ALEA, Lat. Am. J. Probab. Math. Stat. 21, 1601–1624 (2024)
DOI: 10.30757/ALEA.v21-60

# Non-universal moderate deviation principle for the nodal length of arithmetic Random Waves

**Claudio Macci, Maurizia Rossi and Anna Vidotto**

Dipartimento di Matematica
Università degli Studi di Roma Tor Vergata
1, Via della Ricerca Scientifica, 00133 Roma, Italia
*E-mail address*: macci@mat.uniroma2.it
*URL*: https://www.mat.uniroma2.it/~macci

Dipartimento di Matematica e Applicazioni
Università degli Studi di Milano-Bicocca
55, Via Roberto Cozzi, 20125 Milano, Italia
*E-mail address*: maurizia.rossi@unimib.it
*URL*: https://www.unimib.it/maurizia-rossi

Dipartimento di Matematica e Applicazioni
Università di Napoli Federico II
Via Cintia, Complesso Monte S. Angelo, 80126 Napoli, Italia
*E-mail address*: anna.vidotto@unina.it
*URL*: https://sites.google.com/view/annavidotto

**Abstract.** Inspired by the recent work Macci et al. (2021), we prove a non-universal non-central Moderate Deviation Principle for the nodal length of arithmetic random waves (Gaussian Laplace eigenfunctions on the standard flat torus) both on the whole manifold and on shrinking toral domains. Second order fluctuations for the latter were established by Marinucci et al. (2016) and Benatar et al. (2020) respectively, by means of chaotic expansions, number theoretical estimates and full correlation phenomena. Our proof is simple and relies on the interplay between the long memory behavior of arithmetic random waves and the chaotic expansion of the nodal length, as well as on well-known techniques in Large Deviation theory (the contraction principle and the concept of exponential equivalence).

*Received by the editors July 27th, 2023; accepted January 18th, 2024.*

2010 *Mathematics Subject Classification.* 60G60, 60F10, 58J50.

*Key words and phrases.* Arithmetic Random Waves; Nodal length; Moderate Deviation Principle; Wiener chaos.

C. M. has been supported by the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome "Tor Vergata" (CUP E83C18000100006 and CUP E83C23000330006), by University of Rome "Tor Vergata" ("Asymptotic Properties in Probability" (CUP E83C22001780005)) and by the GNAMPA-INdAM Project *Stime asintotiche: principi di invarianza e grandi deviazioni.* M. R. has been supported by the ANR-17-CE40-0008 project *Unirandom.* A. V. has been supported by the co-financing of the European Union - FSE-REACT-EU, PON Research and Innovation 2014–2020, DM 1062/2021. The authors are members of INdAM-GNAMPA..

## 1. Introduction

In recent years there has been a growing interest for the geometry of random waves motivated by theoretical issues in Geometric Analysis, Mathematical Physics, Probability, Number Theory (see among others Yau (1982), Berry (1977), Berry (2002), Wigman (2010), Krishnapur et al. (2013)) as well as applications, for instance in Cosmology, Climate Science and Brain Imaging (see for instance Marinucci and Peccati (2011), Christakos (1992), Caponera et al. (2023), Taylor and Worsley (2007)).

Yau (1982) conjectured that the nodal volume, i.e., the measure of the zero locus, which is a smooth (hyper)surface outside of a codimension-two singular set, of *any* Laplacian eigenfunction $f$ on a closed $C^\infty$-smooth Riemannian manifold $(M, g)$ is comparable to the square root of the corresponding eigenvalue $E$. More precisely,

$$c\sqrt{E} \le \mathrm{Vol}_g(f^{-1}(0)) \le C\sqrt{E}, \tag{1.1}$$

for some constants $C, c > 0$ depending only on the manifold. Brüning (1978) proved the lower bound in (1.1) in dimension two (for the length of nodal lines), while for real analytic metrics in any dimension this conjecture was settled by Donnelly and Fefferman (1988) ten years later. Recently, Logunov (2018) established the lower bound in the smooth case in any dimension, while the optimal upper estimate in (1.1) is still an open problem in full generality.

Inspired by the work by Kac on zeros of random polynomials, Bérard (1985) proposed to investigate *random* eigenfunctions: for compact symmetric spaces of rank one (the round sphere e.g.) he suggested to make use of the multiplicities in the spectrum of the Laplacian in order to endow the eigenspace (which is a *finite* dimensional vector space), say of eigenvalue $E$, with a Gaussian measure and computed the *mean* nodal volume. Consistently with Yau's conjecture, it turned out to be proportional to $\sqrt{E}$ by a constant factor (the volume of the manifold times an explicit constant that only depends on the dimension of the space). Since then, several authors investigated geometric properties of random eigenfunctions, motivated also by Berry's ansatz (Berry, 1977) on universality of high-energy eigenfunctions for "generic" classically chaotic billiards.

Rudnick and Wigman (2008), inspired by Bérard's model, considered random eigenfunctions on the standard two-dimensional flat torus $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, the so-called *arithmetic* random waves. In this case the spectral degeneracy properties are related to the "sums of squares" problem, see e.g. Bombieri and Bourgain (2015). Indeed, the eigenvalues (or energy levels) of the Laplace-Beltrami operator of $\mathbb{T}^2$ are all numbers $E = E_n := 4\pi^2 n$ where $n$ is a sum of two squares, namely

$$n = a^2 + b^2, \quad a, b \in \mathbb{Z}, \tag{1.2}$$

and the multiplicity of $E_n$, say $\mathcal{N}_n$, coincides with the number of pairs $(a, b) \in \mathbb{Z}^2$ lying on the circle of radius $\sqrt{n}$. Landau's Theorem (Landau, 1908) ensures that $\mathcal{N}_n$ grows on average as $\sqrt{\log n}$, however it could be as small as 8 or as big as a power of $\log n$, depending on the chosen subsequence of energy levels.

In order to go deeply into the erratic behavior of lattice points, we define the atomic probability measure $\mu_n$ on the unit circle by projecting points with coordinates $(a, b) \in \mathbb{Z}^2$ satisfying (1.2), attaching to them a Dirac mass and then averaging. It is well known that there exists a density-1 sequence of eigenvalues for which $\{\mu_n\}_n$ converges to the uniform probability measure on the unit circle. (In this case, the corresponding *pointwise* scaling limit of the covariance kernel of arithmetic random waves is indeed those of Berry's random wave model (Berry, 1977, 2002), i.e., the Bessel function of the first kind of order zero.) However, there are other weak-* limits classified in Kurlberg and Wigman (2017). In particular, for every $\eta \in [-1, 1]$ there exists a subsequence of eigenvalues whose corresponding sequence of probability measures $\{\mu_n\}_n$ is such that

$$\widehat{\mu_n}(4) \longrightarrow \eta, \qquad \text{as } \mathcal{N}_n \to +\infty, \tag{1.3}$$

where $\widehat{\mu_n}(4)$ denotes the 4-th Fourier coefficient of $\mu_n$. Note that when the scaling limit is Berry's model, then $\eta = 0$.

For the nodal lines of the $n$-th arithmetic random wave, Rudnick and Wigman proved the length $\mathcal{L}_n$ to be on *average* as big as $\sqrt{E_n}$, the square root of the corresponding eigenvalue (analogously to the spherical case), in accordance with Yau's conjecture. More precisely,

$$\mathbb{E}[\mathcal{L}_n] = \frac{1}{2\sqrt{2}}\sqrt{E_n}. \tag{1.4}$$

The sharp variance, for large eigenvalues (actually as $\mathcal{N}_n \to +\infty$) has been found by Krishnapur et al. (2013) to be

$$\mathrm{Var}(\mathcal{L}_n) \sim \frac{1+\widehat{\mu_n}(4)^2}{512}\frac{E_n}{\mathcal{N}_n^2}. \tag{1.5}$$

Remarkably, its behavior turns out to be non-universal, depending on fine properties of lattice points. In particular, (1.4) and (1.5) together with Markov inequality entail concentration of the nodal length around its mean, giving some information on the constants in (1.1): as $\mathcal{N}_n \to +\infty$, $\mathcal{L}_n/\sqrt{E_n} \longrightarrow^{\mathbb{P}} 1/2\sqrt{2}$, see Section 2.2.1 for more details.

Regarding second order fluctuations of the nodal length of arithmetic random waves, Marinucci et al. (2016) proved a non-universal non-central limit theorem: as $\mathcal{N}_n \to +\infty$ s.t. (1.3) holds, the limiting distribution of $\mathcal{L}_n$ is a linear combination of independent chi-square distributed random variables $X_1^2$ and $X_2^2$, whose coefficients depend on (1.3)

$$\frac{\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n]}{\sqrt{\mathrm{Var}(\mathcal{L}_n)}} \xrightarrow{d} \frac{1}{2\sqrt{1+\eta^2}}(2-(1+\eta)X_1^2-(1-\eta)X_2^2) =: \mathcal{M}_\eta, \tag{1.6}$$

see (2.9) for a complete discussion.

Benatar et al. (2020) studied the nodal length $\mathcal{L}_n$ of high-energy arithmetic random waves restricted to *shrinking domains*, all the way down to Planck scale. In particular, they consider centered balls $B(s_n)$ with vanishing radius $s_n$ such that $s_n\sqrt{E_n} \to +\infty$, see the work by Benatar et al. (2020) for precise assumptions. (Indeed, when valid, Berry's random wave model is applicable to shrinking domains of radius $\approx \frac{C}{\sqrt{E}}$ with $C \gg 0$, see the Introduction in Benatar et al. (2020).) In this work they found that, up to a natural scaling, the variance of the restricted nodal length $\mathcal{L}_{n;s_n}$ obeys the same asymptotic law as the total nodal length $\mathcal{L}_n$, and remarkably global and local nodal lengths are asymptotically fully correlated. In particular this implies $\mathcal{L}_{n;s_n}$ to exhibit the same limiting behavior as $\mathcal{L}_n$ in (1.6).

The present paper is installed within these results as a refinement of the non-central limit theorems for both total and local nodal length of arithmetic random waves established respectively in Marinucci et al. (2016) and Benatar et al. (2020). We are inspired by Macci et al. (2021) where the authors investigate the asymptotic behavior of the nodal length for random Laplace eigenfunctions on the unit two-dimensional round sphere, see Section 3 for a complete discussion.

In the toral case, we are able to prove a non-universal non-central Moderate Deviation Principle (MDP) – see Section 2.3 for the definition – collected in Theorems 3.1 and 3.2 with an explicit rate function depending on (1.3). In particular, we quantify at a logarithmic scale the asymptotic probability of *rare* events such as $\mathbb{P}\Big((\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n])/\sqrt{\mathrm{Var}(\mathcal{L}_n)} \le -y\cdot\alpha_n\Big)$, where $y > 0$ and $\{\alpha_n\}_n$ is a sequence of positive numbers going to infinity slowly enough, in accordance with the growth of $\mathcal{N}_n$. Our analysis provides further information on the nodal length of Laplacian eigenfunctions on the torus; indeed, our Theorem 3.1 implies in particular

$$\mathbb{P}\Big(\mathcal{L}_n \le \mathbb{E}[\mathcal{L}_n] - y\cdot\alpha_n\sqrt{\mathrm{Var}(\mathcal{L}_n)}\Big) = e^{-\alpha_n\left(y\frac{\sqrt{1+\eta^2}}{1+|\eta|}+o(1)\right)}, \tag{1.7}$$

as $\mathcal{N}_n \to +\infty$, where mean and variance are as in (1.4) and (1.5) respectively. A result similar to (1.7) holds for the local nodal length by applying Theorem 3.2. For future research, it would be interesting to prove *exponential* concentration of $\mathcal{L}_n$ around its mean, namely a Large Deviation principle for the nodal length of random eigenfunctions at least in the toral and spherical case.

**Plan of the paper.** Section 2 is devoted to background and notation needed in order to understand the model of interest and the goal of the this paper: we first define arithmetic random waves and present prior results on their nodal length, both on the whole torus and in geodesic balls, and then we recall basic notions in Large Deviation theory. In Section 3 we formulate our main results, i.e., Theorems 3.1 and 3.2, together with a brief outline of their proofs which allows us to make a comparison with the spherical case. The last sections are devoted to the detailed proofs: our argument relies on the interplay between the long memory behavior of arithmetic random waves and the chaotic expansion of the nodal length (Section 4), as well as on well-known techniques in Large Deviation theory such as the contraction principle and the concept of exponential equivalence (Section 5).

## 2. Background and notation

In this section we introduce our random model and recall some basic notions in Large Deviation theory, eventually describing our main results.

*Some conventions.* Given two sequences of positive numbers $\{a_n\}_n$ and $\{b_n\}_n$, we write $a_n \vee b_n$ to indicate the maximum between $a_n$ and $b_n$, $a_n = o(b_n)$ if $a_n/b_n \to 0$ as $n \to +\infty$, $a_n = O(b_n)$ or equivalently $a_n \ll b_n$ if $a_n/b_n$ is asymptotically bounded and $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to +\infty$. Moreover, for random variables $\{X_n\}_n$, $X$ and $Y$ we write $X \stackrel{d}{=} Y$ if $X$ and $Y$ share the same law, and finally $X_n \stackrel{d}{\longrightarrow} X$ if the sequence $X_n$ converges to $X$ in distribution.

2.1. *Arithmetic random waves.* Let $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$ be the standard two-dimensional flat torus and $\Delta$ the Laplacian on $\mathbb{T}^2$. The eigenvalues $E \in \mathbb{R}$ of the Helmholtz equation

$$\Delta f + Ef = 0 \tag{2.1}$$

are all numbers of the form $E_n = 4\pi^2 n$ with $n \in S$, where

$$S = \left\{ n \in \mathbb{Z} : n = a^2 + b^2 \text{ for some } a, b \in \mathbb{Z} \right\}.$$

In order to describe Laplace eigenspaces, for $n \in S$ we denote by $\Lambda_n$ the set of frequencies

$$\Lambda_n := \{\lambda = (\lambda_1, \lambda_2) \in \mathbb{Z}^2 : \lambda_1^2 + \lambda_2^2 = n\}$$

whose cardinality $\mathcal{N}_n := |\Lambda_n|$ equals the number of ways to express $n$ as a sum of two square integers. For $\lambda \in \Lambda_n$, we denote by

$$e_\lambda(x) := \exp(2\pi i \langle \lambda, x \rangle), \quad x = (x_1, x_2) \in \mathbb{T}^2,$$

the complex exponential associated to the frequency $\lambda$. The family $\{e_\lambda\}_{\lambda \in \Lambda_n}$ is an $L^2$-orthonormal basis of the eigenspace of $-\Delta$ corresponding to the eigenvalue $E_n$. In particular, its dimension is $\mathcal{N}_n = |\Lambda_n|$. The number $\mathcal{N}_n$ grows (Landau, 1908) *on average* as $\sqrt{\log n}$, but could be as small as 8 for prime numbers $p \equiv 1 \mod 4$, or as large as a power of $\log n$.

From now on, we assume that every random object considered in this paper is defined on a common complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{E}$ denoting mathematical expectation with respect to $\mathbb{P}$.

**Definition 2.1.** For $n \in S$, the $n$-th *arithmetic random wave* is the random field

$$T_n(x) := \frac{1}{\sqrt{\mathcal{N}_n}} \sum_{\lambda \in \Lambda_n} a_\lambda e_\lambda(x), \quad x \in \mathbb{T}^2, \tag{2.2}$$

where the coefficients $a_\lambda$ are standard complex-Gaussian random variables verifying the following properties: $a_\lambda$ is stochastically independent of $a_\gamma$ whenever $\gamma \notin \{\lambda, -\lambda\}$, and $a_{-\lambda} = \overline{a_\lambda}$ (ensuring that $T_n$ is real-valued).

Let us define, for $n \in S$ such that $\sqrt{n}$ is not an integer,

$$\Lambda_n^+ := \{\lambda = (\lambda_1, \lambda_2) \in \Lambda_n : \lambda_2 > 0\}, \tag{2.3}$$

otherwise $\Lambda_n^+ := \{\lambda = (\lambda_1, \lambda_2) \in \Lambda_n : \lambda_2 > 0\} \cup \{(\sqrt{n}, 0)\}$. We assume that $\{a_\lambda\}_{\lambda \in \Lambda_n^+, n \in S}$ is a family of independent random variables, in particular $\{T_n\}_{n \in S}$ is a family of independent random fields. From (2.2), $T_n$ is stationary, centered Gaussian with covariance function $r_n : \mathbb{T}^2 \longrightarrow [-1, 1]$

$$\mathbb{E}[T_n(x)T_n(y)] = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} e_\lambda(x - y) = \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \cos\left(2\pi \langle x - y, \lambda \rangle\right) =: r_n(x - y), \quad x, y \in \mathbb{T}^2.$$

Note that $r_n(0) = 1$, i.e. $T_n$ has unit variance.

2.2. *Nodal length: prior work.* For $n \in S$, the nodal set $T_n^{-1}\{0\} := \{x \in \mathbb{T}^2 : T_n(x) = 0\}$ is a.s. a smooth curve on the torus, we are interested in the *nodal length* of the random eigenfunctions, i.e. the collection $\{\mathcal{L}_n\}_{n \in S}$ of all random variables of the form

$$\mathcal{L}_n := \text{length}(T_n^{-1}\{0\}). \tag{2.4}$$

2.2.1. *Mean and variance.* Rudnick and Wigman (2008) computed the expected value of $\mathcal{L}_n$:

$$\mathbb{E}[\mathcal{L}_n] = \frac{1}{2\sqrt{2}} \sqrt{E_n}, \tag{2.5}$$

which is well reflecting Yau's conjecture (Yau, 1982). A bound for the variance of $\mathcal{L}_n$ was obtained in Rudnick and Wigman (2008), but the challenging task of attaining an exact asymptotic behaviour for $\mathrm{Var}(\mathcal{L}_n)$ was completely settled in Krishnapur et al. (2013) as follows. Let $\mathcal{S}^1 \subset \mathbb{R}^2$ be the unit circle and set, for $n \in S$, $\mu_n$ to be the probability measure on $\mathcal{S}^1$ defined as

$$\mu_n := \frac{1}{\mathcal{N}_n} \sum_{\lambda \in \Lambda_n} \delta_{\frac{\lambda}{\sqrt{n}}}.$$

Thanks to Erdős and Hall (1999), we know that there exists a density-1 subsequence $\{n_j\}_j \subseteq S$ such that

$$\mu_{n_j} \Rightarrow \frac{d\phi}{2\pi}, \tag{2.6}$$

namely $\mu_{n_j}$ converges to the uniform measure on the unit circle (where $\Rightarrow$ denotes weak-$*$ convergence of probability measures, and $d\phi$ is the Lebesgue measure on $\mathcal{S}^1$). Nevertheless, the sequence $\{\mu_n\}_{n \in S}$ has an infinity of other weak-$*$ adherent points, see Cilleruelo (1993); Krishnapur et al. (2013), which are the so-called *attainable measures*, see Kurlberg and Wigman (2017); Sartori (2018).

Finally, we can state the main result of Krishnapur et al. (2013), which is the following: as $\mathcal{N}_n \to +\infty$,

$$\mathrm{Var}(\mathcal{L}_n) = c_n \frac{E_n}{\mathcal{N}_n^2} \left(1 + O\left(\frac{1}{\mathcal{N}_n^{1/2}}\right)\right), \quad c_n = c(\mu_n) := \frac{1 + \widehat{\mu}_n(4)^2}{512}, \tag{2.7}$$

where $\widehat{\mu}(k) = \int_{\mathcal{S}^1} z^{-k} \, d\mu(z)$, $k \in \mathbb{Z}$, represent the Fourier coefficients of a measure $\mu$ on the unit circle. Moreover, from (2.7) it is clear that, in order for $\mathcal{L}_n$ to exhibit an asymptotic law one has to consider a subsequence $\{n_j\}_j \subseteq S$ such that the limit $\lim_{j\to\infty} |\widehat{\mu}_{n_j}(4)|$ exists. Indeed, if $\{n_j\}_j \subseteq S$ is a subsequence such that $\mathcal{N}_{n_j} \to \infty$ and $\mu_{n_j} \Rightarrow \mu$ for some probability measure $\mu$ on $\mathcal{S}^1$, then

$$\mathrm{Var}(\mathcal{L}_{n_j}) \sim c(\mu)\frac{E_{n_j}}{\mathcal{N}_{n_j}^2}, \quad c(\mu) := \frac{1+\widehat{\mu}(4)^2}{512}. \tag{2.8}$$

Thanks to Krishnapur et al. (2013) and Kurlberg and Wigman (2017), we know that for every $\eta \in [-1,1]$ there exists a subsequence $\{n_j\}_j \subseteq S$ such that, as $j \to \infty$, $\mathcal{N}_{n_j} \to \infty$ and

$$\widehat{\mu}_{n_j}(4) \longrightarrow \eta\,.$$

As a consequence, the possible values of the "asymptotic" constant $c(\mu)$ cover the whole interval $\left[\frac{1}{512}, \frac{1}{256}\right]$. In particular, for the full density subsequence $\{n_j\}_j \subseteq S$ such that the lattice points in $\Lambda_{n_j}$ are asymptotically equidistributed (2.6), $\widehat{\mu}_{n_j}(4) \to 0$. On the other hand, the work by Cilleruelo (1993) has shown that there are *thin* (i.e., with density equal to zero) sequences $\{n_j\}_j \subseteq S$, with $\mathcal{N}_{n_j} \to \infty$, such that $\mu_{n_j}$ converges weakly to an atomic probability measure supported at the four symmetric points $\pm 1$, $\pm i$; hence, $\widehat{\mu}_{n_j}(4)^2 \to 1$ and $c_{n_j} \to 1/256$.

2.2.2. *Asymptotic distribution.* From now on, for a (non-zero) finite variance random variable $X$, we set $\widetilde{X}$ to be its normalized version, that is

$$\widetilde{X} := \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}(X)}}\,.$$

Fix $\eta \in [-1,1]$ and let $\mathcal{M}_\eta$ be the random variable

$$\mathcal{M}_\eta := \frac{1}{2\sqrt{1+\eta^2}}(2 - (1+\eta)X_1^2 - (1-\eta)X_2^2), \tag{2.9}$$

where $X_1, X_2$ are independent standard Gaussians. Note that

$$\mathcal{M}_\eta \overset{\mathrm{d}}{=} \mathcal{M}_{-\eta},$$

moreover $\mathcal{M}_\eta$ is not Gaussian, indeed its support is $\left(-\infty, 1/\sqrt{1+\eta^2}\right]$. Except for $\eta_2 = -\eta_1$, $\mathcal{M}_{\eta_1}$ and $\mathcal{M}_{\eta_2}$ have different laws if $\eta_1 \neq \eta_2$.

**Theorem 2.2** (Theorem 1.1, Marinucci et al. (2016))**.** *Let $\{n_j\}_j \subseteq S$ be a subsequence of $S$ satisfying $\mathcal{N}_{n_j} \to \infty$, such that the sequence $\left\{\left|\widehat{\mu}_{n_j}(4)\right|\right\}_j$ converges, that is:*

$$\left|\widehat{\mu}_{n_j}(4)\right| \to \eta, \tag{2.10}$$

*for some $\eta \in [0,1]$. Then*

$$\widetilde{\mathcal{L}}_{n_j} \overset{\mathrm{d}}{\longrightarrow} \mathcal{M}_\eta, \tag{2.11}$$

*where $\mathcal{M}_\eta$ is defined as in (2.9).*

In Peccati and Rossi (2018) a quantitative version of Theorem 2.2 has been proved. Let us recall the definition of 1-Wasserstein distance (see e.g. Nourdin and Peccati (2012, §C) and the references therein): given two random variables $X, Y$ whose laws are $\mu_X$ and $\mu_Y$, respectively, the Wasserstein distance between $\mu_X$ and $\mu_Y$, written $d_{\mathrm{W}}(X,Y)$, is defined as

$$d_{\mathrm{W}}\left(X,Y\right) := \inf_{(A,B)} \mathbb{E}\left[|A-B|\right],$$

where the infimum runs over all pairs of random variables $(A,B)$ with marginal laws $\mu_X$ and $\mu_Y$, respectively. We will mainly use the dual representation

$$d_{\mathrm{W}}\left(X,Y\right) = \sup_{h\in\mathcal{H}_1} \left|\mathbb{E}\left[h(X) - h(Y)\right]\right|, \tag{2.12}$$

where $\mathcal{H}_1$ denotes the class of Lipschitz functions $h : \mathbb{R} \longrightarrow \mathbb{R}$ whose Lipschitz constant is less or equal than 1. Relation (2.12) implies in particular that, if $d_W(X_n, X) \longrightarrow 0$, then $X_n \xrightarrow{d} X$ (the converse implication is false in general).

**Theorem 2.3** (Theorem 2, Peccati and Rossi (2018)). *Let $\{n_j\}_j \subseteq S$ be a subsequence of $S$ satisfying $\mathcal{N}_{n_j} \to \infty$, such that the sequence $\{|\widehat{\mu}_{n_j}(4)|\}_j$ converges, that is:*

$$|\widehat{\mu}_{n_j}(4)| \to \eta,$$

*for some $\eta \in [0,1]$. Then*

$$d_W\left(\widetilde{\mathcal{L}}_{n_j}, \mathcal{M}_\eta\right) \ll \mathcal{N}_{n_j}^{-1/4} \vee \left||\widehat{\mu}_{n_j}(4)| - \eta\right|^{1/2}. \tag{2.13}$$

2.2.3. *Shrinking domains.* For $0 < s < 1/2$, we set

$$\mathcal{L}_{n;s} := \text{length}\left(T_n^{-1}\{0\} \cap B(s)\right)$$

to be the nodal length of $T_n$ restricted to a radius-$s$ ball $B(s)$, where by the stationarity of $T_n$ we may assume that $B(s)$ is centered. Kac-Rice formula (see e.g. Adler and Taylor (2007, Chapter 11)) immediately gives

$$\mathbb{E}[\mathcal{L}_{n;s}] = \frac{1}{2\sqrt{2}}(\pi s^2) \cdot \sqrt{E_n}.$$

One of the main results in the article by Benatar et al. (2020) is that the variance of the nodal length $\mathcal{L}_{n;s}$ of $T_n$ restricted to balls that are shrinking slightly above the so-called Planck scale has a similar form (2.7), and in particular one has the precise identity Benatar et al. (2020, (3.36))

$$\text{Cov}(\mathcal{L}_{n;s}, \mathcal{L}_n) = (\pi s^2) \cdot \text{Var}(\mathcal{L}_n).$$

In what follows we keep the notation of Benatar et al. (2020) to avoid confusion.

**Theorem 2.4** (Theorem 1.1, Benatar et al. (2020)). *For every $\varepsilon > 0$ there exists a density-1 sequence of numbers*

$$S' = S'(\varepsilon) \subseteq S$$

*so that the following hold.*

(1) *Along $n \in S'$ we have $\mathcal{N}_n \to \infty$, and the set of accumulation points of $\{\widehat{\mu}_n(4)\}_{n\in S'}$ contains the interval $[0,1]$.*

(2) *For $n \in S'$, uniformly for all $s > n^{-1/2+\varepsilon}$ we have*

$$\text{Var}(\mathcal{L}_{n;s}) = c_n \cdot (\pi s^2)^2 \cdot \frac{E_n}{\mathcal{N}_n^2}\left(1 + O_\varepsilon\left(\frac{1}{\mathcal{N}_n^{1/2}}\right)\right), \tag{2.14}$$

*where $c_n$ is defined as in (2.7), and the constant involved in the 'O'-notation depends on $\varepsilon$ only.*

(3) *For random variables $X, Y$ we denote as usual their correlation*

$$\text{Corr}(X,Y) := \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)} \cdot \sqrt{\text{Var}(Y)}}.$$

*Then for every $\varepsilon > 0$ we have that*

$$\sup_{s > n^{-1/2+\varepsilon}} |\text{Corr}(\mathcal{L}_{n;s}, \mathcal{L}_n) - 1| \to 0, \tag{2.15}$$

*i.e. the nodal length $\mathcal{L}_{n;s}$ of $T_n$ restricted to a small ball is asymptotically fully correlated with the total nodal length $\mathcal{L}_n$ of $T_n$, uniformly for all $s > n^{-1/2+\varepsilon}$.*

Under the same assumptions as Theorem 2.4, in view of (2.15), $\mathcal{L}_{n;s}$ obeys the same limiting law (2.11) as the total nodal length. The aim of this paper is to refine the result of Theorem 2.2 and the consequences of Theorem 2.4 by means of Large Deviation theory.

2.3. *Large Deviation Principles.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, and $\{X_n\}_{n\in\mathbb{N}}$ a sequence of random variables taking values in some topological space $\mathcal{X}$. For our purpose we can restrict ourselves to the case where $\mathcal{X}$ is a metric space; we denote by $d$ its metric and by $\mathcal{B}(\mathcal{X})$ its Borel $\sigma$-field.

**Definition 2.5.** We say that $\{X_n\}_{n\in\mathbb{N}}$ satisfies the Large Deviation Principle (LDP) with speed $0 \le s_n \nearrow +\infty$ and (good) rate function $\mathcal{I} : \mathcal{X} \longrightarrow [0, +\infty]$ if the level sets $\{x : \mathcal{I}(x) \le \alpha\}, \alpha \ge 0$ are compact and for all $B \in \mathcal{B}(\mathcal{X})$ we have

$$-\inf_{x\in \mathring{B}} \mathcal{I}(x) \le \liminf_{n\to+\infty} \frac{1}{s_n} \log \mathbb{P}(X_n \in B) \le \limsup_{n\to+\infty} \frac{1}{s_n} \log \mathbb{P}(X_n \in B) \le -\inf_{x\in \bar{B}} \mathcal{I}(x),$$

where $\mathring{B}$ (resp. $\bar{B}$) denotes the interior (resp. the closure) of $B$.

*Remark* 2.6. In this paper, and in several common cases in the literature, there exists $x_0$ such that $\mathcal{I}(x) = 0$ if and only if $x = x_0$. As a consequence, if $X_n$ satisfies the LDP in Definition 2.5, then it converges to $x_0$ as $n \to +\infty$, at least in probability. Roughly speaking, one can also say that for every Borel set $B$ such that $x_0 \notin \bar{B}$, the quantity $\mathbb{P}(X_n \in B)$ decays as $e^{-s_n \mathcal{I}(B)}$, where $\mathcal{I}(B) := \inf \{\mathcal{I}(x) : x \in B\} > 0$. Thus, in some sense, the larger is the rate function $\mathcal{I}$ locally around $x_0$, the faster is the convergence to $x_0$.

Let us state the Gärtner-Ellis Theorem (see point (c) of Dembo and Zeitouni (1998, Theorem 2.3.6)) in $\mathbb{R}^d$; in particular we denote the inner product in $\mathbb{R}^d$ by $\langle \cdot, \cdot \rangle$. This theorem will be used in Section 5.1.1 with $d = 3$.

**Theorem 2.7.** *Let $\{X_n\}_{n\in\mathbb{N}}$ be a family of $\mathbb{R}^d$-valued random variables. Assume that $0 \le s_n \nearrow +\infty$ and, for all $\theta \in \mathbb{R}^d$, the limit*

$$\psi(\theta) := \lim_{n\to+\infty} \frac{1}{s_n} \log \mathbb{E}[e^{s_n \langle \theta, X_n \rangle}]$$

*exists as an extended real number. Further, assume that the origin belongs to the interior $\mathcal{D}(\psi) := \{\theta \in \mathbb{R}^d : \psi(\theta) < +\infty\}$. Then, if the function $\psi$ is essentially smooth and lower semicontinuous, $\{X_n\}_{n\in\mathbb{N}}$ satisfies the LDP with speed $s_n$ and good rate function $\psi^*$ defined by $\psi^*(x) := \sup_{\theta\in\mathbb{R}^d} \{\langle \theta, x \rangle - \psi(\theta)\}$ (that is the* Legendre transform *of $\psi$).*

Let us state the so-called *contraction principle*, see Dembo and Zeitouni (1998, Theorem 4.2.1).

**Theorem 2.8.** *Let $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ be a metric space with its Borel $\sigma$-field, and $f : \mathcal{X} \longrightarrow \mathcal{Y}$ a continuous function. If $\{X_n\}_{n\in\mathbb{N}}$ satisfies a LDP with (good) rate function $\mathcal{I}$, then $\{Y_n := f(X_n)\}_{n\in\mathbb{N}}$ satisfies a LDP with (good) rate function $\mathcal{I}_f : \mathcal{Y} \longrightarrow [0, +\infty]$ defined as*

$$\mathcal{I}_f(y) := \inf_{x\in f^{-1}(y)} \mathcal{I}(x).$$

Theorem 2.8 ensures that the LDP is preserved under continuous transformations. In order to extend the contraction principle beyond the continuous case, it is beneficial to recall the notion of exponential equivalence, see Dembo and Zeitouni (1998, Definition 4.2.10).

**Definition 2.9.** Let $\{X_n\}_{n\in\mathbb{N}}$ and $\{Y_n\}_{n\in\mathbb{N}}$ be two sequences of random variables taking values in $\mathcal{X}$. We say that they are exponentially equivalent at speed $0 \le s_n \nearrow +\infty$ if, for every $\delta > 0$, the set $\Gamma_\delta := \{d(X_n, Y_n) > \delta\} \subseteq \Omega$ is measurable and

$$\limsup_{n\to+\infty} \frac{1}{s_n} \log \mathbb{P}(\Gamma_\delta) = -\infty. \tag{2.16}$$

The following theorem states that, from the point of view of Large Deviations, sequences of random variables that are exponentially equivalent are identical, see also Dembo and Zeitouni (1998, Theorem 4.2.13).

**Theorem 2.10.** *Assume that $\{X_n\}_{n\in\mathbb{N}}$ satisfies the LDP with speed $s_n$ and good rate function $\mathcal{I}$. Then, if $\{X_n\}_{n\in\mathbb{N}}$ and $\{Y_n\}_{n\in\mathbb{N}}$ are exponentially equivalent at speed $s_n$, the same LDP holds for $\{Y_n\}_{n\in\mathbb{N}}$.*

A Moderate Deviation Principle (MDP) is a class of LDPs that allows to *fill the gap* between the following asymptotic regimes:

- a convergence to a constant, which is governed by a suitable LDP;
- a weak convergence to a centered Normal distribution.

Some examples of classes of LDPs of this kind in which the weak convergence is towards a non-Gaussian law can be found e.g. in Giuliano and Macci (2023) (see also the references cited therein). It is worth stressing that actually this kind of phenomena are quite natural for random sums (in particular with some suitable compound Geometric distributions).

Typically a MDP is class of LDPs concerning families of random variables which depend on the choice of certain scalings satisfying some suitable conditions, and these LDPs are governed by the same rate function. As a prototype example we can consider the empirical means of i.i.d. random variables, see e.g. Dembo and Zeitouni (1998, Theorem 3.7.1), where the asymptotic regimes cited above concern the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT); in particular a LDP related to the LLN is provided by the well-known Cramér Theorem, see Dembo and Zeitouni (1998, Theorem 2.2.3).

Actually we cannot rigorously say that the MDP in this paper, see Theorems 3.1 and 3.2, (as well as the MDP in Macci et al. (2021)) allows to fill the gap between two asymptotic regimes as above because for the moment we do not have a LDP for the convergence to a constant for our statistics of the field. In the case of the nodal length, as anticipated in the Introduction, from (2.5) and (2.7) we know that for every $\epsilon > 0$

$$\mathbb{P}\left(\left|\frac{\mathcal{L}_n}{\sqrt{E_n}} - \frac{1}{2\sqrt{2}}\right| > \epsilon\right) = O_\epsilon\left(\frac{1}{\mathcal{N}_n^2}\right), \qquad \text{as } \mathcal{N}_n \longrightarrow +\infty$$

(where the constant involved in the $O$-notation depends on $\epsilon$), hence it is natural to explore exponential concentration for $\mathcal{L}_n/\sqrt{E_n}$ around its mean. Nevertheless we believe that the investigation of a LDP for the sequence of r.v.'s $\{\mathcal{L}_n/\sqrt{E_n}\}_n$ deserves a separate study, eventually together with a LDP for the nodal length of random spherical harmonics. However, if it is possible to fill that gap, we would have a *non-central* MDP where the weak convergence is towards the distribution of the random variable $\mathcal{M}_\eta$ defined in (2.9). Since the limiting distribution of the nodal length is not universal, in particular it depends on the subsequence of energy levels, in accordance with the non-universality of our MDP stated above we say that our MDP is also *non-universal*.

## 3. Main Results

In this section we finally state our main results, which are non-universal non-central MDPs (see Section 2.3) refining Theorem 2.2 and the consequences of Theorem 2.4, specifically a class of LDPs for the nodal length of Arithmetic Random Waves for certain scalings $\{\alpha_{n_j}\}_j$ ($\{n_j\}_j \subset S$) that grow to infinity slowly as $j \to \infty$ (see condition (3.1)), with speed $\alpha_{n_j}$, and an $\eta$-dependent rate function $I_\eta$ (see (2.10)). Moreover, for $\alpha_{n_j} \equiv 1$ (note that in this case the condition (3.1) fails) we have the convergence in distribution to the random variable $\mathcal{M}_\eta$ defined in (2.9), that is, we retrieve Theorems 2.2 and the consequences of Theorem 2.4.

**Theorem 3.1.** *Let $\{n_j\}_j \subseteq S$ be such that $\mathcal{N}_{n_j} \to +\infty$ and $\widehat{\mu}_{n_j}(4) \to \eta \in [-1,1]$ as $j \to +\infty$, and $\{\alpha_{n_j}\}_j$ be any sequence of positive numbers such that, as $j \to +\infty$,*

$$\alpha_{n_j} \longrightarrow +\infty \qquad \text{and} \qquad \alpha_{n_j}/\log \mathcal{N}_{n_j} \longrightarrow 0\,. \tag{3.1}$$

*Then the sequence of random variables $\left\{\widetilde{\mathcal{L}}_{n_j}/\alpha_{n_j}\right\}_j$ satisfies a Moderate Deviation Principle (MDP) with speed $\alpha_{n_j}$ and rate function*

$$I_\eta(y) = \begin{cases} -y\dfrac{\sqrt{1+\eta^2}}{1+|\eta|} & \text{if } y \le 0 \\ +\infty & \text{if } y > 0 \end{cases}. \tag{3.2}$$

**Theorem 3.2.** *Let $\{n_j\}_j \subseteq S$ and $n_j^{-\frac{1}{2}+\varepsilon} < s_{n_j} < 1/2$ be such that $s_{n_j} \to 0$. If $\{\alpha_{n_j},\, j \in \mathbb{N}\}$ is any sequence of positive numbers such that, as $j \to +\infty$,*

$$\alpha_{n_j} \longrightarrow +\infty \qquad \text{and} \qquad \alpha_{n_j}/\log \mathcal{N}_{n_j} \longrightarrow 0\,,$$

*then the sequence of random variables $\left\{\widetilde{\mathcal{L}}_{n_j,s_{n_j}}/\alpha_{n_j}\right\}_j$ satisfies a Moderate Deviation Principle (MDP) with speed $\alpha_{n_j}$ and the same rate function as in* (3.2).

Clearly, the sequences of random variables in Theorems 3.1 and 3.2 converge to zero (this is a known consequence of the weak convergence of $\widetilde{\mathcal{L}}_{n_j}$ and $\widetilde{\mathcal{L}}_{n_j,s}$, combined with the Slutsky Theorem); indeed their common rate function uniquely vanishes at $y=0$, see also Remark 2.6. Then, in both cases, the smaller is $|\eta|$, the larger is the rate function $I_\eta$ in (3.2). Then, considering once more what we said in Remark 2.6, the smaller is $|\eta|$, the faster is the convergence to zero.

*Remark* 3.3. It is worth noticing that the condition on the speed $\alpha_n$ in (3.1) may be suboptimal. However, our technique is flexible enough to deal with other geometric properties of nodal sets of arithmetic random waves, even in higher dimension (Maffucci, 2019; Cammarota, 2019). All these interesting cases were left uncovered by the strategy pursued in Macci et al. (2021), where the spherical counterpart has been investigated and (central) MDPs have been established for the nodal length of random spherical harmonics (Gaussian Laplace eigenfunctions on the round sphere) both on the whole manifold and on shrinking spherical caps. It is worth stressing that on the sphere these objects have Gaussian fluctuations in the high-energy limit, and are independent, in marked contrast with the toral case where they show a non-universal non-central asymptotic behavior, and full correlation. See Section 3.1 for more details on this comparison.

3.1. *On the proof of the main results.* In order to list and motivate the proof steps, we start by recalling that our nodal length $\mathcal{L}_n$ in (2.4) lives in the Wiener chaos (being a finite variance functional of a Gaussian field), in particular it can be written as an orthogonal series, converging in $L^2(\mathbb{P})$, of the form

$$\mathcal{L}_n = \mathbb{E}[\mathcal{L}_n] + \sum_{q=1}^{+\infty} \mathcal{L}_n[q], \tag{3.3}$$

where $\mathcal{L}_n[q]$ denotes the orthogonal projection of $\mathcal{L}_n$ into the so-called $q$-th Wiener chaos. Roughly, this expansion relies on the fact that Hermite polynomials form an orthogonal basis of the space of square integrable functions on the real line with respect to the Gaussian density. It turns out that projections onto *odd* order chaoses vanish, and moreover $\mathcal{L}_n[2]=0$ (related to so-called Berry's cancellation phenomenon, see Krishnapur et al. (2013); Marinucci et al. (2016)).

Our argument relies on the fact that the asymptotic behavior of the (centered) nodal length $\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n]$ is completely determined by $\mathcal{L}_n[4]$, its fourth chaotic component Marinucci et al. (2016), and, moreover, that the dominant term of $\widetilde{\mathcal{L}}_n[4]$ (where $\widetilde{\mathcal{L}}_n[4]$ denotes the standardized fourth chaos of the nodal length), say $\mathcal{M}_n = \mathcal{M}_{\widehat{\mu}_n(4)}$ *abusing notation*, is a continuous function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ of a random vector $\widetilde{S}_n$ (whose components live in the 2-nd Wiener chaos), $\widetilde{S}_n$ converging in law towards a multivariate Gaussian. First we establish a Moderate Deviation Principle for $\widetilde{S}_n/\sqrt{\alpha_n}$ at speed $\alpha_n$, and then, thanks to the celebrated contraction principle (see Dembo and Zeitouni (1998, Theorem 4.2.1) as well as Theorem 2.8), we transfer this result to $\mathcal{M}_n/\alpha_n$ with an explicit rate function (indeed, $f$ is a multivariate polynomial) and speed $\alpha_n$. It remains to deal with the tail of the series (3.3), i.e. with

$$\widetilde{\mathcal{L}}_n - \mathcal{M}_n.$$

To this aim, we prove that $\widetilde{\mathcal{L}}_n/\alpha_n$ and $\mathcal{M}_n/\alpha_n$ – under some additional constraint on the speed $\alpha_n$ – are exponentially equivalent (Definition 2.9), so that they satisfy the same Deviation Principle (Theorem 2.10). Finally, we will take advantage of the full correlation result in Benatar et al. (2020, Theorem 1.5) (see also Theorem 2.4) to check that $\widetilde{\mathcal{L}}_n/\alpha_n$ and $\widetilde{\mathcal{L}}_{n;s}/\alpha_n$, the standardized nodal length restricted to the ball $B(s)$ whose radius is slightly above the Planck scale, are exponentially equivalent, thus sharing the same deviations.

3.2. *The spherical case.* Our proofs are inspired by Macci et al. (2021), where a Moderate Deviation Principle has been established for the nodal length of *random spherical harmonics*, both on the whole sphere and on shrinking spherical caps. However, there are marked differences with the toral case. Indeed, on the sphere the limiting distribution of these two geometric functionals is Gaussian, and they are asymptotically independent (see Todino (2020) and the references therein). On $\mathbb{T}^2$ instead, the high-energy behavior of the total nodal length is non-Gaussian and non-universal (Theorem 2.2), and it is asymptotically fully correlated with the nodal length in shrinking domains slightly above the Planck scale (Theorem 2.4). Moreover, the rate function $J$ in the main results of Macci et al. (2021) is quadratic

$$J(y) = y^2/2, \quad y \in \mathbb{R},$$

while our rate function (3.2) is a line whose angular coefficient depends on $\eta$. However, the condition (3.1) on the MDP speed is comparable in some sense to those on the sphere (see (2.1) and (2.2) in Macci et al. (2021)), indeed $\mathcal{N}_n$ grows on average as $\sqrt{\log n}$, see Landau (1908).

For the nodal length on the sphere, the starting point is a chaotic expansion similar to (3.3) and also in this case the dominant term is the fourth chaotic component, that however behaves much differently than the corresponding term on the torus: it is equivalent in the high-energy limit to the so-called sample trispectrum (i.e., the integral of the fourth Hermite polynomial evaluated at the field), which is *not* true in the toral case, and it is asymptotically Gaussian. To prove a MDP for the total nodal length on the sphere, the authors of Macci et al. (2021) first showed a MDP for the sample trispectrum via the cumulant approach presented in Schulte and Thäle (2016), which is a link between the fourth moment theorem for Gaussian approximation (Nourdin and Peccati, 2012) and the LD theory, and then checked its exponential equivalence with the total nodal length (a similar argument works for shrinking domains).

## 4. Chaotic expansions

In this section we recall the chaotic expansion of the nodal length for arithmetic random waves, that is crucial for the proof of our main results. In particular, at the end of Section 4 we focus on the fourth chaotic component of $\mathcal{L}_n$ and its dominant term $\mathcal{M}_n$.

4.1. *Wiener chaos.* The family of Hermite polynomials $\{H_q\}_{q\in\mathbb{N}}$ is defined as follows: $H_0 \equiv 1$ and

$$H_q(t) := (-1)^q \phi^{-1}(t) \frac{d^q}{dt^q} \phi(t), \qquad t \in \mathbb{R}, q \in \mathbb{N}_{\geq 1}, \tag{4.1}$$

where $\phi$ denotes the standard Gaussian density. It is well known (Nourdin and Peccati, 2012, Proposition 1.4.2) that $\{H_q/\sqrt{q!}\}_{q\in\mathbb{N}}$ is a complete orthonormal system in the space of square integrable real functions with respect to the standard Gaussian measure on the real line.

Arithmetic Random Waves (2.2) are generated (see Definition 2.1) from a family of standard complex-valued Gaussian random variables $\{a_\xi\}_{\xi\in\mathbb{Z}^2}$, defined on $(\Omega, \mathscr{F}, \mathbb{P})$ and stochastically independent, save for the relations $a_{-\xi} = \overline{a_\xi}$.

Let $X$ be the closure in the Hilbert space $L^2(\mathbb{P})$ (with respect to the scalar product $\langle F, G\rangle := \mathbb{E}[(F-\mathbb{E}[F])(G-\mathbb{E}[G])], F, G \in L^2(\mathbb{P})$) of all real finite linear combinations of random variables $\zeta$ of the form

$$\zeta = z\, a_\xi + \overline{z}\, a_{-\xi}\,,$$

where $\xi \in \mathbb{Z}^2$ and $z \in \mathbb{C}$, thus $X$ is a real centered Gaussian Hilbert subspace of $L^2(\mathbb{P})$. Now let $q \in \mathbb{N}$; the $q$-th Wiener chaos $C_q$ associated with $X$ is defined as the closure in $L^2(\mathbb{P})$ of all real finite linear combinations of random variables of the form

$$H_{p_1}(\zeta_1)H_{p_2}(\zeta_2)\cdots H_{p_k}(\zeta_k)$$

for $k \in \mathbb{N}_{\geq 1}$, where $p_1, ..., p_k \in \mathbb{N}$ satisfy $p_1 + \cdots + p_k = q$, and $(\zeta_1, \zeta_2, \ldots, \zeta_k)$ is a standard Gaussian vector extracted from $X$ ($C_0 = \mathbb{R}$ and $C_1 = X$). Note that, from (2.2), for every $n$ the random fields $T_n$, $\nabla T_n$ viewed as collections of Gaussian random variables indexed by $x \in \mathbb{T}^2$ are all lying in $X$.

It turns out that (see e.g. Nourdin and Peccati (2012, Theorem 2.2.4)) $C_q \perp C_{q'}$ in $L^2(\mathbb{P})$ whenever $q \neq q'$, and moreover

$$L^2_X(\mathbb{P}) = \bigoplus_{q=0}^{\infty} C_q,$$

where $L^2_X(\mathbb{P}) := L^2(\Omega, \sigma(X), \mathbb{P})$, that is, every finite-variance real-valued functional $F$ of $X$ admits a unique representation as a series, converging in $L^2_X(\mathbb{P})$, of the form

$$F = \sum_{q=0}^{\infty} F[q], \tag{4.2}$$

$F[q] := \mathrm{proj}(F\,|\,C_q)$ being the orthogonal projection of $F$ onto $C_q$ (in particular, $F[0] = \mathbb{E}[F]$). For a complete discussion on Wiener chaos see Nourdin and Peccati (2012, §2.2) and the references therein.

Since the nodal length $\mathcal{L}_n$ is a finite-variance *explicit* functional (4.3) of the Gaussian field $T_n$, its Wiener-Îto chaos expansion (3.3) can be fruitfully exploited, as we will see in the proofs of our main results in Section 5.

4.2. *Explicit formulas.* The nodal length (2.4) can be formally written as

$$\mathcal{L}_n = \int_{\mathbb{T}} \delta_0(T_n(\theta))\|\nabla T_n(\theta)\|\, d\theta, \tag{4.3}$$

where $\delta_0$ denotes the Dirac delta function and $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^2$ (see Rudnick and Wigman (2008, Lemma 3.1)). Clearly, we mean that the approximating random variables $\mathcal{L}_n^\epsilon$ defined by replacing the Dirac mass $\delta_0$ with $1_{[-\epsilon,\epsilon]}/(2\epsilon)$, for $\epsilon > 0$, in (4.3), converge a.s. and in $L^2(\mathbb{P})$ to $\mathrm{length}(T_n^{-1}\{0\}) = \mathcal{L}_n$.

Note that a straightforward differentiation of the definition (2.2) of $T_n$ yields, for $j = 1, 2$

$$\partial_j T_n(x) = \frac{2\pi i}{\sqrt{\mathcal{N}_n}} \sum_{(\lambda_1,\lambda_2)\in\Lambda_n} \lambda_j a_\lambda e_\lambda(x), \tag{4.4}$$

(here $\partial_j = \frac{\partial}{\partial x_j}$). Hence the random fields $T_n, \partial_1 T_n, \partial_2 T_n$ viewed as collections of Gaussian random variables indexed by $x \in \mathbb{T}^2$ are all lying in $X$, i.e. for every $x \in \mathbb{T}^2$ we have

$$T_n(x),\ \partial_1 T_n(x),\ \partial_2 T_n(x) \in X.$$

We shall often use the following result from Rudnick and Wigman (2008):

**Lemma 4.1** (Rudnick and Wigman (2008), (4.1))**.** *For $j = 1, 2$ we have that*

$$\mathrm{Var}[\partial_j T_n(x)] = \frac{4\pi^2}{\mathcal{N}_n} \sum_{\lambda=(\lambda_1,\lambda_2)\in\Lambda_n} \lambda_j^2 = 4\pi^2 \frac{n}{2},$$

*where the derivatives $\partial_j T_n(x)$ are as in (4.4).*

Accordingly, for $x = (x_1, x_2) \in \mathbb{T}$ and $j = 1, 2$, we will denote by $\partial_j \widetilde{T}_n(x)$ the normalized derivative

$$\partial_j \widetilde{T}_n(x) := \frac{1}{2\pi}\sqrt{\frac{2}{n}} \frac{\partial}{\partial x_j} T_n(x) = \sqrt{\frac{2}{n}} \frac{i}{\sqrt{\mathcal{N}_n}} \sum_{\lambda\in\Lambda_n} \lambda_j\, a_\lambda e_\lambda(x). \tag{4.5}$$

In view of convention (4.5), we formally rewrite (4.3) as

$$\mathcal{L}_n = \sqrt{\frac{4\pi^2 n}{2}} \int_{\mathbb{T}} \delta_0(T_n(x)) \sqrt{(\partial_1 \widetilde{T}_n(x))^2 + (\partial_2 \widetilde{T}_n(x))^2}\, dx.$$

We also introduce two collections of coefficients $\{\alpha_{2n,2m} : n, m \geq 1\}$ and $\{\beta_{2l} : l \geq 0\}$, that are related to the Hermite expansion of the norm $\|\cdot\|$ in $\mathbb{R}^2$ and the (formal) Hermite expansion of the Dirac mass $\delta_0(\cdot)$, respectively. These are given by

$$\beta_{2l} := \frac{1}{\sqrt{2\pi}} H_{2l}(0), \tag{4.6}$$

where $H_{2l}$ denotes the $2l$-th Hermite polynomial, and

$$\alpha_{2n,2m} = \sqrt{\frac{\pi}{2}} \frac{(2n)!(2m)!}{n!m!} \frac{1}{2^{n+m}} p_{n+m}\left(\frac{1}{4}\right), \tag{4.7}$$

where for $N \in \mathbb{N}$ and $x \in \mathbb{R}$

$$p_N(x) := \sum_{j=0}^{N} (-1)^j \cdot (-1)^N \binom{N}{j} \frac{(2j+1)!}{(j!)^2} x^j.$$

**Proposition 4.2** (Marinucci et al. (2016), Proposition 3.2)**.** (a) *For $q = 2$ or $q = 2m + 1$ odd ($m \geq 0$),*

$$\mathcal{L}_n[q] \equiv 0,$$

*that is, the corresponding chaotic projection vanishes.*

(b) *For $q = 0$ or $q \geq 2$*

$$\begin{aligned} &\mathcal{L}_n[2q] \\ &= \sqrt{\frac{4\pi^2 n}{2}} \sum_{u=0}^{q} \sum_{k=0}^{u} \frac{\alpha_{2k,2u-2k}\beta_{2q-2u}}{(2k)!(2u-2k)!(2q-2u)!} \times \\ &\qquad \times \int_{\mathbb{T}} H_{2q-2u}(T_n(x)) H_{2k}(\partial_1 \widetilde{T}_n(x)) H_{2u-2k}(\partial_2 \widetilde{T}_n(x))\, dx. \end{aligned} \tag{4.8}$$

*Consolidating the above, the Wiener-Itô chaotic expansion of $\mathcal{L}_n$ is*

$$\mathcal{L}_n = \mathbb{E}[\mathcal{L}_n] + \sqrt{\frac{4\pi^2 n}{2}} \sum_{q=2}^{+\infty} \sum_{u=0}^{q} \sum_{k=0}^{u} \frac{\alpha_{2k,2u-2k}\beta_{2q-2u}}{(2k)!(2u-2k)!(2q-2u)!} \times$$
$$\times \int_{\mathbb{T}} H_{2q-2u}(T_n(x)) H_{2k}(\widetilde{\partial}_1 T_n(x)) H_{2u-2k}(\widetilde{\partial}_2 T_n(x))\, dx,$$

*in $L^2(\mathbb{P})$.*

4.2.1. *The fourth chaotic component.* For $n \in S$, we set

$$W(n) := \begin{pmatrix} W_1(n) \\ W_2(n) \\ W_3(n) \\ W_4(n) \end{pmatrix} := \frac{1}{n\sqrt{\mathcal{N}_n/2}} \sum_{\lambda=(\lambda_1,\lambda_2)\in\Lambda_n^+} (|a_\lambda|^2 - 1) \begin{pmatrix} n \\ \lambda_1^2 \\ \lambda_2^2 \\ \lambda_1\lambda_2 \end{pmatrix}, \tag{4.9}$$

where $\Lambda_n^+$ is defined as in (2.3). Starting from the formula (4.8) in the case $q = 2$, Marinucci et al. (2016) show the following.

**Lemma 4.3** (Lemma 4.2, Marinucci et al. (2016), Lemma 4, Peccati and Rossi (2018))**.** *We have, for diverging subsequences $\{n_j\} \subseteq S$ such that $\mathcal{N}_{n_j} \to +\infty$ and $\widehat{\mu}_{n_j}(4)$ converges,*

$$\mathcal{L}_{n_j}[4] = \sqrt{\frac{E_{n_j}}{512\,\mathcal{N}_{n_j}^2}} \Big( W_1(n_j)^2 - 2W_2(n_j)^2 - 2W_3(n_j)^2 - 4W_4(n_j)^2 + R(n_j) \Big),$$

*where*

$$R(n_j) = \frac{1}{2\mathcal{N}_{n_j}} \sum_{\lambda\in\Lambda_{n_j}} |a_\lambda|^4$$

*is a sequence of random variables converging in probability to* 1.

It is crucial to note that (see Peccati and Rossi (2018, (46)))

$$W_2(n) + W_3(n) = W_1(n)\,,$$

which implies that the fourth chaotic component of $\mathcal{L}_{n_j}$ can be written as

$$\mathcal{L}_{n_j}[4] = \sqrt{\frac{E_{n_j}}{512\,\mathcal{N}_{n_j}^2}} \Big( (W_2(n_j) + W_3(n_j))^2 - 2W_2(n_j)^2 - 2W_3(n_j)^2 - 4W_4(n_j)^2 + R(n_j) \Big). \tag{4.10}$$

We will use also the following important Lemma from Marinucci et al. (2016).

**Lemma 4.4** (Lemma 4.3, Marinucci et al. (2016))**.** *Assume that the subsequence $\{n_j\}_j \subseteq S$ is such that $\mathcal{N}_{n_j} \to +\infty$ and $\widehat{\mu}_{n_j}(4) \to \eta \in [-1, 1]$. Then, as $n_j \to \infty$, the following CLT holds:*

$$W(n_j) \xrightarrow{\mathrm{d}} Z(\eta) = \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix}, \tag{4.11}$$

*where $Z(\eta)$ is a centered Gaussian vector with covariance*

$$\Gamma = \Gamma(\eta) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{3+\eta}{8} & \frac{1-\eta}{8} & 0 \\ \frac{1}{2} & \frac{1-\eta}{8} & \frac{3+\eta}{8} & 0 \\ 0 & 0 & 0 & \frac{1-\eta}{8} \end{pmatrix}. \tag{4.12}$$

*The eigenvalues of $\Gamma$ are $0, \frac{3}{2}, \frac{1-\eta}{8}, \frac{1+\eta}{4}$, in particular $\Gamma$ is singular.*

The fourth chaotic component $\mathcal{L}_n[4]$ is the dominating term in the series expansion (found in Proposition 4.2) of the total nodal length $\mathcal{L}_n$. Indeed, in Marinucci et al. (2016) it was proved that, as $\mathcal{N}_n \to +\infty$,

$$\mathrm{Var}(\mathcal{L}_n) \sim \mathrm{Var}(\mathcal{L}_n[4]), \tag{4.13}$$

by showing that the asymptotic variance of $\mathcal{L}_n[4]$ equals the right hand side of (2.7). Indeed,

$$\mathrm{Var}(\mathcal{L}_n[4]) = \frac{E_n}{512\mathcal{N}_n^2}\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right). \tag{4.14}$$

The asymptotic equality in (4.13) and the orthogonality properties of Wiener chaoses, that we have seen in Section 4, guarantee that, as $\mathcal{N}_n \to +\infty$,

$$\frac{\mathcal{L}_n - \mathbb{E}[\mathcal{L}_n]}{\sqrt{\mathrm{Var}(\mathcal{L}_n)}} = \frac{\mathcal{L}_n[4]}{\sqrt{\mathrm{Var}(\mathcal{L}_n[4])}} + o_{\mathbb{P}}(1), \tag{4.15}$$

where $o_{\mathbb{P}}(1)$ denotes a sequence converging to 0 in probability.

Moreover, thanks to Peccati and Rossi (2018, Lemma 2), which was proved using a powerful result by Bombieri and Bourgain (2015, Theorem 1) on the "sums of two squares" problem, we know that

$$\mathbb{E}\left[\left|\mathcal{L}_n - \mathcal{L}_n[4]\right|^2\right] = O\left(\frac{E_n}{\mathcal{N}_n^{5/2}}\right). \tag{4.16}$$

## 5. Proofs

In the present section we give the proofs of our main results.

5.1. *Proof of Theorem 3.1.* From now on $\{n_j\}_j \subseteq S$ will denote a sequence of energy levels such that $\mathcal{N}_{n_j} \to +\infty$ and $\widehat{\mu}_{n_j}(4) \to \eta \in [-1,1]$ as $j \to +\infty$: for the sake of notation brevity we will write $n$ instead of $n_j$ in the sequel.

As anticipated in Section 3.1, the proof of Theorem 3.1 is divided into three steps. The first step is a MDP for a random vector, called $S_n$, whose components are linear combinations of independent and centered chi-square random variables, see Section 5.1.1. Then we will show that our functional of interest, which is the nodal length $\mathcal{L}_n$, is *exponentially equivalent* to a simpler functional that we will call $\mathcal{M}_n$, see Section 5.1.2. In the third and final step, see Section 5.1.3, we will prove through a contraction principle that $\mathcal{M}_n$ (and hence $\mathcal{L}_n$), which is a continuous function of $S_n$, satisfies a MDP with rate function as in (3.2).

5.1.1. *MDP for $S_n$.* Recalling the content of Section 4, we define

$$S_n := \sum_{\lambda \in \Lambda_n^+} \left(|a_\lambda|^2 - 1\right) \begin{pmatrix} \lambda_1^2/n \\ \lambda_2^2/n \\ \lambda_1\lambda_2/n \end{pmatrix}; \tag{5.1}$$

in particular $S_n$ is a linear combination of independent and centered chi-square random variables, where the coefficients are three-dimensional. From Lemma 4.4, which is Marinucci et al. (2016, Lemma 4.3), we know that

$$\frac{S_n}{\sqrt{\mathcal{N}_n/2}} = \begin{pmatrix} W_2(n) \\ W_3(n) \\ W_4(n) \end{pmatrix} \xrightarrow{\mathrm{d}} \begin{pmatrix} Z_2 \\ Z_3 \\ Z_4 \end{pmatrix} \sim \mathcal{N}(0,\Sigma), \qquad \Sigma = \Sigma_\eta := \begin{pmatrix} \frac{3+\eta}{8} & \frac{1-\eta}{8} & 0 \\ \frac{1-\eta}{8} & \frac{3+\eta}{8} & 0 \\ 0 & 0 & \frac{1-\eta}{8} \end{pmatrix}. \tag{5.2}$$

Note that $\eta \in [-1,1]$, and $\det(\Sigma) \neq 0$ if and only if $\eta \in (-1,1)$. Note that $\widetilde{S}_n$ introduced in Section 3.1 coincides with $S_n/\sqrt{\mathcal{N}_n/2}$.

Thanks to the Gärtner-Ellis Theorem (see point (c) of Dembo and Zeitouni (1998, Theorem 2.3.6) and Theorem 2.7), in order to establish a MDP for $S_n$, it suffices to prove that, for suitable $\{\alpha_n\}_n$, there exists

$$\lim_{n\to\infty} \frac{1}{\alpha_n} \log \mathbb{E}\left[\exp\left(\alpha_n \left\langle \theta, \frac{S_n}{\sqrt{\alpha_n \mathcal{N}_n/2}}\right\rangle\right)\right] =: \psi(\theta)\,,$$

where $\theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ is a three-dimensional vector, and that

$$\psi(\theta) = \frac{1}{2}\langle \theta, \Sigma\theta\rangle\,. \tag{5.3}$$

This is an easy task and it is immediate to verify that, if

$$\alpha_n \longrightarrow +\infty \qquad \text{and} \qquad \frac{\alpha_n}{\mathcal{N}_n} \longrightarrow 0\,, \tag{5.4}$$

then the random vector

$$\left\{\alpha_n^{-1/2} \frac{S_n}{\sqrt{\mathcal{N}_n/2}}\right\}_n$$

satisfies a MDP with rate function

$$\psi^*(x) = \sup_{\theta\in\mathbb{R}^3} \{\langle \theta, x\rangle - \psi(\theta)\}\,, \quad x \in \mathbb{R}^3. \tag{5.5}$$

Indeed, setting for $\lambda \in \Lambda_n^+$,

$$b_n(\lambda) = \begin{pmatrix} \lambda_1^2/n \\ \lambda_2^2/n \\ \lambda_1\lambda_2/n \end{pmatrix},$$

one has that

$$\begin{aligned}&\frac{1}{\alpha_n} \log \mathbb{E}\left[\exp\left(\alpha_n \left\langle \theta, \frac{S_n}{\sqrt{\alpha_n \mathcal{N}_n/2}}\right\rangle\right)\right] \\ &= \frac{1}{\alpha_n} \sum_{\lambda\in\Lambda_{n_j}^+} \left\{\frac{1}{2}\frac{\langle \theta, b_n(\lambda)\rangle^2}{\mathcal{N}_n/(2\alpha_n)} + \frac{1}{3}\frac{\langle \theta, b_n(\lambda)\rangle^3}{(\mathcal{N}_n/(2\alpha_n))^{3/2}} + o\left(\frac{\langle \theta, b_n(\lambda)\rangle^3}{(\mathcal{N}_n/(2\alpha_n))^{3/2}}\right)\right\},\end{aligned}$$

where we used the fact that $2{\cdot}|a_\lambda|^2$ is distributed as a $\chi^2(2)$ random variable (recall that $a_\lambda = b_\lambda + ic_\lambda$, where $b_\lambda$ and $c_\lambda$ are iid $\sim \mathcal{N}(0, 1/2)$). So our goal becomes proving that

$$\lim_{j\to+\infty} \frac{1}{\mathcal{N}_{n_j}/2} \sum_{\lambda\in\Lambda_{n_j}^+} \langle \theta, b_{n_j}(\lambda)\rangle^2 = \langle \theta, \Sigma\theta\rangle\,, \tag{5.6}$$

and that

$$\frac{1}{\alpha_n} \sum_{\lambda\in\Lambda_{n_j}^+} \frac{\langle \theta, b_n(\lambda)\rangle^3}{(\mathcal{N}_n/\alpha_n)^{3/2}} \longrightarrow 0\,. \tag{5.7}$$

As far as (5.6) is concerned, given the fact that for every $n \in S$ (see Marinucci et al. (2016, Lemma 4.1))

$$\frac{1}{n^2\mathcal{N}_n} \sum_{\lambda\in\Lambda_n} \lambda_1^4 = \frac{1}{n^2\mathcal{N}_n} \sum_{\lambda\in\Lambda_n} \lambda_2^4 = \frac{3+\widehat{\mu}_n(4)}{8} \quad \text{and} \quad \frac{1}{n^2\mathcal{N}_n} \sum_{\lambda\in\Lambda_n} \lambda_1^2\lambda_2^2 = \frac{1-\widehat{\mu}_n(4)}{8}\,,$$

we simply have that

$$\frac{1}{\mathcal{N}_n/2}\sum_{\lambda\in\Lambda_n^+}\langle\theta,b_n(\lambda)\rangle^2=\frac{1}{\mathcal{N}_n/2}\sum_{\lambda\in\Lambda_n^+}\left(\theta_1\frac{\lambda_1^2}{n}+\theta_2\frac{\lambda_2^2}{n}+\theta_3\frac{\lambda_1\lambda_2}{n}\right)^2$$
$$=\theta_1^2\frac{3+\widehat{\mu}_n(4)}{8}+\theta_2^2\frac{3+\widehat{\mu}_n(4)}{8}+\theta_3^2\frac{1-\widehat{\mu}_n(4)}{8}+2\theta_1\theta_2\frac{1-\widehat{\mu}_n(4)}{8}$$

since clearly $\sum_{\lambda\in\Lambda_n}\lambda_1\lambda_2^3=\sum_{\lambda\in\Lambda_n}\lambda_1^3\lambda_2=0$.

As a consequence we just proved that

$$\frac{1}{\mathcal{N}_n/2}\sum_{\lambda\in\Lambda_n^+}\langle\theta,b_n(\lambda)\rangle^2=\langle\theta,\Sigma_n\theta\rangle\,,$$

where

$$\Sigma_n=\begin{pmatrix}\frac{3+\widehat{\mu}_n(4)}{8} & \frac{1-\widehat{\mu}_n(4)}{8} & 0\\ \frac{1-\widehat{\mu}_n(4)}{8} & \frac{3+\widehat{\mu}_n(4)}{8} & 0\\ 0 & 0 & \frac{1-\widehat{\mu}_n(4)}{8}\end{pmatrix}.$$

Therefore, since $\lim_{j\to\infty}\widehat{\mu}_n(4)=\eta$, we obtain (5.6).

As for (5.7), we have

$$\frac{1}{\alpha_{n_j}}\sum_{\lambda\in\Lambda_{n_j}^+}\frac{|\langle\theta,b_{n_j}(\lambda)\rangle|^3}{(\mathcal{N}_{n_j}/\alpha_{n_j})^{3/2}}=\frac{1}{\mathcal{N}_{n_j}^{3/2}\alpha_{n_j}^{-1/2}}\sum_{\lambda\in\Lambda_{n_j}^+}|\langle\theta,b_{n_j}(\lambda)\rangle|^3\le c\frac{\|\theta\|^3}{\mathcal{N}_{n_j}^{1/2}\alpha_{n_j}^{-1/2}}$$

(for some absolute constant $c>0$), which tends to 0 as $j\to\infty$, since $\alpha_{n_j}$ is chosen as in (5.4).

5.1.2. *Exponential equivalence.* In light of (4.10) and (4.14), let us define, slightly abusing notation,

$$\mathcal{M}_n=\mathcal{M}_{\widehat{\mu}_n(4)}:=\left(1+\widehat{\mu}_n(4)^2+\frac{34}{\mathcal{N}_n}\right)^{-1/2}\left((W_2(n)+W_3(n))^2-2W_2(n)^2-2W_3(n)^2-4W_4(n)^2\right). \tag{5.8}$$

**Lemma 5.1.** *Consider as usual $\{n_j\}_j\subseteq S$ such that $\mathcal{N}_{n_j}\to+\infty$ as $j\to+\infty$ and assume that*

$$\alpha_{n_j}\longrightarrow+\infty\quad\text{and}\quad\frac{\alpha_{n_j}}{\log\mathcal{N}_{n_j}}\longrightarrow0\,,$$

*then, for every $\delta>0$,*

$$\limsup_{j\to+\infty}\frac{1}{\alpha_{n_j}}\log\mathbb{P}\left(\alpha_{n_j}^{-1}\left|\widetilde{\mathcal{L}}_{n_j}-\mathcal{M}_{n_j}\right|>\delta\right)=-\infty\,,$$

*i.e. $\alpha_{n_j}^{-1}\widetilde{\mathcal{L}}_{n_j}$ and $\alpha_{n_j}^{-1}\mathcal{M}_{n_j}$ are exponentially equivalent.*

*Proof*: From Lemma 4.3 and (4.14) we can write

$$\widetilde{\mathcal{L}}_n[4]=\mathcal{M}_n+\left(1+\widehat{\mu}_n(4)^2+\frac{34}{\mathcal{N}_n}\right)^{-1/2}R(n)$$

where we recall that

$$R(n)=\frac{1}{2}\frac{1}{\mathcal{N}_n/2}\sum_{\lambda\in\Lambda_n^+}|a_\lambda|^4$$

is a sequence of (independent) random variables converging in probability to 1 (note that $\mathbb{E}[R(n)] = 1$). We have, for $n$ large enough,

$$\mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_n - \mathcal{M}_n\right| > \delta\right) = \mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_n - \widetilde{\mathcal{L}}_n[4] + \left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)^{-1/2} R(n)\right| > \delta\right)$$
$$\leq \mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_n - \widetilde{\mathcal{L}}_n[4]\right| > \frac{\delta}{2}\right) + \mathbb{P}\left(\alpha_n^{-1}\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)^{-1/2} |R(n) - 1| > \frac{\delta}{4}\right)$$

since $\alpha_n^{-1}\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)^{-1/2} \longrightarrow 0$ as $\mathcal{N}_n \longrightarrow +\infty$. As for the first summand, thanks to Markov inequality, we can write

$$\mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_n - \mathcal{M}_n\right| > \delta/2\right) \leq \frac{\alpha_n^{-1}}{\delta/2}\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_n - \mathcal{M}_n\right|\right]$$
$$\leq \frac{\alpha_n^{-1}}{\delta/2}\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_n - \widetilde{\mathcal{L}}_n[4]\right|^2\right]^{\frac{1}{2}}$$

and from (4.16) we have

$$\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_n - \widetilde{\mathcal{L}}_n[4]\right|^2\right] = O\left(\mathcal{N}_n^{-1/2}\right).$$

Since $\alpha_n/\log\mathcal{N}_n \longrightarrow 0$ we deduce

$$\limsup_{n\to+\infty}\frac{1}{\alpha_n}\log\mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_n - \widehat{\mathcal{M}}_n\right| > \delta/2\right) \leq \limsup_{n\to+\infty}\frac{1}{\alpha_n}\log\frac{\alpha_n^{-1}}{\delta/2}\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_n - \mathcal{L}_n[4]\right|^2\right]^{\frac{1}{2}} = -\infty.$$

Regarding the second summand, from Peccati and Rossi (2018, (85)),

$$\mathbb{P}\left(|R(n) - 1| > \frac{\delta}{4}\alpha_n\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)^{1/2}\right) \leq \frac{\mathrm{Var}(R(n))}{\alpha_n^2\delta^2/(16)\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)}$$
$$\ll \frac{1}{\alpha_n^2\mathcal{N}_n}$$

so that, since $\alpha_n/\log\mathcal{N}_n \longrightarrow 0$

$$\limsup_{n\longrightarrow+\infty}\frac{1}{n}\log\mathbb{P}\left(|R(n) - 1| > \frac{\delta}{4}\alpha_n\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right)^{1/2}\right) = -\infty$$

thus concluding the proof of the exponential equivalence. □

5.1.3. *Contraction principle.* In light of the exponential equivalence obtained in the previous section, in order to obtain a MDP for the nodal length of arithmetic random waves, it suffices to prove the following.

**Lemma 5.2.** *The sequence of random variables $\{\alpha_n^{-1}\mathcal{M}_n\}_n$, where $\mathcal{N}_n \longrightarrow +\infty$ such that $\widehat{\mu}_n(4) \longrightarrow \eta \in [-1, 1]$, $\mathcal{M}_n$ is defined as in (5.8), and $\{\alpha_n\}_n$ is a sequence of positive numbers satisfying (5.4), enjoys a MDP with speed $\alpha_n$ and rate function*

$$I_\eta(y) = \begin{cases} -y\frac{\sqrt{1+\eta^2}}{1+|\eta|}, & y \leq 0 \\ +\infty, & y > 0. \end{cases} \tag{5.9}$$

*Proof*: First of all we note that

$$\alpha_n^{-1}\mathcal{M}_n = \left(1+\widehat{\mu}_n(4)^2+\tfrac{34}{\mathcal{N}_n}\right)^{-1/2} f\left(\tfrac{S_n}{\sqrt{\alpha_n \mathcal{N}_n/2}}\right)$$
$$= \left(1+\widehat{\mu}_n(4)^2+\tfrac{34}{\mathcal{N}_n}\right)^{-1/2} f\left(\alpha_n^{-1/2}\left(W_2(n), W_3(n), W_4(n)\right)'\right), \tag{5.10}$$

where $S_n$ is defined as in (5.1) and $f:\mathbb{R}^3 \longrightarrow \mathbb{R}$ is given by

$$f(x_1,x_2,x_3) = (x_1+x_2)^2 - 2x_1^2 - 2x_2^2 - 4x_3^2$$
$$= -(x_1-x_2)^2 - 4x_3^2 .$$

Thanks to the contraction principle (see Dembo and Zeitouni (1998, Theorem 4.2.1) and Theorem 2.8), the sequence of random variables

$$\left\{ f\left(\frac{S_n}{\sqrt{\alpha_n \mathcal{N}_n/2}}\right)\right\}_n$$

satisfies a LDP with speed $\alpha_n$ and rate function

$$I_f(y) = \inf\left\{\psi^*(x):\ f(x)=y\right\}, \quad y\in\mathbb{R},$$

where $\psi^*$ is defined as in (5.5). Since $f$ takes non-positive values, $I_f(y)=+\infty$ whenever $y>0$ and $I_f(0)=0$ since $\psi^*(x)=0$ if and only if $x=0$ and $f(0)=0$.

We are going to explicitly compute $I_f(y)$ for $y<0$. Assume that $\eta\in(-1,1)$, then the matrix $\Sigma$ in (5.2) is non-singular and

$$I_f(y) = \inf\left\{\psi^*(x):\ f(x)=y\right\} = \inf\left\{\frac{1}{2}\langle x, \Sigma^{-1}x\rangle:\ f(x)=y\right\}.$$

Now

$$\Sigma^{-1} = \frac{8}{1+\eta}\begin{pmatrix} \frac{3+\eta}{8} & \frac{\eta-1}{8} & 0 \\ \frac{\eta-1}{8} & \frac{3+\eta}{8} & 0 \\ 0 & 0 & \frac{1+\eta}{1-\eta}\end{pmatrix},$$

and therefore

$$\langle x, \Sigma^{-1}x\rangle = \frac{8}{1+\eta}\left(\frac{3+\eta}{8}(x_1^2+x_2^2) + \frac{\eta-1}{4}x_1x_2 + \frac{1+\eta}{1-\eta}x_3^2\right).$$

As a consequence,

$$\psi^*(x) = \frac{1}{2}\langle x, \Sigma^{-1}x\rangle = \frac{4}{1+\eta}\left(\frac{3+\eta}{8}(x_1^2+x_2^2) + \frac{\eta-1}{4}x_1x_2 + \frac{1+\eta}{1-\eta}x_3^2\right)$$

and

$$I_f(y) = \inf\left\{\frac{4}{1+\eta}\left(\frac{3+\eta}{8}(x_1^2+x_2^2) + \frac{\eta-1}{4}x_1x_2 + \frac{1+\eta}{1-\eta}x_3^2\right):\ -x_1^2-x_2^2+2x_1x_2-4x_3^2=y\right\}. \tag{5.11}$$

Let us now compute $I_f(y)$ for any $\eta\in(-1,1)$. We will use the Lagrange multipliers method with

$$L(x_1,x_2,x_3,\lambda) = h(x_1,x_2,x_3) + \lambda g(x_1,x_2,x_3)$$

where

$$h(x_1,x_2,x_3) = \frac{3+\eta}{2(1+\eta)}(x_1^2+x_2^2) + \frac{(\eta-1)}{1+\eta}x_1x_2 + \frac{4}{1-\eta}x_3^2$$
$$g(x_1,x_2,x_3) = -x_1^2 - x_2^2 + 2x_1x_2 - 4x_3^2 - y .$$

Then, recalling that we take $y<0$, we have the system

$$\begin{cases} h_{x_1}+\lambda g_{x_1}=0\\ h_{x_2}+\lambda g_{x_2}=0\\ h_{x_3}+\lambda g_{x_3}=0\\ g(x_1,x_2,x_3)=0\,, \end{cases} \quad \text{which yields} \quad \begin{cases} \lambda=\dfrac{1}{1+\eta}\\ x_1=\mp\sqrt{-\dfrac{y}{4}}\\ x_2=\pm\sqrt{-\dfrac{y}{4}}\\ x_3=0 \end{cases} \quad \text{or} \quad \begin{cases} \lambda=\dfrac{1}{1-\eta}\\ x_1=0\\ x_2=0\\ x_3=\pm\sqrt{-\dfrac{y}{4}}\,. \end{cases}$$

For the first two solutions we have that

$$\psi^*\left(\sqrt{-\frac{y}{4}},-\sqrt{-\frac{y}{4}},0\right)=\psi^*\left(-\sqrt{-\frac{y}{4}},\sqrt{-\frac{y}{4}},0\right)=\frac{-y}{1+\eta}\,,$$

while for the other two solutions we have that

$$\psi^*\left(0,0,\pm\sqrt{-\frac{y}{4}}\right)=\frac{-y}{1-\eta}\,.$$

As a consequence, the rate function is given by

$$I_f(y)=\frac{-y}{1+|\eta|}\,,\qquad \text{for}\quad y\le 0\quad \text{and}\quad \eta\in(-1,1)\,.$$

If $\eta\in\{-1,1\}$, then $\Sigma$ is singular. In this case, in order to compute $\psi^*(x)$ one has to consider the gradient with respect to $\theta$ to obtain the supremum in (5.5) and it is a known fact that one obtains $\Sigma\theta=x$. Then there are two possibilities:

- if $x\notin \mathrm{Im}(\Sigma)$, then $\Sigma\theta=x$ has no solution, and therefore $\psi^*(x)=+\infty$;
- if $x\in \mathrm{Im}(\Sigma)$, then $\Sigma\theta=x$ has solution $\theta=\widetilde{\Sigma}^{-1}x$ where $\widetilde{\Sigma}^{-1}$ is the inverse of $\Sigma$ restricted to $\mathrm{Im}(\Sigma)$, and therefore $\psi^*(x)=\left[\langle\theta,x\rangle-\frac{1}{2}\langle\theta,\Sigma\theta\rangle\right]_{\theta=\widetilde{\Sigma}^{-1}x}=\frac{1}{2}\langle x,\widetilde{\Sigma}^{-1}x\rangle$.

Then, considering the case when $\eta=1$, we have:

$$\Sigma=\begin{pmatrix} \frac12 & 0 & 0\\ 0 & \frac12 & 0\\ 0 & 0 & 0 \end{pmatrix}$$

and hence $\mathrm{Im}(\Sigma)=\{(x_1,x_2,x_3):\ x_3=0\}$. As a consequence, for $x\in\mathrm{Im}(\Sigma)$, we have that

$$\begin{pmatrix} \frac12 & 0 & 0\\ 0 & \frac12 & 0\\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix}\theta_1\\ \theta_2\\ \theta_3\end{pmatrix}=\begin{pmatrix}x_1\\ x_2\\ 0\end{pmatrix},\ \text{which yields}\ \ \theta=\begin{pmatrix} 2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 0 \end{pmatrix}x,$$

and hence

$$\psi^*(x)=\frac12\langle x,\widetilde{\Sigma}^{-1}x\rangle=x_1^2+x_2^2.$$

So what we have to compute now is the following rate function

$$I_f(y)=\inf_{x\in\mathrm{Im}(\Sigma)}\left\{\frac12\langle x,\widetilde{\Sigma}^{-1}x\rangle:\ f(x)=y\right\}=\inf_{x\in\mathrm{Im}(\Sigma)}\left\{x_1^2+x_2^2:\ -x_1^2-x_2^2+2x_1x_2=y\right\}.$$

We use again the Lagrange method. We have the system

$$\begin{cases} 2x_1+\lambda(-2x_1+2x_2)=0\\ 2x_2+\lambda(-2x_2+2x_1)=0,\\ -x_1^2-x_2^2+2x_1x_2-y=0 \end{cases}\quad \text{which yields}\ \begin{cases} x_1=\pm\sqrt{-\frac{y}{4}}\\ x_2=\mp\sqrt{-\frac{y}{4}}\\ \lambda=\frac12 \end{cases}\quad \text{and}\ \psi^*\left(\pm\sqrt{-\tfrac{y}{4}},\mp\sqrt{-\tfrac{y}{4}},0\right)=-\frac{y}{2}\,.$$

Thus, for $\eta=1$, the rate function is $I_f(y)=-y/2$.

Let us now consider the case $\eta = -1$. We have

$$\Sigma = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}$$

and hence $\mathrm{Im}(\Sigma) = \{(x_1, x_2, x_3) : x_1 = x_2\}$. As a consequence, for $x \in \mathrm{Im}(\Sigma)$,

$$\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \\ x_3 \end{pmatrix}, \quad \text{which yields} \begin{cases} \frac{1}{4}\theta_1 + \frac{1}{4}\theta_2 = x_1 \\ \frac{1}{4}\theta_1 + \frac{1}{4}\theta_2 = x_1 \\ \frac{1}{4}\theta_3 = x_3 \end{cases}, \quad \text{which yields} \begin{cases} \theta_1 + \theta_2 = 4x_1 \\ \theta_1 + \theta_2 = 4x_1 \\ \theta_3 = 4x_3 \end{cases}$$

and therefore

$$\psi^*(x_1, x_1, x_3) = \left[ \langle \theta, x \rangle - \frac{1}{2}\langle \theta, \Sigma\theta \rangle \right]_{\Sigma\theta = x, x_1 = x_2} = 2x_1^2 + 2x_3^2.$$

So what we have to compute now is the following rate function

$$I_f(y) = \inf_{x \in \mathrm{Im}(\Sigma)} \left\{ 2x_1^2 + 2x_3^2 : -4x_3^2 = y \right\}.$$

We use again the Lagrange method. We have the system

$$\begin{cases} 4x_1 + \lambda \cdot 0 = 0 \\ 4x_3 + \lambda(-8x_3) = 0 \\ -4x_3^2 = y\,, \end{cases} \quad \text{which yields} \begin{cases} x_1 = 0 \\ x_3 = \pm\sqrt{-\frac{y}{4}} \\ \lambda = \frac{1}{2} \end{cases} \quad \text{and} \quad \psi^*\left(0, 0, \pm\sqrt{-\frac{y}{4}}\right) = -\frac{y}{2}\,.$$

Thus, for $\eta = -1$, the rate function is $I_f(y) = -y/2$.

Finally, recalling (5.10), since as $n \to +\infty$

$$\left(1 + \widehat{\mu}_n(4)^2 + \frac{34}{\mathcal{N}_n}\right) \longrightarrow 1 + \eta^2,$$

one can prove the desired LDP of $\{\alpha_n^{-1}\mathcal{M}_n\}_n$, with the rate function in (5.9), with some standard computations (in particular, for the proof of the lower bound for open sets, one can use Lemma 19 in Ganesh and Torrisi (2008)). □

*Proof of Theorem 3.1:* Bearing in mind the result of Sections 5.1.1–5.1.3, the proof of Theorem 3.1 immediately follows. Indeed, from Lemma 5.1 the two sequences of random variables $\left\{\widetilde{\mathcal{L}}_n/\alpha_n\right\}_n$ and $\{\mathcal{M}_n/\alpha_n\}_n$ are exponential equivalent at speed $\alpha_n$. From Theorem 2.10 and Lemma 5.2, $\left\{\widetilde{\mathcal{L}}_n/\alpha_n\right\}_n$ enjoys a MDP with speed $\alpha_n$ and rate function (5.9). □

## 5.2. *Proof of Theorem 3.2.*

*Proof*: The proof consists in showing that $\alpha_n^{-1}\widetilde{\mathcal{L}}_{n;s}$ and $\alpha_n^{-1}\widetilde{\mathcal{L}}_n$ are exponentially equivalent at speed $\alpha_n$: this guarantees that they enjoy the same MDP as in Theorem 3.1. In order to do that, it suffices to find an upper bound for

$$\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_{n;s} - \widetilde{\mathcal{L}}_n\right|^2\right] = 1 + 1 - 2\,\mathrm{Corr}(\mathcal{L}_{n;s}, \mathcal{L}_n).$$

Recalling (2.7), (2.14) and the precise identity Benatar et al. (2020, (3.36)), we can write, uniformly for $s > n^{-\frac{1}{2}+\varepsilon}$,

$$\begin{aligned}\mathrm{Corr}(\mathcal{L}_{n;s}, \mathcal{L}_n) &= \frac{\mathrm{Cov}(\mathcal{L}_{n;s}, \mathcal{L}_n)}{\sqrt{\mathrm{Var}\left(\mathcal{L}_{n;s}\right)\mathrm{Var}\left(\mathcal{L}_n\right)}} = \frac{\pi s^2 \mathrm{Var}\left(\mathcal{L}_n\right)}{\sqrt{\mathrm{Var}\left(\mathcal{L}_{n;s}\right)\mathrm{Var}\left(\mathcal{L}_n\right)}} \\ &= \frac{\pi s^2 \sqrt{\mathrm{Var}\left(\mathcal{L}_n\right)}}{\sqrt{\mathrm{Var}\left(\mathcal{L}_{n;s}\right)}} = \frac{\pi s^2 \sqrt{\frac{E_n}{\mathcal{N}_n^2} c_n + O\left(\frac{E_n}{\mathcal{N}_n^{5/2}}\right)}}{\sqrt{(\pi s^2)^2 \left(\frac{E_n}{\mathcal{N}_n^2} c_n + O_\varepsilon\left(\frac{E_n}{\mathcal{N}_n^{5/2}}\right)\right)}} \\ &= 1 + O_\varepsilon\left(\mathcal{N}_n^{-1/2}\right),\end{aligned}$$

where the constant involved in the $O$-notation only depends on $\varepsilon$. As a consequence, we have that

$$\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_{n;s} - \widetilde{\mathcal{L}}_n\right|^2\right] = O_\varepsilon\left(\mathcal{N}_n^{-1/4}\right). \tag{5.12}$$

Now, for every $\delta > 0$, thanks to (5.12) we have

$$\limsup_{n\to+\infty} \frac{1}{\alpha_n} \log \mathbb{P}\left(\alpha_n^{-1}\left|\widetilde{\mathcal{L}}_{n;s} - \widetilde{\mathcal{L}}_n\right| > \delta\right) \le \limsup_{n\to+\infty} \frac{1}{\alpha_n} \log \frac{\alpha_n^{-2}}{\delta^2}\mathbb{E}\left[\left|\widetilde{\mathcal{L}}_{n;s} - \widetilde{\mathcal{L}}_n\right|^2\right] = -\infty$$

and the proof is hence concluded. □

## Acknowledgements

The authors wish to thank the two anonymous referees for their suggestions that improved the presentation of this paper.

## References

Adler, R. J. and Taylor, J. E. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York (2007). ISBN 978-0-387-48112-8. MR2319516.

Benatar, J., Marinucci, D., and Wigman, I. Planck-scale distribution of nodal length of arithmetic random waves. *J. Anal. Math.*, **141** (2), 707–749 (2020). MR4179775.

Bérard, P. Volume des ensembles nodaux des fonctions propres du laplacien. In *Bony-Sjöstrand-Meyer seminar, 1984–1985*, pp. Exp. No. 14 , 10. École Polytech., Palaiseau (1985). MR819780.

Berry, M. V. Regular and irregular semiclassical wavefunctions. *J. Phys. A*, **10** (12), 2083–2091 (1977). MR489542.

Berry, M. V. Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature. *J. Phys. A*, **35** (13), 3025–3038 (2002). MR1913853.

Bombieri, E. and Bourgain, J. A problem on sums of two squares. *Int. Math. Res. Not. IMRN*, (11), 3343–3407 (2015). MR3373053.

Brüning, J. Über knoten von eigenfunktionen des Laplace-Beltrami-operators. *Math. Z.*, (158), 15–21 (1978). DOI: 10.1007/BF01214561.

Cammarota, V. Nodal area distribution for arithmetic random waves. *Trans. Amer. Math. Soc.*, **372** (5), 3539–3564 (2019). MR3988618.

Caponera, A., Marinucci, D., and Vidotto, A. Multiscale CUSUM tests for time-dependent spherical random fields. *ArXiv Mathematics e-prints* (2023). arXiv: 2305.01392.

Christakos, G. *Random field models in earth sciences*. Academic Press, Inc. (1992). ISBN 978-0-12-174230-0. DOI: 10.1016/C2009-0-22238-0.

Cilleruelo, J. The distribution of the lattice points on circles. *J. Number Theory*, **43** (2), 198–202 (1993). MR1207499.

Dembo, A. and Zeitouni, O. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition (1998). ISBN 0-387-98406-2. MR1619036.

Donnelly, H. and Fefferman, C. Nodal sets of eigenfunctions on Riemannian manifolds. *Invent. Math.*, **93** (1), 161–183 (1988). MR943927.

Erdős, P. and Hall, R. R. On the angular distribution of Gaussian integers with fixed norm. *Discrete Math.*, **200** (1-3), 87–94 (1999).

Ganesh, A. J. and Torrisi, G. L. Large deviations of the interference in a wireless communication model. *IEEE Trans. Inform. Theory*, **54** (8), 3505–3517 (2008). MR2451010.

Giuliano, R. and Macci, C. Some examples of noncentral moderate deviations for sequences of real random variables. *Mod. Stoch. Theory Appl.*, **10** (2), 111–144 (2023). MR4573675.

Krishnapur, M., Kurlberg, P., and Wigman, I. Nodal length fluctuations for arithmetic random waves. *Ann. of Math. (2)*, **177** (2), 699–737 (2013). MR3010810.

Kurlberg, P. and Wigman, I. On probability measures arising from lattice points on circles. *Math. Ann.*, **367** (3-4), 1057–1098 (2017). MR3623219.

Landau, E. Uber die einteilung der positiven zahlen nach vier klassen nach der mindestzahl der zu ihrer addition zusammensetzung erforderlichen quadrate. *Archiv der Math. und Physik*, **13** (3), 305–312 (1908). Available at https://www.archive.org/details/archivdermathem37unkngoog.

Logunov, A. Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture. *Ann. of Math. (2)*, **187** (1), 241–262 (2018). MR3739232.

Macci, C., Rossi, M., and Todino, A. P. Moderate deviation estimates for nodal lengths of random spherical harmonics. *ALEA Lat. Am. J. Probab. Math. Stat.*, **18** (1), 249–263 (2021). MR4198876.

Maffucci, R. W. Nodal intersections for arithmetic random waves against a surface. *Ann. Henri Poincaré*, **20** (11), 3651–3691 (2019). MR4019200.

Marinucci, D. and Peccati, G. *Random fields on the sphere. Representation, limit theorems and cosmological applications*, volume 389 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge (2011). ISBN 978-0-521-17561-6. MR2840154.

Marinucci, D., Peccati, G., Rossi, M., and Wigman, I. Non-universality of nodal length distribution for arithmetic random waves. *Geom. Funct. Anal.*, **26** (3), 926–960 (2016). MR3540457.

Nourdin, I. and Peccati, G. *Normal approximations with Malliavin calculus. From Stein's method to universality*, volume 192 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge (2012). ISBN 978-1-107-01777-1. MR2962301.

Peccati, G. and Rossi, M. Quantitative limit theorems for local functionals of arithmetic random waves. In *Computation and combinatorics in dynamics, stochastics and control*, volume 13 of *Abel Symp.*, pp. 659–689. Springer, Cham (2018). MR3967400.

Rudnick, Z. and Wigman, I. On the volume of nodal sets for eigenfunctions of the Laplacian on the torus. *Ann. Henri Poincaré*, **9** (1), 109–130 (2008). MR2389892.

Sartori, A. On the fractal structure of attainable probability measures. *Bull. Pol. Acad. Sci. Math.*, **66** (2), 123–133 (2018). MR3892756.

Schulte, M. and Thäle, C. Cumulants on Wiener chaos: moderate deviations and the fourth moment theorem. *J. Funct. Anal.*, **270** (6), 2223–2248 (2016). MR3460239.

Taylor, J. E. and Worsley, K. J. Detecting sparse signals in random fields, with an application to brain mapping. *J. Amer. Statist. Assoc.*, **102** (479), 913–928 (2007). MR2354405.

Todino, A. P. Nodal lengths in shrinking domains for random eigenfunctions on $S^2$. *Bernoulli*, **26** (4), 3081–3110 (2020). MR4140538.

Wigman, I. Fluctuations of the nodal length of random spherical harmonics. *Comm. Math. Phys.*, **298** (3), 787–831 (2010). MR2670928.

Yau, S. T., editor. *Seminar on Differential Geometry.* Annals of Mathematics Studies, No. 102. Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo (1982). ISBN 0-691-08268-5. MR645728.