Large Deviation Principles for Hypersingular Riesz Gases

Douglas P. Hardin, Thomas Leblé, Edward B. Saff & Sylvia Serfaty

Springer

CONSTRUCTIVE
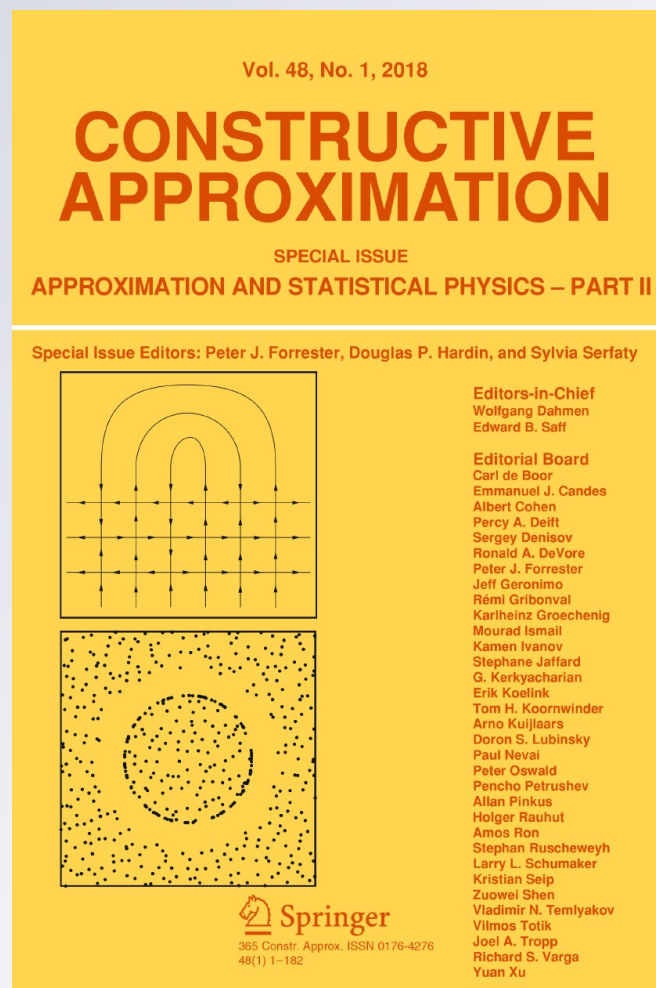APPROXIMATION

CrossMark

Large Deviation Principles for Hypersingular Riesz Gases

Douglas P. Hardin[1] · **Thomas Leblé**[2] ·
Edward B. Saff[1] · **Sylvia Serfaty**[2,3,4]

Abstract We study N-particle systems in $\mathbb{R}^d$ whose interactions are governed by a hypersingular Riesz potential $|x - y|^{-s}$, $s > d$, and subject to an external field. We provide both macroscopic results as well as microscopic results in the limit as $N \to \infty$ for random point configurations with respect to the associated Gibbs measure at scaled inverse temperature β. We show that a large deviation principle holds with a rate function of the form 'β-Energy + Entropy', yielding that the microscopic behavior

✉ Edward B. Saff
edward.b.saff@vanderbilt.edu

Douglas P. Hardin
doug.hardin@vanderbilt.edu

Thomas Leblé
thomasl@math.nyu.edu

Sylvia Serfaty
serfaty@cims.nyu.edu

1 Department of Mathematics, Center for Constructive Approximation, Vanderbilt University, Nashville, TN 37240, USA

2 Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1110, USA

3 Institut Universitaire de France, Paris, France

4 UPMC, CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, Sorbonne Universités, 4, Place Jussieu, 75005 Paris, France

🖄 Springer

(on the scale $N^{-1/d}$) of such N-point systems is asymptotically determined by the minimizers of this rate function. In contrast to the asymptotic behavior in the integrable case $s < d$, where on the macroscopic scale N-point empirical measures have limiting density independent of β, the limiting density for $s > d$ is strongly β-dependent.

Keywords Riesz gases · Gibbs measure · Large deviation principle · Empirical measures · Minimal energy

Mathematics Subject Classification Primary 82D10 · 82B05; Secondary 31C20 · 28A78

1 Introduction and Main Results

1.1 Hypersingular Riesz Gases

Let $d \geq 1$ and s be a real number with $s > d$. We consider a system of N points in the Euclidean space $\mathbb{R}^d$ with *hypersingular* Riesz pairwise interactions, in an external field V. The particles are assumed to live in a *confinement set* $\Omega \subseteq \mathbb{R}^d$. The energy $\mathcal{H}_N(\mathbf{X}_N)$ of the system in a given state $\mathbf{X}_N = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$ is defined to be

$$\mathcal{H}_N(\mathbf{X}_N) := \sum_{1 \leq i \neq j \leq N} \frac{1}{|x_i - x_j|^s} + N^{s/d} \sum_{i=1}^N V(x_i). \tag{1.1}$$

The external field V is a confining potential, growing at infinity, on which we shall make assumptions later. The term *hypersingular* corresponds to the fact that the singularity of the kernel $|x - y|^{-s}$ is nonintegrable with respect to the Lebesgue measure on $\mathbb{R}^d$.

For any $\beta > 0$, the canonical Gibbs measure associated with (1.1) at inverse temperature β and for particles living on Ω is given by

$$d\mathbb{P}_{N,\beta}(\mathbf{X}_N) = \frac{1}{Z_{N,\beta}} \exp\left(-\beta N^{-s/d} \mathcal{H}_N(\mathbf{X}_N)\right) \mathbf{1}_{\Omega^N}(\mathbf{X}_N) d\mathbf{X}_N, \tag{1.2}$$

where $d\mathbf{X}_N$ is the Lebesgue measure on $(\mathbb{R}^d)^N$, $\mathbf{1}_{\Omega^N}(\mathbf{X}_N)$ is the indicatrix function of Ω^N, and $Z_{N,\beta}$ is the "partition function"; i.e., the normalizing factor

$$Z_{N,\beta} := \int_{\Omega^N} \exp\left(-\beta N^{-s/d} \mathcal{H}_N(\mathbf{X}_N)\right) d\mathbf{X}_N. \tag{1.3}$$

We will call the statistical physics system described by (1.1) and (1.2) a *hypersingular Riesz gas*.

For Riesz potentials in the case $s > d$, ground state configurations (or Riesz energy minimizers) of N-particle systems (with or without the external field V) have been extensively studied in the large N limit, see [13,14,16] and the references therein. Furthermore, for the case of positive temperature, the statistical mechanics of Riesz gases have been investigated in [18] but for a different range of the parameter s,

namely $\max(d - 2, 0) \leq s < d$. In that paper, a large deviation principle for the empirical process (which encodes the microsopic behavior of the particles at scale $N^{-1/d}$, averaged in a certain way) was derived. The main goal of the present paper is to extend that work to the hypersingular case. By combining the approaches of the above mentioned papers, we obtain a large deviation principle describing macroscopic as well as microscopic properties for hypersingular Riesz gases.

Studying Riesz interactions for the whole range of s from 0 to infinity is of interest in approximation theory and coding theory, as it connects logarithmic interactions, Coulomb interactions, and (in the limit $s \to \infty$) packing problems, see [13,25]. Investigating such systems with temperature is also a natural question for statistical mechanics, as it improves our understanding of the behavior of systems with long-range vs. short-range interactions (see, for instance, [4,6,9] where the interest of such questions is stressed and [2] and [22, Section 4.2] for additional results). Analyzing the case $s > d$ is also a first step toward the study of physically more relevant interactions, such as the Lennard–Jones potential.

The hypersingular Riesz case $s > d$ and the integrable Riesz case $s < d$ have important differences. For $s < d$ (which can be thought of as long-range) and, more generally, for integrable interaction kernels g (which includes regular interactions) the global, macroscopic behavior can be studied using classical potential theory. Namely, the empirical measure $\frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$ is found to converge rapidly to some equilibrium measure determined uniquely by Ω and V and obtained as the unique minimizer of the potential-theoretic functional

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} g(x - y) \mathrm{d}\mu(x) \mathrm{d}\mu(y) + \int_{\mathbb{R}^d} V \mathrm{d}\mu,$$

which can be seen as a mean-field energy with a nonlocal term. We refer, e.g., to [26] or [27, Chap. 2] for a treatment of this question (among others).

In these integrable cases, if temperature is scaled in the same way as here, the macroscopic behavior is governed by the equilibrium measure and thus is independent of the temperature so that no knowledge of the microscopic distribution of points is necessary to determine the macroscopic distribution. At the next order in energy, which governs the microscopic distribution of the points, a dependency on β appears. As seen in [18], in the Coulomb and potential Riesz cases (it is important in the method that the interaction kernel be reducible to the kernel of a local operator, which is known only for these particular interactions), the microscopic distribution around a point is given by a problem in a whole space with a neutralizing background, fixing the local density as equal to that of the equilibrium measure at that point. The microscopic distribution is found to minimize the sum of a (renormalized) Riesz energy term and a relative entropy term. A crucial ingredient in the proof is a "screening" construction showing that energy can be computed additively over large disjoint microscopic boxes; i.e., interactions between configurations in different large microscopic boxes are negligible to this order.

The hypersingular case can be seen as more delicate than the integrable case due to the absence of an equilibrium measure. The limit of the empirical measure has to be identified differently. In the case of ground state configurations (minimizers),

this was done in [16]. For positive temperature, in contrast with the above described integrable case, we shall show that the empirical limit measure is obtained as a by-product of the study at the microscopic scale and depends on β in quite an indirect way (see Theorem 1.3). The microscopic profiles minimize a full-space version of the problem, giving an energy that depends on the local density. The macroscopic distribution can then be found by a local density approximation, by minimizing the sum of its energy and that due to the confinement potential. Since the energy is easily seen to scale like $N^{1+s/d}$, the choice of the temperature scaling $\beta N^{-s/d}$ is made so that the energy and the entropy for the microscopic distributions carry the same weight of order N. Other choices of temperature scalings are possible, but would lead to degenerate versions of the situation we are examining, with either all the entropy terms asymptotically disappearing for small temperatures, or the effect of the energy altogether disappearing for large temperatures. Note that going to the microscopic scale in order to derive the behavior at the macroscopic scale was already the approach needed in [19] for the case of the "two-component plasma", a system of two-dimensional particles of positive and negative charges interacting logarithmically for which no a priori knowledge of the equilibrium measure can be found.

On the other hand, the hypersingular case is also easier in the sense that the interactions decay faster at infinity, implying that long-range interactions between large microscopic "boxes" are negligible and do not require any sophisticated screening procedures. Our proofs will make crucial use of this "self-screening" property.

To describe the system at the microscopic scale, we define a Riesz energy $\overline{\mathbb{W}}_s$ (see Sect. 2.3.5) for infinite random point configurations that is the counterpart of the renormalized energy of [17,18,23] (defined for $s < d$). It is conjectured to be minimized by lattices for certain low dimensions, but this is a completely open problem with the exception of dimension 1 (see [1] and the discussion following (2.15)).

With any sequence of configurations $\{\mathbf{X}_N\}_N$, we associate an "empirical process" whose limit (a random tagged point process) describes the point configurations $\mathbf{X}_N$ at scale $N^{-1/d}$. Our main result will be that there is a Large Deviation Principle (LDP) for the law of this empirical process with rate function equal to (a variant of) the energy $\beta \overline{\mathbb{W}}_s$ plus the relative entropy of the empirical process with respect to the Poisson point process.

For minimizers of the Riesz energy $\mathcal{H}_N$, we show that the limiting empirical processes must minimize $\overline{\mathbb{W}}_s$, thus describing their microscopic structure.

The question of treating more general interactions than the Riesz ones remains widely open. The fact that the interaction has a precise homogeneity under rescaling is crucial for the hypersingular case treated here. On the other hand, in the integrable case, we do not know how to circumvent the need for expressing the energy via the potential generated by the points, i.e., the need for the Caffarelli–Silvestre representation of the interaction as the kernel of a local operator (achieved by adding a space dimension).

1.2 Assumptions and Notation

1.2.1 Assumptions

In the rest of the paper, we assume that $\Omega \subset \mathbb{R}^d$ is closed with positive d-dimensional Lebesgue measure and that

$$\partial\Omega \text{ is } C^1, \tag{1.4}$$

$$V \text{ is a continuous, non-negative real valued function on } \Omega. \tag{1.5}$$

Furthermore, if Ω is unbounded, we assume that

$$\lim_{|x|\to\infty} V(x) = +\infty, \tag{1.6}$$

$$\exists M > 0 \text{ such that } \int \exp\left(-MV(x)\right) dx < +\infty. \tag{1.7}$$

The assumption (1.4) on the regularity of $\partial\Omega$ is mostly technical, and we believe that it could be relaxed to, e.g., $\partial\Omega$ is locally the graph of, say, a Hölder function in $\mathbb{R}^d$. However, it is unclear to us what the minimal assumption could be (e.g., is it enough to assume that $\partial\Omega$ has zero measure?). An interesting direction would be to study the case where Ω is a p-rectifiable set in $\mathbb{R}^d$ with $p < d$ (see, e.g., [3,16]). Assumption (1.5) is quite mild (in comparison, e.g., with the corresponding assumption in the $s < d$ case, where one wants to ensure some regularity of the so-called equilibrium measure, which is essentially two orders lower than that for V), and we believe it to be sharp for our purposes. Assumption (1.6) is an additional confinement assumption, and (1.7) ensures that the partition function $Z_{N,\beta}$, defined in (1.3), is finite (at least for N large enough). Indeed, the interaction energy is non-negative, hence for N large enough (1.7) ensures that the integral defining the partition function is convergent.

1.2.2 General Notation

We let $\mathcal{X}$ be the space of point configurations in $\mathbb{R}^d$ (see Sect. 2.1 for a precise definition). If X is some measurable space and $x \in X$, we denote by δ_x, the Dirac mass at x.

1.2.3 Empirical Measure and Empirical Processes

Let $\mathbf{X}_N = (x_1, \ldots, x_N)$ in Ω^N be fixed.

- We define the empirical measure $\text{emp}(\mathbf{X}_N)$ as

$$\text{emp}(\mathbf{X}_N) := \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}. \tag{1.8}$$

It is a probability measure on Ω.

- We define $\mathbf{X}'_N$ as the finite configuration rescaled by a factor $N^{1/d}$

$$\mathbf{X}'_N := \sum_{i=1}^{N} \delta_{N^{1/d} x_i}. \tag{1.9}$$

It is a point configuration (an element of $\mathcal{X}$), which represents the N-tuple of particles $\mathbf{X}_N$ seen at microscopic scale.

- We define the *tagged empirical process* $\overline{\mathrm{Emp}}_N(\mathbf{X}_N)$ as

$$\overline{\mathrm{Emp}}_N(\mathbf{X}_N) := \int_\Omega \delta_{\left(x, \, \theta_{N^{1/d} x} \cdot \mathbf{X}'_N\right)} \mathrm{d}x, \tag{1.10}$$

where θ_x denotes the translation by $-x$. It is a positive measure on $\Omega \times \mathcal{X}$.

Let us now briefly explain the meaning of the last definition (1.10). For any $x \in \Omega$, $\theta_{N^{1/d} x} \cdot \mathbf{X}'_N$ is an element of $\mathcal{X}$ that represents the N-tuple of particles $\mathbf{X}_N$ centered at x and seen at microscopic scale (or, equivalently, seen at microscopic scale and then centered at $N^{1/d} x$). In particular, any information about this point configuration in a given ball (around the origin) translates to an information about $\mathbf{X}'_N$ around x. We may thus think of $\theta_{N^{1/d} x} \cdot \mathbf{X}'_N$ as encoding the behavior of $\mathbf{X}'_N$ around x.

The measure

$$\int_\Omega \delta_{\theta_{N^{1/d} x} \cdot \mathbf{X}'_N} \mathrm{d}x \tag{1.11}$$

is a measure on $\mathcal{X}$ that encodes the behavior of $\mathbf{X}'_N$ around each point $x \in \Omega$. We may think of it as the "averaged" microscopic behavior (although it is not, in general, a probability measure, and its mass can be infinite). The measure defined by (1.11) would correspond to what is called the "empirical field".

The tagged empirical process $\overline{\mathrm{Emp}}_N(\mathbf{X}_N)$ is a finer object, because for each $x \in \Omega$ we keep track of the centering point x as well as of the microscopic information $\theta_{N^{1/d} x} \cdot \mathbf{X}'_N$ around x. It yields a measure on $\Omega \times \mathcal{X}$ whose first marginal is the Lebesgue measure on Ω and whose second marginal is the (non-tagged) empirical process defined above in (1.11). Keeping track of this additional information allows one to test $\overline{\mathrm{Emp}}_N(\mathbf{X}_N)$ against functions $F(x, \mathcal{C}) \in C^0(\Omega \times \mathcal{X})$ which may be of the form

$$F(x, \mathcal{C}) = \chi(x) \tilde{F}(\mathcal{C}),$$

where χ is a smooth function localized in a small neighborhood of a given point of Ω, and $\tilde{F}(\mathcal{C})$ is a continuous function on the space of point configurations. Using such test functions, we may thus study the microsopic behavior of the system after a small average (on a small domain of Ω), whereas the empirical process only allows one to study the microscopic behavior after averaging over the whole Ω.

The study of empirical processes, or *empirical fields*, as natural quantities to encode the averaged microscopic behavior appear, e.g., in [11] for particles without interaction or [12] in the interacting case.

1.2.4 Large Deviation Principle

Let us recall that an sequence $\{\mu_N\}_N$ of probability measures on a metric space X is said to satisfy an LDP at speed r_N with rate function $I : X \to [0, +\infty]$ if the following holds for any Borel set $A \subset X$:

$$- \inf_{\mathring{A}} I \leq \liminf_{N \to \infty} \frac{1}{r_N} \log \mu_N(A) \leq \limsup_{N \to \infty} \frac{1}{r_N} \log \mu_N(A) \leq - \inf_{\bar{A}} I,$$

where $\mathring{A}$ (resp. $\bar{A}$) denotes the interior (resp. the closure) of A. The functional I is said to be a *good rate function* if it is lower semi-continuous and has compact sub-level sets. We refer to [10] and [28] for detailed treatments of the theory of large deviations and to [24] for an introduction to the applications of LDP's in the statistical physics setting.

Roughly speaking, an LDP at speed r_N with rate function I expresses the following fact: the probability measures μ_N concentrate around the points where I vanishes, and any point $x \in X$ such that $I(x) > 0$ is not "seen" with probability $1 - \exp(-NI(x))$.

1.3 Main Results

1.3.1 Large Deviations of the Empirical Processes

We let $\overline{\mathfrak{P}}_{N,\beta}$ be the push-forward of the Gibbs measure $\mathbb{P}_{N,\beta}$ (defined in (1.2)) by the map $\overline{\mathrm{Emp}}_N$ defined in (1.10). In other words, $\overline{\mathfrak{P}}_{N,\beta}$ is the law of the random variable "tagged empirical process" when $\mathbf{X}_N$ is distributed following $\mathbb{P}_{N,\beta}$.

The following theorem, which is the main result of this paper, involves the functional $\overline{\mathcal{F}}_\beta = \overline{\mathcal{F}}_{\beta,s}$ defined in (2.26). It is a free energy functional of the type "β Energy + Entropy" (see Sect. 2.2, 2.3, and 2.4 for precise definitions). The theorem expresses the fact that the microscopic behavior of the system of particles is determined by the minimization of the functional $\overline{\mathcal{F}}_\beta$ and that configurations $\mathbf{X}_N$ having empirical processes $\overline{\mathrm{Emp}}(\mathbf{X}_N)$ far from a minimizer of $\overline{\mathcal{F}}_\beta$ have negligible probability of order $\exp(-N)$.

Theorem 1.1 *For any $\beta > 0$, the sequence $\{\overline{\mathfrak{P}}_{N,\beta}\}_N$ satisfies a large deviation principle at speed N with good rate function $\overline{\mathcal{F}}_\beta - \min \overline{\mathcal{F}}_\beta$.*

Corollary 1.2 *The first-order expansion of $\log Z_{N,\beta}$ as $N \to \infty$ is*

$$\log Z_{N,\beta} = -N \min \overline{\mathcal{F}}_\beta + o(N).$$

1.3.2 Large Deviations of the Empirical Measure

As a by-product of our microscopic study, we derive a large deviation principle that governs the asymptotics of the empirical measure (which is a macroscropic quantity). Let us denote by $\mathfrak{emp}_{N,\beta}$ the law of the random variable $\mathrm{emp}(\mathbf{X}_N)$ when $\mathbf{X}_N$ is distributed according to $\mathbb{P}_{N,\beta}$. The rate function $I_\beta = I_{\beta,s}$, defined in Sect. 2.4 (see (2.28)), has the form

$$I_\beta(\rho) = \int_\Omega f_\beta(\rho(x))\rho(x)\,\mathrm{d}x + \beta \int_\Omega V(x)\rho(x)\,\mathrm{d}x + \int_\Omega \rho(x)\log\rho(x)\,\mathrm{d}x \tag{1.12}$$

and is a local density approximation. The function f_β in this expression is determined by a minimization problem over the *microscopic* empirical processes.

Theorem 1.3 *For any $\beta > 0$, the sequence $\{\mathfrak{emp}_{N,\beta}\}_N$ obeys a large deviation principle at speed N with good rate function $I_\beta - \min I_\beta$. In particular, the empirical measure converges almost surely to the unique minimizer of I_β.*

The rate function I_β is quite complicated to study in general. However, thanks to the convexity of f_β and elementary properties of the standard entropy, we may characterize its minimizer in some particular cases (see Sect. 5 for the proof) :

Proposition 1.4 *Let $\mu_{V,\beta}$ be the unique minimizer of I_β.*

1. *If $V = 0$ and Ω is bounded, then $\mu_{V,\beta}$ is the uniform probability measure on Ω for any $\beta > 0$.*
2. *If V is arbitrary and Ω is bounded, $\mu_{V,\beta}$ converges to the uniform probability measure on Ω as $\beta \to 0$.*
3. *If V is arbitrary, $\mu_{V,\beta}$ converges to $\mu_{V,\infty}$ as $\beta \to +\infty$, where $\mu_{V,\infty}$ is the limit empirical measure for energy minimizers as defined in the paragraph below.*

1.3.3 The Case of Minimizers

Our remaining results deal with energy minimizers (in statistical physics, this corresponds to setting $\beta = +\infty$). Let $\{\mathbf{X}_N\}_N$ be a sequence of point configurations in Ω such that for any $N \geq 1$, $\mathbf{X}_N$ has N points and minimizes $\mathcal{H}_N$ on Ω^N.

The macroscopic behavior is known from [16]: there is a unique minimizer $\mu_{V,\infty}$ (the notation differs from [16]) of the functional

$$C_{s,d} \int_\Omega \rho(x)^{1+s/d}\,\mathrm{d}x + \int_\Omega V(x)\rho(x)\,\mathrm{d}x \tag{1.13}$$

among probability densities ρ over Ω ($C_{s,d}$ is a constant depending on s, d defined in (2.10)), and the empirical measure $\mathrm{emp}(\mathbf{X}_N)$ converges to $\mu_{V,\infty}$ as $N \to \infty$. See (5.4) for an explicit formula for $\mu_{V,\infty}$. Note that the formula (1.13) is what one obtains when letting formally $\beta \to \infty$ in the definition of I_β, and resembles some of the terms arising in Thomas–Fermi theory (cf. [21] and [20]).

The notation for the next statement is given in Sects. 2.1 and 2.3. Let us simply say that $\overline{\mathbb{W}}_s$ (resp. $\mathcal{W}_s$) is an energy functional defined for a random point configuration (resp. a point configuration), and that $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$ (resp. $\mathcal{X}_{\mu_V,\infty(x)}$) is some particular subset of random point configurations (resp. of point configurations in $\mathbb{R}^d$). The intensity measure of a random tagged point configuration is defined in Sect. 2.1.7.

Proposition 1.5 *We have:*

1. *$\{\overline{\mathrm{Emp}}(\mathbf{X}_N)\}_N$ converges weakly (up to extraction of a subsequence) to some minimizer $\overline{P}$ of $\overline{\mathbb{W}}_s$ over $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$.*
2. *The intensity measure of $\overline{P}$ coincides with $\mu_{V,\infty}$.*
3. *For $\overline{P}$-almost every $(x, \mathcal{C})$, the point configuration $\mathcal{C}$ minimizes $\mathcal{W}_s(\mathcal{C})$ within the class $\mathcal{X}_{\mu_V,\infty(x)}$.*

The first point expresses the fact that the tagged empirical processes associated with minimizers converge with minimizers of the "infinite-volume" energy functional $\overline{\mathbb{W}}_s$. The second point is a rephrasing of the global result cited above, to which the third point adds some microscopic information.

The problem of minimizing the energy functionals $\overline{\mathbb{W}}_s$, $\mathbb{W}_s$, or $\mathcal{W}_s$ is hard in general. In dimension 1, however, it is not too difficult to show that the "crystallization conjecture" holds, namely that the microscopic structure of minimizers is ordered and converges to a lattice:

Proposition 1.6 *Assume $d = 1$. The unique stationary minimizer of $\mathbb{W}_s$ is the law of $u + \mathbb{Z}$, where u is a uniform choice of the origin in $[0, 1]$.*

In dimension 2, it would be expected that minimizers are given by the triangular (or Abrikosov) lattice; we refer to [1] for a recent review of such conjectures. In large dimensions, it is not expected that lattices are minimizers.

1.4 Outline of the Method

Our LDP result is phrased in terms of the empirical processes associated with point configurations, as in [18], and thus the objects we consider and the overall aproach are quite similar to [18]. It is however quite simplified by the fact that, because the interaction is short-range and we are in the nonpotential case, we do not need to express the energy in terms of the "electric potential" generated by the point configuration. The definition of the limiting microscopic interaction energy $\mathcal{W}_s(\mathcal{C})$ is thus significantly simpler than in [18]; it suffices to take, for $\mathcal{C}$ an infinite configuration of points in the whole space,

$$\mathcal{W}_s(\mathcal{C}) = \liminf_{R \to \infty} \sum_{p,q \in \mathcal{C} \cap K_R, p \neq q} \frac{1}{|p - q|^s},$$

where K_R is the cube of sidelength R centered at the origin. When considering this quantity, there is however no implicit knowledge of the average density of points, in contrast to the situation of [18]. This is then easily extended to an energy for point processes $\overline{\mathbb{W}}_s$ by taking expectations.

As in [18], the starting point of the LDP proof is a Sanov-type result that states that the logarithm of the volume of configurations whose empirical processes lie in a small ball around a given tagged point process $\overline{P}$ can be expressed as $(-N)$ times an entropy denoted $\mathsf{ent}(\overline{P}|\Pi)$. As we shall show, $N^{-1-s/d}\mathcal{H}_N(\mathbf{X}_N) \approx \overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P})$ for a sufficiently large set of configurations $\mathbf{X}_N$ near P, where $\overline{\mathbb{V}}(\overline{P})$ is a term corresponding to the external potential V. Then this will suffice to obtain the LDP since the logarithm of the probability of the empirical field being close to $\overline{P}$ is nearly N times

$$-\beta(\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P})) - \mathsf{ent}(\overline{P}|\Pi),$$

up to an additive constant. The entropy can be expressed in terms of $\overline{P}^x$ (the process centered at x) as

$$\mathsf{ent}(\overline{P}|\Pi) = \int (\mathsf{ent}(\overline{P}^x|\Pi) - 1)\,\mathrm{d}x + 1, \tag{1.14}$$

where $\mathsf{ent}(P|\Pi)$ is a "specific relative entropy" with respect to the Poisson point process Π. Assuming that $\overline{P}^x$ has an intensity $\rho(x)$, then the scaling properties of the energy $\overline{\mathbb{W}}_s$ (the fact that the energy scales like $\rho^{1+s/d}$ where ρ is the density) and of the specific relative entropy ent allow one to transform this into

$$-\int_\Omega \left(\beta\rho^{s/d}\mathbb{W}_s(\overline{P}^x) + \mathsf{ent}(\sigma_{\rho(x)}(\overline{P}^x)|\Pi) + \beta V(x)\right)\rho(x)\,\mathrm{d}x$$
$$-\int_\Omega \rho(x)\log\rho(x)\,\mathrm{d}x,$$

which is the desired rate function. Minimizing over P's of intensity ρ allows one to obtain the rate function I_β of (2.27).

To run through this argument, we encounter the same difficulties as in [18], i.e., the difficulty in replacing $\mathcal{H}_N$ by $\overline{\mathbb{W}}_s$ due to the fact that $\mathcal{H}_N$ is not continuous for the topology on empirical processes that we are considering. The lack of continuity of the interaction near the origin is dealt with by a truncation and regularization argument, similarly as in [18]. The lack of continuity due to the locality of the topology is handled thanks to the short-range nature of the Riesz interaction, by showing that large microscopic boxes effectively do not interact, the "self-screening" property alluded to before, via a shrinking procedure borrowed from [14]. We refer to Sect. 4 for more detail.

2 General Definitions

All the hypercubes considered will have their sides parallel to some fixed choice of axes in $\mathbb{R}^d$. For $R > 0$ we let K_R be the hypercube of center 0 and sidelength R. If $A \subset \mathbb{R}^d$ is a Borel set, we denote by $|A|$ its Lebesgue measure, and if A is a finite set, we denote by $|A|$ its cardinal.

2.1 (Random) (tagged) Point Configurations

2.1.1 Point Configurations

We refer to [8] for further details and proofs of the claims.

- If $A \subset \mathbb{R}^d$, we denote by $\mathcal{X}(A)$ the set of locally finite point configurations in A or equivalently the set of non-negative, purely atomic Radon measures on A giving an integer mass to singletons. We abbreviate $\mathcal{X}(\mathbb{R}^d)$ as $\mathcal{X}$.
- For $\mathcal{C} \in \mathcal{X}$, we will often write $\mathcal{C}$ for the Radon measure $\sum_{p \in \mathcal{C}} \delta_p$.
- The sets $\mathcal{X}(A)$ are endowed with the topology induced by the weak convergence of Radon measures (also known as vague convergence). These topological spaces are Polish, and we fix a distance $d_{\mathcal{X}}$ on $\mathcal{X}$ which is compatible with the topology on $\mathcal{X}$ (and whose restriction on $\mathcal{X}(A)$ is also compatible with the topology on $\mathcal{X}(A)$).
- For $x \in \mathbb{R}^d$ and $\mathcal{C} \in \mathcal{X}$, we denote by $\theta_x \cdot \mathcal{C}$ "the configuration $\mathcal{C}$ centered at x" (or "translated by $-x$"), namely

$$\theta_x \cdot \mathcal{C} := \sum_{p \in \mathcal{C}} \delta_{p-x}. \tag{2.1}$$

We will use the same notation for the action of $\mathbb{R}^d$ on Borel sets: if $A \subset \mathbb{R}^d$, we denote by $\theta_x \cdot A$ the translation of A by the vector $-x$.

2.1.2 Tagged Point Configurations

- When $\Omega \subset \mathbb{R}^d$ is fixed, we define $\overline{\mathcal{X}} := \Omega \times \mathcal{X}$ as the set of "tagged" point configurations with tags in Ω.
- We endow $\overline{\mathcal{X}}$ with the product topology and a compatible distance $d_{\overline{\mathcal{X}}}$.

Tagged objects will usually be denoted with bars (e.g., $\overline{P}$, $\overline{\mathbb{W}}$, ...).

2.1.3 Random Point Configurations

- We denote by $\mathcal{P}(\mathcal{X})$ the space of probability measures on $\mathcal{X}$, i.e., the set of laws of random point configurations.
- The set $\mathcal{P}(\mathcal{X})$ is endowed with the topology of weak convergence of probability measures (with respect to the topology on $\mathcal{X}$), see [18, Remark 2.7].
- We say that P in $\mathcal{P}(\mathcal{X})$ is *stationary* (and we write $P \in \mathcal{P}_{stat}(\mathcal{X})$) if its law is invariant by the action of $\mathbb{R}^d$ on $\mathcal{X}$ as defined in (2.1).

2.1.4 Random Tagged Point Configurations

- When $\Omega \subset \mathbb{R}^d$ is fixed, we define $\overline{\mathcal{M}}(\overline{\mathcal{X}})$ as the space of measures $\overline{P}$ on $\overline{\mathcal{X}}$ such that
  1. The first marginal of $\overline{P}$ is the Lebesgue measure on Ω.

2. For almost every $x \in \Omega$, the disintegration measure $\overline{P}^x$ is an element of $\mathcal{P}(\mathcal{X})$.
- We say that $\overline{P}$ in $\overline{\mathcal{M}}(\mathcal{X})$ is *stationary* (and we write $\overline{P} \in \overline{\mathcal{M}}_{stat}(\mathcal{X})$) if $\overline{P}^x$ is in $\mathcal{P}_{stat}(\mathcal{X})$ for almost every $x \in \Omega$.

Let us emphasize that, in general, the elements of $\overline{\mathcal{M}}(\mathcal{X})$ are *not* probability measures on $\overline{\mathcal{X}}$ (e.g., the first marginal is the Lebesgue measure on Ω).

2.1.5 Density of a Point Configuration

- For $\mathcal{C} \in \mathcal{X}$, we define $\mathrm{Dens}(\mathcal{C})$ (the *density* of $\mathcal{C}$) as

$$\mathrm{Dens}(\mathcal{C}) := \liminf_{R \to \infty} \frac{|\mathcal{C} \cap K_R|}{R^d}. \tag{2.2}$$

- For $m \in [0, +\infty]$, we denote by $\mathcal{X}_m$ the set of point configurations with density m.
- For $m \in (0, +\infty)$, the scaling map

$$\sigma_m : \mathcal{C} \mapsto m^{1/d} \mathcal{C} \tag{2.3}$$

is a bijection of $\mathcal{X}_m$ onto $\mathcal{X}_1$, with inverse $\sigma_{1/m}$.

2.1.6 Intensity of a Random Point Configuration

- For $P \in \mathcal{P}_{stat}(\mathcal{X})$, we define $\mathrm{Intens}(P)$ (the *intensity* of P) as

$$\mathrm{Intens}(P) := \mathbf{E}_P \left[\mathrm{Dens}(\mathcal{C}) \right].$$

- We denote by $\mathcal{P}_{stat,m}(\mathcal{X})$ the set of laws of random point configurations $P \in \mathcal{P}(\mathcal{X})$ that are stationary and such that $\mathrm{Intens}(P) = m$. For $P \in \mathcal{P}_{stat,m}(\mathcal{X})$, the stationarity assumption implies the formula

$$\mathbf{E}_P \left[\int_{\mathbb{R}^d} \varphi \, d\mathcal{C} \right] = m \int_{\mathbb{R}^d} \varphi(x) \, dx, \text{ for any } \varphi \in C_c^0(\mathbb{R}^d).$$

2.1.7 Intensity Measure of a Random Tagged Point Configuration

- For $\overline{P}$ in $\overline{\mathcal{M}}_{stat}(\mathcal{X})$, we define $\overline{\mathrm{Intens}}(\overline{P})$ (the *intensity measure* of $\overline{P}$) as

$$\overline{\mathrm{Intens}}(\overline{P})(x) = \mathrm{Intens}(\overline{P}^x),$$

which really should, in general, be understood in a dual sense: for any $f \in C_c(\mathbb{R}^d)$,

$$\int f \, d\overline{\mathrm{Intens}}(\overline{P}) := \int_\Omega f(x) \mathrm{Intens}(\overline{P}^x) dx.$$

- We denote by $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$ the set of laws of random tagged point configurations $\overline{P}$ in $\overline{\mathcal{M}}(\overline{\mathcal{X}})$ that are stationary and such that

$$\int_{\Omega} \overline{\text{Intens}}(\overline{P})(x)\, dx = 1.$$

- If $\overline{P}$ has intensity measure ρ, we denote by $\overline{\sigma}_{\rho}(\overline{P})$ the element of $\overline{\mathcal{M}}(\overline{\mathcal{X}})$ satisfying

$$\left(\overline{\sigma}_{\rho}(\overline{P})\right)^{x} = \sigma_{\rho(x)}\left(\overline{P}^{x}\right), \text{ for all } x \in \Omega, \tag{2.4}$$

where σ is as in (2.3).

2.2 Specific Relative Entropy

- Let P be in $\mathcal{P}_{stat}(\mathcal{X})$. The *specific relative entropy* $\text{ent}[P|\Pi]$ of P with respect to Π, the law of the Poisson point process of uniform intensity 1, is given by

$$\text{ent}[P|\Pi] := \lim_{R \to \infty} \frac{1}{|K_R|} \text{Ent}\left(P_{|K_R}|\Pi_{|K_R}\right), \tag{2.5}$$

where $P_{|K_R}$ denotes the process induced on (the point configurations in) K_R, and $\text{Ent}(\cdot|\cdot)$ denotes the usual relative entropy (or Kullbak–Leibler divergence) of two probability measures defined on the same probability space.
- It is known (see, e.g., [24]) that the limit (2.5) exists as soon as P is stationary, and also that the functional $P \mapsto \text{ent}[P|\Pi]$ is affine lower semi-continuous with compact sub-level sets (it is a good rate function).
- Let us observe that the empty point process has specific relative entropy 1 with respect to Π.
- If P is in $\mathcal{P}_{stat,m}(\mathcal{X})$, we have (see [18, Lemma 4.2.])

$$\text{ent}[P|\Pi] = \text{ent}[\sigma_m(P)|\Pi]m + m \log m + 1 - m, \tag{2.6}$$

where $\sigma_m(P)$ denotes the push-forward of P by (2.3).

2.3 Riesz Energy of (random) (tagged) Point Configurations

2.3.1 Riesz Interaction

We will use the notation Int (as "interaction") in two slightly different ways:

- If $\mathcal{C}_1, \mathcal{C}_2$ are some fixed point configurations, we let $\text{Int}[\mathcal{C}_1, \mathcal{C}_2]$ be the Riesz interaction between $\mathcal{C}_1$ and $\mathcal{C}_2$.

$$\text{Int}[\mathcal{C}_1, \mathcal{C}_2] := \sum_{p \in \mathcal{C}_1,\, q \in \mathcal{C}_2,\, p \neq q} \frac{1}{|p - q|^s}.$$

- If $\mathcal{C}$ is a fixed point configuration and A, B are two subsets of $\mathbb{R}^d$, we let $\text{Int}[A, B](\mathcal{C})$ be the Riesz interaction between $\mathcal{C} \cap A$ and $\mathcal{C} \cap B$; i.e.,

$$\text{Int}[A, B](\mathcal{C}) := \text{Int}[\mathcal{C} \cap A, \mathcal{C} \cap B] = \sum_{p \in \mathcal{C} \cap A, q \in \mathcal{C} \cap B, p \neq q} \frac{1}{|p - q|^s}.$$

- Finally, if $\tau > 0$, we let Int_τ be the truncation of the Riesz interaction at distances less than τ; i.e.,

$$\text{Int}_\tau[\mathcal{C}_1, \mathcal{C}_2] := \sum_{p \in \mathcal{C}_1, q \in \mathcal{C}_2, |p - q| \geq \tau} \frac{1}{|p - q|^s}. \tag{2.7}$$

2.3.2 Riesz Energy of a Finite Point Configuration

- Let $\omega_N = (x_1, \ldots, x_N)$ be in $(\mathbb{R}^d)^N$. We define its Riesz s-energy as

$$E_s(\omega_N) := \text{Int}[\omega_N, \omega_N] = \sum_{1 \leq i \neq j \leq N} \frac{1}{|x_i - x_j|^s}. \tag{2.8}$$

- For $A \subset \mathbb{R}^d$, we consider the N-*point minimal s-energy*

$$E_s(A, N) := \inf_{\omega_N \in A^N} E_s(\omega_N). \tag{2.9}$$

- The asymptotic minimal energy $C_{s,d}$ is defined as

$$C_{s,d} := \lim_{N \to \infty} \frac{E_s(K_1, N)}{N^{1+s/d}}. \tag{2.10}$$

   The limit in (2.10) exists as a positive real number (see [13,14]).
- By scaling properties of the s-energy, it follows that

$$\lim_{N \to \infty} \frac{E_s(K_R, N)}{N^{1+s/d}} = C_{s,d} R^{-s}. \tag{2.11}$$

2.3.3 Riesz Energy of Periodic Point Configurations

We first extend the definition of the Riesz energy to the case of periodic point configurations.

- We say that $\Lambda \subset \mathbb{R}^d$ is a d-dimensional Bravais lattice if $\Lambda = U\mathbb{Z}^d$, for some nonsingular $d \times d$ real matrix U. A fundamental domain for Λ is given by $\mathbf{D}_\Lambda := U[-\frac{1}{2}, \frac{1}{2})^d$, and the co-volume of Λ is $|\Lambda| := \text{vol}(\mathbf{D}_\Lambda) = |\det U|$.
- If $\mathcal{C}$ is a point configuration (finite or infinite) and Λ a lattice, we denote by $\mathcal{C} + \Lambda$ the configuration $\{p + \lambda \mid p \in \mathcal{C}, \lambda \in \Lambda\}$. We say that $\mathcal{C}$ is Λ-*periodic* if $\mathcal{C} + \Lambda = \mathcal{C}$.

- If $\mathcal{C}$ is Λ-periodic, it is easy to see that $\mathcal{C} = (\mathcal{C} \cap \mathbf{D}_\Lambda) + \Lambda$. The density of $\mathcal{C}$ is thus given by

$$\mathrm{Dens}(\mathcal{C}) = \frac{|\mathcal{C} \cap \mathbf{D}_\Lambda|}{|\Lambda|}.$$

Let Λ be a lattice and $\omega_N = \{x_1, \ldots, x_N\} \subset \mathbf{D}_\Lambda$.

- We define, as in [15] for $s > d$, the Λ-*periodic s-energy of* ω_N as

$$E_{s,\Lambda}(\omega_N) := \sum_{x \in \omega_N} \sum_{\substack{y \in \omega_N + \Lambda \\ y \neq x}} \frac{1}{|x - y|^s}. \tag{2.12}$$

- It follows (cf. [15]) that $E_{s,\Lambda}(\omega_N)$ can be re-written as

$$E_{s,\Lambda}(\omega_N) = N \zeta_\Lambda(s) + \sum_{x \neq y \in \omega_N} \zeta_\Lambda(s, x - y), \tag{2.13}$$

where

$$\zeta_\Lambda(s) := \sum_{0 \neq v \in \Lambda} |v|^{-s}$$

denotes the *Epstein zeta function* and

$$\zeta_\Lambda(s, x) := \sum_{v \in \Lambda} |x + v|^{-s}$$

denotes the *Epstein–Hurwitz zeta function* for the lattice Λ.
- Denoting the *minimum Λ-periodic s-energy* by

$$\mathcal{E}_{s,\Lambda}(N) := \min_{\omega_N \in \mathbf{D}_\Lambda^N} E_{s,\Lambda}(\omega_N), \tag{2.14}$$

it is shown in [15] that

$$\lim_{N \to \infty} \frac{\mathcal{E}_{s,\Lambda}(N)}{N^{1+s/d}} = C_{s,d} |\Lambda|^{-s/d}, \tag{2.15}$$

where $C_{s,d}$ is as in (2.10).

The constant $C_{s,d}$ for $s > d$ appearing in (2.10) and (2.15) is known only in the case $d = 1$, where $C_{s,1} = \zeta_{\mathbb{Z}}(s) = 2\zeta(s)$ and $\zeta(s)$ denotes the classical Riemann zeta function. For dimensions $d = 2, 4, 8,$ and 24, it has been conjectured (cf. [5,7] and references therein) that $C_{s,d}$ for $s > d$ is also given by an Epstein zeta function,

specifically, that $C_{s,d} = \zeta_{\Lambda_d}(s)$ for Λ_d denoting the equilateral triangular (or hexagonal) lattice, the D_4 lattice, the E_8 lattice, and the Leech lattice (all scaled to have co-volume 1) in the dimensions $d = 2, 4, 8$, and 24, respectively[1].

2.3.4 Riesz Energy of an Infinite Point Configuration

- Let $\mathcal{C}$ in $\mathcal{X}$ be an (infinite) point configuration. We define its Riesz s-energy as

$$\mathcal{W}_s(\mathcal{C}) := \liminf_{R \to \infty} \frac{1}{R^d} \sum_{p \neq q \in \mathcal{C} \cap K_R} \frac{1}{|p - q|^s} = \liminf_{R \to \infty} \frac{1}{R^d} \mathrm{Int}[K_R, K_R](\mathcal{C}). \quad (2.16)$$

If $\mathcal{C} = \emptyset$, we define $\mathcal{W}_s(\mathcal{C}) = 0$. The s-energy is non-negative and can be $+\infty$.
- We have, for any $\mathcal{C}$ in $\mathcal{X}$ and any $m \in (0, +\infty)$,

$$\mathcal{W}_s(\sigma_m \mathcal{C}) = m^{-(1+s/d)} \mathcal{W}_s(\mathcal{C}). \quad (2.17)$$

It is not difficult to verify (cf. [7, Lemma 9.1]), that if Λ is a lattice and ω_N is a N-tuple of points in $\mathbf{D}_\Lambda$, we have

$$\mathcal{W}_s(\omega_N + \Lambda) = \frac{1}{|\Lambda|} E_{s,\Lambda}(\omega_N). \quad (2.18)$$

In particular, we have (in view of (2.13))

$$\mathcal{W}_s(\Lambda) = |\Lambda|^{-1} \zeta_\Lambda(s). \quad (2.19)$$

2.3.5 Riesz Energy for Laws of Random Point Configurations

- Let P be in $\mathcal{P}(\mathcal{X})$; we define its Riesz s-energy as

$$\mathbb{W}_s(P) := \liminf_{R \to \infty} \frac{1}{R^d} \mathbf{E}_P \left[\mathrm{Int}[K_R, K_R](\mathcal{C}) \right]. \quad (2.20)$$

- Let $\overline{P}$ be in $\overline{\mathcal{M}}(\mathcal{X})$; we define its Riesz s-energy as

$$\overline{\mathbb{W}}_s(\overline{P}) := \int_\Omega \mathbb{W}_s(\overline{P}^x) \, dx. \quad (2.21)$$

- Let $\overline{P}$ be in $\overline{\mathcal{M}}(\mathcal{X})$ with intensity measure ρ. It follows from (2.17), (2.21), and definition (2.4) that

$$\overline{\mathbb{W}}_s(\overline{P}) = \int_\Omega \rho(x)^{1+s/d} \mathbb{W}_s\left(\left(\overline{\sigma}_\rho(\overline{P})\right)^x \right) dx. \quad (2.22)$$

---

[1] At an April 2018 ICERM workshop, S. Miller announced that, together with H. Cohn, A. Kumar, D. Radchenko and M. Viazovska, the E_8 and Leech lattices are universally optimal, which together with (3.1) verifies the conjecture for $d = 8$ and $d = 24$.

Let us emphasize that we define $\mathbb{W}_s$ as in (2.20) and *not* by $\mathbf{E}_P[\mathcal{W}_s]$. Fatou's lemma easily implies that

$$\mathbf{E}_P[\mathcal{W}_s] \leq \mathbb{W}_s(P), \tag{2.23}$$

and in fact, in the stationary case, we may show that equality holds (see Corollary 3.4).

2.3.6 Expression in Terms of the Two-Point Correlation Function

Let P be in $\mathcal{P}(\mathcal{X})$, and let us assume that the two-point correlation function of P, denoted by $\rho_{2,P}$, exists in some distributional sense. We may easily express the Riesz energy of P in terms of $\rho_{2,P}$ as follows:

$$\mathbb{W}_s(P) = \liminf_{R \to \infty} \frac{1}{R^d} \int_{K_R \times K_R} \frac{1}{|x - y|^s} \rho_{2,P}(x, y) \mathrm{d}x\mathrm{d}y. \tag{2.24}$$

If P is stationary, the expression can be simplified as

$$\mathbb{W}_s(P) = \liminf_{R \to \infty} \int_{[-R,R]^d} \frac{1}{|v|^s} \rho_{2,P}(v) \prod_{i=1}^{d} \left(1 - \frac{|v_i|}{R}\right) \mathrm{d}v\,, \tag{2.25}$$

where $\rho_{2,P}(v) = \rho_{2,P}(0, v)$ (we abuse notation and view $\rho_{2,P}$ as a function of one variable, by stationarity) and $v = (v_1, \ldots, v_d)$. Both (2.24) and (2.25) follow from the definitions and easy manipulations; proofs (in a slightly different context) can be found in [17]. Let us emphasize that the integral in the right-hand side of (2.24) is on two variables, whereas the one in (2.25) is a single integral, obtained by using stationarity and applying Fubini's formula, which gives the weight $\prod_{i=1}^{d} \left(1 - \frac{|v_i|}{R}\right)$.

2.4 The Rate Functions

2.4.1 Definitions

- For P in $\overline{\mathcal{M}}(\mathcal{X})$, we define

$$\overline{\mathbb{V}}(\overline{P}) := \int V(x)\mathrm{d}\big(\overline{\mathrm{Intens}}(\overline{P})\big)(x).$$

This is the energy contribution of the potential V.
- For $\overline{P}$ in $\overline{\mathcal{M}}_{stat,1}(\mathcal{X})$, we define

$$\overline{\mathcal{F}}_\beta(\overline{P}) := \beta\left(\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P})\right) + \int_\Omega \left(\mathrm{ent}[\overline{P}^x|\Pi] - 1\right)\mathrm{d}x + 1. \tag{2.26}$$

It is a free energy functional, the sum of an energy term $\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P})$ weighted by the inverse temperature β and an entropy term.

- If ρ is a probability density, we define $I_\beta(\rho)$ as

$$
I_\beta(\rho) := \int_\Omega \inf_{P \in \mathcal{P}_{stat,\rho(x)}(\mathcal{X})} (\beta \mathbb{W}_s(P) + \mathrm{ent}[P|\Pi] - 1)\, \mathrm{d}x
$$
$$
+ \beta \int_\Omega \rho(x) V(x)\, \mathrm{d}x + 1, \tag{2.27}
$$

which can be written as

$$
I_\beta(\rho) = \int_\Omega \rho(x) \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \rho(x)^{s/d} \mathbb{W}_s(P) + \mathrm{ent}[P|\Pi] \right) \mathrm{d}x
$$
$$
+ \beta \int_\Omega \rho(x) V(x)\, \mathrm{d}x + \int_\Omega \rho(x) \log \rho(x)\, \mathrm{d}x. \tag{2.28}
$$

This last equation may seem more complicated, but note that the inf inside the integral is taken on a fixed set, independent of ρ. The rate function I_β is obtained in Sect. 4.5 as a *contraction* (in the language of large deviation theory, see, e.g., [24, Section 3.1]) of the functional $\overline{\mathcal{F}}_\beta$, and (2.28) follows from (2.27) by scaling properties of $\mathbb{W}_s$ and $\mathrm{ent}[\cdot|\Pi]$.

2.4.2 Properties

Proposition 2.1 *For all $\beta > 0$, the functionals $\overline{\mathcal{F}}_\beta$ and I_β are good rate functions. Moreover, I_β is strictly convex.*

Proof It is proved in Proposition 3.3 that $\overline{\mathbb{W}}_s$ is lower semi-continuous on $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$. As for $\overline{\mathbb{V}}$, we may observe that, if $\overline{P} \in \overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$,

$$
\overline{\mathbb{V}}(\overline{P}) = \int_{\Omega \times \mathcal{X}} (V(x)|\mathcal{C} \cap K_1|)\, \mathrm{d}\overline{P}(x, \mathcal{C}),
$$

and that $(x, \mathcal{C}) \mapsto V(x)|\mathcal{C} \cap K_1|$ is lower semi-continuous on $\overline{\mathcal{X}}$. Thus $\overline{\mathbb{V}}$ is lower semi-continuous on $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$; moreover, it is known that $\mathrm{ent}[\cdot|\Pi]$ is lower semi-continuous (see Sect. 2.2). Thus $\overline{\mathcal{F}}_\beta$ is lower semi-continuous. Since $\overline{\mathbb{W}}_s$ and $\overline{\mathbb{V}}$ are bounded below, the sub-level sets of $\overline{\mathcal{F}}_\beta$ are included in those of $\mathrm{ent}[\cdot|\Pi]$, which are known to be compact (see again Sect. 2.2). Thus $\overline{\mathcal{F}}_\beta$ is a good rate function.

The functional I_β is easily seen to be lower semi-continuous, and since $\mathcal{W}_s$, ent and V are bounded below, the sub-level sets of I_β are included into those of $\int_\Omega \rho \log \rho$ which are known to be compact; thus I_β is a good rate function.

To prove that I_β is strictly convex in ρ, it is enough to prove that the first term in the right-hand side of (2.28) is convex (the second one is clearly affine, and the last one is well known to be strictly convex). We may observe that the map

$$
\rho \mapsto \beta \rho^{1+s/d} \mathbb{W}_s(P) + \rho\, \mathrm{ent}[P|\Pi] - \rho
$$

is convex for all P (because $\mathbb{W}_s(P)$ is non-negative), and the infimum of a family of convex functions is also convex; thus

$$\rho \mapsto \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \rho^{1+s/d} \mathbb{W}_s(P) + \rho \, \mathsf{ent}[P|\Pi] \right)$$

is convex in ρ, which concludes the proof. $\qquad\square$

3 Preliminaries on the Energy

3.1 General Properties

3.1.1 Minimal Energy of Infinite Point Configurations

In this section, we connect the minimization of $\mathcal{W}_s$ (defined at the level of infinite point configurations) with the asymptotics of the N-point minimal energy as presented in Sect. 2.3.2. Let us recall that the class $\mathcal{X}_m$ of point configurations with mean density m has been defined in Sect. 2.1.5.

Proposition 3.1 *We have*

$$\inf_{\mathcal{C} \in \mathcal{X}_1} \mathcal{W}_s(\mathcal{C}) = \min_{\mathcal{C} \in \mathcal{X}_1} \mathcal{W}_s(\mathcal{C}) = C_{s,d}, \tag{3.1}$$

where $C_{s,d}$ is as in (2.10). Moreover, for any d-dimensional Bravais lattice Λ of co-volume 1, there exists a minimizing sequence $\{\mathcal{C}_N\}_N$ for $\mathcal{W}_s$ over $\mathcal{X}_1$ such that $\mathcal{C}_N$ is $N^{1/d}\Lambda$-periodic for $N \geq 1$.

Proof Let Λ be a d-dimensional Bravais lattice Λ of co-volume 1, and for any N let ω_N be a N-point configuration minimizing $E_{s,\Lambda}$. We define

$$\mathcal{C}_N := N^{1/d} \left(\omega_N + \Lambda \right).$$

By construction, $\mathcal{C}_N$ is a $N^{1/d}\Lambda$-periodic point configuration of density 1. Using the scaling property (2.17) and (2.18), we have

$$\mathcal{W}_s(\mathcal{C}_N) = \frac{\mathcal{W}_s(\omega_N + \Lambda)}{N^{1+s/d}} = \frac{E_{s,\Lambda}(\omega_N)}{N^{1+s/d}}.$$

On the other hand, we have by assumption $E_{s,\Lambda}(\omega_N) = \mathcal{E}_{s,\Lambda}(N)$. Taking the limit $N \to \infty$ yields, in light of (2.15), $\lim_{N \to \infty} \mathcal{W}_s(\mathcal{C}_N) = C_{s,d}$. In particular, we have

$$\inf_{\mathcal{C} \in \mathcal{X}_1} \mathcal{W}_s(\mathcal{C}) \leq C_{s,d}. \tag{3.2}$$

To prove the converse inequality, let us consider an arbitrary $\mathcal{C}$ in $\mathcal{X}_1$. We have by definition (see (2.8) and (2.16)) and the scaling properties of E_s,

$$\mathcal{W}_s(\mathcal{C}) = \liminf_{R \to \infty} \frac{E_s\left(\mathcal{C} \cap K_R\right)}{R^d} = \liminf_{R \to \infty} \frac{1}{R^{d+s}} E_s\left(\frac{1}{R}\mathcal{C} \cap K_1\right),$$

and, again by definition (see (2.9)),

$$E_s\left(\frac{1}{R}\mathcal{C} \cap K_1\right) \geq \mathcal{E}_s\left(K_1, |\mathcal{C} \cap K_R|\right).$$

We thus obtain

$$\mathcal{W}_s(\mathcal{C}) \geq \liminf_{R \to \infty} \frac{\mathcal{E}_s\left(K_1, |\mathcal{C} \cap K_R|\right)}{|\mathcal{C} \cap K_R|^{1+s/d}} \left(\frac{|\mathcal{C} \cap K_R|}{R^d}\right)^{1+s/d}.$$

Using the definition (2.10) of $C_{s,d}$ we have

$$\liminf_{R \to \infty} \frac{\mathcal{E}_s\left(K_1, |\mathcal{C} \cap K_R|\right)}{|\mathcal{C} \cap K_R|^{1+s/d}} \geq C_{s,d},$$

and by the definition of density, since $\mathcal{C}$ is in $\mathcal{X}_1$, we have

$$\liminf_{R \to \infty} \left(\frac{|\mathcal{C} \cap K_R|}{R^d}\right)^{1+s/d} = 1.$$

This yields $\mathcal{W}_s(\mathcal{C}) \geq C_{s,d}$, and so (in view of (3.2))

$$\inf_{\mathcal{C} \in \mathcal{A}_1} \mathcal{W}_s(\mathcal{C}) = C_{s,d}. \tag{3.3}$$

It remains to prove that the infimum is achieved. Let us start with a sequence $\{\omega_M\}_{M \geq 1}$ such that ω_M is a M^d-point configuration in K_M satisfying

$$\lim_{M \to \infty} \frac{E_s(\omega_M)}{M^d} = C_{s,d}. \tag{3.4}$$

Such a sequence of point configurations exists by definition of $C_{s,d}$ as in (2.10), and by the scaling properties of E_s. We define a configuration $\mathcal{C}$ inductively as follows:

- Let $r_1, c_1, s_1 = 1$, and let us set $\mathcal{C} \cap K_{r_1}$ to be ω_1.
- Assume that r_N, s_N, c_N and $\mathcal{C} \cap K_{r_N}$ have been defined. We let

$$s_{N+1} = \lceil c_{N+1}r_N + (c_{N+1}r_N)^{\frac{1}{2}} \rceil, \tag{3.5}$$

with $c_{N+1} > 1$ to be chosen later. We also let r_{N+1} be a multiple of s_{N+1} large enough, to be chosen later. We tile $K_{r_{N+1}}$ by hypercubes of sidelength s_{N+1}, and we define $\mathcal{C} \cap K_{r_{N+1}}$ as follows:

- In the central hypercube of sidelength s_{N+1}, we already have the points of $\mathcal{C} \cap K_{r_N}$ (because $r_N \leq s_{N+1}$), and we do not add any points. In particular, this ensures that each step of our construction is compatible with the previous ones.
- In all the other hypercubes, we paste a copy of $\omega_{c_{N+1}r_N}$ "centered" in the hypercube in such a way that

$$\text{all the points are at distance } \geq (c_{N+1}r_N)^{\frac{1}{2}} \text{ from the boundary.} \quad (3.6)$$

This is always possible because $\omega_{c_{N+1}r_N}$ lives, by definition, in a hypercube of sidelength $c_{N+1}r_N$ and because we have chosen s_{N+1} as in (3.5). We claim that the number of points in $K_{r_{N+1}}$ is always less than r_{N+1}^d (as can easily be checked by induction) and is bounded below by

$$\left(\left(\frac{r_{N+1}}{s_{N+1}} \right)^d - 1 \right) (c_{N+1}r_N)^d.$$

Thus it is easy to see that if c_{N+1} is chosen large enough and if r_{N+1} is a large enough multiple of s_{N+1}, then

$$\text{the number of points in } r_{N+1} \text{ is } r_{N+1}^d(1 - o_N(1)). \quad (3.7)$$

Let us now give an upper bound on the interaction energy $\text{Int}[K_{r_{N+1}}, K_{r_{N+1}}](\mathcal{C})$. We recall that we have tiled $K_{r_{N+1}}$ by hypercubes of sidelength s_{N+1}.

- Each hypercube has a self-interaction energy given by $E_s(\omega_{c_{N+1}r_N})$, except the central one, whose self-interaction energy is bounded by $O(r_N^d)$ (as can be seen by induction).
- The interaction of a given hypercube with the union of all the others can be controlled because, by construction (see (3.6)) the configurations pasted in two disjoint hypercubes are far way from each other. We can compare it to

$$\int_{r=(c_{N+1}r_N)^{\frac{1}{2}}}^{+\infty} \frac{1}{r^s} s_{N+1}^d r^{d-1} \mathrm{d}r,$$

and an elementary computation shows that it is negligible with respect to s_{N+1}^d (because $d < s$).
We thus have

$$\text{Int}[K_{r_{N+1}}, K_{r_{N+1}}](\mathcal{C}) \leq \left(\left(\frac{r_{N+1}}{s_{N+1}} \right)^d - 1 \right) E_s(\omega_{c_{N+1}r_N}) + O(r_N^d)$$

$$+ \left(\frac{r_{N+1}}{s_{N+1}} \right)^d o_N\left(s_{N+1}^d \right).$$

We may now use (3.4) and get that

$$\frac{1}{r_{N+1}^d}\text{Int}[K_{r_{N+1}}, K_{r_{N+1}}](\mathcal{C}) \leq C_{s,d} + o_N(1). \tag{3.8}$$

Let $\mathcal{C}$ be the point configuration constructed as above. Taking the limit as $N \to \infty$ in (3.7) shows that $\mathcal{C}$ is in $\mathcal{X}_1$, and (3.8) implies that $\mathcal{W}_s(\mathcal{C}) \leq C_{s,d}$, which concludes the proof of (3.1). $\qquad\square$

3.1.2 Energy of Random Point Configurations

In the following lemma, we prove that for stationary P the lim inf defining $\mathbb{W}_s(P)$ as in (2.20) is actually a limit, and that the convergence is uniform of sublevel sets of $\mathbb{W}_s$ (which will be useful for proving lower semi-continuity).

Lemma 3.2 *Let P be in $\mathcal{P}_{stat}(\mathcal{X})$. The following limit exists in $[0, +\infty]$:*

$$\mathbb{W}_s(P) := \lim_{R \to \infty} \frac{1}{R^d} \mathbf{E}_P\left[\text{Int}[K_R, K_R]\right]. \tag{3.9}$$

Moreover, we have as $R \to \infty$,

$$\left| \mathbb{W}_s(P) - \frac{1}{R^d}\mathbf{E}_P\left[\text{Int}[K_R, K_R]\right] \right| \leq C \left(\mathbb{W}_s(P)^{\frac{2}{1+s/d}} + \mathbb{W}_s(P) \right) o_R(1), \tag{3.10}$$

with $o_R(1)$ depending only on s, d.

Proof We begin by showing that the quantity

$$\frac{1}{n^d}\mathbf{E}_P\left[\text{Int}[K_n, K_n](\mathcal{C})\right]$$

is nondecreasing for integer values of n.

For $n \geq 1$, let $\{\tilde{K}_v\}_{v \in \mathbb{Z}^d \cap K_n}$ be a tiling of K_n by unit hypercubes, indexed by the centers $v \in \mathbb{Z}^d \cap K_n$ of the hypercubes, and let us split $\text{Int}[K_n, K_n]$ as

$$\text{Int}[K_n, K_n] = \sum_{v, v' \in \mathbb{Z}^d \cap K_n} \text{Int}[\tilde{K}_v, \tilde{K}_{v'}].$$

Using the stationarity assumption and writing $v = (v_1, \ldots, v_d)$ and $|v| := \max_i |v_i|$, we obtain

$$\mathbf{E}_P\left[\sum_{v, v' \in \mathbb{Z}^d \cap K_n} \text{Int}[\tilde{K}_v, \tilde{K}_{v'}] \right] = \sum_{v \in \mathbb{Z}^d \cap K_{2n}} \mathbf{E}_P\left[\text{Int}[\tilde{K}_0, \tilde{K}_v]\right] \prod_{i=1}^d (n - |v_i|).$$

We thus get

$$\frac{1}{n^d}\mathbf{E}_P\left[\mathrm{Int}[K_n,K_n]\right] = \sum_{v\in\mathbb{Z}^d\cap K_{2n}}\mathbf{E}_P\left[\mathrm{Int}[\tilde{K}_0,\tilde{K}_v]\right]\prod_{i=1}^d\left(1-\frac{|v_i|}{n}\right), \quad (3.11)$$

and it is clear that this quantity is nondecreasing in n; in particular, the limit as $n\to\infty$ exists in $[0,+\infty]$. We may also observe that $R\mapsto\mathrm{Int}[K_R,K_R]$ is nondecreasing in R. It is then easy to conclude that the limit of (3.9) exists in $[0,+\infty]$.

Let us now quantify the speed of convergence. First, we observe that for $|v|\geq 2$, we have

$$\mathbf{E}_P\left[\mathrm{Int}[\tilde{K}_0,\tilde{K}_v]\right] \leq O\left(\frac{1}{|v-1|^s}\right)\mathbf{E}_P[N_0 N_v],$$

where N_0, N_v denotes the number of points in $\tilde{K}_0$, $\tilde{K}_v$. Indeed, the points of $\tilde{K}_0$ and $\tilde{K}_v$ are at distance at least $|v-1|$ from each other (up to a multiplicative constant depending only on d).

On the other hand, Hölder's inequality and the stationarity of P imply

$$\|N_0 N_v\|_{L^1(P)} \leq \|N_0\|_{L^{1+s/d}(P)}\|N_v\|_{L^{1+s/d}(P)} = \|N_0\|_{L^{1+s/d}(P)}^2,$$

and thus we have $\mathbf{E}_P[N_0 N_v] \leq \mathbf{E}_P[N_0]^{\frac{2}{1+s/d}}$. On the other hand, it is easy to check that for P stationary,

$$\mathbf{E}_P[N_0^{1+s/d}] \leq C\mathbb{W}_s(P)$$

for some constant C depending on d, s. Indeed, the interaction energy in the hypercube $\tilde{K}_0$ is bounded below by some constant times $N_0^{1+s/d}$, and (3.11) shows that

$$\mathbb{W}_s(P) \geq \mathbf{E}_P\left[\mathrm{Int}[\tilde{K}_0,\tilde{K}_0]\right].$$

We thus get

$$\mathbb{W}_s(P) - \sum_{v\in\mathbb{Z}^d\cap K_{2n}}\mathbf{E}_P\left[\mathrm{Int}[\tilde{K}_0,\tilde{K}_v]\right]\prod_{i=1}^d\left(1-\frac{|v_i|}{n}\right)$$

$$\leq \mathbb{W}_s(P)^{\frac{2}{1+s/d}}\left(\sum_{v\in\mathbb{Z}^d\cap K_{2n},|v|\geq 2}\frac{1}{|v-1|^s}\left(1-\prod_{i=1}^d\left(1-\frac{|v_i|}{n}\right)\right)+\sum_{|v|\geq 2n}\frac{1}{|v|^s}\right)$$

$$+\frac{1}{n}\sum_{|v|=1}\mathbf{E}_P\left[\mathrm{Int}[\tilde{K}_0,\tilde{K}_v]\right].$$

It is not hard to see that the parenthesis in the right-hand side goes to zero as $n \to \infty$. On the other hand, we have

$$\sum_{|v|=1} \mathbf{E}_P \left[\operatorname{Int}[\tilde{K}_0, \tilde{K}_v] \right] \leq \mathbb{W}_s(P).$$

Thus we obtain

$$\mathbb{W}_s(P) - \frac{1}{n^d} \mathbf{E}_P \left[\operatorname{Int}[K_n, K_n] \right] \leq \left(\mathbb{W}_s(P)^{\frac{2}{1+s/d}} + \mathbb{W}_s(P) \right) o_n(1),$$

with a $o_n(1)$ depending only on d, s, and it is then not hard to get (3.10). □

For any $R > 0$, the quantity $\operatorname{Int}[K_R, K_R]$ is continuous and bounded below on $\mathcal{X}$; thus the map

$$P \mapsto \frac{1}{R^d} \mathbf{E}_P \left[\operatorname{Int}[K_R, K_R] \right]$$

is lower semi-continuous on $\mathcal{P}(\mathcal{X})$. The second part of Lemma 3.2 shows that we may approximate $\mathbb{W}_s(P)$ by $\frac{1}{R^d} \mathbf{E}_P \left[\operatorname{Int}[K_R, K_R] \right]$ up to an error $o_R(1)$, uniformly on sub-level sets of $\mathbb{W}_s$. The next proposition follows easily.

Proposition 3.3 *1. The functional $\mathbb{W}_s$ is lower semi-continuous on $\mathcal{P}_{stat,1}(\mathcal{X})$.*
2. The functional $\overline{\mathbb{W}}_s$ is lower semi-continuous on $\overline{\mathcal{M}}_{stat,1}(\mathcal{X})$.

We may also prove the following equality (which settles a question raised in Sect. 2.3.5).

Corollary 3.4 *Let P be in $\mathcal{P}_{stat,1}(\mathcal{X})$. Then we have*

$$\mathbb{W}_s(P) = \lim_{R \to \infty} \frac{1}{R^d} \mathbf{E}_P \left[\operatorname{Int}[K_R, K_R](\mathcal{C}) \right] = \mathbf{E}_P \left[\liminf_{R \to \infty} \frac{1}{R^d} \operatorname{Int}[K_R, K_R](\mathcal{C}) \right].$$

Proof As was observed in (2.23), Fatou's lemma implies that

$$\mathbf{E}_P \left[\liminf_{R \to \infty} \frac{1}{R^d} \operatorname{Int}[K_R, K_R](\mathcal{C}) \right] \leq \lim_{R \to \infty} \frac{1}{R^d} \mathbf{E}_P \left[\operatorname{Int}[K_R, K_R](\mathcal{C}) \right] = \mathbb{W}_s(P)$$

(the last equality is by definition). On the other hand, with the notation of the proof of Lemma 3.2, we have for any integer n and any $\mathcal{C}$ in $\mathcal{X}$,

$$\frac{1}{n^d} \operatorname{Int}[K_n, K_n](\mathcal{C}) = \frac{1}{n^d} \sum_{v,v' \in \mathbb{Z}^d \cap K_R} \operatorname{Int}[\tilde{K}_v, \tilde{K}_{v'}],$$

and the right-hand side is dominated under P (as observed in the previous proof); thus the dominated convergence theorem applies. □

3.2 Derivation of the Infinite-Volume Limit of the Energy

The following result is central in our analysis. It connects the asymptotics of the N-point interaction energy $\{\mathcal{H}_N(\mathbf{X}_N)\}_N$ with the infinite-volume energy $\overline{\mathbb{W}}_s(\overline{P})$ of an infinite-volume object: the limit point $\overline{P}$ of the tagged empirical processes $\{\overline{\mathrm{Emp}}_N(\mathbf{X}_N)\}_N$.

Proposition 3.5 *For any $N \geq 1$, let $\mathbf{X}_N = (x_1, \ldots, x_N)$ be in Ω^N, let μ_N be the empirical measure and $\overline{P}_N$ be the tagged empirical process associated with $\mathbf{X}_N$; i.e.,*

$$\mu_N := \mathrm{emp}(\mathbf{X}_N), \quad \overline{P}_N := \overline{\mathrm{Emp}}_N(\mathbf{X}_N),$$

as defined in (1.8) *and* (1.10). *Let us assume that*

$$\liminf_{N \to \infty} \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}} < +\infty.$$

Then, up to extraction of a subsequence,

- *$\{\mu_N\}_N$ converges weakly to some μ in $\mathcal{M}(\Omega)$,*
- *$\{\overline{P}_N\}_N$ converges weakly to some $\overline{P}$ in $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$,*
- *$\mathrm{Intens}(\overline{P}) = \mu$.*

Moreover, we have

$$\liminf_{N \to \infty} \frac{\mathcal{H}_N(x_1, \ldots, x_N)}{N^{1+s/d}} \geq \overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}). \tag{3.12}$$

Proof Up to extracting a subsequence, we may assume that $\mathcal{H}_N(\mathbf{X}_N) = O(N^{1+s/d})$. First, by positivity of the Riesz interaction, we have for $N \geq 1$,

$$\int_\Omega V \, \mathrm{d}\mu_N \leq \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}},$$

and thus $\int_\Omega V \, \mathrm{d}\mu_N$ is bounded. By (1.5) and (1.6) we know that V is bounded below and has compact sub-level sets. An easy application of Markov's inequality shows that $\{\mu_N\}_N$ is tight, and thus it converges (up to another extraction). It is not hard to check that $\{\overline{P}_N\}_N$ converges (up to extraction) to some $\overline{P}$ in $\overline{\mathcal{M}}(\overline{\mathcal{X}})$ (indeed, the average number of points per unit volume is constant, which implies tightness, see, e.g., [18, Lemma 4.1]) whose stationarity is clear (see again, e.g., [18]).

Let $\bar{\rho}$ be the intensity measure of $\overline{P}$ (in the sense of Sect. 2.1.7). We want to prove that $\bar{\rho} = \mu$ (which will in particular imply that $\overline{P}$ is in $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$). It is a general fact that $\bar{\rho} \leq \mu$ (see, e.g., [19, Lemma 3.7]), but it could happen that a positive fraction of the points cluster together, resulting in the existence of a singular part in μ that is missed by $\bar{\rho}$ so that $\bar{\rho} < \mu$. However, in the present case, we can easily bound the moment (under $\overline{P}_N$) of order $1 + s/d$ of the number of points in a given hypercube K_R. Indeed, let $\{\tilde{K}_i\}_{i \in I}$ be a covering of Ω by disjoint hypercubes of sidelength $RN^{-1/d}$, and let $n_i := N\mu_N\left(\tilde{K}_i\right)$ denote the number of points from $\mathbf{X}_N$ in $\tilde{K}_i$. We have, by positivity of the Riesz interaction,

$$\mathcal{H}_N(\mathbf{X}_N) \geq \sum_{i \in I} \mathrm{Int}[\tilde{K}_i, \tilde{K}_i] \geq C \sum_{i \in I} \frac{n_i^{1+s/d} N^{s/d}}{R^s}$$

for some constant $C > 0$ (depending only on s and d) because the minimal interaction energy of n points in $\tilde{K}_i$ is proportional to $\frac{n^{1+s/d} N^{s/d}}{R^s}$ (see (2.10), (2.11)). Since $\mathcal{H}_N(\mathbf{X}_N) = O(N^{1+s/d})$ by assumption, we get that $\sum_{i \in I} n_i^{1+s/d} = O(N)$, with an implicit constant depending only on R. This implies that $x \mapsto N\mu_N(B(x, RN^{-1/d}))$ is uniformly (in N) locally integrable on Ω for all $R > 0$, and arguing as in [19, Lemma 3.7], we deduce that $\bar{\rho} = \mu$.

We now turn to proving (3.12). Using the positivity and scaling properties of the Riesz interaction and a Fubini-type argument, we may write, for any $R > 0$,

$$\mathrm{Int}[\Omega, \Omega](\mathbf{X}_N) \geq N^{1+s/d} \int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \mathrm{Int}[K_R, K_R](\mathcal{C}) \, d\overline{P}_N(x, \mathcal{C}).$$

Of course we have, for any $M > 0$,

$$\int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \mathrm{Int}[K_R, K_R](\mathcal{C}) \, d\overline{P}_N(x, \mathcal{C}) \geq \int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \left(\mathrm{Int}[K_R, K_R](\mathcal{C}) \wedge M \right) d\overline{P}_N(x, \mathcal{C}),$$

and thus the weak convergence of $\overline{P}_N$ to $\overline{P}$ ensures that

$$\int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \mathrm{Int}[K_R, K_R](\mathcal{C}) \, d\overline{P}_N(x, \mathcal{C})$$
$$\geq \int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \left(\mathrm{Int}[K_R, K_R](\mathcal{C}) \wedge M \right) d\overline{P}(x, \mathcal{C}) + o_N(1).$$

Since this is true for all M, we obtain

$$\liminf_{N \to \infty} \frac{\mathrm{Int}[\Omega, \Omega](\mathbf{X}_N)}{N^{1+s/d}} \geq \int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \left(\mathrm{Int}[K_R, K_R](\mathcal{C}) \right) d\overline{P}(x, \mathcal{C}).$$

Sending R to $+\infty$ and using Proposition 3.1, we get

$$\liminf_{N \to \infty} \frac{\mathrm{Int}[\Omega, \Omega](\mathbf{X}_N)}{N^{1+s/d}} \geq \liminf_{R \to \infty} \int_{\Omega \times \mathcal{X}} \frac{1}{R^d} \left(\mathrm{Int}[K_R, K_R](\mathcal{C}) \right) d\overline{P}(x, \mathcal{C}) =: \overline{\mathbb{W}}_s(\overline{P}).$$

$$(3.13)$$

On the other hand, the weak convergence of μ_N to μ and Assumption 1.5 ensure that

$$\liminf_{N \to \infty} \int_{\Omega} V \, d\mu_N \geq \int_{\Omega} V \, d\mu. \tag{3.14}$$

Combining (3.13) and (3.14) gives (3.12). $\qquad\square$

Proposition 3.5 can be viewed as a Γ-lim inf result (in the language of Γ-convergence). We will prove later (e.g., in Proposition 4.5, which is in fact a much stronger statement) the corresponding Γ-lim sup.

4 Proof of the Large Deviation Principles

As in [18], the main obstacle to proving Theorem 1.1 is to deal with the lack of upper semi-continuity of the interaction, namely that there is no upper bound of the type

$$\mathcal{H}_N(\mathbf{X}_N) \lesssim N^{1+s/d} \left(\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) \right)$$

that holds in general under the mere condition that $\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \approx \overline{P}$ (cf. (1.10) for a definition of the tagged empirical process). This yields a problem for proving the large deviation lower bound (in contrast, *lower* semi-continuity holds, and the proof of the large deviations upper bound is quite simple). Let us briefly explain why.

Firstly, due its singularity at 0, the interaction is not uniformly continuous with respect to the topology on the configurations. Indeed, a pair of points at distance ε yields a ε^{-s} energy, but a pair of points at distance 2ε has energy $(2\varepsilon)^{-s}$, with $|\varepsilon^{-s} - (2\varepsilon)^{-s}| \to \infty$, although these two point configurations are very close for the topology on $\mathcal{X}$.

Secondly, the energy is nonadditive: we have in general

$$\mathrm{Int}[\mathcal{C}_1 \cup \mathcal{C}_2, \mathcal{C}_1 \cup \mathcal{C}_2] \neq \mathrm{Int}[\mathcal{C}_1, \mathcal{C}_1] + \mathrm{Int}[\mathcal{C}_2, \mathcal{C}_2].$$

Yet the knowledge of $\overline{\mathrm{Emp}}_N$ (through the fact that $\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)$) yields only *local* information on $\mathbf{X}_N$ and does not allow one to reconstruct $\mathbf{X}_N$ *globally*. Roughly speaking, it is like partitioning Ω into hypercubes and having a family of point configurations, each belonging to some hypercube, but without knowing the precise configuration-hypercube pairing. Since the energy is nonadditive (there are nontrivial hypercube-hypercube interactions in addition to the hypercubes' self-interactions), we cannot (in general) deduce $\mathcal{H}_N(\mathbf{X}_N)$ from the mere knowledge of the tagged empirical process.

In Sect. 4.3, the singularity problem is dealt with by using a regularization procedure similar to that of [18], while the nonadditivity is shown to be negligible due to the short-range nature of the Riesz potential for $s > d$.

4.1 An LDP for the Reference Measure

Let $\mathbf{Leb}_{\Omega^N}$ be the Lebesgue measure on Ω^N, and let $\bar{\mathfrak{Q}}_N$ be the push-forward of $\mathbf{Leb}_{\Omega^N}$ by the "tagged empirical process" map $\overline{\mathrm{Emp}}_N$ defined in (1.10). Let us recall that Ω is not necessarily bounded; hence $\mathbf{Leb}_{\Omega^N}$ may have an infinite mass, and thus there is no natural way of making $\bar{\mathfrak{Q}}_N$ a probability measure.

Proposition 4.1 *Let $\overline{P}$ be in $\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})$. We have*

$$\lim_{\varepsilon \to 0} \liminf_{N \to \infty} \frac{1}{N} \log \bar{\mathfrak{Q}}_N \left(B(\overline{P}, \varepsilon) \right) = \lim_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \bar{\mathfrak{Q}}_N \left(B(\overline{P}, \varepsilon) \right)$$

$$= - \int_{\Omega} \left(\mathrm{ent}[\overline{P}^x | \Pi] - 1 \right) \mathrm{d}x - 1. \quad (4.1)$$

We recall that $\overline{P}^x$ is the disintegration measure of $\overline{P}$ at the point x, or the "fiber at x" (which is a measure on $\mathcal{X}$) of $\overline{P}$ (which is a measure on $\Omega \times \mathcal{X}$), see Sect. 2.1.4.

Proof If Ω is bounded, Proposition 4.1 follows from the analysis of [18, Section 7.2]; see in particular [18, Lemma 7.8]. The only difference is that the Lebesgue measure on Ω used in [18] is normalized, which yields an additional factor of $\log |\Omega|$ in the rate function. The proof extends readily to a nonbounded Ω because the topology of weak convergence on $\overline{\mathcal{M}(\mathcal{X})}$ is defined with respect to test functions that are compactly supported on Ω. $\square$

4.2 An LDP Upper Bound

Proposition 4.2 *Let $\overline{P}$ be in $\overline{\mathcal{M}}_{stat,1}(\mathcal{X})$. We have*

$$\lim_{\varepsilon \to 0} \limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(B(\overline{P}, \varepsilon)) \leq -\overline{\mathcal{F}}_\beta(\overline{P}) + \limsup_{N \to \infty} \left(-\frac{\log Z_{N,\beta}}{N} \right). \quad (4.2)$$

Proof Using the definition of $\overline{\mathfrak{P}}_{N,\beta}$ as the push-forward of $\mathbb{P}_{N,\beta}$ by $\overline{\mathrm{Emp}}_N$, we may write

$$\overline{\mathfrak{P}}_{N,\beta}(B(\overline{P}, \varepsilon)) = \frac{1}{Z_{N,\beta}} \int_{\Omega^N \cap \{\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)\}} \exp\left(-\beta N^{-s/d} \mathcal{H}_N(\mathbf{X}_N) \right) d\mathbf{X}_N.$$

From Propositions 3.5 and 3.3, we know that for any sequence $\mathbf{X}_N$ such that $\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)$, we have

$$\liminf_{N \to \infty} \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}} \geq \overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) + o_\varepsilon(1).$$

We may thus write

$$\limsup_{N \to \infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(B(\overline{P}, \varepsilon)) \leq -\beta \left(\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) \right)$$
$$+ \limsup_{N \to \infty} \int_{\Omega^N \cap \{\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)\}} d\mathbf{X}_N + \limsup_{N \to \infty} \left(-\frac{\log Z_{N,\beta}}{N} \right) + o_\varepsilon(1).$$

Using Proposition 4.1, we know that

$$\limsup_{N \to \infty} \frac{1}{N} \log \int_{\Omega^N \cap \{\overline{\mathrm{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)\}} d\mathbf{X}_N = -\int_\Omega \left(\mathrm{ent}[\overline{P}^x | \Pi] - 1 \right) - 1 + o_\varepsilon(1).$$

We thus obtain, sending $\varepsilon \to 0$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(B(\overline{P}, \varepsilon)) \leq -\beta \left(\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) \right) - \int_\Omega \left(\text{ent}[\overline{P}^x | \Pi] - 1 \right) - 1$$
$$+ \limsup_{N\to\infty} \left(-\frac{\log Z_{N,\beta}}{N} \right),$$

which, in view of the definition of $\overline{\mathcal{F}}_\beta$ as in (2.26), yields (4.2). $\square$

4.3 An LDP Lower Bound

The goal of the present section is to prove a matching LDP lower bound:

Proposition 4.3 *Let $\overline{P}$ be in $\overline{\mathcal{M}}_{stat,1}(\mathcal{X})$. We have*

$$\lim_{\varepsilon\to 0} \liminf_{N\to\infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(B(\overline{P}, \varepsilon)) \geq -\overline{\mathcal{F}}_\beta(\overline{P}) + \liminf_{N\to\infty} \left(-\frac{\log Z_{N,\beta}}{N} \right). \quad (4.3)$$

For $N \geq 1$ and $\delta > 0$, let us define the set $T_{N,\delta}(\overline{P})$ as

$$T_{N,\delta}(\overline{P}) = \left\{ \mathbf{X}_N \mid \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}} \leq \overline{\mathcal{F}}_\beta(\overline{P}) + \delta \right\}. \quad (4.4)$$

We will rely on the following result:

Proposition 4.4 *Let $\overline{P}$ be in $\overline{\mathcal{M}}_{stat,1}(\mathcal{X})$. For all $\varepsilon, \delta > 0$, we have*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbf{Leb}_{\Omega^N} \left(\{\overline{\text{Emp}}_N(\mathbf{X}_N) \in B(\overline{P}, \varepsilon)\} \cap T_{N,\delta}(\overline{P}) \right)$$
$$\geq -\int_\Omega \left(\text{ent}[\overline{P}^x | \Pi] - 1 \right) \mathrm{d}x - 1. \quad (4.5)$$

Proof We may assume that Ω is compact and that the intensity measure of $\overline{P}$, denoted by $\bar{\rho}$, is continuous, compactly supported, and bounded below. Indeed, we can always approximate $\overline{P}$ by random point processes satisfying these additional assumptions. For any $N \geq 1$, we let $\bar{\rho}_N(x) := \bar{\rho}(xN^{-1/d})$ and we let $\Omega_N := N^{1/d}\Omega$.

In fact, for simplicity we will assume that Ω is some large hypercube. The argument below readily extends to the case where Ω can be tiled by small hypercubes, and any C^1 domain can be tiled by small hypercubes up to some "boundary parts" that are negligible for our concerns (a precise argument is given, e.g., in [18, Section 6]).

For $R > 0$, we let $\{\tilde{K}_i\}_{i\in I}$ be a partition of Ω_N by hypercubes of sidelength R. For R, M, we denote by $\overline{P}_{R,M}$ the restriction[2] to K_R of $\overline{P}$, conditioned to the event

$$\left\{ |\mathcal{C} \cap K_R| \leq MR^d \right\}. \quad (4.6)$$

---

[2] That is, $\overline{P}_{R,M} \in \overline{\mathcal{M}}(\Omega \times \mathcal{X}[K_R])$.

Step 1. *Generating microstates.*

For any $\varepsilon > 0$, for any $M, R > 0$, for any $\nu > 0$, for any $N \geq 1$, there exists a family $\mathcal{A} = \mathcal{A}(\varepsilon, M, R, \nu, N)$ of point configurations $\mathcal{C}$ such that:

1. $\mathcal{C} = \sum_{i \in I} \mathcal{C}_i$, where $\mathcal{C}_i$ is a point configuration in $\tilde{K}_i$.
2. $|\mathcal{C}| = N$.
3. The "discretized" empirical process is close to $\overline{P}_{R,M}$:

$$\overline{P}_d(\mathcal{C}) := \frac{1}{|I|} \sum_{i \in I} \delta_{(N^{-1/d} x_i, \, \theta_{x_i} \cdot \mathcal{C}_i)} \text{ belongs to } B(\overline{P}_{R,M}, \nu), \qquad (4.7)$$

     where x_i denotes the center of $\tilde{K}_i$.
4. The associated empirical process is close to $\overline{P}$

$$\overline{P}_c(\mathcal{C}) := \int_{\Omega} \delta_{(x, \, \theta_{N^{1/d} x} \cdot \mathcal{C})} \, dx \text{ belongs to } B(\overline{P}, \varepsilon). \qquad (4.8)$$

     Note that $\overline{P}_c(\mathcal{C}) = \overline{\mathrm{Emp}}_N(N^{-1/d} \mathcal{C})$.
5. The volume of $\mathcal{A}$ satisfies, for any $\varepsilon > 0$,

$$\liminf_{M \to \infty} \liminf_{R \to \infty} \frac{1}{R^d} \lim_{\nu \to 0} \lim_{N \to \infty} \frac{1}{|I|} \log \mathbf{Leb}_{\Omega_N^N}(\mathcal{A}) \geq - \int_{\Omega} \left(\mathrm{ent}[\overline{P}^x | \Pi] - 1 \right) - 1. \qquad (4.9)$$

This is essentially [18, Lemma 6.3] with minor modifications (e.g., the Lebesgue measure in [18] is normalized, which yields an additional logarithmic factor in the formulas).

We will make the following assumption on $\mathcal{A}$:

$$|\mathcal{C}_i| \leq 2MR^d \text{ for all } i \in I. \qquad (4.10)$$

Indeed, for fixed M, when $\overline{P}_d$ is close to $\overline{P}_{R,M}$ (for which (4.6) holds), the fraction of hypercubes on which (4.10) fails to hold as well as the ratio of excess points over the total number of points (namely N) are both small. We may then "redistribute" these excess points among the other hypercubes without affecting (4.8) and changing the energy estimates below only by a negligible quantity.

Step 2. *First energy estimate.*

For any $R, M, \tau > 0$, the map defined by

$$\mathcal{C} \in \mathcal{X}(K_R) : \longrightarrow \mathrm{Int}_\tau[\mathcal{C}, \mathcal{C}] \wedge \frac{(2MR^d)^2}{\tau^s}$$

(where Int_τ is as in (2.7)) is continuous on $\mathcal{X}(K_R)$ and *bounded* (this is precisely the reason for requiring that the number of points are bounded). We may thus write, in view of (4.6), (4.7), and (4.10),

$$\int_{\Omega \times \mathcal{X}(K_R)} \mathrm{Int}_\tau \, \mathrm{d}\overline{P}_d = \int_{\Omega \times \mathcal{X}(K_R)} \mathrm{Int}_\tau \wedge \frac{(2MR^d)^2}{\tau^s} \mathrm{d}\overline{P}_d$$

$$= \int_{\Omega \times \mathcal{X}(K_R)} \mathrm{Int}_\tau \wedge \frac{(2MR^d)^2}{\tau^s} \mathrm{d}\overline{P}_{R,M} + o_\nu(1) = \int_{\Omega \times \mathcal{X}(K_R)} \mathrm{Int}_\tau \, \mathrm{d}\overline{P}_{R,M} + o_\nu(1).$$

Moreover, we have

$$\lim_{M \to \infty} \lim_{R \to \infty} \frac{1}{R^d} \int_{\Omega \times \mathcal{X}(K_R)} \mathrm{Int}_\tau \mathrm{d}\overline{P}_{R,M} = \overline{\mathbb{W}}_s(\overline{P}) + o_\tau(1);$$

thus we see that, with (4.7),

$$\lim_{M \to \infty, R \to \infty} \lim_{\nu \to 0} \lim_{N \to \infty} \frac{1}{N} \sum_{i \in I} \mathrm{Int}_\tau[\mathcal{C}_i, \mathcal{C}_i] = \overline{\mathbb{W}}_s(\overline{P}) + o_\tau(1). \qquad (4.11)$$

Step 3. *Regularization.*

In order to deal with the short-scale interactions that are not captured in Int_τ, we apply the regularization procedure of [18, Lemma 5.11]. Let us briefly present this procedure:

1. We partition Ω_N by small hypercubes of sidelength 6τ.
2. If one of these hypercubes $\mathcal{K}$ contains more than one point or if it contains a point and one of the adjacent hypercubes also contains a point, we replace the point configuration in $\mathcal{K}$ by one with the same number of points but confined in the central, smaller hypercube $\mathcal{K}' \subset \mathcal{K}$ of side length 3τ and that lives on a lattice (the spacing of the lattice depends on the initial number of points in $\mathcal{K}$).

This allows us to control the difference $\mathrm{Int} - \mathrm{Int}_\tau$ in terms of the number of points in the modified hypercubes.

In particular, we replace $\mathcal{A}$ by a new family of point configurations, such that

$$\frac{1}{N} \sum_{i \in I} (\mathrm{Int} - \mathrm{Int}_\tau)[\mathcal{C}_i, \mathcal{C}_i] \leq C\tau^{-s-d} \mathbf{E}_{\overline{P}_d} \left[\left((|\mathcal{C} \cap K_{12\tau}|)^{2+s/d} - 1 \right)_+ \right]. \quad (4.12)$$

The right-hand side of (4.12) should be understood as follows: any group of points that were too close to each other (without any precise control) have been replaced by a group of points with the same cardinality but whose interaction energy is now similar to that of a lattice. The energy of n points in a lattice of spacing $\frac{\tau}{n^{1/d}}$ scales like $n^{2+s/d}\tau^{-s}$, and taking the average over all small hypercubes, is similar to computing $\frac{1}{\tau^d} \mathbf{E}_{\overline{P}_d}$.

As $\nu \to 0$, we may then compare the right-hand side of (4.12) with the same quantity for $\overline{P}$, namely

$$\tau^{-s-d} \mathbf{E}_{\overline{P}} \left[\left((|\mathcal{C} \cap K_{12\tau}|)^{2+s/d} - 1 \right)_+ \right],$$

which can be shown to be $o_\tau(1)$ (following the argument of [18, Section 6.3.3]), because it is in turn of the same order as

$$\mathbf{E}_{\overline{P}}\left[(\mathrm{Int} - \mathrm{Int}_\tau)\left[K_1, K_1\right]\right],$$

which goes to zero as $\tau \to 0$ by dominated convergence.

We obtain

$$\lim_{\tau \to 0} \lim_{M, R \to \infty} \lim_{\nu \to 0} \frac{1}{N} \sum_{i \in I} (\mathrm{Int} - \mathrm{Int}_\tau)\left[\mathcal{C}_i, \mathcal{C}_i\right] = 0, \qquad (4.13)$$

and combining (4.13) with (4.11), we get that

$$\lim_{\tau \to 0} \lim_{M \to \infty, R \to \infty} \lim_{\nu \to 0} \lim_{N \to \infty} \frac{1}{N} \sum_{i \in I} \mathrm{Int}[\mathcal{C}_i, \mathcal{C}_i] \leq \overline{\mathbb{W}}_s(\overline{P}). \qquad (4.14)$$

Step 4. *Shrinking the configurations.* This procedure is borrowed from [14]. It rescales the configuration by a factor less than one (but very close to 1), effectively shrinking it and creating an empty boundary layer around each cube. Thus points belonging to different cubes are sufficiently well-separated so that interactions between the cubes are negligible–a much simpler approach to screening than that in the long-range case.

For $R > 0$, we let $R' := R^{\sqrt{d/s}}$.

It is not true in general that $\mathrm{Int}[\mathcal{C}, \mathcal{C}]$ can be split as the sum $\sum_{i \in I} \mathrm{Int}[\mathcal{C}_i, \mathcal{C}_i]$. However, since the Riesz interaction decays fast at infinity, it is approximately true if the configurations $\mathcal{C}_i$ are separated by a large enough distance. To ensure that, we "shrink" every configuration $\mathcal{C}_i$ in $\tilde{K}_i$; namely, we rescale them by a factor $1 - \frac{R'}{R}$. This operation affects the discrete average (4.7) but not the empirical process; i.e., for any $\varepsilon > 0$, if M, R are large enough and ν small enough, we may still assume that (4.8) holds. The interaction energy in each hypercube $\tilde{K}_i$ is multiplied by $\left(1 - \frac{R'}{R}\right)^{-s} = 1 + o_R(1)$, but the configurations in two distinct hypercubes are now separated by a distance at least R'. Since (4.10) holds, an elementary computation implies that we have, for any i in I,

$$\mathrm{Int}[\mathcal{C}_i, \sum_{j \neq i} \mathcal{C}_j] = M^2 R^d \frac{R^d}{R'^s} O(1),$$

with a $O(1)$ depending only on d, s. We thus get

$$\mathrm{Int}[\mathcal{C}, \mathcal{C}] = \sum_{i \in I} \mathrm{Int}[\mathcal{C}_i, \mathcal{C}_i] + N M^2 \frac{R^d}{R'^s} O(1),$$

but $\frac{R^d}{R'^s} = o_R(1)$ by the choice of R' (and the fact that $d < s$), and thus (in view of (4.14) and the effect of the scaling on the energy)

$$\lim_{\tau \to 0} \lim_{M \to \infty, R \to \infty} \lim_{\nu \to 0} \lim_{N \to \infty} \frac{1}{N} \mathrm{Int}[\mathcal{C}, \mathcal{C}] \leq \overline{\mathbb{W}}_s(\overline{P}). \qquad (4.15)$$

We have thus constructed a large enough (see (4.9)) volume of point configurations in Ω_N whose associated empirical processes converge to $\overline{P}$ and such that

$$\frac{1}{N}\text{Int}[\mathcal{C},\mathcal{C}] \leq \overline{\mathbb{W}}_s(\overline{P}) + o(1).$$

We may view these configurations at the original scale by applying a homothety of factor $N^{-1/d}$; this way we obtain point configurations $\mathbf{X}_N$ in Ω such that

$$\frac{1}{N^{1+s/d}}E_s(\mathbf{X}_N) \leq \overline{\mathbb{W}}_s(\overline{P}) + o(1).$$

It is not hard to see that the associated empirical measure μ_N converges to the intensity measure of $\overline{P}$, and since V is continuous, we also have

$$\frac{1}{N}\int_{\mathbb{R}} V\,\mathrm{d}\mu_N = \overline{\mathbb{V}}(\overline{P}) + o(1).$$

This concludes the proof of Proposition 4.4.

$\square$

We may now prove the LDP lower bound.

Proof of Proposition 4.3 Proposition 4.4 implies (4.3); indeed, we have

$$\overline{\mathfrak{P}}_{N,\beta}(B(\overline{P},\varepsilon)) = \frac{1}{Z_{N,\beta}}\int_{\Omega^N \cap \{\overline{\text{Emp}}_N(\mathbf{X}_N) \in B(\overline{P},\varepsilon)\}} \exp\left(-\beta N^{-s/d}\mathcal{H}_N(\mathbf{X}_N)\right)\mathrm{d}\mathbf{X}_N$$

$$\geq \frac{1}{Z_{N,\beta}}\int_{\Omega^N \cap \{\overline{\text{Emp}}_N(\mathbf{X}_N) \in B(\overline{P},\varepsilon)\} \cap T_{N,\delta}(\overline{P})} \exp\left(-\beta N^{-s/d}\mathcal{H}_N(\mathbf{X}_N)\right)\mathrm{d}\mathbf{X}_N$$

$$\geq \frac{1}{Z_{N,\beta}}\exp\left(-\beta N\left(\overline{\mathcal{F}}_\beta(\overline{P}) + \delta\right)\right)\int_{\Omega^N \cap \{\overline{\text{Emp}}_N(\mathbf{X}_N) \in B(\overline{P},\varepsilon)\} \cap T_{N,\delta}(\overline{P})} \mathrm{d}\mathbf{X}_N,$$

and (4.5) allows us to bound below the last integral as

$$\liminf_{\delta \to 0, \varepsilon \to 0, N \to \infty} \frac{1}{N}\log\int_{\Omega^N \cap \{\overline{\text{Emp}}_N(\mathbf{X}_N) \in B(\overline{P},\varepsilon)\} \cap T_{N,\delta}(\overline{P})} \mathrm{d}\mathbf{X}_N$$

$$\geq -\int_\Omega\left(\text{ent}[\overline{P}^x|\Pi] - 1\right) - 1.$$

$\square$

4.4 Proof of Theorem 1.1 and Corollary 1.2

From Propositions 4.2 and 4.3, the proof of Theorem 1.1 is standard. Exponential tightness of $\overline{\mathfrak{P}}_{N,\beta}$ comes for free (see, e.g., [18, Section 4.1]) because the average number of points is fixed, and we may thus improve the weak large deviation estimates (4.2), into the following: for any $A \subset \overline{\mathcal{M}}_{stat,1}(\mathcal{X})$, we have

$$-\inf_{\mathring{A}} \overline{\mathcal{F}}_\beta + \liminf_{N\to\infty} \left(-\frac{\log Z_{N,\beta}}{N}\right)$$

$$\le \liminf_{N\to\infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(A) \le \limsup_{N\to\infty} \frac{1}{N} \log \overline{\mathfrak{P}}_{N,\beta}(A)$$

$$\le -\inf_{\overline{A}} \overline{\mathcal{F}}_\beta \limsup_{N\to\infty} \left(-\frac{\log Z_{N,\beta}}{N}\right).$$

We easily deduce that

$$\lim_{N\to\infty} \frac{\log Z_{N,\beta}}{N} = -\min_{\overline{\mathcal{M}}_{stat,1}(\overline{\mathcal{X}})} \overline{\mathcal{F}}_\beta,$$

which proves Corollary 1.2, and that the LDP for $\overline{\mathfrak{P}}_{N,\beta}$ holds as stated in Theorem 1.1.

4.5 Proof of Theorem 1.3

Proof Theorem 1.3 follows from an application of the "contraction principle" (see, e.g., [24, Section 3.1]). Let us consider the map $\overline{\mathcal{M}}(\overline{\mathcal{X}}) \to \mathcal{M}(\Omega)$ defined by

$$\widetilde{\mathrm{Intens}} : \overline{P} \mapsto \int_\Omega \delta_x \mathbf{E}_{\overline{P}^x}[\mathcal{C} \cap K_1].$$

It is continuous on $\overline{\mathcal{M}}_{stat}(\overline{\mathcal{X}})$ and coincides with $\overline{\mathrm{Intens}}$. By the contraction principle, the law of $\widetilde{\mathrm{Intens}}(\overline{\mathrm{Emp}}(\mathbf{X}_N))$ obeys a large deviation principle governed by

$$\rho \mapsto \inf_{\overline{\mathrm{Intens}}(\overline{P})=\rho} \overline{\mathcal{F}}_\beta(\overline{P}),$$

which is easily seen to be equal to $I_\beta(\rho)$ as defined in (2.27).

For technical reasons (a boundary effect), it is not true in general that $\widetilde{\mathrm{Intens}}(\overline{\mathrm{Emp}}(\mathbf{X}_N)) = \mathrm{emp}(\mathbf{X}_N)$; however, we have

$$\mathrm{dist}_{\mathcal{M}(\Omega)} \left(\widetilde{\mathrm{Intens}}(\overline{\mathrm{Emp}}(\mathbf{X}_N)), \overline{\mathrm{Emp}}(\mathbf{X}_N)\right) = o_N(1),$$

uniformly for $\mathbf{X}_N \in \Omega$. In particular, the laws of $\widetilde{\mathrm{Intens}}(\overline{\mathrm{Emp}}(\mathbf{X}_N))$ and of $\mathrm{emp}(\mathbf{X}_N)$ are exponentially equivalent (in the language of large deviations); thus any LDP can be transferred from one to the other. This proves Theorem 1.3.   □

5 Additional Proofs: Propositions 1.4, 1.5, and 1.6

5.1 Limit of the Empirical Measure

From Theorem 1.3 and the fact that I_β is strictly convex, we deduce that $\mathrm{emp}(\mathbf{X}_N)$ converges almost surely to the unique minimizer of I_β.

Proof of Proposition 1.4 First, if $V = 0$ and Ω is bounded, I_β can be written as

$$I_\beta(\rho) := \int_\Omega \rho(x) \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \rho(x)^{s/d} \mathbb{W}_s(P) + \mathsf{ent}[P|\Pi] \right) dx$$
$$+ \int_\Omega \rho(x) \log \rho(x) \, dx.$$

We claim that both terms in the right-hand side are minimized when ρ is the uniform probability measure on Ω (we may assume $|\Omega| = 1$ to simplify, without loss of generality). This property is well known for the relative entropy term $\int_\Omega \rho \log \rho$, and we now prove it for the energy term. First, let us observe that

$$\alpha \mapsto \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \alpha^{1+s/d} \mathbb{W}_s(P) + \alpha \, \mathsf{ent}[P|\Pi] \right)$$

is convex in α since it is the infimum over a family of convex functions (recall that $\alpha \mapsto \alpha^{1+s/d}$ is convex in α and that $\mathbb{W}_s$ is always positive). Since $|\Omega| = 1$, we have, by Jensen's inequality,

$$\int_\Omega \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \rho(x)^{1+s/d} \mathbb{W}_s(P) + \rho(x) \mathsf{ent}[P|\Pi] \right) dx$$
$$\geq \inf_{P \in \mathcal{P}_{stat,1}(\mathcal{X})} \left(\beta \left(\int_\Omega \rho(x) \right)^{1+s/d} \mathbb{W}_s(P) + \left(\int_\Omega \rho(x) \right) \mathsf{ent}[P|\Pi] \right),$$

and since $\int_\Omega \rho = 1$, we conclude that I_β is minimal for $\rho \equiv 1$. Thus the empirical measure converges almost surely to the uniform probability measure on Ω, which proves the first point of Proposition 1.4.

Next, let us assume that V is arbitrary and Ω bounded. It is not hard to see that for the minimizer $\mu_{V,\beta}$ of I_β, we have, as $\beta \to 0$,

$$I_\beta(\mu_{V,\beta}) \geq I_\beta(\rho_{\mathsf{unif}}) + O(\beta),$$

where ρ_{unif} is the uniform probability measure on Ω. Moreover, it is also true (as proved above) that the first term in the definition of I_β is minimal for $\rho = \rho_{\mathsf{unif}}$. We thus get that, as $\beta \to 0$,

$$\int_\Omega \mu_{V,\beta} \log \mu_{V,\beta} - \int_\Omega \rho_{\mathsf{unif}} \log \rho_{\mathsf{unif}} = O(\beta);$$

in other words, the relative entropy of $\mu_{V,\beta}$ with respect to ρ_{unif} converges to 0 as $\beta \to 0$. The Csisz–Kullback–Pinsker inequality allows us to bound the square of the total variation distance between $\mu_{V,\beta}$ and ρ_{unif} by the relative entropy (up to a multiplicative constant), and thus $\mu_{V,\beta}$ converges (in total variation) to the uniform probability measure on Ω as $\beta \to 0$. This proves the second point of Proposition 1.4.

Finally for V arbitrary, the problem of minimizing of I_β is, as $\beta \to \infty$, similar to minimizing

$$\beta \left(\int_\Omega \rho(x)^{1+s/d} \min \mathbb{W}_s \, \mathrm{d}x + \int_\Omega \rho(x) V(x) \mathrm{d}x \right).$$

Since $\min \mathbb{W}_s = C_{s,d}$, we recover (up to a multiplicative constant $\beta > 0$) the minimization problem studied in [16], namely the problem of minimizing

$$C_{s,d} \int_\Omega \rho(x)^{1+s/d} \mathrm{d}x + \int_\Omega \rho(x) V(x) \mathrm{d}x,$$

among probability densities, whose (unique) solution is given by $\mu_{V,\infty}$.

In order to prove that $\mu_{V,\beta}$ converges to $\mu_{V,\infty}$ as $\beta \to \infty$, we need to make that heuristic rigorous, which requires an adaptation of [17, Section 7.3, Step 2]. We claim that there exists a sequence $\{P_k\}_{k \geq 1}$ in $\mathcal{P}_{stat,1}(\mathcal{X})$ such that

$$\lim_{k \to \infty} \mathbb{W}_s(P_k) = C_{s,d}, \quad \forall k \geq 1, \, \mathsf{ent}[P_k | \Pi] < +\infty. \tag{5.1}$$

We could think of taking $P_k = P$, where P is some minimizer of $\mathbb{W}_s$ among $\mathcal{P}_{stat,1}(\mathcal{X})$, but it might have infinite entropy (e.g., if P was the law of the stationary process associated with a lattice, as in dimension 1). We thus need to "expand" P (e.g., by making all the points vibrate independently in small balls as described in [17, Section 7.3, Step 2] in the case of the one-dimensional lattice). We may then write that, for any $\beta > 0$ and $k \geq 1$,

$$I_\beta(\mu_{V,\beta}) \leq I_\beta(\mu_{V,\infty}) \leq \beta \left(\int_\Omega \mu_{V,\infty}(x)^{1+s/d} \mathbb{W}_s(P_k) + \int_\Omega \mu_{V,\infty}(x) V(x) \mathrm{d}x \right)$$
$$+ \mathsf{ent}[P_k | \Pi] + \int_\Omega \mu_{V,\infty}(x) \log \mu_{V,\infty}(x)$$
$$\leq \beta \left(\int_\Omega \mu_{V,\infty}(x)^{1+s/d} C_{s,d} + \int_\Omega \mu_{V,\infty}(x) V(x) \mathrm{d}x \right) + \mathsf{ent}[P_k | \Pi] + \beta o_k(1),$$

where we have used (5.1) in the last inequality. Choosing β and k properly so that $k \to \infty$ as $\beta \to \infty$, while assuring that the $\beta o_k(1)$ term goes to zero, we have

$$C_{s,d} \int_\Omega \mu_{V,\beta}(x)^{1+s/d} + \int_\Omega \mu_{V,\beta}(x) V(x) \leq C_{s,d} \int_\Omega \mu_{V,\infty}(x)^{1+s/d} \mathrm{d}x$$
$$+ \int_\Omega \mu_{V,\infty}(x) V(x) \mathrm{d}x + o_{\beta \to \infty}(1).$$

By convexity, this implies that $\mu_{V,\beta}$ converges to $\mu_{V,\infty}$ as $\beta \to \infty$. $\qquad \square$

5.2 The Case of Minimizers

Proof of Proposition 1.5 Let $\{\mathbf{X}_N\}_N$ be a sequence of N-point configurations such that for all $N \geq 1$, $\mathbf{X}_N$ minimizes $\mathcal{H}_N$. From Proposition 3.5, we know that (up to extraction), $\{\overline{\mathrm{Emp}}(\mathbf{X}_N)\}_N$ converges to some $\overline{P} \in \mathcal{M}_{stat,1}(\mathcal{X})$ such that

$$\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) \leq \liminf_{N \to \infty} \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}}, \tag{5.2}$$

and we have, by (2.23), (3.1), and the scaling properties of $\mathbb{W}_s$,

$$\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) \geq C_{s,d} \int_\Omega \rho(x)^{1+s/d} \mathrm{d}x + \int_\Omega V(x)\rho(x)\mathrm{d}x, \tag{5.3}$$

where $\rho = \overline{\mathrm{Intens}}(\overline{P})$. We also know that the empirical measure $\mathrm{emp}(\mathbf{X}_N)$ converges to the intensity measure $\rho = \overline{\mathrm{Intens}}(\overline{P})$.

On the other hand, from [16, Theorem 2.1], we know that $\mathrm{emp}(\mathbf{X}_N)$ converges to some measure $\mu_{V,\infty}$ that is defined as follows: define L to be the unique solution of

$$\int_\Omega \left[\frac{L - V(x)}{C_{s,d}(1 + s/d)} \right]_+^{d/s} \mathrm{d}x = 1,$$

and then let $\mu_{V,\infty}$ be given by

$$\mu_{V,\infty}(x) := \left[\frac{L - V(x)}{C_{s,d}(1 + s/d)} \right]_+^{d/s} \qquad (x \in \Omega). \tag{5.4}$$

It is proved in [16] that $\mu_{V,\infty}$ minimizes the quantity

$$C_{s,d} \int_\Omega \rho(x)^{1+s/d} \, \mathrm{d}x + \int V(x)\rho(x) \, \mathrm{d}x, \tag{5.5}$$

among all probability density functions ρ supported on Ω. It is also proved that

$$\lim_{N \to \infty} \frac{\mathcal{H}_N(\mathbf{X}_N)}{N^{1+s/d}} = C_{s,d} \int_\Omega \mu_{V,\infty}(x)^{1+s/d} \, \mathrm{d}x + \int V(x)\mu_{V,\infty}(x) \, \mathrm{d}x. \tag{5.6}$$

By unicity of the limit, we have $\rho := \overline{\mathrm{Intens}}(\overline{P}) = \mu_{V,\infty}$. In view of (5.2), (5.3), (5.6), and by the fact that $\mu_{V,\infty}$ minimizes (5.5), we get that

$$\overline{\mathbb{W}}_s(\overline{P}) + \overline{\mathbb{V}}(\overline{P}) = C_{s,d} \int_\Omega \mu_{V,\infty}(x)^{1+s/d} \mathrm{d}x + \int_\Omega V(x)\mu_{V,\infty}(x)\mathrm{d}x$$

and that $\overline{P}$ is in fact a minimizer of $\overline{\mathbb{W}}_s + \overline{\mathbb{V}}$. We must also have

$$\overline{\mathbb{W}}_s(\overline{P}) = C_{s,d} \int_\Omega \mu_{V,\infty}(x)^{1+s/d} \, \mathrm{d}x;$$

hence (in view of (2.23)) we get

$$\mathcal{W}_s(\mathcal{C}) = C_{s,d}\mu_{V,\infty}(x)^{1+s/d} = \min_{\mathcal{X}_{\mu_{V,\infty}(x)}} \mathcal{W}_s, \text{ for } \overline{P} - \text{a.e. } (x,\mathcal{C}),$$

which concludes the proof. □

5.3 The One-Dimensional Case

Proposition 1.6 is very similar to the first statement of [17, Theorem 3], and we sketch its proof here.

Proof of Proposition 1.6 First, we use the expression of $\mathbb{W}_s$ in terms of the two-point correlation function, as presented in (2.25):

$$\mathbb{W}_s(P) = \liminf_{R \to \infty} \int_{[-R,R]^d} \frac{1}{|v|^s} \rho_{2,P}(v) \left(1 - \frac{|v|}{R}\right) dv.$$

Then, we split $\rho_{2,P}$ as the sum

$$\rho_{2,P} = \sum_{k=1}^{+\infty} \rho_{2,P}^{(k)},$$

where $\rho_{2,P}^{(k)}$ is the correlation function of the k-th neighbor (which makes sense only in dimension 1). It is not hard to check that

$$\int \rho_{2,P}^{(k)}(x) = 1 \text{ and } \int x\rho_{2,P}^{(k)}(x) = k$$

(the last identity holds because P has intensity 1 and is stationary). Using the convexity of

$$v \mapsto \frac{1}{|v|^s}\left(1 - \frac{|v|}{R}\right),$$

we obtain that for any $k \geq 1$, the following holds:

$$\int \frac{1}{|v|^s}\left(1 - \frac{|v|}{R}\right)\rho_{2,P}^{(k)}dv \geq \int \frac{1}{|v|^s}\left(1 - \frac{|v|}{R}\right)\delta_k(v)dv$$
$$= \int \frac{1}{|v|^s}\left(1 - \frac{|v|}{R}\right)\rho_{2,P_\mathbb{Z}}^{(k)}(v)dv,$$

where $P_\mathbb{Z} = u + \mathbb{Z}$ (with u uniform in $[0,1]$). Thus we have

$$\mathbb{W}_s(P) \geq \mathbb{W}_s(P_\mathbb{Z}),$$

which proves that $\mathbb{W}_s$ is minimal at $P_\mathbb{Z}$. □

References

1. Blanc, X., Lewin, M.: The crystallization conjecture: a review. EMS Surv. Math. Sci. **2**(2), 255–306 (2015)
2. Bloom, T., Levenberg, N., Wielonsky, F.: A large deviation principle for weighted Riesz interactions. Constr. Approx. **47**(1), 119–140 (2018)
3. Borodachov, S.V., Hardin, D.P., Reznikov, A., Saff, E.B.: Optimal discrete measures for Riesz potentials. Trans. Amer. Math. Soc. (2018). https://doi.org/10.1090/tran/7224
4. Bouchet, F., Gupta, S., Mukamel, D.: Thermodynamics and dynamics of systems with long-range interactions. Phys. A **389**(20), 4389–4405 (2010)
5. Brauchart, J.S., Hardin, D.P., Saff, E.B.: The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere. Contemp. Math. **578**, 31–61 (2012)
6. Campa, A., Dauxois, T., Ruffo, S.: Statistical mechanics and dynamics of solvable models with long-range interactions. Phys. Rep. **480**(3–6), 57–159 (2009)
7. Cohn, H., Kumar, A.: Universally optimal distribution of points on spheres. J. Am. Math. Soc. **20**(1), 99–148 (2007)
8. Daley, D.J.: An introduction to the theory of point processes. Vol. I. Probability and its Applications. Elementary theory and methods, 2nd edn. Springer, New York (2003)
9. Dauxois, T., Ruffo, S., Arimondo, E., Wilkens, M.: Dynamics and Thermodynamics of Systems with Long-Range Interactions, volume 602 of Lecture Notes in Physics. Lectures from the conference held in Les Houches. Springer, Berlin (2002)
10. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, volume 38 of Stochastic Modelling and Applied Probability. Springer, Berlin (2010). (Corrected reprint of the second (1998) edition)
11. Föllmer, H., Orey, S.: Large deviations for the empirical field of a Gibbs measure. Ann. Probab. **16**(3), 961–977 (1988)
12. Georgii, H.-O.: Large deviations and maximum entropy principle for interacting random fields on. Ann. Probab. **21**(4), 1845–1875 (1993)
13. Hardin, D.P., Saff, E.B.: Discretizing manifolds via minimum energy points. Not. Am. Math. Soc. **51**(10), 1186–1194 (2004)
14. Hardin, D.P., Saff, E.B.: Minimal Riesz energy point configurations for rectifiable d-dimensional manifolds. Adv. Math. **193**(1), 174–204 (2005)
15. Hardin, D.P., Saff, E.B., Simanek, B.Z.: Periodic discrete energy for long-range potentials. J. Math. Phys. **55**, 123509/27 (2014)
16. Hardin, D.P., Saff, E.B., Vlasiuk, O.V.: Generating point configurations via hypersingular Riesz energy with an external field. SIAM J. Math. Anal. **49**(1), 646–673 (2017)
17. Leblé, T.: Logarithmic, Coulomb and Riesz energy of point processes. J. Stat. Phys. **162**(4), 887–923 (2016)
18. Leblé, T., Serfaty, S.: Large deviation principle for empirical fields of log and Riesz gases. Invent. Math. **210**(3), 645–757 (2017)
19. Leblé, T., Serfaty, S., Zeitouni, O.: Large deviations for the two-dimensional two-component plasma. Commun. Math. Phys. **350**(1), 301–360 (2017)
20. Lieb, E.H.: Thomas–Fermi and related theories of atoms and molecules. Rev. Mod. Phys. **53**, 603–641 (1981)
21. Lieb, E.H., Seiringer, R.: The Stability of Matter in Quantum Mechanics. Cambridge University Press, Cambridge (2010)
22. Mazars, M.: Long ranged interactions in computer simulations and for quasi-2d systems. Phys. Rep. **500**(2), 43–116 (2011)
23. Petrache, M., Serfaty, S.: Next order asymptotics and renormalized energy for Riesz interactions. J. Inst. Math. Jussieu **16**(3), 501–569 (2017)
24. Rassoul-Agha, F., Seppäläinen, T.: A Course on Large Deviations with an Introduction to Gibbs Measures. Graduate Studies in Mathematics, vol. 162. American Mathematical Society, Providence (2015)
25. Saff, E.B., Kuijlaars, A.: Distributing many points on a sphere. Math. Intell. **19**(1), 5–11 (1997)

26. Saff, E.B., Totik, V.: Logarithmic Potentials with External Fields. Grundlehren der mathematischen Wissenchaften, vol. 316. Springer, Berlin (1997)
27. Serfaty, S.: Coulomb Gases and Ginzburg-Landau Vortices. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Europe (2015)
28. Varadhan, S.R.S.: Large Deviations, volume 27 of Courant Lecture Notes in Mathematics. American Mathematical Society, Providence (2016)