Concentration for Coulomb gases and Coulomb transport inequalities

Djalil Chafai, Adrien Hardy, Mylène Maïda

CONCENTRATION FOR COULOMB GASES AND COULOMB TRANSPORT INEQUALITIES

DJALIL CHAFAÏ, ADRIEN HARDY, AND MYLÈNE MAÏDA

Abstract. We study the non-asymptotic behavior of Coulomb gases in dimension two and more. Such gases are modeled by an exchangeable Boltzmann–Gibbs measure with a singular two-body interaction. We obtain concentration of measure inequalities for the empirical distribution of such gases around their equilibrium measure, with respect to bounded Lipschitz and Wasserstein distances. This implies macroscopic as well as mesoscopic convergence in such distances. In particular, we improve the concentration inequalities known for the empirical spectral distribution of Ginibre random matrices. Our approach is remarkably simple and bypasses the use of renormalized energy. It crucially relies on new inequalities between probability metrics, including Coulomb transport inequalities which can be of independent interest. Our work is inspired by the one of Maïda and Maurel-Segala, itself inspired by large deviations techniques. Our approach allows to recover, extend, and simplify previous results by Rougerie and Serfaty.

1. Introduction

The aim of this work is the non-asymptotic study of Coulomb gases around their equilibrium measure. We start by recalling some essential aspects of electrostatics, including the notion of equilibrium measure. We then incorporate randomness and present the Coulomb gas model followed by the natural question of concentration of measure for its empirical measure. This leads us to the study of inequalities between probability metrics, including new Coulomb transport type inequalities. We then state our results on concentration of measure and their applications and ramifications. We close the introduction with some additional notes and comments.

In all this work, we take $d \geq 2$.

1.1. **Electrostatics.** The d-dimensional *Coulomb kernel* is defined by

$$x \in \mathbb{R}^d \mapsto g(x) := \begin{cases} \log \dfrac{1}{|x|} & \text{if } d = 2, \\ \dfrac{1}{|x|^{d-2}} & \text{if } d \geq 3. \end{cases}$$

When $d = 3$, up to a multiplicative constant, $g(x)$ is the electrostatic potential at $x \in \mathbb{R}^3$ generated by a unit charge at the origin according to Coulomb's law. More generally, it is the fundamental solution of Poisson's equation. Namely, according to [Lieb and Loss, 1997, Th. 6.20], if we denote by $\Delta := \partial_1^2 + \cdots + \partial_d^2$ the Laplace

Date: Summer 2017. Compiled August 29, 2017.

2000 *Mathematics Subject Classification.* 26D10; 31A; 31B; 28A35; 60E15; 82B44; 28C15.

Key words and phrases. Coulomb gas; Ginibre ensemble; Random matrix; Transport of measure; Transport inequality; Talagrand inequality; Wasserstein distance; Concentration of measure.

This work was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

operator on $\mathbb{R}^d$ and by δ_0 the Dirac mass at the origin then, in the sense of Schwartz distributions,

$$\Delta g = -c_d\,\delta_0, \tag{1.1}$$

where c_d is a positive constant given by

$$c_d := \begin{cases} 2\pi & \text{if } d = 2, \\ (d-2)|\mathbb{S}^{d-1}| & \text{if } d \geq 3, \end{cases} \quad \text{with} \quad |\mathbb{S}^{d-1}| := \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

The function g is superharmonic on $\mathbb{R}^d$, harmonic on $\mathbb{R}^d \setminus \{0\}$, and belongs to the space $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ of locally Lebesgue-integrable functions.

Let $\mathcal{P}(\mathbb{R}^d)$ be the space of probability mesures on $\mathbb{R}^d$. For any $\mu \in \mathcal{P}(\mathbb{R}^d)$ with compact support, its *Coulomb energy*,

$$\mathcal{E}(\mu) := \iint g(x-y)\mu(\mathrm{d}x)\mu(\mathrm{d}y) \in \mathbb{R} \cup \{+\infty\}, \tag{1.2}$$

is well defined since g is bounded from below on any compact subset of $\mathbb{R}^d$ (actually we even have $g \geq 0$ when $d \geq 3$). Recall that a closed subset of $\mathbb{R}^d$ has *positive capacity* when it carries a compactly supported probability measure with finite Coulomb energy; otherwise it has *null capacity*. We refer for instance to [Helms, 2009, Landkof, 1972] for $d \geq 2$ and [Saff and Totik, 1997] for $d = 2$.

The quantity $\mathcal{E}(\mu)$ represents the electrostatic energy of the distribution μ of charged particles in $\mathbb{R}^d$. Moreover, given any $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with compact support and finite Coulomb energy, the quantity $\mathcal{E}(\mu-\nu)$ is well defined, finite, non-negative, and vanishes if and only if $\mu = \nu$, see [Saff and Totik, 1997, Lem. I.1.8] for $d = 2$ and [Lieb and Loss, 1997, Th. 9.8] for $d \geq 3$. In particular, for such probability measures, the following map is a metric:

$$(\mu, \nu) \mapsto \mathcal{E}(\mu - \nu)^{1/2}. \tag{1.3}$$

One can confine a distribution of charges by using an external potential. More precisely, an *admissible (external) potential* is a map $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that:

1. V is lower semicontinuous;
2. V is finite on a set of positive capacity;
3. V satisfies the growth condition

$$\lim_{|x|\to\infty} \big(V(x) - 2\log|x|\,\mathbf{1}_{d=2}\big) = +\infty.$$

Now, for an admissible potential V, we consider the functional

$$\mathcal{E}_V(\mu) := \iint \Big(g(x-y) + \frac{1}{2}V(x) + \frac{1}{2}V(y)\Big)\,\mu(\mathrm{d}x)\mu(\mathrm{d}y). \tag{1.4}$$

Note that, thanks to the assumptions on V, the integrand of (1.4) is bounded from below. Thus $\mathcal{E}_V(\mu)$ is well defined for every $\mu \in \mathcal{P}(\mathbb{R}^d)$, taking values in $\mathbb{R} \cup \{+\infty\}$. Moreover, when $\mu \in \mathcal{P}(\mathbb{R}^d)$ with $d \geq 3$ or $d = 2$ and $\int_{\mathbb{R}^2} \log(1+|x|)\mu(\mathrm{d}x) < \infty$, the Coulomb energy $\mathcal{E}(\mu)$ is still well defined and in this case one has

$$\mathcal{E}_V(\mu) = \mathcal{E}(\mu) + \int V(x)\mu(\mathrm{d}x). \tag{1.5}$$

It is known that if V is admissible then $\mathcal{E}_V$ has a unique minimizer μ_V on $\mathcal{P}(\mathbb{R}^d)$ called the *equilibrium measure*, which is compactly supported. Moreover, if V has a Lipschitz continuous derivative, then μ_V has a density given by

$$\rho_V = \frac{\Delta V}{2c_d} \tag{1.6}$$

on the interior of its support. For example when $V = |\cdot|^2$ this yields together with the rotational invariance that the probability measure μ_V is uniform on a ball. See for instance [López García, 2010, Prop. 2.13], [Chafaï et al., 2014], and [Serfaty, 2015, Sec. 2.3].

1.2. **Coulomb gas model.** Given $N \geq 2$ unit charges $x_1, \ldots, x_N \in \mathbb{R}^d$ in a confining potential V, the energy of the configuration is given by

$$H_N(x_1, \ldots, x_N) := \sum_{i \neq j} g(x_i - x_j) + N \sum_{i=1}^{N} V(x_i). \tag{1.7}$$

The prefactor N in front of the potential V turns out to be the appropriate scaling in the large N limit, because of the long range nature of the Coulomb interaction. The corresponding ensemble in statistical physics is the Boltzmann–Gibbs probability measure on $(\mathbb{R}^d)^N$ given by

$$\mathrm{d}\mathbb{P}^N_{V,\beta}(x_1, \ldots, x_N) := \frac{1}{Z^N_{V,\beta}} \exp\Big(-\frac{\beta}{2} H_N(x_1, \ldots, x_N)\Big) \mathrm{d}x_1 \cdots \mathrm{d}x_N, \tag{1.8}$$

where $\beta > 0$ is the inverse temperature parameter and where $Z^N_{V,\beta}$ is a normalizing constant known as the partition function. To ensure that the model is well defined, we assume that V is finite on a set of positive Lebesgue measure and

$$\int \mathrm{e}^{-\frac{\beta}{2}(V(x) - \mathbf{1}_{d=2} \log(1+|x|^2))} \mathrm{d}x < \infty. \tag{$\mathbf{H}_\beta$}$$

This indeed implies that $0 < Z^N_{V,\beta} < \infty$. The *Coulomb gas* or *one-component plasma* stands for the exchangeable random particles $x_1, \ldots, x_N$ living in $\mathbb{R}^d$ with joint law $\mathbb{P}^N_{V,\beta}$. We refer to [Forrester, 2010, Serfaty, 2015] and references therein for more mathematical and physical aspects of such models.

It is customary to encode the particles $x_1, \ldots, x_N$ by their empirical measure

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

It is known, at least when V is admissible and continuous, that for any $\beta > 0$ satisfying ($\mathbf{H}_\beta$) we have the weak convergence, with respect to continuous and bounded test functions,

$$\hat{\mu}_N \underset{N\to\infty}{\longrightarrow} \mu_V,$$

with probability one, in any joint probability space. This weak convergence follows from the Γ-convergence of $\frac{1}{N^2} H_N$ towards $\mathcal{E}_V$, or from the large deviation principle at speed N^2 and rate function $\frac{\beta}{2}(\mathcal{E}_V - \mathcal{E}_V(\mu_V))$ satisfied by $\hat{\mu}_N$, see [Serfaty, 2015, Chafaï et al., 2014]. More precisely, the large deviation principle yields the convergence, for any distance d on $\mathcal{P}(\mathbb{R}^d)$ metrizing the weak topology and $r > 0$,

$$\frac{1}{N^2} \log \mathbb{P}^N_{V,\beta}\Big(\mathrm{d}(\hat{\mu}_N, \mu_V) \geq r\Big) \underset{N\to\infty}{\longrightarrow} -\frac{\beta}{2} \inf_{\mathrm{d}(\mu,\mu_V)\geq r} \big(\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)\big).$$

In particular, there exist non-explicit constants $N_0, c, C > 0$ depending on β, V and $r > 0$ such that, for every $N \geq N_0$,

$$\mathrm{e}^{-cN^2} \leq \mathbb{P}^N_{V,\beta}\Big(\mathrm{d}(\hat{\mu}_N, \mu_V) \geq r\Big) \leq \mathrm{e}^{-CN^2}. \tag{1.9}$$

Beyond this large deviation principle, one may look for concentration inequalities, quantifying non-asymptotically the closeness of $\hat{\mu}_N$ to μ_V. Namely, can we improve the large deviation upper bound in (1.9) into an inequality holding for any $N \geq 2$ and with an explicit dependence in r? A quadratic dependence of C in r would yield a sub-Gaussian concentration inequality. Taking for d stronger distances, such as the Wasserstein distances W_p would be of interest too. For the Coulomb gas with $d = 2$ but restricted to the real line, also known as the *one-dimensional log-gas* popularized by random matrix theory (see [Forrester, 2010] and references therein), a concentration inequality for W_1 has been established by Maïda and Maurel-Segala [2014]. But for the usual Coulomb gases in dimension $d \geq 2$, no concentration inequality is available in the literature yet for metrics yielding the weak convergence.

Let us also mention that central limit theorems have been obtained for one dimensional log-gases by Johansson [1998], and for two dimensional Coulomb gases by Rider and Virág [2007], Ameur et al. [2011, 2015] when $\beta = 2$ and recently extended by Leblé and Serfaty [2016], Bauerschmidt et al. [2016] to any $\beta > 0$. As for the higher dimensional Coulomb gases, the fluctuations remain largely unexplored.

A related topic is the meso/microscopic study of Coulomb gases, which has been recently investigated by several authors including Sandier and Serfaty [2015], Bauerschmidt et al. [2015], Leblé [2015] for $d = 2$, Rougerie and Serfaty [2016], Petrache and Serfaty [2015], Leblé and Serfaty [2015] in any dimension. These works provide fine asymptotics for local observables of Coulomb gases, most of them using the machinery of the renormalized energy. For our purpose, an interesting deviation bound can be derived from [Rougerie and Serfaty, 2016]: Under extra regularity conditions for V, for any $d \geq 2$ and any fixed $r, R > 0$,

$$\mathbb{P}^N_{V,\beta}\left(\sup_{\|f\|_{\mathrm{Lip}}\leq 1}\int_{|x|\leq R} f(x)(\hat{\mu}_N - \mu_V)(\mathrm{d}x) \geq r\right) \leq \mathrm{e}^{-C_R r^2 N^2} \tag{1.10}$$

provided $N \geq N_0$, for some constants $C_R, N_0 > 0$ depending on R, β, V; we used the notation

$$\|f\|_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Indeed, getting rid of the constraint $|x| \leq R$ in the integral and of the dependence in R of the constant C_R would yield a concentration inequality in the W_1 metric, and this is exactly what we are aiming at. Although one could try to adapt [Rougerie and Serfaty, 2016]'s proof so as to extend the bound (1.10) to a non-local setting, we shall proceed here differently, without using the electric field machinery developed in the context of the renormalized energy. Our goal is to provide self-contained, short and simple proofs for concentration inequalities of this flavor, as well as providing more explicit constants and ranges of validity when it is possible.

More precisely in this work we provide, under appropriate regularity and growth conditions on V, concentration inequalities for the empirical measure of Coulomb gases in any dimension $d \geq 2$, with respect to both bounded Lipschitz and Wasserstein W_1 metrics. For both metrics, we obtain sub-Gaussian concentration inequalities of optimal order with respect to N, see Theorem 1.5 and Theorem 1.9. Our concentration inequalities are new for general Coulomb gases and Corollary 1.11 even improves the previous concentration bounds for the empirical spectral measure of the Ginibre random matrices, due to Pemantle and Peres [2014] and Breuer and Duits [2014], both results relying on the determinantal structure of the Ginibre ensemble. These inequalities are precise enough to yield weak and W_1 convergence of $\hat{\mu}_N$ to μ_V at mesoscopic scales, see Corollary 1.8. For completeness, we also provide estimates on the probability of having particles outside of arbitrarily large compact sets, see Theorem 1.12.

As for the proofs, a key observation is that the energy (1.7) is a function of the empirical measure $\hat{\mu}_N$. Indeed, with the identity (1.5) in mind, we can write

$$\mathrm{d}\mathbb{P}^N_{V,\beta}(x_1,\dots,x_N) = \frac{1}{Z^N_{V,\beta}} \exp\Big(-\frac{\beta}{2}N^2\,\mathcal{E}^{\neq}_V(\hat{\mu}_N)\Big)\mathrm{d}x_1\cdots\mathrm{d}x_N \tag{1.11}$$

where

$$\mathcal{E}^{\neq}_V(\hat{\mu}_N) := \iint_{x\neq y} g(x-y)\hat{\mu}_N(\mathrm{d}x)\hat{\mu}_N(\mathrm{d}y) + \int V(x)\hat{\mu}_N(\mathrm{d}x). \tag{1.12}$$

The quantity $\mathcal{E}^{\neq}_V(\hat{\mu}_N) - \mathcal{E}_V(\mu_V)$ is an energetic way to measure the closeness of $\hat{\mu}_N$ to μ_V. One may alternatively write it as a regularized L^2 norm of the electric field of $\hat{\mu}_N - \mu_V$, which leads to the notion of renormalized energy. As mentioned previously, it has been successfully used to obtain fine local asymptotics for Coulomb gases. With a concentration inequality involving global observables as a target, we bypass the use of this machinery by using functional inequalities, in the spirit of Talagrand transport inequalities, replacing the usual relative entropy by the Coulomb energy $\mathcal{E}_V$ or the Coulomb metric (1.3). Such inequalities, which we introduce now, may be of independent interest.

1.3. **Probability metrics and Coulomb transport inequality.** In the following, we consider the *bounded Lipschitz* (Fortet–Mourier) distance on $\mathcal{P}(\mathbb{R}^d)$ defined by

$$\mathrm{d}_{\mathrm{BL}}(\mu,\nu) := \sup_{\substack{\|f\|_{\mathrm{Lip}}\leq 1\\ \|f\|_\infty \leq 1}} \int f(x)(\mu-\nu)(\mathrm{d}x), \tag{1.13}$$

which metrizes its weak convergence, see [Dudley, 2002]. Moreover, for any $p \geq 1$, the *Kantorovich* or *Wasserstein distance* of order p between $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\nu \in \mathcal{P}(\mathbb{R}^d)$ is defined by

$$\mathrm{W}_p(\mu,\nu) := \Big(\inf \iint |x-y|^p \pi(\mathrm{d}x,\mathrm{d}y)\Big)^{1/p}, \tag{1.14}$$

where the infimum runs over all probability measures $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with marginal distributions μ and ν, when $|\cdot|^p$ is integrable for both μ and ν. We set $\mathrm{W}_p(\mu,\nu) = +\infty$ otherwise. Convergence in the W_p metric is equivalent to weak convergence plus convergence of moments up to order p. The Kantorovich-Rubinstein dual representation of W_1 [Villani, 2003, Rachev and Rüschendorf, 1998] reads

$$\mathrm{W}_1(\mu,\nu) = \sup_{\|f\|_{\mathrm{Lip}}\leq 1} \int f(x)(\mu-\nu)(\mathrm{d}x). \tag{1.15}$$

In particular, we readily have, for any $\mu,\nu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathrm{d}_{\mathrm{BL}}(\mu,\nu) \leq \mathrm{W}_1(\mu,\nu). \tag{1.16}$$

Our first result states that the Coulomb metric (1.3) locally dominates the W_1 metric, and hence the bounded Lipschitz metric as well. In particular, once restricted to probability measures on a prescribed compact set, the convergence in Coulomb distance implies the convergence in Wasserstein distance.

Theorem 1.1 (Local Coulomb transport inequality)**.** *For every compact subset $D \subset \mathbb{R}^d$, there exists a constant $C_D > 0$ such that, for every $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ supported in D with $\mathcal{E}(\mu) < \infty$ and $\mathcal{E}(\nu) < \infty$,*

$$\mathrm{W}_1(\mu,\nu)^2 \leq C_D \; \mathcal{E}(\mu-\nu).$$

We provide a proof for Theorem 1.1 in Section 2.

As we will see in the proof, a rough upper bound on the optimal constant C_D is given by the volume of the ball of radius four times the one of the domain D.

When $d = 2$ and $D \subset \mathbb{R}$, Theorem 1.1 yields [Popescu, 2013, Th. 1], up to the sharpness of the constant; we can say that Theorem 1.1 extends Popescu's local free transport inequality to higher dimensions.

How to go over the case of compactly supported measures and get rid of the dependence in the domain D of the constant C_D? To avoid restricting the distribution of charges μ to a prescribed bounded domain D, we add a confining external potential to the energy, namely we consider the weighted functional $\mathcal{E}_V$ introduced in (1.4) instead of $\mathcal{E}$. This will provide inequalities when one of the measure is the equilibrium measure μ_V; this is not very restrictive as any compactly supported probability measure such that $g * \mu$ is continuous is the equilibrium measure of a well chosen potential, see for instance [Landkof, 1972]. Our next result is a Coulomb transport type inequality for equilibrium measures.

Theorem 1.2 (Coulomb transport inequality for equilibrium measures)**.** *If the potential $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is admissible, then there exists a constant $C_{\mathrm{BL}}^V > 0$ such that, for every $\mu \in \mathcal{P}(\mathbb{R}^d)$,*

$$\mathrm{d}_{\mathrm{BL}}(\mu,\mu_V)^2 \leq C_{\mathrm{BL}}^V \Big(\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V) \Big). \tag{1.17}$$

If one further assumes that V grows at least quadratically,

$$\liminf_{|x|\to\infty} \frac{V(x)}{|x|^2} > 0, \tag{$\mathbf{H}_{\mathrm{W}_1}$}$$

then there exists a constant $C_{\mathrm{W}_1}^V > 0$ such that, for every $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathrm{W}_1(\mu,\mu_V)^2 \leq C_{\mathrm{W}_1}^V (\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)). \tag{1.18}$$

Theorem 1.2 is proved in Section 3.

Given a potential V, one can obtain rough upper bounds on the optimal constants $C_{\mathrm{BL}}^V, C_{\mathrm{W}_1}^V$ by following carefully the proof of the theorem.

Remark 1.3 (Free transport inequalities)**.** *A useful observation is that the problem of minimizing $\mathcal{E}_V(\mu)$ over $\mu \in \mathcal{P}(S)$, where $S \subset \mathbb{R}^d$ has positive capacity, is equivalent to consider the full minimization problem but setting $V = +\infty$ on $\mathbb{R}^d \setminus S$. In particular, by taking any admissible V on $\mathbb{R}$ and setting $V = +\infty$ on $\mathbb{R}^2 \setminus \mathbb{R}$, Theorem 1.2 with $d = 2$ yields the free transport inequality obtained by Maïda and Maurel-Segala [2014], without the restriction that V is continuous, and Popescu [2013]. Notice that although the lower semicontinuity of V is not assumed in [Popescu, 2013], it is actually a necessary condition for the existence of μ_V in general.*

Remark 1.4 (Optimality of the growth condition for W_1)**.** *Following [Maïda and Maurel-Segala, 2014, Rem. 1], one can check that (1.18) cannot hold if V is admissible but does not satisfy ($\mathbf{H}_{\mathrm{W}_1}$). Indeed, letting ν_n be the uniform probability measure on the ball of $\mathbb{R}^d$ of radius one centered at $(n, 0, \ldots, 0)$, then $\mathrm{W}_1(\nu_n, \mu_V)^2$ grows like n^2 as $n \to \infty$, whereas $\mathcal{E}_V(\nu_n)$ grows like $\int V \mathrm{d}\nu_n = o(n^2)$.*

Equipped with these metric inequalities, we are now in position to state the concentration inequalities.

1.4. **Concentration of measure for Coulomb gases.** Our first result on concentration of measure is the following.

Theorem 1.5 (Concentration of measure for Coulomb gases)**.** *Assume that V is $\mathcal{C}^2$ on $\mathbb{R}^d$ and that its Laplacian ΔV satisfies the following growth constraint*

$$\limsup_{|x|\to\infty}\left(\frac{1}{V(x)}\sup_{\substack{y\in\mathbb{R}^d\\|y-x|<1}}\Delta V(y)\right)<2(d+2). \tag{1.19}$$

If V is admissible then there exist constants $a>0$, $b\in\mathbb{R}$, and a function $\beta\mapsto c(\beta)$ such that, for any $\beta>0$ which satisfies ($\mathbf{H}_\beta$), any $N\geq 2$, and any $r>0$,

$$\mathbb{P}^N_{V,\beta}\Big(\mathrm{d}_{\mathrm{BL}}(\hat{\mu}_N,\mu_V)\geq r\Big)\leq \mathrm{e}^{-a\beta N^2r^2+\mathbf{1}_{d=2}(\frac{\beta}{4}N\log N)+b\beta N^{2-2/d}+c(\beta)N}. \tag{1.20}$$

If there exists $\kappa>0$ such that $\liminf_{|x|\to\infty}\dfrac{V(x)}{|x|^\kappa}>0$, then

$$c(\beta)=\begin{cases}\mathcal{O}(\log\beta), & \text{as }\beta\to 0,\\ \mathcal{O}(\beta), & \text{as }\beta\to\infty.\end{cases} \tag{1.21}$$

If V satisfies ($\mathbf{H}_{\mathrm{W}_1}$), then the same holds true when replacing d_{BL} by W_1.

Theorem 1.5 is proved in Section 4. The proof relies on the Coulomb transport inequality of Theorem 1.2, hence the assumption ($\mathbf{H}_{\mathrm{W}_1}$) in the second part. The constraint (1.19) in Theorem 1.5 that ΔV does not grow faster than V is technical. It allows potentials growing like $|\cdot|^\kappa$ for any $\kappa>0$ or $\exp|\cdot|$, but not $\exp(|\cdot|^2)$. As for the regularity condition, we assume for convenience that V is $\mathcal{C}^2$ but much less is required, see Remark 1.10.

If $\beta>0$ is fixed, then under the last set of assumptions of Theorem 1.5, there exist constants $u,v>0$ depending on β and V only such that, for any $N\geq 2$ and

$$r\geq\begin{cases}v\sqrt{\frac{\log N}{N}} & \text{if } d=2,\\ vN^{-1/d} & \text{if } d\geq 3,\end{cases} \tag{1.22}$$

we have

$$\mathbb{P}^N_{V,\beta}\Big(\mathrm{W}_1(\hat{\mu}_N,\mu_V)\geq r\Big)\leq \mathrm{e}^{-uN^2r^2}. \tag{1.23}$$

Such a bound has been obtained in [Maïda and Maurel-Segala, 2014, Th. 5] for the empirical measure of one-dimensional log-gases.

Remark 1.6 (Sharpness)**.** *The concentration of measure bound (1.23) is of optimal order with respect to N, as one can check using (1.9) and the stochastic dominance*

$$\mathbb{P}^N_{V,\beta}\Big(\mathrm{d}_{\mathrm{BL}}(\hat{\mu}_N,\mu_V)\geq r\Big)\leq\mathbb{P}^N_{V,\beta}\Big(\mathrm{W}_1(\hat{\mu}_N,\mu_V)\geq r\Big).$$

The concentration inequalities provided by Theorem 1.5 are much more precise than the large deviation upper bound in (1.9) since it holds for every $N\geq 2$ and the dependence in r,β is explicit.

Combined with the Borel–Cantelli lemma, Theorem 1.5 directly yields the convergence of $\hat{\mu}_N$ to μ_V in the Wasserstein distance even when one allows β to depend on N, provided it does not go to zero too fast, thanks to (1.21).

Corollary 1.7 (W$_1$ convergence)**.** *Under the last set of assumptions of Theorem 1.5, there exists a constant $\beta_V > 0$ such that the following holds: Given any sequence of positive real numbers (β_N) such that*

$$\beta_N \geq \beta_V \frac{\log N}{N}$$

for every N large enough, then under $\mathbb{P}^N_{V,\beta_N}$,

$$\lim_{N\to\infty} \mathrm{W}_1(\hat{\mu}_N, \mu_V) = 0$$

with probability one, in any joint probability space.

Moreover, if one keeps $\beta > 0$ fixed and lets $r \to 0$ with N, then Theorem 1.5 is precise enough to yield the convergence of $\hat{\mu}_N$ towards μ_V at the mesoscopic scale, that is after zooming on the particle system around any fixed $x_0 \in \mathbb{R}^d$ at the scale N^{-s} for any $0 \leq s < 1/d$. More precisely, set $\tau^{N^s}_{x_0}(x) := N^s(x - x_0)$ and let $\tau^{N^s}_{x_0}\mu$ be the the push-forward of $\mu \in \mathcal{P}(\mathbb{R}^d)$ by the map $\tau^{N^s}_{x_0}$, characterized by

$$\int f(x)\, \tau^{N^s}_{x_0}\mu(\mathrm{d}x) = \int f\big(N^s(x - x_0)\big)\mu(\mathrm{d}x) \tag{1.24}$$

for any Borel function $f : \mathbb{R}^d \to \mathbb{R}$.

Corollary 1.8 (Mesoscopic convergence)**.** *Under the first set of assumptions of Theorem 1.5, for any $\beta > 0$ satisfying $(\mathbf{H}_\beta)$ there exist constants $C, c > 0$ such that, for any $x_0 \in \mathbb{R}^d$, any $s \geq 0$, and any $N \geq 2$, we have when $d = 2$,*

$$\mathbb{P}^N_{V,\beta}\Big(\, \mathrm{d_{BL}}\,(\tau^{N^s}_{x_0}\hat{\mu}_N, \tau^{N^s}_{x_0}\mu_V) \geq CN^s\sqrt{\frac{\log N}{N}}\Big) \leq \mathrm{e}^{-cN\log N}, \tag{1.25}$$

and we have instead when $d \geq 3$,

$$\mathbb{P}^N_{V,\beta}\Big(\, \mathrm{d_{BL}}\,(\tau^{N^s}_{x_0}\hat{\mu}_N, \tau^{N^s}_{x_0}\mu_V) \geq CN^{s-1/d}\Big) \leq \mathrm{e}^{-cN^{2-2/d}}. \tag{1.26}$$

Under the last set of assumptions of Theorem 1.5, the same holds true after replacing $\mathrm{d_{BL}}$ *by* W_1 *.*

The proof of the corollary requires a few lines which are provided at the end of Section 4.

One may also look for a concentration inequality where the constants are explicit in terms of V. We are able to derive such a statement when ΔV is bounded above. Recall that the Boltzmann–Shannon entropy of $\mu \in \mathcal{P}(\mathbb{R}^d)$ is defined by

$$\mathrm{S}(\mu) := -\int \frac{\mathrm{d}\mu}{\mathrm{d}x} \log \frac{\mathrm{d}\mu}{\mathrm{d}x}\, \mathrm{d}x, \tag{1.27}$$

with $\mathrm{S}(\mu) := +\infty$ if μ is not absolutely continuous with respect to the Lebesgue measure. Note that $\mathrm{S}(\mu)$ takes its values in $\mathbb{R} \cup \{+\infty\}$ and is finite if and only if μ admits a density f with respect to the Lebesgue measure such that $f \log f \in L^1(\mathbb{R}^d)$. In particular, if $V \in \mathcal{C}^2(\mathbb{R}^d)$ then $S(\mu_V)$ is indeed finite.

Theorem 1.9 (Concentration for potentials with bounded Laplacian)**.** *Assume that V is $\mathcal{C}^2$ on $\mathbb{R}^d$ and that there exists $D > 0$ with*

$$\sup_{y\in\mathbb{R}^d} \Delta V(y) \leq D. \tag{1.28}$$

Then Theorem 1.5 holds with the following constants:

$$\begin{aligned} a &= \frac{1}{8C^V_{\mathrm{BL}}}, \\ b &= \frac{1}{2}\Big(\frac{1}{C^V_{\mathrm{BL}}} + \mathcal{E}(\lambda_1) + \frac{D}{2(d+2)}\Big), \\ c(\beta) &= \frac{\beta}{2}\int V(x)\mu_V(\mathrm{d}x) - \mathrm{S}(\mu_V) + \log\int \mathrm{e}^{-\frac{\beta}{2}(V(x)-\mathbf{1}_{d=2}\log(1+|x|^2))}\mathrm{d}x, \end{aligned}$$

where the constant C^V_{BL} is as in Theorem 1.2, and where $\mathcal{E}(\lambda_1)$ is the Coulomb energy of the uniform law λ_1 on the unit ball of $\mathbb{R}^d$,

$$\mathcal{E}(\lambda_1) = \begin{cases} \frac{1}{4} & \text{if } d = 2, \\ \frac{2d}{d+2} & \text{if } d \geq 3. \end{cases}$$

Under ($\mathbf{H}_{\mathrm{W}_1}$), the constants a, b, $c(\beta)$ are given by the same formulas after replacing the constant C^V_{BL} by $C^V_{\mathrm{W}_1}$.

Theorem 1.9 is also proved in Section 4.

Remark 1.10 (Regularity assumptions)**.** *The assumption that V is $\mathcal{C}^2$ in theorems 1.5 and 1.9 is made to ease the presentation of the results and may be considerably weakened. As we can check in the proofs, if we assume for instance that V is finite on a set of positive Lebesgue measure and if the Boltzmann–Shannon entropy $\mathrm{S}(\mu_V)$ of the equilibrium measure μ_V is finite, and if $V = \tilde{V} + h$ where $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is superharmonic and $\tilde{V}$ is twice differentiable such that $\Delta\tilde{V}$ satisfies the condition (1.19), resp. (1.28), then same conclusion as in Theorem 1.5, resp. Theorem 1.9, holds for the d_{BL} metric. If moreover V satisfies the growth assumption ($\mathbf{H}_{\mathrm{W}_1}$), then they also hold for W_1.*

A remarkable consequence of Theorem 1.9 is an improvement of a concentration inequality in random matrix theory. Indeed, when $d = 2$, $V = |\cdot|^2$, and $\beta = 2$, the Coulomb gas

$$\mathrm{d}\mathbb{P}^N_{|\cdot|^2,2}(x_1,\dots,x_N) = \frac{1}{\pi^{N^2}\prod_{k=1}^N k!}\prod_{i<j}|x_i - x_j|^2 \mathrm{e}^{-N\sum_{i=1}^N |x_i|^2}\mathrm{d}x_1\cdots\mathrm{d}x_N$$

coincides with the joint law of the eigenvalues of the Ginibre ensemble of $N \times N$ non-Hermitian random matrices, after identifying $\mathbb{C}^N \simeq (\mathbb{R}^2)^N$, see [Ginibre, 1965]. The associated equilibrium μ_V measure is the circular law:

$$\mu_\circ(\mathrm{d}x) := \frac{\mathbf{1}_{\{x\in\mathbb{R}^2 :\, |x|\leq 1\}}}{\pi}\mathrm{d}x.$$

We also refer to [Khoruzhenko and Sommers, 2011], [Mehta, 2004, Forrester, 2010, Ch. 15], [Bordenave and Chafaï, 2012] for more information. The following result is a direct application of Theorem 1.9 together with easy computations for the constants.

Corollary 1.11 (Concentration for eigenvalues of Gaussian random matrices)**.** *Let X and Y be independent real Gaussian random variables with mean 0 and variance 1/2. For any N, let $\mathbf{M}_N$ be the $N \times N$ random matrix with i.i.d. entries distributed as the random complex variable $X + iY$. Let $\mathbb{P}_N$ be the joint law of the eigenvalues $x_1,\dots,x_N$ of $\frac{1}{\sqrt{N}}\mathbf{M}_N$, and consider the empirical spectral measure*

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

Then, for any $N \geq 2$ and any $r > 0$, we have

$$\mathbb{P}_N\Big(\mathrm{W}_1(\hat{\mu}_N, \mu_\circ) \geq r\Big) \leq \mathrm{e}^{-\frac{1}{4C}N^2r^2+\frac{1}{2}N\log N+N[\frac{1}{C}+\frac{3}{2}-\log\pi]}, \tag{1.29}$$

where C is the constant $C^{\mathrm{W}_1}_{|\cdot|^2}$ as in Theorem 1.2 for $d = 2$.

For a fixed test function, similar concentration inequalities for the Ginibre ensemble have been obtained by Pemantle and Peres [2014] and Breuer and Duits [2014], but with asymptotic rate e^{-Nr} instead of $\mathrm{e}^{-N^2r^2}$. Corollary 1.11 provides the almost sure convergence in the Wassertein W_1 metric of the empirical spectral measure to the circular law. We refer to [Meckes and Meckes, 2016] for a survey on empirical spectral measures, focusing on coupling methods and Wasserstein distances, and covering in particular the Ginibre ensemble. We do not know how to deduce the concentration inequality provided by Corollary 1.11 from the Gaussian nature of the entries of $\mathbf{M}_N$. Indeed, the eigenvalues of a non-normal matrix are not Lipschitz functions of its entries, in contrast with the singular values for which the Courant–Fischer formulas and the Hoffman–Wielandt inequality hold, see [Bordenave and Chafaï, 2012] for further details.

It is quite natural to ask about the behavior of the support of the random empirical measure $\hat{\mu}_N$ under $\mathbb{P}^N_{V,\beta}$ when N is large. The following theorem gives an answer.

Theorem 1.12 (Exponential tightness)**.** *Assume V is admissible, finite on a set of positive Lebesgue measure, satisfies the growth assumption*

$$\lim_{|x|\to\infty} \big(V(x) - (2+\varepsilon)\log|x|\,\mathbf{1}_{d=2}\big) = +\infty,$$

for some $\varepsilon > 0$, and that μ_V has finite Boltzmann–Shannon entropy $\mathrm{S}(\mu_V)$. Then, for any β satisfying ($\mathbf{H}_\beta$), there exist constants $r_0 > 0$ and $c > 0$ which may depend on β and V such that, for any N and $r \geq r_0$, we have,

$$\mathbb{P}^N_{V,\beta}(\mathrm{supp}(\hat{\mu}_N) \not\subset B_r) = \mathbb{P}^N_{V,\beta}\big(\max_{1\leq i\leq N}|x_i| \geq r\big) \leq \mathrm{e}^{-cNV_*(r)},$$

where $B_r := \{x \in \mathbb{R}^d : |x| \leq r\}$ and $V_*(r) := \min_{|x|\geq r} V(x)$.

The proof of Theorem 1.12 is given in Section 5.

Since for any admissible V we have $V_*(r) \to +\infty$ as $r \to \infty$, the theorem states that the maximum modulus $\max|x_i|$ is exponentially tight at speed N. In particular, the Borel–Cantelli lemma gives that

$$\limsup_{N\to\infty}\max_{1\leq i\leq N}|x_i| < \infty$$

holds with probability one in any joint probability space.

Theorem 1.12 allows to obtain W_p versions of Theorem 1.5, without much efforts. Indeed, Theorem 1.12 allows to restrict to the event $\{\max_{1\leq i\leq N}|x_i| \leq M\}$ with high probability. On the other hand, the distances d_{BL} and W_p for any $p \geq 1$ are all equivalent on a compact set: for any $p \geq 1$ and any probability measures μ, ν supported in the ball of $\mathbb{R}^d$ of radius $M \geq 1$,

$$\mathrm{W}_p^p(\mu,\nu) \leq (2M)^{p-1}\,\mathrm{W}_1(\mu,\nu) \leq M(2M)^{p-1}\,\mathrm{d}_{\mathrm{BL}}(\mu,\nu).$$

For example, for $p = 2$, this gives, under ($\mathbf{H}_{\mathrm{W}_1}$), an inequality of the form

$$\mathbb{P}^N_{V,\beta}(W_2(\hat{\mu}_N, \mu_V) \geq r) \leq 2\mathrm{e}^{-cN^{3/2}r^2}.$$

The same idea allows to deduce the almost sure convergence of the empirical measure to the equilibrium measure with respect to W_p for any $p \geq 1$.

1.5. **Notes, comments, and open problems.**

1.5.1. *Classical, free, and Coulomb transport inequalities.* The inequalities involving the Wasserstein W_1 metric in Theorem 1.1 and 1.2 can be seen as Coulomb analogues of classical transport inequalities, hence their names. Recall that *Kullback-Leibler divergence* of μ with respect to ν is given by

$$\mathrm{H}(\mu \mid \nu) := \int \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \log \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \,\mathrm{d}\nu, \tag{1.30}$$

with $\mathrm{H}(\mu \mid \nu) := +\infty$ if μ is not absolutely continuous with respect to ν. This functional is also known as relative entropy or free energy, see for instance [Chafaï, 2015]. A probability measure $\nu \in \mathcal{P}(\mathbb{R}^d)$ satisfies the *transport inequality* of order p when for some constant $C > 0$ and every $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathrm{W}_p(\mu,\nu)^2 \leq C\,\mathrm{H}(\mu \mid \nu). \tag{\textbf{T}}$$

A vast vast literature is devoted to the study of transport inequalities, see the surveys [Gozlan and Léonard, 2010] and [Villani, 2003] for more details. Following for instance [Gozlan and Léonard, 2010], we want to emphasize the deep link between (**T**) and Sanov's large deviation principle for the empirical measure of i.i.d random variables with law ν, where the rate function is given by $\mu \mapsto \mathrm{H}(\mu \mid \nu)$ and the speed is N. Similarly, as already mentioned, the empirical measure of a Coulomb gas in any dimension $d \geq 2$ has been shown by Chafaï et al. [2014] to satisfy a large deviation principle, this time with rate function $\mathcal{E}_V - \mathcal{E}_V(\mu_V)$ and speed N^2. Moreover, one can check that if the support of μ is contained in the one of μ_V, then $\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V) = \mathcal{E}(\mu - \mu_V)$, and the squared Coulomb metric $\mathcal{E}(\mu-\nu)$ can also be seen as the counterpart of the Kullback-Leibler divergence $\mathrm{H}(\mu \mid \nu)$. Inequalities of the same flavor for probability measures on the real line, linking the Wassertein W_1 or W_2 metrics with the functionals $\mathcal{E}(\mu - \nu)$ or $\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)$ with $d = 2$, have been previously obtained in the context of free probability by Biane and Voiculescu [2001], Hiai et al. [2004], Ledoux and Popescu [2009], Maïda and Maurel-Segala [2014], Popescu [2013]. In this setting, these functionals are usually referred to as free entropies and are related to the large deviation principle due to Ben Arous and Guionnet [1997] for the one-dimensional log-gas associated to unitarily invariant ensembles in random matrix theory.

There is also a deep connection between the inequality (**T**) and concentration of measure for Lipschitz functions. In particular, Marton [1986] and later Bobkov and Götze [1999] have shown that (**T**) with $p = 1$ is equivalent to sub-Gaussian concentration, while Gozlan [2009] has shown that (**T**) with $p = 2$ is equivalent to dimension-free sub-Gaussian concentration. However, the way we deduce the concentration of measure for Coulomb gases from the Coulomb transport inequality is of different nature, and is inspired from Maïda and Maurel-Segala [2014].

Ledoux and Popescu [2009] and Popescu [2013] provide W_2 free transport inequalities. It is therefore natural to ask about a W_2 variant of Theorem 1.2. In the same spirit, one can study Coulomb versions of related functional inequalities such as the logarithmic Sobolev inequalities (not the Hardy-Littlewood-Sobolev inequalities). It is not likely however that Theorem 1.1 has an extension to other W_p distances, as Popescu [2013] showed this is not true in the free setting.

1.5.2. *Varying potentials and conditional gases.* Theorem 1.5 still holds when V depends mildly on N. This can be useful for conditional Coulomb gases. Indeed, the conditional law $\mathbb{P}^N_{V,\beta}(\cdot \mid x_N)$ is again a Coulomb gas $\mathbb{P}^{N-1}_{\tilde V,\beta}$ with potential

$$\tilde V := \frac{N}{N-1} V + \frac{2}{N-1} g(x_N - \cdot),$$

which is covered by our results since g is superharmonic, see Remark 1.10.

1.5.3. *Weakly confining potentials.* In this work, we consider only admissible potentials and we do not allow the weakly confining potentials considered in [Hardy, 2012], in relation in particular with the Cauchy ensemble of random matrices studied by Forrester and Krishnapur, see [Forrester, 2010, Sec. 2.8]. In this case the equilibrium measure μ_V is no longer compactly supported. The derivation of concentration of measure inequalities for such Coulomb gases is still open.

1.5.4. *Random matrices.* It is likely that the concentration inequality provided by Corollary 1.11 remains valid beyond the Gaussian case. This is still open, even when the entries of $\mathbf{M}_N$ are i.i.d. ± 1 with probability $1/2$.

1.5.5. *Inverse temperature and crossover.* With the parameter β depending on N, Corollary 1.7 states that $\mathrm{W}_1(\hat\mu_N, \mu_V) \to 0$ as $N \to \infty$ provided that $\beta \geq \beta_V \log N/N$. Maybe this still holds as long as $N\beta \to \infty$. However, when β is of order $1/N$, we no longer expect concentration around μ_V but, having in mind for instance [Bodineau and Guionnet, 1999] or [Allez et al., 2012], it is likely that there is still concentration of measure around a limiting distribution which is a crossover between μ_V and a Gaussian distribution.

Acknowledgements. We would like to thank Sylvia Serfaty for her useful comments on this work.

Notations. We denote by $B_R = \{x \in \mathbb{R}^d : |x| \leq R\}$ the centered closed ball of radius R of $\mathbb{R}^d$, by λ_R the uniform probability measure on B_R, by $\mathrm{vol}(B)$ the Lebesgue measure of a Borel set $B \subset \mathbb{R}^d$, and by $\mathrm{supp}(\mu)$ the support of $\mu \in \mathcal{P}(\mathbb{R}^d)$.

2. Proof of Theorem 1.1

2.1. **Core of the proof.** We now give a proof of Theorem 1.1 up to technical lemmas. We postpone their proofs to the next subsection.

The starting point is the Kantorovich–Rubinstein dual representation (1.15) of the Wasserstein W_1 metric, namely

$$\mathrm{W}_1(\mu,\nu) = \sup_{\|f\|_{\mathrm{Lip}} \leq 1} \int f(x)(\mu - \nu)(\mathrm{d}x).$$

A theorem due to Rademacher states that if $\|f\|_{\mathrm{Lip}} < \infty$ then f is differentiable almost everywhere and $\|f\|_{\mathrm{Lip}} = \|\nabla f\|_\infty$, where $\|\cdot\|_\infty$ stands for the L^∞ norm.

The following lemma shows that we can localize the functions in the supremum, provided the measures are supported in a compact set.

Lemma 2.1 (Localization)**.** *For any $D \subset \mathbb{R}^d$ compact, there exists a compact set $D_+ \subset \mathbb{R}^d$ such that:*

(a) $D \subset D_+$
(b) $\mathrm{vol}(D_+) > 0$

(c) *for every $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ supported in D,*

$$\mathrm{W}_1(\mu,\nu) = \sup_{\substack{f \in \mathcal{C}(D_+) \\ \|f\|_{\mathrm{Lip}} \leq 1}} \int f(x)(\mu-\nu)(\mathrm{d}x), \tag{2.1}$$

where $\mathcal{C}(D_+)$ is the set of continuous functions $f : \mathbb{R}^d \to \mathbb{R}$ supported in D_+.

Next, let μ and ν be any probability measures supported in a compact set $D \subset \mathbb{R}^d$ satisfying $\mathcal{E}(\mu) < \infty$ and $\mathcal{E}(\nu) < \infty$, where the Coulomb energy $\mathcal{E}$ has been defined in (1.2). Assume further for now that $\mu - \nu$ has a $\mathcal{C}^\infty$ density h with respect to the Lebesgue measure which is compactly supported; we will remove this extra assumption at the end of this section.

Set $\eta := \mu - \nu$ for convenience. We first gather a few properties of the Coulomb potential of η defined by

$$U^\eta(x) := g * \eta(x) = \int g(x-y)h(y)\mathrm{d}y, \qquad x \in \mathbb{R}^d.$$

Since $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, we see that U^η is well defined. Since moreover $h \in \mathcal{C}^\infty(\mathbb{R}^d)$ so does U^η. We claim that $\nabla U^\eta \in L^2(\mathbb{R}^d)$. Indeed, if we set $\alpha_d := \max(d-2,1)$, then

$$\nabla g(x) = -\alpha_d \frac{x}{|x|^d}.$$

Using that η has compact support and that $\eta(\mathbb{R}^d) = 0$, we have as $|x| \to \infty$,

$$|\nabla U^\eta(x)|^2 = \alpha_d^2 \iint \frac{(x-y)\cdot(x-z)}{|x-y|^d |x-z|^d}\,\eta(\mathrm{d}y)\eta(\mathrm{d}z) = \frac{\alpha_d^2}{|x|^{2d-1}}(1+o(1)),$$

from which our claim follows. Finally, Poisson's equation (1.1) yields

$$\Delta U^\eta(x) = -c_d\, h(x), \qquad x \in \mathbb{R}^d. \tag{2.2}$$

Indeed, (1.1) states that for any test function $\varphi \in \mathcal{C}^\infty(\mathbb{R}^d)$ with compact support,

$$\int \Delta\varphi(y) g(y)\mathrm{d}y = -c_d\, \varphi(0),$$

see [Lieb and Loss, 1997, Th. 6.20], and (2.2) follows by taking $\varphi(y) = h(x-y)$.

Now, take any Lipschitz function $f \in \mathcal{C}(D_+)$, where the compact set D_+ has been introduced in Lemma 2.1. We have by (2.2) that

$$\int f(x)(\mu-\nu)(\mathrm{d}x) = \int f(x)h(x)\mathrm{d}x = -\frac{1}{c_d}\int f(x)\Delta U^\eta(x)\mathrm{d}x. \tag{2.3}$$

Because $\nabla U^\eta, \nabla f \in L^2(\mathbb{R}^d)$ and $\Delta U^\eta \in C^\infty(\mathbb{R}^d)$ with compact support, one can use integration by parts (see for instance [Lieb and Loss, 1997, Th. 7.7]) to obtain

$$-\int f(x)\Delta U^\eta(x)\mathrm{d}x = \int \nabla f(x)\cdot \nabla U^\eta(x)\mathrm{d}x. \tag{2.4}$$

By using the Cauchy–Schwarz inequality in $\mathbb{R}^d$ and then in $L^2(\mathbb{R}^d)$, we have

$$\begin{aligned}\int \nabla f(x)\cdot\nabla U^\eta(x)\mathrm{d}x &\leq \int |\nabla f(x)||\nabla U^\eta(x)|\mathrm{d}x \\ &\leq \|f\|_{\mathrm{Lip}} \int_{D_+} |\nabla U^\eta(x)|\mathrm{d}x \\ &\leq \|f\|_{\mathrm{Lip}} \Big(\mathrm{vol}(D_+)\int |\nabla U^\eta(x)|^2\mathrm{d}x\Big)^{1/2}. \end{aligned} \tag{2.5}$$

By using again an integration by parts,

$$\int |\nabla U^\eta(x)|^2 \mathrm{d}x = -\int U^\eta(x)\Delta U^\eta(x)\mathrm{d}x = c_d \int U^\eta(x)h(x)\mathrm{d}x = c_d\,\mathcal{E}(\eta). \tag{2.6}$$

Finally, by combining (2.1)–(2.6) and Lemma 2.1 we get a proof of Theorem 1.1 under the extra assumption that μ and ν have a $\mathcal{C}^\infty$ density, with $C_D := \mathrm{vol}(D_+)$.

We finally remove this extra assumption by using a density argument. Assume μ and ν are probability measures supported in a compact set $D \subset \mathbb{R}^d$ with $\mathcal{E}(\mu), \mathcal{E}(\nu) < \infty$. Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be such that

$$\varphi \in \mathcal{C}^\infty(B_1), \quad \varphi \geq 0, \quad \int \varphi(x)\mathrm{d}x = 1,$$

and set $\varphi_\varepsilon(x) = \varepsilon^{-d}\varphi(\varepsilon^{-1}x)$. Then both $\varphi_\varepsilon * \mu$ and $\varphi_\varepsilon * \nu$ have a $\mathcal{C}^\infty$ density supported on an ε-neighborhood of D. Thus, by the previous step, if D^δ stands for the δ-neighborhood of the set D, then for every $0 < \varepsilon < \delta < 1$,

$$\mathrm{W}_1(\varphi_\varepsilon * \mu, \varphi_\varepsilon * \nu)^2 \leq \mathrm{vol}(D^\delta_+)\,\mathcal{E}(\varphi_\varepsilon * \mu - \varphi_\varepsilon * \nu). \tag{2.7}$$

The probability measures $\varphi_\varepsilon * \mu$ and $\varphi_\varepsilon * \nu$ converge weakly to μ and ν respectively as $\varepsilon \to 0$. Since all these measures are moreover supported in the compact set D^1, $\mathrm{W}_1(\varphi_\varepsilon * \mu, \varphi_\varepsilon * \nu) \to \mathrm{W}_1(\mu, \nu)$ as $\varepsilon \to 0$. Thus, letting $\delta \to 0$ in (2.7), it is enough to prove that $\mathcal{E}(\varphi_\varepsilon * \mu - \varphi_\varepsilon * \nu) \to \mathcal{E}(\mu - \nu)$ as $\varepsilon \to 0$ in order to complete the proof of the theorem with $C_D := \mathrm{vol}(D_+)$. This is a consequence of the next lemma, which is proven at the end of the section and may be of independent interest. Consider the bilinear form

$$\mathcal{E}(\mu, \nu) = \iint g(x-y)\mu(\mathrm{d}x)\nu(\mathrm{d}y), \tag{2.8}$$

which is well defined for any $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with compact support.

Note that the proof above is not that far from inequalities between Sobolev norms.

Lemma 2.2 (Regularization)**.** *Let μ, ν be two positive finite Borel measures on $\mathbb{R}^d$, with compact support, and satisfying $\mathcal{E}(\mu, \nu) < \infty$. Let $\varphi \in L^\infty(B_1)$ be positive, satisfying $\int \varphi(x)\mathrm{d}x = 1$, and set $\varphi_\varepsilon(x) = \varepsilon^{-d}\varphi(\varepsilon^{-1}x)$. Then, we have*

$$\lim_{\varepsilon \to 0} \mathcal{E}(\varphi_\varepsilon * \mu, \varphi_\varepsilon * \nu) = \mathcal{E}(\mu, \nu).$$

Now, since μ and ν have the same total mass, compact support, $\mathcal{E}(\mu), \mathcal{E}(\nu) < \infty$ and, as mentioned right after (1.2), $\mathcal{E}(\mu - \nu) \geq 0$, we have that $2\,\mathcal{E}(\mu, \nu) \leq \mathcal{E}(\mu) + \mathcal{E}(\nu) < \infty$. Thus Lemma 2.2 applies and gives that

$$\mathcal{E}(\varphi_\varepsilon * \mu - \varphi_\varepsilon * \nu) = \mathcal{E}(\varphi_\varepsilon * \mu) - 2\,\mathcal{E}(\varphi_\varepsilon * \mu, \varphi_\varepsilon * \nu) + \mathcal{E}(\varphi_\varepsilon * \nu)$$

converges to $\mathcal{E}(\mu) - 2\,\mathcal{E}(\mu, \nu) + \mathcal{E}(\nu) = \mathcal{E}(\mu - \nu)$ as $\varepsilon \to 0$ and the proof of Theorem 1.1 is therefore complete up to the technical lemmas 2.1 and 2.2.

2.2. Proof of the technical lemmas 2.1 and 2.2.

Proof of Lemma 2.1. Let $R > 0$ be such that $D \subset B_R$. Given any function f such that $\|f\|_{\mathrm{Lip}} \leq 1$ and $f(0) = 0$, we claim one can find a function $\tilde{f}$ with support in B_{4R} such that $\tilde{f}|_D = f|_D$ and $\|\tilde{f}\|_{\mathrm{Lip}} \leq 1$. Indeed, define the function $\tilde{f}$ as follows: for any $r \geq 0$ and $\theta \in \mathbb{S}^{d-1} := \{x \in \mathbb{R}^d : |x| = 1\}$,

$$\tilde{f}(r\theta) = \begin{cases} f(r\theta), & \text{if } r \leq R, \\ f(R\theta), & \text{if } R \leq r \leq 2R, \\ f(R\theta)\frac{4R-r}{2R}, & \text{if } 2R \leq r \leq 4R, \\ 0, & \text{if } r \geq 4R. \end{cases}$$

By construction the only non trivial point to check is $\|\tilde f\|_{\mathrm{Lip}} \leq 1$, namely

$$|\tilde f(r\theta) - \tilde f(r'\theta')| \leq |r\theta - r'\theta'| \tag{2.9}$$

for every $r' \geq r \geq 0$ and $\theta, \theta' \in \mathbb{S}^{d-1}$. It is enough to prove (2.9) when both r and r' belong to the same interval in the definition of $\tilde f$. But the latter is obvious except when $2R \leq r \leq r' \leq 4R$. In this case, we use that $|f(R\theta)| \leq R$, because $f(0) = 0$ by assumption, in order to get

$$\begin{aligned}
|\tilde f(r\theta) - \tilde f(r'\theta')| &\leq |\tilde f(r\theta) - \tilde f(r'\theta)| + |\tilde f(r'\theta) - \tilde f(r'\theta')| \\
&\leq \frac{f(R\theta)}{2R}|r' - r| + |f(R\theta) - f(R\theta')|\frac{4R - r'}{2R} \\
&\leq \frac{1}{2}|r' - r| + R|\theta - \theta'| \\
&\leq \frac{1}{2}|r' - r| + \frac{1}{2}r|\theta - \theta'| \\
&\leq |r\theta - r'\theta'|.
\end{aligned}$$

We used in the last inequality that $|r\theta - r\theta'| \leq |r\theta - r'\theta'|$, which holds true since $r\theta'$ is the orthogonal projection of $r'\theta'$ onto B_r and $r\theta \in B_r$.

As a consequence, for any μ, ν probability measures supported in D, we obtain from (1.15),

$$\begin{aligned}
\sup_{\substack{f \in \mathcal{C}(B_{4R}) \\ \|f\|_{\mathrm{Lip}} \leq 1}} \int f(x)(\mu - \nu)(\mathrm{d}x) \leq \mathrm{W}_1(\mu, \nu) &= \sup_{\|f\|_{\mathrm{Lip}} \leq 1} \int f(x)(\mu - \nu)(\mathrm{d}x) \\
&= \sup_{\substack{\|f\|_{\mathrm{Lip}} \leq 1 \\ f(0) = 0}} \int f(x)(\mu - \nu)(\mathrm{d}x) \\
&= \sup_{\substack{\|f\|_{\mathrm{Lip}} \leq 1 \\ f(0) = 0}} \int \tilde f(x)(\mu - \nu)(\mathrm{d}x) \\
&\leq \sup_{\substack{\tilde f \in \mathcal{C}(B_{4R}) \\ \|\tilde f\|_{\mathrm{Lip}} \leq 1}} \int \tilde f(x)(\mu - \nu)(\mathrm{d}x).
\end{aligned}$$

Thus,

$$\mathrm{W}_1(\mu, \nu) = \sup_{\substack{f \in \mathcal{C}(B_{4R}) \\ \|f\|_{\mathrm{Lip}} \leq 1}} \int f(x)(\mu - \nu)(\mathrm{d}x), \tag{2.10}$$

which proves the lemma with $D_+ = B_{4R}$. □

Before going further, we show a simple but quite useful lemma, in the spirit of Newton's theorem, see e.g. [Lieb and Loss, 1997, Th. 9.7]. Recall that λ_R is the uniform probability measure on the ball B_R in $\mathbb{R}^d$.

Lemma 2.3 (Superharmonicity). *For every $x \in \mathbb{R}^d$,*

$$\iint g(x + u - v)\lambda_R(\mathrm{d}u)\lambda_R(\mathrm{d}v) \leq g(x).$$

Proof. Since g is superharmonic, we have

$$\int g(x + u)\lambda_R(\mathrm{d}u) \leq g(x), \qquad x \in \mathbb{R}^d.$$

As a consequence,

$$
\begin{aligned}
&\iint g(x+u-v)\lambda_R(\mathrm{d}u)\lambda_R(\mathrm{d}v) \\
&= \int g(x+u) - \underbrace{\Big(g(x+u) - \int g(x+u-v)\lambda_R(\mathrm{d}v)\Big)}_{\geq 0}\lambda_R(\mathrm{d}u) \\
&\leq \int g(x+u)\lambda_R(\mathrm{d}u) \leq g(x).
\end{aligned}
$$

□

Proof of Lemma 2.2. We set for convenience

$$
\begin{aligned}
g_\varepsilon(x-y) \quad &:= \quad \iint g(x-y+u-v)\varphi_\varepsilon(u)\varphi_\varepsilon(v)\mathrm{d}u\mathrm{d}v, \\
&= \quad \iint g(x-y+\varepsilon(u-v))\varphi(u)\varphi(v)\mathrm{d}u\mathrm{d}v,
\end{aligned}
$$

which is positive when $d \geq 3$. When $d = 2$, by dilation we can assume without loss of generality that μ and ν are supported in $B_{1/4}$, so that both $g(x-y)$ and $g_\varepsilon(x-y)$ are positive for every $(x, y) \in \mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$ and every $\varepsilon < 1/4$.

First, Fubini–Tonelli's theorem yields

$$
\begin{aligned}
\mathcal{E}(\varphi_\varepsilon * \mu, \varphi_\varepsilon * \nu) &= \iint \Big(\iint g(u-v)\varphi_\varepsilon(u-x)\varphi_\varepsilon(v-y)\mathrm{d}u\mathrm{d}v\Big)\mu(\mathrm{d}x)\nu(\mathrm{d}y) \\
&= \iint g_\varepsilon(x-y)\mu(\mathrm{d}x)\nu(\mathrm{d}y).
\end{aligned}
$$

For every $x, y \in \mathbb{R}^d$ satisfying $x \neq y$, the map $(u, v) \mapsto g(x-y+\varepsilon(u-v))$ is bounded on $B_1 \times B_1$ for every ε small enough and converges pointwise to $g(x-y)$ as $\varepsilon \to 0$. Moreover, since $\mathcal{E}(\mu, \nu) < \infty$ yields $\mu \otimes \nu(\{x = y\}) = 0$, by dominated convergence, $g_\varepsilon(x-y)$ converges $\mu \otimes \nu$-almost everywhere to $g(x-y)$. Next, we have

$$
\begin{aligned}
g_\varepsilon(x-y) &= \iint g(x-y+u-v)\varphi_\varepsilon(u)\varphi_\varepsilon(v)\mathrm{d}u\mathrm{d}v \\
&\leq \|\varphi\|_\infty^2 \,\mathrm{vol}(B_1)^2 \iint g(x-y+u-v)\lambda_\varepsilon(\mathrm{d}u)\lambda_\varepsilon(\mathrm{d}v), \qquad (2.11)
\end{aligned}
$$

where λ_ε is the uniform probability measure on B_ε. Thus, by Lemma 2.3, we obtain

$$
g_\varepsilon(x-y) \leq \|\varphi\|_\infty^2 \,\mathrm{vol}(B_1)^2\, g(x-y)
$$

and, since $g(x-y)$ is $\mu \otimes \nu$-integrable by assumption, the lemma follows by dominated convergence again. □

3. Proof of Theorem 1.2

We first prove Lemma 3.1, which is a weak version of Theorem 1.2 restricted to compactly supported measures. This is an easy corollary of Theorem 1.1 and the Euler–Lagrange equations which characterize the equilibrium measure. The latter state that, if we consider the Coulomb potential of the equilibrium measure $U^{\mu_V} = g * \mu_V$, then there exists a constant e_V such that

$$
2U^{\mu_V}(x) + V(x) \begin{cases} = e_V & \text{for q.e. } x \in \mathrm{supp}(\mu_V) \\ \geq e_V & \text{for q.e. } x \in \mathbb{R}^d, \end{cases} \qquad (3.1)
$$

where "q.e." stands for quasi-everywhere and means up to a set of null capacity. Conversely, μ_V is the unique probability measure satisfying (3.1). This characterization goes back at least to Frostman [1935] in the case where the potential V is constant on a bounded set and infinite outside. For the general case, see [Saff and Totik, 1997], [Chafaï et al., 2014], [Serfaty, 2015, Th. 2.1], and references therein.

Lemma 3.1. *If $D \subset \mathbb{R}^d$ is compact then for any $\mu \in \mathcal{P}(\mathbb{R}^d)$ supported in D,*

$$\mathrm{d_{BL}}(\mu, \mu_V)^2 \leq \mathrm{W}_1(\mu, \mu_V)^2 \leq C_{D \cup \mathrm{supp}(\mu_V)}\big(\mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V)\big),$$

with $C_{D\cup \mathrm{supp}(\mu_V)}$ as in Theorem 1.1.

Proof. Since the statement is obvious when $\mathcal{E}(\mu) = +\infty$, one can assume $\mathcal{E}(\mu) < \infty$. By integrating (3.1) over μ_V, we first obtain

$$e_V = 2\,\mathcal{E}(\mu_V) + \int V(x)\mu_V(\mathrm{d}x). \tag{3.2}$$

Integrating again (3.1) but over μ, we get together with (3.2),

$$2\,\mathcal{E}(\mu, \mu_V) + \int V(x)\mu(\mathrm{d}x) \geq 2\,\mathcal{E}(\mu_V) + \int V(x)\mu_V(\mathrm{d}x),$$

where we used the notation (2.8). This inequality provides

$$\mathcal{E}(\mu - \mu_V) = \mathcal{E}(\mu) - 2\,\mathcal{E}(\mu, \mu_V) + \mathcal{E}(\mu_V) \leq \mathcal{E}_V(\mu) - \mathcal{E}_V(\mu_V) \tag{3.3}$$

and the result follows from Theorem 1.1 and (1.16). □

We are now in position to prove Theorem 1.2. We follow closely the strategy of the proof of [Maïda and Maurel-Segala, 2014, Lem. 2].

Proof of Theorem 1.2. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$. We can assume without loss of generality that μ has unbounded support and finite Coulomb energy. The strategy consists in constructing a measure $\tilde{\mu}$ with compact support, with smaller Coulomb energy, together with an explicit control on the distance of interest between μ and $\tilde{\mu}$, and then to use Lemma 3.1. To do so, given any $R \geq 1$ consider the probability measures

$$\mu_{\mathrm{core}} := \frac{1}{1-\alpha}\mu|_{B_R}, \qquad \mu_{\mathrm{tail}} := \frac{1}{\alpha}\mu|_{\mathbb{R}^d \setminus B_R} \qquad \text{where} \quad \alpha := \mu(\mathbb{R}^d \setminus B_R),$$

so that $\mu = (1-\alpha)\mu_{\mathrm{core}} + \alpha\mu_{\mathrm{tail}}$. We then define the probability measure

$$\tilde{\mu} := (1-\alpha)\mu_{\mathrm{core}} + \alpha\sigma,$$

where σ stands for the uniform probability measure on the unit sphere $\mathbb{S}^{d-1}$. In particular, $\tilde{\mu}$ is supported in B_R.

First, we write

$$\begin{aligned}\mathcal{E}_V(\mu) - \mathcal{E}_V(\tilde{\mu}) = \alpha\Big\{ \int V(x)(\mu_{\mathrm{tail}} - \sigma)(\mathrm{d}x) + \alpha\,\mathcal{E}(\mu_{\mathrm{tail}}) - \alpha\,\mathcal{E}(\sigma) \\ + 2(1-\alpha)\,\mathcal{E}(\mu_{\mathrm{core}}, \mu_{\mathrm{tail}} - \sigma)\Big\},\end{aligned} \tag{3.4}$$

where we used the notation (2.8).

Note that, for any $x, y \in \mathbb{R}^d$,

$$g(x-y) \geq -(\log(1+|x|) + \log(1+|y|))\mathbf{1}_{d=2}, \tag{3.5}$$

so that

$$\mathcal{E}(\mu_{\mathrm{tail}}) \geq -\int 2\log(1+|x|)\mathbf{1}_{d=2}\,\mu_{\mathrm{tail}}(\mathrm{d}x). \tag{3.6}$$

Moreover, using alternatively,

$$g(x-y) \geq -\log\left(2\max(1+|x|,1+|y|)\right)\mathbf{1}_{d=2}, \tag{3.7}$$

we also obtain

$$\mathcal{E}(\mu_{\mathrm{core}},\mu_{\mathrm{tail}}) \geq -\int \big(\log(1+|x|)+\log 2\big)\mathbf{1}_{d=2}\,\mu_{\mathrm{tail}}(\mathrm{d}x). \tag{3.8}$$

Next, according to [Helms, 2009, Lem. 1.6.1], we have

$$U^\sigma(x) := g*\sigma(x) = \begin{cases} 0 & \text{if } |x|\leq 1 \text{ and } d=2,\\ 1 & \text{if } |x|\leq 1 \text{ and } d\geq 3,\\ g(x) & \text{if } |x|>1. \end{cases}$$

In particular, $U^\sigma(x)\leq 1$ for every $x\in\mathbb{R}^d$ and thus, using that $0\leq\alpha\leq 1$,

$$\alpha\,\mathcal{E}(\mu_{\mathrm{core}},\sigma) = \alpha\int U^\sigma(x)\mu_{\mathrm{core}}(\mathrm{d}x)\leq 1. \tag{3.9}$$

By combining (3.4)–(3.9), we obtain

$$\mathcal{E}_V(\mu)-\mathcal{E}_V(\tilde\mu) \geq \alpha\int\{V(x)-2\log(1+|x|)\mathbf{1}_{d=2}+C\}\mu_{\mathrm{tail}}(\mathrm{d}x) \tag{3.10}$$

for some $C>0$ which is independent on μ.

Next, using that

$$\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu) = \alpha \sup_{\substack{\|f\|_{\mathrm{Lip}}\leq 1\\ \|f\|_\infty\leq 1}}\int f(x)(\mu_{\mathrm{tail}}-\sigma)(\mathrm{d}x)\leq 2\alpha,$$

we obtain

$$\mathcal{E}_V(\mu)-\mathcal{E}_V(\tilde\mu)-\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu)^2 \geq \alpha\int\{V(x)-2\log(1+|x|)\mathbf{1}_{d=2}+C-4\}\mu_{\mathrm{tail}}(\mathrm{d}x).$$

Since V is admissible, the map

$$x\mapsto V(x)-2\log(1+|x|)\mathbf{1}_{d=2}+C-4$$

tends to $+\infty$ as $|x|$ becomes large, it follows that there exists $R>0$ large enough so that this map is positive on $\mathbb{R}^d\setminus B_R$.

Since by construction μ_{tail} is supported on $\mathbb{R}^d\setminus B_R$, this yields

$$\mathcal{E}_V(\mu)-\mathcal{E}_V(\tilde\mu)-\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu)^2\geq 0,$$

and in particular,

$$\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu)^2\leq\mathcal{E}_V(\mu)-\mathcal{E}_V(\mu_V)\quad\text{and}\quad\mathcal{E}_V(\tilde\mu)\leq\mathcal{E}_V(\mu).$$

This yields inequality (1.17) since, by the triangle inequality and Lemma 3.1, we have

$$\begin{aligned}\mathrm{d}_{\mathrm{BL}}(\mu,\mu_V)^2 &\leq (\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu)+\mathrm{d}_{\mathrm{BL}}(\tilde\mu,\mu_V))^2\\ &\leq 2\,\mathrm{d}_{\mathrm{BL}}(\mu,\tilde\mu)^2+2\,\mathrm{d}_{\mathrm{BL}}(\tilde\mu,\mu_V)^2\\ &\leq 2\big(\mathcal{E}_V(\mu)-\mathcal{E}_V(\mu_V)\big)+2C_{B_R\cup\mathrm{supp}(\mu_V)}\big(\mathcal{E}_V(\tilde\mu)-\mathcal{E}_V(\mu_V)\big)\\ &\leq \underbrace{2\left(1+C_{B_R\cup\mathrm{supp}(\mu_V)}\right)}_{C^V_{\mathrm{BL}}}\big(\mathcal{E}_V(\mu)-\mathcal{E}_V(\mu_V)\big).\end{aligned} \tag{3.11}$$

Note that since by construction R is independent on μ, so does C^V_{BL}.

We now turn to the proof of inequality (1.18). Under ($\mathbf{H}_{\mathrm{W}_1}$) we can assume that μ has a finite second moment

$$\int |x|^2\mu(\mathrm{d}x) < \infty$$

and in particular that $\mathrm{W}_1(\mu,\mu_V) < \infty$. Indeed, because of the growth condition ($\mathbf{H}_{\mathrm{W}_1}$), we have otherwise

$$\int V(x)\mu(\mathrm{d}x) = +\infty$$

so that the right hand side of the inequality to prove is infinite. Now, let $\gamma > 0$ be such that

$$\liminf_{|x|\to\infty} \frac{V(x)}{4\gamma|x|^2} > 1,$$

whose existence is ensured by ($\mathbf{H}_{\mathrm{W}_1}$). Using the representation (1.15), we have the upper bound,

$$\begin{aligned}
\mathrm{W}_1(\mu,\tilde\mu) &= \sup_{\|f\|_{\mathrm{Lip}}\le 1} \int f(x)(\mu-\tilde\mu)(\mathrm{d}x)\\
&= \alpha \sup_{\|f\|_{\mathrm{Lip}}\le 1} \int f(x)(\mu_{\mathrm{tail}}-\sigma)(\mathrm{d}x)\\
&= \alpha \sup_{\substack{\|f\|_{\mathrm{Lip}}\le 1\\ f(0)=0}} \int f(x)(\mu_{\mathrm{tail}}-\sigma)(\mathrm{d}x)\\
&\le \alpha \int |x|(\mu_{\mathrm{tail}}+\sigma)(\mathrm{d}x)\\
&\le 2\alpha \int |x|\mu_{\mathrm{tail}}(\mathrm{d}x)\\
&\le 2\alpha \left(\int |x|^2\mu_{\mathrm{tail}}(\mathrm{d}x)\right)^{1/2}.
\end{aligned} \tag{3.12}$$

Combined together with (3.10) this gives

$$\mathcal{E}_V(\mu)-\mathcal{E}_V(\tilde\mu)-\gamma\,\mathrm{W}_1(\mu,\tilde\mu)^2 \ge \alpha\int \{V(x)-4\gamma|x|^2-2\log(1+|x|)\mathbf{1}_{d=2}+C\}\mu_{\mathrm{tail}}(\mathrm{d}x).$$

Therefore, using the definition of γ and ($\mathbf{H}_{\mathrm{W}_1}$), this yields as before that there exists $R\ge 1$ independent on μ such that

$$\gamma\,\mathrm{W}_1(\mu,\tilde\mu)^2 \le \mathcal{E}_V(\mu)-\mathcal{E}_V(\mu_V) \quad\text{and}\quad \mathcal{E}_V(\tilde\mu)\le\mathcal{E}_V(\mu).$$

As in (3.11), this gives the upper bound,

$$\mathrm{W}_1(\mu,\mu_V)^2 \le \underbrace{2\left(\frac{1}{\gamma}+C_{B_R\cup\mathrm{supp}(\mu_V)}\right)}_{C^V_{\mathrm{W}_1}}(\mathcal{E}_V(\mu)-\mathcal{E}_V(\mu_V)),$$

and the proof of the theorem is complete. □

4. Proofs of Theorem 1.5 and Theorem 1.9

We start with the following lower bound on $Z^N_{V,\beta}$. Recall the definition (1.27) of the Boltzmann–Shannon entropy S.

Lemma 4.1 (Normalizing constant)**.** *Assume that V is admissible, finite on set of positive Lebesgue measure, and that $\mathrm{S}(\mu_V)$ is finite. Then for any $N \geq 2$,*

$$Z^N_{V,\beta} \geq \exp\left\{-N^2\frac{\beta}{2}\,\mathcal{E}_V(\mu_V) + N\left(\frac{\beta}{2}\,\mathcal{E}(\mu_V) + \mathrm{S}(\mu_V)\right)\right\}.$$

Note that it is compatible with the convergence $\lim_{N\to\infty}\frac{1}{N^2}\log Z^N_{V,\beta} = -\frac{\beta}{2}\,\mathcal{E}_V(\mu_V)$ from [Chafaï et al., 2014] and with the more refined asymptotic expansion from [Leblé and Serfaty, 2015, Cor. 1.5].

Proof of Lemma 4.1. We start by writing the energy (1.7) as

$$H_N(x_1,\dots,x_N) = \sum_{i\neq j}\left(g(x_i-x_j) + \frac{1}{2}V(x_i) + \frac{1}{2}V(x_j)\right) + \sum_{i=1}^N V(x_i). \tag{4.1}$$

Since the entropy $\mathrm{S}(\mu_V)$ is finite by assumption, the equilibrium measure μ_V has a density ρ_V. If we define the event $E^N_V := \{x\in(\mathbb{R}^d)^N : \prod_{i=1}^N \rho_V(x_i) > 0\}$, then

$$\begin{aligned}
&\log Z^N_{V,\beta}\\
&= \log\int \mathrm{e}^{-\frac{\beta}{2}\sum_{i\neq j}\left(g(x_i-x_j)+\frac12 V(x_i)+\frac12 V(x_j)\right)}\prod_{i=1}^N \mathrm{e}^{-\frac{\beta}{2}V(x_i)}\mathrm{d}x_i\\
&\geq \log\int_{E^N_V} \mathrm{e}^{-\frac{\beta}{2}\sum_{i\neq j}\left(g(x_i-x_j)+\frac12 V(x_i)+\frac12 V(x_j)\right)}\prod_{i=1}^N \mathrm{e}^{-\frac{\beta}{2}V(x_i)}\mathrm{d}x_i\\
&= \log\int_{E^N_V} \mathrm{e}^{-\frac{\beta}{2}\sum_{i\neq j}\left(g(x_i-x_j)+\frac12 V(x_i)+\frac12 V(x_j)\right)-\sum_{i=1}^N\left(\frac{\beta}{2}V(x_i)+\log\rho_V(x_i)\right)}\prod_{i=1}^N \mu_V(\mathrm{d}x_i)\\
&\overset{(*)}{\geq} -\frac{\beta}{2}\int_{E^N_V}\sum_{i\neq j}\left(g(x_i-x_j)+\frac{1}{2}V(x_i)+\frac{1}{2}V(x_j)\right)\prod_{i=1}^N\mu_V(\mathrm{d}x_i)\\
&\qquad - N\int\left(\frac{\beta}{2}V(x)+\log\rho_V(x)\right)\mu_V(\mathrm{d}x)\\
&= -N(N-1)\frac{\beta}{2}\,\mathcal{E}_V(\mu_V) - N\frac{\beta}{2}\int V(x)\mu_V(\mathrm{d}x) + N\,\mathrm{S}(\mu_V)\\
&= -N^2\frac{\beta}{2}\,\mathcal{E}_V(\mu_V) + N\frac{\beta}{2}\,\mathcal{E}(\mu_V) + N\,\mathrm{S}(\mu_V),
\end{aligned}$$

where $(*)$ comes from Jensen's inequality.

□

Next, we use that the energy (1.7) is a function of the empirical measure $\hat\mu_N$, and more precisely that one can express $\mathbb{P}^N_{V,\beta}$ in terms of the partition function $Z^N_{V,\beta}$ and the quantity

$$\mathcal{E}^{\neq}_V(\hat\mu_N) - \mathcal{E}_V(\mu_V),$$

where $\mathcal{E}^{\neq}_V(\hat\mu_N)$ is as in (1.12). The idea is to replace it by its continuous version $\mathcal{E}_V(\hat\mu_N) - \mathcal{E}_V(\mu_V)$, so as to compare it to $\mathrm{d}_{\mathrm{BL}}(\hat\mu_N,\mu_V)$ or $\mathrm{W}_1(\hat\mu_N,\mu_V)$ thanks to the Coulomb transport type inequalities of Theorem 1.2. However, this is meaningless because $\mathcal{E}_V(\hat\mu_N)$ is infinite since $\hat\mu_N$ is atomic, but one can circumvent this problem by approximating $\hat\mu_N$ with a regular measure $\hat\mu_N^{(\varepsilon)}$, as explained in the next lemma. The same strategy has been used by Maïda and Maurel-Segala [2014]. This regularization is also at the heart of the definition of the renormalized energy of Rougerie and Serfaty [2016]. See also [Chafaï et al., 2014, Prop. 2.8 item 3].

Lemma 4.2 (Regularization)**.** *For any $\varepsilon > 0$, let us define*

$$\hat{\mu}_N^{(\varepsilon)} := \hat{\mu}_N * \lambda_\varepsilon,$$

where we recall that λ_ε is the uniform probability measure on the ball B_ε. Then,

$$H_N(x_1,\dots,x_N) \geq N^2\,\mathcal{E}_V(\hat{\mu}_N^{(\varepsilon)}) - N\,\mathcal{E}(\lambda_\varepsilon) - N\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i). \tag{4.2}$$

Note the latter inequality is sharp. Indeed, since g is harmonic on $\mathbb{R}^d \setminus \{0\}$, the inequality (4.3) below is in fact an equality as soon as $|x_i - x_j| \geq \varepsilon$ for every $i \neq j$, and so does (4.2).

Proof. The superharmonicity of g yields, via Lemma 2.3,

$$\begin{aligned} \sum_{i\neq j} g(x_i - x_j) &\geq \sum_{i\neq j} \iint g(x_i - x_j + u - v)\lambda_\varepsilon(\mathrm{d}u)\lambda_\varepsilon(\mathrm{d}v) \\ &= \sum_{i\neq j} \mathcal{E}(\delta_{x_i} * \lambda_\varepsilon, \delta_{x_j} * \lambda_\varepsilon) \\ &= N^2\,\mathcal{E}(\hat{\mu}_N^{(\varepsilon)}) - N\,\mathcal{E}(\lambda_\varepsilon) \\ &= N^2\,\mathcal{E}_V(\hat{\mu}_N^{(\varepsilon)}) - N\,\mathcal{E}(\lambda_\varepsilon) - N\sum_{i=1}^N V*\lambda_\varepsilon(x_i), \end{aligned} \tag{4.3}$$

and (4.2) follows. ☐

We follow this common path to obtain both theorems 1.5 and 1.9 up to dealing with the remaining terms due to the regularization. Indeed, one has to proceed differently depending on the strength of the assumption we make on the Laplacian's potential ΔV. The proofs we provide use milder assumptions than in their statements, as explained in Remark 1.10.

Proofs of Theorem 1.5 and Theorem 1.9. Assume for now that V is admissible (to ensure the existence of μ_V), and finite on a set of positive Lebesgue measure (to ensure that $Z_{V,\beta}^N > 0$), and that the equilibrium measure μ_V has finite Boltzmann–Shannon entropy $\mathrm{S}(\mu_V)$ (to use Lemma 4.1). Let us fix $N \geq 2$, $r > 0$ and $\beta > 0$ satisfying ($\mathbf{H}_\beta$) (to ensure that $Z_{V,\beta}^N < \infty$).

Step 1: Preparation. We start by putting aside a small fraction of the energy to ensure integrability in the next steps. Namely, for any $\eta \in (0,1)$ and any Borel set $A \subset (\mathbb{R}^d)^N$, we write

$$\begin{aligned} \mathbb{P}_{V,\beta}^N(A) &= \frac{1}{Z_{V,\beta}^N}\int_A \mathrm{e}^{-\frac{\beta}{2}H_N(x_1,\dots,x_N)}\prod_{i=1}^N \mathrm{d}x_i \\ &= \frac{1}{Z_{V,\beta}^N}\int_A \mathrm{e}^{-\frac{\beta}{2}(1-\eta)H_N(x_1,\dots,x_N)}\mathrm{e}^{-\frac{\beta}{2}\eta H_N(x_1,\dots,x_N)}\prod_{i=1}^N \mathrm{d}x_i. \end{aligned}$$

Next, if we set

$$C_1 := -\frac{\beta}{2}\,\mathcal{E}(\mu_V) - \mathrm{S}(\mu_V), \tag{4.4}$$

then we deduce from Lemma 4.2 and then Lemma 4.1 that

$$\mathbb{P}^N_{V,\beta}(A) \leq \frac{1}{Z^N_{V,\beta}} \int_A \mathrm{e}^{-\frac{\beta}{2}N(1-\eta)\left(N\,\mathcal{E}_V(\hat\mu_N^{(\varepsilon)})-\mathcal{E}(\lambda_\varepsilon)-\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i)\right)}$$
$$\times\, \mathrm{e}^{-\frac{\beta}{2}\eta H_N(x_1,\ldots,x_N)} \prod_{i=1}^N \mathrm{d}x_i \tag{4.5}$$
$$\leq \mathrm{e}^{NC_1+\frac{\beta}{2}N\,\mathcal{E}(\lambda_\varepsilon)+\frac{\beta}{2}N^2\eta\,\mathcal{E}_V(\mu_V)}\mathrm{e}^{-\frac{\beta}{2}N^2(1-\eta)\inf_A(\mathcal{E}_V(\hat\mu_N^{(\varepsilon)})-\mathcal{E}_V(\mu_V))} \tag{4.6}$$
$$\times \int \mathrm{e}^{\frac{\beta}{2}\left(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i)-\eta H_N(x_1,\ldots,x_N)\right)} \prod_{i=1}^N \mathrm{d}x_i.$$

In the second inequality we also used that $\mathcal{E}(\lambda_\varepsilon) > 0$ as soon as $0 < \varepsilon < 1$. Indeed, by a change of variables we have

$$\mathcal{E}(\lambda_\varepsilon) = \iint g(x-y)\lambda_\varepsilon(\mathrm{d}x)\lambda_\varepsilon(\mathrm{d}y) = \begin{cases} g(\varepsilon)+\mathcal{E}(\lambda_1) & \text{if } d=2,\\ g(\varepsilon)\ \mathcal{E}(\lambda_1) & \text{if } d\geq 3,\end{cases} \tag{4.7}$$

and one can compute explicitly that

$$\mathcal{E}(\lambda_1) = \begin{cases} \frac{1}{4} & \text{if } d=2,\\ \frac{2d}{d+2} & \text{if } d\geq 3.\end{cases} \tag{4.8}$$

Step 2: The remaining integral. We now provide an upper bound on the integral

$$\int \mathrm{e}^{\frac{\beta}{2}\left(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i)-\eta H_N(x_1,\ldots,x_N)\right)} \prod_{i=1}^N \mathrm{d}x_i.$$

Having in mind Remark 1.10, notice that if $V = \tilde V + h$ with h superharmonic then we have, for every $x \in \mathbb{R}^d$,

$$(V*\lambda_\varepsilon - V)(x) \leq (\tilde V * \lambda_\varepsilon - \tilde V)(x).$$

Thus one can assume without loss of generality that $h = 0$. From now, assume further that $\tilde V = V$ is twice differentiable. Using that

$$(V*\lambda_\varepsilon - V)(x) = \int (V(x-\varepsilon u) - V(x))\lambda_1(\mathrm{d}u),$$

and noticing that by symmetry,

$$\int \langle \nabla V(x), u\rangle \lambda_1(\mathrm{d}u) = 0,$$

we obtain by a Taylor expansion that, for any $x \in \mathbb{R}^d$,

$$(V*\lambda_\varepsilon - V)(x) \leq \frac{\varepsilon^2}{2} \sup_{\substack{y\in\mathbb{R}^d\\ |y-x|\leq\varepsilon}} \int \langle \mathrm{Hess}(V)(y)u,u\rangle \lambda_1(\mathrm{d}u), \tag{4.9}$$

where $\mathrm{Hess}(V)(y)$ is the Hessian matrix of the function V at point y. Since the covariance matrix of λ_1 is a multiple of the identity,

$$\begin{aligned}\int \langle \mathrm{Hess}(V)(y)u, u\rangle \lambda_1(\mathrm{d}u) &= \sum_{i,j=1}^{d} \mathrm{Hess}(V)(y)_{i,j} \int u_i u_j \, \lambda_1(\mathrm{d}u)\\ &= \mathrm{Tr}\,\mathrm{Hess}(V)(y) \frac{1}{d} \int |u|^2 \, \lambda_1(\mathrm{d}u)\\ &= \mathrm{Tr}\,\mathrm{Hess}(V)(y) \int_0^1 r^{d+1}\, \mathrm{d}r\\ &= \frac{1}{d+2} \Delta V(y).\end{aligned}$$

Therefore, with $(\Delta V)_+(y) := \max(\Delta V(y), 0)$, we have

$$(V * \lambda_\varepsilon - V)(x) \le \frac{\varepsilon^2}{2(d+2)} \sup_{\substack{y\in\mathbb{R}^d\\ |y-x|\le\varepsilon}} (\Delta V)_+(y). \tag{4.10}$$

Next, we use that $g(x-y) \ge -\frac{1}{2}\log(1+|x|^2) - \frac{1}{2}\log(1+|y|^2)$ when $d=2$, and that $g \ge 0$ when $d \ge 3$, in order to get

$$H_N(x_1, \dots, x_N) \ge N \sum_{i=1}^{N} (V(x_i) - \mathbf{1}_{d=2} \log(1+|x_i|^2)).$$

As a consequence, we obtain with (4.10) that, for any $0 < \epsilon < 1$,

$$\begin{aligned}&\int \mathrm{e}^{\frac{\beta}{2}\left(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i) - \eta H_N(x_1,\dots,x_N)\right)} \prod_{i=1}^N \mathrm{d}x_i\\ &\quad\le \left(\int \mathrm{e}^{-\frac{\beta}{2}N\left(\eta(V(x)-\mathbf{1}_{d=2}\log(1+|x|^2)) - \varepsilon^2 \frac{1}{2(d+2)} \sup_{|y-x|<1}(\Delta V)_+(y)\right)} \mathrm{d}x\right)^N. \end{aligned}\tag{4.11}$$

In the next steps, we will use the Coulomb transport inequalities to bound the term $\inf_A(\mathcal{E}_V(\hat{\mu}_N^{(\varepsilon)}) - \mathcal{E}_V(\mu_V))$ for sets A of interest depending on which growth assumption we make on V. We will moreover bound the remaining integral (4.11) using extra assumptions on ΔV. In the next step, we assume ΔV is bounded above and provide a proof for Theorem 1.9. In the final step, we only assume ΔV grows not faster than V and complete the proof of Theorem 1.5.

Step 3: Proof of Theorem 1.9. Assume in this step there exists $D > 0$ such that

$$\sup_{x\in\mathbb{R}^d} \Delta V(x) \le D. \tag{4.12}$$

We then set

$$\eta := N^{-1}, \qquad \varepsilon := N^{-1/d}. \tag{4.13}$$

and

$$C_2 := \log \int \mathrm{e}^{-\frac{\beta}{2}(V(x) - \mathbf{1}_{d=2}\log(1+|x|^2))} \mathrm{d}x,$$

which is finite by assumption ($\mathbf{H}_\beta$). We obtain from (4.11),

$$\int \mathrm{e}^{\frac{\beta}{2}\left(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i) - \eta H_N(x_1,\dots,x_N)\right)} \prod_{i=1}^N \mathrm{d}x_i \le \mathrm{e}^{NC_2 + \frac{\beta}{2} N^{2-2/d} \frac{D}{2(d+2)}}. \tag{4.14}$$

If we set

$$A := \Big\{ \mathrm{d_{BL}}(\hat\mu_N,\mu_V) \geq r \Big\}, \tag{4.15}$$

then by applying Theorem 1.2 to $\hat\mu_N^{(\varepsilon)}$ we obtain,

$$\inf_A \big(\mathcal{E}_V(\hat\mu_N^{(\varepsilon)}) - \mathcal{E}_V(\mu_V) \big) \geq \frac{1}{C_{\mathrm{BL}}^V} \inf_A \mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\mu_V)^2 \geq \frac{1}{C_{\mathrm{BL}}^V}\Big(\frac{r^2}{2} - \varepsilon^2\Big), \tag{4.16}$$

where we used for the last inequality that

$$\begin{aligned}\frac12 \mathrm{d_{BL}}(\hat\mu_N,\mu_V)^2 &\leq \frac12 \big(\mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\mu_V) + \mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\hat\mu_N)\big)^2 \\ &\leq \mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\mu_V)^2 + \mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\hat\mu_N)^2 \\ &\leq \mathrm{d_{BL}}(\hat\mu_N^{(\varepsilon)},\mu_V)^2 + \varepsilon^2.\end{aligned}$$

Setting $c(\beta) := C_1 + C_2 + \frac{\beta}{2}\mathcal{E}_V(\mu_V)$ and noticing it is the same constant $c(\beta)$ as in Theorem 1.9, we obtain by combining (4.5), (4.7)–(4.8), (4.14), and (4.16),

$$\mathbb{P}^N_{V,\beta}(A) \leq \mathrm{e}^{-\frac{\beta}{4C_{\mathrm{BL}}^V}N(N-1)r^2 + \mathbf{1}_{d=2}(\frac{\beta}{4}N\log N) + \frac{\beta}{2}N^{2-2/d}[\frac{1}{C_{\mathrm{BL}}^V} + \frac{D}{2(d+2)} + \mathcal{E}(\lambda_1)] + Nc(\beta)}. \tag{4.17}$$

Since $N-1 \geq N/2$ for every $N \geq 2$, this gives the desired constants a and b.

If one assumes the stronger growth assumption ($\mathbf{H}_{\mathrm{W}_1}$), then the exact same line of arguments with $A := \{\mathrm{W}_1(\hat\mu_N,\mu_V) \geq r\}$ shows inequality (4.17) holds true after replacing the constant C_{BL}^V by $C_{\mathrm{W}_1}^V$, which completes the proof of Theorem 1.9.

Step 5: Proof of Theorem 1.5. In this last part, we assume that

$$\limsup_{|x|\to\infty} \Bigg(\frac{1}{V(x)} \sup_{\substack{y\in\mathbb{R}^d \\ |y-x|<1}} \Delta V(y) \Bigg) < 2(d+2) \tag{4.18}$$

in place of (4.12), and set instead,

$$\eta := N^{-2/d}, \qquad \varepsilon := N^{-1/d}, \tag{4.19}$$

so that (4.11) now reads

$$\begin{aligned}&\int \mathrm{e}^{\frac{\beta}{2}\big(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i) - \eta H_N(x_1,\ldots,x_N)\big)} \prod_{i=1}^N \mathrm{d}x_i \\ &\quad\leq \Bigg(\int \mathrm{e}^{-\frac{\beta}{2}N^{1-2/d}\big(V(x) - \mathbf{1}_{d=2}\log(1+|x|^2) - \frac{1}{2(d+2)}\sup_{|y-x|<1}(\Delta V)_+(y)\big)} \mathrm{d}x \Bigg)^N.\end{aligned} \tag{4.20}$$

Combining V being admissible and (4.18) yields constants $u > 0$ and $v \in \mathbb{R}$ such that, for every $x \in \mathbb{R}^d$,

$$V(x) - \mathbf{1}_{d=2}\log(1+|x|^2) - \frac{1}{2(d+2)}\sup_{|y-x|<1}(\Delta V)_+(y) \geq 2uV(x) - v.$$

This provides with (4.20),

$$\begin{aligned}\int \mathrm{e}^{\frac{\beta}{2}\big(N(1-\eta)\sum_{i=1}^N (V*\lambda_\varepsilon - V)(x_i) - \eta H_N(x_1,\ldots,x_N)\big)} \prod_{i=1}^N \mathrm{d}x_i &\leq \mathrm{e}^{\frac{\beta}{2}vN^{2-2/d}} \Big(\int \mathrm{e}^{-\beta u V(x)}\mathrm{d}x \Big)^N \\ &\leq \mathrm{e}^{\frac{\beta}{2}vN^{2-2/d} + Nw(\beta)},\end{aligned}$$

where we set

$$w(\beta) := \log \int \mathrm{e}^{-\beta u V(x)} \mathrm{d}x. \tag{4.21}$$

Taking $A := \{\mathrm{d_{BL}}(\hat{\mu}_N, \mu_V) \geq r\}$, this gives with (4.5), (4.7), (4.16), (4.19),

$$\mathbb{P}^N_{V,\beta}(A) \leq \mathrm{e}^{-\frac{\beta N^2 r^2}{4C^V_{\mathrm{BL}}}(1-N^{-\frac{2}{d}}) + \mathbf{1}_{d=2}(\frac{\beta}{4}N\log N) + \frac{\beta}{2}N^{2-\frac{2}{d}}\left[\mathcal{E}_V(\mu_V)+\mathcal{E}(\lambda_1)+\frac{1}{C^V_{\mathrm{BL}}}+v\right] + N(C_1+w(\beta))}.$$

Since $1 - N^{-2/d} \geq 1/2$ for $N \geq 2$, we get inequality (1.20) of Theorem 1.5 with

$$\begin{aligned} a &:= \frac{1}{8C^V_{\mathrm{BL}}}, \\ b &:= \frac{1}{2}\Big(\mathcal{E}_V(\mu_V) + \mathcal{E}(\lambda_1) + \frac{1}{C^V_{\mathrm{BL}}} + v\Big), \\ c(\beta) &:= -\frac{\beta}{2}\mathcal{E}(\mu_V) - S(\mu_V) + w(\beta). \end{aligned}$$

Taking $A := \{\mathrm{W}_1(\hat{\mu}_N, \mu_V) \geq r\}$, we obtain the same inequality for the W_1 metric after replacing C^V_{BL} by $C^V_{\mathrm{W}_1}$ in the definitions of the constants a and b.

Finally, observe that if one assumes $\liminf_{|x|\to\infty} V(x)/|x|^\kappa > 0$ for some $\kappa > 0$, then it is easy to derive the behavior (1.21) of the function $c(\beta)$ from (4.21). □

Proof of Corollary 1.8. Let us give the proof in the bounded Lipschitz case. If $\beta > 0$ is fixed, then Theorem 1.5 yields $u, v > 0$ which only depend on V, β such that, for every $N \geq 2$ and $r > 0$,

$$\mathbb{P}^N_{V,\beta}\Big(\mathrm{d_{BL}}(\hat{\mu}_N, \mu_V) \geq r\Big) \leq \mathrm{e}^{-uN^2r^2 + vNg(N^{-1/d})}. \tag{4.22}$$

Now, by a change of variables, we have for any $\alpha > 0$,

$$\mathbb{P}^N_{V,\beta}\Big(\mathrm{d_{BL}}\,(\tau^{N^s}_{x_0}\hat{\mu}_N, \tau^{N^s}_{x_0}\mu_V) \geq \alpha\Big) \leq \mathbb{P}^N_{V,\beta}\Big(\mathrm{d_{BL}}\,(\hat{\mu}_N, \mu_V) \geq \alpha N^{-s}\Big).$$

The corollary follows by taking

$$\alpha = \begin{cases} CN^{s-1/2}(\log N)^{1/2} & \text{when } d = 2, \\ CN^{s-1/d} & \text{when } d \geq 3, \end{cases}$$

with $C > 0$ large enough so that $c := uC^2 - v > 0$. The proof of the W_1 case is the same. □

5. Proof of Theorem 1.12

We prove the exponential tightness for $\max_{1\leq i\leq N}|x_i|$ formulated in Theorem 1.12, in a similar way as in the proof of [Borot and Guionnet, 2013, Prop. 2.1].

Proof of Theorem 1.12. Since the law $\mathbb{P}^N_{V,\beta}$ is exchangeable, the union bound gives

$$\mathbb{P}^N_{V,\beta}(\max_{1\leq i\leq N}|x_i| \geq r) \leq N\,\mathbb{P}^N_{V,\beta}(|x_N| \geq r). \tag{5.1}$$

Isolating x_N in the joint law, we get for any Borel set $B \subset \mathbb{R}^d$,

$$\mathbb{P}^N_{V,\beta}(x_N \in B) = \frac{Z^{N-1}_{V,\beta}}{Z^N_{V,\beta}} \int_B \mathrm{e}^{-\frac{\beta}{2}NV(x_N)} \int \mathrm{e}^{-\frac{\beta}{2}\sum_{i=1}^{N-1}(g(x_i-x_N)+V(x_i))} \times \mathbb{P}^{N-1}_{V,\beta}(\mathrm{d}x_1,\ldots,\mathrm{d}x_{N-1})\mathrm{d}x_N. \tag{5.2}$$

Since V is admissible, there exists $C \in \mathbb{R}$ such that,

$$g(x-y) + V(x) + \frac{1}{2}V(y) \geq C, \quad \forall x, y \in \mathbb{R}^d.$$

As a consequence, we obtain

$$\begin{aligned}\mathbb{P}^N_{V,\beta}(|x_N| \geq r) &= \frac{Z^{N-1}_{V,\beta}}{Z^N_{V,\beta}} \mathrm{e}^{-\frac{\beta}{2}(N-1)C} \int_{\{|x_N|>r\}} \mathrm{e}^{-\frac{\beta}{4}(N+1)V(x_N)} \mathrm{d}x_N \\ &\leq \frac{Z^{N-1}_{V,\beta}}{Z^N_{V,\beta}} \mathrm{e}^{-\frac{\beta}{2}(N-1)C - \frac{\beta}{4}NV_*(r)} \int \mathrm{e}^{-\frac{\beta}{4}V(x)} \mathrm{d}x, \end{aligned} \tag{5.3}$$

where we recall that $V_*(r) = \min_{|x|\geq r} V(x)$. We now prove that

$$\limsup_{N\to\infty} \frac{1}{N} \log \frac{Z^{N-1}_{V,\beta}}{Z^N_{V,\beta}} < \infty, \tag{5.4}$$

which yields the theorem thanks to (5.1) and (5.3) since $V_*(r) \to +\infty$ as $r \to \infty$.

To prove (5.4), let $L > 0$ large enough so that $\{V < +\infty\} \cap B_L$ has positive Lebesgue measure and use that $\mathbb{P}^N_{V,\beta}(|x_N| \leq L) \leq 1$ together with (5.2) to get

$$\begin{aligned}&\frac{Z^N_{V,\beta}}{Z^{N-1}_{V,\beta}} \\ &\qquad \geq Y_{V,\beta} \int \mathrm{e}^{-(N-1)\frac{\beta}{2}V(x_N) - \frac{\beta}{2}\sum_{i=1}^{N-1} V(x_i) + g(x_i - x_N)} \, \mathbb{P}^{N-1}_{V,\beta} \otimes \eta_{V,\beta}(\mathrm{d}x_1, \ldots, \mathrm{d}x_N),\end{aligned}$$

where $\eta_{V,\beta}$ is the probability measure on $\mathbb{R}^d$ with density $\frac{1}{Y_{V,\beta}} \mathrm{e}^{-\frac{\beta}{2}V} \mathbf{1}_{B_L}$ and $Y_{V,\beta} > 0$ a normalizing constant. Applying Jensen's inequality with respect to the probability measure $\mathbb{P}^{N-1}_{V,\beta} \otimes \eta_{V,\beta}$ then provides

$$\begin{aligned}&\log \frac{Z^N_{V,\beta}}{Z^{N-1}_{V,\beta}} \geq \log Y_{V,\beta} \\ &- \int \Big((N-1)\frac{\beta}{2}V(x_N) + \frac{\beta}{2}\sum_{i=1}^{N-1} V(x_i) + g(x_i - x_N) \Big) \mathbb{P}^{N-1}_{V,\beta} \otimes \eta_{V,\beta}(\mathrm{d}x_1, \ldots, \mathrm{d}x_N) \\ &= \log Y_{V,\beta} - (N-1)\frac{\beta}{2} \int V(x) \, \eta_{V,\beta}(\mathrm{d}x) \\ &- \frac{\beta}{2} \int \left(\sum_{i=1}^{N-1} V(x_i) + g * \eta_{V,\beta}(x_i) \right) \mathbb{P}^{N-1}_{V,\beta}(\mathrm{d}x_1, \ldots, \mathrm{d}x_{N-1}).\end{aligned}$$

Since for an admissible V we have $\int V \, \mathrm{d}\eta_{V,\beta} = \int \mathbf{1}_{\{V<+\infty\}} V \, \mathrm{d}\eta_{V,\beta} < +\infty$ and that $g * \eta_{V,\beta}$ is bounded from above, it is enough to show that

$$\sup_N \int \Big(\frac{1}{N} \sum_{i=1}^N V(x_i) \Big) \mathbb{P}^N_{V,\beta}(\mathrm{d}x_1, \ldots, \mathrm{d}x_N) < \infty \tag{5.5}$$

in order to prove (5.4). To do so, recall there exists by assumption $\varepsilon > 0$ such that $V(x) - (2+\varepsilon) \log |x| \mathbf{1}_{d=2} \to \infty$ as $|x| \to \infty$. In particular, if we set $q := 2/(2+\varepsilon) > 0$, then $g(x-y)+\frac{q}{2}V(x)+\frac{q}{2}V(y)\geq B$ for some $B\in\mathbb{R}$, and thus

$$\begin{aligned}Z^N_{qV,\beta}&=\int \mathrm{e}^{-\frac{\beta}{2}\sum_{i\neq j}(g(x_i-x_j)+\frac{q}{2}V(x_i)+\frac{q}{2}V(x_j))}\prod_{i=1}^N \mathrm{e}^{-\frac{\beta q}{2}V(x_i)}\mathrm{d}x_i\\&\leq \mathrm{e}^{-\frac{\beta}{2}N(N-1)B}\left(\int \mathrm{e}^{-\frac{\beta q}{2}V(x)}\mathrm{d}x\right)^N.\end{aligned}$$

Since Markov's inequality yields for any $R>0$,

$$\mathbb{P}^N_{V,\beta}\left(\frac{1}{N}\sum_{i=1}^N V(x_i)>R\right)\leq \mathrm{e}^{-\frac{\beta q}{2}N^2R}\,\frac{Z^N_{qV,\beta}}{Z^N_{V,\beta}},$$

together with Lemma 4.1 this provides some constant $R_0\in\mathbb{R}$ such that

$$\mathbb{P}^N_{V,\beta}\left(\frac{1}{N}\sum_{i=1}^N V(x_i)>R\right)\leq \mathrm{e}^{-\frac{\beta}{4}N^2(R+R_0)}.$$

This yields (5.5) since V is bounded from below, and the proof of Theorem 1.12 is therefore complete. □

References

R. Allez, J.-P. Bouchaud, and A. Guionnet. Invariant Beta Ensembles and the Gauss-Wigner Crossover. *Physical Review Letters*, 109(9):094102, Aug. 2012. URL http://dx.doi.org/10.1103/PhysRevLett.109.094102. 12

Y. Ameur, H. Hedenmalm, and N. Makarov. Fluctuations of eigenvalues of random normal matrices. *Duke Math. J.*, 159(1):31–81, 2011. ISSN 0012-7094. URL http://dx.doi.org/10.1215/00127094-1384782. 4

Y. Ameur, H. Hedenmalm, and N. Makarov. Random normal matrices and Ward identities. *Ann. Probab.*, 43(3):1157–1201, 2015. ISSN 0091-1798. URL http://dx.doi.org/10.1214/13-AOP885. 4

R. Bauerschmidt, P. Bourgade, M. Nikula, and H.-T. Yau. Local density for two-dimensional one-component plasma. preprint, 2015. URL http://arxiv.org/1510.02074. 4

R. Bauerschmidt, P. Bourgade, M. Nikula, and H.-T. Yau. The two-dimensional Coulomb plasma: quasi-free approximation and central limit theorem. preprint, 2016. URL http://arxiv.org/abs/1609.08582. 4

G. Ben Arous and A. Guionnet. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997. ISSN 0178-8051. URL http://dx.doi.org/10.1007/s004400050119. 11

P. Biane and D. V. Voiculescu. A free probability analogue of the Wasserstein metric on the trace-state space. *Geom. Funct. Anal.*, 11(6):1125–1138, 2001. URL http://dx.doi.org/10.1007/s00039-001-8226-4. 11

S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999. URL http://dx.doi.org/10.1006/jfan.1998.3326. 11

T. Bodineau and A. Guionnet. About the stationary states of vortex systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 35(2):205–237, 1999. ISSN 0246-0203. doi: 10.1016/S0246-0203(99)80011-9. URL https://dx-doi-org.proxy.bu.dauphine.fr/10.1016/S0246-0203(99)80011-9. 12

C. Bordenave and D. Chafaï. Around the circular law. *Probab. Surv.*, 9:1–89, 2012. URL http://dx.doi.org/10.1214/11-PS183. 9, 10

G. Borot and A. Guionnet. Asymptotic Expansion of β Matrix Models in the One-cut Regime. *Comm. Math. Phys.*, 317:447–483, Jan. 2013. URL http://dx.doi.org/10.1007/s00220-012-1619-4. 25

J. Breuer and M. Duits. The Nevai condition and a local law of large numbers for orthogonal polynomial ensembles. *Adv. Math.*, 265:441–484, 2014. ISSN 0001-8708. doi: 10.1016/j.aim.2014.07.026. URL http://dx.doi.org/10.1016/j.aim.2014.07.026. 4, 10

D. Chafaï. From Boltzmann to random matrices and beyond. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):641–689, 2015. URL http://dx.doi.org/10.5802/afst.1459. 11

D. Chafaï, N. Gozlan, and P.-A. Zitt. First-order global asymptotics for confined particles with singular pair repulsion. *Ann. Appl. Probab.*, 24(6):2371–2413, 2014. URL http://dx.doi.org/10.1214/13-AAP980. 3, 11, 17, 20

R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. URL http://dx.doi.org/10.1017/CBO9780511755347. Revised reprint of the 1989 original. 5

P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010. URL http://dx.doi.org/10.1515/9781400835416. 3, 4, 9, 12

O. Frostman. *Potentiel d'Équilibre et Capacité des Ensembles*. PhD thesis, Faculté des sciences de Lund, 1935. 17

J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Mathematical Phys.*, 6:440–449, 1965. 9

N. Gozlan. A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.*, 37(6):2480–2498, 2009. URL http://dx.doi.org/10.1214/09-AOP470. 11

N. Gozlan and C. Léonard. Transport inequalities. A survey. *Markov Process. Related Fields*, 16(4):635–736, 2010. 11

A. Hardy. A note on large deviations for 2D Coulomb gas with weakly confining potential. *Electron. Commun. Probab.*, 17:no. 19, 12, 2012. URL http://dx.doi.org/10.1214/ECP.v17-1818. 12

L. L. Helms. *Potential Theory*. Universitext. Springer, London., 2009. 2, 18

F. Hiai, D. Petz, and Y. Ueda. Free transportation cost inequalities via random matrix approximation. *Probab. Theory Related Fields*, 130(2):199–221, 2004. URL http://dx.doi.org/10.1007/s00440-004-0351-1. 11

K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998. URL http://dx.doi.org/10.1215/S0012-7094-98-09108-6. 4

B. A. Khoruzhenko and H.-J. Sommers. Non-Hermitian ensembles. In *The Oxford handbook of random matrix theory*, pages 376–397. Oxford Univ. Press, Oxford, 2011. 9

N. S. Landkof. *Foundations of modern potential theory*. Springer-Verlag, 1972. Translated from the Russian by A. P. Doohovskoy, Die Grundlehren der mathematischen Wissenschaften, Band 180. 2, 6

T. Leblé. Local microscopic behavior for 2D Coulomb gases. preprint, 2015. URL http://arxiv.org/abs/1510.01506. 4

T. Leblé and S. Serfaty. Large Deviation Principle for Empirical Fields of Log and Riesz Gases. preprint, 2015. URL http://arxiv.org/abs/1502.02970. 4, 20

T. Leblé and S. Serfaty. Fluctuations of Two-Dimensional Coulomb Gases. preprint, 2016. URL http://arxiv.org/abs/1609.08088. 4

M. Ledoux and I. Popescu. Mass transportation proofs of free functional inequalities, and free Poincaré inequalities. *J. Funct. Anal.*, 257(4):1175–1221, 2009. URL http://dx.doi.org/10.1016/j.jfa.2009.03.011. 11

E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. AMS, Providence, 1997. 1, 2, 13, 15

A. López García. Greedy energy points with external fields. In *Recent trends in orthogonal polynomials and approximation theory*, volume 507 of *Contemp. Math.*, pages 189–207. Amer. Math. Soc., Providence, RI, 2010. URL http://dx.doi.org/10.1090/conm/507/09960. 3

M. Maïda and É. Maurel-Segala. Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices. *Probab. Theory Related Fields*, 159(1-2):329–356, 2014. URL http://dx.doi.org/10.1007/s00440-013-0508-x. 4, 6, 7, 11, 17, 20

K. Marton. A simple proof of the blowing-up lemma. *IEEE Trans. Inform. Theory*, 32(3): 445–446, 1986. URL http://dx.doi.org/10.1109/TIT.1986.1057176. 11

E. S. Meckes and M. W. Meckes. Rates of convergence for empirical spectral measures: a soft approach. preprint, 2016. URL http://arxiv.org/abs/1601.03720. 10

M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, third edition, 2004. ISBN 0-12-088409-7. 9

R. Pemantle and Y. Peres. Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combin. Probab. Comput.*, 23(1): 140–160, 2014. ISSN 0963-5483. doi: 10.1017/S0963548313000345. URL http://dx.doi.org/10.1017/S0963548313000345. 4, 10

M. Petrache and S. Serfaty. Next-order asymptotics and renormalized energy for Riesz interactions. *Journal of the Institute of Mathematics of Jussieu*, pages pp. 1–69, 2015. URL http://dx.doi.org/10.1017/S1474748015000201. 4

I. Popescu. Local functional inequalities in one-dimensional free probability. *J. Funct. Anal.*, 264(6):1456–1479, 2013. URL http://dx.doi.org/10.1016/j.jfa.2013.01.008. 6, 11

S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I*. Probability and its Applications (New York). Springer-Verlag, 1998. ISBN 0-387-98350-3. Theory. 5

B. Rider and B. Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN*, 33(2):Art. ID rnm006, 2007. ISSN 1073-7928. URL http://dx.doi.org/10.1093/imrn/rnm006. 4

N. Rougerie and S. Serfaty. Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.*, 69(3):519–605, 2016. URL http://dx.doi.org/10.1002/cpa.21570. 4, 20

E. B. Saff and V. Totik. *Logarithmic potentials with external fields*, volume 316 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. URL http://dx.doi.org/10.1007/978-3-662-03329-6. Appendix B by Thomas Bloom. 2, 17

E. Sandier and S. Serfaty. 2D Coulomb gases and the renormalized energy. *Ann. Probab.*, 43(4):2026–2083, 2015. URL http://dx.doi.org/10.1214/14-AOP927. 4

S. Serfaty. *Coulomb gases and Ginzburg-Landau vortices.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2015. URL http://dx.doi.org/10.4171/152. 3, 17

C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003. ISBN 0-8218-3312-X. URL http://dx.doi.org/10.1007/b12016. 5, 11

(Djalil Chafaï) Université Paris-Dauphine, PSL Research Unversity, France
E-mail address: djalil@chafai.net

(Adrien Hardy) Université de Lille 1, France
E-mail address: adrien.hardy@math.univ-lille1.fr

(Mylène Maïda) Université de Lille 1, France
E-mail address: mylene.maida@math.univ-lille1.fr