Check for updates

Adv. Appl. Probab. **57,** 492–523 (2025)
*doi:*10.1017/apr.2024.61

ON THE ASYMPTOTIC NORMALITY OF PERSISTENT BETTI NUMBERS

JOHANNES KREBS,* *KU Eichstätt-Ingolstadt*
WOLFGANG POLONIK,** *University of California at Davis*

Abstract

Persistent Betti numbers are a major tool in persistent homology, a subfield of topological data analysis. Many tools in persistent homology rely on the properties of persistent Betti numbers considered as a two-dimensional stochastic process $(r, s) \mapsto n^{-1/2}(\beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{X}_n)) - \mathbb{E}[\beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{X}_n))])$. So far, pointwise limit theorems have been established in various settings. In particular, the pointwise asymptotic normality of (persistent) Betti numbers has been established for stationary Poisson processes and binomial processes with constant intensity function in the so-called critical (or thermodynamic) regime; see Yogeshwaran *et al.* (*Prob. Theory Relat. Fields* **167**, 2017) and Hiraoka *et al.* (*Ann. Appl. Prob.* **28**, 2018).

In this contribution, we derive a strong stabilization property (in the spirit of Penrose and Yukich, *Ann. Appl. Prob.* **11**, 2001) of persistent Betti numbers, and we generalize the existing results on their asymptotic normality to the multivariate case and to a broader class of underlying Poisson and binomial processes. Most importantly, we show that multivariate asymptotic normality holds for all pairs (r, s), $0 \le r \le s < \infty$, and that it is not affected by percolation effects in the underlying random geometric graph.

Keywords: Critical regime; strong stabilization; topological data analysis; random geometric complexes; weak convergence

2010 Mathematics Subject Classification: Primary 60D05; 60F05
Secondary 60G55

1. Introduction

In this manuscript we address an important question in topological data analysis (TDA), namely, the study of the weak convergence of persistent Betti numbers

$$\left(n^{-1/2}\left(\beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathcal{X}_n)) - \mathbb{E}\left[\beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathcal{X}_n))\right]\right) : 1 \le i \le \ell\right), \tag{1}$$

where $0 \le q \le d-1$ and $0 \le r_i \le s_i < \infty$ for $1 \le i \le \ell$ $(\ell \in \mathbb{N})$ and where $\mathcal{X}_n$ is either an n-binomial process with a bounded density κ defined on the unit cube $[0,1]^d$ or the corresponding Poisson process with intensity function $n\kappa$ for $n \in \mathbb{N}$.

So far, there exist results on the pointwise asymptotic normality of Betti numbers (i.e., $\ell = 1$) in the case of a homogeneous Poisson process or a binomial process with a constant

Received 2 March 2020; accepted 30 September 2024.
* Postal address: KU Eichstätt-Ingolstadt, Ostenstraße 28, 85072 Eichstätt, Germany. Email address: johannes.krebs@ku.de
** Postal address: Department of Statistics, University of California, Davis, CA 95616, USA. Email address: wpolonik@ucdavis.edu

© The Author(s), 2024. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

density; see Yogeshwaran *et al.* [42]. In the case of a homogeneous Poisson process this result was extended to persistent Betti numbers by Hiraoka *et al.* [16].

Based on the pioneering central limit theorem of Penrose and Yukich [32] for stabilizing functionals on the homogeneous Poisson process, Trinh [37] extends the central limit theorem to strongly stabilizing functionals in the case of an underlying inhomogeneous Poisson process. We will apply the abstract result of Trinh to persistent Betti functions (see below). For this we establish the strong stabilization property of the persistent Betti function, and this is one of our main contributions.

The theory of random geometric complexes is growing rapidly. For pioneering contributions see Kahle [17], Yogeshwaran and Adler [41], or Owada and Adler [27]. We refer the reader to Bobrowski and Kahle [5] for a survey.

Recent contributions in the context of Betti numbers are those of Kahle and Meckes [18], Owada [26], Goel *et al.* [15], Divol and Polonik [10] and Owada and Thomas [28].

TDA is a comparatively young field that has emerged from several contributions in algebraic topology and computational geometry. Milestone contributions which helped to popularize TDA in its early days were made by Edelsbrunner *et al.* [11], Zomorodian and Carlsson [44], and Carlsson [7]. Given a point cloud sampled from a distribution with density f on a d-dimensional manifold, TDA encompasses various techniques aimed at understanding the topology of the manifold and of the density f. TDA methods have been successfully implemented in applied sciences such as biology ([40]), materials science ([21]), and chemistry ([24]). From the mathematical statistician's point of view, a topic of particular interest is the application of TDA to time series; see, e.g., the pioneering works of Seversky *et al.* [35] and Umeda [38], as well as the contributions of Gidea *et al.* to the analysis of financial time series ([12–14]).

Our contribution falls within the area of persistent homology, which is one of the major tools in TDA. We can give only a short introduction to this topic here; a more detailed introduction offering insights into the basic concepts, ideas, and applications of persistent homology can be found in Chazal and Michel [9], Oudot [25], and Wasserman [39].

Given a point cloud in $\mathbb{R}^d$ (a sample of a point process), one first builds simplicial complexes over this point cloud according to a rule that describes the neighborhood relation between points. The two most frequently used simplicial complex models are the Vietoris–Rips complex and the Čech complex, defined below. The topological properties of simplicial complexes considered as geometric structures are characterized by the number of q-dimensional holes they contain—most notably connected components, loops, and cavities (zero-, one- and two-dimensional features). These holes are theoretically defined via a tool from algebraic topology, known as homology. The qth homology of a simplicial complex is determined by a quotient space. Its dimension is the so-called qth Betti number. Intuitively, the qth Betti number counts the number of q-dimensional holes in the simplicial complex.

For a given simplicial complex model, a filtration is an increasing collection of simplicial complexes indexed by a one-dimensional parameter, the so-called filtration parameter, which can be understood as time. Given a filtration on a finite time interval, we can consider the evolution of the qth homology groups, i.e., the dynamic behavior of the Betti numbers. As the underlying simple point process (e.g., a Poisson process on a Euclidean space) is random, these Betti numbers are also random; we are thus considering a stochastic process.

From the applied point of view, the mere knowledge of the evolution of the Betti numbers is often not enough, especially when considering objects obtained from persistence diagrams,

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

such as persistent landscapes. In this context the more general concept of *persistent* Betti numbers is the appropriate tool, and this is the object studied here.

The remainder of this manuscript is organized as follows. In Section 2 we introduce our notation. The main results are presented in Section 3, where we state the property of strong stabilization and present two central limit theorems for persistent Betti numbers. Section 4 offers a short review of important related results in the literature that are also used in our study. The framework of stabilization is treated in detail in Section 5, which also contains further results on the stabilizing properties of persistent Betti numbers. The technical details are given in Section 6 and in Appendix A.

2. Notation

Given a finite subset P of the Euclidean space $\mathbb{R}^d$, the Čech filtration $\mathcal{C}(P) = (\mathcal{C}_r(P) : r \geq 0)$ and the Vietoris–Rips filtration $\mathcal{R}(P) = (\mathcal{R}_r(P) : r \geq 0)$ are defined by

$$\mathcal{C}_r(P) = \left\{\sigma \subseteq P, \bigcap_{x\in\sigma} B(x, r) \neq \emptyset\right\},$$

$$\mathcal{R}_r(P) = \{\sigma \subseteq P, \operatorname{diam}(\sigma) \leq r\},$$

respectively, where $B(x, r) = \{y \in \mathbb{R}^d : \|x - y\| \leq r\}$, $\|\cdot\|$ is the Euclidean norm, and diam is the diameter of a subset of $\mathbb{R}^d$. Each $\sigma \subseteq P$ of size $q + 1$ is a q-simplex, and for each $r \geq 0$, the collections of simplices $\mathcal{C}_r(P)$ and $\mathcal{R}_r(P)$, respectively, form simplicial complexes. Throughout this article, $\mathcal{K}_r(P)$ denotes either the Čech or the Vietoris–Rips complex built from P for some $r \geq 0$. The Čech or the Vietoris–Rips filtration $\mathcal{K}(P)$ is the nested sequence of complexes $\mathcal{K}_r(P)$ as r goes from 0 to $+\infty$.

Consider a filtration $\mathcal{K}(P)$ and a time $r \geq 0$. The chain group generated by the q-dimensional simplices at time r is $C_q(\mathcal{K}_r(P))$. Write $Z_q(\mathcal{K}_r(P))$ for the qth cycle group of the simplicial complex $\mathcal{K}_r(P)$ and $B_q(\mathcal{K}_r(P))$ for the qth boundary group, and let $H_q(\mathcal{K}_r(P))$ be the homology of the simplicial complex $\mathcal{K}_r(P)$ with respect to to the base field $\mathbb{F}_2 = \{0, 1\}$. A q-simplex σ is positive in the filtration $\mathcal{K}(P)$ if, at its filtration time $r(\sigma)$ (the time it enters the complex), its inclusion in the simplicial complex $\mathcal{K}_{r(\sigma)-}$ creates a q-dimensional cycle. Here $\mathcal{K}_{r(\sigma)-}$ is the simplicial complex $\mathcal{K}_{r(\sigma)}$ without the complexes containing σ (as a simplex or as a face). If the q-simplex σ is not positive, then we call it negative.

Let $0 \leq q \leq d - 1$. Then the (r, s)-persistent Betti number of a simplicial complex $\mathcal{K}(P)$ (see [11]) is defined by

$$\begin{aligned}\beta_q^{r,s}(\mathcal{K}(P)) &= \dim \frac{Z_q(\mathcal{K}_r(P))}{Z_q(\mathcal{K}_r(P)) \cap B_q(\mathcal{K}_s(P))} \\ &= \dim Z_q(\mathcal{K}_r(P)) - \dim Z_q(\mathcal{K}_r(P)) \cap B_q(\mathcal{K}_s(P)).\end{aligned}$$

The Betti number $\beta_q^r(\mathcal{K}(P))$ is defined as $\beta_q^{r,r}(\mathcal{K}(P))$, $r \geq 0$; in particular, $\beta_q^r(\mathcal{K}(P)) = \dim Z_q(\mathcal{K}_r(P)) - \dim B_q(\mathcal{K}_r(P))$.

The persistent Betti number $\beta_q^{r,s}(\mathcal{K}(P))$ is closely related to the persistence diagram of the underlying point cloud P; see [16] for further details. It equals the number of q-dimensional topological features (points in the qth persistence diagram) born before time r and still alive at time s (see Figure 1). Persistent Betti numbers are translation-invariant, i.e., $\beta_q^{r,s}(\mathcal{K}(P + v)) =$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

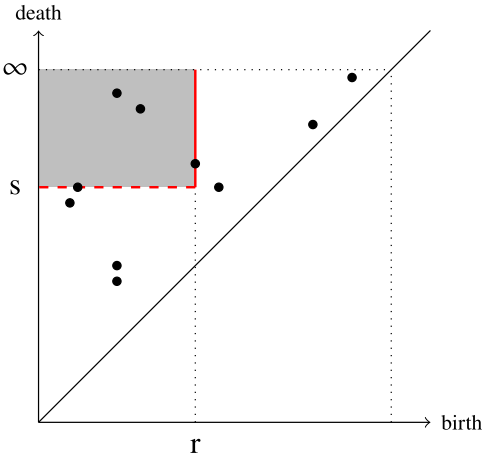


FIGURE 1. The persistent Betti number $\beta_q^{r,s}(\mathcal{K}(P))$ equals the number of points in the gray-shaded rectangle; the point on the dashed red line is not counted, whereas the point on the solid red line is.

$\beta_q^{r,s}(\mathcal{K}(P))$ for each $v \in \mathbb{R}^d$. The add-one cost function

$$\mathfrak{D}_0\beta_q^{r,s}(\mathcal{K}(P)) = \beta_q^{r,s}(\mathcal{K}(P \cup \{0\})) - \beta_q^{r,s}(\mathcal{K}(P))$$

is an important tool in our analysis.

We let $\mathcal{P}$ and $\mathcal{P}'$ be two independent and homogeneous Poisson processes on $\mathbb{R}^d$ with unit intensity, observed on increasing observation windows $B_n = [-2^{-1}n^{1/d}, 2^{-1}n^{1/d}]^d$. Given a function $w \geq 0$, we denote by $\mathcal{P}(w)$ the (inhomogeneous) Poisson process with intensity function w.

We also use the following notation: $\Delta = \{(r, s) : 0 \leq r \leq s < \infty\}$ denotes the domain of the persistent Betti function. We let $Q(x, r) = \{y \in \mathbb{R}^d : |y_i - x_i| \leq r \text{ for } 1 \leq i \leq d\}$ and $Q(x) = (-1/2, 1/2]^d + x$ for $x \in \mathbb{R}^d$ and $r > 0$. For $y, z \in \mathbb{Z}^d$, we write $y \prec z$ if y precedes z in the lexicographic ordering on $\mathbb{Z}^d$, and we write $y \preceq z$ if either $y \prec z$ or $y = z$. If $f : \mathbb{R} \to \mathbb{R}$, we write $\|f\|_\infty$ for the sup-norm of f. We let $\Rightarrow$ denote convergence in distribution of a sequence of random variables. Throughout the article, we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a sufficiently rich probability space, on which all random variables are defined.

3. Main results

We now present the first main result, discussed in detail later in Section 5.

Theorem 1. *Let $\lambda > 0$, $(r, s) \in \Delta$, and $q \in \{0, \ldots, d-1\}$. There is an $\mathcal{F}$-measurable random variable $S_q^{(r,s)} := S_q^{(r,s)}(\mathcal{P}(\lambda))$ which is almost surely (a.s.) finite, such that for all finite sets $A \subseteq \mathbb{R}^d \setminus B(0, S_q^{(r,s)})$, the add-one cost function satisfies*

$$\mathfrak{D}_0\beta_q^{r,s}\Big(\mathcal{K}\big(\big(\mathcal{P}(\lambda) \cap B(0, S_q^{(r,s)})\big) \cup A\big)\Big) \equiv \mathfrak{D}_0\beta_q^{r,s}\Big(\mathcal{K}\big(\mathcal{P}(\lambda) \cap B(0, S_q^{(r,s)})\big)\Big) \quad a.s.$$

Thus, the persistent Betti function is strongly stabilizing on the homogeneous Poisson process $\mathcal{P}(\lambda)$, in the spirit of Penrose and Yukich [32], for each intensity $\lambda \in \mathbb{R}_+$, each pair

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

$(r, s) \in \Delta$, and each dimension q. The proof of Theorem 1 relies on an abstract stabilization result stated in Theorem 4; see Section 5. The proofs of both theorems are given in Subsection 6.1.

By the property of strong stabilization, there are random variables $\Delta_0^{r,s}(\infty)$ taking values in $\mathbb{Z}$ and N_0 taking values in $\mathbb{N}$ such that, for all $n \geq N_0$,

$$\beta_q^{r,s}(\mathcal{K}(\mathcal{P} \cap B_n)) - \beta_q^{r,s}(\mathcal{K}(([\mathcal{P} \setminus Q(0)] \cup [\mathcal{P}' \cap Q(0)]) \cap B_n)) \equiv \Delta_0^{r,s}(\infty);$$

see Lemma 3. Let $\mathcal{F}_0$ be the σ-field generated by $\mathcal{P}$ restricted to $\bigcup_{y \in \mathbb{Z}^d : y \preceq 0} Q(y)$. Define $\gamma((u, v), (r, s)) = \mathbb{E}\left[\mathbb{E}\left[\Delta_0^{u,v}(\infty)|\mathcal{F}_0\right]\mathbb{E}\left[\Delta_0^{r,s}(\infty)|\mathcal{F}_0\right]\right]$. The asymptotic normality in the Poisson sampling scheme can be derived directly from the the strong stabilization stated in Theorem 3 and the abstract result of Trinh [37] via the Cramér–Wold device, as follows.

Theorem 2. *Let $\mathcal{P}_n = \mathcal{P}(n\kappa)$ be a Poisson process with intensity $n\kappa$ on $[0, 1]^d$, where κ is a bounded and measurable density function. Let $X \sim \kappa$ and let $(r_i, s_i) \in \Delta$ for $1 \leq i \leq \ell$ and $\ell \in \mathbb{N}$. Then*

$$\begin{pmatrix} n^{-1/2}\left(\beta_q^{r_1,s_1}(\mathcal{K}(n^{1/d}\mathcal{P}_n)) - \mathbb{E}\left[\beta_q^{r_1,s_1}(\mathcal{K}(n^{1/d}\mathcal{P}_n))\right]\right) \\ \vdots \\ n^{-1/2}\left(\beta_q^{r_\ell,s_\ell}(\mathcal{K}(n^{1/d}\mathcal{P}_n)) - \mathbb{E}\left[\beta_q^{r_\ell,s_\ell}(\mathcal{K}(n^{1/d}\mathcal{P}_n))\right]\right) \end{pmatrix} \Rightarrow \Psi,$$

where $\Psi \sim \mathcal{N}(0, \Sigma)$ has a multivariate normal distribution with mean zero and covariance matrix $\Sigma \geq 0$ given by

$$\Sigma(i, j) = \mathbb{E}\Big[\gamma(\kappa(X)^{1/d}((r_i, s_i), (r_j, s_j)))\Big] \quad (1 \leq i, j \leq \ell).$$

Furthermore,

$$\lim_{n\to\infty} n^{-1}\operatorname{Cov}\Big(\beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathcal{P}_n)), \beta_q^{r_j,s_j}(\mathcal{K}(n^{1/d}\mathcal{P}_n))\Big) = \Sigma(i, j) \qquad (1 \leq i, j \leq \ell).$$

Moreover, for $0 \leq r \leq s < \infty$, define

$$\alpha(r, s) := \mathbb{E}\Big[\mathfrak{D}_0\beta_q^{r,s}\Big(\mathcal{K}\big(\mathcal{P} \cap B(0, S_q^{(r,s)})\big)\Big)\Big], \tag{2}$$

where $S_q^{(r,s)} = S_q^{(r,s)}(\mathcal{P})$ is as in Theorem 1. Then for the binomial sampling scheme the result is as follows.

Theorem 3. *Let $\mathbb{X}_n$ be an n-binomial process with density function κ on $[0, 1]^d$, which is bounded and measurable. Let $X \sim \kappa$ and let $(r_i, s_i) \in \Delta$ for $1 \leq i \leq \ell$ and $\ell \in \mathbb{N}$. Then*

$$\begin{pmatrix} n^{-1/2}\left(\beta_q^{r_1,s_1}(\mathcal{K}(n^{1/d}\mathbb{X}_n)) - \mathbb{E}\left[\beta_q^{r_1,s_1}(\mathcal{K}(n^{1/d}\mathbb{X}_n))\right]\right) \\ \vdots \\ n^{-1/2}\left(\beta_q^{r_\ell,s_\ell}(\mathcal{K}(n^{1/d}\mathbb{X}_n)) - \mathbb{E}\left[\beta_q^{r_\ell,s_\ell}(\mathcal{K}(n^{1/d}\mathbb{X}_n))\right]\right) \end{pmatrix} \Rightarrow \widetilde{\Psi},$$

where $\widetilde{\Psi} \sim \mathcal{N}(0, \widetilde{\Sigma})$ has a multivariate normal distribution with mean zero and covariance matrix $\widetilde{\Sigma} \geq 0$ given by

$$\begin{aligned}\widetilde{\Sigma}(i, j) = {} & \mathbb{E}\Big[\gamma(\kappa(X)^{1/d}((r_i, s_i), (r_j, s_j)))\Big] \\ & - \mathbb{E}\Big[\alpha(\kappa(X)^{1/d}(r_i, s_i))\Big]\mathbb{E}\Big[\alpha(\kappa(X)^{1/d}(r_j, s_j))\Big] \quad (1 \leq i, j \leq \ell).\end{aligned}$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Furthermore,

$$\lim_{n\to\infty} n^{-1}\operatorname{Cov}\left(\beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathbb{X}_n)),\, \beta_q^{r_j,s_j}(\mathcal{K}(n^{1/d}\mathbb{X}_n))\right) = \widetilde{\Sigma}(i,j) \qquad (1 \le i,j \le \ell).$$

We conclude this section with some discussion of the results and the techniques used in the proofs of these theorems, which are given in Subsections 6.2 and 6.3. The univariate central limit theorems for Betti numbers ($r = s$) have already been formulated by Trinh [37] (see also Proposition 2 in this manuscript) under the condition that the parameter r is chosen to be such that no percolation occurs. Here, as we can rely on the strong stabilization property, we can omit this restriction.

For the derivation of the multivariate results in the Poisson sampling scheme, we can rely on the abstract result of Trinh [37, Theorem 3.3] and derive the covariance structure with the help of results from Penrose and Yukich [32]. Multivariate central limit theorems in the spatial context are also studied by Penrose [30]. The results in the binomial setting are established using Trinh [37, Theorem 3.9], which itself relies on a de-Poissonization argument.

Finally, we mention that it is currently unknown whether or not the limiting covariance matrices are strictly positive definite.

4. Related results

Below we discuss some literature closely related to our study. The techniques employed to obtain these results are tools from geometric probability, which studies geometric quantities deduced from simple point processes. A classical result of Steele [36] shows the convergence of the total length of the minimum spanning tree built from an independent and identically distributed (i.i.d.) sample of n points in the unit cube. There are several generalizations of this work; for notable contributions see McGivney and Yukich [22], Yukich [43], Penrose and Yukich [33], and the monograph of Penrose [29].

A different type of contribution, equally important, is that of Penrose and Yukich [32], which considers asymptotic normality of functionals built on Poisson and binomial processes. For completeness, we mention that the study of Gaussian limits (as in [32, 33]) is not limited to the total mass functional. It can be extended to random point measures obtained from the points of a marked point process; see, e.g., [3, 4, 31].

Goel *et al.* [15] prove a convergence result for the expectation of Betti numbers in the critical regime. Their result generalizes directly to persistent Betti numbers, giving us the following well-known result.

Proposition 1. *Let $0 < r \le s < \infty$. Let $\mathcal{X}_n$ be either a Poisson process with intensity $n\kappa$ on $[0,1]^d$ or an n-binomial process with density κ. Then*

$$\lim_{n\to\infty} n^{-1}\mathbb{E}\Big[\beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{X}_n))\Big] = \mathbb{E}\Big[\hat{b}_q(r\kappa(X')^{1/d}, s\kappa(X')^{1/d})\Big],$$

where X' has density κ and where $\hat{b}_q(r,s)$ is the limit of $n^{-1}\mathbb{E}\big[\beta_q^{r,s}(\mathcal{K}((n^{1/d}\mathbb{X}_n^*))\big]$ for a homogeneous Poisson process $\mathbb{X}_n^*$ on $[0,1]^d$ with intensity n.

So far, normality results for (persistent) Betti numbers exist only in a pointwise sense and are rather direct consequences of Theorems 2.1 and 3.1 from [32]. We quote them here in a way that is more in line with our framework. For this we need the notion of the interval of co-existence $I_d(\mathcal{P})$ of a Poisson process $\mathcal{P}$ with unit intensity on $\mathbb{R}^d$. This is determined by the critical radii for percolation of the occupied and the vacant component, respectively, which are

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

defined as follows:

$$r_c(\mathcal{P}) := \inf\{r : \mathbb{P}(\mathcal{C}_r(\mathcal{P}) \text{ percolates}) > 0\}$$

and

$$r_c^*(\mathcal{P}) := \sup\{r : \mathbb{P}(\mathbb{R}^d \setminus \mathcal{C}_r(\mathcal{P}) \text{ percolates}) > 0\}.$$

Both probabilities inside the infimum and supremum are either 0 or 1 by Kolmogorov's 0–1 law. The quantity $r_c(\mathcal{P})$ is called the *critical radius for percolation of the occupied component*, and $r_c^*(\mathcal{P})$ is called the *critical radius for percolation of the vacant component*. The *interval of co-existence*, on which unbounded components of both the (Boolean) model $\mathcal{C}_r(\mathcal{P})$ and its complement co-exist, is defined as follows:

$$I_d(\mathcal{P}) := \begin{cases} (r_c, r_c^*] & \text{if } \mathbb{P}(\mathcal{C}_{r_c}(\mathcal{P}) \text{ percolates}) = 0, \\ [r_c, r_c^*] & \text{otherwise.} \end{cases}$$

We know that in two dimensions, $I_2(\mathcal{P}) = \emptyset$ from [23, Theorem 4.4 and 4.5]. Moreover, from [34, Theorem 1], we know that $I_d(\mathcal{P}) \neq \emptyset$ for each $d \geq 3$. Thus we have the following results.

Proposition 2. *(Pointwise normality of (persistent) Betti numbers.)*

(i) *(Hiraoka et al. [16, Theorem 5.2].) Let $\mathcal{P}|_{[0,n^{1/d}]^d}$ be the restriction of $\mathcal{P}$ to $[0, n^{1/d}]^d$, and let $0 \leq r \leq s < \infty$. Then there is a $\sigma^2(r, s) \in \mathbb{R}_+$ such that*

$$n^{-1/2}\left(\beta_q^{r,s}(\mathcal{K}(\mathcal{P}|_{[0,n^{1/d}]^d})) - \mathbb{E}\left[\beta_q^{r,s}(\mathcal{K}(\mathcal{P}|_{[0,n^{1/d}]^d}))\right]\right) \Rightarrow \Phi_1,$$

where Φ_1 has a normal distribution with mean zero and variance $\sigma^2(r, s) \geq 0$.

(ii) *(Yogeshwaran et al. [42, Theorem 4.7].) Let $\mathcal{K}$ be the Čech filtration, and let $0 \leq r < \infty$ be such that $r \notin I_d(\mathcal{P})$. For each $n \in \mathbb{N}$, let $\mathbb{X}_n$ be an n-binomial process with a uniform density on $[0, 1]^d$. Then there is a $0 < \tau^2(r) \leq \sigma^2(r, r)$ (with σ^2 from (i)) such that*

$$n^{-1/2}\left(\beta_q^r(\mathcal{K}(n^{1/d}\mathbb{X}_n)) - \mathbb{E}\left[\beta_q^r(\mathcal{K}(n^{1/d}\mathbb{X}_n))\right]\right) \Rightarrow \Phi_2,$$

where Φ_2 has a normal distribution with mean zero and variance $\tau^2(r) > 0$.

(iii) *(Trinh [37, Theorem 4.1].) Let κ be a bounded density function with compact support, and let $\mathcal{K}$ be the Čech filtration. Let $0 \leq q \leq d - 1$. Let $r \in (0, (\sup \kappa)^{-1/d} r_c)$. Then*

$$n^{-1/2}\left(\beta_q^r(\mathcal{K}(n^{1/d}\mathcal{P}(n\kappa))) - \mathbb{E}\left[\beta_q^r(\mathcal{K}(n^{1/d}\mathcal{P}(n\kappa)))\right]\right) \Rightarrow \Phi_3,$$

where Φ_3 has a normal distribution with mean zero and variance $\widetilde{\sigma}^2 > 0$; here $\widetilde{\sigma}^2 = \int \sigma^2(\kappa(x)^{1/d}(r, r))\kappa(x)\mathrm{d}x$ with σ^2 from (i). A similar statement is true for the binomial process.

First, we remark that the above statements in their original versions are also valid for more general domains $\tilde{B}_n \subseteq \mathbb{R}^d$ which are not necessarily rectangular. Furthermore, we remark that Hiraoka *et al.* [16] prove their theorem for a general class of filtrations which contains, among others, the Čech and the Vietoris–Rips filtration. Moreover, Theorem 4.7 of Yogeshwaran *et al.* [42] also contains a version of (ii) for Betti numbers of the homogeneous Poisson process,

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

which, in the above list, is contained in the result (i). The result of Trinh [37] is already for Betti numbers from a general density function κ, but the parameter choice for r depends on r_c. Moreover, Trinh [37] points out that in the case $d = 2$ there are no restrictions on the choice of r as $I_2(\mathcal{P})$ is empty; this can be shown with a duality property. Finally, regarding (iii), Yogeshwaran *et al.* [42] remark that the condition $r \notin I_d(\mathcal{P})$ is likely to be superfluous; as already mentioned, we show that this condition can indeed be removed.

5. Background on strong stabilization

In our analysis of the multivariate asymptotic normality of persistent Betti numbers, stabilization properties are crucial. Kesten and Lee [19] introduced stabilization to prove asymptotic normality for the weight of the Euclidean minimal spanning tree. The concept was extended by Penrose and Yukich [32, 33] to treat general functionals defined on Poisson and binomial point processes; it goes as follows. Consider a functional H which is defined on finite subsets of $\mathbb{R}^d$, and define its add-one cost function as

$$\mathfrak{D}_0 H(\mathcal{H}) := H(\mathcal{H} \cup \{0\}) - H(\mathcal{H})$$

for $\mathcal{H} \subseteq \mathbb{R}^d$ finite. The functional H is strongly stabilizing on the homogeneous Poisson process with intensity $\lambda \in (0, \infty)$ on $\mathbb{R}^d$, denoted by $\mathcal{P}(\lambda)$, if there exist a.s. finite random variables S and $\mathfrak{D}_\infty H$ such that for all finite $A \subseteq \mathbb{R}^d \setminus B(0, S)$,

$$\mathfrak{D}_0 H((\mathcal{P}(\lambda) \cap B(0, S)) \cup A) = \mathfrak{D}_\infty H \quad \text{a.s.}$$

Recall that, for $n \in \mathbb{N}$, the sets $B_n = [-n^{1/d}/2, n^{1/d}/2]^d$ denote observation windows, and let $\mathcal{A}$ be the collection $\{B_n + x : x \in \mathbb{R}^d, n \in \mathbb{N}\}$. The functional H is weakly stabilizing on $\mathcal{A}$ (for $\mathcal{P}(\lambda)$) if there is an a.s. finite random variable $\mathfrak{D}'_\infty H$ such that, for any such sequence $(A_n : n \in \mathbb{N})$ from the collection $\mathcal{A}$ with $\lim_{n\to\infty} A_n = \mathbb{R}^d$,

$$\mathfrak{D}_0 H(\mathcal{P}(\lambda) \cap A_n) \to \mathfrak{D}'_\infty H \quad \text{a.s. as } n \to \infty.$$

Recall that the set-theoretic limit $\lim_{n\to\infty} A_n = \mathbb{R}^d$ is equivalent to $\lim_{n\to\infty} \mathbf{1}_{A_n}(x) = 1$ for all $x \in \mathbb{R}^d$. The stabilization of a functional defined on subsets of a point process roughly means that a local change in the point process (e.g., the addition or subtraction of finitely many points) affects the value of the functional only locally. This latter phenomenon can be described in terms of various notions. We consider two radii of stabilization for the persistent Betti function $\beta_q^{r,s}$. Their functionality is related to the classical weak and strong stabilization properties given above. Properties of these radii are addressed in detail below.

Consider a point process P on $\mathbb{R}^d$ without accumulation points, and let Q be finite with circumcenter $z_Q \in \mathbb{R}^d$ and circumradius L_Q. Thus, $Q \subseteq B(z_Q, L_Q)$ for $L_Q \geq 0$ minimal. For short, write $\mathcal{K}_{r,a} = \mathcal{K}_r(P \cap B(z_Q, a))$ and $\mathcal{K}'_{r,a} = \mathcal{K}_r((P \cup Q) \cap B(z_Q, a))$ for $a, r \geq 0$. In the following, the reference case is that $Q \subset Q(0)$, so that $z_Q \in Q(0)$ and $L_Q \leq \sqrt{d}/2$. (Recall that $Q(0)$ is defined in Section 2.)

Radius of weak stabilization: Define the radius of weak stabilization of (r, s) by

$$\begin{aligned}\rho_q^{(r,s)}(P, Q) := \inf\{R > 0 :\ &\dim Z_q(\mathcal{K}'_{r,a}) - \dim Z_q(\mathcal{K}_{r,a}) = \text{const. } \forall a \geq R \text{ and} \\ &\dim Z_q(\mathcal{K}'_{r,a}) \cap B_q(\mathcal{K}'_{s,a}) - \dim Z_q(\mathcal{K}_{r,a}) \cap B_q(\mathcal{K}_{s,a}) = \text{const. } \forall a \geq R\},\end{aligned}$$

$$\rho^{(r,s)}(P, Q) := \max_{0 \leq q \leq d-1} \rho_q^{(r,s)}(P, Q). \tag{3}$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

One can use similar ideas as in the proof of Lemma 5.3 in Hiraoka *et al.* [16] to show that if P a.s. has no accumulation points and if Q is finite, then $\rho^{(r,s)}(P, Q)$ is a.s. finite; we do this in Lemma 5 in the appendix. A similar result was also obtained by Hiraoka *et al.* [16] for the add-one cost function of persistent Betti numbers. Our definition of the radius of weak stabilization implies that for all $0 \le r \le s$ and for all $q \in \{0, \dots, d-1\}$,

$$\beta_q^{r,s}\Big(\mathcal{K}\big((P \cup Q) \cap B(z_Q, R)\big)\Big) - \beta_q^{r,s}\Big(\mathcal{K}\big(P \cap B(z_Q, R)\big)\Big) = \text{const.}$$

as a function in R, for $R \ge \rho_q^{(r,s)}(P, Q)$.

Radius of strong stabilization: Let $r > 0$ be an arbitrary but fixed filtration parameter. Let $\mu(r)$ be an upper bound on the diameter of simplices in the filtration at time r. For the Vietoris–Rips filtration, $\mu(r)$ equals r. For the Čech filtration $\mu(r) = 2r$ is a sharp bound. We choose $a \ge a^*(r) = L_Q + \mu(r)$ sufficiently large so that all simplices containing at least one point of Q have a filtration time of at most $a^*(r)$; recall that L_Q denotes the circumradius of Q.

Given P and Q, let $\sigma_{q,i}^r$, $i = 1, \dots, m_q$, be the q-simplices in $\mathcal{K}'_{r,a} \setminus \mathcal{K}_{r,a}$ contained in the ball $B(z_Q, a)$ that are created up to filtration time r by the addition of the points in Q to the point process P. Without loss of generality, we may assume the simplices are already ordered according to their filtration times; simplices with the same filtration time are ordered at random.

The number R that limits the knowledge of a point process P' to the ball $B(z, R)$ is referred to as the information horizon (with respect to z); i.e., we observe only the process $P' \cap B(z, R) = P'|_{B(z,R)}$ and the corresponding simplicial complexes restricted to $P'|_{B(z,R)}$, i.e., the complexes $\mathcal{K}_r(P'|_{B(z,R)})$, $r \ge 0$.

Let ∂_q denote the qth boundary map. For $i = 1, \dots, m_q$, define the following quantities depending on the filtration parameter $r \ge 0$:

$$C_{q,i}^r(a) = C_q(\mathcal{K}_{r,a}) \oplus \langle \sigma_{q,1}^r, \dots, \sigma_{q,i}^r \rangle \quad \text{and} \quad Z_{q,i}^r(a) = \ker(\partial_q \colon C_{q,i}^r(a) \to B_{q-1,i}^r(a)),$$

where $B_{q-1,i}^r(a) = \partial_q(C_{q,i}^r(a)) \subseteq C_{q-1,i}^r(a)$ is the image of ∂_q. First, we define

$$\begin{aligned}\widetilde{\rho}_q^r(P, Q) := \inf \Big\{ R \ge a^*(r) \;\Big|\; & \text{for each } \sigma_{q,i}^r,\ i \in \{1, \dots, m_q\} : \\ & \text{either } \Big[\exists c \in Z_{q,i}^r(R) : \sigma_{q,i}^r \text{ is contained in } c\Big] \\ & \text{or } \Big[\text{conditional on } (P \cup Q)|_{B(z_Q,R)} \\ & \Big[\forall a \ge R,\ \forall c \in Z_{q,i}^r(a) : \sigma_{q,i}^r \text{ is not contained in } c \Big] \text{ is true } \Big] \Big\}. \end{aligned} \tag{4}$$

This definition means the following. Consider an information horizon $R > \widetilde{\rho}_q^r(P, Q)$; i.e., we observe all points from the point process $(P \cup Q) \cap B(z_Q, R)$. Then, if we include the q-simplex $\sigma_{q,i}^r$ in the simplicial complex, we already have the information either that $\sigma_{q,i}^r$ creates a new q-cycle or that it does not, even when we are additionally including points from an infinite information horizon. In other words, given $(P \cup Q) \cap B(z_Q, R)$, the event of which simplices $\sigma_{q,i}^r$ are ultimately positive is decidable (i.e. computable or recursive). By a simplex being 'ultimately positive' we mean that if the information horizon is large enough, then we see that the simplex is part of a cycle. Similarly, a simplex staying negative means that it never becomes part of a cycle, even if the information horizon is infinite.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Thus, the event of which simplices $\sigma^r_{q,i}$ are ultimately positive being decidable means that having observed $P \cup Q$ up to the information horizon R, i.e., $(P \cup Q) \cap B(z_Q, R)$, where R is 'large enough' ($R > \widetilde{\rho}^r_q(P, Q)$), we know that each potential cycle in $C_q(\mathcal{K}_r(\mathcal{P}(\lambda)|_{B(z_Q,R)})) \oplus \langle \sigma^r_{q,1}, \ldots, \sigma^r_{q,i} \rangle$ containing $\sigma^r_{q,i}$ has already terminated, meaning that $\sigma^r_{q,i}$ is positive, or that $\sigma^r_{q,i}$ will stay negative.

In order to state whether the persistent Betti number for a given pair (r, s) and a given dimension q has stabilized, we need to know whether a new cycle in $Z_q(\mathcal{K}'_{r,\widetilde{\rho}^r_q})$ is also a q-dimensional feature, i.e., is not eventually a boundary in $B_q(\mathcal{K}'_{s,a}) \cap Z_q(\mathcal{K}'_{r,a})$ for some $a > \widetilde{\rho}^r_q$.

For this purpose, apart from the additional q-simplices $\sigma^r_{q,1}, \ldots, \sigma^r_{q,m_q}$, we write $\widetilde{\sigma}^s_{q+1,1}, \ldots, \widetilde{\sigma}^s_{q+1,\widetilde{m}_{q+1}}$ for the additional $(q+1)$-simplices in $\mathcal{K}'_{s,a} \setminus \mathcal{K}_{s,a}$ for $a \geq a^*(s)$.

We can repeat the considerations from above for the $(q+1)$-dimensional simplices and a filtration parameter equal to s. Then by definition, after time $\widetilde{\rho}^s_{q+1}(P, Q)$, we know for each $\widetilde{\sigma}^s_{q+1,i}$ whether it is positive or it is negative for all $a > \widetilde{\rho}^s_{q+1}(P, Q)$.

Consequently, the radius of strong stabilization for the pair $(r, s) \in \Delta$ is as follows:

$$S^{(r,s)}_q(P, Q) = \max\{\widetilde{\rho}^r_q(P, Q), \widetilde{\rho}^s_{q+1}(P, Q)\}. \tag{5}$$

At this stage there is a major difference between the Čech complex and the Vietoris–Rips complex if $q = d - 1$. For the Čech filtration, there are no q-dimensional cycles in d-dimensional Euclidean space for $q \geq d$. Thus, for the Čech filtration, $S^{(r,s)}_{d-1}(P, Q) = \max\{\widetilde{\rho}^r_{d-1}(P, Q), a^*(s)\}$. For the Vietoris–Rips filtration, however, there can be q-dimensional cycles for every possible dimension q; see Bobrowski and Kahle [5]. In particular, $S^{(r,s)}_{d-1}(P, Q) > \max\{\widetilde{\rho}^r_{d-1}(P, Q), a^*(s)\}$ can occur.

In the following, if $Q = \{0\}$, we simply write $\rho^{(r,s)}(P)$, $\widetilde{\rho}^r_q(P)$ or $S^{(r,s)}_q(P)$ for convenience. Next, we show in Theorem 4 that the radius $\widetilde{\rho}^r_q(P, Q)$ is a.s. finite for each $q \geq 0$ and $r \in \mathbb{R}_+$ if P equals a homogeneous Poisson process modulo a finite set of points and $Q \subseteq \mathbb{R}^d$ is finite. In particular, this implies the strong stabilization property of the persistent Betti number $\beta^{r,s}_q$, in the sense that $S^{(r,s)}_q$ from (5) is finite, which in turn leads to Theorem 1.

Theorem 4. *For a Poisson process with constant intensity $\lambda \in \mathbb{R}_+$ and two finite (disjoint) sets $Q_1, Q_2 \subseteq Q(0)$, the radius $\widetilde{\rho}^r_q(\mathcal{P}(\lambda) \cup Q_1, Q_2)$ is a.s. finite for each q and each $r > 0$. In particular, the radius of strong stabilization $S^{(r,s)}_q$ is finite for each q and each $(r, s) \in \Delta$.*

We apply arguments from continuum percolation theory to prove this theorem. These arguments are also used to prove uniqueness of the occupied and vacant component in the Boolean model; see, e.g., Aizenman *et al.* [1, 2] and Burton and Keane [6], as well as the monograph of Meester and Roy [23].

Furthermore, we have the following relation between the two radii.

Lemma 1. *We have $\rho^{(r,s)}_q(P, Q) \leq S^{(r,s)}_q(P, Q)$ for each pair $(r, s) \in \Delta$, $q \in \{0, \ldots, d-1\}$.*

Thus, weak stabilization measured in terms of $\rho^{(r,s)}_q$ is always implied by strong stabilization measured in terms of $S^{(r,s)}_q$. The proof of this lemma is given in Subsection 6.1.

Moreover, our results are not limited to this *static* case, in which we consider only one Poisson process and the persistent Betti number for one pair (r, s): we show in Theorem 5 that Borel probability measures induced by the radius of strong (and of weak) stabilization are tight over a variety of parameter ranges.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

The theorem is divided into three parts. Part (1) concerns uniform stabilization over a variety of homogeneous Poisson processes. These stabilization properties then enable us to derive the results in Parts (2) and (3), where we consider stabilization for the binomial and the Poisson sampling schemes, respectively. The proof of Theorem 5 is deferred to Appendix 7.

Theorem 5. *(Uniform stabilization) For $m \in \mathbb{N}$, let $\mathfrak{Q}_m = \{\{y_1, \dots, y_k\} : y_i \in Q(0), i = 1, \dots, k, k \le m\}$ be the class of sets with at most m points in Q(0). Let $+\infty > \bar{r} \ge \underline{r} > 0$ and $m \in \mathbb{N}$ be arbitrary but fixed. Then we have the following:*

(1) ***Stabilization for the homogeneous Poisson case:*** *The laws of*

$$\{\widetilde{\rho}_q^r(\mathcal{P}(\lambda) \cup Q_1, Q_2)) : \underline{r} \le r \le \bar{r}, \lambda \in \mathbb{R}_+, Q_1, Q_2 \in \mathfrak{Q}_m, q = 0, \dots, d-1\}$$

are tight for each $m \in \mathbb{N}$.

(2) ***Stabilization in the Poisson sampling scheme:*** *Let ν be a probability density on $[0,1]^d$. For $n \in \mathbb{N}$ and $L > 0$, set $B''_{n,L} = \{z \in \mathbb{R}^d : B(z, L) \subseteq [0, n^{1/d}]^d\}$. Consider a specific continuous density κ on $[0,1]^d$.*
Let $\varepsilon > 0$. Then there are $b > 0$, $n_0 \in \mathbb{N}$, and $L \in \mathbb{R}_+$ such that, uniformly in $q = 0, \dots, d-1$ and $r \in [\underline{r}, \bar{r}]$,

$$\sup_{n \ge n_0} \sup_{z \in B''_{n,L}} \sup_{Q_1, Q_2 \in z + \mathfrak{Q}_m} \mathbb{P}(\widetilde{\rho}_q^r(n^{1/d}\mathcal{P}(n\nu) \cup Q_1, Q_2) \ge L) \le \varepsilon$$

for all densities ν on $[0,1]^d$ satisfying $\|\nu - \kappa\|_\infty \le b$.
For each $n \in \mathbb{N}$, let $\mathcal{V}_n, \mathcal{W}_n$ be Poisson processes on $[0, n^{1/d}]^d$ that are independent of $n^{1/d}\mathcal{P}(n\nu)$, and whose intensity functions on $\mathbb{R}^d$ are uniformly bounded in n. Then for each $\varepsilon > 0$, there are $b > 0$, $n_0 \in \mathbb{N}$, and $L > 0$ such that, uniformly in $q = 0, \dots, d-1$ and $r \in [\underline{r}, \bar{r}]$,

$$\sup_{n \ge n_0} \sup_{z \in B''_{n,L}} \mathbb{P}(\widetilde{\rho}_q^r(n^{1/d}\mathcal{P}(n\nu) \cup (\mathcal{V}_n \cap Q(z)), \mathcal{W}_n \cap Q(z)) \ge L) \le \varepsilon$$

for all densities ν on $[0,1]^d$ satisfying $\|\nu - \kappa\|_\infty \le b$.

(3) ***Stabilization in the binomial sampling scheme:*** *Let $\mathbb{X}_n$ be an n-binomial process on $[0,1]^d$ obtained from an i.i.d. sequence $(X_k : k \in \mathbb{N})$ with common density κ. Let X' be a random variable independent of $(X_k : k \in \mathbb{N})$, and with continuous density κ on $[0,1]^d$. Write $Q_{m,n}$ for the point process $n^{1/d}(\mathbb{X}_m - X')$ for $m \in J_n = [n - h(n), n + h(n)]$, where the function h satisfies $h(n) \to \infty$ and $h(n)/n \to 0$ as $n \to \infty$. Then the family $\{\widetilde{\rho}_q^r(Q_{m,n}, \{0\}) : n \in \mathbb{N}, m \in J_n, \underline{r} \le r \le \bar{r}, q = 0, \dots, d-1\}$ is tight.*

Furthermore, with $\widetilde{\Delta} = \{(r,s) \in \Delta : w_1 \le r \le s \le w_2\}$, where $w_2 \ge w_1 > 0$ are arbitrary, all these results remain valid if $\widetilde{\rho}_q^r$, $\underline{r} \le r \le \bar{r}$, is replaced by $\rho_q^{(r,s)}$, $(r,s) \in \widetilde{\Delta}$.

6. Technical results

This section consists of three parts. In the first part we derive the stabilization results. In the second part we prove the asymptotic normality of the finite-dimensional distributions in the case of underlying Poisson processes. In the third part we prove the same limit result in the case of an underlying sequence of binomial processes.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

The next result is crucial for the upcoming proofs. It is a direct consequence of Lemma 2.11 in Hiraoka *et al.* [16]. This so-called geometric lemma enables us to obtain upper bounds on moments.

Lemma 2. *(Corollary of [16, Lemma 2.11]) Let $\mathbb{X} \subseteq \mathbb{Y}$ be two finite point sets of $\mathbb{R}^d$. Then*

$$\left|\beta_q^{r,s}(\mathcal{K}(\mathbb{Y})) - \beta_q^{r,s}(\mathcal{K}(\mathbb{X}))\right| \leq \sum_{j=q}^{q+1} |\mathcal{K}_j(\mathbb{Y}, s) \setminus \mathcal{K}_j(\mathbb{X}, s)|.$$

6.1. Stabilization

We start with the proof of the fundamental Theorem 4, the last step of which uses Proposition 3. While this proposition is stated immediately after the proof of Theorem 4, it is of course helpful to first read the proposition before passing to the last step of the proof.

Subsequently, we derive Theorem 1 and Lemma 1. The proof of Theorem 5 is deferred to Appendix A.

Proof of theorem 4. If $q = 0$, $\widetilde{\rho}_0^r$ is clearly finite. So we can assume that $q > 0$, and we have to study chains that prevent $\widetilde{\rho}_q^r(\mathcal{P}(\lambda) \cup Q_1, Q_2)$ from being finite. We can consider $\mathcal{P} = \mathcal{P}(1)$ because we consider a general positive r. Moreover, we can assume that the circumcenter z_{Q_2} of Q_2 coincides with the origin. We write $a^*(r) = L_{Q_2} + \mu(r)$ with L_{Q_2} being the circumradius of Q_2. (Here, $\mu(r) = r$ for the Vietoris–Rips filtration and $\mu(r) = 2r$ for the Čech filtration.) Clearly, as Q_1, Q_2 are finite, $\widetilde{\rho}_q^r$ is finite if the random geometric graph $G(\mathcal{P}, \mu(r))$ does not percolate. If $G(\mathcal{P}, \mu(r))$ percolates, $\widetilde{\rho}_q^r$ is infinite if and only if for each finite information horizon $R \in \mathbb{R}_+$, $R \geq a^*(r) + \mu(r)$, there is a simplex $\sigma_{q,i} \in \mathcal{K}_r((\mathcal{P} \cup Q_1 \cup Q_2)|_{B(0,R)})$ which intersects the additional points Q_2 and which is negative until R, but we cannot exclude the possibility that it might ultimately become positive. Formally, this means there is a chain

$$\tau = \sum_i \sigma_i, \tag{6}$$

where $\sigma_i \in \bigcup_{n\in\mathbb{N}} \mathcal{K}_r((\mathcal{P} \cup Q_1)|_{B(0,n)})$ are q-simplices, such that the boundary of the restriction of τ to $\mathcal{K}_r((\mathcal{P} \cup Q_1)|_{B(0,R)})$ consists of two disjoint $(q-1)$-cycles which are not boundaries. More precisely, set

$$\tau_R := \tau|_{B(0,R)} := \sum_i \sigma_i \mathbb{1}\left\{\sigma_i \in \mathcal{K}_r((\mathcal{P} \cup Q_1)|_{B(0,R)})\right\}. \tag{7}$$

Then, for each $R \geq a^*(r) + \mu(r)$, we have $\partial\tau_R = e_1 + e_{2,R}$, where

$$\begin{aligned} &e_1, e_{2,R} \in Z_{q-1}(\mathcal{K}_r((\mathcal{P} \cup Q_1)|_{B(0,R)})) \setminus B_{q-1}(\mathcal{K}_r((\mathcal{P} \cup Q_1)|_{B(0,R)})) \\ &\quad \text{such that } e_{2,R} \subseteq B(0,R) \setminus B(0, R - 2\mu(r)) \\ &\quad \text{and } e_1 \subseteq B(0, R_0) \text{ for a certain } R_0 \in \mathbb{R}_+; \end{aligned} \tag{8}$$

here the set inclusions for $e_{2,R}$ (resp. e_1) are to be understood as the usual inclusions between subsets of Euclidean space. The cycle e_1 becomes a boundary of $\tau|_{B(0,R)}$ when we include the additional points of Q_2, i.e., $e_1 = \sum_j \nu_j$, where the ν_j are $(q-1)$-simplices in $\mathcal{K}_r((\mathcal{P} \cup Q_1 \cup Q_2)|_{B(0,a^*(r)+\mu(r))})$ and $e_1 \in B_{q-1}(\mathcal{K}_r((\mathcal{P} \cup Q_1 \cup Q_2)|_{B(0,a^*(r)+\mu(r))}))$. Consequently, $e_{2,R}$ becomes a boundary in this case as well, i.e., $e_{2,R} \in B_{q-1}(\mathcal{K}_r((\mathcal{P} \cup Q_1 \cup Q_2)|_{B(0,R)}))$. See also Figure 2 for an illustration in the special case of such a one-dimensional chain.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

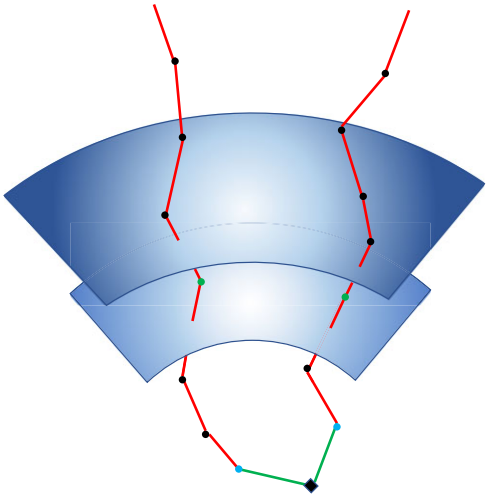


FIGURE 2. Illustration of a chain τ consisting of one-dimensional simplices (red, green) from Poisson points (black, blue, and green dots) and an additional point (black diamond) which is located inside $Q(0)$. The 1-simplices between Poisson points are red; the 1-simplices between a Poisson point and the additional point are green. The layers depict two spheres of $B(z, R)$ and $B(z, R - 2\mu(r))$; e_1 corresponds to the two blue dots (to which the two green 1-simplices are attached), $e_{2,R}$ to the green dots shown between the two layers.

The existence of a τ as above is equivalent to $\widetilde{\rho}_q^r$ being infinite. We show in the remainder of the proof that for a homogeneous Poisson process modulo a finite point process such chains cannot occur. More precisely, we show that the cycle $e_{2,R}$ cannot exist for all $R \geq a^*(r) + \mu(r)$. For this purpose, we can assume that $Q_1 = \emptyset$, because the question of whether such cycles $e_{2,R}$ exist for all $R \geq a^*(r) + \mu(r)$ is an asymptotic property of the Poisson process (this also follows from details given below). Moreover, it suffices to study the case where Q_2 consists of a single point, and by translation-invariance we can assume that this point is the origin, viz., $Q_2 = \{0\}$.

In the remainder of the proof, we study chains which generalize the chain τ given in (6) and show using a Burton–Keane argument that the existence of such chains contradicts the properties of the stationary Poisson process. To this end, we proceed in four steps. We introduce general *maximal* chains in the first step. In the second step, we show that the number of these maximal chains is a.s. constant. In the third and fourth steps, we prove that the number of these maximal chains is a.s. zero.

Step 1—maximal chains. We begin with q-chains τ of the form $\tau = \sum_i \sigma_i$, where the simplices σ_i are in $\bigcup_{n\in\mathbb{N}} \mathcal{K}_r(\mathcal{P}|_{B(0,n)})$ for all i, and there are $y \in \mathbb{R}^d$ ('a shift of the origin') and $R_0 \in \mathbb{R}_+$ such that the following property (P) holds:

(P) For all $R \geq R_0$ and with

$$\tau_{y,R} := \sum_i \sigma_i \mathbb{1}\left\{\sigma_i \in \mathcal{K}_r(\mathcal{P}|_{B(y,R)})\right\},$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

we have that $\partial\tau_{y,R} = e_1 + e_{2,R}$, where

$$e_1 \in Z_{q-1}(\mathcal{K}_r(\mathcal{P}|_{B(y,R)})) \setminus B_{q-1}(\mathcal{K}_r(\mathcal{P}|_{B(y,R)}))$$
$$\text{but } e_1 \in B_{q-1}(\mathcal{K}_r((\mathcal{P}\cup\{y\})|_{B(y,a^*(r)+\mu(r))})) \tag{9}$$

and

$$e_{2,R} \subseteq B(y,R)\setminus B(y,R-2\mu(r))$$
$$\text{and } e_{2,R} \in Z_{q-1}(\mathcal{K}_r(\mathcal{P}|_{B(y,R)})) \setminus B_{q-1}(\mathcal{K}_r(\mathcal{P}|_{B(y,R)})). \tag{10}$$

So, loosely speaking, $\tau_{y,R}$ is a shifted analogue of τ_R from (7). We call a chain τ satisfying (P) a *chain of type (P)*.

Now, we might have two chains τ and τ' of type (P) that have the same local cycle, i.e., their cycles e_1, e'_1 (as characterized by (9)) are homologous (meaning they differ by a boundary only). In this case we say that τ and τ' are equivalent and write $\tau \sim \tau'$; we can then consider the chain τ^* which consists of the union of the simplices of τ and τ'.

This leads to the following notion of maximality: consider an arbitrary but fixed chain τ of type (P) with local cycle e_1. Then take the union over all chains τ' of type (P) equivalent to τ and call the resulting chain the maximal chain $\tau_{\max}$. Formally, given τ with corresponding e_1, the maximal chain is

$$\tau_{max} = \sum_{\sigma\in I}\sigma, \quad \text{where} \quad I = \bigcup_{\tau':\tau'\sim\tau}\ \bigcup_{\sigma\in\tau'}\{\sigma\}.$$

Plainly, if τ and τ' are equivalent, then $\tau_{\max} = {\tau'}_{\max}$.

Step 2—ergodicity and invariance of maximal chains. We show that the number of maximal chains, which we denote by $M(\mathcal{P})$, is a.s. constant. More precisely, we show that

$$M(\mathcal{P}) = m \qquad \text{a.s., for some } m \in \mathbb{N}\cup\{0,+\infty\}.$$

To this end, consider the standard decomposition of the stationary Poisson process $\mathcal{P}$ into a countable sum of independent Poisson processes restricted to the cubes Q_z, viz.,

$$\mathcal{P} = \sum_{z\in\mathbb{Z}^d}\mathcal{P}|_{Q_z} = \sum_{n\in\mathbb{Z}}\left\{\sum_{z'\in\mathbb{Z}^{d-1}}\mathcal{P}|_{Q_{(n,z')}}\right\}. \tag{11}$$

Next, for $m \in \mathbb{N}\cup\{0,+\infty\}$, define the events

$$A_m := \{\omega\in\Omega \mid M(\mathcal{P}(\omega)) = m\}.$$

Consider the shift operator T acting on the first coordinate only. Applying T to a set $P \subseteq \mathbb{R}^d$ gives $T(P) = \{y + (1,0,\dots,0)^t | y \in P\}$. Define the '$k$th translation of A_m' by

$$T^k(A_m) := \{\omega\in\Omega \mid M(T^k(\mathcal{P}(\omega))) = m\}, \quad k\in\mathbb{Z}.$$

Then $M(\mathcal{P}) = M(T(\mathcal{P}))$, and consequently $T(A_m) = A_m$ for each m, i.e., A_m is an invariant event. A standard argument relying on the decomposition (11) into i.i.d. random variables now yields that $\mathbb{P}(A_m) \in \{0,1\}$ for each m. We refer to the book of Klenke [20, Example 20.26] for a formal proof.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Step 3—insertion-tolerance, $M(\mathcal{P}) \notin \mathbb{N}$ a.s. Assume that $M(\mathcal{P}) = m$ with probability 1 for some $m \in \mathbb{N}$. Define the generic 'annulus' (with respect to the maximum norm) $A_{t,s} = [-t, t]^d \setminus (-s, s)^d$ for $s < t$.

We rely on the following decomposition of $\mathcal{P}$: we fix $n \in \mathbb{N}$ large enough—see the paragraph below (12) for details. Denote the restriction of $\mathcal{P}$ to $Q(0, n) = [-n, n]^d$ by $\mathcal{P}_n^\circ$ and its restriction to $\mathbb{R}^d \setminus [-n, n]^d$ by $\mathcal{P}_n^\dagger$. Then $\mathcal{P}_n^\circ$ and $\mathcal{P}_n^\dagger$ are independent. Assume $\mathcal{P}_n^\circ$ is defined on the generic probability space $(\Omega^\circ, \mathcal{A}^\circ, \mathbb{P}^\circ)$ and $\mathcal{P}_n^\dagger$ is defined on $(\Omega^\dagger, \mathcal{A}^\dagger, \mathbb{P}^\dagger)$. Next, for $n \in \mathbb{N}$ and $\varepsilon \in (0, \mu(r))$, consider the event

$$D_{n,\varepsilon} = \Big\{ \text{for each } \tau_{max}\colon \partial(\tau_{max}|_{A_{n+\mu(r),n}}) \text{ contains a cycle } c \text{ inside } A_{n+\mu(r)-\varepsilon,n} \text{ with}$$

$$c \in Z_{q-1}(\mathcal{K}_r(\mathcal{P}|_{Q(0,n+\mu(r))})) \setminus B_{q-1}(\mathcal{K}_r(\mathcal{P}|_{Q(0,n+\mu(r))})) \Big\}. \tag{12}$$

(Here $\tau_{max}|_{A_{n+\mu(r),n}}$ is the restriction of τ_{max} to $A_{n+\mu(r),n}$.) We choose $n \in \mathbb{N}$ sufficiently large so that $D_{n,0}$ occurs with positive probability $\eta > 0$. Next, define the event

$$E_{n,\varepsilon} = \{\omega^\dagger \in \Omega^\dagger \mid \exists\, \widetilde{\omega}^\circ \in \Omega^\circ : (\widetilde{\omega}^\circ, \omega^\dagger) \in D_{n,\varepsilon}\}.$$

Note that $D_{n,\varepsilon}$, and thus also $E_{n,\varepsilon}$, is decreasing in ε, and clearly $\mathbb{P}^\dagger(E_{n,0}) > 0$ as well. Moreover, $E_{n,0} \setminus E_{n,\varepsilon} \downarrow N$ for a $\mathbb{P}^\dagger$-null set N as $\varepsilon \downarrow 0$; hence, we can fix some $\varepsilon^* > 0$ sufficiently small so that $\mathbb{P}^\dagger(E_{n,\varepsilon^*}) > 0$.

For the technical details underlying the remaining arguments in Step 3, which follow, we refer to the proof of Theorem 5(1). First, partition $A_{n,n-\delta}$ with subcubes $(C_i)_{i\in I}$ of edge length δ for $\delta \le \mu(r)(1 - 1/\sqrt{2})/\sqrt{d}$. We can assume that $n/\delta \in \mathbb{N}$. Let G_n denote the event that each C_i contains at least d points of $\mathcal{P}_n^\circ$. Then $\mathbb{P}^\circ(G_n) > 0$, and by construction, E_{n,ε^*} and G_n are independent. Consequently, $\mathbb{P}^\circ \otimes \mathbb{P}^\dagger(E_{n,\varepsilon^*} \cap G_n) = \mathbb{P}^\circ(E_{n,\varepsilon^*})\mathbb{P}^\dagger(G_n) > 0$. However, if both E_{n,ε^*} and G_n occur, there are no maximal chains because each potential feature associated to some τ_{max} already terminates in $A_{n,n-\delta}$. This is because the points of the Poisson process are sufficiently dense inside $A_{n,n-\delta}$. Hence we arrive at a contradiction, and thus $M(\mathcal{P}) \notin \mathbb{N}$ a.s.

Step 4—the Burton–Keane argument, $M(\mathcal{P}) \neq +\infty$ a.s. We begin with general considerations for the volume-boundary argument applied to cycles. The d-dimensional Lebesgue measure of $A_{s+\Delta,s}$ is $2^d\big((s+\Delta)^d - s^d\big) = 2^d \Delta s^{d-1} + o\big(s^{d-1}\big)$ if $s \to \infty$ and if Δ is bounded above by a constant.

Moreover, from the definition of the Čech and Vietoris–Rips filtrations over a point cloud, we have the following: for each $(q-1)$-cycle $\widetilde{e}$ that is contained in the complex corresponding to the filtration parameter r and that is not a boundary (i.e., a non-trivial cycle), there exists a convex set $\mathcal{J}_r$ intersecting the convex hull of $\widetilde{e}$ such that on the one hand $\mathcal{J}_r \cap \mathcal{P} = \emptyset$ and on the other hand the d-dimensional volume of $\mathcal{J}_r$ is positive and bounded away from zero, i.e., $V(\mathcal{J}_r) \ge \delta_r$ for some $\delta_r > 0$ depending on the filtration type but not on the cycle $\widetilde{e}$. Hence, the total number of such cycles that are not boundaries but are located in $A_{s+\Delta,s}$ is of order

$$\frac{2^d \Delta s^{d-1} + o\big(s^{d-1}\big)}{V(\mathcal{J}_r)} = O(s^{d-1}) \tag{13}$$

for $s \to \infty$ and Δ bounded above.

The remainder of this step follows ideas very similar to those in the proof of Theorem 4.6 in Meester and Roy [23], where it is shown that the number of vacant components in the standard Boolean model is a.s. not equal to ∞. To facilitate the comparison, we adopt similar notation.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Also, the remainder of this step crucially relies on Proposition 3 below, which considers the existence of what we call *encounter chains* (see Proposition 3 for their definition).

It follows from Proposition 3 that (under the assumption of infinitely many maximal chains) the event E_m (detailed in the proposition) has positive probability of at least $\eta > 0$, say, for all $m \in \mathbb{N}$ sufficiently large. We assume that $\mathbb{P}(E_m) \geq \eta$ for all m sufficiently large and show that this leads to a contradiction. We can translate E_m by a vector $2mz$ for $z \in \mathbb{Z}^d$ and call this event E_m^{2mz}. Then for each $L \in \mathbb{N}$

$$\begin{aligned} &\mathbb{E}\left[\sum_{z\in\mathbb{Z}^d} \mathbb{1}\left\{E_m^{2mz} \text{ occurs and } Q(2mz, m) \subseteq Q(0, Lm)\right\}\right] \\ &\geq \eta \cdot \#\{z \in \mathbb{Z}^d : Q(2mz, m) \subseteq Q(0, Lm)\} \geq \eta(L-1)^d. \end{aligned} \tag{14}$$

Let $\mathfrak{R}$ be the set of the following encounter configurations in $Q(0,Lm)$: an encounter configuration r lies in $\mathfrak{R}$ if and only if $Q(2mz, m) \subseteq Q(0, Lm)$ such that $r = \tau^*|_{Q(2mz,m)}$ is an encounter configuration for some encounter chain τ^*. Then $\mathbb{E}[\#\mathfrak{R}] \geq \eta(L-1)^d$.

We consider the branch $b = b_r^{(i)}$ for each $r \in \mathfrak{R}$ and $i \in \{1, 2, 3\}$. The boundary ∂b intersected with $A_{Lm+\mu(r),Lm}$ consists of $(q-1)$-cycles e_i that are not boundaries themselves. We can assume that each e_i is minimal in the sense that we cannot decompose the chain e_i into two or more disjoint $(q-1)$-cycles. We define $\mathfrak{V}_b$ as the set that contains all these minimal $(q-1)$-cycles. Furthermore, set $\mathfrak{V} = \bigcup_{r\in\mathfrak{R}} \bigcup_{i=1}^3 \mathfrak{V}_{b_r^{(i)}}$; this is the union of all minimal $(q-1)$-cycles that are not boundaries and that are contained in $A_{Lm+\mu(r),Lm}$. Hence, for a suitable $c \in \mathbb{R}_+$, we have $cL^{d-1} \geq \#\mathfrak{V}$ by (13). Furthermore, for $r \in \mathfrak{R}$ and $i \in \{1, 2, 3\}$, define

$$\mathfrak{C}_r^{(i)} = \{r' \in \mathfrak{R} : r' \subseteq b_r^{(i)}\} \cup \mathfrak{V}_{b_r^{(i)}}$$

as a set of chains. Then clearly $\#\mathfrak{C}_r^{(i)} \geq \#\mathfrak{V}_{b_r^{(i)}} \geq 1 =: \mathfrak{K}$, and the following relation holds for each $r, r' \in \mathfrak{R}$: either $\left[r \cup \bigcup_i \mathfrak{C}_r^{(i)}\right] \cap \left[r' \cup \bigcup_j \mathfrak{C}_{r'}^{(j)}\right] = \emptyset$, or there are i, j which satisfy

$$\mathfrak{C}_r^{(i)} \supseteq \{r'\} \cup \bigcup_{\ell\neq j} \mathfrak{C}_{r'}^{(\ell)} \text{ and } \mathfrak{C}_{r'}^{(j)} \supseteq \{r\} \cup \bigcup_{\ell\neq i} \mathfrak{C}_r^{(\ell)}.$$

Define $\mathfrak{S} := \mathfrak{R} \cup \bigcup_{r\in R} \bigcup_{i=1}^3 \mathfrak{C}_r^{(i)}$. Then the assumptions of Lemma 3.2 in Meester and Roy [23] are satisfied. Consequently, $\#\mathfrak{S} \geq \mathfrak{K}(\#\mathfrak{R} + 2) + \#\mathfrak{R}$. Since $\#\mathfrak{S} = \#\mathfrak{R} + \#\mathfrak{V}$, it follows that $\#\mathfrak{V} \geq \#\mathfrak{R} + 2$. We therefore arrive at the following inequalities:

$$\begin{aligned} cL^{d-1} &\geq \mathbb{E}[\#\mathfrak{V}] \geq \mathbb{E}\left[\#\mathfrak{V}\mathbb{1}\{\mathfrak{R} \neq \emptyset\}\right] \geq \mathbb{E}\left[(\#\mathfrak{R} + 2)\mathbb{1}\{\mathfrak{R} \neq \emptyset\}\right] \\ &= \mathbb{E}\left[\#\mathfrak{R} + 2\mathbb{1}\{\mathfrak{R} \neq \emptyset\}\right] \geq (L-1)^d\eta \end{aligned}$$

for all $L \in \mathbb{N}$. This leads to a contradiction if L is sufficiently large. Therefore, $\eta = 0$ and $M(\mathcal{P})$ equals ∞ with probability 0. This completes the fourth step.

All in all, this contradicts the initial assumption that with positive probability the chain τ from (6) exists and satisfies (**P**). Consequently, $\widetilde{\rho}_q^r$ is a.s. finite. □

To state and prove the upcoming proposition, we rely once more on a suitable decomposition of the homogeneous Poisson process $\mathcal{P}$ with unit intensity. Fix $n \in \mathbb{N}$. We choose a Poisson process $\mathcal{P}_n^\dagger$ on $\mathbb{R}^d \setminus [-n, n]^d$ which is defined on the probability space $(\Omega^\dagger, \mathcal{A}^\dagger, \mathbb{P}^\dagger)$.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

We also choose a complementary Poisson process $\mathcal{P}_n^\circ$ on $Q(0,n)=[-n,n]^d$ which is defined on $(\Omega^\circ, \mathcal{A}^\circ, \mathbb{P}^\circ)$. Obviously, $\mathcal{P}_n^\dagger$ and $\mathcal{P}_n^\circ$ are independent.

Proposition 3. *(Encounters of maximal chains) For $m\in\mathbb{N}$, let E_m denote the event given in the next paragraph. If the numbers of maximal chains is infinite (as is assumed in Step 4 of the proof of Theorem 4), we have that* $\liminf_{m\to\infty}\mathbb{P}(E_m)>0$.

Define $n=m-2\lceil\mu(r)\rceil$ and decompose $\mathcal{P}$ into $\mathcal{P}_n^\circ$ and $\mathcal{P}_n^\dagger$ as above. Set

$E_m:=\Big\{(\omega^\circ,\omega^\dagger)\in\Omega^\circ\times\Omega^\dagger \mid \mathcal{P}_n^\dagger(\omega^\dagger)$ *admits disjoint chains $b^{(1)}, b^{(2)}, b^{(3)}$ which fulfil the following conditions:*

(a) *there is an $\widetilde{\omega}^\circ\in\Omega^\circ$ such that each $b^{(i)}$ can be completed to a maximal chain with elements of $\mathcal{P}_n^\circ(\widetilde{\omega}^\circ)$ and $\mathcal{P}_n^\dagger(\omega^\dagger)$;*

(b) *each $\partial(b^{(i)}|_{A_{m,n}})$ decomposes into a disjoint union of $e_1^{(i)}$ and $e_2^{(i)}$ that are $(q-1)$-cycles but not boundaries in the complex generated by $\mathcal{P}_n^\dagger(\omega^\dagger)$, and $e_1^{(i)}\subseteq A_{n+\mu(r)-\varepsilon,n}$;*

(c) *there exist a chain r in the complex generated by the elements of $\mathcal{P}_n^\circ(\omega^\circ)$ and a chain c in $A_{n+\mu(r),n-\mu(r)}$ generated by the elements of $\mathcal{P}_n^\circ(\omega^\circ)\cup\mathcal{P}_n^\dagger(\omega^\dagger)$ such that $\tau:=r+c+\sum_{i=1}^3 b^{(i)}$ satisfies $\partial(\tau|_{Q(0,m)})=\sum_{i=1}^3 e_2^{(i)}\Big\}$.*

If E_m occurs, we call Q(0,m) an encounter box, Q(0,n) a central box, r an encounter configuration, c an intermediate configuration, $b^{(1)}, b^{(2)}, b^{(3)}$ branches, and τ an encounter chain.

Moreover, we can translate E_m over the vector $y=2mz$ ($y\in\mathbb{Z}^d$) and denote this event by E_m^y. If E_m^y occurs, then Q(y, m) is an encounter box and Q(y, n) a central box.

Proof. We assume the existence of infinitely many maximal chains with probability one, and then we build an event H_m which is contained in E_m. First let F_m be the event that there are (at least) three disjoint maximal chains $\tau^{(1)}, \tau^{(2)}, \tau^{(3)}$ with corresponding $e_1^{(1)}, e_1^{(2)}, e_1^{(3)}$ (as in (9)) which are all located inside $Q(0,n)$, $n=m-2\lceil\mu(r)\rceil$. (If there are more than three such maximal chains, we choose three of them at random.) Then $\liminf_{m\to\infty}\mathbb{P}(F_m)=1>0$ by monotone convergence.

Now, fix $m\in\mathbb{N}$ such that F_m has positive probability. If $(\omega^\circ,\omega^\dagger)\in F_m$, we define the three disjoint branches $b^{(i)}=\tau^{(i)}|_{\mathbb{R}^d\setminus Q(0,n)}$, $1\le i\le 3$. Then, for all $\varepsilon>0$ sufficiently small, the event

$$G_m=\{(\omega^\circ,\omega^\dagger)\in F_m \mid \forall i\colon \partial(b^{(i)}|_{A_{m,n}})=f_1+f_2 \text{ such that } f_1 \text{ and } f_2 \text{ are cycles}$$
$$\text{but not boundaries with respect to } \mathcal{K}_r(\mathcal{P}|_{Q(0,m)\setminus Q(0,n)}) \text{ and } f_1\subseteq A_{n+\mu(r)-\varepsilon,n}\}$$

has positive probability, too.

Now let H_m consist of all $(\omega^\circ,\omega^\dagger)\in\Omega^\circ\times\Omega^\dagger$ such that on the one hand, for this very $\omega^\dagger$, there is an $\widetilde{\omega}^\circ\in\Omega^\circ$ with $(\widetilde{\omega}^\circ,\omega^\dagger)\in G_m$ (for short, '$\omega^\dagger$ enjoys the property P_1') and on the other hand there are the following chains: a chain r, constructed from the elements of $\mathcal{P}_n^\circ(\omega^\circ)$, and a chain c, constructed from the elements of $[\mathcal{P}_n^\circ(\omega^\circ)\cup\mathcal{P}_n^\dagger(\omega^\dagger)]\cap A_{m,n}$ with the property that the chain $\tau=r+c+\sum_{i=1}^3 b^{(i)}$ satisfies $\partial(\tau|_{Q(0,R)})\subseteq A_{R,R-\mu(r)}$ for all $R\ge m$ (for short, 'the pair $(\omega^\circ,\omega^\dagger)$ enjoys the property P_2').

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

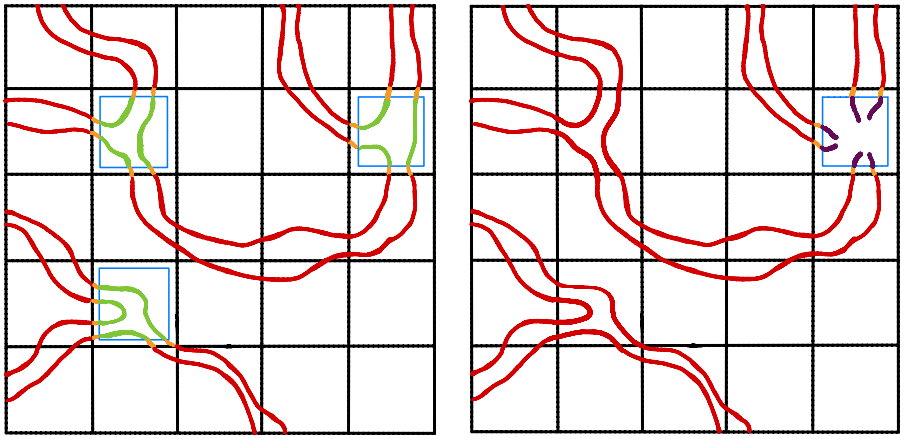


FIGURE 3. Illustration of encounters of maximal chains in two dimensions (in a reduced set-up and not true to scale). The left panel depicts several encounters inside boxes $Q(y, m)$ (for certain $y \in \mathbb{Z}^d$). The (blue) central boxes are located inside the (black) encounter boxes, which are part of the (black) lattice which partitions the plane. Each (green) encounter configuration merges three branches (red and partly in green and orange) through the corresponding (orange) intermediate configuration. The right panel shows a specific central box $Q(y, n)$ (blue). Here a suitable configuration (violet) inside the central box converts the corresponding branches to maximal chains.

Then H_m can be formulated in an abstract way as

$$H_m = \left\{ \left(\omega^\circ, \omega^\dagger\right) \in \Omega^\circ \times \Omega^\dagger : \ \omega^\dagger \text{ has } P_1 \text{ and } (\omega^\circ, \omega^\dagger) \text{ has } P_2 \right\}.$$

We show that H_m has positive probability. Since G_m has positive probability, we have on the one hand

$$\mathbb{P}^\dagger\left(\left\{\omega^\dagger : \omega^\dagger \text{ has } P_1\right\}\right) \geq \mathbb{P}^\circ \otimes \mathbb{P}^\dagger(G_m) > 0$$

and on the other hand

$$\mathbb{P}(H_m) = \int_{\Omega^\dagger} \mathbb{1}\left\{\omega^\dagger \text{ has } P_1\right\} \mathbb{P}^\circ(\{\omega^\circ : (\omega^\circ, \omega^\dagger) \text{ has } P_2\}) \, \mathbb{P}^\dagger(\mathrm{d}\omega^\dagger).$$

Consequently, it remains to show that $\mathbb{P}^\circ(\{\omega^\circ : (\omega^\circ, \omega^\dagger) \text{ has } P_2\}) > 0$ for almost all $\omega^\dagger$ which have P_1.

To this end, we choose an $\omega^\dagger$ which has P_1. Conditionally on $\omega^\dagger$ and relying on classical results for triangulations, there is a finite set P inside $Q(0,n)$ (with all elements in general position) such that r, c, and τ exist as laid out, and all simplices involved in the chain r have a filtration time of at most $\mu(r)/2$. This implies that we can move the vertices of a specific q-simplex σ in $\mathcal{K}_r(P)$ by at most $\mu(r)/(4(q+1))$ and still have a filtration time of σ of at most $3\mu(r)/4$.

Moreover, as $P \cup \mathcal{P}_n^\dagger(\omega^\dagger)|_{Q(0,n+\mu(r))}$ is finite, there is a $\delta > 0$ such that all vertices of P can be moved by at least $\delta > 0$ without adding (resp. removing) another simplex to (resp. in) the complex $\mathcal{K}_r(P \cup \mathcal{P}_n^\dagger(\omega^\dagger)|_{Q(0,n+\mu(r))})$. Note that the probability that the elements of $\mathcal{P}_n^\dagger|_{Q(0,n+\mu(r))}$ do not entail a sharp filtration time of exactly $\mu(r)$ is 1, which is what we implicitly assume in the choice of $\omega^\dagger$. (This effect of local constancy is also studied by Chazal and Divol [8, Lemma 13] for filtration functions generated by (random) point clouds.)

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Thus, for this specific $\delta > 0$, the probability of the following event is positive: there is a realization of $\mathcal{P}_n^\circ$ which has $\#P$ elements and for each $p \in P$ there is an element of $\mathcal{P}_n^\circ$ in the δ-neighborhood of p.

This proves that $\mathbb{P}^\circ(\{\omega^\circ : (\omega^\circ, \omega^\dagger) \text{ has } P_2\}) > 0$ for almost all $\omega^\dagger$ which have P_1. Thus, H_m occurs with positive probability and we arrive at $\mathbb{P}(E_m) \geq \mathbb{P}(H_m) > 0$. This completes the proof. □

Proof of theorem 1. Let $q \in \{0, \ldots, d-1\}$ be fixed. For simplicity, we write $\mathcal{P} = \mathcal{P}(\lambda)$ and $S = S_q^{(r,s)}$. (Recall that S is defined in (5).) S is a.s. finite by Theorem 4. By the definition of $\widetilde{\rho}_q^r$ and $\widetilde{\rho}_{q+1}^s$, the following two functionals do not change when changing the configuration outside $B(0,S)$:

$$A \mapsto \dim \frac{Z_q(\mathcal{K}_r(\mathcal{P} \cup \{0\} \cup A))}{Z_q(\mathcal{K}_r(\mathcal{P} \cup A))} \text{ and } A \mapsto \dim \frac{B_q(\mathcal{K}_s(\mathcal{P} \cup \{0\} \cup A))}{B_q(\mathcal{K}_s(\mathcal{P} \cup A))}, \tag{15}$$

where $A \subseteq \mathbb{R}^d \setminus B(0, S)$ is finite.

In order to conclude the case for the persistent Betti function $\beta_q^{r,s}$, we show that

$$A \mapsto \dim \frac{Z_q(\mathcal{K}_r(\mathcal{P} \cup \{0\} \cup A)) + B_q(\mathcal{K}_s(\mathcal{P} \cup \{0\} \cup A))}{Z_q(\mathcal{K}_r(\mathcal{P} \cup A)) + B_q(\mathcal{K}_s(\mathcal{P} \cup A))}, \tag{16}$$

$A \subseteq \mathbb{R}^d \setminus B(0, S)$ finite, is constant, too. This implies the case for the persistent Betti function by making use of the dimension formula ($\dim(U + V) + \dim(U \cap V) = \dim U + \dim V$ for two finite-dimensional linear spaces U, V) in conjunction with (15).

On the one hand, assume that the map in (16) increases when going from a set A_1 to another set A_2 outside $B(0,S)$. Then there is a q-chain c^* which is not 0 modulo $Z_q(\mathcal{K}_r(\mathcal{P} \cup \{0\} \cup A_2))$ and not 0 modulo $B_q(\mathcal{K}_s(\mathcal{P} \cup \{0\} \cup A_2))$, but is 0 modulo $Z_q(\mathcal{K}_r(\mathcal{P} \cup A_2))$ or 0 modulo $B_q(\mathcal{K}_s(\mathcal{P} \cup A_2))$. This contradicts the conclusion for the mappings in (15).

On the other hand, assume that the map decreases, when going from a set A_1 to A_2. This time there is a q-chain c^* which is 0 modulo $Z_q(\mathcal{K}_r(\mathcal{P} \cup \{0\} \cup A_2))$ or 0 modulo $B_q(\mathcal{K}_s(\mathcal{P} \cup \{0\} \cup A_2))$, but not 0 modulo $Z_q(\mathcal{K}_r(\mathcal{P} \cup A_2))$ and not 0 modulo $B_q(\mathcal{K}_s(\mathcal{P} \cup A_2))$. This is again a contradiction.

Consequently, the persistent Betti number $\beta_q^{r,s}$ does not change with the configuration outside $B(0,S)$. □

Proof of Lemma 1. Clearly, if the right-hand side is infinite, there is nothing to prove. So assume that it is finite. Define $\mathcal{K}_{s,a} = \mathcal{K}_s(P \cap B(z_Q, a))$ and $\mathcal{K}'_{s,a} = \mathcal{K}_s((P \cup Q) \cap B(z_Q, a))$. We show

$$\dim Z_q(\mathcal{K}'_{r,a}) - \dim Z_q(\mathcal{K}_{r,a}) = \text{const.} \quad \text{and} \quad \dim B_q(\mathcal{K}'_{s,a}) - \dim B_q(\mathcal{K}_{s,a}) = \text{const.}$$

for all $a \geq S_q^{(r,s)}(P, Q)$. By definition of $\widetilde{\rho}_q^r(P, Q)$, this is true for differences involving the cycle groups $Z_q(\mathcal{K}'_{r,a})$ and $Z_q(\mathcal{K}_{r,a})$ and $a \geq \widetilde{\rho}_q^r(P, Q)$. Moreover, again by the definition of $S_q^{(r,s)}(P, Q)$, the difference $\dim B_q\mathcal{K}'_{s,a} - \dim B_q\mathcal{K}_{s,a}$ is constant for all $a \geq S_q^{(r,s)}(P, Q)$.

Using the dimension formula $\dim(U + V) + \dim(U \cap V) = \dim U + \dim V$, it only remains to consider the difference $\dim(Z_q(\mathcal{K}'_{r,a}) + B_q(\mathcal{K}'_{s,a})) - \dim(Z_q(\mathcal{K}_{r,a}) + B_q(\mathcal{K}_{s,a}))$.

First, assume that this difference increases at an $a \geq S_q^{(r,s)}(P, Q)$. So, there is a q-chain c^* such that $c^* \neq 0 \bmod B_q(\mathcal{K}'_{s,a-})$ and $c^* \neq 0 \bmod Z_q(\mathcal{K}'_{r,a-})$ but $c^* = 0 \bmod B_q(\mathcal{K}_{s,a-})$ or $c^* = 0 \bmod Z_q(\mathcal{K}_{r,a-})$. This is a contradiction. Second, if the difference decreases at

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

an $a \geq S_q^{(r,s)}(P, Q)$, there is again a q-chain c^* such that $c^* \neq 0 \bmod B_q(\mathcal{K}_{s,a-})$ and $c^* \neq 0 \bmod Z_q(\mathcal{K}_{r,a-})$ but $c^* = 0 \bmod B_q(\mathcal{K}'_{s,a-})$ or $c^* = 0 \bmod Z_q(\mathcal{K}'_{r,a-})$. Again, this is a contradiction. □

6.2. Asymptotic normality for the Poisson process

Recall that $\mathcal{P}$ and $\mathcal{P}'$ denote independent Poisson processes with unit intensity on $\mathbb{R}^d$ and $B_n = [-2^{-1}n^{1/d}, 2^{-1}n^{1/d}]^d$. For $z \in \mathbb{Z}^d$, set $\mathcal{P}''(z) = (\mathcal{P} \setminus Q(z)) \cup (\mathcal{P}' \cap Q(z))$, and let

$$\Delta_z^{r,s}(B_n) = \beta_q^{r,s}(\mathcal{K}(\mathcal{P} \cap B_n)) - \beta_q^{r,s}(\mathcal{K}(\mathcal{P}''(z) \cap B_n)).$$

Lemma 3. *For each $(r, s) \in \Delta$ and each $z \in \mathbb{Z}^d$, there are random variables $\Delta_z^{r,s}(\infty)$ and $N_0 = N_0(z, (r, s)) \in \mathbb{N}$ such that $\Delta_z^{r,s}(B_n) \equiv \Delta_z^{r,s}(\infty)$ a.s. for all $n \geq N_0$.*

Proof. Let $z \in \mathbb{Z}^d$ and $(r, s) \in \Delta$ be arbitrary but fixed. We utilize Lemma 3.1 in [32] to obtain a random variable $\Delta_z^{r,s}(\infty)$ such that

$$\lim_{n\to\infty} \Delta_z^{r,s}(B_n) = \Delta_z^{r,s}(\infty) \text{ with probability 1.}$$

The existence of N_0 follows from the fact that Betti numbers and consequently the $\Delta_z^{r,s}(B_n)$ are integer-valued. □

Proposition 4. *For any two pairs $(r, s), (u, v) \in \Delta$,*

$$\begin{aligned}\gamma((u, v), (r, s)) &:= \mathbb{E}\left[\mathbb{E}\left[\Delta_0^{r,s}(\infty)|\mathcal{F}_0\right]\ \mathbb{E}\left[\Delta_0^{u,v}(\infty)|\mathcal{F}_0\right]\right] \\ &= \lim_{n\to\infty} n^{-1} \operatorname{Cov}\left(\beta_q^{u,v}(\mathcal{K}(\mathcal{P} \cap B_n)), \beta_q^{r,s}(\mathcal{K}(\mathcal{P} \cap B_n))\right).\end{aligned}$$

Proof. Let (r, s) and (u, v) be arbitrary but fixed. Define the two functionals

$$h_1 := \beta_q^{r,s}(\mathcal{K}(\cdot)) \text{ and } h_2 := \beta_q^{u,v}(\mathcal{K}(\cdot)).$$

First, observe that

$$\begin{aligned}&2n^{-1} \operatorname{Cov}(h_1(\mathcal{P} \cap B_n), h_2(\mathcal{P} \cap B_n)) \\ &= n^{-1} \operatorname{Var}[h_1(\mathcal{P} \cap B_n) + h_2(\mathcal{P} \cap B_n)] \\ &\quad - n^{-1} \operatorname{Var}[h_1(\mathcal{P} \cap B_n)] - n^{-1} \operatorname{Var}[h_2(\mathcal{P} \cap B_n)].\end{aligned} \tag{17}$$

In the following we prove the convergence of each of the three terms on the right-hand side to an appropriate limit as $n \to \infty$.

Using the established strong stabilization (Theorem 1), we obtain the convergence of each term in question from [32, Theorem 3.1] once the Poisson bounded moments condition is satisfied, i.e., for $i \in \{1, 2\}$

$$\sup_{A\in\mathcal{B}:0\in A} \mathbb{E}\left[|h_i([\mathcal{P} \cap A] \cup \{0\}) - h_i(\mathcal{P} \cap A)|^4\right] < \infty,$$

where $\mathcal{B} = \{B_n + x : x \in \mathbb{R}^d, n \geq 1\}$. The validity of the Poisson bounded moment condition follows immediately from the geometric lemma. We skip the details and refer to [42] (proof of Lemma 4.1) instead.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Thus, using Theorem 3.1 of [32],

$$\lim_{n\to\infty} n^{-1}\operatorname{Var}[h_1(\mathcal{P}\cap B_n)] = \mathbb{E}\left[\mathbb{E}\left[\Delta_0^{r,s}(\infty)|\mathcal{F}_0\right]^2\right],$$

$$\lim_{n\to\infty} n^{-1}\operatorname{Var}[h_2(\mathcal{P}\cap B_n)] = \mathbb{E}\left[\mathbb{E}\left[\Delta_0^{u,v}(\infty)|\mathcal{F}_0\right]^2\right].$$

Moreover, using the additivity of the add-one cost, the functional h_1+h_2 enjoys the strong stabilization just as h_1 and h_2. Hence, applying Theorem 3.1 of [32] once more, we get

$$\lim_{n\to\infty} n^{-1}\operatorname{Var}[h_1(\mathcal{P}\cap B_n)+h_2(\mathcal{P}\cap B_n)] = \mathbb{E}\left[\mathbb{E}\left[\Delta_0^{r,s}(\infty)+\Delta_0^{u,v}(\infty)|\mathcal{F}_0\right]^2\right].$$

Consequently, by (17),

$$\lim_{n\to\infty} 2n^{-1}\operatorname{Cov}(h_1(\mathcal{P}\cap B_n), h_2(\mathcal{P}\cap B_n)) = 2\mathbb{E}\left[\mathbb{E}\left[\Delta_0^{r,s}(\infty)|\mathcal{F}_0\right]\mathbb{E}\left[\Delta_0^{u,v}(\infty)|\mathcal{F}_0\right]\right]$$

and the claim follows. □

Proof of theorem 2. We show the multivariate asymptotic normality by considering finite linear combinations of persistent Betti numbers. For $a_1,\dots,a_\ell\in\mathbb{R}$ and $n\in\mathbb{N}$, let $H(n^{1/d}\mathcal{P}_n)=\sum_{i=1}^{\ell} a_i\beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathcal{P}_n))$. We will verify that

$$\begin{aligned} &n^{-1}\ \operatorname{Var}\left[H(n^{1/d}\mathcal{P}_n)\right]\to\sigma^2\\ &\text{and } n^{-1/2}\left(H(n^{1/d}\mathcal{P}_n)-\mathbb{E}[H(n^{1/d}\mathcal{P}_n)]\right)\Rightarrow\mathcal{N}(0,\sigma^2) \text{ as } n\to\infty, \end{aligned} \tag{18}$$

where $\sigma^2=\int_{[0,1]^d}\overline{\sigma}^2(\kappa(x))\,\mathrm{d}x$; here $\overline{\sigma}^2(\lambda)$ is the corresponding limiting variance if the Poisson point process $n^{1/d}\mathcal{P}_n$ in the functional H is replaced by a homogeneous Poisson process $\mathcal{P}(\lambda)$ (with intensity λ) restricted to B_n. Now, we have

$$\begin{aligned} \overline{\sigma}^2(\lambda) &= \lim_{n\to\infty} n^{-1}\operatorname{Var}\left[H(\mathcal{P}(\lambda)|_{B_n})\right]\\ &= \sum_{i,j=1}^{\ell} a_i a_j \lim_{n\to\infty} n^{-1}\operatorname{Cov}[\beta_q^{r_i,s_i}(\mathcal{K}(\mathcal{P}(\lambda)|_{B_n})),\ \beta_q^{r_j,s_j}(\mathcal{K}(\mathcal{P}(\lambda)|_{B_n}))]\\ &= \lambda\sum_{i,j=1}^{\ell} a_i a_j \lim_{n\to\infty}(\lambda n)^{-1}\operatorname{Cov}[\beta_q^{\lambda^{1/d}r_i,\lambda^{1/d}s_i}(\mathcal{K}(\mathcal{P}(1)|_{\lambda^{1/d}B_n})),\\ &\qquad\qquad \beta_q^{\lambda^{1/d}r_j,\lambda^{1/d}s_j}(\mathcal{K}(\mathcal{P}(1)|_{\lambda^{1/d}B_n}))]\\ &= \lambda\sum_{i,j=1}^{\ell} a_i a_j \gamma(\lambda^{1/d}(r_i,s_i),\lambda^{1/d}(r_j,s_j)), \end{aligned}$$

where the last equality follows from from Proposition 4, and where γ is the limiting covariance function of an underlying homogeneous Poisson process with unit intensity on $\mathbb{R}^d$. Hence,

$$\sigma^2=\sum_{i,j=1}^{\ell} a_i a_j \int_{[0,1]^d}\gamma(\kappa(x)^{1/d}(r_i,s_i),\kappa(x)^{1/d}(r_j,s_j))\,\kappa(x)\,\mathrm{d}x. \tag{19}$$

Consequently, once the statements in (18) are verified, the proof is complete.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

By Theorem 3.3 of [37], the statements in (18) hold if, apart from the strong stabilization (Theorem 1), the following two moment conditions are satisfied for cubes of the type $W = z + [0, a)^d \subseteq \mathbb{R}^d$ with $z \in \mathbb{R}^d$, $a \in \mathbb{R}_+$, and for each pair (r, s) with $r \le s$:

(1) the Poisson bounded moments condition: for some $p > 2$

$$\sup_n \sup_{y \in \mathbb{R}^d} \sup_{y \in W\colon \text{cube}} \mathbb{E}\left[|\mathfrak{D}_0 \beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{P}_n|_W - y))|^p \right] < \infty;$$

(2) the locally bounded moments condition: for each cube W as above, there is a $p > 2$ such that

$$\sup_n \sup_{y \in \mathbb{R}^d} \mathbb{E}\left[|\beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{P}_n|_{y+W}))|^p \right] < \infty.$$

Both conditions (1) and (2) are immediate consequences of the geometric lemma (Lemma 2). Indeed, to see (1), observe that the geometric lemma gives

$$\begin{aligned} |\mathfrak{D}_0 \beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{P}_n|_W - y))| &\le \sum_{j=q}^{q+1} \mathcal{K}_j((n^{1/d}\mathcal{P}_n|_W) \cup \{y\}, s) \setminus \mathcal{K}_j((n^{1/d}\mathcal{P}_n|_W), s) \\ &\le 2|n^{1/d}\mathcal{P}_n \cap B(y, \mu(s))|^{q+1}. \end{aligned}$$

This last upper bound is stochastically dominated by $|\mathcal{P}(\|\kappa\|_\infty) \cap B(0, \mu(s))|^{q+1}$, which does not depend on n, y, or W, and the moment $\mathbb{E}\left[|\mathcal{P}(\|\kappa\|_\infty) \cap B(0, \mu(s))|^{p(q+1)}\right]$ is finite for each $p, q \ge 0$.

Regarding (2), we have again by Lemma 2 and the translation-invariance that $\beta_q^{r,s}(\mathcal{K}(n^{1/d}\mathcal{P}_n|_{y+W}))$ is stochastically dominated by $2|\mathcal{P}(\|\kappa\|_\infty)|_W|^{q+2}$, which does not depend on y or n. Moreover, for each cube W the moment $\mathbb{E}\left[|\mathcal{P}(\|\kappa\|_\infty)|_W|^{p(q+2)}\right]$ is finite for all $p, q \ge 0$. This completes the proof. □

6.3. Asymptotic normality for the binomial process

For each $n \in \mathbb{N}$, let $(U_{m,n}\colon m \in \mathbb{N})$ be a sequence of binomial processes such that $U_{m,n} = (Y_{1,n}, \dots, Y_{m,n})$ for i.i.d. sequences $(Y_{i,n}\colon i \in \mathbb{N})$ with common marginal density κ.

Proof of theorem 3. Let $\ell \in \mathbb{N}$, $a_1, \dots, a_\ell \in \mathbb{R}$. We apply the abstract Theorem 3.9 in Trinh [37] to the functional $H(n^{1/d}\mathbb{X}_n) = \sum_{i=1}^{\ell} a_i \beta_q^{r_i,s_i}(\mathcal{K}(n^{1/d}\mathbb{X}_n))$. Since we established the Poisson bounded moments condition and the locally bounded moments condition in the proof of Theorem 2, and since the qth persistent Betti number obtained from a finite set is polynomially bounded in the cardinality of this set, it is enough to verify for the add-one cost function that

$$\sup_{n \in \mathbb{N}} \sup_{m \in [n(1-\eta), n(1+\eta)]} \mathbb{E}\left[|\beta_q^{r,s}(\mathcal{K}(n^{1/d}U_{m+1,n})) - \beta_q^{r,s}(\mathcal{K}(n^{1/d}U_{m,n}))|^4 \right] < \infty$$

for some $\eta > 0$. This can be verified using the geometric lemma (Lemma 2), similarly to Lemma 4.1 in Yogeshwaran *et al.* [42]; we omit the details.

Hence the conditions of Theorem 3.9 in Trinh [37] are satisfied, and

$$\frac{\mathrm{Var}\left[H(n^{1/d}\mathbb{X}_n)\right]}{n} \to \tau^2, \qquad \frac{H(n^{1/d}\mathbb{X}_n) - \mathbb{E}\left[H(n^{1/d}\mathbb{X}_n)\right]}{\sqrt{n}} \Rightarrow N(0, \tau^2),$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

where τ^2 is given by the following relation: $\tau^2 = \sigma^2 - \left(\int_{[0,1]^d} \mathbb{E}\left[\overline{\Delta}(\kappa(x))\right] \kappa(x)\, \mathrm{d}x \right)^2$ for σ^2 given in (19) and

$$\begin{aligned}
\mathbb{E}\left[\overline{\Delta}(\lambda)\right] &= \sum_{i=1}^{\ell} a_i\, \mathbb{E}\left[\lim_{n\to\infty} \mathfrak{D}_0 \beta_q^{r_i,s_i}(\mathcal{K}(\mathcal{P}(\lambda)|_{B_n}))\right] \\
&= \sum_{i=1}^{\ell} a_i\, \mathbb{E}\left[\mathfrak{D}_0 \beta_q^{r_i,s_i}(\mathcal{K}(\mathcal{P}(\lambda) \cap B(0, S_q^{(r_i,s_i)}(\lambda))))\right] \\
&= \sum_{i=1}^{\ell} a_i\, \mathbb{E}\left[\mathfrak{D}_0 \beta_q^{\lambda^{1/d} r_i, \lambda^{1/d} s_i}(\mathcal{K}(\mathcal{P}(1) \cap B(0, S_q^{(\lambda^{1/d} r_i, \lambda^{1/d} s_i)}(1))))\right] \\
&= \sum_{i=1}^{\ell} a_i\, \alpha(\lambda^{1/d}(r_i, s_i)).
\end{aligned}$$

Thus, for a random variable X distributed with density κ, we have

$$\begin{aligned}
\tau^2 = \sum_{i,j=1}^{\ell} a_i a_j \Big\{ &\mathbb{E}\left[\gamma(\kappa(X)^{1/d}(r_i, s_i), \kappa(X)^{1/d}(r_j, s_j))\right] \\
&- \mathbb{E}\left[\alpha(\kappa(X)^{1/d}(r_i, s_i))\right] \mathbb{E}\left[\alpha(\kappa(X)^{1/d}(r_j, s_j))\right] \Big\}.
\end{aligned}$$

This completes the proof. □

Appendix A

First we give the proof of Theorem 5.

Proof of theorem 5. Recall that $\mu(r)$ is an upper bound on the diameter of a simplex with filtration time at most $r \geq 0$. Note that the statements for the radius of strong stabilization $\widetilde{\rho}_q^r$, together with the relation

$$\rho^{(r,s)}(P, Q) \leq S_q^{(r,s)}(P, Q) \qquad \text{for each } (r, s) \in \Delta,\, q \in \{0, \ldots, d-1\}$$

from Lemma 1, allow us to conclude the results for the radius of weak stabilization. So it remains to prove the statement for the strong stabilization property. We proceed for each q separately; clearly this is no restriction.

In the following, if Q_1, Q_2, r, and q are fixed, we just write $\widetilde{\rho}(\kappa)$ for $\widetilde{\rho}_q^r(\mathcal{P}(\kappa) \cup Q_1, Q_2)$ if the Poisson process $\mathcal{P}(\kappa)$ has intensity κ.

In the remainder of the proof, we assume without loss of generality that the Poisson process $\mathcal{P}(\kappa)$ on $\mathbb{R}^d$ is coupled to a homogeneous Poisson process $\mathcal{P}$ on $\mathbb{R}^{d+1}$ with intensity 1 via a space–time coupling as follows:

$$\mathcal{P}(\kappa) = \{x |\, \exists\, t : 0 \leq t \leq \kappa(x),\, (x, t) \in \mathcal{P}\}. \tag{20}$$

Proof of (1): Let $r, \varepsilon > 0$ and $q \in \{0, \ldots, d-1\}$ be arbitrary but fixed. We first consider the case of adding exactly one additional point to the Poisson process, so $Q_1 = \emptyset$ and $Q_2 = \{0\}$.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

We show that there is an $L > 0$ such that $\mathbb{P}(\widetilde{\rho}(\lambda) > L) \leq 2\varepsilon$ for each $\lambda \in \mathbb{R}_+$. The generalization then works along the same lines.

We rely on the following ideas: if the intensity λ is sufficiently small, there are no points of the Poisson process in the neighborhood of the additional point 0 with high probability; hence, adding 0 does not create a new cycle. Moreover, if the intensity λ is sufficiently large, then an annulus around the origin is covered by the points of the Poisson process with high probability; hence the impact of adding 0 is only local. Finally, we rely on the space–time coupling for the intermediate intensities.

First, there is a $\underline{\kappa} \in \mathbb{R}_+$ such that

$$\mathbb{P}(|\mathcal{P}(\lambda) \cap B(0, \mu(r))| = 0) \geq 1 - \varepsilon \text{ for all } \lambda \leq \underline{\kappa}.$$

This means that with high probability and for all $q \geq 1$, we have, for all λ below this threshold, that including $\{0\}$ does not create an additional q-simplex. Hence, if $\lambda \leq \underline{\kappa}$, then $\mathbb{P}(\widetilde{\rho}(\lambda) > L) \leq \varepsilon$ for $L \geq \mu(r)$.

Also, there is an intensity $\overline{\kappa}$ such that with high probability all changes in the add-one cost function which are caused by including the origin are limited to a deterministic neighborhood; we carry out the argument simultaneously for the Čech and Vietoris–Rips complexes. Indeed, let $\delta \leq \mu(r)(1 - 1/\sqrt{2})/\sqrt{d}$ be sufficiently small so that $4\mu(r)/\delta \in \mathbb{N}$, and consider a partition of $A := A_{2\mu(r)+\delta,2\mu(r)} = [-2\mu(r) - \delta, 2\mu(r) + \delta]^d \setminus (-2\mu(r), 2\mu(r))^d$ with subcubes $(C_i)_{i\in I}$ of edge length δ. We study $\mathcal{U} = \bigcap_{i\in I}\{\#(\mathcal{P}(\lambda) \cap C_i) \geq d\}$ ('each subcube contains at least d Poisson points'). Then, there is a $\overline{\kappa} \in \mathbb{R}_+$ depending on r and δ such that

$$\mathbb{P}(\mathcal{U}) = \prod_{i\in I} \mathbb{P}(\#(\mathcal{P}(\lambda) \cap C_i) \geq d) \geq 1 - \varepsilon \text{ for all } \lambda \geq \overline{\kappa}.$$

Define $L_0 := (2\mu(r) + \delta)\sqrt{d}$. We show $\{\widetilde{\rho}(\lambda) \leq L_0\} \supseteq \mathcal{U}$, which in turn implies $\mathbb{P}(\widetilde{\rho}(\lambda) > L) \leq \varepsilon$ for $L \geq L_0$ if $\lambda \geq \overline{\kappa}$.

Assume $\mathcal{U}$. Let $\sigma^r_{q,1}, \ldots, \sigma^r_{q,m_q}$ be the q-simplices which contain the origin. Clearly these simplices are all contained in $[-\mu(r), \mu(r)]^d$. In particular, they do not intersect A, which itself is homeomorphic to a $(d-1)$-cycle.

Plainly, we can triangulate A with $(d-1)$-simplices with filtration time at most $2\delta\sqrt{d} < \mu(r)$. Indeed, consider two adjacent cubes C_1, C_2 each containing a $(d-1)$-simplex, say $\sigma_1 = \{x_0, \ldots, x_{d-1}\}$, $\sigma_2 = \{y_0, \ldots, y_{d-1}\}$. (Here adjacent means that the closures of the cubes have nonempty intersection.) Then we can connect σ_1 and σ_2 with the $(d-1)$-simplices $\{x_0, \ldots, x_{i-1}, y_i, \ldots, y_{d-1}\}$ $(1 \leq i \leq d-1)$, and each of these simplices has a filtration time of at most $2\delta\sqrt{d}$.

Moreover, let $\sigma = \{x_0, \ldots, x_{k-1}\}$ be a $(k-1)$-simplex in $A_{2\mu(r)+\mu(r)/2,2\mu(r)-\mu(r)/2}$ for a generic $k \in \mathbb{N}$. Then there is a $(k-1)$-simplex $\sigma^* = \{y_0, \ldots, y_{k-1}\}$ in A such that the Euclidean distance between any two elements of σ and σ^* is at most $\sqrt{(\mu(r)/2)^2 + (\mu(r)/2)^2} + \sqrt{d}\delta$, which in turn is at most $\mu(r)$. Consequently, all k-simplices of the type $\{x_0, \ldots, x_i, y_i, \ldots, y_{k-1}\}$ have a filtration time at most $\mu(r)$.

This shows that, conditional on $\mathcal{U}$, any cycle in $A_{2\mu(r)+\mu(r)/2,2\mu(r)-\mu(r)/2}$ is equivalent to a cycle in $A_{2\mu(r)+\delta,2\mu(r)}$ and thus is a boundary as well. Hence, $\widetilde{\rho}(\lambda) \leq L_0$ in this case. In particular, we have $\mathbb{P}(\widetilde{\rho}(\lambda) > L) \leq \varepsilon$ for all $\lambda \geq \overline{\kappa}$ and all $L \geq L_0 = \sqrt{d}(2r + \delta)$.

It remains to check intensities $\lambda \in [\underline{\kappa}, \overline{\kappa}]$. Assume there is an $\varepsilon > 0$ such that

$$\limsup_{L\to\infty} \sup_{\lambda\in[\underline{\kappa},\overline{\kappa}]} \mathbb{P}(\widetilde{\rho}(\lambda) > L) > 2\varepsilon.$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Then we can find sequences $(L_n)_n$ and $(\lambda_n)_n$ such that $L_n \to \infty$ and $\lambda_n \to \lambda^* \in [\underline{\kappa}, \overline{\kappa}]$ with the property that $\mathbb{P}(\widetilde{\rho}(\lambda_n) > L_n) > \varepsilon$ for all $n \in \mathbb{N}$. However, there is an $L^* \in \mathbb{R}_+$ such that $\mathbb{P}(\widetilde{\rho}(\lambda^*) > L^*) < \varepsilon/4$ as $\widetilde{\rho}$ is a.s. finite by Theorem 4. Also, from the coupling of the Poisson processes (20), and because $\mathcal{P}(\lambda)$ is a simple point process, there are random $\overline{\delta}, \underline{\delta} > 0$ (depending on the choice of L^*) such that $\mathcal{P}(\lambda)$ does not contain any points in $B(0, L^* + \mu(r)) \times [\lambda^* - \underline{\delta}, \lambda^* + \overline{\delta}]$, viz.,

$$\underline{\delta} = \big(\lambda^* - \inf\{\lambda < \lambda^* : \mathcal{P}(\lambda)(B(0, L^* + \mu(r)) \times [\lambda, \lambda^*]) = 0\}\big)/2$$

and $\overline{\delta} = \big(\sup\{\lambda > \lambda^* : \mathcal{P}(\lambda)(B(0, L^* + \mu(r)) \times [\lambda^*, \lambda]) = 0\} - \lambda^*\big)/2$. This means, for all $\lambda \in [\lambda^* - \underline{\delta}, \lambda^* + \overline{\delta}]$,

$$\mathcal{P}(\underline{\lambda})|_{B(0,L^*+\mu(r))} \equiv \mathcal{P}(\lambda)|_{B(0,L^*+\mu(r))} \equiv \mathcal{P}(\overline{\lambda})|_{B(0,L^*+\mu(r))}.$$

Note that $\{\lambda^* - \underline{\lambda} \geq \delta'\} \cap \{\overline{\lambda} - \lambda^* \geq \delta'\} \supseteq \{\mathcal{P}(B(0, L^* + \mu(r)) \times [\lambda^* - 2\delta', \lambda^* + 2\delta']) = 0\}$ for $\delta' > 0$. Consequently, there is an $\delta' \in \mathbb{R}_+$ such that the event $\{\lambda^* - \underline{\lambda} > \delta'\} \cap \{\overline{\lambda} - \lambda^* > \delta'\}$ has probability at least $1 - \varepsilon/4$. Consequently, for all n large enough so that $L_n \geq L^*$ and $|\lambda_n - \lambda^*| \leq \delta'$, we have

$$\begin{aligned}\mathbb{P}(\widetilde{\rho}(\lambda_n) > L_n) &\leq \mathbb{P}\big(\widetilde{\rho}(\lambda_n) > L_n, \widetilde{\rho}(\lambda^*) \leq L^*, \lambda^* - \underline{\lambda} > \delta', \overline{\lambda} - \lambda^* > \delta'\big) \\ &\quad + \mathbb{P}\big(\widetilde{\rho}(\lambda^*) > L^*\big) + \mathbb{P}\big(\{\lambda^* - \underline{\lambda} \leq \delta'\} \cup \{\overline{\lambda} - \lambda^* \leq \delta'\}\big) \leq \varepsilon/2,\end{aligned}$$

because $\{\widetilde{\rho}(\lambda_n) > L_n, \widetilde{\rho}(\lambda^*) \leq L^*, \lambda^* - \underline{\lambda} > \delta', \overline{\lambda} - \lambda^* > \delta'\} = \emptyset$. This contradicts the fact that $\mathbb{P}(\widetilde{\rho}(\lambda_n) > L_n) > \varepsilon$ for all $n \in \mathbb{N}$. Thus, the laws of $\{\widetilde{\rho}(\mathcal{P}(\lambda), \{0\})) : \lambda \in \mathbb{R}_+\}$ are tight.

A similar argument shows also that the laws of $\{\widetilde{\rho}(\mathcal{P}(\lambda) \cup Q_1, Q_2), \lambda \in \mathbb{R}_+, Q_1, Q_2 \in \mathfrak{Q}_m\}$ are tight. At this point, it is essential that we have a uniform upper bound on the parameter $a^*(r)$, as follows:

$$a^*(r) = L_{Q_2} + \mu(r) \leq \sqrt{d} + \mu(\overline{r}) \tag{21}$$

for each $r \leq \overline{r}$ and each $Q_2 \subseteq Q(0)$. So, instead of computing the radius of strong stabilization by taking the infimum in (4) over all $\{R : R \geq a^*(r)\}$, we can first take the infimum over the smaller set $\{R : R \geq \sqrt{d} + \mu(\overline{r})\}$, which does not depend on Q_2 and r, in order to obtain a modified radius of strong stabilization, which is not smaller than the original one in (4). Verifying the claim for this modified radius then implies the claim for the original radius. For the rest of the proof of (1) we argue with this modified radius, which we also denote by $\widetilde{\rho}_q^r$, abusing the notation slightly.

We now sketch the remaining steps. Using the same techniques, we easily see that there are upper and lower bounds $\underline{\kappa}$ and $\overline{\kappa}$ such that intensities $\lambda \notin [\underline{\kappa}, \overline{\kappa}]$ have only a local effect, in the same sense as in the special case for $\{0\}$; i.e., for each $\varepsilon > 0$ there are $\underline{\kappa}, \overline{\kappa}$ and an $L > 0$ such that

$$\sup_{Q_1, Q_2 \in \mathfrak{Q}_m} \sup_{\lambda \notin [\underline{\kappa}, \overline{\kappa}]} \mathbb{P}(\widetilde{\rho}(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L) \leq \varepsilon.$$

For intensities $\lambda \in [\underline{\kappa}, \overline{\kappa}]$, we can repeat the argument from the special case treated above. Indeed, assume the contrary, namely,

$$\limsup_{L \to \infty} \sup_{Q_1, Q_2 \in \mathfrak{Q}_m} \sup_{\lambda \in [\underline{\kappa}, \overline{\kappa}]} \mathbb{P}(\widetilde{\rho}(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L) > 2\varepsilon,$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

for some $\varepsilon > 0$. Then there are sequences $(Q_{n,1})_n$, $(Q_{n,2})_n$, $(\lambda_n)_n$, $(L_n)_n$ with the properties $Q_{n,1} \to Q_1^*$, $Q_{n,2} \to Q_2^*$ (considered as vectors whose entries are elements in $[-2^{-1}, 2^{-1}]^d$) for two admissible elements $Q_1^*, Q_2^* \in \mathfrak{Q}_m$, as well as $\lambda_n \to \lambda^* \in [\underline{\kappa}, \overline{\kappa}]$ and $L_n \to \infty$. These sequences satisfy

$$\mathbb{P}(\widetilde{\rho}(\mathcal{P}(\lambda_n) \cup Q_{n,1}, Q_{n,2}) > L_n) > \varepsilon \qquad \text{for all } n.$$

Now we can argue as before in the special case to obtain a contradiction. We arrive at the following result: for all $\varepsilon > 0$, $m \in \mathbb{N}$, and $r \in \mathbb{R}_+$, there is an $L > 0$ such that

$$\max_{q \in \{0,\dots,d-1\}} \sup_{\lambda \in \mathbb{R}_+} \sup_{Q_1, Q_2 \in \mathfrak{Q}_m} \mathbb{P}(\widetilde{\rho}_q^r(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L) \le \varepsilon.$$

So far, we have been considering a fixed $r \le \overline{r}$. We now prove the general statement given in (1). To this end, we rely once more on (21) and the induced modification of $\widetilde{\rho}_q^r$. Then, for $r \le \overline{r}$ and $\alpha \in \mathbb{R}_+$ arbitrary but fixed,

$$\begin{aligned} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L) &= \mathbb{P}(\widetilde{\rho}_q^r(\alpha^{-1}\mathcal{P}(\lambda) \cup \alpha^{-1}Q_1, \alpha^{-1}Q_2) > \alpha^{-1}L) \\ &= \mathbb{P}(\widetilde{\rho}_q^r(\mathcal{P}(\alpha^d\lambda) \cup \alpha^{-1}Q_1, \alpha^{-1}Q_2) > \alpha^{-1}L), \end{aligned} \tag{22}$$

using the scale-invariance for the first equation and $\mathcal{L}(\alpha^{-1}\mathcal{P}(\lambda)) = \mathcal{L}(\mathcal{P}(\alpha^d\lambda))$ for the second equation. Consequently, for all $\overline{\alpha}$, $\underline{\alpha}$, $\varepsilon \in \mathbb{R}_+$ arbitrary but fixed, with $\underline{\alpha} \le \overline{\alpha}$, there is an $L \in \mathbb{R}_+$ such that

$$\sup_{\alpha \in [\underline{\alpha}, \overline{\alpha}]} \sup_{\lambda \in \mathbb{R}_+} \sup_{Q_1, Q_2 \in \mathfrak{Q}_m} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L) \le \varepsilon. \tag{23}$$

This completes the considerations of Part (1).

Proof of (2): We use a suitable space–time coupling of Poisson processes. Let $r \in \mathbb{R}_+$ and $q \in \{0, \dots, d-1\}$ be arbitrary but fixed. Write $\widetilde{\rho}$ for $\widetilde{\rho}_q^r$. Note that the law of $n^{1/d}\mathcal{P}(n\nu)$ equals the law of $\mathcal{P}(\nu(\,\cdot\,/n^{1/d}))$.

First we show that for all $\varepsilon > 0$, there exist a $b > 0$ and an $L > 0$ such that

$$\sup_{n \in \mathbb{N}} \sup_{z \in B''_{n,L}} \mathbb{P}(\widetilde{\rho}(\mathcal{P}(\nu(\,\cdot\,/n^{1/d}) \cup Q_1, Q_2) > L) \le \varepsilon$$

for all densities ν satisfying $\|\nu - \kappa\|_\infty \le b$, and for all $Q_1, Q_2 \in z + \mathfrak{Q}_m$, where $m \in \mathbb{N}$ is arbitrary but fixed.

From Part (1) of the theorem, for each $\varepsilon > 0$ there is an $L > 0$ such that, for the homogeneous Poisson process $\mathcal{P}(\nu(z/n^{1/d}))$, we have that

$$\mathbb{P}(\widetilde{\rho}(\mathcal{P}(\nu(z/n^{1/d})) \cup Q_1, Q_2) > L) \le \varepsilon$$

uniformly in $z \in [0, n^{1/d}]^d$, $Q_1, Q_2 \in z + \mathfrak{Q}_m$, and $n \in \mathbb{N}$, and for all densities ν. Moreover, with $K > 0$, define the set

$$A_{\nu,n}(K, z) = \Big\{ x \in B(z, K) \;\Big|\; \exists t \in \Big[\nu\Big(\frac{x}{n^{1/d}}\Big) \wedge \nu\Big(\frac{z}{n^{1/d}}\Big),\, \nu\Big(\frac{x}{n^{1/d}}\Big) \vee \nu\Big(\frac{z}{n^{1/d}}\Big) \Big] \text{ and } (x, t) \in \mathcal{P} \Big\}.$$

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

By assumption, the function κ is uniformly continuous with a certain modulus of continuity $\omega(\cdot)$. Hence, $|\kappa(x_1) - \kappa(x_2)| \le \omega(\delta)$ whenever $\|x_1 - x_2\| \le \delta$ and $\omega(\delta) \to 0$ as $\delta \to 0$. Let ν be a density satisfying $\|\nu - \kappa\|_\infty \le b$. Then

$$\begin{aligned}
&\mathbb{P}(A_{\nu,n}(K, z) \neq \emptyset) \\
&\le \mathbb{P}\Big(\exists x \in B(z, K) \;\Big|\; \exists t \in \Big[\kappa\Big(\frac{x}{n^{1/d}}\Big) \wedge \kappa\Big(\frac{z}{n^{1/d}}\Big) - b, \kappa\Big(\frac{x}{n^{1/d}}\Big) \vee \kappa\Big(\frac{z}{n^{1/d}}\Big) + b\Big] \\
&\qquad\qquad \text{and } (x, t) \in \mathcal{P}\Big) \\
&\le \mathbb{P}\Big(\exists x \in B(z, K): \exists t \in \Big[0, \omega(K/n^{1/d}) + 2b\Big] \text{ and } (x, t) \in \mathcal{P}\Big), \qquad (24)
\end{aligned}$$

where the last inequality uses the stationarity of $\mathcal{P}$. Clearly, given a value for K, there are $b > 0$ and $n_0 \in \mathbb{N}$ such that (24) is small uniformly in $z \in B''_{n,K}$, ν in a b-neighborhood of κ, and $n \ge n_0$.

We now come to the conclusion. Let $\varepsilon > 0$ be arbitrary but fixed. We apply the result from Part (1) and choose $L^* \in \mathbb{R}_+$ such that $\mathbb{P}(\widetilde{\rho}(\mathcal{P}(\lambda) \cup Q_1, Q_2) > L^*) \le \varepsilon/2$ is satisfied uniformly in $\lambda \in \mathbb{R}_+$, $Q_1, Q_2 \in z + \mathfrak{Q}_m$, and $z \in \mathbb{R}^d$.

Next, let $K^* = L^* + 2\mu(r)$. Choose $b > 0$ and $n_0 \in \mathbb{N}$ such that $\mathbb{P}(A_{\nu,n}(K^*, z) \neq \emptyset) \le \varepsilon/2$ for all $n \ge n_0$, for all $z \in B''_{n,K^*}$, and for all ν such that $\|\nu - \kappa\|_\infty \le b$. Since by assumption $z \in B''_{n,K^*}$, this implies

$$\begin{aligned}
&\Big\{\widetilde{\rho}(\mathcal{P}(\nu(\cdot/n^{1/d})) \cup Q_1, Q_2) > K^*, \\
&\qquad \widetilde{\rho}(\mathcal{P}(\nu(z/n^{1/d})) \cup Q_1, Q_2) \le L^*, A_{\nu,n}(K^*, z) = \emptyset\Big\} = \emptyset. \qquad (25)
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\mathbb{P}(\widetilde{\rho}(\mathcal{P}(\nu(\cdot/n^{1/d})) \cup Q_1, Q_2) > K^*) \\
&\le \mathbb{P}\big(\widetilde{\rho}(\mathcal{P}(\nu(\cdot/n^{1/d})) \cup Q_1, Q_2) > K^*, \qquad (26) \\
&\qquad \widetilde{\rho}(\mathcal{P}(\nu(z/n^{1/d})) \cup Q_1, Q_2) \le L^*, A_{\nu,n}(K^*, z) = \emptyset\big) + \varepsilon = \varepsilon,
\end{aligned}$$

for all $z \in B''_{n,K^*}$, for $n \ge n_0$, and for all ν such that $\sup|\nu - \kappa| \le b$.

The generalization to an entire parameter range for the filtration parameter now follows from the result in (23) (note that $\overline{\lambda} = \sup\kappa$ is an admissible choice in this equation) and by using a similar ansatz as in the derivation of (26). So, for each $\varepsilon > 0$, $\overline{\alpha} \ge \underline{\alpha} > 0$, there are $L \in \mathbb{R}_+$, $b > 0$, and $n_0 \in \mathbb{N}$ such that for each $0 \le q \le d - 1$

$$\sup_{\underline{\alpha} \le \alpha \le \overline{\alpha}} \sup_{n \ge n_0} \sup_{Q_1, Q_2 \in \mathfrak{Q}_m} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\nu(\cdot/n^{1/d})) \cup Q_1, Q_2) > L) \le \varepsilon, \qquad (27)$$

for all densities ν in a b-neighborhood of κ with respect to the sup-norm. This yields the first result given in Part (2). The second result is now immediate: there is an $m \in \mathbb{N}$ such that with high probability, the number of Poisson points inside $Q(z)$ is at most m. This means we can apply the previous result.

Proof of (3): This time we couple the binomial process to a suitable Poisson process. Again let $0 \le q \le d - 1$ and r be arbitrary but fixed. We first study the radius $\widetilde{\rho}_q^r(n^{1/d}\mathbb{X}_m, n^{1/d}X')$.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

First, note that for all $\varepsilon > 0$ and for all $L > 0$ there is an $n_0 \in \mathbb{N}$ such that $\mathbb{P}(n^{1/d}X' \notin B''_{n,L}) \leq \overline{\kappa} n^{(d-1)/d} L n^{-1} \leq \varepsilon$ for all $n \geq n_0$. So, by independence, for each $s > 0$,

$$\mathbb{P}(\widetilde{\rho}_q^s(n^{1/d}\mathbb{X}_m, n^{1/d}X') > L) \leq \sup_{z \in B''_{n,L}} \mathbb{P}\big(\widetilde{\rho}_q^s(n^{1/d}\mathbb{X}_m, \{z\}) > L\big) + \varepsilon, \quad \forall n \geq n_0.$$

For each $n \in \mathbb{N}$, let $V_{1,n}, V_{2,n}, \ldots$ be i.i.d. with density $\kappa(\cdot/n^{1/d})$. Let $\mathcal{P}(\kappa(\cdot/n^{1/d})) = \{Z_{1,n}, \ldots, Z_{N_n,n}\}$ be the Poisson process from (20) for the intensity function $\kappa(\cdot/n^{1/d})$ and $N_n \sim \mathrm{Poi}(n)$. Then $n^{1/d}\mathbb{X}_m$ has the same distribution as the process

$$U_{m,n} = \Big[\mathcal{P}(\kappa(\cdot/n^{1/d})) \setminus \{Z_{m+1,n}, \ldots, Z_{N_n,n}\}\Big] \cup \{V_{1,n}, \ldots, V_{m-N_n,n}\},$$

where, by convention, $\{Z_{m+1,n}, \ldots, Z_{N_n,n}\}$ is empty if $N_n \leq m$, and $\{V_{1,n}, \ldots, V_{m-N_n,n}\}$ is empty if $N_n \geq m$.

By (27), for each $\overline{\alpha} \geq \underline{\alpha} > 0$ and for each $\varepsilon > 0$, there is an $L > 0$ such that

$$\sup_{\underline{\alpha} \leq \alpha \leq \overline{\alpha}} \sup_{n \in \mathbb{N}} \sup_{z \in B''_{n,L}} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\kappa(\cdot/n^{1/d})), \{z\}) > L) \leq \varepsilon.$$

Note that here we use the fact that we are studying only one density function, namely, κ, so we choose $L > 0$ individually for values of n, $n \leq n_0$, and take the maximum at the end.

Also, for all $\varepsilon > 0$, for all $z \in B_n$, and for all $K > 0$, there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\mathbb{P}(A'_n(K, z)) \geq 1 - \varepsilon$, where

$$A'_n(K, z) = \{[\{Z_{m+1,n}, \ldots, Z_{N_n,n}\} \cup \{V_{1,n}, \ldots, V_{m-N_n,n}\}] \cap B(z, K) = \emptyset\}.$$

Indeed, this result follows from standard calculations, as

$$\mathbb{E}\left[|m - N_n|\right] \leq |m - n| + \mathbb{E}\left[|N_n - n|^2\right]^{1/2} \leq h(n) + n^{1/2}$$

and as the probability that a single point falls in $B(z, K)$ is bounded above by a constant times n^{-1}.

In the last step, we combine these observations as follows. First, let $\varepsilon > 0$ be arbitrary but fixed. Then there is an $L^* > 0$ such that

$$\mathbb{P}(n^{1/d}X' \notin B''_{n,L^*}) \leq \frac{\varepsilon}{3}$$

and

$$\sup_{\underline{\alpha} \leq \alpha \leq \overline{\alpha}} \sup_{n \in \mathbb{N}} \sup_{z \in B''_{n,L^*}} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\kappa(\cdot/n^{1/d})), \{z\}) > L^*) \leq \frac{\varepsilon}{3}.$$

Moreover, with $K^* = L^* + 2\mu(r)$, there is an $n_0 \in \mathbb{N}$ such that for $z \in B_n$ and $n \geq n_0$, $\mathbb{P}(A'_n(K^*, z)^c) \leq \varepsilon/3$. Consequently, similarly to (25), if $n \geq n_0$ and if $L \geq L^*$, then

$$\begin{aligned}
&\mathbb{P}(\widetilde{\rho}_q^{\alpha r}(n^{1/d}\mathbb{X}_m, n^{1/d}X') > L) \\
&\leq \sup_{z \in B''_{n,L^*}} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(U_{m,n}, \{z\}) > L) + \frac{\varepsilon}{3} \\
&\leq \sup_{z \in B''_{n,L^*}} \mathbb{P}(\widetilde{\rho}_q^{\alpha r}(U_{m,n}, \{z\}) > L, A'_n(K^*, z), \widetilde{\rho}_q^{\alpha r}(\mathcal{P}(\kappa(\cdot/n^{1/d})), \{z\}) \leq L^*) + \varepsilon = \varepsilon,
\end{aligned}$$

uniformly in $m \in J_n$ and $\underline{\alpha} \leq \alpha \leq \overline{\alpha}$. This shows (3) and completes the proof. □

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Throughout the remainder of the appendix we consider a more general filtration as in [16]. Examples of these filtrations are the Čech and the Vietoris–Rips filtrations. The following principle will be important.

Lemma 4. *Let $\mathcal{K}$ and $\mathcal{K}'$ be two simplicial complexes. Then $C_q(\mathcal{K}) \cap C_q(\mathcal{K}') = C_q(\mathcal{K} \cap \mathcal{K}')$. Moreover, $Z_q(\mathcal{K}) \cap Z_q(\mathcal{K}') = Z_q(\mathcal{K} \cap \mathcal{K}')$ and $B_q(\mathcal{K}) \cap B_q(\mathcal{K}') \supseteq B_q(\mathcal{K} \cap \mathcal{K}')$.*

Proof. First, we consider the claim concerning the spaces C_q. The inclusion $\supseteq$ is clear, so we prove only $\subseteq$. This inclusion can be deduced from the fact that C_q is a free module over $\mathbb{F}_2$ generated by the corresponding q-simplices in the filtration. We can write $c \in C_q(\mathcal{K}) \cap C_q(\mathcal{K}')$ as $\sum_i a_i \sigma_i$, where σ_i are q-simplices in $\mathcal{K}$, $a_i \in \mathbb{F}_2$, and also as $\sum_j b_j \widetilde{\sigma}_j$, where $\widetilde{\sigma}_j$ are q-simplices in $\mathcal{K}'$, $b_j \in \mathbb{F}_2$. Hence, $\sum_i a_i \sigma_i - \sum_j b_j \widetilde{\sigma}_j = 0$. If $\sigma_i \in \mathcal{K} \setminus \mathcal{K}'$, the coefficient a_i is zero, as this basis element cannot occur in the filtration $\mathcal{K}'$. The same holds in the other direction, if $\widetilde{\sigma}_j \in \mathcal{K}' \setminus \mathcal{K}$, b_j is zero.

The statement $Z_q(\mathcal{K}) \cap Z_q(\mathcal{K}') = Z_q(\mathcal{K} \cap \mathcal{K}')$ follows immediately. Again the inclusion $\supseteq$ is clear and we prove only $\subseteq$. If $c \in Z_q(\mathcal{K}) \cap Z_q(\mathcal{K}')$, then by the above $c \in C_q(\mathcal{K} \cap \mathcal{K}')$ and by assumption $\partial c = 0$. Thus, $c \in Z_q(\mathcal{K} \cap \mathcal{K}')$ as desired. The inclusion concerning the boundary groups is immediate. □

We remark that $B_q(\mathcal{K}) \cap B_q(\mathcal{K}') \not\subseteq B_q(\mathcal{K} \cap \mathcal{K}')$ is possible. For instance, consider a situation where $\beta_0(\mathcal{K}) = \beta_0(\mathcal{K}') = 1$ and where $\mathcal{K} \cap \mathcal{K}' = \{a, b\}$ for two zero-dimensional simplices a, b such that $\mathcal{K} \cap \mathcal{K}'$ is a strict subset of $\mathcal{K}$ and $\mathcal{K}'$. (For example, we can take two 'arc-like' connected components represented by $\mathcal{K}$ and $\mathcal{K}'$ which intersect only at their endpoints.) Then we have $B_0(\mathcal{K} \cap \mathcal{K}') = \{0\}$, but $B_0(\mathcal{K}) \cap B_0(\mathcal{K}')$ contains $a + b$ as a nontrivial element.

In the following, assume that P is a simple point cloud on $\mathbb{R}^d$ without accumulation points and Q is a finite subset of $\mathbb{R}^d$ such that $P \cap Q = \emptyset$ and $Q \subseteq Q(z, L)$ for some $z \in \mathbb{R}^d$ and $L \in \mathbb{R}_+$. Define

$$\mathcal{K}_{s,a} = \mathcal{K}_s(P \cap B(z, a)), \qquad \mathcal{K}'_{s,a} = \mathcal{K}_s((P \cup Q) \cap B(z, a)).$$

Set $a^* = a^*(s) = \mu(s) + L$, where $\mu(s)$ is the upper bound on the diameter of a simplex in the filtration at time s, which is guaranteed by the assumptions of [16] on the filtration. Choose $a_1, a_2 \in \mathbb{R}$ with $a^* \le a_1 \le a_2$ such that $C_0(\mathcal{K}_{s,a_2} \setminus \mathcal{K}_{s,a_1})$ contains exactly one additional basis element (a point) from P, and write

$$C_{q+1}(\mathcal{K}'_{s,a_2} \setminus \mathcal{K}'_{s,a_1}) = \langle \sigma_1, \ldots, \sigma_n \rangle.$$

We can assume without loss of generality that the simplices are already in the right order, i.e.,

$$B_q(\mathcal{K}'_{s,a_2}) = B_q(\mathcal{K}'_{s,a_1}) \oplus \langle \partial\sigma_1, \ldots, \partial\sigma_i \rangle, \tag{28}$$

so that $\partial\sigma_j \neq 0 \mod B_q(\mathcal{K}'_{s,a_1}) \oplus \langle \partial\sigma_1, \ldots, \partial\sigma_{j-1} \rangle$ for $j = 1, \ldots, i$ and $\partial\sigma_j = 0 \mod B_q(\mathcal{K}'_{s,a_1}) \oplus \langle \partial\sigma_1, \ldots, \partial\sigma_i \rangle$ for $j = i+1, \ldots, n$. As a_1 is sufficiently large, we have that each of the simplices σ_j is also contained in $\mathcal{K}_{s,a_2}$. Hence, as $B_q(\mathcal{K}_{s,a})$ is a subspace of $B_q(\mathcal{K}'_{s,a})$, we have that

$$B_q(\mathcal{K}_{s,a_2}) = B_q(\mathcal{K}_{s,a_1}) \oplus \langle \partial\sigma_1, \ldots, \partial\sigma_i \rangle \oplus \langle \partial\sigma_j \colon \text{for } j \in J \rangle, \tag{29}$$

where $J \subseteq \{i+1, \ldots, n\}$ can be empty. In particular,

$$\dim B_q(\mathcal{K}'_{s,a_2}) - \dim B_q(\mathcal{K}_{s,a_2}) = \dim B_q(\mathcal{K}'_{s,a_1}) - \dim B_q(\mathcal{K}_{s,a_1}) + (i - i - \#J).$$

That is, the map $a \mapsto \dim B_q(\mathcal{K}'_{s,a})/B_q(\mathcal{K}_{s,a})$ is non-increasing if $a \ge a^*$.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

Finiteness of the radius of weak stabilization $\rho^{(r,s)}(P, Q)$ follows directly from Lemma 1 and Theorem 4. This argument relies on the finiteness of the radius of strong stabilization (Theorem 4), which, however, is not necessary for the finiteness of $\rho^{(r,s)}(P, Q)$. The following lemma gives a direct proof of the finiteness of $\rho^{(r,s)}(P, Q)$, using ideas similar to those in the proof of Lemma 5.3 of Hiraoka *et al.* [16].

Lemma 5. *The radius $\rho^{(r,s)}(P, Q)$ from (3) is well-defined.*

Proof. It is sufficient to consider $\rho_q^{(r,s)}(P, Q)$ for each $0 \le q \le d-1$. One can use the geometric lemma to show that the nonnegative, integer-valued mappings

$$a \mapsto \dim \frac{Z_q(\mathcal{K}'_{r,a})}{Z_q(\mathcal{K}_{r,a})} \text{ and } a \mapsto \dim \frac{Z_q(\mathcal{K}'_{r,a}) \cap B_q(\mathcal{K}'_{s,a})}{Z_q(\mathcal{K}_{r,a}) \cap B_q(\mathcal{K}_{s,a})} \tag{30}$$

are bounded above. Moreover, the map $Z_q(\mathcal{K}'_{r,a_1})/Z_q(\mathcal{K}_{r,a_1}) \hookrightarrow Z_q(\mathcal{K}'_{r,a_2})/Z_q(\mathcal{K}_{r,a_2})$ is injective for all $0 \le a_1 \le a_2$; this follows from Lemma 4. Thus, there is an $a_1^* \in \mathbb{R}_+$ such that the first mapping in (30), which is integer-valued, is constant for all $a \ge a_1^*$.

Next we show that the second mapping in (30) also becomes constant as $a \to \infty$. To this end, we first show that $a \mapsto \dim B_q(\mathcal{K}'_{s,a})/B_q(\mathcal{K}_{s,a})$ is constant for all $a \ge a_2^*$ for a certain $a_2^* \in \mathbb{R}_+$. This follows from the non-increasing property (see the paragraph right after (29)) and the boundedness from below of this mapping. Now, returning to the second mapping in (30), one can argue as in the proof of Lemma 1 to show that the difference $\dim Z_q(\mathcal{K}'_{r,a}) \cap B_q(\mathcal{K}'_{s,a}) - \dim Z_q(\mathcal{K}_{r,a}) \cap B_q(\mathcal{K}_{s,a})$ is constant for all $a \ge a_1^* \vee a_2^*$. □

Acknowledgements

The authors are very grateful to the associate editor and two referees for their careful reading and their suggestions, which greatly improved the manuscript. Moreover, the authors thank Khanh Duy Trinh for his interest in this paper and for pointing out a mistake in a preliminary version.

Funding information

J. Krebs was supported by the German Research Foundation (DFG), grant numbers KR-4977/1-1 and KR-4977/2-1. W. Polonik was partially supported by the National Science Foundation (NSF), grant number DMS-2015575.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Aizenman, M., Kesten, H. and Newman, C. (1987). Uniqueness of the infinite cluster and related results in percolation. In *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, Springer, New York, pp. 13–20.

[2] Aizenman, M., Kesten, H. and Newman, C. M. (1987). Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Commun. Math. Phys.* **111**, 505–531.

[3] Baryshnikov, Y. and Yukich, J. E. (2005). Gaussian limits for random measures in geometric probability. *Ann. Appl. Prob.* **15**, 213–253.

[4] Baszczyszyn, B., Yogeshwaran, D. and Yukich, J. E. (2019). Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Prob.* 47, 835–895.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

[5] BOBROWSKI, O. AND KAHLE, M. (2018). Topology of random geometric complexes: a survey. *J. Appl. Comput. Topology* **1**, 331–364.
[6] BURTON, R. M. AND KEANE, M. (1989). Density and uniqueness in percolation. *Commun. Math. Phys.* **121**, 501–505.
[7] CARLSSON, G. (2009). Topology and data. *Bull. Amer. Math. Soc.* **46**, 255–308.
[8] CHAZAL, F. AND DIVOL, V. (2019). The density of expected persistence diagrams and its kernel based estimation. *J. Comput. Geom.* **10**, 127–153.
[9] CHAZAL, F. AND MICHEL, B. (2021). An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers Artificial Intellig.* **4**, 1–28.
[10] DIVOL, V. AND POLONIK, W. (2019). On the choice of weight functions for linear representations of persistence diagrams. *J. Appl. Comput. Topology* **3**, 249–283.
[11] EDELSBRUNNER, H., LETSCHER, D. AND ZOMORODIAN, A. (2000). Topological persistence and simplification. In *Proc. 41st Annual Symposium on Foundations of Computer Science*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 454–463.
[12] GIDEA, M. (2017). Topological data analysis of critical transitions in financial networks. In *3rd International Winter School and Conference on Network Science (NetSci-X 2017)*, Springer, Cham, pp. 47–59.
[13] GIDEA, M. *et al.*et al. (2020). Topological recognition of critical transitions in time series of cryptocurrencies. *Physica A* 548, article no. 123843.
[14] GIDEA, M. AND KATZ, Y. (2018). Topological data analysis of financial time series: landscapes of crashes. *Physica A* **491**, 820–834.
[15] GOEL, A., TRINH, K. D. AND TSUNODA, K. (2019). Strong law of large numbers for Betti numbers in the thermodynamic regime. *J. Statist. Phys.* **174**, 865–892.
[16] HIRAOKA, Y., SHIRAI, T. AND TRINH, K. D. (2018). Limit theorems for persistence diagrams. *Ann. Appl. Prob.* **28**, 2740–2780.
[17] KAHLE, M. (2011). Random geometric complexes. *Discrete Comput. Geom.* **45**, 553–573.
[18] KAHLE, M. AND MECKES, E. (2013). Limit theorems for Betti numbers of random simplicial complexes. *Homology Homotopy Appl.* **15**, 343–374.
[19] KESTEN, H. AND LEE, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Prob.* **6**, 495–527.
[20] KLENKE, A. (2013). *Probability Theory: A Comprehensive Course*. Springer, Cham.
[21] LEE, Y. *et al.*et al. (2017). Quantifying similarity of pore-geometry in nanoporous materials. *Nature Commun.* 8, article no. 15396.
[22] MCGIVNEY, K. AND YUKICH, J. (1999). Asymptotics for Voronoi tessellations on random samples. *Stoch. Process. Appl.* **83**, 273–288.
[23] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.
[24] NAKAMURA, T. *et al.*et al. (2015). Persistent homology and many-body atomic structure for medium-range order in the glass. *Nanotechnology* 26, article no. 304001.
[25] OUDOT, S. Y. (2015). *Persistence Theory: From Quiver Representations to Data Analysis*. American Mathematical Society, Providence, RI.
[26] OWADA, T. (2018). Limit theorems for Betti numbers of extreme sample clouds with application to persistence barcodes. *Ann. Appl. Prob.* **28**, 2814–2854.
[27] OWADA, T. AND ADLER, R. J. (2017). Limit theorems for point processes under geometric constraints (and topological crackle). *Ann. Prob.* **45**, 2004–2055.
[28] OWADA, T. AND THOMAS, A. (2020). Limit theorems for process-level Betti numbers for sparse and critical regimes. *Adv. Appl. Prob.* **52**, 1–31.
[29] PENROSE, M. (2003). *Random Geometric Graphs*. Oxford University Press.
[30] PENROSE, M. D. (2005). Multivariate spatial central limit theorems with applications to percolation and spatial graphs. *Ann. Prob.* **33**, 1945–1991.
[31] PENROSE, M. D. (2007). Gaussian limits for random geometric measures. *Electron. J. Prob.* **12**, 989–1035.
[32] PENROSE, M. D. AND YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Prob.* **11**, 1005–1041.
[33] PENROSE, M. D. AND YUKICH, J. E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Prob.* **13**, 277–303.
[34] SARKAR, A. (1997). Co-existence of the occupied and vacant phase in boolean models in three or more dimensions. *Adv. Appl. Prob.* **29**, 878–889.
[35] SEVERSKY, L. M., DAVIS, S. AND BERGER, M. (2016). On time-series topological data analysis: new data and opportunities. In *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 59–67.
[36] STEELE, J. M. (1988). Growth rates of Euclidean minimal spanning trees with power weighted edges. *Ann. Prob.* **16**, 1767–1787.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press

[37] TRINH, K. D. (2019). On central limit theorems in stochastic geometry for add-one cost stabilizing functionals. *Electron. Commun. Prob.* 24, article no. 76.
[38] UMEDA, Y. (2017). Time series classification via topological data analysis. *Inf. Media Technol.* **12**, 228–239.
[39] WASSERMAN, L. (2018). Topological data analysis. *Ann. Rev. Statist. Appl.* **5**, 501–532.
[40] YAO, Y. *et al.*et al. (2009). Topological methods for exploring low-density states in biomolecular folding pathways. J. Chem. Phys. 130, 04B614.
[41] YOGESHWARAN, D. AND ADLER, R. J. (2015). On the topology of random complexes built over stationary point processes. *Ann. Appl. Prob.* **25**, 3338–3380.
[42] YOGESHWARAN, D., SUBAG, E. AND ADLER, R. J. (2017). Random geometric complexes in the thermodynamic regime. *Prob. Theory Relat. Fields* **167**, 107–142.
[43] YUKICH, J. (2000). Asymptotics for weighted minimal spanning trees on random points. *Stoch. Process. Appl.* **85**, 123–138.
[44] ZOMORODIAN, A. AND CARLSSON, G. (2005). Computing persistent homology. *Discrete Comput. Geom.* **33**, 249–274.

https://doi.org/10.1017/apr.2024.61 Published online by Cambridge University Press