Rigidity of the three-dimensional hierarchical Coulomb gas

Sourav Chatterjee[1]

Abstract

A random set of points in Euclidean space is called 'rigid' or 'hyperuniform' if the number of points falling inside any given region has significantly smaller fluctuations than the corresponding number for a set of i.i.d. random points. This phenomenon has received considerable attention in recent years, due to its appearance in random matrix theory, the theory of Coulomb gases and zeros of random analytic functions. However, most of the published results are in dimensions one and two. This paper gives the first proof of hyperuniformity in a Coulomb type system in dimension three, known as the hierarchical Coulomb gas. This is a simplified version of the actual 3D Coulomb gas. The interaction potential in this model, inspired by Dyson's hierarchical model of the Ising ferromagnet, has a hierarchical structure and is locally an approximation of the Coulomb potential. Hyperuniformity is proved at both macroscopic and microscopic scales, with upper and lower bounds for the order of fluctuations that match up to logarithmic factors. The fluctuations have cube-root behavior, in agreement with a well-known prediction for the 3D Coulomb gas. For completeness, analogous results are also proved for the 2D hierarchical Coulomb gas and the 1D hierarchical log gas.

✉ Sourav Chatterjee
souravc@stanford.edu

1   Department of Statistics, Stanford University, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305, USA

1 Introduction and results

1.1 Interacting gases

The probability density of n independent and identically distributed points in $\mathbb{R}^d$ can always be represented as

$$\frac{1}{Z} \exp\left(-\beta \sum_{i=1}^{n} V(x_i)\right)$$

where V is some real-valued function on $\mathbb{R}^d$, β is some positive parameter, and Z is the normalizing constant.

Suppose that we want to introduce some interactions between the points. The simplest way to do that is to introduce a pairwise interaction term in the exponent; the new density is of the form

$$\frac{1}{Z} \exp\left(-\beta \sum_{1 \leq i < j \leq n} w(x_i, x_j) - \beta n \sum_{i=1}^{n} V(x_i)\right),$$

where w is a symmetric real-valued function on $\mathbb{R}^d \times \mathbb{R}^d$, known as the interaction potential. The factor n is put in front of the second term to ensure that the two terms are of comparable size, which is necessary for ensuring that the system has nontrivial properties in the large n limit. A particularly important type of interaction potentials are the Coulomb potentials, defined as

$$w(x, y) = \begin{cases} |x - y| & \text{if } d = 1, \\ -\log |x - y| & \text{if } d = 2, \\ |x - y|^{2-d} & \text{if } d \geq 3, \end{cases}$$

where $|x - y|$ is the Euclidean distance between x and y. With w as above and $V(x) = |x|^2$, we get the so-called Coulomb gases.

The 1D Coulomb gas is a very well-understood exactly solvable system, studied thoroughly by physicists [37,60,66,67] and mathematicians [1,29]. In higher dimensions, much less is known. For $\beta = 1$, the 2D Coulomb gas is an exactly solvable model due to its relationship with the Ginibre ensemble of random matrices [49]. The Ginibre ensemble has received widespread attention from mathematicians [2,3,18,26,27,47,50,80,87]. For general β, however, the 2D Coulomb gas has no representation as an exactly solvable model. Fortunately, a number of results are now known for the case of general β. Large deviation principles for the 2D Coulomb gas were proved in [11,52,78], and extended to general dimensions in [34,83]. Concentration inequalities were proved in [33] and dynamical properties have been recently studied in [17]. The ground state in a related model was studied in [79]. Local properties have been studied in great depth in the recent papers [6,7,62,64,82].

In dimensions three and higher, very precise information about the normalizing constants has been obtained in [63,81]. For a comprehensive survey, see [84]. Further results are provable by the techniques of these papers but have not been written up yet, as I learned from Sylvia Serfaty in a personal communication.

Another widely studied example is the 1D log gas, where $d = 1$, $w(x, y) = -\log|x - y|$ and $V(x) = x^2$. For $\beta = 1, 2$ and 4, the log gases arise as eigenvalues of various random matrix ensembles and are exactly solvable. Precise fluctuation estimates for these special values of β were obtained in [57]. There is now considerable information available about other values of β and more general V [10,23–25,28,85]. Asymptotic series expansions for the normalizing constants were computed in [20–22]. Central limit theorems have been investigated in [9,19,61]. For an introduction to log gases and their connections with random matrices, see [4,36,42]. A recent survey is given in [6].

1.2 Hyperuniformity

If we have a collection of n independent and identically distributed points in $\mathbb{R}^d$, then the number of points that fall in a given set has fluctuations of order $n^{1/2}$ as $n \to \infty$. If a random point process has the property that this order of fluctuations is $o(n^{1/2})$, then it is called 'rigid' or 'hyperuniform'. More generally, a point process is called hyperuniform if its empirical measure has smaller fluctuations than the empirical measure of a collection of i.i.d. random points. In this paper I will use the terms 'rigidity' and 'hyperuniformity' interchangeably, but in general hyperuniformity is probably a more suitable term for the phenomenon described above, since rigidity has also been used to mean other things in the literature.

Sometimes point processes are very rigid, such as eigenvalues of various random matrix ensembles, for which the order of fluctuations may be as small as $O(\sqrt{\log n})$ or even $O(1)$ if one considers integrals of the empirical measure with respect to smooth functions. Rigidity/hyperuniformity has been established for many processes in dimensions one and two. For example, rigidity of the eigenvalues of random unitary matrices was proved in [38,91]. Rigidity of eigenvalues in the standard hermitian random matrix ensembles follow from the results of [23–25,35,43,76,88]. Rigidity of eigenvalues of non-hermitian random matrices has been studied in [18,26,27,47,89]. Another class of 2D processes that exhibit hyperuniformity are zeros of random analytic functions. This has been investigated in [44,45,47,48,54,73–75]. In recent work, rigidity of the 2D Coulomb gas has been established in [6,7,62].

However, no such results for interacting particle systems of the above kind are known in dimensions three and higher. The only random point process which has been shown to be hyperuniform in any dimension $d \geq 3$, as far as I know, is the point process obtained by giving i.i.d. random perturbations to the vertices of $\mathbb{Z}^d$. This is a recent result [77], improving on an earlier work in $d \leq 2$ [53]. The notion of rigidity in these papers is somewhat different than hyperuniformity. For interacting systems such as Coulomb gases, the detailed information about the normalizing constants obtained in [81,84] provide some control on the order of fluctuations in $d \geq 3$, but do not establish that the order of fluctuations is smaller than $n^{1/2}$. There is a remark in [62]

that the 2D techniques of that paper can be extended to higher dimensions for proving rigidity of integrals of smooth functions with respect to the empirical measure of a Coulomb gas, but the details have not yet been written up.

A class of processes that are related to Coulomb gases in dimension one but not in higher dimensions, are the so-called orthogonal polynomial ensembles (see [59] for a survey). These are generalizations of the 1D determinantal point processes arising in random matrix theory, and have nice mathematical structures that allow various exact calculations. A general central limit theorem for orthogonal polynomial ensembles was proved in [86]. Rigidity/hyperuniformity for orthogonal polynomial ensembles (beyond random matrix eigenvalues) have been investigated in recent years, for example in [16,30,31,56] in dimensions one and two, and [5] in dimension three.

There is a considerable amount of work by physicists on hyperuniformity. For example, [72] and [71] give physics proofs of hyperuniformity in 3D Coulomb systems. A non-rigorous computation of covariances in Coulomb systems in all dimensions greater than one was given in [65]. More recently, a physics proof of hyperuniformity of free fermions at zero temperature (a certain kind of determinantal point process) was given in [32] in $d \leq 3$, based on an asymptotic formula for the variance of the number of points falling in a given region. This formula was later extended to arbitrary dimensions in [90]. Similar formulas have been very recently obtained for the 1D log gas (with general β and special V) in [69,70]. For an extensive list of references to the physics literature, see the recent survey [46].

1.3 The hierarchical Coulomb gas model

In this paper, we consider a model of an interacting gas of n particles in the 3D unit cube $[0, 1]^3$, which have joint probability density

$$\frac{1}{Z(n, \beta)} \exp \left(-\beta \sum_{1 \leq i < j \leq n} w(x_i, x_j) \right), \tag{1.1}$$

where $w(x, y)$ is a symmetric potential that behaves like the Coulomb potential $|x - y|^{-1}$ at short distances, and $Z(n, \beta)$ is the normalizing constant. The potential w is defined as follows.

The unit cube in $\mathbb{R}^3$ can be partitioned into 8 sub-cubes of side-length 1/2. Each of these sub-cubes can be further partitioned into 8 sub-cubes of side-length 1/4, and so on, generating a tree of dyadic sub-cubes. For any two distinct points x and y in the unit cube, let $w(x, y) = 2^k$, where k is the smallest number such that x and y belong to distinct dyadic sub-cubes of side-length 2^{-k}. There may be some ambiguity about points on the boundaries of the cubes, but since they form a set of measure zero, they do not matter. This w is our potential, which defines our point process through the density (1.1). Note that w is symmetric but not translation invariant.

For typical x and y which are close together, $w(x, y)$ behaves like a multiple of the Coulomb potential $|x - y|^{-1}$. Indeed, it is not hard to prove that there is a constant C such that for all x and y,

$$w(x, y) \leq \frac{C}{|x - y|}.$$

Conversely, there is a constant $c > 0$ such that for any $0 < \delta < 1$, the average value of $w(x, y)$ over all pairs (x, y) with $|x - y| = \delta$ is bounded below by c/δ.

Replacing the Euclidean distance by a hierarchical distance as above is a famous idea of Dyson [40,41], who formulated and analyzed a hierarchical version of the 1D Ising model with long range interactions. This is now known as 'Dyson's hierarchical model'. Dyson's work has inspired a large body of literature on hierarchical models over the years, and is still an active area of research. The model proposed above is sometimes called the 'hierarchical Coulomb gas'. The 2D hierarchical Coulomb gas has received considerable attention in the mathematical physics literature [14,15,39, 51,58,68]. However, not much is known about this model in dimensions three and higher.

Just as the Coulomb potential is the Green's function for Brownian motion, the potential w can also be realized as the Green's function of a certain continuous time random walk on the unit cube, following a method developed in [12,13] for constructing Markov semigroups on ultrametric spaces. More generally, the prescription given in [12,13] can be used for a large class of hierarchical potentials arising from Dyson-type constructions.

The chief reason why the hierarchical structure of the potential helps in the analysis is that it does an automatic 'coarse-graining' of the interactions. The total interaction between the particles in two disjoint dyadic cubes is determined solely by the numbers of particles in those cubes, rather than their exact locations.

One of our main results, stated in the next subsection, is that if U is a nonempty open subset of the unit cube with a nicely behaved boundary, then the number of points falling in U has fluctuations of order at most $n^{1/3}\sqrt{\log n}$, thereby establishing the hyperuniformity of our point process. This is matched up to a logarithmic factor by a lower bound of order $n^{1/3}$. We also establish microscopic hyperuniformity in a local neighborhood of any given point. Finally, the analogous results in dimensions one and two are established for the sake of completeness.

1.4 Results in 3D

Take any $d \geq 1$, and let U be a nonempty open subset of $\mathbb{R}^d$. Let ∂U be the boundary of U. For each $\epsilon > 0$, let ∂U_ϵ be the set of all points that are at distance $\leq \epsilon$ from ∂U. Let $\mathrm{diam}(U)$ denote the diameter of U. We will say that the boundary of U is *regular* if there is some constant C such that for all $0 < \epsilon \leq \mathrm{diam}(U)$,

$$\mathrm{Leb}(\partial U_\epsilon) \leq C\epsilon, \tag{1.2}$$

where Leb stands for Lebesgue measure.

Now let $d = 3$, and let U be a nonempty open subset of $[0, 1]^3$ whose boundary is regular in the sense defined above. Take any $n \geq 2$ and $\beta > 0$, and consider an interacting gas of n particles behaving according to the model defined above. Let $N(U)$ be the number of particles that fall in U. Our first theorem says that the gas

is macroscopically hyperuniform in the sense that $N(U)$ has fluctuations of order at most $n^{1/3}\sqrt{\log n}$, instead of $n^{1/2}$ as would be the case for a gas of i.i.d. particles.

Theorem 1.1 (Macroscopic hyperuniformity in 3D) *Let U and $N(U)$ be as above. Then*

$$\mathbb{E}(N(U)) = Leb(U)n$$

and

$$\mathrm{Var}(N(U)) \leq C(U, \beta)n^{2/3}\log n,$$

where $C(U, \beta)$ is a constant that depends only on U and β.

The next theorem shows that when ∂U is smooth, $n^{1/3}$ is actually the correct order of fluctuations of $N(U)$, up to possible logarithmic corrections.

Theorem 1.2 (Lower bound in 3D) *Let U be a nonempty connected open subset of $[0, 1]^3$ whose boundary is a smooth, closed, orientable surface. Let $N(U)$ be as in Theorem 1.1. Then $N(U)$ has fluctuations of order at least $n^{1/3}$, in the sense that there are three constants $n_0 \geq 1$, $c_1 > 0$ and $c_2 < 1$, depending only on U and β, such that for any $n \geq n_0$ and any $-\infty < a \leq b < \infty$ with $b - a \leq c_1 n^{1/3}$, we have $\mathbb{P}(a \leq N(U) \leq b) \leq c_2$.*

Incidentally, the $n^{1/3}$ order of fluctuations matches a well-known prediction from physics [55,65,72] for the 3D Coulomb gas model (see also [75]). The 1/3 exponent is also reminiscent of a famous classical result [8] about irregularities in distributions of arbitrary sequences of points in Euclidean space.

Let us now turn our attention to hyperuniformity in the microscopic scale. Take any point $x \in (0, 1)^3$. Blow up the neighborhood of x by a factor of $n^{1/3}$ by applying the blow-up map $y \mapsto n^{1/3}(y - x)$ to the points in our interacting gas. Since the original process had an expected density of n particles per unit volume, the new process has an expected density of one particle per unit volume. Studying the blown up process is the standard way of investigating the local behavior of interacting gases [84].

Let U be a nonempty open subset of $\mathbb{R}^3$ whose boundary is regular in the sense defined above. For each $\lambda > 0$, let λU denote the set $\{\lambda y : y \in U\}$, and let $N_x(\lambda U)$ be the number of points from the blown up process that land in λU. The following theorem shows that for $\lambda \gg 1$, $N_x(\lambda U)$ has fluctuations of order at most $\lambda\sqrt{\log \lambda}$. This is smaller than $\lambda^{3/2}$, the corresponding order of fluctuations for a Poisson point process. This proves the hyperuniformity of our interacting gas at the microscopic scale.

Theorem 1.3 (Microscopic hyperuniformity in 3D) *Let U and $N_x(\lambda U)$ be as above. Then for any λ such that $diam(\lambda U) \geq 1$,*

$$\lim_{n\to\infty} \mathbb{E}(N_x(\lambda U)) = Leb(\lambda U) = \lambda^3 Leb(U),$$

and

$$\limsup_{n \to \infty} \mathrm{Var}(N_x(\lambda U)) \leq C(U, \beta) \lambda^2 \log(4\lambda \, diam(U)),$$

where $C(U, \beta)$ is a constant that depends only on U and β.

Theorems 1.1 and 1.3 are both special cases of a more general theorem (Theorem 2.13 in Sect. 2.4), which gives hyperuniformity at all scales.

Finally, let us consider linear statistics. Any function $f : [0, 1]^3 \to \mathbb{R}$ defines a linear statistic

$$X(f) := \sum_{i=1}^{n} f(X_i), \tag{1.3}$$

where $X_1, \ldots, X_n$ is a realization of our point process. In particular, $N(U)$ is a linear statistic, with f being the indicator function of U. We have the following two theorems about fluctuations of linear statistics when f is continuous. The results are not as definitive as the other results of this section, since the upper and lower bounds do not match.

If f is Lipschitz, we get the following slight improvement of the bound given in Theorem 1.1.

Theorem 1.4 (Upper bound for linear statistics in 3D) *Suppose that $f : [0, 1]^3 \to \mathbb{R}$ is a Lipschitz function with Lipschitz constant L. Let $X_1, \ldots, X_n$ be a realization of points from our model in dimension two. Let $X(f)$ be the linear statistic defined in (1.3). Then*

$$\mathrm{Var}(X(f)) \leq C(\beta)L^2 n^{2/3},$$

where $C(\beta)$ is a constant that depends only on β.

The next theorem gives a lower bound of order $n^{1/6}$ on the order of fluctuations of $X(f)$ when f is a non-constant linear function. This does not match the upper bound from Theorem 1.4, but is nonetheless growing polynomially in n, deviating from the $O(1)$ rate for smooth linear statistics in dimensions one and two [6,7,23–27,35,38,57,62,88,91].

Theorem 1.5 (Lower bound for linear statistics in 3D) *Let $f : [0, 1]^3 \to \mathbb{R}$ be a non-constant linear function, and let $X(f)$ be as in (1.3). Then $X(f)$ has fluctuations of order at least $n^{1/6}$, in the sense that there are three constants $n_0 \geq 1$, $c_1 > 0$ and $c_2 < 1$, depending only on U and β, such that for any $n \geq n_0$ and any $-\infty < a \leq b < \infty$ with $b - a \leq c_1 n^{1/6}$, we have $\mathbb{P}(a \leq X(f) \leq b) \leq c_2$.*

It is not clear whether $n^{1/3}$ or $n^{1/6}$ is the correct order of fluctuations for smooth linear statistics. Theorem 1.2 does not provide any strong evidence in favor of $n^{1/3}$, because, as we will see later for the 2D hierarchical Coulomb gas, linear statistics of smooth functions may have much smaller fluctuations than linear statistics of indicator

functions. However, there is a recent result [5] which shows that $n^{1/3}$ is the correct order of fluctuations for smooth linear statistics of a 3D orthogonal polynomial ensemble. Although orthogonal polynomial ensembles are not related to Coulomb type systems in dimension three, this gives some support in favor of $n^{1/3}$.

1.5 Results in 2D and 1D

In dimension two, we will modify w to mimic the logarithmic potential of the 2D Coulomb gas. This is done by declaring $w(x, y) =$ the minimum k such that x and y belong to distinct dyadic sub-squares of $[0, 1]^2$ of side-length 2^{-k}. We will use the same formula in dimension one as well (with dyadic intervals instead of squares), so that w mimics the logarithmic potential of 1D log gases. With these modifications, we have the following analogs of Theorem 1.1. With $N(U)$ as in Theorem 1.1, it says that $N(U)$ has fluctuations of order at most $n^{1/4} \log n$ in dimension two, and $\log n$ in dimension one.

Theorem 1.6 (Macroscopic hyperuniformity in 2D and 1D) *Consider the model defined above in dimension $d = 1$ or 2. Let U and $N(U)$ be as in Theorem 1.1. Then*

$$\mathbb{E}(N(U)) = Leb(U)n$$

and

$$\mathrm{Var}(N(U)) \le C(U, \beta)n^{(d-1)/d}(\log n)^2,$$

where $C(U, \beta)$ is a constant that depends only on U and β.

The following theorem shows that in dimension two, $N(U)$ has fluctuations of order at least $n^{1/4}$, matching the above upper bound up to a logarithmic factor.

Theorem 1.7 (Lower bound in 2D) *Let U be a nonempty connected open subset of $[0, 1]^2$ whose boundary is a simple, smooth, closed curve. Let $N(U)$ be as in Theorem 1.6. Then $N(U)$ has fluctuations of order at least $n^{1/4}$, in the sense that there are three constants $n_0 \ge 1$, $c_1 > 0$ and $c_2 < 1$, depending only on U and β, such that for any $n \ge n_0$ and any $-\infty < a \le b < \infty$ with $b - a \le c_1 n^{1/4}$, we have $\mathbb{P}(a \le N(U) \le b) \le c_2$.*

Like the $n^{1/3}$ rate in the 3D case, the $n^{1/4}$ rate was also predicted in the physics literature [55,65,72] for the 2D Coulomb gas. The $n^{1/4}$ fluctuation in the special case of $\beta = 1$ in the 2D Coulomb gas (corresponding to the exactly solvable Ginibre ensemble) can be established by standard techniques, as I learned from Paul Bourgade in a personal communication.

We also have the following analog of Theorem 1.3. With $N_x(\lambda U)$ as in Theorem 1.3, it shows that for $\lambda \gg 1$, $N_x(\lambda U)$ has fluctuations of order at most $\lambda^{1/2} \log \lambda$ in dimension two, and $\log \lambda$ in dimension one.

Theorem 1.8 (Microscopic hyperuniformity in 2D and 1D) *Consider the model defined above in dimension $d = 1$ or 2. Let U and $N_x(\lambda U)$ be as in Theorem* 1.3. *Then for any λ such that $diam(\lambda U) \geq 1$,*

$$\lim_{n \to \infty} \mathbb{E}(N_x(\lambda U)) = Leb(\lambda U) = \lambda^d Leb(U),$$

and

$$\limsup_{n \to \infty} Var(N_x(\lambda U)) \leq C(U, \beta)\lambda^{d-1}(\log(7\lambda^d diam(U)^d))^2,$$

where $C(U, \beta)$ is a constant that depends only on U and β.

As before, Theorems 1.6 and 1.8 are special cases of a more general theorem (Theorem 3.10 in Sect. 3.4) that gives hyperuniformity at all scales.

Finally, let us consider linear statistics. It has been proved recently in [6,7,62] that for the 2D Coulomb gas, linear statistics of smooth functions have $O(1)$ fluctuations. For Lipschitz f, the following theorem shows that for our model in dimension two, the fluctuations of $X(f)$ are at most of order $(\log n)^{3/2}$ instead of $n^{1/4}$. Unlike Theorem 1.4, this is a big improvement of the bound from Theorem 1.6, and is within a logarithmic factor of the $O(1)$ bound from [6,7,62].

Theorem 1.9 (Upper bound for linear statistics in 2D and 1D) *Let $d = 1$ or 2. Suppose that $f : [0, 1]^d \to \mathbb{R}$ is a Lipschitz function with Lipschitz constant L. Let $X_1, \ldots, X_n$ be a realization of points from our model in dimension d, and let $X(f)$ be the linear statistic defined in* (1.3). *Then*

$$Var(X(f)) \leq C(\beta)L^2(\log n)^{d+1},$$

where $C(\beta)$ is a constant that depends only on β.

2 Proofs in 3D

The rest of this paper is devoted to proofs. In this section, we will prove the theorems of Sect. 1.4.

2.1 Notation

It is helpful to define some precise notations and terminologies. For a slight technical convenience, we will replace the unit cube by the half-open unit cube $[0, 1)^3$. Clearly, this will not alter the conclusions.

A dyadic sub-interval of the half-open unit interval $[0, 1)$ is an interval of the form $[i2^{-k}, (i + 1)2^{-k})$, where $k \geq 0$ and $0 \leq i \leq 2^k - 1$. A dyadic sub-cube of the half-open unit cube $[0, 1)^3$ is a sub-cube of the form $I_1 \times I_2 \times I_3$, where I_1, I_2 and

I_3 are dyadic sub-intervals of $[0, 1)$ of equal length. Let $\mathcal{D}_k$ be the set of all dyadic sub-cubes of $[0, 1)^3$ of side-length 2^{-k}, and let

$$\mathcal{D} := \bigcup_{k=0}^{\infty} \mathcal{D}_k$$

be the set of all dyadic sub-cubes of $[0, 1)^3$. Then $\mathcal{D}$ has a natural tree structure, with each node having 8 children. We will freely use the terms 'child', 'parent', 'ancestor' and 'descendant' with respect to this tree.

For any two distinct points $x, y \in [0, 1)^3$, let $k(x, y)$ be the smallest k such that x and y belong to distinct elements of $\mathcal{D}_k$. Then our potential w is the function $w(x, y) = 2^{k(x,y)}$. For $x = y$, let $w(x, y) = \infty$.

For each $n \geq 2$, let Σ_n be the set of all n-tuples of points from $[0, 1)^3$. Define the energy of a configuration $(x_1, \ldots, x_n) \in \Sigma_n$ as

$$H_n(x_1, \ldots, x_n) := \sum_{1 \leq i < j \leq n} w(x_i, x_j).$$

For $\beta > 0$, let $\mu_{n,\beta}$ be the probability measure on Σ_n that has density

$$\frac{1}{Z(n, \beta)} e^{-\beta H_n(x_1, \ldots, x_n)}$$

with respect to Lebesgue measure on Σ_n, where $Z(n, \beta)$ is the normalizing constant. The measure $\mu_{n,\beta}$ defines our model of an interacting gas at inverse temperature β.

For certain technical reasons, we will also define the model for $n = 0$ and $n = 1$. When $n = 0$, there are no points. When $n = 1$, there is one point which is uniformly distributed in the cube. We will let $Z(0, \beta) = Z(1, \beta) = 1$ for any β.

2.2 Preliminary calculations

In the following, all integrals are over $[0, 1)^3$ and all double integrals are over $[0, 1)^3 \times [0, 1)^3$, unless otherwise specified.

Lemma 2.1 *For each $x \in [0, 1)^3$,*

$$\int w(x, y) \, dy = \frac{7}{3}.$$

Consequently,

$$\iint w(x, y) \, dx \, dy = \frac{7}{3}.$$

Proof Take any x. For each k, let D_k be the element of $\mathcal{D}_k$ that contains x. It is easy to see that the set of all y with $w(x, y) = 2^k$ is exactly the union of all members of

$\mathcal{D}_k$ that are contained in D_{k-1}, except the one that contains x. The Lebesgue measure of this set is $8^{-k} \cdot 7$. Thus,

$$\int w(x, y)\, dy = 7 \sum_{k=1}^{\infty} 2^k 8^{-k} = \frac{7}{3}.$$

The second assertion is obvious from the first. □

Let us now investigate energy-minimizing configurations of finite size. Henceforth, L_n will denote the minimum possible energy of a configuration of n points. The following result gives upper and lower bounds for L_n.

Theorem 2.2 *There is a positive constant C_1 such that for each $n \geq 2$,*

$$\binom{n}{2}\frac{7}{3} - C_1 n^{4/3} \leq L_n \leq \binom{n}{2}\frac{7}{3}.$$

Proof Let $Y_1, \ldots, Y_n$ be i.i.d. uniform random points from $[0, 1)^3$. Then by symmetry,

$$L_n \leq \mathbb{E}(H_n(Y_1, \ldots, Y_n)) = \sum_{1 \leq i < j \leq n} \mathbb{E}(w(Y_i, Y_j)) = \binom{n}{2}\mathbb{E}(w(Y_1, Y_2)).$$

By Lemma 2.1, $\mathbb{E}(w(Y_1, Y_2)) = 7/3$. This proves the upper bound. For the lower bound, let k be an integer such that

$$n^{-1/3} \leq 2^{-k} \leq 2n^{-1/3}.$$

Take any configuration of n points. For each $D \in \mathcal{D}$, let n_D be the number of points in D. Summing up the contributions to the energy from each cube, it is not difficult to see that

$$H_n(x_1, \ldots, x_n) = \sum_{j=1}^{\infty} \sum_{D \in \mathcal{D}_j} 2^j \binom{n_D}{2} + 2\binom{n}{2} \geq \sum_{j=1}^{k} \sum_{D \in \mathcal{D}_j} 2^j \binom{n_D}{2} + 2\binom{n}{2}$$

$$= \sum_{j=1}^{k} \sum_{D \in \mathcal{D}_j} 2^{j-1} n_D^2 - \sum_{j=1}^{k} 2^{j-1} n + 2\binom{n}{2}.$$

By the Cauchy–Schwarz inequality, for each j,

$$\sum_{D \in \mathcal{D}_j} n_D^2 \geq \frac{1}{|\mathcal{D}_j|}\left(\sum_{D \in \mathcal{D}_j} n_D\right)^2 = \frac{n^2}{8^j}.$$

Thus,

$$H_n(x_1, \ldots, x_n) \geq \frac{n^2}{2} \sum_{j=1}^{k} 4^{-j} - n^{4/3} + 2\binom{n}{2} = \frac{n^2}{6}(1 - 4^{-k}) - n^{4/3} + 2\binom{n}{2}$$

$$\geq \frac{n^2}{6}(1 - 4n^{-2/3}) - n^{4/3} + 2\binom{n}{2}.$$

Since this lower bound holds for any configuration of n points, this completes the proof.   □

2.3 Estimates for the partition function

The following lemma gives important information about the ratio $Z(n+1, \beta)/Z(n, \beta)$. Theorem 2.2 is a crucial ingredient in the proof of this lemma. Recall that $Z(0, \beta) = Z(1, \beta) = 1$. For a measurable function $f : \Sigma_n \to \mathbb{R}$, we will denote its expected value under $\mu_{n,\beta}$ by $\mu_{n,\beta}(f)$.

Lemma 2.3 *There is a constant C_2 such that for any $n \geq 0$ and $\beta > 0$,*

$$e^{-7\beta n/3} \leq \frac{Z(n+1, \beta)}{Z(n, \beta)} \leq e^{-7\beta n/3 + C_2 \beta n^{1/3}}.$$

Proof First suppose that $n \geq 2$. For $x_1, \ldots, x_n, x_{n+1} \in [0, 1)^3$, let

$$f_n(x_1, \ldots, x_n, x_{n+1}) := \sum_{i=1}^{n} w(x_i, x_{n+1}),$$

so that

$$H_{n+1}(x_1, \ldots, x_{n+1}) = f_n(x_1, \ldots, x_n, x_{n+1}) + H_n(x_1, \ldots, x_n).$$

By the above representation and Jensen's inequality,

$$\frac{Z(n+1, \beta)}{Z(n, \beta)} = \iint e^{-\beta f_n(x_1, \ldots, x_n, x_{n+1})} \, dx_{n+1} \, d\mu_{n,\beta}(x_1, \ldots, x_n)$$

$$\geq \exp\left(-\beta \iint f_n(x_1, \ldots, x_n, x_{n+1}) \, dx_{n+1} \, d\mu_{n,\beta}(x_1, \ldots, x_n)\right).$$

But by Lemma 2.1,

$$\iint f_n(x_1, \ldots, x_n, x_{n+1}) \, dx_{n+1} \, d\mu_{n,\beta}(x_1, \ldots, x_n)$$

$$= \sum_{i=1}^{n} \iint w(x_i, x_{n+1}) \, dx_{n+1} \, d\mu_{n,\beta}(x_1, \ldots, x_n)$$

$$= \sum_{i=1}^{n} \int \frac{7}{3} \, d\mu_{n,\beta}(x_1, \ldots, x_n) = \frac{7n}{3}.$$

This gives the desired lower bound. Next, note that

$$\frac{Z(n, \beta)}{Z(n+1, \beta)} = \mu_{n+1,\beta}(e^{\beta f_n(x_1, \ldots, x_n, x_{n+1})}).$$

Therefore by Jensen's inequality and the invariance of $\mu_{n+1,\beta}$ under permutations of coordinates,

$$\frac{Z(n, \beta)}{Z(n+1, \beta)} \geq \exp(\beta \mu_{n+1,\beta}(f(x_1, \ldots, x_{n+1}))) = \exp(\beta n \mu_{n+1,\beta}(w(x_1, x_{n+1})))$$

$$= \exp\left(\frac{\beta n}{\binom{n+1}{2}} \sum_{1 \leq i < j \leq n+1} \mu_{n+1,\beta}(w(x_i, x_j))\right)$$

$$= \exp\left(\frac{\beta n}{\binom{n+1}{2}} \mu_{n+1,\beta}(H_{n+1}(x_1, \ldots, x_{n+1}))\right).$$

But by Theorem 2.2,

$$\mu_{n+1,\beta}(H_{n+1}(x_1, \ldots, x_{n+1})) \geq L_{n+1} \geq \frac{7}{3}\binom{n+1}{2} - C_1(n+1)^{4/3}.$$

This gives the required upper bound and completes the proof of the lemma for $n \geq 2$. When $n = 0$, the bounds hold trivially. When $n = 1$, the lower bound follows from an application of Jensen's inequality and Lemma 2.1. The upper bound can be forced to hold for $n = 1$ by choosing C_2 sufficiently large.   □

Lemma 2.3 is iterated to obtain the following corollary.

Corollary 2.4 *For any $n \geq 0$, $\beta > 0$, and any $k \geq -n$,*

$$\frac{Z(n+k, \beta)}{Z(n, \beta)} \leq \exp\left(-\frac{7\beta nk}{3} - \frac{7\beta k(k-1)}{6} + C_2\beta|k|(n+|k|)^{1/3}\right),$$

where C_2 is the constant from Lemma 2.3.

Proof First suppose that $k \geq 0$. By the upper bound from Lemma 2.3,

$$
\begin{aligned}
\frac{Z(n+k, \beta)}{Z(n, \beta)} &= \prod_{i=0}^{k-1} \frac{Z(n+i+1, \beta)}{Z(n+i, \beta)} \\
&\leq \prod_{i=0}^{k-1} \exp\left(-\frac{7\beta(n+i)}{3} + C_2\beta(n+i)^{1/3}\right) \\
&\leq \exp\left(-\frac{7\beta nk}{3} - \frac{7\beta k(k-1)}{6} + C_2\beta k(n+k)^{1/3}\right).
\end{aligned}
$$

Next, suppose that $k < 0$. Let $l = |k|$. Then by the lower bound from Lemma 2.3,

$$
\begin{aligned}
\frac{Z(n+k, \beta)}{Z(n, \beta)} &= \prod_{i=0}^{l-1} \frac{Z(n-i-1, \beta)}{Z(n-i, \beta)} \leq \prod_{i=0}^{l-1} \exp\left(\frac{7\beta(n-i-1)}{3}\right) \\
&= \exp\left(\frac{7\beta nl}{3} - \frac{7\beta l(l+1)}{6}\right).
\end{aligned}
$$

To complete the proof, note that $l = -k$ and $l(l+1) = k(k-1)$. $\qquad\square$

2.4 Proofs of the upper bounds

Let us now fix some $n \geq 0$ and $\beta > 0$. In the following, $(X_1, \ldots, X_n)$ will denote a random configuration drawn from the measure $\mu_{n,\beta}$. We will assume that $(X_1, \ldots, X_n)$ is defined on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation, variance and covariance with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$, Var and Cov respectively.

Lemma 2.5 *Let $D_1, \ldots, D_8$ denote the 8 elements of $\mathcal{D}_1$, and for each $1 \leq i \leq 8$, let $N_i := |\{j : X_j \in D_i\}|$. Then for each i, $\mathbb{E}(N_i) = n/8$ and*

$$
\mathrm{Var}(N_i) \leq K(\beta)n^{2/3},
$$

where $K(\beta)$ is a non-increasing function of β.

Proof We have already defined universal constants C_1 and C_2 in the previous subsections. In this proof, we will continue to use this convention and denote further universal constants by $C_3, C_4, \ldots$ without explicitly mentioning that they denote universal constants on each occasion.

The identity $\mathbb{E}(N_i) = n/8$ follows by symmetry. We will now prove the claimed bound on the variance. The cases $n = 0$ and $n = 1$ are trivial, so let us assume that $n \geq 2$. First, note that energy of a configuration is the sum of the energies within each D_i, plus the interactions between the D_i's. From this observation it is easy to deduce

the recursive relation

$$Z(n, \beta) = \sum_{\substack{0 \le n_1, \ldots, n_8 \le n \\ n_1 + \cdots + n_8 = n}} \frac{n!}{n_1! n_2! \ldots n_8!} e^{-2\beta \sum_{1 \le i < j \le 8} n_i n_j} \prod_{i=1}^{8} (8^{-n_i} Z(n_i, 2\beta))$$

$$= \sum_{\substack{0 \le n_1, \ldots, n_8 \le n \\ n_1 + \cdots + n_8 = n}} \frac{8^{-n} n!}{n_1! n_2! \ldots n_8!} e^{-2\beta \sum_{1 \le i < j \le 8} n_i n_j} \prod_{i=1}^{8} Z(n_i, 2\beta).$$

Moreover, for any $(n_1, \ldots, n_8)$ occurring in the above sum,

$$\mathbb{P}(N_1 = n_1, \ldots, N_8 = n_8)$$
$$= \frac{8^{-n} n!}{n_1! n_2! \cdots n_8!} e^{-2\beta \sum_{1 \le i < j \le 8} n_i n_j} \frac{\prod_{i=1}^{8} Z(n_i, 2\beta)}{Z(n, \beta)}.$$

Choose nonnegative integers $m_1, \ldots, m_8$ such that $m_1 + \cdots + m_8 = n$ and $|m_i - n/8| \le 1$ for each i. It is not difficult to see that such integers can be found for any n. For convenience, let

$$f(n_1, \ldots, n_8) := \frac{n!}{n_1! n_2! \cdots n_8!},$$
$$g(n_1, \ldots, n_8) := e^{-2\beta \sum_{1 \le i < j \le 8} n_i n_j} = e^{-\beta n^2 + \beta \sum_{i=1}^{8} n_i^2},$$
$$h(n_1, \ldots, n_8) := \prod_{i=1}^{8} Z(n_i, 2\beta).$$

Take any $k_1, \ldots, k_8 \in \mathbb{Z}$ such that $k_1 + \cdots + k_8 = 0$ and $0 \le m_i + k_i \le n$ for each i. Then by Corollary 2.4,

$$\frac{h(m_1 + k_1, \ldots, m_8 + k_8)}{h(m_1, \ldots, m_8)}$$
$$\le \prod_{i=1}^{8} \exp\left(-\frac{14\beta m_i k_i}{3} - \frac{14\beta k_i (k_i - 1)}{6} + 2C_2 \beta |k_i| (n + |k_i|)^{1/3}\right)$$
$$\le \prod_{i=1}^{8} \exp\left(-\frac{14\beta (n k_i / 8 - |k_i|)}{3} - \frac{14\beta k_i (k_i - 1)}{6} + 4C_2 \beta |k_i| n^{1/3}\right)$$
$$\le \exp\left(-\frac{14\beta}{6} \sum_{i=1}^{8} k_i^2 + C_3 \beta n^{1/3} \sum_{i=1}^{8} |k_i|\right).$$

Next, note that

$$
\begin{aligned}
\frac{g(m_1 + k_1, \ldots, m_8 + k_8)}{g(m_1, \ldots, m_8)} &= \exp\left(\beta \sum_{i=1}^{8}(m_i + k_i)^2 - \beta \sum_{i=1}^{8} m_i^2\right) \\
&= \exp\left(\beta \sum_{i=1}^{8}(2m_i k_i + k_i^2)\right) \\
&\leq \exp\left(\beta \sum_{i=1}^{8}(2nk_i/8 + 2|k_i| + k_i^2)\right) \\
&= \exp\left(\beta \sum_{i=1}^{8}(2|k_i| + k_i^2)\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\frac{\mathbb{P}(N_1 = m_1 + k_1, \ldots, N_8 = m_8 + k_8)}{\mathbb{P}(N_1 = m_1, \ldots, N_8 = m_8)} \\
&\leq \frac{f(m_1 + k_1, \ldots, m_8 + k_8)}{f(m_1, \ldots, m_8)} \exp\left(-\frac{4\beta}{3}\sum_{i=1}^{8} k_i^2 + C_4\beta n^{1/3}\sum_{i=1}^{8}|k_i|\right).
\end{aligned}
$$

This shows that there are positive constants C_5 and C_6 such that if

$$
\max_{1 \leq i \leq 8} |k_i| \geq C_5 n^{1/3},
$$

then

$$
\begin{aligned}
&\frac{\mathbb{P}(N_1 = m_1 + k_1, \ldots, N_8 = m_8 + k_8)}{\mathbb{P}(N_1 = m_1, \ldots, N_8 = m_8)} \\
&\leq \frac{f(m_1 + k_1, \ldots, m_8 + k_8)}{f(m_1, \ldots, m_8)} e^{-C_6\beta n^{2/3}}.
\end{aligned} \tag{2.1}
$$

Let A denote the set of all $(n_1, \ldots, n_8)$ such that each n_i is a nonnegative integer, $n_1 + \cdots + n_8 = n$, and

$$
\max_{1 \leq i \leq 8} |n_i - m_i| \geq C_5 n^{1/3}.
$$

Then by (2.1), for any $(n_1, \ldots, n_8) \in A$,

$$
\frac{\mathbb{P}(N_1 = n_1, \ldots, N_8 = n_8)}{\mathbb{P}(N_1 = m_1, \ldots, N_8 = m_8)} \leq \frac{f(n_1, \ldots, n_8)}{f(m_1, \ldots, m_8)} e^{-C_6\beta n^{2/3}}.
$$

Now recall the multinomial formula

$$\sum_{\substack{0 \leq n_1, \ldots, n_8 \leq n \\ n_1 + \cdots + n_8 = n}} f(n_1, \ldots, n_8) = 8^n.$$

A simple calculation using Stirling's formula shows that

$$f(m_1, \ldots, m_8) 8^{-n} \geq C_7 n^{-4}.$$

Thus,

$$\begin{aligned}
\mathbb{P}((N_1, \ldots, N_8) \in A) &\leq \frac{\mathbb{P}((N_1, \ldots, N_8) \in A)}{\mathbb{P}(N_1 = m_1, \ldots, N_8 = m_8)} \\
&= \sum_{(n_1, \ldots, n_8) \in A} \frac{\mathbb{P}(N_1 = n_1, \ldots, N_8 = n_8)}{\mathbb{P}(N_1 = m_1, \ldots, N_8 = m_8)} \\
&\leq e^{-C_6 \beta n^{2/3}} \frac{8^n}{f(m_1, \ldots, m_8)} \leq C_8 n^4 e^{-C_6 \beta n^{2/3}}.
\end{aligned}$$

Therefore for each i,

$$\begin{aligned}
\mathrm{Var}(N_i) &\leq \mathbb{E}(N_i - m_i)^2 \leq C_5^2 n^{2/3} + n^2 \mathbb{P}((N_1, \ldots, N_8) \in A) \\
&\leq C_5^2 n^{2/3} + C_8 n^6 e^{-C_6 \beta n^{2/3}}.
\end{aligned}$$

The above inequality shows that

$$\mathrm{Var}(N_i) \leq K(\beta) n^{2/3},$$

where $K(\beta)$ is a decreasing function of β. □

For any Borel set $A \subseteq [0, 1)^3$, let

$$X(A) := \{X_j : X_j \in A\}.$$

and let $N(A) := |X(A)|$. For each $k \geq 0$, let $\mathcal{F}_k$ be the σ-algebra generated the random variables $\{N(D) : D \in \mathcal{D}_k\}$. Note that $\{\mathcal{F}_k\}_{k \geq 0}$ is a filtration of σ-algebras. This filtration will play an important role in the subsequent discussion.

Lemma 2.6 *Conditional on $\mathcal{F}_k$, the random sets $\{X(D) : D \in \mathcal{D}_k\}$ are mutually independent. Moreover, for any $D \in \mathcal{D}_k$, conditional on $\mathcal{F}_k$, $X(D)$ has the same distribution as a scaled version of a point process from the measure $\mu_{N(D), 2^k \beta}$.*

Proof Take any k. Note that the joint density of $(X_1, \ldots, X_n)$ at a point $(x_1, \ldots, x_n)$ may be written as

$$\frac{1}{Z(n, \beta)} \exp\left(-\beta \sum_{D \in \mathcal{D}_k} H_D(x_1, \ldots, x_n) - \beta R_k(x_1, \ldots, x_n)\right),$$

where $H_D(x_1, \ldots, x_n)$ is the contribution due to the interactions between points in D, and $R_k(x_1, \ldots, x_n)$ is the contribution due to the interactions between points in different members of $\mathcal{D}_k$. The crucial property of the potential w is that $R_k(x_1, \ldots, x_n)$ is a function of $\{n_D : D \in \mathcal{D}_k\}$, where $n_D = |\{j : x_j \in D\}|$. The claims follow easily from this observation. □

Lemma 2.6 allows us to compute conditional means and variances.

Lemma 2.7 *If $D \in \mathcal{D}_k$ and D' is a child of D, then*

$$\mathbb{E}(N(D')|\mathcal{F}_k) = \frac{N(D)}{8}$$

and

$$\mathrm{Var}(N(D')|\mathcal{F}_k) \leq K(\beta)N(D)^{2/3},$$

where K is the function from Lemma 2.5.

Proof The formula for the conditional expectation follows from Lemma 2.6 and symmetry, and the bound on the conditional variance follows from Lemmas 2.5, 2.6, and the observation that $K(2^k\beta) \leq K(\beta)$ since K is a non-increasing function of β. □

The above lemma leads to the following conclusions about unconditional means and variances.

Lemma 2.8 *For any $D \in \mathcal{D}$, $\mathbb{E}(N(D)) = Leb(D)n$ and*

$$\mathrm{Var}(N(D)) \leq 8K(\beta)Leb(D)^{2/3}n^{2/3},$$

where K is the function from Lemma 2.5.

Proof Suppose that $D \in \mathcal{D}_k$. The formula for the expectation follows easily by iterating the formula for the conditional expectation from Lemma 2.7, and observing that $Leb(D) = 8^{-k}$. Next, let D' be the parent of D. Then by Lemma 2.7 and the formula for expected value,

$$\begin{aligned}
\mathbb{E}(N(D)^2) &= \mathbb{E}(N(D)^2 - (\mathbb{E}(N(D)|\mathcal{F}_{k-1}))^2) + \mathbb{E}((\mathbb{E}(N_D|\mathcal{F}_{k-1}))^2) \\
&= \mathbb{E}(\mathrm{Var}(N(D)|\mathcal{F}_{k-1})) + 8^{-2}\mathbb{E}(N(D')^2) \\
&\leq K(\beta)\mathbb{E}(N(D')^{2/3}) + 8^{-2}\mathbb{E}(N(D')^2) \\
&\leq K(\beta)(\mathbb{E}(N(D')))^{2/3} + 8^{-2}\mathbb{E}(N(D')^2) \\
&= K(\beta)4^{-k+1}n^{2/3} + 8^{-2}\mathbb{E}(N(D')^2).
\end{aligned}$$

Iterating this, we get

$$\begin{aligned}
\mathbb{E}(N(D)^2) &\leq K(\beta)n^{2/3}(4^{-k+1} + 8^{-2}4^{-k+2} + 8^{-4}4^{-k+3} + \cdots) + 8^{-2k}n^2 \\
&\leq 8K(\beta)4^{-k}n^{2/3} + 8^{-2k}n^2,
\end{aligned}$$

which completes the proof since $\mathbb{E}(N(D)) = Leb(D)n = 8^{-k}n$. □

Now take any nonempty open set $U \subseteq [0, 1)^3$ with regular boundary. Let $\mathcal{U}$ be the set of all $D \in \mathcal{D}$ such that $D \subseteq U$ but the parent cube of D is not contained in U.

Lemma 2.9 *The set U is the disjoint union of all elements of $\mathcal{U}$.*

Proof Since U is open, each point in U belongs to some dyadic cube that is contained in U. Some ancestor of this cube must belong to $\mathcal{U}$. This shows that U is the union of the members of $\mathcal{U}$. It is easy to see that the elements of $\mathcal{U}$ are disjoint. □

Corollary 2.10 $\mathbb{E}(N(U)) = Leb(U)n$.

Proof Just observe that by Lemmas 2.8 and 2.9,

$$\mathbb{E}(N(U)) = \sum_{D \in \mathcal{U}} \mathbb{E}(N(D)) = \sum_{D \in \mathcal{U}} Leb(D)n = Leb(U)n,$$

where we have implicitly used the fact that $\mathcal{U}$ is a countable collection. □

For each j, let $\mathcal{U}_j := \mathcal{U} \cap \mathcal{D}_j$. Let $\mathcal{V}_j$ denote the set of all $D \in \mathcal{D}_j$ that intersect both U and U^c. Note that $\mathcal{U}_j$ and $\mathcal{V}_j$ do not overlap. For any dyadic cube D, let $p(D)$ denote the proportion of D that belongs to U. Let $M_0 = Leb(U)n$ and for each $j \geq 1$, let

$$M_j := \sum_{i=0}^{j} \sum_{D \in \mathcal{U}_i} N(D) + \sum_{D \in \mathcal{V}_j} p(D)N(D).$$

Lemma 2.11 *The sequence $\{M_j\}_{j \geq 0}$ is a martingale with respect to the filtration $\{\mathcal{F}_j\}_{j \geq 0}$.*

Proof Take any $j \geq 1$. Then

$$\mathbb{E}(M_j | \mathcal{F}_{j-1}) = \sum_{i=0}^{j-1} \sum_{D \in \mathcal{U}_i} N(D) + \sum_{D \in \mathcal{U}_j} \mathbb{E}(N(D) | \mathcal{F}_{j-1})$$

$$+ \sum_{D \in \mathcal{V}_j} p(D)\mathbb{E}(N(D) | \mathcal{F}_{j-1})$$

$$= \sum_{i=0}^{j-1} \sum_{D \in \mathcal{U}_i} N(D) + \sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} p(D)\mathbb{E}(N(D) | \mathcal{F}_{j-1})$$

Take any $D \in \mathcal{V}_{j-1}$. Then each child of D is either a member of $\mathcal{U}_j$, or a member of $\mathcal{V}_j$, or has no intersection with U. Conversely, every member of $\mathcal{U}_j \cup \mathcal{V}_j$ is the child of some member of $\mathcal{V}_{j-1}$. Lastly, note that if $D_1, \ldots, D_8$ are the children of a dyadic cube D, then

$$p(D) = \frac{1}{8} \sum_{i=1}^{8} p(D_i).$$

Combining these observations and applying Lemma 2.7, we get

$$\sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} p(D)\mathbb{E}(N(D)|\mathcal{F}_{j-1}) = \sum_{D \in \mathcal{V}_{j-1}} p(D)N(D),$$

which completes the proof.   □

For the remainder of this section, let $A(U)$ be a constant such that for all $0 < \epsilon \le$ diam(U),

$$\text{Leb}(\partial U_\epsilon) \le A(U)\epsilon. \tag{2.2}$$

By the regularity condition, we can choose $A(U)$ to be finite. The martingale property of M_j and our previous calculations lead to the following conclusion.

Lemma 2.12 *For any $j \ge 1$ such that $\sqrt{3} \cdot 2^{-j+1} \le$ diam(U),*

$$\text{Var}(M_j) \le C(\beta)A(U)n^{2/3} + \text{Var}(M_{j-1}),$$

where $C(\beta)$ is a constant that depends only on β.

Proof By the martingale property,

$$\begin{aligned}
\text{Var}(M_j) &= \mathbb{E}(\text{Var}(M_j|\mathcal{F}_{j-1})) + \text{Var}(\mathbb{E}(M_j|\mathcal{F}_{j-1})) \\
&= \mathbb{E}(\text{Var}(M_j|\mathcal{F}_{j-1})) + \text{Var}(M_{j-1}). \tag{2.3}
\end{aligned}$$

Now,

$$\begin{aligned}
\text{Var}(M_j|\mathcal{F}_{j-1}) &= \text{Var}\left(\sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} p(D)N(D) \middle| \mathcal{F}_{j-1}\right) \\
&= \sum_{D, D' \in \mathcal{U}_j \cup \mathcal{V}_j} p(D)p(D')\text{Cov}(N(D), N(D')|\mathcal{F}_{j-1}). \tag{2.4}
\end{aligned}$$

If D and D' have different parents, then $N(D)$ and $N(D')$ are conditionally independent by Lemma 2.6, and hence the conditional covariance is zero. Otherwise, Lemma 2.7 and the Cauchy–Schwarz inequality imply that

$$|\text{Cov}(N(D), N(D')|\mathcal{F}_{j-1})| \le K(\beta)N(D'')^{2/3},$$

where D'' is the parent of D and D'. Thus, by Lemma 2.8,

$$\begin{aligned}
|\mathbb{E}(\text{Cov}(N(D), N(D')|\mathcal{F}_{j-1}))| &\le K(\beta)(\text{Leb}(D'')n)^{2/3} \\
&= K(\beta)(8^{-j+1}n)^{2/3}.
\end{aligned}$$

On the other hand, each $D \in \mathcal{U}_j \cup \mathcal{V}_j$ has at most 7 sibling cubes that belong to $\mathcal{U}_j \cup \mathcal{V}_j$. Since $p(D)8^{-j} = p(D)\mathrm{Leb}(D) = \mathrm{Leb}(D \cap U)$, this shows that

$$\mathbb{E}(\mathrm{Var}(M_j|\mathcal{F}_{j-1})) \le K(\beta)(8^{-j+1}n)^{2/3} \sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} 7p(D)$$

$$= 28K(\beta)n^{2/3}2^j \sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} \mathrm{Leb}(D \cap U).$$

Note that each element of

$$\bigcup_{D \in \mathcal{U}_j \cup \mathcal{V}_j} (D \cap U)$$

is within distance $\sqrt{3} \cdot 2^{-j+1}$ from ∂U. Since $\sqrt{3} \cdot 2^{-j+1} \le \mathrm{diam}(U)$, inequality (2.2) gives

$$\sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} \mathrm{Leb}(D \cap U) \le A(U)\sqrt{3} \cdot 2^{-j+1}.$$

Consequently,

$$\mathbb{E}(\mathrm{Var}(M_j|\mathcal{F}_{j-1})) \le C(\beta)A(U)n^{2/3},$$

where $C(\beta)$ depends only on β. The proof is completed by plugging this bound into (2.3). $\qquad\square$

We now have all the ingredients for proving the following theorem, which implies Theorems 1.1 and 1.3 and special cases.

Theorem 2.13 (Hyperuniformity at all scales) *Let U and $N(U)$ be as in Theorem* 1.1. *Suppose that $\mathrm{diam}(U) \ge n^{-1/3}$. Let $A(U)$ be the constant defined in* (2.2). *Then*

$$\mathbb{E}(N(U)) = Leb(U)n$$

and

$$\mathrm{Var}(N(U)) \le C(\beta)A(U)n^{2/3} \log(4n^{1/3}diam(U)) + C(\beta)Leb(U)^{2/3}n^{2/3},$$

where $C(\beta)$ is a constant that depends only on β.

Proof Throughout this proof, $C(\beta)$ will denote any constant that depends only on β. The value of $C(\beta)$ may change from line to line or even within a line.

The formula for the expectation follows from Corollary 2.10. It remains to prove the variance bound. Choose k such that

$$\frac{1}{2}n^{-1/3} \le \sqrt{3} \cdot 2^{-k} \le n^{-1/3}.$$

Note that by Lemma 2.9, any point in U either belongs to some $D \in \mathcal{U}_j$ for some $j \leq k$, or belongs to some $D \in \mathcal{U}_j$ for some $j > k$. In the latter case, there is an ancestor of D that belongs to $\mathcal{V}_k$. Thus,

$$U = \left(\bigcup_{j=0}^{k} \mathcal{U}_j \right) \cup \left(\bigcup_{D \in \mathcal{V}_k} (D \cap U) \right),$$

and so

$$N(U) = \sum_{j=0}^{k} \sum_{D \in \mathcal{U}_j} N(D) + \sum_{D \in \mathcal{V}_k} N(D \cap U).$$

Consequently, by Lemmas 2.6, 2.8 and Corollary 2.10,

$$\mathbb{E}(N(U)|\mathcal{F}_k) = \sum_{j=0}^{k} \sum_{D \in \mathcal{U}_j} N(D) + \sum_{D \in \mathcal{V}_k} \mathbb{E}(N(D \cap U)|\mathcal{F}_k)$$

$$= \sum_{j=0}^{k} \sum_{D \in \mathcal{U}_j} N(D) + \sum_{D \in \mathcal{V}_k} p(D)N(D) = M_k.$$

Therefore,

$$\begin{aligned} \mathrm{Var}(N(U)) &= \mathbb{E}(\mathrm{Var}(N(U)|\mathcal{F}_k)) + \mathrm{Var}(\mathbb{E}(N(U)|\mathcal{F}_k)) \\ &= \mathbb{E}(\mathrm{Var}(N(U)|\mathcal{F}_k)) + \mathrm{Var}(M_k). \end{aligned} \tag{2.5}$$

Given $\mathcal{F}_k$, the random variables $\{N(D \cap U) : D \in \mathcal{D}_k\}$ are independent by Lemma 2.6. Therefore, by Lemma 2.8 and Corollary 2.10,

$$\begin{aligned} \mathrm{Var}(N(U)|\mathcal{F}_k) &= \mathrm{Var}\left(\sum_{D \in \mathcal{V}_k} N(D \cap U) \,\middle|\, \mathcal{F}_k \right) = \sum_{D \in \mathcal{V}_k} \mathrm{Var}(N(D \cap U)|\mathcal{F}_k) \\ &\leq \sum_{D \in \mathcal{V}_k} \mathbb{E}(N(D \cap U)^2|\mathcal{F}_k) \\ &\leq \sum_{D \in \mathcal{V}_k} \mathbb{E}(N(D \cap U)|\mathcal{F}_k)N(D) = \sum_{D \in \mathcal{V}_k} p(D)N(D)^2. \end{aligned}$$

By Lemma 2.8 and our choice of k,

$$\mathbb{E}(N(D)^2) = \mathrm{Var}(N(D)) + (\mathbb{E}(N(D)))^2 \leq C(\beta)$$

for all $D \in \mathcal{V}_k$. Also, each element of

$$\bigcup_{D \in \mathcal{V}_k} (D \cap U)$$

is within distance $\sqrt{3} \cdot 2^{-k}$ of ∂U, and $p(D)8^{-k} = \text{Leb}(D \cap U)$. Since

$$\sqrt{3} \cdot 2^{-k} \leq n^{-1/3} \leq \text{diam}(U)$$

by our choice of k and the assumption that $\text{diam}(U) \geq n^{-1/3}$, this gives

$$\mathbb{E}(\text{Var}(N(U)|\mathcal{F}_k)) \leq C(\beta)8^k \sum_{D \in \mathcal{V}_k} \text{Leb}(D \cap U)$$
$$\leq C(\beta)8^k A(U)2^{-k}$$
$$= C(\beta)A(U)4^k \leq C(\beta)A(U)n^{2/3}.$$

Let l be the smallest integer such that $\sqrt{3} \cdot 2^{-l} \leq \text{diam}(U)$. Note that $l \leq k$. Together with (2.5) and Lemma 2.12, the above inequality shows that

$$\text{Var}(N(U)) \leq C(\beta)A(U)n^{2/3}(k - l + 1) + \text{Var}(M_l).$$

By the definition of l, $\mathcal{U}_i$ if empty for all $i < l$. Therefore

$$M_l = \sum_{D \in \mathcal{U}_l \cup \mathcal{V}_l} p(D)N(D).$$

Note that for any $D \in \mathcal{U}_l \cup \mathcal{V}_l$, Lemma 2.8 gives

$$\text{Var}(p(D)N(D)) = p(D)^2 \text{Var}(N(D)) \leq C(\beta)p(D)^2 \text{Leb}(D)^{2/3}n^{2/3}$$
$$\leq C(\beta)(p(D)\text{Leb}(D))^{2/3}n^{2/3}$$
$$= C(\beta)\text{Leb}(D \cap U)^{2/3}n^{2/3} \leq C(\beta)\text{Leb}(U)^{2/3}n^{2/3}.$$

Moreover, it is easy to see that U intersects at most 64 members of $\mathcal{D}_l$, and therefore $|\mathcal{U}_l \cup \mathcal{V}_l| \leq 64$. From these observations, we get

$$\text{Var}(M_l) \leq C(\beta)\text{Leb}(U)^{2/3}n^{2/3}.$$

Finally, note that by the lower bound on $\sqrt{3} \cdot 2^{-k}$ and the upper bound on $\sqrt{3} \cdot 2^{-l}$, we get

$$2^{k-l} \leq 2n^{1/3}\text{diam}(U),$$

and hence $k - l + 1 \leq \log_2(4n^{1/3}\text{diam}(U))$. This completes the proof of the theorem. $\qquad\square$

Proof of Theorem 1.1 This is a direct application of Theorem 2.13. The condition $\mathrm{diam}(U) \geq n^{-1/3}$ is irrelevant because the variance bound can be enforced for small n by adjusting the constant $C(U, \beta)$. □

Proof of Theorem 1.3 Let $V := n^{-1/3}\lambda U + x$. Note that $N_x(\lambda U) = N(V)$. Also, note that

$$\mathrm{Leb}(V) = \lambda^3 n^{-1} \mathrm{Leb}(U),$$
$$A(V) = \lambda^2 n^{-2/3} A(U),$$
$$\mathrm{diam}(V) = \lambda n^{-1/3} \mathrm{diam}(U).$$

In particular, the condition $\mathrm{diam}(V) \geq n^{-1/3}$ is equivalent to $\mathrm{diam}(\lambda U) \geq 1$. The proof is now just an application of Theorem 2.13, and the observation that since $x \in (0, 1)^3$, V is eventually contained in $(0, 1)^3$ as n gets large. □

Finally, let us prove Theorem 1.4.

Proof of Theorem 1.4 Here $C(\beta)$ denotes any constant that depends only on β. Let $f(D)$ be the average value of f in a dyadic square $D \in \mathcal{D}$. For each k, let f_k be the function that is identically equal to $f(D)$ within each $D \in \mathcal{D}_k$. Let

$$W_k := X(f_k).$$

By Lemmas 2.6 and 2.7, it is easy to see that $\{W_k\}_{k \geq 0}$ is martingale with respect to the filtration $\{\mathcal{F}_k\}_{k \geq 0}$. Moreover, for any k,

$$\mathbb{E}(X(f)|\mathcal{F}_k) = X(f_k). \tag{2.6}$$

Now choose k such that

$$n^{-1/3} \leq 2^{-k} \leq 2n^{-1/3}.$$

Then by (2.6) and the martingale property of $\{W_j\}_{j \geq 0}$,

$$\mathrm{Var}(X(f)) = \mathbb{E}(\mathrm{Var}(X(f)|\mathcal{F}_k)) + \sum_{j=1}^{k} \mathbb{E}(\mathrm{Var}(X(f_j)|\mathcal{F}_{j-1})). \tag{2.7}$$

Take any j. For each $D \in \mathcal{D}_{j-1}$, let $c(D)$ denote the set of 8 children of D. By Lemma 2.6 and Lemma 2.7,

$$\mathrm{Var}(X(f_j)|\mathcal{F}_{j-1}) = \mathrm{Var}\left(\sum_{D \in \mathcal{D}_j} f(D)N(D)\middle|\mathcal{F}_{j-1}\right)$$

$$= \sum_{D \in \mathcal{D}_{j-1}} \mathrm{Var}\left(\sum_{D' \in c(D)} f(D')N(D')\middle|\mathcal{F}_{j-1}\right)$$

$$= \sum_{D \in \mathcal{D}_{j-1}} \mathbb{E}\left(\left(\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D)\right)^2 \Bigg| \mathcal{F}_{j-1}\right).$$

Now notice that for any $D \in \mathcal{D}_{j-1}$,

$$\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D)$$

$$= \sum_{D' \in c(D)} (f(D') - f(D))\left(N(D') - \frac{N(D)}{8}\right).$$

Recall that L is the Lipschitz constant of f. For any $D' \in c(D)$,

$$|f(D') - f(D)| \le \sqrt{3}L2^{-j+1}.$$

Thus,

$$\left(\sum_{D' \in c(D)} (f(D') - f(D))\left(N(D') - \frac{N(D)}{8}\right)\right)^2$$

$$\le 4^{-j+2}L^2 \left(\sum_{D' \in c(D)} \left|N(D') - \frac{N(D)}{8}\right|\right)^2$$

$$\le 4^{-j+4}L^2 \sum_{D' \in c(D)} \left(N(D') - \frac{N(D)}{8}\right)^2.$$

Therefore, by Lemma 2.7,

$$\mathbb{E}\left(\left(\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D)\right)^2 \Bigg| \mathcal{F}_{j-1}\right)$$

$$\le 4^{-j+4}L^2 \sum_{D' \in c(D)} \text{Var}(N(D')|\mathcal{F}_{j-1}) \le 4^{-j+6}L^2 K(\beta)N(D)^{2/3}.$$

Consequently, by Lemma 2.8,

$$\mathbb{E}(\text{Var}(X(f_j)|\mathcal{F}_{j-1})) \le C(\beta)L^2 4^{-j} \sum_{D \in \mathcal{D}_{j-1}} \mathbb{E}(N(D)^{2/3})$$

$$\le C(\beta)L^2 4^{-j} \sum_{D \in \mathcal{D}_{j-1}} (\mathbb{E}(N(D)))^{2/3} \le C(\beta)L^2 2^{-j}n^{2/3}.$$

Next, for $D \in \mathcal{D}_k$, let

$$s(D) := \sum_{j\,:\,X_j \in D} f(X_j),$$

so that

$$X(f) = \sum_{D \in \mathcal{D}_k} s(D).$$

Then by Lemma 2.6,

$$\mathrm{Var}(X(f)|\mathcal{F}_k) = \sum_{D \in \mathcal{D}_k} \mathrm{Var}(s(D)|\mathcal{F}_k)$$
$$\leq \sum_{D \in \mathcal{D}_k} \mathbb{E}((s(D) - f(D)N(D))^2|\mathcal{F}_k).$$

By the Lipschitz condition,

$$|s(D) - f(D)N(D)| \leq \sqrt{3}L2^{-k}N(D)$$

for each $D \in \mathcal{D}_k$. Thus, by Lemma 2.8 and our choice of k,

$$\mathbb{E}((s(D) - f(D)N(D))^2) \leq 4^{-k+1}L^2\mathbb{E}(N(D)^2) \leq C(\beta)L^24^{-k}.$$

Consequently,

$$\mathbb{E}(\mathrm{Var}(X(f)|\mathcal{F}_k)) \leq C(\beta)L^24^{-k}|\mathcal{D}_k| \leq C(\beta)L^22^k \leq C(\beta)L^2n^{1/3}.$$

The proof is now easily completed by combining the steps. $\qquad\square$

2.5 Proofs of the lower bounds

Let us now prove Theorem 1.2. We will continue using the notations introduced in the previous sections. We need to prove some simple geometric facts. Let

$$\mathcal{T} := \{z + [0, 1)^3 : z \in \mathbb{Z}^3\}.$$

Our first geometric lemma is very simple.

Lemma 2.14 *Let $\mathcal{T}$ be as above. Take any $D \in \mathcal{T}$ and any $x \in D$. Let δ be the distance of x from the boundary of D. Then any plane through x bifurcates D into two parts, each of which has volume at least $2\pi\delta^3/3$.*

Proof The open ball of radius δ around x is contained in D. Any plane P through x bifurcates this ball into two parts of volume $2\pi\delta^3/3$ each. The proof is completed by observing that these two hemispheres are contained in the two parts of D obtained by bifurcating using P. □

The second lemma is an easy fact about intervals.

Lemma 2.15 *Let I be a closed interval of the real line of length at least $\delta \in [0, 1]$. Then I has a closed subinterval J of length $\delta/4$ such that any integer is at a distance at least $\delta/4$ from J.*

Proof If I contains no integers, then we can take J to be an interval of length $\delta/4$ that is at distance at least $\delta/4$ from each endpoint of I. If I contains an integer n, then at least one of the two intervals $[n, n+\delta/2]$ and $[n-\delta/2, n]$ must be contained in I. In the first case take $J = [n+\delta/4, n+\delta/2]$ and in the second case take $J = [n-\delta/2, n-\delta/4]$. Since $\delta \le 1$, there is no integer within distance $\delta/4$ from J. □

The next lemma is intuitively obvious but a little tedious to prove. The constants are probably not optimal, but that does not matter for us.

Lemma 2.16 *Take any $x \in \mathbb{R}^3$ and a unit vector $u = (u_1, u_2, u_3) \in S^2$. Let P be the plane that contains x and is perpendicular to u. Suppose that*

$$\min\{|u_1|, |u_2|, |u_3|\} \ge 0.1. \tag{2.8}$$

Then there is an element $D \in \mathcal{T}$, within Euclidean distance $\sqrt{402}$ from x, which is bifurcated by the plane P in such a way that each part has volume at least 6×10^{-8}.

Proof Take any $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $u = (u_1, u_2, u_3) \in S^2$ as in the statement of the lemma. Let P_0 be the plane with normal vector u that contains the origin. Define

$$y_1 = \text{sign}(u_1), \quad y_2 = \text{sign}(u_2), \quad y_3 = -\frac{|u_1| + |u_2|}{u_3}.$$

Then $y = (y_1, y_2, y_3) \in P_0$. Also, we have $|y_1| = 1$, $|y_2| = 1$, and by condition (2.8) and the fact that $|u_3| \le 1$,

$$|y_3| = \frac{|u_1| + |u_2|}{|u_3|} \ge |u_1| + |u_2| \ge 0.2.$$

Now consider the set

$$I_1 = \{x_1 + \alpha y_1 : 0 \le \alpha \le 1\}.$$

Since $|y_1| = 1$, I_1 is an interval of length 1. By Lemma 2.15, I_1 has a subinterval of I_2 of length 0.25 such that any integer is at least at a distance 0.25 from I_2. Moreover, since $|y_1| = 1$, I_2 is of the form

$$\{x_1 + \alpha y_1 : a \le \alpha \le b\},$$

where $b - a = 0.25$. Let

$$I_3 := \{x_2 + \alpha y_2 : a \leq \alpha \leq b\}.$$

Since $|y_2| = 1$, I_3 has length 0.25. Thus by Lemma 2.15, I_3 contains a subinterval I_4 of length 0.0625 such that any integer is at a distance at least 0.0625 from I_4. Again, since $|y_2| = 1$, this implies that I_4 is of the form

$$\{x_2 + \alpha y_2 : c \leq \alpha \leq d\},$$

where $a \leq c \leq d \leq b$ and $d - c = 0.0625$. Let

$$I_5 := \{x_3 + \alpha y_3 : c \leq \alpha \leq d\}.$$

Since $|y_3| \geq 0.2$, I_5 has length at least 0.0125. Consequently by Lemma 2.15, I_5 has a subinterval I_6 of length 0.003125 such that any integer is at a distance at least 0.003125 from I_6.

In particular, there is some $\alpha \in [0, 1]$ such that $x_1 + \alpha y_1 \in I_2$, $x_2 + \alpha y_2 \in I_4$ and $x_3 + \alpha y_3 \in I_6$. The distance of $x_i + \alpha y_i$ from the nearest integer is at least 0.003125 for each i. Thus, the distance of the point $x + \alpha y$ from the boundary of the cube $D \in \mathcal{T}$ that contains $x + \alpha y$ is at least 0.003125. By Lemma 2.14 and the fact that $x + \alpha y \in P$, this proves P bifurcates D into two parts, each of which has volume at least 6×10^{-8}. Lastly, note that

$$|(x + \alpha y) - x| \leq |y| = \sqrt{y_1^2 + y_2^2 + y_3^2}$$
$$\leq \sqrt{1 + 1 + \frac{(1+1)^2}{0.1^2}} \leq \sqrt{402},$$

since $|u_1| \leq 1$, $|u_2| \leq 1$ and $|u_3| \geq 0.1$. This completes the proof of the lemma. □

Now recall that the boundary of the set U in the statement of Theorem 1.2 is a smooth, closed, orientable surface. In particular, we can choose a unit normal vector $u(x)$ at each $x \in \partial U$ such that the map $x \mapsto u(x)$ is smooth.

Lemma 2.17 *Take any $x \in \partial U$ such that the normal vector $u(x)$ satisfies (2.8). Then there is some j_0 depending only on U (but not on x), such that for all $j \geq j_0$, there is some $D \in \mathcal{D}_j$ at distance at most $\sqrt{402} \cdot 2^{-j}$ from x, which satisfies*

$$10^{-8} \leq \frac{Leb(D \cap U)}{Leb(D)} \leq 1 - 10^{-8}. \tag{2.9}$$

Proof From the given properties of ∂U, it is clear that ∂U has uniformly bounded curvature. Consequently, there is a constant C depending only on U, such that for any $x \in \partial U$ and any $\epsilon \in (0, 1)$, $B(x, \epsilon) \cap \partial U$ lies inside a slab of width $C\epsilon^2$ around T_x, where $B(x, \epsilon)$ is the Euclidean ball of radius ϵ around x, and T_x is the tangent plane at x. The rest of the proof is an easy application of Lemma 2.16 and scaling. □

The above lemma leads to the following result, which is a key component of the proof of Theorem 1.2.

Lemma 2.18 *There is some $K_1 > 0$ and some $j_1 \geq 1$ depending only on U such that for any $j \geq j_1$, there is a set of at least $K_1 4^j$ cubes $D \in \mathcal{D}_j$ that satisfy (2.9) and the union of these cubes has diameter at most $\mathrm{diam}(U)/3$.*

Proof Let P be the plane through the origin that is perpendicular to the vector $(1, 1, 1)$. Let α_0 be the largest α such that the plane $P_\alpha := (\alpha, \alpha, \alpha) + P$ intersects the closure of U. Let x be a point of intersection. Then $x \in \partial U$, and $P_{\alpha_0} = T_x$. Consequently, there is some $0 < \epsilon < \mathrm{diam}(U)/7$ such that for every $y \in B(x, \epsilon) \cap \partial U$, $u(y)$ satisfies (2.8). Due to the boundedness of the curvature of ∂U, a small enough choice of ϵ guarantees that for any $\delta \in (0, 1)$, there are at least $C\delta^{-2}$ points in $B(x, \epsilon) \cap \partial U$, where C is a positive constant that depends only on U, such that any two points are at distance at least 50δ from each other.

Take $\delta = 2^{-j}$, and choose a collection of points as above. Then by Lemma 2.17, there is an element of $\mathcal{D}_j$ within distance 21δ from each point, that satisfies (2.9). Since the points are separated by distance at least 50δ from each other, these elements of $\mathcal{D}_j$ are distinct. Since $\epsilon < \mathrm{diam}(U)/7$, a large enough choice of j ensures that the union of these cubes has diameter less than $\mathrm{diam}(U)/3$. ∎

Lastly, we need a lemma about our point process. Recall that for any $D \in \mathcal{D}$, $N(D)$ is the number of points landing in D.

Lemma 2.19 *For any $n \geq 1$, $\beta > 0$, $j \geq 0$ and $D \in \mathcal{D}_j$,*

$$\mathbb{P}(N(D) \geq 2) \leq \exp\left(-2^{j+1}\beta + \frac{7\beta}{3}\binom{n}{2}\right).$$

Proof The $n = 1$ case is trivial, so let us take $n \geq 2$. By Jensen's inequality and Lemma 2.1,

$$Z(n, \beta) \geq \exp\left(-\frac{7\beta}{3}\binom{n}{2}\right).$$

On the other hand, if a configuration $x_1, \ldots, x_n$ has two or more points in D, then

$$H_n(x_1, \ldots, x_n) \geq 2^{j+1}.$$

Thus, if A is the set of all such configurations, then

$$\int_A e^{-\beta H_n(x_1, \ldots, x_n)} \, dx_1 \cdots dx_n \leq e^{-2^{j+1}\beta} \mathrm{Leb}(A) \leq e^{-2^{j+1}\beta}.$$

Combining, we get

$$\mathbb{P}(N(D) \geq 2) = \mu_{n,\beta}(A) \leq \exp\left(-2^{j+1}\beta + \frac{7\beta}{3}\binom{n}{2}\right),$$

which completes the proof.   □

Finally, we are ready to prove Theorem 1.2. Recall the filtration $\{\mathcal{F}_k\}_{k\geq 0}$ defined earlier.

Proof of Theorem 1.2 In this proof, the phrase 'n sufficiently large' will mean '$n \geq n_0$, where n_0 depends only on U and β'. Also, C will denote any positive universal constant, $C(\beta)$ will denote any positive constant that depends only on β, and $C(U, \beta)$ will denote any positive constant that depends only on U and β.

Choose k such that

$$n^{-1/3} \leq 2^{-k} \leq 2n^{-1/3}. \tag{2.10}$$

Then for any $D \in \mathcal{D}_k$, Lemma 2.8 gives

$$\mathbb{E}(N(D)^2) \leq K_2(\beta), \tag{2.11}$$

where $K_2(\beta)$ is a positive integer that depends only on β. Let

$$m := 1000 K_2(\beta).$$

Let $j > k$ be the smallest number such that

$$2^{j-k+1} \geq \frac{7}{3}\binom{m}{2} + 1.$$

Note that $0 \leq j - k \leq C(\beta)$.

Take any $D \in \mathcal{D}_k$. Let $\mathcal{D}_j(D)$ denote the set of elements of $\mathcal{D}_j$ that are descendants of D. Take any $D' \in \mathcal{D}_j(D)$. If $N(D) \leq m$, then by Lemmas 2.6 and 2.19,

$$\mathbb{P}(N(D') \geq 2|\mathcal{F}_k) \leq e^{-2^k\beta} \leq e^{-\beta n^{1/3}}.$$

Consequently,

$$\mathbb{P}(N(D) \leq m, \, N(D') \geq 2) = \mathbb{E}(\mathbb{P}(N(D') \geq 2|\mathcal{F}_k); \, N(D) \leq m)$$
$$\leq e^{-\beta n^{1/3}}\mathbb{P}(N(D) \leq m) \leq e^{-\beta n^{1/3}}.$$

In particular, if E is the event

$$\{N(D) \leq m \text{ and } N(D') \geq 2 \text{ for some } D \in \mathcal{D}_k \text{ and some } D' \in \mathcal{D}_j(D)\},$$

then a union bound gives

$$\mathbb{P}(E) \le \sum_{D \in \mathcal{D}_k} \sum_{D' \in \mathcal{D}_j(D)} \mathbb{P}(N(D) \le m,\ N(D') \ge 2)$$

$$\le |\mathcal{D}_j| e^{-\beta n^{1/3}} \le C(\beta) n e^{-\beta n^{1/3}}. \tag{2.12}$$

We will need this inequality later.

Now, if n is sufficiently large, then there is a set $\mathcal{C}' \subseteq \mathcal{D}_j$ that satisfies the conclusions of Lemma 2.18. In particular, $|\mathcal{C}'| \ge C(U, \beta)4^j$. Moreover, since each element of $\mathcal{C}'$ satisfies (2.9), these cubes must lie entirely within distance $\sqrt{3} \cdot 2^{-j}$ from ∂U. If n is large enough, then $\sqrt{3} \cdot 2^{-j} \le \mathrm{diam}(U)$. Therefore by the regularity of ∂U, we have $|\mathcal{C}'| \le C(U, \beta)4^j$.

Let $\mathcal{C}$ denote the set of all members of $\mathcal{D}_k$ who are ancestors of elements of $\mathcal{C}'$. By dropping some elements from $\mathcal{C}'$ if necessary, we can ensure that each member of $\mathcal{C}$ has exactly one descendant in $\mathcal{C}'$. Since $0 \le j - k \le C(\beta)$, this gives the inequalities

$$C_1(U, \beta)4^k \le |\mathcal{C}| = |\mathcal{C}'| \le C_2(U, \beta)4^k, \tag{2.13}$$

where $C_1(U, \beta)$ and $C_2(U, \beta)$ are positive constants that depend only on U and β. Let Q be the union of the elements of $\mathcal{C}$. Recall that by Lemma 2.18 and the relation between $\mathcal{C}$ and $\mathcal{C}'$,

$$\mathrm{diam}(Q) \le \frac{\mathrm{diam}(U)}{3} + 2\sqrt{3} \cdot 2^{-k},$$

which is less than $\mathrm{diam}(U)/2$ if n is sufficiently large. Thus, if n is large enough and $\sqrt{3} \cdot 2^{-k} \le \epsilon \le \mathrm{diam}(Q)$, then

$$\epsilon + \sqrt{3} \cdot 2^{-k} \le 2\epsilon \le 2\,\mathrm{diam}(Q) \le \mathrm{diam}(U).$$

Moreover, each point in Q is at distance at most $\sqrt{3} \cdot 2^{-k}$ from U. Therefore,

$$\mathrm{Leb}(\partial Q_\epsilon) \le \mathrm{Leb}(\partial U_{\epsilon + \sqrt{3} \cdot 2^{-k}}) \le A(U)(\epsilon + \sqrt{3} \cdot 2^{-k}) \le 2A(U)\epsilon.$$

On the other hand, if $0 < \epsilon \le \sqrt{3} \cdot 2^{-k}$, then

$$\mathrm{Leb}(\partial Q_\epsilon) \le \sum_{D \in \mathcal{C}} \mathrm{Leb}(\partial D_\epsilon) \le \sum_{D \in \mathcal{C}} A(D)\epsilon \le C \sum_{D \in \mathcal{C}} 4^{-k}\epsilon = C|\mathcal{C}|4^{-k}\epsilon.$$

Therefore, by (2.13), for $0 < \epsilon \le \sqrt{3} \cdot 2^{-k}$,

$$\mathrm{Leb}(\partial Q_\epsilon) \le C(U, \beta)\epsilon.$$

Combining the two cases, we get $A(Q) \le C(U, \beta)$. Consequently, by Theorem 2.13,

$$\mathrm{Var}(N(Q)) \le C(U, \beta)n^{2/3} \log n, \tag{2.14}$$

provided that n is sufficiently large. Also, by Lemma 2.8 and our choice of k,

$$\mathbb{E}(N(Q)) = \text{Leb}(Q)n = |\mathcal{C}|8^{-k}n \geq |\mathcal{C}|.$$

Thus, by (2.13), (2.14) and Chebychev's inequality,

$$\mathbb{P}\left(\frac{N(Q)}{|\mathcal{C}|} \geq \frac{1}{2}\right) \geq 1 - \frac{4\text{Var}(N(Q))}{|\mathcal{C}|^2} \geq 1 - C(U, \beta)n^{-2/3}\log n. \qquad (2.15)$$

Now let

$$a_1 := \frac{1}{|\mathcal{C}|}\sum_{D \in \mathcal{C}} N(D) = \frac{N(Q)}{|\mathcal{C}|}, \quad a_2 := \frac{1}{|\mathcal{C}|}\sum_{D \in \mathcal{C}} N(D)^2,$$

$$p_1 := \frac{|\{D \in \mathcal{C} : N(D) > 0\}|}{|\mathcal{C}|}, \quad p_2 := \frac{|\{D \in \mathcal{C} : N(D) > m\}|}{|\mathcal{C}|},$$

$$q := \frac{|\{D \in \mathcal{C} : 0 < N(D) \leq m\}|}{|\mathcal{C}|}.$$

By (2.11), $\mathbb{E}(a_2) \leq K_2(\beta)$. Thus,

$$\mathbb{P}(a_2 \geq 2K_2(\beta)) \leq \frac{1}{2}. \qquad (2.16)$$

By the Paley–Zygmund second moment inequality,

$$p_1 \geq \frac{a_1^2}{a_2},$$

and so by (2.15) and (2.16),

$$\mathbb{P}\left(p_1 \geq \frac{1}{8K_2(\beta)}\right) \geq \mathbb{P}\left(a_1 \geq \frac{1}{2}, a_2 \leq 2K_2(\beta)\right) \geq \frac{1}{2} - C(U, \beta)n^{-2/3}\log n.$$

Choose n so large that the above lower bound at least $1/3$. Next, note that by Lemma 2.8 and Markov's inequality,

$$\mathbb{E}(p_2) \leq \frac{1}{m|\mathcal{C}|}\sum_{D \in \mathcal{C}} \mathbb{E}(N(D)) \leq \frac{8}{m},$$

and hence

$$\mathbb{P}\left(p_2 \geq \frac{32}{m}\right) \leq \frac{1}{4}.$$

Since $q = p_1 - p_2$ and

$$\frac{1}{8K_2(\beta)} \geq \frac{64}{m},$$

this gives

$$\mathbb{P}\left(q \geq \frac{32}{m}\right) \geq \mathbb{P}\left(p_1 \geq \frac{64}{m}, \ p_2 \leq \frac{32}{m}\right) \geq \frac{1}{3} - \frac{1}{4} = \frac{1}{12}. \qquad (2.17)$$

Let $\mathcal{C}_0$ be the set of all $D \in \mathcal{C}$ such that $0 < N(D) \leq m$. Let $\mathcal{C}'_0$ be the set of all elements of $\mathcal{C}'$ that are contained in elements of $\mathcal{C}_0$. Let

$$r := \frac{1}{|\mathcal{C}'_0|} \sum_{D \in \mathcal{C}'_0} N(D) \qquad (2.18)$$

if $\mathcal{C}'_0 \neq \emptyset$ and let $r = 0$ otherwise. By Lemma 2.7, if $\mathcal{C}'_0$ is nonempty,

$$\mathbb{E}(r|\mathcal{F}_k) = \frac{1}{8^{j-k}|\mathcal{C}_0|} \sum_{D \in \mathcal{C}_0} N(D) \geq C(\beta), \qquad (2.19)$$

and by Lemmas 2.6 and 2.7,

$$\mathrm{Var}(r|\mathcal{F}_k) \leq \frac{C(\beta)}{|\mathcal{C}'_0|} = \frac{C(\beta)}{|\mathcal{C}|q} \leq \frac{C(\beta)}{n^{2/3}q}. \qquad (2.20)$$

By the last two inequalities and Chebychev's inequality, we see that there is a positive constant $K_3(\beta)$ depending only on β such that if $q \geq 32/m$ and n is sufficiently large, then

$$\mathbb{P}(r \geq K_3(\beta)|\mathcal{F}_k) \geq 1 - C(U, \beta)n^{-2/3}.$$

Therefore by (2.17), if n is sufficiently large,

$$\mathbb{P}\left(r \geq K_3(\beta), \ q \geq \frac{32}{m}\right) \geq \frac{1}{13}. \qquad (2.21)$$

Thus, for sufficiently large n,

$$\mathbb{P}(|\mathcal{C}'_0| \geq K_4(\beta)n^{2/3}) \geq \frac{1}{13},$$

where $K_4(\beta)$ is a positive constant that depends only on β.

Now recall the event E defined earlier. Let E^c denote the complement of E. If E^c happens, then $\mathcal{C}'_0 = \mathcal{C}^*$, where

$$\mathcal{C}^* := \{D \in \mathcal{C}'_0 : N(D) = 1\}. \tag{2.22}$$

Combining this with (2.12), this shows that for sufficiently large n,

$$\begin{aligned}
\mathbb{P}(|\mathcal{C}^*| \geq K_4(\beta)n^{2/3}) &\geq \mathbb{P}(\{|\mathcal{C}'_0| \geq K_4(\beta)n^{2/3}\} \cap E^c) \\
&\geq \mathbb{P}(|\mathcal{C}'_0| \geq K_4(\beta)n^{2/3}) - \mathbb{P}(E) \geq \frac{1}{14}. \tag{2.23}
\end{aligned}$$

By Lemma 2.6, the random variables $\{N(D \cap U) : D \in \mathcal{D}_j\}$ are independent given $\mathcal{F}_j$. If $N(D) = 1$, then the conditional distribution of $N(D \cap U)$ given $\mathcal{F}_j$ is Bernoulli($p(D)$), where $p(D) = \mathrm{Leb}(D \cap U)/\mathrm{Leb}(D)$. Let

$$M := \sum_{D \in \mathcal{C}^*} N(D \cap U).$$

Since $10^{-8} \leq p(D) \leq 1 - 10^{-8}$ for each $D \in \mathcal{C}^*$, the Berry–Esseen theorem for sums of independent random variables shows that for any interval I,

$$\mathbb{P}(M \in I | \mathcal{F}_j) \leq \frac{C(|I| + 1)}{\sqrt{|\mathcal{C}^*|}}, \tag{2.24}$$

where $|I|$ denotes the length of I. Since

$$N(U) = \sum_{D \in \mathcal{D}_j} N(D \cap U) = \sum_{D \in \mathcal{D}_j \setminus \mathcal{C}^*} N(D \cap U) + M,$$

and the two terms in the last expression are independent given $\mathcal{F}_j$, the inequality (2.24) implies that

$$\mathbb{P}(N(U) \in I | \mathcal{F}_j) \leq \frac{C(|I| + 1)}{\sqrt{|\mathcal{C}^*|}}.$$

Therefore by (2.23),

$$\mathbb{P}(N(U) \in I) \leq C(\beta)(|I| + 1)n^{-1/3} + \frac{13}{14}$$

if n is sufficiently large. This completes the proof. $\square$

Finally, let us prove Theorem 1.5. The ingredients are almost all drawn from the proof of Theorem 1.2.

Proof of Theorem 1.5 In this proof, $C(\beta)$ denotes any positive constant that depends only on β, $C(f)$ denotes any positive constant that depends only on f and $C(f, \beta)$ denotes any positive constant that depends only on f and β. Let j and k be defined as in (2.10). Let $f : [0, 1]^3 \to \mathbb{R}$ be a non-constant linear function.

Let $\mathcal{C} := \mathcal{D}_k$, and let a_1, a_2, p_1, p_2 and q be defined as in the proof of Theorem 1.2, with this $\mathcal{C}$. Then $|\mathcal{C}| = 8^k$, and $a_1 = 8^{-k}n \geq 1$. The inequality (2.16) is still valid, and hence we get

$$\mathbb{P}\left(p_1 \geq \frac{1}{8K_2(\beta)}\right) \geq \frac{1}{2}.$$

Proceeding then as in the proof of Theorem 1.2, this gives

$$\mathbb{P}\left(q \geq \frac{32}{m}\right) \geq \frac{1}{4}.$$

Let $\mathcal{C}_0$ be the set of all $D \in \mathcal{C}$ for which $0 < N(D) \leq m$. Construct a set $\mathcal{C}'_0 \subseteq \mathcal{D}_j$ by choosing exactly one descendant of each element of $\mathcal{C}_0$ by some arbitrary deterministic rule. Let r be defined as in (2.18). Then (2.12), (2.19) and (2.20) continue to hold, and therefore so does (2.21) when n is sufficiently large. Since $|\mathcal{C}| \geq n$ in this proof, this shows that for sufficiently large n,

$$\mathbb{P}(|\mathcal{C}^*| \geq K_5(\beta)n) \geq \frac{1}{14}, \tag{2.25}$$

where $\mathcal{C}^*$ is defined as in (2.22) and $K_5(\beta)$ is a positive constant that depends only on β.

For each $D \in \mathcal{D}_j$, let

$$X(f, D) := \sum_{i:X_i \in D} f(X_i).$$

By Lemma 2.6, the random variables $\{X(f, D) : D \in \mathcal{D}_j\}$ are conditionally independent given $\mathcal{F}_j$. Let

$$M := n^{1/3} \sum_{D \in \mathcal{C}^*} X(f, D).$$

Now take any $D \in \mathcal{C}^*$. Recall that D contains exactly one point of our point process, and by Lemma 2.6, the conditional distribution of this point given $\mathcal{F}_j$ is uniform over the cube D. Since f is a linear function, it is easy to see from this observation that for any $D \in \mathcal{C}^*$, the conditional distribution of the random variable

$$n^{1/3}(X(f, D) - \mathbb{E}(X(f, D)))$$

given $\mathcal{F}_j$ is actually non-random, and depends only on f. In particular, since f is also non-constant, this shows that

$$\text{Var}(n^{1/3}X(f, D)|\mathcal{F}_j) = K_6(f)$$

and

$$\mathbb{E}\big(|n^{1/3}X(f, D) - \mathbb{E}(n^{1/3}X(f, D))|^3\big|\mathcal{F}_j\big) = K_7(f),$$

where $K_6(f)$ and $K_7(f)$ are strictly positive constants that depend only on f. Therefore by the Berry–Esseen theorem, for any interval I,

$$\mathbb{P}(M \in I|\mathcal{F}_j) \leq \frac{C(f)(|I| + 1)}{\sqrt{|\mathcal{C}^*|}}, \tag{2.26}$$

where $|I|$ denotes the length of I. Since

$$n^{1/3}X(f) = n^{1/3} \sum_{D \in \mathcal{D}_j} X(f, D) = n^{1/3} \sum_{D \in \mathcal{D}_j \backslash \mathcal{C}^*} X(f, D) + M,$$

and the two terms in the last expression are independent given $\mathcal{F}_j$, the inequality (2.26) implies that

$$\mathbb{P}(n^{1/3}X(f) \in I|\mathcal{F}_j) \leq \frac{C(f)(|I| + 1)}{\sqrt{|\mathcal{C}^*|}}.$$

Therefore by (2.25),

$$\mathbb{P}(n^{1/3}X(f) \in I) \leq \frac{C(f, \beta)(|I| + 1)}{\sqrt{n}} + \frac{13}{14}$$

if n is sufficiently large. This completes the proof. $\qquad\square$

3 Proofs in 2D and 1D

In this section, we will prove the results of Sect. 1.5. The proofs are similar to the proofs in the 3D case, but there are substantial differences, which is why we need a separate section.

3.1 Notation

All notation will remain the same as in the 3D case. For example, $\mathcal{D}_k$ will denote dyadic sub-squares of side-length 2^{-k} in 2D, and dyadic sub-intervals of length 2^{-k} in 1D. The main change is that w is now different, namely, $w(x, y) = k(x, y)$, where $k(x, y)$ is the smallest k such that x and y belong to distinct elements of $\mathcal{D}_k$. The

partition function $Z(n, \beta)$ and the measure $\mu_{n,\beta}$ are defined as before, with this new w instead of the old one. We will denote the dimension by d, which may be 1 or 2.

3.2 Preliminary calculations

First, let us carry out the calculations analogous to those done in Sect. 2.2.

Lemma 3.1 *For each $x \in [0, 1)^d$,*

$$\int w(x, y) \, dy = \frac{2^d}{2^d - 1}.$$

Consequently,

$$\iint w(x, y) \, dx \, dy = \frac{2^d}{2^d - 1}.$$

Proof Take any x. For each k, let D_k be the element of $\mathcal{D}_k$ that contains x. It is easy to see that the set of all y with $w(x, y) = k$ is exactly the union of all members of $\mathcal{D}_k$ that are contained in D_{k-1}, except the one that contains x. The Lebesgue measure of this set is $2^{-dk}(2^d - 1)$. Thus,

$$\int w(x, y) \, dy = (2^d - 1) \sum_{k=1}^{\infty} k 2^{-dk} = \frac{2^d}{2^d - 1}.$$

The second assertion is obvious from the first. □

Let us now investigate energy-minimizing configurations of finite size. As before, L_n will denote the minimum possible energy of a configuration of n points. The following result gives upper and lower bounds for L_n in dimensions one and two.

Theorem 3.2 *There is a positive constant C_1 such that for each $n \geq 2$,*

$$\binom{n}{2} \frac{2^d}{2^d - 1} - C_1 n \log n \leq L_n \leq \binom{n}{2} \frac{2^d}{2^d - 1}.$$

Proof The proof of the upper bound is exactly the same as in Theorem 2.2. For the lower bound, let k be an integer such that

$$n^{-1/d} \leq 2^{-k} \leq 2n^{-1/d}.$$

Take any configuration of n points. For each $D \in \mathcal{D}$, let n_D be the number of points in D. Summing up the contributions to the energy from each cube, we get

$$H_n(x_1, \ldots, x_n) = \sum_{j=1}^{\infty} \sum_{D \in \mathcal{D}_j} \binom{n_D}{2} + \binom{n}{2}$$

$$\geq \sum_{j=1}^{k} \sum_{D \in \mathcal{D}_j} \binom{n_D}{2} + \binom{n}{2} = \frac{1}{2} \sum_{j=1}^{k} \sum_{D \in \mathcal{D}_j} n_D^2 - \frac{nk}{2} + \binom{n}{2}.$$

By the Cauchy–Schwarz inequality, for each j,

$$\sum_{D \in \mathcal{D}_j} n_D^2 \geq \frac{1}{|\mathcal{D}_j|} \left(\sum_{D \in \mathcal{D}_j} n_D \right)^2 = \frac{n^2}{2^{dj}}.$$

Thus,

$$H_n(x_1, \ldots, x_n) \geq \frac{n^2}{2} \sum_{j=1}^{k} 2^{-dj} - \frac{nk}{2} + \binom{n}{2} = \frac{n^2}{2} \frac{1 - 2^{-dk}}{2^d - 1} - \frac{nk}{2} + \binom{n}{2}.$$

By our choice of k, this completes the proof. $\square$

3.3 Estimates for the partition function

Recall that for a measurable function $f : \Sigma_n \to \mathbb{R}$, its expected value under $\mu_{n,\beta}$ is denoted by $\mu_{n,\beta}(f)$.

Lemma 3.3 *There is a constant C_2 such that for any $n \geq 0$ and $\beta > 0$,*

$$\exp\left(-\frac{2^d \beta n}{2^d - 1} \right) \leq \frac{Z(n+1, \beta)}{Z(n, \beta)} \leq \exp\left(-\frac{2^d \beta n}{2^d - 1} + C_2 \log(n+1) \right).$$

Proof The proof of Lemma 2.3 goes through verbatim, the only change being that we need to use Theorem 3.2 instead of Theorem 2.2. $\square$

Corollary 3.4 *For any $n \geq 0$, $\beta > 0$, and any $k \geq -n$,*

$$\frac{Z(n+k, \beta)}{Z(n, \beta)} \leq \exp\left(-\frac{2^d \beta nk}{2^d - 1} - \frac{2^d \beta k(k-1)}{2(2^d - 1)} + C_2 \beta |k| \log(n + |k| + 1) \right),$$

where C_5 is the constant from Lemma 3.3.

Proof Again, the proof of Corollary 2.4 goes through verbatim, except that we need to use Lemma 3.3 instead of Lemma 2.3. $\square$

3.4 Proofs of the upper bounds

Let us now fix some $n \geq 0$ and $\beta > 0$. In the following, $(X_1, \ldots, X_n)$ will denote a random configuration drawn from the measure $\mu_{n,\beta}$. We will assume that $(X_1, \ldots, X_n)$ is defined on some abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation, variance and covariance with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$, Var and Cov respectively.

Lemma 3.5 *Let* $D_1, \ldots, D_{2^d}$ *denote the* 2^d *elements of* $\mathcal{D}_1$, *and for each* $1 \leq i \leq 2^d$, *let* $N_i := |\{j : X_j \in D_i\}|$. *Then for each* i, $\mathbb{E}(N_i) = n/2^d$ *and*

$$\mathrm{Var}(N_i) \leq K(\beta)(\log(n+1))^2,$$

where $K(\beta)$ *is a non-increasing function of* β.

Proof We have already defined universal constants C_1 and C_2 in the previous subsections. In this proof, we will denote further universal constants by $C_3, C_4, \ldots$ without explicitly mentioning that they denote universal constants on each occasion.

The identity $\mathbb{E}(N_i) = n/2^d$ follows by symmetry. We will now prove the claimed bound on the variance. The cases $n = 0$ and $n = 1$ are trivial, so assume that $n \geq 2$. As in the proof of Lemma 2.5, we have a recursion for the partition function, although the recursion is slightly different due to the different nature of the potential:

$$Z(n, \beta)$$

$$= \sum_{\substack{0 \leq n_1, \ldots, n_{2^d} \leq n \\ n_1 + \cdots + n_{2^d} = n}} \frac{n!}{n_1! n_2! \cdots n_{2^d}!} e^{-\beta \sum_{1 \leq i < j \leq 2^d} n_i n_j} \prod_{i=1}^{2^d} (2^{-dn_i} Z(n_i, \beta) e^{-\beta \binom{n_i}{2}})$$

$$= \sum_{\substack{0 \leq n_1, \ldots, n_{2^d} \leq n \\ n_1 + \cdots + n_{2^d} = n}} \frac{2^{-dn} e^{-\beta \binom{n}{2}} n!}{n_1! n_2! \cdots n_{2^d}!} \prod_{i=1}^{2^d} Z(n_i, \beta).$$

Moreover, for any $(n_1, \ldots, n_{2^d})$ occurring in the above sum,

$$\mathbb{P}(N_1 = n_1, \ldots, N_{2^d} = n_{2^d}) = \frac{2^{-dn} e^{-\beta \binom{n}{2}} n! \prod_{i=1}^{2^d} Z(n_i, \beta)}{n_1! n_2! \cdots n_{2^d}!} \frac{1}{Z(n, \beta)}.$$

Choose nonnegative integers $m_1, \ldots, m_{2^d}$ such that $m_1 + \cdots + m_{2^d} = n$ and $|m_i - n/2^d| \leq 1$ for each i. For convenience, let

$$f(n_1, \ldots, n_{2^d}) := \frac{n!}{n_1! n_2! \cdots n_{2^d}!}, \quad h(n_1, \ldots, n_{2^d}) := \prod_{i=1}^{2^d} Z(n_i, \beta).$$

Take any $k_1, \ldots, k_{2^d} \in \mathbb{Z}$ such that $k_1 + \cdots + k_{2^d} = 0$ and $0 \le m_i + k_i \le n$ for each i. Then by Corollary 3.4,

$$
\frac{h(m_1 + k_1, \ldots, m_{2^d} + k_{2^d})}{h(m_1, \ldots, m_{2^d})}
$$

$$
\le \prod_{i=1}^{2^d} \exp\left(-\frac{2^d \beta m_i k_i}{2^d - 1} - \frac{2^d \beta k_i (k_i - 1)}{2(2^d - 1)} + C_2 \beta |k_i| \log(n + |k_i| + 1) \right)
$$

$$
\le \prod_{i=1}^{2^d} \exp\left(-\frac{2^d \beta (nk_i/2^d - |k_i|)}{2^d - 1} - \frac{2^d \beta k_i (k_i - 1)}{2(2^d - 1)} + 2C_2 \beta |k_i| \log n \right)
$$

$$
\le \exp\left(-\frac{2^d \beta}{2(2^d - 1)} \sum_{i=1}^{2^d} k_i^2 + C_3 \beta \log n \sum_{i=1}^{2^d} |k_i| \right).
$$

Therefore,

$$
\frac{\mathbb{P}(N_1 = m_1 + k_1, \ldots, N_{2^d} = m_{2^d} + k_{2^d})}{\mathbb{P}(N_1 = m_1, \ldots, N_{2^d} = m_{2^d})}
$$

$$
\le \frac{f(m_1 + k_1, \ldots, m_{2^d} + k_{2^d})}{f(m_1, \ldots, m_{2^d})} \exp\left(-\frac{2\beta}{3} \sum_{i=1}^{2^d} k_i^2 + C_3 \beta \log n \sum_{i=1}^{2^d} |k_i| \right).
$$

This shows that there are positive constants C_4 and C_5 such that if

$$
\max_{1 \le i \le 2^d} |k_i| \ge C_4 \log n,
$$

then

$$
\frac{\mathbb{P}(N_1 = m_1 + k_1, \ldots, N_{2^d} = m_{2^d} + k_{2^d})}{\mathbb{P}(N_1 = m_1, \ldots, N_{2^d} = m_{2^d})}
$$

$$
\le \frac{f(m_1 + k_1, \ldots, m_{2^d} + k_{2^d})}{f(m_1, \ldots, m_{2^d})} e^{-C_5 \beta (\log n)^2}.
$$

It is now easy to complete the proof by imitating the last part of the proof of Lemma 2.5. □

For a Borel set $A \subseteq [0, 1)^d$, let $X(A)$ and $N(A)$ be defined as before. Also, define $\{\mathcal{F}_k\}_{k \ge 0}$ as before.

Lemma 3.6 *Conditional on $\mathcal{F}_k$, the random sets $\{X(D) : D \in \mathcal{D}_k\}$ are mutually independent. Moreover, for any $D \in \mathcal{D}_k$, conditional on $\mathcal{F}_k$, $X(D)$ has the same distribution as a scaled version of a point process from the measure $\mu_{N(D), \beta}$.*

Proof The proof is the same as the proof of Lemma 2.6, except that β need not be replaced by $2^k \beta$ due to the different nature of the potential. □

Lemma 3.7 *If $D \in \mathcal{D}_k$ and D' is a child of D, then*

$$\mathbb{E}(N(D')|\mathcal{F}_k) = \frac{N(D)}{2^d}$$

and

$$\mathrm{Var}(N(D')|\mathcal{F}_k) \leq K(\beta)(\log(N(D)+1))^2,$$

where K is the function from Lemma 3.5.

Proof The formula for the conditional expectation follows from Lemma 2.6 and symmetry, and the bound on the conditional variance follows from Lemmas 3.5 and 3.6. □

Lemma 3.8 *For any $D \in \mathcal{D}$, $\mathbb{E}(N(D)) = Leb(D)n$ and*

$$\mathrm{Var}(N(D)) \leq C(\beta)(\log(2^d Leb(D)n + 3))^2,$$

where $C(\beta)$ depends only on β.

Proof Suppose that $D \in \mathcal{D}_k$. The formula for the expectation follows easily by iterating the formula for the conditional expectation from Lemma 3.7, and observing that $Leb(D) = 2^{-dk}$. Next, let D' be the parent of D. Then by Lemma 3.7, the formula for expected value, and the concavity of the map $x \mapsto (\log(x+3))^2$ on the nonnegative axis,

$$\begin{aligned}
\mathbb{E}(N(D)^2) &= \mathbb{E}(N(D)^2 - (\mathbb{E}(N(D)|\mathcal{F}_{k-1}))^2) + \mathbb{E}((\mathbb{E}(N_D|\mathcal{F}_{k-1}))^2) \\
&= \mathbb{E}(\mathrm{Var}(N(D)|\mathcal{F}_{k-1})) + 2^{-2d}\mathbb{E}(N(D')^2) \\
&\leq K(\beta)\mathbb{E}((\log(N(D')+1))^2) + 2^{-2d}\mathbb{E}(N(D')^2) \\
&\leq K(\beta)\mathbb{E}((\log(N(D')+3))^2) + 2^{-2d}\mathbb{E}(N(D')^2) \\
&\leq K(\beta)(\log\mathbb{E}(N(D')+3))^2 + 2^{-2d}\mathbb{E}(N(D')^2) \\
&= K(\beta)(\log(2^{-d(k-1)}n+3))^2 + 2^{-2d}\mathbb{E}(N(D')^2).
\end{aligned}$$

Iterating this, we get

$$\mathbb{E}(N(D)^2) \leq K(\beta)\sum_{r=0}^{k-1}(\log(2^{d+rd}Leb(D)n+3))^2 2^{-2rd} + 2^{-2dk}n^2.$$

Now note that for any $r \geq 0$,

$$\begin{aligned}
\frac{\log(2^{d+rd}Leb(D)n+3)}{\log(2^d Leb(D)n+3)} &\leq \frac{\log(2^d Leb(D)n+3) + \log 2^{rd}}{\log(2^d Leb(D)n+3)} \\
&= 1 + \frac{rd\log 2}{\log(2^d Leb(D)n+3)} \leq 1 + \frac{rd\log 2}{\log 3}.
\end{aligned}$$

Thus,

$$\mathbb{E}(N(D)^2) \leq K(\beta)(\log(2^d \text{Leb}(D)n + 3))^2 \sum_{r=0}^{\infty} \left(1 + \frac{rd \log 2}{\log 3}\right)^2 2^{-2rd}$$

$$+ 2^{-2dk} n^2$$

$$\leq C(\beta)(\log(2^d \text{Leb}(D)n + 3))^2 + 2^{-2dk} n^2,$$

where $C(\beta)$ depends only on β. This completes the proof, since $\mathbb{E}(N(D)) = \text{Leb}(D)n = 2^{-dk}n$. $\qquad\square$

Now take any nonempty open set $U \subseteq [0, 1)^d$ with regular boundary, and let $A(U)$ be defined as in (2.2). Define $\mathcal{U}, \mathcal{U}_j, \mathcal{V}_j$ and M_j as in the 3D case. It is easy to see that Lemmas 2.9, 2.11 and Corollary 2.10 remain valid in the 2D and 1D cases.

Lemma 3.9 *For any $j \geq 1$ such that $\sqrt{d} \cdot 2^{-j+1} \leq \text{diam}(U)$,*

$$\text{Var}(M_j) \leq C(\beta)A(U)(\log(2^{-d(j-1)}n + 3))^2 2^{(d-1)j} + \text{Var}(M_{j-1}),$$

where $C(\beta)$ is a constant that depends only on β.

Proof In this proof, $C(\beta)$ will denote any constant that depends only on β. Equations (2.3) and (2.4) are still valid. If $D, D' \in \mathcal{U}_j \cup \mathcal{V}_j$ have different parents, then $N(D)$ and $N(D')$ are conditionally independent by Lemma 3.6, and hence the conditional covariance is zero. Otherwise, Lemma 3.7 and the Cauchy–Schwarz inequality imply that

$$|\text{Cov}(N(D), N(D')|\mathcal{F}_{j-1})| \leq C(\beta)(\log(N(D'') + 1))^2,$$

where D'' is the parent of D and D'. Thus, by Lemma 3.8 and the concavity of the map $x \mapsto (\log(x + 3))^2$ on the nonnegative real axis,

$$|\mathbb{E}(\text{Cov}(N(D), N(D')|\mathcal{F}_{j-1}))| \leq C(\beta)(\log(\text{Leb}(D'')n + 3))^2$$
$$= C(\beta)(\log(2^{-d(j-1)}n + 3))^2.$$

On the other hand, each $D \in \mathcal{U}_j \cup \mathcal{V}_j$ has at most $2^d - 1$ siblings that belong to $\mathcal{U}_j \cup \mathcal{V}_j$. Since $p(D)2^{-dj} = p(D)\text{Leb}(D) = \text{Leb}(D \cap U)$, this shows that

$$\mathbb{E}(\text{Var}(M_j|\mathcal{F}_{j-1})) \leq C(\beta)(\log(2^{-d(j-1)}n + 3))^2 \sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} 2^d p(D)$$

$$= C(\beta)(\log(2^{-d(j-1)}n + 3))^2 2^{d(j+1)} \sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} \text{Leb}(D \cap U).$$

Note that each element of

$$\bigcup_{D \in \mathcal{U}_j \cup \mathcal{V}_j} (D \cap U)$$

is within distance $\sqrt{d} \cdot 2^{-j+1}$ of ∂U. Since $\sqrt{d} \cdot 2^{-j+1} \leq \mathrm{diam}(U)$, inequality (2.2) gives

$$\sum_{D \in \mathcal{U}_j \cup \mathcal{V}_j} \mathrm{Leb}(D \cap U) \leq A(U)\sqrt{d} \cdot 2^{-j+1}.$$

Consequently,

$$\mathbb{E}(\mathrm{Var}(M_j|\mathcal{F}_{j-1})) \leq C(\beta)A(U)(\log(2^{-d(j-1)}n + 3))^2 2^{(d-1)j},$$

where $C(\beta)$ depends only on β. The proof is completed by plugging this bound into (2.3). □

We now have all the ingredients for proving the following analog of Theorem 2.13.

Theorem 3.10 (Hyperuniformity at all scales in 2D and 1D) *Let U and $N(U)$ be as in Theorem 1.1. Suppose that $\mathrm{diam}(U) \geq n^{-1/d}$. Let $A(U)$ be the constant defined in (2.2). Then*

$$\mathbb{E}(N(U)) = Leb(U)n$$

and

$$\mathrm{Var}(N(U)) \leq C(\beta)(A(U)n^{(d-1)/d} + 1)(\log(7 diam(U)^d n))^2,$$

where $C(\beta)$ is a constant that depends only on β.

Proof Throughout this proof, $C(\beta)$ will denote any constant that depends only on β. The value of $C(\beta)$ may change from line to line or even within a line.

The formula for the expectation follows from the d-dimensional version of Corollary 2.10. It remains to prove the variance bound. Choose k such that

$$\frac{1}{2}n^{-1/d} \leq \sqrt{d} \cdot 2^{-k} \leq n^{-1/d}.$$

Equation (2.5) remains valid, as does the inequality

$$\mathrm{Var}(N(U)|\mathcal{F}_k) \leq \sum_{D \in \mathcal{V}_k} p(D)N(D)^2.$$

By Lemma 3.8 and our choice of k,

$$\mathbb{E}(N(D)^2) \leq C(\beta)(\log(2^d \mathrm{Leb}(D)n + 3))^2 + \mathrm{Leb}(D)^2 n^2 \leq C(\beta)$$

for all $D \in \mathcal{V}_k$. Note that each element of

$$\bigcup_{D \in \mathcal{V}_k} (D \cap U)$$

is within distance $\sqrt{d} \cdot 2^{-k}$ of ∂U, and $p(D)2^{-dk} = \text{Leb}(D \cap U)$. Since $\sqrt{d} \cdot 2^{-k} \leq n^{-1/d} \leq \text{diam}(U)$ by our choice of k, this gives

$$\mathbb{E}(\text{Var}(N(U)|\mathcal{F}_k)) \leq C(\beta)2^{dk} \sum_{D \in \mathcal{V}_k} \text{Leb}(D \cap U)$$
$$\leq C(\beta)2^{dk} A(U)2^{-k} \leq C(\beta)A(U)n^{(d-1)/d}.$$

Let l be the smallest integer such that $\sqrt{d} \cdot 2^{-l} \leq \text{diam}(U)$. Note that $l \leq k$. Together with (2.5) and Lemma 3.9, the above inequality shows that

$$\text{Var}(N(U)) \leq C(\beta)A(U) \sum_{j=l+1}^{k} (\log(2^{-d(j-1)}n + 3))^2 2^{(d-1)j} + \text{Var}(M_l)$$

$$\leq C(\beta)A(U)(\log(2^d\text{diam}(U)^dn + 3))^2 \sum_{j=l+1}^{k} 2^{(d-1)j} + \text{Var}(M_l)$$

$$\leq C(\beta)A(U)n^{(d-1)/d}(\log(2^d\text{diam}(U)^dn + 3))^2 + \text{Var}(M_l).$$

By the definition of l, $\mathcal{U}_i$ if empty for all $i < l$. Therefore

$$M_l = \sum_{D \in \mathcal{U}_l \cup \mathcal{V}_l} p(D)N(D).$$

Note that for any $D \in \mathcal{U}_l \cup \mathcal{V}_l$, Lemma 3.8 gives

$$\text{Var}(p(D)N(D)) = p(D)^2\text{Var}(N(D)) \leq C(\beta)(\log(2^d\text{Leb}(D)n + 3))^2$$
$$= C(\beta)(\log(2^d2^{-dl}n + 3))^2$$
$$\leq C(\beta)(\log(2^d\text{diam}(U)^dn + 3))^2.$$

Moreover, it is easy to see that U intersects at most 2^d members of $\mathcal{D}_l$, and therefore $|\mathcal{U}_l \cup \mathcal{V}_l| \leq 2^d$. From these observations, we get

$$\text{Var}(M_l) \leq C(\beta)(\log(2^d\text{diam}(U)^dn + 3))^2.$$

This completes the proof of the theorem. □

Proofs of Theorems 1.6 and 1.8 These are consequences of Theorem 3.10 in the same way as Theorems 1.1 and 1.3 followed from Theorem 2.13. □

Finally, let us prove Theorem 1.9.

Proof of Theorem 1.9 The proof is very similar to the proof of Theorem 1.4, with minor modifications. As usual, $C(\beta)$ denotes any constant that depends only on β. Define

$f(D)$ and W_k as in the proof of Theorem 1.4. Then W_k is again a martingale, and Eq. (2.6) is still valid. Now choose k such that

$$n^{-1/d} \le 2^{-k} \le 2n^{-1/d}.$$

Then (2.7) continues to hold. Take any j. For each $D \in \mathcal{D}_{j-1}$, let $c(D)$ denote the set of 2^d children of D. Proceeding as in the proof of Theorem 1.4, we get

$$\text{Var}(X(f_j)|\mathcal{F}_{j-1})$$
$$= \sum_{D \in \mathcal{D}_{j-1}} \mathbb{E}\left(\left(\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D) \right)^2 \Big| \mathcal{F}_{j-1} \right).$$

Now notice that for any $D \in \mathcal{D}_{j-1}$,

$$\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D)$$
$$= \sum_{D' \in c(D)} (f(D') - f(D)) \left(N(D') - \frac{N(D)}{2^d} \right).$$

Recall that L is the Lipschitz constant of f. For any $D' \in c(D)$,

$$|f(D') - f(D)| \le \sqrt{d} L 2^{-j+1}.$$

As in the proof of Theorem 1.4,

$$\left(\sum_{D' \in c(D)} (f(D') - f(D)) \left(N(D') - \frac{N(D)}{2^d} \right) \right)^2$$
$$\le 4^{-j+3} L^2 \sum_{D' \in c(D)} \left(N(D') - \frac{N(D)}{2^d} \right)^2.$$

Therefore, by Lemma 3.7,

$$\mathbb{E}\left(\left(\sum_{D' \in c(D)} f(D')N(D') - f(D)N(D) \right)^2 \Big| \mathcal{F}_{j-1} \right)$$
$$\le 4^{-j+3} L^2 \sum_{D' \in c(D)} \text{Var}(N(D')|\mathcal{F}_{j-1})$$
$$\le 4^{-j+4} L^2 K(\beta)(\log(N(D) + 1))^2 \le 4^{-j+4} L^2 K(\beta)(\log(n + 1))^2.$$

Consequently,

$$\mathbb{E}(\mathrm{Var}(X(f_j)|\mathcal{F}_{j-1})) \leq C(\beta)L^2(\log n)^2 4^{-j}|\mathcal{D}_{j-1}| \leq C(\beta)L^2 2^{(d-2)j}(\log n)^2.$$

For $D \in \mathcal{D}_k$, let $s(D)$ be defined as in the proof of Theorem 1.4. Then as before, we have

$$\mathrm{Var}(X(f)|\mathcal{F}_k) \leq \sum_{D \in \mathcal{D}_k} \mathbb{E}((s(D) - f(D)N(D))^2|\mathcal{F}_k).$$

By the Lipschitz condition,

$$|s(D) - f(D)N(D)| \leq \sqrt{d}L2^{-k}N(D)$$

for each $D \in \mathcal{D}_k$. Thus, by Lemma 3.8 and our choice of k,

$$\mathbb{E}((s(D) - f(D)N(D))^2) \leq 4^{-k+1}L^2\mathbb{E}(N(D)^2) \leq C(\beta)L^2 4^{-k}.$$

Consequently,

$$\mathbb{E}(\mathrm{Var}(X(f)|\mathcal{F}_k)) \leq C(\beta)L^2 4^{-k}|\mathcal{D}_k| \leq C(\beta)L^2 2^{(d-2)k} \leq C(\beta)L^2 n^{(d-2)/d}.$$

The proof is now easily completed by combining the steps. □

3.5 Proofs of the lower bounds

Let us now prove Theorem 1.7. The proof is similar to the proof of Theorem 1.2, but with some significant changes due to the different nature of the potential. Let

$$\mathcal{T} := \{z + [0, 1)^2 : z \in \mathbb{Z}^2\}.$$

Lemma 3.11 *Let $\mathcal{T}$ be as above. Take any $D \in \mathcal{T}$ and any $x \in D$. Let δ be the distance of x from the boundary of D. Then any line through x bifurcates D into two parts, each of which has volume at least $\pi\delta^2/2$.*

Proof Same as the proof of Lemma 2.14. □

Lemma 3.12 *Take any $x \in \mathbb{R}^2$ and a unit vector $u = (u_1, u_2) \in S^1$. Let L be the line that contains x and is perpendicular to u. Suppose that*

$$\min\{|u_1|, |u_2|\} \geq 0.1. \tag{3.1}$$

Then there is an element $D \in \mathcal{T}$, within Euclidean distance $\sqrt{101}$ from x, which is bifurcated by the line P in such a way that each part has volume at least 6×10^{-5}.

Proof Take any $x = (x_1, x_2) \in \mathbb{R}^2$ and $u = (u_1, u_2) \in S^1$ as in the statement of the lemma. Let L_0 be the line with normal vector u that contains the origin. Define

$$y_1 = \text{sign}(u_1), \quad y_2 = -\frac{|u_1|}{u_2}.$$

Then $y = (y_1, y_2) \in L_0$. Also, we have $|y_1| = 1$, and by condition (3.1) and the fact that $|u_2| \leq 1$,

$$|y_2| = \frac{|u_1|}{|u_2|} \geq |u_1| \geq 0.1.$$

Now consider the set

$$I_1 = \{x_1 + \alpha y_1 : 0 \leq \alpha \leq 1\}.$$

Since $|y_1| = 1$, I_1 is an interval of length 1. By Lemma 2.15, I_1 has a subinterval of I_2 of length 0.25 such that any integer is at least at a distance 0.25 from I_2. Moreover, since $|y_1| = 1$, I_2 is of the form

$$\{x_1 + \alpha y_1 : a \leq \alpha \leq b\},$$

where $b - a = 0.25$. Let

$$I_3 := \{x_2 + \alpha y_2 : a \leq \alpha \leq b\}.$$

Since $|y_2| \geq 0.1$, I_3 has length at least 0.025. Thus by Lemma 2.15, I_3 contains a subinterval I_4 of length 0.00625 such that any integer is at a distance at least 0.00625 from I_4.

In particular, there is some $\alpha \in [0, 1]$ such that $x_1 + \alpha y_1 \in I_2$ and $x_2 + \alpha y_2 \in I_4$. The distance of $x_i + \alpha y_i$ from the nearest integer is at least 0.00625 for each i. Thus, the distance of the point $x + \alpha y$ from the boundary of the square $D \in \mathcal{T}$ that contains $x + \alpha y$ is at least 0.00625. By Lemma 3.11 and the fact that $x + \alpha y \in L$, this proves L bifurcates D into two parts, each of which has volume at least 6×10^{-5}. Lastly, note that

$$|(x + \alpha y) - x| \leq |y| = \sqrt{y_1^2 + y_2^2} \leq \sqrt{1 + \frac{1}{0.1^2}} \leq \sqrt{101},$$

since $|u_1| \leq 1$ and $|u_2| \geq 0.1$. This completes the proof of the lemma. $\square$

Now recall that the boundary of the set U in the statement of Theorem 1.7 is a simple smooth closed curve. In particular, we can choose a unit normal vector $u(x)$ at each $x \in \partial U$ such that the map $x \mapsto u(x)$ is smooth.

Lemma 3.13 *Take any $x \in \partial U$ such that the normal vector $u(x)$ satisfies (3.1). Then there is some j_0 depending only on U (but not on x), such that for all $j \geq j_0$, there is some $D \in \mathcal{D}_j$ at distance at most $\sqrt{101} \cdot 2^{-j}$ from x, such that*

$$10^{-5} \leq \frac{Leb(D \cap U)}{Leb(D)} \leq 1 - 10^{-5}. \tag{3.2}$$

Proof Same as the proof of Lemma 2.17, using Lemma 3.12 instead of Lemma 2.16. □

Lemma 3.14 *There is some $K_1 > 0$ and some $j_1 \geq 1$ depending only on U such that for any $j \geq j_1$, there is a set of at least $K_1 2^j$ squares $D \in \mathcal{D}_j$ that satisfy (3.2) and the union of these squares has diameter at most $diam(U)/3$.*

Proof Same as the proof of Lemma 2.18, with a small adjustment for dimension that replaces $K_1 4^j$ by $K_1 2^j$. □

Lemma 3.15 *Take any $n \geq 1$ and $\beta > 0$, and a Borel set $A \subseteq [0, 1)^2$ with $0 < Leb(A) < 1$. Let $\delta > 0$ be a number such that $\delta \leq Leb(A) \leq 1 - \delta$. Then*

$$c(\beta, n, \delta) \leq \mathrm{Var}(N(A)) \leq n^2,$$

where $c(\beta, n, \delta)$ is a positive real number that depends only on β, n and δ.

Proof The upper bound is trivial since $N(A) \leq n$. For the lower bound, the case $n = 1$ is easy, since in that case $N(A)$ is a Bernoulli($Leb(A)$) random variable. So let us take $n \geq 2$. Trivially, $Z(n, \beta) \leq 1$. Therefore, by Jensen's inequality and Lemma 3.1,

$$\mathbb{P}(N(A) = n) = \mu_{n,\beta}(A^n) \geq \int_{A^n} e^{-\beta H_n(x_1, \ldots, x_n)} \, dx_1 \cdots dx_n$$

$$\geq Leb(A^n) \exp\left(-\frac{\beta}{Leb(A^n)} \int_{A^n} H_n(x_1, \ldots, x_n) \, dx_1 \cdots dx_n\right)$$

$$\geq Leb(A^n) \exp\left(-\frac{4\beta}{3 Leb(A^n)} \binom{n}{2}\right)$$

$$\geq \delta^n \exp\left(-\frac{4\beta}{3\delta^n} \binom{n}{2}\right).$$

Similarly, if $B = [0, 1)^2 \backslash A$, then

$$\mathbb{P}(N(A) = 0) = \mu_{n,\beta}(B^n) \geq \delta^n \exp\left(-\frac{4\beta}{3\delta^n} \binom{n}{2}\right).$$

With the two lower bounds derived above, it is now easy to complete the proof, for example using Chebychev's inequality. □

Proof of Theorem 1.7 In this proof, as in the proof of Theorem 1.2, the phrase 'n sufficiently large' will mean '$n \geq n_0$, where n_0 depends only on U and β'. Also, C will denote any positive universal constant, $C(\beta)$ will denote any positive constant that depends only on β, and $C(U, \beta)$ will denote any positive constant that depends only on U and β.

Choose k such that

$$n^{-1/2} \leq 2^{-k} \leq 2n^{-1/2}.$$

Then for any $D \in \mathcal{D}_k$, Lemma 3.8 gives

$$\mathbb{E}(N(D)^2) \leq L_1(\beta), \tag{3.3}$$

where $L_1(\beta)$ is a positive integer that depends only on β. Let

$$m := 1000 L_1(\beta).$$

If n is sufficiently large, then there is a set $\mathcal{C} \subseteq \mathcal{D}_k$ that satisfies the conclusions of Lemma 3.14. In particular, arguing as in the proof of Theorem 1.2, we get

$$C_1(U, \beta) 2^k \leq |\mathcal{C}| \leq C_2(U, \beta) 2^k, \tag{3.4}$$

where $C_1(U, \beta)$ and $C_2(U, \beta)$ are positive constants that depend only on U and β. Let Q be the union of the elements of $\mathcal{C}$. Proceeding as in the proof of Theorem 1.2, and using Theorem 3.10 instead of Theorem 2.13, we get

$$\mathrm{Var}(N(Q)) \leq C(U, \beta) n^{1/2} (\log n)^2, \tag{3.5}$$

provided that n is sufficiently large. Also, by Lemma 3.8 and our choice of k,

$$\mathbb{E}(N(Q)) = \mathrm{Leb}(Q) n = |\mathcal{C}| 4^{-k} n \geq |\mathcal{C}|.$$

Thus, by (3.4), (3.5) and Chebychev's inequality,

$$\mathbb{P}\left(\frac{N(Q)}{|\mathcal{C}|} \geq \frac{1}{2}\right) \geq 1 - \frac{4\mathrm{Var}(N(Q))}{|\mathcal{C}|^2} \geq 1 - C(U, \beta) n^{-1/2} (\log n)^2. \tag{3.6}$$

Let a_1, a_2, p_1, p_2 and q be defined as in the proof of Theorem 1.2. By the inequality (3.3), $\mathbb{E}(a_2) \leq L_1(\beta)$. Thus,

$$\mathbb{P}(a_2 \geq 2L_1(\beta)) \leq \frac{1}{2}. \tag{3.7}$$

By the Paley–Zygmund second moment inequality,

$$p_1 \geq \frac{a_1^2}{a_2},$$

and so by (3.6) and (3.7),

$$\mathbb{P}\left(p_1 \geq \frac{1}{8L_1(\beta)}\right) \geq \mathbb{P}\left(a_1 \geq \frac{1}{2}, \ a_2 \leq 2L_1(\beta)\right) \geq \frac{1}{2} - C(U, \beta)n^{-1/2}(\log n)^2.$$

Choose n so large that the above lower bound at least $1/3$. Next, note that by Lemma 3.8 and Markov's inequality,

$$\mathbb{E}(p_2) \leq \frac{1}{m|\mathcal{C}|} \sum_{D \in \mathcal{C}} \mathbb{E}(N(D)) \leq \frac{4}{m},$$

and hence

$$\mathbb{P}\left(p_2 \geq \frac{16}{m}\right) \leq \frac{1}{4}.$$

Since $q = p_1 - p_2$ and

$$\frac{1}{8L_1(\beta)} \geq \frac{32}{m},$$

we get

$$\mathbb{P}\left(q \geq \frac{16}{m}\right) \geq \mathbb{P}\left(p_1 \geq \frac{32}{m}, \ p_2 \leq \frac{16}{m}\right) \geq \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Let $\mathcal{C}_0$ be the set of all $D \in \mathcal{C}$ such that $0 < N(D) \leq m$. The above inequality and (3.4) show that if n is sufficiently large, then

$$\mathbb{P}(|\mathcal{C}_0| \geq L_2(\beta)n^{1/2}) \geq \frac{1}{13}, \tag{3.8}$$

where $L_2(\beta)$ is a positive constant that depends only on β. By Lemma 3.6, the random variables $\{N(D \cap U) : D \in \mathcal{D}_k\}$ are independent given $\mathcal{F}_k$. Moreover, for each $D \in \mathcal{C}_0$, $N(D \cap U) \leq m \leq C(\beta)$, and by Lemma 3.6 and Lemma 3.15,

$$\mathrm{Var}(N(D \cap U)|\mathcal{F}_k) \geq L_3(U, \beta),$$

where $L_3(U, \beta)$ is a positive constant that depends only on U and β. (This is the crucial difference with the proof of Theorem 1.2. The scale invariance of the model in dimension two is not valid in dimension three.) Thus, if we let

$$M := \sum_{D \in \mathcal{C}_0} N(D \cap U),$$

then the Berry–Esseen theorem shows that for any interval I,

$$\mathbb{P}(M \in I | \mathcal{F}_k) \leq \frac{C(U, \beta)(|I| + 1)}{\sqrt{|\mathcal{C}_0|}}, \tag{3.9}$$

where $|I|$ denotes the length of I. Since

$$N(U) = \sum_{D \in \mathcal{D}_k} N(D \cap U) = \sum_{D \in \mathcal{D}_k \setminus \mathcal{C}_0} N(D \cap U) + M,$$

and the two terms in the last expression are independent given $\mathcal{F}_k$, the inequality (3.9) implies that

$$\mathbb{P}(N(U) \in I | \mathcal{F}_k) \leq \frac{C(U, \beta)(|I| + 1)}{\sqrt{|\mathcal{C}_0|}}.$$

Therefore by (3.8),

$$\mathbb{P}(N(U) \in I) \leq C(U, \beta)(|I| + 1)n^{-1/4} + \frac{12}{13}$$

if n is sufficiently large. This completes the proof. □

References

1. Aizenman, M., Martin, P.A.: Structure of Gibbs states of one-dimensional Coulomb systems. Commun. Math. Phys. **78**(1), 99–116 (1980)
2. Ameur, Y., Hedenmalm, H., Makarov, N.: Fluctuations of eigenvalues of random normal matrices. Duke Math. J. **159**(1), 31–81 (2011)
3. Ameur, Y., Hedenmalm, H., Makarov, N.: Random normal matrices and Ward identities. Ann. Probab. **43**(3), 1157–1201 (2011)
4. Anderson, G.W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge University Press, Cambridge (2010)
5. Bardenet, R., Hardy, A.: Monte Carlo with Determinantal Point Processes. (2016) arXiv preprint arXiv:1605.00361
6. Bauerschmidt, R., Bourgade, P., Nikula, M., Yau, H.-T.: Local Density for Two-Dimensional One-Component Plasma. (2015) arXiv preprint arXiv:1510.02074
7. Bauerschmidt, R., Bourgade, P., Nikula, M., Yau, H.-T.: The Two-Dimensional Coulomb Plasma: Quasi-free Approximation and Central Limit Theorem. (2016) arXiv preprint arXiv:1609.08582
8. Beck, J.: Irregularities of distribution. I. Acta Math. **159**(1–2), 1–49 (1987)
9. Bekerman, F., Leblé, T., Serfaty, S.: CLT for Fluctuations of β-Ensembles with General Potential (2013). arXiv preprint arXiv:1706.09663
10. Bekerman, F., Lodhia, A.: Mesoscopic Central Limit Theorem for General β-Ensembles (2016). arXiv preprint arXiv:1605.05206
11. Ben Arous, G., Zeitouni, O.: Large deviations from the circular law. ESAIM Probab. Stat. **2**, 123–134 (1998)

12. Bendikov, A.D., Grigor'yan, A.A., Pittet, C., Woess, W.: Isotropic Markov Semigroups on Ultra-Metric Spaces. (Russian) Uspekhi Mat. Nauk, **69**(4), 418, 3–102 (2014); translation in *Russian Math. Surv.*, **69**(4), 589–680

13. Bendikov, A.A., Grigor'yan, A.A., Pittet, Ch.: On a class of Markov semigroups on discrete ultra-metric spaces. Potential Anal. **37**(2), 125–169 (2012)

14. Benfatto, G., Renn, J.: Nontrivial fixed points and screening in the hierarchical two-dimensional Coulomb gas. J. Stat. Phys. **67**(5), 957–980 (1992)

15. Benfatto, G., Gallavotti, G., Nicolò, F.: The dipole phase in the two-dimensional hierarchical Coulomb gas: analyticity and correlations decay. Commun. Math. Phys. **106**(2), 277–288 (1986)

16. Berman, R.J.: Determinantal point processes and fermions on complex manifolds: large deviations and bosonization. Commun. Math. Phys. **327**(1), 1–47 (2014)

17. Bolley, F., Chafaï, D., Fontbona, J.: Dynamics of a Planar Coulomb Gas (2017). arXiv preprint arXiv:1706.08776

18. Borodin, A., Sinclair, C.D.: The Ginibre ensemble of real random matrices and its scaling limits. Commun. Math. Phys. **291**(1), 177–224 (2009)

19. Borodin, A., Gorin, V., Guionnet, A.: Gaussian asymptotics of discrete β-ensembles. Publ. Math. Inst. Hautes Études Sci. **125**, 1–78 (2017)

20. Borot, G., Guionnet, A.: Asymptotic Expansion of β Matrix Models in the Multi-Cut Regime (2013). arXiv preprint arXiv:1303.1045

21. Borot, G., Guionnet, A.: Asymptotic expansion of β matrix models in the one-cut regime. Commun. Math. Phys. **317**(2), 447–483 (2013)

22. Borot, G., Guionnet, A., Kozlowski, K.K.: Large-N asymptotic expansion for mean field models with Coulomb gas interaction. Int. Math. Res. Not. IMRN **2015**(20), 10451–10524 (2015)

23. Bourgade, P., Erdős, L., Yau, H.-T.: Bulk universality of general β-ensembles with non-convex potential. J. Math. Phys. **53**(9), 095221 (2012). 19 pp

24. Bourgade, P., Erdős, L., Yau, H.-T.: Universality of general β-ensembles. Duke Math. J. **163**(6), 1127–1190 (2014)

25. Bourgade, P., Erdős, L., Yau, H.-T.: Edge universality of beta ensembles. Commun. Math. Phys. **332**(1), 261–353 (2014)

26. Bourgade, P., Yau, H.-T., Yin, J.: Local circular law for random matrices. Probab. Theory Relat. Fields **159**(3–4), 545–595 (2014)

27. Bourgade, P., Yau, H.-T., Yin, J.: The local circular law II: the edge case. Probab. Theory Relat. Fields **159**(3–4), 619–660 (2014)

28. Bourgade, P., Erdős, L., Yau, H.-T., Yin, J.: Fixed energy universality for generalized Wigner matrices. Commun. Pure Appl. Math. **69**(10), 1815–1881 (2016)

29. Brascamp, H.J., Lieb, E.H.: Some Inequalities for Gaussian Measures and the Long-Range Order of the One-Dimensional Plasma. Functional Integration and its Applications, pp. 1–14. Clarendon Press, Oxford (1975)

30. Breuer, J., Duits, M.: Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. Commun. Math. Phys. **342**(2), 491–531 (2016)

31. Breuer, J., Duits, M.: Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. J. Am. Math. Soc. **30**(1), 27–66 (2017)

32. Castin, Y.: Basic Theory Tools for Degenerate Fermi Gases (2006). arXiv preprint arXiv:cond-mat/0612613

33. Chafaï, D., Hardy, A., Maïda, M.: Concentration for Coulomb Gases and Coulomb Transport Inequalities (2016). arXiv preprint arXiv:1610.00980

34. Chafaï, D., Gozlan, N., Zitt, P.-A.: First-order global asymptotics for confined particles with singular pair repulsion. Ann. Appl. Probab. **24**(6), 2371–2413 (2014)

35. Costin, O., Lebowitz, J.: Gaussian fluctuation in random matrices. Phys. Rev. Lett. **75**, 69–72 (1995)

36. Deift, P.A.: Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. American Mathematical Society, Providence (1999)

37. Dhar, A., Kundu, A., Majumdar, S.N., Sabhapandit, S., Schehr, G.: Exact extremal statistics in the classical 1D Coulomb gas. Phys. Rev. Lett. **119**, 060601 (2017)

38. Diaconis, P., Evans, S.N.: Linear functionals of eigenvalues of random matrices. Trans. Am. Math. Soc. **353**(7), 2615–2633 (2001)

39. Dimock, J.: The Kosterlitz–Thouless phase in a hierarchical model. J. Phys. A **23**(7), 1207–1215 (1990)

40. Dyson, F.J.: The dynamics of a disordered linear chain. Phys. Rev. **92**(6), 1331–1338 (1953)

41. Dyson, F.J.: Existence of a phase-transition in a one-dimensional Ising ferromagnet. Commun. Math. Phys. **12**(2), 91–107 (1969)
42. Forrester, P.J.: Log-Gases and Random Matrices. Princeton University Press, Princeton (2010)
43. Ghosh, S.: Determinantal processes and completeness of random exponentials: the critical case. Probab. Theory Relat. Fields **163**(3–4), 643–665 (2015)
44. Ghosh, S.: Palm measures and rigidity phenomena in point processes. Electron. Commun. Probab. **21**(85), 1–14 (2016)
45. Ghosh, S., Lebowitz, J.: Number rigidity in superhomogeneous random point fields. J. Stat. Phys. **166**(3–4), 1016–1027 (2017)
46. Ghosh, S., Lebowitz, J.L.: Fluctuations, large deviations and rigidity in hyperuniform systems: a brief survey. Indian J. Pure Appl. Math. **48**(4), 609–631 (2017)
47. Ghosh, S., Peres, Y.: Rigidity and tolerance in point processes: Gaussian zeroes and Ginibre eigenvalues. Duke Math. J. **166**(10), 1789–1858 (2017)
48. Ghosh, S., Zeitouni, O.: Large deviations for zeros of random polynomials with i.i.d. exponential coefficients. Int. Math. Res. Not. IMRN **2016**(5), 1308–1347 (2016)
49. Ginibre, J.: Statistical ensembles of complex, quaternion, and real matrices. J. Math. Phys. **6**, 440–449 (1965)
50. Girko, V.L.: The circular law. Teor. Veroyatnost. i Primenen. **29**(4), 669–679 (1984)
51. Guidi, L.F., Marchetti, D.H.U.: Renormalization group flow of the two-dimensional hierarchical Coulomb gas. Commun. Math. Phys. **219**(3), 671–702 (2001)
52. Hardy, A.: A note on large deviations for 2D Coulomb gas with weakly confining potential. Electron. Commun. Probab. **17**(19), 12 (2012)
53. Holroyd, A.E., Soo, T.: Insertion and deletion tolerance of point processes. Electron. J. Probab. **18**(74), 24 (2013)
54. Hough, J.B., Krishnapur, M., Peres, Y., Virág, B.: Zeros of Gaussian Analytic Functions and Determinantal Point Processes. American Mathematical Society, Providence (2009)
55. Jancovici, B., Lebowitz, J.L., Manificat, G.: Large charge fluctuations in classical Coulomb systems. J. Stat. Phys. **72**(3), 773–787 (1993)
56. Johansson, K., Lambert, G.: Gaussian and Non-Gaussian Fluctuations for Mesoscopic Linear Statistics in Determinantal Processes (2015). arXiv preprint arXiv:1504.06455
57. Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. **91**(1), 151–204 (1998)
58. Kappeler, T., Pinn, K., Wieczerkowski, C.: Renormalization group flow of a hierarchical Sine-Gordon model by partial differential equations. Commun. Math. Phys. **136**(2), 357–368 (1991)
59. König, W.: Orthogonal polynomial ensembles in probability theory. Probab. Surv. **2**, 385–447 (2005)
60. Kunz, H.: The one-dimensional classical electron gas. Ann. Phys. **85**, 303–335 (1975)
61. Lambert, G., Ledoux, M., Webb, C.: Stein's Method for Normal Approximation of Linear Statistics of Beta-Ensembles (2017). arXiv preprint arXiv:1706.10251
62. Leblé, T., Serfaty, S.: Fluctuations of Two-Dimensional Coulomb Gases (2016). arXiv preprint arXiv:1609.08088
63. Leblé, T., Serfaty, S.: Large Deviation Principle for Empirical Fields of Log and Riesz Gases. arXiv preprint arXiv:1502.02970. To appear in Invent. Math. (2015)
64. Leblé, T.: Local Microscopic Behavior for 2D Coulomb Gases (2015). arXiv preprint arXiv:1510.01506
65. Lebowitz, J.L.: Charge fluctuations in coulomb systems. Phys. Rev. A **27**, 1491–1494 (1983)
66. Lenard, A.: Exact statistical mechanics of a one-dimensional system with Coulomb forces. J. Math. Phys. **2**, 682–693 (1961)
67. Lenard, A.: Exact statistical mechanics of a one-dimensional system with Coulomb forces. III. Statistics of the electric field. J. Math. Phys. **4**, 533–543 (1963)
68. Marchetti, D.H.U., Perez, J.F.: The Kosterlitz–Thouless phase transition in two-dimensional hierarchical Coulomb gases. J. Stat. Phys. **55**(1–2), 141–156 (1989)
69. Marino, R., Majumdar, S.N., Schehr, G., Vivo, P.: Phase transitions and edge scaling of number variance in Gaussian random matrices. Phys. Rev. Lett. **112**, 254101 (2014)
70. Marino, R., Majumdar, S.N., Schehr, G., Vivo, P.: Number statistics for β-ensembles of random matrices: applications to trapped fermions at zero temperature. Phys. Rev. E **94**, 032115 (2016)
71. Martin, P.: Sum rules in charged fluids. Rev. Mod. Phys. **60**(4), 1075–1127 (1988)
72. Martin, P., Yalcin, T.: The charge fluctuations in classical Coulomb systems. J. Stat. Phys. **22**(4), 435–463 (1980)

73. Nazarov, F., Sodin, M.: Fluctuations in random complex zeroes: asymptotic normality revisited. Int. Math. Res. Not. IMRN **2011**(24), 5720–5759 (2011)
74. Nazarov, F., Sodin, M., Volberg, A.: Transportation to random zeroes by the gradient flow. Geom. Funct. Anal. **17**(3), 887–935 (2007)
75. Nazarov, F., Sodin, M., Volberg, A.: The Jancovici–Lebowitz–Manificat law for large fluctuations of random complex zeroes. Commun. Math. Phys. **284**(3), 833–865 (2008)
76. Pastur, L.: Limiting laws of linear eigenvalue statistics for Hermitian matrix models. J. Math. Phys. **47**(10), 103303 (2006). 22 pp
77. Peres, Y., Sly, A.: Rigidity and tolerance for perturbed lattices (2014). arXiv preprint arXiv:1409.4490
78. Petz, D., Hiai, F.: Logarithmic energy as an entropy functional. Contemporary Mathematics, vol. 1998. American Mathematical Society, Providence (1998)
79. Radin, C.: The ground state for soft disks. J. Stat. Phys. **26**(2), 365–373 (1981)
80. Rider, B., Virág, B.: The noise in the circular law and the Gaussian free field. Int. Math. Res. Not. IMRN, **2007** no. 2, Art. ID rnm006 (2007)
81. Rougerie, N., Serfaty, S.: Higher-dimensional Coulomb gases and renormalized energy functionals. Commun. Pure Appl. Math. **69**(3), 519–605 (2016)
82. Sandier, E., Serfaty, S.: 2D Coulomb gases and the renormalized energy. Ann. Probab. **43**(4), 2026–2083 (2015)
83. Serfaty, S.: Ginzburg-Landau vortices, Coulomb gases, and renormalized energies. J. Stat. Phys. **154**(3), 660–680 (2014)
84. Serfaty, S.: Coulomb gases and Ginzburg–Landau vortices. European Mathematical Society (EMS), Zürich (2015)
85. Shcherbina, M.: Fluctuations of linear eigenvalue statistics of β matrix models in the multi-cut regime. J. Stat. Phys. **151**(6), 1004–1034 (2013)
86. Soshnikov, A.: Gaussian limit for determinantal random point fields. Ann. Probab. **30**(1), 171–187 (2002)
87. Tao, T., Vu, V.: Random matrices: the circular law. Commun. Contemp. Math. **10**(2), 261–307 (2008)
88. Tao, T., Vu, V.: Random matrices: sharp concentration of eigenvalues. Random Matrices Theory Appl. **2**(3), 1350007 (2013). 31 pp
89. Tao, T., Vu, V.: Random matrices: universality of local spectral statistics of non-Hermitian matrices. Ann. Probab. **43**(2), 782–874 (2015)
90. Torquato, S., Scardicchio, A., Zachary, C.E.: Point processes in arbitrary dimension from fermionic gases, random matrix theory, and number theory. J. Stat. Mech. Theory Exp. **2008**(11), P11019 (2008)
91. Wieand, K.: Eigenvalue distributions of random unitary matrices. Probab. Theory Relat. Fields **123**(2), 202–224 (2002)