



HAL
open science

Extremal life times of persistent loops and holes

Nicolas Chenavier, Christian Hirsch

► **To cite this version:**

Nicolas Chenavier, Christian Hirsch. Extremal life times of persistent loops and holes. 2020. hal-04066214

HAL Id: hal-04066214

<https://hal.science/hal-04066214>

Preprint submitted on 12 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EXTREMAL LIFE TIMES OF PERSISTENT LOOPS AND HOLES

NICOLAS CHENAVIER AND CHRISTIAN HIRSCH

ABSTRACT. Persistent homology captures the appearances and disappearances of topological features such as loops and holes when growing disks centered at a Poisson point process. We study extreme values for the life times of features dying in bounded components and with birth resp. death time bounded away from the threshold for continuum percolation. First, we describe the scaling of the minimal life times for general feature dimensions, and of the maximal life times for holes in the Čech complex. Then, we proceed to a more refined analysis and establish Poisson approximation for large life times of holes and for small life times of loops. Finally, we also study the scaling of minimal life times in the Vietoris-Rips setting and point to a surprising difference to the Čech complex.

1. INTRODUCTION

One of the core challenges in statistics is to provide scientifically rigorous methods to detect structure among noise. While for low-dimensional datasets, we do indeed have a versatile toolkit at our disposal, the situation is radically different in high dimensions. In this context, *topological data analysis (TDA)* is emerging as an exciting novel approach. The simple and effective idea is to leverage invariants from algebraic topology to extract insights from data. While initially TDA started off as an eccentric mathematical idea, it is now applied in a wide variety of disciplines, ranging from astronomy and biology to finance and materials science [8, 16, 17, 18].

A core component of TDA is *persistent homology*. Here, we start growing balls centered at the points of a dataset. At the beginning, all points are in isolated components, but as the balls grow, new topological features such as loops or holes may appear. As the radii increase, these features disappear again, until finally everything is contained in a single connected component without loops or holes. Through this mechanism, we can associate a birth time and a death time with each such feature.

Despite the spectacular successes of TDA in the application domains, major parts of it are still lacking a rigorous statistical foundation. For instance, when analyzing persistent homology, practitioners often look for features living for exceptionally long periods of time, and then draw conclusions if they do occur. However, how can we decide whether the observed long life times come from genuinely interesting phenomena and are not a mere incarnation of chance?

To answer this question, we need to understand how long life times behave under the null model of complete spatial randomness, i.e., a Poisson point process. Figure 1 illustrates cycles of a maximal, typical and minimal life times in a bounded sampling window. In this regard, a breakthrough was achieved in [4], where the order of the longest life time occurring in a large window was determined on a multiplicative scale. Despite the significance of this discovery, it is still difficult to build a rigorous statistical test, knowing only the order of the life time. A further interesting cross-connection is that for 0-dimensional features, the longest life time corresponds to the longest edge in the minimum spanning tree whose asymptotic was established in [15].

In the present paper, we move one step closer to statistical applications and establish Poisson approximation results for extremal life times of loops and holes in large sampling windows. More precisely, in Theorem 2.1 we determine the scaling of the maximal life time of $(d - 1)$ -dimensional features, i.e., holes, in the Čech filtration. Moreover, we show through the Chen-Stein method that after suitable normalization, the locations of exceedances over suitable

Date: March 8, 2021.

2010 Mathematics Subject Classification. 60K35; 82C22.

Key words and phrases. topological data analysis, persistent Betti numbers, Poisson approximation.

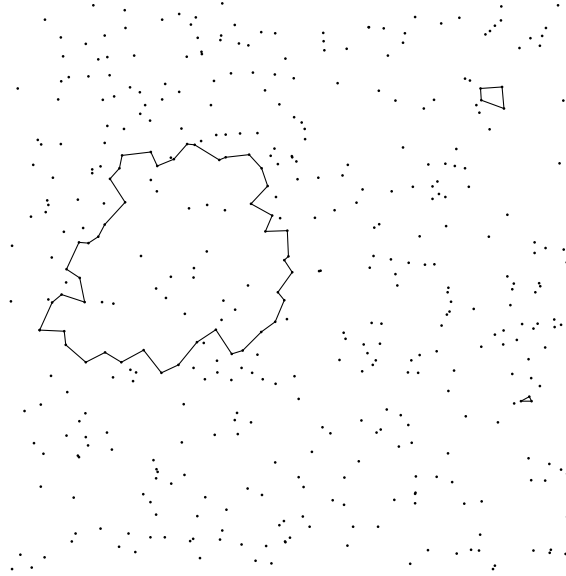


FIGURE 1. Čech-features of a maximal (left), typical (right top) and minimal (right bottom) life times.

thresholds converge to a homogeneous Poisson process. Thus, the investigation is in a similar spirit as the extreme value analysis of geometric characteristics of inradii in tessellations [6, 7]. However, after proceeding from the general set-up of the Chen-Stein method to more concrete tasks, the geometric analyses of the two problems are entirely different.

Another vigorous research stream in spatial extremes concerns limit results for exceptionally small structures [19]. Our second and third main results, concern this setting. More precisely, we determine the minimal life time of features in any dimension both in the Vietoris-Rips (VR) and in the Čech (\check{C}) filtration. Our analysis reveals a striking difference in the scaling of minimal features between the VR- and the \check{C} -filtration. Often the VR-filtration is chosen as an approximation to the \check{C} -filtration, when the latter becomes prohibitively time-consuming to compute. In this respect, our result provides a clear example where this approximation fails on a fundamental level.

The rest of the manuscript is organized as follows. First, in Section 2, we fix the notation and state the main results. Section 3 illustrates in broad strokes how the Chen-Stein method can be used to tackle the assertion. This involves two critical steps, namely identifying the correct scaling and excluding multiple exceedances, which will be established in Sections 4 and 5, respectively. To ease the flow of reading, we have postponed overly technical volume computations into an appendix.

2. MODEL AND MAIN RESULTS

In this section, we describe extremal life times in TDA. First, we review briefly the fundamental concepts of the VR- and \check{C} -filtrations, referring the reader to [5] for an excellent textbook treatment on this topic.

Both the VR- as well as the \check{C} -filtration describe collections of families of simplices in \mathbb{R}^d . A k -simplex $\{x_0, \dots, x_k\} \subset \mathbb{R}^d$ is contained in the VR-filtration at a level $r > 0$ if the distance between any pair of points is smaller than $2r$. It is included in the \check{C} -filtration if $\bigcap_{i \leq k} B_r(x_i) \neq \emptyset$, where $B_r(x_i)$ denotes the Euclidean ball with radius r centered at x_i .

Having explained these basic building blocks, we now introduce more precisely the concept behind birth and death times when dealing with Poisson point processes in the entire Euclidean space. For finite complexes there is a precise algorithmic description of when a birth occurs and when a death occurs. Each time we add a simplex to the filtration it either increases the homology in the dimension of the feature or it decreases the homology in one dimension lower

[5, Algorithm 9]. In the first case, we add a *positive simplex*, which corresponds to the *birth* of a feature. In the second case, we add a *negative simplex* corresponding to the *death* of a feature. Defining the life time $L_{i,k}$ of the i th feature as the difference between the death and the birth time, this algorithm gives rise to a collection of birth-and-death times $\{L_{i,k}\}_{i \geq 1}$. Assigning to the i th feature the center of gravity $Z_{i,k}$ of the simplex at its death time, we may therefore consider features as a marked point process.

Let $\mathcal{P} := \{X_i\}_{i \geq 1}$ be a Poisson point process in \mathbb{R}^d with intensity 1 and $W_n := [-n^{1/d}/2, n^{1/d}/2]^d$ be a sampling window of volume n . In the \check{C} -filtration, persistent homology tracks the topological changes of the Boolean model

$$O_r(\mathcal{P}) := \bigcup_{i \geq 1} B_r(X_i)$$

for increasing values of the radius r . For instance, 1-dimensional features are loops, and $(d-1)$ -dimensional features correspond to holes, i.e., to bounded connected components in the vacant set $\mathbb{R}^d \setminus O_r(\mathcal{P})$.

Before moving to the general setting, we first describe life times for the specific case of $(d-1)$ -dimensional features. When the growing radii cause an existing hole to split into two at a radius r , we say that r is the *birth time* of the new hole. Moreover, the death time of a hole is the smallest radius $r > 0$ when it is covered completely by the Boolean model $O_r(\mathcal{P})$. So far, this definition is still ambiguous: when a new hole is born, it is not clear, which of the newly created vacant components should determine its death time. Therefore, it is a common convention in TDA to impose that features born first die last.

Now, we enumerate the holes as $\{H_i^*\}_{i \geq 1}$ and associate with each such hole its life time $L_i^* > 0$ and the point $Z_i^* \in \mathbb{R}^d$ as the last point that is covered at the death time. In this way, $\{(Z_i^*, L_i^*)\}_{i \geq 1}$ defines a stationary marked point process of holes together with their life times.

Long-range interactions coming from percolation effects make it difficult to study this marked point process in full generality. For holes born in large vacant components, the seemingly innocent convention that features born first die last may induce long-range interactions in order to decide which birth times correspond to which death times. This makes analyzing extremes prohibitively challenging: we are not aware of any form of Poisson approximation reflecting critical phenomena in percolation contexts.

To state a clean Poisson-approximation theorem that applies to the bulk of the considered holes, we restrict $\{H_i^*\}_{i \geq 1}$ to a smaller family of holes $\{H_i\}_{i \geq 1}$ as follows. Let $r_c^V > 0$ denote the critical threshold of vacant continuum percolation and fix a time window $[r_c^V - \varepsilon_c, r_c^V + \varepsilon_c]$ around r_c^V . Then, we consider only holes with birth time outside $[r_c^V - \varepsilon_c, r_c^V + \varepsilon_c]$. As made precise in Section 3 below, the striking advantage of this restriction is that the bounded connected components are of logarithmic size [13].

In the first main result, we describe the limiting behavior of exceedances of life times in the \check{C} -filtration centered in the sampling window W_n . To do it, let \bar{F}_ζ be the survival function of the typical life time of a hole in the \check{C} -filtration, i.e.,

$$\bar{F}_\zeta(\ell) = \mathbb{E}[\#\{Z_i \in W_1 : L_i \geq \ell\}]. \quad (1)$$

We show in Appendix A that the function $\bar{F}_\zeta : (0, \infty) \rightarrow (0, 1)$ is invertible. In what follows, we deal with the asymptotic behavior of the rescaled *point process of exceedances*

$$\Phi_{n, \check{C}}^{\max} := \{(n^{-1/d} Z_i, n \bar{F}_\zeta(L_i))\}_{Z_i \in W_n} \subset W_1 \times (0, \infty).$$

We consider this point process because it can capture both the location as well as the magnitude of large life times of features.

Our first main result shows that the maximal life time of holes in the \check{C} -filtration is **asymptotically equivalent to**

$$\gamma_n := (\kappa_d^{-1} \log n)^{1/d},$$

and in particular is asymptotically independent of the feature dimension $1 \leq k \leq d-1$, where κ_d is the volume of the unit ball in \mathbb{R}^d . Moreover, $\Phi_{n, \check{C}}^{\max}$ is approximately a homogeneous Poisson

point process Φ with intensity 1 in $W_1 \times (0, \infty)$. We write \Rightarrow for convergence in distribution. In the next theorems, we estimate quantities of the form $\bar{F}^{-1}(\tau/n)$. Considering this type of quantities is classical in Extreme Value Theory. Indeed, if $\{Y_i\}_i$ denotes a family of iid random variables with survival function \bar{F} , then $\mathbb{P}(\max\{Y_1, \dots, Y_n\} \leq u_n) \rightarrow \exp(-\tau)$ if and only if $n\bar{F}(u_n) \rightarrow \tau$.

Theorem 2.1 (Maximal life times, \check{C} -filtration). *For every $\tau > 0$,*

$$\lim_{n \rightarrow \infty} \frac{\bar{F}_{\check{C}}^{-1}(\tau/n)}{\gamma_n} = 1 \quad \text{and} \quad \Phi_{n, \check{C}}^{\max} \Rightarrow \Phi.$$

As a corollary, we obtain the scaling of the maximal life time.

Corollary 2.2 (Scaling of maximal life times – \check{C} -filtration). *It holds that*

$$\gamma_n^{-1} \max_{Z_i \in W_n} L_i \xrightarrow{n \rightarrow \infty} 1$$

in probability.

Proof. First, $\max_{Z_i \in W_n} L_i \geq (1 + \varepsilon)\gamma_n$ means that $\{n\bar{F}_{\check{C}}(L_i)\}_{Z_i \in W_n} \cap [0, n\bar{F}_{\check{C}}((1 + \varepsilon)\gamma_n)] \neq \emptyset$. Now, we claim that $n\bar{F}_{\check{C}}((1 + \varepsilon)\gamma_n)$ tends to 0. Indeed, otherwise there exists $\tau > 0$ such that $\bar{F}_{\check{C}}^{-1}(\tau/n) \geq (1 + \varepsilon)\gamma_n$ for infinitely many n , contradicting the scaling in Theorem 2.1. Thus, we conclude by the point-process convergence in Theorem 2.1.

Similarly, $\max_{Z_i \in W_n} L_i \leq (1 - \varepsilon)\gamma_n$ means that $\{n\bar{F}_{\check{C}}(L_i)\}_{Z_i \in W_n} \cap [0, n\bar{F}_{\check{C}}((1 - \varepsilon)\gamma_n)] = \emptyset$. We claim that $n\bar{F}_{\check{C}}((1 - \varepsilon)\gamma_n)$ tends to ∞ . Otherwise, there exists $\tau < \infty$ such that $\bar{F}_{\check{C}}^{-1}(\tau/n) \leq (1 - \varepsilon)\gamma_n$ for infinitely many n , again contradicting Theorem 2.1, and we conclude by the convergence in Theorem 2.1. \square

Remark 2.3. We stress that the point-process convergence in Theorem 2.1 gives much more than just the scaling of the maximal life time: it allows us to describe the joint distribution of life times exceeding different thresholds. A downside in our definition of $\Phi_{n, \check{C}}^{\max}$ is that the normalization of life times of the form $n\bar{F}_{\check{C}}(L_i)$ is rather implicit. This is because even approximating the survival function of the life time of a typical hole sufficiently accurately is challenging. If this could be achieved, then one could proceed as in [6] to obtain a more explicit transformation. In that sense, the leading-order scaling derived in Theorem 2.1 is a promising first step, but a finer analysis would be highly desirable.

Remark 2.4. Theorem 2.1 reveals a striking difference to the multiplicative persistence approach of measuring life times in [4], where the scaling is $(\log n / \log \log n)^{1/(d-1)} \gg \gamma_n$. That work considers the ratio between the death time and the birth time, whereas we work with the difference. The two scalings are different as the multiplicative approach is sensitive with respect to cycles with a very small birth time. Depending on the application context this may or may not be desirable. Hence, it is valuable to know the scaling of both variants and especially how they differ.

Next, we proceed to minimal life times, where we also derive Poisson approximation results. Similarly to the setting of holes, also for general feature dimensions, the long-range correlations close to the critical radius $r_c^{\mathcal{O}}$ of continuum percolation of the Boolean model would pose a formidable obstacle for proving Poisson approximation. Therefore, we henceforth fix again a time window $[r_c^{\mathcal{O}} - \varepsilon_c, r_c^{\mathcal{O}} + \varepsilon_c]$ and consider only features with death time outside $[r_c^{\mathcal{O}} - \varepsilon_c, r_c^{\mathcal{O}} + \varepsilon_c]$ and that die in a bounded connected component. Furthermore, for instance by taking the center point of the $(k + 1)$ -simplex whose insertion kills the k -dimensional feature, we attach to any such feature with life time $L_{i,k}$ a center point $Z_{i,k} \in \mathbb{R}^d$. Hence, we again obtain a stationary marked point process $\{(Z_{i,k}, L_{i,k})\}_{i \geq 1}$.

Fixing a feature dimension $1 \leq k \leq d - 1$, we define the rescaled point process of undershoots

$$\Phi_{n,k,\check{C}}^{\min} := \{(n^{-1/d} Z_{i,k}, nF_{k,\check{C}}(L_{i,k}))\}_{Z_{i,k} \in W_n} \subset W_1 \times (0, \infty),$$

where

$$F_{k,\check{\zeta}}(\ell) = \mathbb{E}[\#\{Z_{i,k} \in W_1 : L_{i,k} \leq \ell\}]. \quad (2)$$

denotes the distribution function of the typical life time in the \check{C} -filtration. In the next result, we show that the minimal life time of k -dimensional features is of order n^{-2} , independent of k . Moreover, the rescaled point process of undershoots is approximately a homogeneous Poisson point process with intensity 1 in $W_1 \times (0, \infty)$. However, again due to involved geometrical configurations occurring in the \check{C} -filtration, we prove the point-process convergence only for a specific feature dimension, namely $k = 1$ (loops).

Theorem 2.5 (Minimal life times, \check{C} -filtration). *Let $1 \leq k \leq d - 1$. Then, for every $\tau > 0$,*

(i)

$$\lim_{n \rightarrow \infty} \frac{\log F_{k,\check{C}}^{-1}(\tau/n)}{\log n} = -2 \quad \text{and} \quad (\log n)^{-1} \min_{Z_{i,k} \in W_n} \log L_{i,k} \Rightarrow -2.$$

(ii) Moreover, $\Phi_{n,k,\check{\zeta}}^{\min} \Rightarrow \Phi$ for $k = 1$.

The construction of $\Phi_{n,k,\check{\zeta}}^{\min}$ outlined above does not only work for the \check{C} - but also for the VR-filtration $\Phi_{n,k,\text{VR}}^{\min}$. The difference in filtration manifests itself prominently in a different scaling for the threshold.

Theorem 2.6 (Minimal life times – VR-filtration). *Let $1 \leq k \leq d - 1$. Then, for every $\tau > 0$,*

$$\lim_{n \rightarrow \infty} \frac{\log F_{k,\text{VR}}^{-1}(\tau/n)}{\log n} = -1 \quad \text{and} \quad \Phi_{n,k,\text{VR}}^{\min} \Rightarrow \Phi.$$

As a corollary, we obtain again the scaling of the minimal life time.

Remark 2.7. A priori, it is not clear whether typical life times in the VR-filtration should be shorter or longer than those in the \check{C} -filtration. However, the specific construction in the proof of Theorem 2.1 suggests it seems plausible that asymptotically the corresponding life-time scale $\bar{F}_{d-1,\text{VR}}^{-1}(\tau/n)$ should be smaller than $\bar{F}_{\check{\zeta}}^{-1}(\tau/n)$ by a constant factor. Loosely speaking, the characteristic of persistent features is a very late death time rather than a very early birth time, and features are dying later in the \check{C} - than in the VR-filtration.

3. THE CHEN-STEIN METHOD

The proofs of the Poisson approximation in Theorems 2.1–2.6 rely fundamentally on the Chen-Stein method in the form of [1, Theorem 1]. To make the presentation self-contained, we sketch now how to tune the general machinery to the setting of extremal life times. The delicate problem-specific conditions are then verified in Sections 4 and 5.

More precisely, proceeding as in [7], we discretize the problem to the Chen-Stein framework in the form of [1, Theorem 1]. For the convenience of the reader, we outline the most important steps for the maximal life times, noting that the arguments carry over seamlessly to the setting of minimal life times.

The first central ingredient is to extract a dependency graph with bounded degrees. It is at this point that the percolation conditions described in Section 2 enter the stage. Loosely speaking, we introduce a discretization that is coarse enough so that the life times of holes in non-adjacent blocks are independent. Hence, the size of the discretization is essentially determined by the diameter of the largest hole that we expect to see in the window W_n . However, due to the convention that features born first should die last, the actual block size that needs to be considered is a bit larger. More precisely, writing $S(x, r)$ for the diameter of the connected component of the vacant set $\mathbb{R}^d \setminus O_r(\mathcal{P})$ containing $x \in \mathbb{R}^d$, we set

$$S_i := \sup_{r \notin [r_c^v - \varepsilon_c, r_c^v + \varepsilon_c]}: S(Z_i, r) < \infty S(Z_i, r)$$

as the diameter of the largest bounded hole containing Z_i .

Now, in Lemma B.1 in the appendix, we show that when defining the event

$$E_n^{\vee} := \{\text{for every radius } r \notin [r_c^{\vee} - \varepsilon_c, r_c^{\vee} + \varepsilon_c] \text{ all vacant bounded connected components centered in } W_n \text{ have diameter at most } (\log n)^2\},$$

then, for all large n ,

$$\mathbb{P}(E_n^{\vee}) \geq 1 - n^{-4d}. \quad (3)$$

Notice that, on the event E_n^{\vee} for each hole with $Z_i \in W_n$, we have $L_i < (\log n)^2$.

Therefore, we partition the sampling window W_n into $\mathcal{N}_n \in O(n/(\log n)^{2d})$ sub-boxes $\{\mathbf{C}_j\}_{j \leq \mathcal{N}_n}$ of side length

$$s_n := 2(\log n)^2.$$

To prove that the point process of exceedances converges to a Poisson point process, we have to estimate the number of **exceedances**. To do it, we first give some notation. For each $j \leq \mathcal{N}_n$ and $B \subset W_1$, we write

$$M_{j,n}(B) := \max_{Z_i \in \mathbf{C}_j \cap (n^{1/d}B)} \{L_i : S_i \leq (\log n)^2\}$$

for the maximum life time of a hole in \mathbf{C}_j with diameter-bound at most $(\log n)^2$. Finally, we denote by

$$N_n(\tau, B) := \Phi_{n, \check{\zeta}}^{\max}(B \times (0, \tau)) \quad \text{and} \quad N'_n(\tau, B) := \#\{j \leq \mathcal{N}_n : n\bar{F}_{\check{\zeta}}(M_{j,n}(B)) \leq \tau\}.$$

the number of exceedances w.r.t. B (resp. the number of sub-boxes intersected by $n^{1/d}B$ for which there exists at least one exceedance with diameter-bound lower than $(\log n)^2$). Setting $s_{n,+} = (3s_n)^d$, the main difficulty is to exclude multiple exceedances in small boxes of size $s_{n,+}$.

Proposition 3.1 (No multiple exceedances). *Let $\tau > 0$.*

- (i) *Let $q_n := \mathbb{P}(N_n(\tau, W_{s_{n,+}}) \geq 2)$. Then $q_n \in O(n^{-1})$.*
- (ii) *The same property holds for the minimal life time in the VR-filtration with $1 \leq k \leq d-1$, and for the minimal lifetime in the \check{C} -filtration with $k = 1$.*

We are now prepared to prove the Poisson approximation in Theorem 2.1.

Theorem 2.1, Poisson approximation. Let $\tau > 0$ and $B \subset W_1$ Borel. According to Kallenberg's theorem (see e.g. Theorem A.1, p.309 in [12]), it is sufficient to prove that,

$$\mathbb{E}[N_n(\tau, B)] = \tau|B| \quad (4)$$

and

$$\mathbb{P}(N_n(\tau, B) = 0) \xrightarrow[n \rightarrow \infty]{} e^{-\tau|B|}. \quad (5)$$

Equation (4) holds since

$$\begin{aligned} \mathbb{E}[N_n(\tau, B)] &= \mathbb{E}[\#\{Z_i \in n^{1/d}B : n\bar{F}_{\check{\zeta}}(L_i) < \tau\}] \\ &= n|B|\mathbb{E}[\#\{Z_i \in W_1 : L_i > (\bar{F}_{\check{\zeta}})^{-1}(\tau/n)\}] \\ &= n|B|\bar{F}_{\check{\zeta}}((\bar{F}_{\check{\zeta}})^{-1}(\tau/n)) \\ &= \tau|B|. \end{aligned}$$

To deal with (5), we first notice that

$$\begin{aligned} \mathbb{P}(N_n(\tau, B) \neq N'_n(\tau, B)) &\leq \mathbb{P}((E_n^{\vee})^c) + \sum_{j \leq \mathcal{N}_n} \mathbb{P}(\#\{Z_i \in \mathbf{C}_j : n\bar{F}_{\check{\zeta}}(L_i) < \tau\} \geq 2) \\ &\leq \mathbb{P}((E_n^{\vee})^c) + \mathcal{N}_n q_n, \end{aligned}$$

where the last uses that \mathcal{P} is stationary. Therefore, thanks to Equation (3) and Proposition 3.1, we have $\mathbb{P}(N_n(\tau, B) \neq N'_n(\tau, B)) \rightarrow 0$. Thus, it is enough to prove that $\mathbb{P}(N'_n(\tau, B) = 0)$ converges to $e^{-\tau|B|}$. We will actually prove that $N'_n(\tau, B)$ converges to a Poisson random variable with parameter $\tau|B|$. To do it, we denote by F_n a Poisson random variable with parameter

$\mathbb{E}[N'_n(\tau, B)]$. With the same computations as [7], we can easily prove (thanks to Proposition 3.1) that

$$\mathbb{E}[F_n] = \mathbb{E}[N'_n(\tau, B)] \underset{n \rightarrow \infty}{\sim} \mathbb{E}[N_n(\tau, B)] = \tau|B|.$$

Thus, it is sufficient to approximate $N'_n(\tau, B)$ by F_n . The main idea to do it is to apply the Chen-Stein method. Indeed, thanks to a result due to Arratia, Goldstein and Gordon (see [1, Theorem 1]), and because each box \mathbf{C}_j is adjacent to at most $3^d - 1$ other boxes, we have

$$\sup_{\mathcal{A} \subset \mathbb{Z}_{\geq 0}} |\mathbb{P}(N'_n(\tau, B) \in \mathcal{A}) - \mathbb{P}(F_n \in \mathcal{A})| \leq 3^d \left(2\tau p_n^{(1)} + \frac{2n}{(2(\log n)^2)^d} \cdot p_n^{(2)} \right), \quad (6)$$

where

$$p_n^{(1)} := \mathbb{P}(n\bar{F}_\zeta(M_{j,n}(W_1)) \leq \tau) \quad \text{and} \quad p_n^{(2)} := \mathbb{P}(n\bar{F}_\zeta(M_{j_1,n}(W_1) \wedge M_{j_2,n}(W_1)) \leq \tau)$$

denote the exceedance probabilities in a single box \mathbf{C}_j or in two adjacent boxes \mathbf{C}_{j_1} and \mathbf{C}_{j_2} . By stationarity and symmetry, notice that $p_n^{(1)}$ and $p_n^{(2)}$ do not depend on j, j_1 or j_2 . The first term of the right-hand side converges to 0 since, by stationarity,

$$p_n^{(1)} \leq \mathbb{E}[\#\{Z_i \in \mathbf{C}_j : L_i > \bar{F}_\zeta^{-1}(\tau/n)\}] = \frac{|\mathbf{C}_j|\tau}{n}.$$

The second one converges to 0 according to Proposition 3.1. This concludes the proof of Theorem 2.1. \square

In a similar way, we obtain the Poisson approximation asserted in Theorem 2.5 and Theorem 2.6, relying on a correspondingly defined $E_n^{\mathcal{O}}$ instead of $E_n^{\mathcal{V}}$.

4. SCALINGS OF \bar{F}_ζ , $F_{k,\text{VR}}$ AND $F_{k,\zeta}$

In the present section, we derive the scalings of \bar{F}_ζ , $F_{k,\text{VR}}$ and $F_{k,\zeta}$ asserted in Theorems 2.1–2.6.

To begin with, we deal with \bar{F}_ζ . As observed in [4, Remark 3.2], the upper bound follows from the results in [10]: according to the scaling of the coverage radius in [9], above γ_n the union of balls covers W_n , which is contractible. In particular, by the nerve theorem [3, Theorem 10.7], the \check{C} -complex has the same Betti numbers as the union of balls, so that any cycle is trivial. The lower bound resorts to a construction inspired by [4, Lemma 5.1].

Scaling of \bar{F}_ζ (proof of Theorem 2.1, scaling).

Henceforth, we write

$$N_{n,\check{C}}^{\max}(\tau) := \Phi_{n,\check{C}}^{\max}(W_1 \times [0, \tau])$$

for the number of features in W_n with life time at least $\bar{F}_\zeta^{-1}(\tau/n)$.

$\bar{F}_\zeta^{-1}(\tau/n)/\gamma_n \leq 1 + \varepsilon$. For $\varepsilon > 0$ write $\ell_n^{\max,+} := (1 + \varepsilon)\gamma_n$ and $\tau_n^- := n\bar{F}_\zeta(\ell_n^{\max,+})$. To prove that $\lim_{n \rightarrow \infty} \tau_n^- = 0$, we have to prove that

$$\mathbb{E}[N_{1,\check{C}}^{\max}(\tau_n^-/n)] = \bar{F}_\zeta(\ell_n^{\max,+}) \in o(1).$$

In order to bound the expected number of holes with death time at least $\ell_n^{\max,+}$ that are centered in W_1 , let $R := \min\{|x| : x \in \mathcal{P}\}$ be the distance of the closest point of \mathcal{P} to the origin. In particular, the death time of any hole centered in W_1 is at most $R + d$, as W_1 is covered completely at the latest at this time. Next, there can be at most one feature with life time exceeding $\ell_n^{\max,+}$ and centered in W_1 . Indeed, if there were at least two such features H_1, H_2 , then at least one of them is born after time $R - d$, as otherwise, H_1 and H_2 would still be the same hole at the corresponding birth time. Combining this with the previous observation on the death time leads to the contradiction that such a hole would have life time at most $2d$. Hence,

$$\mathbb{E}[N_{1,\check{C}}^{\max}(\tau_n^-/n)] = \mathbb{P}(\Phi_{1,\check{C}}^{\max}(W_1 \times [0, \tau_n^-/n]) \neq 0) \leq \mathbb{P}(\mathcal{P} \cap B_{\ell_n^{\max,+} - d}(o) = \emptyset).$$

Invoking the void probability of the Poisson process \mathcal{P} , the right-hand side is at most

$$\exp(-\kappa_d(\ell_n^{\max,+} - d)^d) = \exp(-(1+\varepsilon)^d \log(n)(1 - d/\ell_n^{\max,+})^d) \in o(n^{-1}).$$

$\bar{F}_{\check{c}}^{-1}(\tau/n)/\gamma_n \geq 1 - \varepsilon$. With $\ell_n^{\max,-} := (1 - \varepsilon)\gamma_n$ and $\tau_n^+ := n\bar{F}_{\check{c}}(\ell_n^{\max,-})$, we claim that $\lim_{n \rightarrow \infty} \tau_n^+ = \infty$, i.e.,

$$\lim_{n \rightarrow \infty} n\mathbb{E}[N_{1,\check{c}}^{\max}(\tau_n^+/n)] = \lim_{n \rightarrow \infty} \mathbb{E}[N_{n,\check{c}}^{\max}(\tau_n^+)] = \infty.$$

As indicated at the beginning of this section, we modify a construction of [4, Lemma 5.1]. Loosely speaking, we cover a discretized annulus of radius $\ell_n^{\max,-}$ by Poisson points, see Figure 2.

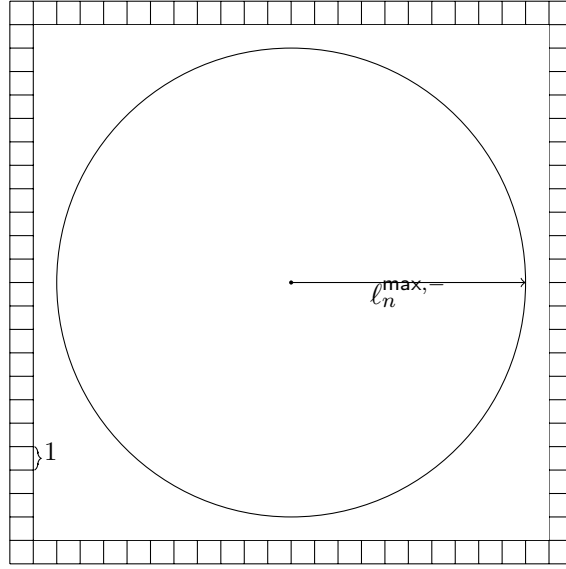


FIGURE 2. Template for maximal life time feature.

More precisely, let

$$E_b := \{\mathcal{P} \cap W(v) \neq \emptyset \text{ for all } v \in \mathbb{Z}^d \text{ with } |v|_\infty = 2\lceil \gamma_n \rceil\}$$

denote the event that each unit block in a d_∞ -annulus of radius $2\lceil \gamma_n \rceil$ contains at least one Poisson point. We also write $m_n = m_{n,\varepsilon}$ for the number of these blocks and note that $m_n \in O((\log n)^{(d-1)/d})$. Moreover, we let

$$E_d := \{\mathcal{P} \cap B_{\ell_n^{\max,-}}(o) = \emptyset\}$$

denote the event that there are no Poisson points in the ball with radius $\ell_n^{\max,-}$. Then, under the event $E_b \cap E_d$ we find a hole with life time at least $\ell_n^{\max,-}$.

Setting $\varepsilon' := 1 - (1 - \varepsilon)^d$ the independence property of the Poisson process gives that

$$\mathbb{P}(E_b \cap E_d) = \mathbb{P}(E_b)\mathbb{P}(E_d) \geq e^{-cm_n} n^{-1+\varepsilon'}$$

for some $c > 0$. Finally, we can take the template described by $E_b \cap E_d$ and shift it to a different location in the window W_n . Since those configurations are of logarithmic extent, we can arrange at least $a_n := n^{1-\varepsilon'/2}$ of them disjointly in W_n . In particular, since $m_n \in O((\log n)^{(d-1)/d})$,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,\check{c}}^{\max}(\tau_n^+)] \geq \lim_{n \rightarrow \infty} a_n e^{-cm_n} n^{-1+\varepsilon'} = \infty.$$

Next, we analyze the scaling of $F_{k,\check{\zeta}}$. Since we now deal with short rather than long life times, the proof structure is converse to that for $\bar{F}_{\check{\zeta}}$. That is, in the lower bound we analyze configurations leading to small life times, whereas the upper bound relies on a specific construction. We also write

$$N_{n,k,\check{\zeta}}^{\min}(\tau) := \Phi_{n,k,\check{\zeta}}^{\min}(W_1 \times [0, \tau])$$

for the number of features in W_n with life time at most $F_{k,\check{\zeta}}^{-1}(\tau/n)$.

Henceforth, let

$$d_{0\dots k} := d(x_0, \dots, x_k) := \inf\{r > 0: \bigcap_{x \in \{x_0, \dots, x_k\}} B_r(x) \neq \emptyset\} \quad (7)$$

denote the *filtration time* of the simplex $\{x_0, \dots, x_k\}$, i.e., the first time it appears in the \check{C} -filtration. In particular, $d_{ij} = |x_i - x_j|/2$.

Moreover, if $x_0, \dots, x_{k-1} \in \mathbb{R}^d$ are affinely independent and $r > d(x_0, \dots, x_{k-1})$, then in every k -dimensional hyperplane containing x_0, \dots, x_{k-1} , there are precisely two k -dimensional balls with radius r containing x_0, \dots, x_{k-1} on their boundary, and we let $D_r(x_0, \dots, x_{k-1})$ denote the union of all such balls. Then, for $r, \ell > 0$, we introduce the crescent

$$D_{r,\ell}(x_0, \dots, x_{k-1}) := D_{r+\ell}(x_0, \dots, x_{k-1}) \setminus D_r(x_0, \dots, x_{k-1}),$$

illustrated in Figure 3 for $k = 2$. **Having** a k -simplex x_0, \dots, x_k with filtration time s means that **there** exist $i_0, \dots, i_m \leq k$ and an m -dimensional ball $B_s(P)$ centered at some $P \in \mathbb{R}^d$ such that $x_0, \dots, x_k \subset B_s(P)$ and $x_{i_0}, \dots, x_{i_m} \in \partial B_s$. In particular, $s \in [r, r + \ell]$ means that $x_{i_m} \in D_{r,\ell}(x_{i_0}, \dots, x_{i_{m-1}})$.

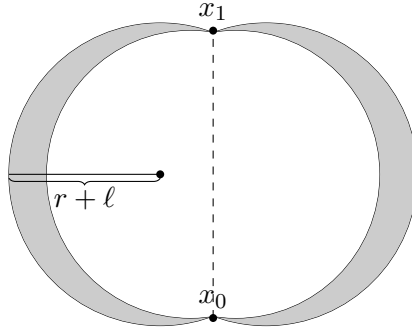


FIGURE 3. The crescent set $D_{r,\ell}(x_0, x_1)$.

Scaling of $F_{k,\check{\zeta}}$ (proof of Theorem 2.5, scaling).

$F_{k,\check{\zeta}}^{-1}(\tau/n) \geq n^{-2-\varepsilon}$ and $\mathbb{P}(\min_{Z_{i,k} \in W_n} L_{i,k} \leq n^{-2-\varepsilon}) \rightarrow 0$. It suffices to prove the first assertion. We write $\ell_n^{\min,-} := n^{-2-\varepsilon}$ and $\tau_n^{\min,-} := nF_{k,\check{\zeta}}(\ell_n^{\min,-})$ and claim that $\lim_{n \rightarrow \infty} \tau_n^{\min,-} = 0$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,k,\check{\zeta}}^{\min}(\tau_n^{\min,-})] = 0.$$

When a \check{C} -feature with life time $\ell > 0$ is born at time $r = d_{0\dots k}$, then there exist Poisson points X_0, \dots, X_k such that $\cup_{i \leq k} B_r(X_i)$ covers the simplex spanned by these points. If a feature dies at time $t_d = d_{0\dots k} + \ell$, then there is $m \leq k$ and Poisson points X'_0, \dots, X'_{m+1} such that the associated simplex is covered at time $d_{0\dots k} + \ell$. After reordering and choosing m to be the smallest such value, there exists $j \leq m$ such that $X'_i = X_i$ for $i \leq j$ and $X'_i \neq X_i$ for $i \geq j + 1$. Then, the previous observations can be succinctly summarized as

$$X'_{m+1} \in D_{d_{0\dots k}, \ell}(X_0, \dots, X_j, X'_{j+1}, \dots, X'_m).$$

Now, we invoke the Slivnyak-Mecke formula from [11, Theorem 4.4]. It shows that the expected number of Poisson points $X_0, \dots, X_j, X'_{j+1}, \dots, X'_{m+1}$ satisfying

$$X'_{m+1} \in D_{d_{0\dots k}, \ell}(X_0, \dots, X_j, X'_{j+1}, \dots, X'_m)$$

is at most

$$\begin{aligned} & \int_{W_{2n} \times B_{s_n}(x_0)^j \times B_{s_n}(x_0)^{m-j}} \int_{D_{d_0 \dots k, \ell}(x_0, \dots, x_m)} 1 dx'_{m+1} d(x_0, \dots, x_j, x'_{j+1}, \dots, x'_m) \\ &= \int_{W_{2n} \times B_{s_n}(x_0)^j \times B_{s_n}(x_0)^{m-j}} |D_{d_0 \dots k, \ell}(x_0, \dots, x_m)| d(x_0, \dots, x_j, x'_{j+1}, \dots, x'_m). \end{aligned}$$

Now, Lemma C.1 in the appendix shows that the volume of the integrand is bounded above by $cs_n^d \sqrt{\ell}$, so that

$$\int_{W_{2n} \times B_{s_n}(x_0)^j \times B_{s_n}(x_0)^{m-j}} |D_{d_0 \dots k, \ell}(x_0, \dots, x_m)| d(x_0, \dots, x_j, x'_{j+1}, \dots, x'_m) \leq c\kappa_d^{2m-j} s_n^{(2m-j+1)d} \int_{W_{2n}} \sqrt{\ell} dx_0,$$

and for $\ell \leq \ell_n^{\min, -}$ the latter expression tends to 0 as $n \rightarrow \infty$.

$F_{k, \check{C}}^{-1}(\tau/n) \leq n^{-2+\varepsilon}$ and $\mathbb{P}(\min_{Z_{i,k} \in W_n} L_{i,k} \geq n^{-2+\varepsilon}) \rightarrow 0$. Next, writing $\ell_n^{\min, +} := n^{-2+\varepsilon}$ and $\tau_n^{\min, +} := nF_{k, \check{C}}(\ell_n^{\min, +})$, we claim that $\lim_{n \rightarrow \infty} \tau_n^{\min, +} = \infty$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,k, \text{VR}}^{\min}(\tau_n^{\min, +})] = \infty.$$

The proof idea is to construct features with a short life time by relying on perturbations from a deterministic template. We first define suitable templates of constant size, which in the final argument will then be copy-pasted throughout the entire window W_n . To lower the technical barriers, we give the proof for $k = d - 1$. The arguments for the lower-dimensional settings are very similar.

First, for affinely independent points $x_0, \dots, x_{d-1} \in \mathbb{R}^d$ let $Q(x_0, \dots, x_{d-1}) \in \mathbb{R}^d$ denote one of the two possible points projecting onto the center O of the $(d-1)$ -dimensional circumball of x_0, \dots, x_{d-1} , and whose distance from that center is given by the radius of that sphere. Figure 4 illustrates this configuration in dimension $d = 2$.

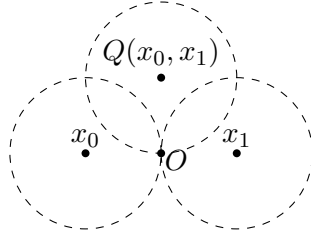


FIGURE 4. Template for \check{C} -feature of minimal life time in dimension $d = 2$.

To fix a distinguished template, we consider the unit vectors $P_i := e_{i+1} \in \mathbb{R}^d$, $0 \leq i \leq d-1$, forming a regular $(d-1)$ -simplex of side length $\sqrt{2}$ with circumradius $\rho_{d-1} = \sqrt{(d-1)/d}$. We claim that at time ρ_{d-1} the faces of the simplex $\{P_0, \dots, P_{d-1}, Q\}$ are all covered. Since ρ_{d-1} is the coverage radius of a regular simplex, it suffices to show by symmetry that $|P_0 - Q| \leq \sqrt{2}$. Indeed, remark that OP_0Q is an isosceles rectangular triangle and write $|P_0 - Q|^2 = 2\rho_{d-1}^2 = 2(d-1)/d$, which gives $|P_0 - Q| = \sqrt{2}\sqrt{(d-1)/d} < \sqrt{2}$.

Next, we consider random perturbations of this template of a small magnitude $\delta \in (0, 1/2)$. More precisely, we define the event

$$E := \{\mathcal{P}(B_\delta(P_i)) = 1 \text{ for all } i \leq d-1\} \cap \{\mathcal{P}(B_{4\delta}(o)) = d+1\}$$

that \mathcal{P} has precisely one point in the δ -neighborhood of each of the points in the template in dimension k and no other points in the ball $B_{4\delta}(o)$. Conditioned on E , we write P'_i for the Poisson point contained in $B_\delta(P_i)$, which is then uniformly distributed at random in $B_\delta(P_i)$. Moreover, under the event E the Poisson process \mathcal{P} contains precisely one further point inside

$B_{4d}(o)$, which we denote by Q' . Hence, writing $d_{0'\dots(d-1)'} = \inf\{r > 0: \bigcap_{x \in \{P'_0, \dots, P'_{d-1}\}} B_r(x) \neq \emptyset\}$, conditioned on E , the probability that this feature has life time at most ℓ is at least

$$\begin{aligned} & \mathbb{P}(Q' \in D_{d_{0'\dots(d-1)'}', \ell}(P'_0, \dots, P'_{d-1}) \cap B_\delta(Q(P'_0, \dots, P'_{d-1}))) \\ &= \frac{1}{\kappa_d((4d)^d - d\delta^d)} \mathbb{E}\left[\left|D_{d_{0'\dots(d-1)'}', \ell}(P'_0, \dots, P'_{d-1}) \cap B_\delta(Q(P'_0, \dots, P'_{d-1}))\right|\right]. \end{aligned}$$

Then, Lemma C.1 in the appendix shows that the volume on the right-hand side is at least $c\sqrt{\ell}$ for a suitable $c > 0$. Now, for a suitable $c' > 0$, we can fit at least $c'n$ such templates into the window W_n . Hence, for $\ell = \ell_n^{\min, +}$,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,k,\zeta}^{\min}(\tau_n^{\min, +})] \geq \lim_{n \rightarrow \infty} cc'n\sqrt{\ell_n^{\min, +}} = \infty,$$

as asserted. To prove that $\mathbb{P}(\min_{Z_{i,k} \in W_n} L_{i,k} \geq n^{-2+\varepsilon}) \rightarrow 0$, we note that by the independence of the Poisson point process, the probability that none of the shifted templates yields a feature with life time at most $\ell_n^{\min, +}$ is at most $(1 - c\sqrt{\ell_n^{\min, +}})^{c'n}$, which tends to 0, since $n\sqrt{\ell_n^{\min, +}} \rightarrow \infty$.

Finally, we move to the scaling of $F_{k, \text{VR}}$. Henceforth,

$$B_{r,\ell}(x) := B_{r+\ell}(x) \setminus B_r(x)$$

denotes the annulus with center $x \in \mathbb{R}^d$, inner radius r and thickness ℓ .

Scaling of $F_{k, \text{VR}}$ (proof of Theorem 2.6, scaling).

$F_{k, \text{VR}}^{-1}(\tau/n) \geq n^{-1-\varepsilon}$. Writing $\ell_n^{\min, -} := n^{-1-\varepsilon}$ and $\tau_n^{\min, -} := nF_{k, \text{VR}}(\ell_n^{\min, -})$, we now assert that $\lim_{n \rightarrow \infty} \tau_n^{\min, -} = 0$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,k, \text{VR}}^{\min}(\tau_n^{\min, -})] = 0.$$

First, we may assume that E_n^{O} occurs. Indeed, by Cauchy-Schwarz,

$$\mathbb{E}[N_{n,k, \text{VR}}^{\min}(\tau_n^{\min, -}) \mathbb{1}\{(E_n^{\text{O}})^c\}] \leq \mathbb{E}[\mathcal{P}(W_{2n})^d \mathbb{1}\{(E_n^{\text{O}})^c\}] \leq (\mathbb{E}[\mathcal{P}(W_{2n})^{2d}]) \mathbb{P}((E_n^{\text{O}})^c)^{1/2}.$$

Since \mathcal{P} is a Poisson point process, the expectation on the right-hand side is of order $O(n^{2d})$, so that it remains to invoke a property which is similar to (3).

For the remaining argument, recall that a simplex belongs to the VR-filtration at time t if the distance between any pair of its vertices is at most $2t$. In particular, when a feature with life time ℓ is born at time t_b , there exist $X_1, X_2 \in \mathcal{P}$ with $d_{12} = t_b$, and when it dies at time $t_d = t_b + \ell$, there are Poisson points X_3, X_4 with $d_{34} = t_b + \ell$. In other words, $X_4 \in B_{2d_{12}, 2\ell}(X_3)$. Moreover, since the event E_n^{O} occurs whp, we may assume that $X_1 \in W_{2n}$ and that $\max_{2 \leq i \leq 4} d_{1i} \leq s_n$.

Here, $\{X_1, \dots, X_4\}$ consists of at least three different points. First, if they are all pairwise distinct, then the Slivnyak-Mecke formula bounds the expected number of undershoots as

$$\begin{aligned} & \int_{W_{2n}} \int_{B_{s_n}(x_1)^2} |B_{2d_{12}, 2\ell_n^{\min, -}}(x_3)| d(x_2, x_3) dx_1 = \kappa_d^2 s_n^d \int_{W_{2n}} \int_{B_{s_n}(x_1)} ((2d_{12} + 2\ell_n^{\min, -})^d - (2d_{12})^d) dx_2 dx_1 \\ & \leq 4^d d \kappa_d^3 s_n^{3d} \int_{W_{2n}} \ell_n^{\min, -} dx_1 \\ & = 8^d d \kappa_d^3 s_n^{3d} n^{-\varepsilon}, \end{aligned}$$

which tends to 0 as $n \rightarrow \infty$. Second, if say $X_2 = X_3$, then we proceed similarly to obtain the bound

$$\begin{aligned} & \int_{W_{2n}} \int_{B_{s_n}(x_1)} |B_{2d_{12}, 2\ell_n^{\min, -}}(x_2)| dx_2 dx_1 = \kappa_d \int_{W_{2n}} \int_{B_{s_n}(x_1)} ((2d_{12} + 2\ell_n^{\min, -})^d - (2d_{12})^d) dx_2 dx_1 \\ & \leq 4^d d \kappa_d^2 s_n^{2d} \int_{W_{2n}} \ell_n^{\min, -} dx_1 \\ & = 8^d d \kappa_d^2 s_n^{2d} n^{-\varepsilon}, \end{aligned}$$

which again tends to 0 as $n \rightarrow \infty$.

$F_{k,\mathbf{VR}}^{-1}(\tau/n) \leq n^{-1+\varepsilon}$. Next, writing $\ell_n^{\min,+} := n^{-1+\varepsilon}$ and $\tau_n^{\min,+} := nF_{k,\mathbf{VR}}(\ell_n^{\min,+})$, we claim that $\lim_{n \rightarrow \infty} \tau_n^{\min,+} = \infty$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{E}[N_{n,k,\mathbf{VR}}^{\min}(\tau_n^{\min,+})] = \infty.$$

To that end, we construct a feature with a short life time for each feature dimension $1 \leq k \leq d-1$. The features depend on a deterministic template illustrated in Figure 5.

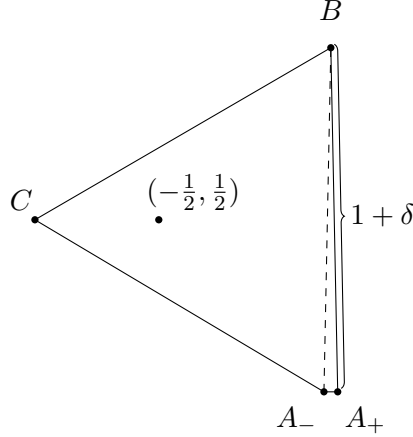


FIGURE 5. Template for VR-feature of minimal life time.

We fix a small value $\delta = \delta(k, d) > 0$. Then, we set

$$A_{\pm} := \pm \sqrt{(1+\delta)^2 - 1} e_1, \quad B := e_2,$$

so that $|A_- - B| = |A_+ - B| = 1 + \delta$. We also fix a point $C := (\eta, 1/2)$, where $\eta < 0$ is chosen, so that $|C - A_+| > 1 + \delta > |C - B|$.

To convey the idea, we sketch how this template gives rise to the desired feature for $k = 1$ before moving to higher k . If we consider the VR-complex on $\{A_-, A_+, B, C\}$ at level $1 + \delta$, then $A_- A_+ B C$ is a loop. Moreover, removing the edge $A_- B$ and the attached higher simplices, the complex does not contain any triangles as by construction $|C - A_+| > 1 + \delta$. In particular, the loop has positive life time. However, after adding the edge $A_- B$, the loop becomes the boundary of the triangles $A_- A_+ B$ and $B C A_-$.

To generalize the construction to higher k , we introduce additional points $\{P_{1,\pm}, \dots, P_{k-1,\pm}\}$ to the complex as follows. First, set

$$P_{i,\pm} := \frac{1}{2}(e_2 - e_1) \pm (0.5 + 4\delta)e_{i+2}.$$

In particular, the $P_{i,\pm}$ are at distance at most $1 + \delta$ from A_-, A_+, B and C . Now, we generalize the above consideration for the VR-complex at level $|A_+ - B|$ but with the edge $\{A_-, B\}$ removed. In particular, the complex contains all k -simplices of the form

$$\{P_{1,\varepsilon_1}, P_{2,\varepsilon_2}, \dots, P_{k-1,\varepsilon_{k-1}}, P, P'\},$$

where $\varepsilon_i \in \{-, +\}$ and $\{P, P'\}$ is one of $\{A_-, A_+\}$, $\{A_+, B\}$, $\{B, C\}$ or $\{C, A_-\}$. Then, these simplices form a cycle. Indeed, when removing the vertex P_{i,ε_i} , then the resulting face is also contained in the k -simplex with P_{i,ε_i} replaced by $P_{i,-\varepsilon_i}$. On the other hand, if for instance $\{P, P'\} = \{A_-, A_+\}$ and if we remove the point A_- , then we find the corresponding face also in the simplex with $\{P, P'\} = \{A_+, B\}$. However, this complex does not contain any $(k+1)$ -simplices. Indeed, $|P_{i,-} - P_{i,+}| > 1 + \delta$ are at distance larger than $1 + \delta$, and for any triple from

$\{A_-, A_+, B, C\}$ at least one edge is also not in the complex. This situation changes drastically if A_-B belongs to the complex. Then, the boundaries of the $(k+1)$ -simplices

$$\{P_{1,\varepsilon_1}, P_{2,\varepsilon_2}, \dots, P_{k-1,\varepsilon_{k-1}}, A_-, A_+, B\} \quad \text{and} \quad \{P_{1,\varepsilon_1}, P_{2,\varepsilon_2}, \dots, P_{k-1,\varepsilon_{k-1}}, B, C, A_-\}$$

yield the cycle constructed before.

It remains to connect this template to random features induced by \mathcal{P} . Similarly as in the Čech case, set $\delta_1 = \delta/8$ and consider the event

$$E' := \{\mathcal{P}(B_{\delta_1}(P)) = 1 \text{ for all } P \in \{A_-, A_+, B, C\} \cup \{P_{i,\pm}\}_{i \leq k-1}\} \cap \{\mathcal{P}(B_3(o)) = 2k+2\}$$

that \mathcal{P} has precisely one point in the δ_1 -neighborhood of each of the template points and no other points in the ball $B_3(o)$. Conditioned on E' , we write A'_-, A'_+, B', C' and $(P_{i,\pm})'$ for the Poisson points lying in the corresponding neighborhoods of the template points. In particular, under E' , the primed points are distributed uniformly at random in the corresponding δ_1 -neighborhoods. Moreover, the construction of the template is sufficiently robust with respect to δ_1 -perturbations to show that the adjacency findings for the VR-complex continue to hold upon replacing the template points by the perturbed points.

Conditioned on E' , the probability that this feature has a short life time is at least

$$\begin{aligned} & \mathbb{P}(0 < |A'_- - B'| - |A'_+ - B'| < \ell_n^{\min,+} | E') \\ &= (\kappa_d \delta_1^d)^{-3} \left| \left\{ (A'_\pm, B') : \max_{P \in \{A_\pm, B\}} |P' - P| \leq \delta_1 \text{ and } 0 < |A'_- - B'| - |A'_+ - B'| < \ell_n^{\min,+} \right\} \right|. \end{aligned}$$

Now, we set $u := A_+ - B'$ and $v := A'_- - B'$, and for $\alpha \in [0, 1]$ and $\eta \in \mathbb{R}^d$

$$A_{+,\alpha,\eta} := B' + (|v| - \alpha \ell_n^{\min,+})(u + \eta) / |u + \eta|,$$

so that $|A_{+,\alpha,\eta} - B'| = |v| - \alpha \ell_n^{\min,+}$. In particular, $|A'_- - B'| - |A_{+,\alpha,\eta} - B'| = \alpha \ell_n^{\min,+} < \ell_n^{\min,+}$.

Hence, writing $\delta_2 := \delta_1/8$, it suffices to show that

$$|A_{+,\alpha,\eta} - A_+| \leq \delta_1$$

for every $A'_- \in B_{\delta_2}(A_-)$, $B' \in B_{\delta_2}(o)$, $\alpha \in [0, 1]$ and $\eta \in B_{\delta_2}(o)$, because then every such point $A_{+,\alpha,\eta}$ is an admissible choice for A'_+ . To that end, we leverage the bound

$$\begin{aligned} |A_{+,\alpha,\eta} - A_+| &\leq \left| -u + |v| \frac{u}{|u|} + |v| \left(\frac{u + \eta}{|u + \eta|} - \frac{u}{|u|} \right) \right| + \ell_n^{\min,+} \\ &\leq ||u| - |v|| + |v| \left| \frac{u + \eta}{|u + \eta|} - \frac{u}{|u|} \right| + \ell_n^{\min,+}. \end{aligned}$$

The first expression is at most $3\delta_2$ and the second one at most $4\delta_2$, so that taken together, the right-hand side is indeed at most δ_1 .

Now, note that

$$|\{A_{+,\alpha,\eta} : \alpha \in [0, 1], \eta \in B_{\delta_2}(o)\}| \geq \ell_n^{\min,+} \kappa_{d-1} \delta_2^{d-1},$$

so that the probability is much larger than n^{-1} , and consequently that the expectation tends to infinity.

5. PROOF OF PROPOSITION 3.1

In this section, we prove that there are no multiple exceedances in $W_{s_{n,+}}$.

Proof of Proposition 3.1 (i). By the scaling derived in Section 4, it suffices to prove that

$$\mathbb{P}(N_{s_{n,+}}^{\max}(s_{n,+}, \bar{F}_{\mathcal{C}}(\ell_n^{\max,-})) \geq 2) \in O(n^{-1}).$$

By the scaling in Section 4, both features H_i and H_j die at time at most $\ell_n^{\max,+}$, so that their birth times are at most $2\varepsilon\gamma_n$. In particular, for the holes H_i and H_j to be contained in different connected components at the birth time of the later hole, say H_i , we have $d(Z_i, \partial H_i) \geq (1-\varepsilon)\gamma_n$ and $d(Z_j, \partial H_j) \geq (1-3\varepsilon)\gamma_n$. Since $d(Z_i, Z_j) \geq d(Z_i, \partial H_i) + d(Z_j, \partial H_j)$ at the birth time of H_i ,

this implies that $|v - v'| \geq (2 - 5\varepsilon)\gamma_n \geq \gamma_n$, for ε small enough, where $W(v) := v + W_1$, $W(v') := v' + W_1$ are the lattice boxes containing Z_i and Z_j , respectively. Therefore,

$$|B_{\ell_n^{\max,-}}(v) \cup B_{\ell_n^{\max,-}}(v')| \geq \kappa_d^{-1} \log(n) |B_{1-\varepsilon}(o) \cup B_{1-\varepsilon}(e_1)|.$$

Then,

$$\begin{aligned} & \mathbb{P}(N_{s_{n,+}}^{\max}(s_{n,+}, \bar{F}_{\check{C}}(\ell_n^{\max,-})) \geq 2) \\ & \leq \sum_{\substack{v, v' \in W_{s_{n,+}} \cap \mathbb{Z}^d \\ |v - v'| \geq \gamma_n}} \mathbb{P}(L_i \wedge L_j \geq \ell_n^{\max,-} \text{ for some } Z_i \in W(v), Z_j \in W(v')) \\ & \leq \sum_{\substack{v, v' \in W_{s_{n,+}} \cap \mathbb{Z}^d \\ |v - v'| \geq \gamma_n}} \mathbb{P}\left((B_{\ell_n^{\max,-}}(v) \cup B_{\ell_n^{\max,-}}(v')) \cap \mathcal{P} = \emptyset\right) \\ & \leq |W_{s_{n,+}}^2| \exp\left(-\kappa_d^{-1} \log(n) |B_{1-\varepsilon}(o) \cup B_{1-\varepsilon}(e_1)|\right). \end{aligned}$$

The last term is in $O(n^{-1})$ provided that $\varepsilon > 0$ is sufficiently small. \square

Next, we move to the minimum in the \check{C} -filtration.

Proof of Proposition 3.1 (ii); \check{C} . Again, it suffices to show that

$$\mathbb{P}(N_{s_{n,+}^+, 1}^{\min, \check{C}}(\ell_n^{\min,+}) \geq 2) \in O(n^{-1}), \quad (8)$$

where we now set $\ell_n^{\min,+} := n^{-2+\varepsilon}$. We need to understand the geometric implications of finding features $Z_i, Z_j \in W_{s_{n,+}}$ with life times shorter than $\ell_n^{\min,+}$. We recall from Section 4 that one such undershoot gives Poisson points $\{X_0, \dots, X_4\}$ such that the triangle $X_2X_3X_4$ is covered by the union of balls centered at X_i , $2 \leq i \leq 4$. First, we observe that if at least one of X_0, X_1 is not contained in $\{X_2, X_3, X_4\}$, then (8) holds even without taking the second feature Z_j into account. Indeed, if for instance $X_1 = X_2$, but $X_0 \notin \{X_2, X_3, X_4\}$, then the Slivnyak-Mecke formula yields the bound

$$\int_{W_{s_{n,+}}^3} |B_{2d_{234}-2\ell_n^{\min,+}, 2\ell_n^{\min,+}}(x_2)| d(x_2, x_3, x_4) \in O(s_n^{4d} \ell_n^{\min,+}).$$

We may hence assume to have points $\{X_0, X_1, X_2\}$ with $d_{01} < d_{012} \leq d_{01} + \ell_n^{\min,+}$ for the first feature and similarly points $\{X'_0, X'_1, X'_2\} \neq \{X_0, X_1, X_2\}$ with $d_{0'1'} < d_{0'1'2'} \leq d_{0'1'} + \ell_n^{\min,+}$. Again, several configurations are possible, where the most challenging ones correspond to the cases where $\{X_0, X_1, X_2\}$ and $\{X'_0, X'_1, X'_2\}$ differ in one variable.

First, assume that $X'_0 = X_0$, $X'_1 = X_1$, but $X'_2 \neq X_2$. Then, invoking Lemma C.1 in Appendix C gives the volume bound

$$\begin{aligned} \left| \{(x_0, x_1, x_2, x'_2) \in W_{s_{n,+}}^4 : d_{012}, d_{01'2'} \in (d_{01}, d_{01} + \ell_n^{\min,+})\} \right| & \leq \int_{W_{s_{n,+}}^2} |D_{d_{01}, \ell_n^{\min,+}}(x_0, x_1)|^2 d(x_0, x_1) \\ & \leq c_{\text{cresc}}^2 s_{n,+}^4 \ell_n^{\min,+}. \end{aligned}$$

Noting that the last line is of order $O(n^{-1})$ concludes the proof of the first case. Second, assuming that $X'_1 = X_1$, $X'_2 = X_2$, but $X'_0 \neq X_0$, Lemmas C.1 and C.3 in Appendix C give that

$$\begin{aligned} & \left| \{(x_0, x_1, x_2, x'_0) \in W_{s_{n,+}}^4 : d_{01} < d_{012} \leq d_{01} + \ell_n^{\min,+} \text{ and } d_{0'1} < d_{0'12} \leq d_{0'1} + \ell_n^{\min,+}\} \right| \\ & \leq \int_{W_{s_{n,+}}^2} \int_{D_{d_{01}, \ell_n^{\min,+}}(x_0, x_1)} |\{x'_0 \in W_{s_{n,+}} : d_{0'1} < d_{0'12} \leq d_{0'1} + \ell_n^{\min,+}\}| dx_2 d(x_0, x_1) \\ & \leq cn^{-1/4} \int_{W_{s_{n,+}}^2} |D_{d_{01}, \ell_n^{\min,+}}(x_0, x_1)| d(x_0, x_1) \\ & \leq cc' s_{n,+}^3 n^{-1/4} \sqrt{\ell_n^{\min,+}}, \end{aligned}$$

which is again in $O(n^{-1})$. Finally consider the case where $X'_1 = X_2$, $X'_2 = X_1$, but $X'_0 \neq X_0$. Then, similarly, Lemmas C.1 and C.3 in Appendix C give that

$$\begin{aligned} & \left| \{(x_0, x_1, x_2, x'_0) \in W_{s_{n,+}}^4 : d_{01} < d_{012} \leq d_{01} + \ell_n^{\min,+} \text{ and } d_{0'2} < d_{0'21} \leq d_{0'2} + \ell_n^{\min,+}\} \right| \\ & \leq \int_{W_{s_{n,+}}^2} \int_{D_{d_{01}, \ell_n^{\min,+}}(x_0, x_1)} |\{x'_0 \in W_{s_{n,+}} : d_{0'2} < d_{0'21} \leq d_{0'2} + \ell_n^{\min,+}\}| dx_2 d(x_0, x_1), \end{aligned}$$

so that we now conclude as in the previous case. \square

Finally, we deal with the minimum in the VR-filtration.

Proof of Proposition 3.1 (ii); VR. Again, it suffices to prove that

$$\mathbb{P}(N_{s_n^+, k}^{\min, \text{VR}}(\ell_n^{\min,+}) \geq 2) \in O(n^{-1}), \quad (9)$$

where we now set $\ell_n^{\min,+} = n^{-1+\varepsilon}$. Now, suppose two features centered at $Z_i, Z_j \in W_{s_{n,+}}$ live shorter than $\ell_n^{\min,+}$. By the VR-filtration, this means that for the feature centered in Z_i , there exist Poisson points $X_0, \dots, X_3 \in W_{s_{n,+}}$ such that

$$d_{01} < d_{23} < d_{01} + \ell_n^{\min,-}.$$

Similarly also the feature centered in Z_j gives rise to Poisson points $X'_0, \dots, X'_3 \in W_{s_{n,+}}$ such that

$$d_{0'1'} < d_{2'3'} < d_{0'1'} + \ell_n^{\min,-}.$$

Again, not all points need to be distinct, but both $\{X_0, \dots, X_3\}$ and $\{X'_0, \dots, X'_3\}$ consist of at least 3 elements. Moreover, $\{X_0, X_1\} \neq \{X'_0, X'_1\}$ and $\{X_2, X_3\} \neq \{X'_2, X'_3\}$ since the features are distinct.

We now distinguish two cases. The first one being that $\{X_0, \dots, X_3\} \neq \{X'_0, \dots, X'_3\}$. Say, for instance $X'_3 \notin \{X_0, \dots, X_3\}$. Then, we may apply the Slivnyak-Mecke formula to see that the expected number of configurations is at most

$$\begin{aligned} & \int_{W_{s_{n,+}}} \mathbb{E} \left[\#\{X_0, \dots, X_3, X'_0, \dots, X'_2 \in \mathcal{P} \cap W_{s_{n,+}} : d_{01} < d_{23} < d_{01} + \ell_n^{\min,-} \text{ and } x_3 \in B_{2d_{01}, 2\ell_n^{\min,-}}(X_2)\} \right] dx'_3 \\ & \leq 2^{d+1} ds_{n,+}^{4d} \ell_n^{\min,-} \mathbb{E} \left[\#\{X_0, \dots, X_3 \in \mathcal{P} \cap W_{s_{n,+}} : d_{01} < d_{23} < d_{01} + \ell_n^{\min,-}\} \right] \\ & \leq 2c(2^d ds_{n,+}^{4d})^2 (\ell_n^{\min,-})^2 \in O(n^{-1}), \end{aligned}$$

for a suitable $c > 0$. It remains to deal with the case $\{X'_0, \dots, X'_3\} = \{X_0, \dots, X_3\}$, which necessarily means that they consist of 4 elements. For instance, it may occur that $X'_0 = X_0$, $X'_1 = X_3$, $X'_2 = X_2$ and $X'_3 = X_1$. Then, $d_{01} < d_{23} < d_{01} + \ell_n^{\min,+}$ and $d_{03} < d_{12} < d_{03} + \ell_n^{\min,+}$. Applying the Slivnyak-Mecke formula, the expected number of such configurations in $W_{s_{n,+}}$ is at most

$$\left| \{(x_0, \dots, x_3) \in W_{s_{n,+}}^4 : |d_{23} - d_{01}| \vee |d_{03} - d_{12}| \leq \ell_n^{\min,+}\} \right|.$$

Lemma C.2 in Appendix C shows that the latter expression is in $O(n^{-1})$, as asserted. \square

REFERENCES

- [1] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statist. Sci.*, 5(4):403–434, 1990.
- [2] C. A. N. Biscio, N. Chenavier, C. Hirsch, and A. M. Svane. Testing goodness of fit for point processes via topological data analysis. *Electron. J. Stat.*, 14(1):1024–1074, 2020.
- [3] A. Björner. Topological methods. In *Handbook of Combinatorics*, pages 1819–1872. Elsevier, Amsterdam, 1995.
- [4] O. Bobrowski, M. Kahle, and P. Skraba. Maximally persistent cycles in random geometric complexes. *Ann. Appl. Probab.*, 27(4):2032–2060, 2017.
- [5] J.-D. Boissonnat, F. Chazal, and M. Yvinec. *Geometric and Topological Inference*. Cambridge University Press, Cambridge, 2018.
- [6] P. Calka and N. Chenavier. Extreme values for characteristic radii of a Poisson-Voronoi tessellation. *Extremes*, 17(3):359–385, 2014.
- [7] N. Chenavier. A general study of extremes of stationary tessellations with examples. *Stochastic Process. Appl.*, 124(9):2917–2953, 2014.
- [8] M. Gidea and Y. A. Katz. Topological data analysis of financial time series: Landscapes of crashes. *Physica A: Statistical Mechanics and its Applications*, 491:820–834, 2018.
- [9] P. Hall. On the coverage of k -dimensional space by k -dimensional spheres. *Ann. Probab.*, 13(3):991–1002, 1985.
- [10] M. Kahle. Random geometric complexes. *Discrete Comput. Geom.*, 45(3):553–573, 2011.
- [11] G. Last and M. D. Penrose. *Lectures on the Poisson Process*. Cambridge University Press, Cambridge, 2016.
- [12] M. R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes*. Springer-Verlag, New York, 1983.
- [13] R. Meester and R. Roy. *Continuum Percolation*. Cambridge University Press, Cambridge, 1996.
- [14] F. Morgan. *Geometric Measure Theory: A Beginner’s Guide*. Elsevier, Amsterdam, 2016.
- [15] M. D. Penrose. The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.*, 7(2):340–361, 1997.
- [16] P. Pranav, H. Edelsbrunner, R. van de Weygaert, G. Vegter, M. Kerber, B. J. T. Jones, and M. Wintraecken. The topology of the cosmic web in terms of persistent Betti numbers. *Monthly Notices of the Royal Astronomical Society*, 465(4):4281–4310, 2016.
- [17] J. W. Rocks, A. J. Liu, and E. Katifori. Revealing structure-function relationships in functional flow networks via persistent homology. *arXiv preprint arXiv:1901.00822*, 2019.
- [18] M. Saadatfar, H. Takeuchi, V. Robins, N. Francois, and Y. Hiraoka. Pore configuration landscape of granular crystallization. *Nature Communications*, 8(1):1–11, 2017.
- [19] M. Schulte and C. Thäle. The scaling limit of Poisson-driven order statistics with applications in geometric probability. *Stochastic Process. Appl.*, 122(12):4096–4120, 2012.

APPENDIX A. INVERTIBILITY OF $F_{k,VR}$ AND $F_{k,\check{C}}$

In this section, we show that $F_{k,VR}, F_{k,\check{C}} : (0, \infty) \rightarrow (0, 1)$ are invertible for every $1 \leq k \leq d-1$.

Lemma A.1 (Invertibility of $F_{k,VR}$ and $F_{k,\check{C}}$). *Let $1 \leq k \leq d-1$. Then, $F_{k,VR}, F_{k,\check{C}} : (0, \infty) \rightarrow (0, 1)$ are invertible.*

Proof. We prove invertibility by showing that $F_{k,\check{C}/VR}$ is a strictly decreasing function, which is continuous and satisfies $\lim_{\ell \rightarrow 0} F_{k,\check{C}/VR}(\ell) = 0$ and $\lim_{\ell \rightarrow \infty} F_{k,\check{C}/VR}(\ell) = 1$.

$\lim_{\ell \rightarrow 0} F_{k,\check{C}/VR}(\ell) = 0$. First, by right-continuity $\lim_{\ell \rightarrow 0} F_{k,\check{C}/VR}(\ell) = F_{k,\check{C}/VR}(0)$, where $F_{k,\check{C}/VR}(0)$ describes the expected number of features with life time 0 centered in W_1 . However, by definition, the life time of a feature is always strictly positive, so that $F_{k,\check{C}/VR}(0) = 0$.

$\lim_{\ell \rightarrow \infty} F_{k,\check{C}/VR}(\ell) = 1$. Now, $\lim_{\ell \rightarrow \infty} 1 - F_{k,\check{C}/VR}(\ell)$ describes the expected number of features in W_1 with an infinite life time. However, this number vanishes. Indeed, in the \check{C} -filtration the feature must have died by the time that a representing cycle is contained in a ball, since the latter is contractible. If we take a \check{C} -boundary chain as a witness for the death of the feature in the \check{C} -filtration, then this chain is also a witness for the death in the VR-filtration. Thus, with probability 1, all life times are finite.

Continuity. Let $\ell > 0$ be arbitrary. To show continuity, we establish that there are no features with life time exactly ℓ . We start by proving the claim for VR-filtration. In that case, writing

$d_{ij} = |X_i - X_j|/2$, there would exist points $X_0, X_1, X_2, X_3 \in \mathcal{P}$ such that $d_{23} = d_{01} + \ell$ with $X_3 \notin \{X_0, X_1, X_2\}$. Then,

$$\begin{aligned} \mathbb{E}\left[\#\{X_3 : X_3 \in \partial B_{2d_{01}+2\ell}(X_2) \text{ for some } X_0, X_1, X_2 \in \mathcal{P}\}\right] &= \int_{\mathbb{R}^d} \mathbb{P}\left(x \in \partial B_{2d_{01}+2\ell}(X_2) \text{ for some } X_0, X_1, X_2 \in \mathcal{P}\right) dx \\ &= \mathbb{E}\left[\left|\bigcup_{X_0, X_1, X_2 \in \mathcal{P}} \partial B_{2d_{01}+2\ell}(X_2)\right|\right] \\ &= 0. \end{aligned}$$

For the \check{C} -filtration, the argumentation is similar. Indeed, we recall from Section 4 that a \check{C} -feature gives Poisson points $X_0, \dots, X_k \in \mathcal{P}$ and $X'_0, \dots, X'_{k+1} \in \mathcal{P}$ with $X'_{k+1} \notin \{X'_0, \dots, X'_k\}$ such that relying on the filtration times from (7), we have $d_{0'\dots(k+1)'} = d_{0\dots k} + \ell$. Hence, similarly as in the VR-filtration,

$$\begin{aligned} &\mathbb{E}\left[\#\{X'_{k+1} : d_{0'\dots(k+1)'} = d_{0\dots k} + \ell \text{ for some } X_0, \dots, X_k, X'_0, \dots, X'_k \in \mathcal{P}\}\right] \\ &= \mathbb{E}\left[\left|\bigcup_{\substack{X_0, \dots, X_k \in \mathcal{P} \\ X'_0, \dots, X'_k \in \mathcal{P}}} \partial D_{d_{0\dots k}}(X'_0, \dots, X'_k)\right|\right] \\ &= 0. \end{aligned}$$

Strict monotonicity. Strict monotonicity means that for $\ell < \ell'$ there is a positive probability to observe a feature with life time in (ℓ, ℓ') . Note that scaling all points by a factor also scales the life time by that factor. Hence, it suffices to show that for some fixed b and every $\varepsilon > 0$, there is a positive probability to have a feature with life time in $(b - \varepsilon, b + \varepsilon)$.

We start with the VR-filtration. Consider a feature described by a cross-polytope $\{\pm e_i\}_{1 \leq i \leq k+1}$. Then, this feature has birth time $1/\sqrt{2}$ and death time 1. Therefore its life time is $b := 1 - 1/\sqrt{2}$. If we allow the vertices of the feature to be perturbed by at most $\varepsilon/2$, the life time is in the interval $(b - \varepsilon, b + \varepsilon)$.

Finally, we deal with the \check{C} -filtration. Consider a feature described by a regular simplex $\{e_i\}_{1 \leq i \leq k+1}$. Then, this feature has birth time $\rho := \sqrt{(k-1)/k}$ and death time $\rho' := \sqrt{k/(k+1)}$. Therefore, its life time is $b := \rho' - \rho$, and we conclude as in the VR-filtration. \square

APPENDIX B. PERCOLATION PROPERTIES

In this section, we show that occupied or vacant components in continuum percolation whose birth time is bounded away from the critical threshold are of poly-logarithmic size with high probability. More precisely, recalling that we center components at their center of gravity, we fix $\varepsilon_c > 0$ and put

$$E_n^{\text{O},*} := \{\text{for every radius } r \notin [r_c^{\text{O}} - \varepsilon_c, r_c^{\text{O}} + \varepsilon_c] \text{ all occupied bounded connected components centered in } W_n \text{ have diameter at most } (\log n)^2\}.$$

and

$$E_n^{\text{V},*} := \{\text{for every radius } r \notin [r_c^{\text{V}} - \varepsilon_c, r_c^{\text{V}} + \varepsilon_c] \text{ all vacant bounded connected components centered in } W_n \text{ have diameter at most } (\log n)^2\},$$

where $r_c^{\text{O}}, r_c^{\text{V}}$ are the critical radii for occupied and vacant continuum percolation, respectively.

We show that the events $E_n^{\text{O},*}$ and $E_n^{\text{V},*}$ occur with high probability. For a *fixed* sub- or supercritical radius r , this follows from classical continuum percolation theory [13, Chapter 4]. However, since in the definitions of $E_n^{\text{O},*}$ and $E_n^{\text{V},*}$ the radius may vary, we need a small discretization argument.

Lemma B.1 ($E_n^{\text{O},*}$ and $E_n^{\text{V},*}$ occur with high probability). *For all sufficiently large $n \geq 1$ it holds that*

$$\mathbb{P}(E_n^{\text{O},*}) \geq 1 - n^{-4d} \quad \text{and} \quad \mathbb{P}(E_n^{\text{V},*}) \geq 1 - n^{-4d}.$$

Proof. We explain how to proceed in the occupied setting, noting that the arguments in the vacant case are almost identical. We say that an event occurs whp if its complement occurs with probability at most n^{-5d} for large n . First, below r_c^{O} all connected components are almost surely bounded, and the size of these components increases with the radius. Hence, we conclude from [13, Lemma 3.3] that the corresponding components are of logarithmic size whp. Second, above r_c^{V} all vacant components are bounded and their size decreases with increasing radius. Since every bounded occupied component lies within a vacant component, we conclude from [13, Lemma 4.1] that again the components are of logarithmic size whp.

Hence, it remains to consider $r \in (r_c^{\text{O}} + \varepsilon_c, r_c^{\text{V}})$, noting that this interval is non-empty only if $d > 2$. Here, we discretize the possible radii that can occur. More precisely, we subdivide the interval $[r_c^{\text{O}} + \varepsilon_c, r_c^{\text{V}}]$ into $K_n \in O(n^{8d})$ parts $\{r_i\}_{i \leq K_n}$ such that $r_{i+1} - r_i \leq n^{-8d}$. Then, by [?, Theorem 10.20], whp for each r_i , all bounded connected components at level r_i centered in W_n are of diameter at most $(\log n)^2$. Hence, it suffices to show that for any $r \in (r_c^{\text{O}} + \varepsilon_c, r_c^{\text{V}})$ the components at level r correspond to the components at level r_i for some $i \leq K_n$. In other words, we claim that whp, for each $i \leq K_n$ there exists at most one pair $\{X_j, X_k\} \subset W_n$ with $|X_j - X_k|/2 \in (r_i, r_{i+1})$.

To that end, note that for fixed i , the Slivnyak-Mecke formula shows that the probability that there exist distinct pairs $\{X_j, X_k\} \subset W_n$ and $\{X'_j, X'_k\} \subset W_n$ with $|X_j - X_k|/2 \in (r_i, r_{i+1})$ and $|X'_j - X'_k|/2 \in (r_i, r_{i+1})$ is of order $O(n^{2(1-8d)})$. Hence, by the union bound, the probability that some $i \leq K_n$ has this property is at most $O(n^{2-8d})$, as asserted. \square

APPENDIX C. VOLUME COMPUTATIONS

In this section, we compute volume bounds for the specific configurations occurring in the proofs in Sections 4 and 5. All computations rely only on findings from elementary geometry, but are still bit tedious when written out in detail.

First, we bound the volumes of crescents. As in the proof of Theorem 2.5, for any affinely independent points $x_0, \dots, x_{d-1} \in \mathbb{R}^d$ let $Q(x_0, \dots, x_{d-1}) \in \mathbb{R}^d$ denote one of the two possible points projecting onto the center of the $(d-1)$ -dimensional circumsphere of x_0, \dots, x_{d-1} , and whose distance from that center is given by the radius of that sphere. As in Sections 4 and 5, we write $d_{i_0 \dots i_k}$ for the filtration time when the simplex $\{x_{i_0}, \dots, x_{i_k}\}$ appears in the \check{C} -filtration.

Lemma C.1 (Volume of crescents). *There exists $c_{\text{cresc}} = c_{\text{cresc}}(d) > 0$ with the following properties. Let $r > 0$, $\ell < 1$, $x_0, \dots, x_{d-1} \in \mathbb{R}^d$ be affinely independent. Then,*

$$|D_{r,\ell}(x_0, \dots, x_{d-1})| \leq c_{\text{cresc}}(r + \ell)^{d-1} \sqrt{\ell}.$$

Moreover, if $d_{0 \dots (d-1)} \geq 1/2$ and $\delta \in (\ell, 1)$, then

$$|D_{d_{0 \dots (d-1)}, \ell}(x_0, \dots, x_{d-1}) \cap B_\delta(Q(x_0, \dots, x_{d-1}))| \geq c_{\text{cresc}}^{-1} \delta^{d-1} \sqrt{\ell}.$$

Proof. To ease notation, set $a := d_{0 \dots (d-1)}$. By rotating and shifting we may assume that x_0, \dots, x_{d-1} are contained in $\{0\} \times \mathbb{R}^{d-1}$ and that their circumcenter is the origin. Furthermore, the set $D_{r,\ell}(x_0, \dots, x_{d-1})$ is rotationally symmetric around the axis $\mathbb{R}e_1$.

Upper bound. Defining $r_+(b), r_-(b) > 0$ for any $b \in \mathbb{R}$ through

$$D_{r,\ell}(-ae_2, ae_2) \cap (\{b\} \times \mathbb{R}) = \{b\} \times ([r_-(b), r_+(b)] \cup [-r_+(b), -r_-(b)]),$$

we obtain by Fubini that

$$|D_{r,\ell}(x_0, \dots, x_{d-1})| = \int_0^{2r+2\ell} |B_{r_+(b)}(o) \setminus B_{r_-(b)}(o)| db = \kappa_{d-1} \int_0^{2r+2\ell} (r_+(b)^{d-1} - r_-(b)^{d-1}) db.$$

Now, $r_+(b)^{d-1} - r_-(b)^{d-1}$ is at most $d(r+\ell)^{d-2}(r_+(b) - r_-(b))$, so that

$$|D_{r,\ell}(x_0, \dots, x_{d-1})| \leq d\kappa_{d-1}(r+\ell)^{d-2} \int_0^{2r+2\ell} (r_+(b) - r_-(b)) db = \frac{d}{2}\kappa_{d-1}(r+\ell)^{d-2} |D_{r,\ell}(-ae_2, ae_2)|.$$

Hence, we have now reduced the upper bound to the special setting where $d = 2$. Here, an elementary geometric argument that is elucidated in [2, Lemma 9.8] concludes the proof.

Lower bound. For a point $x \in \mathbb{R}^{d-1}$ with $|x| \leq \delta$, we let

$$I(x) := \{b \geq 0 : (b, x) \in D_{a,\ell}(x_0, \dots, x_{d-1})\}$$

denote the interval consisting of all points inside $D_{a,\ell}$ projecting onto x . Then, it suffices to show that there exists a constant $c > 0$ such that for each such x we have $|I(x)| \geq c\sqrt{\ell}$. Again, after rotation, we may reduce to the two-dimensional setting, i.e., assume that $x = (0, x, 0, \dots, 0)$ for some $x \leq \delta$.

To derive the lower bound on $|I(x)|$, note that the first coordinate of one of the midpoint of the circle through $\pm ae_2$ with radius $a + \ell$ equals

$$b_0 := \sqrt{(a+\ell)^2 - a^2} = \sqrt{\ell}\sqrt{2a+\ell}.$$

Thus, as illustrated in Figure 6,

$$|I(x)| = \sqrt{(a+\ell)^2 - x^2} + b_0 - \sqrt{a^2 - x^2} = b_0 + \ell \frac{2a+\ell}{\sqrt{(a+\ell)^2 - x^2} + \sqrt{a^2 - x^2}},$$

which is bounded below by a scalar multiple of $\sqrt{\ell}$ since $b_0 \geq \sqrt{2a\ell} \geq \sqrt{\ell}$.

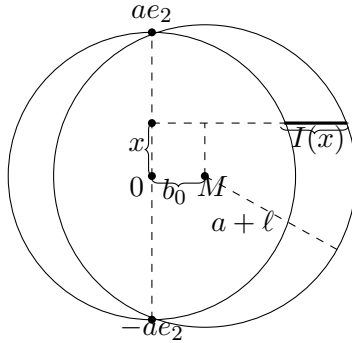


FIGURE 6. Lower bound on crescent volume

□

Next, we bound the volume of almost parallelograms, recalling that $s_{n,+} = 3^d s_n^d$.

Lemma C.2 (Almost parallelogram). *If $\varepsilon > 0$ is sufficiently small, then*

$$\left| \{(x_0, \dots, x_3) \in W_{s_{n,+}}^4 : |d_{23} - d_{01}| \vee |d_{03} - d_{12}| \leq n^{-1+\varepsilon}\} \right| \in O(n^{-1}).$$

Proof. First, by rotating x_3 into the plane spanned by x_0, x_1 and x_2 , we may reduce to the two-dimensional setting. Second, we may assume that $\min_{i \neq j} d_{ij} \geq n^{-1/16}$. Indeed, for instance if $i = 0$ and $j = 1$, then

$$\left| \{(x_0, \dots, x_3) \in W_{s_{n,+}}^4 : d_{01} \leq n^{-1/16} \text{ and } |d_{23} - d_{01}| \leq \ell_n\} \right| \in o(n^{-1}),$$

where we set $\ell_n := n^{-1+\varepsilon}$. Similarly, we may also assume that the angles $\angle x_i x_j x_k$ are at least $n^{-1/16}$ for all pairwise distinct i, j, k . Note that we did not attempt to optimize the exponent $-1/16$.

After these simplifications, it remains to bound the annuli-intersection area $|B_{d_{12}, \ell_n}(x_0) \cap B_{d_{01}, \ell_n}(x_2)|$, as illustrated in Figure 7. To bound this quantity, we rely on the co-area formula from [14, Chapter 3]. More precisely, we have

$$|B_{d_{12}, \ell_n}(x_0) \cap B_{d_{01}, \ell_n}(x_2)| = \int_{d_{01}}^{d_{01} + \ell_n} |\partial B_r(x_0) \cap B_{d_{01}, \ell_n}(x_2)| dr.$$

Hence, if we write P, Q for the endpoints of one of the two arcs $\partial B_r(x_0) \cap B_{d_{01}, \ell_n}(x_2)$ and α for the enclosed angle, it suffices to show that α is of order $O(n^{-1/2})$.

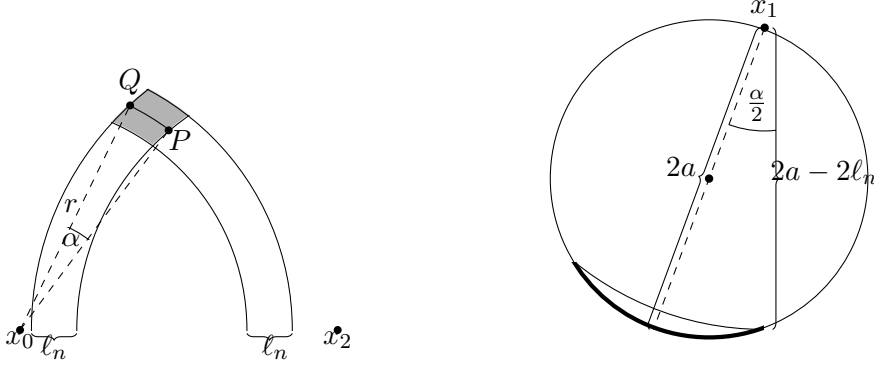


FIGURE 7. Intersection area of annuli (left); Arc describing those x_0 with $d_{01} \geq d_{012} - \ell_n$ (right)

To that end, note that

$$\cos(\angle x_2 x_0 P) = \frac{d_{02}^2 + r^2 - d_{01}^2}{2d_{02}r} \quad \text{and} \quad \cos(\angle x_2 x_0 Q) = \frac{d_{02}^2 + r^2 - (d_{01} + \ell_n)^2}{2d_{02}r}.$$

Since $\alpha = \angle x_2 x_0 Q - \angle x_2 x_0 P$, the mean-value theorem yields some $\alpha' \in [\angle x_2 x_0 P, \angle x_2 x_0 Q]$ with

$$\alpha = \frac{\cos(\angle x_2 x_0 Q) - \cos(\angle x_2 x_0 P)}{\sin(\alpha')} = \ell_n \frac{2d_{01} + \ell_n}{2d_{02}r \sin(\alpha')}.$$

To finish the proof, note that denominator is of order at least $n^{-1/4}$ by the assumptions at the beginning of proof. \square

We conclude the appendix with a final elementary geometric volume bound.

Lemma C.3. *If $\varepsilon > 0$ is sufficiently small, then*

$$|\{(x_0, x_1, x_2) \in W_{s_{n,+}}^3 : d_{01} < d_{012} \leq d_{01} + n^{-2+\varepsilon}\}| \in O(n^{-1/4}).$$

Proof. To ease notation, we put $\ell_n := n^{-2+\varepsilon}$. Arguing similarly as in the proof of Lemma C.2, we may leverage rotational symmetry around the axis formed by $x_1 x_2$ to reduce the proof to the setting $d = 2$. Similarly, we may assume that $d_{12} \geq 2\ell_n$.

Then, for fixed $a > d_{12}$, as observed in the proof of Lemma C.1, the location of all x_0 with $d_{012} = a$ is the union of the two circles with radius a passing through x_1, x_2 . In particular, the location of x_0 with $d_{01} \geq a - \ell_n$ is given by an arc in each of these circles, see Figure 7 (right). Hence, applying the co-area formula [14, Chapter 3] to the level sets of the function $u(x_0) := d_{012}$, we obtain that

$$|\{(x_0, x_1, x_2) \in W_{s_{n,+}}^3 : d_{01} \geq d_{012} - \ell_n\}| = \int_{\ell_n}^{\infty} |\{(x_0, x_1, x_2) \in W_{s_{n,+}}^3 : d_{01} \geq a - \ell_n\} \cap \{d_{012} = a\}| |\nabla u(x_0)|^{-1}$$

Since the gradient of u is bounded away from 0, it suffices to show that the length of these arcs is of order $O(\sqrt{a\ell_n})$. By construction, the angle α associated with one of these arcs satisfies

$\cos(\alpha/2) = (2a - 2\ell_n)/2a$, so that

$$\sin^2(\alpha/2) = 1 - \cos^2(\alpha/2) \leq \frac{\ell_n}{a},$$

Since $\ell_n/a \leq 1/2$, we deduce the asserted $a\alpha \leq a\sqrt{\frac{8\ell_n}{a}} = \sqrt{8a\ell_n}$. □

(Nicolas Chenavier) UNIVERSITÉ DU LITTORAL CÔTE D'OPALE, EA 2597, LMPA, 50 RUE FERDINAND BUISSON, 62228 CALAIS, FRANCE

Email address: `nicolas.chenavier@univ-littoral.fr`

(Christian Hirsch) UNIVERSITY OF GRONINGEN, BERNOULLI INSTITUTE, NIJENBORGH 9, 9747 AG GRONINGEN, THE NETHERLANDS.

Email address: `c.p.hirsch@rug.nl`