



## The spherical ensemble and quasi-Monte-Carlo designs

Downloaded from: <https://research.chalmers.se>, 2026-04-04 17:06 UTC

Citation for the original published paper (version of record):

Berman, R. (2023). The spherical ensemble and quasi-Monte-Carlo designs. *Constructive Approximation*, In Press. <http://dx.doi.org/10.1007/s00365-023-09646-0>

N.B. When citing this work, cite the original published paper.



# The spherical ensemble and quasi-Monte-Carlo designs

Robert J. Berman<sup>1</sup>

Received: 13 December 2021 / Revised: 23 August 2022 / Accepted: 5 March 2023  
© The Author(s) 2023

## Abstract

The spherical ensemble is a well-known ensemble of  $N$  repulsive points on the two-dimensional sphere, which can be realized in various ways (as a random matrix ensemble, a determinantal point process, a Coulomb gas, a Quantum Hall state...). Here we show that the spherical ensemble enjoys remarkable convergence properties from the point of view of numerical integration. More precisely, it is shown that the numerical integration rule corresponding to  $N$  nodes on the two-dimensional sphere sampled in the spherical ensemble is, with overwhelming probability, nearly a quasi-Monte-Carlo design in the sense of Brauchart-Saff-Sloan-Womersley for any smoothness parameter  $s \leq 2$ . The key ingredient is a new explicit sub-Gaussian concentration of measure inequality for the spherical ensemble.

**Keywords** Quasi-Monte-Carlo designs · Coulomb gas · Numerical integration · Concentration of measure

## 1 Introduction

How to optimally distribute  $N$  points on the two-dimensional sphere? This is a question which has a long history and appears in a wide range of areas in pure as well as applied mathematics (see the survey [32] and [8, Section 7]). The notion of optimality depends, of course, on the problem at hand. But a recurrent theme is to distribute the configuration of points  $\mathbf{x}_N := (x_1, \dots, x_N) \in X^N$  so as to minimize

$$\|\delta_N(\mathbf{x}_N) - d\sigma\|$$

---

Communicated by Peter J. Forrester.

---

✉ Robert J. Berman  
robertb@chalmers.se

<sup>1</sup> Chalmers University of Technology and the University of Gothenburg, Göteborg, Sweden

on  $X^N$ , where  $\|\cdot\|$  is a given (semi-)norm on the space of all signed measure on the two-sphere  $X$ ,  $d\sigma$  denotes the standard uniform probability measure on  $X$  and  $\delta_N(\mathbf{x}_N)$  is the *empirical measure* corresponding to  $\mathbf{x}_N$ , i.e. the discrete probability measure on  $X$  defined by

$$\delta_N(x_1, \dots, x_N) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \tag{1.1}$$

More precisely, since finding exact minimizers is usually unfeasible, the aim is typically to distribute the  $N$  points  $(x_1, \dots, x_N)$  so that  $\|\delta_N(\mathbf{x}_N) - d\sigma\|$  achieves the optimal (minimal) rate as  $N \rightarrow \infty$  (as discussed in the introduction of [26, 27, I]). Here we will be concerned with a notion of optimality which naturally appears in the context of numerical integration (cubature) and quasi-Monte-Carlo integration techniques [10, 12], where the norm in question is a Sobolev norm.

### 1.1 Background

#### 1.1.1 (Quasi-)Monte-Carlo integration on cubes

Monte-Carlo integration is a standard probabilistic technique for numerically computing the Lebesgue integral of a given, say continuous, function  $f$  over a domain  $X$  in Euclidean  $\mathbb{R}^d$  (or more generally, a Riemannian manifold  $X$ ). It consists in generating  $N$  random points  $x_1, \dots, x_N$  in  $X$ , with respect to the uniform distribution  $dx$  on  $X$  (assuming for simplicity that  $X$  has unit-volume) and approximating

$$\int_X f dx \approx \frac{1}{N} \sum_{i=1}^N f(x_i)$$

In other words, the points  $x_1, \dots, x_N$  are viewed as as independent  $\mathbb{R}^d$ -valued random variables with identical distribution  $dx$ . By the central limit theorem the error is of the order  $\mathcal{O}(N^{-1/2})$  with high probability:

$$\lim_{N \rightarrow \infty} \mathbb{P} \left( \left| \int_X f dx - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \geq \frac{\lambda}{N^{1/2}} \right) = 1 - \int_{|y| \geq \lambda} e^{-y^2} dx / \pi^{1/2} \tag{1.2}$$

if  $f$  is normalized to have unit variance.

The popular *Quasi-Monte-Carlo method* aims at improving the order of the convergence, by taking  $\mathbf{x}_N := (x_1, \dots, x_N)$  to be a judiciously constructed deterministic sequence of  $N$ -point configurations on  $X$ . In the most commonly studied case when  $X$  is the unit-cube  $[0, 1]^d$  in  $\mathbb{R}^d$  there are well-known explicit so called *low-discrepancy sequences* (e.g. digital nets) constructed using the theory of uniform distribution in

number theory [29], such that

$$\sup_{f \in C(X): V(f) \leq 1} \left| \int_X f dx - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq C_d \frac{(\log N)^d}{N} \tag{1.3}$$

where  $V(f)$  is the *Hardy-Krause variation* of  $f$ , whose general definition is rather complicated, but for  $f$  sufficiently regular it is, when  $d = 2$ , given by

$$V(f) := \int_{[0,1]^2} \left| \frac{\partial^2 f}{\partial x_1 \partial x_2} \right| dx + \int_{[0,1]} \left| \frac{\partial f}{\partial x_1}(x_1, 1) \right| dx_1 + \int_{[0,1]} \left| \frac{\partial f}{\partial x_2}(1, x_2) \right| dx_2 \tag{1.4}$$

This is a consequence of the Koksma-Hlawka inequality, which is the corner stone of the theory of quasi-Monte-Carlo integration on a cube [21, 23, 29].

### 1.1.2 Numerical integration on manifolds

Let us next recall the general setup for numerical integration on manifolds, following [9, 10]. Let  $X$  be a compact manifold that we shall take to be two-dimensional. Given a configuration  $\mathbf{x}_N \in X^N$  of  $N$  points on  $X$  the *worst-case error for the integration rule on  $X$  with node set  $\mathbf{x}_N$  with respect to the smoothness parameter  $s \in ]1, \infty[$*  is defined by

$$\text{wce}(\mathbf{x}_N; s) := \sup_{f: \|f\|_{H^s(X)} \leq 1} \left| \int_X f d\sigma - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \tag{1.5}$$

where  $d\sigma_g$  denotes the normalized volume form defined by  $g$  and  $\|f\|_{H^s(X)}$  denotes the norm in the Sobolev space  $H^s(X)$  of functions with  $s$  fractional derivatives in  $L^2(X)$ . In other words,

$$\text{wce}(\mathbf{x}_N; s) = \|\delta_N(\mathbf{x}_N) - d\sigma_g\|_{H_0^{-s}(X)},$$

where  $\delta_N(\mathbf{x}_N)$  is the empirical measure 1.1 and  $H_0^{-s}(X)$  denotes the Sobolev space of all mean zero distributions on  $X$ , endowed with the Hilbert norm which is dual to the smoothness parameter  $s \in ]1, \infty[$  (see Sect. 2.1). The role of the Hardy-Krause variation norm 1.4 on a Euclidean square will in the present two-dimensional Riemannian setting be played by the Sobolev norm with smoothness parameter 2 :

$$\|f\|_{H^2(X)} := \left( \int_X |\Delta_g f|^2 dV_g \right)^{1/2},$$

where  $\Delta_g$  denotes Laplace operator on  $C^\infty(X)$ . The worst case error  $\text{wce}(\mathbf{x}_N; s)$  is also called the *generalized discrepancy* [12] because of the similarity with the Koksma-Hlawka inequality on a cube. A sequence  $\mathbf{x}_N \in X^N$  is said to be of *convergence order*

$\mathcal{O}(N^{-\kappa})$  with respect to the smoothness parameter  $s$  if

$$\text{wce}(\mathbf{x}_N; s) \leq \mathcal{O}(N^{-\kappa})$$

The optimal convergence order is  $\mathcal{O}(N^{-s/2})$ . More precisely, by [9, Thm 2.14], there exists a positive constant  $c(s)$  such that for any sequence  $\mathbf{x}_N \in X^N$

$$\text{wce}(\mathbf{x}_N; s) \geq c(s)N^{-s/2}. \tag{1.6}$$

### 1.1.3 Quasi-Monte Carlo designs on the two-sphere

Consider now the case when  $X$  is the two-dimensional sphere endowed with the Riemannian metric induced from the standard embedding of  $X$  as the unit-sphere in Euclidean  $\mathbb{R}^3$ . We will denote by  $d\sigma$  the probability measure on  $X$  obtained by normalizing the area form of  $g$ . Following [10] a sequence of  $N$ -point configurations  $\mathbf{x}_N \in X^N$  is said to be a sequence of *Quasi-Monte-Carlo designs (QMC)* wrt the smoothness parameter  $s \in ]1, \infty[$ , if the corresponding worst case errors  $\text{wce}(\mathbf{x}_N; s)$  have optimal convergence order, i.e. if

$$\text{wce}(\mathbf{x}_N; s) = \mathcal{O}(N^{-s/2})$$

In particular, this convergence is faster than the one offered by the standard Monte-Carlo method. Indeed, as recalled above, Monte-Carlo integration gives, with high probability, an error of the order  $N^{-1/2}$  for a fixed function  $f$ , even if the function is smooth.

The notion of a QMC design is modeled on the influential notion of a *spherical  $t$ -design*  $\mathbf{x}_N \in X^N$ , introduced in [13]. In fact, as shown in [10, Thm 6], it follows from the solution of the Korevaar-Meyers conjecture in [7], that there exists a sequence of spherical  $t$ -designs  $X^N$  with  $t$  of the order  $N^{1/2}$ , which is a QMC design for any  $s \in ]1, \infty[$ . Moreover, for a fixed  $s \in ]1, 2[$  reproducing kernel techniques reveal that any sequence of maximizers  $\mathbf{x}_N(s) \in X^N$  of the generalized sum

$$\sum_{i,j \leq N} |x_i - x_j|^{2s-2}$$

is a QMC design wrt the smoothness parameter  $s$  (see [10]).

However, all the sequences of QMC designs discussed above are non-explicit for  $N$  large. Moreover, approximating them numerically is very challenging for large  $N$ , due, in particular, to an abundance of local minima of the corresponding functionals to be minimized on  $X^N$  [19, 36]. One is thus lead to wonder if probabilistic methods can be used to improve the convergence order of the standard Monte-Carlo method by taking the points  $x_1, \dots, x_N$  on the sphere to be appropriately correlated, as in repulsive particle systems? A natural class of such point processes is offered by the class of *determinantal point processes*, whose utility for Monte-Carlo type numerical integration was advocated in [2]. The main aim of the present work is to show that a particular determinantal point process on the two-sphere known as *the spherical ensemble* enjoys

quite remarkable converge properties from the point of view of numerical integration, for any smoothness parameter  $s \in ]1, 2]$ .

### 1.2 Main results for the spherical ensemble

The spherical ensemble first appeared as a *Coulomb gas*, also known as a one-component plasma, in the physics literature (see the monograph [16] and references therein). We recall that the Coulomb gas on the two-sphere  $X$  with  $N$  particles, at inverse temperature  $\beta$ , is defined by the following symmetric probability measure on  $X^N$  :

$$d\mathbb{P}_{N,\beta} := \frac{1}{Z_{N,\beta}} e^{-\beta E^{(N)}} d\sigma^{\otimes N}, \quad E^{(N)} := - \sum_{i \neq j \leq N} \frac{1}{2} \log |x_i - x_j| \quad (1.7)$$

where  $X$  has been embedded as the unit-sphere in Euclidean  $\mathbb{R}^3$ . It represents the microscopic state of  $N$  unit charge particles in thermal equilibrium on  $X$ , interacting by the Coulomb energy  $E^{(N)}$ , subject to a neutralizing uniform back-ground charge. More precisely, the spherical ensemble  $(X, \mathbb{P}_N)$  coincides with Coulomb gas on the sphere at the particular inverse temperature  $\beta = 2$ , for which the Coulomb gas becomes a determinantal point process [16, 22]. The following elegant random matrix realization of the spherical ensemble was exhibited in [24]. Consider two rank  $N$  complex matrices  $A$  and  $B$  and take their entries to be standard normal variables. Then the spherical ensemble coincides with the random point process defined by the eigenvalues of  $AB^{-1}$  in the complex plane, when mapped to the two-sphere, using stereographic projection.

In the present work it is shown that a random  $N$ -point configuration  $\mathbf{x}_N$  in the spherical ensemble is, with overwhelming probability, nearly a Quasi-Monte Carlo design for any  $s \in ]1, 2]$  :

**Theorem 1.1** *Consider the spherical ensemble with  $N$  particles. There exists a constant  $C$  such that for any given  $R$  in  $[(\log N)^{-1/2}, N(\log N)^{-1/2}]$*

$$\mathbb{P}_N \left( wce(\mathbf{x}_N; 2) \leq CR \frac{(\log N)^{1/2}}{N} \right) \geq 1 - \frac{1}{NR^2/C-C}$$

*As a consequence, for any  $s \in ]1, 2]$ , there exists a constant  $C(s)$  such that for any given  $R$  in  $[(\log N)^{-1/2}, N(\log N)^{-1/2}]$*

$$\mathbb{P}_N \left( wce(\mathbf{x}_N; s) \leq C(s)R^{s/2} \frac{(\log N)^{s/4}}{N^{s/2}} \right) \geq 1 - \frac{1}{NR^2/C-C}.$$

Note that for  $N$  large the restrictions on the parameter  $R$  are essentially negligible. That the estimate in the case  $s = 2$  implies the one for  $s < 2$  follows from [10, Lemma 26]. Interestingly, the worst case error in the case  $s = 2$  is similar to the worst case error of low-discrepancy sequences on the cube (formula 1.3). It should, however, be stressed that on the two-sphere there are no explicitly constructed sequences  $(\mathbf{x}_N)$

saturating the optimal rate  $wce(\mathbf{x}_N; 2) = \mathcal{O}(N^{-1})$ , even if logarithmic factors are included (see the discussion in Sect. 1.5.1).

An interesting feature of the proof of the previous theorem is that it yields attractive values on the constants in question. Indeed, for any  $\eta > 2/\log N$  the following explicit bound is obtained:

$$\mathbb{P}_N \left( wce(\mathbf{x}_N; 2) \leq e \left( \frac{1 + \eta}{4\pi} \right)^{1/2} \frac{(\log N)^{1/2}}{N} \right) \geq 1 - \frac{(\frac{1}{2}e\eta \log N)^2}{N^\eta}, \quad (1.8)$$

where  $\log N$  denotes the natural logarithm, assuming that  $N \geq 1000$  (otherwise, Euler’s number  $e$  appearing in the left hand side has to be replaced by a slightly larger constant). For example, when  $N = 1000$  this yields the worst-case-error bound  $wce(\mathbf{x}_N; 2) < 0.003$  with 99.9% confidence (by taking  $\eta = 2$ ). Moreover,  $N = 10000$  yields  $wce(\mathbf{x}_N; 2) < 0.0004$  with 99.999% confidence.

The previous theorem should be compared with the conjecture in [10], supported by numerical simulations, saying that minimizers  $\mathbf{x}_N$  of the logarithmic energy  $E^{(N)}$  (formula 1.7) are QMC designs for  $s \in ]1, 3]$ . However, a practical advantage of the spherical ensemble is that it can be simply generated by employing  $\mathcal{O}(N^3)$  elementary operations (using its random matrix representation), while no polynomial time algorithm for constructing near-minimizers of  $E^{(N)}$  is known [34, Problem 7]. Concerning the sharpness of the inequalities in the previous theorem we note that the restriction to  $s \leq 2$  is necessary, as follows from formula 1.10 below. Moreover, the power  $1/2$  of  $\log N$  appearing in the inequalities for  $s = 2$  can be expected to be optimal.

Theorem 1.1 will be deduced from a new concentration of measure inequality in Sobolev spaces (Theorem 1.3 below), which, in turn, will follow from the following bound on the moment generating function of the square of the random variable  $wce(\mathbf{x}_N; s)$ .

**Theorem 1.2** *For any  $\epsilon > 0$  and  $\alpha \in ]0, 4\pi 2^\epsilon[$*

$$\mathbb{E}_N \left( e^{\alpha N^2 (wce(\mathbf{x}_N; 2+\epsilon))^2} \right) \leq \left( \det \left( I - \frac{\alpha}{2\pi} \Delta_g^{-(1+\epsilon)} \right) \right)^{-1/2} < \infty$$

where  $\Delta_g$  denotes the Laplace operator on  $X$  and  $\det(I - \lambda \Delta_g^{-(1+\epsilon)})$  is the Fredholm (spectral) determinant of its fractional power  $\Delta_g^{-(1+\epsilon)}$  (see Sect. 2.1).

In fact, this is an asymptotic equality, as  $N \rightarrow \infty$  (as will be shown elsewhere [3]). Combining Theorem 1.2 with some spectral theory the following concentration of measure inequality for the spherical ensemble is obtained:

**Theorem 1.3** *There exist explicit constants  $A_1, A_2$  and  $A_3$  such that for any positive integer  $N$  and  $\epsilon > 0$*

$$\mathbb{P}_N \left( \|\delta_N - d\sigma\|_{H_0^{-(2+\epsilon)}(X)} > \delta \right) \leq e^{-A_1 N^2 \delta^2 + \frac{A_2}{\epsilon} + A_3}.$$

### 1.3 Outlook on the the case of general compact surfaces

The results for the two-sphere can be generalized to any compact two-dimensional Riemannian manifold  $X$ . Here we will just highlight the main points, deferring details to [3]. Given a Riemannian surface  $(X, g)$  of strictly positive genus denote by  $g_c$  the unique Riemannian metric on  $X$  with constant curvature which is conformally equivalent to  $g$ . To  $(X, g_c)$  one can attach a canonical  $N$ -particle determinantal point process  $(X^N, d\mathbb{P}_N)$ , which can be viewed as a higher genus generalization of the spherical ensemble [4]. In this general setting the bound in Theorem 1.2 holds up to multiplying the right hand side with a factor of the form  $(1 + e^{-\delta/N})$ , for an explicit positive constant  $\delta$ , depending on the injectivity radius of  $(X, g_c)$ . This is shown in essentially the same way as in the spherical setting, using the general Moser-Trudinger type inequalities in [4] as a replacement for the inequalities 1.11 recalled below. The analogs of Theorems 1.1, 1.3 then follow as before. In particular,

$$\left\| \frac{1}{N} \sum_{i=1}^N \delta_{x_i} - dV_{g_c} \right\|_{H_0^{-2}(X)} = \mathcal{O} \left( \frac{(\log N)^{1/2}}{N} \right)$$

holds with probability  $1 - \mathcal{O}(1/N^\infty)$  (in the sense of Theorem 1.1). Expressed in terms of the original Riemannian metric  $g$  this means that introducing the “weight function”

$$w := dV_g/dV_{g_c}$$

and sampling a configuration  $(x_1, \dots, x_N)$  in the canonical  $N$ -particle ensemble  $(X^N, d\mathbb{P}_N)$

$$\begin{aligned} \mathbb{P}_N \left( \sup_{f: \|f\|_{H_0^2(X)} \leq 1} \left| \int_X f dV_g - \frac{1}{N} \sum_{i=1}^N f(x_i)w(x_i) \right| \leq \mathcal{O} \left( \frac{(\log N)^{1/2}}{N} \right) \right) \\ \geq 1 - \mathcal{O}(1/N^\infty) \end{aligned} \tag{1.9}$$

(in the sense of Theorem 1.1). In fact, the original Riemannian metric  $g$  also induces a determinantal  $N$ -particle point process on  $X$ . In physical terms, the corresponding probability measure  $d\mathbb{P}_g^N$  on  $X^N$  represents the probability density for an integer Quantum Hall state, i.e. an  $N$ -particle state of electrons confined to  $(X, g)$  subject to a constant magnetic field, whose strength is proportional to  $N$ . However, as explained in [3], the error in the corresponding estimate 1.9 will be of the larger order  $\mathcal{O}(1/N^{1/2})$  (unless  $g$  has constant curvature). Accordingly, replacing the original metric  $g$  with the constant curvature one  $g_c$  is analogous to the use of *importance sampling* in the standard Monte-Carlo method. Recall that the latter method amounts to calculating integrals  $\int_X f dV_g$  by taking the points  $x_i$  to be independent realizations of a “target measure”  $\nu$  (taken to be different than  $dV_g$  with the aim of reducing the variance; compare the discussion in [2, Section 1.2])

It should be stressed, however, that one advantage of the spherical setting is that all the constants can be explicitly estimated, while, in the general setting, the constants depend on spectral invariants of  $(X, g_c)$ . Moreover, from a practical point of view the random matrix realization of the spherical ensemble offers a convenient implementation algorithm, while the general algorithm for simulating determinantal point processes [22] has to be employed for a general surface  $X$  (which, loosely speaking, replaces the task of finding the  $N$  eigenvalues with Gram-Schmidt orthogonalization).

### 1.4 Outline of the proof of Theorem 1.1

As shown in [30], for a fixed function  $f$  on  $X$  the following Central Limit Theorem (CLT) holds for the spherical ensemble: for any  $f \in H^1(X)$ , normalized so that  $\int |\nabla_g f|^2 dV_g = 4\pi$ ,

$$\lim_{N \rightarrow \infty} \mathbb{P}_N \left( \left| \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_X f d\sigma \right| \geq \frac{\lambda}{N} \right) = 1 - \int_{|y| \geq \lambda} e^{-y^2} dx / \pi^{1/2} \quad (1.10)$$

A key ingredient in the proof of Theorem 1.1, given in Sect. 2, is the following quantitative refinement of the previous CLT, obtained in [4]:

$$\mathbb{P}_N \left( \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_X f d\sigma \leq \frac{\lambda}{N} \right) \leq e^{-\lambda^2/2}$$

More precisely, the following slightly stronger dual bound on the moment generating function was established in [4] (using complex differential geometry):

$$\mathbb{E}_N \left( \exp N(N+1)\lambda \left( \frac{1}{N} \sum_{i=1}^N f(x_i) - \int_X f d\sigma \right) \right) \leq e^{N(N+1)\lambda^2/2} \quad (1.11)$$

(coinciding when  $N = 1$  with the well-known sharp Moser-Trudinger inequality on the two-sphere). Note, however, that these inequalities only hold for a *fixed* normalized function  $f \in H^1(X)$  and fail drastically for the random variable  $wce(\mathbf{x}_N; 1)$  obtained by taking the sup over all normalized  $f \in H^1(X)$ . Indeed,  $wce(\mathbf{x}_N; 1) = \infty$  on all of  $X^N$  since  $H^1(X)$  contains unbounded functions (recall that  $s = 1$  is the borderline case for the Sobolev embedding of  $H^s(X)$  into  $C(X)$ ).

Here we will interpret the inequality 1.11 as the statement that the random variable

$$Y_N := \sum_{i=1}^N \delta_{x_i} - N d\sigma$$

taking values in the dual Sobolev space  $H^{-(2+\epsilon)}(X)$  is *sub-Gaussian* wrt a canonical Gaussian random variable  $G$  on  $H^{-(2+\epsilon)}(X)$  (see Remark 3.2). Using some basic Gaussian measure theory and spectral theory for the Laplacian we then deduce the

moment bound in Theorem 1.2, which, in turn, implies the concentration of measure inequality Theorem 1.3. Finally, we show that the latter inequality, implies Theorem 1.1, when combined with the results in [9, 10] relating  $wce(\mathbf{x}_N; s)$  corresponding to different values of  $s$ .

### 1.5 Further comparison with previous results

As shown in [20, Thm 1], building on [1], the spherical ensemble satisfies the following asymptotics: for any  $s \in ]1, 2[$  there exists a positive constant  $C(s)$  such that

$$\sqrt{\mathbb{E}(wce(\mathbf{x}_N; s)^2)} = C(s)N^{s/2} + o(N^{s/2}) \tag{1.12}$$

This result should be compared with [10, Thm 24], which says that if  $X$  is partitioned into  $N$  equal area regions whose diameters are bounded by  $CN^{-1/2}$  and a sequence  $\mathbf{x}_N$  of  $N$  point is randomly chosen from  $N$  different regions, then the corresponding  $\sqrt{\mathbb{E}(wce(\mathbf{x}_N; s)^2)}$  is also of the order  $\mathcal{O}(N^{s/2})$ , when  $s \in ]1, 2[$ . However, in contrast to Theorem 1.1, the results in [10, 20], referred to above, do not give any information about the probability that the worst-case-error  $wce(\mathbf{x}_N; s)$  for a random sequence  $\mathbf{x}_N$  in the corresponding ensembles is close to the average worst-case-error. The only previous result in this direction appears to be [1, Thm 1.1], saying that for any  $M > 0$  there exists  $C_M > 0$  such that

$$\mathbb{P}\left(D_{L^\infty}^C(\mathbf{x}_N) \leq C_M \frac{(\log N)^{1/2}}{N^{3/4}}\right) \geq 1 - \frac{1}{NM}. \tag{1.13}$$

where  $D_{L^\infty}^C(\mathbf{x}_N)$  is the  $L^\infty$ -spherical cap discrepancy defined by

$$D_{L^\infty}^C(\mathbf{x}_N) := \sup_{f=1_C} \left| \int_X f d\sigma - \frac{1}{N} \sum_{i=1}^N f(x_i) \right|$$

where the sup is taken over all characteristic  $f$  functions of the form  $f = 1_C$ , where  $C$  is a spherical cap in the two-sphere  $X$ , i.e. the intersection of  $X$  with a half-space in  $\mathbb{R}^3$  (the proof of the inequality 1.13 is based on a variance estimate in [1]). Since

$$wce\left(\mathbf{x}_N; \frac{3}{2}\right) \leq a D_{L^\infty}^C(\mathbf{x}_N) \tag{1.14}$$

for an explicit constant  $a$  [10, Page 16]), the inequality 1.13 implies that

$$\mathbb{P}\left(wce\left(\mathbf{x}_N; \frac{3}{2}\right) \leq C_M \frac{(\log N)^{1/2}}{N^{3/4}}\right) \geq 1 - \frac{1}{NM}. \tag{1.15}$$

This is a bit weaker than the case  $s = 3/2$  of Theorem 1.1 (where the power of  $\log N$  is  $3/8 (< 1/2)$  and moreover the dependence of  $C_M$  on  $M$  is made explicit). We recall

that the inequality 1.14 follows from the fact that  $wce(\mathbf{x}_N; \frac{3}{2})$  is comparable to the  $L^2$ -spherical cap discrepancy

$$D_{L^2}^C(\mathbf{x}_N) := \int \left| \int_X f d\sigma - \frac{1}{N} \sum_{i=1}^N f(x_i) \right|^2 Df, \quad f = 1_C$$

where  $Df$  is a certain probability measure on the space of all spherical caps  $C$  [10, Page 16].

### 1.5.1 Explicit sequences for numerical integration on the sphere

As recalled above,  $wce(\mathbf{x}_N; \frac{3}{2})$  is comparable to the  $L^2$ -spherical cap discrepancy  $D_{L^2}^C(\mathbf{x}_N)$ . In [26, 27] the representation theory of Hecke operators and modular forms was used to obtain an explicit sequence satisfying the bound  $D_{L^2}^C(\mathbf{x}_N) \leq CN^{-1/2} \log N$  (see [26, 27, I, Thm 2.2]). The proof of the bound uses Deligne’s proof of the Weil conjectures and also yields, as explained in [9, Remark 3.10],  $wce(\mathbf{x}_N; s) \leq CN^{-1/2} \log N$  for any  $s > 1$ . However, these rates are only close to optimal as  $s$  approaches 1. A different sequence satisfying  $D_{L^2}^C(\mathbf{x}_N) \leq CN^{-1/2}(\log N)^{1/2}$  was then constructed in [14], by mapping a digital net on the square to  $X$ . Numerical evidence was provided in [14] indicating that the latter sequence has the optimal rate  $\mathcal{O}(N^{-3/4})$ . See also [10, Section 8] for numerical experiments for a range of different classes of point sets on the two-sphere.

### 1.5.2 Concentration of measure

It may be illuminating to compare Theorem 1.1 for  $s = 2$  with the concentration of measure inequalities for *independent* random variables established in [6], which can be viewed as a quantitative refinement of the classical CLT 1.2. In the particular case of standard Monte-Carlo integration on a cube the inequalities in [6] imply that there exists a constant  $C$  such that for any  $R > C$

$$\mathbb{P}_N \left( \sup_{\|\nabla f\|_{L^\infty} \leq 1} \left| \int_X f dx - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq R \frac{(\log N)^{1/2}}{N^{1/2}} \right) \geq 1 - \frac{1}{NR^{2/C}} \quad (1.16)$$

(see also [5] for a simplified proof). This inequality thus exhibits the slower rate  $(\log N)^{1/2}/N^{1/2}$ , since the points  $x_i$  are independent random variables. Moreover, the role of the Sobolev norm  $W^{1,2}$  appearing for  $s = 2$  is in the inequality 1.16 played by the  $W^{1,\infty}$ -norm  $\|\nabla f\|_{L^\infty}$ . The proof uses the dual representation of the  $W^{1,\infty}$ -norm between probability measures as the  $L^1$ -Wasserstein metric (aka Monge-Kantorovich distance) which fits into the general setting of optimal transport theory. We also recall that in the particular case of one dimension the sharp form of the Dvoretzky-Kiefer-Wolfowitz inequality for  $N$  independent real random variables

with the same distribution  $\mu$ , assumed to have a continuous density, says that

$$\mathbb{P}_N \left( d_K \left( \mu, \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \geq \lambda \right) \right) \leq 2e^{-2N\lambda^2}, \quad d_K(\mu, \nu) := \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^x (\mu - \nu) \right|,$$

where  $d_K$  denotes the Kolomgorov distance. As a consequence, if  $\mu$  and the points  $x_i$  are supported in an interval  $X$  of unit-length, then it follows from the fact that the  $L^1$ -Wasserstein metric on  $X$  is bounded from above by  $d_K$  that

$$\mathbb{P}_N \left( \sup_{\|\nabla f\|_{L^\infty} \leq 1} \left| \int_X f \mu - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \geq \frac{\lambda}{N^{1/2}} \right) \leq 2e^{-2\lambda^2}$$

(see the discussion in [5, page 2304-2305]).

For generalized Wigner random matrices of rank  $N$ , whose eigenvalues define a random point with  $N$  points on  $\mathbb{R}$  (which for standard Wigner matrices is a determinantal point process, as well as a Coulomb gas) a concentration of measure type inequality was obtained in [15, Thm 2.2], expressed in terms of the Kolmogorov distance  $d_K$  on an interval. Denoting by  $\mu$  the semi-circle law it yields, by bounding the  $W^{1,\infty}$ -norm by  $d_K$ , the following concentration inequality on an interval  $X$  containing a neighborhood of the support of  $\mu$ ,

$$\mathbb{P}_N \left( \sup_{\|\nabla f\|_{L^\infty} \leq 1} \left| \int_X f \mu - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq \frac{(\log N)^L}{N} \right) \leq C e^{-c(\log N)^\epsilon},$$

for some positive constants  $C, c$  and  $\epsilon$  and an appropriate positive number  $L$  (depending on  $N$ ), assuming that  $N$  is taken sufficiently large. Thus, in this case the distance in question is of the order  $\frac{(\log N)^L}{N}$  with overwhelming probability.

Generalizations of the concentration of measure inequality 1.16 to general Coulomb (and Riesz) gas ensembles ( $d\mathbb{P}_{N,\beta}, \mathbb{R}^N$ ) in Euclidean  $\mathbb{R}^N$  have been obtained in [11, 31] and on compact Riemannian manifolds in [17]. In particular, in the case of the spherical ensemble the inequalities in [17] say that

$$\mathbb{P}_N \left( \sup_{\|\nabla f\|_{L^\infty} \leq 1} \left| \int_X f dx - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq \delta \right) \geq 1 - e^{-\frac{N^2}{4\pi} \frac{\delta^2}{2} + \frac{N}{4\pi} \log N + CN} \tag{1.17}$$

To see the relation to the present  $L^2$ -setting note that the Sobolev inequality shows that, in dimension  $d = 2$ ,

$$\|\nabla f\|_{L^\infty}(X) \leq C_\epsilon \|\nabla f\|_{H^{2+\epsilon}(X)}$$

for any  $\epsilon > 0$ . Hence, the inequality 1.17 implies a concentration inequality for Sobolev  $H^{-(2+\epsilon)}(X)$ -norms which is similar to the inequality in Theorem 1.3. However, the main virtue of the Gaussian estimate in Theorem 1.3 is that there are, apart

from the term proportional to  $-N^2\delta^2$ , no  $N$ -dependent additional terms, for a fixed  $\epsilon > 0$ . This allows one to apply Theorem 1.3 to  $\delta$  of the order  $N^{-1}$  (modulo logarithmic factors, by taking  $\epsilon$  to be of the order  $1/\log N$ ), as in the proof of Theorem 1.1. In contrast, in the inequality 1.17 one can at best take  $\delta$  of the order  $N^{-1/2}$  due to the presence of the terms of order  $N$ .

## 2 Spectral preparations

We will denote by  $\mathbb{P}_N$  and  $\mathbb{E}_N$  the probabilities and expectations, respectively, defined wrt the spherical ensemble with  $N$ -particles  $(X^N, d\mathbb{P}_N)$  (whose definition was recalled in Sect. 1.2). We start with some preliminaries.

### 2.1 Sobolev spaces and spectral theory

Let us first consider a general setup of Sobolev spaces on a compact Riemannian manifold  $(X, g)$ . Denote by  $\langle \cdot, \cdot \rangle_{L^2}$  the corresponding scalar product on  $C^\infty(X)$  :

$$\langle u, v \rangle_{L^2} := \int_X uv dV_g,$$

where  $dV_g$  denotes the Riemannian volume form (we will denote by  $d\sigma_g$  the probability measure obtained by normalizing  $dV_g$ ). We will denote by  $\Delta_g$  the Laplace operator on  $C^\infty(X)$ , with the sign convention which makes  $\Delta_g$  a densely defined positive symmetric operator on  $L^2(X, dV_g)$  :

$$\langle \Delta_g u, v \rangle_{L^2} := \int_X g(\nabla_g u, \nabla_g v) dV_g,$$

where  $\nabla_g u$  denotes the gradient of  $u$  wrt  $g$ . By the spectral theorem, for any  $p \in \mathbb{R}$  the  $p$ th power  $\Delta_g^p$  is a densely defined operator on  $L^2(X, dV_g)$ .

Fix a “smoothness parameter”  $s$ , assumed to be strictly positive:

- $H^s(X)/\mathbb{R}$  is defined as the completion of  $C^\infty(X)/\mathbb{R}$  with respect to the scalar product defined by

$$\langle u, u \rangle_s := \int_X \Delta_g^{s/2} u \Delta_g^{s/2} u dV_g = \int_X u \Delta^s u dV_g \tag{2.1}$$

- $H_0^{-s}(X)$  is defined as the sub-space of all distributions  $v$  on  $X$  such that  $\langle v, 1 \rangle = 0$  satisfying

$$\langle v, v \rangle_{-s} := \sup_{u \in C^\infty(X)} \frac{\langle v, u \rangle}{\langle u, u \rangle_s} < \infty$$

Here we view a distribution  $v$  on  $X$  as an element in the linear dual of the vector space  $C^\infty(X)$ . We endow  $H^s(X)/\mathbb{R}$  and  $H_0^{-s}(X)$  with the Hilbert space structures

defined by the scalar products  $\langle \cdot, \cdot \rangle_s$  and  $\langle \cdot, \cdot \rangle_{-s}$ , respectively. Note that the norm on  $H^s(X)/\mathbb{R}$  is increasing wrt  $s$ , while the norm on  $H_0^{-s}(X)$  is decreasing wrt  $s$ . Moreover, by definition, we have that

$$\|\delta_N(\mathbf{x}_N) - d\sigma_g\|_{H_0^{-s}(X)} = \text{wce}(\mathbf{x}_N; s) \tag{2.2}$$

where  $\text{wce}(\mathbf{x}_N; s)$  is the worst-case error for the integration rule on  $X$  with node set  $\mathbf{x}_N$  with respect to the smoothness parameter  $s \in ]1, \infty[$  (defined by formula 1.5). By the Sobolev embedding theorem  $\text{wce}(\mathbf{x}_N; s)$  is finite precisely when  $s > \dim X/2$ ,

By duality the operator  $\Delta_g$  is also defined on the space of all distributions  $v$  :

$$\langle \Delta_g v, u \rangle := \langle v, \Delta_g u \rangle$$

The following lemma follows directly from the definition of the Hilbert spaces in question:

**Lemma 2.1** *The operator  $\Delta_g$  induces an isometry when restricted to  $C^\infty(X)/\mathbb{R}$*

$$H^s(X)/\mathbb{R} \rightarrow H^{s-2}(X)/\mathbb{R}, \quad H_0^{-s}(X) \rightarrow H_0^{-s-2}(X)$$

Next, recall that, by the spectral theorem, the set of eigenfunctions  $f_i$  of  $\Delta_g$  in  $C^\infty(X)$  form an orthonormal bases for  $L^2(X, dV_g)$ . The following lemma then follows directly by duality:

**Lemma 2.2** *There exists an orthonormal basis  $v_1, v_2, \dots$  in the Hilbert space  $\langle H_0^{-s}(X), \langle \cdot, \cdot \rangle_{-s} \rangle$  such that*

$$v_i = f_i dV_g$$

(acting on  $C^\infty(X)$  by integration) where  $f_i$  runs over all eigenfunctions of the Laplacian on  $C^\infty(X)$  with strictly positive eigenvalues  $\lambda_i$ . As a consequence, if  $f_i$  is normalized so that  $\|f_i\|_{L^2} = 1$  and

$$v = \sum_{i=1}^{\infty} c_i f_i$$

in  $H_0^{-s}(X)$  then

$$\langle v, v \rangle_{-s} := \sum_{i=1}^{\infty} \lambda_i^{-s} c_i^2$$

**Remark 2.3** In the literature different Sobolev space norms on  $H^s(X)/\mathbb{R}$  are often used, for example, obtained by replacing  $\Delta^s$  in the last equality in formula 2.1 with  $(I + \Delta)^s$  (as in [9, 10]) or, more generally, any other elliptic pseudodifferential operator  $P_s$  of order  $s$ . [12] Anyway, the norms on  $H^s(X)/\mathbb{R}$  defined by any two such operators

are quasi-isometric, by elliptic regularity theory (see the discussions in [10, 12]). Hence, when the norm is changed Theorem 1.3 still applies if  $\delta$  is replaced by  $C(\epsilon)\delta$  for a positive constant  $C(\epsilon)$  (and similarly for Theorems 1.2, 1.1).

### 2.1.1 Spectral theory

Recall that the *spectral zeta function* of the Laplacian  $\Delta_g$  is defined by

$$\text{Tr}(\Delta_g^{-p}) := \sum_{i=1}^{\infty} \lambda_i^{-p}$$

which is convergent for  $p > \dim X/2$ . More precisely,

$$\text{Tr}(\Delta_g^{-(d/2+\epsilon)}) = \frac{1}{\Gamma(d/2)} \frac{\text{Vol}(X, g)}{(4\pi)^{d/2}} \frac{1}{\epsilon} + O(1), \quad \epsilon \rightarrow 0^+$$

as follows, for example, from the expansion of the heat kernel, i.e. the Schwartz kernel of  $\text{Tr}(e^{-t\Delta_g})$ . We will prove explicit estimates in the case of the two-sphere below.

The *Fredholm (spectral) determinant* of  $\Delta_g^{-p}$  is the function

$$D(\lambda, p) := \det(I - \lambda\Delta_g^{-p}) := \prod_{i=1}^{\infty} (1 - \lambda\lambda_i^{-p})$$

which is convergent for  $p > \dim X/2$  and  $\lambda \in ]0, \lambda_1^p[$ . Indeed, since the Taylor expansion of  $-\log(1 - \lambda t)$  equals  $\sum_{m=1}^{\infty} \frac{\lambda^m t^m}{m}$ ,

$$-\log \det(I - \lambda\Delta_g^{-p}) := \sum_{m=1}^{\infty} \frac{\text{Tr}(\Delta_g^{-mp})}{m} \lambda^m \tag{2.3}$$

### 2.1.2 The case of the two-sphere

Consider now the case when  $(X, g)$  is the two-sphere, i.e. the unit-sphere in  $\mathbb{R}^3$  endowed with the metric  $g$  induced by the Euclidean metric on  $\mathbb{R}^3$ . Note that under stereographic projection, whereby  $X$  minus the “north pole” is identified with  $\mathbb{R}^2$ , we have

$$\Delta_g u dV_g = - \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) dx \wedge dy, \quad dV_g = \frac{4dx \wedge dy}{(1 + x^2 + y^2)^2}$$

Moreover, the set of non-zero eigenvalues of  $\Delta_g$  are given by all numbers of the form  $l(l + 1)$ , where  $l$  ranges over the positive integers. The eigenvalue corresponding to a given  $l$  has multiplicity  $2l + 1$ .

**Remark 2.4** Another convenient norm on  $H^s(X)/\mathbb{R}$  may be obtained by replacing  $\Delta_g$  with  $\Delta_g + 1/4$  in formula 2.1 (compare the discussion in Remark 2.3). The point is that the eigenvalues of  $\Delta_g + 1/4$  are given by  $(l + 1/2)^2$ . This implies that the corresponding spectral function may be explicitly expressed as  $\text{Tr}((\Delta_g + 4^{-1})^{-p}) := 2^{2p-2}\zeta(2p - 1)$ , where  $\zeta$  is Riemann’s zeta function (see [35, page 453]).

We will use the following refinement of [10, Lemma 26].

**Lemma 2.5** *Let  $(X, g)$  be a  $d$ -dimensional Riemannian manifold of non-negative Ricci curvature. Assume that  $d/2 < s' < s$ . Then*

$$\text{wce}(\mathbf{x}_N; s') \leq c(s', s)\text{wce}(\mathbf{x}_N; s)^{s'/s},$$

where the constant  $c(s', s)$  is given by

$$c(s', s) = \sqrt{\frac{\Gamma(s)}{\Gamma(s')} + \frac{2s^s e^{-s}}{\Gamma(s')(s' - d/2)}}.$$

In particular, when  $d = 2$ ,  $c(2, 2 + \epsilon) < 1.51 < e^{1/2}$  for  $\epsilon \leq 0.15$ .

**Proof** The main difference compared to the case of the  $d$ -dimensional unit-sphere in [10, Lemma 26] is that the constant  $c(s, s')$  in [10, Lemma 26] depends on a non-explicit constant  $c$  such that for  $t \in ]0, \epsilon/2[$  the heat kernel  $K_t$  satisfies  $tK_t \leq c\epsilon K_\epsilon$  on  $X \times X$ . Here we observe, in particular, that one can take  $c = 1$  and allow  $t \in ]0, \epsilon[$ , i.e.

$$t \in ]0, \epsilon] \implies tK_t \leq \epsilon K_\epsilon \tag{2.4}$$

as follows from the Li-Yau parabolic Harnack-inequality on any Riemannian manifold with non-negative Ricci curvature (apply [25, Thm 2.3] to  $u(x) := K_t(x, y)$  for  $y$  fixed and  $\alpha = 1$ ). Setting

$$\mu := N^{-1} \sum_{i=1}^N \delta_{x_i} - dV_g$$

formula 2.2 may be expressed as

$$W(s)^2 := \text{wce}(\mathbf{x}_N; s)^2 = \|\mu\|_{(H^s(X)/\mathbb{R})^*}^2 = \sum_{\lambda_i > 0} \lambda_i^{-s} |\langle \mu, f_i \rangle|^2, \tag{2.5}$$

where, as before,  $\lambda_i$  and  $f_i$  denote the eigenvalues and eigenfunctions of the Laplacian for  $(X, g)$ . Using the identity  $\lambda^{-s} = \int_0^\infty t^{s-1} e^{-t\lambda} dt / \Gamma(s)$  and that  $\int \mu = 0$  it follows that

$$\text{wce}(\mathbf{x}_N; s)^2 = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} g(t) dt, \tag{2.6}$$

where

$$\begin{aligned}
 g(t) &:= \int \int \mathcal{K}_t(x, y) \mu \otimes \mu = \int \int K_t(x, y) \mu \otimes \mu, \quad \mathcal{K}_t(\cdot, y) \\
 &= \sum_{\lambda_i > 0} e^{-t\lambda_i} f_i(x) f_i(y)
 \end{aligned}$$

with  $\mathcal{K}_t$  denoting the heat-kernel  $K_t$  for  $(X, g)$  with the constant term removed (note that the formula in [10] corresponding to 2.6 contains a factor  $e^{-t}$  due to the different definition of the Sobolev norms in [10]). Now, setting

$$\epsilon := \text{wce}(\mathbf{x}_N; s)^{2/s}, \tag{2.7}$$

we split the integral over  $t$  in formula 2.6, with  $s$  replaced by  $s'$ , over the two disjoint regions  $]\epsilon, \infty[$  and  $]0, \epsilon[$  (in [10, Lemma 26] there are three regions and the role of  $\epsilon$  is played by  $\epsilon/2$ , but here we can take  $\epsilon$  since we will use the sharper estimate 2.4). First note that, since  $s' \leq s$ ,

$$\begin{aligned}
 \int_{\epsilon}^{\infty} t^{s'-1} g(t) dt &= \epsilon^{s'} \int_{\epsilon}^{\infty} (t/\epsilon)^{s'} t^{-1} g(t) dt \leq \epsilon^{s'} \int_{\epsilon}^{\infty} (t/\epsilon)^s t^{-1} g(t) dt \\
 &= \epsilon^{s'-s} \int_{\epsilon}^{\infty} t^{s-1} g(t) dt.
 \end{aligned}$$

Hence, bounding the latter integral with the integral over all of  $[0, \infty[$  yields

$$\begin{aligned}
 \frac{1}{\Gamma(s')} \int_{\epsilon}^{\infty} t^{s'-1} g(t) dt &\leq \frac{\Gamma(s)}{\Gamma(s')} \left( \epsilon^{s'-s} \frac{1}{\Gamma(s)} \int_0^{\infty} t^{s-1} g(t) dt \right) =: \frac{\Gamma(s)}{\Gamma(s')} \epsilon^{s'-s} \epsilon^s \\
 &= \frac{\Gamma(s)}{\Gamma(s')} \epsilon^{s'}
 \end{aligned}$$

Turning to the region where  $t \in [0, \epsilon]$ , the estimate 2.4 for the heat-kernel  $K_t$  yields

$$g(t) \leq t^{-d/2} \epsilon^{d/2} g(\epsilon).$$

Hence,

$$\int_0^{\epsilon} t^{s'-1} g(t) dt \leq \left( \int_0^{\epsilon} t^{s'-1-d/2} dt \right) \epsilon^{d/2} g(\epsilon) = \frac{1}{s' - d/2} \epsilon^{s'} g(\epsilon).$$

Finally, we will show that  $g(\epsilon)$  is uniformly bounded (there is a typo in the proof of [10, Lemma 26] saying that  $g(t)$  is uniformly bounded for  $t \in [0, 1[$ , which contradicts the blow-up of the heat-kernel as  $t \rightarrow 0$ ). First, by the very definition of the worst case error  $W(s)$ ,

$$\int \mathcal{H}_t(\cdot, y) \mu(y) \leq W(s) \left\| \int \mathcal{H}_t(\cdot, y) \mu(y) \right\|_{H^s(X)}.$$

The square of the latter Sobolev norm may be expressed as

$$\left\| \sum_{\lambda_i > 0} e^{-t\lambda_i} \langle \mu, f_i \rangle f_i(x) \right\|_{H^s(X)}^2 = \sum_{\lambda_i > 0} \lambda_i^s e^{-2t\lambda_i} |\langle \mu, f_i \rangle|^2.$$

Hence, rewriting  $\lambda^s = \lambda^{-s} \lambda^{2s}$ ,

$$\left\| \int \mathcal{H}_t(\cdot, y) \mu(y) \right\|_s \leq \left( \sum_{\lambda_i > 0} \lambda_i^{-s} |\langle \mu, f_i \rangle|^2 \right)^{1/2} \sup_{\lambda > 0} \lambda^s e^{-t\lambda}$$

By formula 2.5 the first factor equals  $W(s)$  and the second factor is given by

$$\sup_{\lambda > 0} \lambda^s e^{-t\lambda} = t^{-s} C(s), \quad C(s) := s^s e^{-s}$$

Since, by the definition 2.7,  $W(s)^2 := \epsilon^s$ , setting  $t = \epsilon$  thus yields the desired uniform bound:

$$g(\epsilon) \leq 2 \int \mathcal{H}_t(\cdot, y) \mu(y) \leq 2\epsilon^s \epsilon^{-s} C(s) = 2C(s).$$

Finally, adding up the two contributions to the ingral in formula 2.6 concludes the proof of the inequality in the lemma. Setting  $s' = 2$  gives  $c(2, s)^2 = \Gamma(s) + 2s^s e^{-s}$  and hence, if  $\epsilon \leq 0.15$ , then a numerical calculation yields  $c(2, 2 + \epsilon) \leq \sqrt{1.073 + 2(2.15/e)^{2.15}} < 1.52 < e^{1/2}$ .  $\square$

**Lemma 2.6** *On the two-sphere  $(X, g)$  the following inequality holds for any  $\epsilon > 0$*

$$\text{Tr}(\Delta_g^{-(1+\epsilon)}) \leq \frac{1}{\epsilon} + 2, \quad \text{Tr}(\Delta_g^{-2}) = 1$$

and hence

$$-\log \det(I - \lambda \Delta_g^{-(1+\epsilon)}) \leq \lambda \frac{1}{\epsilon} - 4 \log \left( 1 - \frac{\lambda}{2} \right).$$

**Proof** Setting  $\mathcal{Z}(p) := \text{Tr}(\Delta_g^{-p})$ , for  $p \geq 1$ , we have

$$\begin{aligned} \mathcal{Z}(p) &= \sum_{l=1}^{\infty} \frac{2l+1}{l^p(l+1)^p} = 2 \sum_{l=1}^{\infty} \frac{1}{l^{p-1}(l+1)^p} + \sum_{l=1}^{\infty} \frac{1}{l^p(l+1)^p} \\ &\leq 2 \sum_{l=1}^{\infty} \frac{1}{l^{p-1}(l+1)^p} + 1, \end{aligned}$$

where the second sum was estimated by replacing  $p$  with 1 to get a telescoping sum. Next,

$$\sum_{l=1}^{\infty} \frac{1}{l^{p-1}(l+1)^p} \leq \frac{1}{2} + \sum_{l=2}^{\infty} \frac{1}{l^{p-1}(l+1)^p} \leq \frac{1}{2} + \sum_{l=2}^{\infty} \frac{1}{l^{2p-1}} = \frac{1}{2} + \zeta(2p-1) - 1$$

(using the trivial bound  $l+1 \geq l$ ), where  $\zeta(s)$  is the Riemann zeta function. Set  $s = 1 + \delta$ . As is well-known, when  $s > 1$ , a resummation argument gives (cf. [18, formula 3])

$$\zeta(s) \leq \frac{s}{s-1} = 1 + \frac{1}{\delta},$$

Hence, setting  $p = 1 + \epsilon$ , gives  $\zeta(2p-1) = \zeta(1+2\epsilon) \leq 1 + 1/(2\epsilon)$ . All in all, this means that

$$\mathcal{Z}(p) \leq 2 \left( \frac{1}{2} + \left( 1 + \frac{1}{2\epsilon} \right) - 1 \right) + 1 = 2 + \frac{1}{\epsilon},$$

proving the first inequality in the lemma. Next, note that  $\mathcal{Z}(2)$  can be computed as a telescoping sum:

$$\mathcal{Z}(2) = \sum_{l=1}^{\infty} \frac{2l+1}{l^2(l+1)^2} = \sum_{l=1}^{\infty} \frac{(l+1)^2 - l^2}{l^2(l+1)^2} = \sum_{l=1}^{\infty} \frac{1}{l^2} - \frac{1}{(l+1)^2} = 1 + 0.$$

Now, Taylor expanding  $-\log \det(I - \lambda \Delta_g^{-(1+\epsilon)})$  (as in formula 2.3) gives

$$\begin{aligned} -\log \det(I - \lambda \Delta_g^{-(1+\epsilon)}) &= \sum_{m=1}^{\infty} \frac{\lambda^m}{m} \text{Tr} \left( \Delta^{-m(1+\epsilon)} \right) \\ &\leq \lambda \left( \frac{1}{\epsilon} + 2 \right) + \sum_{m=2}^{\infty} \frac{\lambda^m}{m} \text{Tr} \left( \Delta^{-m(1+\epsilon)} \right). \end{aligned}$$

Since the smallest strictly positive eigenvalue of  $\Delta_g$  is equal to 2 rewriting  $\lambda_i = 2(\lambda_i/2)$  and using that that  $\lambda_i/2 \geq 1$  now yields, for  $m \geq 2$ ,

$$\text{Tr} \left( \Delta_g^{-m(1+\epsilon)} \right) \leq 2^2 2^{-m} \text{Tr} \left( \Delta_g^{-2} \right) = 2^2 2^{-m} \cdot 1.$$

As a consequence,

$$\sum_{m=2}^{\infty} \frac{\lambda^m}{m} \text{Tr} \left( \Delta_g^{-m(1+\epsilon)} \right) \leq 2^2 \sum_{m=2}^{\infty} \left( \frac{\lambda}{2} \right)^m \frac{1}{m} = 2^2 \left( -\log \left( 1 - \frac{\lambda}{2} \right) - \frac{\lambda}{2} \right),$$

using again that the Taylor expansion of  $-\log(1-t)$  equals  $\sum_{m=1}^{\infty} \frac{t^m}{m}$ . All in all, this means that  $-\log \det(I - \lambda \Delta_g^{-(1+\epsilon)})$  is bounded from above by

$$\lambda \left( \frac{1}{\epsilon} + 2 \right) - 2\lambda - 2^2 \log \left( 1 - \frac{\lambda}{2} \right) = \lambda \frac{1}{\epsilon} - 4 \log \left( 1 - \frac{\lambda}{2} \right),$$

□

**Remark 2.7** By [28, Prop 5]  $\text{Tr}(\Delta_g^{-(1+\epsilon)}) = 1/\epsilon + 2\gamma - 1 + o(1)$  as  $\epsilon \rightarrow 0^+$ , where  $\gamma = 0.577\dots$  is Euler’s constant. But the point of the previous lemma is to bound the error term explicitly for  $\epsilon$  fixed.

### 3 Proofs of the main results

Fix  $s > 2$  and a positive integer  $N$ . Consider the following  $H_0^{-s}(X)$ -valued random variable on  $(X^N, d\mathbb{P}_N)$  :

$$Y_N := N(\delta_N - d\sigma) : (X^N, d\mathbb{P}_N) \rightarrow H_0^{-s}(X),$$

where  $\delta_N$  denotes the empirical measure 1.1 (the space  $H_0^{-s}(X)$  contains the image of  $Y_N$  for any  $s > 1$ , but the restriction to  $s > 2$  will turn out to be important in the following). To keep things as elementary as possible it will be convenient to consider truncated random variable taking values in finite dimensional approximations of  $H_0^{-s}(X)$  (but a more direct approach could also be used; see Remark 3.2). To this end fix an orthonormal basis  $v_i$  in the Hilbert space  $\langle H_0^{-s}(X), \langle \cdot, \cdot \rangle_{-s} \rangle$ . It will be convenient to take  $v_i$  as in Lemma 2.2 ordered so that  $0 < \lambda_1 \leq \dots \leq \lambda_M$ . Let  $H_{\leq M}^{-s}(X)$  be the  $M$ -dimensional subspace of  $H_0^{-s}(X)$  defined by

$$H_{\leq M}^{-s}(X) := \mathbb{R}v_1 \oplus \dots \oplus \mathbb{R}v_M \subseteq H^{-s}(X)$$

Denote by  $\pi_M$  the orthogonal projection from the Hilbert space  $H_0^{-s}(X)$  onto the  $M$ -dimensional subspace  $H_{\leq M}^{-s}(X)$ .

#### Step 1: $\pi_M(Y_N)$ is a sub-Gaussian random variable

We can view  $\pi_M(Y_N)$  as a random variable on  $(X^N, d\mathbb{P}_N)$  taking values in  $H_{\leq M}^{-s}(X)$  :

$$\pi_M(Y_N) : (X^N, d\mathbb{P}_N) \rightarrow H_{\leq M}^{-s}(X).$$

The first step of the proof is to compare  $\pi_M(Y_N)$  with a *Gaussian* random variable  $G_M$  taking values in  $H_{\leq M}^{-s}(X)$ . To this end we first endow  $H_{\leq M}^{-s}(X)$  with the Hilbert structure define by the scalar product  $\langle \cdot, \cdot \rangle_{-1}$  and denote by  $\gamma_M$  the Gaussian measure on  $\langle H_{\leq M}^{-s}(X), \langle \cdot, \cdot \rangle_{-1} \rangle$ . Concretely, this means that under any linear isometry

of  $\left\langle H_{\leq M}^{-s}(X), \langle \cdot, \cdot \rangle_{-1} \right\rangle$  with  $\mathbb{R}^N$  the measure  $\gamma_M$  corresponds to the standard centered Gaussian measure on  $\mathbb{R}^N$ . Now define  $G_M$  as a random element in  $H_{\leq M}^{-s}(X)$ . In other words,  $G_M$  is the random variable defined by the identity map

$$G_M := I \left( H_{\leq M}^{-s}(X), \gamma_M \right) \rightarrow H_{\leq M}^{-s}(X)$$

The following proposition says that the moment generating function of the random variable  $\pi_M(Y_N)$ , viewed as a function on the linear dual  $(H_{\leq M}^{-s}(X))^*$  of  $H_{\leq M}^{-s}(X)$  is bounded from above by the moment generating function of the scaled Gaussian random variable  $G_M/(4\pi)^{1/2}$ .

**Proposition 3.1** *The following inequality holds:*

$$\mathbb{E}(e^{\langle \pi_M(Y_N), \cdot \rangle}) \leq \mathbb{E} \left( e^{\langle \frac{1}{\sqrt{4\pi}} G_M, \cdot \rangle} \right) \tag{3.1}$$

Equivalently, denoting by  $p_N^{(M)}$  the law of  $\pi_M(Y_N)$ , i.e. the probability measure on  $H_{\leq M}^{-s}(X)$  defined by the push-forward of  $d\mathbb{P}_N$  under the map  $\pi_M(Y_N)$ , we have

$$L[p_N^{(M)}] \leq L[F_*\gamma_M], \quad F(v) := \frac{v}{\sqrt{4\pi}}$$

where  $L[\Gamma]$  denote the Laplace transform of a measure  $\Gamma$  on the finite dimensional vector space  $V := H_{\leq M}^{-s}(X)$ , i.e.  $L[\Gamma]$  is the function on  $V^*$  defined by

$$L[\Gamma](w) = \int_V e^{-w} \Gamma$$

**Proof** Applying the Moser-Trudinger type inequality 1.11 for the spherical ensemble proved in [4] to  $f = w/(N + 1)$  for  $w \in H^1(X)$  gives

$$\mathbb{E}(e^{\langle Y_N, w \rangle}) \leq e^{\frac{N(N+1)}{(N+1)(N+1)} \frac{1}{2} \frac{1}{4\pi} \|w\|_{H^1(X)}^2} \leq e^{\frac{1}{2} \frac{1}{4\pi} \|w\|_{H^1(X)}^2}$$

In particular, taking  $w \in (H_{\leq M}^{-s}(X))^*$ , identified with a subspace of  $H^1(X)/\mathbb{R}$ , gives  $\langle Y_N, w \rangle = \langle \pi_M(Y_N), w \rangle$  and hence it will be enough to verify that

$$e^{\frac{1}{2} \|w\|_{H^1(X)}^2} = \mathbb{E} \left( e^{\langle G_M, w \rangle} \right) (= L[\gamma_M](-w)) \tag{3.2}$$

To this end first note that under the identifications above  $\|w\|_{H^1(X)}^2$  coincides with the dual norm on the Hilbert space dual of  $\left\langle H_{\leq M}^{-s}(X), \langle \cdot, \cdot \rangle_{-1} \right\rangle$ . But then formula 3.2 follows from the well-known fact that if  $\gamma$  is the Gaussian measure on a finite dimensional Hilbert space  $H$ , then

$$L[\gamma](w) = \exp \left( \frac{1}{2} \|w\|_{H^*}^2 \right). \tag{3.3}$$

Indeed, fixing an orthonormal basis in  $H$  this reduces to the basic fact that the Laplace transform of the measure  $e^{-|x|^2/2}dx$  on Euclidean  $\mathbb{R}^N$  is equal to  $e^{-|y|^2/2}$ , which, in turn, follows from “completing the square”.  $\square$

**Remark 3.2** In the terminology introduced by Kahane, the previous inequality says that the random variable  $\pi_M(Y_N)$  is *sub-Gaussian* with respect to the Gaussian random variable  $\frac{1}{\sqrt{4\pi}}G_M$ . In fact, by letting  $M \rightarrow \infty$  this implies that  $Y_N$  is sub-Gaussian with respect to the Laplacian of the Gaussian free field [33], viewed as random variables taking values in  $H_0^{2+\epsilon}(X)$ . This point of view will be elaborated on in [3].

**Step 2: Bounding  $\mathbb{E}(e^{\|\pi_M(Y_N)\|_{-s}^2})$**

We start with the following general

**Lemma 3.3** *Let  $H$  be a finite dimensional Hilbert space and denote by  $\gamma$  the corresponding Gaussian probability measure. If  $\Gamma$  is a measure on  $H$  such that the following inequality of Laplace transforms hold*

$$L[\Gamma] \leq L[\gamma]$$

as functions on the dual vector space  $H^*$ , then

$$\int e^{q\Gamma} \leq \int e^{q\gamma} \tag{3.4}$$

for the squared semi-norm  $q(v)$  defined by any given semi-positive symmetric bilinear form on  $H$ .

**Proof** First observe that the inequality 3.4 holds for any function  $q$  on  $H$  which has the following positivity property:  $e^q$  is the Laplace transform of a positive measure  $\mu_q$  on  $H^*$ , i.e.

$$e^{q(v)} = \int_{w \in H^*} e^{-\langle v, w \rangle} d\mu_q(w)$$

Indeed, changing the order of integration (using Fubini) the integral of  $e^q$  against  $\Gamma$  may be expressed as

$$\int_{v \in H} \left( \int_{w \in H^*} e^{-\langle v, w \rangle} d\mu_q(w) \right) d\Gamma(v) = \int_{w \in H^*} \left( \int_{v \in H} e^{-\langle v, w \rangle} d\Gamma(v) \right) d\mu_q(w).$$

Hence, by assumption,

$$\int_{v \in H} e^{q(v)} d\Gamma(v) \leq \int_{w \in H^*} \left( \int_{v \in H} e^{-\langle v, w \rangle} d\gamma(v) \right) d\mu_q(w),$$

which is equal to  $\int e^q \gamma$  (as seen by changing the order of integration again). All that remains is thus to verify the positivity property in question when  $q$  is a squared seminorm. Identifying  $H$  with Euclidean  $\mathbb{R}^M$  and diagonalizing  $q$  we may as well assume that  $q = \sum |a_i x_i|^2/2$  for  $a_i \geq 0$ . Moreover, by approximation we may as well assume that  $a_i > 0$ . But then it follows from formula 3.3 and scaling the variables that the measure  $d\mu_q = C \exp\left(-\sum |a_i^{-1} y_i|^2/2\right) dy_1 \dots dy_M$  has the required property for a appropriate positive constant  $C$ .  $\square$

In the present situation we get the following

**Proposition 3.4** *For any  $\alpha > 0$  the following inequality holds:*

$$\mathbb{E}\left(e^{\alpha \|\pi_M(Y_N)\|_{-s}^2}\right) \leq \mathbb{E}\left(e^{\frac{\alpha}{4\pi} \|G_M\|_{-s}^2}\right) = \prod_{i \leq M} \left(1 - \frac{\alpha}{2\pi} \lambda_i^{-(s-1)}\right)^{-1/2} \tag{3.5}$$

**Proof** The first inequality follows directly from combining Prop 3.1 and Lemma 3.3 with  $\gamma = F_* \gamma_M q(v) := \alpha \|v\|_{-s}^2$ . To prove the last equality denote by  $v_i$  an orthonormal base in the Hilbert space  $(H_{\leq M}^{-s}(X), \langle \cdot, \cdot \rangle_{-1})$ . Given  $v \in H_{\leq M}^{-s}(X)$  we decompose  $v = \sum_{i=1}^M v_i x_i$  and note that

$$\|v\|_{-s}^2 = \sum_{i=1}^M x_i^2 \lambda_i^{(1-s)}$$

as follows from writing

$$\|v_i\|_{-s}^2 := \langle \Delta^{-s} v_i, v_i \rangle_{L^2} = \langle \Delta^{-1} (\Delta^{1-s} v_i), v_i \rangle_{L^2} = \lambda_i^{(1-s)} \langle \Delta^{-1} v_i, v_i \rangle_{L^2} = \lambda_i^{(1-s)}.$$

Hence,

$$\mathbb{E}\left(e^{\frac{\alpha}{4\pi} \|G_M\|_{-s}^2}\right) = \prod_{i \leq M} \int e^{\frac{\alpha}{4\pi} x_i^2 \lambda_i^{(1-s)}} e^{-x_i^2/2} dx_i / Z_M,$$

where  $Z_M$  is the total integral of  $\prod_{i \leq M} e^{-x_i^2/2} dx_i$ . Finally, changing variables  $x_i \rightarrow \left(1 - \frac{\alpha}{2\pi} \lambda_i^{(1-s)}\right)^{1/2}$  in the corresponding Gaussian integrals then concludes the proof of the proposition.  $\square$

Letting  $M \rightarrow \infty$  and using the monotone convergence theorem now concludes the proof of Theorem 1.2.

**3.1 Proof of Theorem 1.2 and Theorem 1.3**

Set  $s = 2 + \epsilon$  and  $\lambda = \alpha/2\pi$ . By Lemma 2.6

$$\prod_{i \leq M} \left(1 - \lambda \lambda_i^{-(s-1)}\right)^{-1} \leq \det(I - \lambda \Delta_g^{-(1+\epsilon)})^{-1} \leq \exp(f(\lambda)), \tag{3.6}$$

where

$$f(\lambda) := \lambda \frac{1}{\epsilon} - 4 \log \left(1 - \frac{\lambda}{2}\right).$$

Hence, for any fixed positive integer  $M$  Prop 3.4 gives

$$\mathbb{E}_N \left( e^{\alpha \| \pi_M(Y_N) \|_{L^2(2+\epsilon)}^2} \right) \leq \det \left( I - \lambda \Delta_g^{-(1+\epsilon)} \right)^{-1/2} \leq \exp \left( \frac{1}{2} f \left( \frac{\alpha}{2\pi} \right) \right).$$

Letting  $M \rightarrow \infty$  and using the monotone convergence theorem we deduce that

$$\mathbb{E}_N \left( e^{\alpha \| Y_N \|_{L^2(2+\epsilon)}^2} \right) \leq \det \left( I - \lambda \Delta_g^{-(1+\epsilon)} \right)^{-1/2} \leq \exp \left( \frac{1}{2} f \left( \frac{\alpha}{2\pi} \right) \right),$$

proving Theorem 1.2. Finally, by Chebyshev’s inequality, we can write  $\mathbb{P}_N (\| \delta_N - \sigma \|_{H^{-s}} > \delta)$  as

$$\begin{aligned} \mathbb{P}_N \left( \alpha \| Y_N \|_{H^{-(2+\epsilon)}(X)}^2 > \alpha \delta^2 N^2 \right) &\leq e^{-\alpha \delta^2 N^2} \mathbb{E}_N \left( e^{\alpha \| Y_N \|_{L^2(2+\epsilon)}^2} \right) \\ &\leq \exp \left( -\alpha \delta^2 N^2 + \frac{1}{2} f \left( \frac{\alpha}{2\pi} \right) \right) \end{aligned}$$

Finally, since  $\lambda_1 = 2$  on the two-sphere the determinant in the theorem is finite for  $\lambda < 2^{1+\epsilon}$ , i.e. when  $\alpha < 4\pi 2^\epsilon$ .

**3.2 The optimal choice of  $\alpha$**

Setting  $\lambda := \alpha/2\pi$  we have

$$\begin{aligned} \mathbb{P}_N (\| \delta_N - \sigma \|_{H(X)^{-(2+\epsilon)}} > \delta) &\leq \exp \left( \frac{1}{2} \left( -4\pi \delta^2 N^2 \lambda + f(\lambda) \right) \right), \\ f(\lambda) &:= \lambda \frac{1}{\epsilon} - 4 \log \left( 1 - \frac{\lambda}{2} \right). \end{aligned} \tag{3.7}$$

First observe that  $f(\lambda)$  is strictly convex on  $[0, 2[$ ,  $f(0) = 0$  and  $f(2^-) = \infty$ . Hence, the optimal choice of  $\lambda$  satisfies

$$4\pi \delta^2 N^2 = f'(\lambda) = \epsilon^{-1} + \frac{2}{1 - \lambda/2}, \quad \lambda \in ]0, 2[ \tag{3.8}$$

if such a  $\lambda$  exists. Introducing the parameters  $R \in ]0, \infty[$  and  $\eta \in ]-1, \infty[$  determined by

$$R^2 := \delta^2 N^2 \epsilon, \quad \eta := 4\pi R^2 - 1$$

the equation 3.8 for  $\lambda$  becomes

$$\eta = \frac{2\epsilon}{1 - \lambda/2} \iff \lambda = 2 \left( 1 - \frac{2\epsilon}{\eta} \right)$$

assuming that  $R$  is sufficiently large to ensure that  $\lambda \in ]0, 2[$ , i.e. that

$$\eta > 2\epsilon.$$

As a consequence, for this optimal  $\lambda$ , the bracket appearing in the exponent in the estimate 3.7 becomes

$$\begin{aligned} -(4\pi\delta^2 N^2 \epsilon)\epsilon^{-1}\lambda + f(\lambda) &= -(\eta + 1)\epsilon^{-1}\lambda + f(\lambda) \\ &= (-(\eta + 1) + 1)\epsilon^{-1}\lambda + 4 \log \left( \left( 1 - \frac{\lambda}{2} \right)^{-1} \right) \\ &= 2(-\eta) \left( \epsilon^{-1} - \frac{2}{\eta} \right) + 4 \log \frac{\eta}{2\epsilon} \\ &= 2(-\eta)\epsilon^{-1} + 4 \left( 1 + \log \frac{\eta}{2\epsilon} \right) \end{aligned}$$

In particular, taking  $\epsilon := 1/\log N$  gives, for any  $\eta > 2/\log N$ , gives

$$\begin{aligned} \mathbb{P}_N \left( \|\delta_N - \sigma\|_{H(X)^{-(2+1/\log N)}} > R \frac{(\log N)^{1/2}}{N} \right) \\ \leq \exp \left( \frac{1}{2} \left( -4\pi\delta^2 N^2 \lambda + f(\lambda) \right) \right) = \frac{1}{N^\eta} \left( \frac{1}{2} \log N \right)^2 \exp(2(1 + \log \eta)) \end{aligned} \tag{3.9}$$

### 3.3 Proof of Theorem 1.1

By Lemma 2.5 it will be enough to prove the inequality for  $s = 2$ . Consider the real-valued random variable  $W(s) := \text{wce}(\mathbf{x}_N; s)$  on  $(X^N, d\mathbb{P}_N)$ . Applying Theorem 1.3 to  $\delta = R\epsilon^{-1/2}/N$  it will be enough to show the following claim when  $\epsilon := 1/\log N$  under the assumption that  $\epsilon^{1/2} \leq R \leq N\epsilon^{1/2}$ :

$$\text{claim: } W(2 + \epsilon) \leq R \frac{\epsilon^{-1/2}}{N} \implies W(2) \leq CR \frac{\epsilon^{-1/2}}{N}, \quad C := e^{1/2} c(2, 2 + \epsilon),$$

where  $c(2, 2 + \epsilon)$  is defined in Lemma 2.5. To this end recall that, by assumption,  $W(2 + \epsilon) \leq 1$  and hence by Lemma 2.5 we have, since  $\epsilon \leq 1$

$$W(s) \leq cW(2 + \epsilon)^{\frac{s}{2+\epsilon}}, \quad c = c(2, 2 + \epsilon)$$

In particular,

$$W(2) \leq cW(2 + \epsilon)^{\frac{2}{2+\epsilon}} \leq cW(2 + \epsilon)^{(1-\epsilon/2)}$$

using that  $1/(1 + t) \geq 1 - t$  if  $t \geq 0$  and that  $W_N(2 + \epsilon) \leq 1$ . Hence,  $W_N(2 + \epsilon) \leq R^{\frac{\epsilon^{-1/2}}{N}}$  implies that

$$W(2) \leq c \left( R \frac{\epsilon^{-1/2}}{N} \right)^{(1-\epsilon/2)} = cR^{\frac{\epsilon^{-1/2}}{N}} \left( R \frac{\epsilon^{-1/2}}{N} \right)^{-\epsilon/2}.$$

But, by assumption,  $R\epsilon^{-1/2} \geq 1$  and hence

$$\left( R \frac{\epsilon^{-1/2}}{N} \right)^{-\epsilon/2} = \left( R\epsilon^{-1/2} \right)^{-\epsilon/2} N^{\epsilon/2} \leq N^{\epsilon/2} := \left( N^{1/(\log N)} \right)^{1/2} = e^{1/2},$$

since  $\epsilon := 1/\log N$ .

**Remark 3.5** In particular, if  $1/\log N \leq 0.15$  (for, example,  $N = 1000$ ) then Lemma 2.5 gives  $C \leq e$  and hence  $W(2) \leq eR^{\frac{(\log N)^{1/2}}{N}}$ .

### 3.4 Explicit formulation of Theorem 1.1

Combining the previous remark with the explicit bound in formula 3.9 gives the following explicit formulation of the first inequality in Theorem 1.1

$$\mathbb{P}_N \left( \text{wce}(\mathbf{x}_N; 2) \leq e\sqrt{\frac{1 + \eta}{4\pi}} \frac{(\log N)^{1/2}}{N} \right) \geq 1 - \frac{(\frac{1}{2}e\eta \log N)^2}{N^\eta},$$

under the assumption that  $\eta > 2/\log N$  and  $N \geq 1000$ .

**Acknowledgements** Thanks to Klas Modin for comments on a draft of the paper. Also thanks to the referees whose comments helped to improve the exposition. This work was supported by grants from the KAW foundation, the Göran Gustafsson foundation and the Swedish Research Council.

**Funding** Open access funding provided by Chalmers University of Technology.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Alishahi, K., Zamani, M.: The spherical ensemble and uniform distribution of points on the sphere. *Electron. J. Probab.* **20**(23), 27 (2015)
2. Bardenet, R., Hardy, A.: Monte Carlo with determinantal point processes. *Ann. Appl. Probab.* **30**(1), 368–417 (2020) ([arXiv:1605.00361](https://arxiv.org/abs/1605.00361))
3. Berman, R.J.: On convergences rates in Kähler geometry and Coulomb gases. In preparation (2021)
4. Berman, R.J.: Sharp asymptotics for Toeplitz determinants, fluctuations and the Gaussian free field on a Riemann surface. *Int. Math. Res. Not.* **22**, 5031–5062 (2012)
5. Boissard, E.: Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance. *Electron. J. Probab.* **16**(83), 2296–2333 (2011)
6. Bolley, F., Guillin, A., Villani, C.: Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Relat. Fields* **137**(3–4), 541–593 (2007)
7. Bondarenko, Radchenko, D., Viazovsk, M.: Optimal asymptotic bounds for spherical designs. *Annals Math.* **178**(2), 443–452 (2013)
8. Borodachov, S.V., Hardin, D.P., Saff, E.: *B: discrete energy on rectifiable sets*. Springer, Berlin (2019)
9. Brandolini, L., Choirat, C., Colzani, L., Gigante, G., Seri, R., Travaglini, G.: Quadrature rules and distribution of points on manifolds. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **4**(5), 13 (2014)
10. Brauchart, J.S., Saff, E.B., Sloan, I.H., Womersley, R.S.: QMC designs: optimal order quasi Monte Carlo integration schemes on the sphere. *Math. Comp.* **83**(290), 2821–2851 (2014)
11. Chafaï, D., Hardy, A., Maïd, M.: Concentration for Coulomb gases and Coulomb transport inequalities. *J. Funct. Anal.* **275**(6), 1447–1483 (2018)
12. Cui, J., Freeden, W.: Equidistribution on the sphere. *SIAM J. Sci. Comput.* **18**(2), 595–609 (1997)
13. Delsarte, P., Goethals, J.M., Seidel, J.J.: Spherical codes and designs. *Geom. Dedic.* **6**(3), 363–388 (1977)
14. Dick, J., Brauchart, J.S.: Quasi-Monte Carlo rules for numerical integration over the unit sphere  $S^2$ . *Numer. Math.* **121**(3), 473–502 (2012)
15. Erdős, L., Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229**, 1435–1515 (2012)
16. Forrester, P.J.: *Log-Gases and Random Matrices* London. Mathematical Society Monographs 34. Princeton University Press, Princeton (2010)
17. García-Zelada, D.: Concentration for Coulomb gases on compact manifolds. *Elect. Commun. Probab.* **24**(18), 1–18 (2019)
18. Gourdon, X., Sebah, P.: The Riemann Zeta-function  $\zeta(s)$ : generalities. [numbers.computation.free.fr/Constants/Miscellaneous/zetagenralities.pdf](http://numbers.computation.free.fr/Constants/Miscellaneous/zetagenralities.pdf)
19. Gräf, M., Potts, D.: On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms. *Numer. Math.* **119**(4), 699–724 (2011)
20. Hirao, M.: QMC designs and determinantal point processes. In: “Monte Carlo and quasi-Monte Carlo methods”, 331–343, Springer Proc. Math. Stat., 241, Springer, Cham, (2018)
21. Hlawka, E.: Funktionen von beschränkter variation in der theorie der gleichverteilung. *Ann. Math. Pura Appl.* **54**, 325–333 (1961)
22. Hough, J.B., Krishnapur, M., Peres, YI., Virág, B.: Determinantal processes and independence. *Probab. Surv.* **3**, 206–229 (2006)
23. Koksma, J.F.: Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1. *Math. B Zutphen* **11**(43), 7–11 (1942)
24. Krishnapur, M.: From random matrices to random analytic functions. *Ann. Probab.* **37**(1), 314–346 (2009)
25. Li, P., Yau, S.-T.: On the parabolic kernel of the Schrödinger operator. *Acta Math.* **156**(3–4), 153–201 (1986)
26. Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere I. *Comm. Pure Appl. Math.* **39**, 149–186 (1986)

27. Lubotzky, A., Phillips, R., Sarnak, P.: Hecke operators and distributing points on the sphere II. *Comm. Pure Appl. Math.* **40**, 401–420 (1987)
28. Morpurgo, C.: The logarithmic Hardy-Littlewood-Sobolev inequality and extremals of zeta functions on  $S_n$ . *Geom. Funct. Anal.* **6**(1), 146–171 (1996)
29. Niederreiter, H.: Quasi-Monte Carlo methods and pseudo-random numbers. *Bull. Amer. Math. Soc.* **84**(6), 957–1041 (1978)
30. Rider, B., Virág, B.: The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not. IMRN* **2007**, rm006 (2007)
31. Rougerie, N., Serfaty, S.: Higher-dimensional Coulomb gases and renormalized energy functionals. *Comm. Pure Appl. Math.* **69**(3), 519–605 (2016)
32. Saff, E.B., Kuijlaars, A.B.J.: Distributing many points on a sphere. *Math. Intell.* **19**, 5–11 (1997)
33. Sheffield, S.: Gaussian free fields for mathematicians. *Probab. Theory Relat. Fields* **139**(3–4), 521–541 (2007)
34. Smale, S.: *Mathematical Problems for the Next Century, Mathematics: Frontiers and Perspectives*, pp. 271–294. AMS, Providence RI (2000)
35. Voros, A.: Spectral functions, special functions and the Selberg zeta function. *Comm. Math. Phys.* **110**(3), 439–465 (1987)
36. Womersley, R.S.: Efficient spherical designs with good geometric properties. *Contemporary computational mathematics—a celebration of the 80th birthday of Ian Sloan. Vol. 1, 2*, 1243–1285, Springer, Cham, (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.