



Functional strong laws of large numbers for Euler characteristic processes of extreme sample clouds

Andrew M. Thomas¹ · Takashi Owada¹

Received: 19 October 2020 / Revised: 14 February 2021 / Accepted: 28 April 2021 /

Published online: 20 May 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

To recover the topology of a manifold in the presence of heavy tailed or exponentially decaying noise, one must understand the behavior of geometric complexes whose points lie in the tail of these noise distributions. This study advances this line of inquiry, and demonstrates functional strong laws of large numbers for the Euler characteristic process of random geometric complexes formed by random points outside of an expanding ball in $\mathbb{R}^d$. When the points are drawn from a heavy tailed distribution with a regularly varying tail, the Euler characteristic process grows at a regularly varying rate, and the scaled process converges uniformly and almost surely to a smooth function. When the points are drawn from a distribution with an exponentially decaying tail, the Euler characteristic process grows logarithmically, and the scaled process converges to another smooth function in the same sense. All of the limit theorems take place when the points inside the expanding ball are densely distributed, so that the simplex counts outside of the ball of all dimensions contribute to the Euler characteristic process.

Keywords Functional strong law of large numbers · Euler characteristic · Random geometric complex · Topological crackle

AMS 2000 Subject Classification 60F15 · 60G70 · 55U10 · 60F17

✉ Andrew M. Thomas
thoma186@purdue.edu

Takashi Owada
owada@purdue.edu

¹ Department of Statistics, Purdue University, 47907, West Lafayette, IN, USA

1 Introduction

1.1 Heavy tailed noise and annuli structure of homological elements

To recover the topology of a manifold using point cloud data, one needs to have a strong understanding of how the points are perturbed from the manifold. In Niyogi et al. (2008), given a “nice” manifold, it was shown that one can recover the topology of the manifold by a sufficiently dense random sampling of points if the noise is bounded. In Niyogi et al. (2011) it was shown that the recovery is still possible by a sufficiently dense random sampling of points if the noise is standard multivariate Gaussian and the variance is bounded by a function of the reach and dimension of the manifold. However, if the points on the manifold are perturbed by heavy tailed noise, the recovery of topology will be severely impacted because of extraneous homological elements generated by this noise. In Fig. 1, we wish to recover the topology of the circle S^1 from the union of balls around a random point cloud. In Fig. 1b and c, the union of balls recover the essential shape of S^1 , as the size of noise in these cases is sufficiently small. However, in Fig. 1d the noise added to points in S^1 has a heavy tailed Cauchy distribution. Consequently, three extraneous shape elements appear — two distinct components and a tiny one-dimensional cycle. This phenomenon in case (d) raises the question of how the shape of these elements away from the center of S^1 may behave in general; this is roughly the idea of what is called *topological crackle*.

From a more analytic viewpoint, topological crackle is understood as a layered annuli structure of homological elements of different orders. To make our implication more clear, we consider the power-law density,

$$f(x) = \frac{C}{1 + \|x\|^\alpha}, \quad x \in \mathbb{R}^d, \quad (1.1)$$

for some $\alpha > d$ and a normalizing constant $C > 0$. Suppose a random point cloud $\mathcal{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$, $d \geq 2$, is drawn from this density. Let $\text{Ann}(K, L)$ be a closed annulus with inner radius K and outer radius L , and $B(x, t) := \{y \in \mathbb{R}^d : \|y - x\| < t\}$ be an open ball of radius t around $x \in \mathbb{R}^d$ (here $\|\cdot\|$ denotes the Euclidean norm). Then, one can divide $\mathbb{R}^d$ in a way that

$$\mathbb{R}^d = \bigcup_{k=0}^{d+1} \text{Ann}(R_{k,n}, R_{k-1,n}), \quad (1.2)$$

where

$$R_{k,n} = \begin{cases} \infty & k = -1, \\ (Cn)^{\frac{1}{\alpha-d}} & k = 0, \\ (Cn)^{\frac{1}{\alpha-d/(k+2)}} & k \in \{1, \dots, d-1\}, \\ (Cn)^{1/\alpha} & k = d, \\ 0 & k = d+1, \end{cases} \quad (1.3)$$

so that

$$R_{d,n} \ll R_{d-1,n} \ll \dots \ll R_{1,n} \ll R_{0,n}, \quad \text{as } n \rightarrow \infty.$$

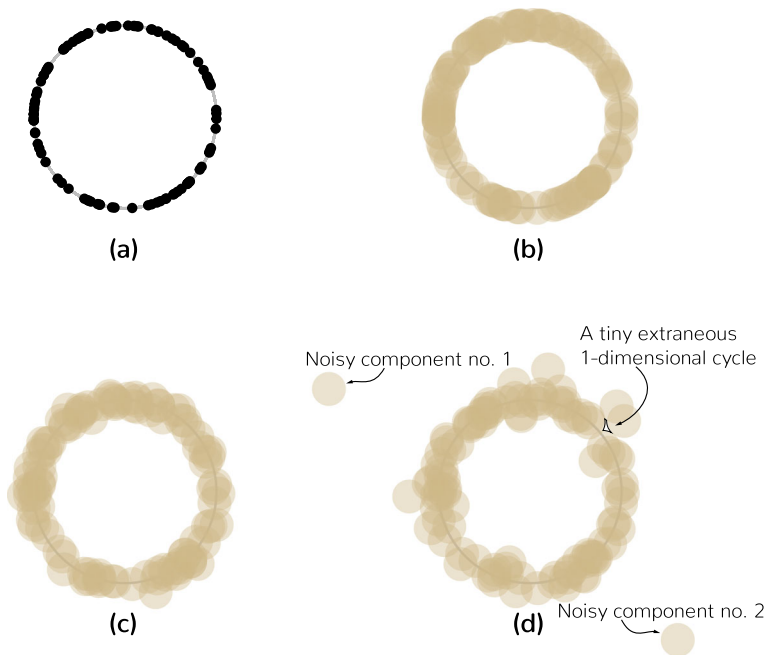


Fig. 1 **a** A random sample $\mathcal{X}_{100} = \{X_1, \dots, X_{100}\}$ of 100 points, uniformly distributed on the unit circle S^1 . **b** The union of balls of radius 0.2 around each point in $\mathcal{X}_{100}$. **c** The union of balls when $\mathcal{X}_{100}$ is perturbed by Gaussian noise. **d** The union of balls when $\mathcal{X}_{100}$ is perturbed by Cauchy noise. Noise is applied in the same manner as Niyogi et al. (2011). Topology is recovered by the union of the balls in cases **b** and **c**, but that is not the case if the heavy tailed Cauchy noise is added

From previous studies (Adler et al. 2014; Owada and Adler 2017; Owada 2018), it is known that the union of balls,

$$U(t) := \bigcup_{X \in \mathcal{X}_n} B(X, t), \quad t \geq 0,$$

asymptotically generates homological elements (i.e., components and cycles as seen in Fig. 1d) in the following way: we have, as $n \rightarrow \infty$,

- Inside $\text{Ann}(R_{0,n}, \infty) = B(0, R_{0,n})^c$ there are finitely many distinct components, but none of the cycles of dimensions $1, 2, \dots, d-1$.
- Inside $\text{Ann}(R_{1,n}, R_{0,n})$ there are infinitely many distinct components and finitely many one-dimensional cycles, but none of the cycles of dimensions $2, 3, \dots, d-1$.

In general, for every $k \in \{2, \dots, d-1\}$,

- Inside $\text{Ann}(R_{k,n}, R_{k-1,n})$ there are infinitely many distinct components and cycles of dimensions $1, \dots, k-1$, and finitely many k -dimensional cycles, but none of the cycles of dimensions $k+1, \dots, d-1$,

and finally,

- Inside $\text{Ann}(R_{d,n}, R_{d-1,n})$ there are infinitely many distinct components and cycles of *all* dimensions $1, \dots, d-1$.

In the literature (Owada 2017), the innermost ball $B(0, R_{d,n})$ is referred to as a *weak core*. Inside the weak core, random points are densely scattered and the homology of the union of unit balls around them is nearly trivial as $n \rightarrow \infty$, i.e., the union has nearly no cycles of all dimensions $1, \dots, d-1$. With this layered structure in mind, topological crackle is formally defined as non-trivial homological elements (i.e., components and cycles) outside of a weak core.

Many of the existing studies in classical extreme value theory (EVT) have focused on the behavior of a random point cloud in the outermost annulus $\text{Ann}(R_{0,n}, \infty)$, or equivalently outside of $B(0, R_{0,n})$. For instance, it is well known that the total number of distinct components outside of $B(0, R_{0,n})$ converges weakly to a Poisson distribution as $n \rightarrow \infty$ (Resnick 1987; 2007; Embrechts et al. 1997). The objective of this paper is to go beyond the studies on the spatial distribution of components and investigate more complicated and higher-dimensional topological features outside of a weak core.

1.2 Topological crackle and Euler characteristic

After the pioneering paper of Adler et al. (2014), the stochastic properties of crackle phenomena have been investigated mostly via the behavior of *Betti numbers*. Loosely speaking, the k th Betti number counts the number of k -dimensional cycles—which can be interpreted as the boundary of a $(k+1)$ -dimensional body. In the related literature, Owada and Adler (2017) studied the case in which the k th Betti number of $U(t)$ outside of $B(0, R_{k,n})$ converges weakly to a Poisson distribution as $n \rightarrow \infty$. Moreover, Owada (2018) established the central limit theorem for the k th Betti numbers, in the case that infinitely many k -dimensional cycles appear outside of an expanding ball. Additionally, Owada (2019) gave a rigorous description of the limiting Betti numbers when the random points are generated by a classical moving average process, and Owada and Bobrowski (2020) discussed the weak convergence of a standard graphical representation of cycles.

In contrast to these previous papers, the objective of this paper is to examine the crackle phenomena from the viewpoint of the *Euler characteristic* of a geometric (simplicial) complex. Among many varieties of geometric complexes (see Ghrist 2014), the *Vietoris-Rips complex* and the *Čech complex* are specific examples that deserve our attention.

Definition 1.1 Given a point set $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$ and a positive number $t > 0$, the Vietoris-Rips complex $\mathcal{R}(\mathcal{X}, t)$ is defined as follows.

- The 0-simplices are the points in $\mathcal{X}$.
- A k -simplex $\sigma = [x_{i_0}, \dots, x_{i_k}]$ is in $\mathcal{R}(\mathcal{X}, t)$ if $\bar{B}(x_{i_p}, t/2) \cap \bar{B}(x_{i_q}, t/2) \neq \emptyset$ for every $0 \leq p < q \leq k$, where $\bar{B}(x, t)$ is the closure of $B(x, t)$.

Definition 1.2 Given the same $\mathcal{X}$ and $t > 0$, the Čech complex $\check{C}(\mathcal{X}, t)$ is defined as follows.

- The 0-simplices are the points in $\mathcal{X}$.
- A k -simplex $\sigma = [x_{i_0}, \dots, x_{i_k}]$ is in $\check{C}(\mathcal{X}, t)$ if a family of balls $\{\bar{B}(x_{i_p}, t/2), p = 0, \dots, k\}$ has a non-empty intersection.

In this paper, we examine two distinct scenarios of noise distributions that experience topological crackle: one where the distribution has a regularly-varying tail and another where the distribution has an exponentially decaying tail. We define a geometric complex that generalizes both $\mathcal{R}(\mathcal{X}, t)$ and $\check{C}(\mathcal{X}, t)$ above, and then establish the functional strong law of large numbers (FSLN) for the corresponding Euler characteristic process for each of the distributional contexts. One of the primary benefits of working with the Euler characteristic comes from the fact that it can be expressed as an alternating sum of Betti numbers of all dimensions (Edelsbrunner and Harer 2010). As a consequence, the Euler characteristic provides a limit theorem containing information on cycles of all different dimensions, whereas Betti numbers can only provide separate limit theorems for cycles of a particular dimension in each individual annulus region of Eq. 1.2. Such global features of the Euler characteristic help to capture the spatial distribution of cycles in the union of annuli $\bigcup_{k=0}^d \text{Ann}(R_{k,n}, R_{k-1,n}) = \mathbb{R}^d \setminus B(0, R_{d,n})$, even though the nature of the distribution of cycles differs from region to region.

In conjunction with the recent development of Topological Data Analysis (TDA), the literature dealing with the asymptotics of the Euler characteristic of random geometric complexes has flourished (Thomas and Owada 2021; Krebs et al. 2020; Bobrowski and Adler 2014; Bobrowski and Mukherjee 2015; Hug et al. 2016). However, none of these studies have paid sufficient attention to the topology of the tail of a probability distribution. In the context of topological crackle as in Fig. 1, ascertaining the topology of noise is an important step in determining how to process the signal of the manifold. In the light of the results in this paper, one can examine how the homology (i.e., components and cycles) of extreme-valued noise distributions evolve. This can be attained by viewing the Euler characteristic as a stochastic process, in which the parameter governing the formation of simplices is taken to be the “time” parameter. The resulting process then relates strongly to persistent homology. Persistent homology is a topological and algebraic structure that tracks the creation and destruction of cycles in different dimensions. It is one of the most widely used and robust tools in the TDA toolbox — see Adler et al. (2010) or Ghrist (2008) for an introduction to persistent homology and Edelsbrunner and Harer (2010) for a more thorough treatment. In particular, Examples 3.2 and 4.2 below provide the FSLNs in the different scenarios of noise distributions for the integrated Euler characteristic process. This process can be viewed as the Euler characteristic of a persistence barcode, which is a well known graphical descriptor of persistent homology (Ghrist 2008; Carlsson 2009).

1.3 Organization of the paper

The remainder of this paper is structured as follows. Section 2 provides a discussion of the background material necessary for this paper. The paper then proceeds to the heavy tailed setup and presents the FSLN for the Euler characteristic in Section 3. The paper continues with a discussion of the intricacies of the exponentially decaying tail case along with the corresponding FSLN in Section 4. The proofs of the main results for both setups are deferred to Section 5. From a technical point of view, the studies most relevant to this paper are Goel et al. (2019) and Thomas and Owada (2021), in which the authors established strong laws of large numbers for topological invariants — such as Betti numbers and the Euler characteristic — in the non-extreme value theoretic setup. In particular, these studies revealed that if the topological invariants are scaled proportionally to the sample size, they converge almost surely to a finite and positive constant. Owing to this fact, the main machinery in their proofs is a direct application of the Borel-Cantelli lemma, together with the calculation of lower-order moments. On the contrary, the main challenge in this paper is that the scaling sequence of the Euler characteristic may grow very slowly (e.g., logarithmically), in which case, a direct application of the Borel-Cantelli lemma does not work. To overcome this difficulty, we need to identify suitable subsequential upper and lower bounds of the Euler characteristic to which one can apply the Borel-Cantelli lemma. This is a standard technique in the theory of random geometric graphs — see Chapter 3 of the monograph of Penrose (2003). It is possible to extend these arguments to our higher-dimensional setup since the geometric complexes such as those in Definitions 1.1 and 1.2 are higher-dimensional analogues of a geometric graph.

As a final remark, we point out that the other types of limit theorems for the Euler characteristic still remain as a future topic. For example, it seems feasible to establish a (functional) central limit theorem for the Euler characteristic via Stein's method for normal approximation (see, e.g., Theorem 2.4 in Penrose 2003). Indeed, Owada (2018) already derived the central limit theorem for the Betti numbers by means of the aforementioned normal approximation technique. We anticipate that the same approach is possible for our Euler characteristic. A similar line of research in the non-extreme value theoretic setup can be found in Thomas and Owada (2021) and Krebs et al. (2020).

2 Preliminaries

The point cloud of interest in this study is the sample $\mathcal{X}_n := \{X_1, \dots, X_n\}$ of n i.i.d random points in $\mathbb{R}^d$, $d \geq 2$ with spherically symmetric density f . Spherical symmetry of f is far from necessary; the results in this paper could be extended to densities with ellipsoidal level sets fairly easily. Because of the imposed spherical symmetry, we can define $f(r) := f(r\theta)$ for all $r \geq 0$ and $\theta \in S^{d-1}$. Denote λ to be

Lebesgue measure on $\mathbb{R}^d$ and $S^{d-1} := \{x \in \mathbb{R}^d : \|x\| = 1\}$. Let us here define the spherical measure

$$\nu_{d-1}(A) := d \cdot \lambda(\{x \in B(0, 1) : x/\|x\| \in A\})$$

for Borel sets $A \subset S^{d-1}$. We denote $\omega_d := \lambda(B(0, 1)) = 2\pi^{d/2}/(d\Gamma(d/2))$, and $s_{d-1} := \nu_{d-1}(S^{d-1}) = d\omega_d$.

Let $\mathcal{F}(\mathbb{R}^d)$ be the collection of all non-empty, finite subsets of $\mathbb{R}^d$. For $\mathcal{X} \in \mathcal{F}(\mathbb{R}^d)$, a *simplicial complex* $\mathcal{K}(\mathcal{X})$ is a collection of subsets of $\mathcal{X}$ such that if $\sigma \in \mathcal{K}(\mathcal{X})$ and $\tau \subset \sigma$ then $\tau \in \mathcal{K}(\mathcal{X})$. Evidently, the Vietoris-Rips complex $\mathcal{R}(\mathcal{X}, t)$ and the Čech complex $\check{\mathcal{C}}(\mathcal{X}, t)$ satisfy this condition. We call $\sigma \in \mathcal{K}(\mathcal{X})$ a *k-simplex* if $|\sigma| = k + 1$.

Subsequently, let $h : \mathcal{F}(\mathbb{R}^d) \rightarrow \{0, 1\}$ be an indicator function satisfying the following conditions.

- H1** $h(\mathcal{X}) \leq h(\mathcal{Y})$ for all $\mathcal{Y} \subset \mathcal{X}$.
- H2** h is translation invariant — that is, for every $\mathcal{X} \in \mathcal{F}(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$, we have $h(\mathcal{X} + y) = h(\mathcal{X})$.
- H3** h is locally determined — that is, there exists $c > 0$ so that $h(\mathcal{X}) = 0$ whenever $\text{diam}(\mathcal{X}) > c$, where $\text{diam}(\mathcal{X}) := \max_{x, y \in \mathcal{X}} \|x - y\|$.

By abusing notation slightly, for $\mathcal{X} = \{x_1, \dots, x_m\} \in \mathcal{F}(\mathbb{R}^d)$, we write $h(\mathcal{X}) = h(x_1, \dots, x_m)$. Moreover, for $\mathcal{X} = \{x_1, \dots, x_m\}$ and $a \in \mathbb{R}$, we write $a\mathcal{X} = \{ax_1, \dots, ax_m\}$. We then define a scaled version of h by

$$h_t(\mathcal{X}) := \begin{cases} h(t^{-1}\mathcal{X}), & t > 0, \\ \mathbf{1}\{|\mathcal{X}| = 1\}, & t = 0, \end{cases}$$

with the additional assumption that

- H4** $t \mapsto h_t(\mathcal{X})$ is right continuous and non-decreasing for each $\mathcal{X} \in \mathcal{F}(\mathbb{R}^d)$.

Given such a scaled indicator h_t , we can construct the geometric (simplicial) complex

$$\mathcal{K}(\mathcal{X}, t) := \{\mathcal{Y} \subset \mathcal{X} : h_t(\mathcal{Y}) = 1\}. \quad (2.1)$$

By virtue of **(H4)** above, Eq. 2.1 induces a *filtration* of geometric complexes over a point set $\mathcal{X}$ — that is,

$$\mathcal{K}(\mathcal{X}, s) \subset \mathcal{K}(\mathcal{X}, t) \text{ for all } 0 \leq s \leq t.$$

Note that if one takes

$$h(\mathcal{X}) = \mathbf{1}\{\text{diam}(\mathcal{X}) \leq 1\}, \quad \mathcal{X} \in \mathcal{F}(\mathbb{R}^d),$$

then (2.1) induces a *Vietoris-Rips filtration*. Moreover, if we define

$$h(\mathcal{X}) = \mathbf{1}\left\{\bigcap_{x \in \mathcal{X}} \bar{B}(x, 1/2) \neq \emptyset\right\}, \quad \mathcal{X} \in \mathcal{F}(\mathbb{R}^d),$$

then (2.1) induces a *Čech filtration*.

As mentioned in the Introduction, the objective of this paper is to study “extreme-value” behavior of random geometric complexes via the Euler characteristic. More

concretely, with a non-random sequence $R_n \rightarrow \infty$, we study the filtration of geometric complexes

$$\mathcal{K}(\mathcal{X}_n \cap B(0, R_n)^c, t), \quad t \geq 0, \quad (2.2)$$

which are distributed increasingly further from the origin as $n \rightarrow \infty$. We now define the Euler characteristic pertaining to Eq. 2.2 by

$$\chi_n(t) := \sum_{k=0}^{\infty} (-1)^k S_{k,n}(t), \quad t \geq 0, \quad (2.3)$$

where $S_{k,n}(t)$ denotes the k -simplex counts in the complex (2.2). Namely,

$$S_{k,n}(t) := \sum_{\mathcal{Y} \subset \mathcal{X}_n, |\mathcal{Y}|=k+1} h_t(\mathcal{Y}) \mathbf{1} \left\{ \min_{y \in \mathcal{Y}} \|y\| \geq R_n \right\}.$$

Note that for every $n \in \mathbb{N} = \{1, 2, \dots\}$, Eq. 2.3 is a finite sum as $S_{k,n}(t) \equiv 0$ for all $k \geq n$. Furthermore, Eq. 2.3 can be seen as a stochastic process in parameter t , with right continuous sample paths and left limits. In the following, we establish the FSLLN for the *Euler characteristic process* $(\chi_n(t), t \geq 0)$ in the space $D[0, \infty)$ of right continuous functions on $[0, \infty)$ with left limits. In particular, we equip $D[0, \infty)$ with the uniform topology.

With the notation in Eq. 2.3, if we set $t = 0$,

$$\chi_n(0) = \sum_{k=0}^{\infty} (-1)^k S_{k,n}(0) = \sum_{i=1}^n \mathbf{1} \{\|X_i\| \geq R_n\} \quad (2.4)$$

represents the number of points outside of an expanding ball $B(0, R_n)$. The asymptotics of Eq. 2.4 can be treated in the standard framework of classical EVT (see, e.g., Resnick 1987, 2007; Embrechts et al. 1997). Unlike this special case, the Euler characteristic process (2.3) intrinsically involves higher-dimensional topological structures, which requires much more complicated machinery to analyze.

From the literature of topological crackle (see Adler et al. 2014; Owada 2017, 2018; Owada and Adler 2017; Owada and Bobrowski 2020), it is known that the behavior of topological invariants significantly depends on the limit value of $nf(R_n)$. The present study focuses exclusively on the case when the limit of $nf(R_n)$ is a positive and finite constant — that is,

$$nf(R_n) \rightarrow \xi \text{ as } n \rightarrow \infty \text{ for some } \xi \in (0, \infty). \quad (2.5)$$

As mentioned in the Introduction, the ball $B(0, R_n)$ with $nf(R_n) \rightarrow 1, n \rightarrow \infty$, is called a weak core. In the special case when the density has a power-law tail as in Eq. 1.1, the radius of a weak core is equal to $(Cn)^{1/\alpha}$ (see Eq. 1.3). Therefore, if R_n is determined by Eq. 2.5, $B(0, R_n)$ coincides with the weak core, up to multiplicative constants. The configuration of points between the outside and inside of a weak core is very different. Inside of the weak core, the homology of the union of balls becomes almost trivial as $n \rightarrow \infty$, i.e., the random points are very densely distributed and nearly every cycle of every dimension becomes filled in (see Adler et al. 2014; Owada 2017). Outside of the weak core, however, the random points are distributed more sparsely, though densely enough so that homology of all feasible dimensions

becomes not only non-trivial, but abundant. As a consequence, by an appropriate scaling as a function of R_n in Eq. 2.5, the simplex counts of all dimensions in Eq. 2.3 will contribute to the limit. In contrast, if R_n grows faster, such that $nf(R_n) \rightarrow 0$ as $n \rightarrow \infty$, then even under an appropriate scaling, the Euler characteristic is dominated asymptotically by the 0-simplices, or the extremal points. In this case, the Euler characteristic simply counts points in $\mathcal{X}_n$ outside $B(0, R_n)$ as $n \rightarrow \infty$.

3 Regularly varying tail case

In this section, we detail the large-sample behavior of Eq. 2.3 of an extreme sample cloud when the distribution of points has a heavy tail. Recall that $\mathcal{X}_n = \{X_1, \dots, X_n\}$ denotes a random sample in $\mathbb{R}^d$ with a spherically symmetric density f . We assume that there exists a tail index $\alpha > d$, such that

$$\lim_{r \rightarrow \infty} \frac{f(rt\theta)}{f(r\theta)} = t^{-\alpha}, \quad \forall t > 0, \quad (3.1)$$

for every (equivalently, some) $\theta \in S^{d-1}$.

Before stating the main result, note that by Eq. 2.5, one can typically take

$$R_n = \xi^{-1/\alpha} (1/f)^{\leftarrow}(n), \quad (3.2)$$

where $(1/f)^{\leftarrow}(x) = \inf\{y : 1/f(y) \geq x\}$ is the (left continuous) inverse of $1/f$. Thus, $(R_n)_{n \geq 1}$ is a regularly varying sequence of exponent (tail index) $1/\alpha$.

Theorem 3.1 *Suppose that f is a spherically symmetric density satisfying Eq. 3.1. Assume that $nf(R_n) \rightarrow \xi$ as $n \rightarrow \infty$ for some $\xi \in (0, \infty)$. Then, the Euler characteristic process in Eq. 2.3 satisfies the following functional SLLN, i.e., as $n \rightarrow \infty$,*

$$\left(\frac{\chi_n(t)}{R_n^d}, t \geq 0 \right) \rightarrow \left(\sum_{k=0}^{\infty} (-1)^k s_k(t), t \geq 0 \right), \quad \text{a.s. in } D[0, \infty). \quad (3.3)$$

where

$$s_k(t) := \frac{s_{d-1} \xi^{k+1}}{(k+1)! (\alpha(k+1) - d)} \int_{(\mathbb{R}^d)^k} h_t(0, y_1, \dots, y_k) dy, \quad t \geq 0, \quad k \geq 1, \quad (3.4)$$

with $s_0(t) \equiv s_{d-1} \xi / (\alpha - d)$. In particular, the limit in Eq. 3.3 is convergent for all $t \geq 0$.

The following example illustrates the uniform convergence that takes place in the above theorem.

Example 3.2 Consider the power-law density defined by

$$f(x) = \frac{2}{\pi \omega_d (1 + \|x\|^{2d})}, \quad x \in \mathbb{R}^d.$$

Define $R_n := (2n/\pi\omega_d)^{1/(2d)}$, so that $nf(R_n) \rightarrow 1$. We consider the Vietoris-Rips complex induced by $h(\mathcal{X}) = \mathbf{1}\{\text{diam}(\mathcal{X}) \leq 1/\sqrt{d}\}$, where diam is calculated here with respect to the ℓ^∞ norm. Then, it follows from Theorem 3.1 that, as $n \rightarrow \infty$,

$$\left(\sqrt{\frac{\pi\omega_d}{2n}}\chi_n(t), t \geq 0\right) \rightarrow \left(\sum_{k=0}^{\infty}(-1)^k s_k(t), t \geq 0\right), \text{ a.s. in } D[0, \infty).$$

The limiting function above can be simplified as follows:

$$\begin{aligned} \sum_{k=0}^{\infty}(-1)^k s_k(t) &= s_{d-1} \sum_{k=0}^{\infty} \frac{(-1)^k (t/\sqrt{d})^{dk}}{(k+1)!(2d(k+1)-d)} \int_{(\mathbb{R}^d)^k} h_{\sqrt{d}}(0, y_1, \dots, y_k) dy \\ &= \omega_d \sum_{k=0}^{\infty} \frac{(-1)^k (t/\sqrt{d})^{dk} (k+1)^d}{(k+1)!(2k+1)}. \end{aligned} \quad (3.5)$$

See Figs. 2 and 3 for actual plots of the limiting function in Eq. 3.5 for $d = 2, 3, 4, 5$.

One of the implications of Theorem 3.1 is that one can immediately obtain various limit theorems of functions of the Euler characteristic process. For every continuous

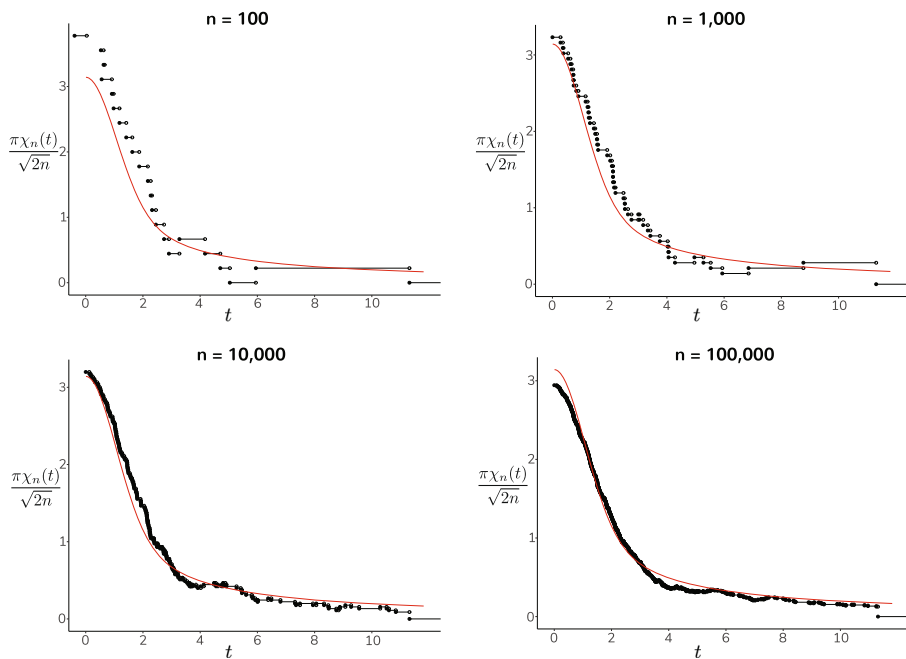


Fig. 2 Plots of random realizations of $\sqrt{\pi\omega_d/(2n)}\chi_n(t)$ for $d = 2$ (in black) in the setup of Example 3.2. In the plots above, as n increases from left to right, the random function converges uniformly to $\sum_{k=0}^{\infty}(-1)^k s_k(t)$ (in red)

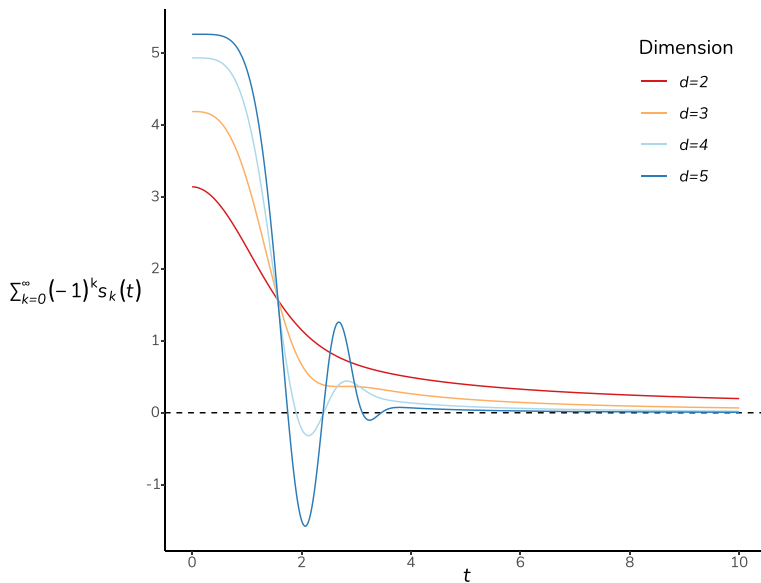


Fig. 3 Plots of $\sum_{k=0}^{\infty} (-1)^k s_k(t)$ at Eq. 3.5 for $d = 2, 3, 4, 5$.

function T on $D[0, \infty)$, it indeed holds that, as $n \rightarrow \infty$,

$$T\left(\sqrt{\frac{\pi\omega_d}{2n}}\chi_n\right) \rightarrow T\left(\sum_{k=0}^{\infty} (-1)^k s_k\right), \text{ a.s.}$$

For example, if $U_{a,b} : D[0, \infty) \rightarrow [0, \infty)$ is defined by $U_{a,b}(x) := \sup_{a \leq x \leq b} |x(t)|$, $0 \leq a < b < \infty$, we have

$$\sqrt{\frac{\pi\omega_d}{2n}} \sup_{a \leq t \leq b} |\chi_n(t)| \rightarrow \sup_{a \leq t \leq b} \left| \sum_{k=0}^{\infty} (-1)^k s_k(t) \right| \text{ a.s.}$$

Furthermore, let $I : D[0, \infty) \rightarrow D[0, \infty)$ be defined by $I(x)(t) := \int_0^t x(r)dr$; then,

$$\left(\sqrt{\frac{\pi\omega_d}{2n}} \int_0^t \chi_n(r)dr, t \geq 0\right) \rightarrow \left(\sum_{k=0}^{\infty} (-1)^k \int_0^t s_k(r)dr, t \geq 0\right), \text{ a.s. in } D[0, \infty).$$

This result is especially important for applications in TDA. Indeed, $\int_0^t \chi_n(r)dr$ represents an alternating sum of the total length of persistence barcodes of all dimensions, up to time t . A persistence barcode is a graphical descriptor of persistent homology, which allows us to visualize the birth time and death time of cycles (Ghrist 2008; Carlsson 2009). In light of the TDA literature (e.g., Section 6 of Bobrowski and Borman 2012), the limit $\int_0^\infty \chi_n(r)dr$ is defined as the Euler characteristic of persistence barcodes of the filtration (2.2). This gives us an estimate of how long the cycles of any dimension live in our extreme sample cloud.

4 Exponentially decaying tail case

In this section, we consider a density of an exponentially decaying tail. We assume that the density f is specified by

$$f(x) = C \exp \{-\psi(\|x\|)\}, \quad x \in \mathbb{R}^d, \quad (4.1)$$

where C is a normalizing constant and $\psi : [0, \infty) \rightarrow [0, \infty)$ is a regularly varying function (at infinity) of an exponent $\tau \in (0, 1]$. Moreover, ψ is assumed to be twice differentiable, such that $\psi'(x) > 0$ for all $x > 0$, and ψ' is eventually non-increasing. Namely, there exists $z_0 > 0$ such that ψ' is non-increasing in $[z_0, \infty)$. Under this setup, let $a(z) := 1/\psi'(z)$; then, a is also regularly varying with index $1 - \tau$ (see, e.g., Proposition 2.5 in Resnick 2007).

Here, it is important to note that the occurrence of topological crackle depends on the limit value of $a(z)$ as $z \rightarrow \infty$ (see Owada and Adler 2017). In particular, Owada and Adler (2017) showed that crackle occurs if and only if

$$\zeta := \lim_{z \rightarrow \infty} a(z) \in (0, \infty]. \quad (4.2)$$

Since the main theme of this study is topological crackle, we do not treat the case $\zeta = 0$. In terms of the regular variation exponent of ψ , we exclude the case $\tau > 1$. So, for instance, the multivariate Gaussian densities do not belong to the scope of our study. Note that (4.2) trivially holds for every $\tau \in (0, 1)$.

Now, we are ready to state the FSLN below. Interestingly, if $\zeta = \infty$ in Eq. 4.2, the limiting function in Eq. 4.4 agrees with Eq. 3.4 up to multiplicative constants.

Theorem 4.1 *Suppose that f is a density specified by Eq. 4.1 with $\tau \in (0, 1]$. If $\tau = 1$, we assume Eq. 4.2. If $d = 2$, we restrict the range of τ to $(0, 1)$. Suppose further that $nf(R_n) \rightarrow \xi$ as $n \rightarrow \infty$ for some $\xi \in (0, \infty)$. Then, we have, as $n \rightarrow \infty$,*

$$\left(\frac{\chi_n(t)}{a(R_n)R_n^{d-1}}, t \geq 0 \right) \rightarrow \left(\sum_{k=0}^{\infty} (-1)^k s_k(t), t \geq 0 \right), \quad \text{a.s. in } D[0, \infty), \quad (4.3)$$

where

$$\begin{aligned} s_k(t) := & \frac{\xi^{k+1}}{(k+1)!} \int_0^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^k} h_t(0, y_1, \dots, y_k) e^{-(k+1)\rho - \zeta^{-1} \sum_{i=1}^k \langle \theta, y_i \rangle} \\ & \times \prod_{i=1}^k \mathbf{1} \left\{ \rho + \zeta^{-1} \langle \theta, y_i \rangle \geq 0 \right\} dy dv_{d-1}(\theta) d\rho, \quad t \geq 0, \quad k \geq 1, \end{aligned} \quad (4.4)$$

with $\langle \cdot, \cdot \rangle$ being the Euclidean inner product and $s_0(t) \equiv s_{d-1}\xi$. In particular, the limit in Eq. (4.3) is convergent for all $t \geq 0$.

Example 4.2 We consider a special case of the density in Eq. 4.1,

$$f(x) = C e^{-\|x\|^\tau / \tau}, \quad x \in \mathbb{R}^d, \quad \tau \in (0, 1].$$

Define $R_n = (\tau \log n + \tau \log C)^{1/\tau}$ so that $nf(R_n) = 1$. Then, $a(z) = z^{1-\tau}$, $z > 0$. According to Theorem 4.1,

$$\left(\frac{\chi_n(t)}{(\tau \log n)^{\frac{d-\tau}{\tau}}}, t \geq 0 \right) \rightarrow \left(\sum_{k=0}^{\infty} (-1)^k s_k(t), t \geq 0 \right), \quad \text{a.s. in } D[0, \infty).$$

where $s_k(t)$ is defined in Eq. 4.4. Moreover, applying the continuous functions $U_{a,b}$ and I from Example 3.2, we have

$$\sup_{a \leq t \leq b} \frac{|\chi_n(t)|}{(\tau \log n)^{\frac{d-\tau}{\tau}}} \rightarrow \sup_{a \leq t \leq b} \left| \sum_{k=0}^{\infty} (-1)^k s_k(t) \right| \quad \text{a.s.}$$

and

$$\left(\frac{\int_0^t \chi_n(r) dr}{(\tau \log n)^{\frac{d-\tau}{\tau}}}, t \geq 0 \right) \rightarrow \left(\sum_{k=0}^{\infty} (-1)^k \int_0^t s_k(r) dr, t \geq 0 \right), \quad \text{a.s. in } D[0, \infty).$$

5 Proofs of main theorems

Throughout the proofs, denote by C^* a positive constant that is independent of n and may vary between (and even within) the lines. Denote by RV_ρ the collection of regularly varying sequences (or functions) at infinity with exponent $\rho \in \mathbb{R}$. For $a, b \in \mathbb{R}$, write $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. For two sequences $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$, $a_n \sim b_n$ means $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$.

First, we present a fundamental result which allows us to extend a pointwise SLLN to a functional SLLN in the space $D[0, \infty)$.

Proposition 5.1 (Proposition 4.2 in Thomas and Owada 2021) *Let $(X_n, n \in \mathbb{N})$ be a sequence of random elements in $D[0, \infty)$ with non-decreasing sample paths. Suppose $a : [0, \infty) \rightarrow \mathbb{R}$ is a deterministic, continuous, and non-decreasing function. If we have*

$$X_n(t) \rightarrow a(t), \quad n \rightarrow \infty, \quad \text{a.s.},$$

for every $t \geq 0$, then it follows that

$$\sup_{t \in [0, T]} |X_n(t) - a(t)| \rightarrow 0, \quad n \rightarrow \infty, \quad \text{a.s.},$$

for every $0 \leq T < \infty$. Hence, it holds that $X_n(t) \rightarrow a(t)$ a.s. in $D[0, \infty)$ under the uniform topology.

By virtue of this proposition, for the proof of Theorem 3.1 it suffices to show that as $n \rightarrow \infty$,

$$\frac{\chi_n(t)}{R_n^d} \rightarrow \sum_{k=0}^{\infty} (-1)^k s_k(t), \quad \text{a.s.}$$

for every $t \geq 0$. Subsequently, we divide the Euler characteristic process into two terms:

$$\chi_n(t) = \sum_{k=0}^{\infty} S_{2k,n}(t) - \sum_{k=0}^{\infty} S_{2k+1,n}(t) =: \chi_n^{(1)}(t) - \chi_n^{(2)}(t). \quad (5.1)$$

In addition, the limiting function can also be decomposed as

$$\sum_{k=0}^{\infty} (-1)^k s_k(t) = \sum_{k=0}^{\infty} s_{2k}(t) - \sum_{k=0}^{\infty} s_{2k+1}(t) =: K_1(t) - K_2(t). \quad (5.2)$$

From Eqs. 5.1 and 5.2, it is now sufficient to prove that for every $t \geq 0$ and $i = 1, 2$,

$$\frac{\chi_n^{(i)}(t)}{R_n^d} \rightarrow K_i(t), \quad n \rightarrow \infty, \text{ a.s.} \quad (5.3)$$

In the case of Theorem 4.1, defining $\chi_n^{(i)}(t)$ and $K_i(t)$ analogously, it suffices to show that for each $t \geq 0$ and $i = 1, 2$,

$$\frac{\chi_n^{(i)}(t)}{a(R_n)R_n^{d-1}} \rightarrow K_i(t), \quad n \rightarrow \infty, \text{ a.s.} \quad (5.4)$$

5.1 Proof of Theorem 3.1

The goal of this subsection is to prove (5.3). We handle the case $i = 1$ only, as the proof is totally the same regardless of $i \in \{1, 2\}$. Let

$$u_m = \lfloor e^{m^\gamma} \rfloor, \quad m = 0, 1, 2, \dots, \quad (5.5)$$

for some $\gamma \in (0, 1)$. Then, for every $n \in \mathbb{N}$, there exists a unique $m = m(n)$ such that $u_m \leq n < u_{m+1}$. Let us also define

$$p_m = \operatorname{argmax}\{u_m \leq \ell \leq u_{m+1} : R_\ell\}, \quad (5.6)$$

$$q_m = \operatorname{argmin}\{u_m \leq \ell \leq u_{m+1} : R_\ell\}. \quad (5.7)$$

It then holds that $R_{p_m} = \max_{u_m \leq \ell \leq u_{m+1}} R_\ell$ and $R_{q_m} = \min_{u_m \leq \ell \leq u_{m+1}} R_\ell$.

Below, we offer a lemma on the asymptotic moments of certain variants of the process $\chi_n^{(1)}(t)$, defined by

$$T_m(t) := \sum_{k=0}^{\infty} \sum_{\substack{\mathcal{Y} \subset \mathcal{X}_{u_{m+1}}, \\ |\mathcal{Y}|=2k+1}} h_t(\mathcal{Y}) \mathbf{1} \left\{ \min_{y \in \mathcal{Y}} \|y\| \geq R_{q_m} \right\}, \quad (5.8)$$

$$U_m(t) := \sum_{k=0}^{\infty} \sum_{\substack{\mathcal{Y} \subset \mathcal{X}_{u_m}, \\ |\mathcal{Y}|=2k+1}} h_t(\mathcal{Y}) \mathbf{1} \left\{ \min_{y \in \mathcal{Y}} \|y\| \geq R_{p_m} \right\} \quad (5.9)$$

Lemma 5.1 *Under the assumptions of Theorem 3.1, we have the following asymptotic results on the first and second moments of $T_m(t)$ and $U_m(t)$.*

$$\lim_{m \rightarrow \infty} R_{q_m}^{-d} \mathbb{E}[T_m(t)] = K_1(t), \quad (5.10)$$

$$\lim_{m \rightarrow \infty} R_{p_m}^{-d} \mathbb{E}[U_m(t)] = K_1(t), \quad (5.11)$$

$$\sup_{m \geq 1} R_{q_m}^{-d} \text{Var}(T_m(t)) < \infty, \quad (5.12)$$

$$\sup_{m \geq 1} R_{p_m}^{-d} \text{Var}(U_m(t)) < \infty. \quad (5.13)$$

Proof We begin by offering the proofs of Eqs. 5.10 and 5.11 by extending the argument in the proof of Proposition 7.2 of Owada (2017). As for Eq. 5.10, it is clear that

$$R_{q_m}^{-d} \mathbb{E}[T_m(t)] = \sum_{k=0}^{\infty} R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \mathbb{E} \left[h_t(X_1, \dots, X_{2k+1}) \mathbf{1} \left\{ \min_{1 \leq i \leq 2k+1} \|X_i\| \geq R_{q_m} \right\} \right],$$

where $X_1, \dots, X_{2k+1}$ are i.i.d random variables with density f . From this, we have

$$\begin{aligned} & R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \mathbb{E} \left[h_t(X_1, \dots, X_{2k+1}) \mathbf{1} \left\{ \min_{1 \leq i \leq 2k+1} \|X_i\| \geq R_{q_m} \right\} \right] \quad (5.14) \\ &= R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \int_{(\mathbb{R}^d)^{2k+1}} h_t(x_1, \dots, x_{2k+1}) \prod_{i=1}^{2k+1} f(x_i) \mathbf{1} \{ \|x_i\| \geq R_{q_m} \} \, d\mathbf{x} \\ &= R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{2k}} h_t(0, y_1, \dots, y_{2k}) f(x) \mathbf{1} \{ \|x\| \geq R_{q_m} \} \\ &\quad \times \prod_{i=1}^{2k} f(x + y_i) \mathbf{1} \{ \|x + y_i\| \geq R_{q_m} \} \, dy dx, \end{aligned}$$

by the change of variables $x_i = x + y_{i-1}$, $i = 1, \dots, 2k+1$ (with $y_0 \equiv 0$) and the translation invariance of h_t . Furthermore, we make the change of variables by $x = R_{q_m} \rho \theta$ with $\rho \geq 1$ and $\theta \in S^{d-1}$, to get that

$$\begin{aligned} & R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \mathbb{E} \left[h_t(X_1, \dots, X_{2k+1}) \mathbf{1} \left\{ \min_{1 \leq i \leq 2k+1} \|X_i\| \geq R_{q_m} \right\} \right] \quad (5.15) \\ &= \binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} \int_1^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) \rho^{d-1} \frac{f(R_{q_m} \rho)}{f(R_{q_m})} \\ &\quad \times \prod_{i=1}^{2k} \frac{f(R_{q_m} \|\rho \theta + y_i / R_{q_m}\|)}{f(R_{q_m})} \mathbf{1} \{ \|\rho \theta + y_i / R_{q_m}\| \geq 1 \} \, dy \, d\nu_{d-1}(\theta) \, d\rho, \end{aligned}$$

where $\mathbf{y} = (y_1, \dots, y_{2k}) \in (\mathbb{R}^d)^{2k}$. Next, for a fixed constant $\eta \in (0, \alpha - d)$, Potter's bounds (see Proposition 2.6 in Resnick 2007) yield that

$$\frac{f(R_{q_m} \rho)}{f(R_{q_m})} \leq 2\rho^{-\alpha+\eta}, \quad (5.16)$$

and

$$\prod_{i=1}^{2k} \frac{f(R_{q_m} \|\rho\theta + y_i/R_{q_m}\|)}{f(R_{q_m})} \mathbf{1}_{\{\|\rho\theta + y_i/R_{q_m}\| \geq 1\}} \leq 2^{2k} \quad (5.17)$$

for sufficiently large m . Since $\int_1^\infty \rho^{d-1-\alpha+\eta} d\rho < \infty$ and $\int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) d\mathbf{y} < \infty$ by property **(H3)** of h , we can see that the regular variation of f , as well as the dominated convergence theorem, ensures that the triple integral in Eq. 5.15 converges to

$$s_{d-1} \int_1^\infty \rho^{d-1-\alpha(2k+1)} d\rho \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) d\mathbf{y} = \frac{s_{d-1}}{\alpha(2k+1) - d} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) d\mathbf{y}.$$

Furthermore, Eq. 2.5 ensures that as $m \rightarrow \infty$,

$$\binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} \sim \frac{(u_{m+1} f(R_{q_m}))^{2k+1}}{(2k+1)!} \sim \frac{\xi^{2k+1}}{(2k+1)!} \left(\frac{u_{m+1}}{q_m}\right)^{2k+1},$$

so that

$$1 \leq \frac{u_{m+1}}{q_m} \leq \frac{u_{m+1}}{u_m} \leq \frac{e^{(m+1)^\gamma - m^\gamma}}{1 - e^{-m^\gamma}} = \frac{e^{m^{\gamma-1}(\gamma + o(1))}}{1 - e^{-m^\gamma}}. \quad (5.18)$$

As $0 < \gamma < 1$, the rightmost term above converges to 1 as $m \rightarrow \infty$. Hence,

$$\binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} \rightarrow \frac{\xi^{2k+1}}{(2k+1)!}, \text{ as } m \rightarrow \infty.$$

Combining all of these results, it follows that

$$R_{q_m}^{-d} \binom{u_{m+1}}{2k+1} \mathbb{E} \left[h_t(X_1, \dots, X_{2k+1}) \mathbf{1}_{\left\{ \min_{1 \leq i \leq 2k+1} \|X_i\| \geq R_{q_m} \right\}} \right] \rightarrow s_{2k}(t), \text{ as } m \rightarrow \infty,$$

which yields (5.10) as desired. The proof of Eq. 5.11 is almost the same, so we omit it here.

Now we will prove Eq. 5.12. We see that

$$\begin{aligned} \mathbb{E}[T_m(t)^2] &= \sum_{k_1=0}^\infty \sum_{k_2=0}^\infty \sum_{\ell=0}^{2(k_1 \wedge k_2)+1} \mathbb{E} \left[\prod_{i=1}^2 \left(\sum_{\substack{\mathcal{Y}_i \subset \mathcal{X}_{u_{m+1}}, \\ |\mathcal{Y}_i|=2k_i+1}} h_t(\mathcal{Y}_i) \mathbf{1}_{\left\{ \min_{y \in \mathcal{Y}_i} \|y\| \geq R_{q_m} \right\}} \right) \mathbf{1}_{\{|\mathcal{Y}_1 \cap \mathcal{Y}_2|=\ell\}} \right] \\ &= \sum_{k_1=0}^\infty \sum_{k_2=0}^\infty \sum_{\ell=0}^{2(k_1 \wedge k_2)+1} \binom{u_{m+1}}{2(k_1+k_2)+2-\ell} \frac{(2(k_1+k_2)+2-\ell)!}{(2k_1+1-\ell)!(2k_2+1-\ell)!\ell!} \\ &\quad \times \mathbb{E} [h_t(X_1, \dots, X_{2k_1+1}) h_t(X_\ell, X_{2k_1+2}, \dots, X_{2(k_1+k_2)+2-\ell}) \\ &\quad \times \mathbf{1}_{\left\{ \min_{1 \leq i \leq 2(k_1+k_2)+2-\ell} \|X_i\| \geq R_{q_m} \right\}}] \\ &=: \sum_{k_1=0}^\infty \sum_{k_2=0}^\infty \sum_{\ell=0}^{2(k_1 \wedge k_2)+1} \binom{u_{m+1}}{2(k_1+k_2)+2-\ell} \frac{(2(k_1+k_2)+2-\ell)!}{(2k_1+1-\ell)!(2k_2+1-\ell)!\ell!} \mathbb{E}[I_{k_1, k_2, \ell}], \end{aligned}$$

where $X_1, \dots, X_{2(k_1+k_2)+2-\ell}$ are i.i.d random points with density f . In the above, if $\ell = 0$, we define

$$h_t(X_1, \dots, X_\ell, X_{2k_1+2}, \dots, X_{2(k_1+k_2)+2-\ell}) := h_t(X_{2k_1+2}, \dots, X_{2(k_1+k_2)+2}).$$

From this, we see that

$$\begin{aligned}
 & R_{q_m}^{-d} \text{Var}(T_m(t)) \\
 &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{\ell=1}^{2(k_1 \wedge k_2)+1} R_{q_m}^{-d} \binom{u_{m+1}}{2(k_1+k_2)+2-\ell} \frac{(2(k_1+k_2)+2-\ell)!}{(2k_1+1-\ell)!(2k_2+1-\ell)!\ell!} \mathbb{E}[I_{k_1,k_2,\ell}] \\
 &\quad + \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} R_{q_m}^{-d} \binom{u_{m+1}}{2(k_1+k_2)+2} \binom{2(k_1+k_2)+2}{2k_1+1} \mathbb{E}[I_{k_1,k_2,0}] \\
 &\quad - \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} R_{q_m}^{-d} \binom{u_{m+1}}{2k_1+1} \binom{u_{m+1}}{2k_2+1} \mathbb{E}[I_{k_1,k_2,0}] \\
 &\leq \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{\ell=1}^{2(k_1 \wedge k_2)+1} R_{q_m}^{-d} \binom{u_{m+1}}{2(k_1+k_2)+2-\ell} \frac{(2(k_1+k_2)+2-\ell)!}{(2k_1+1-\ell)!(2k_2+1-\ell)!\ell!} \mathbb{E}[I_{k_1,k_2,\ell}],
 \end{aligned} \tag{5.19}$$

where the last inequality comes from

$$\begin{aligned}
 & \binom{u_{m+1}}{2(k_1+k_2)+2} \binom{2(k_1+k_2)+2}{2k_1+1} - \binom{u_{m+1}}{2k_1+1} \binom{u_{m+1}}{2k_2+1} \\
 &= \binom{u_{m+1}}{2k_1+1} \left[\binom{u_{m+1}-2k_1-1}{2k_2+1} - \binom{u_{m+1}}{2k_2+1} \right] < 0.
 \end{aligned}$$

For our purposes, we must examine the appropriate upper bounds of $\mathbb{E}[I_{k_1,k_2,\ell}]$ for $\ell \geq 1$. For every $\ell \geq 1$, performing the change of variables, $x_i = x + y_{i-1}$, $i = 1, \dots, 2(k_1+k_2)+2-\ell$ (with $y_0 \equiv 0$), we have that

$$\begin{aligned}
 \mathbb{E}[I_{k_1,k_2,\ell}] &= \int_{(\mathbb{R}^d)^{2(k_1+k_2)+2-\ell}} h_t(x_1, \dots, x_{2k_1+1}) h_t(x_1, \dots, x_\ell, x_{2k_1+2}, \dots, x_{2(k_1+k_2)+2-\ell}) \\
 &\quad \times \prod_{i=1}^{2(k_1+k_2)+2-\ell} f(x_i) \mathbf{1}\{\|x_i\| \geq R_{q_m}\} \, dx \\
 &= \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} h_t(0, y_1, \dots, y_{2k_1}) h_t(0, y_1, \dots, y_{\ell-1}, y_{2k_1+1}, \dots, y_{2(k_1+k_2)+1-\ell}) \\
 &\quad \times f(x) \mathbf{1}\{\|x\| \geq R_{q_m}\} \prod_{i=1}^{2(k_1+k_2)+1-\ell} f(x+y_i) \mathbf{1}\{\|x+y_i\| \geq R_{q_m}\} \, dy \, dx.
 \end{aligned}$$

As in Eq. 5.15, we apply the polar coordinate transform $x = R_{q_m} \rho \theta$ with $\rho \geq 1$ and $\theta \in S^{d-1}$, to obtain that

$$\begin{aligned}
 \mathbb{E}[I_{k_1,k_2,\ell}] &= R_{q_m}^d f(R_{q_m})^{2(k_1+k_2)+2-\ell} \int_1^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} h_t(0, \mathbf{y}_0, \mathbf{y}_1) h_t(0, \mathbf{y}_0, \mathbf{y}_2) \rho^{d-1} \\
 &\quad \times \frac{f(R_{q_m} \rho)}{f(R_{q_m})} \prod_{i=1}^{2(k_1+k_2)+1-\ell} \frac{f(R_{q_m} \|\rho \theta + y_i / R_{q_m}\|)}{f(R_{q_m})} \mathbf{1}\{\|\rho \theta + y_i / R_{q_m}\| \geq 1\} \\
 &\quad d(\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_2) dv_{d-1}(\theta) d\rho,
 \end{aligned}$$

where $\mathbf{y}_0 = (y_1, \dots, y_{\ell-1})$, $\mathbf{y}_1 = (y_\ell, \dots, y_{2k_1})$ and $\mathbf{y}_2 = (y_{2k_1+1}, \dots, y_{2(k_1+k_2)+1-\ell})$. Appealing to Potter's bounds as in Eqs. 5.16 and 5.17, as well as Eq. 2.5, there exists an $N \in \mathbb{N}$ such that for all $m \geq N$,

$$\begin{aligned} & R_{q_m}^{-d} \left(2(k_1 + k_2) + 2 - \ell \right) \mathbb{E}[I_{k_1, k_2, \ell}] \\ & \leq \frac{(u_{m+1} f(R_{q_m}))^{2(k_1+k_2)+2-\ell}}{(2(k_1 + k_2) + 2 - \ell)!} \int_1^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} h_t(0, \mathbf{y}_0, \mathbf{y}_1) h_t(0, \mathbf{y}_0, \mathbf{y}_2) \rho^{d-1} \\ & \quad \times 2\rho^{-\alpha+\eta} \times 2^{2(k_1+k_2)+1-\ell} d(\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_2) dv_{d-1}(\theta) d\rho \\ & \leq \frac{(4\xi)^{2(k_1+k_2)+2-\ell}}{(2(k_1 + k_2) + 2 - \ell)!} \cdot \frac{s_{d-1}}{\alpha - d - \eta} \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} h_t(0, \mathbf{y}_0, \mathbf{y}_1) h_t(0, \mathbf{y}_0, \mathbf{y}_2) d(\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_2). \end{aligned}$$

By virtue of property (H3) of h ,

$$\begin{aligned} \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} h_t(0, \mathbf{y}_0, \mathbf{y}_1) h_t(0, \mathbf{y}_0, \mathbf{y}_2) d(\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_2) & \leq \int_{(\mathbb{R}^d)^{2(k_1+k_2)+1-\ell}} \prod_{i=1}^{2(k_1+k_2)+1-\ell} \mathbf{1}_{\{\|\mathbf{y}_i\| \leq ct\}} d\mathbf{y} \\ & = \left((ct)^d \omega_d \right)^{2(k_1+k_2)+1-\ell}. \end{aligned}$$

Therefore, for all $m \geq N$,

$$\begin{aligned} R_{q_m}^{-d} \left(2(k_1 + k_2) + 2 - \ell \right) \mathbb{E}[I_{k_1, k_2, \ell}] & \leq \frac{(4\xi)^{2(k_1+k_2)+2-\ell} \left((ct)^d \omega_d \right)^{2(k_1+k_2)+1-\ell}}{(2(k_1 + k_2) + 2 - \ell)!} \cdot \frac{s_{d-1}}{\alpha - d - \eta} \\ & \leq \frac{(C^*)^{2(k_1+k_2)+2-\ell}}{(2(k_1 + k_2) + 2 - \ell)!}. \end{aligned} \quad (5.20)$$

Note that the constant C^* does not depend on k_1 , k_2 , and ℓ . Returning to Eq. 5.19 and applying the bound in Eq. 5.20, we have that

$$\begin{aligned} R_{q_m}^{-d} \text{Var}(T_m(t)) & \leq \sum_{k_1=0}^\infty \sum_{k_2=0}^\infty \sum_{\ell=1}^{2(k_1 \wedge k_2)+1} \frac{(C^*)^{2(k_1+k_2)+2-\ell}}{(2k_1+1-\ell)! (2k_2+1-\ell)! \ell!} \\ & \leq 2 \sum_{k_1=0}^\infty \sum_{k_2=0}^{k_1} \sum_{\ell=1}^{k_2+1} \frac{(C^*)^{k_1+k_2+2-\ell}}{(k_1+1-\ell)! (k_2+1-\ell)! \ell!} \\ & \leq 2 \sum_{\ell=1}^\infty \sum_{k_1=\ell-1}^\infty \sum_{k_2=\ell-1}^\infty \frac{(C^*)^{k_1+1-\ell}}{(k_1+1-\ell)!} \cdot \frac{(C^*)^{k_2+1-\ell}}{(k_2+1-\ell)!} \cdot \frac{(C^*)^\ell}{\ell!} = 2e^{3C^*} < \infty. \end{aligned}$$

Since the proof of Eq. 5.13 is very similar to that of Eq. 5.12, we will omit it. $\square$

Proof (Theorem 3.1) By the definition of u_m , p_m , and q_m , we have, for every $n \geq 1$,

$$\frac{U_m(t)}{R_{p_m}^d} \leq \frac{\chi_n^{(1)}(t)}{R_n^d} \leq \frac{T_m(t)}{R_{q_m}^d}.$$

Then, Lemma 5.1 gives that

$$\limsup_{n \rightarrow \infty} \frac{\chi_n^{(1)}(t)}{R_n^d} \leq K_1(t) + \limsup_{m \rightarrow \infty} \frac{T_m(t) - \mathbb{E}[T_m(t)]}{R_{q_m}^d}, \quad \text{a.s.}, \quad (5.21)$$

and

$$\liminf_{n \rightarrow \infty} \frac{\chi_n^{(1)}(t)}{R_n^d} \geq K_1(t) + \liminf_{m \rightarrow \infty} \frac{U_m(t) - \mathbb{E}[U_m(t)]}{R_{p_m}^d}, \quad \text{a.s.} \quad (5.22)$$

Let us continue by showing that

$$R_{q_m}^{-d} (T_m(t) - \mathbb{E}[T_m(t)]) \rightarrow 0, \quad m \rightarrow \infty, \quad \text{a.s.}, \quad (5.23)$$

and

$$R_{p_m}^{-d} (U_m(t) - \mathbb{E}[U_m(t)]) \rightarrow 0, \quad m \rightarrow \infty, \quad \text{a.s.} \quad (5.24)$$

For Eq. 5.23, it follows from Eq. 5.12 in Lemma 5.1 and Chebyshev's inequality that, for every $\epsilon > 0$,

$$\sum_{m=1}^{\infty} \mathbb{P} \left(|T_m(t) - \mathbb{E}[T_m(t)]| > \epsilon R_{q_m}^d \right) \leq \sum_{m=1}^{\infty} \frac{\text{Var}(T_m(t))}{\epsilon^2 (R_{q_m}^d)^2} \leq C^* \sum_{m=1}^{\infty} \frac{1}{R_{q_m}^d}.$$

As $R_n \in \text{RV}_{1/\alpha}$ (see Eq. 3.2), we have that

$$R_{q_m} \geq C^* q_m^{1/(2\alpha)} \geq C^* u_m^{1/(2\alpha)} \geq C^* e^{m^\gamma/(3\alpha)}$$

for all $m \geq 1$. Now, we have $\sum_m R_{q_m}^{-d} \leq C^* \sum_m e^{-dm^\gamma/(3\alpha)} < \infty$, and the Borel-Cantelli lemma concludes (5.23). The proof of Eq. 5.24 is analogous by virtue of Eq. 5.13 in Lemma 5.1. Now, combining (5.21), (5.22), (5.23), and (5.24) completes the proof.

Finally, let us explicitly demonstrate that the limit in Eq. 3.3 is finite for all $t \geq 0$. By virtue of property (H3) of h ,

$$\begin{aligned} \left| \sum_{k=0}^{\infty} (-1)^k s_k(t) \right| &\leq \sum_{k=0}^{\infty} \frac{s_{d-1} \xi^{k+1}}{(k+1)! (\alpha(k+1) - d)} \int_{(\mathbb{R}^d)^k} \prod_{i=1}^k \mathbf{1}_{\{\|y_i\| \leq ct\}} dy \\ &\leq C^* \sum_{k=0}^{\infty} \frac{((ct)^d \xi \omega_d)^k}{k!} = C^* e^{(ct)^d \xi \omega_d} < \infty. \end{aligned}$$

□

5.2 Proof of Theorem 4.1

The goal here is to prove Eq. 5.4. Once again, we deal with the case $i = 1$ only. The proof is essentially similar in character to the proof of Theorem 3.1 but involves more complex machinery. First we take

$$\gamma \in \left(\frac{\tau}{d - \tau}, 1 \right) \quad (5.25)$$

(recall that we have restricted the range of τ to $(0, 1)$ when $d = 2$), and define

$$u_m = \lfloor e^{m^\gamma} \rfloor, \quad m = 0, 1, 2, \dots$$

as in Eq. 5.5. Moreover, let p_m and q_m remain as before—see Eqs. 5.6 and 5.7. Additionally, we also introduce

$$v_m := \operatorname{argmax} \left\{ u_m \leq \ell \leq u_{m+1} : a(R_\ell) R_\ell^{d-1} \right\},$$

$$w_m := \operatorname{argmin} \left\{ u_m \leq \ell \leq u_{m+1} : a(R_\ell) R_\ell^{d-1} \right\}.$$

Let $T_m(t)$ and $U_m(t)$ be defined in the same way as Eqs. 5.8 and 5.9.

The lemma below is analogous to Lemma 5.1 (see also Proposition 7.4 in Owada 2017) that provides the asymptotic moments of $T_m(t)$ and $U_m(t)$.

Lemma 5.2 *Under the assumptions of Theorem 4.1, we have the following.*

$$\lim_{m \rightarrow \infty} \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \mathbb{E} [T_m(t)] = K_1(t), \quad (5.26)$$

$$\lim_{m \rightarrow \infty} \left[a(R_{v_m}) R_{v_m}^{d-1} \right]^{-1} \mathbb{E} [U_m(t)] = K_1(t), \quad (5.27)$$

$$\sup_{m \geq 1} \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \operatorname{Var} (T_m(t)) < \infty, \quad (5.28)$$

$$\sup_{m \geq 1} \left[a(R_{v_m}) R_{v_m}^{d-1} \right]^{-1} \operatorname{Var} (U_m(t)) < \infty. \quad (5.29)$$

Proof We begin by proving (5.26) and (5.27). By the same change of variables as in Eq. 5.14 and the translation invariance of h_t ,

$$\begin{aligned} & \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \mathbb{E} [T_m(t)] \\ &= \sum_{k=0}^{\infty} \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \binom{u_{m+1}}{2k+1} \int_{(\mathbb{R}^d)^{2k+1}} h_t(x_1, \dots, x_{2k+1}) \prod_{i=1}^{2k+1} f(x_i) \mathbf{1} \{ \|x_i\| \geq R_{q_m} \} \, \mathrm{d}x \\ &= \sum_{k=0}^{\infty} \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \binom{u_{m+1}}{2k+1} \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) f(x) \mathbf{1} \{ \|x\| \geq R_{q_m} \} \\ & \quad \times \prod_{i=1}^{2k} f(x + y_i) \mathbf{1} \{ \|x + y_i\| \geq R_{q_m} \} \, \mathrm{d}y \, \mathrm{d}x, \end{aligned}$$

where $\mathbf{y} = (y_1, \dots, y_{2k}) \in (\mathbb{R}^d)^{2k}$. Here, we make the change of variable by $x = (R_{q_m} + a(R_{q_m})\rho)\theta$ with $\rho \geq 0$ and $\theta \in S^{d-1}$. Then, the integral above becomes

$$\begin{aligned} & a(R_{q_m}) \int_0^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) (R_{q_m} + a(R_{q_m})\rho)^{d-1} f(R_{q_m} + a(R_{q_m})\rho) \\ & \times \prod_{i=1}^{2k} f(\|(R_{q_m} + a(R_{q_m})\rho)\theta + y_i\|) \mathbf{1} \{ \|(R_{q_m} + a(R_{q_m})\rho)\theta + y_i\| \geq R_{q_m} \} \, \mathrm{d}y \, \mathrm{d}v_{d-1}(\theta) \, \mathrm{d}\rho, \end{aligned}$$

which implies that

$$\begin{aligned}
 & \left[a(R_{w_m}) R_{w_m}^{d-1} \right]^{-1} \mathbb{E} [T_m(t)] \\
 &= \sum_{k=0}^{\infty} \frac{a(R_{q_m}) R_{q_m}^{d-1}}{a(R_{w_m}) R_{w_m}^{d-1}} \binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} \\
 & \quad \times \int_0^{\infty} \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) \left(1 + \frac{a(R_{q_m})}{R_{q_m}} \rho \right)^{d-1} \frac{f(R_{q_m} + a(R_{q_m})\rho)}{f(R_{q_m})} \\
 & \quad \times \prod_{i=1}^{2k} \frac{f(\|(R_{q_m} + a(R_{q_m})\rho)\theta + y_i\|)}{f(R_{q_m})} \mathbf{1}_{\{\|(R_{q_m} + a(R_{q_m})\rho)\theta + y_i\| \geq R_{q_m}\}} d\mathbf{y} dv_{d-1}(\theta) d\rho.
 \end{aligned} \tag{5.30}$$

For the last expression, we claim that

$$\frac{a(R_{q_m}) R_{q_m}^{d-1}}{a(R_{w_m}) R_{w_m}^{d-1}} \rightarrow 1, \quad m \rightarrow \infty, \tag{5.31}$$

and

$$\binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} \rightarrow \frac{\xi^{2k+1}}{(2k+1)!}, \quad m \rightarrow \infty. \tag{5.32}$$

Because of Eq. 2.5, we have $\psi(R_n) \sim \log(Cn/\xi)$ as $n \rightarrow \infty$. With the assumptions on the density (4.1), Proposition 2.6 in Resnick (2007) gives that $\psi^{\leftarrow} \in \text{RV}_{1/\tau}$. It follows from the uniform convergence of regularly varying sequences [see Proposition 2.4 in Resnick (2007)] that

$$\frac{\psi^{\leftarrow}(\psi(R_n))}{\psi^{\leftarrow}(\log(Cn/\xi))} \sim \left(\frac{\psi(R_n)}{\log(Cn/\xi)} \right)^{1/\tau} \rightarrow 1, \quad \text{as } n \rightarrow \infty.$$

Since $\psi^{\leftarrow}(\psi(R_n)) \sim R_n$ as $n \rightarrow \infty$, the above relation and $\log(Cn/\xi) \sim \log n$, $n \rightarrow \infty$, implies

$$R_n \sim \psi^{\leftarrow}(\log n), \quad n \rightarrow \infty. \tag{5.33}$$

Now we are ready to prove (5.31). By the uniform convergence of regularly varying sequences,

$$\frac{R_{q_m}}{R_{w_m}} \sim \frac{\psi^{\leftarrow}(\log q_m)}{\psi^{\leftarrow}(\log w_m)} \sim \left(\frac{\log q_m}{\log w_m} \right)^{1/\tau}, \quad m \rightarrow \infty. \tag{5.34}$$

Notice that

$$\frac{u_m}{u_{m+1}} \leq \frac{q_m}{w_m} \leq \frac{u_{m+1}}{u_m},$$

so that $u_{m+1}/u_m \rightarrow 1$ as $m \rightarrow \infty$ (see Eq. 5.18), and hence, $q_m/w_m \rightarrow 1$, $m \rightarrow \infty$. Now, Eq. 5.34 implies $R_{q_m}/R_{w_m} \rightarrow 1$ as $m \rightarrow \infty$. Recalling also $a \in \text{RV}_{1-\tau}$ and using the uniform convergence of regularly varying sequences,

$$\frac{a(R_{q_m})}{a(R_{w_m})} \sim \left(\frac{R_{q_m}}{R_{w_m}} \right)^{1-\tau} \rightarrow 1, \quad m \rightarrow \infty;$$

hence, Eq. 5.31 is obtained.

Turning to Eq. 5.32, we note that by Eq. 2.5,

$$\begin{aligned} \binom{u_{m+1}}{2k+1} f(R_{q_m})^{2k+1} &\sim \frac{1}{(2k+1)!} \left(\frac{u_{m+1}}{q_m} \right)^{2k+1} (q_m f(R_{q_m}))^{2k+1} \\ &\sim \frac{\xi^{2k+1}}{(2k+1)!} \left(\frac{u_{m+1}}{q_m} \right)^{2k+1} \rightarrow \frac{\xi^{2k+1}}{(2k+1)!}, \quad m \rightarrow \infty, \end{aligned}$$

where the last convergence is obtained as a result of Eq. 5.18.

Returning to Eq. 5.30, let us next calculate the limits for each term in the integral, while finding their appropriate upper bounds. Under our setup, it is straightforward to see that $a' \in \text{RV}_{-\tau}$ (see, e.g., Proposition 2.5 in Resnick 2007). Therefore, $a(z)/z \rightarrow 0$ as $z \rightarrow \infty$, and for all $\rho > 0$,

$$\left(1 + \frac{a(R_{q_m})}{R_{q_m}} \rho \right)^{d-1} \rightarrow 1, \quad m \rightarrow \infty. \quad (5.35)$$

Note also that Eq. 5.35 is bounded by $2(1 \vee \rho)^{d-1}$ for sufficiently large m .

Next we deal with $f(R_{q_m} + a(R_{q_m})\rho)/f(R_{q_m})$. Write

$$\begin{aligned} \frac{f(R_{q_m} + a(R_{q_m})\rho)}{f(R_{q_m})} &= \exp \{ -\psi(R_{q_m} + a(R_{q_m})\rho) + \psi(R_{q_m}) \} \quad (5.36) \\ &= \exp \left\{ - \int_0^\rho \frac{a(R_{q_m})}{a(R_{q_m} + a(R_{q_m})r)} dr \right\}. \end{aligned}$$

By the uniform convergence of regularly varying functions and $a(z)/z \rightarrow 0$ as $z \rightarrow \infty$, we have for every $r \geq 0$ that

$$\frac{a(R_{q_m})}{a(R_{q_m} + a(R_{q_m})r)} \rightarrow 1, \quad m \rightarrow \infty.$$

Therefore, for every $\rho > 0$,

$$\frac{f(R_{q_m} + a(R_{q_m})\rho)}{f(R_{q_m})} \rightarrow e^{-\rho}, \quad m \rightarrow \infty.$$

Additionally, we define a sequence $(s_\ell(m), \ell \geq 0, m \geq 0)$ by

$$s_\ell(m) = \frac{\psi^\leftarrow(\psi(R_{q_m}) + \ell) - R_{q_m}}{a(R_{q_m})},$$

equivalently, $\psi(R_{q_m} + a(R_{q_m})s_\ell(m)) = \psi(R_{q_m}) + \ell$. Then, Lemma 5.2 in Balkema and Embrechts (2004) implies that for any $\epsilon \in (0, d^{-1})$, there exists a positive integer $N = N(\epsilon)$ such that $s_\ell(m) \leq \epsilon^{-1}e^{\ell\epsilon}$ for all $m \geq N$ and $\ell \geq 0$. Since ψ is increasing,

we can establish the bound of Eq. 5.36 as follows: for $m \geq N$,

$$\begin{aligned} & \exp \left\{ -\psi \left(R_{q_m} + a(R_{q_m})\rho \right) + \psi(R_{q_m}) \right\} \mathbf{1} \{ \rho > 0 \} \\ &= \sum_{\ell=0}^{\infty} \mathbf{1} \{ s_{\ell}(m) < \rho \leq s_{\ell+1}(m) \} \exp \left\{ -\psi \left(R_{q_m} + a(R_{q_m})\rho \right) + \psi(R_{q_m}) \right\} \\ &\leq \sum_{\ell=0}^{\infty} \mathbf{1} \left\{ 0 < \rho \leq \epsilon^{-1} e^{(\ell+1)\epsilon} \right\} e^{-\ell}. \end{aligned}$$

We now discuss the final untreated term from the integral in Eq. 5.30. Let us give a helpful fact about $\| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \|$ for $i \in \{1, \dots, 2k\}$. We have that

$$\begin{aligned} & \| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \| - (R_{q_m} + a(R_{q_m})\rho + \langle \theta, y_i \rangle) \\ &= \frac{\| y_i \|^2 - \langle \theta, y_i \rangle^2}{\| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \| + R_{q_m} + a(R_{q_m})\rho + \langle \theta, y_i \rangle} =: \gamma_m(\rho, \theta, y_i). \end{aligned}$$

In particular, if $\| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \| \geq R_{q_m}$, then

$$|\gamma_m(\rho, \theta, y_i)| \leq \frac{|\| y_i \|^2 - \langle \theta, y_i \rangle^2|}{2R_{q_m} + \langle \theta, y_i \rangle} \rightarrow 0, \quad m \rightarrow \infty. \quad (5.37)$$

This convergence takes place uniformly for $\rho > 0$, $\theta \in S^{d-1}$, and $y_i \in \mathbb{R}^d$ with $\| y_i \| \leq ct$, where c is determined by property **(H3)** of h —see Section 2. Continuing onward, let

$$A_m = \left\{ y \in \mathbb{R}^d : \| (R_{q_m} + a(R_{q_m})\rho) \theta + y \| \geq R_{q_m} \right\};$$

then, for each $i \in \{1, \dots, 2k\}$,

$$\begin{aligned} & \frac{f(\| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \|)}{f(R_{q_m})} \mathbf{1}_{A_m}(y_i) \\ &= \exp \left\{ -\psi \left(R_{q_m} + a(R_{q_m})\rho + \langle \theta, y_i \rangle + \gamma_m(\rho, \theta, y_i) \right) + \psi(R_{q_m}) \right\} \mathbf{1}_{A_m}(y_i) \\ &= \exp \left\{ -\int_0^{\rho + \xi_m(\rho, \theta, y_i)} \frac{a(R_{q_m})}{a(R_{q_m} + a(R_{q_m})r)} dr \right\} \mathbf{1}_{A_m}(y_i), \end{aligned}$$

where $\xi_m(\rho, \theta, y_i) = a(R_{q_m})^{-1} (\langle \theta, y_i \rangle + \gamma_m(\rho, \theta, y_i))$. Note that the last term is bounded by 1, due to the fact that

$$\| (R_{q_m} + a(R_{q_m})\rho) \theta + y_i \| \geq R_{q_m} \Leftrightarrow \rho + \xi_m(\rho, \theta, y_i) \geq 0.$$

Additionally, Eqs. 4.2 and 5.37 yield that

$$\xi_m(\rho, \theta, y_i) \rightarrow \zeta^{-1} \langle \theta, y_i \rangle, \quad m \rightarrow \infty,$$

for all $\rho > 0$, $\theta \in S^{d-1}$, and $y_i \in \mathbb{R}^d$. Thus, as $m \rightarrow \infty$,

$$\exp \left\{ -\int_0^{\rho + \xi_m(\rho, \theta, y_i)} \frac{a(R_{q_m})}{a(R_{q_m} + a(R_{q_m})r)} dr \right\} \rightarrow \exp \left\{ -\rho - \zeta^{-1} \langle \theta, y_i \rangle \right\},$$

and

$$\mathbf{1}_{A_m}(y_i) \rightarrow \mathbf{1}\left\{\rho + \zeta^{-1}\langle\theta, y_i\rangle \geq 0\right\}.$$

Combining all the bounds derived thus far, the integral in Eq. 5.30 is bounded above by

$$\begin{aligned} & 2 \int_0^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) (1 \vee \rho)^{d-1} \sum_{\ell=0}^\infty \mathbf{1}\left\{0 < \rho \leq \epsilon^{-1} e^{(\ell+1)\epsilon}\right\} \\ & e^{-\ell} d\mathbf{y} d\nu_{d-1}(\theta) d\rho \\ & = C^* \int_0^\infty \sum_{\ell=0}^\infty \mathbf{1}\left\{0 < \rho \leq \epsilon^{-1} e^{(\ell+1)\epsilon}\right\} e^{-\ell} (1 \vee \rho)^{d-1} d\rho \\ & \leq C^* \left(\frac{e^\epsilon}{\epsilon}\right)^d \sum_{\ell=0}^\infty e^{-(1-\epsilon d)\ell} < \infty, \end{aligned}$$

as $\epsilon^{-1} e^{(\ell+1)\epsilon} \geq 1$ and $\epsilon d < 1$. Now, by the dominated convergence theorem, we can see that the integral in Eq. 5.30 converges to

$$\int_0^\infty \int_{S^{d-1}} \int_{(\mathbb{R}^d)^{2k}} h_t(0, \mathbf{y}) e^{-(2k+1)\rho - \zeta^{-1} \sum_{i=1}^{2k} \langle\theta, y_i\rangle} \prod_{i=1}^{2k} \mathbf{1}\left\{\rho + \zeta^{-1}\langle\theta, y_i\rangle \geq 0\right\} d\mathbf{y} d\nu_{d-1}(\theta) d\rho.$$

Because of this convergence, as well as Eqs. 5.31 and 5.32, we can get Eq. 5.26 as required.

The proof of Eq. 5.27 is almost the same as above, so we skip it. We can now conclude this lemma by showing Eqs. 5.28 and 5.29. However, the proof has been omitted as it has essentially the same character as the proofs of Eqs. 5.12 and 5.13 in Lemma 5.1—albeit with different upper bounds as discussed above, due to the differing nature of the tail of the probability densities. $\square$

Proof Theorem 4.1 First, for every $n \geq 1$,

$$\frac{U_m(t)}{a(R_{v_m})R_{v_m}^{d-1}} \leq \frac{\chi_n^{(1)}(t)}{a(R_n)R_n^{d-1}} \leq \frac{T_m(t)}{a(R_{w_m})R_{w_m}^{d-1}}.$$

Lemma 5.2 yields that

$$\limsup_{n \rightarrow \infty} \frac{\chi_n^{(1)}(t)}{a(R_n)R_n^{d-1}} \leq K_1(t) + \limsup_{m \rightarrow \infty} \frac{T_m(t) - \mathbb{E}[T_m(t)]}{a(R_{w_m})R_{w_m}^{d-1}},$$

and

$$\liminf_{n \rightarrow \infty} \frac{\chi_n^{(1)}(t)}{a(R_n)R_n^{d-1}} \geq K_1(t) + \liminf_{m \rightarrow \infty} \frac{U_m(t) - \mathbb{E}[U_m(t)]}{a(R_{v_m})R_{v_m}^{d-1}}.$$

Now, the proof can be finished, if one can show that

$$\left[a(R_{w_m})R_{w_m}^{d-1}\right]^{-1} (T_m(t) - \mathbb{E}[T_m(t)]) \rightarrow 0, \quad m \rightarrow \infty, \quad \text{a.s.}, \quad (5.38)$$

$$\left[a(R_{v_m})R_{v_m}^{d-1}\right]^{-1} (U_m(t) - \mathbb{E}[U_m(t)]) \rightarrow 0, \quad m \rightarrow \infty, \quad \text{a.s.} \quad (5.39)$$

By Eq. 5.28 in Lemma 5.2, for every $\epsilon > 0$,

$$\begin{aligned} \sum_{m=1}^{\infty} \mathbb{P} \left(|T_m(t) - \mathbb{E}[T_m(t)]| > \epsilon a(R_{w_m}) R_{w_m}^{d-1} \right) &\leq \frac{1}{\epsilon^2} \sum_{m=1}^{\infty} \frac{\text{Var}(T_m(t))}{\left(a(R_{w_m}) R_{w_m}^{d-1}\right)^2} \\ &\leq C^* \sum_{m=1}^{\infty} \frac{1}{a(R_{w_m}) R_{w_m}^{d-1}}. \end{aligned}$$

Because of the constraint in Eq. 5.25, there exist $\delta_i > 0$, $i = 1, 2$, so that

$$\gamma(d - \tau - \delta_1) \left(\frac{1}{\tau} - \delta_2 \right) > 1.$$

Then, $a \in \text{RV}_{1-\tau}$ implies that

$$a(R_{w_m}) R_{w_m}^{d-1} \geq C^* R_{w_m}^{d-\tau-\delta_1}$$

for all $m \geq 1$. Note that by Eq. 5.33,

$$R_{w_m} \geq C^* \psi^{\leftarrow}(\log w_m) \geq C^* \psi^{\leftarrow}(\log u_m) \geq C^* m^{\gamma(1/\tau-\delta_2)}$$

again for all $m \geq 1$. Therefore,

$$a(R_{w_m}) R_{w_m}^{d-1} \geq C^* m^{\gamma(d-\tau-\delta_1)(1/\tau-\delta_2)},$$

and

$$\sum_{m=1}^{\infty} \frac{1}{a(R_{w_m}) R_{w_m}^{d-1}} \leq C^* \sum_{m=1}^{\infty} \frac{1}{m^{\gamma(d-\tau-\delta_1)(1/\tau-\delta_2)}} < \infty.$$

Now, the Borel-Cantelli lemma completes the proof of Eq. 5.38. The proof of Eq. 5.39 is the same by virtue of Eq. 5.29 in Lemma 5.2.

Before concluding the proof, we show finiteness of the limit in Eq. 4.3. Using property (H3) of h ,

$$\begin{aligned} \left| \sum_{k=0}^{\infty} (-1)^k s_k(t) \right| &\leq \sum_{k=0}^{\infty} \frac{s_{d-1} \xi^{k+1}}{(k+1)!} \int_0^{\infty} \int_{(\mathbb{R}^d)^k} \prod_{i=1}^k \mathbf{1}\{\|y_i\| \leq ct\} e^{-\rho} \mathbf{d}y \, d\rho \\ &\leq C^* \sum_{k=0}^{\infty} \frac{((ct)^d \xi \omega_d)^k}{k!} = C^* e^{(ct)^d \xi \omega_d} < \infty. \end{aligned}$$

□

Acknowledgements The authors would like to thank the anonymous referee and the Associate Editor for their helpful insights, which have made the paper much more accessible. This research is partially supported by the National Science Foundation (NSF) grant, Division of Mathematical Sciences (DMS), #1811428.

References

- Adler, R.J., Bobrowski, O., Borman, M.S., Subag, E., Weinberger, S., et al: Persistent homology for random fields and complexes. In: Borrowing Strength: Theory Powering Applications—a Festschrift for Lawrence D. Brown, Institute of Mathematical Statistics, pp. 124–143 (2010)

- Adler, R.J., Bobrowski, O., Weinberger, S.: Crackle: The homology of noise. *Discret. Comput. Geom.* **52**(4), 680–704 (2014). <https://doi.org/10.1007/s00454-014-9621-6>
- Balkema, G., Embrechts, P.: Multivariate excess distributions. *ETHZ Preprint* (2004)
- Bobrowski, O., Adler, R.J.: Distance functions, critical points, and the topology of random Čech complexes. *Homology Homotopy Appl.* **16**, 311–344 (2014)
- Bobrowski, O., Borman, M.S.: Euler integration of Gaussian random fields and persistent homology. *J. Topol. Anal.* **4**(1), 49–70 (2012)
- Bobrowski, O., Mukherjee, S.: The topology of probability distributions on manifolds. *Probab. Theory Relat. Fields* **161**, 651–686 (2015)
- Carlsson, G.: Topology and data. *Bull. Am. Math. Soc.* **46**, 255–308 (2009)
- Edelsbrunner, H., Harer, J.L.: *Computational Topology. An Introduction*. American Mathematical Society (2010)
- Embrechts, P., Klüppelberg, C., Mikosch, T.: *Modelling Extremal Events: for Insurance and Finance*. Springer, New York (1997)
- Ghrist, R.: Barcodes: The persistent topology of data. *Bull. Am. Math. Soc.* **45**, 61–75 (2008)
- Ghrist, R.: *Elementary Applied Topology*. Createspace (2014)
- Goel, A., Trinh, K., Tsunoda, K.: Strong law of large numbers for Betti numbers in the thermodynamic regime. *J. Stat. Phys.* **174**(4), 865–892 (2019). <https://doi.org/10.1007/s10955-018-2201-z>
- Hug, D., Last, G., Schulte, M.: Second-order properties and central limit theorems for geometric functionals of Boolean models. *Ann. Probab.* **26**, 73–135 (2016)
- Krebs, J., Roycraft, B., Polonik, W.: On approximation theorems for the Euler characteristic with applications to the bootstrap. *arXiv:200507557* (2020)
- Niyogi, P., Smale, S., Weinberger, S.: Finding the homology of submanifolds with high confidence from random samples. *Discret. Comput. Geom.* **39**, 419–441 (2008). <https://doi.org/10.1007/s00454-008-9053-2>
- Niyogi, P., Smale, S., Weinberger, S.: A topological view of unsupervised learning from noisy data. *SIAM J. Comput.* **40**(3), 646–663 (2011). <https://doi.org/10.1137/090762932>
- Owada, T.: Functional central limit theorem for subgraph counting processes. *Electron. J. Probab.* **22**, 38 (2017). <https://doi.org/10.1214/17-EJP30>
- Owada, T.: Limit theorems for Betti numbers of extreme sample clouds with application to persistence barcodes. *Ann. Appl. Probab.* **28**(5), 2814–2854 (2018)
- Owada, T.: Topological crackle of heavy-tailed moving average processes. *Stoch. Process. Appl.* **129**, 4965–4997 (2019)
- Owada, T., Adler, R.J.: Limit theorems for point processes under geometric constraints (and topological crackle). *Ann. Probab.* **45**(3), 2004–2055 (2017). <https://doi.org/10.1214/16-AOP1106>
- Owada, T., Bobrowski, O.: Convergence of persistence diagrams for topological crackle. *Bernoulli* **26**(3), 2275–2310 (2020). <https://doi.org/10.3150/20-BEJ1193>
- Penrose, M.: *Random Geometric Graphs*, vol. 5. Oxford University Press, New York (2003)
- Resnick, S.I.: *Extreme Values Regular Variation and Point Processes*. Springer-Verlag, New York (1987)
- Resnick, S.I.: *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science & Business Media, New York (2007)
- Thomas, A.M., Owada, T.: Functional limit theorems for the Euler characteristic process in the critical regime. *Adv. Appl. Prob.* **53**(1), 57–80 (2021). <https://doi.org/10.1017/apr.2020.46>

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