

Volume distribution of nodal domains of random band-limited functions

Dmitry Beliaev¹  · Igor Wigman²

Received: 14 November 2016 / Revised: 10 October 2017 / Published online: 3 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract We study the volume distribution of nodal domains of families of naturally arising Gaussian random fields on generic manifolds, namely random band-limited functions. It is found that in the high energy limit a typical instance obeys a deterministic universal law, independent of the manifold. Some of the basic qualitative properties of this law, such as its support, monotonicity and continuity of the cumulative probability function, are established.

Mathematics Subject Classification 58J50 · 60F99 · 35P20

1 Introduction

A conjecture of Berry [3] was the original motivation for the study of nodal lines of a random plane wave. Nowadays, the study of the nodal lines of random plane waves and other smooth Gaussian fields is a well developed research area of its own right, also having numerous connections to other areas of mathematics and mathematical physics. We refer to a survey by Nazarov and Sodin [10] for more information. Bogomolny and Schmit [4] argued that some properties of the nodal domains of random plane wave behave similarly to the analogous properties of the critical percolation clusters

✉ Dmitry Beliaev
belyaev@maths.ox.ac.uk
Igor Wigman
igor.wigman@kcl.ac.uk

¹ Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, OX2 6GG Oxford, UK

² Department of Mathematics, King's College London, Strand, London WC2R 2LS, UK

corresponding to the square lattice; these include their number and the distribution of their areas.

It turns out that a lot of techniques could be extended to more general ensembles of random functions that we will discuss later on. In this paper we discuss the distribution of nodal domains' areas and length of their boundaries; we are following the footsteps of Nazarov and Sodin [9, 11], who developed some novel techniques to study the total number of nodal domains of smooth fields, and Sarnak and Wigman [12] who extended their tools to study finer questions of counting nodal domains of a given topological type and their mutual positions (“nestings”). Some of the methods in our paper are similar to these of the aforementioned papers; in particular, we borrow several technical results from these manuscripts. Some marked differences from these will be discussed at the end of the introduction.

Let (\mathcal{M}, g) be a compact smooth Riemannian n -manifold. For a smooth function $f : \mathcal{M} \rightarrow \mathbb{R}$ the nodal domains are the connected components of the complement $\mathcal{M} \setminus f^{-1}(0)$ of the nodal set, and we denote $\Omega(f)$ to be the collection of all nodal domains of f , and for $t > 0$ we denote $\mathcal{N}(f; t)$ to be the number of nodal domains $\omega \in \Omega(f)$ of volume

$$\text{Vol}(\omega) < t,$$

where $\text{Vol} = \text{Vol}_n$ is the n -dimensional volume on \mathcal{M} . We also define

$$\mathcal{N}(f) = \mathcal{N}(f; \infty) = |\Omega(f)|$$

to be the total number of nodal domains of f . In this paper we will investigate the behaviour of $\mathcal{N}(f; t)$ for several classes of random functions f .

Before introducing the most general result we would like to discuss one particular case (whose scaling limit is Berry’s random monochromatic waves) which is easy to explain and is representative in the proof of the general result. It is well-known that the space of spherical harmonics of degree l is of dimension $2l + 1$; let $\{\phi_{l,i}\}_{i=1 \dots 2l+1}$ be an arbitrary L^2 -orthonormal basis. Define a random Gaussian spherical harmonic

$$f_l = \sqrt{\frac{1}{2l + 1}} \sum_{i=1}^{2l+1} c_i \phi_{l,i}, \tag{1.1}$$

where c_i are i.i.d. standard Gaussian variables. The normalizing constant is chosen so that

$$\mathbb{E}[|f(z)|^2] = 1$$

for every $z \in \mathcal{S}^2$.

For the total number of nodal domains of f_l Nazarov and Sodin [9, 11] proved that there exists a constant $c_0 > 0$ (“Universal Nazarov–Sodin Constant”) so that

$$\mathbb{E} \left[\left| \frac{\mathcal{N}(f_l)}{l^2} - 4\pi c_0 \right| \right] \rightarrow 0, \tag{1.2}$$

i.e. that $\frac{\mathcal{N}(f_l)}{l^2}$ converges to the constant $4\pi \cdot c_0 > 0$ in mean, where $4\pi = \text{Vol}_2(\mathcal{S}^2)$ is the surface area of the unit 2-sphere. The constant c_0 in (1.2) will be denoted $c_0 = c(2, 1)$ as a particular case of a more general situation below [see (1.8)]. Our first principal result refines this.

Theorem 1 *Let f_l be the random spherical harmonic of degree l . Then the following holds:*

1. *There exists a monotone non-decreasing function*

$$\Psi = \Psi_{2,1} : (0, \infty) \rightarrow \mathbb{R}_+,$$

so that for all continuity points t of Ψ we have

$$\mathbb{E} \left[\left| \frac{\mathcal{N}(f_l; t/l^2)}{\mathcal{N}(f_l)} - \Psi(t) \right| \right] \rightarrow 0, \quad l \rightarrow \infty. \tag{1.3}$$

(Notation $\Psi_{2,1}$ will be clear from the formulation of Theorem 2.)

2. *Let*

$$t_0 = \pi j_{0,1}^2 = 18.168 \dots$$

where $j_{0,1} \approx 2.4048$ is the first zero of the Bessel function J_0 . Then the function Ψ defined above vanishes on $[0, t_0)$, and is strictly increasing for $t > t_0$.

As a monotone increasing function Ψ is continuous outside a countable set of jumps $\mathcal{T}_0 = \mathcal{T}_{0;2,1} = \{t_k\}_{k=1}^\infty$; hence (1.3) holds for all $t \in \mathbb{R}_{>0} \setminus \mathcal{T}_0$. The question whether we should expect for some genuine “distinguished” numbers $t \in \mathcal{T}_0$ that accumulate a positive proportion of nodal domain areas with strictly positive probability turned out to be quite tricky. On one hand, it is a priori possible to construct examples of spherical harmonics admitting lots of nodal domains of similar volume $\approx t$, for some candidate $t > 0$ for \mathcal{T}_0 . However, we may show that this situation is unstable w.r.t. small perturbations of the given spherical harmonics, by evaluating the derivatives of the volumes of these nodal domains w.r.t. the perturbation, and proving most of them to be bounded away from 0, with high probability.

The latter procedure seems very difficult to implement in order to rigorously rule out atoms, since the typical situation is more complicated, so that, even though the nodal domains of volume $\approx t$ before perturbation have been perturbed sufficiently, there might occur *other* domains, whose volume will be $\approx t$ after the perturbation. Nevertheless the above heuristic could serve as an argument in favour of believing in no “distinguished” areas, accumulating a positive proportions of nodal domain areas, at all, as in the following conjecture (or at least $|\mathcal{T}_0| < \infty$ is finite). This is also our grounds for believing the analogous conjecture in the more general setting (see Conjecture 2 below); if our intuition is correct, then Ψ is also differentiable everywhere, both in the restricted case $\Psi = \Psi_{2,1}$ (see the second part of Conjecture 1), and in the more general scenario.

Conjecture 1 *Let Ψ be the function prescribed by Theorem 1.*

1. The set of jumps of Ψ is empty, i.e. $\mathcal{T}_0 = \emptyset$, and Ψ is continuous.
2. The function Ψ is everywhere differentiable on (t_0, ∞) and its derivative (the probability density of the limiting distribution of nodal areas) is strictly positive.

Let us make a few remarks on Theorem 1. The spherical harmonics have a natural scale $1/l$, hence the natural area scaling is t/l^2 ; the area of a typical nodal domain is of order of magnitude $\frac{1}{l^2}$. Since the spherical harmonics are eigenfunctions of the spherical Laplacian, there is deterministic a lower bound on the nodal domains area. This will follow from the Faber–Krahn inequality which, in the 2-dimensional case, states that for a domain of area A its first Dirichlet eigenvalue of the Laplacian is at least $\pi j_{0,1}^2/A$. This implies that $\Psi(t) = 0$ for $t < t_0$. One may formulate and prove similar results on the distribution of lengths of nodal domains boundaries (equivalently, connected components of the nodal set $f_l^{-1}(0)$), called the *nodal components* of f_l (see also Sect. 6).

All the other principal results of this paper are in the same spirit as Theorem 1 but in different, less specialized, settings. The two main settings are: Euclidian (or “scale invariant”) case, and band-limited ensembles on manifolds; one recovers the former as scaling limits of the band-limited ensembles around every point of the given manifold. Below we briefly describe the various settings of random ensembles of functions treated in this paper.

1.1 Euclidian random fields

Here we are interested in centred Gaussian functions $F : \mathbb{R}^n \rightarrow \mathbb{R}$; it is a well known fact (Kolmogorov’s Theorem) that the distribution of a centred Gaussian field is completely determined by its covariance kernel

$$K(x, y) = \mathbb{E}[F(x) \cdot F(y)].$$

We will be interested in *isotropic* fields i.e. the fields such that

$$K(x, y) = K(|x - y|),$$

which means that F is invariant under translations and rotations. From now on we assume that all our fields on \mathbb{R}^n are isotropic, and moreover, we will also assume that F is normalized so that $K(0) = \mathbb{E}[F^2(x)] = 1$.

It is known that the such covariance kernel K can be expressed as the Fourier transform of a measure ρ , called the *spectral measure*. In many cases it is more convenient to describe the field F in term of its spectral measure instead of the covariance kernel. Given the spectral measure, there is an alternative way of constructing the random function F . It can be constructed as the Gaussian vector in the Hilbert space \mathcal{H} which is Fourier image of the symmetric space $L^2_{\text{sym}}(\rho)$. In particular this means that if $\{\phi_k\}_k$ is an orthonormal basis in \mathcal{H} , then

$$F = \sum \zeta_k \phi_k, \tag{1.4}$$

where ζ_k are i.i.d Gaussian random variables. This series a.s. diverges in \mathcal{H} but, under suitable assumptions on ρ , converges a.s., pointwise or in some other sense, to a well-defined function. The nature of the convergence depends on the properties of ρ ; in cases of our most interest the series (1.4) converges locally in C^k for every k . One of the most important motivational examples is the random plane wave:

Definition 1 The *random plane wave* or the *monochromatic wave* with energy $E = k^2$ is the centred Gaussian field on \mathbb{R}^2 with the covariance kernel

$$K(x, y) = J_0(k|x - y|),$$

where J_0 is the zeroth Bessel function.

Since the spectral measure is supported on the unit circle, the random plane wave is a solution of Helmholtz equation

$$\Delta f + k^2 f = 0 \tag{1.5}$$

on \mathbb{R}^2 . One may think of the random plane wave as a “random” solution of (1.5). It is possible to express the random plane wave in terms of an orthonormal basis in the corresponding Hilbert space via the formula (1.4). A canonical basis for this Hilbert space is given in terms of Bessel functions J_k , and, in polar coordinates $x = re^{i\theta}$, the function F could be written as

$$F(re^{i\theta}) = \Re \sum_{n=-\infty}^{\infty} c_n J_{|n|}(kr) e^{in\theta}, \tag{1.6}$$

where the coefficients c_n are i.i.d. standard complex Gaussians. More details about the construction of the random plane wave and its relation to the random spherical harmonics could be found in [9].

A direct computation shows that the covariance kernel of this function is indeed $K(x, y) = J_0(k|x - y|)$. The spectral measure of F is the normalized Lebesgue measure on the circle of radius k . Since plane waves with different values of k differ by the scaling, it is natural to fix $k = 1$. The description in terms of the spectral measure has a natural generalization to the higher dimensions:

Definition 2 The random plane wave in \mathbb{R}^n is the Gaussian field whose spectral measure is the normalized $(n - 1)$ -dimensional Lebesgue measure on $S^{n-1} \subset \mathbb{R}^n$.

In Theorem 5 below we will show that under relatively mild conditions on the spectral measure, an analogue of Theorem 1 holds for F . Theorem 5 will be proved in great generality, though the most important, relevant for Theorem 1, is the case of the random monochromatic plane-wave, of significant importance by itself.

1.2 Band-limited functions

Let (\mathcal{M}, g) be a compact Riemannian n -manifold, then the eigenfunctions of the Laplacian $\{\phi_i\}_{i \geq 1}$ form an orthonormal basis of $L^2(\mathcal{M})$. We denote the square roots of eigenvalues by $0 = t_0 \leq t_1 \leq t_2 \dots$ i.e. satisfying

$$\Delta \phi_i + t_i^2 \phi_i = 0.$$

For a fixed $\alpha \in [0, 1)$ and $T \gg 0$ a large spectral parameter we define the α -band-limited functions (corresponding to T)

$$f(x) = f_{\alpha;T}(x) = \sum_{\alpha T \leq t_j \leq T} c_j \phi_j(x). \tag{1.7}$$

where c_j are independent real Gaussian variables of mean 0 and variance 1. For $\alpha = 1$ we define $f_{1;T}$ by

$$f(x) = f_{1;T}(x) = \sum_{T-\eta(T) \leq t_j \leq T} c_j \phi_j(x),$$

where η is a function growing to infinity slower than T , i.e. $\eta(T) \rightarrow \infty$ as $T \rightarrow \infty$ and $\eta(T) = o(T)$.

The random spherical harmonic (1.1) defined above is an $\alpha = 1$ band-limited function on the unit sphere S^2 with $\alpha = 1$ and $\eta(T) = O(T^{1/2})$. For the total number of nodal domains of the band limited functions Nazarov and Sodin [11, 13] proved that for every $\alpha \in [0, 1]$, $n \geq 2$ there exist a constant $c(n, \alpha) > 0$ (“Nazarov–Sodin constant”¹) satisfying

$$\mathbb{E} \left[\left| \frac{\mathcal{N}_\Omega(f)}{T^n} - c(n, \alpha) \cdot \text{Vol}_n(\mathcal{M}) \right| \right] \rightarrow 0. \tag{1.8}$$

Sarnak and Wigman [12] refined the latter result (1.8) for counting the number of nodal domains (or components) of f of a given topological class; they also found an elegant way to formulate it in terms of convergence of random probability measures consolidating all topological types into a universal deterministic probability measure that conserves all the topologies. Gayet and Welshinger [8] proved lower (and upper) bounds for the expected number of domains of a given topological class in a different ensemble of random polynomials in the high degree limit. We will show below that a result similar to Theorem 1 holds for band-limited functions.

Theorem 2 *Let $f = f_{\alpha;T}$ be a band-limited function on some n -dimensional compact manifold \mathcal{M} . There exists a monotone non-decreasing function*

$$\Psi = \Psi_{n;\alpha} : (0, \infty) \rightarrow \mathbb{R}_+,$$

¹ Note the different normalization as compared to [12].

so that for all continuity points t of Ψ we have

$$\mathbb{E} \left[\left| \frac{\mathcal{N}_\Omega(f; t/T^{n/2})}{\mathcal{N}_\Omega(f)} - \Psi(t) \right| \right] \rightarrow 0, \quad T \rightarrow \infty. \quad (1.9)$$

Importantly, the function Ψ prescribed by Theorem 2 depends on dimension n , but not on the manifold.

1.3 Some properties of the limiting distribution $\Psi_{n;\alpha}$

Here we investigate the most basic property of $\Psi_{n;\alpha}$, i.e. its monotonicity. We need to distinguish between $\alpha < 1$, where the corresponding distribution function $\Psi_{n;\alpha}$ is strictly positive and increasing everywhere (Theorem 3), and $\alpha = 1$, where the behaviour of the function $\Psi_{n;1}$ is more complicated (Theorem 4).

Theorem 3 *For every $n \geq 2$, $\alpha < 1$ the function $\Psi_{n;\alpha}$ as above is strictly increasing on \mathbb{R}_+ .*

Theorem 4 *For every $n \geq 2$ the function $\Psi_{n;1}$ vanishes on $[0, t_0]$ where*

$$t_0 = t_0(n) = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} j_{n/2-1,1}^n$$

is the volume of the ball of radius $j_{n/2-1,1}$ – the first zero of the Bessel function $J_{n/2-1}$. Moreover, $\Psi_{n;1}$ is strictly increasing on (t_0, ∞) .

Motivated by similar arguments to Conjecture 1 (see the couple of paragraphs preceding Conjecture 1) it is only natural to conjecture the following:

Conjecture 2 *The function $\Psi_{n;\alpha}$ is continuous, everywhere differentiable. For $\alpha < 1$ the derivative $\frac{d\Psi_{n;\alpha}(t)}{dt} > 0$ is everywhere positive, whereas for $\alpha = 1$ the same holds for $t > t_0$.*

1.4 Main ideas and the plan of the paper

Our general strategy is similar to [11, 12]; for a start, in Sect. 2.1 we give the necessary background on the behaviour of band-limited functions in the semiclassical limit. Roughly speaking, we show that on small scale bigger than $1/T$ the rescaled version $f_{\alpha;T}$ of the band-limited function is well-approximated locally around every point $x \in \mathcal{M}$ by its limit $\mathfrak{g}_{n,\alpha}$ defined on the tangent space $T_x\mathcal{M}$. Importantly, $\mathfrak{g}_{n,\alpha}$ is universal, i.e. it does not depend on manifold or x , and it is sufficiently explicit for us to conduct local computations. This will allow us to prove the main results for $\mathfrak{g}_{n,\alpha}$, and then use the effective convergence to deduce the results for the band-limited functions. In Sect. 2.2 we formulate several Kac–Rice type results that will yield universal upper bounds on various local quantities like the number of nodal domains or nodal components.

In Sect. 3 we discuss the behaviour of Gaussian fields on \mathbb{R}^n . We formulate and prove Theorem 5, a Euclidean version of Theorem 2, holding under very mild conditions on a random field, including all of the $\mathfrak{g}_{n,\alpha}$. The proof of Theorem 2 is based on *a priori* upper bounds via Kac–Rice, and an ergodic theorem, yielding the existence of the scaling limit of the volume distribution. Towards the end of Sect. 3 where Theorems 6 and 7 on the distribution function $\Psi_{n,\alpha}$ are proved; these are Euclidean analogues of Theorems 3 and 4, covering the generic case $\alpha < 1$, and $\alpha = 1$ respectively. The proofs of Theorems 6 and 7 are rather similar. Here our first goal is establishing the existence of a *deterministic* function with nodal domain containing the origin of volume approximating the given number. Then we show that the same holds for functions approximating the postulated one in C^1 -norm (see Lemma 10); finally we claim that the latter happens with positive probability. The translation invariance of the underlying random fields yields that, as the claimed result holds near origin with positive probability, it holds with positive *density*.

In Sect. 4 we prove that since functions $f_{\alpha;T}$ and $\mathfrak{g}_{n,\alpha}$ are close to each other (after coupling and rescaling), their respective numbers of nodal domains of restricted volume are close with high probability (this is quantified in Proposition 1). In Sect. 5 we prove the main theorems of the paper. The proofs are based on *semi-locality* of nodal domains, that is, that most of the nodal domains of $f_{\alpha;T}$ are neither too small nor too long, i.e. that the semi-local approximation by $\mathfrak{g}_{n,\alpha}$ captures most of the nodal domains of $f_{\alpha;T}$. This allows to infer the results on $f_{\alpha;T}$ from the analogous results on $\mathfrak{g}_{n,\alpha}$. Finally, in Sect. 6 we make some final remarks about the proof, in particular we explain that with some minor modification our methods imply similar results for the distribution of the surface volume of the nodal components or even for the joint distribution of the volumes of nodal domains and boundary volumes.

Despite the fact that our general approach follows the footsteps of Nazarov and Sodin [11, 13] and Sarnak and Wigman [12] (and Canzani and Sarnak [7] for the full support statement in the monochromatic case $\alpha = 1$), our case offers new significant challenges on both the Euclidean stage and the inference of the Riemannian one, and also for proving the full support statement on the volume distribution. First, as we are interested in the number of nodal domains of a given approximate volume, we need to refine the techniques to control the volumes of the perturbed nodal domains rather than their mere number; though it shares some similarities with [12], the continuous variable t makes the analysis more challenging as compared to the purely discrete (and hence atomic) case in [12]. While the more refined technical work is a relatively standard application of techniques in differential geometry (Lemma 10), if some $t > 0$ happens to be an atom of the limiting distribution, it does not guarantee that t maintains its mass after the perturbation. In fact, as a result of the perturbation its mass might spread in an infinitesimal neighbourhood $(t - \epsilon, t + \epsilon)$ of t ; this is the reason why the main results are only formulated for continuity points of the limit distribution function ψ .

Second, as part of the full support statement for the limit distribution (Theorems 3 and 4), we need to construct a deterministic function in some space of functions with a nodal domain of a prescribed volume. In the monochromatic case this is only possible if it obeys the Faber–Krahn inequality; in this case we show that the converse is also valid, whence we are building on [7] and refine it for our needs. Finally, while an analogue

of the ergodic approach of Nazarov and Sodin [11, 13] yields the mere *existence* of $\Psi(t)$ for every given $t > 0$ the properties of Ψ are far from being obvious. In fact, it could be that the atoms of the limit distribution corresponding to Ψ concentrate all the probability.

2 Necessary background

2.1 Semiclassical properties of band-limited functions and their scaling limits

In this section we introduce a few facts about band-limited functions. The results are stated without proofs, for more detailed discussion we refer the readers to [12, Section 2.1] and references therein.

For the band-limited function $f_{\alpha;T}$ we have the covariance function

$$K_{\alpha;T}(x, y) := \mathbb{E} [f_{\alpha;T}(x) \cdot f_{\alpha;T}(y)] = \sum_{\alpha T \leq t_j \leq T} \phi_j(x)\phi_j(y),$$

with the same conventions as above for $\alpha = 1$. The following semiclassical approximation, independent of \mathcal{M} , holds (see [12, Section 2.1] with case $\alpha = 1$ due to Canzani and Hanin [5, 6]):

$$\widetilde{K}_{\alpha}(T; x, y) := \frac{1}{D_{\alpha}(T)} K_{\alpha}(T; x, y) = B_{n,\alpha}(T \cdot d(x, y)) + O(T^{-1}), \tag{2.1}$$

uniformly for $x, y \in \mathcal{M}$, where $d(x, y)$ is the (geodesic) distance in \mathcal{M} between x and y ,

$$D_{\alpha}(T) = \frac{1}{\text{Vol}(\mathcal{M})} \int_{\mathcal{M}} K_{\alpha}(T; x, x) d \text{Vol}(x),$$

and for $w \in \mathbb{R}^n$

$$B_{n,\alpha}(w) = B_{n,\alpha}(|w|) = \frac{1}{|A_{\alpha}|} \int_{A_{\alpha}} e^{2\pi i \langle w, \xi \rangle} d\xi \tag{2.2}$$

where $A_{\alpha} = \{w : \alpha \leq |w| \leq 1\}$ and in the case $\alpha = 1$ the n -dimensional measure $d\xi$ is replaced by $(n - 1)$ -dimensional surface measure on the unit sphere.

We may differentiate both sides of (2.1) to obtain asymptotic expression for *finitely* many derivatives of K_{α} . By appropriately normalizing f_{α} we may assume w.l.o.g that \widetilde{K}_{α} is the covariance of f_{α} and we will neglect this difference from this point on.

Around each point $x \in \mathcal{M}$ we define the scaled random fields $f_{x;T}$ (we drop α from notations) on a big ball (after scaling) lying on the Euclidian tangent space $\mathbb{R}^n \cong T_x \mathcal{M}$ with the use of the exponential map and an isometry $I_x : \mathbb{R}^n \rightarrow T_x \mathcal{M}$, $\Phi_x = \exp_x \circ I_x : \mathbb{R}^n \rightarrow \mathcal{M}$ via

$$f_{x;T}(u) := f_T(\Phi_x(u/T)), \tag{2.3}$$

with the covariance function

$$K_{x;T}(u, v) := \mathbb{E}[f_{x;T}(u) \cdot f_{x;T}(v)] = K_T(\Phi_x(u/T), \Phi_x(v/T)).$$

Observe that locally Φ_x is almost an isometry: for each ξ there exists $r_0 = r_0(\xi)$ such that if $\mathcal{D} \subseteq B_x(r_0)$ is a smooth domain in \mathcal{M} , then

$$|\text{Vol}_{\mathcal{M}}(\mathcal{D}) - \text{Vol}_{\mathbb{R}^n}(\Phi_x^{-1}(\mathcal{D}))| < \xi, \tag{2.4}$$

uniformly w.r.t. $x \in \mathcal{M}$ (r_0 is assumed to be sufficiently small so that the exponential map is $1 - 1$). By the scaling (2.3) we obviously have

$$\mathcal{N}\left(f, \frac{t}{T^n}; x, \frac{R}{T}\right) \approx \mathcal{N}(f_{x;T}, t; x, R),$$

the precise meaning will be given in (4.3). This means that studying $f_{\alpha;T}$ and $f_{x;T}$ is essentially equivalent. Finally, from (2.1) we see that the covariance kernel of $f_{x;T}$ converges to

$$r_{n,\alpha}(u, v) = B_{n,\alpha}(|u - v|).$$

This suggests to define the local scaling limit $\mathfrak{g}_{n,\alpha}$ to be a Gaussian function in \mathbb{R}^n with this covariance kernel. Alternatively, it could be defined by its spectral measure which by (2.2) is the normalized Lebesgue measure on A_α (or the normalized surface area on A_1 for $\alpha = 1$). It is important to point out that the scaling limit is universal: it does not depend on x or \mathcal{M} , but only on n and α .

From the above it follows that for T sufficiently large, the covariance kernels of $f_{x;T}$ are asymptotic to those of $\mathfrak{g}_{n,\alpha}$. However it is not obvious that this implies that with high probability $f_{x;T}$ and $\mathfrak{g}_{n,\alpha}$ are close (which is essential for our approach). One obstacle we need to overcome is that the random fields $f_{x;T}$ and $\mathfrak{g}_{n,\alpha}$ are defined on different probability spaces. However, it is still possible to couple $f_{x;T}$ and $\mathfrak{g}_{n,\alpha}$ so that $f_{x;T} - \mathfrak{g}_{n,\alpha}$ is small in C^1 on a ball of arbitrarily big radius. To establish the latter statement (in a suitable sense to be formulated rigorously) one would restrict the respective fields on a sufficiently dense grid and use the approximation of the multivariate distribution to approximate the finite-dimensional Gaussian variables. One then extends the respective fields to be defined smoothly w.r.t. the continuous variable on the relevant domain, still maintaining the C^1 -approximation property. The above procedure was carried out rigorously in [13, Lemma 4].

Lemma 1 ([13, Lemma 4]) *Given $R > 0$ and $b > 0$, there exists $T_0 = T_0(R, b)$ such that for all $T \geq T_0$ we have*

$$\mathbb{E}\left[\|f_{x;T} - \mathfrak{g}_{n,\alpha}\|_{C^1(\overline{B}(2R))}\right] < b.$$

2.2 The Kac–Rice premise

In this section we collect a number of *local* results required below. The Kac–Rice formula is a powerful tool for computing moments of local quantities, and in principle it is capable of expressing *any* moment of a local quantity of a given random Gaussian field F in terms of explicit, albeit complicated, Gaussian expectations, or, equivalently, in terms of the covariance kernel.

Let $m \leq n$, $F : \mathcal{D} \rightarrow \mathbb{R}^m$ be a smooth random field on a domain $\mathcal{D} \subseteq \mathbb{R}^n$, and $\mathcal{Z}(F; \overline{\mathcal{D}})$ be either the $(n - m)$ -volume of $F^{-1}(0)$ (for $m < n$), or the number of the discrete zeros (for $m = n$). For example, if $H : \mathcal{M} \rightarrow \mathbb{R}$ is a random field and $F = \nabla H|_{\mathcal{D}} : \mathcal{M} \rightarrow \mathbb{R}^n$ is its gradient restricted to a coordinate patch, then $\mathcal{Z}(F, \overline{\mathcal{D}})$ counts the number of critical points of H on \mathcal{D} . We set $J_F(x)$ to be the (random) Jacobi matrix of F at x , and define the “zero density” of F at $x \in \mathcal{D}$ as the conditional Gaussian expectation

$$K_1(x) = K_{1;F}(x) = \phi_{F(x)}(0) \cdot \mathbb{E}[|\det J_F(x)| | F(x) = 0], \tag{2.5}$$

where $\phi_{F(x)}(0, \dots, 0)$ is the m -variate Gaussian density of $F(x)$ evaluated at $(0, \dots, 0) \in \mathbb{R}^m$; if the distribution of $F(x)$ in non-degenerate with covariance matrix $D = D(x)$, then we have

$$\phi_{F(x)}(0, \dots, 0) = \frac{1}{(2\pi)^{m/2} \sqrt{\det D}}.$$

With the above notation the Kac–Rice formula (meta-theorem) states that, under some non-degeneracy condition on F ,

$$\mathbb{E}[\mathcal{Z}(F; \mathcal{D})] = \int_{\mathcal{D}} K_1(x) dx.$$

Concerning the sufficient conditions that guarantee that (2.5) holds, a variety of results is known [1, 2]. The following version of Kac–Rice merely requires the non-degeneracy of the values of F (vs. the non-degeneracy of $(F, J_F(x))$ in the appropriate sense, as in some more classical sources), to our best knowledge, the mildest sufficient condition. All random fields considered in this paper satisfy these assumptions.

Lemma 2 (Standard Kac–Rice [2, Theorem 6.2]) *Let $F : \mathcal{D} \rightarrow \mathbb{R}^m$ be an a.s. C^2 -smooth Gaussian field, such that for every $x \in \mathcal{D}$ the distribution of the random vector $F(x) \in \mathbb{R}^m$ is non-degenerate Gaussian, with zero almost surely not a critical value. Then*

$$\mathbb{E}[\mathcal{Z}(F; \mathcal{D})] = \int_{\mathcal{D}} K_1(x) dx \tag{2.6}$$

with the zero density $K_1(x)$ as in (2.5).

Condition (iv) of [2, Theorem 6.2] that zero is not a critical level is automatically satisfied by all C^2 -smooth Gaussian fields; this is known as Bulinskaya’s Lemma, see [2, Proposition 6.12].

The following lemma is an upper bound for either the number of critical points of a random field or its restriction to a hypersphere as a result of an application of Kac–Rice on coordinate patches.

Lemma 3 ([12, Corollary 2.3]) *Let $\mathcal{D} \subseteq \mathbb{R}^m$ be a domain and $F : \mathcal{D} \rightarrow \mathbb{R}$ an a.s. C^2 -smooth stationary Gaussian random field, such that for $x \in \mathcal{D}$ the distribution of $\nabla F(x)$ is non-degenerate Gaussian.*

1. For $r > 0$ let $\mathcal{A}(F; r)$ be the number of critical points of F inside $B(r) \subseteq \mathcal{D}$. Then

$$\mathbb{E}[\mathcal{A}(F; r)] = O(\text{Vol}(B(r))),$$

where the constant involved in the ‘ O ’-notation depends on the law of F only.

2. For $r > 0$ let $\tilde{\mathcal{A}}(F; r)$ be the number of critical points of the restriction $F|_{\partial B(r)}$ of F to the sphere $\partial B(r) \subseteq \mathcal{D}$. Then

$$\mathbb{E}[\tilde{\mathcal{A}}(F; r)] = O(\text{Vol}(\partial B(r))),$$

where the constant involved in the ‘ O ’-notation depends on the law of F only.

The following estimate is the upper bound part of the (precise) Kac–Rice formula applied to the band limited functions.

Lemma 4 ([13, Lemma 2] and [12, Lemma 7.8]) *For $x \in \mathcal{M}$, $r > 0$ let $\mathcal{N}_\Omega(f_{\alpha;T}; x, r)$ be the number of nodal domains of $f_{\alpha;T}$ entirely lying in $B(x, r) \subseteq \mathcal{M}$: the geodesic ball of radius r centred at x . Then*

$$\mathbb{E}[\mathcal{N}_\Omega(f_{\alpha;T}; x, r)] = O(r^n \cdot T^n),$$

with constant involved in the ‘ O ’-notation depending on \mathcal{M} and α only.

3 Distribution of nodal domain areas for Euclidian fields

3.1 Notation and statement of the main result on Euclidian fields

3.1.1 Notation and basic setup

Let f be a smooth function, $t > 0$, and $R > 0$. We denote $\mathcal{N}(f, t; R)$ to be the number of domains $\omega \in \Omega(f)$ of f lying entirely in $B(R)$ of volume

$$\text{Vol}_n(\omega) \leq t;$$

note that

$$\mathcal{N}(f; R) = \mathcal{N}(f, \infty; R)$$

is the total number of nodal domains lying inside $B(R)$, as considered by Nazarov and Sodin [9].

We are interested in the asymptotic distribution² of the nodal domain volumes, that is, the asymptotic behaviour of $\mathcal{N}(F, t; R)$ as $R \rightarrow \infty, t > 0$ fixed; throughout this section we will tacitly assume that F is stationary, so that its spectral measure makes sense. We would like to establish the limit

$$\Psi_\rho(t) := \lim_{R \rightarrow \infty} \frac{\mathcal{N}(F, t; R)}{\mathcal{N}(F; R)} \tag{3.1}$$

in mean, and, moreover, that $\Psi_\rho(t)$ is a distribution function, i.e.

$$\lim_{t \rightarrow \infty} \Psi_\rho(t) = 1.$$

The latter will follow once having established the former (in the proper sense) via the obvious *deterministic* upper bound

$$\mathcal{N}(F; R) - \mathcal{N}(F, t; R) \leq \frac{\text{Vol } B(R)}{t}$$

for the number of domains of volume greater than t .

Following Nazarov and Sodin [11] we assume that the spectral measure ρ of F satisfies the following axioms:

- ($\rho 1$) The measure ρ has no atoms (if and only if the action of the translations is ergodic by Grenander–Fomin–Maruyama, see [13, Theorem 3]).
- ($\rho 2$) For some $p > 4$,

$$\int_{\mathbb{R}^n} |\lambda|^p d\rho(\lambda) < \infty$$

(this ensures the a.s. C^2 -smoothness of F).

- ($\rho 3$) The spectral support $\text{supp } \rho$ does not lie in a linear hyperplane. (This ensures that the random Gaussian field, together with its gradient is not degenerate.)

For this model Nazarov and Sodin [11] proved that there exists a constant $c(\rho) \geq 0$ (the “Nazarov–Sodin constant”) so that

$$\frac{\mathcal{N}(F; R)}{\text{Vol } B(R)} \rightarrow c(\rho) \tag{3.2}$$

both in mean and a.s.

² To make sense of the distribution of $\mathcal{N}(F, \cdot; \cdot)$ it is essential to show that the latter is a *random variable*, i.e. a measurable function on the sample space. Fortunately, the proof of a similar statement, given in [12, Appendix A], is sufficiently robust to cover our case and all the other similar quantities of this manuscript; from this point on we will neglect any issue of measurability.

It is shown in [11] that under the additional mild condition the constant $c(\rho)$ is strictly positive. We do not want to discuss these technicalities, so instead we will use the assumption

(ρ_4) The Nazarov–Sodin constant $c(\rho)$ is positive.

Sometimes we will invoke a stronger axiom:

(ρ_4^*) The support of the spectral measure ρ has non-empty interior.

3.1.2 Existence of limiting distribution Ψ_ρ

Theorem 5 *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be a stationary random field whose spectral measure ρ satisfies the axioms (ρ_1) – (ρ_3), then*

$$\frac{\mathcal{N}(F, t; R)}{\text{Vol}(B(R))}$$

converges in mean as $R \rightarrow \infty$. If we additionally assume the axiom (ρ_4), then the limit

$$\lim_{R \rightarrow \infty} \frac{\mathcal{N}(F; R)}{\text{Vol}(B(R))} = c(\rho) > 0$$

does not vanish, so that we can define the normalized limit

$$\Psi_\rho(t) := \lim_{R \rightarrow \infty} \frac{\mathcal{N}(F, t; R)}{c(\rho) \cdot \text{Vol}(B(R))}. \tag{3.3}$$

Since the total number of nodal domains of F lying inside $B(R)$ was proven to be asymptotic to

$$\mathcal{N}(F; R) \sim c(\rho) \cdot \text{Vol } B(R),$$

[see (3.2)], (3.3) may be equivalently read as

$$\frac{\mathcal{N}(F, t; R)}{\mathcal{N}(F; R)} \rightarrow \Psi_\rho(t),$$

[cf. (3.1)]; this limit could be proven in mean, see the proof of Theorem 2 in Sect. 5.1. Theorem 5 in particular implies that for every $t > 0$ the expected number $\mathcal{N}(F, t; R)$ of nodal domains of F of volume at most t lying in $B(R)$ is

$$\mathbb{E}[\mathcal{N}(F, t; R)] = c(\rho) \cdot \Psi_\rho(t) \cdot \text{Vol}(B(R))(1 + o_{R \rightarrow \infty}(1)), \tag{3.4}$$

with concentration: for every $\epsilon > 0$

$$\lim_{R \rightarrow \infty} \mathcal{P} \left\{ \left| \frac{\mathcal{N}(F, t; R)}{\text{Vol } B(R)} - c(\rho) \cdot \Psi_\rho(t) \right| > \epsilon \right\} = 0, \tag{3.5}$$

via Chebyshev’s inequality.

3.1.3 Some properties of $\Psi_\rho(t)$

Theorem 6 (Lower bound on domains in the generic case) *Assume that the spectral measure of F satisfies axioms $(\rho 1) - (\rho 3)$ and $(\rho 4^*)$. Then $\Psi_\rho(0) = 0$, and $\Psi_\rho(\cdot)$ is strictly increasing on $[0, \infty)$; in particular, for every $t > 0$ we have $\Psi_\rho(t) > 0$.*

For the random plane wave (RPW) the situation is slightly different. Its spectral measure does not satisfy $(\rho 4^*)$ but satisfies $(\rho 4)$. The corresponding function $\Psi = \Psi_{RPW}$ vanishes up to a certain explicit threshold. The precise statement is given in the following theorem.

Theorem 7 *Let $F = F_{RPW}$ be the random plane wave in \mathbb{R}^n . As its spectral measure satisfies axioms $(\rho 1) - (\rho 4)$, the function $\Psi_{RPW} = \Psi_\rho$ defined as in Theorem 5 exists. Define*

$$t_0 = t_0(n) = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} j_{n/2-1,1}^n$$

as in Theorem 4. Then the following holds:

1. For every nodal domain ω of F_{RPW} we have $\text{Vol}(\omega) \geq t_0$, and hence, in particular, we have

$$\Psi_{RPW}(t) = 0, \quad t < t_0$$

2. Moreover, for $n = 2$ the function Ψ is strictly increasing on $[t_0, \infty)$

The remaining part of Sect. 3 is dedicated to the proofs of Theorems 5, 6 and 7.

3.2 Integral-Geometric Sandwich

Let $\Gamma \subseteq \mathbb{R}^n$ be a hypersurface (a curve for $n = 2$). For $u \in \mathbb{R}^n$, $r > 0$ and $t > 0$ we denote $\mathcal{N}(\Gamma, t; u, r)$ the number of domains of Γ of n -dimensional volume bounded by t lying entirely in the radius- r ball $B_u(r)$ centred at u . Similarly, define $\mathcal{N}^*(\Gamma, t; u, r)$ by relaxing the condition to domains merely intersecting $\overline{B_u(r)}$. We use the shortcuts

$$\mathcal{N}(g, t; u, r) := \mathcal{N}(g^{-1}(0), t; u, r),$$

respectively

$$\mathcal{N}^*(g, t; u, r) := \mathcal{N}^*(g^{-1}(0), t; u, r),$$

and $\mathcal{N}(\cdot, \cdot; r) := \mathcal{N}(\cdot, \cdot; 0, r)$ (resp. $\mathcal{N}^*(\cdot, \cdot; r) := \mathcal{N}^*(\cdot, \cdot; 0, r)$), consistent to Sect. 3.1.1. Finally, let $\mathcal{N}(g; u, r) = \mathcal{N}(g, \infty; u, r)$ be the total number of domains lying inside $B_0(r)$.

Lemma 5 (cf. [13, Lemma 1]) *Let Γ be a closed hypersurface. Then for $0 < r < R$, $t > 0$,*

$$\begin{aligned} \frac{1}{\text{Vol}(B(r))} \int_{B(R-r)} \mathcal{N}(\Gamma, t; u, r) du &\leq \mathcal{N}(\Gamma, t; R) \\ &\leq \frac{1}{\text{Vol}(B(r))} \int_{B(R+r)} \mathcal{N}^*(\Gamma, t; u, r) du. \end{aligned} \tag{3.6}$$

Proof We follow along the lines of the proof of [13, Lemma 1]: For a connected component γ of Γ denote $\mathcal{A}(\gamma)$ to be the area of the domain having γ as its outer boundary.

Let γ be any connected component of Γ . Define

$$G_*(\gamma) = \bigcap_{v \in \gamma} B_v(r) = \{u : \gamma \subseteq B_u(r)\}$$

and

$$G^*(\gamma) = \bigcup_{v \in \gamma} \overline{B_v(r)} = \{u : \gamma \cap \overline{B_u(r)} \neq \emptyset\}.$$

We have for every $v \in \gamma$,

$$G_*(\gamma) \subseteq B_v(r) \subseteq G^*(\gamma),$$

and thus, in particular,

$$\text{Vol}(G_*(\gamma)) \leq \text{Vol}(B_v(r)) \leq \text{Vol}(G^*(\gamma)). \tag{3.7}$$

Summing up (3.7) for all connected components $\gamma \subseteq \Gamma$ lying inside $B(R)$, corresponding to domains of volume at most t , we obtain

$$\sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \subseteq B(R)}} \text{Vol}(G_*(\gamma)) \leq \text{Vol}(B(r)) \cdot \mathcal{N}(\Gamma, t; R) \leq \sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \subseteq B(R)}} \text{Vol}(G^*(\gamma)). \tag{3.8}$$

Writing the volume as an integral

$$\text{Vol}(G_*(\gamma)) = \int_{G_*(\gamma)} du,$$

and exchanging the order of summation and integration we obtain

$$\sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \subseteq B(R)}} \text{Vol}(G_*(\gamma)) \geq \int_{B(R-r)} \left[\sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \subseteq B_u(r)}} 1 \right] du = \int_{B(R-r)} \mathcal{N}(\Gamma, t; u, r) du, \tag{3.9}$$

since if $u \in B(R - r)$ then $B_u(r) \subseteq B(R)$. Similarly,

$$\sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \subseteq B(R)}} \text{Vol}(G^*(\gamma)) \leq \int_{B(R+r)} \left[\sum_{\substack{\gamma: \mathcal{A}(\gamma) \leq t \\ \gamma \cap \overline{B_u(r)} \neq \emptyset}} 1 \right] du = \int_{B(R+r)} \mathcal{N}^*(\Gamma, t; u, r) du, \tag{3.10}$$

since if $\gamma \subseteq B(R)$ and for some u , $\overline{B_u(r)} \cap \gamma \neq \emptyset$, then necessarily $u \in B(R + r)$. The statement of the present lemma then follows from substituting (3.9) and (3.10) into (3.8), and dividing both sides by $\text{Vol } B(R)$. \square

3.3 Proof of Theorem 5

Proof Let $t > 0$ be given, and fix $r > 0$; we apply (3.6) to $\Gamma = F^{-1}(0)$:

$$\begin{aligned} & \left(1 - \frac{r}{R}\right)^n \frac{1}{\text{Vol } B(R - r)} \int_{B(R-r)} \frac{\mathcal{N}(F, t; u, r)}{\text{Vol}(B(r))} du \leq \frac{\mathcal{N}(F, t; R)}{\text{Vol}(B(R))} \\ & \leq \left(1 + \frac{r}{R}\right)^n \cdot \frac{1}{\text{Vol } B(R + r)} \int_{B(R+r)} \frac{\mathcal{N}^*(F, t; u, r)}{\text{Vol}(B(r))} du \\ & \leq \left(1 + \frac{r}{R}\right)^n \frac{1}{\text{Vol } B(R + r)} \int_{B(R+r)} \frac{\mathcal{N}(F, t; u, r) + \mathcal{C}(u, r; t, F)}{\text{Vol}(B(r))} du, \end{aligned}$$

where $\mathcal{C}(u, r; t, F)$ is the number of domains $\omega \in \Omega(F)$ intersecting $\partial B_u(r)$, of volume

$$\text{Vol}(\omega) \leq t,$$

bounded by the total number of critical points of the restriction $F|_{\partial B_u(r)}$ of F to the hypersphere $\partial B_u(r)$, and

$$\text{Vol}(B(R \pm r)) = \text{Vol}(B(R)) \cdot \left(1 \pm \frac{r}{R}\right)^n.$$

We rewrite $\mathcal{N}(F, t; u, r) = \mathcal{N}(\tau_u F, t; r)$, where τ_u is the translation operator

$$(\tau_u F)(x) = F(u + x).$$

Choose r so small that $(1 \pm r/R)^n$ are ϵ -close to 1, then we have

$$\begin{aligned} & (1 - \epsilon) \frac{1}{\text{Vol } B(R - r)} \int_{B(R-r)} \frac{\mathcal{N}(\tau_u F, t; r)}{\text{Vol}(B(r))} du \leq \frac{\mathcal{N}(F, t; r)}{\text{Vol}(B(R))} \\ & \leq (1 + \epsilon) \frac{1}{\text{Vol } B(R + r)} \int_{B(R+r)} \frac{\mathcal{N}(\tau_u F, t; r) + \mathcal{C}(\tau_u F, t; r)}{\text{Vol}(B(r))} du. \tag{3.11} \end{aligned}$$

Note that for every r, t the functional

$$F \mapsto \Upsilon_{r;t}(F) := \frac{\mathcal{N}(F, t; r)}{\text{Vol}(B(r))}$$

(and its translations) is measurable, and since the number of nodal domains is bounded by the number of critical points, this functional has finite expectation by Lemma 3, part (1). It then follows from the ergodic theory that both

$$\frac{1}{\text{Vol } B(R+r)} \int_{B(R+r)} \frac{\mathcal{N}(\tau_u F, t; r)}{\text{Vol}(B(r))} du$$

and

$$\frac{1}{\text{Vol } B(R-r)} \int_{B(R-r)} \frac{\mathcal{N}(\tau_u F, t; r)}{\text{Vol}(B(r))} du$$

converge to (the same) limit in L^1

$$\frac{1}{\text{Vol } B(R)} \int_{B(R)} \frac{\mathcal{N}(\tau_u F, t; r)}{\text{Vol}(B(r))} du \rightarrow c_{r;t}(\rho) := \mathbb{E}[\Upsilon_{r;t}].$$

(Initially only the existence of the limit is known; it equals the mean value by L^1 -convergence.)

Observe that, if we get rid of

$$\mathcal{C}(\tau_u F, t; r)$$

on the rhs of (3.11) then, up to $\pm\epsilon$, both the lhs and the rhs of (3.11) would converge to the same limit $c_{r;t}(\rho)$ (in either L^1 or a.s.). We will be able to get rid of $\mathcal{C}(\tau_u F, t; r)$ asymptotically for r large; it will yield that as $r \rightarrow \infty$, we have the limit $c_{r;t}(\rho) \rightarrow c_t(\rho)$ satisfying

$$\frac{\mathcal{N}(F, t; R)}{\text{Vol}(B(R))} \rightarrow c_t(\rho). \tag{3.12}$$

Setting

$$\Psi_\rho(t) := \frac{c_t(\rho)}{c(\rho)}$$

with $c(\rho)$ the Nazarov–Sodin constant (3.2) will ensure that $\Psi_\rho(t)$ obeys (3.3) of Theorem 5. That Ψ_ρ is a distribution function will then follow easily:

$$t \mapsto c_t(\rho)$$

is monotone nondecreasing by (3.12); $\Psi_\rho(\infty) = 1$ since $c_\infty(\rho) = c(\rho)$ is the Nazarov–Sodin constant.

To see that indeed we can get rid of $\mathcal{C}(\tau_u F, t; r)$ we use the same ergodic argument as before, now applied to the map

$$F \mapsto \frac{\mathcal{C}(F, t; r)}{\text{Vol}(B(r))},$$

yielding the L^1 limit

$$\frac{1}{\text{Vol } B(R+r)} \int_{B(R+r)} \frac{\mathcal{C}(\tau_u F, t; r)}{\text{Vol}(B(r))} \rightarrow a_{r;t}(\rho),$$

whence

$$a_{r;t}(\rho) = \mathbb{E} \left[\frac{\mathcal{C}(\tau_u F, t; r)}{\text{Vol}(B(r))} \right] = O\left(\frac{1}{r}\right)$$

by the second part of Lemma 3. Hence (3.11) implies

$$\mathbb{E} \left[\left| \frac{\mathcal{N}(F, t; R)}{\text{Vol}(B(R))} - c_{r;t}(\rho) \right| \right] = O\left(\epsilon + \frac{1}{r}\right);$$

this proves the existence of the limits

$$c_t(\rho) := \lim_{r \rightarrow \infty} c_{r;t}(\rho),$$

and, as it was mentioned above,

$$\Psi_\rho(t) := \frac{c_t(\rho)}{c(\rho)},$$

is the distribution function satisfying the L^1 -convergence (3.3) we were after in Theorem 5. □

Remark 1 This proof also shows a.s. convergence but we will not need that for the rest (i.e. the Riemannian case).

3.4 Proofs of Theorems 6 and 7

Proof of Theorem 6 Since the spectral measure satisfies conditions $(\rho 1) - (\rho 3)$, the existence of the limiting distribution Ψ is guaranteed by Theorem 5. We only have to show that Ψ is strictly increasing.

The proof of Theorem 6 will consist of three steps. First we will construct a deterministic function with nodal domain of given area. After that we will show that the

probability that there is a random function with almost the same nodal domain is positive. Finally we show that this implies that the expected density of such nodal domains is positive.

The first step is essentially the same as the argument in [12, Proposition 5.2]. By condition $(\rho 4^*)$ the interior of the support of ρ is non-empty. Let us assume that $B(\zeta_0, r_0) \subset \text{supp } \rho$. We are going to show that this implies that for every compact Q the functions from $\mathcal{H}(\rho)$ are dense in $C^k(Q)$ for every k .

Let ϕ_n be a sequence of positive C^∞ functions with supports converging to ζ_0 which approximate δ -function at ζ_0 . Functions $\psi_n(\zeta) = \phi_n(\zeta) + \phi_n(-\zeta)$ and all their derivatives belong to $L^2_{\text{sym}}(\rho)$ and hence their Fourier transforms belong to \mathcal{H} . The Fourier transform of ψ_n converges to $e^{2\pi i \langle \zeta_0, x \rangle}$ in any $C^k(Q)$ and its derivatives $\partial^\alpha \psi_n$ converge to $(-2\pi x)^\alpha e^{2\pi i \langle \zeta_0, x \rangle}$, where $x^\alpha = \prod x_i^{\alpha_i}$. This proves that we can approximate any polynomial (times a fixed plane wave). Since the polynomials are dense in $C^k(Q)$, we have that \mathcal{H} is dense in $C^k(Q)$.

In particular, for given $t > 0$ and $\epsilon > 0$ we can take a C^1 function such that its nodal domain around the origin is inside some disc $B(0, r)$ and has area t . Moreover, we can assume that its gradient is bounded away from zero in some neighbourhood of the boundary of this nodal domain. By approximating this function by a function from \mathcal{H} and applying Lemma 10 we obtain a function f_t from \mathcal{H} such that its nodal domain around the origin has area ϵ -close to t and its gradient is bounded away from zero on the nodal line. □

Without loss of generality we may assume that \mathcal{H} norm of f_t is 1 and extend it to an orthonormal basis ϕ_i . In this basis F could be written as

$$F = c_0 f_t + \phi,$$

where c_0 is a standard Gaussian random variable and ϕ is a Gaussian function independent of c_0 (spanned by all additional basis functions ϕ_i). Since f_t is non-degenerate, the probability that

$$\max\{c_0|f|, c_0|\nabla f_t|\} > 1$$

in some (fixed) neighbourhood of the nodal domain containing the origin is strictly positive. Let r be such that the nodal domain of f_t is in $B(0, r)$, then if $C^1(B(0, 2r))$ norm of ϕ is sufficiently small (see Lemma 10 for the precise statement), then the volume of the nodal domain of F is ϵ -close to that of f_t . It is a standard fact (cf. [9]) that the probability that ϕ have small norm is strictly positive. This completes the proof that the area of the nodal domain containing the origin is 2ϵ close to t is strictly positive.

We can pack \mathbb{R}^2 with infinitely many copies of the disc of radius $2r$ such that this covering is locally uniformly finite. By translation invariance, for each disc the probability that there is a nodal domain of area 2ϵ -close to t is strictly positive (independently of the disc). This implies that

$$\frac{\mathcal{N}(F, t + 2\epsilon; R) - \mathcal{N}(F, t - 2\epsilon; R)}{\text{Vol}(B(R))} > c > 0$$

for some positive c and all sufficiently large R . This prove that $\Psi(t+2\epsilon) - \Psi(t-2\epsilon) > c$ which means that Ψ is strictly increasing for $t > t_0$. \square

Proof of Theorem 7 The proof of this theorem is almost exactly the same as the proof of Theorem 6, but the first deterministic step is different. The main difference is that the spectral measure is supported on the sphere, hence its support has empty interior and the space \mathcal{H} is not dense in $C^k(Q)$.

The fact that Ψ vanishes for $t < t_0$ is completely deterministic and follows from the Faber–Krahn inequality. Recall that the random plane wave is a solution to $\Delta f + f = 0$. Let ω be a nodal component of a plane wave F . Since F has a constant sign in ω it must be the first eigenfunction of Laplacian in ω . This implies that the first eigenvalue in ω is 1. By the Faber–Krahn inequality the volume of ω is greater than the volume of the ball for which the main eigenvalue is also equal to 1. It is a well known fact that $|x|^{1-n/2} J_{n/2-1}(|x|)$ is an eigenfunction of the Laplacian in \mathbb{R}^n with eigenvalue 1. Its nodal domain containing the origin is a ball of radius $j_{n/2-1,1}$ and area t_0 . This proves that the volume of every nodal domain is bounded below by t_0 .

To prove the last part of the theorem we first claim that for every $t \geq t_0$ there is a domain Ω_t such that $\text{Vol}(\Omega_t) = t$ and the main eigenvalue of the Laplacian is 1. We define Ω_L to be the union of the cylinder of length L and radius $j_{n/2-1,1}$ with two hemispherical caps at both ends. For $L = 0$ this is the ball of the radius $j_{n/2-1,1}$ and as $L \rightarrow \infty$ it converges to the infinite cylinder. Since Ω_L depends on L continuously, the main eigenvalue λ_L changes continuously from $\lambda_0 = 1$ to $\lambda_\infty = (j_{(n-1)/2-1,1}/j_{n/2-1,1})^2 > 0$. By rescaling Ω_L by $\sqrt{\lambda_L}$ we obtain a family of domains such that their main eigenvalue is 1 and their volumes change continuously from t_0 to infinity. Let us relabel this family and say that Ω_t is the element of the family with volume t .

After that we follow the strategy from [7]. By Whitney approximation theorem we can construct a domain with analytic boundary which is an arbitrary small perturbation of Ω_t . By continuity of the main eigenvalue we can assume that it is arbitrary close to 1. By rescaling the domain we obtain a domain with analytic boundary, main eigenvalue 1 and volume ϵ close to t . Abusing notation we call this domain Ω_t as well.

Let $g_t > 0$ be the main eigenfunction of Laplacian in Ω_t . Since the boundary is analytic, it could be extended to a small neighbourhood B of $\partial\Omega_t$ in such a way that $\inf \|\nabla g_t\| > 0$ in a small tubular neighbourhood of $\partial\Omega_t$. By Lax–Malgrange Theorem for every $b > 0$ there is a function f_t which is a *global* solution to (1.5) such that $\|f_t - g_t\|_{C^1(B)} < b$. If b is sufficiently small, then by Lemma 10 there is a nodal domain of f_t with volume ϵ close to $\text{Vol}(\Omega_t)$. This proves that there is a *deterministic* non-degenerate plane wave with a nodal domain such that its volume is 2ϵ -close to t . The rest of the proof is exactly the same as in Theorem 6. \square

4 Local results

4.1 Local setting and statement of the local result

Recall that $g_{n,\alpha}$ is the “scaling limit” of $f_{n,\alpha;T}$ as $T \rightarrow \infty$, around every point $x \in \mathcal{M}$ in the following sense (cf. Sect. 2.1). We know that for T large, their covariance

functions $K_{\alpha,T}$ and $r_{n,\alpha}$ are, after scaling, asymptotic to each other, uniformly on a large ball $B(R) \subseteq \mathbb{R}^n$. By Lemma 1 they could be coupled so that

$$\|f_{x;T} - \mathfrak{g}\|$$

is arbitrarily small in C^1 -norm on $B(R)$ (or $B(2R)$).

For $u \in \mathcal{M}$, $r > 0$ sufficiently small, $t > 0$ and $g : \mathcal{M} \rightarrow \mathbb{R}$ a smooth (deterministic) function denote $\mathcal{N}(g, t; u, r)$ (resp. $\mathcal{N}^*(g, t; u, r)$) the number of domains $\omega \in \Omega(g)$ of $g^{-1}(0)$ of volume

$$\text{Vol}(\omega) \leq t$$

lying entirely inside (resp. intersecting) the geodesic ball $B_u(r) \subseteq \mathcal{M}$. Theorem 8, to follow immediately, states that, with the coupling as above, unless t is a discontinuity point of the limiting volume distribution (i.e. atoms of the Ψ_ρ obtained from an application of Theorem 5 on $\mathfrak{g}_{n,\alpha}$), after appropriate rescaling, the nodal domain volumes of $f_{\alpha;T}$ are locally point-wise approximated by the ones of $\mathfrak{g}_{n,\alpha}$. Otherwise, if t_0 is an atom of the said Ψ_ρ (its existence is unlikely, by Conjecture 2), then the corresponding probability mass may spread in an interval $(t_0 - \epsilon, t_0 + \epsilon)$ ($\epsilon > 0$ arbitrarily small) beyond our control (see the proof of Proposition 1 below).

Theorem 8 (cf. [13, Theorem 5]) *Let $f(x) = f_{\alpha;T}(x)$ be the random band limited functions (1.7), $\Psi = \Psi_{n,\alpha}$ the distribution function prescribed by Theorem 5 applied on $\mathfrak{g}_{n,\alpha}$. Then for every continuity point $t > 0$ of $\Psi_{n,\alpha}(\cdot)$, $x \in \mathcal{M}$ and $\epsilon > 0$ we have*

$$\lim_{R \rightarrow \infty} \limsup_{T \rightarrow \infty} \mathcal{P} \left\{ \left| \frac{\mathcal{N}\left(f, \frac{t}{T^n}; x, \frac{R}{T}\right)}{c(n, \alpha) \cdot \text{Vol}(B(R))} - \Psi(t) \right| > \epsilon \right\} = 0, \tag{4.1}$$

where $c(n, \alpha) > 0$ is the Nazarov–Sodin constant of $\mathfrak{g}_{n,\alpha}$.

The rest of this section is dedicated to giving a proof of Theorem 8.

4.2 Proof of Theorem 8

We formulate the following proposition that will imply Theorem 8. The proof is postponed till Sect. 4.4 after some preparatory work in Sect. 4.3.

Proposition 1 *Let $x \in \mathcal{M}$, $R > 0$, small numbers $\delta, \xi, \eta > 0$ be given, and $t > 0$. Then, for $R > 0$ sufficiently big, outside of an event of probability at most δ , for all $T > T_0(R, \delta, \xi)$ we have*

$$\begin{aligned} \mathcal{N}(\mathfrak{g}_{n,\alpha}, t - \xi; R - 1) - \eta \cdot R^n &\leq \mathcal{N}(f_{x;T}, t; R) \\ &\leq \mathcal{N}(\mathfrak{g}_{n,\alpha}, t + \xi; R + 1) + \eta \cdot R^n. \end{aligned} \tag{4.2}$$

Proof of Theorem 8 assuming Proposition 1 Let $x \in \mathcal{M}, \epsilon > 0$ be fixed, and $\xi, \eta > 0$ be given small numbers. We observe that by the definition (2.3) of the scaled fields $f_{x;T}$ and (2.4), for T sufficiently big (depending on R) we have

$$\mathcal{N}(f_{x;T}, t - \xi; R) \leq \mathcal{N}\left(f, \frac{t}{T^n}; x, \frac{R}{T}\right) \leq \mathcal{N}(f_{x;T}, t + \xi; R) \tag{4.3}$$

Hence Proposition 1 together with (4.3) holding for T sufficiently big (depending on R) imply that for T sufficiently big, outside an event of arbitrarily small probability,

$$\begin{aligned} \mathcal{N}(\mathfrak{g}_{n,\alpha}, t - 2\xi; R - 1) - \eta \cdot R^n &\leq \mathcal{N}\left(f, \frac{t}{T^n}; x, \frac{R}{T}\right) \\ &\leq \mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1) + \eta \cdot R^n. \end{aligned}$$

Choose

$$\eta < \frac{c(n, \alpha) \cdot \epsilon \cdot \text{Vol } B(1)}{2} \tag{4.4}$$

so that

$$\begin{aligned} &\mathcal{P}\left\{\left|\frac{\mathcal{N}\left(f, \frac{t}{T^n}; x, \frac{R}{T}\right)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t)\right| > \epsilon\right\} \\ &\leq \mathcal{P}\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t)\right| > \epsilon - \frac{\eta}{c(n, \alpha) \cdot \text{Vol } B(1)}\right\} \\ &\quad + \mathcal{P}\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t - 2\xi; R - 1)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t)\right| > \epsilon - \frac{\eta}{c(n, \alpha) \text{Vol } B(1)}\right\}. \end{aligned} \tag{4.5}$$

Now, bearing in mind our choice (4.4), we have that $\epsilon - \frac{\eta}{c(n,\alpha) \cdot \text{Vol } B(1)} > \frac{\epsilon}{2}$, so that

$$\begin{aligned} &\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t)\right| > \epsilon - \frac{\eta}{c(n, \alpha) \cdot \text{Vol } B(1)}\right\} \subseteq \\ &\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t + 2\xi)\right| > \frac{\epsilon}{2} - (\Psi(t + 2\xi) - \Psi(t))\right\}, \end{aligned}$$

and thus

$$\begin{aligned} &\mathcal{P}\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1)}{c(n, \alpha) \cdot \text{Vol } B(R)} - \Psi(t)\right| > \epsilon - \frac{\eta}{c(n, \alpha) \cdot \text{Vol } B(1)}\right\} \\ &\leq \mathcal{P}\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t + 2\xi; R + 1)}{\text{Vol } B(R)} - \Psi(t + 2\xi)\right| > \frac{\epsilon}{2} - (\Psi(t + 2\xi) - \Psi(t))\right\} \rightarrow 0 \end{aligned} \tag{4.6}$$

as $R \rightarrow \infty$ by Theorem 5, and the continuity of $\Psi(\cdot)$ at t . Similarly, as $R \rightarrow \infty$,

$$\mathcal{P}\left\{\left|\frac{\mathcal{N}(\mathfrak{g}_{n,\alpha}, t - 2\xi; R - 1)}{c(n, \alpha) \text{Vol } B(R)} - \Psi(t)\right| > \epsilon - \frac{\eta}{c(n, \alpha) \cdot \text{Vol } B(1)}\right\} \rightarrow 0. \tag{4.7}$$

The above proves that for each $R > 0$ sufficiently large there exists $T_0 = T_0(R)$, such that (4.5) holds for $T > T_0$; the latter could be made arbitrarily small by (4.6) and (4.7). This is precisely the statement of Theorem 8. \square

4.3 Some preparatory results towards the proof of Proposition 1

In course of the proof of Proposition 1 we will need to exclude some exceptional events. Let $\delta > 0$ be a small parameter that will control the probabilities of the discarded events, $b, \beta > 0$ be small parameters that will control the quality of the various approximations, $\eta > 0$ will control the number of discarded domains, and $M, Q > 0$ large parameters. Given R and T big we define

$$\begin{aligned} \Delta_1 &= \Delta_1(R, T; b) = \{\|f_{x;T} - g_{n,\alpha}\|_{C^1(\overline{B}(2R))} \geq b\}, \\ \Delta_2 &= \Delta_2(R, T; M) = \left\{ \|f_{x;T}\|_{C^2(\overline{B}(2R))} \geq M \right\}, \\ \Delta_3 &= \Delta_3(R; M) = \left\{ \|g_{n,\alpha}\|_{C^2(\overline{B}(2R))} \geq M \right\} \end{aligned}$$

and the “unstable event”

$$\Delta_4(R; \beta) = \left\{ \min_{u \in \overline{B}(2R)} \max\{|g_{n,\alpha}(u)|, |\nabla g_{n,\alpha}(u)|\} \leq 2\beta \right\}.$$

Moreover let Δ_5 be the event

$$\begin{aligned} \Delta_5(R; \eta, Q) &= \left\{ \left| \left\{ \text{components } \omega \in \Omega(g_{n,\alpha}) : \omega \subseteq B(2R), \text{Vol}_{n-1}(\partial\omega) > Q \right\} \right| > \eta R^n \right\} \end{aligned}$$

that there is a significant number of nodal domains of $g_{n,\alpha}$ entirely lying in $B(2R)$ whose boundary volume is at least Q . The boundary of such a domain is comprised of a number of nodal components; below we will argue that, with high probability, each of the boundary components is of bounded volume and their total number is bounded (see the proof of Lemma 9).

The following is a simple corollary of Lemma 1.

Lemma 6 *For a given sufficiently big $R > 0$ and small $b, \delta > 0$ there exists $T_0 = T_0(R, b, \delta)$ so that for $T > T_0$ the probability of Δ_1*

$$\mathcal{P}(\Delta_1(R, T; b)) < \delta. \tag{4.8}$$

The following lemma yields a bound on the probabilities $\mathcal{P}(\Delta_2)$ and $\mathcal{P}(\Delta_3)$.

Lemma 7 1. *For every $R, \delta > 0$ there exists $M = M(R, \delta) > 0$ so that*

$$\mathcal{P}(\Delta_3(R; M)) < \delta. \tag{4.9}$$

2. For $R, \delta > 0$ there exist $M(R, \delta) > 0$ and $T_0 = T_0(R)$ so that

$$\mathcal{P}(\Delta_2(R, T; M)) < \delta \tag{4.10}$$

for all $T > T_0$.

Proof For (4.9) we may choose M to be

$$M = \delta^{-1} \mathbb{E}[\|\mathfrak{g}_{n,\alpha}\|_{C^2(\overline{B(2R)})}],$$

and the latter expectation is finite by [1, Theorem 2.1.1]. The estimate (4.9) with this value of M follows from Chebyshev’s inequality.

In order to establish (4.10) we observe that by “Sudakov-Fernique comparison inequality” [1, Theorem 2.2.3] and (2.1) applied to both $K_{x,T}$ and its derivatives, for all M_1 such that

$$\mathbb{E}[\|\mathfrak{g}_{n,\alpha}\|_{C^2(\overline{B(2R)})}] < M_1$$

there exists $T_0 = T_0(R, M_1)$ such that for all $T > T_0$

$$\mathbb{E} \left[\|f_{x,T}\|_{C^2(\overline{B(2R)})} \right] < M_1.$$

Hence we can set $M = \delta^{-1} M_1$ and (4.10) will follow from Chebyshev’s inequality as before.

The event Δ_4 means that the nodal set of $\mathfrak{g}_{n,\alpha}$ is relatively unstable. For small β the probability of this event can be made arbitrary small. This is Lemma 7 from [11]:

Lemma 8 (cf. [11]) *For $R > 0$ and $\delta > 0$ there exist $\beta > 0$ such that*

$$\mathcal{P}(\Delta_4(R, \beta)) < \delta. \tag{4.11}$$

Finally, for Δ_5 we have the following bound:

Lemma 9 *For every $\eta > 0$ and $\delta > 0$ there exist $Q > 0$ and $R_0 = R_0(\eta, \delta)$ such that for all $R > R_0$ the probability of Δ_5 is*

$$\mathcal{P}(\Delta_5(R; \eta, Q)) < \delta. \tag{4.12}$$

Proof In what follows we argue that most of the domains have a bounded number of components lying in the boundary, and with high probability, each has bounded $n - 1$ -volume. First, using Nazarov–Sodin’s (3.2), there exists a number A sufficiently big so that the probability that the total number of nodal domains of $\mathfrak{g}_{n,\alpha}$ entirely lying inside $B(2R)$ is bigger than $A \cdot R^n$ is

$$\mathcal{P}(\Delta_6(A)) := \mathcal{P} \{ \mathcal{N}(\mathfrak{g}_{n,\alpha}; 2R) > A \cdot R^n \} < \delta/2, \tag{4.13}$$

so we may exclude this unlikely event $\Delta_6 = \Delta_6(A)$.

Now we are going to show that the number of boundary components of most of the nodal domains of $g_{n,\alpha}$ lying in $B(2R)$ is bounded; outside of Δ_6 there are less than $A \cdot R^n$. To this end we use the *nesting graph* introduced in [12]. Let $G = (V, E)$ be the graph with the set of vertexes V being the collection of all nodal domains

$$V = \{\omega \in \Omega(g_{n,\alpha}) : \omega \subseteq B(2R)\}$$

of $g_{n,\alpha}$ lying entirely in $B(2R)$, and an edge $e = e(\gamma) \in E$ connects between two domains in V , if they have a common boundary component γ . The graph G is a subgraph of the nesting *tree*, same graph with no restriction of the domains to be contained in $B(2R)$; though G may fail to be a tree, it has no cycles [12, Section 2.4]; a degree $d(\omega)$ of a vertex $\omega \in V$ in V precisely equals to the number of nodal components lying the boundary of ω we are to bound. We then have

$$\sum_{\omega \in V} d(\omega) = 2|E| \leq 2(|V| - 1) < 2\mathcal{N}(g_{n,\alpha}; 2R),$$

equivalently

$$\frac{1}{\mathcal{N}(g_{n,\alpha}; 2R)} \sum_{\omega \in V} d(\omega) < 2.$$

Hence, by Chebyshev’s inequality, outside $\Delta_6(A)$ as in (4.13), the number of those $\omega \in V$ with $d(\omega) > L$ is at most

$$|\{\omega \in V : d(\omega) > L\}| \leq \frac{2}{L} \cdot \mathcal{N}(g; 2R) \leq \frac{2}{L} \cdot AR^n, \tag{4.14}$$

and below we will choose L sufficiently big so that

$$\frac{2A}{L} < \frac{\eta}{2}. \tag{4.15}$$

Next we show that, with high probability, the $(n - 1)$ -volume of most of the components is bounded. Let $\mathcal{Z}_{g_{n,\alpha}}(2R)$ be the nodal volume of $g_{n,\alpha}$ inside $B(2R)$

$$\mathcal{Z}_{g_{n,\alpha}}(2R) = \text{Vol}_{n-1}(g_{n,\alpha}^{-1}(0) \cap B(2R)) = \sum_{e(\gamma) \in E} \text{Vol}_{n-1}(e). \tag{4.16}$$

Then, by a standard application of Lemma 2 and the stationarity of $g_{n,\alpha}$, the expectation of the nodal volume is given by

$$\mathbb{E}[\mathcal{Z}_{g_{n,\alpha}}(2R)] = c_0 \cdot R^n, \tag{4.17}$$

where $c_0 = c_0(g_{n,\alpha}) > 0$ is a positive constant, that could be evaluated explicitly in terms of n and α . Hence, by (4.16) and (4.17), and Chebyshev’s inequality, outside of

event of probability $\delta/2$ the number of those components with large $(n - 1)$ -volume is bounded:

$$\mathcal{P}(\Delta_7(Q/L)) := \mathcal{P} \left\{ \left| \left\{ \gamma \in E : \text{Vol}_{n-1}(\gamma) > \frac{Q}{L} \right\} \right| > \frac{\eta}{4} \cdot R^n \right\} < \frac{\delta}{2}, \tag{4.18}$$

provided that Q/L is sufficiently big. Since each component is lying in the boundary of at most two domains $\omega \in V$, it follows that outside of Δ_7 for all but at most $\frac{\eta}{2} \cdot R^n$ domains $\omega \in V$, for all components $\gamma \in E$ lying in the boundary of ω we have

$$\text{Vol}_{n-1}(\gamma) \leq \frac{Q}{L}. \tag{4.19}$$

Now, given $\delta > 0$ we choose $A > 0$ sufficiently big so that (4.13) is satisfied. This forces a choice of L via (4.15), so that the r.h.s. of (4.14) is $< \frac{\eta}{2} \cdot R^n$, and then we take Q sufficiently big so that (4.18) is satisfied and

$$\mathcal{P}(\Delta_6 \cup \Delta_7) \leq \mathcal{P}(\Delta_6) + \mathcal{P}(\Delta_7) < \delta.$$

From the above, outside of $\Delta_6 \cup \Delta_7$ for all but $\eta \cdot R^n$ nodal domains $\omega \in V$ we have that the number of boundary components of ω is $< L$, and the $(n - 1)$ -volume of each one of them is bounded by $\frac{Q}{L}$, and hence the $(n - 1)$ -volume of $\partial\omega$ is $< Q$ as claimed by this lemma. □

4.4 Proof of Proposition 1

To prove Proposition 1 we need the following lemma which is the uniform version of an obvious statement that the volume of a nodal domain depends continuously on the function as long as 0 is not a critical value.

Lemma 10 (cf. [13, Lemmas 6–7]) *Let $R > 0$ and g and h be two (deterministic) C^2 -smooth functions on $B(2R) \subset \mathbb{R}^{n+1}$. We assume that:*

1. *For some $\beta > 0$ we have*

$$\min_{\overline{B(2R)}} \max\{|g|, |\nabla g|\} > \beta.$$

2. *For some $M > 0$ the C^2 -norms of both g and h are bounded*

$$\|g\|_{C^2(B(2R))}, \|h\|_{C^2(B(2R))} < M.$$

3. *We have*

$$\|g - h\|_{C^1(B(R))} < b$$

for some $b > 0$.

Then if b is sufficiently small (depending on β and M only) there exists an injective map $\gamma \mapsto \gamma^h$ between connected components $\gamma \subseteq B(R - 1)$ of $g^{-1}(0)$ and connected components $\gamma^h \subseteq B(R)$ of $h^{-1}(0)$ with the following properties:

1. For every γ as above the components γ and γ^h are “uniformly close”: there exists a smooth bijective map $\psi_\gamma : \gamma \rightarrow \gamma^h$ so that for all $x \in \gamma$ we have

$$\|\psi_\gamma(x) - x\|_\infty = O_{\beta, M}(b) \tag{4.20}$$

with constants involved in the ‘ O ’-notation depending on β and M only.

2. Let $\mathcal{G}_{\gamma, \gamma^h}$ be the region enclosed between γ and γ^h . Its (signed or not) volume satisfies

$$|\text{Vol}_n(\mathcal{G}_{\gamma, \gamma^h})| = \text{Vol}_{n-1}(\gamma) \cdot O_{\beta, M}(b). \tag{4.21}$$

The proof of Lemma 10 is postponed until immediately after the proof of Proposition 1.

Proof of Proposition 1 assuming Lemma 10 We are going to show that the small exceptional event is

$$\Delta = \bigcup_{i=1}^5 \Delta_i.$$

Outside Δ we have

$$\begin{aligned} \|f_{x;T} - \mathfrak{g}_{n,\alpha}\|_{C^1(\overline{B(2R)})} &< b, \\ \|f_{x;T}\|_{C^2(\overline{B(2R)})}, \|\mathfrak{g}_{n,\alpha}\|_{C^2(\overline{B(2R)})} &< M, \end{aligned} \tag{4.22}$$

and assuming $b < \beta$ we also have

$$\min_{\overline{B(2R)}} \max\{|f_{x;T}|, |\nabla f_{x;T}|\} > \beta; \quad \min_{\overline{B(2R)}} \max\{|\mathfrak{g}_{n,\alpha}|, |\nabla \mathfrak{g}_{n,\alpha}|\} > \beta. \tag{4.23}$$

Hence the conditions of Lemma 10 are satisfied with $h = f_{x;T}$, $g = \mathfrak{g}_{n,\alpha}$ (or the other way round) and b, β as above.

By Lemma 10 for each nodal domain of $f_{x;T}$ lying in $B(R)$ there is a unique nodal domain of $\mathfrak{g}_{n,\alpha}$ which is $O(b)$ close, hence lying in $B(R + 1)$. Reversing the roles of f and \mathfrak{g} we see that the same holds for the nodal domains of \mathfrak{g} . We are going to prove the first inequality of (4.2), the proof of the second one is identical.

Let us consider the nodal domains of $\mathfrak{g}_{n,\alpha}$ that are inside $B(R - 1)$ and of area at most $t - \xi$. Their total number is $\mathcal{N}(\mathfrak{g}_{n,\alpha}, t - \xi, R - 1)$. For each of these nodal domains \mathcal{D} there is a unique nodal domain \mathcal{D}' of $f_{x;T}$ which is $O(b)$ close and lies in $B(R)$. By the estimate (4.21) of Lemma 10 we have

$$|\text{Vol}(\mathcal{D}) - \text{Vol}(\mathcal{D}')| \leq O_{\beta, M}(b) \text{Vol}_{n-1}(\partial \mathcal{D}).$$

Since we excluded Δ_5 , at most ηR^n of nodal domains \mathcal{D} have boundary volume exceeding Q . For all other domains $\text{Vol}(\mathcal{D}') < t$ provided that b is so small that have $O(b)Q < \xi$. Hence, their number is bounded by $\mathcal{N}(f_{x;T}, t, R)$. This means that

$$\mathcal{N}(g_{n,\alpha}, t - \xi, R - 1) \leq \eta R^n + \mathcal{N}(f_{x;T}, t, R)$$

i.e. the first inequality of (4.2).

Finally we have to show that the probability of Δ could be made arbitrary small, using Lemmas 6–9. The argument is straightforward but the order in which we have to choose all constants is a bit fiddly.

Let $t, \xi, \delta, \eta > 0$ be given. First we use Lemma 9 to choose R and Q sufficiently large so that $\mathcal{P}(\Delta_5) < \delta$. From now on this value of R is fixed. By Lemma 8 there is $\beta > 0$ such that $\mathcal{P}(\Delta_4) < \delta$. By Lemma 7 we can choose M and T_0 large enough so that probabilities of Δ_2 and Δ_3 are bounded by δ . Fix some $0 < b < \beta$ sufficiently small, applying Lemma 6 and (possibly) increasing T_0 we make sure that $\mathcal{P}(\Delta_1) < \delta$ as well.

All in all, for the above choice of all constants and for all $T > T_0$ we have that the probability of the exceptional event $\Delta := \cup_{i=1}^5 \Delta_i$ is bounded by 5δ . Replacing δ by $\delta/5$ we complete the proof of the proposition. \square

Proof of Lemma 10 Let $\gamma \subseteq B(R)$ be a connected component of $g^{-1}(0)$ and $x \in \gamma$ be any point on γ . By the Implicit Function Theorem we can introduce local coordinates (z, y) such that the surface near x can be parameterized as a graph of a smooth function $p(z)$. In other words γ around x is given by $\phi(z) = (z, p(z))$ where p is a smooth function on an open domain $U \subset \mathbb{R}^{n-1}$. Let

$$N(x) = N(z) = \nabla g(x) / |\nabla g(x)|$$

be a unit normal vector to γ at x . Since the second derivatives of g are uniformly bounded, there is a number $r_0 = r_0(\beta, M) > 0$ depending on β and M only so that

$$N(x) \cdot \nabla g(x) > \beta/2$$

on an r_0 -neighbourhood of x in \mathbb{R}^n .

Now fix $x \in \gamma$ and consider

$$\zeta(r) = h(x + rN(x)) = g(x + rN(x)) + f(x + rN(x)),$$

where $f = h - g$. Obviously

$$|\zeta(0)| = |h(x)| < b,$$

and

$$\zeta'(r) > \beta/2 - b$$

for all $|r| < r_0$. For sufficiently small b we then have

$$\zeta' > \beta/4 \tag{4.24}$$

for all $|r| < r_0$. This means that if $4b/\beta < r_0$, which is true for sufficiently small b , there is a unique r with $|r| < 4b/\beta < r_0$ such that $\zeta(r) = 0$. We denote it by $r(x)$ or $r(z)$ with $x = \phi(z)$. It is important to notice that r is uniformly bounded in terms of β and M and for fixed β and M it is $O(b)$.

All in all, for every b sufficiently small (depending on β and M only) the map

$$x \mapsto x + r(x)N(x)$$

maps γ onto γ^h , moreover,

$$r(x) = O_{\beta, M}(b).$$

Since $\mathcal{G}_{\gamma, \gamma^h}$ is the domain between two surfaces that are $O_{\beta, M}(b)$ close, the volume of this domain (signed or unsigned) is bounded by $\text{Vol}_{n-1}(\gamma)O_{\beta, M}(b)$. This completes the proof of Lemma 10. □

5 Global results

5.1 Proofs of the main results: Theorems 1, 2, 3, and 4

Notation 9 (*Global notation*)

1. For $y \in \mathbb{R}$ denote

$$|y|_+ = \max\{0, y\},$$

and

$$|y|_- = \max\{0, -y\},$$

so that

$$|\cdot| = |\cdot|_+ + |\cdot|_-.$$

2. Let $\mathcal{N}(f; t)$ be the total number of nodal domains $\omega \in \Omega(f)$ of f of volume

$$\text{Vol}_n(\omega) < t.$$

Proposition 2 *Let $\{f = f_{\alpha, T}\}_{T>0}$ be the random fields (1.7), $c(n, \alpha)$ the Nazarov–Sodin constant of $\mathfrak{g}_{n, \alpha}$, and $t > 0$ a continuity point of*

$$\Psi(\cdot) = \Psi_{n, \alpha}(\cdot),$$

as in the formulation of Theorem 8. Then the following holds:

$$\mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}_n(\mathcal{M}) \cdot T^n} - \Psi(t) \right|_{\pm} \right] \rightarrow 0, \tag{5.1}$$

that is, (5.1) is claimed for both $|\cdot|_+$ and $|\cdot|_-$, as in Notation 5.1.

Proposition 2 will be proved in Sect. 5.2.2 for $|\cdot|_+$ only, with the proof for $|\cdot|_-$ following along similar, but easier, lines, with no need to excise the very small and very long domains (see Sect. 5.2.1).

Proof of Theorem 2 assuming Proposition 2 Combining both estimates (5.1) of Proposition 2 (i.e. $|\cdot|_+$ and $|\cdot|_-$) implies that for every $t > 0$ continuity point of $\Psi(\cdot)$ we have

$$\mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n} - \Psi(t) \right| \right] \rightarrow 0, \tag{5.2}$$

which is not quite the same as the statement (1.9) of Theorem 2 as the denominator needs to be replaced by $\mathcal{N}(f)$ rather than

$$c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n.$$

To this end we use the triangle inequality to write

$$\begin{aligned} \mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{\mathcal{N}(f)} - \Psi(t) \right| \right] &\leq \mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{\mathcal{N}(f)} - \frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n} \right| \right] \\ &\quad + \mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n} - \Psi(t) \right| \right]. \end{aligned} \tag{5.3}$$

Since the second summand of the r.h.s. of (5.3) vanishes by (5.2), we only need to take care of the first one. We have

$$\begin{aligned} &\mathbb{E} \left[\left| \frac{\mathcal{N}(f, t/T^n)}{\mathcal{N}(f)} - \frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n} \right| \right] \\ &= \mathbb{E} \left[\frac{\mathcal{N}(f, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot \mathcal{N}(f)} \cdot \left| \frac{\mathcal{N}(f)}{T^n} - c(n, \alpha) \text{Vol}(\mathcal{M}) \right| \right] \rightarrow 0, \end{aligned} \tag{5.4}$$

by (1.8), as it is obvious that

$$\frac{\mathcal{N}_\Omega(f, t/T^n)}{\mathcal{N}(f)} \leq 1.$$

We finally substitute (5.2) and (5.4) into (5.3) to yield the statement (1.9) of Theorem 2. □

Proof of Theorem 1 The first assertion of Theorem 1 is a particular case of the statement of Theorem 2 with $\mathcal{M} = \mathcal{S}^2$ the round 2-sphere, and $\alpha = 1$. The second assertion

of Theorem 1 is the content of Theorem 4 (which, in this case, follows directly from Theorem 7). □

Proofs of Theorems 3 and 4 By Theorem 2 the distribution function is universal and from (2.2) we know the spectral measures. In the case $\alpha < 1$ it satisfies the axioms $(\rho 1) - (\rho 3)$ and $(\rho 4^*)$ and Theorem 3 follows directly from Theorem 6. In the case $\alpha = 1$ the limiting field is the random plane wave and Theorem 4 follows from Theorem 7. □

5.2 Proof of Proposition 2

5.2.1 Excising the very small and very long domains

Definition 3 Let $\xi, D > 0$ be parameters.

1. A domain $\omega \in \Omega(f)$ of $f = f_{\alpha;T}$ is called ‘ ξ -small’ if its n -dimensional Riemannian volume in \mathcal{M} is

$$\text{Vol}_{\mathcal{M}}(\omega) < \xi T^{-n}.$$

- Let $\mathcal{N}_{\xi-sm}(f)$ be the total number of ξ -small domains (components) of f in \mathcal{M} .
2. For $D > 0$, a nodal domain $\omega \in \Omega(f)$ is D -long if its diameter is $> D/T$. Let $\mathcal{N}_{D-long}(f)$ be the total number of the D -long domains of f .
 3. Given parameters $D, \xi > 0$ a nodal domain $\omega \in \Omega(f)$ is (D, ξ) -normal (or simply normal), if it is not ξ -small nor D -long. For $t > 0$ let $\mathcal{N}_{norm}(f, t)$ be the total number of (ξ, D) -normal domains of f of volume $< t$, and for $x \in \mathcal{M}, r > 0$ let $\mathcal{N}_{norm}(f_L, t; x, r)$ (resp. $\mathcal{N}_{norm}^*(f_L, t; x, r)$) be the number of those contained in $B_x(r)$ (resp. intersecting $B_x(r)$). Finally, if t is omitted, then it is assumed to be infinite $t = \infty$, i.e. no restriction on the domain volume is imposed.

By the definition of normal domains, for every $t \in \mathbb{R} \cup \{\infty\}$ we have

$$\mathcal{N}_{norm}^*(f, t; x, r) \leq \mathcal{N}_{norm}\left(f, t; x, r + \frac{D}{T}\right) \tag{5.5}$$

(as we discarded the very long ovals), and

$$\mathcal{N}_{norm}(f; x, r) \leq \xi^{-1} T^n \text{Vol}_{\mathcal{M}}(B_x(r)), \tag{5.6}$$

by the natural volume estimate.

Lemma 11 (cf. [13, Lemma 9], see also [12] for a more detailed proof) *There exist constants $c, C > 0$ so that we have the following estimate on the number of ξ -small components:*

$$\limsup_{T \rightarrow \infty} \frac{\mathbb{E}[\mathcal{N}_{\xi-sm}(f_{\alpha;T})]}{T^n} \leq C \cdot \xi^c.$$

Lemma 12 (cf. [13, Lemma 8]) *There exists a constant $C > 0$ such that the following bound holds for the number of D -long components:*

$$\limsup_{T \rightarrow \infty} \frac{\mathbb{E}[\mathcal{N}_{D\text{-long}}(f_{\alpha;T})]}{T^n} \leq C \cdot \frac{1}{D}.$$

5.2.2 Proof of Proposition 2

Proposition 3 *Let $\xi, D > 0$ be fixed, and*

$$\Psi = \Psi_{n,\alpha}.$$

Then for every point of continuity $t > 0$ of $\Psi(\cdot)$ the following holds:

$$\mathbb{E} \left[\left| \frac{\mathcal{N}_{norm}(f_{\alpha;T}, t/T^n)}{c(n, \alpha) \text{Vol}(\mathcal{M}) \cdot T^n} - \Psi(t) \right|_+ \right] \rightarrow 0. \tag{5.7}$$

The proof of Proposition 3 will be given in Sect. 5.3.

Proof of Proposition 2 assuming Proposition 3 The estimate (5.1) for $|\cdot|_+$ follows directly from (5.7), lemmas 12–11, and the triangle inequality for $|\cdot|_+$. The proof of (5.1) for $|\cdot|_-$ follows along the same lines, except that in this case we do not need to excise the very small and very long domains (which makes the proof somewhat simpler). \square

5.3 Proof of Proposition 3

We need to formulate a couple of auxiliary lemmas.

Lemma 13 (cf. [13] Lemma 1, and Lemma 5 in the scale-invariant case) *Given $\epsilon > 0$, there exists $\eta > 0$ such that for every $r < \eta, t > 0$*

$$\begin{aligned} (1 - \epsilon) \int_{\mathcal{M}} \frac{\mathcal{N}_{norm}(f_{\alpha;T}, t; x, r)}{\text{Vol}(B(r))} dx &\leq \mathcal{N}_{norm}(f_{\alpha;T}, t) \\ &\leq (1 + \epsilon) \int_{\mathcal{M}} \frac{\mathcal{N}_{norm}^*(f_{\alpha;T}, t; x, r)}{\text{Vol}(B(r))} dx \end{aligned} \tag{5.8}$$

The proof of Lemma 13 is very similar to the one of Lemma 5 and is omitted here.

Proof of Proposition 3 For convenience in course of this proof we will assume that \mathcal{M} is unit volume

$$\text{Vol}(\mathcal{M}) = 1, \tag{5.9}$$

and recall that $f = f_{\alpha;T}$ are the band-limited random fields (1.7), and that we use the shorthand

$$\Psi(t) = \Psi_{n,\alpha}(t)$$

as in the formulation of Proposition 3. Let $\epsilon > 0$ be a small number. To bound the l.h.s. of (5.7) we let $R > 0$ be sufficiently big, so that both $R/T < \eta$ as in Lemma 13, and D/R is sufficiently small, so that the following holds

$$\begin{aligned} \left| \frac{\text{Vol}_{\mathcal{M}}(B_x(R/T))}{\text{Vol}(B(R/T))} - 1 \right| &< \epsilon, \\ \frac{\text{Vol}(B(R + D))}{\text{Vol}(B(R))} &< 1 + \epsilon \end{aligned} \tag{5.10}$$

uniformly for $x \in \mathcal{M}$.

Now apply Lemma 13 with $r = \frac{R}{T}$ and t replaced by $\frac{t}{T^n}$; by the triangle inequality for $|\cdot|_+$ we have

$$\begin{aligned} &\mathbb{E} \left[\left| \frac{\mathcal{N}_{norm}(f_{\alpha}; T, t/T^n)}{c(n, \alpha) \cdot T^n} - \Psi(t) \right|_+ \right] \\ &\leq \mathbb{E} \left[\int_{\mathcal{M}} \left| (1 + 2\epsilon) \frac{\mathcal{N}_{norm}^*(f_{\alpha}; T, t/T^n; x, R/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ dx \right] \\ &\leq \mathbb{E} \left[\int_{\mathcal{M}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha}; T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ dx \right] \\ &\quad + O \left(\epsilon \cdot \int_{\mathcal{M}} \frac{\mathbb{E}[\mathcal{N}_{norm}(f_{\alpha}; T; x, (R + D)/T)]}{\text{Vol } B(R + D)} dx \right), \end{aligned} \tag{5.11}$$

by (5.5) and (5.10). Observe that the integrand

$$\frac{\mathbb{E}[\mathcal{N}_{norm}(f_{\alpha}; T; x, (R + D)/T)]}{\text{Vol } B(R + D)}$$

is uniformly bounded by Lemma 4. Hence (5.11) is

$$\begin{aligned} &\mathbb{E} \left[\left| \frac{\mathcal{N}_{norm}(f_{\alpha}; T, t/T^n)}{c(n, \alpha) \cdot T^n} - \Psi(t) \right|_+ \right] \\ &\leq \mathbb{E} \left[\int_{\mathcal{M}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha}; T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ dx \right] + O(\epsilon). \end{aligned}$$

It is then sufficient to prove that

$$\begin{aligned} & \mathbb{E} \left[\int_{\mathcal{M}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha;T}, t/T^n; x, (R+D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R+D)} - \Psi(t) \right|_+ dx \right] \\ &= \int_{\Delta} \int_{\mathcal{M}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha;T}, t/T^n; x, (R+D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R+D)} - \Psi(t) \right|_+ dx d\mathcal{P}(\omega) \rightarrow 0, \end{aligned} \tag{5.12}$$

where Δ is the underlying probability space, and \mathcal{P} is the probability measure on Δ . Now consider the event

$$\Delta_{T,t;x,R} = \left\{ \left| \frac{\mathcal{N}(f_{\alpha;T}, t/T^n; x, R/T)}{c(n, \alpha) \cdot \text{Vol } B(R+D)} - \Psi(t) \right| > \epsilon \right\}.$$

Then, recalling the assumption of Proposition 3 on t (i.e. that $\Psi(\cdot)$ is continuous at t), Theorem 8 implies that for all $x \in \mathcal{M}$

$$\lim_{R \rightarrow \infty} \limsup_{T \rightarrow \infty} \mathcal{P}(\Delta_{T,t;x,R}) = 0. \tag{5.13}$$

We claim that the above implies that there exists a sequence $\{R_j\}_{j \rightarrow \infty}$ of values $R = R_j \rightarrow \infty$ so that the limit (5.13) is almost uniform w.r.t. $x \in X$, that is, for every $\rho > 0$ there exists $\mathcal{M}_\rho \subseteq \mathcal{M}$ of volume

$$\text{Vol } \mathcal{M}_\rho > 1 - \rho,$$

such that

$$\lim_{R \rightarrow \infty} \overline{\lim}_{T \rightarrow \infty} \sup_{x \in \mathcal{M}_\rho} \mathcal{P}(\Delta_{T,t;x,R+D}) = 0. \tag{5.14}$$

To see (5.14) we first apply an Egorov-type theorem on the limit in (5.13) w.r.t. $R \rightarrow \infty$: working with the sets

$$E_{n,k} = \bigcup_{R > n \text{ integer}} \left\{ x \in \mathcal{M} : \mathcal{P}(\Delta_{T,t;x,R+D}) > \frac{1}{k} \text{ for } T = T_j \rightarrow \infty \right\}$$

yields that for some M_ρ with

$$\text{Vol}(M_\rho) > 1 - \frac{\rho}{2}$$

we have

$$\lim_{R \rightarrow \infty} \sup_{x \in \mathcal{M}_\rho} \overline{\lim}_{T \rightarrow \infty} \mathcal{P}(\Delta_{T,t;x,R+D}) = 0;$$

this is not quite the same as the claimed result (5.14), as the order of $\sup_{x \in \mathcal{M}_\rho}$ and the \limsup w.r.t. $T \rightarrow \infty$ is wrong. We use an Egorov-type argument once again,

w.r.t. the limit $\overline{\lim}_{T \rightarrow \infty}$ to mollify this. Fix an integer $r > 0$, and let $R = R(r) > 0$ sufficiently big so that

$$\sup_{x \in \mathcal{M}_\rho} \overline{\lim}_{T \rightarrow \infty} \mathcal{P}(\Delta_{T,t;x,R+D}) < \frac{1}{r}. \tag{5.15}$$

Define the monotone decreasing sequence of sets

$$F_m = \bigcup_{T > m} \left\{ x \in \mathcal{M}_\rho : \mathcal{P}(\Delta_{T,t;x,R+D}) > \frac{2}{r} \right\}.$$

Since, by (5.15),

$$\bigcap_{m \geq 1} F_m = \emptyset,$$

we may find $m = m(r)$ sufficiently big so that $\text{Vol}(F_{m(r)}) < \frac{\rho}{2^{r+1}}$. Therefore the claimed result (5.14) holds on

$$\mathcal{M}_\rho \setminus \bigcup_{r \geq 1} F_{m(r)},$$

i.e. further excising the set

$$\bigcup_{r \geq 1} F_{m(r)}$$

of volume $< \frac{\rho}{2}$ from \mathcal{M}_ρ .

We then write the integral (5.12) as

$$\begin{aligned} & \int_{\Delta} \int_{\mathcal{M}} \left| \frac{\mathcal{N}_{norm}(f_\alpha; T, t/T^n; x, (R+D)/T)}{c(n, \alpha) \cdot \text{Vol } B((R+D)/T)} - \Psi(t) \right|_+ dx d\mathcal{P}(\omega) \\ &= \int_{\mathcal{M} \cap \Delta_{T,t;x,R+D}} + \int_{\mathcal{M} \setminus \Delta_{T,t;x,R+D}}. \end{aligned} \tag{5.16}$$

First, on $\Delta \setminus \Delta_{T,t;x,R+D}$, the integrand of (5.16) is

$$\begin{aligned} & \left| \frac{\mathcal{N}_{norm}(f_\alpha; T, t/T^n; x, (R+D)/T)}{c(n, \alpha) \cdot \text{Vol } B((R+D)/T)} - \Psi(t) \right|_+ \\ & \leq \left| \frac{\mathcal{N}(f_\alpha; T, t/T^n; x, (R+D)/T)}{c(n, \alpha) \cdot \text{Vol } B((R+D)/T)} - \Psi(t) \right| \leq \epsilon, \end{aligned}$$

and hence the contribution of this range is

$$\begin{aligned} & \int_{\mathcal{M} \setminus \Delta_{T,t;x,R+D}} \int_{\Delta_{T,t;x,R+D}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B((R + D)/T)} - \Psi(t) \right|_+ d\mathcal{P}(\omega) dx \\ & \leq \int_{\mathcal{M} \setminus \Delta_{T,t;x,R+D}} \int_{\Delta_{T,t;x,R+D}} \epsilon d\mathcal{P}(\omega) dx \leq \epsilon. \end{aligned} \tag{5.17}$$

On $\Delta_{T,t;x,R+D}$ we use the volume estimate (5.6) yielding uniformly on $x \in \mathcal{M}$

$$\begin{aligned} & \int_{\Delta_{T,t;x,R+D}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ d\mathcal{P}(\omega) \\ & \leq \int_{\Delta_{T,t;x,R+D}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} \right|_+ d\mathcal{P}(\omega) \\ & \leq \int_{\Delta_{T,t;x,R+D}} \xi^{-1} T^n \frac{\text{Vol}_{\mathcal{M}}(B_x((R + D)/T))}{c(n, \alpha) \cdot \text{Vol } B(R + D)} d\mathcal{P}(\omega) \\ & \leq (1 + \epsilon) \xi^{-1} \mathcal{P}(\Delta_{T,t;x,R+D}) \end{aligned} \tag{5.18}$$

(where, by the obvious properties of $|\cdot|_+$, to obtain the first inequality we omitted the non-negative $\Psi(t)$). Similarly to the above, uniformly for $\omega \in \Delta$

$$\int_{\mathcal{M} \setminus \mathcal{M}_\rho} \left| \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ dx \leq (1 + \epsilon) \xi^{-1} \rho. \tag{5.19}$$

The uniform estimates (5.18) and (5.19) imply that

$$\begin{aligned} & \int_{\mathcal{M} \setminus \Delta_{L,t;x,R+D}} \int_{\Delta_{L,t;x,R+D}} \left| \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ \mathcal{P}(\omega) dx \\ & \leq (1 + \epsilon) \xi^{-1} (\sup_{x \in \mathcal{M}_\rho} \mathcal{P}(\Delta_{T,t;x,R+D}) + \rho). \end{aligned}$$

Upon substituting the latter estimate and (5.17) into (5.16), and then to the integral (5.12), we finally obtain

$$\begin{aligned} & \mathbb{E} \left| \int_{\mathcal{M}} \frac{\mathcal{N}_{norm}(f_{\alpha};T, t/T^n; x, (R + D)/T)}{c(n, \alpha) \cdot \text{Vol } B(R + D)} - \Psi(t) \right|_+ dx \\ & \leq \epsilon + (1 + \epsilon) \xi^{-1} (\sup_{x \in \mathcal{M}_\rho} \mathcal{P}(\Delta_{T,t;x,R+D}) + \rho), \end{aligned}$$

which could be made arbitrarily small for each sufficiently small choice of ξ excising the very small components, and using (5.14). This concludes the proof of (5.12), sufficient to yield the conclusion of Proposition 3. \square

6 Final remark: volumes of the nodal components

6.1 Volume distribution of boundary components

We would like to point out that the methods we are using are rather general and with minimal change one may prove other results. These in particular include similar results about the $(n - 1)$ dimensional volumes of nodal components or boundaries of nodal domains instead of the n dimensional volume of the nodal domains. The analogue of Theorem 5 is proved in exactly the same way using Kac–Rice for the bound and ergodic theory for the existence of the limit. Going from the planar case to the Riemannian one is also straightforward. All these results could be rewritten line-by-line, with the definition of \mathcal{N} essentially the only change.

The only difference is in the results stating that the limit functions are strictly increasing i.e. Theorems 6 and 7. The main idea is still the same: we have to create a deterministic example with a nodal component of (approximately) given size, and then show that small perturbations do not significantly alter its size with positive probability. For $\alpha < 1$ the argument is essentially the same as for the volume of nodal domains. In this case the corresponding Hilbert space \mathcal{H} is dense in any $C^k(Q)$, hence we can construct any non-degenerate example. For $\alpha = 1$, the same family of domains works for the nodal sets, the minimal volume will be the volume of the surface volume of the sphere of radius $j_{n/2-1,1}$ instead of the the n dimensional volume of the corresponding ball. (This follows from the isoperimetric inequality.) The only real difference is in Lemma 10: controlling the C^1 norm is not sufficient in order to control the change of the boundary volume. But it is not very difficult to show that if the perturbation has a small C^2 -norm b , then the ratio of boundary volumes of the original and perturbed domains is $1 + O(b)$ where the constant in $O(b)$ depends only on the C^2 norms.

6.2 Coupling between domain boundary volumes

Finally, we point that it is possible to consider the *joint* distribution of

$$(\text{Vol}_n(\omega), \text{Vol}_{n-1}(\partial\omega))$$

where ω is a nodal domain. The existence of the scaling limit is established via an identical argument to the above. Constructing a deterministic example in the case $\alpha < 1$ also follows along the lines of the argument in this manuscript. This would give that the limiting distribution $\Psi(t, l)$ is strictly increasing in both t and l for all $t > 0$ and $l > l_0(t)$, where $l_0(t)$ is the surface area of a sphere with volume t (follows from the isoperimetric inequality).

The case $\alpha = 1$ is more complicated. From the isoperimetric inequality, for each $t > t_0$, the size of the boundary must be $\geq l_0(t)$. But it is clear that this is not the best

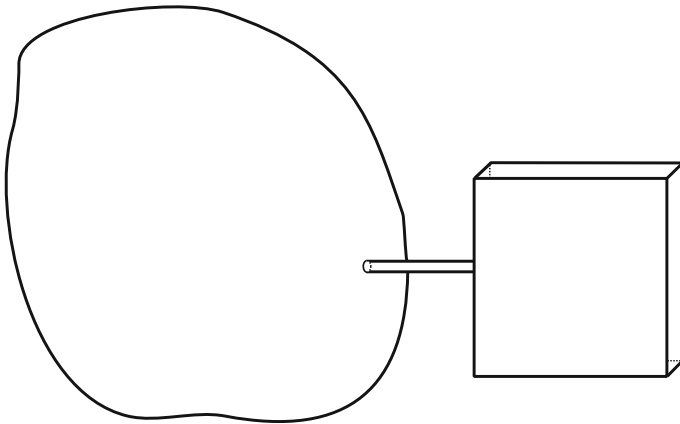


Fig. 1 Connecting a thin plate by a thin tube

possible lower bound, as otherwise the domain would be a ball and its main eigenvalue would be larger than 1 unless $t = t_0$. On the other hand, there *exists* a deterministic infimum of the boundary volume over all domains that have volume $t \geq t_0$ and the main eigenvalue 1. We denote this infimum by $l_1(t)$. There exists a domain for which the boundary is *arbitrary close* to $l_1(t)$. Next we attach to this domain a very thin tube and at the end of this tube we attach a very thin plate, see Fig. 1. This only slightly perturbs the volume, does not alter the main eigenvalue significantly, but increases the boundary volume a lot. After that we take a small perturbation to make the domain real analytic and use Lax-Malgrange to approximate the main eigenfunction by a plane wave. This way we can construct a nodal domain with eigenvalue 1, volume arbitrary close to $t \geq t_0$ and boundary volume arbitrary close to $l \geq l_1(t)$.

Acknowledgements We would like to thank P. Sarnak for raising the question of nodal volumes and his interest in our work, and M. Sodin for many useful discussions. We are grateful to the anonymous referees for the very careful reading of the original manuscript and for helping us to improve the readability of the presented research. The first author was partially funded by Engineering & Physical Sciences Research Council (EPSRC) Fellowship Ref. EP/M002896/1. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013), ERC Grant Agreement No. 335141 (I.W.)

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer Monographs in Mathematics. Springer, New York (2007)
2. Azařs, J.-M., Wschebor, M.: Level Sets and Extrema of Random Processes and Fields. Wiley, Hoboken (2009)
3. Berry, M.: Regular and irregular semiclassical wavefunctions. J. Phys. A: Math. General **10**(12), 2083 (1977)

4. Bogomolny, E., Schmit, C.: Percolation model for nodal domains of chaotic wave functions. *Phys. Rev. Lett.* **88**, 114102 (2002)
5. Canzani, Y., Hanin, B.: High frequency eigenfunction immersions and supremum norms of random waves. *Electr. Res. Announc. Math. Sci.* **22**, 76–86 (2015)
6. Canzani, Y., Hanin, B.: Scaling limit for the kernel of the spectral projector and remainder estimates in the pointwise weyl law. *Anal. PDE* **8**(7), 1707–1731 (2015)
7. Canzani, Y., Sarnak, P.: On the Topology of the Zero Sets of Monochromatic Random Waves. [arXiv:1412.4437](https://arxiv.org/abs/1412.4437) (2014)
8. Gayet, D., Welschinger, J.-Y.: Lower estimates for the expected Betti numbers of random real hypersurfaces. *J. Lond. Math. Soc. (2)* **90**(1), 105–120 (2014)
9. Nazarov, F., Sodin, M.: On the number of nodal domains of random spherical harmonics. *Am. J. Math.* **131**(5), 1337–1357 (2009)
10. Nazarov, F., Sodin, M.: Random complex zeroes and random nodal lines. In: *Proceedings of the International Congress of Mathematicians, vol. III*, pp. 1450–1484. Hindustan Book Agency, New Delhi (2010)
11. Nazarov, F., Sodin, M.: Asymptotic Laws for the Spatial Distribution and the Number of Connected Components of Zero Sets of Gaussian Random Functions. [arXiv:1507.02017](https://arxiv.org/abs/1507.02017) (2015)
12. Sarnak, P., Wigman, I.: Topologies of Nodal Sets of Random Band Limited Functions. [arXiv:1510.08500](https://arxiv.org/abs/1510.08500) (2015)
13. Sodin, M.: Lectures on random nodal portraits. In: *Probability and statistical physics in St. Petersburg, volume 91 of Proceedings of Symposia in Pure Mathematics*, pp. 395–422. American Mathematical Society, Providence, RI (2016)