\infty})N^{2-\frac{1}{d}} + \|\nabla h^{\mu^t}\|_{L^\infty}\|\mu^t\|_{L^\infty}N^{1+\frac{s+1}{d}} + \|\nabla^2 h^{\mu^t}\|_{L^\infty}\|\mu^t\|_{L^\infty})N^{1+\frac{s}{d}}\right] \text{ if } \mathsf{s} < \mathsf{d}-1.\end{cases}$$

Since $\mathsf{s} < \mathsf{d}$ and $\sigma > \mathsf{s} - \mathsf{d} + 1$, this implies by Gronwall's lemma and in view of (1.17) that for every $t \le T$,

$$F_N(X_N^t, \mu^t) \le \left(F_N(X_N^0, \mu^0) + C_1 N^\beta\right)e^{C_2 t} \quad \text{for some } \beta < 2.$$

In view Proposition 3.6 below, this proves the main theorem.

3. FORMULATION VIA THE ELECTRIC POTENTIAL

3.1. **The extension representation for the fractional Laplacian.** In general, the kernel g is not the convolution kernel of a local operator, but rather of a fractional Laplacian. Here we use the *extension representation* popularized by [CaffSi]: by adding one space variable

$y \in \mathbb{R}$ to the space $\mathbb{R}^{\mathsf{d}}$, the nonlocal operator can be transformed into a local operator of the form $-\operatorname{div}(|z|^{\gamma}\nabla\cdot)$.

In what follows, k will denote the dimension extension. We will take $\mathsf{k} = 0$ in the Coulomb cases for which g itself is the kernel of a local operator. In all other cases, we will take $\mathsf{k} = 1$.

For now, points in the space $\mathbb{R}^{\mathsf{d}}$ will be denoted by x, and points in the extended space $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ by X, with $X = (x, z)$, $x \in \mathbb{R}^{\mathsf{d}}$, $z \in \mathbb{R}^{\mathsf{k}}$. We will often identify $\mathbb{R}^{\mathsf{d}} \times \{0\}$ and $\mathbb{R}^{\mathsf{d}}$ and thus $(x_i, 0)$ with x_i.

If γ is chosen such that

$$\text{(3.1)} \qquad \qquad \mathsf{d} - 2 + \mathsf{k} + \gamma = \mathsf{s},$$

then, given a probability measure μ on $\mathbb{R}^{\mathsf{d}}$, the g-potential generated by μ, defined in $\mathbb{R}^{\mathsf{d}}$ by

$$h^{\mu}(x) := \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(x - \tilde{x}) \, d\mu(\tilde{x})$$

can be extended to a function h^{μ} on $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ defined by

$$\mathsf{h}^{\mu}(X) := \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(X - (\tilde{x}, 0)) \, d\mu(\tilde{x}),$$

and this function satisfies

$$\text{(3.2)} \qquad \qquad -\operatorname{div}(|z|^{\gamma}\nabla\mathsf{h}^{\mu}) = \mathsf{c}_{\mathsf{d},\mathsf{s}}\mu\delta_{\mathbb{R}^{\mathsf{d}}}$$

where by $\delta_{\mathbb{R}^{\mathsf{d}}}$ we mean the uniform measure on $\mathbb{R}^{\mathsf{d}} \times \{0\}$. The corresponding values of the constants $\mathsf{c}_{\mathsf{d},\mathsf{s}}$ are given in [PS, Section 1.2]. In particular, the potential g seen as a function of $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ satisfies

$$\text{(3.3)} \qquad \qquad -\operatorname{div}(|z|^{\gamma}\nabla\mathsf{g}) = \mathsf{c}_{\mathsf{d},\mathsf{s}}\delta_0.$$

To summarize, we will take

- $\mathsf{k} = 0, \gamma = 0$ in the Coulomb cases. *The reader only interested in the Coulomb cases may thus just ignore the k and the weight $|z|^{\gamma}$ in all the integrals.*
- $\mathsf{k} = 1, \gamma = \mathsf{s} - \mathsf{d} + 2 - \mathsf{k}$ in the Riesz cases and in the one-dimensional logarithmic case (then we mean $\mathsf{s} = 0$). Note that our assumption $(\mathsf{d} - 2)_+ \leq \mathsf{s} < \mathsf{d}$ implies that γ is always in $(-1, 1)$. We refer to [PS, Section 1.2] for more details.

We now make a remark on the regularity of h^{μ}:

Lemma 3.1. *Assume μ is a probability density in $C^{\theta}(\mathbb{R}^{\mathsf{d}})$ for some $\theta > \mathsf{s} - \mathsf{d} + 2$, then we have*

$$\text{(3.4)} \qquad \qquad \|\nabla h^{\mu}\|_{L^{\infty}(\mathbb{R}^{\mathsf{d}})} \leq C\left(\|\mu\|_{C^{\theta-1}(\mathbb{R}^{\mathsf{d}})} + \|\mu\|_{L^1(\mathbb{R}^{\mathsf{d}})}\right),$$

and

$$\text{(3.5)} \qquad \qquad \|\nabla^2 h^{\mu}\|_{L^{\infty}(\mathbb{R}^{\mathsf{d}})} \leq C\left(\|\mu\|_{C^{\theta}(\mathbb{R}^{\mathsf{d}})} + \|\mu\|_{L^1(\mathbb{R}^{\mathsf{d}})}\right).$$

Proof. As is well known, g is (up to a constant) the kernel of $\Delta^{\frac{\mathsf{d}-\mathsf{s}}{2}}$, hence $h^{\mu} = \mathsf{c}_{\mathsf{d},\mathsf{s}}\Delta^{\frac{\mathsf{s}-\mathsf{d}}{2}}\mu$ and the relations follow (cf. also [Du, Lemma 2.5]). $\qquad \square$

3.2. **Electring rewriting of the energy.** We briefly recall the procedure used in [RS, PS] for truncating the interaction or, equivalently, spreading out the point charges. It will also be crucial to use the variant introduced in [LSZ, LS2] where we let the truncation distance depend on the point.

For any $\eta \in (0, 1)$, we define

$$(3.6) \qquad \mathsf{g}_\eta := \min(\mathsf{g}, \mathsf{g}(\eta)), \quad \mathsf{f}_\eta := \mathsf{g} - \mathsf{g}_\eta$$

and

$$(3.7) \qquad \delta_0^{(\eta)} := -\frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}} \mathrm{div}\,(|z|^\gamma \nabla \mathsf{g}_\eta),$$

which is a positive measure supported on $\partial B(0, \eta)$.

Remark 3.2. *This nonsmooth truncation of g_η can be replaced with no change by a smooth one such that*

$$\mathsf{g}_\eta(x) = \mathsf{g}(x) \text{ for } |x| \geq \eta, \quad \mathsf{g}_\eta(x) = cst \text{ for } |x| \leq \eta - \varepsilon \qquad \varepsilon < \frac{1}{2}\eta$$

and this way $\delta_0^{(\eta)}$ gets replaced by a probability measure with a regular density supported in $B(0, \eta) \backslash B(0, \eta - \varepsilon)$. We make this modification whenever the integrals against the singular measures may not be well-defined.

We will also let

$$(3.8) \qquad \mathsf{f}_{\alpha,\eta} := \mathsf{f}_\alpha - \mathsf{f}_\eta = \mathsf{g}_\eta - \mathsf{g}_\alpha,$$

and we observe that $\mathsf{f}_{\alpha,\eta}$ has the sign of $\alpha - \eta$, vanishes outside $B(0, \max(\alpha, \eta))$, and satisfies

$$(3.9) \qquad \mathsf{g} * (\delta_x^{(\eta)} - \delta_x^{(\alpha)}) = \mathsf{f}_{\alpha,\eta}(\cdot - x)$$

and

$$(3.10) \qquad -\mathrm{div}\,(|z|^\gamma \nabla \mathsf{f}_{\alpha,\eta}) = \mathsf{c}_{\mathsf{d},\mathsf{s}}(\delta_0^{(\eta)} - \delta_0^{(\alpha)}).$$

For any configuration $X_N = (x_1, \ldots, x_N)$, we define for any i the minimal distance

$$(3.11) \qquad \mathsf{r}_i = \min\left(\frac{1}{4} \min_{j \neq i} |x_i - x_j|, N^{-\frac{1}{\mathsf{d}}}\right).$$

For any $\vec{\eta} = (\eta_1, \ldots, \eta_N) \in \mathbb{R}^N$ and measure μ, we define the electric potential

$$(3.12) \qquad H_N^\mu[X_N] = \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{g}(x - y) d\left(\sum_{i=1}^N \delta_{x_i} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(y)$$

and the truncated potential

$$(3.13) \qquad H_{N,\vec{\eta}}^\mu[X_N] = \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{g}(x - y) d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(y),$$

where we will quickly drop the dependence in X_N. We note that

$$(3.14) \qquad H_{N,\vec{\eta}}^\mu[X_N] = H_N^\mu[X_N] - \sum_{i=1}^N \mathsf{f}_\eta(x - x_i).$$

These functions are viewed in the extended space $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ as described in the previous subsection, and solve

$$(3.15) \qquad -\operatorname{div}\left(|z|^{\gamma}\nabla H_N^{\mu}\right) = \mathsf{c}_{\mathsf{d},\mathsf{s}}\left(\sum_{i=1}^{N}\delta_{x_i} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right) \quad \text{in } \mathbb{R}^{\mathsf{d}+\mathsf{k}},$$

and

$$(3.16) \qquad -\operatorname{div}\left(|z|^{\gamma}\nabla H_{N,\vec{\eta}}^{\mu}\right) = \mathsf{c}_{\mathsf{d},\mathsf{s}}\left(\sum_{i=1}^{N}\delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right) \quad \text{in } \mathbb{R}^{\mathsf{d}+\mathsf{k}}.$$

The following proposition shows how to express F_N in terms of the truncated electric fields $\nabla H_{N,\vec{\eta}}^{\mu}$. In addition, we show that the quantities

$$\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\eta}}^{\mu}|^2 - \mathsf{c}_{\mathsf{d},\mathsf{s}}\sum_{i=1}^{N}\mathsf{g}(\eta_i)$$

converge almost monotonically (i.e. up to a small error) to F_N, while the discrepancy between the two can serve to control the energy of close pairs of points.

Proposition 3.3. *Let μ be a bounded probability density on $\mathbb{R}^{\mathsf{d}}$ and X_N be in $(\mathbb{R}^{\mathsf{d}})^N$. We may re-write $F_N(X_N,\mu)$ as*

$$(3.17) \qquad F_N(X_N,\mu) := \frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\lim_{\eta\to 0}\left(\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\eta}}^{\mu}|^2 - \mathsf{c}_{\mathsf{d},\mathsf{s}}\sum_{i=1}^{N}\mathsf{g}(\eta_i)\right),$$

and for any $\vec{\eta}$ we have the bound

$$(3.18) \quad \sum_{i\neq j}\left(\mathsf{g}(x_i - x_j) - \mathsf{g}(\eta_i)\right)_{+}$$

$$\leq F_N(X_N,\mu) - \left(\frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\eta}}^{\mu}|^2 - \sum_{i=1}^{N}\mathsf{g}(\eta_i)\right) + CN\|\mu\|_{L^{\infty}}\sum_{i=1}^{N}\eta_i^{\mathsf{d}-\mathsf{s}},$$

for some C depending only on d and s.

The proof, which is an adaptation and improvement of [PS,LS2], is postponed to Section 5.

What makes our main proof work is the ability to find some choice of truncation $\vec{\eta}$ such that $\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\eta}}^{\mu}|^2$ (*without* the renormalizing term $-\mathsf{c}_{\mathsf{d},\mathsf{s}}\sum_{i=1}^{N}\mathsf{g}(\eta_i)$) is controlled by $F_N(X_N,\mu)$ and the balls $B(x_i,\eta_i)$ are disjoint. In view of (3.18) the former could easily be achieved by taking the η_i's large enough, say $\eta_i = N^{-1/\mathsf{d}}$, but the balls would not necessarily be disjoint. Instead the choice of $\eta_i = \mathsf{r}_i$ where r_i are the minimal distances as in (3.11) allows to fulfill both requirements, as seen in the following

Corollary 3.4. *Under the same assumptions, we have*
$$(3.19)$$
$$\sum_{i=1}^{N}\mathsf{g}(\mathsf{r}_i) \leq C\left(F_N(X_N,\mu) + (1 + \|\mu\|_{L^{\infty}})N^{1+\frac{\mathsf{s}}{\mathsf{d}}} + \left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)}\right) + C\left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)}$$
and
$$(3.20) \qquad \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\mathsf{r}}}^{\mu}|^2 \leq C\left(F_N(X_N,\mu) + (1 + \|\mu\|_{L^{\infty}})N^{1+\frac{\mathsf{s}}{\mathsf{d}}} + \left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)}\right)$$
for some C depending only on s,d.

Proof. Let us choose $\eta_i = N^{-1/\mathsf{d}}$ for all i in (3.18) and observe that for each i, by definition (3.11) there exists $j \neq i$ such that $(\mathsf{g}(|x_i - x_j|) - \mathsf{g}(N^{-1/\mathsf{d}}))_+ = (\mathsf{g}(4\mathsf{r}_i) - \mathsf{g}(N^{-1/\mathsf{d}}))_+$. We may thus write that

(3.21)
$$\sum_{i=1}^{N}(\mathsf{g}(4\mathsf{r}_i) - \mathsf{g}(N^{-\frac{1}{\mathsf{d}}}))_+ \leq F_N(X_N, \mu) - \frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{\eta}}|^2 + N\mathsf{g}(N^{-\frac{1}{\mathsf{d}}}) + O(N\|\mu\|_{L^\infty})N^{\frac{\mathsf{s}}{\mathsf{d}}}.$$

from which (3.19) follows.

Let us next choose $\eta_i = \mathsf{r}_i$ in (3.18). Using that $\mathsf{r}_i \leq N^{-1/\mathsf{d}}$, this yields

$$0 \leq F_N(X_N, \mu) - \frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla H_{N,\vec{r}}|^2 + \sum_{i=1}^{N}\mathsf{g}(\mathsf{r}_i) + O(N\|\mu\|_{L^\infty})N^{\frac{\mathsf{s}}{\mathsf{d}}}.$$

Combining with (3.19), (3.20) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

From (3.20) we directly obtain that F_N is bounded below:

Corollary 3.5. *Under the same assumptions we have*

(3.22)
$$F_N(X_N, \mu) \geq -\left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)} - CN^{1+\frac{\mathsf{s}}{\mathsf{d}}}$$

for some $C > 0$ depending only on d, s and $\|\mu\|_{L^\infty}$.

3.3. **Coerciveness of the modulated energy.** Here we prove that the modulated energy does metrize the convergence of μ_N^t to μ^t.

Proposition 3.6. *For any $0 < \alpha \leq 1$, there exists $\lambda > 0$ and $C > 0$ depending only on α, d, s, such that for any $X_N \in (\mathbb{R}^{\mathsf{d}})^N$, any probability density μ, and any $\xi \in C^\infty(\mathbb{R}^{\mathsf{d}})$, we have*

(3.23)
$$\left|\int_{\mathbb{R}^{\mathsf{d}}}\xi d\left(\sum_{i=1}^{N}\delta_{x_i} - N\mu\right)\right| \leq C\|\xi\|_{C^{0,\alpha}(\mathbb{R}^{\mathsf{d}})}N^{1-\frac{\lambda}{\mathsf{d}}}$$

$$+ C\|\xi\|_{\dot{H}^{\frac{\mathsf{d}-\mathsf{s}}{2}}(\mathbb{R}^{\mathsf{d}})}\left(F_N(X_N, \mu) + (1 + \|\mu\|_{L^\infty})N^{1+\frac{\mathsf{s}}{\mathsf{d}}} + 2\left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)}\right)^{\frac{1}{2}}.$$

In particular, if $\frac{1}{N^2}F_N(X_N, \mu) \to 0$ as $N \to \infty$, we have that

(3.24)
$$\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i} \rightharpoonup \mu \quad \text{in the weak sense.}$$

Proof. Let ξ be a smooth test function on $\mathbb{R}^{\mathsf{d}}$. Let $\bar{\xi}$ denote an extension of ξ to $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ satisfying

$$-\text{div}\left(|z|^{\gamma}\nabla\bar{\xi}\right) = 0 \text{ in } \{z \neq 0\}.$$

By [FKS], $|z|^{\gamma}$ being a Muckenhoupt A_2 weight, the function $\bar{\xi}$ is in $C^{0,\lambda}(\mathbb{R}^{\mathsf{d}+\mathsf{k}})$ for some $\lambda > 0$ depending on the other parameters, with $\|\bar{\xi}\|_{C^{0,\lambda}} \leq C\|\xi\|_{C^{0,\alpha}}$. This can also be seen from the Poisson kernel representation given in [CaffSi]. In addition, we also have (this can be seen in Fourier, see [CaffSi, Section 3.2]

(3.25)
$$\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}}|z|^{\gamma}|\nabla\bar{\xi}|^2 = C\|\xi\|_{\dot{H}^{\frac{\mathsf{d}-\mathsf{s}}{2}}(\mathbb{R}^{\mathsf{d}})}^2.$$

Using (3.16) let us write for any probability density μ

$$(3.26) \quad \int_{\mathbb{R}^d} \xi \, d \left(\sum_{i=1}^N \delta_{x_i} - N\mu \right) = \int_{\mathbb{R}^d} \xi d \left(\sum_{i=1}^N \delta_{x_i} - \delta_{x_i}^{(r_i)} \right) - \frac{1}{c_{d,s}} \int_{\mathbb{R}^{d+k}} \bar{\xi} \operatorname{div} (|z|^\gamma \nabla H_{N,\vec{r}}^\mu [X_N]).$$

For the first term in the right-hand side we use the Hölder continuity of $\bar{\xi}$ and the fact that $\delta_{x_i}^{(r_i)}$ is supported in $B(x_i, r_i)$ and write

$$(3.27) \quad \left| \int_{\mathbb{R}^d} \xi d \sum_{i=1}^N \left(\delta_{x_i} - \delta_{x_i}^{(r_i)} \right) \right| \le C \|\xi\|_{C^{0,\alpha}(\mathbb{R}^d)} \sum_{i=1}^N r_i^\lambda \le C \|\xi\|_{C^{0,\alpha}(\mathbb{R}^d)} N^{1-\frac{\lambda}{d}}.$$

For the second term, we integrate by parts and use the Cauchy-Schwarz inequality to write

$$(3.28) \quad \left| \int_{\mathbb{R}^{d+k}} \bar{\xi} \operatorname{div} (|z|^\gamma \nabla H_{N,\vec{r}}^\mu [X_N]) \right| = \left| \int_{\mathbb{R}^{d+k}} |z|^\gamma \nabla \bar{\xi} \cdot \nabla H_{N,\vec{r}}^\mu \right|$$
$$\le \left(\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla \bar{\xi}|^2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla H_{N,\vec{r}}^\mu|^2 \right)^{\frac{1}{2}}$$

In view of (3.20) and (3.25) we thus find

$$(3.29) \quad \left| \int_{\mathbb{R}^{d+k}} \bar{\xi} \operatorname{div} (|z|^\gamma \nabla H_{N,\vec{r}}^\mu [X_N]) \right|$$
$$\le C \|\xi\|_{\dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d)} \left(F_N(X_N, \mu) + (1 + \|\mu\|_{L^\infty}) N^{1+\frac{s}{d}} + 2 \left(\frac{N}{d} \log N \right) \mathbf{1}_{(1.4)} \right)^{\frac{1}{2}}.$$

Inserting (3.27) and (3.29) into (3.26) we conclude the result. $\qquad\square$

Remark 3.7. *In a density formulation aiming at proving propagation of chaos, arguing exactly as in [RS, Lemma 8.4] for instance, we may deduce from this result and the main theorem the convergence of the k-marginal densities in the dual of some Sobolev space, with rate k/N times the right-hand side of* (3.23).

3.4. **Proof of Lemma 2.3.** Using the Cauchy-Schwarz inequality we have the bound

$$\iint_{\triangle^c} \mathsf{F} * (\mu^t - \mu_N^t) \cdot \nabla \mathsf{g}(x - y) d(\mu_N^t - \mu^t)(y) d\mu_N^t(x)$$
$$\le \left(\int_{\mathbb{R}^d} |\mathsf{F} * (\mu^t - \mu_N^t)|^2 d\mu_N^t \right)^{\frac{1}{2}} \left(\int \left| P.V. \int_{\mathbb{R}^d \setminus \{x\}} \nabla \mathsf{g}(x - y) d(\mu_N^t - \mu^t)(y) \right|^2 d\mu_N^t(x) \right)^{\frac{1}{2}}$$

thus to prove the lemma, it suffices to show that

$$(3.30) \quad N \|\mathsf{F} * (\mu^t - \mu_N^t)\|_{L^\infty} \le C \|\mathsf{F}\|_{C^{0,\alpha}(\mathbb{R}^d)} N^{1-\frac{\lambda}{d}}$$
$$+ C \|\mathsf{F}\|_{\dot{H}^{\frac{d-s}{2}}(\mathbb{R}^d)} \left(F_N(X_N, \mu) + (1 + \|\mu\|_{L^\infty}) N^{1+\frac{s}{d}} + 2 \left(\frac{N}{d} \log N \right) \mathbf{1}_{(1.4)} \right)^{\frac{1}{2}},$$

which is a direct consequence of (3.23).

4. Proof of Proposition 1.1

4.1. **Stress-energy tensor.**

Definition 4.1. For any functions h, f in $\mathbb{R}^{d+k}$ such that $\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla h|^2$ and $\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla f|^2$ are finite, we define the stress tensor $[h, f]$ as the $(d + k) \times (d + k)$ tensor

$$(4.1) \qquad [h, f] = |z|^\gamma \left(\partial_i h \partial_j f + \partial_i h \partial_j f \right) - |z|^\gamma \nabla h \cdot \nabla f \delta_{ij}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

We note that

Lemma 4.2. *If h and f are regular enough, we have*

$$(4.2) \qquad \operatorname{div}[h, f] = \operatorname{div}\left(|z|^\gamma \nabla h\right) \nabla f + \operatorname{div}\left(|z|^\gamma \nabla f\right) \nabla h - \nabla |z|^\gamma \nabla h \cdot \nabla f$$

where $\operatorname{div} T$ here denotes the vector with components $\sum_i \partial_i T_{ij}$, with j ranging from 1 to $d + k$.

Proof. This is a direct computation. Below, all sums range from 1 to $d + k$.

$$\sum_i \partial_i [h, f]_{ij}$$

$$= \sum_i \left[\partial_i(|z|^\gamma \partial_i h)\partial_j f + \partial_i(|z|^\gamma \partial_i f)\partial_j h + |z|^\gamma \partial_{ij} h \partial_i f + |z|^\gamma \partial_{ij} f \partial_i h \right] - \partial_j \left(|z|^\gamma \sum_i \partial_i h \partial_i f \right)$$

$$= \operatorname{div}\left(|z|^\gamma \nabla h\right) + \operatorname{div}\left(|z|^\gamma \nabla f\right) - \nabla h \cdot \nabla f \partial_j |z|^\gamma.$$

$\square$

In view of (4.2), we have

Lemma 4.3. *Let $\psi : \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz, and if $k = 1$ let $\hat{\psi}$ be an extension of it to a map from $\mathbb{R}^{d+k}$ to $\mathbb{R}^{d+k}$, whose last component identically vanishes, which tends to 0 as $|z| \to \infty$ and has the same pointwise and Lipschitz bounds as ψ.* [2] *For any measures μ, ν on $\mathbb{R}^{d+k}$, if $-\operatorname{div}\left(|z|^\gamma \nabla h^\mu\right) = c_{d,s}\mu$ and $-\operatorname{div}\left(|z|^\gamma \nabla h^\nu\right) = c_{d,s}\nu$, and assuming that $\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla h^\mu|^2$ and $\int_{\mathbb{R}^{d+k}} |z|^\gamma |\nabla h^\nu|^2$ are finite and the left-hand side in (4.3) is well-defined, we have*

$$(4.3) \qquad \iint_{\mathbb{R}^{d+k} \times \mathbb{R}^{d+k}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla g(x - y) d\mu(x) d\nu(y) = \frac{1}{c_{d,s}} \int_{\mathbb{R}^{d+k}} \nabla \hat{\psi}(x) : [h^\mu, h^\nu].$$

Proof. If μ is smooth enough then we may use (4.2) to write

$$\iint_{\mathbb{R}^{d+k} \times \mathbb{R}^{d+k}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla g(x - y) d\mu(x) d\nu(y) = \int_{\mathbb{R}^{d+k}} \hat{\psi} \cdot \left(\nabla h^\mu d\nu + \nabla h^\nu d\mu \right)$$

$$= -\frac{1}{c_{d,s}} \int_{\mathbb{R}^{d+k}} \hat{\psi} \cdot \operatorname{div}[h^\mu, h^\nu]$$

since the last component of $\hat{\psi}$ vanishes identically. Integrating by parts, we obtain

$$\iint_{\mathbb{R}^{d+k} \times \mathbb{R}^{d+k}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla g(x - y) d\mu(x) d\nu(y) = \frac{1}{c_{d,s}} \int_{\mathbb{R}^{d+k}} \nabla \hat{\psi} : [h^\mu, h^\nu].$$

By density, we may extend this relation to all measures μ, ν such that both sides of (4.3) make sense. $\square$

---

[2] Such an extension exists, for instance by solving the ∞-Laplacian in a strip, which provides an "absolutely minimal Lipschitz extension"

4.2. **Proof of Proposition 1.1.** We now proceed to the proof. Given the Lipschitz map $\psi : \mathbb{R}^{\mathsf{d}} \to \mathbb{R}^{\mathsf{d}}$, we choose an extension $\hat{\psi}$ to $\mathbb{R}^{\mathsf{d}+\mathsf{k}}$ which satisfies the same conditions as in Lemma 4.3.

Step 1: *renormalizing the quantity and expressing it with the stress-energy tensor.* Clearly,

$$(4.4) \quad \iint_{\triangle^c} (\psi(x) - \psi(y)) \cdot \nabla \mathsf{g}(x-y) d(\sum_{i=1}^N \delta_{x_i} - N\mu)(x) d(\sum_{i=1}^N \delta_{x_i} - N\mu)(y)$$

$$= \lim_{\eta \to 0} \Big[\iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d(\sum_{i=1}^N \delta_{x_i}^{(\eta)} - N\mu \delta_{\mathbb{R}^{\mathsf{d}}})(x) d(\sum_{i=1}^N \delta_{x_i}^{(\eta)} - N\mu \delta_{\mathbb{R}^{\mathsf{d}}})(y)$$

$$- \sum_{i=1}^N \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\eta)}(x) d\delta_{x_i}^{(\eta)}(y)\Big].$$

Applying Lemma 4.3, in view of (3.13), (3.16) we find that

$$(4.5) \quad \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d(\sum_{i=1}^N \delta_{x_i}^{(\eta)} - N\mu \delta_{\mathbb{R}^{\mathsf{d}}})(x) d(\sum_{i=1}^N \delta_{x_i}^{(\eta)} - N\mu \delta_{\mathbb{R}^{\mathsf{d}}})(y)$$

$$= \frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \nabla \hat{\psi} : [H_{N,\vec{\eta}}^{\mu}, H_{N,\vec{\eta}}^{\mu}].$$

Step 2: *analysis of the diagonal terms.* The main point is to understand how they vary with η_i. Let $\vec{\alpha}$ be such that $\alpha_i \geq \eta_i$ for every i.

We may write that

$$(4.6) \quad \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\eta_i)}(x) d\delta_{x_i}^{(\eta_i)}(y)$$

$$- \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\alpha_i)}(x) d\delta_{x_i}^{(\alpha_i)}(y)$$

$$= \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla \mathsf{g}(x-y) d(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)})(x) d(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)})(y)$$

$$+ 2 \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\alpha_i)}(x) d(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)})(y).$$

We claim that

$$(4.7) \quad \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} (\hat{\psi}(x) - \hat{\psi}(y)) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\alpha_i)}(x) d(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)})(y) = 0.$$

Assuming this, inserting it to (4.6) and using (3.9) and Lemma 4.3, we conclude that

$$(4.8) \quad \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\eta)}(x) d\delta_{x_i}^{(\eta)}(y)$$

$$- \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\hat{\psi}(x) - \hat{\psi}(y)\right) \cdot \nabla \mathsf{g}(x-y) d\delta_{x_i}^{(\alpha_i)}(x) d\delta_{x_i}^{(\alpha_i)}(y)$$

$$= \frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \nabla \hat{\psi} : [\mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i), \mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i)].$$

Step 3: *proof of* (4.7). Let us write the quantity in (4.7) as

$$(4.9) \quad 2\iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}}(\hat\psi(x)-\hat\psi(y))\cdot\nabla\mathsf{g}(x-y)d\delta_{x_i}^{(2\alpha_i)}(x)d(\delta_{x_i}^{(\eta_i)}-\delta_{x_i}^{(\alpha_i)})(y)$$

$$+2\iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}}(\hat\psi(x)-\hat\psi(y))\cdot\nabla\mathsf{g}(x-y)d\left(\delta_{x_i}^{(\alpha_i)}-\delta_{x_i}^{(2\alpha_i)}\right)(x)d\left(\delta_{x_i}^{(\eta_i)}-\delta_{x_i}^{(\alpha_i)}\right)(y).$$

In view of (3.7), $\nabla\mathsf{g}*\delta_{x_i}^{(2\alpha_i)}=\nabla\mathsf{g}_{2\alpha_i}(\cdot-x_i)$ and in view of (3.9), $\nabla\mathsf{g}*(\delta_{x_i}^{(\eta_i)}-\delta_{x_i}^{(\alpha_i)})=\nabla\mathsf{f}_{\alpha_i,\eta_i}(x-x_i)$. We may thus rewrite the first term in (4.9) as

$$2P.V.\int_{\mathbb{R}^{d+k}}\hat\psi\cdot\nabla\mathsf{f}_{\alpha_i,\eta_i}(\cdot-x_i)d\delta_{x_i}^{(2\alpha_i)}+2P.V.\int_{\mathbb{R}^{d+k}}\hat\psi\cdot\nabla\mathsf{g}_{2\alpha_i}(\cdot-x_i)d\left(\delta_{x_i}^{(\eta_i)}-\delta_{x_i}^{(\alpha_i)}\right).$$

But $\nabla\mathsf{f}_{\alpha_i,\eta_i}(\cdot-x_i)$ is supported in $B(x_i,\alpha_i)$ while $\delta_{x_i}^{(2\alpha_i)}$ is supported on $\partial B(x_i,2\alpha_i)$, and in the same way $\nabla\mathsf{g}_{\alpha_i}(\cdot-x_i)$ vanishes in $B(x_i,2\alpha_i)$ where $\delta_{x_i}^{(\eta_i)}-\delta_{x_i}^{(\alpha_i)}$ is supported, so we conclude that the first term in (4.9) is zero. The second term in (4.9) is equal by (3.9) and Lemma 4.3 to

$$\frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\int_{\mathbb{R}^{d+k}}\nabla\hat\psi:[\mathsf{f}_{2\alpha_i,\alpha_i}(\cdot-x_i),\mathsf{f}_{\alpha_i,\eta_i}(\cdot-x_i)]$$

and it is zero, since $\mathsf{f}_{2\alpha_i,\alpha_i}$ and $\mathsf{f}_{\alpha_i,\eta_i}$ have disjoint supports. This finishes the proof of (4.7).

Step 4: *combining* (4.5) *and* (4.8). The following lemma allows to recombine the terms obtained at different values of η_i while making only a small error.

Lemma 4.4. *Assume that* $\mu\in C^\sigma(\mathbb{R}^\mathsf{d})$ *with* $\sigma>\mathsf{s}-\mathsf{d}+1$ *if* $\mathsf{s}\geq\mathsf{d}-1$. *Assume* $\mu\in L^\infty(\mathbb{R}^\mathsf{d})$ *or* $\mu\in C^\sigma(\mathbb{R}^\mathsf{d})$ *with* $\sigma>0$ *if* $\mathsf{s}<\mathsf{d}-1$. *If for each* i *we have* $\eta_i<\alpha_i\leq\mathsf{r}_i$, *then*

$$\int_{\mathbb{R}^{d+k}}\nabla\hat\psi:[H_{N,\vec\eta}^\mu,H_{N,\vec\eta}^\mu]=\int_{\mathbb{R}^{d+k}}\nabla\hat\psi:[H_{N,\vec\alpha}^\mu,H_{N,\vec\alpha}^\mu]+\sum_{i=1}^N\int_{\mathbb{R}^{d+k}}\nabla\hat\psi:[\mathsf{f}_{\alpha_i,\eta_i}(\cdot-x_i),\mathsf{f}_{\alpha_i,\eta_i}(\cdot-x_i)]+\mathcal{E}$$

with

$$(4.10) \quad |\mathcal{E}|\leq C\|\nabla\psi\|_{L^\infty}\left(F_N(X_N,\mu)+\left(\frac{N}{\mathsf{d}}\log N\right)\mathbf{1}_{(1.4)}+(1+\|\mu\|_{L^\infty})N^{2-\frac{\mathsf{d}-\mathsf{s}}{\mathsf{d}(\mathsf{d}+1)}}\right)$$

$$+CN\min\left(\|\psi\|_{L^\infty}\|\mu\|_{L^\infty}\sum_{i=1}^N\alpha_i^{\mathsf{d}-\mathsf{s}-1}+\|\nabla\psi\|_{L^\infty}\|\mu\|_{L^\infty}\sum_{i=1}^N\alpha_i^{\mathsf{d}-\mathsf{s}},\|\psi\|_{W^{1,\infty}}\|\mu\|_{C^\sigma}\sum_{i=1}^N\alpha_i^{\mathsf{d}-\mathsf{s}+\sigma-1}\right)$$

$$+CN\begin{cases}\|\nabla\psi\|_{L^\infty}(1+\|\mu\|_{C^\sigma})\sum_{i=1}^N\alpha_i & \text{if }\mathsf{s}\geq\mathsf{d}-1\\ \|\nabla\psi\|_{L^\infty}(1+\|\mu\|_{L^\infty})\sum_{i=1}^N\alpha_i & \text{if }\mathsf{s}<\mathsf{d}-1\end{cases}$$

where C *depends only on* s,d.

Assuming this, and combining (4.4), (4.5) and (4.8) we find that for any $\alpha_i\leq\mathsf{r}_i$,

$$\iint_{\triangle^c}(\psi(x)-\psi(y))\cdot\nabla\mathsf{g}(x-y)d(\sum_{i=1}^N\delta_{x_i}-N\mu)(x)d(\sum_{i=1}^N\delta_{x_i}-N\mu)(y)=\frac{1}{\mathsf{c}_{\mathsf{d},\mathsf{s}}}\int_{\mathbb{R}^{d+k}}\nabla\hat\psi:[H_{N,\vec\alpha}^\mu,H_{N,\vec\alpha}^\mu]$$

$$-\sum_{i=1}^N\iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}}\left(\hat\psi(x)-\hat\psi(y)\right)\cdot\nabla\mathsf{g}(x-y)d\delta_{x_i}^{(\alpha_i)}(x)d\delta_{x_i}^{(\alpha_i)}(y)+O(\mathcal{E})$$

where $\mathcal{E}$ is as in (4.10). Using the Lipschitz character of ψ and the expression of g, we find that the second term on the right-hand side can be bounded by [3]

$$C\|\nabla\psi\|_{L^\infty}\sum_{i=1}^N \iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}} \mathsf{g}(x-y)d\delta_{x_i}^{(\alpha_i)}(x)d\delta_{x_i}^{(\alpha_i)}(y)$$

$$= C\|\nabla\psi\|_{L^\infty}\sum_{i=1}^N \int_{\mathbb{R}^{d+k}} \mathsf{g}_{\alpha_i}(\cdot - x_i)d\delta_{x_i}^{(\alpha_i)} = C\|\nabla\psi\|_{L^\infty}\sum_{i=1}^N \mathsf{g}(\alpha_i)$$

where we have used (3.7). Choosing finally $\alpha_i = \mathsf{r}_i \leq N^{-1/\mathsf{d}}$, bounding pointwise $[H_{N,\vec{\mathsf{r}}}^\mu, H_{N,\vec{\mathsf{r}}}^\mu]$ by $2|z|^\gamma|\nabla H_{N,\vec{\mathsf{r}}}^\mu|^2$ and using (3.20), while using (3.19) to bound $\sum_{i=1}^N \mathsf{g}(\mathsf{r}_i)$, we conclude the proof of Proposition 1.1.

Proof of Lemma 4.4. First, we observe from (3.14) that $[H_{N,\vec{\alpha}}^\mu, H_{N,\vec{\alpha}}^\mu]$ and $[H_{N,\vec{\eta}}^\mu, H_{N,\vec{\eta}}^\mu]$ only differ in the balls $B(x_i, \alpha_i)$ which are disjoint since $\alpha_i \leq \mathsf{r}_i$, and that in each $B(x_i, \alpha_i)$ we have

$$H_{N,\vec{\eta}}^\mu = H_{N,\vec{\alpha}}^\mu + \mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i).$$

We thus deduce that

$$(4.11) \quad \int_{B(x_i,\alpha_i)} \nabla\hat{\psi} : \left([H_{N,\vec{\eta}}^\mu, H_{N,\vec{\eta}}^\mu] - [H_{N,\vec{\alpha}}^\mu, H_{N,\vec{\alpha}}^\mu]\right)$$

$$= \int_{B(x_i,\alpha_i)} \nabla\hat{\psi} : \left([\mathsf{f}_{\alpha_i,\eta_i}, \mathsf{f}_{\alpha_i,\eta_i}](\cdot - x_i) + 2[\mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i), H_{N,\vec{\alpha}}^\mu]\right)$$

$$= \int_{\mathbb{R}^{d+k}} \nabla\hat{\psi} : \left([\mathsf{f}_{\alpha_i,\eta_i}, \mathsf{f}_{\alpha_i,\eta_i}](\cdot - x_i) + 2[\mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i), H_{N,\vec{\alpha}}^\mu]\right).$$

There only remains to control the second part of the right-hand side. By Lemma 4.3, we have

$$\int_{\mathbb{R}^{d+k}} \nabla\hat{\psi} : [\mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i), H_{N,\vec{\alpha}}^\mu]$$

$$= \mathsf{c}_{\mathsf{d},\mathsf{s}} \iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}} (\hat{\psi}(x) - \hat{\psi}(y))\cdot\nabla\mathsf{g}(x-y)d\left(\sum_{j=1}^N \delta_{x_j}^{(\alpha_j)} - N\mu\delta_{\mathbb{R}^d}\right)(x)d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)(y).$$

In view of (4.7), we just need to bound the sum over i of

$$(4.12)$$

$$\mathsf{c}_{\mathsf{d},\mathsf{s}} \iint_{\mathbb{R}^{d+k}\times\mathbb{R}^{d+k}} (\hat{\psi}(x) - \hat{\psi}(y))\cdot\nabla\mathsf{g}(x-y)d\left(\sum_{j:j\neq i}\delta_{x_j}^{(\alpha_j)} - N\mu\delta_{\mathbb{R}^d}\right)(x)d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)(y)$$

$$= \mathsf{c}_{\mathsf{d},\mathsf{s}}\int_{\mathbb{R}^{d+k}} \hat{\psi}\cdot\nabla\mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i)d\left(\sum_{j:j\neq i}\delta_{x_j}^{(\alpha_j)} - N\mu\delta_{\mathbb{R}^d}\right)$$

$$+ \mathsf{c}_{\mathsf{d},\mathsf{s}}\int_{\mathbb{R}^{d+k}} \hat{\psi}\cdot\left(\sum_{j:j\neq i}\nabla\mathsf{g}_{\alpha_j}(x-x_j) - N\nabla\mathsf{h}^\mu\right)d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right),$$

where we used (3.9).

_____

[3]In the case (1.4) we bound instead $|x-y||\nabla\mathsf{g}(x-y)|$ by 1, which yields an even better control.

Step 1: *first term in* (4.12). Since $\mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i)$ is supported in $B(x_i, \alpha_i)$, $\delta_{x_j}^{(\alpha_j)}$ in $B(x_j, \alpha_j)$ and the balls are disjoint, one type of terms vanishes and there remains

$$-N\mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}}} \psi \cdot \nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i) d\mu.$$

Thanks to the explicit form of $\mathsf{f}_{\alpha, \eta}$ we have

$$\mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i) = \begin{cases} \mathsf{g}(x - x_i) - \mathsf{g}(\alpha_i) & \text{for } \eta_i < |x - x_i| \le \alpha_i \\ \mathsf{g}(\eta_i) - \mathsf{g}(\alpha_i) & \text{for } |x - x_i| \le \eta_i \end{cases}$$

and

$$\nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i) = \nabla \mathsf{g}(x - x_i) \mathbf{1}_{\eta_i \le |x - x_i| \le \alpha_i}.$$

It follows that

$$(4.13) \qquad \int_{\mathbb{R}^{\mathsf{d}}} |\mathsf{f}_{\alpha}| \le C\alpha^{\mathsf{d}-\mathsf{s}}, \qquad \int_{\mathbb{R}^{\mathsf{d}}} |\mathsf{f}_{\alpha_i, \eta_i}| \le C\alpha_i^{\mathsf{d}-\mathsf{s}}, \qquad \int_{\mathbb{R}^{\mathsf{d}}} |\nabla \mathsf{f}_{\alpha_i, \eta_i}| \le C\alpha_i^{\mathsf{d}-\mathsf{s}-1}.$$

Indeed, it suffices to observe that

$$(4.14) \qquad \int_{B(0,\eta)} \mathsf{f}_{\alpha} = C \int_0^{\alpha} (\mathsf{g}(r) - \mathsf{g}(\alpha)) r^{\mathsf{d}-1} \, dr = -\frac{C}{\mathsf{d}} \int_0^r \mathsf{g}'(r) r^{\mathsf{d}} \, dr,$$

with an integration by parts.

We may always write

$$(4.15) \quad \left| \int_{\mathbb{R}^{\mathsf{d}}} \psi \cdot \nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i)(\mu - \mu(x_i)) \right| \le C\|\psi\|_{L^{\infty}} \|\mu\|_{C^{0,1}} \int_{\eta_i}^{\alpha_i} \frac{r^{\mathsf{d}}}{r^{\mathsf{s}+1}} dr$$
$$\le C\|\psi\|_{L^{\infty}} \|\mu\|_{C^{0,1}} \alpha_i^{\mathsf{d}-\mathsf{s}},$$

and, integrating by parts and using (4.13),

$$(4.16) \quad \left| \int_{\mathbb{R}^{\mathsf{d}}} \psi \cdot \nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i)\mu(x_i) \right| \le \|\mu\|_{L^{\infty}} \|\nabla\psi\|_{L^{\infty}} \int_{\mathbb{R}^{\mathsf{d}}} |\mathsf{f}_{\alpha_i, \eta_i}| \le \|\mu\|_{L^{\infty}} \|\nabla\psi\|_{L^{\infty}} \alpha_i^{\mathsf{d}-\mathsf{s}}.$$

Alternatively, we may use the simpler bound derived from (4.13),

$$(4.17) \qquad \left| \int_{\mathbb{R}^{\mathsf{d}}} \psi \cdot \nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i) d\mu \right| \le C\|\psi\|_{L^{\infty}} \|\mu\|_{L^{\infty}} \alpha_i^{\mathsf{d}-\mathsf{s}-1}.$$

A standard interpolation argument yields that $\|g\|_{(C^{\sigma})^*} \le \|g\|_{(C^1)^*}^{\sigma} \|g\|_{(C^0)^*}^{1-\sigma}$ so interpolating between (4.15)–(4.16) and (4.17), we obtain

$$\left| \int_{\mathbb{R}^{\mathsf{d}}} \psi \cdot \nabla \mathsf{f}_{\alpha_i, \eta_i}(\cdot - x_i) d\mu \right| \le C\|\psi\|_{L^{\infty}}^{1-\sigma} \|\psi\|_{W^{1,\infty}}^{\sigma} \|\mu\|_{C^{\sigma}} \alpha_i^{\mathsf{d}-\mathsf{s}+\sigma-1}.$$

We conclude that the sum over i of the first terms in (4.12) is bounded by both

$$(4.18) \qquad C\|\psi\|_{L^{\infty}}^{1-\sigma} \|\psi\|_{W^{1,\infty}}^{\sigma} \|\mu\|_{C^{\sigma}} \sum_i \alpha_i^{\mathsf{d}-\mathsf{s}+\sigma-1} \quad \text{and} \quad C\|\psi\|_{L^{\infty}} \|\mu\|_{L^{\infty}} \sum_i \alpha_i^{\mathsf{d}-\mathsf{s}-1}.$$

Step 2: *second term in* (4.12). We may rewrite the integral as

$$(4.19) \quad -N \int_{\mathbb{R}^{d+k}} \hat{\psi}(x_i) \cdot \nabla_{\mathbb{R}^d} h^\mu d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)$$

$$+ \sum_{j:j\neq i} \int_{\mathbb{R}^{d+k}} \hat{\psi}(x_j) \cdot \nabla g_{\alpha_j}(x - x_j) d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)$$

$$+ \sum_{j:j\neq i} \int_{\mathbb{R}^{d+k}} (\hat{\psi} - \hat{\psi}(x_j)) \cdot \nabla g_{\alpha_j}(x - x_j) d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)$$

$$+ O\left(N\|\nabla\psi\|_{L^\infty} \int_{\mathbb{R}^{d+k}} |x - x_i| \, |\nabla_{\mathbb{R}^d} h^\mu| \, d\left(\delta_{x_i}^{(\eta_i)} + \delta_{x_i}^{(\alpha_i)}\right)\right),$$

where we used that the last component of $\hat{\psi}$ vanishes, so that only the derivatives along the $\mathbb{R}^d$ directions appear.

Substep 2.1: *first term of* (4.19). We may write that $\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)} = -\frac{1}{c_{d,s}} \text{div}\,(|z|^\gamma \nabla f_{\alpha_i,\eta_i}(\cdot - x_i))$ and integrate by parts twice to get

$$\frac{1}{c_{d,s}} \int_{\mathbb{R}^{d+k}} |z|^\gamma \nabla \left(\hat{\psi}(x_i) \cdot \nabla_{\mathbb{R}^d} h^\mu\right) \cdot \nabla f_{\alpha_i,\eta_i}(\cdot - x_i) = \int_{\mathbb{R}^d} (\psi(x_i) \cdot \nabla\mu) f_{\alpha_i,\eta_i}.$$

Here, we used that $-\text{div}\,(|z|^\gamma \nabla h^\mu) = c_{d,s}\mu\delta_{\mathbb{R}^d}$ and took the $\hat{\psi}(x_i) \cdot \nabla_{\mathbb{R}^d}$ of this relation. In view of (4.13), this is then bounded by

$$C\|\psi\|_{L^\infty} \|\nabla\mu\|_{L^\infty} \alpha_i^{d-s}.$$

Alternatively, we may integrate by parts in $\mathbb{R}^d$ to bound it by

$$C\|\mu\|_{L^\infty}\left(\|\psi\|_{L^\infty} \int_{\mathbb{R}^d} |\nabla f_{\alpha_i,\eta_i}| + \|\nabla\psi\|_{L^\infty} \alpha_i^{d-s}\right).$$

Interpolating as above, we conclude with (4.13) that the sum over i of these terms is bounded by both

$$(4.20) \quad C\|\mu\|_{C^\sigma}\left(\|\psi\|_{L^\infty}^\sigma \|\nabla\psi\|_{L^\infty}^{1-\sigma} \alpha_i^{d-s} + \|\psi\|_{L^\infty} \sum_i \alpha_i^{d-s+\sigma-1}\right)$$

$$\text{and } \|\mu\|_{L^\infty}\left(\|\psi\|_{L^\infty} \sum_i \alpha_i^{d-s-1} + \|\nabla\psi\|_{L^\infty} \sum_i \alpha_i^{d-s}\right).$$

Substep 2.2: *second term of* (4.19). Arguing in the same way as for the first term, using that $-\text{div}\,(|z|^\gamma \nabla g_{\alpha_j}(\cdot - x_j)) = c_{d,s}\delta_{x_j}^{(\alpha_j)}$ and the disjointness of the balls, we find that this term vanishes.

Substep 2.3: *third term in* (4.19). We separate the sum into two pieces and bound this term by

$$(4.21) \quad \sum_{j\neq i, |x_i-x_j|\geq N^{-\frac{\varepsilon}{d}}} \int_{\mathbb{R}^{d+k}} (\hat{\psi} - \hat{\psi}(x_j)) \cdot \nabla g_{\alpha_j}(\cdot - x_j) d\left(\delta_{x_i}^{(\eta_i)} - \delta_{x_i}^{(\alpha_i)}\right)$$

$$+ \|\nabla\psi\|_{L^\infty} \sum_{j\neq i, |x_i-x_j|\leq N^{-\frac{\varepsilon}{d}}} \int_{\mathbb{R}^{d+k}} |x - x_j| |\nabla g_{\alpha_j}(x - x_j)| d\left(\delta_{x_i}^{(\eta_i)} + \delta_{x_i}^{(\alpha_i)}\right)(x),$$

for some $\varepsilon > 0$ to be determined. For the first term of (4.21), we may use that $|x_i - x_j| \geq N^{-\frac{\varepsilon}{\mathsf{d}}}$ to write

$$\|\nabla\left((\hat{\psi} - \hat{\psi}(x_j)) \cdot \nabla\mathsf{g}_{\alpha_j}(x - x_j)\right)\|_{L^\infty(B(x_i, \mathsf{r}_i))} \leq C\|\nabla\psi\|_{L^\infty} N^{\frac{\varepsilon(\mathsf{s}+1)}{\mathsf{d}}}$$

and using that $\eta_i \leq \alpha_i \leq N^{-\frac{1}{\mathsf{d}}}$ we may thus bound the sum of such terms by

$$C\|\nabla\psi\|_{L^\infty} N^{\frac{\varepsilon(\mathsf{s}+1)}{\mathsf{d}} + 2 - \frac{1}{\mathsf{d}}}.$$

Since $|\nabla\mathsf{g}_\alpha| \leq |\nabla\mathsf{g}|$, we may bound the second term in (4.21) by

$$(4.22) \qquad \|\nabla\psi\|_{L^\infty} \sum_{j \neq i, |x_i - x_j| \leq N^{-\frac{\varepsilon}{\mathsf{d}}}} |x_i - x_j|^{-\mathsf{s}}.$$

To bound this, let us choose $\eta_i = 2N^{-\frac{\varepsilon}{\mathsf{d}}}$ in (3.18) to obtain that

$$\sum_{i \neq j} \left(\mathsf{g}(x_i - x_j) - \mathsf{g}(2N^{-\frac{\varepsilon}{\mathsf{d}}})\right)_+ \leq F_N(X_N, \mu) + N\mathsf{g}(2N^{-\frac{\varepsilon}{\mathsf{d}}}) + C\|\mu\|_{L^\infty} N^{2 - \frac{\varepsilon(\mathsf{d}-\mathsf{s})}{\mathsf{d}}}.$$

In the cases (1.3), it follows that

$$\sum_{i \neq j, |x_i - x_j| \leq N^{-\frac{\varepsilon}{\mathsf{d}}}} \mathsf{g}(x_i - x_j) \leq C\left(F_N(X_N, \mu) + N^{1 + \frac{(\mathsf{d}-\mathsf{s})\varepsilon}{\mathsf{d}}} + \|\mu\|_{L^\infty} N^{2 - \frac{\varepsilon(\mathsf{d}-\mathsf{s})}{\mathsf{d}}}\right).$$

In the cases (1.4), it follows that

$$\sum_{i \neq j, |x_i - x_j| \leq N^{-\frac{\varepsilon}{\mathsf{d}}}} \log 2 \leq F_N(X_N, \mu) + \frac{N}{\mathsf{d}} \log N + C(1 + \|\mu\|_{L^\infty})N,$$

and this suffices to bound (4.22) as well. Choosing $\varepsilon = \frac{1}{\mathsf{d}+1}$, we conclude in all cases that the sum over i of the third terms in (4.19) is bounded by

$$(4.23) \qquad C\|\nabla\psi\|_{L^\infty}\left(F_N(X_N, \mu) + (1 + \|\mu\|_{L^\infty})N^{2 - \frac{\mathsf{d}-\mathsf{s}}{\mathsf{d}(\mathsf{d}+1)}} + \left(\frac{N}{\mathsf{d}} \log N\right) \mathbf{1}_{(1.4)}\right).$$

Substep 2.4: *fourth term in* (4.19). We may bound it by $O\left(\|\nabla\psi\|_{L^\infty} \alpha_i \|\nabla_{\mathbb{R}^\mathsf{d}} \mathsf{h}^\mu\|_{L^\infty(\mathbb{R}^{\mathsf{d}+\mathsf{k}})}\right)$. But since $\mathsf{h}^\mu = \mathsf{g} * \mu$, it is straightforward to check that $\|\nabla_{\mathbb{R}^\mathsf{d}} \mathsf{h}^\mu\|_{L^\infty(\mathbb{R}^{\mathsf{d}+\mathsf{k}})} \leq \|\nabla h^\mu\|_{L^\infty(\mathbb{R}^\mathsf{d})}$. Using (3.4), we conclude the sum of these terms is bounded by

$$(4.24) \qquad \begin{cases} C\sum_i \alpha_i \|\nabla\psi\|_{L^\infty}(1 + \|\mu\|_{L^\infty}) & \text{if } \mathsf{s} < \mathsf{d} - 1 \\ C\sum_i \alpha_i \|\nabla\psi\|_{L^\infty}(1 + \|\mu\|_{C^\sigma}) & \text{if } \mathsf{s} \geq \mathsf{d} - 1 \text{ and } \sigma > \mathsf{s} - \mathsf{d} + 1. \end{cases}$$

Combining the bounds (4.18), (4.20), (4.23), (4.24) concludes the proof of the lemma. $\square$

5. Proof of Proposition 3.3

We drop the superscripts μ. First, $\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} |z|^\gamma |\nabla H_{N, \vec{\eta}}|^2$ is a convergent integral and

$$(5.1)$$
$$\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} |z|^\gamma |\nabla H_{N, \vec{\eta}}|^2 = \mathsf{c}_{\mathsf{d}, \mathsf{s}} \iint_{\mathbb{R}^{\mathsf{d}+\mathsf{k}} \times \mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{g}(x - y) d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^\mathsf{d}}\right)(x) d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^\mathsf{d}}\right)(y).$$

Indeed, we may choose R large enough so that all the points of X_N are contained in the ball $B_R = B(0, R)$ in $\mathbb{R}^{\mathsf{d+k}}$. By Green's formula and (3.16), we have

$$\int_{B_R} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2 = \int_{\partial B_R} |z|^\gamma H_{N,\vec{\eta}} \frac{\partial H_N}{\partial n} + \mathsf{c}_{\mathsf{d,s}} \int_{B_R} H_{N,\vec{\eta}} \, d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right).$$

Since $\int d(\sum_i \delta_{x_i} - N\mu) = 0$, the function $H_{N,\vec{\eta}}$ decreases like $1/|x|^{\mathsf{s+1}}$ and $\nabla H_{N,\vec{\eta}}$ like $1/|x|^{\mathsf{s+2}}$ as $|x| \to \infty$, hence the boundary integral tends to 0 as $R \to \infty$, and we may write

$$\int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2 = \mathsf{c}_{\mathsf{d,s}} \int_{\mathbb{R}^{\mathsf{d+k}}} H_{N,\vec{\eta}} \, d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)$$

and thus by (3.16), (5.1) holds. We may next write that

$$\lim_{\eta \to 0} \Big[\iint_{\mathbb{R}^{\mathsf{d+k}} \times \mathbb{R}^{\mathsf{d+k}}} \mathsf{g}(x-y) d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(x) d\left(\sum_{i=1}^N \delta_{x_i}^{(\eta_i)} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(y) - \sum_{i=1}^N \mathsf{g}(\eta_i) \Big]$$

$$= \iint_{\triangle^c} \mathsf{g}(x-y) d\left(\sum_{i=1}^N \delta_{x_i} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(x) d\left(\sum_{i=1}^N \delta_{x_i} - N\mu\delta_{\mathbb{R}^{\mathsf{d}}}\right)(y)$$

and we deduce in view of (5.1) that (3.17) holds.

We next turn to the proof of (3.18), adapted from [PS]. From (3.16) applied with $\vec{\eta} = \vec{\alpha}$ and in view of (3.13), we have $\nabla H_{N,\vec{\eta}} = \nabla H_{N,\vec{\alpha}} + \sum_{i=1}^N \nabla \mathsf{f}_{\alpha_i,\eta_i}(\cdot - x_i)$ thus

$$\int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2 = \int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma |\nabla H_{N,\vec{\alpha}}|^2 + \sum_{i,j} \int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \cdot \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_j)$$

$$+ 2\sum_{i=1}^N \int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \cdot \nabla H_{N,\vec{\alpha}}.$$

Using (3.10), we first write

$$\sum_{i,j} \int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \cdot \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_j)$$

$$= -\sum_{i,j} \int_{\mathbb{R}^{\mathsf{d+k}}} \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \mathrm{div}\left(|z|^\gamma \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_j)\right) = \mathsf{c}_{\mathsf{d,s}} \sum_{i,j} \int_{\mathbb{R}^{\mathsf{d+k}}} \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) d(\delta_{x_j}^{(\eta_i)} - \delta_{x_j}^{(\alpha_i)}).$$

Next, using (3.16), we write

$$2\sum_{i=1}^N \int_{\mathbb{R}^{\mathsf{d+k}}} |z|^\gamma \nabla \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \cdot \nabla H_{N,\vec{\alpha}} = -2\sum_{i=1}^N \int_{\mathbb{R}^{\mathsf{d+k}}} \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \mathrm{div}\left(|z|^\gamma \nabla H_{N,\vec{\alpha}}\right)$$

$$= 2\mathsf{c}_{\mathsf{d,s}} \sum_{i=1}^N \int_{\mathbb{R}^{\mathsf{d+k}}} \mathsf{f}_{\alpha_i,\eta_i}(x-x_i) \, d\Big(\sum_{j=1}^N \delta_{x_j}^{(\alpha_i)} - \mu\delta_{\mathbb{R}^{\mathsf{d}}}\Big).$$

These last two equations add up to give a right-hand side equal to

$$(5.2) \quad \sum_{i \neq j} \mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{f}_{\alpha_i,\eta_i}(x - x_i) d(\delta_{x_j}^{(\alpha_i)} + \delta_{x_j}^{(\eta_j)}) - 2\mathsf{c}_{\mathsf{d},\mathsf{s}} \sum_{i=1}^{N} \int \mathsf{f}_{\alpha_i,\eta_i}(x - x_i) d\mu \delta_{\mathbb{R}^{\mathsf{d}}}$$

$$+ N\mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{f}_{\alpha_i,\eta_i} d(\delta_0^{(\alpha_i)} + \delta_0^{(\eta_i)}) \, .$$

We then note that $\int \mathsf{f}_{\alpha_i,\eta_i} d(\delta_0^{(\alpha_i)} + \delta_0^{(\eta_i)}) = - \int \mathsf{f}_{\eta_i} d\delta_0^{(\alpha_i)} = -(\mathsf{g}(\alpha_i) - \mathsf{g}(\eta_i))$ by definition of f_η and the fact that $\delta_0^{(\alpha)}$ is a measure supported on $\partial B(0, \alpha)$ and of mass 1. Secondly, by (4.13) we may bound $\int_{\mathbb{R}^{\mathsf{d}}} \mathsf{f}_{\alpha_i,\eta_i}(x - x_i)\mu \delta_{\mathbb{R}^{\mathsf{d}}}$ by $C\|\mu\|_{L^\infty} \alpha_i^{\mathsf{d}-\mathsf{s}}$.

Thirdly, we observe that since $\mathsf{f}_{\alpha_i,\eta_i} \leq 0$, the first term in (5.2) is nonpositive and we may bound it above by

$$\sum_{i \neq j} \mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \mathsf{f}_{\alpha_i,\eta_i} d\delta_{x_j}^{(\alpha_j)} \leq \sum_{i \neq j} \mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\mathsf{g}_{\eta_i}(x - x_i) - \mathsf{g}_{\alpha_i}(x - x_i) \right) d\delta_{x_j}^{(\alpha_j)}$$

$$\leq \sum_{i \neq j} \mathsf{c}_{\mathsf{d},\mathsf{s}} \int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} \left(\mathsf{g}(\eta_i) - \mathsf{g}_{\alpha_i}(|x_i - x_j| + \alpha_j) \right)_-$$

where we used the fact that g_α is radial decreasing. Combining the previous relations yields

$$- CN\|\mu\|_{L^\infty} \sum_{i=1}^{N} \eta_i^{\mathsf{d}-\mathsf{s}} + \mathsf{c}_{\mathsf{d},\mathsf{s}} \sum_{i \neq j} \left(\mathsf{g}_{\alpha_i}(|x_i - x_j| + \alpha_j) - \mathsf{g}(\eta_i) \right)_+$$

$$\leq \left(\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} |z|^\gamma |\nabla H_{N,\vec{\alpha}}|^2 - \mathsf{c}_{\mathsf{d},\mathsf{s}} \sum_{i=1}^{N} \mathsf{g}(\alpha_i) \right) - \left(\int_{\mathbb{R}^{\mathsf{d}+\mathsf{k}}} |z|^\gamma |\nabla H_{N,\vec{\eta}}|^2 - \mathsf{c}_{\mathsf{d},\mathsf{s}} \sum_{i=1}^{N} \mathsf{g}(\eta_i) \right)$$

and letting all $\alpha_i \to 0$ finishes the proof in view of (3.17).

Appendix A. Mean-field limit for monokinetic Vlasov systems, with Mitia Duerinckx

In this appendix, we turn to examining the mean-field limit of solutions of Newton's second-order system of ODEs, that is,

$$(A.1) \quad \begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = -\frac{1}{N} \nabla_{x_i} \mathcal{H}_N(x_1, \ldots, x_N), & i = 1, \ldots, N \\ x_i(0) = x_i^0, \quad v_i(0) = v_i^0, \end{cases}$$

where $\mathcal{H}_N$ is the interaction energy defined in (1.3). We then consider the phase-space empirical measure

$$f_N^t := \frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i^t, v_i^t)} \in \mathcal{P}(\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}}),$$

where $\mathcal{P}$ denotes probability measures, associated to a solution $Z_N^t := ((x_1^t, v_1^t), \ldots, (x_N^t, v_N^t))$ of the flow (A.1). If the points (x_i^0, v_i^0), which themselves depend on N, are such that f_N^0 converges to some regular measure f^0, then formal computations indicate that for $t > 0$, f_N^t

should converge to the solution of the Cauchy problem with initial data f^0 for the following Vlasov equation,

$$(A.2) \qquad \partial_t f + v \cdot \nabla_x f + (\nabla \mathsf{g} * \mu) \cdot \nabla_v f = 0, \qquad \mu^t(x) := \int_{\mathbb{R}^{\mathsf{d}}} f^t(x,v)\, dv.$$

In the case of a smooth interaction kernel g, such a mean-field result was first established in the 1970s by a weak compactness argument [NW, BH], while Dobrushin [Do] gave the first quantitative proof in 1-Wasserstein distance. In recent years much attention has been given to the physically more relevant case of singular Coulomb-like kernels, but only very partial results have been obtained. In [HJ2, HJ1], exploiting a Grönwall argument on the ∞-Wasserstein distance between f_N and f, Hauray and Jabin treated all interaction kernels g satisfying $|\nabla g(x)| \leq C|x|^{-s}$ and $|\nabla^2 g(x)| \leq C|x|^{-s-1}$ for some $s < 1$. In [JW1], Jabin and Wang introduced a new approach, allowing them to treat all interaction kernels with bounded gradient. The same problem has been addressed in [BP, La, LP], leading to results that require a small N-dependent cutoff of the interaction kernel.

In this appendix, we show that the method presented in the article allows to unlock the mean-field problem with Coulomb-like interaction in the simpler case of monokinetic data, that is, if there exists a regular velocity field $u^0 : \mathbb{R}^{\mathsf{d}} \to \mathbb{R}^{\mathsf{d}}$ such that $v_i^0 \approx u^0(x_i^0)$ for $i = 1, \ldots, N$, which implies that the solutions remain monokinetic, at least for short times. Justifying a complete mean-field result for non-monokinetic solutions in the Coulomb case in dimensions $\mathsf{d} \geq 2$ remains one of the main open problems in the field. It is expected to be rendered difficult by the possibility of concentration and filamentation in both space and velocity variables.

In the monokinetic setting, we focus on the (spatial) empirical measure

$$\mu_N^t := \frac{1}{N} \sum_{i=1}^N \delta_{x_i^t} \in \mathcal{P}(\mathbb{R}^{\mathsf{d}}),$$

associated to a solution Z_N^t of the flow (A.1). If the points x_i^0 are such that μ_N^0 converges to some regular measure μ^0, then formal computations (see for instance [LZ]) lead to expecting that for $t > 0$, μ_N^t converges to the solution μ^t of the Cauchy problem with initial data (μ^0, u^0) of the following monokinetic version of $(A.2)^4$,

$$(A.3) \qquad \partial_t \mu + \mathrm{div}\,(\mu u) = 0, \qquad \partial_t u + u \cdot \nabla u = -\nabla \mathsf{g} * \mu.$$

In the Coulomb case, these equations are known as the pressureless (repulsive) Euler-Poisson system. Since for the second-order system (A.1) the total energy splits into a potential and a kinetic part, we naturally introduce a modulated *total* energy taking both parts into account: for $Z_N := ((x_1, v_1), \ldots, (x_N, v_N))$ and for a couple $(\mu, u) \in \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \times C(\mathbb{R}^{\mathsf{d}})$, we set

$$H_N(Z_N, (\mu, u)) := N \sum_{i=1}^N |u(x_i) - v_i|^2$$

$$+ \iint_{\mathbb{R}^{\mathsf{d}} \times \mathbb{R}^{\mathsf{d}} \backslash \triangle} \mathsf{g}(x-y)\, d\Big(\sum_{i=1}^N \delta_{x_i} - N\mu\Big)(x)\, d\Big(\sum_{i=1}^N \delta_{x_i} - N\mu\Big)(y),$$

---

[4]Formally, the pair (μ, u) indeed satisfies the system (A.3) if and only if the monokinetic ansatz $f^t(x,v) := \mu^t(x)\delta_{v=u^t(x)}$ satisfies the Vlasov equation (A.2).

and we view $H_N(Z_N^t, (\mu^t, u^t))$ as a good notion of distance between μ_N^t and μ^t that is adapted to the monokinetic setting. The condition $H_N(Z_N^t, (\mu^t, u^t)) = o(N^2)$ indeed entails the weak convergence $\mu_N^t \rightharpoonup \mu^t$ as in Proposition 3.6, while due to the kinetic part it also implies that the flow Z_N^t remains monokinetic, that is, $v_i^t \approx u^t(x_i^t)$. In parallel with Theorem 1, our main result is a Gronwall inequality on the time-derivative of $H_N(Z_N^t, (\mu^t, u^t))$, which implies a quantitative rate of convergence of μ_N^t to μ^t in that metric.

Theorem 2. *Assume that* g *is of the form* (1.3) *or* (1.4). *Assume that* (A.3) *admits a solution* (μ, u) *in* $[0, T] \times \mathbb{R}^d$ *for some* $T > 0$, *with* $\mu \in L^\infty([0, T]; \mathcal{P} \cap L^\infty(\mathbb{R}^d))$ *and* $u \in L^\infty([0, T]; W^{1,\infty}(\mathbb{R}^d)^d)$. *In the case* $s \geq d - 1$, *also assume* $\mu \in L^\infty([0, T]; C^\sigma(\mathbb{R}^d))$ *for some* $\sigma > s - d + 1$. *Let* Z_N^t *solve* (A.1). *Then there exist constants* C_1, C_2 *depending only on controlled norms of* (μ, u) *and on* d, s, σ, *and there exists an exponent* $\beta < 2$ *depending only on* d, s, σ, *such that for every* $t \in [0, T]$ *we have*

$$H_N(Z_N^t, (\mu^t, u^t)) \leq \left(H_N(Z_N^0, (\mu^0, u^0)) + C_1 N^\beta \right) e^{C_2 t}.$$

In particular, if $\mu_N^0 \rightharpoonup \mu^0$ *and is such that*

$$\lim_{N \to \infty} \frac{1}{N^2} H_N(Z_N^0, \mu^0) = 0,$$

then the same is true for every $t \in [0, T]$, *hence* $\mu_N^t \rightharpoonup \mu^t$ *in the weak sense.*

Remark A.1. *We briefly comment on the regularity assumptions on the solution* (μ, u) *of the mean-field system* (A.3) *in the Coulomb case, that is, the pressureless (repulsive) Euler-Poisson system. The local-in-time existence of a smooth solution is easily established by standard Besov techniques, e.g. [BCD, Section 3.2]: for non-integer* $\sigma > 0$, *given initial data* $\mu^0 \in \mathcal{P} \cap C^\sigma(\mathbb{R}^d)$ *and* $u^0 \in C^{\sigma+1}(\mathbb{R}^d)$, *there exists some* $T > 0$ *such that* (A.3) *admits a unique solution* (μ, u) *in* $[0, T] \times \mathbb{R}^d$ *with* $\mu \in L^\infty([0, T]; \mathcal{P} \cap C^\sigma(\mathbb{R}^d))$ *and* $u \in L^\infty([0, T]; C^{\sigma+1}(\mathbb{R}^d))$. *While global existence of a weak solution was proved in [NT], together with a partial uniqueness result, the solution can in general remain smooth only locally in time due to possible wave breakdown [Pe, En]. Global smoothness was however established in some cases [CW, Gu]. The precise configuration of initial data happens to play a decisive role: a critical threshold phenomenon was discovered by Engelberg, Liu, and Tadmor [ELT], showing in dimension* 1 *that global smoothness holds whenever the initial data do not cross some intrinsic critical threshold, while finite-time breakdown occurs otherwise. In particular, global smoothness is expected to hold for a large set of initial configurations, but only partial results in this vein are still available in dimension* $d > 1$, *cf. [En, LT1, LT2, LL].*

Before turning to the proof of Theorem 2, we start with a weak-strong stability principle for the mean-field system (A.3), analogous to (1.21) for the mean-field equation (1.13) in the first-order case.

Lemma A.2. *Let* (μ_1^t, u_1^t) *and* (μ_2^t, u_2^t) *denote two smooth solutions of* (A.3) *in* $[0, T] \times \mathbb{R}^d$. *Then there is a constant* C *depending only on* d *such that, in terms of*

$$H((\mu_1^t, u_1^t), (\mu_2^t, u_2^t)) := \int_{\mathbb{R}^d} |u_1^t - u_2^t|^2 \mu_1^t + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(x - y) \, d(\mu_1^t - \mu_2^t)(x) \, d(\mu_1^t - \mu_2^t)(y),$$

we have

$$H((\mu_1^t, u_1^t), (\mu_2^t, u_2^t)) \leq e^{C \int_0^t \|\nabla u_2^u\|_{L^\infty} du} H((\mu_1^0, u_1^0), (\mu_2^0, u_2^0)).$$

Proof. Let $h_i := \mathsf{g} * \mu_i$. We compute

$$
\begin{aligned}
\partial_t H((\mu_1, u_1), (\mu_2, u_2)) &= \int_{\mathbb{R}^\mathsf{d}} |u_1 - u_2|^2 \partial_t \mu_1 + 2 \int_{\mathbb{R}^\mathsf{d}} (u_1 - u_2) \cdot (\partial_t u_1 - \partial_t u_2)\, \mu_1 \\
&\quad + 2 \int_{\mathbb{R}^\mathsf{d}} (h_1 - h_2)(\partial_t \mu_1 - \partial_t \mu_2) \\
&= \int_{\mathbb{R}^\mathsf{d}} \nabla |u_1 - u_2|^2 \cdot \mu_1 u_1 - 2 \int_{\mathbb{R}^\mathsf{d}} (u_1 - u_2) \cdot (u_1 \cdot \nabla u_1 - u_2 \cdot \nabla u_2)\, \mu_1 \\
&\quad - 2 \int_{\mathbb{R}^\mathsf{d}} \nabla(h_1 - h_2) \cdot (u_1 - u_2)\, \mu_1 + 2 \int_{\mathbb{R}^\mathsf{d}} \nabla(h_1 - h_2) \cdot (\mu_1 u_1 - \mu_2 u_2).
\end{aligned}
$$

Decomposing $2(u_1 - u_2) \cdot (u_1 \cdot \nabla u_1 - u_2 \cdot \nabla u_2) = u_1 \cdot \nabla |u_1 - u_2|^2 + 2\nabla u_2 : (u_1 - u_2) \otimes (u_1 - u_2)$, we find after straightforward simplifications,

$$
\begin{aligned}
&\partial_t H((\mu_1, u_1), (\mu_2, u_2)) \\
&\qquad = -2 \int_{\mathbb{R}^\mathsf{d}} \mu_1 \nabla u_2 : (u_1 - u_2) \otimes (u_1 - u_2) + 2 \int_{\mathbb{R}^\mathsf{d}} u_2 \cdot \nabla(h_1 - h_2)(\mu_1 - \mu_2).
\end{aligned}
$$

In the Coulomb case, we recognize from (1.25) in the second right-hand side term the divergence of the stress-energy tensor $[h_1 - h_2, h_1 - h_2]$, hence

$$
\begin{aligned}
&\partial_t H((\mu_1, u_1), (\mu_2, u_2)) \\
&= -2 \int_{\mathbb{R}^\mathsf{d}} \mu_1 \nabla u_2 : (u_1 - u_2) \otimes (u_1 - u_2) - \frac{1}{\mathsf{c_d}} \int_{\mathbb{R}^\mathsf{d}} u_2 \cdot \operatorname{div}[h_1 - h_2, h_1 - h_2] \\
&\leq C \|\nabla u_2\|_{L^\infty} \int_{\mathbb{R}^\mathsf{d}} |u_1 - u_2|^2 \mu_1 + \frac{C}{\mathsf{c_d}} \|\nabla u_2\|_{L^\infty} \int_{\mathbb{R}^\mathsf{d}} |\nabla(h_1 - h_2)|^2 \\
&= C \|\nabla u_2\|_{L^\infty} H((\mu_1, u_1), (\mu_2, u_2)),
\end{aligned}
$$

and the result follows by Gronwall's lemma. In the Riesz case, replacing (1.25) by (3.2) and by (4.1)–(4.2), the same conclusion follows. $\qquad\square$

Taking inspiration from the above calculations for the weak-strong stability principle, we now turn to the proof of Theorem 2.

Proof of Theorem 2. Let $h := \mathsf{g} * \mu$. We decompose

$$
\partial_t H_N(Z_N, (\mu, u)) = 2N \sum_{i=1}^N (u(x_i) - v_i) \cdot ((\partial_t u)(x_i) + \dot{x}_i \cdot \nabla u(x_i) - \dot{v}_i)
$$

$$
+ 2 \sum_{i \neq j} \dot{x}_i \cdot \nabla \mathsf{g}(x_i - x_j) + 2N^2 \int_{\mathbb{R}^\mathsf{d}} h \partial_t \mu - 2N \sum_{i=1}^N \int_{\mathbb{R}^\mathsf{d}} \dot{x}_i \cdot \nabla \mathsf{g}(x_i - y)\, d\mu(y) - 2N \sum_{i=1}^N \int_{\mathbb{R}^\mathsf{d}} \mathsf{g}(x_i - \cdot)\, \partial_t \mu.
$$

Inserting equations (A.1) and (A.3), we obtain

$$
\partial_t H_N(Z_N, (\mu, u))
$$
$$
= 2N \sum_{i=1}^N (u(x_i) - v_i) \cdot \Big(-(u \cdot \nabla u)(x_i) - \nabla h(x_i) + v_i \cdot \nabla u(x_i) + \frac{1}{N} \sum_{j:j\neq i} \nabla \mathsf{g}(x_i - x_j) \Big)
$$
$$
+ 2 \sum_{i\neq j} v_i \cdot \nabla \mathsf{g}(x_i - x_j) + 2N^2 \int_{\mathbb{R}^\mathsf{d}} \nabla h \cdot u\mu
$$
$$
- 2N \sum_{i=1}^N v_i \cdot \nabla h(x_i) - 2N \sum_{i=1}^N \int_{\mathbb{R}^\mathsf{d}} \nabla \mathsf{g}(\cdot - x_i) \cdot u \, d\mu,
$$

and hence, after straightforward simplifications,

$$
\partial_t H_N(Z_N, (\mu, u)) = -2N \sum_{i=1}^N \nabla u(x_i) : (u(x_i) - v_i) \otimes (u(x_i) - v_i)
$$
$$
+ N^2 \iint_{\mathbb{R}^\mathsf{d} \times \mathbb{R}^\mathsf{d} \setminus \triangle} (u(x) - u(y)) \cdot \nabla \mathsf{g}(x - y) \, d(\mu_N - \mu)(x) \, d(\mu_N - \mu)(y).
$$

Applying Proposition 1.1 to estimate the second right-hand side term, we are led to

$$
\partial_t H_N(Z_N, (\mu, u))
$$
$$
\leq C \|\nabla u\|_{L^\infty} \Big(H_N(Z_N, (\mu, u)) + (1 + \|\mu\|_{L^\infty}) N^{1+\frac{\mathsf{s}}{\mathsf{d}}} + N^{2-\frac{\mathsf{s}+1}{2}} + \Big(\frac{N}{\mathsf{d}} \log N \Big) \mathbf{1}_{(1.4)} \Big)
$$
$$
+ C \begin{cases} (1 + \|\mu\|_{C^\sigma}) \|\nabla u\|_{L^\infty} N^{2-\frac{1}{\mathsf{d}}} + \|u\|_{W^{1,\infty}} \|\mu\|_{C^\sigma} N^{1+\frac{\mathsf{s}+1-\sigma}{\mathsf{d}}}, & \text{if } \mathsf{s} \geq \mathsf{d} - 1, \\ (1 + \|\mu\|_{L^\infty}) \|\nabla u\|_{L^\infty} N^{2-\frac{1}{\mathsf{d}}} + \|u\|_{L^\infty} \|\mu\|_{L^\infty} N^{1+\frac{\mathsf{s}+1}{\mathsf{d}}} \\ \qquad\qquad + \|\nabla u\|_{L^\infty} \|\mu\|_{L^\infty} N^{1+\frac{\mathsf{s}}{\mathsf{d}}}, & \text{if } \mathsf{s} < \mathsf{d} - 1, \end{cases}
$$

and the conclusion follows by Gronwall's lemma. $\qquad \square$

References

[AS] L. Ambrosio, S. Serfaty, A gradient flow approach for an evolution problem arising in superconductivity, *Comm. Pure Appl. Math* **61** (2008), No. 11, 1495–1539.

[BCD] H. Bahouri, J.-Y. Chemin, R. Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2011.

[BaCh] F. Bach, L. Chizat, On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport, *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (2018), 3040–3050.

[BK] H. Bae, J. Kelliher, The vortex patches of Serfati, arXiv:1409.5169.

[BO] R. Berman, M. Onnheim, Propagation of chaos for a class of first order models with singular mean field interactions, *SIAM J. Math. Anal.* **51**, No 1, 159–196.

[BerCon] A. Bertozzi, P. Constantin, Global regularity for vortex patches, *Comm. Math. Phys.* **152** (1993), No. 1, 19–28.

[BIK] P. Biler, C. Imbert, G. Karch, Barenblatt profiles for a nonlocal porous medium equation, *C. R. Acad. Sci. Paris, Ser. I* **349** (2011), 641–645.

[BP] N. Boers, P. Pickl, On mean field limits for dynamical systems, *J. Stat. Phys.* **164** (2016), No. 1, 1–16.

[BCF] F. Bolley, D. Chafai, J. Fontbona, Dynamics of a planar Coulomb gas, *Ann. Appl. Probab.* **28** (2018), No. 5, 3152–3183.

[BFT] M. Bossy, O. Faugeras, D. Talay, Clarification and complement to "Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons", *J. Math. Neurosci.* **5** Art. 19, (2015), 23 pp.

[BH] W. Braun, K. Hepp, The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles, *Comm. Math. Phys.* **56** (1977), No. 2, 101–113.

[BJW1] 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. *C. R. Math. Acad. Sci. Paris* **357** (2019), no. 9, 708–720.

[BJW2] D. Bresch, P.E. Jabin, Z. Wang, Mean field limit and quantitative estimates with a large class of singular kernels, in preparation.

[CaffSi] L. Caffarelli, L. Silvestre, An extension problem related to the fractional Laplacian. *Comm. PDE* **32** (2007), No. 8, 1245–1260.

[CSV] L. Caffarelli, F. Soria, J.L. Vázquez, Regularity of solutions of the fractional porous medium flow. *J. Eur. Math. Soc.* **15** (2013), No. 5, 1701-1746.

[CV] L. Caffarelli, J.L. Vázquez. Nonlinear porous medium flow with fractional potential pressure. *Arch. Rat. Mech. Anal.* **202** (2011), No. 2, 537–565.

[CCS] J. A. Carrillo, Y. P. Choi, S. Salem, Propagation of chaos for the VPFP equation with a polynomial cut-off, *Commun. Contemp. Math.* **21** (2019), No. 4, 1850039, 28 pp.

[CCH] 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*, Springer (2014), p. 1–46.

[CFP] 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. in Math.* **231** (2012), 306–327.

[Ch1] J. Y Chemin, Persistance de structures géométriques dans les fluides incompressibles bidimensionnels, *Ann. Sci. Ecole Norm. Sup.* **26** (1993), No. 4, 517–542.

[Ch2] J.-Y. Chemin, *Perfect incompressible fluids.* Oxford Lecture Series in Mathematics and its Applications, 14. T Oxford University Press, 1998.

[CW] G.- Q. Chen, D. Wang, Convergence of shock capturing schemes for the compressible Euler-Poisson equations, *Comm. Math. Phys.* **179** (1996), No. 2, 333–364.

[De] J.-M. Delort, Existence de nappes de tourbillon en dimension deux, *J. Amer. Math. Soc* **4** (1991), 553–586.

[Do] R. L. Dobrushin, Funktsional. Anal. i Prilozhen. *Fjournal Akademiya Nauk SSSR. Funktsional'nyui Analiz i ego Prilozheniya* **13** (1979), No. 2, 48–58, 96.

[DZ] Q. Du, P. Zhang, Existence of weak solutions to some vortex density models, *SIAM J. Math. Anal.* **34** (2003), No. 6, 1279–1299.

[Du] M. Duerinckx, Mean-field limits for some Riesz interaction gradient flows, *SIAM J. Math. Anal.* **48** (2016), No. 3, 2269–2300.

[En] S. Engelberg, Formation of singularities in the Euler and Euler-Poisson equations, *Physica D* **98**, 67–74.

[ELT] S. Engelberg, H. Liu, E. Tadmor, Critical thresholds in Euler-Poisson equations, *Indiana Univ. Math. J.* **50** (2001), 109–157.

[FKS] E. Fabes, C. Kenig, R. Serapioni, The local regularity of solutions of degenerate elliptic equations, *Comm. PDE* **7** (1982), No. 1, 77–116.

[FHM] N. Fournier, M. Hauray, S. Mischler, Propagation of chaos for the 2D viscous vortex model, *J. Eur. Math. Soc.* **16** (2014), No. 7, 1423–1466.

[Go] 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, *Lect. Notes Appl. Math. Mech.* **3** (2016), 1-144.

[GP] F. Golse, T. Paul, Mean-Field and Classical Limit for the N-Body Quantum Dynamics with Coulomb Interaction, arXiv:1912.06750.

[GHL] J. Goodman, T. Hou, J. Lowengrub, Convergence of the point vortex method for the 2-D Euler equations. *Comm. Pure Appl. Math.* **43** (1990), No. 3, 415–430.

[Gu] Y. Guo, Smooth irrotational flows in the large to the Euler-Poisson system in $\mathbf{R}^{3+1}$, *Comm. Math. Phys.* **195** (1998), No. 2, 249–265.

[Hau] M. Hauray, Wasserstein distances for vortices approximation of Euler-type equations, *Math. Models Methods Appl. Sci.* **19** (2009), No. 8, 1357–1384.

[HM] M. Hauray, S. Mischler, On Kac's chaos and related problems, *JFA* **266** (2014), No. 10, 6055–6157.

[HJ1] M. Hauray, P-E. Jabin, Particle approximation of Vlasov equations with singular forces: propagation of chaos, *Ann. Sci. Ec. Norm. Super.* (4) **48** (2015), No. 4, 891–940.

[HJ2] M. Hauray, P-E. Jabin, *N*-particles approximation of the Vlasov equations with singular potential, *Arch. Ration. Mech. Anal.* **183** (2007), No. 3, 489–524.

[He] F. Hélein, *Harmonic Maps, Conservation Laws and Moving Frames*, Cambridge University Press, 2002.

[IP] A. D. Ionescu, B. Pausader, The Euler-Poisson system in 2D: global stability of the constant equilibrium solution, *Inter. Math. Research Not.* **4** (2013), 761–826.

[Jab] P. E. Jabin, A review of the mean field limits for Vlasov equations. *Kinet. Rel. Mod.* **7** (2014), 661–711.

[JW1] 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, 3588–3627.

[JW2] P. E. Jabin, Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels, *Invent. Math.* **214**, No 1, 523–591.

[La] D. Lazarovici, The Vlasov-Poisson dynamics as the mean field limit of extended charges, *Comm. Math. Phys.* **347**, No. 1, (2016), 271–289.

[LP] D. Lazarovici, P. Pickl, A mean field limit for the Vlasov-Poisson system. *Arch. Ration. Mech. Anal.* **225** (2017), No. 3, 1201–1231.

[LS1] T. Leblé, S. Serfaty, Large Deviation Principle for Empirical Fields of Log and Riesz gases, *Invent. Math.* **210**, No. 3, 645–757.

[LS2] T. Leblé, S. Serfaty, Fluctuations of Two-Dimensional Coulomb Gases, *GAFA* **28**, No. 2, (2018), 443–508.

[LSZ] T. Leblé, S. Serfaty, O. Zeitouni, Large deviations for the two-dimensional two-component plasma. *Comm. Math. Phys.* **350** (2017), No. 1, 301–360.

[LL] Y. K. Lee, H. L. Liu, Thresholds in three-dimensional restricted Euler-Poisson equations, *Physica D* **262** (2013), 59–70.

[LWJB] X. Li, J. G. Wöhlbier, S. Jin, J. H. Booske, Eulerian method for computing multivalued solutions of the Euler-Poisson equations and applications to wave breaking in klystrons, *Phys. Rev. E* **70** (2004), 016502.

[LZ] F. H. Lin, P. Zhang, On the hydrodynamic limit of Ginzburg-Landau vortices, *Disc. Cont. Dyn. Systems* **6** (2000), 121–142.

[LT1] H. Liu, E. Tadmor. Spectral dynamics of the velocity gradient field in restricted fluid flows. Comm. Math. Phys. **228** (2002), No. 3, 435–466.

[LT2] H. Liu, E. Tadmor, Critical thresholds in 2D restricted Euler-Poisson equations, *SIAM J. Appl. Math.* **63** (2003), No. 6, 1889–1910.

[LLY] L. Lo, J-G. Liu, P. Yu, On mean field limit for Brownian particles with Coulomb interaction in 3D, *J. Math. Phys.* **60** (2019), 111501.

[MP] C. Marchioro, M. Pulvirenti, *Mathematical theory of incompressible nonviscous fluids.* Applied Mathematical Sciences, 96. Springer-Verlag, New York, 1994.

[MZ] N. Masmoudi, P. Zhang, Global solutions to vortex density equations arising from sup-conductivity, *Ann. I.H.P. Anal. non lin.* **22** (2005), No. 4, 441–458.

[MMn] S. Mei, A. Montanari, P.M. Nguyen, A mean field view of the landscape of two-layer neural networks. *Proc. Nat. Acad. Sci*, **115** (2018), No. 33, E7665–E7671.

[NW] H. Neunzert, J. Wick, Die Approximation der Lösung von Integro-Differentialgleichungen durch endliche Punktmengen, *Numerische Behandlung nichtlinearer Integrodifferential- und Differentialgleichungen,* Springer, 1974, 275–290.

[NT] T. Nguyen, A. Tudorascu, Pressureless Euler/Euler-Poisson Systems via Adhesion Dynamics and Scalar Conservation Laws. *SIAM J. Math. Anal.* **40** (2008) No. 2, 754–775.

[Pe] B. Perthame, Nonexistence of global solutions to Euler-Poisson equations for repulsive forces, *Japan J. Appl. Math.* **7**, No. 2 (1990), 363–367.

[PS] M. Petrache, S. Serfaty, Next Order Asymptotics and Renormalized Energy for Riesz Interactions, *J. Inst. Math. Jussieu* **16** (2017), No. 3, 501–569.

[RVE] G. Rostkoff, E. Vanden Eijnden, Trainability and Accuracy of Neural Networks: An Interacting Particle System Approach, arXiv:1805.00915.

[RS] N. Rougerie, S. Serfaty, Higher Dimensional Coulomb Gases and Renormalized Energy Functionals, *Comm. Pure Appl. Math* **69** (2016), 519–605.

[SR] L. Saint Raymond, *Hydrodynamic limits of the Boltzmann equation.* Lecture Notes in Mathematics 1971, Springer, (2009).

[Scho1] S. Schochet, The weak vorticity formulation of the 2-D Euler equations and concentration-cancellation, *Comm. PDE* **20** (1995), No. 5-6, 1077–1104.

[Scho2] S. Schochet, The point-vortex method for periodic weak solutions of the 2-D Euler equations. *Comm. Pure Appl. Math.* **49** (1996), No. 9, 911–965.

[Si] P. Serfati, Une preuve directe d'existence globale des vortex patches 2D, *C. R. Acad. Sci. Paris Ser. I Math.* **318** (1994), No. 6, 515–518.

[Sy] S. Serfaty, Mean Field Limits for the Gross-Pitaevskii and Parabolic Ginzburg-Landau Equations, *J. Amer. Math. Soc.* **30** (2017), No. 3, 713–768.

[SV] S. Serfaty, J. L. Vázquez, A Mean Field Equation as Limit of Nonlinear Diffusions with Fractional Laplacian Operators, *Calc Var. PDE* **49** (2014), No. 3-4, 1091–1120.

[Wo] W. Wolibner, Un théorème sur l'existence du mouvement plan d'un fluide parfait, homogène, incompressible, pendant un temps infiniment long. *Math. Z.* **37** (1933), No. 1, 698–726.

[XZ] W. Xiao, X. Zhou, Well-posedness of a porous medium flow with fractional pressure in Sobolev spaces, *Electron. J. Differential Equations* **2017**, Paper No. 238, 7 pp.

[Yu] V.I Yudovich, Non-stationary flow of an ideal incompressible liquid. *Zh. Vych. Mat.* **3** (1963), 1032–1066.

MITIA DUERINCKX

Université Paris-Saclay, CNRS, Laboratoire de Mathématiques d'Orsay, 91405 Orsay, France & Université Libre de Bruxelles, Département de Mathématique, 1050 Brussels, Belgium.

Email: mduerinc@ulb.ac.be

SYLVIA SERFATY

Courant Institute, New York University, 251 Mercer street, New York, NY 10012, USA.

Email: serfaty@cims.nyu.edu