COULOMB GASES UNDER CONSTRAINT:
SOME THEORETICAL AND NUMERICAL RESULTS

DJALIL CHAFAÏ, GRÉGOIRE FERRÉ, AND GABRIEL STOLTZ

Abstract. We consider Coulomb gas models for which the empirical measure typically concentrates, when the number of particles becomes large, on an equilibrium measure minimizing an electrostatic energy. We study the behavior when the gas is conditioned on a rare event. We first show that the special case of quadratic confinement and linear constraint is exactly solvable due to a remarkable factorization, and that the conditioning has then the simple effect of shifting the cloud of particles without deformation. To address more general cases, we perform a theoretical asymptotic analysis relying on a large deviations technique known as the Gibbs conditioning principle. The technical part amounts to establishing that the conditioning ensemble is an l-continuity set of the energy. This leads to characterizing the conditioned equilibrium measure as the solution of a modified variational problem. For simplicity, we focus on linear statistics and on quadratic statistics constraints. Finally, we numerically illustrate our predictions and explore cases in which no explicit solution is known. For this, we use a Generalized Hybrid Monte Carlo algorithm for sampling from the conditioned distribution for a finite but large system.

Contents

1. Introduction 1
2. From conditioning to shifting: quadratic confinement with linear constraint 5
3. A general conditioning framework 6
3.1. Gibbs conditioning 6
3.2. Linear statistics 8
3.3. Quadratic statistics 10
4. Numerical illustration 12
4.1. Description of the algorithm 12
4.2. Numerical results 16
Acknowledgements 21
Appendix A. Proof of Theorem 2.1 21
Appendix B. Proofs of Section 3.1 23
Appendix C. Proof of Theorem 3.7 25
Appendix D. Proof of Theorem 3.14 28
References 30

1. Introduction

This section contains the main elements of the considered model, some motivations and the plan of the paper. We consider here the so called Coulomb gas model that, in addition to its physical interest, shows an interesting behaviour in the limit of a large number of particles, see for instance [12, 47, 32]. The model consists of a set of random particles $X_{n,1},\ldots,X_{n,n}$ for $n \geq 2$, where each $X_{n,i}$ belongs to $\mathbb{R}^d$ for some physical dimension $d \geq 2$. The particles interact through the Coulomb kernel $g: \mathbb{R}^d \rightarrow \mathbb{R}$ defined by

$$g(x) = \begin{cases} 
\frac{1}{|x|}, & \text{if } d = 2, \\
\frac{1}{(d-2)|x|^{d-2}}, & \text{if } d \geq 3. 
\end{cases}$$

Date: Spring 2019, compiled July 15, 2019.
2000 Mathematics Subject Classification. 65C05 (Primary); 82C22; 60G57.
Key words and phrases. Coulomb gases; random matrices; large deviations; conditioning; Gibbs principle; numerical simulation; constrained dynamics.
This denomination comes from the equation satisfied by the interaction \(g\). Indeed, denoting by \(\delta_0\) the Dirac mass at 0, \(g\) solves in the sense of distributions the following Poisson problem:

\[
-\Delta g = c_d \delta_0, \quad \text{with} \quad c_d = \text{surface}(\{x \in \mathbb{R}^d : |x| = 1\}) = 2\pi^{d/2}/\Gamma(d/2). \tag{1.1}
\]

Note that \(\lim_{|x| \to +\infty} g(x) = 0\) if \(d \geq 3\), while \(\lim_{|x| \to +\infty} g(x) = +\infty\) if \(d = 2\). In (1.1) \(\Delta = \sum_{i=1}^d \partial_i^2\) denotes the Laplacian operator in \(\mathbb{R}^d\). In addition to this pair interaction, the particles are subject to a confining potential \(V : \mathbb{R}^d \to \mathbb{R}\) assumed to be lower semi-continuous and such that

\[
\lim_{|x| \to +\infty} (V(x) - 2I_{d=2 \log |x|}) > -\infty. \tag{1.2}
\]

Following [13] or [40], under this assumption, we can define the electrostatic energy on \(\mathcal{P}(\mathbb{R}^d)\) by

\[
\mathcal{E}(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \left( g(x - y) + \frac{V(x) + V(y)}{2} \right) \mu(dx)\mu(dy). \tag{1.3}
\]

This makes sense in \(\mathbb{R} \cup \{+\infty\}\) since the integrand is bounded from below thanks to the assumption (1.2) on \(V\). Moreover for all \(\mu \in \mathcal{P}(\mathbb{R}^d)\) such that \(\int \log(1 + |x|)1_{d=2}(dx) < \infty\), we have

\[
\mathcal{E}(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^d} (g(x - y)) \mu(dx)\mu(dy) + \int_{\mathbb{R}^d} V(x) \mu(dx). \tag{1.4}
\]

The functional \(\mathcal{E}\) has a unique minimizer on \(\mathcal{P}(\mathbb{R}^d)\) called the equilibrium measure [13] [10]

\[
\mu_* = \arg\min_{\mathcal{P}(\mathbb{R}^d)} \mathcal{E}. \tag{1.5}
\]

It has compact support, and if moreover \(V\) has a Lipschitz continuous derivative then it has density

\[
\frac{\Delta V}{2c_d} \tag{1.6}
\]

on the interior of its support. In particular if \(V\) is proportional to \(|\cdot|^2\), then \(\mu_*\) is uniform on a ball. The compactness of the support of \(\mu_*\) comes from the strong confinement assumption (1.2). Note that it is possible to consider weakly confining potentials for which the equilibrium measure still exists but is no longer compactly supported, see for instance the spherical ensemble in [27] [14].

Let \(X_n := (X_{n,1}, \ldots, X_{n,n})\) be a random vector of \((\mathbb{R}^d)^n\) with law

\[
P_n(dx) = \frac{e^{-\beta_n H_n(x_1, \ldots, x_n)}}{Z_n} dx_1 \cdots dx_n, \tag{1.7}
\]

where \(\beta_n > 0\) satisfies

\[
\lim_{n \to +\infty} \frac{\beta_n}{n} = +\infty, \tag{1.8}
\]

and

\[
H_n(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^n V(x_i) + \frac{1}{n^2} \sum_{i \neq j} g(x_i - x_j). \tag{1.9}
\]

This makes sense only if

\[
Z_n = \int_{(\mathbb{R}^d)^n} e^{-\beta_n H_n(x_1, \ldots, x_n)} dx_1 \cdots dx_n < \infty, \tag{1.10}
\]

which is the case when \(V\) satisfies

\[
\int_{\mathbb{R}^d} e^{-\beta_n (V(x) - 2\log(1 + |x|))} dx < \infty. \tag{1.11}
\]

This model is standard in mathematical physics: \(P_n\) is a Boltzmann–Gibbs measure modelling a gas of particles, called here a Coulomb gas, at inverse temperature \(\beta_n\) and with Hamiltonian \(H_n\). The law \(P_n\) is exchangeable in the sense that \(H_n\) is symmetric in \(x_1, \ldots, x_n\). Indeed, it depends on \(x_1, \ldots, x_n\) only via the empirical measure, namely, \(P_n\) almost surely,

\[
H_n = \int_{\mathbb{R}^d} V(x) \mu_n(dx) + \int_{\mathbb{R}^d} g(x - y)\mu_n(dx)\mu_n(dy) \quad \text{with} \quad \mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \tag{1.12}
\]

where the double integration runs over \(\{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d : x \neq y\}\). A heuristic reasoning suggests that, if \(\beta_n \to +\infty\) fast enough, under \(P_n\) the empirical measure \(\mu_n\) should concentrate in the
limit \( n \to +\infty \) on the equilibrium measure \( \mu_n \) that minimizes the energy \( \mathcal{E} \) in (1.4)-(1.5). This is intuited from the Laplace principle given the expression (1.7) for \( P_n \), where \( H_n \) is defined by (1.12). This intuition can be made rigorous through a large deviations principle (LDP), which can be established in this case and many others, see for instance [5] and the references therein. In particular, the case \( d = 2 \) with quadratic confinement \( V \) corresponds to the well-known Ginibre ensemble for random matrices [22]. We could also consider more general interactions, such as Riesz kernels [24], discontinuous or weak confinement, but we stick to this setting for ease of presentation. The technical requirements needed for extending our proofs will be pointed out throughout the paper, and cases not covered by the theoretical analysis will be investigated numerically in Section 3.

As large deviations are concerned with probabilities of rare fluctuations, it is possible to consider the empirical measure of the random gas conditioned on such a fluctuation. There has been a number of works on the behaviour of such gases conditioned on having an unusual proportion of the particles lying in some region of the space. As an example, for \( d = 2 \) and \( V \) quadratic, [3] reformulates the conditioned equilibrium measure through an obstacle problem. On the other hand [24] consider the rare situation in which there is a “hole” in the distribution, in other words no particle around zero. Finally [39] consider the one dimensional Wigner situation in which an abnormal proportion of particles lie on one side of the real line. Explicit expressions can be obtained in the latter case. The study of such conditionings is motivated by questions arising in theoretical physics, see for instance the references in [20].

While the above mentioned works bring substantial contributions to the understanding of conditioned random gas distributions, they also motivate further questions. Indeed, one may consider more general constraints, like conditioning on the barycenter of the cloud being far away from the origin. This may be of interest for both theoretical and practical purposes (if one wants to filter out noise conditioned on some rare event [9]). Moreover, the numerical methods proposed in [24] do not seem adapted to sampling the empirical distribution conditioned on some event – since this event is typically rare, naive sampling is generally inefficient. The goal of this paper is therefore to investigate some theoretical results on such conditioned Coulomb gases, as well as providing an efficient algorithm to sample conditioned distributions.

Mathematically, our aim is to consider the particles \( Y_n = (Y_{n,1}, \ldots, Y_{n,n}) \) in \((\mathbb{R}^d)^n\) such that
\[
Y_n \sim \text{Law}(X_n \mid \xi_n(X_n) \leq 0), \tag{1.13}
\]
where \( \xi_n : (\mathbb{R}^d)^n \to \mathbb{R} \), and to consider the limiting behaviour of the empirical measure
\[
\frac{1}{n} \sum_{i=1}^n \delta_{Y_{n,i}},
\]
as \( n \to +\infty \), depending on the confinement potential \( V \) and the constraint \( \xi_n \). Instead of an inequality constraint like (1.13), we may instead consider an equality constraint
\[
Y_n \sim \text{Law}(X_n \mid \xi_n(X_n) = 0).
\]
We will generally consider inequality constraints since they naturally lead to a Gibbs conditioning principle. Equality constraints could be considered as well by an additional limiting procedure, see [15] and the discussion in Section 3. We could also consider \( \xi_n \) to be \( \mathbb{R}^m \)-valued for some \( m \geq 2 \) but we restrict to one dimensional constraints for ease of exposition. The cases studied in [15] correspond to the choice
\[
\xi_n(x_1, \ldots, x_n) = \mu_n(1_U) - c,
\]
for some measurable set \( U \subset \mathbb{R}^d \) and constant \( c \in \mathbb{R} \). We will study in this paper more general linear statistics of the form
\[
\xi_n(x_1, \ldots, x_n) = \mu_n(\varphi) - c, \tag{1.14}
\]
for some constraint function \( \varphi : \mathbb{R}^d \to \mathbb{R} \) satisfying growth conditions, see Section 3.2. A particular case of interest is when the constraint function \( \varphi \) is itself linear, namely:
\[
\varphi(x) = x \cdot v - c, \tag{1.15}
\]
for \( v \in \mathbb{R}^d \) and \( c \in \mathbb{R} \). Indeed, when \( \varphi \) is chosen according to (1.15) and \( V \) is quadratic, the equilibrium measure under conditioning is the unconditioned one translated in the direction of \( v \). We provide a simple proof of this result in Section 2. We next turn to more general constraints in...
Section 3 proving first an abstract Gibbs conditioning principle in Section 3.1. When considering linear statistics, we prove in Section 3.2 that conditioning $P_n$ boils down to modifying the confinement potential $V$. In Section 3.3 we consider the case of quadratic statistics, which amounts to modifying the interaction kernel $g$.

In order to validate our theoretical results and explore cases in which explicit solutions are not available, we also propose a method for sampling the law of $Y_n$ for a fixed $n$. Based on the Hamiltonian Monte Carlo (HMC) method used in [11] for sampling Gibbs measures associated to Coulomb and Log-gases, we describe and implement the generalized Hamiltonian Monte Carlo algorithm proposed in [35] for sampling probability measures on submanifolds. The method, detailed in Section 4.1 shows remarkable performance, and the results presented in Section 4.2 are in agreement with the theory.

**Notations.** We introduce some notation used throughout the paper. For all $d \geq 1$ we denote by $|x| = (x_1^2 + \cdots + x_n^2)^{1/2}$ the Euclidean norm and by $x \cdot y = x_1 y_1 + \cdots + x_n y_n$ the scalar product on $\mathbb{R}^d$. We denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$ and, for all $p \geq 1$, by $\mathcal{P}_p(\mathbb{R}^d)$ those probability measures having finite $p$-moments in the sense that $|x|^p$ is integrable. For any measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, the support of $\mu$ is defined as $\text{supp}(\mu) = \mathbb{R}^d \setminus A$, where $A$ is the largest open set such that $\mu(A) = 0$ (which may be empty). For all measurable $f : \mathbb{R}^d \to \mathbb{R}$, we define

$$
\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)| \quad \text{and} \quad \|f\|_{\text{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.
$$

We define the bounded-Lipschitz distance on $\mathcal{P}(\mathbb{R}^d)$ by

$$
d_{\text{BL}}(\mu, \nu) = \sup_{\|f\|_{\text{Lip}} \leq 1} \int_{\mathbb{R}^d} f \, d(\mu - \nu).
$$

For all $p \geq 1$, we define the $p$-Wasserstein distance on $\mathcal{P}(\mathbb{R}^d)$ by

$$
d_{W_p}(\mu, \nu) = \left( \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) \right)^{1/p},
$$

where $\Pi(\mu, \nu)$ is the set of probability measures on the product space $\mathbb{R}^d \times \mathbb{R}^d$ with marginal distributions $\mu$ and $\nu$. For $p \geq 1$, the application $p \mapsto d_{W_p}$ is monotonic. Moreover, following [50], the Kantorovich–Rubinstein duality theorems state that

$$
d_{W_1}(\mu, \nu) = \sup_{\|f\|_{\text{Lip}} \leq 1} \int_{\mathbb{R}^d} f \, d(\mu - \nu) \quad \text{and} \quad d_{W_p}(\mu, \nu)^p = \sup_{f \in L^1(\mu), \, g \in L^1(\nu)} \left( \int_{\mathbb{R}^d} f \, d\mu - \int_{\mathbb{R}^d} g \, d\nu \right).
$$

For any $p \geq 1$, we say that a function $f$ is dominated by $|x|^p$ when

$$
\|f\|_{\infty,p} = \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{1 + |x|^p} < \infty.
$$

The $p$-Wasserstein topology is the one induced on $\mathcal{P}_p(\mathbb{R}^d)$ by $d_{W_p}$. If $(\nu_n)$_n is a sequence in $\mathcal{P}(\mathbb{R}^d)$ then $\lim_{n \to \infty} d_{W_p}(\nu_n, \nu) = 0$ if and only if $\lim_{n \to \infty} \int f \, d\nu_n = \int f \, d\nu$ for all bounded continuous $f : \mathbb{R}^d \to \mathbb{R}$. For all $p \geq 1$ and all sequence $(\nu_n)$_n in $\mathcal{P}_p(\mathbb{R}^d)$, we have $\lim_{n \to \infty} d_{W_p}(\nu_n, \nu) = 0$ if and only if $\lim_{n \to \infty} d_{BL}(\nu_n, \nu) = 0$ and $\lim_{n \to \infty} \int |x|^p \, d\nu_n = \int |x|^p \, d\nu$. In other words $d_{BL}$ metrizes weak convergence, while $d_{W_p}$ metrizes weak convergence plus $p$-moment convergence of the $p$-moment, see [50]. Moreover, for any $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, we denote by $\mu \ast \nu$ the convolution of $\nu$ with $\mu$. This probability measure is defined by its action over test functions $\varphi$ through

$$
(\mu \ast \nu)(\varphi) = \int_{\mathbb{R}^d} \varphi(x - y) \mu(dy) \nu(dx).
$$

We write $X \sim P$ to say that the random variable $X$ has law $P$, and $X \overset{d}{=} Y$ to say that the random variables $X$ and $Y$ have same law. A sequence of random variables $(X_1, \ldots, X_n)$ is exchangeable if for any permutation $\sigma$ of $\{1, \ldots, n\}$ it holds $(X_1, \ldots, X_n) \overset{d}{=} (X_{\sigma(1)}, \ldots, X_{\sigma(n)})$.\footnote{Or Monge, or Kantorovich, or transportation distance.}
We say that a sequence $Z_1, \ldots, Z_n$ of random variables taking value in a metric space $(Z,d)$ satisfies a large deviations principle at speed $\beta_n$ if, for any $A \subset Z$, it holds
\[
-\inf_{\mu} I(\mu) \leq \liminf_{n \to +\infty} \frac{1}{\beta_n} \log P(Z_n \in A) \leq \limsup_{n \to +\infty} \frac{1}{\beta_n} \log P(Z_n \in A) \leq -\inf_{\mu} I(\mu),
\]
where the interior and closure are taken with respect to the topology induced by $d$, while $I : Z \to [0, +\infty]$ is lower semicontinuous, and called the rate function. If $I$ has compact level sets for the topology induced by $d$, we say that $I$ is a good rate function.

We finally recall some elements of potential theory. The interaction energy
\[
J(\mu, \nu) = \int_{\mathbb{R}^d \times \mathbb{R}^d} g(x-y)\mu(dx)\nu(dy)
\]
is well defined for any signed measures $\mu, \nu$ with compact supports and takes values in $\mathbb{R} \cup \{+\infty\}$, see [30, Chapter I]. We use the abuse of notation $J(\mu) = J(\mu, \mu)$ for the associated quadratic form. Moreover, for any compact set $K \subset \mathbb{R}^d$, $J$ attains its infimum over probability measures supported on $K$. This value is called the capacity of the set $K$ [30, Chapter II]. A measurable set $A$ has positive capacity if it contains a compact set $K$ and a measure $\mu$ with $\text{supp}(\mu) \subset K$ and such that $J(\mu) < +\infty$. Otherwise, the set is said to have null capacity. A property is said to hold quasi-everywhere if it is satisfied on a set whose complement has null capacity. Although inner and outer capacities should be considered, we know these notions coincide for Borel sets on $\mathbb{R}^d$, see [30, Theorem 2.8]. Denoting by $\mathcal{P}_c(\mathbb{R}^d)$ the set of compactly supported probability measures, in accordance with [30, Theorems 1.15 and 1.16], for any $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ with $J(\mu) < +\infty$ and $J(\nu) < +\infty$, it holds $J(\mu - \nu) = 0$ if and only if $\mu = \nu$.

2. From conditioning to shifting: quadratic confinement with linear constraint

This section is devoted to the particular case where $V(x) = |x|^2$ and the constraint is chosen according to (1.14–1.15). The following theorem states that this special case is exactly solvable: the conditioning has the effect of a shift without deformation, due to a remarkable factorization. The proof, presented in Appendix A, is quite elegant. It is inspired from the seemingly unrelated work [15]. The result by itself appears as a special case of the general variational approach presented in Section 3.

Theorem 2.1 (From conditioning to shifting). Let $d,n \geq 2$ and $V = |x|^2$, so that (1.10) holds. Let $X_n := (X_{n,1}, \ldots, X_{n,n})$ and $P_n$ be as in (1.7). Then the equilibrium measure $\mu_*$ is the uniform law on the centered ball of $\mathbb{R}^d$ of radius 1. Moreover, almost surely and for all $p \geq 1$, it holds
\[
\lim_{n \to \infty} d_{W_p} \left( \frac{1}{n} \sum_{i=1}^{n} \delta_{X_{n,i}}, \mu_* \right) = 0,
\]
regardless of the way we define the random variables $X_n$ on the same probability space. Now let $v \in \mathbb{R}^d$ with $|v| = 1$, $c \in \mathbb{R}$, choose $\varphi(x) := x \cdot v - c$ and consider $Y_n := (Y_{n,1}, \ldots, Y_{n,n})$ with
\[
Y_n \sim \text{Law} \left( X_n \left| \frac{\varphi(X_{n,1}) + \cdots + \varphi(X_{n,n})}{n} = 0 \right. \right).
\]
Then
\[
Y_n \overset{d}{=} X_n + \left( c - \frac{X_{n,1} + \cdots + X_{n,n}}{n} \right) (v, \ldots, v).
\]
Moreover, denoting by $\mu^v := \delta_c * \mu_*$, we have that almost surely and for all $p \geq 1$, it holds
\[
\lim_{n \to \infty} d_{W_p} \left( \frac{1}{n} \sum_{i=1}^{n} \delta_{Y_{n,i}}, \mu^v \right) = 0,
\]
regardless of the way we define the random variables $Y_n$ on the same probability space.

The proof of Theorem 2.1 relies crucially on the quadratic nature of the confinement potential, but remains valid whatever the pair interaction, beyond Coulomb gases, as far as it is translation invariant. More general linear projections can be used. Indeed, if we choose $\varphi(x) = p(x) - c$ where $p$ is a linear projection over a subspace $E \subset \mathbb{R}^d$ of dimension $m$ and $c \in \mathbb{R}^m$, the result still holds.
Remark 2.2 (Ginibre random matrices and Hermite ensembles). When $d = 2$ and $\beta_n = (\beta/2)n^2$ for some $\beta > 0$, the probability distribution $P_n$ in Theorem 2.1 is a Coulomb gas known as the $\beta$ Ginibre or $\beta$ Hermite ensemble in random matrix theory. It appears also in the Laughlin fractional quantum Hall effect. The case $d = 2$ and $\beta = 2$ is even more remarkable and is known as the complex Ginibre ensemble. More precisely, let $M$ be an $n \times n$ random matrix with independent and identically distributed complex Gaussian entries with independent real and imaginary parts of mean 0 and variance $1/(2n)$. Its density is proportional to $e^{-n\text{Tr}(MM^*)}$. Its eigenvalues have law $P_n$ with $d = 2$, $\beta_n = n^2$, $V(x) = |x|^2$, see for instance [22, Chapter 15]. Let $\mathbb{R}^2 = \mathbb{C}$, $v \in \mathbb{R}^2$ with $|v| = 1$ and $c \in \mathbb{R}$. The assumptions of Theorem 2.1 are satisfied and the constraint in terms of matrices reads $\text{Tr}(M) \cdot v = nc$, where we identify again $\mathbb{C}$ with $\mathbb{R}^2$. More precisely (2.2) holds and the conditioned equilibrium measure reads

$$
\mu^\varphi(dz) = \frac{1_{|z-cv| \leq 1}}{\pi} dz.
$$

Since the entries of $M$ are independent, we have in particular the decomposition $M = M - \text{diag}(M) + \text{diag}(M)$ where $M - \text{diag}(M)$ and $\text{diag}(M)$ are independent. In this case we could probably deduce the desired result on $\mu^\varphi$ from the Tao and Vu universality theorem on the circular law [18] together with the Gibbs conditioning principle for independent Gaussian random variables to handle the diagonal part $\text{diag}(M)$ conditioned on the value of $\text{Tr}(M) = \text{Tr}(\text{diag}(M))$. Some numerical experiments are provided in Section 4.2. Note that such fixed trace random matrix models appear in the Physics literature, see for instance [1, 31].

In practice, we would like to consider a non-quadratic confinement and a non-linear constraint function $\varphi$. The numerical applications presented in Section 4.2 show a much wider range of behaviour than shifting the equilibrium measure. It turns out that the conditioning mechanism is an instance of the Gibbs conditioning principle from large deviations theory. The purpose of the next section is to provide proofs in this direction, which allow to derive the conditioned equilibrium measure in more general contexts, of which Theorem 2.1 appears as a particular case.

3. A general conditioning framework

As is known from the seminal work of Ben Arous and Guionnet [5], large deviations theory provides a natural framework to study the concentration of empirical measures of the spectrum of random matrices and, beyond, of singularly interacting particles systems. We refer in particular to [12, 19, 7] and references therein for recent accounts. Since large deviations theory is concerned with estimating probabilities of rare events, conditioning on such a rare event is a natural direction to follow. This procedure is generally referred to as Gibbs conditioning principle or maximum entropy principle. This principle is explained for instance in [14, Section 6.3] and [15, Section 7.3].

When no conditioning is considered, we know that under mild assumptions the empirical measure associated to the Gibbs measure (1.7) satisfies a LDP with rate function $\mathcal{E}$. When the random gas is considered under conditioning on an appropriate rare event, the Gibbs conditioning principle states that the resulting conditioned empirical measures concentrate on a minimizer of $\mathcal{E}$ under constraint. Proofs of this fact in our context are presented in Section 3.1. Next, Section 3.2 studies the corresponding constrained minimization problem for linear statistics, while Section 3.3 is concerned with quadratic statistics.

3.1. Gibbs conditioning. The goal of this section is to present an abstract Gibbs conditioning principle and apply it to the Coulomb gas model. Most works considered hitherto Gibbs principles associated to Sanov’s theorem [18, 43, 57], in other words in absence of interaction, showing that by conditioning the empirical measure, the resulting equilibrium measure minimizes the rate function under constraint. The same strategy can actually be applied to any exchangeable system satisfying a large deviations principle provided the conditioning set is an $\mathcal{I}$-continuity set, following for instance [10, 13, Section 1.2] and [14, Section 5.3]. This is the purpose of the next proposition, which can be of independent interest, and whose proof is postponed to Appendix B.

**Proposition 3.1** (A Gibbs conditioning principle). Suppose that $Z_1, \ldots, Z_n$ are random variables taking values in a metric space $(Z, d)$ satisfying a large deviations principle at speed $(\beta_n)_n$, and
with good rate function $I$. Consider a closed set $B$ which is $I$-continuous in the sense that
\[
\inf_B I = \inf_B I < +\infty.
\] (3.1)

Then, the set of minimizers
\[
\mathcal{F}_B = \left\{ z \in \mathcal{Z} : I(z) = \inf_B I \right\}
\] (3.2)
is a non-empty closed subset of $B$. Moreover, for any $\varepsilon > 0$, setting
\[
A_\varepsilon = \left\{ z \in \mathcal{Z} : d(z, \mathcal{F}_B) > \varepsilon \right\},
\]
there exists $c_\varepsilon > 0$ such that
\[
\limsup_{n \to +\infty} \frac{1}{\beta_n} \log \mathbb{P} \left( Z_n \in A_\varepsilon \mid Z_n \in B \right) \leq -c_\varepsilon.
\] (3.3)

In particular if we define a random variable $Z'_n \sim \text{Law}(Z_n \mid Z_n \in B)$ for all $n$ then almost surely
\[
\lim_{n \to +\infty} d(Z'_n, \mathcal{F}_B) = 0
\] regardless of the way we define the $Z'_n$’s on the same probability space.

Finally, in the particular case where $\mathcal{F}_B = \{z_B\}$ is a singleton then almost surely $\lim_{n \to +\infty} Z'_n = z_B = \min_B I$.

In words, (3.3) shows that the variables $Z_n$ conditioned on being in $B$ concentrate on a minimizer of $I$ over $B$. It is alluring to consider a more general LDP for the conditioned empirical measure as in [29], but the arguments proposed in [29] do not fit the case of the singular rate functional (1.4). Our strategy is then to restrict to an $I$-continuity set $B$ satisfying (3.1), which allows to use the lower bound of the LDP, very similarly to [14].

In order to use Proposition 3.1 in the Coulomb gas setting, we start by recalling a LDP associated with the Coulomb gas model (see [12] [19]). In order to consider unbounded constraints in what follows, we make the following assumption.

**Assumption 3.2 (Growth condition).** There exist $a > 0$, $R \in \mathbb{R}$ and $q > 1$ such that
\[
\forall x \in \mathbb{R}^d, \quad V(x) \geq a|x|^q - R.
\]

The above growth condition not only ensures that $V$ satisfies (1.2), but also allows to consider a finer topology for the LDP, see [19] Theorem 1.8. It could certainly be relaxed under appropriate modifications. In particular, Assumption 3.2 shows that [19] Assumption C’1] is satisfied for any function of the form $|x|^p$ for $1 < q < p < q$, so [19] Lemma 1.1 applies. Assumption 3.2 thus has the following consequence. Consider an exponent $p \in (1, q]$. Then, under $P_n$ defined in (1.7)–(1.8) the empirical measure
\[
\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}
\]
satisfies a large deviations principle (1.17) in the $p$-Wasserstein topology at speed $(\beta_n)_n$ and with the following good rate function:
\[
\mathcal{E}_* = \mathcal{E} - \inf_{\mathcal{P}_p(\mathbb{R}^d)} \mathcal{E}, \quad \text{with} \quad \inf_{\mathcal{P}_p(\mathbb{R}^d)} \mathcal{E} > -\infty,
\]
where $\mathcal{E}$ is defined in (1.3). The energy $\mathcal{E}$ has additional nice properties, which we recall below for convenience.

**Proposition 3.3 (Properties of the electrostatic energy).** Let $\mathcal{E}$ be as in (1.3). Suppose that Assumption 3.2 holds, and take some $p \in (1, q)$. Then, denoting by $D_\mathcal{E} = \{ \mu \in \mathcal{P}(\mathbb{R}^d) : \mathcal{E}(\mu) < +\infty \}$ the domain of $\mathcal{E}$, the following properties are satisfied:

- $D_\mathcal{E}$ is convex and $\mathcal{E}$ is convex on $D_\mathcal{E}$;
- $D_\mathcal{E} \subset \mathcal{P}_p(\mathbb{R}^d)$ and there exists a unique $\mu_\star \in \mathcal{P}_p(\mathbb{R}^d)$ such that
  \[
  \mathcal{E}(\mu_\star) = \inf_{\mathcal{P}_p(\mathbb{R}^d)} \mathcal{E} = \inf_{\mathcal{P}_p(\mathbb{R}^d)} \mathcal{E} = \inf_{\mathcal{P}_p(\mathbb{R}^d)} \mathcal{E};
  \] (3.4)
- the minimizer $\mu_\star \in \mathcal{P}_p(\mathbb{R}^d)$ is unique and satisfies the Euler–Lagrange conditions (where $C_\star = \mathcal{E}(\mu_\star)$)
  \[
  \begin{cases}
  2g * \mu_\star + V = C_\star, & \text{quasi-everywhere in supp}(\mu_\star), \\
  2g * \mu_\star + V \geq C_\star, & \text{quasi-everywhere}.
  \end{cases}
  \] (3.5)
The domain is not empty since it contains for instance measures with a smooth density over a compact support. For convenience, we recall a proof of these classical results in Appendix [3].

**Remark 3.4** (Going beyond Coulomb gases and convexity). The LDP presented here holds for a much larger range of models than the Coulomb gas setting, see for instance [12]. However, the assumptions in [12] do not ensure the convexity of the rate function $\mathcal{E}$, which poses problems when it comes to identifying the equilibrium measure – we thus stick to this setting here. In practice, the convexity of $\mathcal{E}$ is derived from a Bochner-type positivity of the interaction potential, see [12].

We are now in position to apply Proposition [3.1] to the Coulomb gas model.

**Theorem 3.5** (Gibbs conditioning for Coulomb gases). Let $\mathcal{E}$ be as in (1.3). Suppose that Assumption [3.2] holds and take some $p \in (1, q)$. Consider a closed set $B \subset \mathcal{P}(\mathbb{R}^d)$ such that

$$\inf_B \mathcal{E} = \inf_B \mathcal{E} < +\infty,$$

where the interior is taken with respect to the $p$-Wasserstein topology. Then the set of minimizers

$$\mathcal{E}_B = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \mathcal{E}(\mu) = \inf_B \mathcal{E} \right\}$$

is a non-empty closed subset of $B$. Moreover, if $X_n \sim P_n$ is as in (1.7), and if $Y_n = (Y_{n,1}, \ldots, Y_{n,n})$ is such that

$$Y_n \sim \text{Law}(X_n \mid \mu_n \in B), \quad \text{with} \quad \mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_{n,i}},$$

then almost surely it holds

$$\lim_{n \to \infty} d_{W_p} \left( \frac{1}{n} \sum_{i=1}^{n} \delta_{Y_{n,i}}, \mathcal{E}_B \right) = 0,$$

regardless of the way we define the random variables $Y_n$ on the same probability space.

The proof of Theorem [3.5] which can be found in Appendix [3], is an instance of the Gibbs conditioning principle of Proposition [3.1] where the conditioned empirical measure concentrates almost surely on a minimizer of $\mathcal{E}$ over $B$.

Next, rather than aiming at the greatest generality, we consider the case of linear and quadratic statistics constraints, for which $I$-continuity can be proved and the resulting equilibrium measure can be identified in terms of a modified version of (3.5).

### 3.2. Linear statistics

As explained in the introduction, the case of linear statistics is of particular importance. This motivates focusing first on conditioning sets $B$ of the form

$$B = \left\{ \nu \in \mathcal{P}_n(\mathbb{R}^d) : \nu(\varphi) \leq 0 \right\},$$

for some measurable function $\varphi : \mathbb{R}^d \to \mathbb{R}$. This kind of constraint was studied for example in [3] [24] [25] [33] [40]. In particular, the Ginibre case with $\varphi = 1_U - c$ for a measurable set $U \subset \mathbb{R}^2$ and $c \in \mathbb{R}$ is considered in [3]. The choice $\varphi(x) = c - x \cdot v$ for $v \in \mathbb{R}^d$ has been treated in Section 2 for the related equality constraint. We consider here more general potentials $V$ and constraint functions $\varphi$. The next assumption on $\varphi$ ensures that $B$ is suitable for conditioning.

**Assumption 3.6.**

- Assumption [3.2] holds for some $q > 1$;
- $\|\varphi\|_{\text{Lip}} < +\infty$ and thus $\|\varphi\|_{\infty,p} < +\infty$ for all $p \in (1, q)$;
- there exists $\mu_- \in D_S$ such that $\mu_-(\varphi) < 0$;
- there exists $\mu_+ \in D_S$ such that $\mu_+(\varphi) > 0$.

The existence of $\mu_-$ means that the set $B$ has non empty interior, while that of $\mu_+$ implies that $B \neq \mathcal{P}_n(\mathbb{R}^d)$, so that the constraint is not trivial. Since the Gibbs principle relies on $B$ being an $I$-continuity set, we provide a fine analysis of the minimization of $\mathcal{E}$ over the set $B$ defined in (3.8). We prove in particular that the minimizer is unique with compact support, and we characterize it through an integral equation similar to (3.5) with an additional Lagrange multiplier. The proof of this result is presented in Appendix [3].
Theorem 3.7 (Variational characterization). Let \( \mu_* \in \mathcal{P}_c(\mathbb{R}^d) \) be the unconstrained equilibrium measure as in Proposition 3.3, and let \( B \) the set defined in (3.5). Suppose that Assumption 3.6 holds, for some \( q > 1 \) and \( p \in (1, q) \). Then \( B \) is closed in the \( p \)-Wasserstein topology and
\[
\inf_B \mathcal{E} = \inf_B \mathcal{E} < +\infty.
\]
Moreover
\[
\mathcal{E}_B = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \mathcal{E}(\mu) = \inf_B \mathcal{E} \right\} = \{ \mu^\varphi \},
\]
where \( \mu^\varphi \) has compact support and is solution to, for some \( \alpha \geq 0 \),
\[
\begin{align*}
2g \ast \mu^\varphi + V + \alpha \varphi &= C_\varphi, \quad \text{quasi-everywhere in } \text{supp}(\mu^\varphi), \\
2g \ast \mu^\varphi + V + \alpha \varphi &\geq C_\varphi, \quad \text{quasi-everywhere},
\end{align*}
\]
with \( C_\varphi = \mathcal{E}(\mu^\varphi) \). Finally, one of the two following conditions holds:
- \( \mu_* \in B \) and \( \alpha = 0 \);
- \( \mu_* \notin B \), in which case \( \mu^\varphi(\varphi) = 0 \) and \( \alpha > 0 \). In other words, the constraint is saturated and the Lagrange multiplier is active.

We now have the following consequence of Theorems 3.5 and 3.7

Corollary 3.8 (From conditioning to confinement deformation). Suppose that Assumption 3.6 holds with \( q > 1 \) and \( p \in (1, q) \). Consider a Coulomb gas \( X_n = (X_{n,1}, \ldots, X_{n,n}) \sim P_n \) as in (1.7). Introduce \( Y_n = (Y_{n,1}, \ldots, Y_{n,n}) \) with law given by
\[
Y_n \sim \text{Law} \left( X_n \left| \frac{1}{n} \sum_{i=1}^n \varphi(X_{n,i}) \leq 0 \right. \right).
\]
Let \( \mu^\varphi \) be as in Theorem 3.7. Then almost surely it holds
\[
\lim_{n \to \infty} d_{W_p} \left( \frac{1}{n} \sum_{i=1}^n \delta_{Y_{n,i}}, \mu^\varphi \right) = 0,
\]
regardless of the way we define the random variables \( Y_n \) on the same probability space.

Theorem 3.7 and Corollary 3.8 show that conditioning on a linear statistics is equivalent to changing the confinement potential \( V \) into \( V + \alpha \varphi \) where \( \alpha \geq 0 \) is a constant determined by the constraint. If \( \mu_* \in B \), \( \alpha = 0 \) and the conditioning produces no effect. Note also that the global Lipschitz condition on \( \varphi \) in Assumption 3.6 could possibly be relaxed. For instance, if \( \|\varphi\|_{\infty, p} < +\infty \) for some \( p \in (1, q) \) we expect that a minimizing measure has compact support, and that assuming \( \varphi \) locally Lipschitz suffices to prove Theorem 3.7. We leave these refinements to further studies.

Remark 3.9 (Equality constraints). When considering conditioning principles, one is often interested in equality constraints. It is not obvious at first sight to consider a set \( B \) defined by an equality constraint, since its interior may well be empty. A common strategy is to use a limiting procedure by introducing nested sets [13]. This is unnecessary here since we observe in Theorem 3.7 that either the equilibrium measure lies in \( B \), either the constraint is saturated.

Remark 3.10 (Projection). The conditioned equilibrium measure \( \mu^\varphi \) can be interpreted as an instance of entropic projection. These projections have been studied for a long time in the context of the Sanov theorem, in other words independent particles or equivalently product measures without interaction at all, see [30] and the references therein. Theorem 3.7 is therefore a precise study of such a projection in the context of Coulomb gases under linear statistics constraint, where the entropy is replaced by the electrostatic energy \( \mathcal{E} \). These remarks also apply to Section 3.3.

Remark 3.11 (Formula for constrained equilibrium measure under regularity assumptions). Suppose that Assumption 3.6 holds and that \( V \) and \( \varphi \) have Lipschitz continuous derivatives. Then the conditioned equilibrium measure \( \mu^\varphi \) which appears in Theorem 3.7 and in Corollary 3.8 satisfies
\[
\begin{align*}
\mu^\varphi &= \frac{\Delta V + \alpha \Delta \varphi}{2c_d}, \quad \text{almost everywhere in } \text{supp}(\mu^\varphi), \\
\mu^\varphi &= 0, \quad \text{almost everywhere outside } \text{supp}(\mu^\varphi),
\end{align*}
\]
Indeed, it suffices to apply the Laplacian to both sides of (3.10) and use (1.1); see for example [46] Proposition 2.22] for the technical details. We mention that a density is non-negative, hence
$$\text{supp}(\mu^\varphi) \subset \{ x \in \mathbb{R}^d \mid \Delta V + \alpha \Delta \varphi \geq 0 \}.$$ 

The constraint may therefore significantly change the support of the equilibrium measure.

It is now possible to come back to the translation phenomenon described in Section 2 through the energetic approach considered in this section.

**Alternative proof of Theorem 2.1.** Using (3.12) under the assumptions of Theorem 2.1 we have \( \Delta V = 2d \) and \( \Delta \varphi = 0 \), so that \( \mu^\varphi \) is constant and equal to \( d/c_d \) on its support. It then remains to show that this support is indeed a ball of correct center and radius. For this, we observe that, since \( |\nu| = 1 \),
$$V(x) + \alpha \varphi(x) = |x|^2 + \alpha (e - x \cdot v) = \left| x - \frac{\alpha v}{2} \right|^2 + \frac{\alpha^2}{4} + \alpha c,$$
so that the effective confining potential is quadratic with variance \( 1/2 \) and center \( x_0 = \alpha v/2 \). By radial symmetry around \( x_0 \), \( \mu^\varphi \) must be a uniform distribution on a ball \( B(x_0, r) \) centered at \( x_0 \) with radius \( r > 0 \). In order to find the value of \( \alpha \), we write the constraint
$$|B(x_0, r)|^{-1} \int_{B(x_0, r)} x \cdot v \, dx = c.$$
The left hand side of the above equation reads, by symmetry,
$$|B(x_0, r)|^{-1} \int_{B(x_0, r)} (x - x_0) \cdot v \, dx + x_0 \cdot v = |B(x_0, r)|^{-1} \int_{B(0, r)} x \cdot v \, dx + x_0 \cdot v = x_0 \cdot v.$$
Since \( x_0 = \alpha v/2 \) and \( |\nu|^2 = 1 \) we obtain
$$\alpha = 2c,$$
which leads to \( x_0 = cv \). Finally, the value of \( \mu^\varphi \) over its support is \( d/c_d \), where \( c_d \) is the surface of the sphere in dimension \( d \). Since the volume of the sphere of radius \( r \) is equal to \( r c_d/d \), we obtain that \( r = 1 \) and we reach the conclusion of Theorem 2.1. \( \square \)

### 3.3. Quadratic statistics.

Once the linear statistics case has been studied, it is natural to turn to more general constraints. Considering second order statistics is a first step in this direction, which motivates to consider sets of the form, for \( p > 1 \),
$$B = \{ \nu \in \mathcal{P}_p(\mathbb{R}^d) : Q(\nu) \leq 0 \},$$
where \( Q \) is the “quadratic form”
$$Q : \mu \in \mathcal{P}_p(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x, y) \mu(dx)\mu(dy),$$
and \( \psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) is a prescribed function. For any \( \mu \in \mathcal{P}(\mathbb{R}^d) \), we denote by
$$U^\psi_\mu : x \in \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} \psi(x, y) \mu(dy)$$
the “potential” generated by \( \mu \) for the interaction \( \psi \), whenever this makes sense. We now make some assumptions on the interaction \( \psi \) for the functional \( Q \) to define an \( I \)-continuity set \( B \) in (3.13).

**Assumption 3.12.**
- Assumption 3.2 holds for some \( q > 1 \);
- there is \( C_{\text{Lip}} > 0 \) such that, for any \( \mu \in \mathcal{P}(\mathbb{R}^d) \),
  $$\|U^\psi_\mu\|_{\text{Lip}} \leq C_{\text{Lip}},$$
and thus \( \|U^\psi_\mu\|_{\infty, p} < +\infty \) for all \( p \in (1, q) \);
- \( \psi \) is symmetric, i.e. \( \psi(x, y) = \psi(y, x) \) for all \( x, y \in \mathbb{R}^d \);
- \( Q \) is convex;
- there exists \( \mu_- \in D_F \) such that \( Q(\mu_-) < 0 \);
- there exists \( \mu_+ \in D_F \) such that \( Q(\mu_+) > 0 \).

Before turning to the minimization under constraint, let us present a class of functions \( \psi \) for which (3.15) holds.
Corollary 3.15 which led to a change of confinement. From Theorem 3.14, we obtain the following result.

\[ \|\phi\|_{Lip} < +\infty \text{ and thus } \|\phi\|_{Lip} < +\infty. \]

Then (3.15) holds with \( C_{Lip} = \|\phi\|_{Lip}. \)

Proof of Proposition 3.13. For all \( \nu \in \mathcal{P}(\mathbb{R}^d) \) and \( x, x' \in \mathbb{R}^d \), it holds

\[ |U^\nu(x) - U^\nu(x')| \leq \int_{\mathbb{R}^d} |\phi(x-y) - \phi(x'-y)| \nu(dy) \leq \|\phi\|_{Lip} \int_{\mathbb{R}^d} |x-y-(x'-y)| \nu(dy) = \|\phi\|_{Lip} |x-x'|. \]

We thus obtain (3.15) with \( C_{Lip} = \|\phi\|_{Lip}. \) \( \square \)

Remark 3.16 (Higher order constraints, convexity, and regularity) From the proof of Theorem 3.14, the Gibbs principle holds for a set \( B \) of the form (3.11) for \( C \) convex and lower semicontinuous. However, we would not be able to say much on the solution in such an abstract setting. In particular, higher order statistics could be considered, leading to higher order convolutions, but checking the convexity of the associated functional becomes less convenient. By lack of applications in mind, we do not consider these higher order constraints here.

Proposition 3.13 (Sufficient condition for (3.15)). Assume that \( \psi(x, y) = \phi(x-y) \) for a function \( \phi: \mathbb{R}^d \to \mathbb{R} \) satisfying \( \|\phi\|_{Lip} < +\infty \) and thus \( \|\phi\|_{Lip} < +\infty \). Then (3.15) holds for some \( q > 1 \).

Remark 3.16 (Higher order constraints, convexity, and regularity). From the proof of Theorem 3.14, the Gibbs principle holds for a set \( B \) of the form (3.11) for \( C \) convex and lower semicontinuous. However, we would not be able to say much on the solution in such an abstract setting. In particular, higher order statistics could be considered, leading to higher order convolutions, but checking the convexity of the associated functional becomes less convenient. By lack of applications in mind, we do not consider these higher order constraints here.
Remark 3.17 (Formula for constrained equilibrium measure under regularity assumptions). Suppose that Assumption 3.12 holds, that $V$ has Lipschitz continuous derivatives and that $\psi$ is $C^2(\mathbb{R}^d)$. Then the conditioned equilibrium measure $\mu^\psi$ which appears in Theorem 3.14 and in Corollary 3.15 satisfies
\[
\begin{cases}
    c_d\mu^\psi - \alpha \int_{\mathbb{R}^d} \Delta \psi(\cdot, y) \mu^\psi(dy) = \frac{\Delta V}{2}, & \text{almost everywhere in } \text{supp}(\mu^\psi), \\
    \mu^\psi = 0, & \text{almost everywhere outside } \text{supp}(\mu^\psi).
\end{cases}
\] (3.19)

Indeed, (3.19) follows by applying the Laplacian on both sides of (3.17). Note that the expression (3.19) is not explicit as in the linear constraint case of Remark 3.14 because we are not able to invert the convolution associated to $\Delta \psi$ in general.

4. Numerical illustration

In this section we consider the problem of sampling from conditioned distributions of the form
\[
\text{Law}(X_n \mid \xi_n(X_n) = 0),
\] (4.1)
where $\xi_n : (\mathbb{R}^d)^n \to \mathbb{R}^m$ for some $m \geq 1$, and $X_n$ is distributed according to $P_n$ defined in (1.7) for $n$ fixed. We drop the index $n$ on $\xi_n$ in what follows to shorten the notation, and consider constraints taking values in $\mathbb{R}^m$ for generality. Note that we consider equality rather than inequality constraints since we have seen in Sections 3.2 and 3.3 that inequality constraints are either satisfied by the equilibrium measure or saturated.

The first contribution of this section is to propose in Section 4.1 an algorithm for sampling from (4.1). In a second step, we present in Section 4.2 some numerical applications, where we illustrate the predictions of Sections 2 and 3. This is also the opportunity to explore conjectures which are not proved in the present paper.

4.1. Description of the algorithm. The description of the constrained Hamiltonian Monte Carlo algorithm used for sampling follows several steps. We first make precise the structure of the measure (4.1) in Section 4.1.1. Section 4.1.2 next introduces a constrained Langevin dynamics used for sampling, while Section 4.1.3 gives the details of the numerical integration.

4.1.1. Dirac and Lebesgue measures on submanifolds. Let us first describe more precisely the structure of the constrained measure (4.1) by introducing the submanifold $\mathcal{M}_z$ associated with the $z$-level set of $\xi$ for $z \in \mathbb{R}^m$, namely
\[
\mathcal{M}_z = \{ x \in (\mathbb{R}^d)^n : \xi(x) = z \}. \quad (4.2)
\]
We use the shorthand notation $\mathcal{M} = \mathcal{M}_0$. To define the conditioned measure, we rely on the following disintegration of (Lebesgue) measure formula: for any bounded continuous test function $\varphi$,
\[
\int_{(\mathbb{R}^d)^n} \varphi(x) \, dx = \int_{\mathbb{R}^m} \int_{\mathcal{M}_z} \varphi(x) \delta_{\xi(x) - z}(dx) \, dz.
\] (4.3)
This defines for any $z \in \mathbb{R}^m$ the conditioned measure $\delta_{\xi(x) - z}(dx)$, see [34] Section 2.3.2. Since $P_n$ is given by (1.7), the constrained measure (4.1) can be written with the conditioned measure $\delta_{\xi(x)}(dx)$ associated with $\mathcal{M}$ as: for any bounded continuous $\varphi$,
\[
\mathbb{E}[\varphi(X_n) \mid \xi_n(X_n) = 0] = \frac{1}{Z_n^\xi} \int_{\mathcal{M}} \varphi(x) e^{-\beta_n H_n(x)} \delta_{\xi(x)}(dx),
\] (4.4)
where $Z_n^\xi$ is a normalizing constant. The measure of interest is therefore
\[
P_n^\xi(dx) = \frac{e^{-\beta_n H_n(x)}}{Z_n^\xi} \delta_{\xi(x)}(dx).
\] (4.5)
In order to obtain a better understanding of (4.5), we relate the conditioned measure proportional to $\delta_{\xi(x)}(dx)$ to the Lebesgue measure induced on the submanifold $\mathcal{M}$ by the canonical Euclidean scalar product, which we denote by $\sigma_{\mathcal{M}}(dx)$. We use to this end the co-area formula [2, 21, 34]. We denote by $\nabla \xi = (\nabla \xi_1, \ldots, \nabla \xi_m) \in \mathbb{R}^{dn \times m}$ and introduce the Gram matrix:
\[
G(x) = \nabla \xi(x)^T \nabla \xi(x) \in \mathbb{R}^{m \times m},
\] (4.6)
where the superscript $T$ denotes matrix transposition. In what follows, we assume that the Gram matrix $[4.6]$ is non-degenerate in the sense that $G(x)$ is invertible for $x$ in a neighborhood of $M$ (see [34 Proposition 2.1]).

**Proposition 4.1.** The measures $\delta_{\xi(x)}(dx)$ and $\sigma_M(dx)$ are related by

$$\delta_{\xi(x)}(dx) = |\det G(x)|^{-\frac{1}{2}} \sigma_M(dx).$$  

(4.7)

In particular it holds

$$P_n^t(dx) = \frac{e^{-\beta_n H_n^t(x)}}{Z_n^t} \sigma_M(dx),$$

(4.8)

where

$$H_n^t(x) = H_n(x) + U_n(x), \quad U_n(x) = -\frac{1}{2\beta_n} \log |\det G(x)|.$$  

(4.9)

**Remark 4.2** (Parametrization invariance). It seems at first sight that the definition of the conditioned measure in (4.3) depends on the choice of parametrization of $\xi$, but it does not. To illustrate this point, we consider for simplicity that $m = 1$ and $M = \{x \in (\mathbb{R}^d)^n : \xi(x) = 0\}$. First, the induced Lebesgue measure on $M$ does not depend on the parametrization of $\xi$. Consider next a smooth function $F : \mathbb{R} \to \mathbb{R}$ such that $F(0) = 0$ and $F'(0) \neq 0$, and the change of parametrization $M = \{x \in (\mathbb{R}^d)^n : F(\xi(x)) = 0\}$. The gradient of the constraint at $x \in (\mathbb{R}^d)^n$ is then $\nabla F(\xi(x)) = F'(\xi(x))\nabla \xi(x)$. Since $\xi(x) = 0$ for $x \in M$, the right hand side of (4.7) is changed only by a multiplicative factor $|F'(0)|^2 \neq 0$. Therefore, the conditioned probability measure (4.3) is left unchanged.

The aim is therefore to sample from (4.8), which is not an easy task except in very particular situations, like the one studied in Section 2. The attempts in [24][25][40] show that a naive approach is not efficient in general, since the conditioning event is rare. Actually, sampling probability distributions under constraint is a long standing problem in molecular dynamics and computational statistics. Concerning molecular simulation, one can be interested in fixing some degrees of freedom of a system like bond lengths, or the value of a so-called reaction coordinate, typically for free energy computations—we refer e.g. to [33][17][35] for more details. An example of application in computational statistics is for instance Approximate Bayesian Computations, see [40][42].

For sampling measures on submanifolds, a naive penalization of the constraint is not a good idea in general, since the dynamics used to sample it (such as (4.10) below) are difficult to integrate because of the stiffness of the penalized energy. Moreover, our problem is made harder by the singularity of the pair interaction in the Hamiltonian (1.9). It is known that Hybrid Monte Carlo schemes (relying on a second order discretization of an underdamped Langevin dynamics with a Metropolis–Hastings acceptance rule) provide efficient methods for sampling such probability distributions, see [11] and references therein. An issue when combining a Metropolis–Hastings rule with a projection on a submanifold is that reversibility may be lost, which introduces a bias. A recent strategy has been to introduce a reversibility check in addition to the standard acceptance-rejection rule, which makes the HMC scheme under constraint reversible [52][35]. Note that [35] proposes an interesting alternative to the scheme used here, which is however not compatible with a Metropolis selection procedure in its current form. We thus present the algorithm as written in [35], with some simplifications and adaptations to our context, for which we introduce next the constrained Langevin dynamics.

**4.1.2. Constrained Langevin dynamics.** We define here an underdamped Langevin dynamics over the submanifold $M$, whose invariant measure has a marginal in position which coincides with (4.8). We motivate using this dynamics by first considering the problem of sampling from the unconstrained measure $P_n$. For a given $\gamma > 0$, we define

$$
\begin{align*}
\text{d}X_t &= Y_t \, \text{d}t, \\
\text{d}Y_t &= -\nabla H_n(X_t) \, \text{d}t - \gamma Y_t \, \text{d}t + \sqrt{\frac{2\gamma}{\beta_n}} \, \text{d}W_t,
\end{align*}
$$

(4.10)

where $(W_t)_{t \geq 0}$ is a $dn$-dimensional Wiener process. In this dynamics, $(X_t)_{t \geq 0}$ stands for a position, while $(Y_t)_{t \geq 0}$ represents a momentum variable. Let us mention that the long time convergence of the law of this process towards $P_n$ (a difficult problem due to the singularity of the Hamiltonian)
can be proved through Lyapunov function techniques, see [38] for a recent account. In practice, the singularity of \( g \) also makes the numerical integration of (4.10) difficult, and a Metropolis–Hastings selection rule can be used to stabilize the numerical discretization, see [11] and references therein. The algorithm described below makes precise how to adapt this strategy to sample measures constrained to the submanifold \( \mathcal{M} \).

Since we aim at sampling from (4.18), it is natural to consider the dynamics (4.10) with positions constrained to the submanifold (4.2), that is

\[
\begin{align*}
\text{d}X_t &= Y_t \, \text{d}t, \\
\text{d}Y_t &= -\nabla H_n(X_t) \, \text{d}t - \gamma Y_t \, \text{d}t + \sqrt{2\gamma} \, \text{d}W_t + \nabla \xi(X_t) \, \text{d}\theta_t, \\
\xi(X_t) &= 0,
\end{align*}
\tag{4.11}
\]

where \((\theta_t)_{t \geq 0} \in \mathbb{R}^m\) is a Lagrange multiplier enforcing the dynamics to stay on \( \mathcal{M} \). Let us emphasize that the position constraint induces a hidden constraint on the momenta in (4.11), which reads

\[\forall t \geq 0, \quad \nabla \xi(X_t)^T Y_t = 0.\]

The above relation is obtained by taking the derivative of \( t \mapsto \xi(X_t) \) along the dynamics (4.11). This implies that momenta are tangent to the submanifold's zero level set, which is a natural geometric constraint [34, 35]. However, the dynamics (4.11) does not sample from the conditioned measure (4.1), as shown in the following proposition [34, 35].

**Proposition 4.3 (Invariant measure).** The dynamics (4.11) has a unique invariant measure with marginal distribution in position given by

\[
\frac{e^{-\beta_n H_n(x)}}{Z_n} \sigma_{\mathcal{M}}(\text{d}x),
\]

where \( Z_n \) is a normalization constant.

Although (4.11) does not sample from \( P_n^X \), we have seen in Section 4.1.1 how to fix this problem. More precisely, Proposition 4.1 shows that the dynamics (4.11) run with the modified Hamiltonian \( H_n^\gamma \) defined in (4.9) samples from \( P_n^X \).

However, in practice, it may be preferable not to use the gradient of \( U_n \), since it involves the Hessian of the constraint \( \xi \) and may be cumbersome to compute. Therefore, we will not run the dynamics (4.11) with the modified Hamiltonian \( H_n^\gamma \) but with \( H_n \), and perform some reweighting to correct for the bias arising from the missing factor \( |\det G(x)|^{-\frac{1}{2}} \). As explained in Remark 4.8 below in the context of a HMC discretization, this ensures that we are sampling from the correct target distribution while only moderately increasing the rejection rate.

### 4.1.3. Discretization

In order to make a practical use of (4.11) combined with Proposition 4.1 we need to define a discretization scheme. We present below the strategy proposed by [35], which relies on a second order discretization of (4.11) with a Metropolis–Hastings selection and a reversibility check.

As discussed after (4.11), momenta are tangent to the level sets of the submanifold. We introduce

\[
\Pi_{\mathcal{M}} = \text{Id} - \nabla \xi(x) G^{-1}(x) \nabla \xi(x)^T \in (\mathbb{R}^d)^n,
\tag{4.12}
\]

whose action is to project the momentum orthogonally to the submanifold \( \mathcal{M} \). We next define the RATTLE scheme, which is a second order discretization of the Hamiltonian part of (4.11).

A particular feature of this numerical integrator is to be reversible up to momentum reversal, at least for sufficiently small time steps (which is crucial in order to simplify the Metropolis–Hastings selection rule, see step (4) in Algorithm 4.5 below). We assume in Algorithm 4.4 that this condition holds, and we refer to [26, Section VII.1.4] and [35] for more details.

**Algorithm 4.4 (RATTLE).** Starting from a configuration \((x_m, y_m) \) with \( x_m \in \mathcal{M} \) and \( \nabla \xi(x_m)^T y_m = 0 \),

1. \( y_{m+\frac{1}{2}} = y_m - \frac{\Delta t}{2} \nabla H_n(x_m); \)
2. \( x_{m+\frac{1}{2}} = x_m + \Delta t y_{m+\frac{1}{2}}; \)

3. \( \theta_{m+1} = \theta_m + \int_{t_m}^{t_{m+\frac{1}{2}}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
4. \( y_{m+1} = y_{m+\frac{1}{2}} - \frac{\Delta t}{2} \nabla H_n(x_{m+\frac{1}{2}}); \)
5. \( z_{m+1} = \Pi_{\mathcal{M}} \left( y_{m+1} \right); \)
6. \( \theta_{m+1} = \theta_{m+1} + \int_{t_{m+\frac{1}{2}}}^{t_{m+1}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
7. \( x_{m+1} = x_{m+\frac{1}{2}} + \Delta t z_{m+1}; \)
8. \( y_{m+1} = y_{m+\frac{1}{2}} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
9. \( \theta_{m+2} = \theta_{m+2} + \int_{t_{m+1}}^{t_{m+2}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
10. \( y_{m+2} = y_{m+1} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
11. \( z_{m+2} = \Pi_{\mathcal{M}} \left( y_{m+2} \right); \)
12. \( \theta_{m+2} = \theta_{m+2} + \int_{t_{m+2}}^{t_{m+1}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
13. \( x_{m+2} = x_{m+1} + \Delta t z_{m+2}; \)
14. \( y_{m+2} = y_{m+1} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
15. \( \theta_{m+3} = \theta_{m+3} + \int_{t_{m+2}}^{t_{m+3}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
16. \( y_{m+3} = y_{m+2} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
17. \( z_{m+3} = \Pi_{\mathcal{M}} \left( y_{m+3} \right); \)
18. \( \theta_{m+3} = \theta_{m+3} + \int_{t_{m+3}}^{t_{m+2}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
19. \( x_{m+3} = x_{m+2} + \Delta t z_{m+3}; \)
20. \( y_{m+3} = y_{m+2} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
21. \( \theta_{m+4} = \theta_{m+4} + \int_{t_{m+3}}^{t_{m+4}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
22. \( y_{m+4} = y_{m+3} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
23. \( z_{m+4} = \Pi_{\mathcal{M}} \left( y_{m+4} \right); \)
24. \( \theta_{m+4} = \theta_{m+4} + \int_{t_{m+4}}^{t_{m+3}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
25. \( x_{m+4} = x_{m+3} + \Delta t z_{m+4}; \)
26. \( y_{m+4} = y_{m+3} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
27. \( \theta_{m+5} = \theta_{m+5} + \int_{t_{m+4}}^{t_{m+5}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
28. \( y_{m+5} = y_{m+4} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
29. \( z_{m+5} = \Pi_{\mathcal{M}} \left( y_{m+5} \right); \)
30. \( \theta_{m+5} = \theta_{m+5} + \int_{t_{m+5}}^{t_{m+4}} \frac{\nabla \xi(x_s)}{\sqrt{2\gamma}} \, \text{d}s; \)
31. \( x_{m+5} = x_{m+4} + \Delta t z_{m+5}; \)
32. \( y_{m+5} = y_{m+4} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}); \)
(3) compute the Lagrange multiplier \( \theta_m \in \mathbb{R}^m \) associated with \( x_{m+\frac{1}{2}} \) to enforce the constraint, using Algorithm 4.6 below (if convergence has been reached);

(4) project as \( x_{m+1} = x_{m+\frac{1}{2}} + \nabla \xi(x_m)\theta_m \) and \( y_{m+\frac{1}{2}} = y_{m+\frac{1}{2}} + \nabla \xi(x_m)\theta_m / \Delta t \);

(5) \( y_{m+\frac{3}{4}} = y_{m+\frac{1}{2}} - \frac{\Delta t}{2} \nabla H_n(x_{m+1}) \);

(6) \( y_{m+1} = \Pi_{M^+} y_{m+\frac{3}{4}} \) where the projector \( \Pi_{M^+} \) is defined in (4.12).

Finally, return \( (x_{m+1}, y_{m+1}) \).

We can now present the algorithm used to sample the conditioned distribution by integrating the Ornstein–Uhlenbeck process in (4.11) projected orthogonally to the submanifold with \( \Pi_{M^+} \). Next, we evolve the configuration \( (x_m, y_m) \) with a RATTLE step, leading to \( (\hat{x}_{m+1}, \hat{y}_{m+1}) \). However, reversibility may be lost in the procedure for two reasons: either it is not possible to perform one step of RATTLE starting from \( (\hat{x}_{m+1}, -\hat{y}_{m+1}) \), or the image of \( (\hat{x}_{m+1}, -\hat{y}_{m+1}) \) differs from \( (x_m, -\hat{y}_m) \). In both cases, the RATTLE move is rejected, and the configuration is updated as \( (x_m, -\hat{y}_m) \) (mind the fact that momenta need to be reversed here). Finally, a Metropolis–Hastings acceptance rule corrects for the time step bias in the sampling. The full algorithm reads as follows [33].

**Algorithm 4.5** (Constrained HMC with reversibility check). Fix \( T > 0, \Delta t > 0, \gamma > 0, K_{\max} \geq 1 \), \( N_{\text{iter}} = \lceil T/\Delta t \rceil \) and choose an initial configuration \( (x_0, y_0) \) with \( x_0 \in M \) and \( \nabla \xi(x_0)^{T} y_0 = 0 \) (possibly obtained by projection). Set also thresholds \( \varepsilon_{\text{rev}}, \varepsilon_{N} > 0 \), and define \( \eta_{\Delta t} = e^{-\gamma \Delta t} \). For \( m = 0, \ldots, N_{\text{iter}} - 1 \), run the following steps:

1. resample the momenta as
   \[
   \tilde{y}_m = \Pi_{M^+} \left( \eta_{\Delta t} y_m + \sqrt{1 - \eta_{\Delta t}^2} G_m \right),
   \]
   where \( G_m \) are independent \( d_n \)-dimensional standard Gaussian random variables;

2. perform one RATTLE step with Algorithm 4.4 starting from \( (x_m, \tilde{y}_m) \), providing \( (\hat{x}_{m+1}, \hat{y}_{m+1}) \)
   if the Newton algorithm with \( K_{\max} \), \( \varepsilon_{N} \) has converged; otherwise set \( (x_{m+1}, y_{m+1}) = (x_m, -\hat{y}_m) \) and increment \( m \);

3. compute a RATTLE backward step from \( (\hat{x}_{m+1}, -\hat{y}_{m+1}) \), providing \( (x_{m+1}^{\text{rev}}, y_{m+1}^{\text{rev}}) \)
   if the Newton algorithm with \( K_{\max} \), \( \varepsilon_{N} \) has converged. If the Newton algorithm has not converged or if \( |x_m - x_{m+1}^{\text{rev}}| > \varepsilon_{\text{rev}}, \) reject the move by setting \( (x_{m+1}, y_{m+1}) = (x_m, -\hat{y}_m) \) and increment \( m \);

4. compute the Metropolis–Hastings ratio
   \[
   p_m = 1 \wedge \exp \left[ -\beta_n \left( H_n^\xi(\hat{x}_{m+1}) + \frac{|\hat{y}_{m+1}|^2}{2} - H_n^\xi(x_m) - \frac{|\hat{y}_m|^2}{2} \right) \right],
   \tag{4.13}
   \]
   and set
   \[
   (x_{m+1}, y_{m+1}) = \begin{cases} 
   (\hat{x}_{m+1}, \hat{y}_{m+1}) \text{ with probability } p_m, \\
   (x_m, -\hat{y}_m) \text{ with probability } 1 - p_m.
   \end{cases}
   \]

A particularity of our implementation with respect to [35] is that we run the dynamics with the Hamiltonian \( H_n \) while the Metropolis–Hastings ratio (4.13) (step (4) of Algorithm 4.5) is computed with the modified Hamiltonian \( H_n^\xi \). As pointed out in Remark 4.8 below, the modification induced by the correction term \( U_n \) in (4.9) is generally small. Therefore, considering \( H_n \) for the dynamics allows to avoid the computation of the Hessian of \( \xi \), while the selection rule corrects for this small error.

In order for our description to be complete, we define how to project the position on \( M \) (step (3) in Algorithm 4.4). We use a variant of Newton’s algorithm defined below.

**Algorithm 4.6** (Newton algorithm). Consider a tolerance threshold \( \varepsilon_{N} > 0 \) and a maximal number of steps \( K_{\max} \geq 1 \). Starting from an initial position \( x^0 \notin M \) and a Lagrange multiplier \( \theta^0 = 0 \in \mathbb{R}^m \), the projection procedure reads as follows: while \( k \leq K_{\max} \),
(1) compute \( M_k = \nabla \xi(x^0)^T \nabla \xi(x^k) \in \mathbb{R}^{m \times m} \);

(2) set \( \theta^{k+1} = \theta^k - M_k^{-1} \xi(x^k) \);

(3) define the new position \( x^{k+1} = x^k + \nabla \xi(x^0)^T \theta^{k+1} \);

(4) if \( \max(\|\theta^{k+1} - \theta^k\|, |\xi(x^k)|) \leq \varepsilon_N \), the algorithm has converged, else go back to step (1).

If the algorithm has converged in \( k \leq K_{\text{max}} \) steps, return the value \( \theta^k \) of the Lagrange multiplier.

We emphasize that a fixed direction \( \nabla \xi(x^0) \) is considered for projection, which is needed to preserve the reversibility property of the final algorithm \[33\]. The procedure works provided the matrix \( M_k \) defined in step (1) is indeed invertible at each step of the inner loop, and we refer to \[33\] for more details. We also mention that we can consider different stopping criteria for the Newton algorithm in step (4) in Algorithm \[4.5\], for instance by using the relative error \( |x^{k+1} - x^k| \).

We are now ready to use Algorithm \[4.5\] to sample from the constrained distribution, challenge the theoretical results of Sections \[2\] and \[3\] and explore conjectures.

**Remark 4.7** (Rejection sources). For a standard HMC scheme, rejection is only due to the Metropolis–Hastings selection (step (4) in Algorithm \[4.5\]). Here, rejection can be due to the following reasons:

- the Newton algorithm in Step (2) (forward move) has not converged;
- the Newton algorithm in Step (3) (backward or reversed move) has not converged;
- the reversibility check in Step (3) has failed;
- the Metropolis rule in Step (4) has rejected the step.

In any case, the first step resamples the momentum variable according to the Ornstein–Uhlenbeck process part in \[4.10\], and rejection comes with a reversal of momenta. Let us also mention that, when the ratio \[4.13\] is computed with \( H_n \) (up to an additive constant), the Metropolis rejection rate should scale as \( \Delta t^3 \). This will be the case in our situation when \( \xi \) and \( \varphi \) are linear since in this case the additional term \( U_n \) in \[4.9\] is constant. This rate of decay is confirmed by numerical simulations (see Figure \[2\]).

**Remark 4.8** (Correction term). Proposition \[4.6\] shows that the Hamiltonian of the system must be modified in order for the constrained dynamics \[4.11\] to sample from the probability distribution \[4.5\]. However, in Algorithm \[4.5\] we run the dynamics with \( H_n \) and perform the selection with \( H_{\xi} \). This is motivated by the following scaling argument. Consider

\[
\xi(x) = \frac{1}{n} \sum_{i=1}^{n} \varphi(x_i)
\]

for some real-valued smooth function \( \varphi \), which corresponds to the linear constraint situation described in Section \[3.2\]. In this case, the corrector term in \[4.9\] reads

\[
U_n(x) = -\frac{1}{2 \beta_n} \log \left( \sum_{i=1}^{n} |\nabla \varphi(x_i)| \right),
\]

up to an additive constant. This means that the correction term in \[4.19\] scales like \( O(\log(n)/n^2) \) when \( \beta_n = \beta n^2 \), whereas the remainder of the Hamiltonian is \( O(1) \). As a result, the correction is much smaller than the Hamiltonian energy \( H_n \), and we may neglect it in the dynamics. This allows to avoid computing the Hessian of the constraint \( \xi \) at the price of a small increase in the rejection rate.

4.2. Numerical results.

4.2.1. Linear statistics with linear constraint: the influence of confinement. Since one motivation for our work was to study the trace constraint with quadratic confinement, as detailed in Section \[2\] we first consider the model presented in Theorem \[2.1\] with \( d = 2 \), \( \beta_n = n^2 \) and

\[
\forall x \in \mathbb{R}^2, \quad \varphi(x) = c - x \cdot v, \quad v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.
\] (4.14)

We run Algorithm \[4.5\] setting \( n = 300 \), \( T = 10^6 \), \( \Delta t = 0.5 \), \( \gamma = 1 \) and \( \varepsilon_N = \varepsilon_{\text{rev}} = 10^{-12} \) with \( K_{\text{max}} = 20 \). In all the simulations in dimension 2, the initial configuration is drawn uniformly over \([-1,1]^2\). We first set \( V(x) = 2|x|^2 \) and \( c = 1 \), so that according to Theorem \[2.1\] the conditional
law of the empirical measure $\mu_n$ under $P_n$ with the constraint $\mu_n(\varphi) = 0$ should converge in the limit of large $n$ towards a disk of radius $1/\sqrt{2}$, centered at $(1, 0)$ in $\mathbb{R}^2$. The simulations presented in Figure 1 show a very good agreement with the expected result.

Figure 1. Study of the quadratic confinement for $n = 300$ without constraint (left) and with the constraint (4.14) (right). We see that the constrained measure is a disk of radius $1/\sqrt{2}$ centered at $(1, 0)$.

In this simple case, the Hamiltonian in (4.9) is only modified by a constant, so we expect the Metropolis rejection rate (step (5) in Algorithm 4.5) to scale like $O(\Delta t^3)$ when $\Delta t \to 0$. In Figure 2, we plot this rate in log-log coordinates (setting here $n = 20$ to reduce the computation time). The slope is indeed close to 3, which confirms our expectation.

Figure 2. Study of the rejection rate of the Metropolis–Hastings selection rule with $n = 20$ (step (5) in Algorithm 4.5) in log-log coordinate. The slope of the linear fit is about 2.9.

In order to show that the translation phenomenon is specific to the quadratic confinement, we first consider the case of a quartic confinement potential, namely $V(x) = |x|^4/2$ subject to the constraint (4.14) with $c = 0.5$. This choice for $V$ together with $\varphi$ defined in (4.14) satisfies Assumption 3.6 so that Theorem 3.7 applies. However, no analytic solution is a priori available because the rotational symmetry is lost. The unconstrained equilibrium measure in Figure 3 (left) shows a depletion of the density around $(0, 0)$. In Figure 3 (right), we observe that the shape of the distribution is significantly modified by the constraint, and does not possess any rotational invariance. As could have been expected, the particles close to the origin feel a weaker confinement, so the distribution is more concentrated near the outer edge.

Another interesting case is when the confinement is weaker than quadratic, e.g. $V(x) = \frac{4}{3}|x|^\frac{3}{2}$, for which Theorem 3.7 still applies. The results are shown in Figure 4, considering again the
Figure 3. Study of the quartic confinement for $n = 300$ without constraint (left) and with the constraint \((4.14)\) (right). The shape of the equilibrium measure is significantly distorted by the constraint.

constraint \((4.14)\) with $c = 0.5$. We observe that the shape of the distribution also significantly changes by spreading in the direction of the constraint. This can be interpreted as follows: since the confinement is stronger at the origin, the more likely way to observe a fluctuation of the barycenter (or less costly in terms of energy) is in this case to spread the distribution.

Quite interestingly, for both potentials the distribution obtained as $c \to +\infty$ seems to reach a limiting ellipsoidal shape, under an appropriate rescaling (figures not shown here). Studying more precisely these limiting shapes and the rate at which they appear is an interesting open problem.

Figure 4. Study of the weak confinement for $n = 300$ without constraint (left) and with the constraint \((4.14)\) (right). The constraint now spreads the equilibrium measure to the right.

4.2.2. Other constraints in dimension two. In order to illustrate the efficiency of our algorithm in situations richer than the linear constraint with a linear function $\varphi$, we now present two other cases. First, we keep a linear constraint with $V(x) = 2|x|^2$, but set

$$\forall x \in \mathbb{R}^2, \quad \varphi(x) = c - \frac{\cos(5x_1) + \cos(5x_2)}{2},$$

where $x_1$ and $x_2$ denote here the first and second coordinates of $x \in \mathbb{R}^2$. This choice is motivated by Remark \([3.11]\) since the Laplacian of $\varphi$ takes positive and negative values, we expect the particles to concentrate in some regions of $\mathbb{R}^2$, possibly leading to a phase separation. Note also that, in order for the two last conditions in Assumption \([3.6]\) to be satisfied, we need to choose $c \in (-1, 1)$. We set again $n = 300$ but $\Delta t = 0.4$ to reduce the rejection rate. The other parameters are the same as in Section \([4.2.1]\). We plot in Figure 5 the result of the simulation for $c = 0.2$ and $c = 0.5$. 

The particles concentrate in the regions where the cosines are higher, which seems to lead to a phase separation for the second value of $c$.

Figure 5. Study of the cosine constraint for $n = 300$ with $c = 0.2$ (left) and $c = 0.5$ (right). A phase separation appears as the particles are constrained to stay in the local maxima of the cosines.

In order to illustrate the results of Section 3.3 we consider a quadratic constraint $\xi : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$ of the form

$$\forall x \in (\mathbb{R}^d)^n, \quad \xi(x) = \frac{1}{n^2} \sum_{i,j=1}^n \psi(x_i,x_j),$$

(4.15)

with, for $x,y \in \mathbb{R}^2$,

$$\psi(x,y) = \phi(x-y), \quad \text{and} \quad \phi(x) = c - |x|.$$  

(4.16)

A motivation for this choice is to modify the rigidity of the gas by constraining the particles to be closer or further one from another in average. In order to make this rigidity anisotropic, we also consider

$$\forall x \in \mathbb{R}^2, \quad \phi(x) = c - |x_1|.$$  

(4.17)

The choice (4.17) modifies the rigidity only in one direction. For illustration we take $V(x) = |x|^4/2$, $n = 50$, $\Delta t = 0.5$, $T = 10^6$ (we take a lower number of particles because the constraint makes the dynamics quite stiff). We set $c = 1$ for (4.16) and $c = 0.5$ for (4.17), which forces the particles to move away from each other. These choices for $\psi$ satisfy the conditions of Proposition 3.13, and the application $Q$ defined in (3.14) can be proved to be convex, so Assumption 3.12 is satisfied and Theorem 3.14 applies. The distribution obtained for the constraint (4.16), presented in Figure 6 (left), shows that the more likely way for the particles to be repelled by the constraint induced by $\psi$ is to move away from the center and concentrate on the edge, compared to Figure 3 (left). For the constraint (4.17), we clearly observe in Figure 6 (right) the effect of anisotropy.

4.2.3. A one dimensional example. We consider the Gaussian Unitary Ensemble (GUE), which is a degenerate two-dimensional Coulomb gas for which the particles are confined on the real axis. Its corresponds in a sense to (4.17) with $d = 1$, $V(x) = 2|x|^2$ but $g(x) = -\log |x|$, and $\beta_n = n^2$. It is known that the equilibrium measure is then the Wigner semi-circle law, and we refer for instance to [5] for a large deviations study. We can apply Theorem 3.1 for the linear constraint (4.14). In this case, the Wigner semi-circle law is indeed translated by a factor $c$ (figure not shown here).

Next, in order to illustrate a case which is not covered by our analysis, we want to sample the spectrum of those matrices whose determinant is equal to $\pm 1$. In our context, this corresponds to the configurations $x \in (\mathbb{R}^d)^n$ with $\prod_{i=1}^n |x_i| = 1$. By taking the logarithm, this constraint is actually of the form (4.17) (by Remark 4.2 the conditioned probability measure (4.1) does not depend the parametrization) with

$$\forall x \in (\mathbb{R}^d)^n, \quad \xi(x) = \frac{1}{n} \sum_{i=1}^n \log |x_i|.$$  

(4.18)
Figure 6. Study of the quartic confinement for \( n = 50 \) with the quadratic statistics constraint \((4.15)-(4.16)\) where \( c = 1 \) (left), and with the constraint \((4.17)\) with \( c = 0.5 \) (right). This has to be compared to the unconstrained distribution in Figure 3 (left).

We plot in Figure 7 the distribution for \( n = 300, \ T = 10^5 \) and \( \Delta t = 0.05 \) for the unconstrained log-gas, and with the constraint \((4.18)\) for \( \Delta t = 0.01 \) (starting with particles equally spaced over the interval \([-1, 1]\)). We observe what looks like a symmetrized Marchenko–Pastur distribution. Actually Remark 3.11 suggests that the effective potential of the constrained distribution is \(|\cdot|^2 - \alpha \log |\cdot|\) for some \( \alpha > 0 \), which is not that far from the Laguerre potential \(|\cdot| - \alpha \log |\cdot|\).

Figure 7. Study of the one dimensional log-gas for \( n = 300 \) without constraint (top) and with the constraint \((4.14)\) (bottom). This corresponds to a deformation of the semi-circle distribution.
The PhD of Grégoire Ferré is supported by the Labex Bézout ANR-10-LABX-58-01. The work of Gabriel Stoltz was funded in part by the Agence Nationale de la Recherche, under grant ANR-14-CE23-0012 (COSMOS). Gabriel Stoltz is supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement number 614492, and also benefited from the scientific environment of the Laboratoire International Associé between the Centre National de la Recherche Scientifique and the University of Illinois at Urbana-Champaign.

Acknowledgements

This section is devoted to the proof of Theorem 2.1. The following lemma is some sort of quantitative Wasserstein version of [8] Lemma C.1.

**Lemma A.1** (Translation). Let $d \geq 1$, $p \geq 1$, $\mu_1, \mu_2 \in \mathcal{P}_p(\mathbb{R}^d)$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then for all $m_1, m_2 \in \mathbb{R}^d$,

$$d_{W_p}(\mu_1 \ast \delta_{m_1}, \mu_2 \ast \delta_{m_2}) \leq 2^{p-1}|m_1 - m_2|^p + 2^{p-1}d_{W_p}(\mu_1, \mu_2).$$

Moreover, if $m_i = a(\int \varphi \, d\mu_i)$ for $i = 1, 2$ and $a \in \mathbb{R}^d$, then

$$d_{W_p}(\mu_1 \ast \delta_{m_1}, \mu_2 \ast \delta_{m_2}) \leq 2 \frac{a^p}{p} (1 + \|a\| \|\varphi\|_{Lip_p}^p) \frac{1}{p} d_{W_p}(\mu_1, \mu_2).$$

Note that the right hand side is infinite if $\varphi$ is not Lipschitz.

**Proof.** We have, by using the infimum formulation of the distance $d_{W_p}$ (see (1.16)),

$$d_{W_p}(\mu_1 \ast \delta_{m_1}, \mu_2 \ast \delta_{m_2}) \leq 2^{p-1}|m_1 - m_2|^p + 2^{p-1}d_{W_p}(\mu_1, \mu_2),$$

where we used the convexity inequality $|u + v|^p \leq 2^{p-1}(|u|^p + |v|^p)$ valid for all $u, v \in \mathbb{R}^d$. Then, since $p \mapsto d_{W_p}$ is monotonically non-increasing for $p \geq 1$, it holds

$$|m_1 - m_2| = \left| a \left( \int_{\mathbb{R}^d} \varphi \, d(\mu_1 - \mu_2) \right) \right| \leq \|a\| \|\varphi\|_{Lip_p} d_{W_p}(\mu_1, \mu_2),$$

which is the claimed estimate. 

The following lemma is a $d$-dimensional version of the factorization lemma in [15]. It expresses a non obvious independence between the center of mass and the shape of the cloud of particles distributed according to $P_n$. As noticed in [15], it reminds the structure of certain continuous spins systems such as in [10] [41].

**Lemma A.2** (Factorization). Suppose that the assumptions of Theorem 2.1 are satisfied, and define $u = (v, \ldots, v) \in (\mathbb{R}^d)^n$. Let $\pi$ and $\pi^\perp$ be the orthogonal projections in $(\mathbb{R}^d)^n$ on the linear subspaces

$$L = Ru \quad \text{and} \quad L^\perp = \{ x \in (\mathbb{R}^d)^n : x \cdot u = 0 \}.$$ 

Then, abriding $X_n$ into $X$, the following properties hold:

- for all $x \in (\mathbb{R}^d)^n$, denoting $s(x) = \frac{2x + x_n}{n} \in \mathbb{R}^d$, we have

  $$\pi(x) = (s(x) \cdot v) u = \left( (s(x) \cdot v) v, \ldots, (s(x) \cdot v) v \right),$$

  $$\pi^\perp(x) = x - \pi(x) = x - (s(x) \cdot v) u = (x_1 - (s(x) \cdot v) v, \ldots, x_n - (s(x) \cdot v) v);$$

- $\pi(X)$ and $\pi^\perp(X)$ are independent random vectors;

- $\pi(X)$ is Gaussian with law $\mathcal{N} \left( 0, \frac{\nu^2}{2^{n-1}} \right)$, $u$, so that $s(X) \cdot v$ has law $\mathcal{N} \left( 0, \frac{\nu^2}{2^{n-1}} \right)$;

- $\pi^\perp(X)$ has law of density proportional to $x \in L^\perp \mapsto e^{-\frac{\nu^2}{2^{n-1}} s_n(x)}$ with respect to the trace of the Lebesgue measure on the linear subspace $L^\perp$ of $\mathbb{R}^{dn-1}$.

**Proof of Lemma A.2.** Since $|v| = 1$, we have $|u| = \sqrt{n}$, so the orthonormal projection on $L$ reads, for $x \in (\mathbb{R}^d)^n$,

$$\pi(x) = \frac{x \cdot u}{|u|^2} u = \left( \frac{1}{n} \sum_{i=1}^{n} x_i \cdot v \right) u = (s(x) \cdot v) u.$$
The expression of \( \pi^\perp \) follows easily. For all \( x \in (\mathbb{R}^d)^n \), from \( x = \pi(x) + \pi^\perp(x) \) we get
\[
|x|^2 = |\pi(x)|^2 + |\pi^\perp(x)|^2.
\]
On the other hand, for all \( i, j \in \{1, \ldots, n\} \) it holds
\[
\begin{align*}
x_i - x_j = \pi(x)_i + \pi^\perp(x)_i - \pi(x)_j - \pi^\perp(x)_j \\
&= s(x) + \pi^\perp(x)_i - s(x) - \pi^\perp(x)_j \\
&= \pi^\perp(x)_i - \pi^\perp(x)_j.
\end{align*}
\]
Since \( V(x) = |x|^2 \), it follows that, for all \( x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n \),
\[
H_n(x) = \frac{1}{n} |x|^2 + \frac{1}{n^2} \sum_{i \neq j} g(x_i - x_j) = \frac{1}{n} |\pi(x)|^2 + H_n(\pi^\perp(x)).
\]
Now, let \( u_1, \ldots, u_d \) be an orthogonal basis of \( (\mathbb{R}^d)^n = \mathbb{R}^d \) with \( u_1 = u/\sqrt{n} \in L \). For all \( x \in (\mathbb{R}^d)^n \) we write \( x = \sum_{i=1}^d t_i(u_i) \). We have \( \pi(x) = t_1(x)u_1 = (s(x) \cdot v)u \) and \( \pi^\perp(x) = \sum_{i=2}^d t_i(x)u_i \). Then we have, for all bounded measurable \( f : L \to \mathbb{R} \) and \( g : L^\perp \to \mathbb{R} \),
\[
\begin{align*}
\mathbb{E}[f(\pi(X))g(\pi^\perp(X))] &= Z^{-1} \int_{(\mathbb{R}^d)^n} f(\pi(x))g(\pi^\perp(x))e^{-\beta_n H_n(\pi^\perp(x))} dx_1 \cdots dx_n \\
&= Z^{-1} \left( \int_{\mathbb{R}} f(t')e^{-\beta_n H_n(t')} dt' \right) \left( \int_{\mathbb{R}^d} g(t''e^{-\beta_n H_n(t'')} dt'' \right),
\end{align*}
\]
where \( t' = t_1u_1, dt' = dt_1, t'' = \sum_{i=2}^d t_iu_i \) and \( dt'' = \prod_{i=2}^d dt_i \). This concludes the proof of the last two points of the lemma. \( \square \)

We can now turn to the proof of Theorem 2.1.

**Proof of Theorem 2.1.** Thanks to Lemma \ref{lem:approx}, we have, denoting again by \( u = (v_1, \ldots, v) \in (\mathbb{R}^d)^n \),
\[
\text{Law} \left( X_n \left| \frac{1}{n} \sum_{i=1}^n \varphi(X_i) = 0 \right. \right) = \text{Law} \left( X_n \left| \frac{1}{n} \sum_{i=1}^n X_{n,i} \cdot v = c \right. \right)
\]
\[
= \text{Law} \left( X_n \left| s(X_n) \cdot v = c \right. \right)
\]
\[
= \text{Law} \left( X_n \left| (s(X_n) \cdot v)u = cu \right. \right)
\]
\[
= \text{Law} \left( X_n \left| \pi(X_n) = cu \right. \right)
\]
\[
= \text{Law} \left( \tilde{X}_n \right),
\]
where \( \tilde{X}_n = cu + \pi^\perp(X_n) = cu + X_n - \pi(X_n) \). We also have
\[
\tilde{X}_n = \left( (c - s(X_n) \cdot v) + X_{n,1} + \cdots + (c - s(X_n) \cdot v) + X_{n,n} \right)
\]
where \( s(X_n) = \frac{X_{n,1} + \cdots + X_{n,n}}{n} \). In other words (recall that \( \varphi(x) = x \cdot v - c \))
\[
\tilde{X}_n = \left( X_{n,1} - \frac{1}{n} \sum_{i=1}^n \varphi(X_{n,i})v, \ldots, X_{n,n} - \frac{1}{n} \sum_{i=1}^n \varphi(X_{n,i})v \right),
\]
so that
\[
\tilde{\mu}_n = \frac{1}{n} \sum_{i=1}^n \delta_{\tilde{X}_{n,i}} = \mu_n * \delta_m \quad \text{where} \quad \mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_{n,i}} \quad \text{and} \quad m_n = v \int \varphi d\mu_n.
\]
Thanks to the assumptions on \( V \) and \( g \) we know that the equilibrium measure \( \mu_* \) is the uniform distribution on a ball of radius \( 1 \). Now we note that \( \|\varphi\|_{\text{Lip}} \leq 1 \) and \( \int \varphi d\mu_* = c \), so that by Lemma \ref{lem:Wasserstrom} denoting by \( \mu^p = \delta_{c_p} * \mu \), for all \( p \geq 1 \), there exists \( c_p > 0 \) with
\[
d_{W_p}(\tilde{\mu}_n, \mu^p) \leq c_p d_{W_p}(\mu_n, \mu_*). \quad (A.2)
\]
On the other hand, the large deviations principle (see [12] Proof of Theorem 1.1(4)) and [19] for the fact that the condition (1.8) ensures that $(\beta_n)_{n}$ diverges fast enough, gives, for any $\varepsilon > 0$,
\[ \sum_n P(\text{d}_{\text{BL}}(\mu_n, \mu_*) \geq \varepsilon) < \infty, \]  
(A.3)
(alternatively we could use the concentration of measure [13] Theorem 1.5] and get the result for $d_{W_1}$ as well). This summable convergence in probability towards a non-random limit, known as complete convergence [51], is equivalent, via Borel–Cantelli lemmas, to stating that almost surely,\[ \lim_{n \to \infty} d_{\text{BL}}(\mu_n, \mu_*) = 0, \] regardless of the way we defined the random variables $X_n$ and thus the random measures $\mu_n$ on the same probability space.

In order to upgrade the convergence from $d_{\text{BL}}$ to $d_{W_p}$ for all $p \geq 1$, we note that, from [13] Theorem 1.12], there exists $r_0 > 0$ such that, for all $r \geq r_0$,
\[ \sum_n P\left( \max_{1 \leq k \leq n} |X_{n,k}| \geq r \right) < \infty. \]  
(A.4)
Now for all $p \geq 1$ and all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ supported in the ball of $\mathbb{R}^d$ of radius $r \geq 1$, we have\[ d_{W_p}^{\mu}(\mu, \nu) \leq (2r)^{p-1} d_{W_1}(\mu, \nu) \leq r(2r)^{p-1} d_{\text{BL}}(\mu, \nu). \]
Also, by combining (A.3) and (A.4), we obtain that for all $p \geq 1$ and all $\varepsilon > 0$,
\[ \sum_n P(\text{d}_{W_p}(\mu_n, \mu_*) \geq \varepsilon) < \infty. \]  
(A.5)
By the Borel–Cantelli lemma, for all $p \geq 1$, almost surely,\[ \lim_{n \to \infty} d_{W_p}(\mu_n, \mu_*) = 0, \] regardless of the way we define the random variables $X_n$ on the same probability space. Finally, since $p \mapsto d_{W_p}$ is monotonic in $p$, we can make the almost sure event valid for all $p$ by taking the intersection of all the almost sure events obtained for integer values of $p$.

By combining (A.5) with (A.2), we obtain that for all $p \geq 1$ and $\varepsilon > 0$,
\[ \sum_n P(\text{d}_{W_p}(\mu_n, \mu^*) \geq \varepsilon) < \infty. \]  
(A.6)
Now if $Y_n$ is a random vector of $(\mathbb{R}^d)^n$ such that \( \text{Law}(Y_n) = \text{Law}(X_n | \varphi(X_{n,1}) + \cdots + \varphi(X_{n,n}) = 0) \) then, denoting by $\mu^Y_n = \frac{1}{n} \sum_{i=1}^n \delta_{Y_{n,i}}$, using (A.1) and the fact that $\mu^Y$ is deterministic, we get\[ d_{W_p}(\mu^Y_n, \mu^*) \overset{d}{=} d_{W_p}(\tilde{\mu}_n, \mu^*). \]
Therefore, from (A.6) we get, for all $p \geq 1$ and all $\varepsilon > 0$,
\[ \sum_n P(\text{d}_{W_p}(\mu^Y_n, \mu^*) \geq \varepsilon) < \infty. \]
By the Borel–Cantelli lemma, for all $p \geq 1$, almost surely,\[ \lim_{n \to \infty} d_{W_p}(\mu^Y_n, \mu^*) = 0, \] regardless of the way we define the random variables $Y_n$ on the same probability space. Finally, since $d_{W_p}$ is monotonic in $p$, we can make the almost sure event valid for all $p$ by taking the intersection of all the almost sure events obtained for integer values of $p$.

Note that the above proof relies crucially, via Lemma A.2, on the quadratic nature of $V$. However the Coulomb nature of $g$ is less crucial and the result should remain essentially valid provided that the convergence to the equilibrium measure holds, for instance at the level of generality of the assumptions of the large deviations principle in [12].

**Appendix B. Proofs of Section 3.1**

We start with the proof of the abstract Gibbs conditioning principle.

**Proof of Proposition 3.1** The set $\mathcal{J}_B$ defined in (3.2) is not empty because the infimum is finite by (3.1), $B$ is closed and $I$ has compact level sets (which also implies that $I$ is lower semicontinuous), so the infimum is attained at least for one measure. Moreover, $\mathcal{J}_B$ is closed by lower semicontinuity of $I$. Now, since
\[ \frac{1}{\beta_n} \log P \left( Z_n \in A_{\varepsilon} \mid Z_n \in B \right) = \frac{1}{\beta_n} \log P \left( Z_n \in A_{\varepsilon} \cap B \right) - \frac{1}{\beta_n} \log P \left( Z_n \in B \right), \]
the result follows from an upper bound on $\mathbb{P}(A_\varepsilon \cap B)$ and a lower bound on $\mathbb{P}(B)$. The upper bound of the large deviations principle implies that

$$\limsup_{n \to +\infty} \frac{1}{\beta_n} \log \mathbb{P}(Z_n \in A_\varepsilon \cap B) \leq - \inf_{A_\varepsilon \cap B} I. \quad (B.1)$$

Assume first that $A_\varepsilon \cap B \neq \emptyset$. Since $A_\varepsilon = \{ z \in \mathcal{Z}, d(z, \mathcal{F}_B) > \varepsilon \}$, the lower semi-continuity of $I$ shows that (see [12, Section 2.5]) there exists $c_\varepsilon > 0$ for which

$$\inf_{A_\varepsilon \cap B} I \geq c_\varepsilon + \inf_B I,$$

so that

$$\limsup_{n \to +\infty} \frac{1}{\beta_n} \log \mathbb{P}(Z_n \in A_\varepsilon \cap B) \leq - \inf_B I - c_\varepsilon. \quad (B.2)$$

If $A_\varepsilon \cap B = \emptyset$, the infimum in the right hand side of (B.1) is equal to $+\infty$ so that (B.2) still holds. The lower bound for the set $B$ reads

$$\liminf_{n \to +\infty} \frac{1}{\beta_n} \log \mathbb{P}(Z_n \in B) \geq - \inf_B I.$$

Since $B$ satisfies (3.1), it holds

$$\limsup_{n \to +\infty} \frac{1}{\beta_n} \log \mathbb{P}(Z_n \in B) \leq \inf_B I,$$

which, together with (B.2), leads to (3.3).

Finally, if $Z_n^\varepsilon \sim \text{Law}(Z_n \mid Z_n \in B)$ and if we define the $Z_n^\varepsilon$’s on the same probability space then, for all $\varepsilon > 0$, by the Borel–Cantelli lemma, $\sum_n \mathbb{P}(Z_n^\varepsilon \in A_\varepsilon) < \infty$ and thus, almost surely, $Z_n^\varepsilon \not\in A_\varepsilon$ for large enough $n$. Since the set $A_\varepsilon$ depends on $\varepsilon > 0$, by taking $\varepsilon \to 0$ with $\varepsilon \in \mathbb{Q}$, we obtain that almost surely, $\lim_{n \to +\infty} d(Z_n^\varepsilon, \mathcal{F}_B) = 0$. \hfill $\Box$

We next recall elements of proof for the properties of the rate function $\mathcal{E}$.

**Proof of Proposition 3.3** Consider a probability measure $\mu \in D_\mathcal{E}$, so

$$\int_{\mathbb{R}^d} V(x) \mu(dx) + \int_{\mathbb{R}^d \times \mathbb{R}^d} g(x-y) \mu(dx) \mu(dy) < +\infty.$$ 

Since $V$ satisfies Assumption 3.2 and therefore beats $g$ at infinity (in particular when $d = 2$), we have

$$\int_{\mathbb{R}^d} |x|^p \mu(dx) < +\infty,$$

for $1 < p < q$. Thus $D_\mathcal{E} \subset \mathcal{P}_p(\mathbb{R}^d)$.

In order to show the convexity of $\mathcal{E}$ it is sufficient to study that of $\mu \mapsto J(\mu)$ as defined in (1.18). When $d \geq 3$, this is a consequence of [12] Lemma 3.1. In the case $d = 2$, convexity over $D_\mathcal{E}$ is also shown in [12, Section 3], Note that convexity is in general due to a Bochner-type positivity of the interaction kernel.

Finally, the existence of a unique minimizer with compact support solving (3.5) follows from [12, Theorem 1.2] for $d \geq 3$ and [13, Chapter I, Theorem 1.3] for $d = 2$. Since the minimizer of $\mathcal{E}$ over $\mathcal{P}(\mathbb{R}^d)$ has compact support, the three problems in (3.4) clearly coincide. \hfill $\Box$

We finally present the proof of Theorem 3.5 which is a consequence of Proposition 3.1 and the Borel–Cantelli lemma.

**Proof of Theorem 3.5** Under $P_n$, the empirical measure $\mu_n$ associated to $X_n$ satisfies a LDP in the $p$-Wasserstein topology with good rate function $\mathcal{E}$. Since $B$ is assumed to be a closed continuity set for the $p$-Wasserstein topology, the set $\mathcal{E}_B$ defined in (3.7) is closed and non-empty by Proposition 3.1.

For simplicity we denote by

$$\mu_n^V = \frac{1}{n} \sum_{i=1}^n \delta_{Y_n,i}.$$
the empirical measure associated to $Y_n$, where $Y_n \sim \text{Law}(X_n \mid \mu_0 \in B)$. For any $\varepsilon > 0$, we define the set $A_\varepsilon$ as in Proposition 3.1. Then, there exists $c_\varepsilon > 0$ such that
\[
\sum_n P(d_{W_p}(\mu_n^Y, \mathcal{E}_B) > \varepsilon) = \sum_n P(d_{W_p}(\mu_n, \mathcal{E}_B) > \varepsilon \mid \mu_n \in B)
\]
\[
= \sum_n P(\mu_n \in A_\varepsilon \mid \mu_n \in B)
\]
\[
\leq C \sum_n e^{-\beta_n c_\varepsilon} < +\infty,
\]
for some $C > 0$. Since $\beta_n \gg n$, the Borell–Cantelli lemma implies that
\[
\lim_{n \to \infty} d_{W_p}(\mu_n^Y, \mathcal{E}_B) = 0,
\]
almost surely in any probability space, which concludes the proof. □

Appendix C. Proof of Theorem 3.7

The proof is decomposed into four steps. We first show that under Assumption 3.6, the set $B$ is an $I$-continuity set for the electrostatic energy $\mathcal{E}$. We next show that any minimizer of $\mathcal{E}$ over $B$ has a compact support, and hence the minimizer is actually unique. The last two steps characterize the minimizer through (3.10).

Step 1: $I$-continuity. Let us first show that $B$ is closed for the $p$-Wasserstein topology by showing that $B^c = \{\mu \in \mathcal{P}_p(\mathbb{R}^d) \mid \mu(\varphi) = 0\}$ is open. Take $\mu \in B^c$ and $\nu$ such that $d_{W_p}(\mu, \nu) \leq \frac{\varepsilon}{\bar{\beta}_p}$ for some $\varepsilon > 0$. By definition of the $p$-Wasserstein distance it holds
\[
\sup_{f \in L^1(\mu), g \in L^1(\nu)} \left( \int_{\mathbb{R}^d} f \, d\mu - \int_{\mathbb{R}^d} g \, d\nu \right) \leq \varepsilon.
\]
(C.1)

Since $\|\varphi\|_{\infty, p} < +\infty$, for any $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ it holds $\varphi \in L^1(\nu) \cap L^1(\mu)$. Moreover, $\|\varphi\|_{\text{Lip}} < +\infty$ and $\varphi$ cannot be a constant function because this would contradict the existence of $\mu_{\pm}$ in Assumption 3.6 so $\|\varphi\|_{\text{Lip}} > 0$. As a result, for $|x - y| \geq 1$ we have
\[
|\varphi(x) - \varphi(y)| \leq \|\varphi\|_{\text{Lip}}|x - y| \leq \|\varphi\|_{\text{Lip}}|x - y|^p.
\]

Therefore, $\varphi/\|\varphi\|_{\text{Lip}}$ satisfies the inf-convolution condition in (C.1) and we may pick $f = g = \varphi/\|\varphi\|_{\text{Lip}}$ so that
\[
\int_{\mathbb{R}^d} \frac{\varphi}{\|\varphi\|_{\text{Lip}}} \, d\mu - \int_{\mathbb{R}^d} \frac{\varphi}{\|\varphi\|_{\text{Lip}}} \, d\nu \leq \varepsilon,
\]
which becomes
\[
\nu(\varphi) \geq \mu(\varphi) - \varepsilon\|\varphi\|_{\text{Lip}} > 0 \quad \text{for} \quad \varepsilon < \frac{\mu(\varphi)}{\|\varphi\|_{\text{Lip}}}.
\]
As a result, $B^c$ is open and $B$ is closed for the $p$-Wasserstein topology.

We now prove that $B$ is an $I$-continuity set, namely that (3.9) holds. By the same reasoning as above, the existence of $\mu_- \in D_\mathcal{E}$ such that $\mu_-(\varphi) < 0$ ensures that $\mu_- \in \hat{B}$ so
\[
\inf_{\hat{B}} \mathcal{E} < +\infty.
\]
Since $B$ is closed and $\mathcal{E}$ has compact level sets (in particular it is lower semicontinuous), there exists $\bar{\mu}$ such that
\[
\mathcal{E}(\bar{\mu}) = \inf_{\hat{B}} \mathcal{E}.
\]
If $\bar{\mu}(\varphi) < 0$, it holds $\bar{\mu} \in \hat{B}$ and the proof is complete. Thus we may assume that $\bar{\mu}(\varphi) = 0$ and, by considering a minimizing sequence, for any $\varepsilon > 0$ we may find $\mu_\varepsilon \in \hat{B}$ such that
\[
\mathcal{E}(\mu_\varepsilon) \leq \inf_{\hat{B}} \mathcal{E} + \varepsilon.
\]
(C.2)

For $t \in [0, 1]$, we introduce $\mu_t = t\mu_\varepsilon + (1 - t)\bar{\mu} \in D_\mathcal{E}$. Since $\bar{\mu}(\varphi) = 0$ it holds $\mu_t \in \hat{B}$ for any $t \in (0, 1]$. By convexity of $\mathcal{E}$ on its domain we have
\[
\mathcal{E}(\mu_t) \leq t\mathcal{E}(\mu_\varepsilon) + (1 - t)\mathcal{E}(\bar{\mu}).
\]
We now proceed by contradiction by assuming that $\mathcal{E}(\tilde{\mu}) = \inf_{\tilde{B}} \mathcal{E} - \eta$ for some $\eta > 0$. Recalling (C.2) we have, for some $\varepsilon > 0$ and any $t \in (0, 1]$,\[ \mathcal{E}(\mu_t) \leq t \left( \inf_{\tilde{B}} \mathcal{E} + \varepsilon \right) + (1 - t) \left( \inf_{\tilde{B}} \mathcal{E} - \eta \right) = \inf_{\tilde{B}} \mathcal{E} + t\varepsilon - (1 - t)\eta. \]

Considering\[ t < \frac{\eta}{\varepsilon + \eta}, \]

we obtain that $\mu_t \in \tilde{B}$ with\[ \mathcal{E}(\mu_t) < \inf_{\tilde{B}} \mathcal{E}, \]

which is a contradiction. Therefore, (3.9) holds true.

**Step 2: the minimizer is unique and has compact support.** We now show that any minimizer $\mu^*$ has a compact support, before turning to uniqueness. We detail the proof for $d \geq 3$ following [12] by highlighting the necessary modifications, and leave the proof for $d = 2$ to the reader (which is deduced from [15] Chapter I, Theorem 1.3). We introduce\[ \zeta = \inf_{\tilde{B}} \mathcal{E} \]

and, for any compact $K$,

\[ \zeta_K = \inf_{\tilde{B}_K} \mathcal{E}, \quad \text{where} \quad \tilde{B}_K = \{ \mu \in \tilde{B} \mid \text{supp}(\mu) \subset K \}. \]

By Assumption 3.6 $B_K$ is non empty for $K$ large enough (consider $\mu_-(\cdot \mathbf{1}_K)/\mu_-(K)$ for $K$ large enough). By Assumption 3.2 for any constant $C$ the set

\[ K = \{ x \in \mathbb{R}^d, V(x) \leq C \} \]

is compact. In all what follows, we assume that $V \geq 0$. Since $V$ is lower bounded and defined up to a constant, there is no loss of generality in this assumption.

Let us show that $\zeta = \zeta_K$ for $C$ large enough. Since the infimums on $B$ and $\tilde{B}$ coincide, we can consider a measure $\mu \in \tilde{B}$ such that $\mu(\varphi) < 0$ and $\mathcal{E}(\mu) \leq \zeta + 1$. If $\mu(K) = 1$, the measure has compact support and we are done, so we assume that $\mu(K) < 1$. The goal of the following computations is to build a measure $\mu_K \in \tilde{B}$ supported in $K$ such that $\mathcal{E}(\mu_K) < \mathcal{E}(\mu)$; this contradiction will show that $\zeta$ and $\zeta_K$ are equal. Let us first show that $\mu(K) > 0$ for $C$ large enough. Indeed,

\[ \zeta + 1 \geq \mathcal{E}(\mu) = \int_K V \, d\mu + \int_{K^c} V \, d\mu + J(\mu) \geq C(1 - \mu(K)), \]

which shows that $\mu(K) > 0$ if $C > \zeta + 1$. We may therefore define the restriction

\[ \mu_K(\cdot) = \frac{\mu(K \cap \cdot)}{\mu(K)}. \]

Since $\mu(K) < 1$, we define similarly $\mu_K^c$. The measure $\mu$ then reads

\[ \mu = \mu(K)\mu_K + (1 - \mu(K))\mu_K^c. \]

Moreover, we chose $\mu$ such that $\mu(\varphi) < 0$, so it holds $\mu_K \in B_K$ for $C$ large enough. Using the positivity of $V$ and $J$ (since $d \geq 3$ and so $g \geq 0$) and $\mu(K) < 1$, we obtain that

\[ \mathcal{E}(\mu) \geq \mu(K) \int_{\mathbb{R}^d} V \, d\mu_K + (1 - \mu(K)) \int_{\mathbb{R}^d} V \, d\mu_K^c + \mu(K)^2 J(\mu_K) \]

\[ \geq \mu(K)^2 J(\mu_K) + \mu(K)^2 \int_{\mathbb{R}^d} V \, d\mu_K + (1 - \mu(K))C \]

\[ \geq \mu(K)^2 \mathcal{E}(\mu_K) + (1 - \mu(K))C. \]

Let us proceed by contradiction by assuming that $\mathcal{E}(\mu_K) \geq \mathcal{E}(\mu)$, which leads to

\[ \mathcal{E}(\mu) \geq \mu(K)^2 \mathcal{E}(\mu) + (1 - \mu(K))C. \]

Since $\mathcal{E}(\mu) \leq \zeta + 1$ we obtain

\[ (\zeta + 1)(1 - \mu(K)^2) \geq (1 - \mu(K))C. \]
Simplifying by $1 - \mu(K)$ we have
\[ 2(\zeta + 1) \geq C, \]
which is absurd for $C > 2(\zeta + 1)$. Since $\mu_K \in B$ for $C$ large enough, this shows that $\zeta$ and $\zeta_K$ coincide and that any minimizer has compact support.

In the above proof, the only modification with respect to previous works (see for instance [12]) is to check that the restricted measure $\mu_K$ satisfies the constraint for $C$ large enough. This is done by picking the measure $\mu$ close to the minimum and such that $\mu(\varphi) < 0$. The same strategy can be used in the situation where $d = 2$ by writing
\[ E(\mu) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \left( \frac{V(x) + V(y)}{2} - \log |x - y| \right) \mu(dx)\mu(dy), \]
and adapting [15] Chapter I, Theorem 1.3 since $V(x) + V(y)$ dominate $\log |x - y|$ at infinity by Assumption 3.2.

In order to show that the minimizer is unique, it suffices to notice that for two probability measures $\mu, \nu$ with compact support, it holds $J(\nu - \mu) = 0$ if and only if $\mu = \nu$, see [10] Theorems 1.15 and 1.16.

**Step 3: Lagrange multiplier.** We now turn to a first step towards the expression of $\mu^\varphi$ involving a Lagrange multiplier $\alpha$. We adapt the proof of [1] Theorem 3.1 by introducing the following subset of $\mathbb{R}^2$:
\[ R = \{ (E(\mu) - E(\mu^\varphi) + a_0, \mu(\varphi) + a_1) : a_0 > 0, a_1 > 0, \mu \in D_E \}. \quad (C.3) \]
Since $E$ is convex on its domain $D_E$ (which is convex), $R$ is a non void convex subset of $\mathbb{R}^2$ that does not contain $(0, 0)$ (recall also that $D_E \subset P_\mu(\mathbb{R}^d)$ so the constraint takes finite values). Separating $R$ from $(0, 0)$ with a hyperplane (see [3] Corollary 1.41]), this ensures the existence of $(a_0, a_1) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ such that, for any $\mu \in D_E$ and $a_0, a_1 > 0$ it holds
\[ a_0(E(\mu) - E(\mu^\varphi) + a_0) + a_1(\mu(\varphi) + a_1) \geq 0. \]
By taking $a_1 \rightarrow -\mu^\varphi(\varphi) \geq 0$ and $\mu = \mu^\varphi$ in the above equation, we obtain that $a_0 \geq 0$. Then, choosing $\mu = \mu^\varphi$, $a_0 \rightarrow 0$ and $a_1 > -\mu^\varphi(\varphi) \geq 0$ we find $a_1 \geq 0$. Taking $a_0, a_1 \rightarrow 0$, we obtain
\[ \forall \mu \in D_E, \quad a_0E(\mu^\varphi) \leq a_0E(\mu) + a_1\mu(\varphi). \quad (C.4) \]
We now prove that $a_0 > 0$ by contradiction, using a kind of qualification of constraint argument. If $a_0 = 0$, (C.4) becomes
\[ \forall \mu \in D_E, \quad 0 \leq a_1\mu(\varphi). \]
Since $a_1 \neq 0$ in this case, the above equation contradicts Assumption 3.6 by taking $\mu = \mu_\varphi$, so $a_0 > 0$ and we may renormalize (C.4) into
\[ \forall \mu \in D_E, \quad E(\mu^\varphi) \leq E(\mu) + a\mu(\varphi), \quad (C.5) \]
where we set $\alpha = a_1/a_0 \geq 0$.

Finally, we show that either $\mu^\varphi = \mu_\varphi$, in which case $\alpha = 0$, or $\mu^\varphi(\varphi) = 0$ and $\alpha > 0$. First, if $\mu_\varphi \in B$, $\mu_\varphi$ satisfies the constraint and we know from Proposition 3.3 that it solves (3.10) with $\alpha = 0$. Otherwise, it holds $\mu_\varphi(\varphi) > 0$. Assume that $\mu^\varphi(\varphi) < 0$ and take $\mu = \mu^\varphi$ in (C.5), so
\[ a\mu^\varphi(\varphi) \geq 0, \]
which implies $\alpha = 0$. Therefore, (C.5) implies that $\mu^\varphi$ is the global minimizer $\mu_\varphi$, which is in contradiction with $\mu_\varphi(\varphi) = \mu^\varphi(\varphi) < 0$, so the minimizer actually saturates the constraint and $\alpha > 0$.

**Step 4: potential equation.** In order to derive the equation for $\mu^\varphi$, we follow [12] Section 4] by introducing the modified potential and electrostatic energy, for $\mu \in \mathcal{P}(\mathbb{R}^d)$,
\[ V_\alpha = V + \alpha\varphi, \quad E_\alpha(\mu) = \int_{\mathbb{R}^d} V_\alpha(x)\mu(dx) + J(\mu). \]
Since the case when $\alpha = 0$ corresponds to no-conditioning and we already know the equation satisfied by the equilibrium measure in this case, we restrict our attention to the situation in which $\alpha > 0$ and $\mu^\varphi(\varphi) = 0$. We define next, for any $\mu \in D_E$,
\[ \forall t \in (0, 1), \quad \psi(t) = E_\alpha(((1 - t)\mu^\varphi + t\mu). \]
Because of (C.5) and the convexity of $\psi$, it holds $\psi'(0) \geq 0$, so that
\[
0 \leq \psi'(0) = \int_{\mathbb{R}^d} V_\alpha \, d(\mu - \mu^\varphi) + 2J(\mu^\varphi, \mu - \mu^\varphi)
\]
\[
\leq \int_{\mathbb{R}^d} V_\alpha \, d\mu + 2J(\mu^\varphi, \mu) - \left( \int_{\mathbb{R}^d} V_\alpha \, d\mu^\varphi + 2J(\mu^\varphi) \right)
\]
\[
\leq \int_{\mathbb{R}^d} (V_\alpha + 2U_{\mu^\varphi}) \, d\mu - C_\varphi,
\]
where we set $U_{\mu^\varphi} = \mu^\varphi \ast g$ and
\[
C_\varphi = \int_{\mathbb{R}^d} V_\alpha \, d\mu^\varphi + 2J(\mu^\varphi) = \int_{\mathbb{R}^d} (V + 2U_{\mu^\varphi}) \, d\mu^\varphi,
\]
since $\mu^\varphi(x) = 0$. The above inequality may be rewritten as
\[
\forall \mu \in D_\mathcal{E}, \quad \int_{\mathbb{R}^d} (V_\alpha + 2U_{\mu^\varphi} - C_\varphi) \, d\mu \geq 0,
\]
which proves the second line of (3.10) (by definition of quasi-everywhere).

Let us now prove the first line in (3.10) by contradiction. Assume that there is $x \in \text{supp}(\mu^\varphi)$ such that $V_\alpha(x) + 2U_{\mu^\varphi}(x) < C_\varphi$. Since $\mu^\varphi$ has compact support, $U_{\mu^\varphi}$ is lower semi-continuous [30, page 59]. Since $V$ is lower semi-continuous and $\varphi$ is Lipschitz hence continuous, $V_\alpha + 2U_{\mu^\varphi}$ is lower semi-continuous. There exists therefore a neighborhood $\mathcal{U}$ of $x$ and $\varepsilon > 0$ such that
\[
\forall x \in \mathcal{U}, \quad V_\alpha(x) + 2U_{\mu^\varphi}(x) \geq C_\varphi + \varepsilon.
\]
Integrating with respect to $\mu^\varphi$ and using $\mu^\varphi(\varphi) = 0$ leads to
\[
C_\varphi = \int_{\mathbb{R}^d} (V_\alpha + 2U_{\mu^\varphi}) \, d\mu^\varphi = \int_{\mathcal{U}} (V + 2U_{\mu^\varphi}) \, d\mu^\varphi + \int_{\mathbb{R}^d \setminus \mathcal{U}} (V + 2U_{\mu^\varphi}) \, d\mu^\varphi
\]
\[
\geq (C_\varphi + \varepsilon)\mu^\varphi(\mathcal{U}) + \int_{\mathbb{R}^d \setminus \mathcal{U}} (V + 2U_{\mu^\varphi}) \, d\mu^\varphi.
\]
Since $V_\alpha + 2U_{\mu^\varphi} \geq C_\varphi$ quasi-everywhere and $\mu^\varphi \in D_\mathcal{E}$, the above inequality becomes
\[
C_\varphi = \int_{\mathbb{R}^d} (V_\alpha + 2U_{\mu^\varphi}) \, d\mu^\varphi \geq C_\varphi + \varepsilon\mu^\varphi(\mathcal{U}).
\]
We reach a contradiction by noting that $\mu^\varphi(\mathcal{U}) > 0$ since $\mathcal{U}$ is a neighborhood of $x \in \text{supp}(\mu^\varphi)$ (using the definition of the support), which proves the first line of (3.10).

**APPENDIX D. PROOF OF THEOREM 3.14**

We outline the proof of Theorem 3.14 which follows the same lines as in the linear case.

**Proof.** We show below that $\mathcal{B}$ defined in (3.13) is closed for the $p$-Wasserstein topology under Assumption 3.12. For this, we show that $\mathcal{B}^c$ is open by picking $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ such that $Q(\mu) > 0$ and using again that, for $\varepsilon > 0$ and $\nu \in \mathcal{P}_p(\mathbb{R}^d)$ such that $d_{W_p}(\mu, \nu) \leq \varepsilon$, it holds, by (1.16),
\[
\sup_{f \in L^1(\mu), g \in L^1(\nu)} \left( \int_{\mathbb{R}^d} f \, d\mu - \int_{\mathbb{R}^d} g \, d\nu \right) \leq \varepsilon.
\]
First, by Assumption 3.12, we note that for any $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ it holds $\|U^\psi_\mu\|_{\infty, p} < +\infty$ and $\|U^\psi_\nu\|_{\infty, p} < +\infty$. Therefore, $U^\psi_\mu \in L^1(\mu) \cap L^1(\nu)$ and $U^\psi_\nu \in L^1(\mu) \cap L^1(\nu)$ for any probability measures $\mu, \nu$ with moments of order $p$. Next, by (3.13), it holds $\|U^\psi_\mu\|_{\text{Lip}} \leq C_{\text{Lip}}$ and $\|U^\psi_\nu\|_{\text{Lip}} \leq C_{\text{Lip}}$.

We assume for now that these norms are non-zero, so we may first choose $f = g = U^\psi_\nu / \|U^\psi_\nu\|_{\text{Lip}}$, which leads to
\[
Q(\mu) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x, y) \nu(dx) \mu(dy) \leq \varepsilon \|U^\psi_\nu\|_{\text{Lip}}.
\]
Symmetrically we take $f = g = -U^\psi_\mu / \|U^\psi_\mu\|_{\text{Lip}}$, which leads to
\[
\int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(y, x) \nu(dx) \mu(dy) - Q(\nu) \leq \varepsilon \|U^\psi_\mu\|_{\text{Lip}}.
\]
By summing (D.2) and (D.3) and using the symmetry of $\psi$, we obtain

$$Q(\nu) \geq Q(\mu) - \varepsilon (\|U^\psi_\mu\|_{\text{Lip}} + \|U^\psi_\nu\|_{\text{Lip}}) \geq Q(\mu) - 2C_{\text{Lip}}\varepsilon. \quad (D.4)$$

This shows that $Q(\nu) > 0$ for $\varepsilon < Q(\mu)/(2C_{\text{Lip}})$. To finish the argument, we consider the cases where the Lipschitz norm of the potentials generated by $\mu$ and $\nu$ may be zero. Suppose first that $\|U^\psi_\mu\|_{\text{Lip}} = 0$. This implies the existence of $c_\mu \in \mathbb{R}$ such that

$$\forall x \in \mathbb{R}^d, \quad \int_{\mathbb{R}^d} \psi(x, y)\mu(dy) = c_\mu. \quad (D.5)$$

Integrating the above equation with respect to $\mu$ shows that $c_\mu = Q(\mu) > 0$. As a result, if $\|U^\psi_\mu\|_{\text{Lip}} = 0$ and $\nu$ is such that $\|U^\psi_\nu\|_{\text{Lip}} > 0$ we can consider (D.3), which becomes (integrating (D.5) with respect to $\nu$)

$$Q(\nu) \geq Q(\mu) - \varepsilon \|U^\psi_\nu\|_{\text{Lip}} \geq Q(\mu) - \varepsilon C_{\text{Lip}}.$$  

In this case, $Q(\nu) > 0$ for $\varepsilon < Q(\mu)/C_{\text{Lip}}$. Then, if $\|U^\psi_\nu\|_{\text{Lip}} = 0$ it holds, for some $c_\nu \in \mathbb{R}$,

$$\forall x \in \mathbb{R}^d, \quad \int_{\mathbb{R}^d} \psi(x, y)\nu(dy) = c_\nu. \quad (D.6)$$

Integrating with respect to $\mu$ and using the symmetry of $\psi$ we obtain that $c_\nu = c_\mu > 0$. Integrating next (D.6) with respect to $\nu$ shows that $Q(\nu) = c_\nu = Q(\mu) > 0$. Finally, if $\|U^\psi_\mu\|_{\text{Lip}} > 0$ but $\|U^\psi_\nu\|_{\text{Lip}} = 0$, (D.6) holds with $c_\nu = Q(\nu)$ so that (D.2) becomes

$$Q(\nu) \geq Q(\mu) - \varepsilon C_{\text{Lip}},$$

and the same conclusion follows. As a result, in any case the measures $\nu$ such that $d_{W_1}(\mu, \nu) \leq \varepsilon \frac{1}{2}$ for $\varepsilon < Q(\mu)/(2C_{\text{Lip}})$ belong to $B^c$ so that $B^c$ is open and $B$ is closed in the $p$-Wasserstein topology.

We next show that $B$ is an $I$-continuity set. The existence of $\mu_- \in D_\mathcal{E}$ such that $Q(\mu_-) < 0$ ensures that

$$\inf_{\mathcal{B}} \mathcal{E} < +\infty.$$ 

Since $\mathcal{E}$ has compact level sets and $B$ is closed, there exists $\tilde{\mu}$ such that

$$\mathcal{E}(\tilde{\mu}) = \inf_{\mathcal{B}} \mathcal{E}.$$ 

If $Q(\tilde{\mu}) < 0$, $I$-continuity is proven, so we may assume that $Q(\tilde{\mu}) = 0$. Like in the linear case, we may take $\mu_\varepsilon \in \tilde{B}$ such that

$$\mathcal{E}(\mu_\varepsilon) \leq \inf_{\mathcal{B}} \mathcal{E} + \varepsilon,$$

and consider the convex combination $\mu_t = t\mu_\varepsilon + (1-t)\tilde{\mu}$ for $t \in (0,1)$. The convexity of $Q$ shows that, for any $t \in (0,1)$ it holds

$$Q(\mu_t) \leq tQ(\mu_\varepsilon) + (1-t)Q(\tilde{\mu}) < 0,$$

so that $\mu_t \in \tilde{B}$. Proceeding by contradiction by supposing that $\mathcal{E}(\tilde{\mu}) < \inf_{\tilde{B}} \mathcal{E}$, we obtain that for $t > 0$ small enough it holds $\mu_t \in \tilde{B}$ and

$$\mathcal{E}(\mu_t) < \inf_{\mathcal{B}} \mathcal{E},$$

which is a contradiction, proving that $B$ is an $I$-continuity set.

One can next follow step 2 of the proof of Theorem 3.7 to show that the minimizer $\mu^\psi$ is unique with compact support.

At this stage, the remaining statements in Theorem 3.14 can be proved as for Theorem 3.7. In particular, we can introduce a set similar to (C.3) by setting

$$R = \{ (\mathcal{E}(\mu) - \mathcal{E}(\mu^\psi) + a_0, Q(\mu) + a_1) : a_0 > 0, a_1 > 0, \mu \in D_\mathcal{E} \}.$$ 

The set $R$ is convex by convexity of $Q$, so the same convex separation theorem can be used, and we can show that there exists $\alpha \geq 0$ such that

$$\forall \mu \in D_\mathcal{E}, \quad \mathcal{E}(\mu^\psi) \leq \mathcal{E}(\mu) + \alpha Q(\mu).$$
In this procedure, we use the existence of $\mu_\alpha$ from Assumption 3.12 in order to reproduce the qualification of constraint argument. This leads to computations where the interaction energy $J$ is replaced by

$$J_\alpha(\mu, \nu) = J(\mu, \nu) + \alpha \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x, y) \mu(dx) \mu(dy),$$

from which (3.17) follows by mimicking step 4 of the proof of Theorem 3.7. □

As a final comment, let us insist on the importance of the convexity of $Q$ for the above proof to be valid.

### References

[1] G. Akemann & M. Cikovic – “Products of random matrices from fixed trace and induced Ginibre ensembles”, *J. Phys. A* 51 (2018), no. 18, p. 184002, 34. [6]

[2] L. Ambrosio, N. Fusco & D. Pallara – *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Mathematical Monographs, vol. 254, Clarendon Press Oxford, 2000. [12]

[3] S. N. Armstrong, S. Serfaty & O. Zeitouni – “Remarks on a constrained optimization problem for the Ginibre ensemble”, *Potential Anal.* 41 (2014), no. 3, p. 945–958. [8 15]

[4] V. Barbui & T. Pfeutpan – *Convexity and Optimization in Banach Spaces*, Mathematics and its Applications, vol. 10, Springer Science & Business Media, 2012. [27]

[5] G. Ben Arous – “Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy”, *Probab. Theory Relat. Fields* 108 (1997), no. 4, p. 517–542. [3 6 19]

[6] C. Berg, J. P. R. Christensen & P. Ressel – *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*, Graduate Texts in Mathematics, vol. 100, Springer, 1984. [11]

[7] R. J. Berman – “The Coulomb gas, potential theory and phase transitions”, preprint arXiv:1811.10249 (2018). [9]

[8] C. Bordenave, P. Caputo & D. Chafaï – “Circular law theorem for random Markov matrices”, *Probab. Theory Relat. Fields* 152 (2012), no. 3-4, p. 751–779. [21]

[9] J.-P. Bouchaud & M. Potters – “Financial applications of random matrix theory: a short review”, preprint arXiv:0910.1205 (2009). [8]

[10] D. Chafaï – “Glauber versus Kawasaki for spectral gap and logarithmic Sobolev inequalities of some unbounded conservative spin systems”, *Markov Process. Relat. Fields* 9 (2003), no. 3, p. 341–362. [24]

[11] D. Chafaï & G. Ferré – “Simulating Coulomb and log-gases with Hybrid Monte Carlo algorithms”, *J. Statist. Phys.* 174 (2018), no. 3, p. 692–714. [11 16 27]

[12] D. Chafaï, N. Gozlan & P.-A. Zitt – “First-order global asymptotics for confined particles with singular pair repulsion”, *Ann. Appl. Probab.* 24 (2014), no. 6, p. 2371–2413. [1 3 4 6 7 8 13 21 24 27]

[13] D. Chafaï, A. Hardy & M. Maida – “Concentration for Coulomb gases and Coulomb transport inequalities”, *J. Funct. Anal.* 275 (2018), no. 6, p. 1447–1483. [23]

[14] D. Chafaï & S. Péché – “A note on the second order universality at the edge of Coulomb gases on the plane”, *J. Stat. Phys.* 156 (2014), no. 2, p. 368–383. [2]

[15] D. Chafaï & J. Lehec – “On Poisson–Carré and logarithmic Sobolev inequalities for a class of singular Gibbs measures”, preprint arXiv:1805.00708v2 (2018). [2 21]

[16] I. Csiszar – “Sanov property, generalized I-projection and a conditional limit theorem”, *Ann. Probab.* 12 (1984), no. 3, p. 768–793. [6]

[17] E. Darve – “Thermodynamic integration using constrained and unconstrained dynamics”, in *Free energy calculations* (C. Chipot & A. Pohorille, eds.), Springer Series in Chemical Physics, vol. 86, Springer, 2007, p. 119–170. [19]

[18] A. Dembo & O. Zeitouni – *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition. [5 8 9]

[19] P. Dupuis, V. Laschos & K. Ramanan – “Large deviations for empirical measures generated by Gibbs measures with singular energy functionals”, preprint arXiv:1511.06928v1 (2015). [2 3 6 8 23]

[20] L. Erdős & H.-T. Yau – *A Dynamical Approach to Random Matrix Theory*, Courant Lecture Notes in Mathematics, vol. 28, American Mathematical Society, Providence, RI, 2017. [6]

[21] L. Evans & R. Gariepy – *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, vol. 5, CRC Press, 1992. [12]

[22] P. J. Forrester – *Log-gases and Random Matrices*, London Mathematical Society Monographs Series, vol. 34, Princeton University Press, Princeton, NJ, 2010. [5 6]

[23] D. García-Zelada – “A large deviation principle for empirical measures on Polish spaces: Application to singular Gibbs measures on manifolds”, preprint arXiv:1703.02001 (2017). [27]

[24] S. Ghosh & A. Nishry – “Gaussian complex zeros on the hole event: the emergence of a forbidden region”, preprint arXiv:1609.00081v2 (2018). [3 5 13]

[25] S. Ghosh & A. Nishry – “Point processes, hole events, and large deviations: random complex zeros and Coulomb gases”, *Constr. Approx.* 48 (2018), no. 1, p. 101–136. [3 8 13]
