Séminaire Laurent Schwartz

EDP et applications

Année 2022-2023

Exposé n° VI (13 décembre 2022)

Thomas Leblé

HYPERUNIFORMITY OF THE TWO-DIMENSIONAL ONE-COMPONENT PLASMA

https://doi.org/10.5802/slsedp.161

Institut des hautes études scientifiques
Le Bois-Marie • Route de Chartres
F-91440 BURES-SUR-YVETTE
http://www.ihes.fr/

Centre de mathématiques Laurent Schwartz
CMLS, École polytechnique, CNRS, Université Paris-Saclay
F-91128 PALAISEAU CEDEX
http://www.math.polytechnique.fr/

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2266-0607

Séminaire Laurent-Schwartz — EDP et applications
Institut des hautes études scientifiques, 2022-2023
Exposé n° VI, 1-16

Hyperuniformity of the two-dimensional one-component plasma

Thomas Leblé

1 Introduction

1.1 Point processes and number variance

We study certain two-dimensional *point processes*, which are random finite or locally finite collection of points in $\mathbb{R}^2$. When $\mathbf{X}$ is such a random variable, it is natural to consider the counting statistics i.e. the (random) number of points $|\mathbf{X} \cap \Lambda|$ falling in various regions $\Lambda \subset \mathbb{R}^2$, as these are arguably the most fundamental "observables" for point processes. We will adopt the following convention:

1. Either $\mathbf{X}$ is a.s. finite, with a deterministic total number of points N all contained in a certain "big box" Σ_N which is specified when defining $\mathbf{X}$, and has area N.
2. Or $\mathbf{X}$ is a.s. infinite and satisfies (with $\mathfrak{D}_R$ the disk of center 0 and radius R)
$$\lim_{R\to\infty} \frac{1}{|\mathfrak{D}_R|}\mathbb{E}[|\mathbf{X} \cap \mathfrak{D}_R|] = 1. \tag{1.1}$$

One can think of both cases as imposing that the average number of points per unit volume be equal to 1 *at large scales*. However, it does not mean that if Λ is a bounded region (contained in Σ_N in the first case, or simply in $\mathbb{R}^2$ in the second one) then we must have $|\mathbf{X} \cap \Lambda| = |\Lambda|$ or even that $\mathbb{E}\left[|\mathbf{X} \cap \Lambda|\right] = |\Lambda|$. We call the random difference $|\mathbf{X} \cap \Lambda| - |\Lambda|$ the *discrepancy of* $\mathbf{X}$ *in* Λ, we denote it by $\mathsf{Dis}(\mathbf{X}, \Lambda)$, and we raise two basic questions:

1. Centeredness of discrepancies? For Λ bounded, is $\mathbb{E}[|\mathbf{X} \cap \Lambda|]$ approximately equal to $|\Lambda|$, in other words are discrepancies approximately centered? What is the size of the error when Λ is, say, some large disk or large square?

2. Number variance growth? How does the *number variance*, i.e. the variance of $|\mathbf{X} \cap \Lambda|$ (note that this is of course also the variance of the discrepancy) grow when Λ is chosen to be, say, some large disk or large square?

There are of course many more questions that one could ask about number statistics, let us mention two others:

3. Rigidity properties? The notion of number-rigidity (which is only relevant for infinite point processes) was introduced by S. Ghosh, see [GP17]. The point process is said to be *number-rigid* when for all bounded regions Λ, the number statistics $|\mathbf{X}\cap\Lambda|$ is a measurable function of the "exterior configuration" $\mathbf{X}\cap(\mathbb{R}^2\setminus\Lambda)$ - in other words, one can hide the point configuration in Λ and correctly guess the number of points falling in Λ by observing only the point configuration *outside* Λ.

In fact, it is sometimes possible (even in truly random situations) to guess the value of other statistics, like the center of mass, or higher moments of the point configuration, or even the full configuration (one says that the process is *fully rigid*), see [GL17].

4. Decay of correlations? This question could be phrased in many different ways, the simplest one is to ask whether, when Λ is a bounded domain and Λ_t corresponds to Λ shifted by t in a certain direction, the covariance between $|\mathbf{X}\cap\Lambda|$ and $|\mathbf{X}\cap\Lambda_t|$ tends to 0 as $t\to\infty$.

Another approach would be to ask about the k-point correlation functions ρ_k for $k\geq 1$ (assuming they all exist as functions) and their hypothetical "clustering" properties, in particular whether $\rho_2(x,y)-\rho_1(x)\rho_1(y)$ tends to 0 as $|x-y|\to\infty$ (and if yes at which speed).

1.2 A quick panorama of related point processes

Although our goal here is to say something about the "two-dimensional one-component plasma" (2DOCP) which is a Gibbsian point process coming from statistical physics, let us start by presenting a few other interesting two-dimensional point processes and see how they behave in regard to the four questions we just raised.

A. The Poisson point process By definition, the Poisson point process (with intensity measure constant equal to 1) is such that:

i) In each bounded region Λ, the number of points $|\mathbf{X}\cap\Lambda|$ follows a Poisson distribution of parameter $|\Lambda|$.

ii) If Λ,Λ' are disjoint then the configurations $\mathbf{X}\cap\Lambda$ and $\mathbf{X}\cap\Lambda'$ are independent.

From the first property we immediately see that:

1. The discrepancies are **exactly centered**.
2. The number variance in Λ grows exactly as the area of Λ. In particular, in a disk of radius R the number variance is **proportional to** R^2.

Moreover there is absolutely **no rigidity** in the sense of Ghosh-Peres: the configuration outside Λ is indeed completely independent from the configuration

inside Λ. For the same reason there is **extremely good decay of correlations** (in every possible sense).

Physically speaking, one can think of the Poisson point process as a stereotype of what an "infinite temperature" system could look like. It represents complete disorder and lack of interaction between the points.

B. Stationary lattices and their perturbations On the opposite end of the spectrum, let us take for $\mathbf{X}$ the *stationary lattice* i.e. $\mathbb{Z}^2$ with a random choice of the origin uniformly in $[0,1]^2$ (or another unimodular lattice with origin chosen uniformly in a fundamental domain).

Thanks to the stationarization it is easy to check that the discrepancies are **exactly centered**, but the number variance computation is not completely obvious - in particular it does depend quite heavily on the shape of the region Λ. In a disk of radius R however, one finds that **the number variance is** $\mathcal{O}(R)$, in sharp contrast with the Poisson case (proportional to R^2). It will be interesting to keep in mind that $\mathcal{O}(R)$ is in fact *the slowest possible growth*, see [Bec87] for a precise statement.

Regarding "rigidity" à la Ghosh-Peres, a stationary lattice is obviously number-rigid, in fact it is "fully rigid" i.e. one can completely deduce $\mathbf{X} \cap \Lambda$ from observing $\mathbf{X} \cap (\mathbb{R}^2 \setminus \Lambda)$. Regarding correlations, there is very long-range order and absolutely no decay.

One can add more randomness to this model by displacing each lattice point with some random perturbation. When the perturbations are i.i.d. and have finite first moment, then the number variance remains $\mathcal{O}(R)$, see [GS75].

In physical terms, lattices can be thought of as describing "ground states" or zero-temperature systems, with a lot of order. Perturbed lattices could (at least very unformally) correspond to small positive temperatures.

C. Zeroes of the Gaussian Entire Function We now turn to a more truly random object. Let $(a_n)_{n\geq 1}$ be a family of independent standard complex Gaussian variables. Almost surely the random series $\sum_{n\geq 0} \frac{a_n}{\sqrt{n!}} z^n$ converges for all $z \in \mathbb{C}$ and forms an entire function over the complex plane called the "Gaussian Entire Function" (GEF). Its zero set is a random, infinite, locally finite collection of points whose distribution is translation-invariant (although that last property is not obvious).

Due to translation-invariance, **discrepancies are again exactly centered**. Moreover it is known (see [FH99, SZ08]) that the number variance in a disk of radius R is **of order** $\mathcal{O}(R)$ and thus that the zeroes of the GEF have the "best" number variance, comparable to the case of a lattice. Moreover, they display somewhat amusing rigidity properties: not only is the point process **number-rigid**, but observing $\mathbf{X} \cap (\mathbb{R}^2 \setminus \Lambda)$ allows one to **also recover the center of mass** of $\mathbf{X} \cap \Lambda$.

D. The Ginibre ensemble Starting with a family $(a_{i,j})_{1\leq i,j\leq n}$ of independent standard complex Gaussian variables, form the associated square matrix and consider its spectrum. It is a.s. formed of n distinct complex eigenvalues, whose joint distribution is known explicitely [Gin65] and admits a limit as $n \to \infty$, called the (infinite) Ginibre ensemble. It is a prominent member of the family of "determinantal point processes" for which we refer to [HKPV06]: in short, this means that the system is in some sense "integrable" and opens the way to many exact - but possibly difficult - computations.

The Ginibre ensemble turns out to have a translation-invariant distribution and thus **exactly centered discrepancies**. The number variance in large disks grows, again, **as the radius of the disk** (and not the area, like Poisson).

This point process is known to also be **number-rigid** but compared to the zeroes of the GEF it does not possess the extra "center of mass"-rigidity property. Moreover this is one of the rare settings in which the two-point correlation is known, and it exhibits **a very fast decay** - namely $\rho_2(x,y) - 1 = -e^{-c|x-y|^2}$ (for some $c > 0$).

2 The 2DOCP

The two-dimensional "one-component plasma" (also called log-gas, or Coulomb gas, or jellium) is a Gibbs measure whose density has the following form:

$$\mathtt{d}\mathbb{P}_N^\beta(\mathrm{X}_N) \propto \exp(-\beta \mathsf{F}_N(\mathrm{X}_N))\ \mathtt{d}\mathrm{X}_N, \tag{2.1}$$

where $\mathrm{X}_N = (x_1, \dots, x_N)$ denotes a N-tuple of points, $\mathtt{d}\mathrm{X}_N$ is the product Lebesgue measure, and $\mathsf{F}_N(\mathrm{X}_N)$ is the appropriately defined "energy" in the "state" X_N. The parameter $\beta > 0$ is called the "inverse temperature" and the "proportional to" sign $\propto$ means that we need to normalize the density by the right constant (called the partition function) in order to get a probability density.

Here for convenience we impose that the N points be contained the disk $\Sigma_N := \mathfrak{D}(0, \frac{N}{\pi})$ (e.g. by setting $\mathsf{F}_N(\mathrm{X}_N) = +\infty$ otherwise) and then define their energy as:

$$\mathsf{F}_N(\mathrm{X}_N) := \frac{1}{2} \iint_{\Sigma_N \times \Sigma_N, x \neq y} -\log|x-y| \mathtt{d}\mathbf{f}_N(x) \mathtt{d}\mathbf{f}_N(y), \quad \mathbf{f}_N := \sum_{i=1}^N \delta_{x_i} - \mathtt{1}_{\Sigma_N} \mathtt{d}x. \tag{2.2}$$

The signed measure $\mathbf{f}_N$ is the difference between a positive, purely atomic measure of mass N representing the positions of the N particles $x_1, \dots, x_N$, and a negative uniform density on the disk Σ_N (physicists speak about a "uniform neutralizing background"). The charged system encoded by $\mathbf{f}_N$ interacts with itself through the two-dimensional Coulomb kernel $-\log$ which is solution to the Poisson equation $-\Delta(-\log) = 2\pi\delta_0$ and is thus the interaction potential of two-dimensional electrostatics. Note that we remove the diagonal to avoid point charges self-interactions, which are infinite. Let us mention two things:

- There is a (coincidental) close link between the 2DOCP at the specific value $\beta = 2$ and the Ginibre ensemble of random matrices described above: the joint distribution of random eigenvalues and random particles are in fact the same (up to a minor difference due to us choosing a "strict" confinement of the particles within Σ_N when defining the 2DOCP). This allows for exact computations and very precise results in the $\beta = 2$ case.
- Existence of an "infinite-volume limit" i.e. a limit in law of the finite point process defined by $\mathbb{P}_N^\beta$ as $N \to \infty$ is taken for granted in the physics literature (which speak of an "infinitely extended" one-component plasma see e.g. [JLM93]) but is only mathematically proven in the special case $\beta = 2$. Available results in that direction is existence of *limit points* at arbitrary $\beta > 0$ (see [AS21]), and description of those limit points as actual Gibbs states.

There is an extensive literature devoted to the 2DOCP (and closely related two-dimensional models like Riesz gases, where the logarithmic/Coulomb kernel $-\log|x-y|$ is replaced by a power law $|x-y|^{-s}$ with $s > 0$), both on the physics and mathematics side, see refer to the surveys [Ser18] and [Lew22] for references. The fact that the pairwise interaction potential is long-range makes the analysis of the 2DOCP particularly interesting for physicists and mathematicians alike.

Questions related to discrepancies ("charge fluctuations") in classical Coulomb systems like the 2DOCP have been of particular interest for physicists, see e.g. the series of papers [MY80, Leb83, Mar88, JLM93, LWL00] devoted to this topic. The main take-away of their studies is the following "cancellation of charge fluctuations" statement:

> For all $\beta > 0$, as $R \to \infty$, the number variance in a large disk of radius R is $o(R^2)$, it is even $\mathcal{O}(R)$, and in fact it is equivalent to $c_\beta R$ for some constant c_β depending on β.

Early on, this was fairly well understood in the special case $\beta = 2$ thanks to explicit computations.

Hyperuniformity The term "hyperuniform(ity)" has been coined in the theoretical chemistry literature by S. Torquato (see [TS03, Tor18] for surveys), an alternative terminology due to J. Lebowitz is "superhomoge- neous/ity". A system is hyperuniform when *the number variance in a large ball is asymptotically negligible with respect to the volume of said ball.* In other words, the physicists' claim can be rephrased as the fact that a 2DOCP is *hyperuniform at all positive temperatures*, and even that it is "Type I" hyperuniform (in the classification of Torquato) i.e. that the number variance in large disks scales like the perimeter of the disk (which we recall is the "best" possible growth).

For a Poisson point process the number variance scales exactly like the volume, so "hyperuniform" systems are those who do "better than Poisson" in that regard. We have seen several examples in Section 1.2: stationary lattices and their i.i.d. perturbations, the zeroes of the GEF, as well as the points of the Ginibre ensemble. Although differing with respect to the rest of their rigidity properties, they are all "type I" hyperuniform.

3 JLM law & NSV theorem

3.1 The JLM law

In [JLM93], Jancovici-Lebowitz-Manificat study the behavior of discrepancies in the 2DOCP and state that for all $\alpha > \frac{1}{2}$:

$$\mathbb{P}\left[\text{Discrepancy of size } R^\alpha \text{ in a disk of radius } R\right] \sim \exp(-R^{\varphi(\alpha)}) \quad \textbf{(“JLM law”)}$$

where the rate $\varphi(\alpha) > 0$ is an explicit piecewise affine function of α. Note that this implies Type I hyperuniformity for the 2DOCP and is in fact a much more precise statement on the tail probability of “large charge fluctuations” (i.e. fluctuations that are much larger than the predicted standard deviation, which is of order $R^{\frac{1}{2}}$). In the sequel we will focus on the most difficult regime, which is $\alpha \in (\frac{1}{2}, 1)$. The argument in [JLM93] is a physical one:

1. By general properties of charged systems, any excess of charges of size R^α within the disk $\mathfrak{D}_R$ must concentrate over a thin annulus of width $R^{\alpha-1}$ near the boundary of the disk.

2. This will be compensated immediately outside $\mathfrak{D}_R$ by an annulus of comparable width carrying an opposite *deficit* of charges.

3. The free energy of such a double electrical layer can then be estimated by a scaling argument.

Although the first step can be made mathematically rigorous (in a weaker form, see below), it is very hard to treat thin layers with sub-microscopic ($R^{\alpha-1} \ll 1$ here) length scale. The statement of [JLM93] has been checked (in the mathematical literature) for $\beta = 2$ through explicit computations, see [Shi06, FL21] but the general $\beta > 0$ case remains open.

3.2 The NSV theorem

Remarkably, although the original prediction of [JLM93] deals with Coulomb gases, it was first verified in [NSV08] for a different model, namely the zeros of the Gaussian Entire Function! The overall strategy developed by Nazarov-Sodin-Volberg is surprisingly simple and can be summarized as follows:

1. Show that any discrepancy of size R^α must be located near the boundary of the disk, say in an annulus of width ≈ 1.

2. Cut the annulus into R “rectangular” pieces of size ≈ 1 and group them modulo M for some $M \gg 1$. By the pigeonhole principle, one of the families mod M must carry a total discrepancy of size at least R^α/M.

3. Such a family is made of R/M pieces of size 1, with two neighboring pieces distant by $\approx M \gg 1$. One then has to show:

(a) That the point processes induced on each piece are (approximately) independent.

(b) That discrepancies are (approximately) centered on each piece.

(c) That the number variance on each piece is of size ≈ 1.

4. Finally, one concludes using a standard concentration inequality for sums of independent random variables.

Step 3. and its substeps are the central part of their argument. Point 3) (a) is surprising because the points must in fact be very much correlated at large scale (uncorrelated systems can hardly be hyperuniform, think of a Poisson point process) and clearly, the large distance between pieces will play a crucial role. Point 3) (b) seems innocent, especially in view of Section 1.2 where all the processes had discrepancies which were exactly centered, and it is in fact easy for [NSV08] to obtain - see however the discussion below for the 2DOCP, where it is a major roadblock. Finally, although Point 3) (c) might also seem innocent since it amounts to saying that the number variance in a piece of size 1 is... of size 1, one should keep in mind that this small piece is part of a much larger/infinite system which could influence its behavior in a dramatic way.

Our goal is to adapt the strategy of [NSV08] in the case of the 2DOCP which, although it bears some resemblance with the zeroes of GEF, is still of a very different nature.

4 Existing and missing tools for the 2DOCP

Recall the "fluctuation measure" $\mathbf{f}_N := \sum_{i=1}^N \delta_{x_i} - \mathtt{1}_{\Sigma_N} \mathtt{d}x$ introduced in (2.2).

Energy. Thanks to the pioneering analysis of Sandier-Serfaty [SS15], the global logarithmic energy $\mathsf{F}_N(\mathrm{X}_N)$ is known to be of order N with overwhelming probability. Defining the proper notion of a "local" logarithmic energy (say at scale $1 \leq \ell \leq \sqrt{N}$) requires some care, in particular taking the logarithmic interaction of the particles in the sub-region does not lead to the "right" quantity. The correct definition is quite lengthy and we refer to [Leb17, BBNY17, AS21], which establish *local laws* stating that the local energy at every mesoscopic scale $1 \ll \ell \ll \sqrt{N}$ and even at large enough microscopic scale $\ell \geq C$ (a crucial improvement from [AS21]) is $\mathcal{O}(\ell^2)$ with high probability.

Discrepancy-energy inequalities. A leading intuition in the study of systems like the 2DOCP is that:

1. Typical realizations X_N have a "reasonable" energy $\mathsf{F}_N(\mathrm{X}_N)$.

2. In order for $\mathsf{F}_N(\mathrm{X}_N)$ to be "reasonable", the signed measure $\mathbf{f}_N$ should be "close" to 0 i.e. the discrete measure $\sum_{i=1}^N \delta_{x_i}$ should "well" approximate the uniform background $\mathtt{1}_{\Sigma_N} \mathtt{d}x$.

3. In particular, any event that implies a significant difference (a discrepancy!) between the point charges and the Lebesgue measure on the disk should have some kind of energy cost, and thus a quantifiable probabilistic price.

This can be verified, for instance, in the case of charge fluctuations: there are discrepancy-energy inequalities which roughly speaking are of the form:

$$|\mathsf{Dis}(\mathbf{X}, \mathfrak{D}_\ell)|^2 \leq C \times \text{Local energy in } \mathfrak{D}_{2\ell} \tag{4.1}$$

(see [SS15, RS16, AS21] for increasingly sharp statements). On the other hand the local laws imply that the local energy at scale ℓ is of order ℓ^2 in expectation, and one can thus deduce that the number variance in a disk is *bounded by the area* (up to some multiplicative constant). It is interesting to note that getting to this result, i.e. knowing that the 2DOCP is at least as "good" as (let alone *strictly better than*) a Poisson point process in terms of number variance, already required significant work.

Fluctuations of linear statistics - the Lipschitz case. By definition, the discrepancy in a domain Ω can can be written as $\mathsf{Dis}(\mathbf{X}, \Omega) = \int \mathtt{1}_\Omega(x)\mathtt{df}_N(x)$ i.e. as the so-called "linear statistics" associated to the indicator function of Ω (which is not a smooth test function). More generally, for a given test function φ, one can consider the corresponding linear statistics $\int \varphi(x)\mathtt{df}_N(x)$, which also measures a kind of "discrepancy" between the point charges and the background measure, as seen through the "eyes" of φ. When φ is of class C^1 and compactly supported within Σ_N, it is not hard (thanks to a simple integration by parts) to state a "linear statistics-energy" inequality of the form:

$$\left|\int \varphi(x)\mathtt{df}_N(x)\right|^2 \leq C \times \|\varphi\|_{H^1}^2 \times \text{Local energy in } \operatorname{supp}\varphi, \tag{4.2}$$

which is reminiscent of (4.1). Such a bound on fluctuations of C^1 (or Lipschitz) test functions already appeared in [SS15].

Fluctuations of linear statistics - the smooth case. However, when φ is sufficiently smooth (say of class C^4 and compactly supported) then (4.2) is in fact off by the "energy" term, indeed one can prove ([LS18, BBNY19, Ser23]) that the linear statistics is then of order 1, with a standard deviation given by the H^1 norm of the test function up to some multiplicative constant depending on β. This remarkable property (sometimes presented as a "central limit theorem *without* normalization") is a sign of strong rigidity within the system *at all temperatures*.

Interestingly, this last result can be used to prove the first step of the NSV strategy described in Section 3.2: indeed, if the excess of charges is spread uniformly in the disk instead of being localized near the boundary, then it is not hard to construct some smooth test function φ which "detects" this excess of charges, and to which one can apply the precise controls on smooth fluctuations (if however the excess (or default) of charges within the disk is located in a

thin strip near the boundary, then essentially it can only be detected by the indicator function or another test function with sharp cut-off, for which the results of [LS18, BBNY19, Ser23] are irrelevant).

Missing tools. Since the first step of the strategy can be reduced to a matter of smooth linear statistics, and the second step is simply an application of the pigeonhole principle, it remains to implement the three points forming the third and main step of Section 3.2, which all require a new viewpoint.

5 Approximate conditional independence

It is of course not true that the restrictions of the 2DOCP to two regions A, B in Σ_N become independent when A and B are disjoint or even very distant. Clearly the Gibbs measure $\mathbb{P}_N^\beta$ (2.1) couples A and B through the energy term F_N which contains the interaction between (the sub-systems in) A and B, namely:

$$\iint_{A\times B} -\log|x-y|\mathtt{d}\mathbf{f}_N(x)\mathtt{d}\mathbf{f}_N(y). \tag{5.1}$$

Let us take A and B as two disks of radius S. When the distance $\mathrm{dist}(A,B)$ between A and B is much larger than the characteristic size S of A and B, one can perform the following Taylor's expansion for $x \in A, y \in B$:

$$-\log|x-y| = -\log \mathrm{dist}(A,B) + \frac{1}{\mathrm{dist}(A,B)} \times \mathcal{O}(S),$$

and insert this into (5.1) to get:

$$\begin{aligned}\iint_{A\times B} -\log|x-y|\mathtt{d}\mathbf{f}_N(x)\mathtt{d}\mathbf{f}_N(y) = -\log \mathrm{dist}(A,B) \times \mathsf{Dis}(\mathbf{X},A) \times \mathsf{Dis}(\mathbf{X},B) \\ + \frac{1}{\mathrm{dist}(A,B)} \times \mathcal{O}(S) \times \mathcal{O}(S^2) \times \mathcal{O}(S^2),\end{aligned} \tag{5.2}$$

where we have used the fact that the total variation of $\mathbf{f}_N$ within A and B, which is bounded by the number of points (positive part) plus the surface (negative background), is $\mathcal{O}(S^2)$ - this requires to control the typical number of points down to the scale S, which is indeed guaranteed by the local laws mentioned at the beginning of Section 4.

The take-away message from (5.2) is that the interaction between regions A and B can be written as a first part that depends *only on the relative position of A and B and the number of points in both zones, but not on the precise arrangement of the points within them* plus an error term that can be made small if $\mathrm{dist}(A,B)$ is sufficiently large compared to the size S of A and B. So one can hope that: *after conditioning on the number of points in A and B, the regions become approximately independent.* This idea relies simply on a Taylor's expansion but does not seem to have appeared previously.

In fact, the precise statement of this conditional independence leads one to study the restriction of the 2DOCP to certain regions after conditioning not only on the number of points in each region, but also on the state of the system outside all the regions under consideration: in the previous example one would condition on $|\mathbf{X}\cap A|, |\mathbf{X}\cap B|$ *and* on the exterior configuration $\mathbf{X}\cap(\Sigma_N\setminus(A\cup B))$. This naturally leads to an extension of the definition (2.1) of the Gibbs measure in order to consider generalized 2DOCP's with two modifications:

1. The number of points is still fixed, but does not exactly coincide with the area of the region (because although by a priori controls on the discrepancy we know that $|\mathsf{Dis}(\mathbf{X}, A)|$ will typically not be very large, we cannot ensure that $|\mathbf{X}\cap A|$ will be exactly equal to $|A|$).

2. On top of the logarithmic interactions between the points, the energy contains the effect of a one-body potential V which is harmonic but may be singular near the boundary of the support. This potential reflects the influence of the exterior configuration.

Such generalized 2DOCP's have appeared previously in the literature under the name of "conditional" or "local" measures, see e.g. [BBNY17]. Combining their analysis with the techniques of [AS21] we are able to prove that those systems basically retain all the good properties known for the standard 2DOCP, at least when looked at in the bulk i.e. sufficiently far from the boundary (where the external potential V is not well controlled).

The approximate conditional independence together with the study of generalized 2DOCP's as described above allow to treat points (a) and (c) of the strategy's third step. The last property that remains to be checked might seem the simplest one, as it is "only" a matter of estimating an expectation in order to show that discrepancies are almost centered.

6 Centeredness of fluctuations and approximate translation-invariance

6.1 Origin of the problem

First, let us explain where this needs come from. The very last step in the strategy of NSV consists in using the following standard concentration inequality (Hoeffding's inequality): if $X_1, \dots, X_n$ are independent centered random variables bounded by b then:

$$\mathbb{P}[|\sum_{i=1}^{n} X_i| \geq t] \leq \exp\left(-\frac{t^2}{2nb^2}\right).$$

In particular as soon as $t \gg \sqrt{n}b$ this yields a small probability. Now, if for each i we have some doubt δ about the expectation of X_i, these errors sum up and we need to take t larger than $\max(\sqrt{n}b, n\delta)$ before getting an interesting

estimate. In our computation, n corresponds to the radius of the disk for which we are controlling the number variance and the X_i's are discrepancies on small pieces of size ≈ 1, with b of order 1 thanks to the local laws. We see that ideally, one would get an interesting probability estimate as soon as t (the global discrepancy carried by the smaller pieces) exceeds $\sqrt{n}$ - which corresponds the threshold posited by physicists - but unfortunately our uncertainties about the centeredess of discrepancies within each piece will add up and might compromise the entire strategy by raising the threshold to some quantity proportional to n.

Now that we understand why it is important to control the expectation of discrepancies, let us explain why it might *not be easy*. All the examples reviewed in Section 1.2 had their discrepancies exactly centered for one simple reason: they are all infinite, *translation-invariant* point processes. The property of being translation-invariant implies that for additive quantities like (the expectation of) the number of points in domains one can compare small-scale properties to (the average of) large-scale ones and in particular the intensity assumption (1.1) (which is by definition an infinite-volume limit) translates into the fixed lengthscale identity $\frac{1}{|\mathfrak{D}_R|}\mathbb{E}[|\mathbf{X} \cap \mathfrak{D}_R|] = 1$ (for any $R > 0$).

However the "conditional measures" introduced in the previous section are finite systems and cannot be perfectly translation-invariant because of the presence of boundaries, one might however hope that some translation-invariance remains true in the bulk. This yields the question of how to prove (quasi)-invariance properties for Gibbs measures.

6.2 Continuous symmetries in spin systems

Such questions arise naturally in the slightly different context of so-called spin systems with continuous symmetries. The typical example is the situation where the spins (indexed by the lattice $\mathbb{Z}^2$) are *angles* in $\mathbb{S}^1$ (the unit circle) and the interaction energy of a spin configuration is invariant under rotation of all the spins by a same angle. It is then well-known (see e.g. [DS75, FP81]) that "in dimension 2, continuous symmetries cannot be broken" (of course there are minimal assumptions to put for this theorem, known as "Mermin-Wagner" to hold, in particular the interactions must be sufficiently short-ranged), which can mean two things:

1. The infinite-volume Gibbs measures ("Gibbs states" in the sense of the Dobrushin-Lanford-Ruelle (DLR) formalism) are all invariant under the continuous symmetry.

2. The finite-volume Gibbs measures with arbitrary boundary conditions are "almost invariant" under the continuous symmetry.

While the first conclusion is unambiguous, the second one requires a speficic, model-dependent quantified statement.

The classical way to prove the first statement is to construct for an arbitrary angle θ_0 and for each $L \geq 1$, a "spin wave" or "localized rotation" $\theta : \mathbb{Z}^2 \to \mathbb{S}^1$ such that:

- $\theta(x) = \theta_0$ if $|x| \leq L$
- $\theta(x) = 0$ if $|x|$ is large enough,

in such a way that θ has *a bounded energy cost* i.e. that rotating the spin located at x by $\theta(x)$ (for $x \in \mathbb{Z}^2$) changes the interaction energy of the entire spin configuration by some constant *bounded uniformly in* L. This is enough to prove invariance by rotation of all solutions of the DLR equations.

Morally speaking, one might wonder how showing that something (here, the energy cost) is bounded might yield to the conclusion that two things are equal. This is due to the structure of extremal Gibbs states (every Gibbs state can be decomposed as a mixture of extremal ones and it is enough to prove invariance of those ones), which are mutually singular. The following observation is a good toy model of the argument: if μ and ν are two Dirac measures and if we prove that the Radon-Nikodym derivative $\frac{\mathtt{d}\mu}{\mathtt{d}\nu} \leq C$ ("a certain quantity is bounded") then in fact $\mu = \nu$ ("a certain equality holds").

Bounded versus small energy cost When trying to prove almost-invariance of *finite-volume* Gibbs measures however, making a spin wave construction with bounded energy cost yields *almost no information* and a quick computation reveals that in order to derive any sort of invariance, one would need to build localized rotations which have a *small* energy cost: $o(1)$ and not simply $\mathcal{O}(1)$. This surprising difference in the two approaches (infinite or finite-volume) is noted in [Sim14] which mysteriously mentions in passing (as an empirical fact) that in all known cases, when one can construct spin waves with bounded cost *one can also construct spin waves with vanishing energy cost.*

The role of the H^1 norm One informal way to understand B. Simon's remark is to observe that that in "natural" cases the energy cost of a localized rotation θ (or more precisely of the average effect of two localized rotations θ and $-\theta$, which is enough to consider thanks to a useful trick) is given by a quantity of the type $\int_{\mathbb{R}^2} |D\theta(x)|^2 \mathtt{d}x$ (here for convenience let us assume we have extended θ as a smooth function on $\mathbb{R}^2$ instead of just looking at it on $\mathbb{Z}^2$).

In dimension 2, since the H^1 (sometimes denoted by $\dot{W}^{1,2}$) norm of θ is scale-invariant it is not hard to construct spin waves with a cost bounded uniformly in L by simply rescaling a given function. But in fact, since in dimension 2 the Sobolev space $\dot{W}^{1,2}$ does *not* embed into L^∞, it is also possible to find functions with arbitrarily small H^1 norm and which are yet constant (and non zero) over a large box. Note however that those functions have a very slow decay, and in order to make the energy cost smaller than ε requires to "dampen" the function over a domain of diameter $e^{1/\varepsilon}$.

6.3 Almost translation-invariance

The problem of translation-invariance in particle systems is closely related to rotation-invariance in spin systems, in fact the same method of proof can be

employed after replacing "localized rotations" by "localized translations" as described in [FP81, FP86, Geo99] - in which case one seeks to construct $\psi : \mathbb{R}^2 \to \mathbb{R}^2$ such that:

- $\psi(x) = \psi_0$ if $|x| \leq L$,
- $\psi(x) = 0$ if $|x|$ is large enough,

in such a way that translating the particle located at x by $\psi(x)$ (for $x \in \mathbb{R}^2$) changes the interaction energy of the entire point configuration by some constant *bounded uniformly in* L. Note that in this situation one should also require that the map $x \mapsto x + \psi(x)$ has Jacobian equal to 1 (this came for free with localized rotations). This would be enough to show that *infinite-volume* 2DOCP's are translation-invariant, but does not yield anything useful in finite volume.

To get some approximate translation-invariance, one must rely on the observation of B. Simon and find a localized translation with an energy cost that is *small, not merely bounded*. To do that, as explained above, it is enough to show that the dominant term in the energy cost of volume-preserving perturbations of the identity map ($x \mapsto x + \psi(x)$) applied to all particles is given by the Sobolev $\dot{W}^{1,2}$ norm of ψ, which requires a careful revisit of the heavy computations of [Ser23]. One can then use localized translations that are spread on a sufficiently large region - of size $e^{1/\varepsilon}$ in order to get an energy error of size ε, which in means that conversely, in a disk of radius T, one can hope to have approximate translation-invariance near the origin with an "error" of size $\frac{1}{\log T}$ (which has a slow decay with respect to T).

Using this approximate translation-invariance, one can estimate expectations of discrepancies and get something of the type $\mathbb{E}[\mathsf{Dis}(\mathbf{X}, \mathfrak{D}_r)] = o(r)$ with a quantitative but very limited gain compared to the linear bound that one gets from a priori controlling the second moment. This is enough for our purposes, although one could hope to get a much better gain, at least for high enough temperature (in the hypothetical "liquid" phase one would expect a lot of translation-invariance).

7 Conclusion

The result we are able to get is a statement of hyperuniformity in the bulk of the 2DOCP, giving a quantitative bound on the number variance which is negligible with respect to the volume.

Theorem 1. Let $\delta > 0$ be fixed. For all N and R large enough (both depending on β and δ), for all x in Σ_N such that x is "in the bulk" in the following sense:

$$\mathrm{dist}(\mathfrak{D}(x, R), \partial\Sigma_N) \geq \delta\sqrt{N},$$

we have in the disk $\mathfrak{D}(x, R)$ of center x and radius R:

$$\mathbb{P}_N^\beta \left(\left\{ |\mathsf{Dis}(\mathbf{X}_N, \mathfrak{D}(x, R))| \geq R(\log R)^{-0.3} \right\} \right) \leq \exp\left(-\log^{1.5} R \right).$$

Since good exponential tails for Dis at values higher than R are already known (see e.g. [AS21, Thm 1]), this new sub-algebraic tail valid for values of the discrepancy between $R\log^{-0.3} R$ and R imply that:

$$\mathrm{Var}[\mathsf{Dis}(\mathbf{X}_N, \mathfrak{D}(x,R))] \leq \mathbb{E}_N^\beta\left[\mathsf{Dis}^2(\mathbf{X}_N, \mathfrak{D}(x,R))\right] = \mathcal{O}\left(\frac{R^2}{\log^{0.6} R}\right) = o(R^2),$$

so the 2DOCP is thus indeed *hyperuniform*. Recall that the full physical prediction says not only that the number variance is negligible with respect to the area of the disk, but that it should even be *comparable to the perimeter* ($\mathcal{O}(R)$ and not only $o(R^2)$). In conclusion, hyperuniformity does hold at all temperatures, however our upper bound on the number variance remains far from the conjectured sharp estimate.

References

[AS21] Scott Armstrong and Sylvia Serfaty. Local laws and rigidity for coulomb gases at any temperature. *Ann. Probab.*, 49(1):46–121, 2021.

[BBNY17] Roland Bauerschmidt, Paul Bourgade, Miika Nikula, and Horng-Tzer Yau. Local density for two-dimensional one-component plasma. *Comm. Math. Phys.*, 356(1):189–230, 2017.

[BBNY19] Roland Bauerschmidt, Paul Bourgade, Miika Nikula, and Horng-Tzer Yau. The two-dimensional Coulomb plasma: quasi-free approximation and central limit theorem. *Adv. Theor. Math. Phys.*, 23(4):841–1002, 2019.

[Bec87] József Beck. Irregularities of distribution. i. *Acta Mathematica*, 159(1):1–49, 1987.

[DS75] R. L. Dobrushin and S. B. Shlosman. Absence of breakdown of continuous symmetry in two-dimensional models of statistical physics. *Comm. Math. Phys.*, 42(1):31–40, Feb 1975.

[FH99] Peter J Forrester and Graeme Honner. Exact statistical properties of the zeros of complex random polynomials. *J. Phys. A*, 32(16):2961, 1999.

[FL21] Marcel Fenzl and Gaultier Lambert. Precise Deviations for Disk Counting Statistics of Invariant Determinantal Processes. *Int. Math. Res. Not.*., 2022(10):7420–7494, 01 2021.

[FP81] Jürg Fröhlich and Charles Pfister. On the absence of spontaneous symmetry breaking and of crystalline ordering in two-dimensional systems. *Comm. Math. Phys*, 81:277–298, 1981.

[FP86] Jürg Fröhlich and Charles-Edouard Pfister. Absence of crystalline ordering in two dimensions. *Comm. Math. Phys*, 104:697–700, 1986.

[Geo99] Hans-Otto Georgii. Translation invariance and continuous symmetries in two-dimensional continuum systems. In *Mathematical results in statistical mechanics (Marseilles, 1998)*, pages 53–69. World Sci. Publ., River Edge, NJ, 1999.

[Gin65] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.*, 6(3):440–449, 1965.

[GL17] Subhroshekhar Ghosh and Joel L Lebowitz. Fluctuations, large deviations and rigidity in hyperuniform systems: a brief survey. *Indian J. Pure Appl. Math.*, 48(4):609–631, 2017.

[GP17] Subhroshekhar Ghosh and Yuval Peres. Rigidity and tolerance in point processes: Gaussian zeros and ginibre eigenvalues. *Duke Math. J.*, 166(10):1789–1858, 2017.

[GS75] Peter Gacs and Domokos Szász. On a problem of cox concerning point processes in $\mathbb{R}^k$ of "controlled variability". *Ann. Probab.*, 3(4):597–607, 1975.

[HKPV06] J Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. Determinantal processes and independence. *Probab. Surv.*, 3:206–229, 2006.

[JLM93] B Jancovici, Joel L Lebowitz, and G Manificat. Large charge fluctuations in classical coulomb systems. *J. Statist. Phys.*, 72(3-4):773–787, 1993.

[Leb83] Joel Lebowitz. Charge fluctuations in coulomb systems. *Phys. Rev. A*, 27(3):1491, 1983.

[Leb17] Thomas Leblé. Local microscopic behavior for 2D Coulomb gases. *Probab. Theory Related Fields*, 169(3-4):931–976, 2017.

[Lew22] Mathieu Lewin. Coulomb and riesz gases: The known and the unknown. *J. Math. Phys.*, 63(6), 2022.

[LS18] Thomas Leblé and Sylvia Serfaty. Fluctuations of two dimensional Coulomb gases. *Geom. Funct. Anal.*, 28(2):443–508, 2018.

[LWL00] D. Levesque, J.-J. Weis, and Joel Lebowitz. Charge fluctuations in the two-dimensional one-component plasma. *J. Statist. Phys.*, 100(1):209–222, 2000.

[Mar88] Ph A Martin. Sum rules in charged fluids. *Rev. Modern Phys.*, 60(4):1075, 1988.

[MY80] Ph. A. Martin and T. Yalcin. The charge fluctuations in classical coulomb systems. *J. Statist. Phys.*, 22(4):435–463, 1980.

[NSV08] Fedor Nazarov, Mikhail Sodin, and Alexander Volberg. The Jancovici–Lebowitz–Manificat law for large fluctuations of random complex zeroes. *Comm. Math. Phys.*, 284(3):833–865, 2008.

[RS16] Nicolas Rougerie and Sylvia Serfaty. Higher-dimensional coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.*, 69(3):519–605, 2016.

[Ser18] Sylvia Serfaty. Systems of points with coulomb interactions. *Eur. Math. Soc. Newsl*, 110:16–21, 2018.

[Ser23] Sylvia Serfaty. Gaussian fluctuations and free energy expansion for Coulomb gases at any temperature. *Ann. Inst. H. Poincaré Probab. Statist.*, 59(2):1074 – 1142, 2023.

[Shi06] Tomoyuki Shirai. Large deviations for the fermion point process associated with the exponential kernel. *J. Statist. Phys.*, 123(3):615–629, 2006.

[Sim14] Barry Simon. The statistical mechanics of lattice gases, volume i. In *The Statistical Mechanics of Lattice Gases, Volume I*. Princeton University Press, 2014.

[SS15] Etienne Sandier and Sylvia Serfaty. 2D Coulomb gases and the renormalized energy. *Ann. Probab.*, 43(4):2026–2083, 2015.

[SZ08] Bernard Shiffman and Steve Zelditch. Number variance of random zeros on complex manifolds. *Geom. Funct. Anal.*, 18(4):1422–1475, 2008.

[Tor18] Salvatore Torquato. Hyperuniform states of matter. *Phys. Rep.*, 745:1–95, 2018.

[TS03] Salvatore Torquato and Frank H Stillinger. Local density fluctuations, hyperuniformity, and order metrics. *Phys. Rev. E*, 68(4):041113, 2003.