The One-Dimensional Log-Gas Free Energy Has a Unique Minimizer

MATTHIAS ERBAR
Universität Bonn

MARTIN HUESMANN
Universität Wien

AND

THOMAS LEBLÉ
Courant Institute

Abstract

We prove that, at every positive temperature, the infinite-volume free energy of the one-dimensional log-gas, or beta-ensemble, has a unique minimizer, which is the Sine-beta process arising from random matrix theory. We rely on a quantitative displacement convexity argument at the level of point processes, and on the screening procedure introduced by Sandier-Serfaty. © 2021 The Authors. *Communications on Pure and Applied Mathematics* published by Wiley Periodicals LLC.

1 Introduction and Main Result

1.1 The One-Dimensional Log-Gas

The one-dimensional log-gas in finite volume can be defined as a system of particles interacting through a repulsive pairwise potential proportional to the logarithm of the distance and confined by some external field.

For a fixed value of $\beta > 0$, called the *inverse temperature* parameter, and for $N \geq 1$, we consider the probability measure $\mathbb{P}_{N,\beta}$ on $\vec{X}_N = (x_1, \ldots, x_N) \in \mathbb{R}^N$ defined by the density

$$d\mathbb{P}_{N,\beta}(\vec{X}_N) := \frac{1}{Z_{N,\beta}} \exp\left(-\beta\left(\sum_{i<j} -\log|x_i - x_j| + \sum_{i=1}^{N} N\frac{x_i^2}{2}\right)\right), \tag{1.1}$$

with respect to the Lebesgue measure on $\mathbb{R}^N$. The quantity $Z_{N,\beta}$ is a normalization constant, the *partition function*. We call $\mathbb{P}_{N,\beta}$ the *canonical Gibbs measure* of the log-gas.

Part of the motivation for studying log-gases comes from random matrix theory (RMT), for which $\mathbb{P}_{N,\beta}$ describes the joint law of N eigenvalues in certain classical models: the Gaussian orthogonal, unitary, and symplectic ensemble, respectively,

Communications on Pure and Applied Mathematics, Vol. LXXIV, 0615–0675 (2021)

for $\beta = 1, 2, 4$, and the "tridiagonal model" discovered in [6] for arbitrary β. We refer to the book [8] for a comprehensive presentation of the connection between log-gases and random matrices. Log-gases are also interesting from a statistical physics point of view, as a toy model with singular, long-range interaction.

Questions about such systems usually deal with the large-N limit (also called thermodynamic, or infinite-volume limit) of the system, as encoded by certain observables. For example, in order to understand the "global" behavior, one may look at the empirical measure $\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$, and ask about the typical behavior of this *random* probability measure on $\mathbb{R}$ as N tends to infinity. By now, this is fairly well understood; we refer, e.g., to the recent lecture notes [23] and the references therein. In the present paper, we are rather interested in the asymptotic behavior at *microscopic* scale.

1.2 The Sine-Beta Process

Let $\mathcal{C}_{N,0}$ be the point configuration

$$\mathcal{C}_{N,0} := \sum_{i=1}^{N} \delta_{Nx_i},$$

where $\vec{X}_N = (x_1, \dots, x_N)$ is distributed according to $\mathbb{P}_{N,\beta}$. The limit in law of $\mathcal{C}_{N,0}$ as $N \to \infty$ was constructed in [24] and named the Sine_β process. We refer to [12] for a different construction of a process that turns out to be the same, and to [25, 26] for recent developments concerning Sine_β. This process is the universal behavior of log-gases (in the bulk) in the sense that replacing the $\frac{x_i^2}{2}$ term in (1.1) by a general potential $V(x_i)$ yields the same microscopic limit, up to a scaling on the average density of points (our convention is that Sine_β has intensity 1) and mild assumptions on V; see [3, 4].

In [5], a different description of Sine_β is given using the Dobrushin-Landford-Ruelle (DLR) formalism, but the question of whether Sine_β is the unique solution to DLR equations is left open. The main result of the present paper answers positively to a slightly different uniqueness question, phrased in terms of the log-gas free energy.

1.3 The Log-Gas Free Energy

In [16], the infinite-volume free energy of the log-gas (and of other related systems) was introduced as the weighted sum $\mathcal{F}_\beta := \beta\mathcal{W} + \mathcal{E}$, where the functionals $\mathcal{W}, \mathcal{E}$ and the free energy $\mathcal{F}_\beta$ are defined on the space of stationary random point processes. The functional $\mathcal{W}$ corresponds to the "renormalized energy" introduced in [20] and adapted to this context in [18, 21], and $\mathcal{E}$ is the usual *specific relative entropy*. Both terms are defined below; see Section 2.

The free energy $\mathcal{F}_\beta$ appears in [16] as the rate function for a large-deviation principle concerning the behavior of log-gases at the microscopic level. If $\vec{X}_N = (x_1, \dots, x_N)$ is an N-tuple of particles distributed according to the Gibbs measure

(1.1) of a log-gas, they are known to typically arrange themselves on an interval approximately given by $[-2, 2]$. For x in this interval, we let $\mathcal{C}_{N,\mathrm{x}}$ be the point configuration $(x_1, \dots, x_N)$ "seen from x", namely $\mathcal{C}_{N,\mathrm{x}} := \sum_{i=1}^{N} \delta_{N(x_i - \mathrm{x})}$. We may then consider the *empirical field* of the system in the state $\vec{X}_N$, defined as

$$\mathrm{Emp}_N(\vec{X}_N) := \frac{1}{4} \int_{-2}^{2} \delta_{\mathcal{C}_{N,\mathrm{x}}} d\mathrm{x}.$$

The empirical field $\mathrm{Emp}_N(\vec{X}_N)$ is a probability measure on (finite) point configurations in $\mathbb{R}$, and it was proven in [16] that its law satisfies a large-deviation principle, at speed N, with a rate function built using $\mathcal{F}_\beta$. We refer to the paper cited above for a precise statement; here it suffices to say that *understanding the minimizers* of $\mathcal{F}_\beta$ gives an *understanding of the typical microscopic behavior* of a finite N log-gas at temperature β when N is large.

For any β in $(0, +\infty)$, the functional $\mathcal{F}_\beta$ is known to be lower semicontinuous, with compact sublevel sets. In particular, it admits a compact subset of minimizers. However, the question of uniqueness of minimizers for $\mathcal{F}_\beta$ remained open, and we address it in this paper.

1.4 Main Result

THEOREM 1.1. *For any β in $(0, +\infty)$, the free energy $\mathcal{F}_\beta$ has a unique minimizer.*

Since it was proven in [16, cor. 1.2] that Sine_β minimizes $\mathcal{F}_\beta$, we deduce the following:

COROLLARY 1.2. *For any β in $(0, +\infty)$, the* Sine_β *process is the unique minimizer of $\mathcal{F}_\beta$.*

This provides a variational characterization of Sine_β. A weaker variational property of Sine_β had been used, e.g., in [14] to prove the convergence of Sine_β to a Poisson point process as $\beta \to 0$, retrieving a result of [1].

Let us note that uniqueness also holds for $\beta = 0$ (the unique minimizer of the specific relative entropy $\mathcal{E}$ is the Poisson point process) and for $\beta = +\infty$, in the sense that the energy term $\mathcal{W}$ has a unique stationary minimizer; see [13].

1.5 Elements of Proof and Plan of the Paper

The proof of Theorem 1.1 goes by contradiction. We assume that $\mathcal{F}_\beta$ admits two distinct minimizers, and using convexity arguments we construct a stationary point process whose free energy is strictly less than $\min \mathcal{F}_\beta$, which is of course absurd. Since the free energy $\mathcal{F}_\beta$ is affine, it is not strictly convex for the usual linear interpolation of probability measures. We use instead the notion of *displacement convexity*, which was introduced in [17] to remedy situations where energy functionals are not convex in the usual sense. This idea was suggested to us by Alice Guionnet, and we warmly thank her for her insight.

Strategy of the Proof

The proof goes by contradiction. We start with two stationary point processes $\mathsf{P}^0, \mathsf{P}^1$ such that $\mathsf{P}^0 \neq \mathsf{P}^1$, and assume that both are minimizers of $\mathcal{F}_\beta$. We cannot argue via displacement convexity directly on the level of $\mathsf{P}^0, \mathsf{P}^1$ since they are probability measures on *infinite* point configurations. Optimal transport theory for random stationary measures as initiated in [7, 10, 11] is not yet developed well enough to be directly applicable. Instead, we use transport theory between *finite* measures together with a careful approximation argument relying on screening of electric fields. More precisely, we write

$$\mathcal{F}_\beta(\mathsf{P}) = \lim_{R\to\infty} \frac{1}{|\Lambda_R|}(\beta \mathcal{W}_R(\mathsf{P}) + \mathcal{E}_R(\mathsf{P})),$$

where $\mathcal{W}_R, \mathcal{E}_R$ are quantities (the energy and the relative entropy) depending on the restriction of P to the line segment $\Lambda_R := [-R, R]$.

(1) Let $\mathsf{P}^0_R, \mathsf{P}^1_R$ be the restriction of $\mathsf{P}^0, \mathsf{P}^1$ to Λ_R. Assume that there are almost surely $2R$ points in $[-R, R]$. We may thus see $\mathsf{P}^0_R, \mathsf{P}^1_R$ as probability measures on $[-R, R]^{2R}$, apply classical results about optimal transportation, and find an optimal transport map T_R that pushes P^0_R onto P^1_R.

In fact, it is not true that $\mathsf{P}^0, \mathsf{P}^1$ have almost surely $2R$ points in Λ_R, so we first perform a version of the "screening procedure" of Sandier-Serfaty. It has for effect to modify the configurations near the extremities of $[-R, R]$ in order to enforce the correct number of points, while not changing too much the average energy, nor the entropy. The screening procedure requires some conditions in order to be applied; we will need to guarantee that they are verified with high probability under $\mathsf{P}^0, \mathsf{P}^1$.

(2) We let $\mathsf{P}^{\mathsf{h}}_R$ be the half-interpolate of $\mathsf{P}^0_R, \mathsf{P}^1_R$ along the displacement T_R, i.e., the push-forward of P^0_R by $\frac{1}{2}(\mathrm{Id} + \mathsf{T}_R)$. A result of [17] ensures that the relative entropy is displacement convex, so

$$\mathcal{E}_R[\mathsf{P}^{\mathsf{h}}_R] \leq \frac{1}{2}\big(\mathcal{E}_R(\mathsf{P}^0) + \mathcal{E}_R(\mathsf{P}^1)\big).$$

Moreover, the interaction potential $-\log|x-y|$ is strictly convex, so again by a result of [17], the energy $\mathcal{W}_R$ is also displacement convex. More precisely, we have

$$\mathcal{W}_R\big[\mathsf{P}^{\mathsf{h}}_R\big] \leq \frac{1}{2}\big(\mathcal{W}_R(\mathsf{P}^0) + \mathcal{W}_R(\mathsf{P}^1)\big) - \mathrm{Gain}_R,$$

where $\mathrm{Gain}_R > 0$ is some quantitative positive gain due to the *strict* convexity of the interaction. With some work, using the fact that $\mathsf{P}^0, \mathsf{P}^1$ are stationary, we are able to show that the gain is at least proportional to R.

(3) We turn $\mathsf{P}^{\mathsf{h}}_R$ into a process on the full line by pasting independent copies of itself on disjoint intervals of length $2R$. The relative entropy is additive, and we can show that the interaction of two independent copies is almost

zero. Thanks to the quantitative convexity estimate, we obtain a global candidate P^{h} for which

$$\mathcal{F}_\beta(\mathsf{P}^{\mathsf{h}}) < \frac{1}{2}\big(\mathcal{F}_\beta(\mathsf{P}^0) + \mathcal{F}_\beta(\mathsf{P}^1)\big),$$

which is absurd; hence the minimizer of $\mathcal{F}_\beta$ is unique.

Plan of the Paper

- In Section 2, we define formally the logarithmic energy and the specific relative entropy, which are the two ingredients of the log-gas free energy.
- In Section 3, we present the screening procedure.
- In Section 4, we state a key discrepancy estimate and collect auxiliary results controlling the typical size of the gaps and positions of points for processes of finite logarithmic energy.
- In Section 5, we implement the first step of the proof, by performing the screening procedure on a large box for any given point process with finite energy.
- In Section 6, we recall results from optimal transportation and the theory of displacement convexity.
- In Section 7, we apply displacement interpolation to build a local candidate: the interpolate process in a large box, whose local energy and entropy are better than the two processes it was constructed from.
- In Section 8, we use the interpolate process to build a global candidate, which is a stationary point process whose free energy is strictly smaller than the minimum of $\mathcal{F}_\beta$, which is the desired contradiction.
- Finally, in Appendix A, we gather the proofs of some technical lemmas, and in Appendix B we sketch a proof of the screening procedure adapted to the present setting.

1.6 Preliminary Definitions and Notation

Notation

- For $R > 0$, we denote by Λ_R the interval $[-R, R]$.
- If I is an interval of $\mathbb{R}$, we let $|I|$ be its length.
- If A is a subset of $\mathbb{R}$ or $\mathbb{R}^2$, we denote the boundary of A by ∂A.
- We let $\diamond$ be the diagonal in $\mathbb{R} \times \mathbb{R}$.
- We will use the symbol $\preceq$ to denote an inequality that holds up to some universal (positive) multiplicative constant.
- Points in $\mathbb{R} \times \mathbb{R}$ are denoted by uppercase letters, e.g., $X = (x, y)$.
- If μ is a signed measure on some interval $I \subset \mathbb{R}$, we will sometimes treat it as a signed measure on $I \times \mathbb{R}$, where it is understood that, $\varphi : I \times \mathbb{R} \to \mathbb{R}$ being a test function, we define

$$\int_{I\times\mathbb{R}} \varphi(x, y)\, d\mu(x, y) := \int_I \varphi(x, 0)\, d\mu(x). \tag{1.2}$$

- If μ is a probability measure on X and F is a map from X to Y, we let $F_*\mu$ be the push-forward of μ by F, which is a probability measure on Y.

Point Configurations and Processes

- A *point configuration* $\mathcal{C}$ on a set $S \subset \mathbb{R}$ is defined as a purely atomic Radon measure on S, giving integer mass to singletons and finite mass to any compact set. Any configuration can be written as $\mathcal{C} = \sum_{i\in I} \delta_{p_i}$, where $(p_i)_i$ is a collection of points in S and I is finite or countable. Here, multiple points are allowed. A configuration is simple if no multiple points occur, i.e., $\mathcal{C}(\{x\}) \in \{0, 1\}$ for all x. In this case we also consider $\mathcal{C}$ as a locally finite collection of points, i.e., the set $\{p_i, i \in I\}$. Abusing notation we write $\mathcal{C} = \sum_{p\in\mathcal{C}} \delta_p$. For all random configurations we consider, multiple points do not occur almost surely.
- For any interval $I \subset \mathbb{R}$, we let $\mathrm{Conf}(I)$ be the space of point configurations supported on I. It is endowed with the initial topology for the family of maps $\mathcal{C} \mapsto \int \varphi\, d\mathcal{C}$ for φ continuous and compactly supported on $\mathbb{R}$ (topology for which it is a Polish space), and with the associated Borel sigma-algebra.
- Given a subset $J \subset I$ and $\mathcal{C} \in \mathrm{Conf}(I)$, we denote by $\mathcal{C}|_J$ the restriction of $\mathcal{C}$ to J (viewed as a measure).
- If $u \in \mathbb{R}$ and $\mathcal{C} \in \mathrm{Conf}(\mathbb{R})$, we let $\mathcal{C}-u$ be the translation of $\mathcal{C}$ by u, which is the point configuration corresponding to shifting all the points of $\mathcal{C}$ by $-u$, or equivalently, for any test function φ
$$\int \varphi\, d(\mathcal{C}-u) := \int \varphi(\cdot + u) d\mathcal{C}.$$
- We will use two ways of enumerating points of a configuration:
  - In a fixed interval $[-R, R]$, we will write $-R \le \mathsf{z}_0 \le \mathsf{z}_1 \le \cdots \le \mathsf{z}_k \le \cdots \le R$, enumerating points from the leftmost one.
  - On $\mathbb{R}$, we will write $\cdots \le x_{-k} \le \cdots \le x_{-1} < 0 \le x_0 \le x_1 \le \cdots \le x_k \le \cdots$, enumerating points starting from the origin.

  We will need to pass from one enumeration to the other, for example in (5.9).
- A *point process* on I is a probability measure on $\mathrm{Conf}(I)$. For $J \subset I$ and a point process P on I we denote by P_J its restriction to J, i.e., $\mathsf{P}_J = (|_J)_*\mathsf{P}$ the point process on J obtained as the image under the restriction map $|_J : \mathrm{Conf}(I) \to \mathrm{Conf}(J)$.
- We let Π be the Poisson point process on $\mathbb{R}$ of intensity 1.
- If $\mathcal{C}$ is a finite point configuration, we let $|\mathcal{C}|$ be the number of points of $\mathcal{C}$. If $\mathcal{C}$ is a point configuration and I an interval, we let $\mathcal{C}_I$ be the restriction of $\mathcal{C}$ to I. In particular, $|\mathcal{C}_I|$ denotes the number of points of $\mathcal{C}$ in I. We then define the *discrepancy* of $\mathcal{C}$ in I as the difference:

$$\mathrm{Discr}_I(\mathcal{C}) := |\mathcal{C}_I| - |I|. \tag{1.3}$$

2 Definition of the Free Energy

In this section we recall the definition of the free energy $\mathcal{F}_\beta$. To this end, we first recall the notion of electric field associated to a point configuration and its electric energy.

Let us recall that, in the sense of distributions, the following identity holds:

$$-\mathrm{div}(\nabla(-\log|\cdot|)) = 2\pi\delta_0 \quad \text{in } \mathbb{R}\times\mathbb{R},$$

which corresponds to the fact that $X \mapsto \frac{1}{2\pi}\log|X|$ is the fundamental solution of Laplace's equation $\Delta f = \delta_0$ in dimension two.

DEFINITION 2.1 (Electric fields). Let I be an interval of $\mathbb{R}$.

- We call *electric fields on I* the set of all vector fields E in $L^p_{\mathrm{loc}}(I\times\mathbb{R}, \mathbb{R}\times\mathbb{R})$, for some $1 < p < 2$ fixed.
- Let $\mathcal{C}$ be a finite point configuration in I and E an electric field on I. We say that E *is compatible with $\mathcal{C}$ in I* provided

$$-\mathrm{div}(\mathrm{E}) = 2\pi(\mathcal{C} - dx) \quad \text{in } I\times\mathbb{R}, \tag{2.1}$$

in the sense of distributions, and using the convention of (1.2). See (2.4) for an example.

- If E is compatible with $\mathcal{C}$ in I, for $\eta\in(0,1)$ we define the η-truncation of the electric field E as

$$\mathrm{E}_\eta(X) := \mathrm{E}(X) - \sum_{p\in\mathcal{C}\cap I} \nabla\mathsf{f}_\eta(X-(p,0)), \tag{2.2}$$

where f_η is the function

$$\mathsf{f}_\eta(x) = \max\left(-\log\left(\frac{|x|}{\eta}\right), 0\right).$$

- In particular, if $\mathcal{C} = \sum_{x\in\mathcal{C}}\delta_x$ is a point configuration in I and E is compatible with $\mathcal{C}$, the truncation E_η satisfies the equation (compare with (2.1))

$$-\mathrm{div}(\mathrm{E}_\eta) = 2\pi\Big(\sum_{x\in\mathcal{C}}\sigma_{x,\eta} - dx\Big) \quad \text{in } I\times\mathbb{R}, \tag{2.3}$$

where $\sigma_{x,\eta}$ is the "smeared out charge" as in [19], the uniform measure on the circle of radius η around x (the measure $\sigma_{x,\eta}$ is truly supported on $\mathbb{R}\times\mathbb{R}$).

DEFINITION 2.2 (The local electric field). Let $R > 0$, let $\mathcal{C}$ be a point configuration in Λ_R.

- We define the *local electric field generated by $\mathcal{C}$* as

$$X\in\mathbb{R}\times\mathbb{R} \mapsto \mathrm{E}^{\mathrm{loc}}(X;\mathcal{C};\Lambda_R) := \int_{\Lambda_R} -\nabla\log|X-(u,0)|(d\mathcal{C}(u) - du). \tag{2.4}$$

- For any $\eta > 0$, we define the local energy of $\mathcal{C}$ (truncated at η) as

$$\mathrm{W}_\eta^{\mathrm{elec}}(\mathcal{C}; \Lambda_R) := \frac{1}{2\pi} \int_{\mathbb{R}\times\mathbb{R}} \left|\mathrm{E}_\eta^{\mathrm{loc}}\right|^2 + |\mathcal{C}| \log \eta. \tag{2.5}$$

Remark 2.3. $\mathrm{E}^{\mathrm{loc}}$ is a model case for the definition of electric fields. Indeed, $\mathrm{E}^{\mathrm{loc}}$ is an electric field compatible with $\mathcal{C}$ in Λ_R in the sense of (2.1), and it is in L^p_{loc} for all $p \in (1,2)$ but fails to be in L^2 near each point of $\mathcal{C}$.

DEFINITION 2.4 (The intrinsic energy). Let $\mathcal{C}$ be a point configuration in Λ_R. We define the *intrinsic energy of* $\mathcal{C}$ as the quantity

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R) := \iint_{(\Lambda_R\times\Lambda_R)\setminus\diamond} -\log|x-y|(d\mathcal{C}(x)-dx)(d\mathcal{C}(y)-dy), \tag{2.6}$$

where $\diamond$ denotes the diagonal in $\Lambda_R \times \Lambda_R$.

LEMMA 2.5 (Monotonicity estimates). *Let $\mathcal{C}$ be a finite point configuration in Λ_R with exactly $|\Lambda_R|$ points.*

(1) *The limit* $\lim_{\eta\to 0} \mathrm{W}_\eta^{\mathrm{elec}}(\mathcal{C}; \Lambda_R)$ *exists and satisfies*

$$\lim_{\eta\to 0} \mathrm{W}_\eta^{\mathrm{elec}}(\mathcal{C}; \Lambda_R) = \mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R). \tag{2.7}$$

(2) *We have, for $\eta > 0$,*

$$\mathrm{W}_\eta^{\mathrm{elec}}(\mathcal{C}; \Lambda_R) \le \mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R) + |\Lambda_R|\mathrm{ErMono}(\eta), \tag{2.8}$$

where ErMono *is some function independent of $\mathcal{C}$ satisfying*

$$\lim_{\eta\to 0} \mathrm{ErMono}(\eta) = 0. \tag{2.9}$$

(3) *Conversely, we have, for $\eta > 0$*

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R) \le \mathrm{W}_\eta^{\mathrm{elec}}(\mathcal{C}; \Lambda_R) + |\Lambda_R|\mathrm{ErMono}(\eta) + \mathrm{ErTrun}(\mathcal{C}, \eta; \Lambda_R), \tag{2.10}$$

where ErMono *is as above, and* ErTrun *satisfies*

$$\begin{aligned} &\mathrm{ErTrun}(\mathcal{C}, \eta; \Lambda_R) \\ &\quad \le \iint_{(x,y)\in\Lambda_R\times\Lambda_R, 0<|x-y|<2\eta} -\log|x-y|d\mathcal{C}(x)d\mathcal{C}(y). \end{aligned} \tag{2.11}$$

PROOF. The existence of the limit and (2.7) is proven, e.g., in [21]. The monotonicity estimates (2.8), (2.10), and (2.11) follow from [18, lemma 2.3](by sending the parameter α appearing there to 0). □

DEFINITION 2.6 (Electric energy of a point process). We introduce the first component of the free energy functional: the *electric energy of a point process*.

- Let $\mathcal{C}$ be a point configuration on $\mathbb{R}$. We define the global electric energy of $\mathcal{C}$ as

$$\widetilde{\mathrm{W}}^{\mathrm{elec}}(\mathcal{C}) := \inf_{E}\left(\lim_{\eta\to 0}\left(\limsup_{R\to\infty}\frac{1}{|\Lambda_R|}\frac{1}{2\pi}\int_{\Lambda_R\times\mathbb{R}}|\mathrm{E}_\eta|^2+\log\eta\right)\right), \tag{2.12}$$

where the inf is taken over electric fields E that are compatible with $\mathcal{C}$ in $\mathbb{R}$, in the sense of Definition 2.1. As usual, if there is no such electric field, we let $\widetilde{\mathrm{W}}^{\mathrm{elec}}(\mathcal{C}) := +\infty$.
- Let P be a point process. We define the electric energy of P as the quantity

$$\mathcal{W}(\mathsf{P}) := \mathbb{E}_{\mathsf{P}}\big[\widetilde{\mathrm{W}}^{\mathrm{elec}}(\mathcal{C})\big]. \tag{2.13}$$

We refer to [16, sec. 2.7] for more details. The following remark is a consequence of [16, lemmas 2.3 and 3.8].

Remark 2.7. If $\mathcal{C}$ is a point configuration such that $\widetilde{\mathrm{W}}^{\mathrm{elec}}(\mathcal{C})$ is finite, then the infimum in (2.12) is attained for exactly one electric field E, and in fact all other compatible fields have infinite energy. If P is a point process with finite energy, there is a measurable choice $\mathcal{C}\mapsto \mathrm{E}$ of compatible electric fields such that

$$\mathcal{W}(\mathsf{P}) = \mathbb{E}_{\mathsf{P}}\left[\lim_{\eta\to 0}\left(\limsup_{R\to\infty}\frac{1}{|\Lambda_R|}\frac{1}{2\pi}\int_{\Lambda_R\times\mathbb{R}}|\mathrm{E}_\eta|^2+\log\eta\right)\right].$$

Moreover, if P is stationary, we can write for any $R > 1$,

$$\mathcal{W}(\mathsf{P}) = \lim_{\eta\to 0}\mathbb{E}_{\mathsf{P}}\left[\frac{1}{|\Lambda_R|}\frac{1}{2\pi}\int_{\Lambda_R\times\mathbb{R}}|\mathrm{E}_\eta|^2+\log\eta\right]. \tag{2.14}$$

Remark 2.8. An alternative definition for the logarithmic energy of a point process P was introduced in [14], inspired by [2]. We define the intrinsic energy of P as

$$\mathcal{W}^{\mathrm{int}}(\mathsf{P}) := \liminf_{L\to\infty}\frac{1}{|\Lambda_L|}\mathbb{E}_{\mathsf{P}}\big[\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_L)\big]. \tag{2.15}$$

The energies $\mathcal{W}$ and $\mathcal{W}^{\mathrm{int}}$ are related to each other; in particular, it is proven in [14, prop. 5.1.] that if P is a stationary process with small discrepancies, for example, such that $\sup_L \mathbb{E}_{\mathsf{P}}[\mathrm{Discr}^2_{\Lambda_L}] < \infty$, we have

$$\mathcal{W}(\mathsf{P}) \le \mathcal{W}^{\mathrm{int}}(\mathsf{P}), \tag{2.16}$$

and in fact one can obtain some weak form of the converse inequality via a recovery sequence, we refer to [14] for more details.

DEFINITION 2.9 (Specific relative entropy). Let P be a stationary point process. The following limit exists in $[0,+\infty]$ and defines the *specific relative entropy* of P with respect to the Poisson point process:

$$\mathcal{E}[\mathsf{P}] := \lim_{R\to\infty}\frac{1}{|\Lambda_R|}\mathrm{Ent}[\mathsf{P}_{|\Lambda_R}|\Pi_{|\Lambda_R}]. \tag{2.17}$$

It is a lower semicontinuous, affine function of P. Its unique minimizer—among stationary processes—is the Poisson process itself, for which $\mathcal{E}[\Pi] = 0$. We refer, e.g., to [9, sec. 6.9.2] for elementary properties of the specific entropy.

DEFINITION 2.10 (The free energy $\mathcal{F}_\beta$). Let $\beta > 0$ be fixed. For any stationary point process P, we define the *free energy* $\mathcal{F}_\beta(\mathsf{P})$ as the weighted sum

$$\mathcal{F}_\beta(\mathsf{P}) := \beta\mathcal{W}(\mathsf{P}) + \mathcal{E}(\mathsf{P}). \tag{2.18}$$

It is lower semicontinuous and affine, and has compact sublevel sets. In particular, it has a minimum and admits a compact set of minimizers (our point is to prove that there is only one).

3 The Screening Procedure

The screening technique has been introduced in [22], followed by several adaptations in, e.g., [18, 19]. In this section, we state a version of the procedure suitable for us. In particular, the stationary nature of our problem allows us to bypass the usual step of "finding a good boundary", and we work with a fixed good boundary. In Claim 3.4, we add a technical remark concerning the "new" points produced by the screening construction that is needed here.

DEFINITION 3.1 (Screened fields). Let $\mathcal{C}$ be a point configuration in an interval I, and let E be an electric field, compatible with $\mathcal{C}$ in I; i.e., (2.1) holds. We say that E is screened and write $\mathrm{E} \in \mathsf{Screen}(\mathcal{C}, I)$ when

$$\mathrm{E} \cdot \vec{\nu} = 0 \quad \text{on } \partial(I \times \mathbb{R}), \tag{3.1}$$

where $\vec{\nu}$ is the outer unit normal vector.

The original motivation for working with *screened* electric fields is that one can paste such fields together: since their normal component agree on the boundary, they form a globally compatible electric field. In the present paper, we use this screening property rather incidentally, only to compare the "screened" energy with the "local" energy in the proof of Proposition 5.7. Here we mostly use the screening procedure as a way to transform a configuration with possibly more or less than $2R$ points in $[-R, R]$ into a configuration with exactly $2R$ points, in such a way that we keep the "old" configuration on a large segment, and that we do not augment much the energy.

PROPOSITION 3.2 (Screening). *Let $s \in \left(0, \frac{1}{4}\right)$ and let $M > 1$. There exists a universal constant $\eta_{\mathrm{scr}} > 0$ and $R_0 = R_0(s, M) > 0$ such that for any $\eta \in (0, \eta_{\mathrm{scr}})$ and any integer $R \geq R_0$, the following holds*:

$$\textit{Put } R' := R(1-s),\ \mathsf{Old} := \Lambda_{R'},\ \textit{and } \mathsf{New} := \Lambda_R \setminus \mathsf{Old}.$$

Let $\mathcal{C}$ be a point configuration in Λ_R, let E be compatible with $\mathcal{C}$ in Old, *and let E_η be the truncation as in* (2.2).

Assume that

$$\mathsf{M}_{\mathrm{scr}} := \int_{\{-R',R'\}\times[-R,R]} |\mathrm{E}_\eta|^2 \leq M, \tag{3.2}$$

$$\mathsf{e}_{\mathrm{scr}} := \frac{1}{s^4 R} \int_{\Lambda_R \times (\mathbb{R}\setminus(-\frac{1}{2}s^2 R, \frac{1}{2}s^2 R))} |\mathrm{E}|^2 \leq 1, \tag{3.3}$$

and that furthermore

$$|\mathcal{C}_{[-R'-2\eta,-R'+2\eta]}| = 0, \quad |\mathcal{C}_{[R'-2\eta,R'+2\eta]}| = 0. \tag{3.4}$$

Then, there exists a probability measure $\Phi^{\mathrm{scr}}_{\mathcal{C},s,\eta,R}$ on point configurations in Λ_R such that for $\Phi^{\mathrm{scr}}_{\mathcal{C},s,\eta,R}$-a.e. configuration $\mathcal{C}^{\mathrm{scr}}$, the following holds:

- *The number of points is given by $|\mathcal{C}^{\mathrm{scr}}_{\Lambda_R}| = |\Lambda_R|$.*
- *The configurations $\mathcal{C}$ and $\mathcal{C}^{\mathrm{scr}}$ coincide on* Old.
- *There is no point at distance less than $\frac{1}{10}$ of $\{-R, R\}$.*
- *The pairs of points of $\mathcal{C}^{\mathrm{scr}}$ that are at distance less than η form a subset of the pairs of points of $\mathcal{C}$ with the same property; in particular,*

$$\begin{aligned} &\iint_{0<|x-y|<\eta} -\log|x-y|\, d\mathcal{C}^{\mathrm{scr}}(x)\, d\mathcal{C}^{\mathrm{scr}}(y) \\ &\quad \leq \iint_{0<|x-y|<\eta} -\log|x-y|\, d\mathcal{C}(x)\, d\mathcal{C}(y). \end{aligned} \tag{3.5}$$

- *There exists an electric field $\mathrm{E}^{\mathrm{scr}}$ in $\mathsf{Screen}(\mathcal{C}^{\mathrm{scr}}, \Lambda_R)$ such that*

$$\int_{\Lambda_R\times\mathbb{R}} \left|\mathrm{E}^{\mathrm{scr}}_\eta\right|^2 \leq \int_{\Lambda_R\times[-R,R]} |\mathrm{E}_\eta|^2 + \mathrm{ErrScr}, \tag{3.6}$$

  with an error term ErrScr *bounded by*

$$\mathrm{ErrScr} \preceq |\log \eta| M s R. \tag{3.7}$$

Moreover, the restriction of $\Phi^{\mathrm{scr}}_{\mathcal{C},s,\eta,R}$ to New *can be written as the point process given by*

$$\mathcal{C}^{\mathrm{scr}}_{\mathsf{New}} := \sum_{i=1}^{n} \delta_{p_i + r_i \eta}, \tag{3.8}$$

where $n := |\mathcal{C}^{\mathrm{scr}}_{\mathsf{New}}|$ is the number of points of $\mathcal{C}^{\mathrm{scr}}$ in New*, the p_i's are points in* New*, and the r_i's are i.i.d. uniform random variables in $\left[-\frac{1}{4}, \frac{1}{4}\right]$. The number n and the points p_i are deterministic for $\Phi^{\mathrm{scr}}_{\mathcal{C},s,\eta,R}$; namely, they depend only on $\mathcal{C}$.*

Remark 3.3. In the screening procedure of [18], designed for being applied to a finite gas, the condition (3.2) is replaced by a bound of the form

$$\frac{1}{R} \int_{\Lambda_R\times[-R,R]} |\mathrm{E}_\eta|^2 \leq M.$$

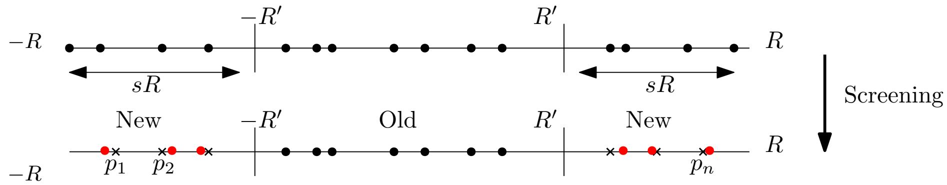

FIGURE 3.1. A before/after illustration of the screening procedure. The crosses correspond to the deterministic positions p_i, which are *not* points of the configuration, but around which we place "new" points (in red).

Then a mean value argument is used to find a "good boundary" on which an estimate of the type (3.2) holds. Taking advantage of the stationary character of the infinite gas, we may skip this step and impose a fixed good boundary. It shortens the procedure a bit, and turns out to be very convenient for the proof of entropy estimates below.

If (3.4) is not mandatory and not used in [18], it simplifies the technical details of the proof.

A Technical Remark on the Screening Construction

The set New consists of two intervals $[-R, -R']$ and $[R', R]$, with $R - R' = sR$. We state here the results for the "left side" $[-R, -R']$; they extend readily to the other side.

CLAIM 3.4 (Position of the new points). *Let $k_{\max}$ be the number of points of $\mathcal{C}^{\mathrm{scr}}$ in $[-R, -R']$ and for $1 \le k \le k_{\max}$, let z_k be the k^{th} point of $\mathcal{C}^{\mathrm{scr}}$ in $[-R, R']$, starting from the leftmost one. We also write*

$$\bar{\mathsf{z}}_k = -R + k - 1/2,$$

which corresponds to the "ideal" position of z_k if the points were regularly spaced.

- *The number $k_{\max}$ satisfies*

$$|k_{\max} - sR| \preceq M^{1/2} s R^{1/2}. \tag{3.9}$$

- *For k such that $s^2 R \preceq |sR - k|$ (for points that are close to $-R$ and far from the boundary point $-R'$ between New and Old), we have*

$$|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq \frac{k}{R^{1/2}}, \tag{3.10}$$

  so in particular we have $|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq \frac{k}{2}$ and $|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq sR^{1/2}$. It also implies

$$|\mathrm{Discr}_{[-R, -R+k]}| \preceq \frac{k}{R^{1/2}}, \tag{3.11}$$

- *For k such that $|sR - k| \preceq s^2 R$ (for points that are close to the boundary point $-R'$ between New and Old and far from $-R$), we have*

$$|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq M^{1/2} s R^{1/2}. \tag{3.12}$$

It also implies

$$|\mathrm{Discr}_{[-R,-R+k]}| \preceq M^{1/2} s R^{1/2}. \tag{3.13}$$

Proposition 3.2 is based on the existing screening procedure. A sketch of the proof with, in particular, a justification of Claim 3.4, is given in Appendix B.

4 Discrepancy Estimate, Gaps

4.1 A Discrepancy Estimate

Let us recall that the discrepancy $\mathrm{Discr}_I(\mathcal{C})$ is defined as $|\mathcal{C}_I| - |I|$ for an interval I, hence $\mathbb{E}_{\mathsf{P}}\big[\mathrm{Discr}_I^2\big]$ can be thought of as the variance, under P, of the number of points in I. It is observed in [16, lemma 3.2], that if $\mathcal{W}(\mathsf{P})$ is finite, we have

$$\limsup_{R\to\infty} \frac{1}{|\Lambda_R|} \mathbb{E}_{\mathsf{P}}\big[\mathrm{Discr}^2_{\Lambda_R}\big] < \infty. \tag{4.1}$$

Moreover, in [16, remark 3.3], it was proven that if P is a stationary point process on $\mathbb{R}$ with finite energy, then

$$\liminf_{R\to\infty} \frac{1}{|\Lambda_R|} \mathbb{E}_{\mathsf{P}}\big[\mathrm{Discr}^2_{\Lambda_R}\big] = 0. \tag{4.2}$$

In fact, an examination of the proof allows for a stronger statement.

LEMMA 4.1 (The variance of the number of points). *Assume that* $\mathcal{W}(\mathsf{P})$ *is finite. Then*

$$\lim_{R\to\infty} \frac{1}{|\Lambda_R|} \mathbb{E}_{\mathsf{P}}\big[\mathrm{Discr}^2_{\Lambda_R}\big] = 0. \tag{4.3}$$

We give the proof of Lemma 4.1 in Appendix A.1. The gain from (4.2) to (4.3) may seem minor, but it turns out to be crucial here (it is also used in a central way in [15]).

4.2 Points, Gaps, and Positions

DEFINITION 4.2 (Points and gaps, counted from the origin). For any point configuration $\mathcal{C}$ on $\mathbb{R}$, let us enumerate the points of $\mathcal{C}$, counted from the origin, as

$$\cdots < x_{-k} < \cdots < x_{-2} < x_{-1} < 0 \le x_0 < x_1 < \cdots < x_k < \cdots. \tag{4.4}$$

For $\mathcal{C}$ fixed and $k \in \mathbb{Z}$, we will write $x_k(\mathcal{C})$ to denote the k^{th} point of $\mathcal{C}$ in the sense of this enumeration. The enumeration (4.4) is well-defined if all the points are simple, which is an event of full measure for all the processes considered here.

As shown on Figure 4.1 below, we define the sequence of gaps "counted from 0" for $\mathcal{C}$ as the sequence $\{\Gamma_k\}_{k\in\mathbb{Z}}$, where

$$\Gamma_k = x_{k+1} - x_k. \tag{4.5}$$

If $\mathcal{C}$ is a finite configuration, we let $\Gamma_k = +\infty$ after the last point is reached in either direction.

DEFINITION 4.3 (Position of the first point after translation). For a given point configuration $\mathcal{C}$ and a real number u, we let $\mathrm{Pos}(\mathcal{C}; u)$ be the integer such that

$$\mathrm{Pos}(\mathcal{C}; u) = m \iff x_0(\mathcal{C} - u) = x_m(\mathcal{C}) \iff \Gamma_0(\mathcal{C} - u) = \Gamma_m(\mathcal{C}). \tag{4.6}$$

In plain words, $\mathrm{Pos}(\mathcal{C}; u)$ is the position in $\mathcal{C}$ of the first point (counted from 0) of the translated configuration $\mathcal{C} - u$.

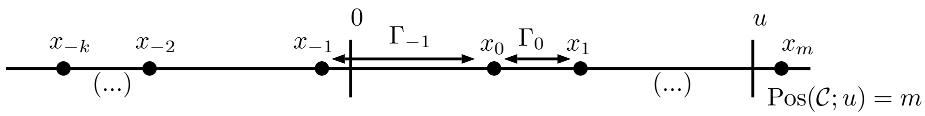

FIGURE 4.1.

LEMMA 4.4 (Gaps are of order 1 in average). *Let* $0 < \eta < 1/2$, $R > 0$, *let* $\mathcal{C}$ *be a configuration in* Λ_R *that has at least* $R/2$ *points*[1] *in* $[0, R]$ *and in* $[-R, 0]$. *We have*

$$\sum_{i=-\frac{R}{2}}^{\frac{R}{2}} (\Gamma_i)^2 \preceq R + \int_{[-R,R]\times\mathbb{R}} |\mathrm{E}_\eta|^2. \tag{4.7}$$

This result says that gaps are "typically of order 1", because in view of (2.14) we expect the second term in the right-hand side of (4.7) to be of order R. We give the proof of Lemma 4.4 in Appendix A.2.

Intuitively, since the processes have intensity 1, the first point of a configuration, counted from the origin, is close to 0, and the k^{th} point is at a distance $\approx k$. Following this heuristic, the first point of $C - u$ is $\approx u$, and more generally the v^{th} point of $\mathcal{C} - u$ is $\approx u + v$ so that $x_{\mathrm{Pos}(\mathcal{C},u)} \approx u$ and more generally that $x_{\mathrm{Pos}(C,u)+v} \approx u + v$. The following lemma shows that these heuristics are correct.

LEMMA 4.5 (Deviations estimates for the positions). *Let* P *be a stationary point process and assume that* $\mathcal{W}(\mathsf{P})$ *is finite. It holds, for any* u, v *in* $\mathbb{Z}$

$$\lim_{w\to+\infty} \mathsf{P}\big(x_{\mathrm{Pos}(\mathcal{C},u)+v}(\mathcal{C}) \geq (u+v) + w\big) = 0. \tag{4.8}$$

Similarly, we have

$$\lim_{w\to+\infty} \mathsf{P}\big(x_{\mathrm{Pos}(\mathcal{C},u)+v}(\mathcal{C}) \leq (u+v) - w\big) = 0. \tag{4.9}$$

PROOF. The fact that $x_{\mathrm{Pos}(\mathcal{C},u)+v}(\mathcal{C}) \geq (u+v) + w$ implies that there are less than v points between u and $u + v + w$, and hence

$$\big\{x_{\mathrm{Pos}(\mathcal{C},u)+v}(\mathcal{C}) \geq w + u + v\big\} \subset \big\{\mathrm{Discr}_{[u,u+v+w]} \geq w.\big\}$$

[1] One could, for example, enforce that $R/2$ is an integer, but more generally for x real, having "at least x points" means "having at least $\lceil x \rceil$ points", $\sum_{k=0}^{x}$, means $\sum_{k=0}^{\lfloor x \rfloor}$, etc.

In view of (4.1), and since P is stationary, the probability of the latter event tends to 0 as $w \to \infty$, for v fixed, independently of u. The proof of (4.9) is identical. □

In particular:

- Taking $u, v = 0$, so that $\mathrm{Pos}(\mathcal{C}, u) + v = 0$ and sending $w \to \infty$, we get

$$\lim_{w\to\infty} \mathsf{P}(x_0(\mathcal{C}) \geq w) = 0. \tag{4.10}$$

  This shows that indeed the first point of a configuration is typically at a bounded distance from 0.
- Taking v fixed, $w = u - v$, and sending $u \to \infty$, we obtain

$$\lim_{u\to\infty} \mathsf{P}(x_{\mathrm{Pos}(\mathcal{C},u)+v}(\mathcal{C}) \geq 2u) = 0. \tag{4.11}$$

  This shows that for v fixed and u large, the v^{th} point of $\mathcal{C} - u$ cannot be much further than $2u$. This estimate is far from being sharp, but it is enough for our purposes.

4.3 Distinct Stationary Processes Have Distinct Gap Distributions

PROPOSITION 4.6 (Distinct gap distributions I). *Let $\mathsf{P}^0, \mathsf{P}^1$ be two stationary point processes such that $\mathcal{W}(\mathsf{P}^0)$ and $\mathcal{W}(\mathsf{P}^1)$ are finite, and assume that $\mathsf{P}^0 \neq \mathsf{P}^1$. Then there exists $c > 0$, an integer $r \geq 1$, and a function $H : \mathbb{R}_+^{2r+1} \to \mathbb{R}$ such that*

- *H is compactly supported, bounded by 1, and 1-Lipschitz with respect to the $\|\cdot\|_1$ norm on $\mathbb{R}_+^{2r+1}$, i.e.,*

$$\|H\|_\infty \leq 1, \quad |H(a_{-r}, \dots, a_r) - H(b_{-r}, \dots, b_r)| \leq \sum_{i=-r}^{r} |a_i - b_i|. \tag{4.12}$$

- *H detects the difference of distribution of the gaps of P^0 and P^1, namely,*

$$\mathbb{E}_{\mathsf{P}^0}[H(\Gamma_{-r}, \dots, \Gamma_r)] - \mathbb{E}_{\mathsf{P}^1}[H(\Gamma_{-r}, \dots, \Gamma_r)] > c. \tag{4.13}$$

We postpone the proof of Proposition 4.6 to Appendix A.3.

PROPOSITION 4.7 (Distinct gap distributions II). *Let $\mathsf{P}^0, \mathsf{P}^1$ be two stationary point processes such that $\mathcal{W}(\mathsf{P}^0)$ and $\mathcal{W}(\mathsf{P}^1)$ are finite, and assume that $\mathsf{P}^0 \neq \mathsf{P}^1$. There exists $\mathfrak{g} > 0$ such that for R large enough, any coupling Q of the restrictions of $\mathsf{P}^0, \mathsf{P}^1$ to Λ_R, and any random variable S bounded by $R^{1/2}$, we have*

$$\mathbb{E}_{\mathsf{Q}}\left[\sum_{i=-R/2}^{R/2} \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_{i+S}(\mathcal{C}^1)\right|^2}{|\Gamma_i(\mathcal{C}^0)|^2 + |\Gamma_{i+S}(\mathcal{C}^1)|^2}\right] \geq \mathfrak{g} R. \tag{4.14}$$

We give the proof of Proposition 4.7 in Section A.4. Intuitively, the reason that (4.14) holds true is the following: we know that the gaps are typically of order 1 (see Lemma 4.4), so the denominator is of order 1, and $\mathsf{P}^0, \mathsf{P}^1$ are stationary so the difference between their gaps distribution on a given interval should be proportional to its length.

Since the quantity

$$\frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_{i+s}(\mathcal{C}^1)\right|^2}{|\Gamma_i(\mathcal{C}^0)|^2 + |\Gamma_{i+s}(\mathcal{C}^1)|^2}$$

is bounded, the following extension of Proposition 4.7 is straightforward.

COROLLARY 4.8. *In the situation of Proposition* 4.7, *if* A, B *are events such that*

$$\mathsf{P}^0(A) \geq 1 - \frac{\mathfrak{g}}{100}, \quad \mathsf{P}^1(B) \geq 1 - \frac{\mathfrak{g}}{100},$$

and if $\widetilde{\mathsf{Q}}$ *is a coupling of the restrictions to* Λ_R *of* P^0 *conditioned to* A *and* P^1 *conditioned to* B, *then*

$$\mathbb{E}_{\widetilde{\mathsf{Q}}}\left[\sum_{i=-R/2}^{R/2} \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_{i+s}(\mathcal{C}^1)\right|^2}{|\Gamma_i(\mathcal{C}^0)|^2 + |\Gamma_{i+s}(\mathcal{C}^1)|^2}\right] \geq \frac{1}{2}\mathfrak{g}R.$$

PROOF. Indeed, we can lift the coupling $\widetilde{\mathsf{Q}}$ of the conditioned processes (restricted to Λ_R) to a coupling of P^0, P^1 (restricted to Λ_R) by defining

$$\mathsf{Q} := \mathsf{P}^0(A)\mathsf{P}^1(B) \cdot \widetilde{\mathsf{Q}} + 1_{(A\times B)^c} \cdot \mathsf{P}^0\big|_{\Lambda_R} \otimes \mathsf{P}^1\big|_{\Lambda_R},$$

and we apply Proposition 4.7 to Q. Letting

$$G := \sum_{i=-R/2}^{R/2} \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_{i+s}(\mathcal{C}^1)\right|^2}{|\Gamma_i(\mathcal{C}^0)|^2 + |\Gamma_{i+s}(\mathcal{C}^1)|^2},$$

which is a quantity bounded by R, we have $\mathbb{E}_{\mathsf{Q}}[G] \geq \mathfrak{g}R$, but also

$$\mathbb{E}_{\widetilde{\mathsf{Q}}}[G] = \mathbb{E}_{\mathsf{Q}}[1_{A\times B}G]\frac{1}{\mathsf{P}^0(A)\mathsf{P}^1(B)} \geq \mathbb{E}_{\mathsf{Q}}[1_{A\times B}G] \geq \mathbb{E}_{\mathsf{Q}}[G] - \mathsf{Q}((A\times B)^c)R,$$

and using the assumptions on $\mathsf{P}^0(A)$, $\mathsf{P}^1(B)$ we get the result. □

5 Large Box Approximation

In this section, we introduce an approximation scheme: given a point process P and parameters $\varepsilon, s > 0$ for R large enough, we are able, after conditioning on a likely event GE_R, to restrict P in Λ_R, to modify the configurations near the edges of Λ_R and to produce a "local candidate" $\widetilde{\mathsf{P}}_R$ in Λ_R that

(1) "looks like" P (in a sense that is controlled by the parameter s),
(2) has energy and entropy comparable to those of P, up to small errors (controlled by ε),
(3) has various good discrepancies bounds (controlled by s), and
(4) has almost surely $2R$ points in Λ_R.

Points 1 and 4 are obtained through the screening procedure described in Section 3. Points 2 and 3 are guaranteed by conditioning on a good event and using properties that hold for processes with finite energy, as mentioned in Section 4. The first step is to check that certain conditions, related to the applicability of the screening procedure or to the good controls on the discrepancy that we want to enforce are often satisfied under P if P has finite energy.

5.1 Conditions That Are Often Satisfied

Let P be a stationary point process such that $\mathcal{W}(\mathsf{P})$ is finite. We recall that to P-almost every configuration is associated an electric field E such that (2.14) holds. We will show that under P with high probability a number of useful estimates on the point configuration and its electric field are satisfied.

The first control is related to the condition (3.2) in the statement of the screening procedure.

CLAIM 5.1 (Good energy). *For all η in $(0,1)$, for all $\delta > 0$, for all M large enough (depending on P, η, δ), for all $s \in \left(0, \frac{1}{4}\right)$ for all $R > 1$, letting $R' = R(1-s)$ (as in Proposition 3.2), the event*

$$\mathsf{GEnergy}(R, s, M, \eta) := \left\{ \int_{\{-R', R'\} \times [-R, R]} |\mathrm{E}_\eta|^2 \leq M \right\}$$

satisfies

$$\mathsf{P}(\mathsf{GEnergy}(R, s, M, \eta)) \geq 1 - \delta.$$

PROOF. Since P is stationary and has finite energy we have, in view of (2.14), for any fixed η, for any $x > 0$,

$$\mathbb{E}_{\mathsf{P}}\left[\int_{\{-x, x\} \times \mathbb{R}} |\mathrm{E}_\eta|^2 \right] = \mathbb{E}_{\mathsf{P}}\left[\frac{1}{R} \int_{\Lambda_R \times \mathbb{R}} |\mathrm{E}_\eta|^2 \right] < +\infty,$$

and the claim follows by applying Markov's inequality. $\square$

The following control is related to the condition (3.3) in the statement of the screening procedure.

CLAIM 5.2 (Good decay of the field on the vertical axis). *For all $\delta > 0$, for all $s > 0$, for all R large enough (depending on P, δ, s), the event*

$$\mathsf{GDecay}(R, s) := \left\{ \frac{1}{s^4 R} \int_{[-R, R] \times (\mathbb{R} \setminus (-s^2 R, s^2 R))} |\mathrm{E}|^2 \leq 1 \right\}$$

satisfies

$$\mathsf{P}(\mathsf{GDecay}(R, s)) \geq 1 - \delta.$$

PROOF. For any R, for any $t > 0$, since P is stationary we have

$$\mathbb{E}_{\mathsf{P}}\left[\frac{1}{R} \int_{[-R, R] \times (\mathbb{R} \setminus (-t, t))} |\mathrm{E}|^2 \right] = \mathbb{E}_{\mathsf{P}}\left[\int_{[-1, 1] \times (\mathbb{R} \setminus (-t, t))} |\mathrm{E}|^2 \right] < +\infty.$$

On the other hand, by dominated convergence we have

$$\lim_{t \to +\infty} \mathbb{E}_{\mathsf{P}}\left[\int_{[-1, 1] \times (\mathbb{R} \setminus (-t, t))} |\mathrm{E}|^2 \right] = 0.$$

Therefore, for any given δ, s, we see that for $t \geq t_0$ large enough (depending on P, δ, s), it holds that

$$\mathbb{E}_{\mathsf{P}}\left[\int_{[-1,1]\times(\mathbb{R}\setminus(-t,t))} |\mathrm{E}|^2\right] \leq \delta s^4.$$

Then for any $R \geq R_0 := \frac{t_0}{s^2}$ (depending on P, δ, s), we have

$$\begin{aligned}\mathsf{P}\left(\frac{1}{s^4 R}\int_{[-R,R]\times(\mathbb{R}\setminus(-s^2R,s^2R))} |\mathrm{E}|^2 > 1\right)\\ \leq \frac{1}{s^4}\mathbb{E}_{\mathsf{P}}\left[\int_{[-1,1]\times(\mathbb{R}\setminus(-s^2R,s^2R))} |\mathrm{E}|^2\right] \leq \delta,\end{aligned} \tag{5.1}$$

which proves the claim. □

DEFINITION 5.3 (The "good field" event). We introduce the event $\mathsf{GField}(R, s, M, \eta)$ as the set of all configurations $\mathcal{C}$ such that there exists a field E satisfying the following three conditions:

(1) E is compatible with $\mathcal{C}$ in $\mathsf{Old} := \Lambda_{R(1-s)} = \Lambda_{R'}$.
(2) E satisfies the "good energy" estimate

$$\int_{\{-R',R'\}\times[-R,R]} |\mathrm{E}_\eta|^2 \leq M.$$

(3) E satisfies the "good decay" estimate

$$\frac{1}{s^4 R}\int_{[-R,R]\times(\mathbb{R}\setminus(-s^2R,s^2R))} |\mathrm{E}|^2 \leq 1.$$

By construction, the event $\mathsf{GField}(R, s, M, \eta)$ depends only on the restriction of $\mathcal{C}$ to $\Lambda_{R'}$ and the global field (compatible with $\mathcal{C}$ on the whole real line) is always a candidate. In view of the definitions above combined with Claim 5.1 and Claim 5.2, we have the following:

CLAIM 5.4 (Good field with high probability).

$$\mathsf{P}(\mathsf{GField}(R, s, M, \eta)) \geq \mathsf{P}(\mathsf{GEnergy}(R, s, M, \eta) \cap \mathsf{GDecay}(R, s)),$$

In particular, for all $\delta > 0$, for all M large enough, for all $s > 0$, and for all R large enough, we have

$$\mathsf{P}(\mathsf{GField}(R, s, M, \eta)) \geq 1 - \delta.$$

The next claim concerns estimates on the discrepancy, the first condition (5.2) is related to the assumption (3.4) in the screening procedure.

CLAIM 5.5 (Good discrepancy estimates). *Let us introduce the following events:*

$$\mathsf{GD}_1(R,s,\eta) := \left\{\left|\mathcal{C}_{[-R'-2\eta,-R'+2\eta]}\right| = \left|\mathcal{C}_{[R'-2\eta,R'+2\eta]}\right| = 0\right\} \tag{5.2}$$

$$\mathsf{GD}_2(R,s) := \left\{\left|\mathrm{Discr}_{[-R',0]}\right| + \left|\mathrm{Discr}_{[0,R']}\right| \leq sR^{1/2}\right\}, \tag{5.3}$$

$$\mathsf{GD}_3(R,s) := \left\{\sum_{i=-R'}^{R'} \left|\mathrm{Discr}_{[-R',i]}\right| \leq s^2 R^{3/2}, \right\}. \tag{5.4}$$

and we let $\mathsf{GDiscr}(R,s,\eta) := \mathsf{GD}_1(R,s,\eta) \cap \mathsf{GD}_2(R,s) \cap \mathsf{GD}_3(R,s)$.

Then for all $\delta > 0$, *for all* $\eta > 0$ *small enough* (*depending on* δ), *for all* $s \in \left(0, \frac{1}{4}\right)$, *for all* R *large enough (depending on* P, δ, s), *we have*

$$\mathsf{P}(\mathsf{GDiscr}(R,s,\eta)) \geq 1 - \delta.$$

PROOF. For GD_1, we recall that P has intensity 1, so for all s, R and $\eta > 0$.

$$\mathbb{E}_{\mathsf{P}}\left[\left|\mathcal{C}_{[-R'-2\eta,-R'+2\eta]}\right|\right] = 4\eta.$$

Of course the number of points in a given interval is an integer-valued random variable. We thus have

$$\mathsf{P}\left[\left|\mathcal{C}_{[-R'-2\eta,-R'+2\eta]}\right| > 0\right] = \mathsf{P}\left[\left|\mathcal{C}_{[-R'-2\eta,-R'+2\eta]}\right| \geq 1\right] \leq 4\eta,$$

and we then take η small enough (depending only on δ).

For GD_2, we use Lemma 4.1 and Markov's inequality. For GD_3, we note that by Lemma 4.1 we have

$$\frac{1}{R^{3/2}} \mathbb{E}_{\mathsf{P}}\left[\sum_{i=-R'}^{R'} |\mathrm{Discr}_{[-R',i]}|\right] \leq \frac{1}{R} \sum_{i=-R'}^{R'} \sqrt{\frac{1}{R}\mathbb{E}_{\mathsf{P}}\left[|\mathrm{Discr}_{[-R',i]}|^2\right]} \longrightarrow 0 \quad \text{as } R \to \infty,$$

and use again Markov's inequality. $\square$

CLAIM 5.6 (Good truncation error). *For all* $\delta, \varepsilon > 0$, *for all* $\eta \in (0,1)$ *small enough* (*depending on* P, δ, ε), *and for all* $R > 1$, *the event*

$$\mathsf{GTrunc}(R,\eta,\varepsilon) := \left\{\iint_{\{x,y \in \Lambda_R \times \Lambda_R, 0<|x-y|<2\eta\}} -\log|x-y| \, d\mathcal{C}(x) d\mathcal{C}(y) \leq \frac{\varepsilon}{100} R\right\} \tag{5.5}$$

satisfies

$$\mathsf{P}(\mathsf{GTrunc}(R,\eta,\varepsilon)) \geq 1 - \delta.$$

PROOF. By stationarity we have, for $R > 1$,

$$\mathbb{E}_{\mathsf{P}}\left[\iint_{\{(x,y)\in\Lambda_R\times\Lambda_R, 0<|x-y|<2\eta\}} -\log|x-y|d\mathcal{C}(x)d\mathcal{C}(y)\right] \\ \preceq R \times \mathbb{E}_{\mathsf{P}}\left[\iint_{\{(x,y)\in\Lambda_1\times\Lambda_1, 0<|x-y|<2\eta\}} -\log|x-y|d\mathcal{C}(x)d\mathcal{C}(y)\right].$$

From [16, lemma 3.5.], we know that

$$\lim_{\eta\to 0} \mathbb{E}_{\mathsf{P}}\left[\iint_{\{(x,y)\in\Lambda_1\times\Lambda_1, 0<|x-y|<2\eta\}} -\log|x-y|d\mathcal{C}(x)d\mathcal{C}(y)\right] = 0,$$

and we use Markov's inequality to conclude. □

5.2 Large Box Approximation

We now state the approximation result.

PROPOSITION 5.7 (Large box approximation). *Let* P *be a stationary process such that* $\mathcal{F}_\beta(\mathsf{P})$ *is finite. Let* $\varepsilon > 0$ *be fixed. There exists an "energy threshold"* M*, depending on* P *and* ε*, such that for any* $s > 0$ *small enough, for all integers* R *large enough* (*depending on* $\mathsf{P}, \varepsilon, s$)*, there exists an event* GE_R *depending only on the restriction of configurations to* Λ_R *such that*

$$\mathsf{P}(\mathsf{GE}_R) \geq 1 - \varepsilon, \tag{5.6}$$

and there exists a probability measure $\widetilde{\mathsf{P}}_R$ *on* $\mathrm{Conf}(\Lambda_R)$ *such that*:

- *the configurations have exactly* $2R$ *points in* Λ_R*,* $\widetilde{\mathsf{P}}_R$*-almost surely, and*
- *the restrictions to* $\mathsf{Old} = \Lambda_{R(1-s)}$ *of the process* $\widetilde{\mathsf{P}}_R$ *and of the process* P *conditioned on* GE_R *coincide. More precisely, if* $|_{\mathsf{Old}}$ *denotes the restriction of a configuration to* Old*, then we have*

$$|_{\mathsf{Old}*}\widetilde{\mathsf{P}}_R = |_{\mathsf{Old}*}\mathsf{P}[\cdot|\mathsf{GE}_R].$$

Moreover, the following inequalities hold:

$$\frac{1}{|\Lambda_R|}\mathbb{E}_{\widetilde{\mathsf{P}}_R}\left[\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_R)\right] \leq \mathcal{W}(\mathsf{P}) + \varepsilon, \tag{5.7}$$

$$\frac{1}{|\Lambda_R|}\mathrm{Ent}\left[\widetilde{\mathsf{P}}_R|\Pi_{|\Lambda_R}\right] \leq \mathcal{E}(\mathsf{P}) + \varepsilon. \tag{5.8}$$

For $\widetilde{\mathsf{P}}_R$*-a.e. configuration* $\mathcal{C}^{\mathrm{scr}}$:

- *The points in* New *follow the conclusions of Claim* 3.4.
- *Letting* S *be the random variable such that*

$$x_0(\mathcal{C}^{\mathrm{scr}}) = \mathsf{z}_{R+S(\mathcal{C}^{\mathrm{scr}})}(\mathcal{C}^{\mathrm{scr}}), \tag{5.9}$$

where $x_0(\mathcal{C}^{\mathrm{scr}})$ *denotes the first point of* $\mathcal{C}^{\mathrm{scr}}$ *to the right of* 0 *as in* (4.4) *and* $\mathsf{z}_k(\mathcal{C}^{\mathrm{scr}})$ *denotes the* k^{th} *point of* $\mathcal{C}^{\mathrm{scr}}$ *to the right of* $-R$*, as in Claim* 3.4*, we have*

$$|S| \leq sM^{1/2}R^{1/2} \quad \widetilde{\mathsf{P}}_R\text{-a.s.} \tag{5.10}$$

• *We have*

$$\sum_{i=-R}^{R} \left|\mathrm{Discr}_{[-R,i]}(\mathcal{C}^{\mathrm{scr}})\right| \preceq s^2 R^{3/2}. \tag{5.11}$$

Let us comment on (5.9), (5.10). For a given configuration with $2R$ points in Λ_R, we consider two different ways of enumerating points: with z_0 starting from the left, as in Claim 3.4, and with x_0 being the first positive point, as in (4.4). There should be roughly R points in $[-R, 0]$, so x_0 should correspond to z_R. The "shift" S defined in (5.9) quantifies the deviation to this guess, and (5.10) affirms that this deviation is small.

The proof relies on the screening procedure of Proposition 3.2. First, we need to show that we can apply this procedure with high probability.

PROOF. The proof is split into five steps. We first construct the event GE_R, from which we build the approximate process $\widetilde{\mathsf{P}}_R$. In the remaining steps we prove the claimed properties (5.7) to (5.11).

Step 1. FINDING A GOOD EVENT. Let $\varepsilon > 0$ be given.

(a) *The truncation.* First, since P is stationary, we know from (2.14) that, for any $R > 1$,

$$\lim_{\eta\to 0} \mathbb{E}_{\mathsf{P}}\left[\frac{1}{|\Lambda_R|}\frac{1}{2\pi}\int_{\Lambda_R\times\mathbb{R}} |\mathrm{E}_\eta|^2 + \log\eta\right] = \mathcal{W}(\mathsf{P}),$$

and the limit is uniform in R because the quantity (the expectation) whose limit is taken is independent of R, by stationarity. In particular, for η small enough (depending only on P), for all $R > 1$, we have

$$\left|\mathbb{E}_{\mathsf{P}}\left[\frac{1}{|\Lambda_R|}\frac{1}{2\pi}\int_{\Lambda_R\times\mathbb{R}} |\mathrm{E}_\eta|^2 + \log\eta\right] - \mathcal{W}(\mathsf{P})\right| \leq \frac{\varepsilon}{100}, \tag{5.12}$$

Another important feature for choosing a suitable truncation is that the error terms in the monotonicity estimates of Lemma 2.5 should be small. By Claim 5.6, we may take some $\eta > 0$ (depending on P, ε) such that

$$\mathsf{P}(\mathsf{GTrunc}(R, \eta, \varepsilon)) \geq 1 - \frac{\varepsilon}{100}.$$

and by taking η smaller if needed, we may also impose that

$$\mathrm{ErMono}(\eta) < \frac{\varepsilon}{100}, \quad \eta < \eta_{\mathrm{scr}}, \tag{5.13}$$

where ErMono is the error term in (2.8), (2.10), and η_{scr} is the constant in Proposition 3.2.

(b) *The energy threshold.* We now fix an energy threshold M. We take it high enough such that most configurations have an energy at most M. By Claim 5.4, for M large enough depending on P, ε, for any $s \in (0, 1/4)$, for R large enough (depending on $\mathsf{P}, s, M, \varepsilon$) we have

$$\mathsf{P}(\mathsf{GField}(R, s, M, \eta)) \geq 1 - \frac{\varepsilon}{100}. \tag{5.14}$$

(c) *The size of the screening zone.* Next, we fix the size of the screening zone. The screening procedure affects a small region near the endpoints and has an energy cost proportional to this size, controlled by s. We choose $s \in (0, 1/4)$ small enough such that for all R

$$\mathrm{ErrScr} \leq \frac{\varepsilon R}{100}, \tag{5.15}$$

where ErrScr is the error term in (3.6). This choice of s depends only on η, M.

(d) *The discrepancy.* Claim 5.5 ensures that, for R large enough, depending on $\mathsf{P}, \varepsilon, s$, we have

$$\mathsf{P}(\mathsf{GDiscr}(R, s, \eta)) \geq 1 - \frac{\varepsilon}{100}.$$

(e) *The good event.* Let ε, η, M, s be as chosen in the previous paragraphs, and for all $R > 0$ we let GE_R be the intersection of the previous events, i.e.,

$$\mathsf{GE}_R := \mathsf{GTrunc}(R, \eta, \varepsilon) \cap \mathsf{GField}(R, s, M, \eta) \cap \mathsf{GDiscr}(R, s, \eta).$$

By a union bound, choosing R large enough depending only on $\mathsf{P}, \varepsilon, s$ we obtain that $\mathsf{P}(\mathsf{GE}_R) \geq 1 - \varepsilon$, which ensures that (5.6) holds. By Claim 5.4, GField is measurable with respect to the restriction to Λ_R. Moreover, GTrunc, GDiscr clearly depend only on $\mathcal{C}$ in Λ_R. Therefore, GE_R depends only on the restriction to Λ_R.

Step 2. DEFINING THE MODIFIED PROCESS. Let $\mathcal{C}$ be in GE_R. By definition of $\mathsf{GField}(R, s, M, \eta)$, there exists an electric field E that is compatible with $\mathcal{C}$ on $\mathsf{Old} \subset \Lambda_R$ and satisfying

$$\mathsf{M}_{\mathrm{scr}} := \int_{\{-R', R'\} \times [-R, R]} |\mathrm{E}_\eta|^2 \leq M,$$
$$\mathsf{e}_{\mathrm{scr}} := \frac{1}{s^4 R} \int_{[-R,R] \times \mathbb{R} \setminus (-s^2 R, s^2 R)} |\mathrm{E}|^2 \leq 1.$$

In the following, we choose a field E that satisfies these conditions *and has minimal energy*, i.e., $\mathsf{M}_{\mathrm{scr}}$ is minimal. In particular, this electric field E depends only on $\mathcal{C}$ in Old.

The assumption (3.4) is also satisfied by definition of $\mathsf{GDiscr}(R, s, \eta)$, which is included in $\mathsf{GD}_1(R, s, \eta)$ as defined in (5.2).

Hence for all R large enough (depending on $\mathsf{P}, \varepsilon, s$) we may apply the screening procedure to $\mathcal{C}$. We let $\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})$ be the resulting probability measure on $\mathrm{Conf}(\Lambda_R)$. We define $\widetilde{\mathsf{P}}_R$ as the mixture of the $\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})$ for $\mathcal{C}$ in GE_R, weighted by $\mathsf{P}(\mathcal{C})$, i.e.,

$$\widetilde{\mathsf{P}}_R := \frac{1}{\mathsf{P}(\mathsf{GE}_R)} \int \Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C}) 1_{\mathcal{C}\in\mathsf{GE}_R}\, d\mathsf{P}(\mathcal{C}). \tag{5.16}$$

Equivalently, if F is a bounded function on $\mathrm{Conf}(\Lambda_R)$, we let

$$\mathbb{E}_{\widetilde{\mathsf{P}}_R}[F] = \frac{1}{\mathsf{P}(\mathsf{GE}_R)} \int \mathbb{E}_{\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})}[F] 1_{\mathcal{C}\in\mathsf{GE}_R}\, d\mathsf{P}(\mathcal{C}). \tag{5.17}$$

We may already check the following:

- By construction, the screened configurations all have $2R$ points in Λ_R.
- Still by construction, for a given $\mathcal{C} \in \mathsf{GE}_R$, all the configurations $\mathcal{C}^{\mathrm{scr}}$ in $\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})$ coincide with $\mathcal{C}$ inside $\mathsf{Old} = \Lambda_{R'}$. From the definition (5.17), we thus see that the restrictions to $\Lambda_{R'}$ of the process $\widetilde{\mathsf{P}}_R$ and of the process P conditioned to GE_R coincide, in other words,

$$|_{\Lambda_{R'}*}\widetilde{\mathsf{P}}_R = |_{\Lambda_{R'}*}\mathsf{P}[\cdot|\mathsf{GE}_R].$$

Step 3. ENERGY ESTIMATE. We prove (5.7). Let $\mathcal{C}$ be in GE_R. By the screening procedure, for all $\mathcal{C}^{\mathrm{scr}}$ in the support of $\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})$, there exists a screened electric field whose energy, in view of the controls (3.6), (3.7) on the energy after screening, and by the choice of s as in (5.15), is bounded as follows:

$$\int_{\Lambda_R\times\mathbb{R}} \left|\mathrm{E}^{\mathrm{scr}}_\eta\right|^2 \le \int_{\Lambda_R\times[-R,R]} |\mathrm{E}_\eta|^2 + \frac{\varepsilon R}{100}. \tag{5.18}$$

Let $\mathrm{E}^{\mathrm{loc}}$ be the local electric field generated by $\mathcal{C}^{\mathrm{scr}}$ in Λ_R, as in (2.4). By "minimality of the local energy"[2] (see, e.g., [16, lemma 3.10]), we have

$$\int_{\mathbb{R}\times\mathbb{R}} \left|\mathrm{E}^{\mathrm{loc}}_\eta\right|^2 \le \int_{\Lambda_R\times\mathbb{R}} |\mathrm{E}^{\mathrm{scr}}_\eta|^2. \tag{5.19}$$

Combining (5.18) and (5.19), we get

$$\int_{\mathbb{R}\times\mathbb{R}} \left|\mathrm{E}^{\mathrm{loc}}_\eta\right|^2 \le \int_{\Lambda_R\times[-R,R]} |\mathrm{E}_\eta|^2 + \frac{\varepsilon R}{100}. \tag{5.20}$$

Next, the monotonicity inequality (2.10) reads

$$\begin{aligned}\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R) \le \frac{1}{2\pi} \int_{\mathbb{R}\times\mathbb{R}} \left|\mathrm{E}^{\mathrm{loc}}_\eta\right|^2 &+ |\mathcal{C}^{\mathrm{scr}}| \log \eta + |\mathcal{C}^{\mathrm{scr}}| \mathrm{ErMono}(\eta) \\ &+ \mathrm{ErTrun}(\mathcal{C}^{\mathrm{scr}}, \eta; \Lambda_R).\end{aligned}$$

By construction we have $\mathrm{ErTrun}(\mathcal{C}, \eta; \Lambda_R) \le \frac{\varepsilon R}{100}$, and we know by (3.5) that the screening procedure does not augment the truncation error. Moreover, η has been

[2] This is the only moment where we really use the fact that the new electric field that we have produced is *screened*.

chosen such that $\mathrm{ErMono}(\eta) \leq \frac{\varepsilon}{100}$. Using the fact that $|\mathcal{C}^{\mathrm{scr}}| = |\Lambda_R| = 2R$, we obtain

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R) \leq \frac{1}{2\pi} \int_{\mathbb{R}\times\mathbb{R}} \left|\mathrm{E}_\eta^{\mathrm{loc}}\right|^2 + 2R \log \eta + \frac{4R\varepsilon}{100}. \tag{5.21}$$

Combining (5.20) and (5.21) yields

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R) \leq \frac{1}{2\pi} \int_{\Lambda_R \times [-R,R]} |\mathrm{E}_\eta|^2 + 2R \log \eta + \frac{5R\varepsilon}{100}. \tag{5.22}$$

Moreover, we obtain, in view of the control (5.22) on $\mathrm{W}^{\mathrm{int}}$ and the definition (5.17) of $\widetilde{\mathsf{P}}_R$,

$$\begin{aligned} &\mathbb{E}_{\widetilde{\mathsf{P}}_R}\left[\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R)\right] \\ &\quad = \frac{1}{\mathsf{P}(\mathsf{GE}_R)} \int \mathbb{E}_{\Phi^{\mathrm{scr}}_{s,\eta,R}(\mathcal{C})}\left[\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R)\right] 1_{\mathsf{GE}_R}(\mathcal{C}) d\,\mathsf{P}(\mathcal{C}) \\ &\quad \leq \frac{1}{\mathsf{P}(\mathsf{GE}_R)} \int \left(\frac{1}{2\pi} \int_{\Lambda_R \times [-R,R]} |\mathrm{E}_\eta|^2 + 2R \log \eta \right) 1_{\mathsf{GE}_R}(\mathcal{C}) d\,\mathsf{P}(\mathcal{C}) + \frac{5R\varepsilon}{100}. \end{aligned} \tag{5.23}$$

Then since $\mathsf{P}(\mathsf{GE}_R) \geq 1 - \varepsilon$, and (5.12) holds, we write

$$\begin{aligned} &\frac{1}{\mathsf{P}(\mathsf{GE}_R)} \int \left(\frac{1}{2\pi} \int_{\Lambda_R \times [-R,R]} |\mathrm{E}_\eta|^2 + 2R \log \eta \right) 1_{\mathsf{GE}_R}(\mathcal{C}) d\,\mathsf{P}(\mathcal{C}) \\ &\quad \leq |\Lambda_R|(1+\varepsilon)(\mathcal{W}(\mathsf{P}) + \varepsilon), \end{aligned}$$

so we obtain

$$\frac{1}{|\Lambda_R|} \mathbb{E}_{\widetilde{\mathsf{P}}_R}\left[\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathrm{scr}}; \Lambda_R)\right] \leq \mathcal{W}(\mathsf{P}) + \varepsilon \mathcal{W}(\mathsf{P}) + O(\varepsilon).$$

Up to replacing ε with a smaller ε', we may ensure that $\varepsilon' \mathcal{W}(\mathsf{P}) + O(\varepsilon') \leq \varepsilon$ and so (5.7) holds.

Step 4. Entropy estimate. We prove (5.8). Let us first recall a general disintegration principle.

LEMMA 5.8 (Relative entropy and disintegration). *Let X, Y be Polish spaces, let μ, π be two probability measures on X, and let $T : X \to Y$ be a measurable map. Let $\bar{\mu} = T_*\mu$ and $\bar{\pi} = T_*\pi$, and let $\mu(\cdot|T = y)$ and $\pi(\cdot|T = y)$ denote the regular conditional probabilities*; *i.e., we have the disintegration*

$$\mu(A) = \int_Y \mu(A|T = y) d\bar{\mu}(y), \qquad \pi(A) = \int_Y \pi(A|T = y) d\bar{\pi}(y) \quad \forall A.$$

Then we have that

$$\mathrm{Ent}[\mu|\pi] = \mathrm{Ent}[\bar{\mu}|\bar{\pi}] + \int_Y \mathrm{Ent}\left[\mu(\cdot|T = y)|\pi(\cdot|T = y)\right] d\bar{\mu}(y). \tag{5.24}$$

This can be verified by a direct computation. Since the relative entropy w.r.t. a probability is nonnegative, we have in particular

$$\mathrm{Ent}[\mu|\pi] \geq \mathrm{Ent}[\bar{\mu}|\bar{\pi}]. \tag{5.25}$$

To prove (5.8), we will repeatedly use (5.24) at the different steps of the construction of $\widetilde{\mathsf{P}}_R$. Let us briefly summarize some important properties from its construction:

- First, we condition P to the "good event" GE_R, with $\mathsf{P}(\mathsf{GE}_R) \geq 1 - \varepsilon$, then we restrict to Λ_R. More precisely, we consider the process

$$\check{\mathsf{P}}_R = |_{\Lambda_R *}\mathsf{P}[\cdot|\mathsf{GE}_R], \tag{5.26}$$

  namely, the restriction to Λ_R of the process P conditioned to GE_R.
- Then we apply the screening procedure. Given a configuration $\mathcal{C}$ in GE_R, we find (see Figure 3.1):
  - A number $n = n(\mathcal{C})$ such that $n(\mathcal{C}) + |\mathcal{C}|_{\mathsf{Old}}| = 2R$.
  - Points $p_1(\mathcal{C}), \dots, p_n(\mathcal{C})$ in New with the segments $[p_i - \eta, p_i + \eta]$ being disjoint and not intersecting $[-R', R']$.
- We define $\Phi^{\mathrm{scr}}(\mathcal{C})$ by adding $n(\mathcal{C})$ points placed independently, one into each interval of the form $[p_i - \eta, p_i + \eta]$.
- We define $\widetilde{\mathsf{P}}_R$ as the weighted average of $\Phi^{\mathrm{scr}}(\mathcal{C})$ for $\mathcal{C}$ in GE_R, as in (5.16).

We now prove (5.8) in two steps.

Step 4a. We first estimate the entropy of $\check{\mathsf{P}}_R$ and claim that, for some constant C (independent of P, R)

$$\mathrm{Ent}[\check{\mathsf{P}}_R|\Pi_{\Lambda_R}] \leq (1 + 2\varepsilon)\mathrm{Ent}[\mathsf{P}_{\Lambda_R}|\Pi_{\Lambda_R}] + C. \tag{5.27}$$

Indeed, if ρ is the density of $\mathsf{P}_{\Lambda_R} := |_{\Lambda_R *}\mathsf{P}$ with respect to Π_{Λ_R}, the density of $\check{\mathsf{P}}_R$ is given by $\mathtt{1}_E \rho/\mathsf{P}(E)$, where we set $E = \mathsf{GE}_R$ for brevity. Thus we have

$$\begin{aligned}\mathrm{Ent}[\check{\mathsf{P}}_R|\Pi_{\Lambda_R}] &= \int_E \frac{\rho}{\mathsf{P}(E)} \log\left(\frac{\rho}{\mathsf{P}(E)}\right) d\Pi_{\Lambda_R} \\ &= -\log \mathsf{P}(E) + \frac{1}{\mathsf{P}(E)}\left(\mathrm{Ent}\big[\mathsf{P}_{\Lambda_R}|\Pi_{\Lambda_R}\big] - \int_{E^c} \rho \log \rho \, d\Pi_{\Lambda_R}\right).\end{aligned}$$

By Jensen's inequality, we have

$$\frac{1}{\Pi_{\Lambda_R}(E^c)} \int_{E^c} \rho \log \rho \, d\Pi_{\Lambda_R} \geq \frac{\mathsf{P}_{\Lambda_R}(E^c)}{\Pi_{\Lambda_R}(E^c)} \log \frac{\mathsf{P}_{\Lambda_R}(E^c)}{\Pi_{\Lambda_R}(E^c)} \geq -e^{-1},$$

since $r \log r \geq -e^{-1}$. Thus we obtain

$$\begin{aligned}\mathrm{Ent}[\check{\mathsf{P}}_R|\Pi_{\Lambda_R}] \leq \mathrm{Ent}[\mathsf{P}_{\Lambda_R}|\Pi_{\Lambda_R}] + \frac{1 - \mathsf{P}(E)}{\mathsf{P}(E)} \mathrm{Ent}[\mathsf{P}_{\Lambda_R}|\Pi_{\Lambda_R}] - \log \mathsf{P}(E) \\ + \frac{1}{e} \frac{\Pi_{\Lambda_R}(E^c)}{\mathsf{P}_{\Lambda_R}(E)}.\end{aligned}$$

Using the fact that $\mathsf{P}_{\Lambda_R}(E) = \mathsf{P}(E) \geq 1 - \varepsilon$, and assuming, for example, $\varepsilon \leq \frac{1}{2}$, yields (5.27).

Step 4b. Let us now compare $\mathrm{Ent}[\check{\mathsf{P}}_R|\Pi_{\Lambda_R}]$ and $\mathrm{Ent}[\widetilde{\mathsf{P}}_R|\Pi_{\Lambda_R}]$.

Let $\mathrm{Conf}(\Lambda_R)$ denote the set of point configurations in Λ_R and $\mathrm{Conf}(\mathsf{Old})$ denote the set of point configurations on $\mathsf{Old} = \Lambda_{R'}$. Let $T : \mathrm{Conf}(\Lambda_R) \to \mathrm{Conf}(\mathsf{Old})$ be the map given by the restriction $\mathcal{C} \mapsto \mathcal{C}\big|_{\mathsf{Old}}$. Let

$$\Pi^{\mathcal{C}}_{\Lambda_R} = \Pi_{\Lambda_R}(\cdot|T = \mathcal{C})$$

denote the disintegration of Π_{Λ_R} w.r.t. T and similarly define $\check{\mathsf{P}}^{\mathcal{C}}_R, \widetilde{\mathsf{P}}^{\mathcal{C}}_R$. Using the disintegration formula (5.24) for the map T, we have

$$\begin{aligned}\mathrm{Ent}[\check{\mathsf{P}}_R|\Pi_{\Lambda_R}] &= \mathrm{Ent}[T_*\check{\mathsf{P}}_R|T_*\Pi_{\Lambda_R}] \\ &\quad + \int_{\mathrm{Conf}(\mathsf{Old})} \mathrm{Ent}\big[\check{\mathsf{P}}^{\mathcal{C}}_R\big|\Pi^{\mathcal{C}}_{\Lambda_R}\big]d(T_*\check{\mathsf{P}}_R)(\mathcal{C}), \\ \mathrm{Ent}[\widetilde{\mathsf{P}}_R|\Pi_{\Lambda_R}] &= \mathrm{Ent}[T_*\widetilde{\mathsf{P}}_R|T_*\Pi_{\Lambda_R}] \\ &\quad + \int_{\mathrm{Conf}(\mathsf{Old})} \mathrm{Ent}\big[\widetilde{\mathsf{P}}^{\mathcal{C}}_R\big|\Pi^{\mathcal{C}}_{\Lambda_R}\big]d(T_*\widetilde{\mathsf{P}}_R)(\mathcal{C}),\end{aligned} \tag{5.28}$$

Note that $T_*\Pi_{\Lambda_R} = \Pi_{\mathsf{Old}}$. Further, we have by construction that $T_*\widetilde{\mathsf{P}}_R = T_*\check{\mathsf{P}}_R$; hence the first terms in the decomposition of the entropy of $\check{\mathsf{P}}_R$ and $\widetilde{\mathsf{P}}_R$ coincide, and it remains to compare the conditional entropies.

By definition of the Poisson point process, a random configuration drawn from $\Pi_{\Lambda_R}(\cdot|T = \mathcal{C})$ consists of $\mathcal{C}$ in Old plus a sample of points in $\mathsf{New} = \Lambda_R \setminus \mathsf{Old}$ drawn from Π_{New}, the Poisson process in New. Denoting by $\widetilde{\mathsf{P}}^{\mathcal{C},\mathsf{New}}_R$ the image of $\widetilde{\mathsf{P}}_R(\cdot|T = \mathcal{C})$ under restriction to New we have

$$\mathrm{Ent}\big[\widetilde{\mathsf{P}}^{\mathcal{C}}_R\big|\Pi^{\mathcal{C}}_{\Lambda_R}\big] = \mathrm{Ent}\big[\widetilde{\mathsf{P}}^{\mathcal{C},\mathsf{New}}_R\big|\Pi_{\mathsf{New}}\big].$$

By construction, a random configuration drawn from $\widetilde{\mathsf{P}}_R(\cdot|T = \mathcal{C})$ consists of $\mathcal{C}$ in Old plus $n(\mathcal{C})$ points placed independently uniformly within the intervals $(p_i(\mathcal{C})-\eta, p_i(\mathcal{C})+\eta)$. To calculate the entropy with respect to the Poisson process in New, let us introduce another disintegration concerning the number of points placed in New. For a set $I \subset \mathbb{R}$ we set

$$\mathrm{Conf}(I)^{(n)} := \{\mathcal{C} \in \mathrm{Conf}(I), |\mathcal{C}| = n\}.$$

Given a probability σ on $\mathrm{Conf}(I)$ we consider the restrictions

$$\sigma_n := \frac{1}{\sigma(n)}\sigma\bigg|_{\mathrm{Conf}(I)^{(n)}} \qquad \text{where we denote } \sigma(n) := \sigma\big(\mathrm{Conf}(I)^{(n)}\big). \tag{5.29}$$

By definition of the Poisson point process, we have $\Pi_I(n) = e^{-|I|}(|I|^n/n!)$, and the law of $\Pi_{I,n}$ is that of n independent points placed uniformly in I. Using again the disintegration formula (5.24), we get

$$\mathrm{Ent}[\sigma|\Pi_I] = \sum_{n=0}^{\infty} \sigma(n)\bigg[\log\bigg(\frac{\sigma(n)}{\Pi_I(n)}\bigg) + \mathrm{Ent}[\sigma_n|\Pi_{I,n}]\bigg]. \tag{5.30}$$

Recalling that $\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}}$ is the law of $n(\mathcal{C})$ independent points, each placed uniformly into $n(\mathcal{C})$ intervals of size 2η, we find that for $n = n(\mathcal{C})$ the entropy of $(\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}})_n$ with respect to $\Pi_{\mathsf{New},n}$ is given by the entropy of the uniform distribution on the set

$$B_{n,\eta} := \bigcup_{\sigma \in S_n} (p_{\sigma(1)} - \eta, p_{\sigma(1)} + \eta) \times \cdots \times (p_{\sigma(n)} - \eta, p_{\sigma(n)} + \eta)$$

relative to the uniform distribution on New^n, i.e.,

$$\mathrm{Ent}\big[\big(\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}}\big)_n | \Pi_{\mathsf{New},n}\big] = \log \frac{|\mathsf{New}|^n}{|B_{n,\eta}|} = \log \frac{|\mathsf{New}|^n}{n!(2\eta)^n}.$$

By construction we have $\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}}(m) = 1$ for $m = n(\mathcal{C})$ and 0 otherwise. Thus we obtain from (5.30) that for $n = n(\mathcal{C})$

$$\begin{aligned}\mathrm{Ent}\big[\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}} | \Pi_{\mathsf{New}}\big] &= \log \frac{1}{e^{-|\mathsf{New}|}|\mathsf{New}|^n/n!} + \mathrm{Ent}\big[\big(\widetilde{\mathsf{P}}_R^{\mathcal{C},\mathsf{New}}\big)_n \big| \Pi_{\mathsf{New},n}\big] \\ &= \log \frac{e^{|\mathsf{New}|}n!}{|\mathsf{New}|^n} + \log \frac{|\mathsf{New}|^n}{n!(2\eta)^n} = |\mathsf{New}| - n(\mathcal{C}) \log 2\eta \\ &= 2sR - (2R - \mathcal{C}(\mathsf{Old})) \log 2\eta.\end{aligned}$$

Finally, note that $\mathrm{Ent}[\check{\mathsf{P}}_R^{\mathcal{C}} | \Pi_{\Lambda_R}^{\mathcal{C}}] \geq 0$. Thus combining (5.28) and (5.31) we finally obtain

$$\mathrm{Ent}\big[\widetilde{\mathsf{P}}_R \big| \Pi_{\Lambda_R}\big] \leq \mathrm{Ent}\big[\check{\mathsf{P}}_R \big| \Pi_{\Lambda_R}\big] + 2sR - \log 2\eta \mathbb{E}_{\check{\mathsf{P}}_R}[2R - |\mathcal{C}_{\mathsf{Old}}|].$$

By construction (see (5.3)), we have

$$|2R - |\mathcal{C}_{\mathsf{Old}}|| \leq 2sR + sR^{1/2} \leq 3sR.$$

We thus obtain

$$\mathrm{Ent}[\widetilde{\mathsf{P}}_R | \Pi_{\Lambda_R}] \leq \mathrm{Ent}[\check{\mathsf{P}}_R | \Pi_{\Lambda_R}] + 5|\log \eta| sR. \tag{5.31}$$

Combining (5.27) and (5.31), we obtain (for R large, the constant error term in (5.27) can be absorbed in the dominant error terms)

$$\frac{1}{|\Lambda_R|}\mathrm{Ent}[\widetilde{\mathsf{P}}_R | \Pi_{\Lambda_R}] \leq \frac{1}{|\Lambda_R|}\mathrm{Ent}[\mathsf{P}_{\Lambda_R} | \Pi_{\Lambda_R}] + \frac{2\varepsilon}{|\Lambda_R|}\mathrm{Ent}[\mathsf{P}_{\Lambda_R} | \Pi_{\Lambda_R}] + \frac{5}{2}|\log \eta| s.$$

First, we observe that the error term $|\log \eta| s$ is of the same order as ErrScr and s was chosen small enough so that ErrScr is small. On the other hand the limit defining the specific relative entropy is non-decreasing (it follows from a super-additivity argument, see e.g., [9, Cor. 6.77] where S there is the opposite of our entropy), so

$$\frac{1}{|\Lambda_R|}\mathrm{Ent}[\mathsf{P}_{\Lambda_R} | \Pi_{\Lambda_R}] + \frac{2\varepsilon}{|\Lambda_R|}\mathrm{Ent}\big[\mathsf{P}_{\Lambda_R} | \Pi_{\Lambda_R}\big] \leq \mathcal{E}(\mathsf{P}) + 2\varepsilon\mathcal{E}(\mathsf{P}),$$

and up to taking a smaller ε' we may ensure that $2\varepsilon'\mathrm{Ent}[\mathsf{P}_{\Lambda_R} | \Pi_{\Lambda_R}] + 5|\log \eta| s \leq \varepsilon$; hence we obtain (5.8).

Step 5. ADDITIONAL BOUNDS. Since (5.4) holds, and since Claim 3.4 ensures that there are $sR \pm sM^{1/2}R^{1/2}$ points in $[-R, -R']$, we see that there are $R \pm sM^{1/2}R^{1/2}$ points of $\mathcal{C}^{\mathrm{scr}}$ in $[-R, 0]$ and thus (5.10) holds.

Finally, (5.11) follows from combining the definition of GD_3 in Old with the discrepancy estimates in New (3.11), (3.13). □

6 Displacement Convexity

6.1 Labeling on the Orthant

In this section, it is more convenient to treat the laws of random $2R$-point configurations in Λ_R as probability measures on $2R$-tuples. In order to identify a configuration and a $2R$-tuple, we introduce the orthant $\mathcal{O}_R$:

$$\mathcal{O}_R := \left\{(\mathsf{z}_1, \dots, \mathsf{z}_{2R}) \in (\Lambda_N)^{2R} \mid \mathsf{z}_1 \le \dots \le \mathsf{z}_{2R}\right\}. \tag{6.1}$$

- We denote by Conf_R the set of point configurations in Λ_R with exactly $2R$ points.
- We let B_R be the Bernoulli point process with $2R$ points in Λ_R, which is a probability measure on Conf_R.
- We let $\mathbf{L}_R$ be the normalized Lebesgue measure on $\mathcal{O}_R$.

DEFINITION 6.1 (Label map). We define the label map π_R as

$$\pi_R : \begin{cases} \mathrm{Conf}_R \longrightarrow \mathcal{O}_R \subset (\Lambda_R)^{2R} \\ \mathcal{C} = \sum_{i=1}^{2R} \delta_{\mathsf{z}_i} \mapsto (\mathsf{z}_1, \dots, \mathsf{z}_{2R}),\ \mathsf{z}_1 \le \dots \le \mathsf{z}_{2R}. \end{cases} \tag{6.2}$$

LEMMA 6.2 (π_R is essentially bijective). *The label map π_R is well-defined and is a bijection from* Conf_R *to $\mathcal{O}_R$, up to a subset of measure zero* (*for* B_R) *in the source and a subset of measure zero (for $\mathbf{L}_R$) in the target.*

PROOF. The label map π_R as in (6.2) is well-defined and injective on the set of simple configurations when all points are distinct. The image of this set in $\mathcal{O}_R$ is the set of strictly ordered $2R$-tuples $\mathsf{z}_1 < \mathsf{z}_2 < \dots < \mathsf{z}_{2R}$. It is clear that the first set has full measure in Conf_R (for B_R), and the second set has full measure in $\mathcal{O}_R$ (for $\mathbf{L}_R$). □

The following fact is easy to check.

LEMMA 6.3 (Effect of π_R on B_R). B_R *and* $\mathbf{L}_R$ *are images of each other under* π_R *and its inverse, i.e.,* $\pi_{R*}\mathsf{B}_R = \mathbf{L}_R$ *and* $(\pi_R^{-1})_*\mathbf{L}_R = \mathsf{B}_R$.

LEMMA 6.4 (Effect of π_R on the entropy). *Let* P *be a point process on* Λ_R *with almost surely* $2R$ *points, i.e.,* $\mathsf{P}(\mathrm{Conf}_R) = 1$, *and let* $\widehat{\mathsf{P}}$ *be its image under* π_R, *i.e.,* $\widehat{\mathsf{P}} = (\pi_R)_*\mathsf{P}$. *Then, there exists a constant* c_R *depending only on* R *such that*

$$\mathrm{Ent}[\mathsf{P}|\Pi_{\Lambda_R}] = \mathrm{Ent}[\widehat{\mathsf{P}}|\mathbf{L}_R] + c_R. \tag{6.3}$$

PROOF. Since P has almost surely $2R$ points and noting that the restriction of Π_{Λ_R} to Conf_R is B_R, we infer from the disintegration formula (5.30) that

$$\mathrm{Ent}[\mathsf{P}|\Pi_{\Lambda_R}] = \log \frac{e^{2R}(2R)!}{|2R|^{2R}} + \mathrm{Ent}[\mathsf{P}|\mathsf{B}_R].$$

Since π_R is essentially bijective, we have that $\mathsf{P} = (\pi_R^{-1})_* \widehat{\mathsf{P}}$. By Lemma 6.3, $\mathbf{L}_R = (\pi_R)_* \mathsf{B}_R$ and $\mathsf{B}_R = (\pi_R^{-1})_* \mathbf{L}_R$. Now, the claim follows immediately from (5.24). $\square$

6.2 The Optimal Transportation Map and Convexity of the Entropy

Let $\widehat{\mathsf{P}}^0, \widehat{\mathsf{P}}^1$ be two probability measures on $\mathcal{O}_R$, with finite relative entropy with respect to $\mathbf{L}_R$. In particular, they are absolutely continuous with respect to $\mathbf{L}_R$.

PROPOSITION 6.5 (Existence of a transportation map). *There exists a map* $\mathsf{T}_R : \mathcal{O}_R \to \mathbb{R}^{2R}$ *satisfying the following*:

- *The push-forward of* $\widehat{\mathsf{P}}^0$ *by* T_R *is* $\widehat{\mathsf{P}}^1$.
- *There exists a convex function* $\varphi : \mathcal{O}_R \to \mathbb{R}$ *such that* $\mathsf{T}_R = \nabla\varphi$.

PROOF. This follows from [27, theorem 2.12], because $\widehat{\mathsf{P}}^0$ and $\widehat{\mathsf{P}}^1$ are both compactly supported probability measures on $\mathbb{R}^{2R}$, absolutely continuous with respect to the Lebesgue measure. $\square$

The first item expresses the fact that T_R transports $\widehat{\mathsf{P}}^0$ onto $\widehat{\mathsf{P}}^1$. The fact that T_R is the gradient of a convex function expresses the *optimality* of T_R for the quadratic cost. We will use the optimal character of T_R only once, to argue that the relative entropy is displacement convex.

DEFINITION 6.6 (Displacement interpolation). For any $t \in [0, 1]$, we introduce the displacement interpolate $\widehat{\mathsf{P}}^t$ as

$$\widehat{\mathsf{P}}^t := ((1-t)\mathrm{Id} + t\mathsf{T}_R)_* \widehat{\mathsf{P}}^0, \tag{6.4}$$

which is consistent with the previously defined $\widehat{\mathsf{P}}^0, \widehat{\mathsf{P}}^1$ in the cases $t = 0$ and $t = 1$.

DEFINITION 6.7 (Half-interpolate). Let $\mathrm{X}^0 = (\mathsf{z}_1^0, \dots, \mathsf{z}_{2R}^0)$ be a $2R$-tuple in $\mathcal{O}_R$, and let $\mathrm{X}^1 = (\mathsf{z}_1^1, \dots, \mathsf{z}_{2N}^1)$ in $\mathcal{O}_R$ be the image of X^0 by the transportation map T_R. We introduce X^{h} as

$$\mathrm{X}^{\mathsf{h}} := \frac{\mathrm{Id} + \mathsf{T}_R}{2}(\mathrm{X}^0) := \left(\frac{\mathsf{z}_1^0 + \mathsf{z}_1^1}{2}, \dots, \frac{\mathsf{z}_{2R}^1 + \mathsf{z}_{2R}^1}{2}\right). \tag{6.5}$$

LEMMA 6.8 (Effect of the transport on the discrepancy). *Let* $R > 0$*, let* $\mathcal{C}^0, \mathcal{C}^1$ *be two point configurations with* $2R$ *points in* Λ_R*, and let* $\mathcal{C}^{\mathsf{h}}$ *be the half-interpolate of* $\mathcal{C}^0$ *and* $\mathcal{C}^1$*. For any* r *in* $[-R, R]$*, we have*

$$\left|\mathrm{Discr}_{[-R,r]}(\mathcal{C}^{\mathsf{h}})\right| \le \max\left(\left|\mathrm{Discr}_{[-R,r]}(\mathcal{C}^0)\right|, \left|\mathrm{Discr}_{[-R,r]}(\mathcal{C}^1)\right|\right) \tag{6.6}$$

PROOF. The construction of $\mathcal{C}^{\mathsf{h}}$ clearly implies, for any $r \in [-R, R]$,

$$\min\big(\big|\mathcal{C}^0_{[-R,r]}\big|, \big|\mathcal{C}^1_{[-R,r]}\big|\big) \leq \big|\mathcal{C}^{1/2}_{[-R,r]}\big| \leq \max\big(\big|\mathcal{C}^0_{[-R,r]}\big|, \big|\mathcal{C}^1_{[-R,r]}\big|\big),$$

and the results follow from the definition of Discr. □

In particular, since the configurations that we construct all satisfy (5.11), this is also the case for all the configurations obtained by interpolation.

LEMMA 6.9 (Displacement convexity of the entropy). *The map* $t \mapsto \mathrm{Ent}[\widehat{\mathsf{P}}^t|\mathbf{L}_R]$ *is convex on* $[0, 1]$.

PROOF. This lemma is a well-known "displacement convexity" result, which was proven in the pioneering paper [17, theorem 2.2]. □

Remark 6.10. This is the only moment where we use the fact that T_R is the *optimal* transport map for the quadratic cost.

In particular, we obtain

$$\mathrm{Ent}[\widehat{\mathsf{P}}^{1/2}|\mathbf{L}_R] \leq \frac{1}{2}\big(\mathrm{Ent}\big[\widehat{\mathsf{P}}^1|\mathbf{L}_R\big] + \mathrm{Ent}\big[\widehat{\mathsf{P}}^2|\mathbf{L}_R\big]\big). \tag{6.7}$$

If $\mathsf{P}^0_R, \mathsf{P}^{\mathsf{h}}_R, \mathsf{P}^1_R$ are the push-forward of $\widehat{\mathsf{P}}^0, \widehat{\mathsf{P}}^{\frac{1}{2}}, \widehat{\mathsf{P}}^1$ by π_R^{-1}, using Lemma 6.4 we see that

$$\mathrm{Ent}\big[\mathsf{P}^{\mathsf{h}}_R\big|\Pi_{\Lambda_R}\big] \leq \frac{1}{2}\big(\mathrm{Ent}\big[\mathsf{P}^0_R\big|\Pi_{\Lambda_R}\big] + \mathrm{Ent}\big[\mathsf{P}^1_R\big|\Pi_{\Lambda_R}\big]\big). \tag{6.8}$$

Since $x \mapsto x\ln(x)$ is a strictly convex function on $(0, +\infty)$, the results of [17] imply that the inequality in (6.7), and thus in (6.8), is strict unless $\mathsf{P}^0_R = \mathsf{P}^1_R$. Inspecting the proof could possibly yield some quantitative bound, but here we rely instead on the energy term to get a tractable convexity inequality along the displacement interpolation.

6.3 Convexity Inequality for the Energy

Although [17, sec. 3] deals with the displacement convexity of energies similar to ours, the setting is different and we cannot directly apply these results. Moreover, we crucially need a quantitative convexity estimate, which is the aim of this section.

We start by stating an elementary inequality:

LEMMA 6.11 (Quantitative convexity for $-\log$). *Let* $x, y > 0$. *We have*

$$-\log\left(\frac{x+y}{2}\right) \leq \frac{-\log x - \log y}{2} - \frac{(x-y)^2}{8(x^2+y^2)}. \tag{6.9}$$

PROOF. The function $f : x \mapsto -\log(x)$ is convex on $(0, +\infty)$, since its second derivative is given by $f''(x) = \frac{1}{x^2} > 0$. The result follows by applying the following elementary inequality

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2} - \frac{1}{8}(x-y)^2 \inf_{[x,y]} f''(c). \qquad \square$$

As an immediate consequence, if $(\mathsf{z}^0_1, \dots, \mathsf{z}^0_{2R})$ and $(\mathsf{z}^1_1, \dots, \mathsf{z}^1_{2R})$ are two $2R$-tuples of points in $\mathcal{O}_R$, we have

$$\begin{aligned} &\sum_{1 \le i < j \le 2R} \frac{-\log(\mathsf{z}^0_j - \mathsf{z}^0_i) - \log(\mathsf{z}^1_j - \mathsf{z}^1_i)}{2} + \log\left(\frac{\mathsf{z}^0_j + \mathsf{z}^1_j}{2} - \frac{\mathsf{z}^0_i + \mathsf{z}^1_i}{2}\right) \\ &\quad \ge \frac{1}{8} \sum_{1 \le i < j \le 2R} \frac{\left((\mathsf{z}^0_j - \mathsf{z}^0_i) - (\mathsf{z}^1_j - \mathsf{z}^1_i)\right)^2}{(\mathsf{z}^0_j - \mathsf{z}^0_i)^2 + (\mathsf{z}^1_j - \mathsf{z}^1_i)^2}. \end{aligned} \tag{6.10}$$

Let $\mathrm{X} = (\mathsf{z}_1, \dots, \mathsf{z}_{2R})$ be a $2R$-tuple in $\mathcal{O}_R$, and let $\mathcal{C} = \pi_R^{-1}(\mathrm{X})$. We define the gaps of X, counted from the left, as

$$\mathrm{G}_i := \mathsf{z}_{i+1} - \mathsf{z}_i \quad (1 \le i \le 2R - 1). \tag{6.11}$$

We denote by $\mathrm{G}_i(\mathrm{X})$ the i^{th} such gap of a given $\mathrm{X} \in \mathcal{O}_R$. Moreover, we introduce the following notation for the interaction of X:

$$\mathrm{Int}[\mathrm{X}] := \sum_{1 \le i < j \le 2R} -\log|\mathsf{z}_i - \mathsf{z}_j| = \iint_{x<y} -\log|x-y| \, d\mathcal{C}(x) \, d\mathcal{C}(y). \tag{6.12}$$

PROPOSITION 6.12 (Quantitative convexity along displacement). *Let* $\mathrm{X}^0, \mathrm{X}^1, \mathrm{X}^{\mathsf{h}}$ *be as in Definition* 6.7. *We have*

$$\mathrm{Int}[\mathrm{X}^{\mathsf{h}}] \le \frac{\mathrm{Int}[\mathrm{X}^0] + \mathrm{Int}[\mathrm{X}^1]}{2} - \frac{1}{8} \sum_{i=1}^{2R-1} \frac{|\mathrm{G}_i(\mathrm{X}^0) - \mathrm{G}_i(\mathrm{X}^1)|^2}{(\mathrm{G}_i(\mathrm{X}^0))^2 + (\mathrm{G}_i(\mathrm{X}^1))^2}. \tag{6.13}$$

PROOF. This is a straightforward consequence of (6.10) by writing

$$\sum_{1 \le i < j \le 2R} \frac{\left((\mathsf{z}^0_j - \mathsf{z}^0_i) - (\mathsf{z}^1_j - \mathsf{z}^1_i)\right)^2}{(\mathsf{z}^0_j - \mathsf{z}^0_i)^2 + (\mathsf{z}^1_j - \mathsf{z}^1_i)^2} \ge \sum_{i=1}^{2R-1} \frac{\left((\mathsf{z}^0_{i+1} - \mathsf{z}^0_i) - (\mathsf{z}^1_{i+1} - \mathsf{z}^1_i)\right)^2}{(\mathsf{z}^0_{i+1} - \mathsf{z}^0_i)^2 + (\mathsf{z}^1_{i+1} - \mathsf{z}^1_i)^2},$$

and using the notation (6.11). □

For $R > 0$, we introduce the "background potential" as

$$\mathsf{V}_R(t) := \int_{-R}^{R} \log|t - s| \, ds. \tag{6.14}$$

An explicit computation yields

$$\begin{aligned} \mathsf{V}_R(t) &= \int_{-R}^{t} \log(t-s) \, ds + \int_{t}^{R} \log(s-t) \, ds \\ &= [(R+t)\log(R+t) - (R+t) + (R-t)\log(R-t) - (R-t)]. \end{aligned}$$

and thus

$$\mathsf{V}'_R(t) = \log(R+t) + \log(R-t), \quad \mathsf{V}''_R(t) = \frac{1}{R+t} + \frac{1}{R-t}. \tag{6.15}$$

As a straightforward consequence, we obtain the following:

LEMMA 6.13. V_R *is a convex function on* Λ_R.

LEMMA 6.14 (Intrinsic energy, logarithmic interaction, background field). *Let $\mathcal{C}$ be in* Conf_R*, and let* $\mathrm{X} = \pi_R(\mathcal{C})$*. We have*

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R) = 2\mathrm{Int}[\mathrm{X}] + 2\sum_{i=1}^{2R} \mathsf{V}_R(\mathsf{z}_i) + \mathrm{const}_R, \tag{6.16}$$

where Int *is as in* (6.12)*,* V_R *is as above, and* const_R *is a constant depending only on* R.

PROOF. It follows from the definition (2.6) of $\mathrm{W}^{\mathrm{int}}$, by expanding the quadratic term $(d\mathcal{C} - dx)(d\mathcal{C} - dy)$,

$$\begin{aligned}
&\mathrm{W}^{\mathrm{int}}(\mathcal{C}; \Lambda_R) \\
&\quad = \iint_{\Lambda_R \times \Lambda_R \setminus \diamond} -\log|x-y|(d\mathcal{C} - dx)(d\mathcal{C} - dy) \\
&\quad = 2\iint_{x<y} -\log|x-y| d\mathcal{C}(x) d\mathcal{C}(y) + 2\iint_{\Lambda_R \times \Lambda_R \setminus \diamond} \log|x-y| d\mathcal{C}(x) dy \\
&\qquad + \iint_{\Lambda_R \times \Lambda_R \setminus \diamond} -\log|x-y| dx dy,
\end{aligned}$$

and using definition (6.14). $\square$

DEFINITION 6.15 (Gain and background field contribution). Assume $\mathcal{C}^0, \mathcal{C}^1$ is in Conf_R, let $\mathrm{X}^0, \mathrm{X}^1$ be their image by π_R, and let X^{h} be as in (6.5).

- We define the "gain" term $\mathsf{Gain}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R)$ as

$$\mathsf{Gain}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R) := \sum_{i=1}^{2R-1} \frac{|\mathrm{G}_i(\mathrm{X}^0) - \mathrm{G}_i(\mathrm{X}^1)|^2}{(\mathrm{G}_i(\mathrm{X}^0))^2 + (\mathrm{G}_i(\mathrm{X}^1))^2}. \tag{6.17}$$

- We define the "background field contribution" term $\mathsf{BF}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R)$ as

$$\mathsf{BF}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R) := 2\sum_{i=1}^{2R} \left(\mathsf{V}_R(\mathsf{z}_i^{\mathsf{h}}) - \frac{1}{2}\big(\mathsf{V}_R(\mathsf{z}_i^0) + \mathsf{V}_R(\mathsf{z}_i^1)\big) \right). \tag{6.18}$$

PROPOSITION 6.16 (Convexity of the interaction energy). *Let* $\mathcal{C}^0, \mathcal{C}^1$ *be in* Conf_R*; we have*

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathsf{h}}; \Lambda_R) \le \frac{1}{2}\big(\mathrm{W}^{\mathrm{int}}(\mathcal{C}^0; \Lambda_R) + \mathrm{W}^{\mathrm{int}}(\mathcal{C}^1; \Lambda_R)\big) - \frac{1}{4}\mathsf{Gain}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R). \tag{6.19}$$

PROOF. We combine (6.13) and (6.16), and we use the notation introduced in (6.17) and (6.18). We obtain

$$\begin{aligned}
\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathsf{h}}; \Lambda_R) \le{} & \frac{1}{2}\big(\mathrm{W}^{\mathrm{int}}(\mathcal{C}^0; \Lambda_R) + \mathrm{W}^{\mathrm{int}}(\mathcal{C}^1; \Lambda_R)\big) \\
& - \frac{1}{4}\mathsf{Gain}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R) + \mathsf{BF}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R).
\end{aligned}$$

Since V_R has been observed to be a convex function on Λ_R (see Lemma 6.13), from the definition of BF in (6.18) we readily see that $\mathsf{BF}(\mathcal{C}^0, \mathcal{C}^1; \Lambda_R) \leq 0$, so in particular (6.19) holds.[3] □

7 The Interpolate Process

Let $\mathsf{P}^0, \mathsf{P}^1$ be two minimizers of $\mathcal{F}_\beta$, and assume that $\mathsf{P}^0 \neq \mathsf{P}^1$. Let $\mathfrak{g}$ be the "gain" given by Proposition 4.7, i.e., $1/R$ times the right-hand side of (4.14).

DEFINITION 7.1 (Interpolate process). For any $s > 0$ and for R large enough (depending on $s, \mathsf{P}^0, \mathsf{P}^1$), we apply the "large box approximation" of Proposition 5.7 to $\mathsf{P}^0, \mathsf{P}^1$ with $\varepsilon = \frac{\mathfrak{g}}{100}$. We let $\widetilde{\mathsf{P}}^0_R, \widetilde{\mathsf{P}}^1_R$ be the processes on Λ_R obtained this way. Further, we let $\hat{\mathsf{P}}^0_R = (\pi_R)_* \widetilde{\mathsf{P}}^0_R$, $\hat{\mathsf{P}}^1_R = (\pi_R)_* \widetilde{\mathsf{P}}^1_R$ be the correponding measures on $\mathcal{O}_R$ obtained by push-orward via the labeling map π_R. We let $\hat{Q}_R$ be an optimal coupling of $\hat{\mathsf{P}}^0_R$ and $\hat{\mathsf{P}}^1_R$ given by the optimal transport map T_R, and let $\widehat{\mathsf{P}^{\mathsf{h}}}_R$ be the half-interpolate measure on $\mathcal{O}_R$ obtained from displacement interpolation. Finally, we let $\mathsf{P}^{\mathsf{h}}_R = (\pi_R^{-1})_* \widehat{\mathsf{P}^{\mathsf{h}}}_R$ be the corresponding half-interpolate process on Λ_R and let $Q_R = (\pi_R^{-1} \times \pi_R^{-1})_* \hat{Q}_R$ be the corresponding coupling of $\widetilde{\mathsf{P}}^0_R, \widetilde{\mathsf{P}}^1_R$.

Next, we show that the energy gain is proportional to the size of the segment.

LEMMA 7.2 (The energy gain). *Taking R large enough, we have*

$$\mathbb{E}_{\mathsf{Q}_R}[\mathsf{Gain}(\cdot, \cdot; \Lambda_R)] \geq \frac{1}{2}\mathfrak{g}R. \tag{7.1}$$

PROOF. Let $\mathcal{C}^{0,\mathrm{scr}}, \mathcal{C}^{1,\mathrm{scr}}$ be in the support of $\widetilde{\mathsf{P}}^0_R, \widetilde{\mathsf{P}}^1_R$. Let S^0, S^1 be the quantities, as in (5.9), relating the two ways of enumerating points, i.e., such that

$$x_0(\mathcal{C}^{0,\mathrm{scr}}) = \mathsf{z}^0_{R+S^0}, \quad x_0(\mathcal{C}^{1,\mathrm{scr}}) = \mathsf{z}^1_{R+S^1},$$

which means $\mathcal{C}^{0,\mathrm{scr}}$ (resp., $\mathcal{C}^{1,\mathrm{scr}}$) has $R + S^0$ (resp., $R + S^1$) points in $[-R, 0]$.

From (5.10) we know that (up to choosing s small enough with respect to M) S^0, S^1 are bounded by $\frac{1}{2}R^{1/2}$, so $S = S^1 - S^0$ is bounded by $R^{1/2}$. We may write, using the definition (6.17) and switching indices:

$$\mathsf{Gain}(\mathcal{C}^{0,\mathrm{scr}}, \mathcal{C}^{1,\mathrm{scr}}; \Lambda_R) \geq \sum_{i=-R/2}^{R/2} \frac{\left|\Gamma_i(\mathcal{C}^{0,\mathrm{scr}}) - \Gamma_{i+S}(\mathcal{C}^{1,\mathrm{scr}})\right|^2}{(\Gamma_i(\mathcal{C}^{0,\mathrm{scr}}))^2 + (\Gamma_{i+S}(\mathcal{C}^{1,\mathrm{scr}}))^2}.$$

Note that the right-hand side only depends on the restrictions of $\mathcal{C}^{0,\mathrm{scr}}, \mathcal{C}^{1,\mathrm{scr}}$ to $\mathsf{Old} = \Lambda_{R'}$. Let us denote by $\bar{\mathsf{Q}}_R$ the image of the coupling Q_R under restriction to Old. Since the screening procedure does not change the configurations in Old by

[3] In fact, the contribution of the background potential can be shown to be negligible with respect to R.

construction, $\bar{\mathsf{Q}}_R$ is a coupling of the restrictions to Old of the stationary processes P^0 and P^1 conditioned on an event of probability $\geq 1 - \frac{\mathfrak{g}}{100}$.

$$\mathbb{E}_{\mathsf{Q}_R}[\mathsf{Gain}(\cdot,\cdot;\Lambda_R)] \geq \mathbb{E}\left[\sum_{i=-R/2}^{R/2} \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_{i+S}(\mathcal{C}^1)\right|^2}{(\Gamma_i(\mathcal{C}^0))^2 + (\Gamma_{i+S}(\mathcal{C}^1))^2}\right],$$

where $(\mathcal{C}^0, \mathcal{C}^1)$ is distributed according to $\bar{\mathsf{Q}}_R$. In view of Corollary 4.8, the last expression is bounded below by $\frac{1}{2}\mathfrak{g}R$, which yields the result. □

PROPOSITION 7.3 (Energy of the interpolate process). *For s small enough (depending on $\mathsf{P}^0, \mathsf{P}^1$), for R large enough (depending on $\mathsf{P}^0, \mathsf{P}^1, s$) we have*

$$\frac{1}{|\Lambda_R|}\mathbb{E}_{\mathsf{P}^{\mathsf{h}}_R}\big[\mathrm{W}^{\mathrm{int}}(\mathcal{C}^{\mathsf{h}};\Lambda_R)\big] \leq \frac{1}{2}\big(\mathcal{W}(\mathsf{P}^0) + \mathcal{W}(\mathsf{P}^1)\big) - \frac{\mathfrak{g}}{4}. \tag{7.2}$$

PROOF. We combine

- (6.19), from the transportation argument, that expresses $\mathbb{E}_{\mathsf{P}^{\mathsf{h}}_R}[\mathrm{W}^{\mathrm{int}}]$ in terms of $\mathbb{E}_{\mathsf{P}^0_R}[\mathrm{W}^{\mathrm{int}}]$, $\mathbb{E}_{\mathsf{P}^1_R}[\mathrm{W}^{\mathrm{int}}]$ and the "gain" term;
- Lemma 7.2, saying that the "gain" is proportional to the volume;
- (5.7), which controls the energies $\mathbb{E}_{\mathsf{P}^0_R}[\mathrm{W}^{\mathrm{int}}]$, $\mathbb{E}_{\mathsf{P}^1_R}[\mathrm{W}^{\mathrm{int}}]$ in terms of $\mathcal{W}(\mathsf{P}^0)$, $\mathcal{W}(\mathsf{P}^1)$. □

8 Conclusion: Proof of Theorem 1.1

In this section we conclude the proof of Theorem 1.1 by contradiction. Assuming $\mathcal{F}_\beta$ has two distinct minimizers, we contruct a stationary point process with strictly smaller free energy. To this end, we start from the interpolate process constructed in the last section, which is supported in a large segment $[-R, R]$, into a stationary point process on the whole real line. We proceed in two steps: first we paste independent copies of the interpolate process, and then we average over translations in the original segment. This construction was used in [14] for similar purposes.

Let $\mathsf{P}^0, \mathsf{P}^1$ be two stationary point processes such that $\mathcal{F}_\beta(\mathsf{P}^0), \mathcal{F}_\beta(\mathsf{P}^1) < \infty$, and let $\mathsf{P}^{\mathsf{h}}_R$ be the interpolate process on $\Lambda_R = [-R, R]$ contructed in Section 7; see Definition 7.1. Let $\{K_i := \Lambda_R - 2Ri\}_{i\in\mathbb{Z}}$ be a tiling of $\mathbb{R}$ by translates of the interval Λ_R, and let $\{\mathcal{C}^{\mathsf{h}}_i\}_{i\in\mathbb{Z}}$ be independent random variables identically distributed according to $\mathsf{P}^{\mathsf{h}}_R$. We let Paste_R be the map

$$\mathrm{Paste}_R := \{\mathcal{C}_i\}_{i\in\mathbb{Z}} \mapsto \sum_{i\in\mathbb{Z}}(\mathcal{C}_i - 2iR)$$

that creates a configuration on $\mathbb{R}$ from a collection of configurations on Λ_R by copying one on each interval K_i. We let $\mathsf{P}^{\mathrm{tot}}$ be the push-forward of the law of $\{\mathcal{C}^{\mathsf{h}}_i\}_{i\in\mathbb{Z}}$ by Paste_R. We let P^{av} be the law of the point process defined by averaging $\mathsf{P}^{\mathrm{tot}}$ over translations in Λ_R. More precisely, we let P^{av} be the law of the point process defined by duality as follows: for any bounded measurable test function F,

$$\mathbb{E}_{\mathsf{P}^{\mathrm{av}}}[F] := \mathbb{E}_{\mathsf{P}^{\mathrm{tot}}}\left[\frac{1}{|\Lambda_R|}\int_{-R}^{R} F(\mathcal{C}-t)\,dt\right]. \tag{8.1}$$

Now, by construction, P^{av} is a stationary process of intensity 1, and there are always $2L \pm 4R$ points in any interval of length $2L$.

PROPOSITION 8.1 (The new process is a better candidate). *The process* P^{av} *satisfies*

$$\mathcal{F}_\beta(\mathsf{P}^{\mathrm{av}}) < \frac{1}{2}\big(\mathcal{F}_\beta(\mathsf{P}^0) + \mathcal{F}_\beta(\mathsf{P}^1)\big). \tag{8.2}$$

With this we can conclude the proof of our main result.

PROOF OF THEOREM 1.1. Assume now that $\mathsf{P}^0, \mathsf{P}^1$ are two distinct minimizers of $\mathcal{F}_\beta$. Proposition 8.1 gives that we have constructed a stationary point process P^{av} that satisfies

$$\mathcal{F}_\beta(\mathsf{P}^{\mathrm{av}}) < \frac{1}{2}\big(\mathcal{F}_\beta(\mathsf{P}^0) + \mathcal{F}_\beta(\mathsf{P}^1)\big) = \min \mathcal{F}_\beta,$$

which is absurd. Hence, the minimizer of $\mathcal{F}_\beta$ is unique, which proves the main theorem. $\square$

The rest of this section is devoted to proving Proposition 8.1.

PROOF OF PROPOSITION 8.1. Recall that $\mathcal{F}_\beta(\mathsf{P}^{\mathrm{av}}) := \beta\mathcal{W}(\mathsf{P}^{\mathrm{av}}) + \mathcal{E}(\mathsf{P}^{\mathrm{av}})$. The proof will be carried out in several steps in which we evalutate to the terms separately.

Step 1. THE SPECIFIC RELATIVE ENTROPY.

CLAIM 8.2 (The specific relative entropy of P^{av}). *We have*

$$\mathcal{E}(\mathsf{P}^{\mathrm{av}}) \le \frac{1}{2}\big(\mathcal{E}(\mathsf{P}^0) + \mathcal{E}(\mathsf{P}^1)\big) + \frac{\mathfrak{g}}{50}.$$

PROOF. By construction, the process $\mathsf{P}^{\mathrm{tot}}$ is made of independent copies of P^{h} on Λ_R, so we have

$$\mathcal{E}(\mathsf{P}^{\mathrm{av}}) = \lim_{M\to\infty} \frac{1}{|\Lambda_M|}\mathrm{Ent}\big[\mathsf{P}^{\mathrm{av}}_{\Lambda_M} \big| \Pi_{\Lambda_M}\big] = \frac{1}{|\Lambda_R|}\mathrm{Ent}\big[\mathsf{P}^{\mathrm{h}}_R \big| \Pi_{\Lambda_R}\big].$$

Using (6.8) and (5.8) (with $\varepsilon = \frac{\mathfrak{g}}{100}$) yields the claim. $\square$

Step 2. THE INTRINSIC ENERGY.

CLAIM 8.3 (The intrinsic energy of P^{av}). *We have*

$$\mathcal{W}^{\mathrm{int}}(\mathsf{P}^{\mathrm{av}}) \le \frac{1}{2}\big(\mathcal{W}(\mathsf{P}^0) + \mathcal{W}(\mathsf{P}^1)\big) - \frac{\mathfrak{g}}{10}.$$

The proof of Claim 8.3 will be carried out in Steps 2a–2g below.

Step 2a. DECOMPOSING THE INTERACTION ENERGY. Since $\mathcal{W}^{\mathrm{int}}$ is defined in (2.15) as a lim inf, it is enough to show that

$$\liminf_{M\to\infty} \frac{1}{|\Lambda_{MR}|} \mathbb{E}_{\mathsf{P}^{\mathsf{h}}_R}\big[\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_{MR})\big] \le \frac{1}{2}\big(\mathcal{W}(\mathsf{P}^0)+\mathcal{W}(\mathsf{P}^1)\big) - \frac{\mathfrak{g}}{10}.$$

By the definition (8.1), P^{av} is a uniform mixture of processes that we will denote by $\mathsf{P}^{\mathrm{tot},z}$ for $z \in \Lambda_R$, where $\mathsf{P}^{\mathrm{tot},z}$ is the process $\mathsf{P}^{\mathrm{tot}}$ translated by z. For $z \in [-R, R]$ and $M \ge 10$ fixed, a configuration of $\mathsf{P}^{\mathrm{tot},z}$ in Λ_{MR} consists of the following:

- $M-2$ "full" configurations $\mathcal{C}_1, \dots, \mathcal{C}_{M-2}$ obtained from independent copies of P^{h}, and supported in the intervals $\Lambda_R - z - 2Ri$ for $i \approx -M/2 \dots M/2$,
- Two "partial" configurations $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_{\mathrm{right}}$, where $\mathcal{C}_{\mathrm{left}}$ is supported in $[-MR, -MR+R-z]$ and $\mathcal{C}_{\mathrm{right}}$ in $[MR-R-z, MR]$.

When we compute the interaction energy $\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_{MR})$, we obtain

$$\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_{MR}) = A+B+C+E+F+G: \tag{8.3}$$

- A is the sum of the interactions of a configuration $\mathcal{C}_i$ with itself, for $i = 1, \dots, M-2$;
- B is the interaction between $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_{\mathrm{right}}$;
- C is the interaction between $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_1$ and between $\mathcal{C}_{\mathrm{right}}$ and $\mathcal{C}_{M-2}$;
- D is the interaction of $\mathcal{C}_{\mathrm{left}}$ with itself and of $\mathcal{C}_{\mathrm{right}}$ with itself;
- E is the sum of the interactions between $\mathcal{C}_i$ and $\mathcal{C}_{i+k}$ for $i = 1, \dots, M-2$ and $2 \le k \le M-2-i$, i.e., between *nonneighboring* "full" configurations;
- F is all of the interactions between two neighboring "full" configurations $\mathcal{C}_i, \mathcal{C}_{i+1}$;
- G is the sum of the interactions between $\mathcal{C}_{\mathrm{left}}$ and all nonneighboring "full" configurations ($\mathcal{C}_i$ for $i = 2, \dots M-2$), and the interactions between $\mathcal{C}_{\mathrm{right}}$ and all nonneighboring "full" configurations ($\mathcal{C}_i$ for $i = 1, \dots, M-3$).

Step 2b. THE TERM A (SELF-INTERACTION OF FULL CONFIGURATIONS). Taking the expectation under $\mathsf{P}^{\mathrm{tot},z}$, in view of (7.2), we have

$$\begin{aligned}\frac{1}{M|\Lambda_R|}\mathbb{E}_{\mathsf{P}^{\mathrm{tot},z}}(A) &= \frac{M-2}{M}\frac{1}{|\Lambda_R|}\mathbb{E}_{\mathsf{P}^{\mathsf{h}}_R}[\mathrm{W}^{\mathrm{int}}(\mathcal{C};\Lambda_R)] \\ &\le \frac{1}{2}\big(\mathcal{W}(\mathsf{P}^0)+\mathcal{W}(\mathsf{P}^1)\big) - \frac{\mathfrak{g}}{4} + o_M(1).\end{aligned}$$

It remains to study all the other terms and to show that they are either negligible with respect to M as $M \to \infty$ or only yield a perturbation of order MR, that can be made arbitrarily small through the choice of s.

Step 2c. THE TERMS B, C, D. We may already observe that the interaction between $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_{\mathrm{right}}$ is bounded by $O(R^2 \log M)$, and is thus $o(M)$. So are the interactions between $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_1$ or between $\mathcal{C}_{\mathrm{right}}$ and $\mathcal{C}_{M-2}$. We may also bound the self-interaction of $\mathcal{C}_{\text{left}}, \mathcal{C}_{\text{right}}$ by a quantity independent of M. We thus have

$$\frac{1}{M|\Lambda_R|}\mathbb{E}_{\mathsf{P}^{\mathrm{tot},z}}(B + C + D) = o_M(1).$$

Step 2d. A PRIORI BOUND ON FLUCTUATIONS. To control the other "pairwise" interactions we rely on the following bound expressing fluctuations in terms of discrepancies.

LEMMA 8.4 (Controlling fluctuations via discrepancies). *Let* $[a, b]$ *be an interval of* $\mathbb{R}$*, let* g *be a* C^1 *function on* $[a, b]$*, and let* $\mathcal{C}$ *be a point configuration on* $[a, b]$*. We have*

$$\begin{aligned}\int_a^b g(x)(d\mathcal{C} - dx) \\ \preceq \sum_{k=a}^{b} [\|g'\|_\infty(|\mathrm{Discr}_{[a,k]}(\mathcal{C})| + |\mathrm{Discr}_{[k,k+1]}(\mathcal{C})| + 1) \\ + \|g\|_\infty|\mathrm{Discr}_{[a,b]}(\mathcal{C})|].\end{aligned} \tag{8.4}$$

In particular, if $[a, b] = \Lambda_R$ and $\mathcal{C}$ is a configuration with $2R$ points in Λ_R, the last term in the right-hand side of (8.4) vanishes.

PROOF. This follows by splitting $[a, b]$ into intervals of length 1, using a Taylor expansion of g on each interval and using a summation by parts. We refer to [15, prop. 1.6] for details in a similar situation. □

We use the notation $\widetilde{D}^{\mathrm{Left}}_{R,k}(\mathcal{C})$ for the summand in (8.4), i.e.,

$$\widetilde{D}^{\mathrm{Left}}_{R,k}(\mathcal{C}) := |\mathrm{Discr}_{[-R,k]}(\mathcal{C})| + |\mathrm{Discr}_{[k,k+1]}(\mathcal{C})| + 1. \tag{8.5}$$

Since the double integral defining $\mathrm{W}^{\mathrm{int}}$ (see (2.6)) involves the "fluctuation" terms $(d\mathcal{C} - dx)(d\mathcal{C} - dy)$, using (8.4) we can derive the following control on the interaction between two configurations living in two nonneighboring copies of Λ_R:

LEMMA 8.5. *Let* $\mathcal{C}^a, \mathcal{C}^b$ *be two configurations with* $2R$ *points supported on the intervals* $\Lambda_R^a := \Lambda_R - 2Ra$, $\Lambda_R^b := \Lambda_R - 2Rb$ *respectively, with* $|a - b| \geq 2$*. Then*

$$\begin{aligned}\iint_{\Lambda_R^a \times \Lambda_R^b} -\log|x - y|(d\mathcal{C}^a(x) - dx)(d\mathcal{C}^b(y) - dy) \\ \preceq \frac{1}{|a-b|^2 R^2} \sum_{k=0}^{2R}\sum_{j=0}^{2R} \widetilde{D}^{\mathrm{Left}}_{R,k}(\mathcal{C}^a + 2Ra)\widetilde{D}^{\mathrm{Left}}_{R,j}(\mathcal{C}^b + 2Rb),\end{aligned} \tag{8.6}$$

where $\widetilde{D}^{\mathrm{Left}}_{R,k}(\mathcal{C})$ *is as in* (8.5).

PROOF. To show (8.6), we apply 8.4 twice (once for each variable) to $h(x, y) = -\log|x - y|$, bounding the second derivative of h, for x in Λ_R^a and y in Λ_R^b, by $\frac{1}{|a-b|^2 R^2}$. □

Step 2e. THE TERM E. Combining the result of Lemma 8.5 with the discrepancy bounds (5.11) (which are still valid for the interpolate configurations, as observed in Lemma 6.8), we may bound the interaction between $\mathcal{C}_i$ and $\mathcal{C}_{i+k}$, for $k \geq 2$, by

$$\iint_{\Lambda_R^i \times \Lambda_R^{i+k}} -\log|x - y|(d\mathcal{C}_i(x) - dx)(d\mathcal{C}_{i+k}(y) - dy) \preceq \frac{1}{k^2 R^2}(s^2 R^{3/2})^2.$$

For a given i, we thus have

$$\sum_{k\geq 2} \iint_{\Lambda_R^i \times \Lambda_R^{i+k}} -\log|x - y|(d\mathcal{C}_i(x) - dx)(d\mathcal{C}_{i+k}(y) - dy) \preceq s^4 R,$$

and summing again over $i = 1, \dots, M - 2$, we bound the term E in (8.3) by

$$E \preceq s^4 MR,$$

which is a contribution of order MR that can be made arbitrarily small by taking s small.

The remaining interactions, between neighbors $\mathcal{C}_i, \mathcal{C}_{i+1}$ and between $\mathcal{C}_{\text{left}}, \mathcal{C}_{\text{right}}$ and the $\mathcal{C}_i$'s, also yield arbitrarily small contributions. The argument is similar: we use Lemma 8.4 and the discrepancy estimates, with two small modifications.

Step 2f. THE TERM F. To treat neighbors, here is a sketch of the argument: take two configurations $\mathcal{C}^l$ (on the left) and $\mathcal{C}^r$ (on the right) living in $[-R, 0]$ and $[0, R]$ respectively. In view of Lemma 8.4, we write their interaction as

$$\begin{aligned} \text{Interaction} &= \iint_{[-R,0]\times[0,R]} -\log|x - y|(d\mathcal{C}^l - dx)(d\mathcal{C}^r - dy) \\ &\preceq \sum_{i,j=1}^{R} \frac{1}{(i + j)^2} |\mathrm{Discr}_{[0,-i]}(\mathcal{C}^l)| \cdot |\mathrm{Discr}_{[0,j]}(\mathcal{C}^r)|, \end{aligned} \tag{8.7}$$

where we have only kept the leading-order discrepancy in (8.4). The term $\frac{1}{(i+j)^2}$ comes from the fact that the second derivative $\partial_{xy} - \log|x - y|$ is controlled, for x in $[-i - 1, -i]$ and y in $[j, j + 1]$, by $\frac{1}{(i+j)^2}$. We are looking for a bound of the type Interaction $\preceq o_s(1)R$, since then we have $O(M)$ pairs of neighbors, each yielding a contribution $o_s(1)R$, so the term F would be of order $o_s(1)MR$, which is enough for our purposes.

Note that, to simplify, when writing (8.7) we have assumed that the configurations were separated by a distance 1 (there is no $i, j = 0$ term); in reality the contribution of the terms at distance ≤ 1 is bounded by $O(1)$ and we can forget about them. The key point is to use the fact that the second moment of the discrepancy in a segment is very small compared to the size of the segment, as expressed by Lemma 4.1. In particular, we may write

$$|\mathrm{Discr}_{[0,-i]}(\mathcal{C}^l)| \cdot |\mathrm{Discr}_{[0,j]}(\mathcal{C}^r)| \ll \sqrt{i}\sqrt{j}. \tag{8.8}$$

Then, we estimate the double sum of (8.7) as follows:

$$\begin{aligned}
&\sum_{i,j=1}^{R} \frac{1}{(i+j)^2} |\mathrm{Discr}_{[0,-i]}(\mathcal{C}^l)| \cdot |\mathrm{Discr}_{[0,j]}(\mathcal{C}^r)| \\
&\quad \preceq \sum_{i=1}^{R}\sum_{j=i}^{R} \frac{1}{(i+j)^2} |\mathrm{Discr}_{[0,-i]}(\mathcal{C}^l)| \cdot |\mathrm{Discr}_{[0,j]}(\mathcal{C}^r)| \\
&\quad \ll \sum_{i=1}^{R}\sum_{j=i}^{R} \frac{1}{(i+j)^2}\sqrt{i}\sqrt{j} \preceq \sum_{i=1}^{R}\sum_{j=i}^{2i} \frac{i}{i^2} + \sum_{i=1}^{R}\sqrt{i}\sum_{j=2i}^{R}\frac{1}{j^2}\sqrt{j},
\end{aligned}$$

and thus a direct computation yields Interaction $\ll R$.

Step 2g. THE TERM G. If we apply Lemma 8.4 to $\mathcal{C}_{\mathrm{left}}$ (or $\mathcal{C}_{\mathrm{right}}$), the boundary term in (8.4) is not 0, but bounded by $2R\|g\|_\infty$. In particular, when estimating the interaction between $\mathcal{C}_{\mathrm{left}}$ and $\mathcal{C}_i$ for $i \geq 2$, we obtain a term similar to (8.6) above, plus a boundary term:

$$\begin{aligned}
&\iint -\log|x-y|(d\mathcal{C}_{\mathrm{left}}(x)-dx)(d\mathcal{C}_i(y)-dy) \\
&\quad \preceq \frac{1}{i^2R^2}\left(\sum_{k=0}^{2R}\widetilde{D}^{\mathrm{Left}}_{R,k}(\mathcal{C}_{\mathrm{left}})\right)\left(\sum_{j=0}^{2R}\widetilde{D}^{\mathrm{Left}}_{R,j}(\mathcal{C}_i)\right) + \frac{1}{iR}2R\sum_{j=0}^{2R}\widetilde{D}^{\mathrm{Left}}_{R,j}(\mathcal{C}_i).
\end{aligned}$$

Using the discrepancy bound $\sum_{j=0}^{2R}\widetilde{D}^{\mathrm{Left}}_{R,j}(\mathcal{C}_i) \preceq R^{3/2}$, the new term in the right-hand side is controlled by $\frac{1}{i}R^{3/2}$, and the sum of these terms over $i = 2,\dots,M-2$ is thus bounded by

$$R^{3/2}\sum_{i=2}^{M-2}\frac{1}{i} \preceq R^{3/2}\log M,$$

which is negligible with respect to M.

Step 3. THE ELECTRIC ENERGY.

CLAIM 8.6 (The electric energy of P^{av}). *We have*

$$\mathcal{W}(\mathsf{P}^{\mathrm{av}}) < \frac{1}{2}\big(\mathcal{W}(\mathsf{P}^0) + \mathcal{W}(\mathsf{P}^1)\big) - \frac{\mathfrak{g}}{10}.$$

PROOF. The proof follows from the previous claim and the "electric-intrinsic" inequality (2.16), which applies here because, by construction, the discrepancy in any interval is bounded by $4R$ (see the remark following immediately (8.1)). $\square$

Combining Claims 8.2 and 8.6 yields (8.2) and finishes the proof of Proposition 8.1. □

Appendix A Miscellaneous Proofs

A.1 Proof of Lemma 4.1

We follow the argument developed in [16, sec. 8.5] with parameters $\mathsf{d} = 1$, $ß = 0$, $\mathsf{k} = 1$, and $\gamma = 0$, but the proof below is self-contained. In what follows, equation numbers with a **bold typeface** refer to the corresponding equations in that paper.

We obtain (we could use (2.14)) with $\eta_0 = \frac{1}{4}$ (the precise value of η_0 is, in fact, irrelevant):

$$\frac{1}{2\pi}\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1\times\mathbb{R}} |\mathrm{E}_{\eta_0}|^2\right] \leq \mathcal{W}(P) + C. \tag{A.1}$$

For any $T > 0$, we let $H_{R,T}$ be the rectangle $H_{R,T} := \Lambda_R \times [-T, T]$. Let us emphasize that here, in contrast to **(8.5)**, we do not yet fix T with respect to R. The integration by parts as in **(8.6)** still holds, and we get

$$\int_{\partial H_{R,T}} \mathrm{E}_{\eta_0} \cdot \vec{\nu} = -2\pi(\mathrm{Discr}_{\Lambda_R}(\mathcal{C}) + r_{\eta_0}), \tag{A.2}$$

where r_{η_0} is an error term bounded by the number of points in a η_0-neighborhood of $\{-R, +R\}$. It is easy to see that $\mathbb{E}_{\mathsf{P}}[r_{\eta_0}^2]$ is bounded by a constant independent of R, and thus, since we are aiming for a $o(R)$ bound, this term is negligible; for simplicity we will forget it.

The main improvement on the existing proof is to observe that the choice $T \in (R, 2R)$ as in **(8.5)** is valid, but slightly suboptimal. We replace it by the following claim.

CLAIM A.1. *There exists* $f : [0, +\infty) \to [0, +\infty)$ *satisfying*

$$\lim_{x\to\infty} f(x) = +\infty, \quad \lim_{x\to\infty} \frac{f(x)}{x} = 0, \tag{A.3}$$

and such that

$$\lim_{x\to\infty} x\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1\times\{-f(x), f(x)\}} |\mathrm{E}_{\eta_0}|^2\right] = 0. \tag{A.4}$$

PROOF. Let $\mathrm{Tail}(x)$ be the quantity

$$\mathrm{Tail}(x) = \mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1\times(\mathbb{R}\setminus(-x,x))} |\mathrm{E}_{\eta_0}|^2\right]. \tag{A.5}$$

The map $x \mapsto \mathrm{Tail}(x)$ is continuous, positive, nonincreasing, and satisfies

$$\lim_{x\to+\infty} \mathrm{Tail}(x) = 0.$$

So, introducing the map

$$u \mapsto \frac{u}{\sqrt{\mathrm{Tail}(u)}},$$

it is continuous on $[0, +\infty)$, increasing, is equal to 0 at 0 and tends to $+\infty$ for $u \to +\infty$. Thus, for any given $x > 0$ there exists (a unique) $u > 0$ such that

$$\frac{u}{\sqrt{\mathrm{Tail}(u)}} = x. \tag{A.6}$$

Now, by a mean value argument, we may find $v \in [u, 2u]$ such that

$$\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1 \times \{-v,v\}} |\mathrm{E}_{\eta 0}|^2\right] \leq \frac{1}{u}\mathrm{Tail}(u).$$

We define $f(x)$ as the smallest such real number v. It is easy to check that the properties (A.3) are satisfied, and moreover we have

$$x\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1 \times \{-f(x),f(x)\}} |\mathrm{E}_{\eta 0}|^2\right] \leq \frac{x}{u}\mathrm{Tail}(u) = \sqrt{\mathrm{Tail}(u)},$$

where we have used (A.6). This quantity (seen as depending on x) tends to 0 as $x \to \infty$, which proves (A.4). □

We now take $T = f(R)$ in (A.2), and we apply the Cauchy-Schwarz inequality to the left-hand side. Since we are dealing with two (slightly) different lengths R, $f(R)$, it is important to be more precise than in **(8.8)** and to split $\partial H_{R,f(R)}$ as

$$\partial H_{R,T} = \Lambda_R \times \{-f(R), f(R)\} \cup \{-R, R\} \times [-f(R), f(R)].$$

We obtain

$$\left(\int_{\partial H_{R,f(R)}} \mathrm{E}_{\eta 0} \cdot \vec{v}\right)^2 \preceq R \int_{\Lambda_R \times \{-f(R),f(R)\}} |\mathrm{E}_{\eta 0}|^2 + f(R) \int_{\{-R,R\} \times [-f(R),f(R)]} |\mathrm{E}_{\eta 0}|^2.$$

By stationarity, we have, in view of (A.1)

$$\frac{1}{2\pi}\mathbb{E}_{\mathsf{P}}\left[\int_{\{-R,R\}\times[-f(R),f(R)]} |E_{\eta 0}|^2\right] = \frac{1}{2\pi}\mathbb{E}_{P}\left[\int_{\Lambda_1 \times [-f(R),f(R)]} |E_{\eta 0}|^2\right] \leq \mathcal{W}(P) + C,$$

and also

$$\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_R \times \{-f(R),f(R)\}} |E_{\eta 0}|^2\right] = R\mathbb{E}_{\mathsf{P}}\left[\int_{\Lambda_1 \times \{-f(R),f(R)\}} |E_{\eta 0}|^2\right]. \tag{A.7}$$

According to (A.3), $f(R) = o(R)$, and from (A.4) we see that the right-hand side of (A.7) is $o_R(1)$. We thus obtain

$$\mathbb{E}_{\mathsf{P}}\left[\left(\int_{\partial H_{R,f(R)}} E_{\eta 0} \cdot \vec{v}\right)^2\right] = Ro_R(1) + o(R)(\mathcal{W}(P) + C),$$

and thus, in view of (A.2), we obtain $\mathbb{E}_{\mathsf{P}}[\mathrm{Discr}_R^2(\mathcal{C})] = o(R)$, which proves (4.3).

A.2 Proof of Lemma 4.4

The argument is essentially a re-interpretation of the discrepancy controls as given in [18, lemma 2.2]. It is enough to show that

$$\sum_{i=-\frac{R}{2}}^{\frac{R}{2}} 1_{\Gamma_i \geq 10} (\Gamma_i)^2 \preceq \int_{[-R,R]\times\mathbb{R}} |\mathrm{E}_\eta|^2.$$

Let i be such that $\Gamma_i \geq 10$. Let $m = \frac{x_i + x_{i+1}}{2}$, and for $\ell > 0$ let H_ℓ be the rectangle $[m-\ell, m+\ell] \times [-\ell, \ell]$. By a mean value argument, we can find $\ell \in [\frac{1}{4}\Gamma_i, \frac{1}{2}\Gamma_i]$ such that

$$\int_{\partial H_\ell} |\mathrm{E}_\eta|^2 \preceq \frac{1}{\Gamma_i} \int_{[x_i, x_{i+1}]\times\mathbb{R}} |\mathrm{E}_\eta|^2. \tag{A.8}$$

On the other hand, using (2.3) and an integration by parts, we have

$$\int_{\partial H_\ell} \mathrm{E}_\eta \cdot \vec{n} = + \int_{[m-\ell,m+\ell]\times[-\ell,\ell]} \mathrm{div}(\mathrm{E}_\eta) = 4\pi\ell.$$

Indeed, by construction there is no point of $\mathcal{C}$ between $m-\ell$ and $m+\ell$.

Let us observe that $\Gamma_i \preceq \ell$. Thus, using the Cauchy-Schwarz inequality and (A.8), we obtain

$$|\Gamma_i|^2 \leq \Gamma_i \int_{\partial H_\ell} |\mathrm{E}_\eta|^2 \preceq \int_{[x_i, x_{i+1}]\times\mathbb{R}} |\mathrm{E}_\eta|^2.$$

The result follows by summing on i and observing that

$$\sum_{i=-R/2}^{R/2} \int_{[x_i, x_{i+1}]\times\mathbb{R}} |\mathrm{E}_\eta|^2 \leq \int_{\Lambda_R\times\mathbb{R}} |\mathrm{E}_\eta|^2,$$

since by assumption there are at least $R/2$ points on both sides of Λ_R.

A.3 Proof of Proposition 4.6

We recall that $\mathrm{Conf}(\mathbb{R})$ denotes the space of point configuration on $\mathbb{R}$. Since $\mathsf{P}^0 \neq \mathsf{P}^1$, there exists a continuous function $F : \mathrm{Conf}(\mathbb{R}) \to \mathbb{R}$ with $\|F\|_\infty = 1$ and $c > 0$ such that

$$\mathbb{E}_{\mathsf{P}^0}[F] - \mathbb{E}_{\mathsf{P}^1}[F] = c. \tag{A.9}$$

Furthermore, without loss of generality we may assume that F is *local* in the sense that there exists $N > 0$ such that for any $\mathcal{C}$ in $\mathrm{Conf}(\mathbb{R})$,

$$F(\mathcal{C}) = F(\mathcal{C} \cap \Lambda_N). \tag{A.10}$$

Indeed, by dominated convergence, we have

$$\lim_{R\to\infty} (\mathbb{E}_{\mathsf{P}^0}[F(\cdot \cap \Lambda_R)] - \mathbb{E}_{\mathsf{P}^1}[F(\cdot \cap \Lambda_R)]) = \mathbb{E}_{\mathsf{P}^0}[F] - \mathbb{E}_{\mathsf{P}^1}[F].$$

For any $M > 0$, we define the function $F_M : \mathrm{Conf}(\mathbb{R}) \to \mathbb{R}$ as

$$F_M(\mathcal{C}) := \frac{1}{2M} \int_{-M}^{M} F(\mathcal{C} + t)\,dt.$$

Since $\mathsf{P}^0, \mathsf{P}^1$ are stationary, we have, for M arbitrary, in view of (A.9),

$$\mathbb{E}_{\mathsf{P}^0}[F_M] - \mathbb{E}_{\mathsf{P}^1}[F_M] = c > 0. \tag{A.11}$$

Also, since F is local and satisfies (A.10), we have

$$F_M(\mathcal{C}) = F_M(\mathcal{C} \cap \Lambda_{M+N}). \tag{A.12}$$

Recall that, in Definition 4.2, we denote by $x_0(\mathcal{C})$ the first nonnegative point of $\mathcal{C}$. Since $\mathsf{P}^0, \mathsf{P}^1$ are assumed to have finite energy, the discrepancy estimate (4.10) holds, and we have

$$\lim_{T \to \infty} \mathsf{P}^0(x_0(\mathcal{C}) > T) + \mathsf{P}^1(x_0(\mathcal{C}) > T) = 0,$$

so we may choose T (depending on $c, \mathsf{P}^0, \mathsf{P}^1$) such that

$$\mathsf{P}^0(x_0(\mathcal{C}) > T) + \mathsf{P}^1(x_0(\mathcal{C}) > T) \leq \frac{c}{100}. \tag{A.13}$$

Once T is fixed, we choose M such that

$$\frac{T}{M} \leq \frac{c}{100}. \tag{A.14}$$

We may also impose that $M > N$, where N is as in (A.10), so in view of (A.12) we have for any $\mathcal{C}$ in $\mathrm{Conf}(\mathbb{R})$,

$$F_M(\mathcal{C}) = F_M(\mathcal{C} \cap \Lambda_{2M}). \tag{A.15}$$

CLAIM A.2. *If* $x_0(\mathcal{C}) \leq T$, *and* (A.14) *holds, we have*

$$|F_M(\mathcal{C}) - F_M(\mathcal{C} - x_0(\mathcal{C}))| \leq \frac{c}{100} \tag{A.16}$$

PROOF. We use the definition of F_M, the fact that F is bounded by 1, and the choice of M with respect to T as in (A.14), and compute

$$\begin{aligned} &|F_M(\mathcal{C}) - F_M(\mathcal{C} - x_0(\mathcal{C}))| \\ &\quad \leq \frac{1}{2M}\left(\int_{M - x_0(\mathcal{C})}^{M} |F(\mathcal{C} + t)|\,dt + \int_{-M - x_0(\mathcal{C})}^{-M} |F(\mathcal{C} + t)|\,dt\right) \\ &\quad \leq \frac{2x_0(\mathcal{C})}{2M}\|F\|_\infty \leq \frac{T}{M} \leq \frac{c}{100}. \end{aligned}$$ □

Since $\mathsf{P}^0, \mathsf{P}^1$ have intensity 1, for $r \geq 1$ large enough (depending on $M, c, \mathsf{P}^0, \mathsf{P}^1$), we have

$$\mathsf{P}^0(|\mathcal{C} \cap \Lambda_{8M}| > r) + \mathsf{P}^1(|\mathcal{C} \cap \Lambda_{8M}| > r) \leq \frac{c}{100}. \tag{A.17}$$

For r fixed such that (A.17) holds, we consider the map on $\mathbb{R}^{2r+1}$ defined by

$$H(a_{-r}, \dots, a_r) := F_M\left(\delta_0 + \sum_{i=1}^{r} (\delta_{p_i} + \delta_{p_{-i}})\right), \tag{A.18}$$

where we let the p_i's be

$$p_i := \sum_{k=1}^{i} a_k \ (i \geq 1), \quad p_{-i} := \sum_{k=1}^{i} -a_{-k} \ (i \geq 1). \tag{A.19}$$

In other words, $H(a_{-r}, \dots, a_r)$ is obtained by applying F_M to the configuration made of one point at 0 and $2r$ points located at p_i for $|i| = 1, \dots, r$. The idea is that if $a_{-r}, \dots, a_r$ are the first gaps of $\mathcal{C}$, then this configuration is made of the first points of $\mathcal{C} - x_0(\mathcal{C})$. Since F_M is bounded by 1, clearly so is H.

CLAIM A.3. *We have*

$$\mathbb{E}_{\mathsf{P}^0}[H(\Gamma_{-r}(\mathcal{C}), \dots, \Gamma_r(\mathcal{C}))] - \mathbb{E}_{\mathsf{P}^1}[H(\Gamma_{-r}(\mathcal{C}), \dots, \Gamma_r(\mathcal{C}))] > 0. \tag{A.20}$$

PROOF. First of all, since $\mathsf{P}^0, \mathsf{P}^1$ have finite energy, the configurations have almost surely infinitely many points in $\mathbb{R}_+$ and $\mathbb{R}_-$; hence the gaps are almost surely all finite and $H(\Gamma_{-r}(\mathcal{C}), \dots, \Gamma_r(\mathcal{C}))$ is well-defined almost surely.

Since H is bounded by 1, and since (A.13), (A.17) hold, we have

$$\begin{aligned} &|\mathbb{E}_{\mathsf{P}^0}[H(\Gamma_{-r}, \dots, \Gamma_r)] - \mathbb{E}_{\mathsf{P}^0}[1_{x_0(\mathcal{C}) \leq T} 1_{|\mathcal{C} \cap \Lambda_{8M}| \leq r} H(\Gamma_{-r}, \dots, \Gamma_r)]| \\ &\quad \leq \frac{c}{100} + \frac{c}{100}, \end{aligned} \tag{A.21}$$

and the same holds for P^1. Knowing that $x_0(\mathcal{C}) \leq T$, we have by (A.16)

$$|F_M(\mathcal{C} - x_0(\mathcal{C})) - F_M(\mathcal{C})| \leq \frac{c}{100},$$

and since F_M satisfies (A.15), we have

$$F_M(\mathcal{C} - x_0(\mathcal{C})) = F_M((\mathcal{C} - x_0(\mathcal{C})) \cap \Lambda_{2M}).$$

Of course, the point configuration $\mathcal{C} - x_0(\mathcal{C})$ can be written in terms of the gaps of $\mathcal{C}$ as

$$\mathcal{C} - x_0(\mathcal{C}) = \delta_0 + \sum_{i=1}^{+\infty} (\delta_{p_i} + \delta_{p_{-i}}),$$

where the p_i's are as in (A.19). Knowing that, moreover, $|\mathcal{C} \cap \Lambda_{8M}| \leq r$, we see that

$$(\mathcal{C} - x_0(\mathcal{C})) \cap \Lambda_{2M} = \left(\delta_0 + \sum_{i=1}^{r} \delta_{p_i} + \delta_{p_{-i}}\right) \cap \Lambda_{2M},$$

and thus, using again (A.12),

$$F_M((\mathcal{C} - x_0(\mathcal{C})) \cap \Lambda_{2M}) = F_M\left(\left(\delta_0 + \sum_{i=1}^{r}(\delta_{p_i} + \delta_{p_{-i}})\right) \cap \Lambda_{2M}\right)$$
$$= F_M\left(\delta_0 + \sum_{i=1}^{r}(\delta_{p_i} + \delta_{p_{-i}})\right) = H(\Gamma_{-r}, \dots, \Gamma_r).$$

We thus obtain by (A.16)

$$\begin{aligned}&\big|\mathbb{E}_{\mathsf{P}^0}[1_{x_0(\mathcal{C})\le T} 1_{|\mathcal{C}\cap\Lambda_{8M}|\le r} H(\Gamma_{-r}, \dots, \Gamma_r)] \\ &- \mathbb{E}_{\mathsf{P}^0}[1_{x_0(\mathcal{C})\le T} 1_{|\mathcal{C}\cap\Lambda_{8M}|\le r} F_M(\mathcal{C})]\big| \\ &\le \frac{c}{100},\end{aligned}$$

which easily yields, in view of (A.21),

$$|\mathbb{E}_{\mathsf{P}^0}[H(\Gamma_{-r}, \dots, \Gamma_r)] - \mathbb{E}_{\mathsf{P}^0}[F_M(\mathcal{C})]| \le \frac{5c}{100}, \tag{A.22}$$

and the same goes for P^1, which proves the claim. □

Finally, we use a density argument in $L^1(\mathbb{R}_+^{2r+1})$ to find a test function on $\mathbb{R}_+^{2r+1}$ that satisfies (A.20) *and* is compactly supported and Lipschitz with respect to the $\|\cdot\|_1$ norm. By possibly reducing the lower bound in (A.20), we can assume the test function to be 1-Lipschitz.

A.4 Proof of Proposition 4.7

The proof will be carried out in a number of steps. We recall that $\mathrm{Conf}(\mathbb{R})$ denotes the space of point configuration on $\mathbb{R}$.

Step 1. DETECTING THE LOCAL DIFFERENCE. Let $c > 0$, $r \ge 1$, and a function $H : \mathbb{R}_+^{2r+1} \to \mathbb{R}$ as given by Proposition 4.6 such that

$$\mathbb{E}_{\mathsf{P}^0}[H(\Gamma_{-r}, \dots, \Gamma_r)] - \mathbb{E}_{\mathsf{P}^1}[H(\Gamma_{-r}, \dots, \Gamma_r)] \ge c. \tag{A.23}$$

The function H is compactly supported, so let L be such that H is supported in $[0, L]^{2r+1}$. Without loss of generality, we can take $L > 10$. To clarify notation, let us define $\widetilde{H}(\mathcal{C})$ for a configuration $\mathcal{C}$ in $\mathrm{Conf}(\mathbb{R})$ as

$$\widetilde{H}(\mathcal{C}) := H(\Gamma_{-r}(\mathcal{C}), \dots, \Gamma_r(\mathcal{C})). \tag{A.24}$$

Strictly speaking, $\widetilde{H}$ is not defined everywhere on $\mathrm{Conf}(\mathbb{R})$, but it is well-defined on the set of configurations with at least $r + 1$ points on $\mathbb{R}_+$ and $\mathbb{R}_-$, because it ensures that the gaps $\Gamma_{-r}(\mathcal{C}), \dots, \Gamma_r(\mathcal{C})$ are all finite.

Step 2. A TEST FUNCTION TO DETECT THE GLOBAL DIFFERENCE.
For $R > 1$, we define the function $\breve{H}_R$ on $\mathrm{Conf}(\mathbb{R})$ as

$$\breve{H}_R : \mathcal{C} \mapsto \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C} - t)\,dt. \tag{A.25}$$

Clearly, $\breve{H}_R$ is bounded by $\frac{R}{10}$. Since $\mathsf{P}^0, \mathsf{P}^1$ are stationary, we have of course, using (A.23),

$$\mathbb{E}_{\mathsf{P}^0}[\breve{H}_R] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R] \geq \frac{cR}{10}. \tag{A.26}$$

Strictly speaking, $\breve{H}_R$ is not defined everywhere on Conf($\mathbb{R}$). It is well-defined on the set configurations with at least $r+1$ points on $[\frac{R}{10}, +\infty)$ and on $(-\infty, 0]$, because it guarantees that the gaps $\Gamma_{-r}(\mathcal{C}-t), \dots, \Gamma_r(\mathcal{C}-t)$ are finite for all $t \in [0, \frac{R}{10}]$.

Now, if Q is a coupling of P^0 and P^1, we may rewrite (A.26) as

$$\mathbb{E}_{\mathsf{Q}}\left[\int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^0 - t)dt - \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^1 - t)dt\right] \geq \frac{cR}{10}. \tag{A.27}$$

Since we work only with a coupling of the restrictions of P^0 and P^1 to Λ_R, we have to restrict the argument above to an event of high probability, ensuring in particular that there are at least $r+1$ points in $[-R, 0]$ and $[R/10, R]$, so that $\breve{H}_R$ depends only on the configuration in Λ_R; see (A.39) for a precise formulation.

Step 3. STRATEGY OF THE REST OF THE PROOF. To simplify, let us assume that $\widetilde{H}(\mathcal{C})$ depends only on the first gap of $\mathcal{C}$ (the one with index 0, between $x_0(\mathcal{C})$ and $x_1(\mathcal{C})$; see Definition 4.2). Since H is Lipschitz, for any two configurations $\mathcal{C}^0, \mathcal{C}^1$ and any t, we thus would have

$$|\widetilde{H}(\mathcal{C}^0 - t) - \widetilde{H}(\mathcal{C}^1 - t)| \leq |\Gamma_0(\mathcal{C}^0 - t) - \Gamma_0(\mathcal{C}^1 - t)|.$$

As t goes from 0 to $\frac{R}{10}$, the first gap of $\mathcal{C}^0 - t$ would successively correspond to the first gap of $\mathcal{C}^0$, then the second one, etc., up to a gap of order $\approx \frac{R}{10}$ in $\mathcal{C}^0$, and similarly for $\mathcal{C}^1 - t$. If the gaps of $\mathcal{C}^0 - t$ and $\mathcal{C}^1 - t$ were always "aligned", i.e., for any t, the index k_0 such that $\Gamma_0(\mathcal{C}^0 - t) = \Gamma_{k_0}(\mathcal{C}^0)$ and the index k_1 (defined similarly for $\mathcal{C}^1$) were equal, we would write, using a Fubini-type argument,

$$\int_0^{\frac{R}{10}} |\widetilde{H}(\mathcal{C}^0 - t) - \widetilde{H}(\mathcal{C}^1 - t)|dt \lessapprox \sum_{k=0}^{\frac{R}{10}} |\Gamma_k(\mathcal{C}^0) - \Gamma_k(\mathcal{C}^1)|, \tag{A.28}$$

and thus, in view of (A.25) and (A.26), we would get a lower bound of order R on a certain "gap difference". Getting a lower bound of the type

$$\sum_{k=0}^{\frac{R}{10}} \left|\Gamma_k(\mathcal{C}^0) - \Gamma_k(\mathcal{C}^1)\right| \geq cR \tag{A.29}$$

would be enough for our purposes.

Compared to this situation, the quantity $\widetilde{H}(\mathcal{C})$ depends on more than one gap. However, it only depends on a finite number of gaps (here, at most $2r+1$), and the strategy can be easily adapted. The major complication comes from aligning gaps, which we describe it below.

We want to detect a difference between the gaps of $\mathcal{C}^0$ and $\mathcal{C}^1$. In the statement of Proposition 4.7, we include the possibility of a fixed shift S, but let us take $S = 0$ and try to transform (A.27) into a bound of the type (A.29). Let us consider a typical situation[4] where $x_0(\mathcal{C}^0) > 0$ and $x_0(\mathcal{C}^1) > 0$. Taking the time $t = 0$ in the integrals, we can bound

$$\begin{aligned} |\widetilde{H}(\mathcal{C}^0 - 0) - \widetilde{H}(\mathcal{C}^1 - 0)| &\le \sum_{i=-r}^{r} |\Gamma_i(\mathcal{C}^0 - 0) - \Gamma_i(\mathcal{C}^1 - 0)| \\ &= \sum_{i=-r}^{r} |\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)|. \end{aligned} \tag{A.30}$$

The right-hand side of (A.30) appears in the sum in (A.29), and the left-hand side of (A.30) appears in the integral of (A.27), so (A.30) can be used to transform the lower bound in (A.27) into a lower bound of the type (A.29). Now let us increase t. We have again, using the assumption on H,

$$|\widetilde{H}(\mathcal{C}^0 - t) - \widetilde{H}(\mathcal{C}^1 - t)| \le \sum_{i=-r}^{r} |\Gamma_i(\mathcal{C}^0 - t) - \Gamma_i(\mathcal{C}^1 - t)|,$$

and for t small we still have

$$\Gamma_i(\mathcal{C}^0 - t) = \Gamma_i(\mathcal{C}^0), \quad \Gamma_i(\mathcal{C}^1 - t) = \Gamma_i(\mathcal{C}^1).$$

However, these identities cease to hold as soon as we encounter a point of $\mathcal{C}^0$ or $\mathcal{C}^1$, i.e., as soon as $t = T_0 := \min(x_0(\mathcal{C}^0), x_0(\mathcal{C}^1))$. Indeed, assuming that the first point encountered is $x_0(\mathcal{C}^0)$, we have, for t slightly larger than T_0 but smaller than $x_0(\mathcal{C}^1)$,

$$\Gamma_i(\mathcal{C}^0 - t) = \Gamma_{i+1}(\mathcal{C}^0), \quad \text{but still } \Gamma_i(\mathcal{C}^1 - t) = \Gamma_i(\mathcal{C}^1),$$

so comparing naively the two integrands with the Lipschitz control of H gives us a lower bound

$$|\widetilde{H}(\mathcal{C}^0 - t) - \widetilde{H}(\mathcal{C}^1 - t)| \le \sum_{i=-r}^{r} |\Gamma_{i+1}(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)|. \tag{A.31}$$

The left-hand side of (A.31) is still present in the integrals of (A.27), however the right-hand side of (A.31) does *not* appear in the sum (A.29), and thus (A.31) is useless for us. To remedy this misalignment, we need to shift the configuration $\mathcal{C}^1$ by $x_0(\mathcal{C}^1) - x_0(\mathcal{C}^0)$, i.e., to add a quantity $x_0(\mathcal{C}^1) - x_0(\mathcal{C}^0)$ to the "proper time" t of $\mathcal{C}^1$. Indeed, we have, for t slightly larger than $T_0 + x_0(\mathcal{C}^1) - x_0(\mathcal{C}^0) = x_0(\mathcal{C}^1)$,

$$\Gamma_i(\mathcal{C}^1 - t) = \Gamma_{i+1}(\mathcal{C}^1),$$

and thus comparing $\mathcal{C}^0 - t$ and $\mathcal{C}^1 - t - (x_0(\mathcal{C}^1) - x_0(\mathcal{C}^0))$ yields again a summand from (A.29). Each time that we encounter a "k^{th} point", no matter whether

[4] For technical reasons, when writing the formal proof below, we actually enforce that $x_0(\mathcal{C}^0) = x_0(\mathcal{C}^1) = 0$, which twists the enumeration a little bit, but does not modify the general strategy.

it comes from $\mathcal{C}^0$ or from $\mathcal{C}^1$, we need to shift the "proper time" of the other configuration in order to "align the gaps". Doing this, we effectively "lose" portions of the interval $[0, \frac{R}{10}]$ on which we integrate, which possibly deteriorates the lower bound of (A.27). In fact, it turns out that the loss can be expressed in terms of the "gap difference" itself.

Another technicality occurs when one tries to make the Fubini argument yielding (A.28) rigorously, and we will need to estimate the time "spent" on each gap Γ_i for $i \in [0, \frac{R}{10}]$.

Step 4. A GOOD EVENT. Let us introduce the following events:

$$\begin{aligned}
\mathsf{EG}_1 &:= \left\{\left|\mathcal{C}_{[\frac{R}{10},R]}\right| \geq r+1\right\} \cap \left\{\left|\mathcal{C}_{[-R,0]}\right| \geq r+1\right\},\\
\mathsf{EG}_2 &:= \left\{x_{\mathrm{Pos}(\mathcal{C};\frac{R}{10})+r} \leq \frac{R}{4}\right\},\\
\mathsf{EG}_3 &:= \left\{x_{R^{1/2}}(\mathcal{C}) \leq 2R^{1/2}\right\} \cap \left\{x_{-R^{1/2}}(\mathcal{C}) \geq -2R^{1/2}\right\},\\
\mathsf{EG}_4 &:= \left\{\left|\mathcal{C}_{[-\frac{R}{10},\frac{R}{10}]}\right| \leq \frac{R}{3}\right\},
\end{aligned}$$

and EGap_R as the intersection

$$\mathsf{EGap}_R := \mathsf{EG}_1 \cap \mathsf{EG}_2 \cap \mathsf{EG}_3 \cap \mathsf{EG}_4. \tag{A.32}$$

By construction, if $\mathcal{C}$ is in EGap_R, we know that

- Event EG_1. There are at least $r+1$ points in $[\frac{R}{10}, R]$ and in $[-R, 0]$, so $\breve{H}_R(\mathcal{C}|_{\Lambda_R})$ is well-defined.
- Event EG_2. The r^{th} gap of $\mathcal{C} - \frac{R}{10}$ (which is the rightmost gap that we consider when applying $\breve{H}_R$ to a configuration $\mathcal{C}$) corresponds to points in $[0, \frac{R}{4}]$. It implies that $\breve{H}_R(\mathcal{C})$ depends only on $\mathcal{C}|_{\Lambda_{R/4}}$, namely

$$\breve{H}_R(\mathcal{C}) = \breve{H}_R(\mathcal{C}|_{\Lambda_{R/4}}). \tag{A.33}$$

- Event EG_3. The $R^{1/2\text{th}}$ point on each side is at distance at most $2R^{1/2}$. Since by assumption S is a random variable bounded by $R^{1/2}$, it yields

$$|x_S(\mathcal{C})| \preceq R^{1/2}. \tag{A.34}$$

  We may note that we also have the very rough bound

$$x_0(\mathcal{C}) \preceq R^{1/2}. \tag{A.35}$$

- Event EG_4. There are at most $\frac{R}{3}$ points in $[-\frac{R}{10}, \frac{R}{10}]$, and in particular

$$|x_k(\mathcal{C})| \leq \frac{R}{10} \implies |k| \leq \frac{R}{3}. \tag{A.36}$$

Since r is fixed, the discrepancy estimates (4.1) and (4.11) guarantee that

$$\lim_{R\to\infty} \mathsf{P}^0(\mathsf{EGap}_R) = \lim_{R\to\infty} \mathsf{P}^1(\mathsf{EGap}_R) = 1,$$

so for R large enough, we have

$$\mathsf{P}^0(\mathsf{EGap}_R) \geq 1 - \frac{c}{100}, \quad \mathsf{P}^1(\mathsf{EGap}_R) \geq 1 - \frac{c}{100}. \tag{A.37}$$

Step 5. THE QUANTITY TO COMPUTE. For R large enough, we may write, since $\breve{H}_R$ is bounded by $\frac{R}{10}$,

$$\mathbb{E}_{\mathsf{P}^0}[\breve{H}_R] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R] \leq \mathbb{E}_{\mathsf{P}^0}[\breve{H}_R 1_{\mathsf{EGap}_R}] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R 1_{\mathsf{EGap}_R}] + \frac{2c}{100} \times \frac{R}{10}. \tag{A.38}$$

On the other hand, since Q is a coupling of the restrictions of P^0, P^1 to Λ_R, in view of (A.33) we have

$$\begin{aligned}
&\mathbb{E}_{\mathsf{P}^0}\big[\breve{H}_R 1_{\mathsf{EGap}_R}\big] - \mathbb{E}_{\mathsf{P}^1}\big[\breve{H}_R 1_{\mathsf{EGap}_R}\big] \\
&\quad = \mathbb{E}_{\mathsf{Q}}\big[\breve{H}_R(\mathcal{C}^0) 1_{\mathsf{EGap}_R}(\mathcal{C}^0) - \breve{H}_R(\mathcal{C}^1) 1_{\mathsf{EGap}_R}(\mathcal{C}^1)\big].
\end{aligned}$$

We now turn to compute the quantity

$$\begin{aligned}
&\mathbb{E}_{\mathsf{Q}}\big[1_{\mathsf{EGap}_R}(\mathcal{C}^0)\breve{H}_R(\mathcal{C}^0) - 1_{\mathsf{EGap}_R}(\mathcal{C}^1)\breve{H}_R(\mathcal{C}^1)\big] \\
&\quad = \mathbb{E}_{\mathsf{Q}}\bigg[1_{\mathsf{EGap}_R}(\mathcal{C}^0) \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^0 - t)\,dt \\
&\qquad - 1_{\mathsf{EGap}_R}(\mathcal{C}^1) \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^1 - t)\,dt\bigg].
\end{aligned} \tag{A.39}$$

Let $\mathcal{C}^0, \mathcal{C}^1$ be fixed, and assume both belong to EGap_R.

Step 6. THE INITIAL SHIFT. Let v_0, v_1 be defined as

$$v_0 = x_0(\mathcal{C}^0), \quad v_1 = x_S(\mathcal{C}^1). \tag{A.40}$$

In view of (A.34), (A.35), we have

$$|v_0| + |v_1| \preceq R^{1/2}. \tag{A.41}$$

By the definition (4.4), (4.5) of the enumeration of points and gaps, the choice of v_0, v_1 as in (A.40) ensures that for all i

$$\begin{aligned}
&x_i(\mathcal{C}^0 - v_0) = x_i(\mathcal{C}^0) - v_0, \quad x_i(\mathcal{C}^1 - v_1) = x_{i+S}(\mathcal{C}^1) - v_1, \\
&\Gamma_i(\mathcal{C}^0 - v_0) = \Gamma_i(\mathcal{C}^0), \qquad \Gamma_i(\mathcal{C}^1 - v_1) = \Gamma_{i+S}(\mathcal{C}^1).
\end{aligned} \tag{A.42}$$

Remark A.4 (Notational choice). In what follows, we will write, for simplicity, $\mathcal{C}^0 = \mathcal{C}^0 - v_0$ and $\mathcal{C}^1 = \mathcal{C}^1 - v_1$. This amounts to doing a translation in each integral of (A.39), and the error is of order $|v_0| + |v_1|$, which is bounded as in (A.41), and thus negligible for our purposes (we pursue a lower bound of order R). It places us in an ideal situation where $S = 0$, and where, at time $t = 0$, both configurations have a point at 0.

Step 7. THE PROPER TIMES.

- Let $T_0 = 0$, and recall (see previous paragraph) that $x_0(\mathcal{C}^0) = x_0(\mathcal{C}^1) = 0$. We define the functions
$$t^0 1_0(t) = t, \quad t_0^1(t) = t.$$
- We let T_1 be the first time at which we encounter a new point, more precisely:
$$T_1 := \min\bigl\{t: t_0^0(t) - t_0^0(T_0) = \Gamma_0(\mathcal{C}^0) \text{ or } t_0^1(t) - t_0^1(T_0) = \Gamma_0(\mathcal{C}^1)\bigr\}. \tag{A.43}$$
- We then define the functions
$$\begin{aligned} t_1^0(t) &= t_0^0(t) + \bigl(\Gamma_0(\mathcal{C}^0) - \bigl(t_0^0(T_1) - t_0^0(T_0)\bigr)\bigr), \\ t_1^1(t) &= t_0^1(t) + \bigl(\Gamma_0(\mathcal{C}^1) - \bigl(t_0^1(T_1) - t_0^1(T_0)\bigr)\bigr). \end{aligned} \tag{A.44}$$
Of course, we have in fact $t_0^0(T_0) = t_0^1(T_0) = 0$, $T_1 = \min(\Gamma_0(\mathcal{C}^0), \Gamma_0(\mathcal{C}^1))$, and one of the quantities
$$\Gamma_0(\mathcal{C}^0) - T_1, \quad \Gamma_0(\mathcal{C}^1) - T_1,$$
is equal to 0, while the other one is the "shift"
$$\Gamma_0(\mathcal{C}^0) - \Gamma_0(\mathcal{C}^1) \quad \text{or} \quad \Gamma_0(\mathcal{C}^1) - \Gamma_0(\mathcal{C}^0),$$
that we must apply to the configuration with a larger first gap.
- Assume that T_k, t_k^0, t_k^1 have been defined for some $k \geq 1$. We let the time T_{k+1} be given by
$$T_{k+1} := \min\bigl\{t,\ t_k^0(t) - t_k^0(T_k) = \Gamma_k(\mathcal{C}^0) \text{ or } t_k^1(t) - t_k^1(T_k) = \Gamma_k(\mathcal{C}^1)\bigr\}, \tag{A.45}$$
and we introduce the functions
$$\begin{cases} t_{k+1}^0(t) := t_k^0(t) + \bigl(\Gamma_k(\mathcal{C}^0) - \bigl(t_k^0(T_{k+1}) - t_k^0(T_k)\bigr)\bigr), \\ t_{k+1}^1(t) := t_k^1(t) + \bigl(\Gamma_k(\mathcal{C}^1) - \bigl(t_k^1(T_{k+1}) - t_k^1(T_k)\bigr)\bigr). \end{cases} \tag{A.46}$$
Once again, by definition one of the shifts is equal to 0, and the other one is given by
$$\Gamma_k(\mathcal{C}^0) - \Gamma_k(\mathcal{C}^1) \quad \text{or} \quad \Gamma_k(\mathcal{C}^1) - \Gamma_k(\mathcal{C}^0). \tag{A.47}$$
- Let $k_{\max}$ be defined as
$$k_{\max} := \min\left\{k : T_k \geq \frac{R}{10}\right\}. \tag{A.48}$$
We will only consider $k \leq k_{\max}$; in other words, we stop the definition when $T_k \geq \frac{R}{10}$ and let $k_{\max}$ be the number of steps. Since (A.36) holds, we have $k_{\max} \leq \frac{R}{3}$.

CLAIM A.5. *For all k, we have*
$$\begin{aligned} x_0(\mathcal{C}^0 - t_k^0(T_k)) &= x_k(\mathcal{C}^0) - t_k^0(T_k) = 0, \\ x_0(\mathcal{C}^1 - t_k^1(T_k)) &= x_k(\mathcal{C}^1) - t_k^1(T_k) = 0, \end{aligned} \tag{A.49}$$

and for t in (T_k, T_{k+1}), we have, using the notation of (4.6),

$$\mathrm{Pos}\big(\mathcal{C}^0, t_k^0(t)\big) = \mathrm{Pos}\big(\mathcal{C}^1, t_k^1(t)\big) = k+1, \tag{A.50}$$

which in particular implies the following gap alignment identity for $t \in (T_k, T_{k+1})$:

$$\Gamma_0\big(\mathcal{C}^0 - t_k^0(t)\big) = \Gamma_{k+1}(\mathcal{C}^0), \quad \Gamma_0\big(\mathcal{C}^1 - t_k^1(t)\big) = \Gamma_{k+1}(\mathcal{C}^1). \tag{A.51}$$

PROOF. We recall that we enforced $x_0(\mathcal{C}^0) = x_1(\mathcal{C}^0) = 0$, which is (A.49) for $k = 0$. It allows us to rewrite the definition (A.43) as

$$T_1 = \min\big(t, t = \Gamma_0(\mathcal{C}^0) \text{ or } t = \Gamma_0(\mathcal{C}^1)\big);$$

thus in fact $\Gamma_0(\mathcal{C}^0) = x_1(\mathcal{C}^0)$ and $\Gamma_0(\mathcal{C}^1) = x_1(\mathcal{C}^1)$, and T_1 is *the first positive time at which a point of $\mathcal{C}^0$ or $\mathcal{C}^1$ is encountered.*

For $0 < t < T_1$, the configuration $\mathcal{C}^0 - t$ has a "first negative point" given by $x_0(\mathcal{C}^0) - t = -t$ and a "first positive point" given by $x_1(\mathcal{C}^0) - t$, and the same holds for $\mathcal{C}^1 - t$. Thus (A.50) holds for $k = 0$.

Assume that (A.49) holds for some $k \geq 0$. Without loss of generality, assume that

$$T_{k+1} = \min\big(t, t_k^0(t) - t_k^0(T_k) = \Gamma_k(\mathcal{C}^0)\big).$$

Since by the induction hypothesis $t_k^0(T_k) = x_k(\mathcal{C}^0)$, we see that $t_{k+1}^0(T_{k+1})$ must be given by $x_k(\mathcal{C}^0) + \Gamma_k(\mathcal{C}^0)$, which is equal to $x_{k+1}(\mathcal{C}^0)$. We also obtain, following the definition,

$$t_{k+1}^1(T_{k+1}) = t_k^1(T_{k+1}) + \big(\Gamma_k(\mathcal{C}^1) - \big(t_k^1(T_{k+1}) - t_k^1(T_k)\big)\big),$$

and using the induction hypothesis $t_k^1(T_k) = x_k(\mathcal{C}^1)$, we obtain

$$t_{k+1}^1(T_{k+1}) = x_k(\mathcal{C}^1) + \Gamma_k(\mathcal{C}^1) = x_{k+1}(\mathcal{C}^1),$$

which proves (A.49) at rank $k+1$. We also deduce that (A.50) holds at rank k. The claim is thus proven by induction. $\square$

CLAIM A.6 (Estimate on the time loss). *For a bounded function $F(t)$, we have*

$$\begin{aligned} &\left| \int_0^{\frac{R}{10}} F(t)dt - \sum_{k=0}^{k_{\max}-1} \int_{T_k}^{T_{k+1}} F(t_k^0(t))dt \right| \\ &\quad \leq \|F\|_\infty \sum_{k=0}^{k_{\max}} \big|\Gamma_k(\mathcal{C}^0) - \Gamma_k(\mathcal{C}^1)\big|, \end{aligned} \tag{A.52}$$

and similarly when replacing t^0 by t^1.

PROOF. The intervals (T_k, T_{k+1}) are disjoint, and the change of variable $t \mapsto t_k^0(t)$ has speed 1. The time loss

$$\left[0, \frac{R}{10}\right] \setminus \bigcup_{k=0}^{k_{\max}-1} \{t_k^0((T_k, T_{k+1}))\}$$

comes from the time shifts (A.47), and its length is thus bounded by the sum of all possible shifts, which yields the right-hand side of (A.52). □

Step 8. USING THE LIPSCHITZ BOUND ON SHIFTED INTERVALS.

CLAIM A.7. *For all* $k \leq k_{\max}$, *for* $t \in (T_k, T_{k+1})$, *we have*

$$\begin{aligned} &\left|\widetilde{H}\big(\mathcal{C}^0 - t_k^0(t)\big) - \widetilde{H}\big(\mathcal{C}^1 - t_k^1(t)\big)\right| \\ &\quad \preceq L \sum_{i=-r}^{r} \frac{\left|\Gamma_{i+1+k}(\mathcal{C}^0) - \Gamma_{i+1+k}(\mathcal{C}^1)\right|}{\Gamma_{i+1+k}(\mathcal{C}^0) + \Gamma_{i+1+k}(\mathcal{C}^1)}. \end{aligned} \tag{A.53}$$

PROOF. We first use the fact that H is 1-Lipschitz with respect to the $\|\cdot\|_1$ norm on $\mathbb{R}_+^{2r+1}$, and the identity (A.50) to obtain

$$\left|\widetilde{H}\big(\mathcal{C}^0 - t_k^0(t)\big) - \widetilde{H}\big(\mathcal{C}^1 - t_k^1(t)\big)\right| \leq \sum_{i=-r}^{r} \left|\Gamma_{i+k+1}(\mathcal{C}^0) - \Gamma_{i+k+1}(\mathcal{C}^1)\right|. \tag{A.54}$$

To improve (A.54) into (A.53), we recall that the function H is bounded by 1 and supported in $[0, L]^{2r+1}$ for some $L > 10$.

For any i between $-r$ and r, let us distinguish cases:

- The gaps $\Gamma_{i+k+1}(\mathcal{C}^0)$ and $\Gamma_{i+k+1}(\mathcal{C}^1)$ are both smaller than $2L$, in which case we certainly have

$$\left|\Gamma_{i+k+1}(\mathcal{C}^0) - \Gamma_{i+k+1}(\mathcal{C}^1))\right| \leq 4L \frac{\left|\Gamma_{i+k+1}(\mathcal{C}^0) - \Gamma_{i+k+1}(\mathcal{C}^1)\right|}{\Gamma_{i+k+1}(\mathcal{C}^0) + \Gamma_{i+k+1}(\mathcal{C}^1))}.$$

- Both are larger than L, in which case the left-hand side of (A.53) is 0.
- One of these quantities is smaller than L (say, without loss of generality, the first one) and the other one is larger than $2L$, in which case the left-hand side of (A.53) is bounded by 1, and the right-hand side contains the term

$$4L \frac{\left|\Gamma_{i+k+1}(\mathcal{C}^0) - \Gamma_{i+k+1}(\mathcal{C}^1)\right|}{\Gamma_{i+k+1}(\mathcal{C}^0) + \Gamma_{i+k+1}(\mathcal{C}^1)} \geq 4L \frac{\Gamma_{i+k+1}(\mathcal{C}^1) - L}{L + \Gamma_{i+k+1}(\mathcal{C}^1)},$$

  but clearly $\frac{4L(x-L)}{x+L} \geq 2$ for $x \geq 2L$ and $L > 10$, so the right-hand side of (A.53) is bounded below by 2, and the inequality holds. □

Step 9. SHIFTING INTEGRALS AND A FUBINI ARGUMENT.

CLAIM A.8.

$$\begin{aligned} &\left|\int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^0 - t)\,dt - \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^1 - t)\,dt\right| \\ &\quad \leq \sum_{i=-\infty}^{+\infty} L\mathrm{Weight}_i \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)\right|}{\Gamma_i(\mathcal{C}^0) + \Gamma_i(\mathcal{C}^1)} + \mathsf{Error}, \end{aligned} \tag{A.55}$$

where Weight_i *satisfies*

$$\mathrm{Weight}_i \preceq \sum_{j=i-r-2}^{i+r} \min\bigl(\Gamma_j(\mathcal{C}^0), \Gamma_j(\mathcal{C}^1)\bigr), \tag{A.56}$$

and the range condition

$$\mathrm{Weight}_i = 0 \quad \text{if } |i| \geq R/2, \tag{A.57}$$

and Error *is bounded by*

$$\mathsf{Error} \preceq \sum_{i=0}^{R/2} \bigl|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)\bigr|. \tag{A.58}$$

PROOF. We use Claim A.6 with $F = \widetilde{H}(\mathcal{C}^0 - \cdot)$ and $\widetilde{H}(\mathcal{C}^1 - \cdot)$, and write

$$\begin{aligned} &\int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^0 - t)dt - \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^1 - t)dt \\ &\quad = \sum_{k=0}^{k_{\max}-1} \int_{T_k}^{T_{k+1}} \bigl(\widetilde{H}(\mathcal{C}^0 - t_k^0(t)) - \widetilde{H}(\mathcal{C}^1 - t_k^1(t))\bigr)dt + \mathsf{Error}, \end{aligned} \tag{A.59}$$

with Error bounded as in (A.58). □

Next, we use (A.53) from Claim A.7, and write, for every k,

$$\begin{aligned} &\left|\int_{T_k}^{T_{k+1}} \widetilde{H}(\mathcal{C}^0 - t_k^0(t)) - \widetilde{H}(\mathcal{C}^1 - t_k^1(t))dt\right| \\ &\quad \preceq L(T_{k+1} - T_k) \sum_{i=-r}^{r} \frac{\bigl|\Gamma_{i+k+1}(\mathcal{C}^0) - \Gamma_{i+k+1}(\mathcal{C}^1)\bigr|}{\Gamma_{i+k+1}(\mathcal{C}^0) + \Gamma_{i+k+1}(\mathcal{C}^1)}. \end{aligned} \tag{A.60}$$

Combining (A.59) and (A.60), a Fubini argument yields (A.55), with weights

$$\mathrm{Weight}_i \preceq \sum_k 1_{k \in [i-r-1, i+r-1], 0 \leq k \leq k_{\max}} (T_{k+1} - T_k).$$

This is a telescopic sum, and we may write

$$\mathrm{Weight}_i \preceq T_{i+r} - T_{\min\{i-r-1,0\}}.$$

By definition of the times T_k, we see that

$$T_{i+r} - T_{\min\{i-r-1,0\}} \leq \sum_{j=i-r-1}^{i+r} \min\bigl(\Gamma_j(\mathcal{C}^0), \Gamma_j(\mathcal{C}^1)\bigr), \tag{A.61}$$

which yields (A.56). Moreover, Weight_i is 0 if $i + r - 1 < 0$ or $i + r - 1 > k_{\max}$, which in particular (since $k_{\max} \leq \frac{R}{3}$, taking R large enough with respect to r) implies (A.57).

Step 10. FINAL COMPUTATION I, Let us introduce the quantity

$$\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1) := \sum_{i=-\frac{R}{2}}^{\frac{R}{2}} \frac{\left(\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)\right)^2}{(\Gamma_i(\mathcal{C}^0))^2 + (\Gamma_i(\mathcal{C}^1))^2}, \tag{A.62}$$

and let us recall that the goal of this proof is to obtain a lower bound on the expectation, under the coupling Q, of $\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)$ that would be proportional to R.

CLAIM A.9. *There exists a constant $\mathsf{C}_{L,r}$ depending only on L, r such that*

$$\begin{aligned} &\left| \int_0^{R/10} \widetilde{H}(\mathcal{C}^0 - t)dt - \int_0^{R/10} \widetilde{H}(\mathcal{C}^1 - t)dt \right| \\ &\quad \leq \mathsf{C}_{L,r} \Bigg(\sum_{i=-R/2}^{R/2} \Gamma_k(\mathcal{C}^0)^2 + \Gamma_k(\mathcal{C}^1)^2 \Bigg)^{1/2} \left(\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)\right)^{1/2}. \end{aligned} \tag{A.63}$$

PROOF OF CLAIM A.9. We write the result of Claim A.8 as follows:

$$\begin{aligned} &\left| \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^0 - t)dt - \int_0^{\frac{R}{10}} \widetilde{H}(\mathcal{C}^1 - t)dt \right| \\ &\preceq \sum_{i=-R/2}^{R/2} L \Bigg(\sum_{j=i-r-2}^{i+r} \min\left(\Gamma_j(\mathcal{C}^0), \Gamma_j(\mathcal{C}^1)\right) \Bigg) \frac{\left|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)\right|}{\Gamma_i(\mathcal{C}^0) + \Gamma_i(\mathcal{C}^1)} \\ &\quad + \sum_{i=0}^{R/2} \left|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)\right|. \end{aligned} \tag{A.64}$$

For the first term in the right-hand side of (A.64), we use the Cauchy-Schwarz inequality and the trivial bound

$$\min\left(\Gamma_j(\mathcal{C}^0), \Gamma_j(\mathcal{C}^1)\right)^2 \leq \Gamma_j(\mathcal{C}^0)^2 + \Gamma_j(\mathcal{C}^1)^2,$$

while for the last term in the right-hand side, we write

$$|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)| = \frac{|\Gamma_i(\mathcal{C}^0) - \Gamma_i(\mathcal{C}^1)|}{\Gamma_i(\mathcal{C}^0) + \Gamma_i(\mathcal{C}^1)} \left(\Gamma_i(\mathcal{C}^0) + \Gamma_i(\mathcal{C}^1)\right),$$

and use the Cauchy-Schwarz inequality. In both cases, we obtain a term bounded by the right-hand side of (A.63) up to a multiplicative constant depending only on L, r. □

Step 11. FINAL COMPUTATION II. Taking the expectation of (A.63) under the coupling Q (to simplify, we forget about the conditioning on the event EGap_R, which does not affect the estimate; see, e.g., (A.38). We obtain, replacing the left-hand side of (A.63) as in (A.39),

$$\begin{aligned}&\mathbb{E}_{\mathsf{P}^0}[\breve{H}_R] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R] \\ &\preceq \mathbb{E}_{\mathsf{Q}}\left[\left(\sum_{i=-\frac{R}{2}}^{R/2} \Gamma_k(\mathcal{C}^0)^2 + \Gamma_k(\mathcal{C}^1)^2\right)^{1/2} \left(\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)\right)^{1/2}\right].\end{aligned} \tag{A.65}$$

Using the Cauchy-Schwarz inequality, Lemma 4.4, (2.14), and the fact that $\mathsf{P}^0, \mathsf{P}^1$ have finite energy, the right-hand side of (A.65) can be bounded in order to yield

$$\mathbb{E}_{\mathsf{P}^0}[\breve{H}_R] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R] \preceq R^{1/2}\mathbb{E}_{\mathsf{Q}}\left[\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)\right]^{\frac{1}{2}},$$

and if we use (A.26) to bound below $\mathbb{E}_{\mathsf{P}^0}[\breve{H}_R] - \mathbb{E}_{\mathsf{P}^1}[\breve{H}_R]$ we deduce that

$$R \preceq R^{1/2}\mathbb{E}_{\mathsf{Q}}\left[\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)\right]^{\frac{1}{2}},$$

so $R \preceq \mathbb{E}_{\mathsf{Q}}[\mathrm{gain}_R(\mathcal{C}^0, \mathcal{C}^1)]$, which concludes the proof.

Appendix B The Screening Procedure

We present a sketch of the screening argument as developed in [18, sec. 6]. For the notation of that paper, our setting is $d = 1$, $s = 0$, $k = 1$, $\gamma = 0$, $g(x) = -\log|x|$, $c_{d,s} = 2\pi$, and our s is their ε. We also use the fact that, in the present case, the background measure has a constant intensity. We do not claim to make any serious improvement on the procedure.

SKETCH OF PROOF OF PROPOSITION 3.2. We recall that $R' = R(1-s)$ and that by assumptions (3.2), (3.3) we have

$$\int_{\{-R', R'\}\times[-R,R]} |\mathrm{E}_\eta|^2 = \mathsf{M}_{\mathrm{scr}} \le M, \tag{B.1}$$

$$\frac{1}{s^4 R}\int_{\Lambda_R\times\mathbb{R}\setminus(-s^2R, s^2R)} |E|^2 = \mathsf{e}_{\mathrm{scr}} \le 1, \tag{B.2}$$

and, that by assumption (3.4), the smeared-out charges $\sigma_{x,\eta}$ for x in $\mathcal{C}$ do not intersect the boundary $\Lambda_{R'} \times \mathbb{R}$.

By a mean value argument, there exists ℓ in $[s^2R, 2s^2R]$ such that

$$\int_{\Lambda_R\times\{-\ell,\ell\}} |E|^2 \le s^2\mathsf{e}_{\mathrm{scr}} \le s^2. \tag{B.3}$$

Subdividing the domain

As depicted on Figure B.1, we decompose $\Lambda_R \times [-R, R]$ into three parts:

$$D_0 = \Lambda_{R'} \times [-\ell, \ell], \quad D_\partial = (\Lambda_R \times [-\ell, \ell]) \setminus D_0,$$
$$D_1 = (\Lambda_R \times [-R, R]) \setminus (D_0 \cup D_\partial).$$

Roughly speaking, here is what the screening procedure entails:

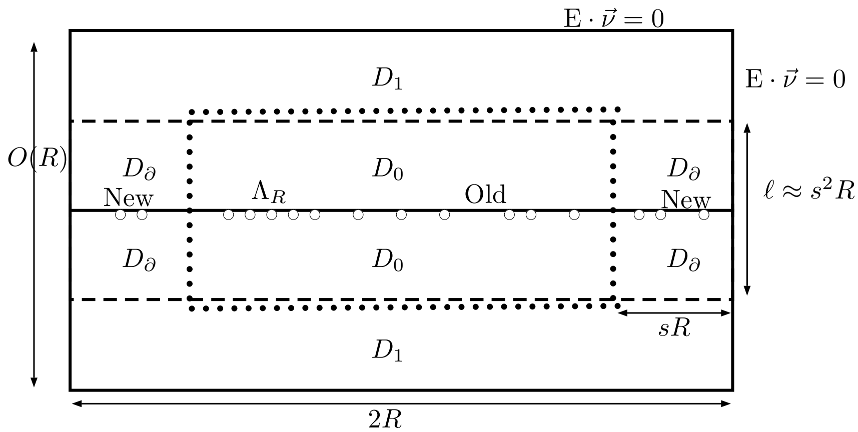

FIGURE B.1. Sketch of the situation.

- The point configuration will be kept in Old and the existing electric field will be kept in D_0.
- In D_∂, we throw away the field and configuration. We will place a correct number of new points and define an electric field whose normal component coincides with the existing one on the vertical dotted lines and reaches 0 on the vertical full line. There will be, however, a nonvanishing component of the field at the level of the dashed line.
- In D_1, we manipulate the electric field, starting from the horizontal dashed line, in order to reach a zero normal component on the exterior (full line).

We let $N_{\text{int}} = |\mathcal{C} \cap \Lambda_{R'}| = |\mathcal{C} \cap D_0|$ be the number of points of $\mathcal{C}$ inside D_0– these will not be touched, and we will place $|\Lambda_R| - N_{\text{int}}$ points in New in order for the final configuration to have $|\Lambda_R|$ points.

We also define U_0 as the quantity

$$U_0 := \frac{1}{2(|\Lambda_R| - |\Lambda_{R'}|)} \int_{[-R',R'] \times \{-\ell,\ell\}} \mathrm{E} \cdot \vec{\nu}. \tag{B.4}$$

The integral in (B.4) corresponds to the integral of the normal component of the existing field on the horizontal dotted lines in Figure B.1.

The screening procedure enforces that the normal component of the constructed field on the dashed line part of the boundary of D_∂ exactly compensates U_0, so that the total flux on the boundary of both rectangles that form D_1 is 0.

CLAIM B.1 (Size of U_0). *We have*

$$U_0 \preceq R^{-1/2}. \tag{B.5}$$

PROOF. Since $R' = R(1-s)$ and by definition of U_0, we have

$$U_0 \preceq \frac{1}{sR} \int_{[-R',R']\times\{-\ell,\ell\}} |E|,$$

and combining the Cauchy-Schwarz inequality with (B.3) we obtain (B.5). $\square$

We split $\Lambda_R \setminus \Lambda_{R'}$ into intervals H_i whose lengths belong to $[\ell/2, 2\ell]$, and we let $\widetilde{H}_i = H_i \times [-\ell, \ell]$. We denote by H_{left} (resp., H_{right} the interval exactly to the left (resp.,, to the right) or $-R'$ (resp., R').

For any interval in this decomposition, we let m_i be such that

$$2\pi(m_i - 1)|H_i| = \int_{(\{-R',R'\}\times[-\ell,\ell])\cap\partial\widetilde{H}_i} \mathrm{E}_\eta \cdot \vec{\nu} - 2U_0|H_i|. \tag{B.6}$$

The first term in the right-hand side of (B.6) is only present if H_i is H_{left} or H_{right}; the second term is always present.

CLAIM B.2 (Estimate on m_i). *We quantify how close to* 1 *the number m_i defined in* (B.6) *is.*

- *If H_i is not H_{left} or H_{right}, we have then*

$$|m_i - 1| \preceq R^{-1/2}. \tag{B.7}$$

- *If H_i is one of the two intervals H_{left}, H_{right} that have an intersection with $\partial\Lambda_{R'}$, we have*

$$|m_i - 1| \preceq R^{-1/2} + \mathsf{M}_{\text{scr}}^{1/2} R^{-1/2} s^{-1}. \tag{B.8}$$

In particular, in both cases, for R large enough, we have

$$|m_i - 1| \preceq \frac{1}{2}. \tag{B.9}$$

PROOF. In the first case, we have $|m_i - 1| \preceq U_0$ and we apply (B.5). In the second case, we add the contribution of the integral

$$\frac{1}{|H_i|} \int_{\partial D_0 \cap \partial\widetilde{H}_i} \mathrm{E}_\eta \cdot \vec{\nu} \preceq \ell^{-1}\ell^{1/2} \left(\int_{\partial D_0} |\mathrm{E}_\eta|^2 \right)^{1/2} \preceq (s^2 R)^{-1/2} \mathsf{M}_{\text{scr}}^{1/2},$$

where we have used the Cauchy-Schwarz inequality and (B.1). $\square$

Then each interval H_i is divided into subintervals of length $\frac{1}{m_i}$, and in each of these intervals exactly one point of the new screened configuration is placed. More precisely, a point is placed randomly at a distance less than $\frac{\eta}{4}$ from the center of the interval. This randomness is important, because it “creates volume” and yields the description of the “new points” as in (3.8).

The rest of the screening procedure

For the end of the screening procedure (constructing the screened electric field and estimating its energy), we refer to the proof of [18]. Let us emphasize that, in our setting, some technicalities become irrelevant and that the construction and estimates could be written in a much more concise way. Here is a short sketch thereof:

(1) First, as the k^{th} point z_k is being "placed" in the subinterval I_k, we define the electric field by solving
$$-\mathrm{div}\,\mathrm{E} = 2\pi(\delta_{\mathsf{z}_k} - dx)$$
in $I_k \times [-\ell, \ell]$ with some boundary condition. These boundary conditions are chosen in a compatible fashion for two neighboring subintervals. We also impose a zero boundary condition on the left for the subinterval that contains $-R$ (and similarly for the rightmost subinterval), and to match the existing boundary condition given by the pre-existing field E for the two subintervals that share an endpoint with Old. Estimating the "energy" created this way is an additional task, but we find that it only yields a small error compared to the total energy.

(2) At this stage, the electric field is defined on $\mathsf{Old}\times[-\ell, \ell]$, the region denoted by D_0 in Figure B.1, by keeping the pre-existing field; and on $\mathsf{New} \times [-\ell, \ell]$, the region denoted by D_∂ in Figure B.1, by defining it on each subinterval, as in the previous step. It remains to define it on the region denoted by D_1 in Figure B.1. Of course, we do not place any points here, but we tile this region by small rectangles of side length $\approx \ell$, and on each of these we solve
$$-\mathrm{div}\,\mathrm{E} = 0$$
with an appropriate choice of (mutually compatible) boundary conditions, which allow us to pass from whatever boundary condition exists at the frontier of D_∂, D_0, to the desired Neumann condition $\mathrm{E}\cdot\vec{\nu} = 0$ on the boundary on the big rectangle in Figure B.1. There again, one must estimate the energy of these "patching" fields.

(3) We then obtain the desired screened electric field. $\square$

PROOF OF CLAIM 3.4. By construction, for any H_i we place exactly one point at the center of each subinterval of length $\frac{1}{m_i}$. Since the length of H_i is in $[\ell/2, 2\ell]$, with $\ell \preceq s^2 R$, and since $|m_i - 1| \leq \frac{1}{2}$ (see (B.9)) the number of subintervals in each interval is $\preceq s^2 R$. In particular,

$$\#\text{ points in } H_{\text{left}} \preceq s^2 R. \tag{B.10}$$

In view of Claim B.2, the distance between the position of the k^{th} point (starting from the leftmost one) and $-R + k - \frac{1}{2}$ is bounded, up to a universal multiplicative constant, by

- $kR^{-1/2}$ as long as the point does not belong to H_{left}.

- $kR^{-1/2} + (k-k_0)\mathsf{M}_{\text{scr}}^{1/2}R^{-1/2}s^{-1}$ if the point belongs to H_{left}, where k_0 is the index of the last point outside H_{left}. By (B.10), we may write
$$kR^{-1/2} + (k-k_0)\mathsf{M}_{\text{scr}}^{1/2}R^{-1/2}s^{-1} \preceq \left(1+\mathsf{M}_{\text{scr}}^{1/2}\right)sR^{1/2}.$$

We thus have, as claimed in (3.10),
$$|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq kR^{-1/2},$$
as long as z_k is not in H_{left}, that is, for k such that
$$s^2R \preceq |sR-k|.$$
When z_k is in H_{left}, that is, for k such that
$$|sR-k| \preceq s^2R,$$
we have, as claimed in (3.12),
$$|\mathsf{z}_k - \bar{\mathsf{z}}_k| \preceq \mathsf{M}_{\text{scr}}^{1/2}sR^{1/2}.$$

Let us recall that $k_{\max}$ is the index of the first point such that $\mathsf{z}_k \geq -R+sR$. We deduce that $k_{\max}$ is equal to sR, up to an error of order $\mathsf{M}_{\text{scr}}^{1/2}sR^{1/2}$, which yields (3.9).

The inequalities (3.10), (3.12) can then be converted into discrepancy estimates (3.11) and (3.13). ☐

Acknowledgment. ME was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2047 – 390685813, *Hausdorff Center for Mathematics*, as well as the Collaborative Research Center 1060 – 211504053, *The Mathematics of Emergent Effects* at Bonn University. MH has been partially funded by the Vienna Science and Technology Fund (WWTF) through project VRG17-005 and has been partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044 –390685587, Mathematics Münster: Dynamics–Geometry–Structure. Open Access funding enabled and organized by Projekt DEAL.

Bibliography

[1] Allez, R.; Dumaz, L. From sine kernel to Poisson statistics. *Electron. J. Probab.* **19** (2014), no. 114, 25 pp. doi:10.1214/EJP.v19-3742

[2] Borodin, A.; Serfaty, S. Renormalized energy concentration in random matrices. *Comm. Math. Phys.* **320** (2013), no. 1, 199–244. doi:10.1007/s00220-013-1716-z

[3] Bourgade, P.; Erdös, L.; Yau, H.-T. Bulk universality of general β-ensembles with non-convex potential. *J. Math. Phys.* **53** (2012), no. 9, 095221, 19. doi:10.1063/1.4751478

[4] Bourgade, P.; Erdös, L.; Yau, H.-T. Universality of general β-ensembles. *Duke Math. J.* **163** (2014), no. 6, 1127–1190. doi:10.1215/00127094-2649752

[5] Dereudre, D.; Hardy, A.; Leblé, T.; Maïda, M. DLR equations and rigidity for the Sine-beta process. *Comm. Pure Appl. Math.*, forthcoming. arXiv:1809.03989 [math.PR]

[6] Dumitriu, I.; Edelman, A. Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002), no. 11, 5830–5847. doi:10.1063/1.1507823
[7] Erbar, M.; Huesmann, M. Curvature bounds for configuration spaces. *Calc. Var. Partial Differ. Equ.* **54** (2015), no. 1, 397–430. doi:10.1007/s00526-014-0790-1
[8] Forrester, P. J. *Log-gases and random matrices.* London Mathematical Society Monographs Series, 34. Princeton University Press, Princeton, N.J., 2010. doi:10.1515/9781400835416
[9] Friedli, S.; Velenik, Y. *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press, Cambridge, 2018.
[10] Huesmann, M. Optimal transport between random measures. *Ann. Inst. Henri Poincaré, Probab. Stat.* **52** (2016), no. 1, 196–232. doi:10.1214/14-AIHP634
[11] Huesmann, M.; Sturm, K.-T. Optimal transport from Lebesgue to Poisson. *Ann. Probab.* **41** (2013), no. 4, 2426–2478. doi:10.1214/12-AOP814
[12] Killip, R.; Stoiciu, M. Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.* **146** (2009), no. 3, 361–399. doi:10.1215/00127094-2009-001
[13] Leblé, T. A uniqueness result for minimizers of the 1D log-gas renormalized energy. *J. Funct. Anal.* **268** (2015), no. 7, 1649–1677. doi:10.1016/j.jfa.2014.11.023
[14] Leblé, T. Logarithmic, Coulomb and Riesz energy of point processes. *J. Stat. Phys.* **162** (2016), no. 4, 887–923. doi:10.1007/s10955-015-1425-4
[15] Leblé, T. CLT for fluctuations of linear statistics in the Sine-beta process. Preprint, 2018. arXiv:1809.03448 [math.PR]
[16] Leblé, T.; Serfaty, S. Large deviation principle for empirical fields of log and Riesz gases. *Invent. Math.* **210** (2017), no. 3, 645–757. doi:10.1007/s00222-017-0738-0
[17] McCann, R. J. A convexity principle for interacting gases. *Adv. Math.* **128** (1997), no. 1, 153–179. doi:10.1006/aima.1997.1634
[18] Petrache, M.; Serfaty, S. Next order asymptotics and renormalized energy for Riesz interactions. *J. Inst. Math. Jussieu* **16** (2017), no. 3, 501–569. doi:10.1017/S1474748015000201
[19] Rougerie, N.; Serfaty, S. Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.* **69** (2016), no. 3, 519–605. doi:10.1002/cpa.21570
[20] Sandier, É.; Serfaty, S. From the Ginzburg-Landau model to vortex lattice problems. *Comm. Math. Phys.* **313** (2012), no. 3, 635–743. doi:10.1007/s00220-012-1508-x
[21] Sandier, É.; Serfaty, S. 1D log gases and the renormalized energy: crystallization at vanishing temperature. *Probab. Theory Related Fields* **162** (2015), no. 3-4, 795–846. doi:10.1007/s00440-014-0585-5
[22] Sandier, É.; Serfaty, S. 2D Coulomb gases and the renormalized energy. *Ann. Probab.* **43** (2015), no. 4, 2026–2083. doi:10.1214/14-AOP927
[23] Serfaty, S. Microscopic description of log and Coulomb gases. Preprint, 2017. arXiv:1709.04089 [math-ph]
[24] Valkó, B.; Virág, B. Continuum limits of random matrices and the Brownian carousel. *Invent. math.* **177** (2009), no. 3, 463–508. doi:10.1007/s00222-009-0180-z
[25] Valkó, B.; Virág, B. Operator limit of the circular beta ensemble. Preprint, 2017. arXiv:1710.06988 [math.PR]
[26] Valkó, B.; Virág, B. The Sine_β operator. *Invent. math.* **209** (2017), no. 1, 275–327. doi:10.1007/s00222-016-0709-x
[27] Villani, C. *Topics in optimal transportation.* Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, R.I., 2003. doi:10.1090/gsm/058

MATTHIAS ERBAR
Fakultät für Mathematik
Universität Bielefeld
Universitätsstraße 25
33615 Bielefeld
GERMANY
E-mail: erbar@math.uni-bielefeld.de

THOMAS LEBLÉ
Université de Paris
CNRS, MAP5 UMR 8145
F-75006 Paris
FRANCE
E-mail: thomasleble@gmail.com

MARTIN HUESMANN
Institut für Mathematische Stochastik
Westfälische Wilhelms-Universität Münster
Orléans Ring 10
48149 Münster
GERMANY
E-mail: martin.huesmann@uni-muenster.de

Received February 2019.