Next-order asymptotic expansion for N-marginal optimal transport with Coulomb and Riesz costs

Codina Cotar[*] and Mircea Petrache [†]

November 6, 2018

Abstract: Motivated by a problem arising from Density Functional Theory, we provide the sharp next-order asymptotics for a class of multimarginal optimal transport problems with cost given by singular, long-range pairwise interaction potentials. More precisely, we consider an N-marginal optimal transport problem with N equal marginals supported on $\mathbb{R}^d$ and with cost of the form $\sum_{i \neq j} |x_i - x_j|^{-s}$. In this setting we determine the second-order term in the $N \to \infty$ asymptotic expansion of the minimum energy, for the long-range interactions corresponding to all exponents $0 < s < d$. We also prove a small oscillations property for this second-order energy term. Our results can be extended to a larger class of models than power-law-type radial costs, such as non-rotationally-invariant costs. The key ingredient and main novelty in our proofs is a robust extension and simplification of the Fefferman-Gregg decomposition [F85], [Ge89], extended here to our class of kernels, and which provides a unified method valid across our full range of exponents. Our first result generalizes a recent work of Lewin, Lieb and Seiringer [LLS17], who dealt with the second-order term for the Coulomb case $s = 1, d = 3$.

Keywords: Density functional theory (DFT), Hohenberg-Kohn functional, N-body optimal transport, Coulomb and Riesz costs, exchange-correlation functional, finite exchangeability, N-representability, positive definite kernels, Fefferman-Gregg decomposition, Lieb-Oxford bound, Uniform Electron Gas, Jellium

1   Introduction

1.1   The next-order asymptotics

Let $\mathsf{c}(x, y) = g(x - y) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a "pairwise interaction" cost function, and consider the space $\mathcal{P}_{sym}^N(\mathbb{R}^d)$ of probability measures on $(\mathbb{R}^d)^N$ which are invariant under the action induced by the permutation group S_N acting on the N coordinates $(x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$. We consider the following N-marginals optimal transport (OT) minimization problem, for given $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $N \geq 2$:

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) := \inf \left\{ \int_{(\mathbb{R}^d)^N} \sum_{1 \leq i \neq j \leq N} \mathsf{c}(x_i, x_j) d\gamma_N(x_1, \ldots, x_N) : \ \gamma_N \in \mathcal{P}_{sym}^N(\mathbb{R}^d), \gamma_N \mapsto \mu \right\} . \qquad (1.1)$$

The notation $\gamma_N \mapsto \mu$ means that γ_N has one-body density μ (physics terminology) or equivalently equal $\mathbb{R}^d$-marginals μ (probability terminology),

$$\gamma_N(\underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{\text{i-1 times}} \times A_i \times \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{\text{N-i times}}) = \int_{A_i} \mu(x) \, dx \text{ for all Borel } A_i \subseteq \mathbb{R}^d \text{ and all } i = 1, \ldots, N . \qquad (1.2)$$

Even though our techniques work for more general costs, we will consider especially the case of

$$\mathsf{c}(x, y) = |x - y|^{-s} \quad \text{if} \quad 0 < s < d, \ \ x, y \in \mathbb{R}^d , \qquad (1.3)$$

*University College London, Statistical Science Department, London, United Kingdom, `c.cotar@ucl.ac.uk`
†Pontificia Universidad Catolica de Chile, Santiago, Chile, `decostruttivismo@gmail.com`

1

and we will prove the following asymptotic expansion for $F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu)$. For all $0 < s < d$, if μ has density $\rho \in L^{1+s/d}(\mathbb{R}^d)$, then we have (see Theorem 1.1 below)

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) = N^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{\rho(x)\rho(y)}{|x-y|^s} dx\, dy - C(s,d) N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x) dx + o(N^{1+s/d}) \qquad \text{as} \quad N \to \infty. \quad (1.4)$$

Moreover we will prove that the *strictly positive* constant $C(s,d)$ is independent of the choice of the marginal $d\mu(x) = \rho(x)dx$ and therefore can be interpreted as arising in an independent model problem. Furthermore, we will derive a small oscillations principle (with respect to N) for

$$N^{-1-s/d}\left(F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) - N^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{\rho(x)\rho(y)}{|x-y|^s} dx\, dy \right),$$

which could also be interpreted as a rough third order asymptotic bound for $F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu)$ with c as in (1.3).

Our methods are extendable to more general costs, as non-rotationally-invariant ones. See Remark 1.3 below.

For $d = 3, s = 1$, the $F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu)$ was introduced in the physics literature in the context of Density Functional Theory (DFT) by Seidl, Perdew, Levy, Gori-Giorgi, and Savin [Sd99, SPL99, SGGS07], without them being aware of its meaning in optimal transport. Namely, for $s = 1, d = 3$, (1.1) is a natural semiclassical limit to the famous *Hohenberg-Kohn (HK)* functional from quantum mechanics, originally introduced by Hohenberg-Kohn in [HK64], and rigorously proved by Levy and Lieb in [Ly79], [Li83]. The connection to optimal transport was mathematically established later by [CFK11] and [BdPGG12] for $N = 2$, later further extended to $N = 3$ in [BdP17], and recently independently proved for all $N \geq 2$ by [CFK17] and [Lw17].

In the process of establishing our two main results, we were required to prove, as key new tools for them, a set of additional secondary results of independent interest and of possible use to other settings, leading to generalized versions of the Fefferman-Gregg decomposition and positive definiteness criteria, as described in Section 4.1. What we use is a decomposition of positive definite kernels following the strategy established for $s = 1, d = 3$ by Fefferman [F85], and extended by Gregg [Ge89], in $d = 3$, with methods which work there for $0 < s < 3$ (but which in general dimension work only for $0 < s < 2 + [(d-1)/2]$, where $[\cdot]$ denotes the integer part). The decomposition is based on a separation of the kernel into contributions from different ranges, coupled with a good packing strategy for the domain. We extend the range of validity of the Fefferman-Gregg decomposition, simplify the proofs, and apply the strategy to our new problem.

We note here that, together with our main result (1.4), the extended Fefferman-Gregg decomposition also allows to establish the equality of the next-order term for the two minimization problems below:

1. The energy-minimization for Coulomb and Riesz gases

$$\mathcal{E}_{N,\mathsf{c}}(V) := \inf\left\{ \sum_{1 \leq i \neq j \leq N} \mathsf{c}(x_i - x_j) + N \sum_{i=1}^{N} V(x_i) \text{ for } x_1, \ldots, x_N \in \mathbb{R}^d : x_1, \ldots, x_N \in \mathbb{R}^d \right\}, \quad (1.5)$$

   for c as in (1.3), and where $V : \mathbb{R}^d \to \mathbb{R}$ is a suitable external "confining" potential assumed to be bounded below, lower semicontinuous, such that $\{x : V(x) < \infty\}$ has non-zero c-capacity, and such that $V(x) \to \infty$ as $x \to \infty$. The asymptotic expansion of $\mathcal{E}_{N,\mathsf{c}}(V)$ can be written as

$$\mathcal{E}_{N,\mathsf{c}}(V) = N^2 \mathcal{I}_{\mathsf{c},V}(\mu_V) - C_{\mathrm{Jel}}(s,d) N^{1+s/d} \int \rho^{1+s/d}(x) dx + o(N^{1+\frac{s}{d}}) \qquad \text{as} \quad N \to \infty,$$

   where μ_V is the unique minimizer of

$$\mathcal{I}_{\mathsf{c},V}(\mu) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathsf{c}(x-y) d\mu(x) d\mu(y) + \int_{\mathbb{R}^d} V(x) d\mu(x),$$

   and where the constant $C_{\mathrm{Jel}}(s,d) > 0$ corresponds to the minimum energy of a unit density Jellium-type problem, as described for example in [PS].

2. The minimum of the optimal transport problem considered in (1.4). Here, if we write $C_{\mathrm{UEG}}(s,d) := C(s,d)$, the next-order term can then be written as

$$C_{\mathrm{UEG}}(s,d)N^{1+s/d}\int \rho^{1+s/d}(x)dx,$$

where $C_{\mathrm{UEG}}(s,d)$ can be interpreted as the energy of a Uniform Electron Gas, as described in [CP17].

As a remark about terminology, note that the model of classical electrons corresponds to the Coulomb-type interaction potentials with $s=1, d \leq 3$, while other more general "Uniform Gas" problems appear for pairwise interactions mediated by general kernels $\mathsf{c}(x-y)$ in general dimensions.

This equality of the above constants, $C_{\mathrm{Jel}}(s,d) = C_{\mathrm{UEG}}(s,d)$, is proved for $d-2 < s < d$ in [CP17].

The Fefferman-Gregg decomposition also allows to find the precise next-order term for the Jellium problem for $d-2 \leq s < d$, giving an alternative proof for the main result in [PS], as we will prove in future work.

1.2 Comments about the decomposition techniques used in this paper

1.2.1 Relevance of the decomposition in the proof

The importance of positive-definiteness in our problem is crucial, and was emphasized in [P15]: the property of our kernels of being balanced-positive-definite is *equivalent* to the convergence of the renormalized minimum energies $N^{-2}F_N^{\mathrm{OT}}[\rho]$ to the mean field $\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathsf{c}(x,y)d\mu(x)d\mu(y)$.

This ingredient was not present in the problem considered by [F85] and [Ge89], which was of different nature, and is thus specific to our optimal-transport problem.

In the decomposition of our power-law kernels, we can only hope to have next-order bounds for the $E_N^{\mathrm{xc}}(\rho)$-energies of the decomposed kernels when the first-order contribution is cancelled by subtracting the mean field. By the above result [P15] we thus have the binding constraint that *all the kernels in our decomposition must be positive definite*. This explains why we put much emphasis in the positive-definite kernel decomposition below.

1.2.2 Link to previous work on the Fefferman-Gregg technique

As described in the abstract of [Ge89], that paper treats inverse-power kernels in dimension $d=3$, with s in a range around 1. This is based on a decomposition of the kernels into positive-definite finite-range parts, together with a very tame error, as done for the case $s=1, d=3$ by [F85], based on the "Swiss cheese" packings present in [LeLi72] and [H89]. The key observation in [Ge89] is that the kernels with $s \neq d-2$ do not satisfy a version of Newton's lemma. This forbids the spherical averaging methods from [LeLi72] to apply to $s \neq d-2$.

The construction in [Ge89] is formulated in $d=3$ for $0 < s < 2$, due to other constraints coming from the kinetic part of the energy considered there, but the kernel decomposition methods hold for $0 < s < 3$, and in general dimension they allow to treat exponents $0 < s < 2+[(d-1)/2]$, not covering the whole range $0 < s < d$. We thus optimized here several steps of that strategy, to that aim, as described below. Note that the range $0 < s < 2$ appearing in [Ge89] is due to $-\Delta_{\mathbb{R}^{Nd}} + \sum_{1 \leq i \neq j \leq N}|x_i - x_j|^{-s}$ being a self-adjoint operator only for $0 < s < 2$. This is also true in general dimension d, as follows from Reed-Simon Vol. II, p. 169 and Thm. X.19. (The last result is stated in $d=3$ but extends/can be adapted to general d, but the range of self-adjointness of the Hamiltonian, $0 < s < 2$, is the one that stays valid for general $d \geq 1$.)

A related property which seems to never have emphasized before, concerning the methods of [F85] and [Ge89], is that they work also for *non-rotationally-invariant kernels*. The lack of rotation invariance of the *background field* instead was one of the motivations for [F85, Ge89], as compared to the previous work [LeLi72] where radial symmetrization was a crucial ingredient in the proofs.

The fact that the decomposition is available for non-rotationally-invariant interaction kernels, is a property that seems not to have been exploited in applications yet, and forms an important point of advantage of this type of decomposition, as compared to other methods such as [GS95] used in [LLS17].

The functions called Q_s^i in [Ge89] do not provide bounds on d partial derivatives, in general dimension d, and derivative bounds for self-convolutions of characteristic functions of balls $1_B * 1_B$ can be improved by only one extra derivative by their method, in order to use the positive-definiteness criteria of Lemma A.1. This would only allow to use the lemma for $0 < s < 2 + [(d-1)/2]$, and not for the whole range $0 < s < d$. Therefore we introduce a better mollification to produce the functions which we denote by $Q_{i,\eta}$ here. The functions $Q_{i,\eta}$ then have the derivative bounds required in Lemma A.1 in general dimension d.

In [Ge89] an averaging procedure on ball packings used in the decomposition (needed there to pass from bounds on 2 derivatives of the kernels, to bounds on 3 derivatives) is done by changing the radii of balls from a fixed packing. In [Ge89] a new packing is constructed by the Swiss cheese lemma for each dilation factor, and the volume-fractions covered by the balls from each family only have the rough bounds coming from that lemma. This produces some extra error terms in the final superposition of the kernels, error terms whose control is then not discussed in [Ge89], and is made harder by the fact that explicit decomposition formulas such as (4.19) below are not available in that case.

To solve these problems, we use a different procedure: we dilate and contract the *whole family* of balls given in the Swiss cheese lemma, thereby producing covers which can be fit to our mollification procedure directly, by the algebraic dependence of the $Q_{i,\eta}$ on the dilation parameter of our families. In particular, the dilated families automatically cover fixed volume-fractions, and we avoid some of the error terms from [Ge89].

1.3 Link to previous related works

The result in (1.4) has been very recently proved for the specific case of $s = 1, d = 3$, in [LLS17]. Note that the method used therein, by means of the Graf-Schenker decomposition, does not extend past that specific $s = 1, d = 3$ situation to our general class of costs, and it is mentioned in that paper that it would be interesting to consider the general situation for exponents and dimensions $0 < s < d$, treated here. As it turns out, two points are different between our methods:

- We replace the Graf-Schenker decomposition used in [LLS17] by the generalized Fefferman-Gregg-type approach based on positive definiteness, approach which can be seen as a natural continuation of the new understanding from [CFP15], [P15] of the crucial role of positive definiteness in asymptotics problems such as the one considered here.

- A general difference between this paper and [LLS17] is that we rely on and are inspired by Optimal Transport tools, such as duality and the method [BCdP16] for giving some separation condition on the points in the optimizer. We hope that the reader will profit from comparing our approach and the Statistical Physics framework of [LLS17], both of which highlight different aspects of the theory.

Moreover, we rely in our proofs on probability ideas such as convergence approximations.

We note that our tools are also applicable, by similar arguments, to other models than power-law-type radial costs, such as *anisotropic costs* and costs with *radial oscillations* (see Remark 1.3, Examples (b), (c)). We also can transfer the study done here to the case of natural interaction kernels in *curved spaces*, e.g. on compact Riemannian manifolds (see Remark 1.3, Example (a)).

We remark that in $d = 1$ the next order term can be directly described by a very elegant computation for a very large class of costs (as was explained to us by Simone Di Marino [dM]), by using the explicit "monotone rearrangement" description of the optimal transport plan from [CdPdM13]. Further results and a detailed review of optimal transport results with Coulomb costs and other repulsive costs can be found in [dMGN15].

1.4 Range of validity and future work

Screening and effective energy localization. Concerning the use of statistical mechanics methods, the form of our problem (1.3) creates a certain number of complications. As a main comparison, recall that in the study of the next order for *classical* Coulomb/Riesz gas energies (so far understood only for $s \in [d-2, d)$: see [PS] and the references therein), the source of localization of energy needed for obtaining the next-order term

was directly provided via the boundary value problems connected to the elliptic PDE's coming from the kernels under consideration. However the methods from [PS] are at the moment not extendable outside $d - 2 \leq s < d$, and the next-order Jellium-type energy considered in [PS] may be infinite for $0 < s < d - 2$.

In order to achieve the full range of exponents $0 < s < d$, we apply the more robust strategy, initiated by Fefferman in [F85], and which consists, roughly speaking, in *truncating the kernels instead of truncating the solutions* of the equations, and this is done precisely in such a way as to preserve positive definiteness, at least up to an asymptotically negligible error. The careful positivity-preserving kernel truncation method used here can be then interpreted as a rough and robust counterpart of the above PDE methods, which is available for operators which are more general, namely *positive definite*, rather than elliptic. In this precise sense, we could say that the Fefferman truncation method which we extend and apply here mainly gives some way of an answer to the basic question: *What is the analogue of a Neumann boundary value problem, if we replace the Laplacian with a general positive definite operator?* This question seems to have appeared and to have been answered already before, in a quite different form, in the theory of finite range decompositions: see [BT06] and the references therein.

Comparison to the Coulomb/Riesz gas asymptotics. As noted in the appendix [LwLi15], there seems to be a discrepancy in $s = 1, d = 3$, between the computations for the value of the Jellium energy and that of the Uniform Electron Gas (UEG) energy, on the specific example of lattice-like configurations.

Extending this consideration to the *minimizers* of the corresponding UEG and Jellium energies would lead to the interesting consequence that there would be, for $s = d - 2$, a *gap between the constant* $\mathsf{C}_{\mathrm{Jel}}(d - 2, d)$ appearing in the next-order term expansions for the large-N expansions of the Coulomb gas minimum energy (described in [SS15], [PS], as a minimum jellium energy) and the constant $C(d - 2, d)$ whose existence is found in our main result (1.4) (which corresponds to an UEG energy minimization problem, as appearing e.g. in Proposition 2.5 for the case $\Lambda = [0, 1]^d$).

For the Coulomb case $s = d - 2$, the precise estimates needed for proving the presence of a gap (beyond computations for special examples such as lattices) seems to require a precise understanding of the Jellium and UEG minimizers, which goes beyond the state-of-the-art results currently available. However for the exponents $s \in (d - 2, d)$, for which both the values $\mathsf{C}_{\mathrm{Jel}}(s, d)$ and $C(s, d)$ are computable due to our result (1.4) and to [PS], these values agree as proved recently in [CP17].

Towards a general localization theory for positive definite operators. The first paper where positive-definite truncations for the Coulomb interaction potential were constructed/used seems to be [F85], then refined and extended to more general $d = 3$ interactions by [Ge89] and [H89]. A different, simpler construction by means of the Yukawa potential was introduced by Conlon, Lieb and Yau [CLY89]. Later on a much simpler construction, which is specific to the case $d = 3$ and to the Coulomb cost appeared in [GS95].

Well-behaved truncations of operators of positive type are a key tool used in the recent renormalization group results such as [BGM04], [BBS15], [AKM13]. In these cases the setting is often $\mathbb{Z}^d$ rather than $\mathbb{R}^d$, and the kernels that have to be localized are often the Coulomb kernels, interpreted as infinite-dimensional positive definite matrices. General decompositions have been studied for example in [BT06] and [B13].

These results so far profit of explicit representations of the operators involved, however one should be able to extend the truncation theory such as presented here to both discrete settings and more general positive definite operators. We plan to pursue this direction in future work.

1.5   Main results

In our statements and proofs below, we define a generalized version $E_{N,\mathsf{c}}^{\mathrm{xc}}(\mu)$ of the so-called "exchange-correlation" energy for a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ and a cost function $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$ such that the integrals below are finite

$$E_{N,\mathsf{c}}^{\mathrm{xc}}(\mu) := F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) - N^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathsf{c}(x, y) d\mu(x) d\mu(y) . \tag{1.6}$$

The above definition, which is the starting point for our paper, could be stated for general finite positive measures μ, under the condition that we apply the right normalization for the second term. Therefore the requirement

5

$\mu \in \mathcal{P}(\mathbb{R}^d)$ constitutes our choice of normalization. For our costs of particular interest as in (1.3), we will also use for $0 < s < d$ the notation $F^{\mathrm{OT}}_{N,s}(\mu)$, respectively $E^{\mathrm{xc}}_{N,s}(\mu)$.

We can extend the definition of $F^{\mathrm{OT}}_{N,\mathsf{c}}(\mu)$ also to $N \in \{0,1\}$, by taking $F^{\mathrm{OT}}_{N,\mathsf{c}}(\mu) := 0$ in these two cases. Moreover, for all $N \in \mathbb{R}_{>0}, N \geq 2$, let us define the "grand-canonical optimal transport"

$$F^{\mathrm{OT}}_{\mathrm{GC},N,\mathsf{c}}(\mu) := \inf \left\{ \sum_{n=2}^{\infty} \alpha_n F^{\mathrm{OT}}_{n,\mathsf{c}}(\mu_n) \;\middle|\; \begin{array}{l} \sum_{n=0}^{\infty} \alpha_n = 1, \ \sum_{n=1}^{\infty} n\alpha_n \mu_n = N\mu, \\[2mm] \mu_n \in \mathcal{P}(\mathbb{R}^d), \ \alpha_n \geq 0, \quad \text{for all} \quad n \in \mathbb{N} \end{array} \right\}, \tag{1.7a}$$

and the "grand-canonical exchange correlation energy"

$$E^{\mathrm{xc}}_{\mathrm{GC},N,\mathsf{c}}(\mu) := F^{\mathrm{OT}}_{\mathrm{GC},N,\mathsf{c}}(\mu) - N^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathsf{c}(x,y) d\mu(x) d\mu(y). \tag{1.7b}$$

Here again, the classical definition of (1.7) is usually given only for $N \geq 2$, though one can define the quantities for all $N > 0$. The reason for introducing α_0 and α_1 in (1.7a), even though they do not appear in the sums of optimal transport problems there, is that they naturally appear in the proof of Lemma 4.8 below (see Step 4 therein). Furthermore, our convention for $F^{\mathrm{OT}}_{N,\mathsf{c}}$ for $N \in \{0,1\}$ implies that for $N \in (0,1]$ $F_{\mathrm{GC},N}(\mu) = 0$, as we can take a competitor of the form $\alpha_n = 0, n \geq 2, \alpha_0 = 1 - N, \alpha_1 = N, \mu_n = \mu, n \geq 1$. Also, the same convention implies that in (1.7a) we may equivalently minimize over sums $\sum_{n=0}^{\infty} \alpha_n F^{\mathrm{OT}}_{n,\mathsf{c}}(\mu_n)$.

Theorem 1.1. *Fix $d \geq 1$ and $0 < s < d$ and let c be the corresponding cost as in (1.3). Assume that $\mu \in \mathcal{P}(\mathbb{R}^d)$ has density $\rho \in L^{1+s/d}(\mathbb{R}^d)$. Then there exists $C(s,d) > 0$, which depends only on s and d, such that we have (where both limits exist for the full sequence)*

$$\lim_{\substack{N \to \infty \\ \mathbb{N} \in \mathbb{N}_+}} N^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{N,s}(\mu) = \lim_{\substack{N \to \infty \\ N \in \mathbb{R}_+}} N^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu) = -C(s,d) \int_{\mathbb{R}^d} \rho^{1+\frac{s}{d}}(x) dx . \tag{1.8}$$

Remark 1.2. *After a first draft of the present paper was completed, the very important case $s = 1, d = 3$, of Theorem 1.1, appeared in [LLS17, Thm. 4.1] (for the case of continuous slowly-varying densities in $L^{4/3}(\mathbb{R}^3)$). In [LLS17], a positive function $\rho \in L^1(\mathbb{R}^d)$ is called slowly-varying, if the oscillations at infinity are controlled in the sense that there holds $\sum_{k \in \mathbb{Z}^d} \max_{x \in [0,1)^d + k} \rho(x) < \infty$. The proof of the upper bound of Theorem 1.1 is based on classical tools such as subadditivity, and is similar to [LLS17], to which we often refer the reader for the proofs. The more difficult sharp lower bound from Section 4 requires different and more robust kernel decomposition techniques compared to [LLS17], making our methods extendable to non-isotropic and oscillating kernels.*

Remark 1.3 (extensions to more general costs)**.** *We note the following possible extensions of Theorem 1.1:*

(a) *If (M, g) is a compact Riemannian manifold, then we may consider the optimal transport problem for the interaction cost $G(x, y)$ given by the Green function of the Laplacian of (M, g). This cost is positive definite and has the same homogeneity near zero as the Coulomb potential $|x - y|^{2-d}$ where $d = \dim(M)$. This type of minimization has been studied in [BCC17]. For the case of embedded submanifolds $M \subset \mathbb{R}^D$ this fits in the same framework as the celebrated Smale's 7th problem [Sm98]. For applying our methods to this example it is not essential that the operator under consideration be of second order, and it is possible to consider also interactions given by the Green functions of other other fractional or higher order positive definite operators.*

(b) *The costs $\mathsf{c}(x, y)$ to which our method applies need not be rotation-invariant, as the decomposition method described here uses only translation-invariance unlike the Graf-Schenker approach [GS95]. This allows to extend our results to the costs of the form*

$$\mathsf{c}(x,y) := g(x - y), \ \text{with} \ g(x) = |x|^{-s} f(x/|x|), \ \text{where} \ f \in C^0(\mathbb{S}^{d-1}) .$$

Non-rotation-invariant costs such as the above are not treatable by any previous methods, to our knowledge.

(c) *Other kernels treatable by our methods are those of the form*

$$\mathsf{c}(x,y) := g(x - y), \ \text{with} \ g(x) = l(|x|) f(x/|x|), \ \text{with} \ l(r) = r^{-s-3}(\sin(r) - r\cos(r)) \ \text{and} \ f \in C^0(\mathbb{S}^{d-1}) ,$$

again for $0 < s < d$, which are including radial oscillations too. Kernels with radial oscillations may occur in physics e.g. in the study of Friedel oscillations.

(d) *With considerable more effort regarding the Fefferman-Gregg adaptation than the relatively straightforward cases (a)-(c) above, the proof could potentially also be adapted to other cost cases such as*

$$\mathsf{c}(x,y) := \prod_{i=1}^{d} |x_i - y_i|^{-\alpha}, \quad for \ x = (x_1, \ldots, x_d), y = (y_1, \ldots, y_d) \in \mathbb{R}^d, \ and \ 0 < \alpha < 1 \ ,$$

and to a corresponding version including radial oscillations as well.

In the more general cases (a)-(c), we note that we have $\lim_{\epsilon \downarrow 0} \frac{\mathsf{c}(y + \epsilon x, y)}{g_0(\epsilon x)} = 1$ for all y (and this should be transferred to a local chart at y, in case (a)), locally uniformly for $x \neq 0$, where the function $g_0 : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ is homogeneous of degree $-s$ for some $0 < s < d$, i.e. for all $x \neq 0$ and all $\lambda > 0$ there holds $g_0(\lambda x) = \lambda^{-s} g_0(x)$. In that setting, we expect the next-order term limit treated here to be

$$\lim_{N \to \infty} N^{-1-s/d} E_{N,\mathsf{c}}^{\mathrm{xc}}(\mu) = -C(g_0, s) \int_{\mathbb{R}^d} (\rho(x))^{1+\frac{s}{d}} dx, \quad where \quad C(g_0, s) > 0 \ . \tag{1.9}$$

Moreover, $E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)$ satisfies the small oscillations property detailed in Theorem 1.4 below, which is not only crucial for settling the continuity of $s \to C(s,d)$ in [CP17], but it is also interesting in its own right. As a counterpart for the Jellium problem, oscillation bounds for minimizers of the Coulomb gas energy appeared in [RoSer] for the log-interaction in $d = 2$ and were extended in [PRN] to Coulomb interactions $s = d - 2$ in general dimension $d \geq 2$, and to $d - 2 < s < d$ under an extra hypothesis. Note that in the setting of [RoSer, PRN], spatial oscillations and uniformity of the asymptotic spatial distribution of optimum configurations were controlled, whereas here we control oscillations of the energy values.

Theorem 1.4 (Small oscillations property of $E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)$). *Fix $0 < \epsilon < d/2$ and let $\epsilon \leq s \leq d - \epsilon$. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a probability measure with density of the form $\rho(x) = \sum_{i=1}^{k} \alpha_i 1_{\Lambda_i}(x)$, where $\Lambda_1, \ldots, \Lambda_k$ are hyperrectangles with disjoint interiors.*

Then there exists $C(\rho, d, \epsilon) > 0$ such that for all $N, \widetilde{N} \in \mathbb{R}_{>0}$, $N \geq \widetilde{N} \geq 2$, we have

$$\left| \frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} - \frac{E_{\mathrm{GC},\widetilde{N},s}^{\mathrm{xc}}(\mu)}{\widetilde{N}^{1+s/d}} \right| \leq \frac{C(\rho, d, \epsilon)}{\log \tilde{N}} \ , \tag{1.10}$$

where $C(\rho, d, \epsilon)$ depends on ρ, d, ϵ but is independent of the choice of the parameter $s \in [\epsilon, d - \epsilon]$.

Remark 1.5. *Though for simplicity of calculations we only prove Theorem 1.4 for the setting described therein, we may extend the above small-oscillations bound to general Borel sets Λ_i with ϕ-regular boundary (see Definition 2.8 below), and also to more general densities, such as any positive Riemann-integrable function, i.e. any function that can be approximated well by piecewise constant functions on hyper-rectangles. This includes in particular continuous densities with compact support and ϕ-regular boundary.*

However, with the present strategy, this may come at the cost of a less tame oscillation bound than the one of order $1/(\log \widetilde{N})$ of (1.10) above. Note that the order can be improved for $s = 1, d = 3$, to an s-dependent order $1/N^{\alpha}, \alpha > 0$, by applying the Graf-Shenker decomposition [GS95] instead of the Fefferman-Gregg one.

The above small oscillation result connects to the conjecture about the *optimal Lieb-Oxford bounds*, initially formulated for the case $s = 1, d = 3$, but actually open in the whole range $0 \leq s < d$, $d \geq 2$ (where for $s = 0$ we consider the kernel $\mathsf{c}(x,y) = -\log|x - y|$ instead), regarding the precise value of the optimal constant $C_{\mathrm{LO}}(s,d)$ such that for all $N \geq 2$ and for all measures μ with density $\rho \in L^{1+s/d}(\mathbb{R}^d)$ there holds

$$\frac{E_{N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} \geq -C_{\mathrm{LO}}(s,d) \int_{\mathbb{R}^d} \rho^{1+s/d}(x) dx. \tag{1.11}$$

As mentioned in [LLS17], the fact that the optimal constant is realized in the limit $N \to \infty$ would fit with numerical results.

1.5.1 Fefferman-Gregg decomposition

We next state the generalized Fefferman-Gregg decomposition, originally introduced by Fefferman [F85] for $s = 1, d = 3$, extended by Gregg [Ge89] to $0 < s < 2 + [(d-1)/2]$, and further extended by us to $0 < s < d$. This is the main new tool for the proofs of Theorem 1.1 and of Theorem 1.4, as explained in the next section.

Proposition 1.6 (Fefferman-Gregg decomposition). *Let $M \in \mathbb{N}_+$, $0 < \epsilon < d/2$ and $\epsilon \leq s \leq d - \epsilon$. Then there exists a constant C depending only on d, ϵ, a family Ω of ball packings F_ω of $\mathbb{R}^d, \omega \in \Omega$, a radius $R_1 > 0$ and a probability measure $\mathbb{P}$ on Ω such that the cost $|x_1 - x_2|^{-s}$ can be decomposed as follows:*

$$\frac{1}{|x_1 - x_2|^s} = \frac{M}{M+C} \left\{ \int_\Omega \left(\sum_{A \in F_\omega} \frac{1_A(x_1)1_A(x_2)}{|x_1 - x_2|^s} \right) d\mathbb{P}(\omega) + w(x_1 - x_2) \right\} , \qquad (1.12)$$

where

1. *w is positive definite;*

2. *if $\mu \in \mathcal{P}(\mathbb{R}^d)$ has density $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$ then*

$$E_{N,w}^{\mathrm{xc}}(\mu) \geq E_{\mathrm{GC},N,w}^{\mathrm{xc}}(\mu) \geq -\frac{C(w,d,\epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M} R_1^{-s}(N-1)$$

for some $C(w,d,\epsilon) > 0$, which depends only on w, d and ϵ.

1.5.2 Summary of constructions for Proposition 1.6

While we stated here the proposition in a more self-contained way, we note some further parameters which will intervene and allow to fit the decomposition within the rest of the proof:

- The parameters M and $R_1 > 0$ above, as well as the further scale parameter $\ell > 0$ introduced below are going to satisfy the relative constraints of the Swiss cheese Lemma 4.2.

- The packing family Ω and $\mathbb{P}$ will therefore later depend on l and will be denoted Ω_l, respectively $\mathbb{P}_l$. Each family Ω_l will be composed of $(l\mathbb{Z})^d$-periodic packings of balls F_ω^l, obtained from Lemma 4.2.

- We will perform the precise choice of $\Omega_l, \mathbb{P}_l$ in (4.21), and the constant C figuring in Proposition 1.6 will be fixed in (4.23).

Remark 1.7. *As already stated in the introduction, the decomposition (1.12) can be obtained for much more general costs, such as for example, the costs from Remark 1.3. In this more general setting, the decomposition is of the form*

$$\mathsf{c}(x_1, x_2) = \frac{M}{M+C} \left\{ \int_\Omega \left(\sum_{A \in F_\omega} 1_A(x_1)1_A(x_2)\mathsf{c}(x_1, x_2) \right) d\mathbb{P}(\omega) + w(x_1, x_2) \right\} , \qquad (1.13)$$

with w positive definite and (at least for the costs (a)-(c) from Remark 1.3, and using the asymptotic profile $g(x) = |x|^{-s} f(x/|x|)$ as defined there)

$$E_{N,w}^{\mathrm{xc}}(\mu) \geq E_{\mathrm{GC},N,w}^{\mathrm{xc}}(\mu) \geq -\frac{C_{\mathrm{rough}}(f,s,d)}{M} N^{1+\frac{s}{d}} \int_{\mathbb{R}^d} (\rho(x))^{1+\frac{s}{d}} dx - \frac{C_{\mathrm{rough}}(f,s,d)}{M} R_1^{-s}(N-1) , \qquad (1.14)$$

for some $C_{\mathrm{rough}}(f,s,d)$ depending only on f, s, d, and C depends only on s, d. For an explanation of the above, see Remark 4.5.

1.6 Strategy of the proof of Theorem 1.1

The case of uniform marginals

We start by individuating the constant $C(s,d)$ in (1.8) in the main theorem. To determine its value, it suffices to consider the case $\rho(x) = 1_\Lambda/|\Lambda|$ for a specific Borel-measurable set Λ of positive measure. We treat this case in Section 2, and this is the result of Proposition 2.5. The proof is by a classical method based on subadditivity and on the scaling properties of our functionals, which is direct consequence of the $(-s)$-homogeneity of our kernels.

Marginals with piecewise constant density

The first main difficulty is to prove the sharp upper/lower bound from (1.8) for the case of μ with density ρ which is the sum of characteristic functions of a finite union of disjoint hyperrectangles. (Such densities ρ are called "piecewise constant functions" in an analysis terminology and the measures μ are called "mixtures of uniform measures" in a probability terminology; throughout the paper, we use the former terminology.)

While the proof of the upper bound will be an easy consequence of the strategy above, the lower bound is a lot more involved and requires to use Proposition 1.6. The construction and main principle leading to this proposition are as described at an informal level below:

Use of the Fefferman-Gregg decomposition

1. We construct a family of packings, here denoted by $\{F_\omega^l\}_{\omega \in \Omega_l}$, each F_ω^l consisting of balls of small radii $R_1, \ldots, R_M$, forming a geometric series, and then periodized at a slightly larger scale $\sim l$.

   The parameters $M, l, R_1, \ldots, R_M$ are fixed in the Swiss cheese Lemma 4.2.

   For the purposes of our schematic explanation, only the parameter l will be relevant, and the other parameters, as well as the specific choice of F_ω^l, are used in the proof of Proposition 1.6. For the related explanation see Section 4.1.

2. Separately for each cover F_ω^l, we perform the following decomposition of the kernel c:

$$c(x_1, \ldots, x_N) := \sum_{1 \le i \ne j \le N} \mathsf{c}(x_i, x_j) = \sum_{A \in F_\omega^l} \sum_{1 \le i \ne j \le N} \mathsf{c}(x_i, x_j) + err_\omega^l(x_1, \ldots, x_N), \qquad (1.15)$$

   where err_ω^l is an error term, to be carefully estimated. The precise error estimate will be given later.

3. Equation (1.15) translates to an "averaged" inequality, in which the expectation is taken with respect to a suitable probability measure on Ω_l (for the explicit description, see (4.21) below):

$$\int_{\Omega_l} \left(\sum_{A \in F_\omega^l} E_{N\mu(A)}^{\mathrm{xc}} \left(\frac{\mu|_A}{\mu(A)} \right) \right) d\mathbb{P}_l(\omega) \le E_N^{\mathrm{xc}}(\mu) + err, \qquad (1.16)$$

   where in order to present the heuristics better we forget for a moment about the complications coming from the fact that $N\mu(A)$ may not be integer or that $\mu(A)$ may be zero.

4. Since we are considering the case that the density ρ is constant on each of a disjoint union of regular enough sets (hyperrectangles, in this case), it turns out that the contribution of those $A \in F_\omega^l$ on which ρ is not constant, becomes negligible in the limit $l \to 0$. Suppose for simplicity of exposition that for all $A \in F_\omega^l$ we have $\rho|_A$ constant, equal to $|A|^{-1}\mu(A)$. Then for each $\omega \in \Omega_l$ the expression (1.16) approximates a Riemann sum. Indeed for each term in the sum (1.16) we find

$$E_{N\mu(A)}^{\mathrm{xc}} \left(\frac{\mu|_A}{\mu(A)} \right) \simeq -(N\mu(A))^{1+s/d} C(s,d)|A|^{-s/d} = -C(s,d)N^{1+s/d} \int_A (\rho|_A)^{1+s/d} dx, \qquad (1.17)$$

and then, using the fact that our packings are asymptotically precise as $l \to 0$, we find that uniformly in $\omega \in \Omega_l$, the sums appearing (1.16) are asymptotic to the desired integral:

$$\int_{\Omega_l} \left(\sum_{A \in F_\omega^l} E_{N\mu(A)}^{\mathrm{xc}} \left(\frac{\mu|_A}{\mu(A)} \right) \right) d\mathbb{P}_l(\omega) = -(1 + o(1)) N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x) dx, \quad \text{as} \quad l \to 0, N \to \infty.$$

(1.18)

The way to treat the error terms err from (1.16), is to first note that they correspond to the same type of minimization problem as the starting one, but with cost w as in Proposition 1.6, i.e. $err = E_{N,w}^{\mathrm{xc}}(\mu)$.

For the err-term to really give "small error" contribution with the next-order scaling $N^{1+s/d}$, and not for example a leading-order contribution of order N^2, one must use, in a sharp way, the *screening* (or *charge nullity*) behavior of minimizers γ_N, quantifying that γ_N,s F_N^{OT}-type energy is locally *cancelled to leading order* by the corresponding energy coming from the mean field.

The criteria for this cancellation to happen was recently formalized in [CFP15] in the study of the first order term in the asymptotics of F_N^{OT}: the *positive definiteness* of the kernel plays the main role, and in [P15] it was proved that the *necessary and sufficient* condition for the convergence to the mean field is that the kernel should be balanced positive definite. This necessary and sufficient criterion is precisely why it is a posteriori essential to require w to be positive definite in Proposition 1.6, a property without which we would not be able to obtain the desired bounds in the second point of that proposition, which allow us to conclude (1.18) from (1.16).

The appearance of the functional E_{GC}

In the above description of the core of our proof, we did not consider the technical problem given by the fact that while using truncations of our cost to obtain localization, the number of marginals in the so-created "truncated transport plans" is actually not constant. This and the fact that $N\mu(A)$ appearing in (1.16) is not necessarily integer, justifies here, like it did in [LLS17] and in previous works, the introduction of the grand-canonical version $E_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{xc}}$ of $E_{N,\mathsf{c}}^{\mathrm{xc}}$, in which we fix the number of marginals only "in average". This functional is then used as a technical tool and as a replacement of $E_{N,\mathsf{c}}^{\mathrm{xc}}$ throughout the paper.

Approximation of more general densities ρ by piecewise constant ones

The passage from piecewise constant marginals to our class of densities follows then by approximating the respective density by piecewise constant ones.

Small oscillations for E_{GC}

The proof of Theorem 1.4 is based again on the Fefferman-Gregg decomposition, which in this case can be done for two distinct values of $N, \widetilde{N}$ at the same time. Here we look first once more to piecewise constant marginals as for Theorem 1.1, as the arising further approximation constants have a tame dependence on the L^p-norms of the marginals and can be controlled well.

The core of the strategy consists of splitting our domain via Fefferman-Gregg decompositions. Then we ensure that all but an asymptotically negligible proportion of the small balls in the decomposition are completely contained in parts where the marginal is constant, and thus "cut out" uniform marginals. We next compare the uniform-case asymptotics separately to balls coming from the decompositions done at N and at $\widetilde{N}$ marginals, and use the precise asymptotic valid in that case to ensure the oscillation bound.

Note here that the small oscillation bound strategy cannot work for the initial functional E^{xc}, because in order to bound a difference as appearing in (1.10), we require *matching upper and lower bounds* of each of the two quantities, in terms of the same Fefferman-Gregg-sums. The (rigorous version of the) lower bound as in (1.16) figures a sum of $E_{\mathrm{GC}}^{\mathrm{xc}}$-terms, and must be complemented by an (asymptotically) matching upper bound, which

we can only obtain via the classical tool of subadditivity. As a sharp subadditivity bound for E^{xc} only furnishes upper bounds via E^{xc}-terms, and a priori we have only the strict inequality $E^{\mathrm{xc}}_{\mathrm{GC}} < E^{\mathrm{xc}}$, this forces us in order to find a sharp *matching upper bound* to work with $E^{\mathrm{xc}}_{\mathrm{GC}}$.

1.6.1 Plan of the paper

In **Section 2** we establish some preliminary definitions and useful formulas.

The main result in **Section 3** is the optimal sharp upper bound in the next-order asymptotics, for the case of probability measures μ whose density is piecewise constant on a union of hyperrectangles with disjoint interiors.

In **Section 4** we provide the sharp lower bound matching the one from Section 3.1. This is the core result of our paper. In **Section** 4.2.1 we introduce the "grand-canonical version" $E^{\mathrm{xc}}_{\mathrm{GC},N}$ of E^{xc}_N from (1.7b), and we give some preliminary lemmas, to be used in the rest of paper. We present the Fefferman-Gregg decomposition for our kernel in a self-contained way in **Section 4.1**. The proof of the sharp lower bound result of Proposition 4.1 for piecewise constant densities is the object of **Section 4.2**. In **Section** 4.3 we show the proof of Theorem 1.1.

In **Section** 5 we give the proof to our second main result, Theorem 1.4.

In **Appendix A** we provide detailed proofs of the results pertaining to our decomposition of kernels present in Section 4. **Appendix B** contains a brief remark to the sharp function spaces of densities to be expected in our generalized setting, in **Appendix C** we give a statement of a Lieb-Oxford bound which is uniform in s. **Appendix D** presents the proofs of some key properties of $E^{\mathrm{xc}}_{\mathrm{GC},N}$ stated in Section 4.2.1, and Appendix **Section** E gives some helpful optimal transport results.

2 Preliminaries and notation

We will use when convenient the notation $\mathsf{c}(x,y) := g(x-y)$. We assume from now on that $\gamma_N \in \mathcal{P}^N_{sym}(\mathbb{R}^d)$ is a solution to (1.1); such a solution exists for our costs c of interest by standard arguments as given e.g in Villani [V09]. We note that if $\mathsf{c}(x,y) = |x-y|^{-s}$ for μ with density $\rho \in L^1(\mathbb{R}^d)$, we have the following function space sharp requirements that arise naturally in relation with this type of kernel:

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} |x-y|^{-s}\rho(y)dy < \infty \quad \text{for} \quad \rho \in L^{\frac{d}{d-s},1}(\mathbb{R}^d), \tag{2.1a}$$

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^{-s}\rho(x)\rho(y)dx\,dy < \infty \quad \text{for} \quad \rho \in L^{\frac{2d}{2d-s},2}(\mathbb{R}^d), \tag{2.1b}$$

where the spaces $L^{p,q}(\mathbb{R}^d)$ are the Lorentz spaces (see Appendix Section B). Note in particular that for $0 < s < d$ there follows $\frac{d}{d-s} > 1 + \frac{s}{d} > \frac{2d}{2d-s}$, therefore the requirement (2.1a) is stronger than $\rho \in L^{1+s/d}(\mathbb{R}^d)$ and (2.1b) is less restrictive than $\rho \in L^{1+s/d}(\mathbb{R}^d)$, at least for $\rho \in L^1(\mathbb{R}^d)$.

Remark 2.1. *Let $\mathsf{c}(x,y) = |x-y|^{-s}$ and set $0 < s < d$. Then for all $N \geq 2$ and all $\mu \in \mathcal{P}(\mathbb{R}^d)$ with $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$, we have for some $c_{\mathrm{LO}}(s,d) > 0$ which does not depend on N and μ*

$$-c_{\mathrm{LO}}(s,d) \int_{\mathbb{R}^d} \rho^{1+\frac{s}{d}}(x)dx \leq N^{-1-\frac{s}{d}}\left[F^{\mathrm{OT}}_{N,s}(\mu) - N^2 \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(x-y)d\mu(x)d\mu(y) \right] \leq 0\,. \tag{2.2}$$

Due to (2.1), the space $\rho \in L^1(\mathbb{R}^d) \cap L^{\frac{2d}{2d-s},2}(\mathbb{R}^d)$ gives a sharp condition ensuring that the mean field energy for ρ is finite, and if also $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$ then the next-order term is finite too and (2.2) follows by Lemma 16 from [LNP16], which result extends the Lieb-Oxford inequality from $s = 1, d = 3$, to $0 < s < d$.

An optimal transport result that we will use in the paper is the following many-marginals Monge-Kantorovich duality result, proved in [dP15] for the Coulomb cost, and later extended to more general costs in [BCdP16].

Following [BCdP16], we assume that

$$\mathsf{c}(x,y) = g(x-y) = l(|x-y|) \quad \text{where} \quad l : [0,\infty) \to [0,+\infty] \quad \text{is} \quad \begin{cases} \text{continuous on } (0,\infty) \,, \\ \text{strictly decreasing} \,, \\ \text{such that } \lim_{t\to 0^+} l(t) = +\infty \,. \end{cases} \tag{2.3}$$

Proposition 2.2. *Let* c *satisfy conditions* (2.3). *Then for any* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *and any* $N > 1$, *the equality*

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) = \sup_{f \in \mathcal{I}_\mu} \left\{ N \int_{\mathbb{R}^d} f(x) d\mu(x) : \sum_{i=1}^N f(x_i) \leq \sum_{1 \leq i \neq j \leq N} \mathsf{c}(x_i, x_j) \right\}$$

holds, where $\mathcal{I}_\mu$ *denotes the set of* μ-*integrable functions and the pointwise inequality is satisfied everywhere.*

In preparation for the proof of Proposition 2.5 below, we first need the following subadditivity result, proved in Lemma 2.5 from [LLS17]. We observe here that such subadditivity methods have been used previously in similar settings (see, e.g. [HLS09a] or [HLS09b]). Note that, unlike [LLS17], we work in our statements with probability measures.

We recall here that our convention is that $F_{1,\mathsf{c}}^{\mathrm{OT}}(\mu) = 0$ and thus $E_{1,\mathsf{c}}^{\mathrm{xc}}(\mu) = -\int_{\mathbb{R}^d \times \mathbb{R}^d} \mathsf{c}(x,y) d\mu(x) d\mu(y)$.

Proposition 2.3. *Let* $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$. *Consider* k *probability measures* $\mu_1, \ldots, \mu_k$, *with densities respectively equal to* $\rho_1, \ldots, \rho_k$, *such that the quantities below are well-defined and finite (for* $\mathsf{c}(x,y) = |x-y|^{-s}$, *let* $\rho_i \in L^{1+\frac{s}{d}}(\mathbb{R}^d), i = 1, \ldots, k$). *Fix* $M_1, \ldots, M_k \in \mathbb{N}_+$, *and let* μ *be the probability measure with density* $\frac{\sum_{i=1}^k M_i \rho_i}{\sum_{i=1}^k M_i}$. *Then the following subadditivity relation holds:*

$$E_{\sum_{i=1}^k M_i, \mathsf{c}}^{\mathrm{xc}}(\mu) := E_{\sum_{i=1}^k M_i, \mathsf{c}}^{\mathrm{xc}} \left(\frac{\sum_{i=1}^k M_i \mu_i}{\sum_{i=1}^k M_i} \right) \leq \sum_{i=1}^k E_{M_i, \mathsf{c}}^{\mathrm{xc}}(\mu_i). \tag{2.4}$$

Proof. The proof for $k = 2$ can be found in [LLS17]. Given the result for $k = 2$, the general case follows by induction on k: for the inductive step one can apply the $k = 2$ case to the measures $\mu = \mu_k, \mu' = \frac{\sum_{i=1}^{k-1} M_i \mu_i}{\sum_{i=1}^{k-1} M_i}$ and the numbers $M = M_k, M' = \sum_{i=1}^{k-1} M_i$. $\square$

We state in Lemma 2.4 the different scalings of $F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu)$ under change of μ only. The proof follows directly via Proposition 2.2, and will be omitted. See also Corollary 2.6 from [CMC10] for an adaptation to the case of finite measures of the usual two-marginals Monge-Kantorovich duality.

Lemma 2.4. *Let* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *be a probability measure with a density* ρ *such that the quantities defined below are all finite. Let* $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$.

(a) *If we replace* μ *by* $\beta\mu$ *for* $\beta > 0$, *then we get*

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\beta\mu) = \beta F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu), \quad \text{which converges to} \quad \beta \int_{\mathbb{R} \times \mathbb{R}^d} c(x,y) d\mu(x) d\mu(y) \quad \text{as} \quad N \to \infty. \tag{2.5}$$

(b) *Set* $0 < s < d$ *and let* $\mathsf{c}(x,y) = |x-y|^{-s}$. *If we replace for* $\alpha > 0$, μ *by* μ_α *defined as* $d\mu_\alpha(x) = \rho_\alpha(x) dx$ *with* $\rho_\alpha(x) = \rho(\alpha x)$, *then we get for such* c

$$F_{N,s}^{\mathrm{OT}}(\mu_\alpha) = \alpha^{-d+s} F_{N,s}^{\mathrm{OT}}(\mu), \quad \text{which converges to} \quad \alpha^{-d+s} \int_{\mathbb{R} \times \mathbb{R}^d} c(x,y) d\mu(x) d\mu(y) \quad \text{as} \quad N \to \infty. \tag{2.6}$$

Note that the transformation $\mu \mapsto \alpha^d \mu_\alpha$ *maintains the property of* μ *of being a probability measure. Moreover, the above checks that*

$$E_{N,s}^{\mathrm{xc}}(\alpha^d \mu_\alpha) = \alpha^s E_{N,s}^{\mathrm{xc}}(\mu) \,. \tag{2.7}$$

The convergence to the mean field from (2.5) and (2.6) follows from [CFP15] and [P15]. We emphasise that, in view of (2.5), it will be crucial for us to ensure that the correct normalizations are used in the definition of $E_N^{\mathrm{xc}}(\mu)$, a fact that is automatically ensured because we work with probability measures μ rather than with positive measures.

Before we proceed, we need to introduce regularity of sets. As in [F64], we say that a set Λ has a ϕ-*regular boundary* if for some $t_0 > 0$ and for a continuous function $\phi : [0, t_0) \to \mathbb{R}^+$ with $\phi(0) = 0$ there holds

$$\forall 0 \le t \le t_0 \ , \ \left| \left\{ x : \ d(x, \partial \Lambda) \le |\Lambda|^{1/d} t \right\} \right| \le \phi(t) |\Lambda| \ . \tag{2.8}$$

We will mostly work with bounded Borel-measurable sets with ϕ-regular boundary for some ϕ, to make our estimates quantitative. Note that if a set has ϕ-regular boundary then it is Jordan-measurable. We state next the following result, proved via classical subadditivity reasonings in [LLS17, Thm. 2.6]) (see also [RR67, Prop. 2] and [R69, Prop. 7.2.4]), and pointed out to us by M. Lewin in January 2016 in IHP.

Proposition 2.5 (Uniform electron gas). *Fix $0 < s < d$ and let $\mathsf{c}(x, y) = |x - y|^{-s}$. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a uniform measure, with density $\rho(x) := |\Lambda|^{-1} 1_\Lambda(x)$ supported on a Borel set with ϕ-regular boundary $\Lambda \subset \mathbb{R}^d$. Then there exists*

$$-\infty < \lim_{N \to \infty} N^{-1 - \frac{s}{d}} \left[F_{N,s}^{\mathrm{OT}}(\mu) - N^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{1}{|x - y|^s} d\mu(x) d\mu(y) \right] = \lim_{N \to \infty} N^{-1 - \frac{s}{d}} E_{N,s}^{\mathrm{xc}}(\mu) < 0 \ .$$

Moreover, there exists $0 < C(s, d) < \infty$, which is independent of Λ, such that

$$\lim_{N \to \infty} N^{-1 - \frac{s}{d}} E_{N,s}^{\mathrm{xc}}(\mu) = -C(s, d) \int_\Lambda \rho^{1 + \frac{s}{d}}(x) dx = -C(s, d) |\Lambda|^{-s/d} \ . \tag{2.9}$$

Even though Proposition 2.5 is stated for sets with ϕ-regular boundary, in fact for some of our results we only need to consider the more restricted case of hyper-rectangles with disjoint interiors, like in Theorem 3.1 below.

Remark 2.6. *Note that Proposition 2.5 can be adapted, with some more work, and by means of Proposition 2.2, to the costs examples from Remark 1.3, to derive in this case a limiting result as in 1.9.*

3 Optimal constant for piecewise constant ρ

As already explained in the introduction, before we proceed to the proof of our main statement for the general marginals case, we will need to understand first the case of marginals with piecewise constant densities. Even though stated only for $0 < s < d$, the result actually holds in much greater generality of the costs provided that one can show existence of the corresponding limit to the one in Proposition 2.5, and that one can extend the Fefferman-Gregg decomposition to these costs. In particular, Proposition 2.5 can be shown to hold for the costs from Remark 1.3.

Theorem 3.1 (Optimal constant for marginals with piecewise constant density). *Fix $0 < s < d$ and let $\mathsf{c}(x, y) = |x - y|^{-s}$. Set $k \ge 1$. For $i = 1, \ldots, k$, let $\mu_i \in \mathcal{P}(\mathbb{R}^d)$ be the uniform measure on $\Lambda_i \subset \mathbb{R}^d$ with density ρ_i, where the $\Lambda_i, i = 1, \ldots, k$ are Borel sets with ϕ-regular boundaries and disjoint interiors. Then there exists $C(s, d) > 0$, depending only on s and d, such that*

$$\lim_{N \to \infty} N^{-1 - s/d} E_{N,s}^{\mathrm{xc}} \left(\sum_{i=1}^k \alpha_i \mu_i \right) = \lim_{N \to \infty} N^{-1 - s/d} E_{\mathrm{GC}, N, s}^{\mathrm{xc}} \left(\sum_{i=1}^k \alpha_i \mu_i \right) = C(s, d) \int_{\mathbb{R}^d} \left(\sum_{i=1}^k \alpha_i \rho_i(x) \right)^{1 + s/d} dx \ . \tag{3.1}$$

The theorem follows immediately in view of Proposition 3.2 and of Proposition 4.1 below.

13

3.1 Optimal upper bound for piecewise constant ρ

Proposition 3.2 (Optimal upper bound for piecewise constant marginals). *Fix $0 < s < d$ and let $\mathsf{c}(x,y) = |x-y|^{-s}$. Set $k \geq 1$. For $i = 1, \ldots, k$, let $\mu_i \in \mathcal{P}(\mathbb{R}^d)$ be the uniform measure on Λ_i with density $\rho_i(x) = \frac{1_{\Lambda_i}(x)}{|\Lambda_i|}$, where $\Lambda_i \subset \mathbb{R}^d, i = 1, \ldots, k$ are Borel sets with ϕ-regular boundary and disjoint interiors. Let $\alpha_i \in \mathbb{R}_{>0}, i = 1, \ldots k$, be such that $\sum_{i=1}^{k} \alpha_i = 1$. Then if $C(s,d) > 0$ is the optimal constant from Proposition 2.5, we have*

$$
\begin{aligned}
\limsup_{N\to\infty} N^{-1-s/d} E^{\mathrm{xc}}_{N,s}\left(\sum_{i=1}^{k} \alpha_i \mu_i\right) &\leq -C(s,d) \sum_{i=1}^{k} \alpha_i^{1+s/d}\left(\int_{\Lambda_i} \rho_i^{1+s/d}(x)dx\right) \\
&= -C(s,d) \int_{\mathbb{R}^d}\left(\sum_{i=1}^{k} \alpha_i \rho_i(x)\right)^{1+s/d} dx \ .
\end{aligned} \tag{3.2}
$$

Remark 3.3. *The proof relies only on Propositions 2.3 and 2.5, and thus extends to more general costs.*

Proof. We would like to use the formula in Proposition 2.3. However, this will not be an immediate application to our case due to the form $\sum_{i=1}^{k} \alpha_i \mu_i$ of our measure, so we will first need to re-write the terms in our probability measure to bring them to a suitable form. In view of (2.5), and to preserve convergence to the mean field, we need to work with probability measures rather than just with positive measures.

We write $\alpha_i = \frac{q_i^N}{N}, q_i^N \in \mathbb{R}_{>0}, i = 1, \ldots, k$, with $\sum_{i=1}^{k} q_i^N = N$. We then write, denoting $\bar{N} := \sum_{i=1}^{k}[q_i^N]$,

$$
\begin{aligned}
\sum_{i=1}^{k} \alpha_i \mu_i &= \sum_{i=1}^{k} \frac{q_i^N}{N} \mu_i = \sum_{i=1}^{k} \frac{[q_i^N]}{N} \mu_i + \sum_{i=1}^{k} \frac{q_i^N - [q_i^N]}{N} \mu_i = \frac{\bar{N}}{N} \sum_{i=1}^{k} \frac{[q_i^N]}{\bar{N}} \mu_i + \frac{N - \bar{N}}{N} \sum_{i=1}^{k} \frac{q_i^N - [q_i^N]}{N - \bar{N}} \mu_i \\
&= \frac{\bar{N}}{N} \mu' + \frac{N - \bar{N}}{N} \mu'' \ ,
\end{aligned} \tag{3.3}
$$

with the obvious definitions for μ' and μ''.

From a double application of Proposition 2.3 we have

$$
E^{\mathrm{xc}}_{N,s}\left(\sum_{i=1}^{k} \alpha_i \mu_i\right) \leq E^{\mathrm{xc}}_{\bar{N},s}(\mu') + E^{\mathrm{xc}}_{N-\bar{N},s}(\mu'') \leq E^{\mathrm{xc}}_{\bar{N},s}(\mu') \leq \sum_{i=1}^{k} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \ , \tag{3.4}
$$

where for the second inequality we used that $E^{\mathrm{xc}}_{N-\bar{N}}(\mu'') \leq 0$. Equation (3.4) leads to

$$
\begin{aligned}
\limsup_{N\to\infty} N^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{N,s}\left(\sum_{i=1}^{k} \alpha_i \mu_i\right) &\leq \sum_{i=1}^{k} \limsup_{N\to\infty} N^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \\
&= \sum_{i=1}^{k} \limsup_{N\to\infty}\left(\frac{[q_i^N]}{N}\right)^{1+\frac{s}{d}} ([q_i^N])^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \\
&= \sum_{i=1}^{k} \alpha_i^{1+\frac{s}{d}} \limsup_{N\to\infty} ([q_i^N])^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \ ,
\end{aligned}
$$

where for the second equality we utilised that $[q_i^N]/N \leq \alpha_i \leq ([q_i^N]+1)/N, i = 1, \ldots, k$.

Fix $\delta > 0$. By Proposition 2.5 for large enough N we have for all $i = 1, \ldots, k$,

$$
([q_i^N])^{-1-\frac{s}{d}} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \leq -C(s,d) \int_{\Lambda_i} \rho_i^{1+\frac{s}{d}}(x)dx + \delta \ ,
$$

where $C(s,d) > 0$ is the constant from Proposition 2.5. Thus, for any fixed $\delta > 0$

$$
\limsup_{N\to\infty} N^{-1-\frac{s}{d}} \sum_{i=1}^{k} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \leq \sum_{i=1}^{k} \alpha_i^{1+\frac{s}{d}}\left(-C(s,d) \int_{\Lambda_i} \rho_i^{1+\frac{s}{d}}(x)dx + \delta\right).
$$

14

Taking $\delta \to 0$ in the last line in the above produces

$$\lim_{N\to\infty} N^{-1-\frac{s}{d}} \sum_{i=1}^{k} E^{\mathrm{xc}}_{[q_i^N],s}(\mu_i) \leq -C(s,d) \sum_{i=1}^{k} \alpha_i^{1+\frac{s}{d}} \int_{\Lambda_i} \rho_i^{1+\frac{s}{d}}(x)dx = \int_{\mathbb{R}^d} \left(\sum_{i=1}^{k} \alpha_i \rho_i(x) \right)^{1+s/d} dx, \qquad (3.5)$$

where we used the fact that $\rho_i(\Lambda_j) = 0$ for all $1 \leq i \neq j \leq k$ due to the disjointness assumption on the Λ_i and to the definition of ρ_i. $\square$

4   Optimal lower bound for piecewise constant ρ

The main result of this section is the following:

Proposition 4.1 (Optimal lower bound for piecewise constant marginals). *Fix $0 < s < d$ and let $\mathsf{c}(x,y) = |x-y|^{-s}$. Set $k \geq 1$. For $i = 1, \ldots, k$, let $\mu_i \in \mathcal{P}(\mathbb{R}^d)$ be the uniform measure on $\Lambda_i \subset \mathbb{R}^d$ with density $\rho_i(x) = \frac{1_{\Lambda_i}(x)}{|\Lambda_i|}$, where Λ_i are Borel sets with ϕ-regular boundary and disjoint interiors. Let $\alpha_i \in \mathbb{R}_{>0}, i = 1, \ldots k$, be such that $\sum_{i=1}^{k} \alpha_i = 1$. Then if $C(s,d) > 0$ is the optimal constant from Proposition 2.5, we have*

$$\liminf_{N\to\infty} N^{-1-s/d} E^{\mathrm{xc}}_{N,s} \left(\sum_{i=1}^{k} \alpha_i \mu_i \right) \geq -C(s,d) \sum_{i=1}^{k} \alpha_i^{1+s/d} \int_{\Lambda_i} \rho_i^{1+s/d}(x)dx$$

$$= -C(s,d) \int_{\cup_{i=1}^{k} \Lambda_i} \left(\sum_{i=1}^{k} \alpha_i \rho_i(x) \right)^{1+s/d} dx . \qquad (4.1)$$

In order to prove the statement of Proposition 4.1, we will use the result in Proposition 2.5 above, together with the Fefferman-Gregg decomposition introduced below.

4.1   Fefferman-Gregg decomposition and positive definiteness

The main aim of this section is to introduce the setting and tools required to prove Proposition 1.6 stated in the introduction.

4.1.1   The translations of $\mathbb{R}^d$ and their action

We consider the group of translations of $\mathbb{R}^d$ and its action

$$\tau_y x := x + y . \qquad (4.2)$$

This action extends to functions, as usual: if $f : (\mathbb{R}^d)^k \to V$ where V is a vector space, then we define $\tau_y(f)(x_1, \ldots, x_k) := f(\tau_y x_1, \ldots, \tau_y x_k)$. In particular we have the following properties:

1. For $y \in \mathbb{R}^d, f_1, f_2 : (\mathbb{R}^d)^k \to V$ we have $\tau_y(f_1 \cdot f_2) = (\tau_y f_1) \cdot (\tau_y f_2)$ if $\cdot$ is a scalar product on V

2. With the same notations as above, $\tau_y(f_1 * f_2) = (\tau_y f_1) * (\tau_y f_2)$.

The first property above uses just the definition of the action and is valid more in general, whereas the second property uses the fact that translations are linear and measure-preserving transformations.

For $f : (\mathbb{R}^d)^k \to V$, where V is a vector space, we define, whenever the following integral converges (in particular for compactly supported f, for example),

$$\langle f \rangle(x_1, \ldots, x_k) := \langle f(x_1, \ldots, x_k) \rangle := \int_{\mathbb{R}^d} \tau_y f(x_1, \ldots, x_k)dy = \int_{\mathbb{R}^d} f(x_1 + y, \ldots, x_k + y)dy .$$

In the case $k = 1$, we find that the above integral is defined for any $f \in L^1(\mathbb{R}^d, V)$ and equals the integral of f, by using the fact that translations are measure-preserving:

$$\langle f \rangle(x) = \int_{\mathbb{R}^d} f(x + y) dy = \int_{\mathbb{R}^d} f(y') dy' .$$

By using the property 2. of our group action, and denoting

$$f_-(x) := f(-x) ,$$

we then find that, for $f_1, f_2 : \mathbb{R}^d \to V$ such that $(f_1)_- * f_2 \in L^1(\mathbb{R}^d, V)$, the following holds:

$$
\begin{aligned}
\langle f_1(x_1) f_2(x_2) \rangle &= \int_{\mathbb{R}^d} f_1(x_1 + y) f_2(x_2 + y) dy = \int_{\mathbb{R}^d} f_1(x_1 - x_2 + y') f_2(y') dy' \\
&= ((f_1)_- * f_2) (x_2 - x_1) ,
\end{aligned}
$$

which gives a function depending only of $x_1 - x_2$.

In the case of $f_1 = f_2 = 1_A$, which is the indicator function of a set A, we also find, as $x_1 - x_2 = \tau_y x_1 - \tau_y x_2$ for all $y \in \mathbb{R}^d$ and $(1_A) - = 1_{(-A)}$, that for any function $g_0 : \mathbb{R}^d \to \mathbb{R}$ there holds

$$
\begin{aligned}
\langle 1_A(x_1) 1_A(x_2) g_0(x_1 - x_2) \rangle &= g_0(x_1 - x_2) \langle 1_A(x_1) 1_A(x_2) \rangle = g_0(x_1 - x_2) 1_{(-A)} * 1_A(x_2 - x_1) \\
&:= g_0(x_1 - x_2) h_A(x_1 - x_2) .
\end{aligned}
$$

Note that in particular, as a consequence of the fact that we performed an averaging $\langle \cdot \rangle$, there holds $h_{\tau A} = h_A$, i.e. the above function does not change under translations of A. In case $A = B_r(x)$ is a ball, by also using the fact that for $B_r := B_r(0)$ we have $B_r = -B_r$, there holds

$$\langle 1_{B_r(x)}(x_1) 1_{B_r(x)}(x_2) \rangle = \langle 1_{B_r}(x_1) 1_{B_r}(x_2) \rangle = h_{B_r}(x_1 - x_2) = 1_{B_r} * 1_{B_r}(x_1 - x_2) . \tag{4.3}$$

The main idea developed by Fefferman [F85], Gregg [Ge89] and [CLY89], and later perfected by Graf and Schenker [GS95] (based on the Yukawa potentials decomposition of [CLY89]) for the case where one averages also over rotations $SO(d)$, is to reorder the above integrals such that sums over good packings or over tilings occur. For this, consider a lattice $\mathcal{L} \subset \mathbb{R}^d$ and let $\Omega_{\mathcal{L}}$ be a fundamental domain for $\mathcal{L}$. Then we can write

$$\int_{\mathbb{R}^d} f(y) dy = \sum_{p \in \mathcal{L}} \int_{\Omega_{\mathcal{L}}} f(p + y') dy'. \tag{4.4}$$

We may use this principle for reordering the $\langle f \rangle$ integrals as follows:

$$
\begin{aligned}
\langle f(x_1, \ldots, x_k) \rangle &= \int_{\mathbb{R}^d} f(x_1 + y, \ldots, x_k + y) dy = \sum_{p \in \mathcal{L}} \int_{\Omega_{\mathcal{L}}} f(x_1 + p + y', \ldots x_k + p + y') dy' \\
&= \int_{\Omega_{\mathcal{L}}} \sum_{p \in \mathcal{L} + y'} f(x_1 + p, \ldots, x_k + p) dy'.
\end{aligned} \tag{4.5}
$$

4.1.2 Localization procedures on packings

Note that we don't need our kernel to be rotation-invariant, i.e. we don't require c to have the form $\mathsf{c}(x, y) = l(|x - y|)$: this stronger requirement would be necessary only for the Graf-Schenker decomposition, in which averages over rotations appear, and not for the one we describe here.

In order to obtain the asymptotic lower bound of our energies for kernels of the form $\mathsf{c}(x, y) = g(x - y)$, we desire to find a way in which to "localize" the interaction energies. Roughly stated, this means that if F is a packing of $\mathbb{R}^d$ by disjoint sets, then we would like to find a decomposition of our kernel of the form

$$g(x - y) = \sum_{A \in F} 1_A(x) 1_A(y) g(x - y) + err(x, y), \tag{4.6}$$

16

where the error term $err(x, y)$ can be well controlled. The kernels $1_A(x)1_A(y)g(x - y)$ then detect only interactions between points $x, y \in A$ and therefore, provided that the error term is well-behaved, a decomposition like (4.6) allows to reduce a study of our energies over the whole $\mathbb{R}^d$ to studies done "locally", separately on each $A \in F$. In fact in the concrete situations we face we will rather consider decompositions of the following averaged form, which is slightly more complicated than (4.6):

$$g(x - y) = \int_\Omega \left(\sum_{A \in F_\omega} 1_A(x)1_A(y)g(x - y) \right) d\mathbb{P}(\omega) + err(x, y), \tag{4.7}$$

where $(\Omega, \mathbb{P})$ is a probability space, and for each $\omega \in \Omega$ we decompose g along a separate packing F_ω. To have that $\mathbb{P}$ is a probability measure is not essential, but it simplifies our situation, because it implies that bounds done separately for each $\omega \in \Omega$ directly give the same bound for the integral. *Crucial* in our arguments will be to construct a decomposition such that both $err(x, y)$ and the integral term are positive (semi-)definite. This will allow us to cancel the mean field term in our calculations, for which a weaker form of positive definiteness is both a necessary and sufficient assumption, as shown in [P15].

To pass to more concrete calculations, we look at periodic packings which are all related to a basic one by isometries. Consider again a lattice $\mathcal{L}$ (which below will always be $\mathcal{L} = (\ell\mathbb{Z})^d$ for some suitably chosen $\ell > 0$) with fundamental domain $\Omega_\mathcal{L}$ as in the previous section, and suppose now that $A_1, \ldots, A_n$ are pairwise disjoint sets such that $A_i \subset \Omega_\mathcal{L}, 1 \le i \le n$. Observe that the following is a packing of $\mathbb{R}^d$ by copies of these sets:

$$F := \{A_i - p : p \in \mathcal{L}, \ 1 \le i \le n\}. \tag{4.8}$$

If we particularize the formula (4.5) to the case $k = 2$ and $f(x_1, x_2) = 1_{A_i}(x_1)1_{A_i}(x_2)$, we suppose that $\Omega_\mathcal{L}$ is symmetric so that $\Omega_\mathcal{L} = -\Omega_\mathcal{L}$, and we sum over $i = 1, \ldots, n$, then keeping in mind (4.8), we find:

$$\sum_{i=1}^n \langle 1_{A_i}(x_1)1_{A_i}(x_2) \rangle \overset{(4.5)}{=} \sum_{i=1}^n \int_{\Omega_\mathcal{L}} \sum_{p \in \mathcal{L}} (\tau_y 1_{A_i})(x_1 + y)(\tau_y 1_{A_i})(x_2 + y) dy$$

$$\overset{(4.8)}{=} \int_{\Omega_\mathcal{L}} \sum_{A \in F} 1_{A-y}(x_1)1_{A-y}(x_2) dy = \int_{\Omega_\mathcal{L}} \sum_{A \in F+y} 1_A(x_1)1_A(x_2) dy \tag{4.9}$$

where $F + y = \{A + y : A \in F\}$. Note that if the A_i are balls of the form $B_{r_i}(y_i)$ with M possible values $r_i \in \{R_1, \ldots, R_M\}$ then, due to the formula (4.3) and grouping together the n balls into families having same radius R_i we find, with $Y_i := \{y \in \mathbb{R}^d : B_r(y) \in F, r = R_i\}$:

$$\frac{1}{|\Omega_\mathcal{L}|} \int_{\Omega_\mathcal{L}} \sum_{A \in F+y} 1_A(x_1)1_A(x_2)g(x_1 - x_2) dy = \frac{1}{|\Omega_\mathcal{L}|} \sum_{i=1}^n \langle 1_{A_i}(x_1)1_{A_i}(x_2)g(x_1 - x_2) \rangle$$

$$= g(x_1 - x_2) \frac{1}{|\Omega_\mathcal{L}|} \sum_{i=1}^M \sum_{y \in Y_i} \langle 1_{B_{R_i}(y)}(x_1)1_{B_{R_i}(y)}(x_2) \rangle \tag{4.10}$$

$$\overset{(4.3)}{=} g(x_1 - x_2) \sum_{i=1}^M \frac{\#Y_i}{|\Omega_\mathcal{L}|} 1_{B_{R_i}} * 1_{B_{R_i}}(x_1 - x_2) = g(x_1 - x_2) \sum_{i=1}^M c_i \frac{1_{B_{R_i}} * 1_{B_{R_i}}}{|B_{R_i}|}(x_1 - x_2) ,$$

where c_i is the proportion of $\Omega_\mathcal{L}$ covered by those balls in F that have radius R_i, i.e.

$$c_i := \frac{\#Y_i |B_{R_i}|}{|\Omega_\mathcal{L}|} = \frac{\left| \bigcup_{x \in Y_i} B_{R_i}(x) \right|}{|\Omega_\mathcal{L}|} . \tag{4.11}$$

Following (4.9) and its link to (4.7), we see that in order to study the error introduced by the localization, in the case $g(x) = |x|^{-s}$ one faces the problem of how to bound a contribution of the form

$$\frac{1}{|x_1 - x_2|^s} \left(1 - \sum_{i=1}^n c_i \langle 1_{A_i}(x_1)1_{A_i}(x_2) \rangle \right) , \tag{4.12}$$

where $c_i \in \mathbb{R}$ are coefficients chosen depending on the precise details of the strategy that one follows. The two main approaches that proved successful in order do this are as follows:

17

- Fefferman [F85] for $d = 3, s = 1$ (later extended in [Ge89] to $0 < s < 3$) used formulas based on (4.10). At the same time the kernel error analogous to (4.12) is positive definite, and it has a bound by $M^{-1}|x_1 - x_2|^{-s} + O(M^{-2})$, where M is the total number of scales used in the technical part of the construction. The A_i's are here chosen to form a so-called "Swiss cheese" decomposition of $\Omega_{\mathcal{L}}$. One of the first appearances of this type of decomposition seems to be in [LeLi72, Sec. III], but see also the references therein. The parameter M appearing in the bound denotes the total number of scales used in this packing.

- Graf and Schenker [GS95] (see also [HLS09a], [HLS09b]) consider the situation where the A_i form a partition of $\Omega_{\mathcal{L}}$, and are all isometric to a fixed simplex, in which case the sum in (4.12) reduces to the case where only a single term appears, as the quantity h_A is invariant under isometry. In this case, it becomes crucial that at the same time (4.12) is positive definite and bounded by a value that decreases to zero as the scale of A decreases to zero. Unfortunately these conditions seem to hold only in the case $s \leq 1, d = 3$ and are explicitly proven in [GS95] only for $s = 1, d = 3$. On the other hand, the above Fefferman method will be extended here to all d and all $0 < s < d$.

The main structural differences between our setting and that of [F85] and [Ge89] are the following:

- The role of positive definiteness of g is more stringent here than in [F85, Ge89], because the mean field $\frac{\mu \otimes \mu}{\mu^2(\mathbb{R}^d)}$ is a minimizer of $F_{\infty,g}^{\mathrm{OT}}$ only in that case, by [P15]. We don't know how to explicitly describe the minimizers of $F_{\infty,g}^{\mathrm{OT}}(\mu)$ for g not positive definite. The energy $E_{N,g}^{\mathrm{xc}}(\mu) = F_{N,g}^{\mathrm{OT}}(\mu) - \frac{N(N-1)}{2} F_{\infty,g}^{\mathrm{OT}}(\mu)$ is thus giving actual next-order terms only if g is positive definite: thus only by having positive definite error terms in our kernel decompositions can we expect the first order effects to be cancelled by the associated mean fields. In this way the result of [P15] is a crucial ingredient to keep in mind in our strategy.

- The competitors for $F_{\infty,g}^{\mathrm{OT}}(\mu)$ are automatically competitors for $F_N^{\mathrm{OT}}(\mu)$: therefore the next-order energy $E_{N,g}^{\mathrm{xc}}$ is negative, and we look for a lower bound for it. On the contrary, [F85, Ge89] had a positive energy, and were interested in an upper bound for it.

4.1.3 The packing lemma and the construction of $\Omega_\ell, \mathbb{P}_\ell, F_\omega^\ell$

In order to be more specific about the form that (4.12) will have, we recall the Swiss cheese lemma, first introduced in [LeLi72] and used in [F85, H89, Ge89], for $d = 3$. The main idea therein was to decompose regions in space into sets of disjoint balls with geometrically increasing radii. We need to adapt the computations from [H89] to the case of general dimension d. In the previous works the authors work at scales $R_1 > 1$, whereas here we only require $R_1 > 0$, since the argument in [H89] easily adapts to this more general setting.

Lemma 4.2 ("Swiss cheese" lemma in general dimension)**.** *For each $d \geq 2$ define $C_d := \frac{2^{d+1}}{|B_1|}$, where B_1 is the unit ball centred at zero $\{x \in \mathbb{R}^d : |x| \leq 1\}$. Consider a sequence of real numbers $0 < R_1 < \cdots < R_M$ such that $R_{k+1} > (1 + 4\sqrt{d}|B_1|)R_k$ for all k. Assume that Q is a cube of side length $l > 8\sqrt{d}|B_1|(M + C_d)R_M$. Then there exists a family $\mathcal{B}$ of disjoint balls $B_r(x) \subset Q$, with $x \in Q$, such that*

- *The balls in $\mathcal{B}$ have radii $r \in \{R_1, \ldots, R_M\}$.*

- *For each $i = 1, \ldots, M$, if*
$$Y_i := \{x \in Q : \exists B_r(x) \in \mathcal{B}, r = R_i\} ,$$
  then there holds
$$\frac{1}{M + C_d + 1} < c_i := \frac{\left|\bigcup_{x \in Y_i} B_{R_i}(x)\right|}{|Q|} < \frac{1}{M + C_d} . \tag{4.13}$$

The proof of the lemma follows the same ideas as in [Ge89, Sec. 4] or in [H89, Sec. 4.3], namely it is based on packing on a cube, then periodized, and it proceeds by induction on i. We refer to [Ge89] for the details.

We will use the ball packing given by Lemma 4.2 for the case $Q = \Omega_{\mathcal{L}} = [-\frac{l}{2}, \frac{l}{2}]^d$ in (4.10), and then we define a packing of the whole of $\mathbb{R}^d$ by extending by $\mathcal{L}$-periodicity, for $\mathcal{L} = (\ell\mathbb{Z})^d$. We use a special case of the

18

definition (4.8) for the case $\{A_1, \ldots, A_n\} = \mathcal{B}$. The dependence on $\mathcal{L}$ is not recorded in the notation for the sake of lightness, and we define

$$F_{\mathcal{B}} := \{B - p : B \in \mathcal{B}, p \in \mathcal{L}\} . \tag{4.14}$$

We note that the formulas (4.13) and (4.11) giving the coefficients c_i agree, for the choice of packing family $\mathcal{B}$ as in Lemma 4.2.

However the basic formula (4.10), in which $1_{B_{R_i}} * 1_{B_{R_i}}(x)$ appears, cannot be directly used, for regularity reasons. Indeed, the relevant positive definiteness criterion, formulated in Corollary A.2, requires a control on derivatives up to order d of our kernel approximants (see (A.2)). We consider first the expression

$$1_{B_r} * 1_{B_r}(x) = \begin{cases} \dfrac{\pi^{\frac{d-1}{2}}}{2^{d-1}\Gamma\left(\frac{d+1}{2}\right)} \int_{|x|}^{2r} (4r^2 - y^2)^{\frac{d-1}{2}} dy & \text{for } |x| \leq 2r , \\ 0 & \text{for } |x| > 2r , \end{cases} \tag{4.15}$$

The above convolutions of balls have only $1 + [(d-1)/2]$ continuous derivatives, a fact which forbids the desired bounds. Similarly to the procedure in [Ge89], we therefore apply a mollification to (4.15). Note that the mollification in [Ge89] does not allow full regularity: the averaging over the radii allows only to add the control on one more extra derivative to the above-mentioned $1 + [(d-1)/2]$ derivatives of $1_{B_r} * 1_{B_r}$. While this is enough to cover the range $0 < s < 3$ in dimension $d = 3$, this is unsatisfactory for general d and therefore we proceed differently than [Ge89], with a reasoning that covers all the cases $0 < s < d$ at once.

To start with, fix a small parameter η, say $\eta \in (0, 1/2]$, a choice that will play no special role in the computations below. For $t \in [1 - \eta, 1 + \eta]$, consider a positive function $\rho_\eta \in C_0^\infty([1 - \eta, 1 + \eta])$ such that $\int \rho_\eta(t)dt = 1$, and set

$$Q_{0,\eta}(x) := \int_{1-\eta}^{1+\eta} \frac{\left(1_{B_t} * 1_{B_t}\right)(x)}{|B_t|} \rho_\eta(t)dt . \tag{4.16}$$

Then $Q_{0,\eta}$ is smooth outside the origin, because it is radial and in radial coordinates we may use the smoothness of ρ_η to control the radial partial derivatives. Also note that $Q_{0,\eta}(0) = 1$ because $1_{B_t} * 1_{B_t}(0) = |B_t|$ and $\int \rho_\eta(t)dt = 1$. We then define, for $i = 1, \ldots, M$, and R_i as in Lemma 4.2 (but note that here B_{tR_i} is for each i the ball of radius tR_i centered at 0)

$$Q_{i,\eta}(x) := Q_{0,\eta}\left(\frac{x}{R_i}\right) = \int_{1-\eta}^{1+\eta} \frac{\left(1_{B_{tR_i}} * 1_{B_{tR_i}}\right)(x)}{|B_{tR_i}|} \rho_\eta(t)dt . \tag{4.17}$$

As a consequence of the bounds and support properties of the above integrands, for all $x \in \mathbb{R}^d$ we have

$$Q_{i,\eta}(x) \leq 1_{B_{2(1+\eta)R_i}}(x) . \tag{4.18}$$

Lemma 4.2 with the above choices of radii $R_i, 1 \leq i \leq M$, and $Q = [-\ell/2, \ell/2]^d$, produces a family of disjoint balls satisfying in particular estimate (4.13), which for $\mathcal{L} := (\ell\mathbb{Z})^d$ is then extended by $\mathcal{L}$-periodicity to a disjoint packing F as in (4.13).

For $t \in [1 - \eta, 1 + \eta]$, we consider the dilation $t\mathcal{B}$ by t of the packing $\mathcal{B}$, and the dilations $t\mathcal{L}, tQ$ of the lattice $\mathcal{L}$ and of the cube Q. Moreover if $B = B_r(x)$ is a ball in $\mathbb{R}^d$ then we denote $tB := B_{tr}(tx)$. Then

$$t\mathcal{B} := \{tB : B \in \mathcal{B}\} \qquad t\mathcal{L} := \{tp : p \in \mathcal{L}\} , \qquad tQ := [-t\ell/2, t\ell/2]^d = \Omega_{t\mathcal{L}} .$$

Then we claim that the covered volume ratios c_i appearing in (4.11) and in (4.13) do not depend on the choice of $t > 0$. Indeed, the packing $t\mathcal{B}$ obtained after the dilation by t is still made of balls, this time with set of radii $tR_1, \ldots, tR_M$, which now cover $tQ = t\Omega_{\mathcal{L}}$. Moreover the dilation preserves the property of the packing of being made of disjoint balls, and for $t\mathcal{B}$ the ball center sets Y_i are replaced by new center sets

$$\begin{aligned} tY_i : \ &= \ \{tx : x \in Y_i\} = \{tx \in \mathbb{R}^d : \exists B_r(x) \in \mathcal{B}, r = R_i\} = \{tx \in \mathbb{R}^d : \exists B_r(tx) \in t\mathcal{B}, r = tR_i\} \\ &= \ \{x \in \mathbb{R}^d : \exists B_r(x) \in t\mathcal{B}, r = tR_i\} . \end{aligned}$$

After the above replacements, the values c_i appearing in the formula (4.13) don't change:

$$\frac{\left|\bigcup_{x \in tY_i} B_{tR_i}(tx)\right|}{|tQ|} = \frac{\#(tY_i)|B_{tR_i}|}{|tQ|} = \frac{\#(Y_i)t^d|B_{R_i}|}{t^d|Q|} = \frac{\#(Y_i)|B_{R_i}|}{|Q|} = \frac{\left|\bigcup_{x \in Y_i} B_{R_i}(x)\right|}{|Q|} .$$

Then we define, analogously to (4.14),

$$F_{tB} := \{B - p : \ B \in t\mathcal{B}, p \in t\mathcal{L}\} = \{tB - tp : \ B \in \mathcal{B}, p \in \mathcal{L}\} = tF_{\mathcal{B}} \ .$$

We now use formula (4.10) for the ball covers corresponding to the packing of Lemma 4.2 for the choice $Q = \Omega_{\mathcal{L}}$, then we superpose dilated versions of these packings in order to obtain a contribution like (4.17) in place of the contribution of each B_{R_i} in (4.10). We find the following formula, where we use the fact that $\Omega_{\mathcal{L}} = [-\ell/2, \ell/2]^d, |\Omega_{\mathcal{L}}| = l^d$ and $t\Omega_{\mathcal{L}} = [-t\ell/2, t\ell/2]^d, |t\Omega_{\mathcal{L}}| = (tl)^d$:

$$\sum_{i=1}^{M} g(x_1 - x_2)c_i Q_{i,\eta}(x_1 - x_2) = \int_{1-\eta}^{1+\eta} \left(\frac{1}{(lt)^d} \int_{[-\frac{lt}{2}, \frac{lt}{2}]^d} \sum_{B \in tF_{\mathcal{B}} + y} 1_B(x_1) 1_B(x_2) g(x_1 - x_2) dy \right) \rho_\eta(t) dt \ . \quad (4.19)$$

We observe that the left hand side of (4.19) is of the form that appears in (4.7), namely

Lemma 4.3. *We have*

$$\sum_{i=1}^{M} g(x_1 - x_2)c_i Q_{i,\eta}(x_1 - x_2) = \int_{\Omega} \left(\sum_{A \in F_\omega} 1_A(x_1) 1_A(x_2) g(x_1 - x_2) \right) d\,\mathbb{P}_l(\omega) \ , \quad (4.20)$$

for the choices

$$\Omega = \Omega_l := \left\{ (t, y) \in [1 - \eta, 1 + \eta] \times \mathbb{R}^d : \ y \in \left[-\frac{lt}{2}, \frac{lt}{2} \right]^d \right\},$$

$$d\,\mathbb{P}(t, y) = d\,\mathbb{P}_l(t, y) := \frac{\rho_\eta(t)}{(lt)^d} 1_{\Omega_l}(t, y) \ dt \ dy \ , \quad (4.21)$$

and to each $\omega := (t, y) \in \Omega_l$ we associate an isometry y and a packing F_ω given by

$$F_\omega = F_\omega^l := tF_{\mathcal{B}} + y = \{tA + y : \ A \in F_{\mathcal{B}}\} \ . \quad (4.22)$$

In the above we use the notation (4.2) regarding the definition and action of isometries, while the parameterization of isometries by y, and of dilations by t is encoded via the set Ω_l. Finally note that the measure $\mathbb{P}_l$ is defined precisely so as to equate (4.19) to (4.20), and $l, \mathcal{B}$ are as defined from Lemma 4.2.

Note that differently than in Gregg's work [Ge89], here we fix once and for all the packing $\mathcal{B}$ depending only on the choices of $R_1, \ldots, R_M, l$ from that lemma, and do not need to produce a new packing separately anew for each dilation t. It is obvious that if $\mathcal{B}$ is a packing by disjoint balls then $g(t\mathcal{B})$ is still a disjoint packing for all isometries g and all dilations t.

Remark 4.4 (the case of simplices). *We may directly define $Q_{A,\eta}$ as in (4.17) for the case of more general sets A rather than balls B_{R_i}, by replacing B_r by $tA := \{tx : \ x \in A\}$. In particular this can be done for the case of simplices $A = \triangle$ if we include also an averaging over $SO(d)$ when defining $\langle f \rangle$, as done in [GS95], where the singularity for the function $h_\triangle$ appearing therein, has the same kind of discontinuity as $1_B * 1_B$. The packing from Lemma 4.2 could then be replaced by a tiling, again by families of simplices of sizes decreasing like a geometric series. However, it seems hard to prove positive definiteness and boundedness at zero of the error terms based on the less explicit formula for $h_\triangle$, outside the case $s = 1, d = 3$.*

It would be interesting to have a more thorough investigation of the influence of general packing strategies on the study of positive definiteness, but that would go beyond the scope of the present work.

4.1.4 Control of the error terms in the localization estimate

Proof of Proposition 1.6

With the decompositions from the previous section we come back to the study of the final formula of type (4.7) which will appear in our proof. We consider now only the case $c(x, y) = |x - y|^{-s}, 0 < s < d$. We use the formula

(4.20) with choices (4.21), and in this case we write, following the formalism (4.7), with $\Omega = \Omega_l, \mathbb{P} = \mathbb{P}_l, F_\omega = F_\omega^l$ as in (4.21), (4.22)

$$err(x_1, x_2) := \frac{1}{|x_1 - x_2|^s} - \int_{\Omega_l} \left(\sum_{A \in F_\omega^l} \frac{1_A(x_1) 1_A(x_2)}{|x_1 - x_2|^s} \right) d\mathbb{P}_l(\omega) \ .$$

In fact, we will find it useful in our computations to work with a slightly different error term from the one above, for which we can show both positive-definiteness, and also a rough next-order lower bound needed in our proof. More precisely, we introduce for $0 < \epsilon < d/2$ and $\epsilon \le s \le d - \epsilon$

$$w(x_1 - x_2) := \left(1 + \frac{C}{M} \right) |x_1 - x_2|^{-s} - \int_{\Omega_l} \left(\sum_{A \in F_\omega^l} \frac{1_A(x_1) 1_A(x_2)}{|x_1 - x_2|^s} \right) d\mathbb{P}_l(\omega) \ , \tag{4.23}$$

with $C = C(\rho_\eta, d, \epsilon) > 0$ a constant depending only on the choice of ρ_η, on d and on ϵ, the form of which will be made explicit in Lemma A.4 in the Appendix.

Proposition 1.6 becomes an immediate consequence of (4.7), due to Lemma A.6 which proves the positivity and boundedness properties of w, and due to Lemma A.7 which proves asymptotic lower bounds for w.

$\square$

Remark 4.5 (continuation of Remarks 1.3 and 1.7)**.** *We briefly mention here how the construction of the decomposition of Proposition 1.6 should be adapted for the examples from Remark 1.3.*

(a) *For the case of a general compact d-dimensional Riemannian manifold (M, g) the main difference to our setting is that we don't have a group action by translations. However in this case we can still adapt Lemma 4.2 and find a good packing by metric balls of scale $l > 0$. Then for $l \ll 1$ we may use the fact that at small scales near each $x_0 \in M$, the manifold (M, g) is closer and closer to $(\mathbb{R}^d, g(x_0))$. This allows to reduce the packing construction to one on $\mathbb{R}^d$ and use translation averaging locally, up to an error depending on the modulus of continuity of g. Regarding Lemma A.6, in this case it can be directly implemented on manifolds, by using the decomposition in dyadic scales and the same type of positive definiteness criteria as in the appendix. Lemma A.7 is provable by a similar localization argument, reducing it to a problem on $(\mathbb{R}^d, g(x_0))$ up to small error, and all the tools [HS02], [LNP16] used in the appendix can be extended to this case.*

(b), (c) *For the non-homogeneous kernels as in these cases, the packing arguments go precisely as in the present section, allowing to obtain a kernel decomposition. Concerning Lemma A.6, we may still use the positive definiteness criteria as in Lemma A.1, by replacing the smoothed kernels $Q_{i,\eta}$ with non-isotropic analogues. For adapting Lemma A.7 the main difference is that the proof of a rough lower bound as in Lemma A.7 cannot be directly done by using the result of [HS02]. This however is a robust result, and in our case we could for example apply an ad-hoc version of [HS02] which uses decompositions of the product-type costs from (b), (c) which use non-isotropic elementary kernels modelled on the kernels at hand, rather than the isotropic $1_{B_r} * 1_{B_r}$ as in [HS02].*

We also note that in all cases (a), (b), (c), the kernels $\mathsf{c}(x, y) = g(x - y)$ under consideration have asymptotic homogeneity $-s$ near the singularity at $x = y$ (i.e. there exists a function $g_0(x) : \mathbb{R}^d \to \mathbb{R}$ such that $g_0(\lambda x) = \lambda^{-s} g_0(x)$ and $\lim_{x \to 0} g(x)/g_0(x) = 1$), and this property can be quantified in order to allow the rough lower bound formula (1.14).

4.2 Optimal lower bound for Coulomb and Riesz costs by Fefferman decomposition

Before proving Proposition 4.1, we will need a number of helpful lemmas and corollaries.

Recall that we denoted for any cost function $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$ and for any $N \ge 2$ by

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) = \inf \left\{ \int_{(\mathbb{R}^d)^N} \sum_{i,j=1, i \ne j}^N \mathsf{c}(x_i, x_j) d\gamma_N(x_1, \ldots, x_N) : \ \gamma_N \in \mathcal{P}_{sym}^N(\mathbb{R}^d), \gamma_N \mapsto \mu \right\} \ ,$$

and $F_{0,c}^{OT}(\mu) = F_{1,c}^{OT}(\mu) = 0$, with a corresponding formula for $E_{N,c}^{xc}(\mu)$. Moreover, for the particular case of the cost $c(x,y) = |x - y|^{-s}$ we use the notations $F_{N,s}^{OT}(\mu)$ and $E_{N,s}^{xc}(\mu)$. Note that for the purposes of the present section we don't need to impose that c be positive, and this is imposed just for consistency with the rest of the paper.

4.2.1 Splitting the cost

The following lemma uses in the Optimal Transport framework the "grand-canonical" formalism of [LLS17], and will help us obtain our main result for very general densities, by means of our strategy involving the use of measures with piecewise constant density. We thank M. Lewin and S. Di Marino for pointing out at the end of May 2017 in Jyväskylä a flaw in the corresponding "splitting-the-cost" lemma which we had used in the preliminary version of the paper. In order to settle this flaw it turned out to be sufficient, as suggested by M. Lewin in Jyväskylä, to use the functional $F_{GC,N,c}^{OT}$ allowing fluctuations in the number of marginals, rather than using $F_{N,c}^{OT}$ throughout the proof. We also introduce in (4.24) the auxiliary minimization problem $F_{GCB,N,c}^{OT}$ in which we impose an upper bound on these fluctuations.

To the best of our knowledge, the results in Lemma 4.8, Lemma 4.10 and Corollary 4.11 below are new and of independent interest.

We will find it useful in some of the proofs below to work with another quantity, $F_{GCB,M,c}^{OT}(\mu, \bar{M})$, due to its finite summation structure. That is, for all $M \in \mathbb{R}_{>0}, \bar{M} \in \mathbb{N}_+$ such that $\bar{M} \geq M$, and for $\mu \in \mathcal{P}(\mathbb{R}^d)$, let

$$F_{GCB,M,c}^{OT}\left(\mu, \bar{M}\right) := \inf \left\{ \sum_{n=2}^{\bar{M}} \alpha_n F_{n,c}^{OT}(\mu_n) \ \middle| \ \begin{array}{c} \sum_{n=0}^{\bar{M}} \alpha_n = 1, \ \sum_{1 \leq n \leq \bar{M}} n\alpha_n \mu_n = M\mu, \\ \mu_n \in \mathcal{P}(\mathbb{R}^d), \ \alpha_n \geq 0, \quad \text{for} \quad n = 0, 1, \ldots, \bar{M} \end{array} \right\}, \tag{4.24a}$$

or equivalently, as summing zero contributions has no effect on the sums,

$$F_{GCB,M,c}^{OT}\left(\mu, \bar{M}\right) = \inf \left\{ \sum_{n=2}^{\infty} \alpha_n F_{n,c}^{OT}(\mu_n) \ \middle| \ \begin{array}{c} \sum_{n=0}^{\infty} \alpha_n = 1, \ \sum_{n=1}^{\infty} n\alpha_n \mu_n = M\mu, \\ \mu_n \in \mathcal{P}(\mathbb{R}^d), \ \alpha_n \geq 0, \quad \text{for} \quad n = 0, 1, \ldots, \bar{M} \\ \alpha_n = 0, \quad \text{for} \quad n > \bar{M} \end{array} \right\}. \tag{4.24b}$$

We then define

$$E_{GCB,M,c}^{xc}\left(\mu, \bar{M}\right) := F_{GCB,M,c}^{OT}\left(\mu, \bar{M}\right) - M^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) d\mu(x) d\mu(y). \tag{4.24c}$$

Remark 4.6. *Definition* (4.24a) *says that there are M marginals "on average", and the total number of marginals usable in this decomposition is bounded above by $\bar{M}$. Removing this last constraint we defined $F_{GC,M,c}^{OT}(\mu)$ in* (1.7a)*, which is the precise analogue of [LLS17, eq. (3.1)]. For the GCB-problem in* (4.24a)*, as the possible number of marginals used is bounded by $\bar{M}$, the existence of minimizers can be proven like for the $F_{N,c}^{OT}(\mu)$-problems, based on the same techniques which work for the $N = 2$ case (see [V09, Thm. 4.1]), and therefore in case of lower semicontinuous cost c and μ as in the statement of Lemma 4.8 the infimum is realized. The proof then extends to the GC-problem by using the fact that the measures given by $\lambda_k(\{n\}) := \lambda_{n,k}$ corresponding to the GC-optimizers form a tight sequence of probability measures on $\mathbb{N}$, as will be shown in detail in the proofs of Lemma E.1 and Lemma E.2.*

Remark 4.7. *We note here the following special properties related to the problems* (1.7) *and* (4.24):

1. *For the case $\bar{N} = N \in \mathbb{N}_+, N \geq 2$, the relaxation* (4.24a) *satisfies*

$$F_{GCB,N,c}^{OT}(\mu, N) = F_{N,c}^{OT}(\mu), \tag{4.25}$$

   as in this case the condition $\sum_{n=1}^{N} n\alpha_n = N$ can be realized only if $\alpha_n = 0$ for all $n < N$ and $\alpha_N = 1$. The (4.25) *together with* (4.24a) *imply in turn that for all $\bar{M}, \bar{N}, N \in \mathbb{N}_+, \bar{N} \geq N \geq 2$, we have*

$$F_{GC,N,c}^{OT}(\mu) \leq F_{GCB,N,c}^{OT}(\mu, \bar{N} + \bar{M}) \leq F_{GCB,N,c}^{OT}(\mu, \bar{N}) \leq F_{N,c}^{OT}(\mu). \tag{4.26}$$

2. If $\mu = \frac{N'\mu' + N''\mu''}{N' + N''}$ where μ', μ'' are probability measures, and $N', N'' \in \mathbb{R}_{>0}$, then

$$E^{\text{xc}}_{\text{GC}, N'+N'', \mathsf{c}}(\mu) \leq E^{\text{xc}}_{\text{GC}, N', \mathsf{c}}(\mu') + E^{\text{xc}}_{\text{GC}, N'', \mathsf{c}}(\mu''). \qquad (4.27)$$

If, for example, $0 < N' \leq 1$, then

$$E^{\text{xc}}_{\text{GC}, N'+N'', \mathsf{c}}(\mu) \leq E^{\text{xc}}_{\text{GC}, N'', \mathsf{c}}(\mu''). \qquad (4.28)$$

Equation (4.27) was stated in [LLS17, eqn. (3.2)]; we provide a sketch of the proof in Appendix D.

3. We now consider the GCB-problem, in the equivalent definition (4.24b). If $\mu = \frac{N'\mu' + N''\mu''}{N' + N''}$ where μ', μ'' are probability measures, and $N', N'' \in \mathbb{R}_{>0}, \bar{N}', \bar{N}'' \in \mathbb{N}_+, \bar{N}' \geq N', \bar{N}'' \geq N''$, then

$$E^{\text{xc}}_{\text{GCB}, N'+N'', \mathsf{c}}(\mu, \bar{N}' + \bar{N}'') \leq E^{\text{xc}}_{\text{GCB}, N', \mathsf{c}}(\mu', \bar{N}') + E^{\text{xc}}_{\text{GCB}, N'', \mathsf{c}}(\mu'', \bar{N}''). \qquad (4.29)$$

If, for example, $0 < N' \leq 1$, then

$$E^{\text{xc}}_{\text{GCB}, N'+N'', \mathsf{c}}(\mu, \bar{N}' + \bar{N}'') \leq E^{\text{xc}}_{\text{GCB}, N'', \mathsf{c}}(\mu'', \bar{N}''). \qquad (4.30)$$

The proof of (4.29) and (4.30) can be found in Appendix D.

4. If $N \in \mathbb{R}_{>0}, \bar{N} \in \mathbb{N}_+, \bar{N} \geq N$ then

$$E^{\text{xc}}_{\text{GC}, N, \mathsf{c}}(\mu) \leq E^{\text{xc}}_{\text{GCB}, N, \mathsf{c}}(\mu, \bar{N}) \leq 0. \qquad (4.31)$$

For $N \in \mathbb{N}_+$ this follows from (4.26) since $E^{\text{xc}}_{N, \mathsf{c}}(\mu) \leq 0$. For the general $N \in \mathbb{R}_+$ case, if $N < 2$ then we use our convention that $F^{\text{OT}}_N = 0$ for $N = 0, 1$, while for $N \geq 2$ we apply (4.28) to $N' + N'' = N, N'' = [N], N' = N - [N]$, and $\mu' = \mu'' = \mu$, to get $E^{\text{xc}}_{\text{GC}, N, \mathsf{c}}(\mu) \leq E^{\text{xc}}_{\text{GC}, [N], \mathsf{c}}(\mu) \leq 0$. A similar argument holds for $E^{\text{xc}}_{\text{GCB}, N, \mathsf{c}}(\mu, \bar{N})$.

5. For $\alpha > 0$ let μ_α be the measure given by $d\mu_\alpha(x) = \rho_\alpha(x)dx$ with $\rho_\alpha(x) = \rho(\alpha x)$. Then we have

$$E^{\text{xc}}_{\text{GCB}, N, s}(\alpha^d \mu_\alpha) = \alpha^s E^{\text{xc}}_{\text{GCB}, N, s}(\mu) \quad \text{and} \quad E^{\text{xc}}_{\text{GC}, N, s}(\alpha^d \mu_\alpha) = \alpha^s E^{\text{xc}}_{\text{GC}, N, s}(\mu) . \qquad (4.32)$$

The proof follows similarly to the one of (2.7) from Lemma 2.4, the key difference being that since the the optimizers μ_n are not assured to possess densities, one needs to use the change of measure

$$\alpha^d \mu_\alpha = F_\# \mu \quad \text{and} \quad \alpha^d \mu_{n, \alpha} = F_\# \mu_n, \quad \text{where} \quad F(x) = \alpha x, \quad x \in \mathbb{R}^d.$$

Lemma 4.8. Let $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$. Take $\mu_1, \ldots, \mu_k \in \mathcal{P}(\mathbb{R}^d)$, with densities $\rho_1, \ldots, \rho_k$, such that the quantities below are finite (e.g. if $\mathsf{c}(x, y) = |x - y|^{-s}, 0 < s < d$, we can require $\rho_1, \ldots, \rho_k \in L^{\frac{2d}{2d-s}, 2}(\mathbb{R}^d)$), and assume that the μ_j have essentially disjoint supports, i.e. that $\mu_i(A)\mu_j(A) = 0$ for every Borel set A and for all $1 \leq i \neq j \leq k$. Fix $M_1, \ldots, M_k \in \mathbb{R}_{>0}$ such that $\bar{M} := \sum_{i=1}^{k} M_i \in \mathbb{N}_+, \bar{M} \geq 2$, and denote by $\mathsf{c}_{ii}(x, y) := 1_{\Lambda_i}(x)1_{\Lambda_i}(y)\mathsf{c}(x, y)$ for $i = 1, \ldots, k$ and $x, y \in \mathbb{R}^d$. Let μ be the probability measure with density $\sum_{i=1}^{k} M_i \rho_i / \bar{M}$.

Then the following holds:

$$F^{\text{OT}}_{\bar{M}, \sum_{i=1}^{k} \mathsf{c}_{ii}}(\mu) = F^{\text{OT}}_{\bar{M}, \sum_{i=1}^{k} \mathsf{c}_{ii}}\left(\frac{\sum_{i=1}^{k} M_i \mu_i}{\bar{M}}\right) \geq \sum_{i=1}^{k} F^{\text{OT}}_{\text{GCB}, M_i, \mathsf{c}}(\mu_i, \bar{M}) \geq \sum_{i=1}^{k} F^{\text{OT}}_{\text{GC}, M_i, \mathsf{c}}(\mu_i) . \qquad (4.33)$$

Before we prove the above lemma, note that

$$F^{\text{OT}}_{N, \mathsf{c}}(\mu) = \sup \left\{ N \int f(x)d\mu(x) \, \middle| \, \begin{array}{l} \text{for all } x_1, \ldots, x_N \in \text{spt}(\mu), \\ \text{there holds } \sum_{i=1}^{N} f(x_i) \leq \sum_{1 \leq i \neq j \leq N} \mathsf{c}(x_i, x_j) \end{array} \right\} . \qquad (4.34)$$

The above duality is a bit different than the usual version (see, for example, Theorem 5.9 from [V09]), due to the fact that we test only configurations that are all contained in the support of μ rather than general configurations

in $(\mathbb{R}^d)^N$. This restriction can be applied by *firstly* noting that any plan γ_N such that $\gamma_N \mapsto \mu$ has support in $(\mathrm{spt}(\mu))^N$, therefore

$$F_{N,\mathsf{c}}^{\mathrm{OT}}(\mu) = \inf\left\{\int_{(\mathrm{spt}(\mu))^N} \sum_{1 \leq i \neq j \leq N} \mathsf{c}(x_i, x_j) d\gamma_N(x_1, \ldots, x_N), \gamma_N \in \mathcal{P}_{sym}((\mathrm{spt}(\mu))^N), \ \gamma_N \mapsto \mu\right\},$$

and *secondly*, noting that the proof of duality in [dP15] carries through without changes once we replace the space $\mathbb{R}^d$ by the closed subset $\mathrm{spt}(\mu) \subset \mathbb{R}^d$, proving (4.34).

From (4.34), we immediately get

$$F_{M_i,\mathsf{c}}^{\mathrm{OT}}(\mu_i) = F_{M_i,\mathsf{c}_{ii}}^{\mathrm{OT}}(\mu_i) = F_{M_i,\sum_{j=1}^k \mathsf{c}_{jj}}^{\mathrm{OT}}(\mu_i) . \tag{4.35}$$

Proof of Lemma 4.8: We consider below for simplicity of exposition just the case with $k = 2$, and we denote $\Lambda_1 =: A, \Lambda_2 =: B$ and $\mathsf{c}_{11} =: \mathsf{c}_A, \mathsf{c}_{22} =: \mathsf{c}_B$, moreover we replace in the notation M_1, M_2 by M, N and μ_1, μ_2 by μ, ν.

Step 1. To begin with, we decompose the minimization problem defining $F_{\mathsf{c}_A+\mathsf{c}_B,M+N}^{\mathrm{OT}}\left(\frac{M\mu+N\nu}{M+N}\right)$ according to the cardinality of points in the configurations used in our transport plans, and belonging respectively to A, B, in the spirit of [LLS17, Sec. 3]. Note that the Borel sets $\{B_{N_A,N_B} : N_A, N_B \in \mathbb{N}, N_A + N_B = M + N\}$ form a partition of such configurations, where

$$B_{N_A,N_B} := \left\{(x_1, \ldots, x_{M+N}) \in (A \cup B)^{M+N} \ \middle| \ \begin{array}{l} \#\{i : \ 1 \leq i \leq M+N : \ x_i \in A\} = N_A \\ \#\{i : \ 1 \leq i \leq M+N : \ x_i \in B\} = N_B \end{array}\right\} .$$

These sets are symmetric under permutations of the $\mathbb{R}^d$-coordinates in $(\mathbb{R}^d)^N$. We observe here that there are $M + N + 1$ sets B_{N_A,N_B} such that $N_A + N_B = M + N$. As explained in more detail in Step 4 below, this fact will be crucial as to why the extra parameters α_0 and α_1 are introduced in (4.24a) and (1.7a).

Denote now the symmetric probability measures with fixed numbers of points in A, B as follows:

$$\mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B}) := \left\{\gamma_{N_A|N_B} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{M+N}) : \ \gamma_{N_A|N_B}(B_{N_A,N_B}) = 1\right\} . \tag{4.36}$$

Note that for any $\gamma_{N_A|N_B} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B})$ its marginal is in the class

$$\mathcal{P}_{N_A|N_B}(\mathbb{R}^d) := \left\{\bar{\mu} \in \mathcal{P}(\mathbb{R}^d) : \ \bar{\mu}(A) = \frac{N_A}{M+N}, \ \bar{\mu}(B) = \frac{N_B}{M+N}\right\} . \tag{4.37}$$

Thus for $\bar{\mu} \in \mathcal{P}_{N_A|N_B}(\mathbb{R}^d)$, the infimum in the following optimal transport problem is over a nonempty set:

$$F_{N_A|N_B,\mathsf{c}}^{\mathrm{OT}}(\bar{\mu}) := \inf\left\{\int_{(\mathbb{R}^d)^{M+N}} \sum_{1 \leq i \neq j \leq M+N} \mathsf{c}(x_i, x_j) d\gamma_{N_A|N_B}(x_1, \ldots, x_{N_A+N_B}) \ \middle| \ \begin{array}{l} \gamma_{N_A|N_B} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B}), \\ \gamma_{N_A|N_B} \mapsto \bar{\mu} \end{array}\right\} .$$

By linearity of the marginal map $\gamma_{N_A|N_B} \mapsto \bar{\mu}$ given by $d\bar{\mu}(x) := \int d\gamma_{N_A|N_B}(x, x_2, \ldots, x_{M+N})$, we can generalize the above definition to marginals equal to general positive measures $\bar{\mu}$, such that for any $\alpha > 0$

$$F_{N_A|N_B,\mathsf{c}_A+\mathsf{c}_B}^{\mathrm{OT}}(\alpha\bar{\mu}) = \alpha F_{N_A|N_B,\mathsf{c}_A+\mathsf{c}_B}^{\mathrm{OT}}(\bar{\mu}) .$$

Now we connect these definitions to the unconstrained problem. If $\gamma_{M+N} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{M+N})$ is a symmetric transport plan with marginal μ_{M+N}, then we can decompose

$$\gamma_{M+N} = \sum_{\substack{N_A+N_B=M+N \\ N_A,N_B \in \mathbb{N}}} \gamma_{M+N}|_{B_{N_A,N_B}}, \quad \mu_{M+N} = \sum_{\substack{N_A+N_B=M+N \\ N_A,N_B \in \mathbb{N}}} \bar{\mu}_{N_A,N_B}, \tag{4.38a}$$

where $\bar{\mu}_{N_A,N_B}$ is the marginal of the restriction $\gamma_{M+N}|_{B_{N_A,N_B}}$:

$$\bar{\mu}_{N_A,N_B} := \int_{(\mathbb{R}^d)^{M+N-1}} d(\gamma_{M+N}|_{B_{N_A,N_B}})(\cdot, x_2, \ldots, x_{M+N}) . \tag{4.38b}$$

Here $\gamma_{M+N}|_{B_{N_A,N_B}}$ are not probability measures, but they belong to the set $\mathcal{M}^+_{sym}((\mathbb{R}^d)^{N_A|N_B}) := \{\alpha\gamma_{N_A|N_B} : \gamma_{N_A|N_B} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B}), \alpha > 0\}$ and similarly their marginals $\bar{\mu}_{N_A,N_B}$ belong to the set of measures $\mathcal{M}^+_{N_A|N_B}(\mathbb{R}^d) := \{\alpha\mu : \mu \in \mathcal{P}_{N_A|N_B}(\mathbb{R}^d), \alpha > 0\}$.

Step 2. We now claim that

$$F^{\mathrm{OT}}_{M+N,\mathsf{c}}\left(\tfrac{M\mu+N\nu}{M+N}\right)$$

$$= \min\left\{ \sum_{\substack{N_A+N_B=M+N\\N_A,N_B\in\mathbb{N}}} F^{\mathrm{OT}}_{N_A|N_B,\mathsf{c}}(\mu_{N_A,N_B}) \;\middle|\; \begin{array}{l} \mu_{N_A,N_B} \in \mathcal{M}^+_{N_A|N_B}(\mathbb{R}^d)\,,\\[4pt] \sum_{\substack{N_A,N_B\in\mathbb{N}\\N_A+N_B=M+N}} \mu_{N_A,N_B} = \tfrac{M\mu+N\nu}{M+N} \end{array} \right\}. \tag{4.39}$$

Indeed, the plans $\gamma_{M+N}|_{B_{N_A,N_B}}$ with marginals $\bar{\mu}_{N_A,N_B}$ as in the decomposition (4.38) form a possible decomposition as in the minimum from (4.39), and vice-versa, a set of minimizing measures μ_{N_A,N_B} on the right in (4.39) together with the optimal transport plans realizing $F^{\mathrm{OT}}_{N_A|N_B,\mathsf{c}}(\mu_{N_A,N_B})$ form by superposition a competitor to the minimization problem in (4.39). By the linearity of the map

$$\gamma_{M+N} \mapsto \int_{(\mathbb{R}^d)^{M+N}} \sum_{1\le i\ne j\le M+N} \mathsf{c}(x_i,x_j)d\gamma_{M+N}(x_1,\ldots,x_{M+N})\,,$$

we can then directly compare the costs, and conclude the proof of (4.39).

Step 3. If $\mu_{N_A,N_B} \in \mathcal{P}(\mathbb{R}^d)$ is the marginal of a plan $\gamma_{N_A|N_B} \in \mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B})$ with $N_A, N_B \ne 0$, then μ_{N_A,N_B} splits due to the definition (4.37) as

$$\mu_{N_A,N_B} = \frac{N_A}{M+N}\mu_{N_A} + \frac{N_B}{M+N}\mu_{N_B}, \tag{4.40}$$

where $\mu_{N_A} \in \mathcal{P}(A), \mu_{N_B} \in \mathcal{P}(B)$. We then claim that

$$F^{\mathrm{OT}}_{N_A|N_B,\mathsf{c}_A+\mathsf{c}_B}(\mu_{N_A,N_B}) = \begin{cases} F^{\mathrm{OT}}_{N_A,\mathsf{c}_A}(\mu_{N_A}) + F^{\mathrm{OT}}_{N_B,\mathsf{c}_B}(\mu_{N_B}), & \text{if}\quad N_A,N_B\in\{2,\ldots,M+N-2\}\\[4pt] F^{\mathrm{OT}}_{N_A,\mathsf{c}_A}(\mu_{N_A}), & \text{if}\quad N_B\in\{0,1\}\\[4pt] F^{\mathrm{OT}}_{N_B,\mathsf{c}_B}(\mu_{N_B}), & \text{if}\quad N_A\in\{0,1\}. \end{cases} \tag{4.41}$$

Indeed, define first the "space of precisely split $(M+N)$-points configurations" in $\mathbb{R}^d$ given by

$$\mathcal{C}_{N_A|N_B}(\mathbb{R}^d) := \left\{ C_{\vec{x}} := \{x_1,\ldots,x_{M+N}\} \subset \mathbb{R}^d : \#(C_{\vec{x}}\cap A) = N_A, \#(C_{\vec{x}}\cap B) = N_B \right\}. \tag{4.42}$$

Analogously to this, we define $\mathcal{C}_N(\mathbb{R}^d) := \{\{x_1,\ldots,x_N\} \subset \mathbb{R}^d\}$. Note that in all cases we consider the point configurations $\{x_1,\ldots,x_k\}$ as *multisets*.

Then, if under the isomorphism $\mathcal{P}_{sym}((\mathbb{R}^d)^{N_A|N_B}) \simeq \mathcal{P}(\mathcal{C}_{N_A|N_B}(\mathbb{R}^d))$ induced by the map

$$(\mathbb{R}^d)^{M+N} \ni \vec{x} := (x_1,\ldots,x_{M+N}) \mapsto \{x_1,\ldots,x_{M+N}\} =: C_{\vec{x}} \in \mathcal{C}_{M+N}(\mathbb{R}^d)\,,$$

we identify a plan $\gamma_{N_A|N_B}$ having marginal μ_{N_A,N_B}, to a measure $\bar{\gamma}_{N_A|N_B} \in \mathcal{P}(\mathcal{C}_{N_A|N_B}(\mathbb{R}^d))$, then we may take the operations $r_A : C_{\vec{x}} \mapsto C_{\vec{x}} \cap A$ and the similarly defined r_B and consider the pushforward

$$\gamma_{N_A} \simeq \bar{\gamma}_{N_A} := (r_A)_\# \bar{\gamma}_{N_A|N_B} \in \mathcal{P}(\mathcal{C}_{N_A}(\mathbb{R}^d)) \simeq \mathcal{P}_{sym}((\mathbb{R}^d)^{N_A})\,, \tag{4.43}$$

We next use the following lemma, which we prove later, and which translates the formula (4.43) more precisely. In particular, we can, and will, use (4.44) as an equivalent definition of γ_{N_A}, and we will not be further using (4.43) below.

Lemma 4.9. *The measure γ_{N_A} from (4.43) is given, in terms of duality with test functions $f \in C_b((\mathbb{R}^d)^{N_A})$, by the formula*

$$\int f(x_1, \ldots, x_{N_A}) d\gamma_{N_A}(x_1, \ldots, x_{N_A})$$

$$= \frac{1}{N_A! N_B!} \sum_{\sigma \in S_{M+N}} \int f\left(x_{\sigma(1)}, \ldots, x_{\sigma(N_A)}\right) \prod_{i=1}^{N_A} 1_A\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N}) . \quad (4.44)$$

Next, we will use (4.44) for a function $f(x_1)$ depending only on the first variable x_1. Note that in particular for any configuration $(x_1, \ldots, x_{M+N}) \in B_{N_A, N_B}$ there exist *precisely* $N_A! N_B!$ permutations $\sigma \in S_{M+N}$ such that $x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A$, which shows that γ_A is a probability measure. Then for each $\sigma \in S_{M+N}$ the set $\{\sigma(j) : j = 1, \ldots, M+N\}$ is in bijection with $\{1, \ldots, M+N\}$, and we have by the symmetry of $\gamma_{N_A|N_B}$ that

$$\frac{1}{N_A! N_B!} \sum_{\sigma \in S_{M+N}} \int f\left(x_{\sigma(1)}\right) \prod_{i=1}^{N_A} 1_A\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{M+N} \sum_{j=1}^{M+N} \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} f(x_{\sigma(j)}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} \frac{1}{M+N} \sum_{j=1}^{M+N} f(x_{\sigma(j)}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} \frac{1}{M+N} \sum_{j=1}^{M+N} f(x_j) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{M+N} \sum_{j=1}^{M+N} f(x_j) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N}) = \int f d\mu_{N_A, N_B}. \quad (4.45)$$

The above chain of equalities is then an expression of the marginal μ_{N_A, N_B} of $\gamma_{N_A|N_B}$ in duality with an arbitrary test function f. The marginal for $\mu|_A$ is given by using the above for the modified test function $f(x) 1_A(x)$. Using this choice instead of f, together with the self-evident fact that for each $(x_1, \ldots, x_{M+N}) \in B_{N_A, N_B}$, a cyclic permutation of the indices $1, \ldots, N_A$, induces a bijection on the set of permutations $\{\sigma \in S_{M+N} : x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A\}$, and we have from the second and last lines in (4.45)

$$\int f 1_A d\mu_{N_A, N_B} = \int \frac{1}{M+N} \sum_{j=1}^{M+N} \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} f(x_{\sigma(j)}) 1_A(x_{\sigma(j)}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} \frac{1}{M+N} \sum_{j=1}^{N_A} f(x_{\sigma(j)}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \frac{1}{N_A! N_B!} \sum_{\substack{\sigma \in S_{M+N}, \\ x_{\sigma(1)}, \ldots, x_{\sigma(N_A)} \in A}} \frac{N_A}{M+N} f(x_{\sigma(1)}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \frac{N_A}{M+N} \int f(x_1) d\gamma_{N_A}(x_1, \ldots, x_{N_A}) = \frac{N_A}{M+N} \int f d\mu_{N_A}. \quad (4.46)$$

This shows that $\mu_{N_A, N_B}|_A = \frac{N_A}{M+N} \mu_{N_A}$ and (together with the analogue statement for γ_B) we obtain that the marginals of γ_A, γ_B are μ_{N_A}, μ_{N_B}, as they appear in (4.40), respectively. We now find, using the support and

symmetry properties (4.36) of $\gamma_{N_A|N_B}$, that

$$\int f(x_1, \ldots, x_{M+N}) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int f(x_{\sigma(1)}, \ldots, x_{\sigma(M+N)}) \prod_{i=1}^{N_A} 1_A\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int f(x_{\sigma(1)}, \ldots, x_{\sigma(M+N)}) \prod_{i=N_A+1}^{N+M} 1_B\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N}). \qquad (4.47)$$

To prove (4.47), note that for any $(N + M)$-ple $(x_1, \ldots, x_{N+M})$ which belongs to any of the sets B_{N_A,N_B} from (4.36), there exists precisely one partition $J_A := \{i_1, \ldots, i_{N_A}\}, J_B := \{i_{N_A+1}, \ldots, i_{N+M}\}$ of $\{1, \ldots, N + M\}$ such that $x_i \in A$ if $i \in J_A$, and $x_i \in B$ if $i \in J_B$. Only the terms in the sum on the right in (4.47) corresponding to permutations $\sigma \in S_{M+N}$ which send $\{1, \ldots, N_A\}, \{N_A+1, \ldots, N+M\}$ into respectively J_A, J_B are nonzero, and the number of those σ is $N_A!N_B!$. Together with the invariance under symmetrization of $\gamma_{N_A|N_B}$, this proves (4.47). Note that, even if this is not directly evident from the formula, the roles of A and B for the right hand side of formula (4.47) are interchangeable.

Now by using (4.47) for the choice $f(x_1, \ldots, x_{M+N}) = \sum_{1 \leq i \neq j \leq M+N}(c_A(x_i, x_j) + c_B(x_i, x_j))$ and (4.44)

$$\int \sum_{1 \leq i \neq j \leq M+N} (c_A(x_i, x_j) + c_B(x_i, x_j)) \, d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int \sum_{1 \leq i \neq j \leq M+N} c_A(x_{\sigma(i)}, x_{\sigma(j)}) \prod_{i=1}^{N_A} 1_A\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$+ \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int \sum_{1 \leq i \neq j \leq M+N} c_B(x_{\sigma(i)}, x_{\sigma(j)}) \prod_{i=N_A+1}^{M+N} 1_B\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int \sum_{1 \leq i \neq j \leq N_A} c_A(x_{\sigma(i)}, x_{\sigma(j)}) \prod_{i=1}^{N_A} 1_A\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$+ \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} \int \sum_{N_A+1 \leq i \neq j \leq M+N} c_B(x_{\sigma(i)}, x_{\sigma(j)}) \prod_{i=N_A+1}^{M+N} 1_B\left(x_{\sigma(i)}\right) d\gamma_{N_A|N_B}(x_1, \ldots, x_{M+N})$$

$$= \int \sum_{1 \leq i \neq j \leq N_A} c_A(x_i, x_j) d\gamma_{N_A}(x_1, \ldots, x_{N_A}) + \int \sum_{1 \leq i \neq j \leq N_B} c_B(x_i, x_j) d\gamma_{N_B}(x_1, \ldots, x_{N_B}), \qquad (4.48)$$

where for the first equality we have applied the interchangeable versions of (4.47), as well as the symmetry under permutations of the above integrands, for the second equality the support properties (4.36) of c_A and c_B, and for third equality we used (4.44).

Then (4.48) and (4.42), directly prove (4.41).

Step 4. For the case of the split cost $c = c_A + c_B$ we claim that

$$F_{M+N, c_A+c_B}^{\mathrm{OT}}\left(\frac{M\mu + N\nu}{M+N}\right)$$

$$= \min\left\{\sum_{n=2}^{M+N} \alpha_n F_{n, c_A}^{\mathrm{OT}}(\mu_n) + \sum_{n=0}^{M+N-2} \alpha_n F_{M+N-n, c_B}^{\mathrm{OT}}(\nu_n) \left|\begin{array}{l} \sum_{n=1}^{M+N} n\alpha_n\mu_n = M\mu \,, \\[4pt] \sum_{n=0}^{M+N-1} (M+N-n)\alpha_n\nu_n = N\nu \,, \\[4pt] \alpha_n \geq 0, \mu_n \in \mathcal{P}(A), \nu_n \in \mathcal{P}(B) \end{array}\right.\right\}. \qquad (4.49)$$

We now prove (4.49). For each decomposition as in (4.39), namely if

$$\frac{M\mu + N\nu}{M+N} = \sum_{\substack{N_A+N_B=M+N \\ N_A,N_B \in \mathbb{N}}} \mu_{N_A,N_B}, \quad \mu_{N_A,N_B} \in \mathcal{M}_{N_A|N_B}^+(\mathbb{R}^d) \quad \text{for all} \quad N_A, N_B \,, \qquad (4.50)$$

27

and taking into account that there are $M + N + 1$ terms (N_A, N_B) with $N_A + N_B = M + N$, $N_A, N_B \in \mathbb{N}$, we can perform a reparameterization as follows:

$$\text{There exist} \begin{cases} \alpha_n \geq 0, \ \sum_{n=0}^{M+N} \alpha_n = 1 \, , \\[2mm] \mu_n \in \mathcal{P}(A), \nu_n \in \mathcal{P}(B) \end{cases} \text{such that} \begin{cases} N_A = n, \quad N_B = M + N - n \, , \\[2mm] \mu_{N_A, N_B} = \alpha_n \left(\frac{N_A}{M+N} \, \mu_n + \frac{N_B}{M+N} \, \nu_n \right) \, . \end{cases} \tag{4.51}$$

Assuming (4.51), we may apply to each μ_{N_A, N_B} equations (4.39) and (4.41) of the previous step together with the essentially disjoint supports hypothesis on $\mu, \nu,$, to obtain (4.49).

To prove (4.51), with the notation in (4.50), we first note that $N_A + N_B = M + N$ implies that if we write $N_A = n$ then $N_B = M + N - n$.

Next, we define $\alpha_{N_A} := \mu_{N_A, N_B}(A \cup B) \geq 0$, and by applying equation (4.50), we find $\sum_{n=0}^{M+N} \alpha_n = 1$. In case $\alpha_n \neq 0$, the measures μ_n, ν_n are obtained from the renormalized measure $(\alpha_n)^{-1} \mu_{N_A, N_B} \in \mathcal{P}_{N_A | N_B}(\mathbb{R}^d)$ like in (4.40) from the previous step, via the definition (4.37). Then the bottom equation on the right of (4.51) follows directly. This completes the proof of (4.51), and thus we completed the proof of (4.49) too.

Step 5. Now comparing the right hand side of (4.49) to the decompositions coming from definition (4.24a) of the $F_{\mathrm{GCB},\mathsf{c},N}$-problems, we see that $\{\alpha_n, \mu_n\}_{n \leq M+N}$ and $\{\alpha_n, \nu_n\}_{n \leq M+N}$ form competitors for the minimum problems $F_{\mathrm{GCB},\mathsf{c}_A,M}(\mu, M + N)$ and $F_{\mathrm{GCB},\mathsf{c}_B,N}(\nu, M + N)$, respectively. This allows to bound the left hand side of (4.49) from below by the sum of these GCB-problems, which concludes the proof. $\square$

Proof of Lemma 4.9: **Step 1.** *Notations.* We recall the notation $C_{\vec{x}} \overset{r_A}{\mapsto} C_{\vec{x}} \cap A$ for the restriction map and we recall that $\gamma_{N_A} = (r_A)_\# \gamma_{N_A | N_B}$. We further introduce a notation for the map that transforms an ordered $M + N$-ple into a multiset of cardinality $M + N$, which is defined by

$$i_{M+N} : (\mathbb{R}^d)^{M+N} \to \mathcal{C}_{M+N}(\mathbb{R}^d), \quad i_{M+N}(x_1, \ldots, x_{M+N}) = \{x_1, \ldots, x_{M+N}\}. \tag{4.52}$$

Then we use the notation

$$\bar{\gamma} := (i_{M+N})_\# \gamma, \quad \text{for} \quad \gamma \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^{M+N}). \tag{4.53}$$

We introduce the following equivalence relation

$$(x_1, \ldots, x_{M+N}) \sim (y_1, \ldots, y_{M+N}) \quad \text{if} \quad \exists \sigma \in S_{M+N}, \ \forall j = 1, \ldots, M+N, x_j = y_{\sigma(j)}. \tag{4.54}$$

We also denote $(y_{\sigma(1)}, \ldots, y_{\sigma(M+N)})$ by $\sigma \vec{y}$ for brevity. Then for the quotient under the above equivalence relation $(\mathbb{R}^d)^{M+N} / \sim$ has a measurable set of representatives $\Omega_{M+N} \subset (\mathbb{R}^d)^{M+N}$, i.e. we can find a measurable set Ω_{M+N} such that for every $\vec{x} \in (\mathbb{R}^d)^{M+N}$ there exists a unique $\vec{y} \in \Omega_{M+N}$ for which there exist $\sigma \in S_{M+N}$ such that $\vec{x} = \sigma \vec{y}$. We define the map

$$\rho_{M+N} : \mathcal{C}_{M+N} \to (\mathbb{R}^d)^{M+N}, \quad \rho_{M+N}(\{x_1, \ldots, x_{M+N}\}) := \vec{y} \in \Omega_{M+N} \text{ s.t. } \vec{x} \sim \vec{y}. \tag{4.55}$$

By the preceding discussion, the above choice of $\vec{y}$ is unique, and thus the map ρ_{M+N} is well defined, and furthermore it is measurable.

We next define the involution $\mathcal{P}((\mathbb{R}^d)^{M+N}) \to \mathcal{P}((\mathbb{R}^d)^{M+N})$ with image $\mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^{M+N})$, given by the symmetrization operation

$$\gamma \in \mathcal{P}((\mathbb{R}^d)^{M+N}) \mapsto \gamma^{\mathrm{sym}} \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^{M+N}), \tag{4.56}$$

given in duality with test functions $f \in C_b((\mathbb{R}^d)^{M+N})$ by the formula

$$\int f(\vec{x}) d\gamma^{\mathrm{sym}}(\vec{x}) := \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} f(\sigma \vec{x}) d\gamma(\vec{x}). \tag{4.57}$$

Step 2. *Claim:* There holds $(\rho_{M+N})_\# \bar{\gamma} \in \mathcal{P}(\Omega_{M+N}) \subset \mathcal{P}((\mathbb{R}^d)^{M+N})$, and for all $\gamma \in \mathcal{P}_{\mathrm{sym}}((\mathbb{R}^d)^{M+N})$ there holds, with ρ_{M+N} as defined in (4.55),

$$\left((\rho_{M+N})_\# \bar{\gamma} \right)^{\mathrm{sym}} = \gamma. \tag{4.58}$$

It is enough to verify that for every test function $f \in C_b((\mathbb{R}^d)^{M+N})$ the duality with the above two measures gives the same value. Indeed, by repeatedly using the formula $\forall f \in C_b, \int f dg_{\#}\mu = \int f \circ g d\mu$ for the pushforward of a measure and the definitions introduced above, we have:

$$
\int f(\vec{x}) d\left((\rho_{M+N})_{\#}\,\bar{\gamma}\right)^{\text{sym}}(\vec{x})
$$
$$
= \int \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} f(\sigma\vec{x}) d\left((\rho_{M+N})_{\#}\,\bar{\gamma}\right)(\vec{x})
$$
$$
= \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} \int f\left(\sigma\left(\rho_{M+N}\left(\{x_1,\dots,x_{M+N}\}\right)\right)\right) d\bar{\gamma}\left(\{x_1,\dots,x_{M+N}\}\right)
$$
$$
= \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} \int f\left(\sigma\left(\rho_{M+N} \circ i_{M+N}(\vec{x})\right)\right) d\gamma(\vec{x}) = \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} \int f(\sigma\vec{y}) d\gamma(\vec{x})
$$
$$
= \frac{1}{(M+N)!} \sum_{\sigma \in S_{M+N}} \int f(\sigma\vec{x}) d\gamma(\vec{x}) = \int f(\vec{x}) d\gamma(\vec{x}), \tag{4.59}
$$

where, to justify the above equalities, we observe that $\{x_1,\dots,x_{M+N}\}$ is S_{M+N}-invariant, then we observed that $\rho_{M+N} \circ i_{M+N}(\vec{x}) = \vec{y}$, the unique element in the S_{M+N}-orbit of $\vec{x}$ that lies in Ω_{M+N}, and thus $f \circ \rho_{M+N} \circ i_{M+N}$ is constant on this orbit, which is equal to the S_{M+N}-orbit of $\vec{y}$, and finally we used the fact that $\gamma = \gamma^{\text{sym}}$.

Step 3. *Conclusion of the proof.* We thus have proved (4.58), and this gives an inverse operation to the pushforward $(i_{M+N})_{\#} : \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^{M+N}) \to \mathcal{P}(\mathcal{C}_{M+N}(\mathbb{R}^d))$. We denote this inverse and characterize it as in (4.58), as follows

$$
I_{M+N} : \mathcal{P}(\mathcal{C}_{M+N}(\mathbb{R}^d)) \xrightarrow{\cong} \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^{M+N}), \quad I_{M+N}(\bar{\gamma}) := ((\rho_{M+N})_{\#}\bar{\gamma})^{\text{sym}} \quad \text{for} \quad \bar{\gamma} \in \mathcal{P}(\mathcal{C}_{M+N}(\mathbb{R}^d)).
$$

The link expressed in (4.43), between γ_{N_A} and $\gamma_{N_A|N_B}$ is thus given via the maps i_{M+N} and I_{M+N}. More precisely, (4.43) defines γ_{N_A} as
$$
\gamma_{N_A} := I_{N_A}\left((r_A)_{\#}(i_{M+N})_{\#}\gamma_{N_A|N_B}\right). \tag{4.60}
$$
We now prove that this γ_{N_A} satisfies (4.44). Indeed, starting from the definition (4.60) we have

$$
\int f(\vec{x}_{N_A}) d\gamma_A(\vec{x}_{N_A}) = \int f(\vec{x}_{N_A}) dI_{N_A}\left((r_A)_{\#}(i_{M+N})_{\#}\gamma_{N_A|N_B}\right)
$$
$$
= \int f(\vec{x}_{N_A}) d\left((\rho_{N_A})_{\#}(r_A)_{\#}(i_{M+N})_{\#}\gamma_{N_A|N_B}\right)^{\text{sym}}
$$
$$
= \frac{1}{N_A!} \sum_{\sigma \in S_{N_A}} \int f\left(\sigma\left(\rho_{N_A} \circ r_A \circ i_{M+N}(\vec{x}_{M+N})\right)\right) d\gamma(\vec{x}_{M+N})
$$
$$
= \frac{1}{N_A!} \sum_{\sigma \in S_{N_A}} \int f\left(\sigma\left(\rho_{N_A}(\{x_1,\dots,x_{M+N}\} \cap A)\right)\right) d\gamma(\vec{x}_{M+N}). \tag{4.61}
$$

Next, we claim that for each $\vec{x}_{M+N} \in B_{N_A, N_B}$ as defined in Step 1 of the proof of Lemma 4.8 (and thus for $\gamma_{N_A|N_B}$-almost every $\vec{x}_{M+N}$) there holds

$$
\frac{1}{N_A!} \sum_{\sigma \in S_{N_A}} f\left(\sigma\left(\rho_{N_A}(\{x_1,\dots,x_{M+N}\} \cap A)\right)\right) = \frac{1}{N_A!N_B!} \sum_{\sigma \in S_{M+N}} f(x_{\sigma(1)},\dots,x_{\sigma(N_A)}) \prod_{j=1}^{N_A} 1_A(x_{\sigma(j)}). \tag{4.62}
$$

Indeed, for a fixed $\vec{x}_{M+N} \in B_{N_A, N_B}$ let $i_1 < \dots < i_{N_A}$ be the indices such that $x_{i_j} \in A$. Then consider first the right hand side of (4.62). There are precisely $N_A!N_B!$ possible permutations $\sigma \in S_{M+N}$ which send the set $\{i_1,\dots,i_{N_A}\}$ into the set $\{1,\dots,N_A\}$, and they form precisely the set of terms on which the left-hand side in (4.62) is nonzero. This set of permutations can be written as $\sigma_0 S_{N_A}$, where $\sigma_0 \in S_{M+N}$ is any fixed

permutation such that $\sigma_0(k) = i_k$ for $1 \leq k \leq N_A$ and S_{N_A} are the permutations of $\{1, \ldots, N_A\}$ which leave fixed the remaining indices. This last subgroup acts transitively on the first N_A indices, so the right hand side of (4.62) equals, for $f : (\mathbb{R}^d)^{N_A} \to \mathbb{R}$ and with ρ_{N_A} defined similarly to (4.55),

$$\frac{1}{N_A!} \sum_{\sigma \in S_{N_A}} f(x_{i_{\sigma(1)}}, \ldots, x_{i_{\sigma(N_A)}}). \tag{4.63}$$

Now, for the left hand side of (4.62), note that $\rho_{N_A}(\{x_1, \ldots, x_{M+N}\} \cap A)$ is of the form $(x_{k_1}, \ldots, x_{k_{N_A}})$, and, by the transitivity action of S_{N_A} on the indices, the average on left hand side of (4.62) is identified to (4.63), as desired. As now (4.62) is proved, we may substitute its right hand side into (4.61), which then lends the right hand side of (4.43), proving the claim (4.44) of the lemma. $\qquad\square$

Lemma 4.8 allows to show the following useful result.

Lemma 4.10. *Let* $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$. *Consider* $\mu_1, \ldots, \mu_k \in \mathcal{P}(\mathbb{R}^d)$ *with densities* $\rho_1, \ldots, \rho_k$, *such that the integrals needed to define the quantities below are finite, such that the* μ_j *have essentially disjoint supports, i.e. that for every Borel set* A *and every* $1 \leq i \neq j \leq k$ *there holds* $\mu_i(A)\mu_j(A) = 0$. *Fix* $M_1, \ldots, M_k \in \mathbb{R}_{>0}$, *let* $\bar{M} := \sum_{i=1}^k M_i \in \mathbb{R}_{>0}$, *and denote by* $\mathsf{c}_{ii}(x, y) := 1_{\Lambda_i}(x)1_{\Lambda_i}(y)\mathsf{c}(x, y)$ *for* $i = 1, \ldots, k$, *and* $x, y \in \mathbb{R}^d$. *Let* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *have density* $\sum_{i=1}^k M_i\rho_i/\bar{M}$. *Set* $M' \in \mathbb{N}_+, M' \geq \max\{\bar{M}, 1\}$. *Then*

$$F^{\mathrm{OT}}_{\mathrm{GCB},\bar{M},\sum_{i=1}^k \mathsf{c}_{ii}} (\mu, M') \geq \sum_{i=1}^k F^{\mathrm{OT}}_{\mathrm{GCB},M_i,\mathsf{c}_{ii}} (\mu_i, M') = \sum_{i=1}^k F^{\mathrm{OT}}_{\mathrm{GCB},M_i,\mathsf{c}} (\mu_i, M') \ . \tag{4.64}$$

and

$$F^{\mathrm{OT}}_{\mathrm{GC},\bar{M},\sum_{i=1}^k \mathsf{c}_{ii}} (\mu) = \sum_{i=1}^k F^{\mathrm{OT}}_{\mathrm{GC},M_i,\mathsf{c}_{ii}} (\mu_i) = \sum_{i=1}^k F^{\mathrm{OT}}_{\mathrm{GC},M_i,\mathsf{c}} (\mu_i) \ . \tag{4.65}$$

Proof. Note that the thesis is trivially valid if $M' < 2$ due to the convention that $F^{\mathrm{OT}}_N = 0$ for $N = 0, 1$, and thus we assume from now on that $M' \geq 2$. We prove the lemma for the case of a decomposition into two measures, and by applying the lemma repeatedly we can extend it to the case of k measures.

By definition of $F^{\mathrm{OT}}_{\mathrm{GCB},\bar{M},\mathsf{c}_{11}+\mathsf{c}_{22}}(\mu, M')$, there exist $\alpha_n \geq 0$ and $\mu_n \in \mathcal{P}(\mathbb{R}^d)$, with $0 \leq n \leq M'$, with

$$\sum_{n=0}^{M'} \alpha_n = 1, \quad \sum_{n=1}^{M'} n\alpha_n\mu_n = \bar{M}\mu, \quad F^{\mathrm{OT}}_{\mathrm{GCB},\bar{M},\mathsf{c}_{11}+\mathsf{c}_{22}}(\mu, M') = \sum_{n=2}^{M'} \alpha_n F^{\mathrm{OT}}_{n,\mathsf{c}_{11}+\mathsf{c}_{22}}(\mu_n) \ . \tag{4.66}$$

By Lemma 4.8 for each $n \in \{1, \ldots, M'\}$, there exist positive real numbers $M_n^1, M_n^2 > 0$, and measures $\mu_n^1 \in \mathcal{P}(\Lambda_1), \mu_n^2 \in \mathcal{P}(\Lambda_2)$, such that

$$M_n^1 + M_n^2 = n, \quad n\mu_n = M_n^1\mu_n^1 + M_n^2\mu_n^2, \quad F^{\mathrm{OT}}_{n,\mathsf{c}_{11}+\mathsf{c}_{22}}(\mu_n) \geq F^{\mathrm{OT}}_{\mathrm{GCB},M_n^1,\mathsf{c}} (\mu_n^1, n) + F^{\mathrm{OT}}_{\mathrm{GCB},M_n^2,\mathsf{c}} (\mu_n^2, n) \ . \tag{4.67}$$

Let $\alpha_{n,m}^1, \alpha_{n,m}^2 \geq 0$, and $\mu_{n,m}^1 \in \mathcal{P}(\Lambda_1), \mu_{n,m}^2 \in \mathcal{P}(\Lambda_2), m \in \{0, \ldots, M'\}$, be such that for $j = 1, 2$

$$F^{\mathrm{OT}}_{\mathrm{GCB},M_n^j,\mathsf{c}}(\mu_n^j, n) = \sum_{m=2}^n \alpha_{n,m}^j F^{\mathrm{OT}}_{m,\mathsf{c}}(\mu_{n,m}^j), \quad \sum_{m=0}^n \alpha_{n,m}^j = 1, \quad \sum_{m=1}^n m\alpha_{n,m}^j\mu_{n,m}^j = M_n^j\mu_n^j \ . \tag{4.68}$$

Up to adding extra coefficients $\alpha_{n,n+1}^j = \cdots = \alpha_{n,M'}^j = 0$, we may extend the sums in (4.68) up to $m = M'$. Now, we sum (4.68) multiplied by the coefficients from (4.66), and we find from (4.66), (4.67) and (4.68)

$$\sum_{m=0}^{M'} \sum_{n=0}^{M'} \alpha_n\alpha_{n,m}^1 = 1, \quad \sum_{m=1}^{M'} m \sum_{n=1}^{M'} \alpha_n\alpha_{n,m}^1\mu_{n,m}^1 = \sum_{n=1}^{M'} \alpha_n M_n^1\mu_n^1 = \bar{M}\mu|_{\Lambda_1} = M_1\mu_1 \ ,$$

$$\sum_{m=0}^{M'} \sum_{n=0}^{M'} \alpha_n\alpha_{n,m}^2 = 1, \quad \sum_{m=1}^{M'} m \sum_{n=1}^{M'} \alpha_n\alpha_{n,m}^2\mu_{n,m}^2 = \sum_{n=1}^{M'} \alpha_n M_n^2\mu_n^2 = \bar{M}\mu|_{\Lambda_2} = M_2\mu_2 \ ,$$

$$\text{and} \quad F^{\mathrm{OT}}_{\mathrm{GCB},N,\mathsf{c}_{11}+\mathsf{c}_{22}}(\mu, M') \geq \sum_{m=2}^{M'} \sum_{n=2}^{M'} \alpha_n\alpha_{n,m}^1 F^{\mathrm{OT}}_{m,\mathsf{c}}(\mu_{n,m}^1) + \sum_{m=2}^{M'} \sum_{n=2}^{M'} \alpha_n\alpha_{n,m}^2 F^{\mathrm{OT}}_{m,\mathsf{c}}(\mu_{n,m}^2) \ . \tag{4.69}$$

By defining $\beta_m^1 = \sum_{n=0}^{M'} \alpha_n \alpha_{n,m}^1$, $\beta_m^2 = \sum_{n=0}^{M'} \alpha_n \alpha_{n,m}^2$, and by using the inequalities below and their analogues for μ_2 (where the inequality below holds by the Monge-Kantorovich duality) for $\beta_m^1 > 0$

$$\frac{\sum_{n=2}^{M'} \alpha_n \alpha_{n,m}^1 F_{m,\mathsf{c}}^{\mathrm{OT}}(\mu_{n,m}^1)}{\sum_{n=0}^{M'} \alpha_n \alpha_{n,m}^1} = \frac{1}{\beta_m^1} \sum_{n=2}^{M'} \alpha_n \alpha_{n,m}^1 F_{m,\mathsf{c}}^{\mathrm{OT}}(\mu_{n,m}^1) \geq F_{m,\mathsf{c}}^{\mathrm{OT}}\left(\frac{1}{\beta_m^1} \sum_{n=2}^{M'} \alpha_n \alpha_{n,m}^1 \mu_{n,m}^1\right), \tag{4.70}$$

we find that the measures $(\beta_m^1)^{-1} \sum_{n=2}^{M'} \alpha_n \alpha_{n,m}^1 \mu_{n,m}^1$ and coefficients β_m^1 form a competitor for the definition of $F_{\mathrm{GCB},M_1,\mathsf{c}}^{\mathrm{OT}}(\mu_1, M')$, due to the first line in (4.69), and a similar construction and statement apply also for μ_2. Thereby (4.70) together with (4.69) and (4.35) prove the first statement of the lemma.

The inequality $\geq$ in (4.65) follows by similar arguments, hence its proof will be omitted. The inequality $\leq$ in (4.65) follows from (4.27). $\qquad\square$

As an immediate consequence of Lemma 4.8 and 4.10, we have the next corollary.

Corollary 4.11. *Let* $\mathsf{c} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{>0} \cup \{+\infty\}$, *and set* $N \in \mathbb{N}_+, N \geq 2$, *and* $\bar{N} \in \mathbb{R}_{>0}$. *For* $k \in \mathbb{N}_+$, *consider the probability measures* $\mu_1, \ldots, \mu_k$, *with densities* $\rho_1, \ldots, \rho_k$, *such that the integrals required to define the quantities below are finite, and the measures* μ_j *have essentially disjoint supports, i.e that for all Borel sets* A *and all* $1 \leq i \neq j \leq k$ *there holds* $\mu_i(A)\mu_j(A) = 0$. *Let* $\alpha_j \in \mathbb{R}_{>0}, j = 1, \ldots, k$, *be such that* $\sum_{j=1}^k \alpha_j = 1$. *Let* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *have density* $\rho = \sum_{j=1}^k \alpha_j \rho_j$. *Denote by* $\mathsf{c}_{ii} := \mathsf{c}1_{\Lambda_i \times \Lambda_i}, i = 1, \ldots, k$. *Then*

$$E_{N,\sum_{i=1}^k \mathsf{c}_{ii}}^{\mathrm{xc}}\left(\sum_{j=1}^k \alpha_j \mu_j\right) \geq \sum_{j=1}^k E_{\mathrm{GC},N\alpha_j,\mathsf{c}}^{\mathrm{xc}}(\mu_j) \quad and \quad E_{\mathrm{GC},\bar{N},\sum_{i=1}^k \mathsf{c}_{ii}}^{\mathrm{xc}}\left(\sum_{j=1}^k \alpha_j \mu_j\right) \geq \sum_{j=1}^k E_{\mathrm{GC},\bar{N}\alpha_j,\mathsf{c}}^{\mathrm{xc}}(\mu_j). \tag{4.71}$$

We note here that in [LLS17, Thm. 3.1, Cor. 3.4] also the result analogous to Proposition 2.5 for $E_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{xc}}$ was proved. More precisely, as a consequence of [LLS17, Thm. 3.1], for any measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that $d\mu(x) = \rho(x)dx$ with $\rho(x) = |\Lambda|^{-1}1_\Lambda(x)$ and Λ a Borel set with ϕ-regular boundary, we have

$$\lim_{N\to\infty,N\in\mathbb{N}} N^{-1-s/d} E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) = \lim_{N\to\infty} N^{-1-s/d} E_{N,s}^{\mathrm{xc}}(\mu) = -C(s,d)\int_{\mathbb{R}^d} \rho^{1+s/d} = -C(s,d)|\Lambda|^{-s/d}. \tag{4.72}$$

In the above cited result from [LLS17], (4.72) is stated with $N \in \mathbb{N}$. In the next lemma we extend it to the case $N \in \mathbb{R}_{>0}$.

Lemma 4.12. *For any* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *with density* ρ *of form* $\rho(x) = |\Lambda|^{-1}1_\Lambda(x)$, *where* Λ *is a Borel set with* ϕ-*regular boundary, it holds that*

$$\lim_{N\to\infty,N\in\mathbb{R}_{>0}} N^{-1-s/d} E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) = -C(s,d)\int_{\mathbb{R}^d} \rho^{1+s/d} = -C(s,d)|\Lambda|^{-s/d}. \tag{4.73}$$

Proof. To show (4.73), fix a large non-integer $N \in \mathbb{R}_{>0}$. We have from Remark 4.7 2) and 4) that

$$\begin{aligned} E_{\mathrm{GC},[N]+2,s}^{\mathrm{xc}}(\mu) &\leq E_{\mathrm{GC},[N]+2,s}^{\mathrm{xc}}(\mu) - E_{\mathrm{GC},[N]+2-N,s}^{\mathrm{xc}}(\mu) \\ &\leq E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \leq E_{\mathrm{GC},[N]-1,s}^{\mathrm{xc}}(\mu) + E_{\mathrm{GC},N+1-[N],s}^{\mathrm{xc}}(\mu) \leq E_{\mathrm{GC},[N]-1,s}^{\mathrm{xc}}(\mu). \end{aligned} \tag{4.74}$$

Combining now (4.72) and (4.74) produces (4.73). $\qquad\square$

4.2.2 Proof of the lower bound

Before we proceed, we remind the reader that the formula we will be working with below, as explained in (4.7), (4.20), (4.21), (A.21a), (A.21b) and (A.21c), is

$$\frac{1}{|x-y|^s} = \frac{M}{M+C}\left\{\int_{\Omega_l}\left(\sum_{A\in F_\omega^l} \frac{1_A(x)1_A(y)}{|x-y|^s}\right)d\mathbb{P}_l(\omega) + w(x-y)\right\}, \tag{4.75}$$

where $l > 0$, and where w satisfies the properties in Proposition 1.6 above.

Denote for each $\omega \in \Omega_l$ by

$$\Lambda_{\mathrm{nc}}(l, \omega) := \Lambda \setminus \bigcup_{A \in F_\omega^l} (\Lambda \cap A) . \tag{4.76a}$$

This is the part left not covered in the packing of $\Lambda \subset \mathbb{R}^d$ by balls given by the Swiss cheese Lemma 4.2. Moreover, let

$$\Lambda_{\mathrm{err}}(l, \omega) := \bigcup \left\{ A \in F_\omega^l : \ A \cap \partial \Lambda \neq \varnothing \right\} . \tag{4.76b}$$

The (4.76b) will be substituted in the proof of Proposition 4.1, where $\Lambda = \cup_{i=1}^k \Lambda_i$, by

$$\Lambda_{\mathrm{err}}(l, \omega) := \bigcup \left\{ A \in F_\omega^l : \ A \cap \bigcup_{i=1}^k \partial \Lambda_i \neq \varnothing \right\} , \tag{4.77}$$

for which the proof of Lemma 4.14 below also works. From (4.76b), it immediately follows that

$$\Lambda \setminus (\Lambda_{\mathrm{err}}(l, \omega) \cup \Lambda_{\mathrm{nc}}(l, \omega)) = \bigcup \left\{ A \in F_\omega^l : \ A \subset \Lambda \right\} , \tag{4.78}$$

and from (4.77) that

$$\Lambda \setminus (\Lambda_{\mathrm{err}}(l, \omega) \cup \Lambda_{\mathrm{nc}}(l, \omega)) = \bigcup \left\{ A \in F_\omega^l : \ A \subset \Lambda_i, \quad \text{for some} \quad i = 1, \ldots, k \right\} . \tag{4.79}$$

We will next state a result involving the uniform decay in ω of $|\Lambda_{\mathrm{nc}}|, |\Lambda_{\mathrm{err}}|$ as $M \to \infty$.

Lemma 4.13. *If Λ has finite volume, then for M, l as in Lemma 4.2 and for Ω_l as in (4.21) above, for each $\omega \in \Omega_l$, we have*

$$\sup_{\omega \in \Omega_l} |\Lambda_{\mathrm{nc}}(l, \omega)| \leq \frac{C_d}{M} |(\Lambda)_{2l\sqrt{d}}|, \quad \text{where} \quad (\Lambda)_r := \{x \in \mathbb{R}^d : \ \mathrm{dist}(x, \Lambda) \leq r\}, \tag{4.80}$$

and if Λ has ϕ-regular boundary in the sense of (2.8) for a continuous $\phi : [0, t_0) \to \mathbb{R}^+$ with $\phi(0) = 0$ then

$$\sup_{\omega \in \Omega_l} |\Lambda_{\mathrm{err}}(l, \omega)| \leq \phi(\ell |\Lambda|^{-1/d}) |\Lambda|. \tag{4.81}$$

Proof. We observe first that by the property of having ϕ-regular boundary from (2.8), the total volume of cubes $A \in F_\omega^l$ which touch $\partial \Lambda$ for the case (4.76b) (respectively $\partial \Lambda_i$ for the case (4.77)), is smaller than

$$|\{x : \ d(x, \partial \Lambda) \leq (1 + \eta)\ell\}| \leq \phi(\ell |\Lambda|^{-1/d}) |\Lambda| \quad \text{and} \quad |\{x : \ d(x, \partial \Lambda_i) \leq (1 + \eta)\ell\}| \leq \phi(\ell (\min_{1 \leq i \leq k} |\Lambda_i|)^{-1/d}) |\Lambda|,$$

as for all $\omega \in \Omega_l$ these sets are included in cubes of size at most $(1 + \eta)l$, due to (4.21). This proves (4.81).

Applying Lemma 4.2, we find that uniformly among $\omega \in \Omega_l$ of the form (4.21), we can estimate for the case (4.76b) the contribution of $\Lambda_{\mathrm{nc}}(l, \omega)$ as being at most $C_d M^{-1} |K|$ restricted to each cube $K = K_\omega$ that meets Λ, and taking among the cubes from the cheese lemma coverings used for $\omega \in \Omega_l$. Each such cube K_ω remains within the neighborhood $(\Lambda)_{2l\sqrt{d}}$ of thickness $2l\sqrt{d}$ of Λ, as the diameter of K is $tl\sqrt{d} \leq 2l\sqrt{d}$, since $t \in [1 - \eta, 1 + \eta]$. Summing all the contributions of all such cubes gives (4.80). $\qquad \square$

We introduce here the following normalization notation and convention for measures, which will be used at several instances below. If μ is a Borel measure, $A \subset \mathbb{R}^d$ is a Borel set and $\mu(A) > 0$, then we denote

$$\hat{\mu}_A := \frac{\mu|_A}{\mu(A)} , \tag{4.82}$$

and if $\mu(A) = 0$ we set $\hat{\mu}_A = 0$. Also note that automatically in this case there holds $\mu|_A = \mu(A) \hat{\mu}_A$, which will also be used in several instances below.

Next, we will show

Lemma 4.14. *Fix $0 < \epsilon < d/2$, set $\epsilon \leq s \leq d - \epsilon$ and let $\mathsf{c}(x,y) = |x-y|^{-s}$. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a probability measure with density $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$ supported on a Borel set $\Lambda \subset \mathbb{R}^d$ with ϕ-regular boundary. Set $N \in \mathbb{N}, N \geq 2$. Let $l > 0$ and $M > 0$ be as in Lemma 4.2. Then there holds*

$$
\begin{aligned}
E_{N,s}^{\mathrm{xc}}(\mu) \;\geq\;& E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \\
\geq\;& \frac{M}{M+C}\Bigg\{ \int_{\Omega_l}\Bigg(\sum_{\substack{A \in F_\omega^l \\ A \subset \Lambda, \mu(A) > 0}} E_{\mathrm{GC},N\mu(A),s}^{\mathrm{xc}}(\hat{\mu}_A) + E_{\mathrm{GC},N\mu(\Lambda_{\mathrm{err}}(l,\omega)),s}^{\mathrm{xc}}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \Bigg) d\,\mathbb{P}_l(\omega) \\
& - \frac{C(w,d,\epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M} R_1^{-s}(N-1) \Bigg\}\,,
\end{aligned} \tag{4.83}
$$

for some $C, C(w,d,\epsilon)$, which depend only on ϵ and d. Furthermore, the second inequality in (4.83) holds also for $N \in \mathbb{R}_+$. To state (4.83) for the case (4.79), the summation under the first integral in (4.83) is taken over $A \subset \Lambda_i$ and over $i = 1,\ldots,k$.

Proof. **Step 1.** We will show here the following inequality, independent of the specific properties of μ:

$$
E_{N,s}^{\mathrm{xc}}(\mu) \geq E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \geq E_{N,\mathsf{c}-\frac{M}{M+C}w}^{\mathrm{xc}}(\mu) - \frac{C(w,d,\epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M} R_1^{-s}(N-1)\,, \tag{4.84}
$$

where $C, C(w,d,\epsilon) > 0$ are the constants from Proposition 1.6, and where from (4.75)

$$
\mathsf{c} - \frac{M}{M+C}w = \frac{M}{M+C}\int_{\Omega_l}\Bigg(\sum_{A \in F_\omega^l} \frac{\mathbf{1}_A(x)\mathbf{1}_A(y)}{|x-y|^s} \Bigg) d\,\mathbb{P}_l(\omega). \tag{4.85}
$$

The proof of (4.84) follows immediately from the definitions of $E_{N,s}^{\mathrm{xc}}(\mu)$ and $E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)$, by making use of Proposition 1.6 and of the re-expression $\mathsf{c}(x,y) = \mathsf{c}(x,y) - \frac{M}{M+C}w(x-y) + \frac{M}{M+C}w(x-y)$.

Step 2. We will show here that

$$
\begin{aligned}
E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \geq\;& \frac{M}{M+C}\Bigg\{ \int_{\Omega_l}\Bigg(\sum_{\substack{A \in F_\omega^l \\ A \subset \Lambda, \mu(A) > 0}} E_{\mathrm{GC},N\mu(A),s}^{\mathrm{xc}}(\hat{\mu}_A) + E_{\mathrm{GC},N\mu(\Lambda_{\mathrm{err}}(l,\omega)),s}^{\mathrm{xc}}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \Bigg) d\,\mathbb{P}_l(\omega) \\
& - \frac{C(w,d,\epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M} R_1^{-s}(N-1) \Bigg\}\,.
\end{aligned} \tag{4.86}
$$

To start the proof, we observe that $\mu(\Lambda \setminus (\Lambda_{\mathrm{err}}(l,\omega) \cup \Lambda_{\mathrm{nc}}(l,\omega))) > 0$ holds whenever there exists at least one $A \in \mathcal{F}_\omega^l$ such that $A \subset \Lambda \setminus (\Lambda_{\mathrm{err}}(l,\omega) \cup \Lambda_{\mathrm{nc}}(l,\omega))$ and $\mu(A) > 0$. Denote for all $A \in F_\omega^l$ by

$$
\mathsf{c}_A(x,y) := \frac{\mathbf{1}_A(x)\mathbf{1}_A(y)}{|x-y|^s}\,.
$$

With this notation, we have in view of (4.85)

$$
F_{N,\mathsf{c}-\frac{M}{M+C}w}^{\mathrm{OT}}(\mu) \geq \frac{M}{M+C}\int_{\Omega_l} F_{n,\sum_{A\in F_\omega^l}\mathsf{c}_A}^{\mathrm{OT}}(\mu)\, d\,\mathbb{P}_l(\omega)\,, \tag{4.87}
$$

where for the inequality we interchanged the integration order, took the minimum inside the integral, and used the definition of $F_{n,\sum_{A\in F_\omega^l}\mathsf{c}_A}^{\mathrm{OT}}(\mu)$.

Directing next our attention briefly to the mean field term, we get

$$
\begin{aligned}
\int_{\mathbb{R}^{2d}}\Bigg(\mathsf{c}(x,y) - \frac{M}{M+C}w(x,y)\Bigg)\, d\mu(x)d\mu(y)\, d\,\mathbb{P}_l(\omega) & \\
= \frac{M}{M+C}\int_{\Omega_l}\int_{\mathbb{R}^{2d}}& \sum_{A\in F_\omega^l}\mathsf{c}_A(x,y)d\mu(x)d\mu(y)\, d\,\mathbb{P}_l(\omega)\,. \tag{4.88}
\end{aligned}
$$

From (4.87) and (4.88), we obtain in (4.84)

$$
\begin{aligned}
&E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \\
&\geq \frac{M}{M+C}\left(\int_{\Omega_l} E_{\mathrm{GC},N,\sum_{A\in F_\omega^l} \mathsf{c}_A}^{\mathrm{xc}}(\mu)\, d\,\mathbb{P}_l(\omega) - \frac{C(w,d,\epsilon)}{M} N^{1+s/d}\int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M} R_1^{-s}(N-1)\right).
\end{aligned}
$$
(4.89)

Furthermore

$$
\mu = \mu|_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))} + \mu|_{\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega)}\ ,
$$

With the notation (4.82), we now have on the right-hand side of (4.89) by Corollary 4.11 (using the notations $E_{\mathrm{GC}}^{\mathrm{xc}}[N,\mathsf{c}](\mu) := E_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{xc}}(\mu)$, in order to make the formulas easier to read):

$$
\begin{aligned}
E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N,\sum_{A\in F_\omega^l}\mathsf{c}_A\Big](\mu) &\geq E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega)),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda_{\mathrm{err}}(l,\omega)}}\mathsf{c}_A\Big](\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega)}) \\
&\quad +E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda}}\mathsf{c}_A\Big]\big(\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}\big)\ .
\end{aligned}
$$
(4.90)

By means of (4.65) we get for the first term on the right hand side of (4.90),

$$
E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega)),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda_{\mathrm{err}}(l,\omega)}}\mathsf{c}_A\Big]\big(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega)}\big) = E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda_{\mathrm{err}}(l,\omega)),\mathsf{c}\Big](\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)})\ ,
$$

which integrated against $d\mathbb{P}_\ell(\omega)$ gives the second term in (4.86). The simplification in the last equation appears because the cost $\sum_{A\in F_\omega^l} c_A$ is zero on $\Lambda_{\mathrm{nc}}(l,\omega)$ due to (4.76a).

By notation (4.82) and via (4.78) (respectively (4.79)), we get

$$
\begin{aligned}
\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))} &= \sum_{\substack{A\in F_\omega^l\\A\subset\Lambda,\mu(A)>0}}\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}(A)\frac{\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}\big|_A}{\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}(A)} \\
&= \sum_{\substack{A\in F_\omega^l\\A\subset\Lambda,\mu(A)>0}}\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}(A)\,\hat{\mu}_A.
\end{aligned}
$$

For the second term on the r.h.s. in (4.90) we have

$$
\begin{aligned}
&E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda}}\mathsf{c}_A\Big]\big(\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}\big) \\
&= E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda}}\mathsf{c}_A\Big]\Big(\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda,\mu(A)>0}}\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}(A)\hat{\mu}_A\Big).
\end{aligned}
$$
(4.91)

We now obtain on the right-hand side of (4.91) via another application of Corollary 4.11

$$
\begin{aligned}
&E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))),\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda}}\mathsf{c}_A\Big]\Big(\sum_{\substack{A\in F_\omega^l\\A\subset\Lambda,\mu(A)>0}}\hat{\mu}_{\Lambda\setminus(\Lambda_{\mathrm{err}}(l,\omega)\cup\Lambda_{\mathrm{nc}}(l,\omega))}(A)\hat{\mu}_A\Big) \\
&\hspace{9cm}\geq \sum_{\substack{A\in F_\omega^l\\A\subset\Lambda,\mu(A)>0}}E_{\mathrm{GC}}^{\mathrm{xc}}\Big[N\mu(A),\mathsf{c}\Big](\hat{\mu}_A)\ .
\end{aligned}
$$

34

Integrated against $d\mathbb{P}_\ell(\omega)$, this gives the first term in (4.86) and concludes the proof of Step 2, for our choice of $\mathsf{c}(x,y) = |x-y|^{-s}$.

$\square$

We are now ready to prove Proposition 4.1 concerning our sharp lower bound.

Proof of Proposition 4.1: Assume the quantities below are as in Lemma 4.2. Assume also that $[-\ell/2, \ell/2]^d \subset \Lambda_i, \ell \ll |\Lambda_i|, i = 1, \ldots, k$ (or else we can re-scale the μ_i in view of (2.7)). By Lemma 4.14 we have

$$E_{\text{GC},N,s}^{\text{xc}}(\mu) \geq \frac{M}{M+C}\Bigg\{\sum_{i=1}^k \int_{\Omega_l} \Bigg(\sum_{\substack{A \in F_\omega^l \\ A \subset \Lambda_i \\ \mu(A) > 0}} E_{\text{GC},N\mu(A),s}^{\text{xc}}(\hat{\mu}_A) + E_{\text{GC},N\mu(\Lambda_{\text{err}}(l,\omega)),s}^{\text{xc}}(\hat{\mu}_{\Lambda_{\text{err}}(l,\omega)})\Bigg) d\mathbb{P}_l(\omega)$$
$$- \frac{C(w,d,\epsilon)}{M}N^{1+s/d}\int_{\mathbb{R}^d}\rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M}R_1^{-s}(N-1)\Bigg\}. \quad (4.92)$$

In the next two steps we will calculate separately each of the terms $E_{\text{GC},N\mu(\Lambda_{\text{err}}(l,\omega)),s}^{\text{xc}}(\hat{\mu}_{\Lambda_{\text{err}}(l,\omega)})$ and $E_{\text{GC},N\mu(A),s}^{\text{xc}}(\hat{\mu}_A)$ appearing in (4.92). In order to prove that the boundary terms have a small total volume, we will need to work in the regime $l \ll 1, M \gg 1$. Lemma 4.2 applies in particular if $R_M > (1 + 4\sqrt{d}|B_1^d|)^M R_1$, and $l > 8\sqrt{d}|B_1|(M + C_d)R_M$. Thus, for Lemma 4.2 to apply, the extra constraint linking M, l, R_1, which can be formulated in two equivalent ways:

$$l > C(M+C)R_M > C(M+C)C^M R_1 \quad \Leftrightarrow \quad \log\frac{l}{R_1} > \log C + \log(M+C) + M\log C, \quad (4.93)$$

where $C := \max\{1 + 4\sqrt{d}|B_1^d|, 8\sqrt{d}|B_1|, C_d\}$ depends only on d. If R_1, l, M are such that

$$M < \frac{\log(l/R_1)}{3\log C} \quad \Leftrightarrow \quad R_1 < C^{-3M} l, \quad (4.94)$$

then there exists $M_d > 0$ depending only on d such that (4.93) holds for all $M \geq M_d$. Indeed, note that if

$$M \geq \max\left\{1, \frac{\log(M+C)}{\log C}\right\}, \quad (4.95)$$

then $3M\log C$ is larger than the right hand side of the second equation in (4.93), and as a consequence (4.94) implies (4.93) for such M. It suffices then to take M_d to be the smallest value of $M \geq 1$ such that (4.95) holds. It is easy to verify that for any $M \geq M_d$ (4.95) also holds, that the value M_d depends only on d because C above depends only on d.

To find l, M, R_1 satisfying condition (4.94) and also such that $l \ll 1 \ll M$, it suffices to fix $R_1 > 0$ separately for each choice of l, M and small enough. Note that the above choices must be performed depending on N as well, so as to optimize our asymptotic estimates.

A suitable choice turns out to be as follows. We take for $N \geq C^{18d(M_d+1)}$

$$R_1 := N^{-\frac{1}{2d}}, \quad l := N^{-\frac{1}{3d}}, \quad M := \left[\frac{\log N}{18\, d\log C}\right] - 1, \quad (4.96)$$

For clarity of exposition, we will substitute the choice (4.96) for l, R_1 only at the very end in our estimates.

Step 1. We will show here that

$$\liminf_{N\to\infty} N^{-1-s/d} E_{\text{GC},N,s}^{\text{xc}}(\mu) \geq \liminf_{N\to\infty} N^{-1-s/d}\sum_{i=1}^k \int_{\Omega_l}\Bigg(\sum_{\substack{A \in F_\omega^l \\ A \subset \Lambda_i}} E_{\text{GC},N\alpha_i|A|/|\Lambda_i|,s}^{\text{xc}}(\hat{\mu}_A)\Bigg) d\mathbb{P}_l(\omega). \quad (4.97)$$

Note that, since $l \ll 1$, and since all sets $A \in F^l_\omega$ are disjoint balls of scales $R_j < l$, for given $A \in F^l_\omega$, either the set A intersects the boundaries $\partial \Lambda_i, i = 1, \ldots k$, or it is included exactly in one of the $\Lambda_i, i = 1, \ldots, k$. Thus, from (4.92) and in view of the properties $M \gg 1$ and $R_1 \gg N^{-\frac{1}{d}}$ following from (4.96), we find

$$\liminf_{N \to \infty} N^{-1-s/d} E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu) \geq \liminf_{N \to \infty} N^{-1-s/d} \sum_{i=1}^{k} \int_{\Omega_l} \left(\sum_{\substack{A \in F^l_\omega \\ A \subset \Lambda_i}} E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda \cap A),s}(\hat{\mu}_A) \right) d\,\mathbb{P}_l(\omega)$$

$$+ \liminf_{N \to \infty} N^{-1-s/d} \int_{\Omega_l} E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda_{\mathrm{err}}(l,\omega)),s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \, d\,\mathbb{P}_l(\omega) \ . \quad (4.98)$$

We next need to consider separately the two terms on the r.h.s in (4.98). We observe first that $\mu(A) = \alpha_i \frac{|A|}{|\Lambda_i|}$ if $A \subset \Lambda_i$ for some $i = 1, \ldots, k$. Therefore, in this case

$$E^{\mathrm{xc}}_{\mathrm{GC},N\mu(A),s}(\hat{\mu}_A) = E^{\mathrm{xc}}_{\mathrm{GC},N\alpha_i|A|/|\Lambda_i|,s}(\hat{\mu}_A) \ .$$

We also have

$$\mu(\Lambda_{\mathrm{err}}(l,\omega)) = \sum_{i=1}^{k} \alpha_i \frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|} \ ,$$

and thus

$$E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda_{\mathrm{err}}(l,\omega)),s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) = E^{\mathrm{xc}}_{\mathrm{GC},\sum_{i=1}^{k} N\alpha_i \frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|},s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \ .$$

Therefore, (4.98) becomes

$$\liminf_{N \to \infty} N^{-1-s/d} E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu) \geq \liminf_{N \to \infty} N^{-1-s/d} \sum_{i=1}^{k} \int_{\Omega_l} \left(\sum_{\substack{A \in F^l_\omega \\ A \subset \Lambda_i}} E^{\mathrm{xc}}_{\mathrm{GC},N\alpha_i|A|/|\Lambda_i|,s}(\hat{\mu}_A) \right) d\,\mathbb{P}_l(\omega)$$

$$+ \liminf_{N \to \infty} N^{-1-s/d} \int_{\Omega_l} E^{\mathrm{xc}}_{\mathrm{GC},\sum_{i=1}^{k} N\alpha_i|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|/|\Lambda_i|,s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \, d\,\mathbb{P}_l(\omega) \ . \quad (4.99)$$

Consider now the second term on the right hand side of (4.99). Using (4.82) we obtain

$$E^{\mathrm{xc}}_{\mathrm{GC},\sum_{i=1}^{k} N\alpha_i|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|/|\Lambda_i|,s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)})$$

$$\geq -c_{\mathrm{LO}}(d,\epsilon) \int_{\Omega_l} \left(\sum_{i=1}^{k} N\alpha_i \frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|} \right)^{1+s/d} \frac{\sum_{i=1}^{k} \alpha_i^{1+s/d} \left(\frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|} \right)^{1+s/d}}{\left(\sum_{i=1}^{k} \alpha_i \frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|} \right)^{1+s/d}} d\,\mathbb{P}_l(\omega)$$

$$= -c_{\mathrm{LO}}(d,\epsilon) N^{1+s/d} \int_{\Omega_l} \sum_{i=1}^{k} \alpha_i^{1+s/d} \left(\frac{|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|}{|\Lambda_i|} \right)^{1+s/d} d\,\mathbb{P}_l(\omega)$$

$$\geq -c_{\mathrm{LO}}(d,\epsilon) N^{1+s/d} \int_{\Omega_l} \sum_{i=1}^{k} \phi(\ell|\Lambda_i|^{-1/d}) \frac{\alpha_i^{1+s/d}}{|\Lambda_i|^{s/d}} d\,\mathbb{P}_l(\omega)$$

$$= -c_{\mathrm{LO}}(d,\epsilon) C_\rho N^{1+s/d} \ell \int_{\mathbb{R}^d} \rho^{1+s/d}(x) dx, \quad (4.100)$$

where $C_\rho > 0$. For the first inequality we applied the lower bound (C.1), which is analogous to [LLS17, (3.3)] but uniform in $s \in (\epsilon, d - \epsilon)$. For the second inequality, we applied the domain boundary regularity estimates (4.81), in which we observe that for a fixed hyperrectangle Λ_i, we can take $\phi(t) = C_{\Lambda_i} t$.

From (4.100), it follows that

$$\int_{\Omega_l} E^{\mathrm{xc}}_{\mathrm{GC},\sum_{i=1}^{k} N\alpha_i|\Lambda_i \cap \Lambda_{\mathrm{err}}(l,\omega)|/|\Lambda_i|,s}(\hat{\mu}_{\Lambda_{\mathrm{err}}(l,\omega)}) \, d\,\mathbb{P}_l(\omega) \geq -C_\rho C(d,\epsilon) \ell N^{1+s/d} \ .$$

Together with (4.99), the above proves (4.97).

36

Step 2. We will show here the statement of the theorem.

Firstly, since for $A \subset \Lambda_i, i = 1, \ldots, k,$ we have

$$\hat{\mu}_A = \frac{\alpha_i}{|\Lambda_i|} 1_A \bigg/ \frac{\alpha_i |A|}{|\Lambda_i|} = \frac{1_A}{|A|} ,$$

we get from (4.97)

$$\liminf_{N \to \infty} N^{-1-s/d} E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \ge \liminf_{N \to \infty} N^{-1-s/d} \sum_{i=1}^{k} \int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l \\ A \subset \Lambda_i}} E_{\mathrm{GC},N\alpha_i|A|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_A}{|A|} \right) \right) d\mathbb{P}_l(\omega) . \quad (4.101)$$

Next we follow in some more detail the dependency of the sets $A \in F_\omega^l$. More precisely, fix $i = 1, \ldots, k$. Any $A \in F_\omega^l$ is of the form $B_R(x)$ for some choice $R \in \{tR_1, \ldots, tR_M\}$, $t \in [1 - \eta, 1 + \eta]$ and $x \in \mathbb{R}^d$. In this case we find, in view of (4.32)

$$E_{\mathrm{GC},N\alpha_i|A|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_A}{|A|} \right) = E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_R}}{|B_R|} \right) ,$$

where we recall that B_R is the ball of radius R centred at 0. The proof of this statement follows by a change of variables idea, done by means of Proposition 2.2, and will be omitted. Furthermore, by the same type of argument we can further reduce to the unit ball B_1:

$$E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_R}}{|B_R|} \right) = R^{-s} E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) .$$

Plugging this in (4.101), we get

$$\liminf_{N \to \infty} N^{-1-s/d} E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)$$

$$\ge \liminf_{N \to \infty} N^{-1-s/d} \sum_{i=1}^{k} \int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l : \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^{-s} E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) \right) d\mathbb{P}_l(\omega) . \quad (4.102)$$

Next, we will use that by (4.73) we have for every fixed $i = 1, \ldots, k$

$$\lim_{N \to \infty} \left(N\alpha_i R^d / |\Lambda_i| \right)^{-1-s/d} E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) = -C(s,d)|B_1| , \quad \text{where} \quad C(s,d) > 0 , \quad (4.103)$$

which will allow us to take the limits in (4.102), for the sums under the integral, uniformly in N, t and $A \in F_\omega^l$. Using (4.103) in (4.102), we get for $i = 1, \ldots, k$, for arbitrary $\delta > 0$ and large N

$$\int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l : \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^{-s} E_{\mathrm{GC},N\alpha_i|B_R|/|\Lambda_i|,s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) \right) d\mathbb{P}_l(\omega)$$

$$\ge \int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l : \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^{-s} |B_1| \left(N\alpha_i R^d / |\Lambda_i| \right)^{1+s/d} (-C(s,d) - \delta) \right) d\mathbb{P}_l(\omega)$$

$$\ge \int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l : \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} N^{1+s/d} \frac{\alpha_i^{1+s/d}}{|\Lambda_i|^{1+s/d}} R^d |B_1| \right) (-C(s,d) - \delta) \, d\mathbb{P}_l(\omega)$$

$$= (-C(s,d) - \delta) N^{1+s/d} \left(\frac{\alpha_i}{|\Lambda_i|} \right)^{1+s/d} \int_{\Omega_l} \left(\sum_{\substack{A \in F_\omega^l : \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^d |B_1| \right) d\mathbb{P}_l(\omega) , \quad (4.104)$$

37

where for the first inequality we used (4.103) and in order to remove the integer part and obtain the second inequality, we used the fact that in view of (4.96) we have $R_1 \gg N^{-1/d}$ and thus $((1-\eta)R_1)^d N \gg 1$ and thus automatically due to the choices of t, R_j, also $R^d N \gg 1$ for any $R \in \{tR_1, \dots, tR_M\}$ and any $t \in [1-\eta, 1+\eta]$.

For bounding (4.104) we use the fact that $C(s, d) > 0$ and the following bound:

$$\sup_{\omega \in \Omega_l} \sum_{\substack{A \in F^l_\omega: \ \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^d |B_1| = \sup_{\omega \in \Omega_l} \sum_{\substack{A \in F^l_\omega \\ A \subset \Lambda_i}} |A| \le |\Lambda_i| \, . \tag{4.105}$$

Making use of (4.104) and (4.105) in (4.102) we find that with the choices (4.96) we have

$$\liminf_{N \to \infty} N^{-1-s/d} E^{\mathrm{xc}}_{\mathrm{GC}, N, s}(\mu) \ge \liminf_{N \to \infty} (-C(s, d) - \delta) \sum_{i=1}^k \left(\frac{\alpha_i}{|\Lambda_i|} \right)^{1+s/d} \int_{\Omega_l} \left(\sum_{\substack{A \in F^l_\omega: \ \exists x \in \mathbb{R}^d, \\ A = B_R(x) \subset \Lambda_i}} R^d |B_1| \right) d\, \mathbb{P}_l(\omega)$$

$$\ge (-C(s, d) - \delta) \sum_{i=1}^k \left(\frac{\alpha_i}{|\Lambda_i|} \right)^{1+s/d} |\Lambda_i| = (-C(s, d) - \delta) \int_{\mathbb{R}^d} \left(\sum_{i=1}^k \alpha_i \rho_i(x) \right)^{1+s/d} dx \, .$$

Taking now $\delta \to 0$ in the above proves the statement of our Proposition. $\qquad\square$

4.3 Proof of Theorem 1.1

Proof of Theorem 1.1: **Step 1.** Let ρ be continuous and with compact support $\mathrm{supp}(\mu) =: \Lambda \subset \mathbb{R}^d$. Fix $l > 0$ such that $[-\frac{l}{2}, \frac{l}{2}]^d$ is much smaller than Λ. Take as in Lemma 4.2 a Swiss cheese packing of $[-\frac{l}{2}, \frac{l}{2}]^d$ by balls of radii $R_1, \dots, R_M$, extended by periodicity to the whole $\mathbb{R}^d$. We also assume below that R_1, l, M, satisfy (4.94), but without the specific N-dependent choices (4.96). At the end of the proof we will further take $l \to 0, M \to \infty$, and the discussion is going to be independent of the choice of $R_1 > 0$, as long as $R_1 < C^{-3M} l$ as in (4.94).

We write, recalling definition (4.76a) for the second equality (now used for one single covering family, i.e. without the ω-dependence in (4.76a))

$$\rho = \rho|_{\cup_{j=1}^M \cup_{\substack{A \in \mathcal{B}^l_{R_j} \\ A \cap \Lambda \ne \varnothing}} A} + \rho|_{\Lambda \setminus \left(\cup_{j=1}^M \cup_{\substack{A \in \mathcal{B}^l_{R_j} \\ A \cap \Lambda \ne \varnothing}} A \right)} = \rho|_{\Lambda \setminus \Lambda_{\mathrm{nc}}} + \rho|_{\Lambda_{\mathrm{nc}}},$$

where we denoted by $\mathcal{B}^l_{R_j}$ the set of balls of radius R_j from the Swiss cheese packing. Define for all $x \in \Lambda$ the sequences

$$\rho^l_{\min}(x) := \frac{1}{k^l_M} \sum_{j=1}^M \sum_{\substack{A \in \mathcal{B}^\delta_{R_j} \\ A \cap \Lambda \ne \varnothing}} \left(\min_{x \in A \cap \Lambda} \rho(x) \right) 1_{A \cap \Lambda}(x) \quad \text{and} \quad \rho^l_{\max}(x) := \frac{1}{m^l_M} \sum_{j=1}^M \sum_{\substack{A \in \mathcal{B}^\delta_{R_j} \\ A \cap \Lambda \ne \varnothing}} \left(\max_{x \in A \cap \Lambda} \rho(x) \right) 1_{A \cap \Lambda}(x),$$
(4.106)

where k^l_M and m^l_M are normalization constants required to make $\rho^l_{\min}, \rho^l_{\max}$ probability densities and where we set $\rho^l_{\min} = 0$ if $\min_{x \in A \cap \Lambda} \rho(x) = 0$ for all $A \in \bigcup_{j=1}^M \mathcal{B}^\delta_{R_j}$. Then

$$\sum_{j=1}^M \sum_{\substack{A \in \mathcal{B}^\delta_{R_j} \\ A \cap \Lambda \ne \varnothing}} \left(\min_{x \in A \cap \Lambda} \rho(x) \right) 1_{A \cap \Lambda}(x) \le \rho(x)|_{\cup_{j=1}^M \cup_{\substack{A \in \mathcal{B}^\delta_{R_j} \\ A \cap \Lambda \ne \varnothing}} A} = \rho|_{\Lambda \setminus \Lambda_{\mathrm{nc}}} \le \sum_{j=1}^M \sum_{\substack{A \in \mathcal{B}^\delta_{R_j} \\ A \cap \Lambda \ne \varnothing}} \left(\max_{x \in A \cap \Lambda} \rho(x) \right) 1_{A \cap \Lambda}(x), \tag{4.107}$$

and we claim that $\rho^l_{\min}, \rho^l_{\max}$ converge strongly in $L^{1+s/d}(\mathbb{R}^d)$ to ρ as $l \to 0, M \to \infty$.

We prove only the convergence $\lim_{l \to 0} \|\rho - \rho^l_{\min}\|_{L^{1+s/d}(\mathbb{R}^d)} = 0$, as the convergence of $\rho^l_{\max}$ is proved similarly. Note first that as ρ is continuous and compactly-supported, therefore it is uniformly continuous. Thus, for

all $\delta > 0$ there exists $l_\delta > 0$ such that for all $l \leq l_\delta$ and all A in a given packing of $[-\frac{l}{2}, \frac{l}{2}]^d$, we have $\max_{x \in A \cap \Lambda} \rho(x) - \min_{x \in A \cap \Lambda} \rho(x) \leq \delta$. Hence

$$m_M^l - k_M^l \leq \delta \sum_{j=1}^{M} \sum_{\substack{A \in \mathcal{B}_{R_j}^\delta \\ A \cap \Lambda \neq \varnothing}} \int_{\mathbb{R}^d} 1_{A \cap \Lambda}(x) \leq \delta |\Lambda| \rightarrow 0 \text{ as } l, \delta \rightarrow 0.$$

Furthermore, for $l \leq l_\delta$ we have

$$\left(\int_{\mathbb{R}^d} \left(\rho(x) - k_M^l \rho_{\min}^l(x) \right)^{1+s/d} dx \right)^{\frac{1}{1+s/d}}$$

$$\leq \left(\int_{\mathbb{R}^d} \left(\sum_{j=1}^{M} \sum_{\substack{A \in \mathcal{B}_{R_j}^\delta \\ A \cap \Lambda \neq \varnothing}} \left(\max_{x \in A \cap \Lambda} \rho(x) - \min_{x \in A \cap \Lambda} \rho(x) \right) 1_{A \cap \Lambda}(x) \right)^{1+s/d} dx \right)^{\frac{1}{1+s/d}} + \left(\int_{\Lambda_{\mathrm{nc}}(l)} \left(\max_{y \in \Lambda} \rho(y) \right)^{1+s/d} dx \right)^{\frac{1}{1+s/d}}$$

$$\leq \delta |\Lambda|^{1+s/d} + \left(\max_{y \in \Lambda} \rho(y) \right) |\Lambda_{\mathrm{nc}}(l)|^{\frac{1}{1+s/d}}, \tag{4.108}$$

which tends to 0 as $l, \delta \rightarrow 0$ and $M \rightarrow \infty$, in view of (4.80) and of the boundedness of Λ. As ρ and $\rho_{\min}^l$ are probability densities, (4.108) gives in particular that $k_M^l \rightarrow 1$ as $l \rightarrow 0$. Then, by means of (4.108), we obtain

$$\left\| \rho - \rho_{\min}^l \right\|_{L^{1+s/d}} \leq \left(\int_{\mathbb{R}^d} \left(\rho(x) - k_M^l \rho_{\min}^l(x) \right)^{1+s/d} dx \right)^{\frac{1}{1+s/d}} + (1 - k_M^l) \left(\int_{\mathbb{R}^d} \left(\rho_{\min}^l(x) \right)^{1+s/d} dx \right)^{\frac{1}{1+s/d}},$$

which tends to 0 as $l \rightarrow 0$ and $M \rightarrow \infty$. This proves the claim.

Step 2. We assume here that ρ is continuous and $\rho \in L^{1+s/d}(\mathbb{R}^d)$, and we prove the lower bound. The proof of the upper bound follows similarly from an argument similar to (3.2) and by means of Step 1. Take $0 < l \ll 1$. Then for a fixed compact Borel measurable set Λ with nonempty interior, such that $\Lambda \subseteq \mathrm{supp}(\mu)$ and $\mu(\Lambda) > 0$, we denote $\widetilde{\Lambda} := \mathrm{supp}(\mu) \setminus \Lambda$ and, with the notations (4.76a), (4.76b), we use the splitting

$$\mu = \mu|_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)} + \mu|_{\Lambda_{\mathrm{nc}}(l,\omega)} + \mu|_{\widetilde{\Lambda} \setminus \widetilde{\Lambda}_{nc}(l,\omega)} + \mu|_{\widetilde{\Lambda}_{nc}(l,\omega)}. \tag{4.109}$$

By applying the same arguments as in Lemma 4.14 and with the notation (4.82), we have by means of (4.89) and (4.109) (and with the convention that $E_{\mathrm{GC},k,\mathsf{c}}^{\mathrm{xc}}(0) = 0$)

$$E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \geq \frac{M}{M+C} \left(\int_{\Omega_l} E_{\mathrm{GC},N\mu(\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)), \sum_{\substack{A \in F_\omega^l \\ A \cap \Lambda \neq \varnothing}} \mathsf{c}_A}^{\mathrm{xc}}(\hat{\mu}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)}) \, d\mathbb{P}_l(\omega) \right.$$

$$+ \int_{\Omega_l} E_{\mathrm{GC},N\mu(\widetilde{\Lambda} \setminus \widetilde{\Lambda}_{nc}(l,\omega)), s}^{\mathrm{xc}}(\hat{\mu}_{\widetilde{\Lambda} \setminus \widetilde{\Lambda}_{nc}(l,\omega)}) \, d\mathbb{P}_l(\omega) - \frac{C(w,d,\epsilon)}{M} R_1^{-s} (N-1)$$

$$\left. - \frac{C(w,d,\epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x) dx \right). \tag{4.110}$$

We now apply the construction (4.107) from Step 1, to the first term in (4.110), with $\hat{\mu}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)}$ (respectively $\hat{\mu}_\Lambda$) instead of μ and with respect to the ball covering F_ω^l, and we denote by $(\hat{\rho}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)})_{\max}^{l,\omega}$ (respectively $(\hat{\rho}_\Lambda)_{\max}^{l,\omega}$) the so-obtained densities. Explicitly, we have by using the notations (4.82) and (4.106)

$$(\hat{\rho}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)})_{\max}^{l,\omega}(x) := \frac{1}{\hat{m}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(l,\omega)}^{l,\omega}} \sum_{\substack{A \in F_\omega^l \\ A \cap \Lambda \neq \varnothing}} \left(\max_{x \in A \cap \Lambda} \hat{\rho}_{\Lambda \setminus \Lambda_{\mathrm{nc}}(\omega,l)}(x) \right) 1_{A \cap \Lambda}(x)$$

$$= \frac{1}{\hat{m}_\Lambda^{l,\omega}} \sum_{\substack{A \in F_\omega^l \\ A \cap \Lambda \neq \varnothing}} \left(\max_{x \in A \cap \Lambda} \hat{\rho}_\Lambda(x) \right) 1_{A \cap \Lambda}(x) = (\hat{\rho}_\Lambda)_{\max}^{l,\omega}(x), \tag{4.111}$$

where $\hat{m}^{l,\omega}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)}, \hat{m}^{l,\omega}_{\Lambda}$ are normalization factors. Explicitly, we have

$$
\hat{m}^{l,\omega}_{\Lambda\setminus\Lambda_{nc}(l,\omega)} : = \sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} |A\cap\Lambda| \max_{y\in A\cap\Lambda} \hat{\rho}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)}(y)
$$

$$
= \frac{1}{\mu(\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega))} \sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} |A\cap\Lambda| \max_{y\in A\cap\Lambda} \rho(y) = \frac{\mu(\Lambda)}{\mu(\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega))}\hat{m}^{l,\omega}_{\Lambda} \geq 1 \qquad (4.112)
$$

We find directly from the definitions that $(\hat{\rho}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)})^{l,\omega}_{\max} \geq \frac{1}{\hat{m}^{l,\omega}_{\Lambda\setminus\Lambda_{nc}(l,\omega)}}\hat{\mu}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)}$. By (4.29) of item 2 of Remark 4.7, applied with $\mu'' = \hat{\mu}_{\Lambda\setminus\Lambda_{\mathrm{nc}}}$, $N'' = N\mu(\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega))$ and $N'+N'' = N\mu(\Lambda)\hat{m}^{l,\omega}_{\Lambda}$, and by applying (4.31) to the resulting μ' we get

$$
E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)),\sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} \mathsf{c}_A}(\hat{\mu}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)}) \geq E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda)\hat{m}^{l,\omega}_{\Lambda},\sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} \mathsf{c}_A}((\hat{\mu}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)})^{l,\omega}_{\max}). \qquad (4.113)
$$

By plugging (4.113) in (4.110), while the other terms remain unchanged, we obtain

$$
\begin{aligned}
E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu) \geq{} & \frac{M}{M+C}\Bigg(\int_{\Omega_l} E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\Lambda)\hat{m}^{l,\omega}_{\Lambda},\sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} \mathsf{c}_A}((\hat{\mu}_{\Lambda\setminus\Lambda_{\mathrm{nc}}(l,\omega)})^{l,\omega}_{\max})d\,\mathbb{P}_l(\omega) \\
& + \int_{\Omega_l} E^{\mathrm{xc}}_{\mathrm{GC},N\mu(\widetilde{\Lambda}\setminus\widetilde{\Lambda}_{nc}(l,\omega)),s}(\hat{\mu}_{\widetilde{\Lambda}\setminus\widetilde{\Lambda}_{nc}(l,\omega)})\, d\,\mathbb{P}_l(\omega) - \frac{C(w,d,\epsilon)}{M}R_1^{-s}\,(N-1) \\
& - \frac{C(w,d,\epsilon)}{M}N^{1+s/d}\int_{\mathbb{R}^d}\rho^{1+s/d}(x)dx \Bigg) .
\end{aligned} \qquad (4.114)
$$

By using (4.112) and applying now to the first term in (4.114) a similar argument as the one to get (4.102), but this time without any need to consider separately the terms with $A\subset\Lambda$ and those with $A\cap\partial\Lambda\neq\varnothing$, as l is now independent of N, we get

$$
\begin{aligned}
E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu) \geq{} & \frac{M}{M+C}\Bigg\{ \int_{\Omega_l} \bigg(\sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} E^{\mathrm{xc}}_{\mathrm{GC},N|A\cap\Lambda|(\max_{x\in A\cap\Lambda}\rho(x)),s}\bigg(\frac{1_{A\cap\Lambda}}{|A\cap\Lambda|}\bigg) \bigg) d\,\mathbb{P}_l(\omega) \qquad (4.115) \\
& -c_{\mathrm{LO}}(d,\epsilon)N^{1+s/d}\int_{\mathbb{R}^d\setminus\Lambda}\rho^{1+s/d}(x)dx - \frac{C(w,d,\epsilon)}{M}R_1^{-s}\,(N-1) - \frac{C(w,d,\epsilon)}{M}N^{1+s/d}\int_{\mathbb{R}^d}\rho^{1+s/d}(x)dx \Bigg\},
\end{aligned}
$$

where for the second term in (4.115) we applied (C.1) and used the fact that $\widetilde{\Lambda}\subseteq\mathbb{R}^d\setminus\Lambda$. By (4.73), we have for the first term in (4.115), using also (4.111) in order to transfer the estimate to $(\hat{\rho}_\Lambda)^{l,\omega}_{\max}$

$$
\begin{aligned}
& \liminf_{N\to\infty}\frac{1}{N^{1+s/d}}\int_{\Omega_l} \bigg(\sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} E^{\mathrm{xc}}_{\mathrm{GC},N|A\cap\Lambda|(\max_{x\in A\cap\Lambda}\rho(x)),s}\bigg(\frac{1_{A\cap\Lambda}}{|A\cap\Lambda|}\bigg) \bigg) d\,\mathbb{P}_l(\omega) \\
\geq{} & -C(s,d)\int_{\Omega_l} \sum_{\substack{A\in F^l_\omega \\ A\cap\Lambda\neq\varnothing}} |A\cap\Lambda| \bigg(\max_{y\in A\cap\Lambda}\rho(y) \bigg)^{1+s/d} d\mathbb{P}_l(\omega) \\
={} & -C(s,d)\int_{\Omega_l} (\mu(\Lambda))^{1+s/d}(\hat{m}^{l,\omega}_{\Lambda})^{1+s/d}\int_{\mathbb{R}^d}\big((\hat{\mu}_\Lambda)^{l,\omega}_{\max}\big)^{1+s/d} dx\, d\,\mathbb{P}_l(\omega).
\end{aligned}
$$

Therefore, dividing in (4.115) by $N^{1+s/d}$ and taking limits directly gives the bound

$$
\begin{aligned}
\liminf_{N\to\infty}\frac{E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu)}{N^{1+s/d}} \geq{} & -\frac{M}{M+C}\Bigg\{ C(s,d)\int_{\Omega_l} (\mu(\Lambda))^{1+s/d}(\hat{m}^{l,\omega}_{\Lambda})^{1+s/d}\int_{\mathbb{R}^d}\big((\hat{\mu}_\Lambda)^{l,\omega}_{\max}\big)^{1+s/d} dx\, d\,\mathbb{P}_l(\omega) \\
& +c_{\mathrm{LO}}(d,\epsilon)\int_{\mathbb{R}^d\setminus\Lambda}\rho^{1+s/d}(x)dx + \frac{C(w,d,\epsilon)}{M}\int_{\mathbb{R}^d}\rho^{1+s/d}(x)dx \Bigg\}.
\end{aligned} \qquad (4.116)
$$

40

To take the limits $l \to 0, M \to \infty$ in the above, for the first integral we made use of Step 1, in particular that $\hat{m}_\Lambda^{l,\omega} \to 1$ and $(\hat{\rho}_\Lambda)_{\max}^{l,\omega} \to \hat{\rho}_\Lambda$ in $L^{1+s/d}(\mathbb{R}^d)$ as $l \to 0$, and we used the Dominated Convergence Theorem (whose conditions are satisfied since $\rho|_\Lambda$ is bounded and Λ is bounded). In conclusion, from (4.116) we get

$$\liminf_{N\to\infty} \frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} \;\geq\; -C(s,d)\int_\Lambda \rho^{1+s/d}(x)dx - c_{\mathrm{LO}}(d,\epsilon)\int_{\mathbb{R}^d\setminus\Lambda} \rho^{1+s/d}(x)dx. \tag{4.117}$$

Taking now $\Lambda = [-R,R]^d \cap \mathrm{supp}(\mu)$ and by taking $R \to \infty$ we have $\rho|_\Lambda \leq \rho$ and $\rho|_\Lambda(x) \to \rho(x)$ for all $x \in \mathbb{R}^d$, therefore again by dominated convergence the first integral in (4.117) converges to $\int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx$, and similarly the second integral in (4.117) tends to zero. This allows to conclude Theorem 1.1.

Step 3. The extension to general marginals $\rho \in L^{1+s/d}(\mathbb{R}^d)$ follows now from Lusin's Theorem due to the measurability of ρ and via (4.117) from Step 2. More precisely, if $\nu(A) := \int_A \rho^{1+s/d}(x)dx$ then for every $n \in \mathbb{N}$ there exists a compact set $\Lambda_n \subset \mathbb{R}^d$ such that $\rho|_{\Lambda_n}$ is continuous and $\nu(\mathbb{R}^d \setminus \Lambda_n) < 2^{-n}$. Then we have from (4.117), with Λ_n instead of Λ

$$\liminf_{N\to\infty} \frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} \geq -C(s,d)\int_{\Lambda_n} \rho^{1+s/d}(x)dx - c_{\mathrm{LO}}(d,\epsilon)\int_{\mathbb{R}^d\setminus\Lambda_n} \rho^{1+s/d}(x)dx.$$

Then $\nu(\mathbb{R}^d \setminus \Lambda_n) = \int_{\mathbb{R}^d\setminus\Lambda_n} \rho^{1+s/d}(x)dx \to 0$ as $n \to \infty$, which produces the desired result. $\qquad\square$

5   Small oscillations property of $E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)/N^{1+s/d}$

Proof of Theorem 1.4

In order to prove (1.10), we will use again the Fefferman-Gregg decomposition. Let μ be a piecewise constant density of form $\sum_{j=1}^k \alpha_j \mu_j$, where for all $j = 1,\ldots,k,$, $\mu_j \in \mathcal{P}(\mathbb{R}^d)$ is a uniform measure on $\Lambda_j \subset \mathbb{R}^d$, with $\Lambda_j, j = 1,\ldots,k$, hyper-rectangles with disjoint nonempty interiors, and $\alpha_j \in \mathbb{R}_{>0}, j = 1,\ldots k$, $\sum_{j=1}^k \alpha_j = 1$.

We prove below that

$$\frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} \;-\; \frac{E_{\mathrm{GC},\widetilde{N},s}^{\mathrm{xc}}(\mu)}{\widetilde{N}^{1+s/d}} \;\geq\; -\frac{C(\Lambda_1,\ldots,\Lambda_k,\alpha_1,\ldots,\alpha_k,d,\epsilon)}{\log\widetilde{N}}\;, \tag{5.1}$$

as the other direction can be easily argued similarly.

Step 1. At first, from Lemma 4.14 and eqn. (4.99) and (4.100) from Step 1 in the proof of Proposition 4.1, we get, if B_1 is the unit ball in $\mathbb{R}^d$ centered at zero,

$$\begin{aligned}
E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \;\geq\; & \frac{M}{M+C}\sum_{j=1}^k \int_{\Omega_l}\left(\sum_{\substack{A\in F_\omega^l,\exists x\in\mathbb{R}^d \\ A=B_R(x)\subset\Lambda_j}} R^{-s}E_{\mathrm{GC},N\alpha_j|B_R|/|\Lambda_j|,s}^{\mathrm{xc}}\left(\frac{1_{B_1}}{|B_1|}\right)\right)d\,\mathbb{P}_l(\omega) \\
& -\frac{C'(w,\epsilon,d)}{M}R_1^{-s}(N-1) - \frac{C(w,d,\epsilon)}{M}N^{1+s/d}\int_{\mathbb{R}^d}\rho^{1+s/d}(x)dx \\
& -c_{\mathrm{LO}}(d,\epsilon)N^{1+s/d}\int_{\Omega_l}\left(\int_{\Lambda_{\mathrm{err}}(l,\omega)}\rho^{1+s/d}(x)dx\right)d\,\mathbb{P}_l(\omega),
\end{aligned} \tag{5.2}$$

where any $A \in F_\omega^l$ is a ball $B_R(x)$ for some choice $R \in \{tR_1,\ldots,tR_M\}$, $t \in [1-\eta,1+\eta]$ and some $x \in \mathbb{R}^d$. Furthermore, for each l,ω the set $\Lambda_{\mathrm{err}}(l,\omega)$ is the union of the balls from F_ω^l which intersect $\bigcup_{j=1}^k \partial\Lambda_j$, which in turn are contained in cubes of sidelength tl that intersect $\partial\Lambda_j$ for some $1 \leq j \leq k$, which cubes are all included in the set $\bigcup_{j=1}^k (\partial\Lambda_j)_{2l\sqrt{d}}$.

To proceed, we now re-write for each $j = 1, \ldots, k$, in the first term in (5.2)

$$\int_{\Omega_l} \left(\sum_{\substack{B_R(x) \in F_\omega^l, \\ B_R(x) \subset \Lambda_j}} R^{-s} E_{\mathrm{GC}, N\alpha_j |B_R| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) \right) d\,\mathbb{P}_l(\omega)$$

$$= \int_{\Omega_l} \left(\sum_{\substack{B_R(x) \in F_\omega^l, \\ B_R(x) \subset \Lambda_j}} R^{-s} (N\alpha_j R^d |B_1| / |\Lambda_j|)^{1+s/d} \frac{E_{\mathrm{GC}, N\alpha_j |B_R| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(N\alpha_j R^d |B_1| / |\Lambda_j|)^{1+s/d}} \right) d\,\mathbb{P}_l(\omega)$$

$$= \left(\frac{\alpha_j}{|\Lambda_j|} \right)^{1+s/d} \sum_{i=1}^{M} \int_{\Omega_l} \left(\sum_{\substack{B_{tR_i}(x) \in F_\omega^l, \\ B_{tR_i}(x) \subset \Lambda_j}} (tR_i)^{-s} (N(tR_i)^d |B_1|)^{1+s/d} \frac{E_{\mathrm{GC}, N\alpha_j |B_{tR_i}| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(N\alpha_j (tR_i)^d |B_1| / |\Lambda_j|)^{1+s/d}} \right) d\,\mathbb{P}_l(\omega)$$

$$= \left(\frac{\alpha_j}{|\Lambda_j|} \right)^{1+s/d} \sum_{i=1}^{M} \int_{\Omega_l} \frac{E_{\mathrm{GC}, N\alpha_j |B_{tR_i}| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(N\alpha_j (tR_i)^d |B_1| / |\Lambda_j|)^{1+s/d}} \left(\sum_{\substack{B_{tR_i}(x) \in F_\omega^l, \\ B_{tR_i}(x) \subset \Lambda_j}} (tR_i)^{-s} (N(tR_i)^d |B_1|)^{1+s/d} \right) d\,\mathbb{P}_l(\omega).$$

$$(5.3)$$

We now use the fact that by Lemma 4.2 there holds for each $1 \le i \le M$, $1 \le j \le k$ and $\omega \in \Omega_l$, by a reasoning similar to the one leading to (4.80)

$$\sum_{\substack{B_{tR_i}(x) \in F_\omega^l, \\ B_{tR_i}(x) \subset \Lambda_j}} (tR_i)^d |B_1| \le \frac{|\{y : \mathrm{dist}(y, \Lambda_j) \le 2l\sqrt{d}\}|}{M + C_d} =: \frac{|(\Lambda_j)_{2l\sqrt{d}}|}{M + C_d}. \tag{5.4}$$

To obtain the above bound, note that the balls from F_ω^l which have radius tR_i cover at most $(M + C_d)^{-1}$ of each cube $K \in \mathcal{F}_\omega$, where $\mathcal{F}_\omega$ is a covering of $\mathbb{R}^d$ by cubes of sidelength tl with $t < 2$, and disjoint interiors. The cubes of such covering that have nonempty intersection with Λ_j stay within distance $2l\sqrt{d}$ of Λ_j and thus their total volume is at most $|(\Lambda_j)_{2l\sqrt{d}}|$. These considerations directly lead to (5.4).

Applying (5.4), we obtain

$$\sum_{\substack{B_{tR_i}(x) \in F_\omega^l, \\ B_{tR_i}(x) \subset \Lambda_j}} (tR_i)^{-s} (N(tR_i)^d |B_1|)^{1+s/d} \le N^{1+s/d} |B_1|^{s/d} \frac{|(\Lambda_j)_{2l\sqrt{d}}|}{M + C_d}. \tag{5.5}$$

In view of (4.31), from (5.3), (5.5) we get for $1 \le j \le k$

$$\int_{\Omega_l} \left(\sum_{\substack{B_R(x) \in F_\omega^l, \\ B_R(x) \subset \Lambda_j}} R^{-s} E_{\mathrm{GC}, N\alpha_j |B_R| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right) \right) d\,\mathbb{P}_l(\omega)$$

$$\ge N^{1+s/d} |B_1|^{s/d} \frac{|(\Lambda_j)_{2l\sqrt{d}}|}{M + C_d} \int_{\Omega_l} \sum_{i=1}^{M} \frac{E_{\mathrm{GC}, N\alpha_j |B_{tR_i}| / |\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(N\alpha_j (tR_i)^d |B_1| / |\Lambda_j|)^{1+s/d}} d\,\mathbb{P}_l(\omega). \tag{5.6}$$

We now move to finding an upper bound for the last term in (5.2). Similarly to the reasoning used to justify (5.4), $\Lambda_{\mathrm{err}}(l, \omega) \cap \Lambda$ is contained in $\bigcup_{j=1}^{k} (\partial \Lambda_j)_{2l\sqrt{d}}$, and thus we have for the last term in (5.2)

$$\int_{\Lambda_{\mathrm{err}}(l,\omega)} \rho^{1+s/d}(x) dx \le \int_{\bigcup_{j=1}^{k} (\partial \Lambda_j)_{2l\sqrt{d}}} \rho^{1+s/d}(x) dx$$

$$\le \int_{\bigcup_{j=1}^{k} (\partial \Lambda_j)_{2l\sqrt{d}}} \rho^{1+\epsilon/d}(x) dx + \int_{\bigcup_{j=1}^{k} (\partial \Lambda_j)_{2l\sqrt{d}}} \rho^{1+(d-\epsilon)/d}(x) dx. \tag{5.7}$$

42

Thus, by using the estimates from (5.4), we obtain in (5.2) by means of (5.6) and of (5.7) for some $c'(d,\epsilon) > 0$

$$
\frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} \geq |B_1|^{s/d}\frac{M}{M+C}\sum_{j=1}^k\left(\frac{\alpha_j}{|\Lambda_j|}\right)^{1+s/d}\frac{|(\Lambda_j)_{2l\sqrt{d}}|}{M+C_d}\sum_{i=1}^M\int_{\Omega_l}\frac{E_{\mathrm{GC},N\alpha_j|B_{tR_i}|/|\Lambda_j|,s}^{\mathrm{xc}}\left(\frac{1_{B_1}}{|B_1|}\right)}{(N\alpha_j(tR_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}d\,\mathbb{P}_l(\omega)
$$

$$
-\frac{C}{M}\frac{N-1}{N^{1+\frac{s}{d}}}R_1^{-s} - c'(d,\epsilon)\sum_{s'\in\{\epsilon,d-\epsilon\}}\left(\int_{\bigcup_{j=1}^k(\partial\Lambda_j)_{2l\sqrt{d}}}\rho^{1+s'/d}(x)dx + \frac{1}{M}\int_{\mathbb{R}^d}\rho^{1+s'/d}(x)dx\right). \quad (5.8)
$$

Step 2. We will next get a lower bound for $E_{\mathrm{GC},\tilde{N},s}^{\mathrm{xc}}(\mu)$ in function of $E_{\mathrm{GC},N|B_{t\tilde{R}_i}|,s}^{\mathrm{xc}}(\mu)$, where $\widetilde{N} > 0$ and $(\tilde{R}_i)$ is a new balls cover.

To begin with, using as in the proof of Lemma 4.14, for another cover and for any $\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)$ as in (4.77), we can write

$$
\mu = \mu(\Lambda \setminus \widetilde{\Lambda}_{\mathrm{err}}(l,\omega))\hat{\mu}_{\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)} + \mu(\widetilde{\Lambda}_{\mathrm{err}}(l,\omega))\hat{\mu}_{\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)} \quad (5.9)
$$

and

$$
\hat{\mu}_{\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)} = \sum_{\substack{\widetilde{A}\in\widetilde{F}_\omega^l \\ \tilde{A}\subset\Lambda,\mu(\tilde{A})>0}}\hat{\mu}_{\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)}(\widetilde{A})\,\hat{\mu}_{\widetilde{A}} + \sum_{\substack{\widetilde{A}\in\widetilde{F}_\omega^l \\ \tilde{A}\subset\Lambda,\mu(\tilde{A})=0}}\hat{\mu}_{\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)}|_{\widetilde{A}}\,. \quad (5.10)
$$

Applying Remark 4.7 (4) twice, firstly to (5.9) and then to (5.10), and then using (4.31), we get

$$
\begin{aligned}
E_{\mathrm{GC},\widetilde{N},s}^{\mathrm{xc}}(\mu) &\leq E_{\mathrm{GC},\widetilde{N}\mu(\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)),s}^{\mathrm{xc}}\left(\hat{\mu}_{\Lambda\setminus\widetilde{\Lambda}_{\mathrm{err}}(l,\omega)}\right) \\
&\leq \sum_{j=1}^k\sum_{\substack{\widetilde{A}\in\widetilde{F}_\omega^l \\ \tilde{A}\subset\Lambda_j,\mu(\tilde{A})>0}}E_{\mathrm{GC},\widetilde{N}\mu(\widetilde{A}),s}^{\mathrm{xc}}(\hat{\mu}_{\tilde{A}}) = \sum_{j=1}^k\sum_{\substack{\widetilde{A}\in\widetilde{F}_\omega^l \\ \tilde{A}\subset\Lambda_j,\mu(\tilde{A})>0}}E_{\mathrm{GC},\widetilde{N}\alpha_j|A|/|\Sigma_j|,s}^{\mathrm{xc}}\left(\frac{1_{\tilde{A}}}{|\tilde{A}|}\right) \\
&= \sum_{j=1}^k\sum_{\substack{\tilde{A}=B_{\tilde{R}}(x)\in\widetilde{F}_\omega^l, \\ B_{\tilde{R}}(x)\subset\Lambda_j}}E_{\mathrm{GC},\widetilde{N}\alpha_j|B_{\tilde{R}}|/|\Lambda_j|,s}^{\mathrm{xc}}\left(\frac{1_{B_{\tilde{R}}}}{|B_{\tilde{R}}|}\right) = \sum_{j=1}^k\sum_{\substack{B_{\tilde{R}}(x)\in\widetilde{F}_\omega^l, \\ B_{\tilde{R}}(x)\subset\Lambda_j}}\widetilde{R}^{-s}E_{\mathrm{GC},\widetilde{N}\alpha_j|B_{\tilde{R}}|/|\Lambda_j|,s}^{\mathrm{xc}}\left(\frac{1_{B_1}}{|B_1|}\right)\,.
\end{aligned}
$$
$$
(5.11)
$$

Performing this procedure for every ω, and then integrating, we get

$$
E_{\mathrm{GC},\widetilde{N},s}^{\mathrm{xc}}(\mu) \leq \sum_{j=1}^k\int_{\Omega_l}\sum_{\substack{B_{\tilde{R}}(x)\in\widetilde{F}_\omega^l, \\ B_{\tilde{R}}(x)\subset\Lambda_j}}\widetilde{R}^{-s}E_{\mathrm{GC},\widetilde{N}\alpha_j|B_{\tilde{R}}|/|\Lambda_j|,s}^{\mathrm{xc}}\left(\frac{1_{B_1}}{|B_1|}\right)d\,\mathbb{P}_l(\omega). \quad (5.12)
$$

By Lemma 4.2 and similarly to (5.4), for each $1 \leq i \leq \tilde{M}$, each $1 \leq j \leq k$ and each $\omega \in \Omega_l$ we have, by a reasoning similar to the one used to justify (5.4),

$$
\sum_{\substack{\tilde{B}_{t\tilde{R}_i}(x)\in\tilde{F}_\omega^l, \\ B_{t\tilde{R}_i}(x)\subset\Lambda_j}}(t\tilde{R}_i)^d|B_1| \geq \frac{|\{y:\ \mathrm{dist}(y,\mathbb{R}^d\setminus\Lambda_j) > 2l\sqrt{d}\}|}{\widetilde{M}+C_d+1} := \frac{|(\Lambda_j)_{2l\sqrt{d}}^*|}{\widetilde{M}+C_d+1}. \quad (5.13)
$$

To justify (5.13), we use the same notation as in the paragraph following (5.4). As $\mathcal{F}_\omega$ forms a covering of $\mathbb{R}^d$, any point $y \in (\Lambda_j)_{2l\sqrt{d}}^*$ is in some cube $K_y \in \mathcal{F}_\omega$. As the sidelength of K_y is smaller than $2l$ and $y \in (\Lambda_j)_{2l\sqrt{d}}^*$, we have $K_y \subset B(y,2l\sqrt{d}) \subset \Lambda_j$; this proves that $(\Lambda_j)_{2l\sqrt{d}}^*$ is covered by cubes from $\mathcal{F}_\omega$ that are completely contained in Λ_j. The balls appearing in the sum on the left in (5.13) cover at least $(M+C_d+1)^{-1}$ of K for each $K \in \mathcal{F}_\omega$ completely contained in Λ_j, and as the $K \in \mathcal{F}_\omega$ completely contained in Λ_j cover $(\Lambda_j)_{2l\sqrt{d}}^*$, (5.13) follows.

Now (5.12) and (5.13) give

$$
\begin{aligned}
\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N},s}(\mu)}{\widetilde{N}^{1+s/d}} \;&\leq\; \sum_{j=1}^{k}\int_{\Omega_l}\sum_{\substack{B_{\widetilde{R}}(x)\in\widetilde{F}^l_\omega,\\ B_{\widetilde{R}}(x)\subset\Lambda_j}}\widetilde{R}^{-s}\,\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N}\alpha_j|B_{\widetilde{R}}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{\widetilde{N}^{1+s/d}}\,d\,\mathbb{P}_l(\omega)\\[2mm]
&=\; \sum_{j=1}^{k}\sum_{i=1}^{\widetilde{M}}\int_{\Omega_l}\sum_{\substack{B_{t\widetilde{R}_i}(x)\in\widetilde{F}^l_\omega,\\ B_{t\widetilde{R}_i}(x)\subset\Lambda_j}}(t\widetilde{R}_i)^{-s}\big((t\alpha_j\widetilde{R}_i)^d|B_1|/|\Lambda_j|\big)^{1+s/d}\,\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N}\alpha_j|B_{t\widetilde{R}_i}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{(\widetilde{N}\alpha_j(t\widetilde{R}_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}\,d\,\mathbb{P}_l(\omega)\\[2mm]
&=\; \sum_{j=1}^{k}\left(\frac{\alpha_j}{|\Lambda_j|}\right)^{1+s/d}\sum_{i=1}^{\widetilde{M}}\int_{\Omega_l}\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N}\alpha_j|B_{t\widetilde{R}_i}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{(\widetilde{N}\alpha_j(t\widetilde{R}_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}\sum_{\substack{B_{t\widetilde{R}_i}(x)\in\widetilde{F}^l_\omega,\\ B_{t\widetilde{R}_i}(x)\subset\Lambda_j}}(t\widetilde{R}_i)^d|B_1|^{1+s/d}\,d\,\mathbb{P}_l(\omega)\\[2mm]
&\leq\; |B_1|^{s/d}\sum_{j=1}^{k}\left(\frac{\alpha_j}{|\Lambda_j|}\right)^{1+s/d}\frac{|(\Lambda_j)^*_{2l\sqrt{d}}|}{\widetilde{M}+C_d+1}\sum_{i=1}^{\widetilde{M}}\int_{\Omega_l}\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N}\alpha_j|B_{t\widetilde{R}_i}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{(\widetilde{N}\alpha_j(t\widetilde{R}_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}\,d\,\mathbb{P}_l(\omega). \qquad (5.14)
\end{aligned}
$$

Step 3. We start by subtracting (5.14) from (5.8). Therefore we have

$$
\begin{aligned}
&\frac{E^{\mathrm{xc}}_{\mathrm{GC},N,s}(\mu)}{N^{1+s/d}}-\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N},s}(\mu)}{\widetilde{N}^{1+s/d}}\\[2mm]
&\geq\; -\frac{C}{M}\frac{N-1}{N^{1+\frac{s}{d}}}R_1^{-s}-c'(d,\epsilon)\sum_{s'\in\{\epsilon,d-\epsilon\}}\left(\int_{\bigcup_{j=1}^{k}(\partial\Lambda_j)_{2l\sqrt{d}}}\rho^{1+s'/d}(x)dx+\frac{1}{M}\int_{\mathbb{R}^d}\rho^{1+s'/d}(x)dx\right)\\[2mm]
&\quad +|B_1|^{s/d}\sum_{j=1}^{k}\left(\frac{\alpha_j}{|\Lambda_j|}\right)^{1+s/d}\Bigg[|(\Lambda_j)_{2l\sqrt{d}}|\frac{M}{(M+C)(M+C_d)}\sum_{i=1}^{M}\int_{\Omega_l}\frac{E^{\mathrm{xc}}_{\mathrm{GC},N\alpha_j|B_{tR_i}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{(N\alpha_j(tR_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}d\,\mathbb{P}_l(\omega)\\[2mm]
&\qquad -\frac{|(\Lambda_j)^*_{2l\sqrt{d}}|}{\widetilde{M}+C_d+1}\sum_{i=1}^{\widetilde{M}}\int_{\Omega_l}\frac{E^{\mathrm{xc}}_{\mathrm{GC},\widetilde{N}\alpha_j|B_{t\widetilde{R}_i}|/|\Lambda_j|,s}\left(\frac{1_{B_1}}{|B_1|}\right)}{(\widetilde{N}\alpha_j(t\widetilde{R}_i)^d|B_1|/|\Lambda_j|)^{1+s/d}}\,d\,\mathbb{P}_l(\omega)\Bigg].\\[2mm]
&:=\; -\frac{C}{M}\frac{N-1}{N^{1+\frac{s}{d}}}R_1^{-s}-c'(d,\epsilon)\sum_{s'\in\{\epsilon,d-\epsilon\}}I(s',\rho,l,M)+II(s,\rho,l,M,\widetilde{M},N,\widetilde{N}) \qquad (5.15)
\end{aligned}
$$

In order to treat the more complicated term $II(s,\rho,l,M,\widetilde{M},N,\widetilde{N})$ in (5.15), in the next paragraphs we fix $l,M,R_1,N,\widetilde{M},\widetilde{R}_1,\widetilde{N}$, satisfying (4.93) and (4.94), (where we define then $R_i=(c(d))^{i-1}R_1$ and $\widetilde{R}_j=(c(d))^{j-1}\widetilde{R}_1$ for $2\leq i\leq M$ and $2\leq j\leq\widetilde{M}$, where $c(d)$ is the constant coming from the Lemma 4.2, depending only on d).

Firstly, we may take $M=\widetilde{M}$, and choose $N,R_1,\widetilde{N},\widetilde{R}_1$ such that $N\geq\widetilde{N}$, and

$$
\widetilde{N}\widetilde{R}_i^d=NR_i^d,\quad\text{for all}\quad i=1,\ldots,M. \qquad (5.16)
$$

Furthermore, we also need $\widetilde{N}\widetilde{R}_i^d\ll\widetilde{N},i=1,\ldots,M$. This can be ensured by imposing the precise relation $\widetilde{R}_1=R_1(N/\widetilde{N})^{\frac{1}{d}}\geq R_1$ and $NR_M^d\ll\widetilde{N}$. The required conditions from Lemma 4.2 for N,M,R_1,l and for $\widetilde{N},M,\widetilde{R}_1,l$ can be written as

$$
l\ll 1\ll M,\quad (M+C)C^M\max\{R_1,\widetilde{R}_1\}=(M+C)C^M\widetilde{R}_1<l,
$$

where we recall that C is a dimensional constant, independent of the choice of the function c.

As we assumed $N \geq \widetilde{N}$ and $C > 1$, we find that, due to (5.16), the above condition is thus equivalent to the Lemma 4.2 condition on the parameters $\widetilde{N}, \widetilde{R}_1, l, M$ and is therefore achievable. The assumptions in (4.96) combined with (5.16) give

$$\widetilde{R}_1 := \widetilde{N}^{-\frac{1}{2d}}, \quad R_1 = \widetilde{N}^{\frac{1}{2d}} N^{-\frac{1}{d}}, \quad l := \widetilde{N}^{-\frac{1}{3d}}, \quad M = \widetilde{M} := \left[\frac{\log \widetilde{N}}{18 d \log C} \right] - 1 . \tag{5.17}$$

The requirements (5.17) together with the requirement $M \geq M_d$ as in the discussion preceding (4.96) turn out to be sufficient for our purposes.

The choices (5.16) and (5.17) imply that for all $i = 1, \ldots, M$, and for all $j = 1, \ldots, k$, we have

$$\int_{\Omega_l} \frac{E_{\mathrm{GC}, N\alpha_j |B_{tR_i}|/|\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(N\alpha_j (tR_i)^d |B_1|/|\Lambda_j|)^{1+s/d}} \, d\,\mathbb{P}_l(\omega) = \int_{\Omega_l} \frac{E_{\mathrm{GC}, \widetilde{N}\alpha_j |B_{t\widetilde{R}_i}|/|\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(\widetilde{N}\alpha_j (t\widetilde{R}_i)^d |B_1|/|\Lambda_j|)^{1+s/d}} \, d\,\mathbb{P}_l(\omega). \tag{5.18}$$

Also note that the $2l\sqrt{d}$-outer layer of Λ_j differs from its $2l\sqrt{d}$-inner layer precisely by the $2l\sqrt{d}$-neighborhood of $\partial\Lambda_j$, thus with the notations introduced in (5.4) and (5.13) we have

$$\left| (\Lambda_j)_{2l\sqrt{d}} \right| = \left| (\Lambda_j)_{2l\sqrt{d}}^* \right| + \left| (\partial\Lambda_j)_{2l\sqrt{d}} \right| . \tag{5.19}$$

Using (5.19) and using the fact that we chose $M = \widetilde{M}$ in (5.17), we can write, for $C > 1$ and $\widetilde{M} \geq M_d$ large enough so that the below leading-order terms in $\widetilde{M}$ below dominate the other terms

$$\begin{aligned}
&|(\Lambda_j)_{2l\sqrt{d}}| \frac{M}{(M+C)(M+C_d)} - \frac{|(\Lambda_j)_{2l\sqrt{d}}^*|}{\widetilde{M}+C_d+1} \\
&= |(\Lambda_j)_{2l\sqrt{d}}^*| \left(\frac{\widetilde{M}}{(\widetilde{M}+C)(\widetilde{M}+C_d)} - \frac{1}{\widetilde{M}+C_d+1} \right) + |(\partial\Lambda_j)_{2l\sqrt{d}}| \frac{\widetilde{M}}{(\widetilde{M}+C)(\widetilde{M}+C_d)} \\
&\leq -(C-1) \frac{|(\Lambda_j)_{2l\sqrt{d}}^*|}{\widetilde{M}^2} + \frac{|(\partial\Lambda_j)_{2l\sqrt{d}}|}{\widetilde{M}} \leq \frac{|(\partial\Lambda_j)_{2l\sqrt{d}}|}{\widetilde{M}}.
\end{aligned} \tag{5.20}$$

Inserting (5.20) in the II-term from (5.15) and keeping in mind the non-positivity (4.31), the bound (C.1) and the fact that $\mathbb{P}_l$ is a probability measure, we get

$$\begin{aligned}
&II(s, \rho, l, M, \widetilde{M}, N, \widetilde{N}) \\
&\geq \frac{\max_{1 \leq j \leq k} |(\partial\Lambda_j)_{2l\sqrt{d}}|}{\widetilde{M}} \sum_{j=1}^{k} |B_1|^{s/d} \left(\frac{\alpha_j}{|\Lambda_j|} \right)^{1+s/d} \sum_{i=1}^{\widetilde{M}} \int_{\Omega_l} \frac{E_{\mathrm{GC}, \widetilde{N}\alpha_j |B_{t\widetilde{R}_i}|/|\Lambda_j|, s}^{\mathrm{xc}} \left(\frac{1_{B_1}}{|B_1|} \right)}{(\widetilde{N}\alpha_j (t\widetilde{R}_i)^d |B_1|/|\Lambda_j|)^{1+s/d}} d\,\mathbb{P}_l(\omega) \\
&\geq -c_{\mathrm{LO}}(d, \epsilon) \left(\max_{1 \leq j \leq k} |(\partial\Lambda_j)_{2l\sqrt{d}}| \right) \sum_{j=1}^{k} |B_1|^{s/d} \left(\frac{\alpha_j}{|\Lambda_j|} \right)^{1+s/d} := -l C_1(\Lambda_1, \ldots, \Lambda_k, \alpha_1, \ldots, \alpha_k, d, \epsilon), \quad (5.21)
\end{aligned}$$

where the crucial ingredient used is the fact that Λ_j are ϕ-regular and in particular $|(\partial\Lambda_j)_r| \leq C_{\Lambda_j} r$ for $r \in (0, 1)$, and the constant C_1 above depends on these bounds C_{Λ_j} as well.

In order to bound the first term of the last line in (5.15), we directly use (5.17) and write

$$R_1^{-s} N^{-\frac{s}{d}} = \left(\frac{\widetilde{N}}{N} \right)^{-\frac{s}{d}} \widetilde{N}_1^{\frac{s}{2d}} N^{-\frac{s}{d}} = \widetilde{N}^{-\frac{s}{2d}}. \tag{5.22}$$

In order to estimate the I-term in (5.15) we treat separately the two integrals appearing in it. For the integration on the neighborhood of the boundaries, we use again the bound $|(\partial\Lambda_j)_r| \leq C_{\Lambda_j} r$ and the explicit form of ρ.

For the integral of $\rho^{1+s'/d}$ with $s' \in \{\epsilon, d-\epsilon\}$, again we can use the explicit form of ρ, this time in an even more direct way. Explicitly, we obtain

$$c'(d,\epsilon) \sum_{s' \in \{\epsilon, d-\epsilon\}} I(s', \rho, l, M)$$

$$\leq c'(d,\epsilon) \left(2l\sqrt{d} \cdot \left(\max_{1 \leq j \leq k} C_{\Lambda_j} \right) \sum_{s' \in \{\epsilon, d-\epsilon\}} \max_{1 \leq j \leq k} \frac{\alpha_i^{1+s'/d}}{|\Lambda_j|^{1+s'/d}} + \frac{1}{M} \sum_{s' \in \{\epsilon, d-\epsilon\}} \sum_{j=1}^{k} \frac{\alpha_i^{1+s'/d}}{|\Lambda_j|^{s'/d}} \right)$$

$$:= \left(l + \frac{1}{M} \right) C_2(\Lambda_1, \ldots, \Lambda_k, \alpha_1, \ldots, \alpha_k, d, \epsilon). \tag{5.23}$$

Summing now the estimates (5.21), (5.22) and (5.23) and taking $C'(\Lambda_1, \ldots, \Lambda_k, \alpha_1, \ldots, \alpha_k, d, \epsilon)$ to be the sum of the analogous constants from those equations, we obtain from (5.15) that (recalling again that we chose $\widetilde{M} = M$)

$$\frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}} - \frac{E_{\mathrm{GC},\widetilde{N},s}^{\mathrm{xc}}(\mu)}{\widetilde{N}^{1+s/d}} \geq \frac{C'(w,\epsilon,d)}{\widetilde{M}} \widetilde{N}^{-\frac{s}{2d}} - \left(l + \frac{1}{\widetilde{M}} \right) C'(\Lambda_1, \ldots, \Lambda_k, \alpha_1, \ldots, \alpha_k, d, \epsilon). \tag{5.24}$$

By the choices (5.17) we now see that $\widetilde{M}^{-1} > l$ and thus its contribution dominates the bound in (5.24), allowing to conclude. □

An immediate consequence of Theorem 1.4, and of the Cauchy criterion of uniform convergence is:

Corollary 5.1 (Uniform convergence of $E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)$ with respect to s). *Fix $0 < \epsilon < d/2$ and let $\epsilon \leq s \leq d - \epsilon$. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be a probability measure with density of the form $\rho(x) = \sum_{i=1}^{k} \alpha_i 1_{\Lambda_i}(x)$ where $\Lambda_1, \ldots, \Lambda_k$ are Borel sets with ϕ-regular boundary and disjoint interiors. Then the sequence of functions*

$$f_s(N) := \frac{E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu)}{N^{1+s/d}}$$

converges as $N \to \infty$ uniformly with respect to the parameter $s \in [\epsilon, d - \epsilon]$.

Remark 5.2. *As detailed also in the introduction, both Theorem 1.4 and Corollary 5.1 can be shown for more general densities, such as any positive Riemann-integrable density, including in particular continuous densities with compact support and ϕ-regular boundary. The proof follows by a similar argument as the proof of Theorem 1.4, combined with an adaptation of the small oscillations proof for piece-wise constant densities and using also the small oscillations inequality proved above for the uniform density on $1_{B_1}/|B_1|$.*

A   Fefferman decomposition and positive definiteness

A.1   Derivative bounds and positive definite control of the error

We present here the main ingredients for the generalization of Fefferman's and Gregg's approach to effectively localizing our interaction kernels. We restrict to the case $c(x - y) = |x - y|^{-s}, 0 < s < d$, though the methods below can be applied to more general costs of the form $c(x - y)$, under suitable regularity assumptions.

A.1.1   Fefferman-Gregg positive definiteness criterion

In this subsection we present a generalization to arbitrary $0 < s < d$ of the main lemma and of the kernel decomposition from [Ge89], given there for $d = 3$ and $0 < s < 3$.

We recall the multi-index notation. We will denote by $\beta = (\beta_1, \ldots, \beta_d) \in \mathbb{N}^d$ a multiindex and its length will be defined as $|\beta| := \beta_1 + \cdots + \beta_d$. For a function $f : \mathbb{R}^d \to \mathbb{R}$, for $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$ we then use the partial differentiation and monomial notations

$$\partial_x^\beta f(x) := \partial_{x_1}^{\beta_1} \partial_{x_1}^{\beta_2} \cdots \partial_{x_d}^{\beta_d} f(x), \quad x^\beta := x_1^{\beta_1} x_2^{\beta_1} \cdots x_d^{\beta_d}, \quad \xi^\beta := \xi_1^{\beta_1} \xi_2^{\beta_1} \cdots \xi_d^{\beta_d}. \tag{A.1}$$

In particular $\partial_x^0 f(x) = f(x)$.

Lemma A.1. *Let $0 < s < d$ and consider a kernel $g : \mathbb{R}^d \to \mathbb{R}$ which is d times differentiable on $\mathbb{R}^d \setminus \{0\}$. Assume that there exists $\widetilde{C}(d, \epsilon) > 0$, depending only on d, ϵ, if $0 < \epsilon < d/2$ and $\epsilon \leq s \leq d - \epsilon$, such that for all $x \neq 0$, for $|\beta| = 0$ and for all β with $|\beta| = d$, there holds*

$$|\partial_x^\beta g(x)| \leq \widetilde{C} |x|^{-s-|\beta|} . \tag{A.2}$$

Then there exists a constant $c(d, \epsilon) > 0$, also depending only on d, ϵ for our choices of s, d, ϵ, and such that for all $\xi \neq 0$ there holds

$$|\hat{g}(\xi)| \leq c(d, \epsilon) |\xi|^{s-d} . \tag{A.3}$$

The proof of the lemma follows the same steps as in the beginning of the proof of the main lemma in [Ge89, p. 257-258, down to the line after eq. (3)], with the extra verification that the bounds in that paper are explicit and depend on s in a polynomial way, without exploding in the interval $s \in (\epsilon, d - \epsilon)$. The details will be omitted. As an immediate consequence of the above lemma, we have

Corollary A.2 (positive definiteness criterion). *Under the hypotheses of Lemma A.1, there exists a constant $C > 0$ depending only on d, ϵ, such that $|x|^{-s} \pm C g(x)$ is positive definite. In particular, if $c(d, \epsilon)$ is the constant appearing in Lemma A.1 and the Fourier transform of $|x|^{-s}$ is $c_{s,d} |\xi|^{d-s}$ then we can choose $C = \frac{\inf_{s \in (\epsilon, d-\epsilon)} c_{s,d}}{c(d, \epsilon)}$.*

We will now prove that estimates of the type (A.2) hold for two functions built from $Q_{i,\eta}$ and which will be relevant in our constructions. This is reminiscent of, and can be seen as a generalization of the bounds [Ge89, p. 272, for the functions called $k_{\dagger}^i$], which are there stated in $d = 3$ and not proved explicitly. Due to this, and in order to provide an explicit proof in general dimension, we include here the full proof. In preparation for the more difficult result of Lemma A.4 below, which really uses the geometric growth of the R_i from Lemma 4.2, we prove an easier bound:

Lemma A.3. *Let $0 < \epsilon < d/2$ and $\epsilon \leq s \leq d - \epsilon$, and let $Q_{i,\eta}$ be defined as in (4.17), for R_i defined as in Lemma 4.2. Then there exists $C(\rho_\eta, d, \epsilon) > 0$ depending only on the choice of ρ_η and of d, ϵ, such that the bounds (A.2) hold for $|\beta| \leq d + 1$ for*

$$\frac{1}{C(\rho_\eta, d, \epsilon)} \frac{Q_{i,\eta}(x)}{|x|^s} . \tag{A.4}$$

Proof. We first note that

$$\partial_x^\beta \left(\frac{Q_{i,\eta}(x)}{|x|^s} \right) = \frac{1}{R_i^{s+|\beta|}} \left. \partial_x^\beta \left(\frac{Q_{0,\eta}(x')}{|x'|^s} \right) \right|_{x' = \frac{x}{R_i}} . \tag{A.5}$$

therefore, looking at (A.2), we see that it suffices to prove the bound for $R_i = 1$ (i.e. for the case where we replace $Q_{i,\eta}$ by $Q_{0,\eta}$), and the one for general R_i will directly follow from it via (A.5). We note that due to the mollification (4.16), partial derivatives of $Q_{0,\eta}$ up to order $|\beta| = d$ satisfy supremum bounds depending only on ρ_η, d, and the partial derivatives up to order $|\beta| = d$ of $C^{-1} |x|^{-s}$ satisfy (A.2) whenever $C > c(d, \epsilon)$ for some constant depending only on d, ϵ. The claim follows by triangular inequality, if we distribute the $|\beta| \leq d$ partial derivatives on the two terms of the product $|x|^{-s} Q_{0,\eta}(x)$. $\square$

The above estimate (A.4) will not be directly helpful for us, because as we sum it over the M terms corresponding to $i = 1, \ldots, M$, we will obtain a bound dependent on M. The following subtler result uses crucially the geometric growth of the R_i from Lemma 4.2, and requires that for each i we subtract a tamer kernel which effectively cancels the tail behaviour of $|x|^{-s}(1 - Q_{i,\eta}(x))$ at infinity.

Lemma A.4. *Let $0 < \epsilon < d/2$ and $\epsilon \leq s \leq d - \epsilon$, and let $Q_{i,\eta}$ be defined as in (4.17), for R_i defined as in Lemma 4.2. Then there exists $C(\rho_\eta, d, \epsilon) > 0$ depending only on the choice of ρ_η and of d, ϵ, such that the bounds (A.2) hold for $|\beta| \leq d + 1$ for*

$$\frac{1}{C(\rho_\eta, d, \epsilon)} \sum_{i=1}^{M} \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy \right) . \tag{A.6}$$

We then prove the following:

Proposition A.5. *With the above notations, the kernels*

$$\frac{C(\rho_\eta, d, \epsilon)}{M}|x|^{-s} \pm \frac{1}{M}\sum_{i=1}^{M}\left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y)dy\right)^{-1}\int_{\mathbb{R}^d}\frac{Q_{i,\eta}(y)}{|x-y|^s}dy\right) \tag{A.7}$$

and

$$\frac{1}{M}\sum_{i=1}^{M}\left[\left(\int_{\mathbb{R}^d} Q_{i,\eta}(y)dy\right)^{-1}\int_{\mathbb{R}^d}\frac{Q_{i,\eta}(y)}{|x-y|^s}dy\right] , \tag{A.8}$$

as well as

$$\frac{C(\rho_\eta, d, \epsilon)}{|x|^s} \pm \frac{Q_{i,\eta}(x)}{|x|^s}, \qquad \text{for every } 1 \le i \le M , \tag{A.9}$$

are positive definite.

Proof of Proposition A.5: The first claim follows from lemmas A.4 and A.1. As the bounds required in these lemmas do not depend on sign, both signs are allowed. The second claim follows by using the fact that the Fourier transform transforms convolutions into products, that $|x - y|^{-s}$ is positive definite, and the structure (4.17) of $Q_{i,\eta}$ where the integral over r can be commuted with the Fourier transform, by linearity. The last claim follows directly from Lemma A.3 via lemma A.1, exactly like the first claim. □

Proof of Lemma A.4: We will estimate first separately the terms in the large parenthesis from (A.6) at fixed $i = 1, \dots, M$, then we will combine the estimates in the end. We will distinguish below between terms in B_{3R_i} and terms outside of B_{3R_i}, because the convolution $1_{B_{3R_i/2}} * 1_{B_{3R_i/2}}$ has support in B_{3R_i}.

Step 1: *The first term on B_{3R_i}.*

We will prove a separate bound for each of the two terms in the sum (A.6) and then use the triangle inequality to sum all the terms. By keeping in mind the definition (4.17), we find that

$$\partial_x^\beta\left(\frac{1 - Q_{i,\eta}(x)}{|x|^s}\right) = \frac{1}{R_i^{s+|\beta|}}\partial_x^\beta\left(\frac{1 - Q_{0,\eta}(x')}{|x'|^s}\right)\Bigg|_{x'=\frac{x}{R_i}} . \tag{A.10}$$

We note that $f_\eta(|x|) := 1 - Q_{0,\eta}(x)$ is a radial increasing function which vanishes at zero. Due to the definition (4.16) and to the explicit form of the function (4.15) of which $Q_{0,\eta}$ is a superposition, f_η has a finite right derivative at zero: Indeed, the function appearing in (4.15) can be directly differentiated in $|x|$ and the averaging in (4.16) preserves these bounds. Due to the definition (4.16) and to the explicit form of the function (4.15) of which $Q_{0,\eta}$ is a superposition, f_η has a finite right derivative at zero: Indeed, the function appearing in (4.15) can be directly differentiated in $|x|$ and the averaging in (4.16) preserves these bounds. Therefore by Taylor expansion, and since f_η only depends on ρ_η and d and we assumed that $\epsilon \le s \le d - \epsilon$, we have the following control for $|\beta| = 0$ and $x \ne 0$:

$$\frac{1 - Q_{0,\eta}(x)}{|x|^s} \le |x|^{1-s}\lim_{r\downarrow 0}\frac{f_\eta(r)}{r} + C_{\rho_\eta, d}|x|^{1-s} = C(\rho_\eta, d, \epsilon)|x|^{1-s} . \tag{A.11}$$

Similarly, based on the formulas (4.15) and (4.16) and on the boundedness properties of ρ_η, we see that the norms of all the derivatives $\partial_x^\beta Q_{0,\eta}$ have a finite bound near zero, and thus all the partial derivatives are *uniformly* bounded on $\mathbb{R}^d \setminus \{0\}$. This means that for $0 < |\beta| \le d+1$, again using the bound (A.11) for the term where all the partial derivatives fall on $|x|^{-s}$ as we apply the Leibnitz rule to distribute derivatives over the product $Q_{0,\eta}(x)|x|^{-s}$, we find the following bounds valid for $x \ne 0$:

$$\begin{aligned}
\left|\partial_x^\beta\left(|x|^{-s}(1 - Q_{0,\eta}(x))\right)\right| &\le \sum_{\beta'+\beta''=\beta}\left|\partial_x^{\beta'}(|x|^{-s})\right|\left|\partial_x^{\beta''}(1 - Q_{0,\eta}(x))\right| \\
&\le C_1(\beta)(|x|^{-s-|\beta|}|(1 - Q_{0,\eta}(x))| \\
&\quad + C_2(\beta)|x|^{1-s-|\beta|}\sum_{|\beta'|=1}\left|\partial_x^{\beta'}Q_{0,\eta}(x)\right| + O(|x|^{2-s-|\beta|}) \quad \text{as} \quad |x| \to 0 \\
&\le C(\rho_\eta, d, \epsilon)|x|^{1-s-|\beta|} + O(|x|^{2-s-|\beta|}) \quad \text{as} \quad |x| \to 0 ,
\end{aligned} \tag{A.12}$$

48

where the constants $C_1(\beta), C_2(\beta)$ are the necessary bounds valid for all $\epsilon \leq s \leq d - \epsilon$, which we can verify to be depending only on d, ϵ for $0 < \epsilon < d/2$. The bounds (A.12) and (A.11) can be applied to the r.h.s. of (A.10) to give

$$\left| \partial_x^\beta \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} \right) \right| \leq C(\rho_\eta, d, \epsilon) \frac{1}{R_i} |x|^{1-s-|\beta|} + o(|x|^{1-s-|\beta|}) \quad \text{as} \quad |x| \to 0 . \tag{A.13}$$

Also note that the expansion in powers of $|x|$ from the bounds (A.12) and (A.11) still holds on the whole of B_3 (i.e. on the whole support of $Q_{0,\eta}$), as all the growth behavior of different powers of $|x|$ does not change for $|x| \in (0,1)$, whereas for $|x| \in [1,3]$ the powers of $|x|$ contributing to leading order to the estimates are equivalent up to a constant factor depending only on ϵ, d. Similarly, for each fixed $\beta \geq 0$ the derivatives $\partial_x^\beta Q_{0,\eta}$ are thus all uniformly bounded in x, with a constant depending only on $d, |\beta|$. Then we can appeal to the same scaling reasoning as in passing from (A.12) to (A.13), which leads to the fact that also on the whole of B_{3R_i} the bounds (A.13) extend, with a finite constant depending on $d, \epsilon, |\beta|$. As we are concerned only with the finitely many (and whose number depends only on d) choices $0 \leq |\beta| \leq d + 1$, the constant we obtain effectively depends only on s, d. Summarizing this reasoning, at the cost of increasing the value of the constant $C(\rho_\eta, d, \epsilon)$, we still have for $0 \leq |\beta| \leq d + 1$

$$\left| \partial_x^\beta \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} \right) \right| \leq C(\rho_\eta, d, \epsilon) \frac{1}{R_i} |x|^{1-s-|\beta|} \quad \text{for} \quad |x| \leq 3R_i . \tag{A.14}$$

For the remaining term in (A.6) we have a similar procedure, but it is better controlled near 0, as we now show. First note that $\int_{\mathbb{R}^d} Q_{i,\eta}(x) dx = C(\rho_\eta, d) R_i^d$ and $\int_{B_R(x)} |y|^{-s} dy \leq \int_{B_R} |y|^{-s} dy \leq c(d, \epsilon) R^{d-s}$. Using these, for $0 \leq |\beta| \leq d + 1$ we have:

$$\left| \partial_x^\beta \left(\left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy \right) \right| \leq \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{|\partial_x^\beta Q_{i,\eta}(x-y')|}{|y'|^s} dy'$$

$$\leq C(\rho_\eta, d, \epsilon) \frac{\sup_y |\partial_x^\beta Q_{0,\eta}(y)|}{R_i^{|\beta|}} \frac{1}{R_i^d} \int_{B_{2(1+\eta)R_i}(x)} \frac{dy}{|y|^s} \leq \frac{C(\rho_\eta, d, \epsilon)}{R_i^{s+|\beta|}} \quad \text{for} \quad |x| < 3R_i . \tag{A.15}$$

Note that in the range $|x| < 3R_i$ the bound (A.15) is stronger than (A.14).

Step 3: *Bound outside B_{3R_i}, and final estimates at fixed i.*

For $|x| > 3R_i$ we note that $Q_{i,\eta}(x) = 0$ and the first term from (A.6) simplifies as $(1 - Q_{i,\eta}(x))|x|^{-s} = |x|^{-s}$ in this range. Thus, using the fact that the integral below is with respect to a probability measure in order to pass the term $|x|^{-s}$ inside the integral, and noting that the x-derivatives can be equally passed inside the integral, we are required to bound the following:

$$\left| \int_{B_{2(1+\eta)R_i}} \frac{Q_{i,\eta}(y)}{\int Q_{i,\eta}(z) dz} \partial_x^\beta \left(\frac{1}{|x|^s} - \frac{1}{|x-y|^s} \right) dy \right|$$

$$\leq C(\rho_\eta, d, \epsilon) C(\beta) \int_{B_{3R_i}} \frac{Q_{i,\eta}(y)}{\int Q_{i,\eta}(z) dz} \left| \frac{1}{|x|^{s+|\beta|}} - \frac{1}{|x-y|^{s+|\beta|}} \right| dy . \tag{A.16}$$

By using the fact that $|y| \leq 3R_i$ and the Taylor expansion to first order of $|x|^{-s-|\beta|}$ near x, we have the bound by $R_i |x|^{-1-s-|\beta|}$, for $0 \leq |\beta| \leq d + 1$ in (A.16) for $|x| > 3R_i$. Adding this bound to the bounds (A.14) and (A.15), we find

$$\left| \partial_x^\beta \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy \right) \right|$$

$$\leq \frac{1}{|x|^{s+|\beta|}} \cdot \begin{cases} C(\rho_\eta, d, \epsilon) \frac{|x|}{R_i} & \text{for } |x| \leq 3R_i , \\ C(\rho_\eta, d, \epsilon) \frac{R_i}{|x|} & \text{for } |x| \geq 3R_i . \end{cases} \tag{A.17}$$

Step 4: *Summing over i.*

49

For $r > 0$ let $i_r \in \{0, \ldots, M\}$ be defined as follows:

$$i_r := \begin{cases} 0 \text{ if } r < 3R_1 \ , \\ \text{such that } r \in (3R_{i_r}, 3R_{i_r+1}] \text{ if } R_1 < r \le R_M \ , \\ M \text{ if } r > R_M \ . \end{cases} \tag{A.18}$$

Then by using (A.17) we find that

$$\left| \partial_x^\beta \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy \right) \right| \le \frac{C(\rho_\eta, d, \epsilon)}{|x|^{s+|\beta|}} \left(\sum_{i=1}^{i_{|x|}} \frac{R_i}{|x|} + \sum_{i=i_{|x|}}^{M} \frac{|x|}{R_i} \right) \ . \tag{A.19}$$

Denote by $\bar{C} := 1 + 4\sqrt{d}|B_1| > 1$ the constant appearing in Lemma 4.2 and such that $R_{i+1} > \bar{C}R_i$ for all $1 \le i \le M$. Then

$$\text{for } i \le i_{|x|} \text{ there holds} \qquad |x| \ge 3R_{i_{|x|}} \ge 3\bar{C}^{i_{|x|}-i}R_i, \ \text{ therefore } \frac{R_i}{|x|} \le \frac{1}{3}\frac{1}{\bar{C}^{i_{|x|}-i}} \ , \tag{A.20a}$$

$$\text{for } i > i_{|x|} \text{ there holds} \qquad |x| \le 3R_{i_{|x|}+1} \le 3\bar{C}^{i_{|x|}-i+1}R_i, \ \text{ therefore } \frac{|x|}{R_i} \le 3\bar{C}\frac{1}{\bar{C}^{i-i_{|x|}}} \ . \tag{A.20b}$$

Since $\sum_{i \ge 0} \bar{C}^{-i} \le \frac{2\bar{C}-1}{\bar{C}-1}$ (which up to change of index can be used in both above cases), and $\bar{C}$ depends only on d, we find from (A.19) and (A.20) that the sum over $i = 1, \ldots, M$ of the terms appearing in (A.17) is bounded by $C(\rho_\eta, d, \epsilon)|x|^{-s-|\beta|}$, and this holds for all $0 \le |\beta| \le d+1$, which proves the claim. $\qquad \square$

A.1.2 Properties of w

We recall now that from (4.23) we have

$$w(x_1 - x_2) := \left(1 + \frac{C}{M}\right)|x_1 - x_2|^{-s} - \int_{\Omega_l} \left(\sum_{A \in F_\omega^l} \frac{1_A(x_1)1_A(x_2)}{|x_1 - x_2|^s} \right) d\mathbb{P}_l(\omega)$$

$$\overset{(4.20)}{=} \left(1 + \frac{C}{M}\right)|x_1 - x_2|^{-s} - \sum_{i=1}^{M} c_i \frac{Q_{i,\eta}(x_1 - x_2)}{|x_1 - x_2|^s}$$

$$= \frac{C}{M}|x_1 - x_2|^{-s} + \frac{1}{M}\sum_{i=1}^{M} \frac{1 - Q_{i,\eta}(x_1 - x_2)}{|x_1 - x_2|^s} + \sum_{i=1}^{M} \left(\frac{1}{M} - c_i\right) \frac{Q_{i,\eta}(x_1 - x_2)}{|x_1 - x_2|^s}.$$

$$= \frac{C}{2M}|x_1 - x_2|^{-s} + \frac{1}{M}\sum_{i=1}^{M} \left(\frac{1 - Q_{i,\eta}(x_1 - x_2)}{|x_1 - x_2|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y)dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x_1 - x_2 - y|^s}dy \right)$$

$$+ \frac{1}{M}\sum_{i=1}^{M} \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y)dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x_1 - x_2 - y|^s}dy + \frac{C}{2M}|x_1 - x_2|^{-s}$$

$$+ \sum_{i=1}^{M} \left(\frac{1}{M} - c_i\right) \frac{Q_{i,\eta}(x_1 - x_2)}{|x_1 - x_2|^s}$$

where for $0 < \epsilon < d/2$ and $\epsilon \le s \le d - \epsilon$ we can fix $C = C(\rho_\eta, d, \epsilon) > 0$ depending only on the choice of ρ_η, on d and on ϵ, the form of which will be used in Proposition 1.6 above and is made explicit in Lemma A.4.

In order to facilitate the discussions that follow, and to be able to analyse the properties of w, we will find it useful to introduce the following notations:

$$w_\#(x) := \frac{C}{2M}|x|^{-s} + \frac{1}{M}\sum_{i=1}^{M} \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y)dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s}dy \right) \ , \tag{A.21a}$$

$$w_1(x) = \frac{1}{M} \sum_{i=1}^{M} \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy \ , \tag{A.21b}$$

and

$$w_2(x) := \frac{C}{2M} |x|^{-s} + \sum_{i=1}^{M} \left(\frac{1}{M} - c_i \right) \frac{Q_{i,\eta}(x)}{|x|^s} \ . \tag{A.21c}$$

Then

$$w(x) = w_\#(x) + w_1(x) + w_2(x) \ .$$

We observe that, as explained in Lemma A.7, the form chosen for $w_\#$ and w_1 is used as much for positive definiteness as to derive the rough next order lower bound. More precisely, the w_1 term is key to us being able to control the tail of $w_\#$ in such a way that the error from $w_\#$ is of order $1/M$. Moreover, we have

Lemma A.6. *Let $0 < \epsilon < d/2$ and $\epsilon \le s \le d - \epsilon$. Then there exists a constant $C = C(\rho_\eta, d, \epsilon)$ depending only on the choice of ρ_η, on d and on ϵ such that for $w_\#, w_1$ and w_2 defined as in (A.21a), (A.21b) and (A.21c), the following properties hold:*

1. *$w_\#(x)$ is positive definite.*

2. *w_1 is continuous, positive definite and $w_1(x) \le \frac{C}{M} R_1^{-s}$.*

3. *w_2 is positive definite.*

Proof. The first point and the positive definiteness part of the second point are precisely the first two statements of Proposition A.5.

The boundedness of $w_1(x)$ follows by noting that due to the scaling (4.17), there holds $\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy = C(\rho_\eta, d) R_i^d$ and due to the upper bound (4.18) and to the fact that $\eta \le 1/2$ we have for each fixed i that

$$\frac{1}{\int Q_{i,\eta}(z) dz} \int \frac{Q_{i,\eta}(x-y)}{|y|^s} dy \ \le \ \frac{1}{C(\rho_\eta, d) R_i^d} \int_{B_{3R_i}(x)} |y|^{-s} dy$$

$$\le \ \frac{1}{C(\rho_\eta, d) R_i^d} \int_{B_{3R_i}} |y|^{-s} dy \ = \ \frac{C(\rho_\eta, d, \epsilon)}{R_i^s} \ .$$

As R_i forms a geometric series, summing this bound over i and dividing by M gives the desired bound on w_1, up to increasing $C(\rho_\eta, d, \epsilon)$ to a value which for $0 < \epsilon < d/2$ and $\epsilon \le s \le d - \epsilon$ still depends only on ρ_η, d, ϵ.

For proving the positive definiteness in the last point we recall the bounds from the Swiss cheese lemma, (4.13), from which it follows that for each i we have

$$\frac{1}{M} - c_i \in \left[\frac{1}{M} - \frac{1}{M + C_d}, \frac{1}{M} - \frac{1}{M + C_d + 1} \right] \subset \frac{1}{M^2} \left[\frac{C_d}{1 + C_d}, C_d + 1 \right] \ . \tag{A.22}$$

Together with (A.9) from Proposition (A.5) this proves that for every $i = 1, \ldots, M$ and up to increasing the value of $C = C(\rho_\eta, d, \epsilon)$ to a value still depending only on ρ_η, d, ϵ, we have that $M^{-2} C |x|^{-s} \pm (M^{-1} - c_i) |x|^{-s} Q_{i,\eta}(x)$ is positive definite. The constants are increased only a finite number of times (with the upper bound on that number depending only on d). Up to replacing g from Lemma A.1 by g/C, we obtain that a bound on the derivatives of g by a $C |x|^{-s-|\beta|}$ implies that $C' |x|^{-s} - g(x)$ is positive definite, where $C' = c(d, \epsilon) C$ and $c(d, \epsilon)$ is bounded and depending only on d, ϵ if $0 < \epsilon < d$ and if s is in the range $(\epsilon, d - \epsilon)$ as before. In other words, use of the Lemma 4.2 just introduces yet another constant, depending only on d and on ϵ.

Summing the terms over $i = 1, \ldots, M$, we find the last claim of our lemma, and this concludes the proof. $\square$

A.2 Rough next-order lower bound for $w_\#$ and w_2

The present section produces a lower bound for the $E_{N,c}^{xc}(\mu)$ energy in the case of costs d of a special form, inspired by the above constructions. This is based on the proof of a rough fractional Lieb-Oxford inequality from

[LNP16, Appendix], which itself is inspired by [LSY94, Lem. 5.3] (see also [Li79], [Li83], [LO81]. Translated to our notation, the bound proved in [LNP16, Lem. 16] states that for cost $\mathsf{c}(x,y) = |x-y|^{-s}$ for $0 < s < d$, and for any transport plan $\gamma_N \in \mathcal{P}_{sym}((\mathbb{R}^d)^N), \gamma_N \mapsto \mu$, for $d\mu(x) = \rho(x)dx$ with $\rho \in L^{1+s/d}(\mathbb{R}^d)$, there holds

$$\int \sum_{\substack{1 \leq i,j \leq N \\ i \neq j}} \mathsf{c}(x_i, x_j) d\gamma_N(x_1, \ldots, x_N) - N^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathsf{c}(x,y)\rho(x)\rho(y)dx\,dy$$

$$\geq -CN^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx \ . \quad (A.23)$$

The proof of (A.23) done in [LNP16] is based on the following radial decomposition formula, that seems to first have been used by Fefferman-De la Llave [FL86] in a statistical mechanics context:

$$\frac{1}{|x-y|^s} = c(s,d) \int_0^\infty \int_{\mathbb{R}^d} 1_{B_r}(x-u)1_{B_r}(y-u)du\frac{dr}{r^{d+1+s}} \quad \text{for} \quad x \neq y \ , \quad (A.24)$$

which is well-defined for $s > 0$ (the reason why (A.23) only holds for $0 < s < d$, is that for $s \geq d$ the energy in (A.23) is $-\infty$ as $|x-y|^{-s}$ stops being integrable near $x = y$).

Even though the lemma below can be shown for much more general costs, we restrict ourselves to showing it only for $w_\#, w_1$ and w_2. We are now ready to show:

Lemma A.7. *Let $0 < \epsilon < d/2$ and $\epsilon \leq s \leq d - \epsilon$, and let $w_\#, w_1$ and w_2 be defined as in (A.21a), (A.21b) and (A.21c). Then for all $\mu \in \mathcal{P}(\mathbb{R}^d)$ with density $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$ we have*

$$E_{N,w_\#}^{\text{xc}}(\mu) \quad \geq \quad -\frac{C(w_\#, d, \epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx \ , \quad (A.25)$$

$$E_{N,w_2}^{\text{xc}}(\mu) \quad \geq \quad -\frac{C(w_2, d, \epsilon)}{M} N^{1+s/d} \int_{\mathbb{R}^d} \rho^{1+s/d}(x)dx \quad (A.26)$$

and

$$E_{N,w_1}^{\text{xc}}(\mu) \geq -\frac{C(\rho_\eta, d, \epsilon)}{M} R_1^{-s}(N-1) \ , \quad (A.27)$$

for some constants $C(w_\#, d, \epsilon), C(w_2, d, \epsilon) > 0$, which are independent of the choices of ρ, N and $s \in [\epsilon, d - \epsilon]$, and $C(\rho_\eta, d, \epsilon)$ is the one of Lemma A.6.

Similar statements as above, and with the same lower bounds constants, hold also for the $E_{\text{GC},N}^{\text{OT}}$ versions.

Proof. The proof of the inequality involving w_1 follows immediately from Lemma 3.6 in [FLS15] due to its boundedness, and will be omitted. We focus next on the remaining two inequalities, whose proof relies partly on an adaptation of the proof of [LNP16, Lem. 16], the main ideas of which we briefly sketch it below. A key factor in our reasoning is that, since w_1 cancels the tail behaviour of $w_\#$, we preserve the factor $1/M$.

To begin with, we will write $w_\#$ and w_2 in a form similar to (A.24). Since the reasoning is the same for both costs, we will only detail below the argument for $w_\#$. To this purpose, we will first apply [HS02, Thm. 1 and eqn. (11)] to $w_\#(x)$. This is possible due to the fact that, by the arguments in Lemma A.5, $w_\#$ satisfies all the conditions of [HS02, eqn. (11)].

More precisely, we are going to use the following general representation as proved in [HS02, Thm. 1]: Let $V : \mathbb{R}^d \to \mathbb{R}$ be a radial function which is $d + 1$ times differentiable away from $x = 0$. Assume also that $\lim_{|x| \to \infty} |x|^m \partial_{|x|}^m V(x) = 0$ for all $0 \leq m \leq [d/2] + 1$. Then

$$V(x) = \int_0^\infty \int_{\mathbb{R}^d} 1_{B_{r/2}}(u)1_{B_{r/2}}(x-u)f(r)du\,dr \ , \quad (A.28)$$

where

$$f(r) = \frac{(-1)^{d+1}}{\Gamma([d/2]+2)} \frac{2}{(\pi r^2)^{(d-1)/2}} \int_r^\infty V^{(d+1)}(v)v(v^2-r^2)^{(d-3)/2}dv \ , \quad (A.29)$$

and where by abuse of notation $V(v) = V(|x|)$, with $|x| = v$. Let

$$w_*(|x|) := \sum_{i=1}^{M} \left(\frac{1 - Q_{i,\eta}(x)}{|x|^s} - \left(\int_{\mathbb{R}^d} Q_{i,\eta}(y) dy \right)^{-1} \int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x - y|^s} dy \right) ,$$

which is a radial function since $Q_{i,\eta}(x)$ is a radial function by (4.17), and $\int_{\mathbb{R}^d} \frac{Q_{i,\eta}(y)}{|x-y|^s} dy$ is radial since it is a convolution of radial functions. In view of Lemma A.4 and of [HS02, Thm. 1], we have for w_* the representation

$$w_*(|x|) = \int_0^\infty \int_{\mathbb{R}^d} 1_{B_{r/2}}(u) 1_{B_{r/2}}(x - u) f_2(r) du \, dr$$

for which we use the expression of the type (A.29) for $f_2(r)$, depending on the $d+1$-th derivative of w_*. More precisely we have for a constant $c(d) > 0$ depending only on the dimension,

$$f_2(r) = (-1)^{d+1} c(d) \int_r^\infty w_*^{(d+1)}(v) v (v^2 - r^2)^{\frac{d-3}{2}} dv . \tag{A.30}$$

If now $f_1(r)$ is the weight f corresponding to $V(x) = C|x|^{-s}$ as obtained from (A.29), we also have

$$f_1(r) = c(d) C \int_r^\infty \frac{1}{v^{s+d+1}} v (v^2 - r^2)^{\frac{d-3}{2}} dv ,$$

because of the fact that the $d+1$-th derivative of $|x|^{-s}$ has sign $(-1)^{d+1}$. Therefore

$$w_\#(x, y) = \frac{C}{M}|x - y|^{-s} + \frac{1}{M} w_*(x - y) = \frac{1}{M} \int_0^\infty 1_{B_{r/2}}(x - u) 1_{B_{r/2}}(y - u)(f_1(r) + f_2(r)) du dr ,$$

where

$$f_1(r) + f_2(r) = c(d) \int_r^\infty \left(C v^{-s-d-1} + (-1)^{d+1} w_*(v)^{(d+1)} \right) v (v^2 - r^2)^{\frac{d-3}{2}} dv . \tag{A.31}$$

Note here that due to the bound (A.6) for $|\beta| \le d+1$, we find that up to enlarging the above constant C by a factor depending only on d, ϵ for our choice of s, there holds for a constant $\widetilde{C} > 0$ depending only on d, ϵ, C,

$$\widetilde{C} v^{-s-d-1} \ge C v^{-s-d-1} + (-1)^{d+1} w_*(v)^{(d+1)} \ge C v^{-s-d-1} - |w_*(v)^{(d+1)}| \ge 0 , \tag{A.32}$$

and thus, due to the fact that the weight $v(v^2 - r^2)^{(d-3)/2}$ appearing in (A.31) is positive and to the fact that $f_1(r)$ can be equivalently re-expressed also via (A.24), we find from (A.32) that for some constant $\bar{C}(d, \epsilon)$ depending only on $d, \epsilon, w_\#, \rho_\eta$ there holds

$$0 \le f_1(r) + f_2(r) \le c(d) \widetilde{C} \int_r^\infty \frac{1}{v^{s+d+1}} v (v^2 - r^2)^{\frac{d-3}{2}} dv \le \frac{\bar{C}(d, \epsilon)}{r^{d+s+1}} . \tag{A.33}$$

Fix now $\mu \in \mathcal{P}(\mathbb{R}^d)$ with density $\rho \in L^{\frac{d}{d-s}, 1}(\mathbb{R}^d)$ as well. Since $f_1 + f_2 \ge 0$, we can proceed next as in the proof of [LNP16, Lem. 16]. As in [LNP16, (79) and (80)], we have

$$\int_{\mathbb{R}^{2d}} w_\#(x, y) d\rho(x) d\rho(y) = \frac{1}{M} \int_0^\infty \int_{\mathbb{R}^d} h_{r,\rho}^2(u)(f_1(r) + f_2(r)) \, du \, dr ,$$

where

$$h_{r,\rho} := \rho * 1_{B_{r/2}} .$$

Moreover, for any $\gamma_N \in \mathcal{P}(\mathbb{R}^{Nd}), \gamma_N \mapsto \mu$, we have

$$\int_{\mathbb{R}^{Nd}} \sum_{1 \le i \ne j \le N} w_\#(x_i, x_j) \, d\gamma_N(x_1, \ldots, x_N) = \frac{1}{M} \int_0^\infty \int_{\mathbb{R}^d} k_{r,\rho}(u)(f_1(r) + f_2(r)) \, du \, dr ,$$

where

$$k_{r,\rho}(u) := \int_{\mathbb{R}^{Nd}} \sum_{1 \le i \ne j \le N} 1_{B_{r/2}}(x_i - u) 1_{B_{r/2}}(x_j - u) \, d\gamma_N(x_1, \ldots, x_N) .$$

By the same Cauchy-Schwarz reasoning used in [LNP16] to obtain [LNP16, eqn. (81)], we get

$$
\begin{aligned}
k_{r,\rho}(u) &= \int_{\mathbb{R}^{Nd}} \left(\sum_{i=1}^{N} 1_{B_{r/2}}(x_i - u) \right)^2 d\gamma_N(x_1, \ldots, x_N) - \int_{\mathbb{R}^{Nd}} \left(\sum_{i=1}^{N} 1_{B_{r/2}}(x_i - u) \right) d\gamma_N(x_1, \ldots, x_N) \\
&\geq \left[\int_{\mathbb{R}^{Nd}} \left(\sum_{i=1}^{N} 1_{B_{r/2}}(x_i - u) \right) d\gamma_N(x_1, \ldots, x_N) \right]^2 - \int_{\mathbb{R}^{Nd}} \left(\sum_{i=1}^{N} 1_{B_{r/2}}(x_i - u) \right) d\gamma_N(x_1, \ldots, x_N) \\
&= N^2 \left[\int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\rho(x)\,dx \right]^2 - N \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\rho(x)dx = N^2 h_{r,\rho}^2(u) - N h_{r,\rho}(u) \\
&\geq N^2 h_{r,\rho}^2(u) - \min(N h_{r,\rho}(u), N^2 h_{r,\rho}^2(u)), \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{(A.34)}
\end{aligned}
$$

where for the last inequality we used that $k_{r,\rho}(u) \geq 0$. We therefore have

$$
\begin{aligned}
\int \sum_{1 \leq i,j \leq N} & w_{\#}(x_i, x_j) d\gamma_N(x_1, \ldots, x_N) - N^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} w_{\#}(x, y)\rho(x)\rho(y)dx \\
&\geq -\frac{1}{M} \int_0^\infty \int_{\mathbb{R}^d} \min(N h_{r,\rho}(u), N^2 h_{r,\rho}^2(u))(f_1(r) + f_2(r))\,du\,dr \\
&\geq -\frac{\bar{C}(d,\epsilon)}{M} \int_0^\infty \int_{\mathbb{R}^d} \min(N h_{r,\rho}(u), N^2 h_{r,\rho}^2(u)) \frac{1}{r^{s+d+1}}\,du\,dr \ ,
\end{aligned}
$$

where for the last inequality we applied (A.33).

The proof now follows exactly by the same Hardy-Littlewood arguments as in [LNP16], and will be omitted.

To show the equivalent inequalities to (A.25), (A.26) and (A.27) for the $E_{\mathrm{GC},N}^{\mathrm{xc}}$ versions, we restrict again for simplicity to $E_{\mathrm{GC},N,w_{\#}}^{\mathrm{xc}}(\mu)$. We will make use here of the definition of $F_{\mathrm{GC}}^{\mathrm{OT}}(\mu)$ as detailed in (1.7a) above. More precisely, let $(\lambda_0, \lambda_1, \ldots, \lambda_n, \ldots)$, $(\mu_1, \mu_2, \ldots, \mu_n, \ldots)$, be an optimizer for $F_{\mathrm{GC}}^{\mathrm{OT}}(\mu)$. Then by (A.34) and using the same notations as before, we have

$$
\begin{aligned}
k_{r,\rho}(u) &\geq \sum_{n=2}^\infty \alpha_n n^2 \left[\int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \right]^2 - \sum_{n=2}^\infty \alpha_n n \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \\
&\geq \sum_{n=1}^\infty \alpha_n n^2 \left[\int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \right]^2 - \sum_{n=1}^\infty \alpha_n n \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \\
&\geq \left(\sum_{n=1}^\infty \alpha_n n \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \right)^2 - \sum_{n=1}^\infty \alpha_n n \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_n(x) \\
&= N^2 \left[\int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\rho(x)\,dx \right]^2 - N \int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\rho(x)dx = N^2 h_{r,\rho}^2(u) - N h_{r,\rho}(u).
\end{aligned}
$$

For the second inequality in the above we used that $\int_{\mathbb{R}^d} 1_{B_{r/2}}(x - u)\,d\mu_1(x) \leq 1$, for the third inequality we used that $\sum_{n=0}^\infty \alpha_n = 1$, and for the first equality we applied $\sum_{n=1}^\infty n\alpha_n \mu_n = N\mu$. The argument proceeds now the same as for the $E_{N,w_{\#}}^{\mathrm{xc}}(\mu)$ term above. $\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

B   Optimal function spaces estimates

The kernel $g(x) = |x|^{-s}$ on $\mathbb{R}^d$ is in the Lorentz space $L^{\frac{d}{s},\infty}(\mathbb{R}^d)$ (for the case where the second exponent is ∞ the space $L^{p,\infty}(\mathbb{R}^d)$ is also called weak-$L^p(\mathbb{R}^d)$, or Marcinkiewicz space). Recall that the space $L^{p,p}(\mathbb{R}^d)$) equals $L^p(\mathbb{R}^d)$ while for $q > p$ the space $L^{p,q}(\mathbb{R}^d)$ is slightly smaller than L^p but includes all $L^{p+\epsilon}(\mathbb{R}^d), \epsilon > 0$, while for $q < p$ it is slightly smaller than $L^p(\mathbb{R}^d)$ but includes all $L^{p-\epsilon}(\mathbb{R}^d), \epsilon > 0$.

We now provide references to (2.1):

- *Translation of the condition that $f * g$ belongs to $L^\infty(\mathbb{R}^d)$.* By a refined Hölder-Young inequality the first one gives $f \in L^{\frac{d}{d-s},1}(\mathbb{R}^d)$ (which is nothing but the dual space of $L^{\frac{d}{s},\infty}(\mathbb{R}^d)$, cf. [Ga08] thm. 1.4.17 (v)).

- *Translation of the condition that $f \cdot (f * g)$ belongs to $L^1(\mathbb{R}^d)$.* Clearly then the requirement on $f * g$ is less restrictive than to be in $L^\infty(\mathbb{R}^d)$. By the refined multilinear estimate [Ga08] ex. 1.4.18 (which is proved by interpolation methods starting from the analogue more classical $L^p(\mathbb{R}^d)$-space version of the inequality) one gets this time, using again the hypothesis $f \in L^{\frac{d}{s},\infty}(\mathbb{R}^d)$, the requirement $f \in L^{\frac{2d}{2d-s},2}(\mathbb{R}^d)$, under which condition we have $\|f \cdot (g * f)\|_{L^1(\mathbb{R}^d)} \le \|f\|_{L^{\frac{2d}{2d-s},2}(\mathbb{R}^d)} \|g * f\|_{L^{\frac{2d}{s},2}(\mathbb{R}^d)} \le \|f\|^2_{L^{\frac{2d}{2d-s},2}(\mathbb{R}^d)} \|g\|_{L^{\frac{d}{s},\infty}(\mathbb{R}^d)}$.

C   Lieb-Oxford bound uniform in s

Lemma C.1. *Let $N \ge 2$. Fix $0 < \epsilon < d/2$, and let $\epsilon \le s \le d - \epsilon$. Then for all $\mu \in \mathcal{P}(\mathbb{R}^d)$ with $\rho \in L^{1+\frac{s}{d}}(\mathbb{R}^d)$, we have for some $-\infty < c_{\mathrm{LO}}(d,\epsilon) < 0$ which does not depend on N and μ*

$$c_{\mathrm{LO}}(d,\epsilon) \int_{\mathbb{R}^d} \rho^{1+\frac{s}{d}}(x)dx \le N^{-1-\frac{s}{d}} E_{N,s}^{\mathrm{xc}}(\mu) \le 0 \quad and \quad c_{\mathrm{LO}}(d,\epsilon) \int_{\mathbb{R}^d} \rho^{1+\frac{s}{d}}(x)dx \le N^{-1-\frac{s}{d}} E_{\mathrm{GC},N,s}^{\mathrm{xc}}(\mu) \le 0. \tag{C.1}$$

Proof. We begin with the proof of the first inequality in (C.1). To start the proof, a careful analysis of the argument in [LNP16, Lem. 16] reveals that there exists $c_{\mathrm{LO}}(d,\epsilon) < 0$ such that for all $\epsilon \le s \le d - \epsilon$.

$$c_{\mathrm{LO}}(s,d) > c_{\mathrm{LO}}(d,\epsilon). \tag{C.2}$$

To see why (C.2) is true, we note that by the last inequality in the proof of [LNP16, Lem. 16]

$$c_{\mathrm{LO}}(s,d) = -\frac{d\bar{c}_{s,d} M_{s,d}}{2s(d-s)} |B_1|^{1+s/d},$$

for some $\bar{c}_{s,d}, M_{s,d} > 0$. Let us discuss the different constants appearing in the above formula:

- The constant $\bar{c}_{s,d}$ in [LNP16, Lem. 16] comes from the Fefferman-de Llave representation and the allowed range of the exponent s is $0 < s < d$. Moreover for $\epsilon \le s \le d - \epsilon$ it holds that $0 < \sup_{s \in [\epsilon, d-\epsilon]} \bar{c}_{s,d} < \infty$ (see [FL86] for the Fefferman-de Llave representation for the Coulomb potential, [LiLo, Thm. 9.8] for homogenous potentials, and [HS02, Thm. 1] for more general potentials).

- The $M_{s,d}$ comes from the application in the proof of [LNP16, Lem. 16] of the Hardy-Littlewood maximal inequality. As explained for example in [Tao], the standard Vitali-type covering argument used to establish the inequality gives $M_{s,d} = dC^d s^{-1}$ for some $C > 1$ independent of s and d.

The proof of the second inequality in (C.1) uses the first bound from (C.1) by applying the same arguments used in Lemma A.7 to lower-bound the $E_{\mathrm{GC}}^{\mathrm{xc}}$ terms therein. $\qquad\square$

D   Proof of Remark 4.7

Proof of 2).

Towards proving (4.27), we use choices $\{\alpha_n', \mu_n', \gamma_n'\}_{n \ge 0}, \{\alpha_n'', \mu_n'', \gamma_n''\}_{n \ge 0}$, that realize the minimum for $F_{\mathrm{GC},N',\mathsf{c}}(\mu')$ and $F_{\mathrm{GC},N'',\mathsf{c}}(\mu'')$, respectively. Since the corresponding optimal transport problems do not exist for $n = 0, 1$, we take by convention $\gamma_0' = \gamma_0'' = 1, \gamma_1' = \mu_1', \gamma_1'' = \mu_1''$, in order to construct a competitor for the left-hand side in (4.27). Let then $\gamma_{i,n-i} := \gamma_i' \otimes \gamma_{n-i}'' \in \mathcal{P}^n(\mathbb{R}^d)$ for all $n \ge 2$, and $0 \le i \le n$. We symmetrise $\gamma_{i,n-i}$ by the formula

$$\widetilde{\gamma}_{i,n-i}(B) = \frac{1}{n!} \sum_{\sigma \in S_n} \gamma_{i,n-i}(\{(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) : (x_1, \ldots, x_n) \in B\}), \tag{D.1}$$

for all Borel sets $B \subseteq (\mathbb{R}^d)^n$, where S_n is the permutation group on n symbols. Note that $\widetilde{\gamma}_{i,n-i} \mapsto \frac{i\mu_i' + (n-i)\mu_{n-i}''}{n}$. Define now for all $n \geq 0$

$$\alpha_n := \sum_{i=0}^{n} \alpha_i' \alpha_{n-i}'', \tag{D.2a}$$

and if $\alpha_n \neq 0$ let

$$\gamma_n := \frac{1}{\alpha_n} \left(\sum_{i=0}^{n} \alpha_i' \alpha_{n-i}'' \widetilde{\gamma}_{i,n-i} \right) \in \mathcal{P}_{sym}^n(\mathbb{R}^d), \ \mu_n := \frac{1}{\alpha_n} \left(\sum_{i=0}^{n} \alpha_i' \alpha_{n-i}'' \frac{i\mu_i' + (n-i)\mu_{n-i}''}{n} \right). \tag{D.2b}$$

Then $\gamma_n \in \mathcal{P}^n(\mathbb{R}^d)$ and has marginal μ_n, therefore by re-arranging the terms in the summations we have

$$
\begin{aligned}
\sum_{n=1}^{\infty} n\alpha_n\mu_n &= \sum_{n=0}^{\infty}\sum_{i=0}^{n} \alpha_i'\alpha_{n-i}'' \left(i\mu_i' + (n-i)\mu_{n-i}'' \right) = \sum_{k=0}^{\infty} \sum_{\substack{i,n=0 \\ n-i=k}}^{\infty} \alpha_i'\alpha_{n-i}'' \left(i\mu_i' + (n-i)\mu_{n-i}'' \right) \\
&= \sum_{k=0}^{\infty} \left(\alpha_k'' \sum_{n=1}^{\infty} n\alpha_n'\mu_n' + k\alpha_k''\mu_k'' \sum_{n=0}^{\infty} \alpha_n' \right) = N'\mu' + N''\mu'',
\end{aligned} \tag{D.3}
$$

and similarly $\sum_{n=0}^{\infty} \alpha_n = 1$. From the definition of $\gamma_{i,n-i}$ and (D.1), we have for all $n \geq 2$

$$
\int_{(\mathbb{R}^d)^n} \sum_{k \neq l=1}^{n} c(x_k, x_l) \, d\widetilde{\gamma}_{i,n-i}(x_1, \ldots, x_n)
$$
$$
= F_i(\mu_i') + F_{n-i}(\mu_{n-i}'') + 2i(n-i) \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) d\mu_i'(x) d\mu_{n-i}''(y). \tag{D.4}
$$

Due to (D.3) use the choice (D.2) as a competitor for $F_{\mathrm{GC},N'+N'',\mathsf{c}}(\mu)$. Combining (D.4) and using linearity, we get

$$
\begin{aligned}
F_{\mathrm{GC},N'+N'',\mathsf{c}}^{\mathrm{OT}}(\mu) &\leq \sum_{n=0}^{\infty}\sum_{i=0}^{n} \alpha_i'\alpha_{n-i}'' \int_{(\mathbb{R}^d)^n} \sum_{1 \leq k \neq l \leq n} c(x_k, x_l) \, d\widetilde{\gamma}_{i,n-i}(x_1, \ldots, x_n) \\
&= \sum_{n=2}^{\infty} \alpha_n' F_n(\mu_n') + \sum_{n=2}^{\infty} \alpha_n'' F_n(\mu_n'') + 2\sum_{n=0}^{\infty}\sum_{i=0}^{n} \alpha_i'\alpha_{n-i}'' i(n-i) \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y) d\mu_i'(x) d\mu_{n-i}''(y) \\
&= F_{\mathrm{GC},N',\mathsf{c}}^{\mathrm{OT}}(\mu') + F_{\mathrm{GC},N'',\mathsf{c}}^{\mathrm{OT}}(\mu'') + 2N'N'' \int_{\mathbb{R} \times \mathbb{R}^d} \frac{1}{|x-y|^s} d\mu'(x) d\mu''(y).
\end{aligned}
$$

The (4.27) follows by applying the formula for the mean field term

$$
\begin{aligned}
(N'+N'')^2 \int_{\mathbb{R} \times \mathbb{R}^d} \frac{1}{|x-y|^s} d\mu(x) d\mu(y) &= (N')^2 \int_{\mathbb{R} \times \mathbb{R}^d} \frac{1}{|x-y|^s} d\mu'(x) d\mu'(y) + (N'')^2 \int_{\mathbb{R} \times \mathbb{R}^d} \frac{1}{|x-y|^s} d\mu''(x) d\mu''(y) \\
&\quad + 2N'N'' \int_{\mathbb{R} \times \mathbb{R}^d} \frac{1}{|x-y|^s} d\mu'(x) d\mu''(y).
\end{aligned}
$$

As already noted in the introduction, for the $N' \leq 1$ case, it suffices to take instead for $F_{\mathrm{GC},N',\mathsf{c}}(\mu')$ a competitor of the form $\alpha_n' = 0, n \geq 2, \alpha_1' = N', \alpha_0' = 1 - N', \mu_n' = \mu', n \geq 1$, and then proceed as above.

Proof of 3).

To show (4.29), we can apply the same reasoning as in the previous point with minor modifications, as follows. Take $\{\alpha_n', \mu_n', \gamma_n'\}_{n \geq 0}, \{\alpha_n'', \mu_n'', \gamma_n''\}_{n \geq 0}$, which realize the minima of $F_{\mathrm{GCB},N',\mathsf{c}}(\mu', \bar{N}'), F_{\mathrm{GCB},N'',\mathsf{c}}(\mu'', \bar{N}'')$, respectively. Here for $n > \bar{N}'$ we may arbitrarily assign μ_n', γ_n' because the coefficients α_n' are zero due to (4.24b), and similarly for $n > \bar{N}''$. Then we use the definitions (D.2) verbatim, and in (D.2a) one checks that, since α_n is the sum of terms of the form $\alpha_{n'}' \alpha_{n''}''$, with $n' + n'' = n, n' \geq 0, n'' \geq 0$, this term can only be nonzero if both $\alpha_{n'}'$ and $\alpha_{n''}''$ are nonzero, i.e. if $n' \leq \bar{N}', n'' \leq \bar{N}''$, for some of the terms. From this it follows that $n \leq \bar{N}' + \bar{N}''$ is a necessary condition for α_n to be nonzero. We claim that the choices (D.2) give a competitor for $E_{\mathrm{GCB},N'+N'',\mathsf{c}}(\mu, \bar{N}' + \bar{N}'')$. Indeed, the vanishing conditions on the α_n are what we just verified, and the remaining conditions from (4.24b) are checked exactly like for the GC-problem, treated in the previous point. By the same passages as for the GC-problem, we then obtain (4.29).

E   Some helpful optimal transport results

Here we assume that c satisfies (2.3), which we recall here:

$$\mathsf{c}(x,y) = g(x-y) = l(|x-y|) \quad \text{where} \quad l : [0,\infty) \to [0,\infty) \quad \text{is} \quad \begin{cases} \text{continuous on } (0,\infty) \\ \text{strictly decreasing ,} \\ \text{such that } \lim_{t\to 0^+} l(t) = +\infty \; . \end{cases} \tag{E.1a}$$

Hypotheses (E.1) are satisfied by $g(x) = |x|^{-s}$ for $s > 0$, but we note here that (E.1a) is not satisfied by $g(x) = -\log|x|$.

Lemma E.1 (Lower semi-continuity of $F_{\mathrm{GCB},N,\mathsf{c}}^{\mathrm{OT}}$ and $F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}$). *Let c be as in (E.1a) and let $(\mu_k)_{k\geq 1}, \mu \in \mathcal{P}(\mathbb{R}^d)$, such that the integrals required to define the quantities below are finite. Further suppose that μ_k converges weakly to μ and that $\sup_{k\geq 1} F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k) < \infty$ (it suffices that $\sup_{k\geq 1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathsf{c}(x,y)\,\mathrm{d}\mu_k(x)\mathrm{d}\mu_k(y) < \infty$ for this purpose). Fix $\bar{N} \geq N \in \mathbb{R}_+, N \geq 2$. Then there holds*

$$\liminf_{k\to\infty} F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k) \geq F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu) \quad \text{and} \quad \liminf_{k\to\infty} F_{\mathrm{GCB},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k, \bar{N}) \geq F_{\mathrm{GCB},N,\mathsf{c}}^{\mathrm{OT}}(\mu, \bar{N}). \tag{E.2}$$

Proof. We do the proof only for the GC-problem, as the GCB-problem works similarly, and requires the only extra fact that the condition $\lambda_{n,k} = 0$ for $n > \bar{N}$ passes to weak limits.

Step 0. *Notation and setting.*

Let $((\lambda_{n,k})_{n\geq 0}, (\gamma_{n,k})_{n\geq 2})$ be a minimizer for the $F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k)$ problem. Set:

$$\lambda_k := \sum_{n\geq 0} \lambda_{n,k}\delta_n \in \mathcal{P}(\mathbb{N}_{\geq 0}), \quad \text{where} \quad \mathbb{N}_{\geq 0} := \{0,1,2,\ldots\} \; . \tag{E.3}$$

We define next the disjoint union space

$$X_{\mathrm{GC}} := \coprod_{n\geq 2} X_n \quad \text{where we denote} \quad X_n := (\mathbb{R}^d)^n, n \geq 2. \tag{E.4}$$

Noting that the Borel sets of a disjoint union $\mathcal{B}(\coprod_{n\geq 2} X_n)$ are generated by $\cup_{n\geq 2}\mathcal{B}(X_n)$, we may build measures $\gamma_{\mathrm{GC},N,k} \in \mathcal{M}_+(\coprod_{n\geq 2} A_n)$ by defining

$$\gamma_{\mathrm{GC},N,k}(A_n) := \lambda_{n,k}\gamma_{n,k}(A_n) \quad \text{for all choices of Borel sets} \quad A_n \subseteq X_n, \; n \geq 2. \tag{E.5}$$

The two measures $\gamma_{\mathrm{GC},N,k}, \lambda_k$, defined above are related by the disintegration relation

$$\forall n \geq 2, \quad \gamma_{\mathrm{GC},N,k}(X_n) = \lambda_k(\{n\}), \tag{E.6a}$$

and if $\gamma_{n,k}^{(m)}, m \geq 1$, is the m-marginal measure of $\gamma_{n,k}$, the constraint from the definition of $F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k)$ gives the following relation on μ_k:

$$\sum_{n\geq 1} n\lambda_{n,k}\gamma_{n,k}^{(1)} = N\mu_k. \tag{E.6b}$$

Given two measures $\lambda_k^* \in \mathcal{P}(\mathbb{N}_{\geq 2})$ and $\gamma_k^* \in \mathcal{M}_+(X_{\mathrm{GC}})$, the requirement that they satisfy (E.5) is equivalent to saying that $(\lambda_k^*, \gamma_k^*)$, encode a competitor for $F_{\mathrm{GC},N,\mathsf{c}}^{\mathrm{OT}}(\mu_k)$, given by $(\vec{\lambda}_k^*, \vec{\gamma}_k^*)$ with

$$\lambda_{n,k}^* := \lambda_k^*(\{n\}) \quad \text{for all} \quad n \geq 2, \tag{E.7a}$$

and for $\lambda_{n,k}^* > 0$,

$$\gamma_{n,k}^*(A_n) := (\lambda_{n,k}^*)^{-1} \gamma_k^*(A_n), \quad A_n \subseteq X_n, n \geq 2. \tag{E.7b}$$

Provided (E.7) holds, we further define

$$\lambda_k^*(\{0\}) := \lambda_{0,k}^*, \; \lambda_k^*(\{1\}) := \lambda_{1,k}^*, \quad \text{where} \quad \lambda_{1,k}^* := N - \sum_{n=2}^{\infty} n\lambda_{n,k}^*, \quad \lambda_{0,k}^* := 1 - \sum_{n=1}^{\infty} \lambda_{n,k}^*, \tag{E.8a}$$

and if $\lambda_{1,k}^* > 0$, then we also define

$$\left(\gamma_{1,k}^*\right)^{(1)}(A_1) := N\mu(A_1) - \sum_{n \geq 2} n\lambda_{n,k}\left(\gamma_{n,k}^*\right)^{(1)}(A_1) \quad \text{for} \quad A_1 \in \mathcal{B}(\mathbb{R}^d). \tag{E.8b}$$

Assume μ has support Λ and fix $\delta > 0$. Let $\Lambda_R := \text{supp}(\mu) \cap B_R$, with $R = R(\delta) > 0$ chosen such that

$$\mu(\mathbb{R}^d \setminus \Lambda_R) < \delta, \quad \text{and} \quad \forall k \geq 1, \quad \mu_k(\mathbb{R}^d \setminus \Lambda_R) < \delta. \tag{E.9}$$

This choice is possible because μ, μ_k, are finite measures. Denote now for all $n \geq 2$ by

$$\Lambda_{R,\alpha}^n := \left\{ (x_1, \ldots, x_n) \in \Lambda_R^n : \min_{1 \leq i \neq j \leq n} |x_i - x_j| \geq \alpha \right\}, \tag{E.10}$$

and let

$$\Lambda_{R,\alpha} := \left(\coprod_{2 \leq n < n_\alpha} \Lambda_R^n \right) \coprod \left(\coprod_{n \geq n_\alpha} \Lambda_{R,\alpha}^n \right) \subset X_{\text{GC}}. \tag{E.11}$$

Step 1. *Mass bound uniform in k.* Denoting the 2-marginals of the $\gamma_{n,k}$ by $\gamma_{n,k}^{(2)}$, we obtain

$$F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu_k) = \sum_{n=2}^{\infty} n(n-1)\lambda_{n,k} \int_{\mathbb{R}^{2d}} \mathsf{c}(x,y) \, d\gamma_{n,k}^{(2)}(x,y). \tag{E.12}$$

Due to our hypothesis assumptions, and in view of (4.31), we have

$$\begin{aligned}
\sup_{k \geq 1} F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu_k) &= \sup_{k \geq 1} \sum_{n=2}^{\infty} n(n-1)\lambda_{n,k} \int_{\mathbb{R}^{2d}} \mathsf{c}(x,y) \, d\gamma_{n,k}^{(2)}(x,y) \\
&\leq N^2 \sup_{k \geq 1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathsf{c}(x,y) \, d\mu_k(x) d\mu_k(y) := N^2 C.
\end{aligned} \tag{E.13}$$

Step 2. *The $\Lambda_{R,\alpha}^n$ contain the mass up to small error.*

Let $\alpha \in (0,1[$. Then we have

$$\gamma_{n,k}(X_n \setminus \Lambda_{R,\alpha}^n) = \gamma_{n,k}(X_n \setminus \Lambda_R^n) + \gamma_{n,k}(\Lambda_R^n \setminus \Lambda_{R,\alpha}^n). \tag{E.14}$$

For the first term in (E.14) we note that $X_n \setminus \Lambda_R^n$ is included in the union of the sets $(\mathbb{R}^d)^i \times (\mathbb{R}^d \setminus \Lambda_R) \times (\mathbb{R}^d)^{n-1-i}$ for $i = 0, \ldots, n-1$, each of which projects to $\mathbb{R}^d \setminus \Lambda_R$, and thus has $\gamma_{n,k}$-measure at most ϵ due to (E.9). This means that

$$\gamma_{n,k}(X_n \setminus \Lambda_R^n) \leq \sum_{i=0}^{n-1} \gamma_{n,k}\left((\mathbb{R}^d)^i \times (\mathbb{R}^d \setminus \Lambda_R) \times (\mathbb{R}^d)^{n-1-i}\right) \leq n\delta. \tag{E.15}$$

For the second term in (E.14) we estimate

$$\begin{aligned}
\gamma_{n,k}\left(X_n \setminus \Lambda_{R,\alpha}^n\right) &= \gamma_{n,k}\left(\left\{(x_1, \ldots, x_n) \in \Lambda_R^n : \min_{1 \leq i \neq j \leq n} |x_i - x_j| < \alpha\right\}\right) \\
&= n(n-1)\gamma_{n,k}^{(2)}\left(\{(x,y) \in \Lambda_R \times \Lambda_R : |x - y| < \alpha\}\right) \\
&= n(n-1)\gamma_{n,k}^{(2)}\left(\{(x,y) \in \Lambda_R \times \Lambda_R : \mathsf{c}(x,y) > l(\alpha)\}\right) \\
&\leq n(n-1)l(\alpha) \int_{\Lambda_R \times \Lambda_R} \mathsf{c}(x,y) \, d\gamma_{n,k}^{(2)}(x,y).
\end{aligned} \tag{E.16}$$

For the second equality in the above, we used that $\gamma_{n,k}$ is a symmetric measure and the definition of $\gamma_{n,k}^{(2)}$, for the third equality we applied the properties of l from (E.1a), and for the inequality we used Markov's inequality. Then (E.14), (E.15) and (E.16) give

$$\gamma_{n,k}\left(X_n \setminus \Lambda_{R,\alpha}^n\right) \leq n\delta + n(n-1)\, l(\alpha) \int_{\Lambda_R \times \Lambda_R} \mathsf{c}(x,y) \, d\gamma_{n,k}^{(2)}(x,y) \tag{E.17}$$

Due to (E.13), we now have from (E.17) that

$$
\sup_{k\geq 1}\left[\sum_{n\geq 2}\lambda_{n,k}\gamma_{n,k}\left(X_n\setminus\Lambda_{R,\alpha}^n\right)\right] \leq \sup_{k\geq 1}\left[\sum_{n\geq 2}n(n-1)\lambda_{n,k}l(\alpha)\int_{\Lambda_R\times\Lambda_R}\mathsf{c}(x,y)\,d\gamma_{n,k}^{(2)}(x,y)\right]+\sup_{k\geq 1}\left[\sum_{n\geq 2}n\lambda_{n,k}\delta\right]
$$

$$
\leq \ l(\alpha)\ N^2 C + N\delta. \tag{E.18}
$$

Step 3. $\gamma_{n,k}$ *give no mass to the* $\Lambda_{R,\alpha}^n$ *for large* n.

We note that, by definition (E.10) of the $\Lambda_{R,\alpha}^n$, if

$$
(x_1,\ldots,x_n)\in\Lambda_{R,\alpha}^n, \tag{E.19}
$$

then the balls $B_{\alpha/2}(x_i)$, with $i=1,\ldots,n$, are disjoint and contained in the enlargement of B_R given by $B_{R+\alpha/2}$, and so by volume comparison we have

$$
n\ 2^{-d}\alpha^d = \sum_{i=1}^n \frac{|B_{\alpha/2}(x_i)|}{|B_1|} \leq \frac{|B_{R+\alpha/2}|}{|B_1|} = \left(R+\frac{\alpha}{2}\right)^d, \tag{E.20}
$$

and thus

$$
n\leq n_\alpha := \left(\frac{R}{\alpha/2}+1\right)^d. \tag{E.21}
$$

Therefore for $n>n_\alpha$ we have that condition (E.19) cannot hold. Therefore

$$
\text{For all }n>n_\alpha\text{ there holds}\quad\mathrm{supp}(\gamma_{n,k})\subset X_n\setminus\Lambda_{R,\alpha}^n. \tag{E.22}
$$

Step 4. *Tightness of the* $(\lambda_k)_{k\geq 1}$.

Due to the result (E.22) of Step 3, together with the result (E.18) of Step 2, we find that

$$
\sup_{k\geq 1}\lambda_k\left(\mathbb{N}_{\geq 0}\setminus\{0,1,2,3,\ldots,n_\alpha\}\right) \overset{\text{(E.3),(E.6a)}}{=} \sup_{k\geq 1}\left[\sum_{n\geq n_\alpha}\lambda_{n,k}\right] \overset{\text{(E.22)}}{=} \sup_{k\geq 1}\left[\sum_{n\geq n_\alpha}\lambda_{n,k}\gamma_{n,k}\left(X_n\setminus\Lambda_{R,\alpha}^n\right)\right]
$$

$$
\leq \sup_{k\geq 1}\left[\sum_{n\geq 2}\lambda_{n,k}\gamma_{n,k}\left(X_n\setminus\Lambda_{R,\alpha}^n\right)\right]
$$

$$
\overset{\text{(E.18)}}{\leq} l(\alpha)\cdot N^2 C + N\delta. \tag{E.23}
$$

As the set $\{0,1,2,3,\ldots,n_\alpha\}\subset\mathbb{N}_{n\geq 0}$ is finite, and thus compact, this shows in particular that the family of probability measures

$$
\lambda_k = \sum_{n\geq 0}\lambda_{n,k}\delta_n,\quad k\geq 1,
$$

is tight, and thus a fixed subsequence of $k\uparrow\infty$ which realizes the $\liminf$ from (E.2) has a further subsequence which we denote $(k_j)_j$, converging to ∞ along which the $\liminf$ in (E.2) is realized, in the sense that $\liminf_{k\to\infty}F_{GC,N,s}(\mu_k)=\lim_{j\to\infty}F_{GC,N,s}(\mu_{k_j})$, and such that furthermore the measures λ_{k_j} converge weakly, i.e.

$$
\mathcal{P}(\mathbb{N}_{\geq 0})\ni\lambda_{k_j}\overset{*}{\rightharpoonup}\lambda^*\in\mathcal{P}(\mathbb{N}_{\geq 0}). \tag{E.24}
$$

Step 5. *Tightness of* $(\gamma_{\mathrm{GC},N,k})_{k\geq 1}$.

Due to (E.15) coupled with the result (E.23) of Step 4, coupled with the links (E.5) and (E.6a) between $\gamma_{\mathrm{GC},N,k}$ and $\gamma_{n,k},\lambda_{n,k}$, and by using (E.6b), we find that

$$
\sup_{k\geq 1}\gamma_{\mathrm{GC},N,k}\left(X_{\mathrm{GC}}\setminus\coprod_{2\leq n<n_\alpha}\Lambda_R^n\right) \leq \sup_{k\geq 1}\sum_{2\leq n<n_\alpha}\gamma_{\mathrm{GC},N,k}\left(X_n\setminus\Lambda_R^n\right) + \sup_{k\geq 1}\gamma_{\mathrm{GC},N,k}\left(\coprod_{n\geq n_\alpha}X_n\right)
$$

$$
\leq \sup_{k\geq 1}\sum_{2\leq n<n_\alpha}n\lambda_{n,k}\epsilon + l(\alpha)\cdot N^2 C + N\delta
$$

$$
\leq l(\alpha)\cdot N^2 C + 2N\delta. \tag{E.25}
$$

59

Due to the fact that for any choice of $\delta > 0$ the set B_R is compact and that the product of finitely many compact sets is compact, we find that the set $\coprod_{2 \le n < n_\alpha} \Lambda_R^n$ is compact, and thus by the arbitrariness of $\alpha, \delta > 0$, the bound (E.25) shows that the sequence obtained from the previous step, denoted $\gamma_{\mathrm{GC},N,k_j}$, is tight. Thus, up to extracting a further subsequence, we find a measure $\gamma^* \in \mathcal{M}_+(X_{\mathrm{GC}})$ to which the $\gamma_{\mathrm{GC},N,k_j}$ converge weak-$*$, namely we have

$$\gamma_{\mathrm{GC},N,k_j} \xrightarrow{*} \gamma^* \in \mathcal{M}_+(X_{\mathrm{GC}}). \tag{E.26}$$

Step 6. $(\lambda^*, \gamma^*_{\mathrm{GC},N})$ *is a competitor for* $F^{\mathrm{OT}}_{\mathrm{GC},N,\mathsf{c}}(\mu)$.

Due to the facts (E.24) and (E.26), we obtain that the constraint (E.6a) holds also for (γ^*, λ^*). By testing the weak convergence from (E.24) against the function $f \in C_b(\mathbb{N}_{\ge 0})$, defined to be equal to one at n and zero elsewhere (which is a continuous function on the discrete space $\mathbb{N}_{\ge 0}$), we find that for each $n \in \mathbb{N}_{\ge 0}$ there holds along our subsequence

$$\lambda_{n,k_j} \overset{\text{(E.3),(E.7)}}{=} \lambda_{k_j}(\{n\}) \to \lambda^*(\{n\}) =: \lambda_n^* \in [0,1], \tag{E.27}$$

Condition (E.6a), together with the fact that by (E.26) we have for all $n \ge 2$ that $\gamma_{\mathrm{GC},N,k_j}|_{X_n} \xrightarrow{*} \gamma^*|_{X_n}$, means that for each $n \in \mathbb{N}_{\ge 2}$ such that $\lambda_n^* > 0$, there holds

$$\gamma_{\mathrm{GC},N,k_j}|_{X_n} \overset{\text{(E.5),(E.7)}}{=} \lambda_{n,k_j} \gamma_{n,k_j} \xrightarrow{*} \gamma^*|_{X_n}, \tag{E.28}$$

which implies in turn via (E.27) that whenever $\lambda_n^* \neq 0$ we have

$$\gamma_{n,k_j} \xrightarrow{*} (\lambda_n^*)^{-1} \gamma^*|_{X_n} := \gamma_n^*. \tag{E.29}$$

It follows from (E.27) and (E.29) that $\gamma_n^* \in \mathcal{P}(X_n)$. It remains to show that (E.6b) holds with $(\lambda_n^*, \gamma_n^*)$.

For the case of $\mathsf{c} > 0$ and $f \in C_c^0(\mathbb{R}^d)$ with support contained in the ball $B(0,R)$, we have $\mathsf{c}(x,y) > \min_{x',y' \in B(0,R)} \mathsf{c}(x',y') = m_R > 0$, valid for all $x,y \in \mathrm{supp}(f)$, and thus we have the pointwise bound

$$\mathsf{c}(x',y') \ge \frac{m_R}{\sup_{y \in \mathrm{supp}(f)} |f(y)|} |f(x')| := C_f |f(x')| \quad \text{for all} \quad x', y' \in \mathbb{R}^d. \tag{E.30}$$

By integrating in (E.30) and summing over i, j, and by the marginal property of $\gamma_{n,k}$, there holds for $k \in \mathbb{N}$

$$\sum_{n=2}^\infty \lambda_{n,k} \int_{\mathbb{R}^{Nd}} \sum_{i,j=1,i\neq j}^n \mathsf{c}(x_i,x_j) d\gamma_{n,k}(x_1,\ldots,x_n)$$

$$\ge \quad C_f \sum_{n=2}^\infty \lambda_{n,k} n(n-1) \int_{\mathbb{R}^{Nd}} |f(x_1)| d\gamma_{n,k}(x_1,\ldots,x_n) = C_f \sum_{n=2}^\infty n(n-1)\lambda_{n,k} \int |f(x)| d\mu_{n,k}(x). \tag{E.31}$$

In view of (E.31) we have that

$$C_f \sum_{n=2}^\infty n(n-1)\lambda_{n,k} \int |f| \, d\mu_{n,k} \le F^{\mathrm{OT}}_{\mathrm{GC},N,\mathsf{c}}(\mu_k) \le \sup_k F^{\mathrm{OT}}_{N,\mathsf{c}}(\mu_k) < N^2 \sup_{k \ge 1} \int_{\mathbb{R}^{2d}} \mathsf{c}(x,y) d\mu_k(x) d\mu_k(y) < \infty,$$

thus we obtain that at fixed N, f for $(\vec{\lambda}_k, \vec{\gamma}_k)$ optimizing $F^{\mathrm{OT}}_{\mathrm{GC},N,\mathsf{c}}(\mu_k)$

$$\sup_{k \ge 1} \sum_{n=2}^\infty n(n-1)\lambda_{n,k} \int |f(x)| \, d\mu_{n,k}(x) < \infty. \tag{E.32}$$

Therefore, as $\mu_{n,k}$ are positive measures, the integrals

$$\sum_{n=2}^\infty n\lambda_{n,k} \int f\mu_{n,k}$$

are uniformly summable as $k \to \infty$ and thus equation (E.6b) also passes to the weak limit in duality with $f \in C_c^0(\mathbb{R}^d)$, and holds with λ_n^*, γ_n^* as in (E.27), (E.29). Moreover, since the property of being a symmetric

60

probability measure passes to weak limits, we find from (E.29) also that $\gamma_n^* \in \mathcal{P}_{\text{sym}}\left((\mathbb{R}^d)^n\right)$. Thus $\gamma_{\text{GC},N}^* := \sum_{n \geq 2} \lambda_n^* \gamma_n^* \in \mathcal{P}(X_{\text{GC}})$ is a competitor to $F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu)$.

Step 7. *Conclusion of the proof.*

We start by noting that for all $n \geq 2$, we have

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \mathsf{c}(x,y)\, \lambda_n^* d(\gamma_n^*)^{(2)}(x,y) \leq \liminf_{j \to \infty} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathsf{c}(x,y)\, \lambda_{n,k_j}\, d\gamma_{n,k_j}^{(2)}(x,y) =: \liminf_{j \to \infty} f_j(n). \tag{E.33}$$

The above follows by applying Lemma 6.1 from [SS15] (or [BCdP16, Thm. 3.9]). The assumptions of the lemma are then verified on the space $X = \mathbb{R}^d \times \mathbb{R}^d$ by $\gamma_{n,k_j}^{(2)}$ and c.

For all $k \geq 1$, we apply Fatou's theorem for the atomic measure ν on $\mathbb{N}$ with $\nu(\{n\}) = n(n-1)$ for all $n \in \mathbb{N}$ and for $f_j(n)$ defined in (E.33), so by summing first over n the positive terms coming from (E.33)

$$\sum_{n=2}^{\infty} \int_{\mathbb{R}^d \times \mathbb{R}^d} n(n-1)\lambda_n^* \mathsf{c}(x,y)\, d(\gamma_n^*)^{(2)}(x,y) \leq \sum_{n=2}^{\infty} \liminf_{j \to \infty} \left(n(n-1) \int_{\mathbb{R}^d \times \mathbb{R}^d} \lambda_{n,k_j} \mathsf{c}(x,y)\, d\gamma_{n,k_j}^{(2)}(x,y) \right)$$

$$= \int \liminf_{j \to \infty} f_j(n) d\nu(n) \leq \liminf_{j \to \infty} \int f_j(n) d\nu(n) = \liminf_{j \to \infty} \left(\sum_{n=2}^{\infty} n(n-1) \int_{\mathbb{R}^d \times \mathbb{R}^d} \lambda_{n,k_j} \mathsf{c}(x,y)\, d\gamma_{n,k_j}^{(2)}(x,y) \right)$$

$$= \liminf_{j \to \infty} F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu_{k_j}). \tag{E.34}$$

Then, as by Step 6 we have that $\gamma_{\text{GC},N}^*$ is a competitor for $F_{\text{GC},N}^{\text{OT}}(\mu)$, we get

$$F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu) \leq \sum_{n=2}^{\infty} \int_{\mathbb{R}^d \times \mathbb{R}^d} n(n-1)\, \lambda_n^*\, \mathsf{c}(x,y)\, d(\gamma_n^*)^{(2)}(x,y) \leq \liminf_{j \to \infty} F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu_{k_j}), \tag{E.35}$$

where for the second inequality we applied (E.34). The (E.35) thus implies that (E.2) holds. $\qquad\square$

We now show

Lemma E.2. *(Existence of an optimal solution for $F_{\text{GC},N,\mathsf{c}}^{\text{OT}}$) Let c be as in (E.1a) and let $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that the integrals required to define the quantities involved in the definition of $F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu)$ are finite. Then $F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu)$ has at least one solution.*

Proof. It suffices to take a sequence of competitors $(\lambda_{n,k}, \gamma_{n,k})$ to $F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu)$ such that $\sum_{n \geq 2} \lambda_{n,k} F_n^{\text{OT}}(\mu_{n,k}) \to F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu)$, and then proceed similarly to the proof in Lemma E.1: there exists a tight subsequence converging to a competitor $(\lambda_n^*, \gamma_n^*)$, and the thesis follows by the same arguments as in Lemma E.1. This will allow us to say that

$$F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu) = \liminf_{k \to \infty} \sum_{n \geq 2} \lambda_{n,k} F_n^{\text{OT}}(\mu_{n,k}) \geq F_{\text{GC},N,\mathsf{c}}^{\text{OT}}(\mu^*),$$

which will imply in particular that $(\lambda_n^*, \gamma_n^*)$ is a minimizer. $\qquad\square$

Acknowledgments

References

[AKM13] S. ADAMS, R., KOTECKY, S. MÜLLER. Finite range decomposition for families of gradient Gaussian measures. *JFA* **264**, 169-206 (2013).

[B13] R. BAUERSCHMIDT. A simple method for finite range decomposition of quadratic forms and Gaussian fields. *Probab. Theory Related Fields* **157**, 3-4, 817-845 (2013).

[BBS15] R. BAUERSCHMIDT, D.C. BRYDGES, AND G. SLADE. Logarithmic correction for the susceptibility of the 4-dimensional weakly self-avoiding walk: a renormalisation group analysis. *Commun. Math. Phys.* **337**, 2, 817-877 (2015).

[BCC17] C. BELTRÁN, N. CORRAL, J. G. CRIADO DEL REY. Discrete and Continuous Green Energy on Compact Manifolds. *preprint.* https://arxiv.org/abs/1702.00864.(2017).

[BBL11] R.D. BENGURIA, G.A. BLEY, M. LOSS A new estimate on the indirect Coulomb energy. *Int Journ. Quant Chem.* **112**, 6, 1579-1584 (2012).

[BdP17] U. BINDINI, L. DE PASCALE. Optimal transport with Coulomb cost and the semiclassical limit of Density Functional Theory, *preprint.* (2017).

[B87] Y. BRENIER. Decomposition polaire et rearrangement monotone des champs de vecteurs. *C.R. Acad. Sci. Pair. Ser. I Math.* **305**, 805-808 (1987).

[BGM04] D. C. BRYDGES, G. GUADAGNI, P. K. MITTER. Finite range Decomposition of Gaussian Processes. *JSP* **115**, 415-449 (2004).

[BT06] D. C. BRYDGES, A. TALARCZYK. Finite range decompositions of positive-definite functions. *Journ. Funct. Anal.* **236**, 2, 682-711 (2006).

[BCdP16] G. BUTTAZZO, T.CHAMPION, L. DE PASCALE. Continuity and estimates for multimarginal optimal transportation problems with singular costs, *preprint* (2016).

[BdPGG12] G. BUTTAZZO, L. DE PASCALE, P. GORI-GIORGI. Optimal transport formulation of electronic density-functional theory. *Phys. Rev. A* **85**, 062502 (2012).

[CFM02] L. CAFFARELLI, M. FELDMAN, R. MCCANN. Constructing optimal maps for Monges transport problem as a limit of strictly convex costs, *Journal of the American Mathematical Society* **15**, no. 1, 1-26 (2002).

[CMC10] L. CAFFARELLI, R. MCCANN. Free boundaries in optimal transport and Monge-Ampére obstacle problems, *Annals of Mathematics* **171**, 673-730 (2010).

[CdPdM13] M. COLOMBO, L. DE PASCALE, S. DI MARINO Multimarginal optimal transport maps for 1-dimensional repulsive costs. *Canad. J. Math., to appear* (2013).

[CLY89] J.G. CONLON, E.H. LIEV, H.-T. YAU The Coulomb Gas at Low Temperature and Low Density. *Comm. Math. Phys.* **125**, 153-180 (1989).

[CFK11] C. COTAR, G. FRIESECKE, C. KLÜPPELBERG. Density functional theory and optimal transportation with Coulomb cost, *Comm. Pure. Appl. Math.* **66**, 4, 548-599 (2013).

[CFK17] C. Cotar, G. Friesecke, C. Klüppelberg. Smoothing of transport plans with fixed marginals and rigorous semiclassical limit of the Hohenberg-Kohn functional. *Arch. Rat. Mech. Anal. (to appear)* arxiv:1706.05676 (2017).

[CFP15] C. Cotar, G. Friesecke, B. Pass. Infinite-body optimal transport with Coulomb cost. *Calc. Var. PDEs.* **54**, 1, 717-742 (2015).

[CP17] C.Cotar, M. Petrache. Equality of the Jellium and Uniform Electron Gas next-order asymptotic terms for Riesz potentials. *preprint* (2017).

[dP15] L. de Pascale. Optimal Transport with Coulomb cost. Approximation and duality. *ESAIM: Math. Mod. num. Anal.* **49**, 6, 1643-1657 (2015).

[dMGN15] S. di Marino, A. Gerolin, L. Nenna. Optimal transportation theory for repulsive costs, *preprint* (2015).

[dM] S. Di Marino. Personal communication.

[D30] P.A.M. Dirac. Note on Exchange Phenomena in the Thomas Atom. *Mathematical Proceedings of the Cambridge Philosophical Society* **26**, 376-385 (1930).

[DS58] N. Dunford, J.T. Schwartz. Linear operators, Part I, *Wiley-Interscience* (1958).

[F85] C. Fefferman. The thermodynamic limit for a crystal. *Comm. Math. Phys.* **98**, 289-311 (1985).

[FL86] C.L. Fefferman, R. de la Llave. Relativistic stability of matter I. *Revista Matematica Iberoamericana* **2**, 119-161 (1986).

[F64] M. E. Fisher. The free energy of a macroscopic system. *Arch. Ration. Mech. Anal.*, **17**, 377-410 (1964).

[FLS15] S. Fournais, M. Lewin, J. Solovej. The semi-classical limit of large fermionic systems. *preprint.* arXiv:1510.01124 (2015).

[GM96] W. Gangbo, R. McCann. The geometry of optimal transportation, *Acta Math.* **177**, 113-161 (1996).

[GS98] W. Gangbo, A. Swiech. Optimal Maps for the Multidimensional Monge-Kantorovich Problem. *Comm. Pure Applied Math.* **1**, 23-45 (1998).

[Ga08] L. Grafakos. Modern Fourier Analysis. *Springer.* 524pp (2008).

[GS95] G. M. Graf, D. Schenker. *On the molecular limit of Coulomb gases. Comm. Math. Phys.* **174**, 1, 215-227 (1995).

[Ge89] J.N. Gregg. The Existence of the Thermodynamic Limit in Coulomb-like Systems. *Comm. Math. Phys.* **123**, 255-276 (1989).

[HLS09a] C. Hainzl, M. Lewin, J. P. Solovej. The thermodynamic limit of quantum Coulomb systems. Part I. General theory. *Advances in Math.* **221**, 454-487 (2009).

[HLS09b] C. Hainzl, M. Lewin, J. P. Solovej. The thermodynamic limit of quantum Coulomb systems. Part II. Applications. *Advances in Math.* **221**,488-546 (2009).

[HK64] P. Hohenberg, W. Kohn. Inhomogeneous electron gas, *Phys. Rev. B* **136**, 864-871 (1964).

[HS02] C. Hainzl, R. Seiringer. General Decomposition of Radial Functions on Rn and Applications to N-Body Quantum Systems. *Letters in Mathematical Physics.* **61**, 1, 75-84 (2002).

[H89] W. Hughes. Thermodynamics for Coulomb Systems: A Problem at Vanishing Particle Densities. *Jour. Stat. Phys.* **41**, 5/6 (1985).

[KS65] W. Kohn, L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A* **140**, 1133-1138 (1965).

[Ly79] M. LEVY. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. USA* **76**, 12, 6062-6065 (1979).

[Lw15talk] M. LEWIN. https://team.inria.fr/mokaplan/files/2015/09/Dauphine.pdf (2015).

[Lw17] M. LEWIN. Semi-classical limit of the Levy-Lieb functional in Density Functional Theory. *preprint.* https://arxiv.org/abs/1706.02199 (2017).

[LwLi15] M. LEWIN, E. H. LIEB. Improved Lieb-Oxford exchange-correlation inequality with gradient correction. *Phys. Rev. A* **91**, 2, 022507 (2015).

[LLS17] M. LEWIN, E. H. LIEB, R. SEIRINGER. Statistical Mechanics of the Uniform Electron Gas. *preprint.* (2017).

[LeLi72] J.L. LEBOWITZ, E. LIEB. The constitution of matter: Existence of thermodynamics for systems composed of electrons and nuclei. In Statistical Mechanics, 17-24, Springer Berlin Heidelberg (1972).

[Li79] E.H. LIEB. A Lower Bound for Coulomb Energies. *Physics Letters* **70 A**, 444-446 (1979).

[Li83] E.H. LIEB. Density functionals for Coulomb systems, *International Journal of Quantum Chemistry* **24**, 243-277 (1983).

[LiLo] E. H. LIEB AND M. LOSS, *Analysis, vol. 14 of Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, second ed., (2001).

[LSY94] E. H. LIEB, J. P. SOLOVEJ, J. YNGVASON. Asymptotics of heavy atoms in high magnetic fields. II. Semiclassical regions. *Comm. Math. Phys.* **161**, 1, 77-124 (1994). Asymptotics of heavy atoms in high magnetic fields. II. Semiclassical regions

[LO81] E.H. LIEB, S. OXFORD. Improved lower bound on the indirect Coulomb energy, *International Journal of Quantum Chemistry* **19**, 427-439 (1981).

[LNP16] D. LUNDHOLM, P. THANH, F. PORTMANN. Fractional Hardy-Lieb-Thirring and Related Inequalities for Interacting Systems. *Arch. Rat. Mech. Anal.* **219**, 3, 1343-1382 (2016).

[P15] M. PETRACHE. Decorrelation as an avatar of convexity. *preprint.* https://arxiv.org/abs/1507.00782 (2015).

[PRN] M. PETRACHE, S. ROTA NODARI. Equidistribution of jellium energy for Coulomb and Riesz Interactions. *Constructive Approximation* **47**, 1, 163–210 (2018).

[PS] M. PETRACHE, S. SERFATY. Next-order asymptotics and renormalized energy for Riesz interactions. *J. Inst. Math. Jussieu.* **16**, 501-569 (2017).

[RR67] D. ROBINSON, D. RUELLE. Mean entropy of states in classical statistical mechanics *Comm. Math. Phys.* **5**, 4, 288-300 (1967).

[RoSer] S. ROTA-NODARI, S. SERFATY. Renormalized energy equidistribution and local charge balance in 2D Coulomb systems. *Int. Math. Res. Not. IMRN.* **11**, 3035-3093 (2015).

[R69] D. RUELLE. Statistical Mechanics: Rigorous Results, *World Scientific* (1969).

[SS15] E. SANDIER, S. SERFATY. 2D Coulomb gases and the renormalized energy, to appear in *Annals of Probability* **43**, 4, 2026-2083 (2015).

[Sd99] M. SEIDL. Strong-interaction limit of density-functional theory. *Physical Review A* **60**, 6, 4387 (1999).

[SGGS07] M. SEIDL, P. GORI-GIORGI, A. SAVIN. Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Phys. Rev. A*, **75**, 042511 (2007).

[SPL99] Michael Seidl, John P. Perdew, and Mel Levy. Strictly correlated electrons in density-functional theory. *Phys. Rev. A* **59**, 51-54 (1999).

[Sm98] S. Smale. Mathematical problems for the next century. *The Mathematical Intelligencer.* **20**, 2, 7-15 (1998).

[Tao] T. Tao, https://terrytao.wordpress.com/tag/hardy-littlewood-maximal-inequality/.

[V09] C. Villani. Optimal Transport Old and New. *Springer.* 1000pp (2009).