The Annals of Applied Probability
2018, Vol. 28, No. 5, 2740–2780
https://doi.org/10.1214/17-AAP1371

LIMIT THEOREMS FOR PERSISTENCE DIAGRAMS[1]

BY YASUAKI HIRAOKA*, TOMOYUKI SHIRAI†,2 AND KHANH DUY TRINH†,3

Tohoku University and Kyushu University†*

The persistent homology of a stationary point process on $\mathbf{R}^N$ is studied in this paper. As a generalization of continuum percolation theory, we study higher dimensional topological features of the point process such as loops, cavities, etc. in a multiscale way. The key ingredient is the persistence diagram, which is an expression of the persistent homology. We prove the strong law of large numbers for persistence diagrams as the window size tends to infinity and give a sufficient condition for the support of the limiting persistence diagram to coincide with the geometrically realizable region. We also discuss a central limit theorem for persistent Betti numbers.

1. Introduction.

1.1. *Background.* The prototype of this work dates back to the random geometric graphs. In those original settings, a set V of points is randomly scattered in a space according to some probability distribution, and a graph with the vertices V is constructed by assigning edges whose distances are less than a certain threshold value $r \geq 0$. Then some characteristic features in the graph such as connected components and loops are broadly and thoroughly studied (see, e.g., [30]). Furthermore, the random geometric graphs provide mathematical models for applications such as mobile wireless networks [25, 27], epidemics [34], and so on.

Recently, the concept of random topology has emerged and rapidly grown as a higher dimensional generalization of random graphs [3, 23]. One of the simple models studied in random topology is a simplicial complex, which is given by a collection of subsets closed under inclusion. Obviously, a graph is regarded as a one-dimensional simplicial complex consisting of singletons as vertices and doubletons as edges.

In geometric settings, a simplicial complex is built over randomly distributed points in a space by a certain rule respecting the nearness of multiple points, like random geometric graphs. Two standard simplicial complex models constructed

---

Received December 2016; revised August 2017.
[1]Supported in part by JST CREST Mathematics (15656429).
[2]Supported in part by JSPS Grant-in-Aid (JP26610025, JP26287019, JP16H06338).
[3]Supported in part by JSPS Grant-in-Aid for Young Scientists (B) (JP16K17616).
MSC2010 subject classifications. Primary 60K35, 60B10; secondary 55N20.
Key words and phrases. Point process, persistence diagram, persistent Betti number, random topology.

from the points are Čech complexes and Rips complexes, which are also determined by a threshold value r measuring the nearness of points. Then, in such an extended geometric object, it is natural to study higher dimensional topological features such as cavities (2 dim.) and more general q-dimensional holes, beyond connected components (0 dim.) and loops (1 dim.).

In algebraic topology, q-dimensional holes are usually characterized by using the so-called homology. Here, the qth homology of a simplicial complex is given by a vector space and its dimension is called the Betti number which counts the number of q-dimensional holes. Hence, in the setting of random simplicial complexes, the Betti numbers become random variables through a random point configuration, and studying the asymptotic behaviors of the randomized Betti numbers is a significant problem for understanding global topological structures embedded in the random simplicial complexes (e.g., [22, 29, 37–39]).

On the other hand, another type of generalizations has been recently attracting much attention in applied topology. In that setting, we are interested in how persistent the holes are for changing the threshold parameter $r \in \mathbf{R}$. Namely, we deal with one parameter filtration of simplicial complexes obtained by increasing the parameter r and characterize robust or noisy holes in that filtration. The persistent homology [10, 40] is a tool invented for this purpose, and especially, its expression called persistence diagram is now applied to a wide variety of applied areas (see, e.g., [4, 11, 16, 28, 35]). From this point of view, there have been some works on a functional of persistence diagram, called lifetime sum or total persistence, for random complexes (that are not geometric in the sense above) such as Linial–Meshulam processes and random cubical complexes (e.g., [17–19]).

Therefore, it is natural to further extend the results on random geometric simplicial complexes to this generality, and the purpose of this paper is to show several of these extensions. In particular, we are interested in asymptotic behaviors of persistence diagrams themselves defined on stationary point processes. These subjects are mathematically meaningful in their own right, but are also interesting for practical applications. For example, the paper [16] studies topological and geometric structures of atomic configurations in glass materials by comparing persistence diagrams with those of disordered states. By regarding atomic configurations in disordered states as random point processes, further understanding of those persistence diagrams will be useful for characterizing geometry and topology of glass materials, which is one of the important research topics in current physics.

1.2. *Prior work.* Let Φ be a stationary point process on $\mathbf{R}^N$ with all finite moments, that is,

$$\mathbb{E}[\Phi(A)^k] < \infty \qquad \text{for all bounded Borel sets } A \text{ and any } k = 1, 2, \ldots. \tag{1.1}$$

Here, $\Phi(A)$ denotes the number of points in A. For simplicity, we always assume that Φ is *simple*, that is,

$$\mathbb{P}\big(\Phi(\{x\}) \le 1 \text{ for every } x \in \mathbf{R}^N\big) = 1.$$

We denote by Φ_{Λ_L} the restriction of Φ on $\Lambda_L = [-\frac{L}{2}, \frac{L}{2})^N$.

Let $C(\Phi_{\Lambda_L}, r)$ be the Čech complex built over the points Φ_{Λ_L} with parameter $r > 0$ (see Section 2.1 for the definition). The 0th Betti number $\beta_0(C(\Phi_{\Lambda_L}, r))$ for Poisson point processes, which is closely related to the binomial processes, has been studied in an extensive literature (cf. [30]) from various points of view such as the geometric percolation theory and computational geometry. Recently, the limiting behaviors of higher Betti numbers $\beta_q(C(\Phi_{\Lambda_L}, r))$ $(q = 1, 2, \dots, N-1)$ over general stationary point processes have also been widely investigated [38, 39]. Among them, we here restate the most related results.

THEOREM 1.1 ([39], Lemma 3.3 and Theorem 3.5). *Assume that Φ is a stationary point process on $\mathbf{R}^N$ having all finite moments. Then, for each $0 \le q \le N-1$, there exists a constant $\hat{\beta}_q^r \ge 0$ such that*

$$\frac{\mathbb{E}[\beta_q(C(\Phi_{\Lambda_L}, r))]}{L^N} \to \hat{\beta}_q^r \qquad \text{as } L \to \infty.$$

In addition, if Φ is ergodic, then

$$\frac{\beta_q(C(\Phi_{\Lambda_L}, r))}{L^N} \to \hat{\beta}_q^r \qquad \text{almost surely as } L \to \infty.$$

THEOREM 1.2 ([39], Theorem 4.7). *Assume that Φ is a homogeneous Poisson point process on $\mathbf{R}^N$ with unit intensity. Then, for each $0 \le q \le N-1$, there exists a constant $\sigma_r^2 > 0$ such that*

$$\frac{\beta_q(C(\Phi_{\Lambda_L}, r)) - \mathbb{E}[\beta_q(C(\Phi_{\Lambda_L}, r))]}{L^{N/2}} \xrightarrow{d} \mathcal{N}(0, \sigma_r^2) \qquad \text{as } L \to \infty.$$

Here, $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2, and $\xrightarrow{d}$ denotes the convergence in distribution of random variables.

The purpose of this paper is to extend Theorem 1.1 to the setting on persistence diagrams and Theorem 1.2 to persistent Betti numbers.

1.3. *Main results*. In this paper, we study the following simplicial complex model for the point process Φ which is a generalization of the Čech complex and the Rips complex.

Let $\mathscr{F}(\mathbf{R}^N)$ be the collection of all finite (nonempty) subsets in $\mathbf{R}^N$. We can identify $\mathscr{F}(\mathbf{R}^N)$ with the set $\bigsqcup_{k=1}^{\infty} (\mathbf{R}^N)^k / \sim$, where $\sim$ is the equivalence relation induced by permutations of coordinates. For a function f on $\mathscr{F}(\mathbf{R}^N)$, there exists a permutation invariant function $\tilde{f}_k$ on $(\mathbf{R}^N)^k$ for each $k \geq 1$ such that $f(\{x_1, \dots, x_k\}) = \tilde{f}_k(x_1, \dots, x_k)$. We say that f is measurable if so is $\tilde{f}_k$ on $(\mathbf{R}^N)^k$ for each $k \geq 1$.

Let $\kappa : \mathscr{F}(\mathbf{R}^N) \to [0, \infty]$ be a measurable function satisfying:

(K1) $0 \leq \kappa(\sigma) \leq \kappa(\tau)$, if σ is a subset of τ;

(K2) κ is translation invariant, that is, $\kappa(\sigma + x) = \kappa(\sigma)$ for any $x \in \mathbf{R}^N$, where $\sigma + x := \{y + x : y \in \sigma\}$;

(K3) there is an increasing function ρ: $[0, \infty] \to [0, \infty]$ with $\rho(t) < \infty$ for $t < \infty$ such that

$$\|x - y\| \leq \rho\big(\kappa\big(\{x, y\}\big)\big),$$

where $\|x\|$ denotes the Euclidean norm in $\mathbf{R}^N$.

Without loss of generality, we can assume $\kappa(\{x\}) = 0$ because of the translation invariance.

Given such a function κ, we construct a filtration $\mathbb{K}(\Xi) = \{K(\Xi, t) : 0 \leq t < \infty\}$ of simplicial complexes from a finite point configuration $\Xi \in \mathscr{F}(\mathbf{R}^N)$ by

$$K(\Xi, t) = \big\{\sigma \subset \Xi : \kappa(\sigma) \leq t\big\}, \tag{1.2}$$

that is, $\kappa(\sigma)$ is the birth time of a simplex σ in the filtration $\mathbb{K}(\Xi)$. Although we do not explicitly show the dependence on κ in the notation $\mathbb{K}(\Xi)$ because the function κ is fixed in the paper, we here call it the *κ-filtration* over Ξ.

EXAMPLE 1.3. Two important examples of κ which we have in mind are

$$\kappa_C\big(\{x_0, x_1, \dots, x_q\}\big) = \inf_{w \in \mathbf{R}^N} \max_{0 \leq i \leq q} \|x_i - w\|, \tag{1.3}$$

$$\kappa_R\big(\{x_0, x_1, \dots, x_q\}\big) = \max_{0 \leq i < j \leq q} \frac{\|x_i - x_j\|}{2}, \tag{1.4}$$

which define the Čech filtration $\mathbb{C}(\Phi) = \{C(\Phi, t)\}_{t \geq 0}$ and the Rips filtration $\mathbb{R}(\Phi) = \{R(\Phi, t)\}_{t \geq 0}$, respectively. Both κ's satisfy Assumption (K3) with $\rho(t) = 2t$. See also Section 2.1 for these filtrations.

For Theorem 1.9 below, we also remark that both κ_C and κ_R are 1-Lipshitz continuous on $\mathscr{F}(\mathbf{R}^N)$ with respect to the Hausdorff distance d_H. See Appendix C for the definition of d_H.

For the filtration $\mathbb{K}(\Xi)$, we denote its qth persistence diagram by

$$D_q(\Xi) = \big\{(b_i, d_i) \in \Delta : i = 1, \dots, n_q\big\},$$

which is given by a multiset on $\Delta = \{(x, y) \in \overline{\mathbf{R}}^2 : 0 \leq x < y \leq \infty\}$ determined from the unique decomposition of the persistent homology (see (2.2) for the definition). The pair (b_i, d_i) indicates the persistence of the ith homology class, that is, it appears at b_i and disappears at d_i, and $d_i = \infty$ means that the ith homology class persists forever.

In this paper, we deal with the persistence diagram $D_q(\Xi)$ as the counting measure

$$\xi_q(\Xi) = \sum_{(b_i,d_i)\in D_q(\Xi)} \delta_{(b_i,d_i)},$$

rather than as a multiset, where $\delta_{(x,y)}$ is the Dirac measure at $(x, y) \in \overline{\mathbf{R}}^2$.

For each $L > 0$, we define a random filtration built over the points Φ_{Λ_L} and denote it by $\mathbb{K}(\Phi_{\Lambda_L}) = \{K(\Phi_{\Lambda_L}, t)\}_{t\geq 0}$. We write $\xi_{q,L}$ for the point process $\xi_q(\Phi_{\Lambda_L})$ and $\mathbb{E}[\xi_{q,L}]$ for its mean measure (see Section 3 for the precise definition of mean measure).

EXAMPLE 1.4. The top three panels in Figure 1 show point processes with negative (Ginibre), zero (Poisson) and positive (Poisson cluster) correlations, respectively (see [2] for more examples and correlation properties of point processes including the above). All point processes consist of 1,000,000 points with the density $1/2\pi$, and only restricted areas of them are visualized. The bottom shows the corresponding normalized persistence diagrams $\xi_{1,L}/L^2$ of the Čech filtrations applied to the above, respectively.

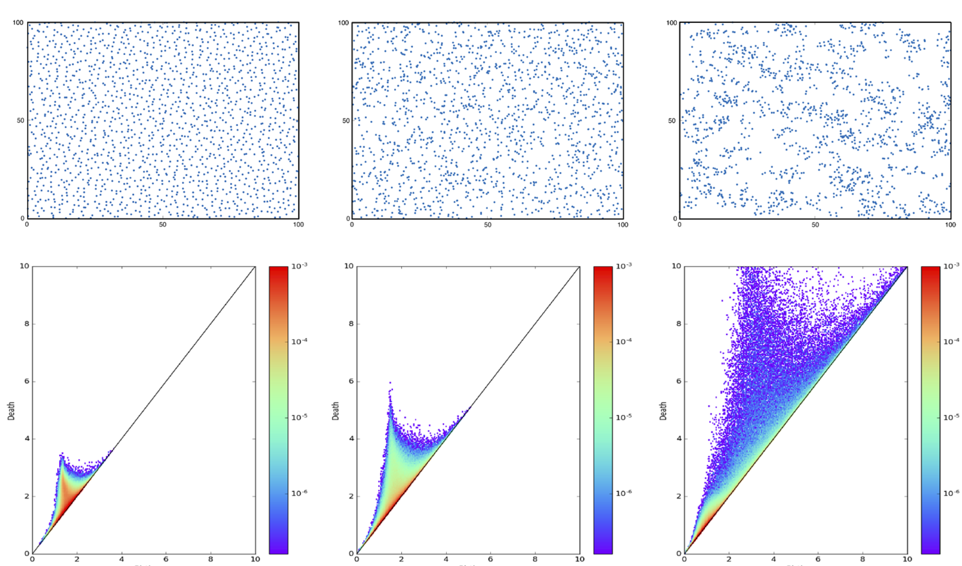

FIG. 1. *Top: Point processes with negative (Ginibre), zero (Poisson), and positive (Poisson cluster) correlations. In these three point processes, the number of points and the density are set to be* $1{,}000{,}000$ *and* $1/2\pi$, *respectively. Bottom: The normalized persistence diagrams* $\xi_{1,L}/L^2$ *of the above.*

One of the main results in this paper is as follows.

THEOREM 1.5. *Assume that Φ is a stationary point process on $\mathbf{R}^N$ having all finite moments. Then, for each $q \geq 0$, there exists a unique Radon measure ν_q on Δ such that*

$$\frac{1}{L^N}\mathbb{E}[\xi_{q,L}] \xrightarrow{v} \nu_q \qquad \text{as } L \to \infty. \tag{1.5}$$

Here, $\xrightarrow{v}$ denotes the vague convergence of measures on Δ. In addition, if Φ is ergodic, then almost surely,

$$\frac{1}{L^N}\xi_{q,L} \xrightarrow{v} \nu_q \qquad \text{as } L \to \infty. \tag{1.6}$$

We call the limiting Radon measure ν_q the qth persistence diagram of a stationary ergodic point process Φ. In nonergodic case, by using the ergodic decomposition (cf. [14]), the right-hand side in (1.6) is replaced by the random measure $\nu_{q,\omega}$ which is measurable with respect to the translation invariant σ-field $\mathcal{I}$ defined in Section 3.

REMARK 1.6. The set Δ is topologically the same as the triangle

$$\{(x, y) \in \mathbf{R}^2 : 0 \leq x < y \leq 1\}$$

with open boundary $\partial\Delta = \{(x, x) \in \mathbf{R}^2 : 0 \leq x \leq 1\}$. Although we do not consider the mass on $\partial\Delta$, intuitively speaking, the (virtual) mass on $\partial\Delta$ comes from configurations of special forms such as three vertices of a right triangle. With the vague convergence, we do not see the mass escaping towards the boundary $\partial\Delta$ in the limit $L \to \infty$. In applications, the mass appearing near the boundary is considered to be fragile under perturbation while the one away from the boundary is considered to be robust.

The limiting measure ν_q may be trivial. Indeed, for Čech complexes, $\nu_q = 0$ for $q \geq N$. This is just because there is no configuration in $\mathbf{R}^N$ that realizes the qth homology class for $q \geq N$. In order to characterize the support of ν_q, we introduce the notion of realizability of a point in a persistence diagram.

DEFINITION 1.7. We say that a point $(b, d) \in \Delta$ is *realizable by* $\Xi \in \mathscr{F}(\mathbf{R}^N)$ in the qth persistent homology if (b, d) is contained in the qth persistence diagram of the κ-filtration over Ξ, that is, $\xi_q(\Xi)(\{(b, d)\}) \geq 1$. If such Ξ exists for (b, d), we call (b, d) a *realizable* point. We denote by $R_q = R_q(\kappa)$ the set of all realizable points in the qth persistent homology of the κ-filtration.

EXAMPLE 1.8. For $\alpha > 0$ and $\sigma \in \mathscr{F}(\mathbf{R}^N)$, we define $\alpha\sigma \in \mathscr{F}(\mathbf{R}^N)$ by $\alpha\sigma = \{\alpha x \in \mathbf{R}^N : x \in \sigma\}$. It is easy to see that if κ is homogeneous in the sense that $\kappa(\alpha\sigma) = \alpha\kappa(\sigma)$ for every $\alpha > 0$ and $\sigma \in \mathscr{F}(\mathbf{R}^N)$, then $R_q(\kappa)$ forms a cone in Δ. Since both κ_C and κ_R given in Example 1.3 are homogeneous, we can see that $R_q(\kappa_C)$ and $R_q(\kappa_R)$ are cones for every $q \geq 0$. In particular, for Čech complexes, we have

$$R_q(\kappa_C) = \begin{cases} \{0\} \times (0, \infty] & \text{if } q = 0, \\ \{(b, d) : 0 < b < d < \infty\} & \text{if } q = 1, 2, \dots, N-1, \\ \varnothing & \text{if } q = N, N+1, \dots. \end{cases} \tag{1.7}$$

The sketch of the proof is given at the end of Section 2.2.

It is clear that $\operatorname{supp} \nu_q \subset \overline{R_q(\kappa)}$. Indeed, if $x \notin \overline{R_q(\kappa)}$, there exists $\varepsilon > 0$ such that $\xi_{q,L}(B_\varepsilon(x)) = 0$, where $B_\varepsilon(x)$ is the open ε-neighborhood of x. It follows from the vague convergence (1.5) that $\nu_q(B_\varepsilon(x)) = 0$. Therefore, $x \notin \operatorname{supp} \nu_q$. In Theorem 4.7, we give sufficient conditions for a point in $R_q(\kappa)$ to be in the support of ν_q. The following result, as a consequence of that general theorem, states that $\operatorname{supp} \nu_q$ coincides with $\overline{R_q(\kappa)}$ under conditions that κ is Lipschitz continuous and all local densities of the point process Φ are almost surely positive with respect to the Lebesgue measures.

THEOREM 1.9. *Let Φ be a stationary point process on $\mathbf{R}^N$ and Θ its probability distribution. Assume that for every compact set $\Lambda \subset \mathbf{R}^N$, the restriction $\Theta|_\Lambda$ on Λ is absolutely continuous with respect to $\Pi|_\Lambda$ and the Radon–Nikodym density $d\Theta|_\Lambda / d\Pi|_\Lambda$ is strictly positive $\Pi|_\Lambda$-almost surely, where Π is the distribution of a homogeneous Poisson point process on $\mathbf{R}^N$. In addition, assume that κ on $\mathscr{F}(\mathbf{R}^N)$ is Lipschitz continuous with respect to the Hausdorff distance. Then* $\operatorname{supp} \nu_q = \overline{R_q(\kappa)}$ *for every* $q \geq 0$.

EXAMPLE 1.10. All finite configurations are allowed to appear in a point process if the positivity assumption in Theorem 1.9 holds. There are many "natural" stationary point processes satisfying the assumption. Homogeneous Poisson point processes, a certain class of Gibbs point processes, Ginibre point processes and the zeros of the Gaussian entire function $X(z) = \sum_{n=0}^{\infty} (n!)^{-1/2} a_n z^n$ with $\{a_n\}_{n \geq 0}$ being i.i.d. complex standard Gaussian random variables, etc. are such examples. Thus if κ is Lipschitz continuous with respect to the Hausdorff distance, then $\operatorname{supp} \nu_q = \overline{R_q(\kappa)}$ for such point processes. In particular, for Čech filtrations, $\operatorname{supp} \nu_q = \Delta$ for $q = 1, \dots, N-1$ and $\operatorname{supp} \nu_0 = \{0\} \times (0, \infty]$. On the other hand, the shifted lattice considered in Example 4.3 does not satisfy the assumption and $\operatorname{supp} \nu_q$ for the Čech filtration turns out to be a singleton in Δ.

See Example 4.9 for more explanation about positivity. One can also refer to [7] and references therein for Gibbs point processes and other concrete examples.

For the proof of Theorem 1.5, we exploit a general theory of Radon measures for the vague convergence (cf. [1, 24]). In particular, we show that the convergence of the values of measures on the class $\{A_{r,s} = [0,r] \times (s,\infty] : 0 \le r \le s < \infty\}$ is enough to ensure the vague convergence of random measures in Theorem 1.5. The value of $\xi_{q,L}$ on $A_{r,s}$ is nothing but the persistent Betti number

$$\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L})) = \xi_{q,L}([0,r] \times (s,\infty]) = \left|\{(b_i,d_i) : 0 \le b_i \le r \le s < d_i\}\right|.$$

Here, $|A|$ denotes the cardinality of a finite set A. Later, $|A|$ is also used to denote the Lebesgue measure of a set $A \in \mathbf{R}^N$. The meaning is clear from the context. Hence Theorem 1.5 follows from the following strong law of large numbers for persistent Betti numbers.

THEOREM 1.11. *Assume that Φ is a stationary point process having all finite moments. Then, for any $0 \le r \le s < \infty$ and $q \ge 0$, there exists a constant $\hat{\beta}_q^{r,s}$ such that*

$$\frac{\mathbb{E}[\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L}))]}{L^N} \to \hat{\beta}_q^{r,s} \qquad \text{as } L \to \infty.$$

In addition, if Φ is ergodic, then

$$\frac{\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L}))}{L^N} \to \hat{\beta}_q^{r,s} \qquad \text{almost surely as } L \to \infty.$$

Note that, for $r = s$, the persistent Betti number becomes the usual Betti number, that is, $\beta_q^{r,r}(\mathbb{K}(\Phi_{\Lambda_L})) = \beta_q(K(\Phi_{\Lambda_L}, r))$. Hence, this result is a generalization of Lemma 3.3 and Theorem 3.5 in [39]. The positivity of the limiting persistent Betti number $\hat{\beta}_q^{r,s}$ is related to the previous problem of characterizing the support of ν_q. In particular, $\hat{\beta}_q^{r,s} \ge \nu_q([0,r) \times (s,\infty]) > 0$, if $\operatorname{supp} \nu_q$ touches $[0,r) \times (s,\infty]$.

Note also that when $q = 0$, all the measures $\xi_{0,L}$ are supported on $\{0\} \times (0,\infty]$ and $\beta_0^{r,s}(\mathbb{K}(\Phi_{\Lambda_L})) = \beta_0(K(\Phi_{\Lambda_L}, s))$ just counts the number of connected components in the geometric graph $G(\Phi_{\Lambda_L}, s) = (V, E)$, where

$$V = \Phi_{\Lambda_L}, \qquad E = \{(x,y) \in V \times V : \kappa(x,y) \le s\}.$$

In this case, the limiting measure ν_0 is also supported on $(0,\infty]$ with the following explicit formula:

$$\nu_0((s,\infty]) = \lambda \mathbb{E}^0[h_s(0, \Phi \cup \{0\})],$$

where λ is the intensity of Φ, $\mathbb{E}^0$ is the reduced Palm measure at 0, and $h_s(x,\Phi)$ is the reciprocal of the size of the connected components containing x in $G(\Phi, s)$.

Refer to Section 13.7 in [30] for more about the law of large numbers as well as the central limit theorem for β_0 of the Čech or Rips complex built over Poisson point processes and binomial point processes.

For Poisson point processes, we also generalize the central limit theorem in [39] for Betti numbers to persistent Betti numbers as follows.

THEOREM 1.12. *Let Φ be a homogeneous Poisson point process on $\mathbf{R}^N$ with unit intensity. Then for any $0 \le r \le s < \infty$ and $q \ge 0$, there exists a constant $\sigma_{r,s}^2 = \sigma_{r,s}^2(q)$ such that*

$$\frac{\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L})) - \mathbb{E}[\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L}))]}{L^{N/2}} \xrightarrow{d} \mathcal{N}(0, \sigma_{r,s}^2) \qquad \text{as } L \to \infty.$$

We remark that the proof of the central limit theorem for (usual) Betti numbers in [39] uses the Mayer–Vietoris exact sequence to estimate the effect of one point adding on the Betti number. However, in the setting of persistent homology, although we can obtain the Mayer–Vietoris exact sequence for each parameter r, we do not have the exactness property with regard to the parameter change. Hence, the same technique may not be applicable to the case of persistent Betti numbers. Instead, we give an alternative (and elementary) proof for the generalization. Remark also that by establishing the strong stabilization, the central limit theorem for Betti numbers of Čech complexes built over binomial point processes is also established in [39]. In this case, the positivity of the limiting variance is also proved under a certain condition on radius parameter r. The positivity problem for the limiting variance is left open in case of persistent Betti numbers of general κ-complexes built over homogeneous Poisson point processes.

The organization of this paper is given as follows. Necessary concepts and properties of persistent homology and random measures are explained in Section 2 and Section 3, respectively. Theorem 4.7 which characterizes the support of limiting persistence diagrams is stated and proved in Section 4.3. The proofs of Theorems 1.5, 1.9, 1.11 and 1.12 are given in Sections 4.2, 4.3, 4.1 and 5 in order. In Section 6, we summarize the conclusions of the paper and show some future works.

2. Geometric models and persistent homology.

In this section, we assume fundamental properties about simplicial complexes and their homology. For details, the reader may refer to Appendix B or [9, 15].

2.1. *Geometric models for point processes.*

Let $\kappa : \mathscr{F}(\mathbf{R}^N) \to [0, \infty]$ be a function satisfying the three conditions explained in Section 1, where $\mathscr{F}(\mathbf{R}^N)$ is the collection of all finite subsets in $\mathbf{R}^N$. For such a function κ, the κ-filtration $\mathbb{K}(\Phi) = \{K(\Phi, t)\}_{t \ge 0}$ can be defined in the same way as in (1.2) for an infinite point configuration (or a point process) $\Phi \subset \mathbf{R}^N$ as well as for a finite one.

We remark that all vertices (i.e., 0-simplices) exist at $t = 0$. Also, all simplices in $K(\Phi, t)$ possessing a point x must lie in the ball $\bar{B}_{\rho(t)}(x)$ since $\{x, x_1, \dots, x_q\} \in K(\Phi, t)$ with Assumption (K3) implies that $\|x - x_i\| \le \rho(\kappa(x, x_i)) \le \rho(t)$ for all i. Here, $\bar{B}_r(x) = \{y \in \mathbf{R}^N : \|y - x\| \le r\}$ is the closure of $B_r(x)$ which denotes the open ball of radius r centered at x. Hence, for each parameter t, the presence of simplices containing x is localized in $\bar{B}_{\rho(t)}(x)$.

This geometric model includes some of the standard models studied in random topology. For instance, the Čech complex $C(\Phi, t)$ is a simplicial complex with the vertex set Φ and, for each parameter t, it is defined by

$$\sigma = \{x_0, \dots, x_q\} \in C(\Phi, t) \quad \Longleftrightarrow \quad \bigcap_{i=0}^{q} \bar{B}_t(x_i) \neq \varnothing$$

for q-simplices. Similarly, the Rips complex $R(\Phi, t)$ with a parameter t is defined by

$$\sigma = \{x_0, \dots, x_q\} \in R(\Phi, t) \quad \Longleftrightarrow \quad \bar{B}_t(x_i) \cap \bar{B}_t(x_j) \neq \varnothing \qquad \text{for } 0 \leq i < j \leq q.$$

It is clear that these geometric models are generated by the functions given in Example 1.3. We note that $R(\Phi, t/2) \subset C(\Phi, t) \subset R(\Phi, t)$ since $\kappa_R \leq \kappa_C \leq 2\kappa_R$.

2.2. *Persistent homology.* Let $\mathbb{K} = \{K_r : r \geq 0\}$ be a (right continuous) filtration of simplicial complexes, that is, $K_r \subset K_s$ for $r \leq s$ and $K_r = \bigcap_{r<s} K_s$. In this paper, the homology $H_q(K)$ of a simplicial complex K is defined on an arbitrary field $\mathbf{F}$. For $r \leq s$, we denote the linear map on homologies induced from the inclusion $K_r \hookrightarrow K_s$ by $\iota_r^s\colon H_q(K_r) \to H_q(K_s)$. The *$q$th persistent homology* $H_q(\mathbb{K}) = (H_q(K_r), \iota_r^s)$ of $\mathbb{K}$ is defined by the family of homologies $\{H_q(K_r) : r \geq 0\}$ and the induced linear maps ι_r^s for all $r \leq s$.

A *homological critical value* of $H_q(\mathbb{K})$ is a number $r > 0$ such that the linear map $\iota_{r-\varepsilon}^{r+\varepsilon}\colon H_q(K_{r-\varepsilon}) \to H_q(K_{r+\varepsilon})$ is not isomorphic for any sufficiently small $\varepsilon > 0$. The persistent homology $H_q(\mathbb{K})$ is said to be *tame* if $\dim H_q(K_r) < \infty$ for any $r \geq 0$ and the number of homological critical values is finite. A tame persistent homology $H_q(\mathbb{K})$ has a nice decomposition property.

THEOREM 2.1 ([40]). *Assume that $H_q(\mathbb{K})$ is a tame persistent homology. Then there uniquely exist indices $p \in \mathbf{Z}_{\geq 0}$ and $b_i, d_i \in \overline{\mathbf{R}}_{\geq 0} = \mathbf{R}_{\geq 0} \sqcup \{\infty\}$ with $b_i < d_i$, $i = 1, 2, \dots, p$, such that the following isomorphism holds*:

$$H_q(\mathbb{K}) \simeq \bigoplus_{i=1}^{p} I(b_i, d_i). \tag{2.1}$$

Here, $I(b_i, d_i) = (U_r, f_r^s)$ consists of a family of vector spaces

$$U_r = \begin{cases} \mathbf{F}, & b_i \leq r < d_i, \\ 0 & \textit{otherwise}, \end{cases}$$

and the identity map $f_r^s = \mathrm{id}_{\mathbf{F}}$ for $b_i \leq r \leq s < d_i$.

Each summand $I(b_i, d_i)$ in (2.1) is called a generator of the persistent homology and (b_i, d_i) is called its birth-death pair. From the unique decomposition in Theorem 2.1, we define the qth persistence diagram as a multiset in $\overline{\mathbf{R}}_{\geq 0}^2$,

$$D_q(\mathbb{K}) = \{(b_i, d_i) \in \overline{\mathbf{R}}_{\geq 0}^2 : i = 1, \dots, p\}. \tag{2.2}$$

By denoting the multiplicity of the point (b,d) in (2.2) by $m_{b,d} \in \mathbf{N}_0 = \{0,1,2,\dots\}$, we can also express the decomposition (2.1) as

$$H_q(\mathbb{K}) \simeq \bigoplus_{(b,d)} I(b,d)^{m_{b,d}}.$$

Later, we identify a persistence diagram $D_q(\mathbb{K})$ as an integer-valued Radon measure $\xi = \sum_{(b,d)} m_{b,d}\delta_{(b,d)}$ rather than as a multiset.

Intuitively speaking, the persistent homology $H_q(\mathbb{K})$ characterizes topological features (components, rings, cavities, etc.) in $\mathbb{K}$ in a multiscale way, and indeed, the interval decompositions (2.1) provide this viewpoint. Namely, each interval $I(b,d)$ means that a topological feature appears at the scale $r = b$, persists for $b \le r < d$, and disappears at $r = d$. Then the persistence diagram $D_q(\mathbb{K})$ is widely used for a compact visualization of this multiscale characterization.

Although our target object $\mathbb{K}(\Phi)$ is built over infinite points, all persistent homologies studied in this paper are defined on the geometric models with finite points. Hence, the persistent homology becomes tame, and the persistence diagrams are well defined.

EXAMPLE 2.2. In Figure 2, two (1-dim)cycles appear at times 1 and 2 and disappear at times 3 and 4. The representation corresponding to $H_1(\mathbb{K})$ is given as

$$\begin{aligned} 0 &\to \mathbf{F}(c_1 + c_2) \to \mathbf{F}(c_1) \oplus \mathbf{F}(c_2) \to \mathbf{F}(c_1) \oplus \mathbf{F}(c_2)/\mathbf{F}(c_1) \\ &\to \mathbf{F}(c_1) \oplus \mathbf{F}(c_2)/\mathbf{F}(c_1) \oplus \mathbf{F}(c_2) \simeq 0, \end{aligned}$$

where $c_1 = \langle 12\rangle + \langle 23\rangle + \langle 31\rangle$ and $c_2 = \langle 13\rangle + \langle 34\rangle + \langle 41\rangle$ and each arrow is the linear map induced by inclusion. As pairs of birth-death times, we have $(1,4)$ and $(2,3)$ since the decomposition of the representation is given by

$$\begin{aligned} H_1(\mathbb{K}) = \big(&0 \to \mathbf{F}(c_1 + c_2) \to \mathbf{F}(c_1 + c_2) \to \mathbf{F}(c_1 + c_2) \to 0\big) \\ &\oplus \big(0 \to 0 \to \mathbf{F}(c_1) \to 0 \to 0\big). \end{aligned}$$

REMARK 2.3. More generally, a persistence module $\mathbb{U} = (U_a, f_a^b)$ on $\overline{\mathbf{R}}_{\ge 0}$ is defined by a sequence of general vector spaces $U_a, a \ge 0$, and linear maps $f_a^b \colon U_a \to U_b$ for $a \le b$ satisfying $f_a^c = f_b^c \circ f_a^b$. Under the same definition of the tameness, we can similarly define its persistence diagrams.

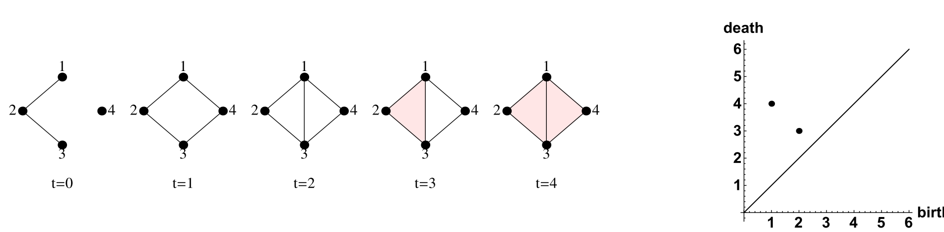

FIG. 2. *A filtration of simplicial complexes and the* 1*st persistence diagram*

REMARK 2.4. There is another definition of persistent homology as graded modules over a monoid ring for the continuous parameter (resp., a polynomial ring for the discrete parameter); see, for example, [18].

REMARK 2.5. The persistent homology $H_q(\mathbb{K})$ defined over the whole Φ is not tame in general while $H_q(\mathbb{K}_L)$ defined over a restriction Φ_{Λ_L} is tame. Theorem 1.5 informally says that

$$\frac{1}{L^N} H_q(\mathbb{K}_L) \overset{\simeq}{\to} H_q(\mathbb{K}) = \int_\Delta^\oplus I(x, y)\nu_q(dx\, dy),$$

where $\mathbb{K}_L = \{K(\Phi_{\Lambda_L}, t)\}_{t \geq 0}$, and $\int_\Delta^\oplus$ denotes the direct integral of interval representations (cf. [33]).

REMARK 2.6. In our paper, we use the persistence diagram for representing topological information obtained from filtrations. People sometimes use the so-called barcode representation in which each persistence interval $I(b, d)$ is represented as a barcode $[b, d]$ (cf. [38]). We consider the marginal measure of persistence diagram on death times (also on birth times), that is, the induced measure $\xi^{(\text{death})}$ obtained from a measure ξ on Δ by the projection $\Delta \ni (x, y) \mapsto y \in (0, \infty]$. The marginal measure $\xi_{q,L}^{(\text{death})}$ of a persistence diagram $\xi_{q,L}$ induces a (scaled) right-continuous step function $f_{q,L}(t) = L^{-N}\xi_{q,L}^{(\text{death})}([0, t])$, which corresponds to the one obtained by simulation in [38]. The function $f_{q,L}(t)$ is also expected to converge to a limit $f_{q,\infty}(t)$ as $L \to \infty$, however, it does not necessarily coincide with $f_q(t) := \nu_q^{(\text{death})}([0, t])$ because of the mass escaping to $\partial\Delta$.

In Example 1.8, we showed the set $R_q(\kappa_C)$ of the realizable points in (1.7) for the Čech filtration. Here, we give a brief sketch of the proof. The cases $q = 0$ and $q \geq N$ are easily derived. For $q = 1, \dots, N-1$, we show that any birth-death pair (b, d) with $0 < b < d < \infty$ is realizable by explicitly constructing the points $\Xi \in \mathscr{F}(\mathbf{R}^N)$ realizing (b, d) (see Figure 3 for $q = 1$). Indeed, let $S_d^q \subset \mathbf{R}^N$ be a q-dimensional sphere with radius d and take a $(q-1)$-dimensional sphere S_b^{q-1} with radius b so that $S_d^q = H_+ \sqcup S_b^{q-1} \sqcup H_-$, where H_+ (resp., H_-) is the upper (resp., lower) hemisphere with $\partial H_\pm = S_b^{q-1}$ and H_+ is chosen to be the smaller one. We choose points Ξ_+ on S_b^{q-1} and Ξ_- on H_- such that:

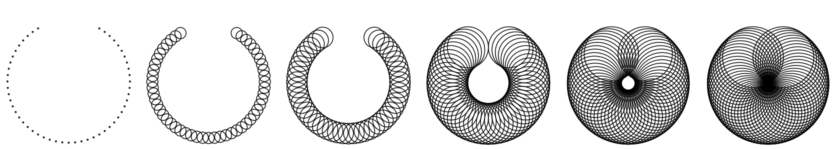

FIG. 3. *For $(b, d) = (5, 10)$ when $q = 1$, the set $\bigcup_{x\in\Xi} \bar{B}_r(x)$ is drawn for $r = 0, 1, 2, 5, 8, 10$. A cycle appears at $r = 5$ and disappears at $r = 10$.*

(i) $\bigcup_{x\in\Xi_-} \bar{B}_r(x)$ covers $\bar{H}_-$ earlier than $r = b$;

(ii) $\bigcup_{x\in\Xi_+} \bar{B}_b(x)$ covers S_b^{q-1} and is contractive;

(iii) $\bigcup_{x\in\Xi} \bar{B}_r(x)$ provides the generator of q-dimensional homology homeomorphic to S_d^q, where $\Xi = \Xi_+ \sqcup \Xi_-$.

Then the birth-death pair of the generator $\bigcup_{x\in\Xi} \bar{B}_r(x)$ is exactly (b, d).

2.3. *Persistent Betti numbers.* For a filtration $\mathbb{K}$, the (r, s)-persistent Betti number [10] is defined by

$$\beta_q^{r,s}(\mathbb{K}) = \dim \frac{Z_q(K_r)}{Z_q(K_r) \cap B_q(K_s)} \qquad (r \le s), \tag{2.3}$$

where $Z_q(K_r)$ and $B_q(K_r)$ are the qth cycle group and boundary group, respectively. We remark that this is equal to the rank of $\iota_r^s : H_q(K_r) \to H_q(K_s)$, because

$$\operatorname{im} \iota_r^s \simeq \frac{\frac{Z_q(K_r)}{B_q(K_r)}}{\frac{Z_q(K_r)\cap B_q(K_s)}{B_q(K_r)}} \simeq \frac{Z_q(K_r)}{Z_q(K_r) \cap B_q(K_s)}.$$

Thus, from the decomposition of the persistent homology, we have

$$\beta_q^{r,s}(\mathbb{K}) = \sum_{b \le r, d > s} m_{b,d}.$$

This means that the (r, s)-persistent Betti number $\beta_q^{r,s}(\mathbb{K})$ counts the number of birth-death pairs in the persistence diagram $D_q(\mathbb{K})$ located in the gray region of Figure 4.

LEMMA 2.7. *Let $\mathbb{U} = (U_a, f_a^b)$ be a persistence module on $\overline{\mathbf{R}}_{\ge 0}$ and let $\mathbb{V} = (V_a, g_a^b)$ be its truncation on the interval $[r, s]$, meaning that*

$$V_a = \begin{cases} U_r, & a \le r, \\ U_a, & r \le a \le s, \\ U_s, & a \ge s, \end{cases} \qquad g_a^b = \begin{cases} f_a^b, & r \le a \le b \le s, \\ f_a^s, & r \le a \le s \le b, \\ f_r^b, & a \le r \le b \le s, \\ f_r^s, & a \le r \le s \le b. \end{cases} \tag{2.4}$$

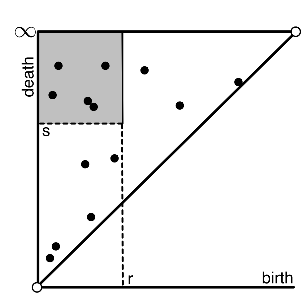

FIG. 4. $\beta_q^{r,s}(\mathbb{K})$ *counts the number of generators in the gray region.*

For interval decompositions $\mathbb{U} \simeq \oplus I(b,d)^{m_{b,d}}$ *and* $\mathbb{V} \simeq \oplus I(b,d)^{n_{b,d}}$, *let*

$$\beta^{r,s}(\mathbb{U}) = \sum_{b \le r, d > s} m_{b,d}, \qquad \beta^{0,\infty}(\mathbb{V}) = n_{0,\infty}.$$

Then $\beta^{r,s}(\mathbb{U}) = \beta^{0,\infty}(\mathbb{V})$.

PROOF. This is because $\beta^{r,s}(\mathbb{U}) = \operatorname{rank} f_r^s = \beta^{0,\infty}(\mathbb{V})$. □

Here, we recall the following basic facts in linear algebra for later use.

LEMMA 2.8. *Let* A, B, U, V *be subspaces of a vector space satisfying* $A \subset U$ *and* $B \subset V$. *Then*

$$\dim \frac{U \cap V}{A \cap B} \le \dim \frac{U}{A} + \dim \frac{V}{B}.$$

PROOF. It follows from the formulas $\dim(U \cap V) + \dim(U + V) = \dim U + \dim V$ and $\dim(U/A) = \dim U - \dim A$. □

LEMMA 2.9. *Let* $D = [AB]$ *be a matrix composed by submatrices* A *and* B. *Let* ℓ *be the number of columns in* B. *Then*

$$\operatorname{rank} D \le \operatorname{rank} A + \ell, \qquad \dim\ker D \le \dim\ker A + \ell.$$

PROOF. Let $B = [b_1 \cdots b_\ell]$, where b_i is the ith column vector of B, and set $D^{(i)} = [Ab_1 \cdots b_i]$. Then, for each i, we have

$$\operatorname{rank} D^{(i)} \le \operatorname{rank} D^{(i-1)} + 1, \qquad \dim\ker D^{(i)} \le \dim\ker D^{(i-1)} + 1.$$

Hence, in total, we have the desired inequalities. □

Now, we show a basic estimate on the persistent Betti number for nested filtrations $\mathbb{K} \subset \tilde{\mathbb{K}}$. First, we note the following property.

LEMMA 2.10. *Let* $\mathbb{K}$ *be a filtration. For a fixed* $a > 0$, *let* $\tilde{\mathbb{K}} = \{\tilde{K}_t : t \ge 0\}$ *be a filtration given by*

$$\tilde{K}_t = \begin{cases} K_t, & t < a, \\ K_t \cup \sigma, & t \ge a, \end{cases}$$

where σ *is a new simplex added on* K_a. *Then* $\beta_q^{r,s}(\tilde{\mathbb{K}}) = \beta_q^{r,s}(\mathbb{K})$ *for* $\dim \sigma \ne q, q+1$. *For* $\dim \sigma = q, q+1$,

$$\left|\beta_q^{r,s}(\tilde{\mathbb{K}}) - \beta_q^{r,s}(\mathbb{K})\right| \le \begin{cases} 0, & \tilde{K}_r = K_r \text{ and } \tilde{K}_s = K_s, \\ 1 & \text{otherwise}. \end{cases}$$

PROOF. We first note that

$$\begin{aligned}\beta_q^{r,s}(\tilde{\mathbb{K}}) - \beta_q^{r,s}(\mathbb{K}) &= \dim Z_q(\tilde{K}_r) - \dim Z_q(\tilde{K}_r) \cap B_q(\tilde{K}_s) \\ &\quad - \big(\dim Z_q(K_r) - \dim Z_q(K_r) \cap B_q(K_s)\big) \\ &= \dim \frac{Z_q(\tilde{K}_r)}{Z_q(K_r)} - \dim \frac{Z_q(\tilde{K}_r) \cap B_q(\tilde{K}_s)}{Z_q(K_r) \cap B_q(K_s)}.\end{aligned}$$

Hence, the statement is trivial for $\dim \sigma \neq q, q+1$. Furthermore, when $\tilde{K}_r = K_r$ and $\tilde{K}_s = K_s$, we also have $\beta_q^{r,s}(\tilde{\mathbb{K}}) = \beta_q^{r,s}(\mathbb{K})$.

Let $\dim \sigma = q$. Then it follows from Lemma 2.9 that

$$\dim \frac{Z_q(\tilde{K}_r)}{Z_q(K_r)} = \dim Z_q(\tilde{K}_r) - \dim Z_q(K_r) = 0 \text{ or } 1,$$

$$\dim \frac{B_q(\tilde{K}_s)}{B_q(K_s)} = 0.$$

Also, from Lemma 2.8, we have

$$0 \le \dim \frac{Z_q(\tilde{K}_r) \cap B_q(\tilde{K}_s)}{Z_q(K_r) \cap B_q(K_s)} \le \dim \frac{Z_q(\tilde{K}_r)}{Z_q(K_r)} + \dim \frac{B_q(\tilde{K}_s)}{B_q(K_s)} \le \dim \frac{Z_q(\tilde{K}_r)}{Z_q(K_r)}.$$

Therefore, $|\beta_q^{r,s}(\tilde{\mathbb{K}}) - \beta_q^{r,s}(\mathbb{K})| \le 1$. The statement for $\dim \sigma = q+1$ is similarly proved. □

LEMMA 2.11. *Let $\mathbb{K} = \{K_t\}_{t \ge 0}$ and $\tilde{\mathbb{K}} = \{\tilde{K}_t\}_{t \ge 0}$ be filtrations with $K_t \subset \tilde{K}_t$ for $t \ge 0$. Then*

$$\big|\beta_q^{r,s}(\tilde{\mathbb{K}}) - \beta_q^{r,s}(\mathbb{K})\big| \le \sum_{j=q,q+1} \big(|\tilde{K}_{s,j} \setminus K_{s,j}| + \big|\{\sigma \in K_{s,j} \setminus K_{r,j} : \tilde{t}_\sigma \le r\}\big|\big),$$

where $\tilde{K}_{t,j}$ (or $K_{t,j}$) is the set of j-simplices in $\tilde{K}_t$ (or K_t), and $\tilde{t}_\sigma$ (or t_σ) is the birth time of σ in the filtration $\tilde{\mathbb{K}}$ (or $\mathbb{K}$).

PROOF. We first decompose $\tilde{K}_s \setminus K_r = Y \sqcup Y^c$ by

$$Y = (\tilde{K}_s \setminus K_s) \sqcup \{\sigma \in K_s \setminus K_r : \tilde{t}_\sigma \le r\}, \qquad Y^c = \{\sigma \in K_s \setminus K_r : r < \tilde{t}_\sigma \le t_\sigma\}.$$

We use the same notation Y_j for the set of j-simplices in Y. For the simplices in $\tilde{K}_s \setminus K_r = \{\sigma_i\}_{i=1}^L$, we assign their indices so that the birth times are in increasing order $\tilde{t}_{\sigma_1} \le \cdots \le \tilde{t}_{\sigma_L}$ and $K_r \cup \{\sigma_1, \dots, \sigma_\ell\}$ becomes a simplicial complex for each ℓ. We note that $\tilde{t}_\sigma \le t_\sigma$. Furthermore, it suffices to consider the truncations of $\mathbb{K}$ and $\tilde{\mathbb{K}}$ on $[r, s]$ from Lemma 2.7.

Now, we inductively construct a sequence of filtrations $\mathbb{K} = \mathbb{K}^0 \subset \mathbb{K}^1 \subset \cdots \subset \mathbb{K}^L = \tilde{\mathbb{K}}$. The filtration $\mathbb{K}^i = \{K_t^i : t \geq 0\}$ is given by adding a simplex σ_i to $\mathbb{K}^{i-1}$ at $\tilde{t}_{\sigma_i}$, that is,

$$K_t^i = \begin{cases} K_t^{i-1}, & t < \tilde{t}_{\sigma_i}, \\ K_t^{i-1} \cup \{\sigma_i\}, & t \geq \tilde{t}_{\sigma_i}. \end{cases}$$

Then it follows from Lemma 2.10 that $|\beta_q^{r,s}(\mathbb{K}^i) - \beta_q^{r,s}(\mathbb{K}^{i-1})| \leq 1$ for $\sigma_i \in Y$, since $K_r^i \neq K_r^{i-1}$ or $K_s^i \neq K_s^{i-1}$ holds. On the other hand, Lemma 2.10 implies $\beta_q^{r,s}(\mathbb{K}^i) = \beta_q^{r,s}(\mathbb{K}^{i-1})$ for $\sigma \in Y^c$. Therefore,

$$\left|\beta_q^{r,s}(\tilde{\mathbb{K}}) - \beta_q^{r,s}(\mathbb{K})\right| \leq \sum_{i=1}^{L} \left|\beta_q^{r,s}(\mathbb{K}^i) - \beta_q^{r,s}(\mathbb{K}^{i-1})\right| \leq |Y_q| + |Y_{q+1}|,$$

which completes the proof of Lemma 2.11. □

REMARK 2.12. Let $\Phi, \tilde{\Phi} \in \mathscr{F}(\mathbf{R}^N)$ with $\Phi \subset \tilde{\Phi}$, and t_σ and $\tilde{t}_\sigma$ be the birth times of the simplex σ in the κ-filtrations $\mathbb{K}(\Phi)$ and $\mathbb{K}(\tilde{\Phi})$, respectively. Then it is obvious that $\tilde{t}_\sigma = t_\sigma$ if $\sigma \subset \Phi \subset \tilde{\Phi}$. Hence, for the estimate $|\beta_q^{r,s}(\mathbb{K}(\tilde{\Phi})) - \beta_q^{r,s}(\mathbb{K}(\Phi))|$, the second term obtained in Lemma 2.11 does not appear under this setting.

3. General theory of random measures. In this section, we give a brief account of random measures (cf. [24]) and prove Proposition 3.4 which provides a sufficient condition for the law of large numbers for random measures to hold. The notion of convergence-determining class for vague convergence plays an important role in Proposition 3.4. We discuss it separately in Appendix A.

Let S be a locally compact Hausdorff space with countable basis and $\mathcal{S}$ be the Borel σ-algebra on S. It is well known that S is a Polish space, that is, a complete separable metrizable space. If needed, we take a metric ρ which makes S complete and separable. We denote by $\mathscr{B}(S)$ the ring of all relatively compact sets in $\mathcal{S}$. A measure μ on $(S, \mathcal{S})$ is said to be a *Radon measure* if $\mu(B) < \infty$ for every $B \in \mathscr{B}(S)$. Let $\mathfrak{M}(S)$ be the set of all Radon measures on $(S, \mathcal{S})$ and $\mathcal{M}(S)$ be the σ-algebra generated by the mappings $\mathfrak{M}(S) \ni \mu \mapsto \mu(B) \in [0, \infty)$ for every $B \in \mathscr{B}(S)$.

We say that a sequence $\{\mu_n\}_{n\geq 1} \subset \mathfrak{M}(S)$ converges to $\mu \in \mathfrak{M}(S)$ *vaguely* (or *in the vague topology*) if $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for every continuous function f with compact support, where $\langle \mu, f\rangle = \int_S f(x)\, d\mu(x)$. In this case, we write $\mu_n \xrightarrow{v} \mu$. The space $\mathfrak{M}(S)$ equipped with the vague topology again becomes a Polish space and its Borel σ-algebra coincides with $\mathcal{M}(S)$.

We denote by $\mathfrak{N}(S)$ the subset in $\mathfrak{M}(S)$ of all integer-valued Radon measures on S. Each element in $\mathfrak{N}(S)$ can be expressed as a sum of delta measures, that is,

$\mu = \sum_i \delta_{x_i} \in \mathfrak{N}(S)$. We note that the set $\mathfrak{N}(S)$ is a closed subset of $\mathfrak{M}(S)$ in the vague topology.

An $\mathfrak{M}(S)$-valued [resp., $\mathfrak{N}(S)$-valued] random variable $\xi = \xi_\omega$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *random measure* (resp., *point process*) on S. If $\lambda_1(A) := \mathbb{E}[\xi(A)] < \infty$ for all $A \in \mathscr{B}(S)$, then λ_1 defines a Radon measure and is referred to as the mean measure, or the intensity measure of ξ. Sometimes we denote it by $\mathbb{E}[\xi]$.

In this paper, two kinds of point processes will appear. One is point processes on $\mathbf{R}^N$ as spatial point data and the other is point processes on $\Delta = \{(x, y) \in \overline{\mathbf{R}}^2 : 0 \le x < y \le \infty\}$ as persistence diagrams. The former will be denoted by the upper case letters like Φ and the latter by the lower case letters like ξ.

The point process Φ on $\mathbf{R}^N$ is called *stationary*, if the distribution $\mathbb{P}\Phi^{-1}$ is invariant under translations, that is, $\mathbb{P}\Phi_x^{-1} = \mathbb{P}\Phi^{-1}$ for any $x \in \mathbf{R}^N$, where Φ_x is the translated point process defined by $\Phi_x(B) = \Phi(B - x)$ for $B \in \mathscr{B}(\mathbf{R}^N)$. For $A \subset \mathfrak{M}(\mathbf{R}^N)$, let $A_x = \{\mu_x : \mu \in A\}$ be a set of translated measures defined by $\mu_x(B) = \mu(B - x)$. Given a point process Φ, let $\mathcal{I}$ be the translation invariant σ-field in $\mathfrak{N}(\mathbf{R}^N)$, that is, the class of subsets $I \subset \mathfrak{N}(\mathbf{R}^N)$ satisfying

$$\mathbb{P}\Phi^{-1}\big((I \setminus I_x) \cup (I_x \setminus I)\big) = 0$$

for all $x \in \mathbf{R}^N$. Then Φ is called *ergodic* if $\mathcal{I}$ is trivial, that is, for every $I \in \mathcal{I}$, $\mathbb{P}\Phi^{-1}(I) \in \{0, 1\}$.

From now on and until the end of this section, we fix a space S and write $\mathscr{B}$ and $\mathfrak{M}$ for $\mathscr{B}(S)$ and $\mathfrak{M}(S)$, respectively. For a subset $A \subset S$, we denote by ∂A and A° the boundary and interior of A, respectively. For a measure $\mu \in \mathfrak{M}$, let $\mathscr{B}_\mu := \{B \in \mathscr{B} : \mu(\partial B) = 0\}$ be the class of relatively compact continuity sets of μ.

LEMMA 3.1 ([24], 15.7.2). *Let $\mu, \mu_1, \mu_2, \ldots \in \mathfrak{M}$. Then the following statements are equivalent*:

(i) $\mu_n \xrightarrow{v} \mu$;
(ii) $\mu_n(B) \to \mu(B)$ *for all* $B \in \mathscr{B}_\mu$;
(iii) $\limsup_{n\to\infty} \mu_n(F) \le \mu(F)$ *and* $\liminf_{n\to\infty} \mu_n(G) \ge \mu(G)$ *for all closed* $F \in \mathscr{B}$ *and open* $G \in \mathscr{B}$.

LEMMA 3.2 ([24], 15.7.5). *A subset $\mathscr{C}$ in $\mathfrak{M}$ is relatively compact in the vague topology iff*

$$\sup_{\mu \in \mathscr{C}} \mu(B) < \infty \qquad \text{for every } B \in \mathscr{B}.$$

A class $\mathscr{A} \subset \mathscr{B}$ is called a *convergence-determining class* (for vague convergence) if for every $\mu \in \mathfrak{M}$ and every sequence $\{\mu_n\} \subset \mathfrak{M}$, the condition

$$\mu_n(A) \to \mu(A) \qquad \text{for all } A \in \mathscr{A} \cap \mathscr{B}_\mu$$

implies the vague convergence $\mu_n \xrightarrow{v} \mu$. A class $\mathscr{A}_\mu \subset \mathscr{B}_\mu$ is called a convergence-determining class for μ if for any sequence $\{\mu_n\} \subset \mathfrak{M}$, the condition

$$\mu_n(A) \to \mu(A) \qquad \text{as } n \to \infty \text{ for all } A \in \mathscr{A}_\mu,$$

implies that $\mu_n \xrightarrow{v} \mu$. By definition, a class $\mathscr{A}$ is a convergence-determining class if and only if for any $\mu \in \mathfrak{M}$, $\mathscr{A}_\mu = \mathscr{A} \cap \mathscr{B}_\mu$ is a convergence-determining class for μ.

We say that a class $\mathscr{C}$ has the *finite covering property* if any subset $B \in \mathscr{B}$ can be covered by a finite union of $\mathscr{C}$-sets.

LEMMA 3.3. *Let $\mathscr{A}$ be a convergence-determining class with finite covering property. Let $\{\mu_n\}$ be a sequence of measures in $\mathfrak{M}$. If $\mu_n(A)$ converges to a finite limit for any $A \in \mathscr{A}$, then there exists a measure μ to which the sequence $\{\mu_n\}$ converges vaguely.*

PROOF. For any relatively compact set $B \in \mathscr{B}$, we can find a finite cover $\{A_i\}_{i=1}^m \subset \mathscr{A}$ of B so that

$$\limsup_{n\to\infty} \mu_n(B) \le \limsup_{n\to\infty} \mu_n\left(\bigcup_{i=1}^m A_i\right) \le \lim_{n\to\infty} \sum_{i=1}^m \mu_n(A_i) < \infty.$$

Therefore, the sequence $\{\mu_n\}_{n\ge 1}$ is relatively compact by Lemma 3.2, and hence, there is a subsequence $\{\mu_{n_k}\}$ and $\mu \in \mathfrak{M}$ such that $\mu_{n_k} \xrightarrow{v} \mu$, that is, $\mu_{n_k}(A) \to \mu(A)$ for every $A \in \mathscr{B}_\mu$. This together with the assumption implies that $\mu_n(A) \to \mu(A)$ for every $A \in \mathscr{A} \cap \mathscr{B}_\mu$. Consequently, μ_n converges to μ vaguely from the definition of convergence-determining class. The proof is complete. □

PROPOSITION 3.4. *Let $\mathscr{A}$ be a convergence-determining class with finite covering property and the property that for every $\mu \in \mathfrak{M}$, it contains a countable convergence-determining class for μ. Let $\{\xi_n\}$ be a sequence of random measures on S, that is, a sequence of $\mathfrak{M}$-valued random variables. Assume that*:

(i) $\mathbb{E}[\xi_n] \in \mathfrak{M}$ *for all n, and that*

(ii) *for every $A \in \mathscr{A}$, there exists $c_A \in [0, \infty)$ such that $\mathbb{E}[\xi_n(A)] \to c_A$ as $n \to \infty$.*

Then there exists a unique measure $\mu \in \mathfrak{M}$ such that the mean measure $\mathbb{E}[\xi_n]$ converges vaguely to μ and $\mu(A) = c_A$ for $A \in \mathscr{A} \cap \mathscr{B}_\mu$.

Assume further that for every $A \in \mathscr{A}$,

$$\xi_n(A) \to c_A \qquad \textit{almost surely as } n \to \infty.$$

Then $\{\xi_n\}$ converges vaguely to μ almost surely.

PROOF. By Lemma 3.3, there exists a unique measure μ such that $\mathbb{E}[\xi_n]$ converges vaguely to μ as $n \to \infty$, and hence $\mu(A) = c_A$ for $A \in \mathscr{A} \cap \mathscr{B}_\mu$.

Now let $\mathscr{A}_\mu \subset \mathscr{A}$ be a countable convergence-determining class for μ. Then almost surely

$$\xi_n(A) \to \mu(A) \qquad \text{as } n \to \infty, \text{ for all } A \in \mathscr{A}_\mu,$$

which implies that the sequence $\{\xi_n\}$ converges vaguely to μ almost surely. The proof is complete. □

4. Convergence of persistence diagrams.

4.1. *Proof of Theorem* 1.11. Let Φ be a stationary point process on $\mathbf{R}^N$ having all finite moments. Let $F_q(\Phi_A, r)$ be the number of q-simplices in $K(\Phi_A, r)$ and $F_q(\Phi, r; A)$ be the number of q-simplices in $K(\Phi, r)$ with at least one vertex in $A \subset \mathbf{R}^N$. Recall that every q-simplex in $K(\Phi, r)$ containing x must lie in the closed ball $\bar{B}_{\rho(r)}(x)$. Therefore, similar to [39], Lemma 3.1, there exists a constant $C_{q,r}$ such that

$$\mathbb{E}\big[F_q(\Phi_A, r)\big] \le \mathbb{E}\big[F_q(\Phi, r; A)\big] \le C_{q,r}|A|$$

for all bounded Borel sets A, where $|A|$ is the Lebesgue measure of A.

We divide Λ_{mM} into m^N rectangles that are congruent to Λ_M and write as follows:

$$\Lambda_{mM} = \bigsqcup_{i=1}^{m^N} (\Lambda_M + c_i),$$

where c_i is the center of the ith rectangle. We compare $\mathbb{K}(\Phi_{\Lambda_{mM}})$ with a smaller filtration $\mathbb{K}^\circ(\Phi_{\Lambda_{mM}}) := \bigsqcup_{i=1}^{m^N} \mathbb{K}(\Phi_{\Lambda_M + c_i})$.

Let $\psi(L) = \mathbb{E}[\beta_q^{r,s}(\mathbb{K}(\Phi_{\Lambda_L}))]$ for $r \le s$. By Lemma 2.11, we have

$$\big|\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_{mM}})\big) - \beta_q^{r,s}\big(\mathbb{K}^\circ(\Phi_{\Lambda_{mM}})\big)\big| \le \sum_{j=q}^{q+1} \sum_{i=1}^{m^N} F_j(\Phi_{(\partial\Lambda_M)^{(\rho(s))} + c_i}, s). \tag{4.1}$$

Here, for $A \subset \mathbf{R}^N$, we write $A^{(t)} = \{x \in \mathbf{R}^N : \inf_{y \in A} \|x - y\| \le t\}$. Since $\mathbb{E}[F_j(\Phi_{(\partial\Lambda_M)^{(\rho(s))} + c_i}, s)] = O(|(\partial\Lambda_M)^{(\rho(s))} + c_i|) = O(M^{N-1})$ as $M \to \infty$, we have

$$\frac{\psi(mM)}{(mM)^N} = \frac{\psi(M)}{M^N} + O(M^{-1}). \tag{4.2}$$

Moreover, for $L > L'$,

$$\big|\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_L})\big) - \beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_{L'}})\big)\big| \le \sum_{j=q}^{q+1} F_j(\Phi_{\Lambda_L}, s; \Lambda_L \setminus \Lambda_{L'})$$

and

$$\mathbb{E}\big[F_j(\Phi_{\Lambda_L}, s; \Lambda_L \setminus \Lambda_{L'})\big] = O\big(|\Lambda_L \setminus \Lambda_{L'}|\big) = O\big((L-L')L^{N-1}\big).$$

Then, for fixed $M > 0$, taking $m \in \mathbb{N}$ such that $mM \le L < (m+1)M$, we see that

$$\frac{\psi(L)}{L^N} = \frac{\psi(mM)}{(mM)^N} + O\big(ML^{-1}\big). \tag{4.3}$$

It follows from (4.2) and (4.3) that $\{L^{-N}\psi(L)\}_{L\ge 1}$ is a Cauchy sequence by taking sufficient large M first and then L, which completes the first part of the proof.

Let us assume now that Φ is ergodic. Since the arguments are similar to those in the proof of Theorem 3.5 in [39], we only sketch main ideas. By the multidimensional ergodic theorem, we see that almost surely as $m \to \infty$,

$$\frac{1}{m^N}\beta_q^{r,s}\big(\mathbb{K}^\circ(\Phi_{\Lambda_{mM}})\big) = \frac{1}{m^N}\sum_{i=1}^{m^N}\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_M + c_i})\big) \to \mathbb{E}\big[\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_M})\big)\big],$$

and for $j = q, q+1$,

$$\frac{1}{m^N}\sum_{i=1}^{m^N} F_j(\Phi_{(\partial\Lambda_M)^{(\rho(s))}+c_i}, s) \to \mathbb{E}\big[F_j(\Phi_{(\partial\Lambda_M)^{(\rho(s))}}, s)\big] = O\big(M^{N-1}\big).$$

Remark here that the above equations hold for all except a countable set of M (cf. [32], Theorem 1). Therefore, it follows from (4.1) that

$$\limsup_{m\to\infty} \frac{\pm 1}{(mM)^N}\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_{mM}})\big) \le \frac{\pm 1}{M^N}\mathbb{E}\big[\beta_q^{r,s}\big(\mathbb{K}(\Phi_{\Lambda_M})\big)\big] + O\big(M^{-1}\big).$$

The rest of the proof is similar to the last step in the first part by noting that the following laws of large numbers for $F_j(\Phi_{\Lambda_L}, s)$, $j = q, q+1$, hold (cf. [39], Lemma 3.2),

$$\frac{F_j(\Phi_{\Lambda_L}, s)}{L^j} \to \hat{F}_j(s) \qquad \text{almost surely as } L \to \infty.$$

This completes the second part of the proof. $\square$

COROLLARY 4.1. *Let Φ be a stationary point process on $\mathbf{R}^N$ having all finite moments, and $\xi_{q,L}$ be the point process on Δ corresponding to the qth persistence diagram for $\mathbb{K}(\Phi_{\Lambda_L})$. Then, for every rectangle of the form $R = (r_1, r_2] \times (s_1, s_2], [0, r_1] \times (s_1, s_2] \subset \Delta$, there exists a constant $C_R \in [0, \infty)$ such that*

$$\frac{1}{L^N}\mathbb{E}\big[\xi_{q,L}(R)\big] \to C_R \qquad \textit{as } L \to \infty.$$

In addition, if Φ is ergodic, then

$$\frac{1}{L^N}\xi_{q,L}(R) \to C_R \qquad \textit{almost surely as } L \to \infty.$$

PROOF. It is a direct consequence of Theorem 1.11 because for $R = (r_1, r_2] \times (s_1, s_2]$,

$$\xi_{q,L}(R) = \beta_q^{r_2,s_1}(\mathbb{K}(\Phi_{\Lambda_L})) - \beta_q^{r_2,s_2}(\mathbb{K}(\Phi_{\Lambda_L})) + \beta_q^{r_1,s_2}(\mathbb{K}(\Phi_{\Lambda_L})) - \beta_q^{r_1,s_1}(\mathbb{K}(\Phi_{\Lambda_L})),$$

and for $R = [0, r_1] \times (s_1, s_2]$,

$$\xi_{q,L}(R) = \beta_q^{r_1,s_1}(\mathbb{K}(\Phi_{\Lambda_L})) - \beta_q^{r_1,s_2}(\mathbb{K}(\Phi_{\Lambda_L})). \qquad \square$$

4.2. *Proof of Theorem* 1.5. Let $S = \Delta = \{(x, y) \in \overline{\mathbf{R}}^2 : 0 \le x < y \le \infty\}$. Set

$$\mathscr{A} = \{(r_1, r_2] \times (s_1, s_2], [0, r_1] \times (s_1, s_2] \subset \Delta : 0 \le r_1 < r_2 \le s_1 < s_2 \le \infty\}.$$

We will show in Corollary A.3 that $\mathscr{A}$ is a convergence-determining class which satisfies the condition in Proposition 3.4. Theorem 1.5 then follows from Proposition 3.4 and Corollary 4.1. $\square$

DEFINITION 4.2. We call the limiting Radon measure $\nu_q \in \mathfrak{M}(\Delta)$ in Theorem 1.5 the qth persistence diagram for a stationary ergodic point process Φ on $\mathbf{R}^N$.

EXAMPLE 4.3. Let Φ be a randomly shifted $\mathbf{Z}^N$-lattice with intensity 1, that is, $\Phi = \mathbf{Z}^N + U$, where U is a uniform random variable on the unit cube $[0, 1]^N$. Then Φ is a stationary ergodic point process in $\mathbf{R}^N$. We compute the limiting persistence diagram ν_q of the Čech filtration $\mathbb{C}(\Phi) = \{C(\Phi, r)\}_{r \ge 0}$ for $q = 1, 2, \dots, N - 1$.

For this purpose, we introduce a filtration $\bar{\mathbb{C}}(L) = \{\bar{C}(L, r)\}_{r \ge 0}$ of cubical complexes by

$$\bar{C}(L, r) = \begin{cases} \mathrm{CL}(L, q), & \dfrac{\sqrt{q}}{2} \le r < \dfrac{\sqrt{q+1}}{2}, \\ \mathrm{CL}(L, N), & r \ge \dfrac{\sqrt{N}}{2}, \end{cases}$$

where $\mathrm{CL}(L, N)$ is the cubical complex consisting of all the elementary cubes in $[0, L] \times \cdots \times [0, L] \subset \mathbf{R}^N$, and $\mathrm{CL}(L, q)$ is the q-dimensional skeleton of $\mathrm{CL}(L, N)$. Here, a cube $Q = I_1 \times \cdots \times I_N \subset \mathbf{R}^N$ consisting of $I_k = [a, a]$ or $I_k = [a, a + 1]$ for some $a \in \mathbf{Z}$ is called an elementary cube [21]. From the stationarity of Φ and the homotopy equivalence between $\bar{C}(L, r)$ and $C(\mathbf{Z}^N \cap [0, L]^N, r)$, it suffices to compute the persistence diagram by using the filtration $\bar{\mathbb{C}}(L)$. We also note that $(\sqrt{q}/2, \sqrt{q+1}/2)$ is the only birth-death pair for the qth persistence diagram. Therefore, all we need to verify is the multiplicity of that pair with respect to L.

The Euler–Poincaré formula for $X = \mathrm{CL}(L, q)$ is given by

$$\sum_{k=0}^{q}(-1)^k|X_k| = \sum_{k=0}^{q}(-1)^k\beta_k(X). \tag{4.4}$$

The number $|X_k|$ of k-cells in X is given by (see, e.g., [17])

$$|X_k| = \sum_{p=k}^{N}\binom{p}{k}S_p(L),$$

where $S_p(x_1, \dots, x_N)$ is the elementary symmetric polynomial of degree p and $S_p(L)$ is an abbreviation for $S_p(L, \dots, L)$. On the other hand, since X is homotopy equivalent to a wedge sum of q-spheres, we have $\beta_0 = 1$ and $\beta_k = 0$ for $k = 1, \dots, q-1$. Then it follows from (4.4) that

$$\beta_q(X) = \sum_{k=0}^{q}(-1)^{k+q}\sum_{p=k}^{N}\binom{p}{k}S_p(L) + (-1)^{q+1},$$

and hence

$$\frac{\beta_q(X)}{L^N} = \sum_{k=0}^{q}(-1)^{k+q}\binom{N}{k} + O(L^{-1}) = \binom{N-1}{q} + O(L^{-1}).$$

Therefore, the limiting persistence diagram is given by

$$\nu_q = \binom{N-1}{q}\delta_{(\sqrt{q}/2, \sqrt{q+1}/2)}.$$

4.3. *The support of ν_q.* In this section, we give some sufficient conditions both on κ and Φ to ensure the positivity of the limiting measure ν_q. We use the following stability result on persistence diagrams of κ-filtrations (cf. [6, 8]). Here, the qth persistence diagram of the κ-filtration on $\Xi \in \mathscr{F}(\mathbf{R}^N)$ is simply denoted by $D_q(\Xi)$.

LEMMA 4.4. *Assume that κ is Lipschitz continuous with respect to the Hausdorff distance, that is, there exists a constant c_κ such that*

$$|\kappa(\sigma) - \kappa(\sigma')| \le c_\kappa d_H(\sigma, \sigma')$$

for $\sigma, \sigma' \in \mathscr{F}(\mathbf{R}^N)$. Then, for $\Xi, \Xi' \in \mathscr{F}(\mathbf{R}^N)$,

$$d_B(D_q(\Xi), D_q(\Xi')) \le c_\kappa d_H(\Xi, \Xi').$$

Here, d_B and d_H denote the bottleneck distance and the Hausdorff distance, respectively.

See Appendix C for the detail, where we recall definitions of d_B and d_H and give a proof of a generalization of this lemma.

Next, we introduce the notion of *marker* which is a finite point configuration for finding a specified point in Δ.

DEFINITION 4.5. Let Λ be a bounded Borel set in $\mathbf{R}^N$ and $(b,d) \in \Delta$. We say that $\Xi \in \mathscr{F}(\mathbf{R}^N)$ is a *(b,d)-marker in* Λ for the qth persistent homology (PH_q) if (i) $\Xi \subset \Lambda$ and (ii) for any $\Phi \in \mathscr{F}(\mathbf{R}^N)$

$$\xi_q(\Phi_{\Lambda^c} \sqcup \Xi)\big(\{(b,d)\}\big) \ge \xi_q(\Phi_{\Lambda^c})\big(\{(b,d)\}\big) + 1. \tag{4.5}$$

Here, Λ^c denotes the complement of Λ in $\mathbb{R}^N$. For a subset $A \subset \Delta$, we also say that Ξ is an A-marker in Λ if there exists $(b,d) \in A$ such that Ξ is a (b,d)-marker in Λ.

EXAMPLE 4.6. (i) Assume that a point $(b,d) \in \Delta$ is realizable by $\Xi \in \mathscr{F}(\mathbf{R}^N)$. Then there exists $M_0 > 0$ such that Ξ is a (b,d)-marker in Λ_M for any $M \ge M_0$ because Ξ is enough isolated from Λ_M^c for sufficiently large M.

(ii) We note that $(1/2, \sqrt{2}/2) \in \Delta$ is realized in PH_1 by

$$\big\{(0,0),(1,0),(0,1),(1,1)\big\} \in \mathscr{F}(\mathbf{R}^2)$$

in the Čech or Rips filtration. It is easy to see that for each $c \in \Lambda_1$, $\Xi_M = \sum_{x \in \mathbb{Z}^2 \cap \Lambda_M} \delta_{x+c}$ is a $(1/2, \sqrt{2}/2)$-marker in Λ_M for any sufficiently large M, for example, $M = 3$.

THEOREM 4.7. *Let*

$$A_{q,\varepsilon,(b,d)} := \bigcup_{M=1}^{\infty} \{\Phi_{\Lambda_M} \textit{ is a } B_\varepsilon\big((b,d)\big)\textit{-marker in } \Lambda_M \textit{ for } \mathrm{PH}_q\}$$

and

$$S_{q,\varepsilon} := \big\{(b,d) \in \Delta : \mathbb{P}(A_{q,\varepsilon,(b,d)}) > 0\big\}, \qquad S_q := \bigcap_{\varepsilon > 0} S_{q,\varepsilon}.$$

Then, $S_q \subset \operatorname{supp} \nu_q$.

Before proving Theorem 4.7, we give a lower bound for ν_q.

LEMMA 4.8. *For a closed set* $A \subset \Delta$,

$$\nu_q(A) \ge \frac{1}{M^N} \mathbb{P}(\Phi_{\Lambda_M} \textit{ is an } A\textit{-marker in } \Lambda_M \textit{ for } \mathrm{PH}_q). \tag{4.6}$$

PROOF. Let Λ be a bounded Borel set in $\mathbf{R}^N$ and $(b,d) \in \Delta$. If one could find disjoint subsets $\Lambda^{(1)}, \dots, \Lambda^{(k)} \subset \Lambda$ such that $\Phi_{\Lambda^{(i)}}$ is a (b,d)-marker in $\Lambda^{(i)}$ for each i, then

$$\xi_q(\Phi_\Lambda)\bigl(\{b,d\}\bigr) \ge k. \tag{4.7}$$

Indeed, by using (4.5) successively, we have

$$\xi_q(\Phi_\Lambda)\bigl(\{(b,d)\}\bigr) \ge \xi_q(\Phi_{\Lambda \setminus \bigcup_{j=1}^k \Lambda^{(j)}})\bigl(\{(b,d)\}\bigr) + k \ge k.$$

For $L > M > 0$ and $m = \lfloor L/M \rfloor$, we claim that

$$\xi_q(\Phi_{\Lambda_L})(A) \ge \sum_{i=1}^{m^N} \mathbf{1}_{\{\Phi_{\Lambda_M + c_i} \text{ is an } A\text{-marker in } \Lambda_M + c_i \text{ for } \mathrm{PH}_q\}}, \tag{4.8}$$

where $c_i \in \Lambda_L, i = 1, 2, \dots, m^N$ are chosen so that $\Lambda_L \supset \bigsqcup_{i=1}^{m^N}(\Lambda_M + c_i)$. If the right-hand side of (4.8) is equal to k, we have disjoint subsets $I_j \subset \{c_1, c_2, \dots, c_{m^N}\}$, $j = 1, 2, \dots, J$ with $\sum_{j=1}^J |I_j| = k$ such that for every $j = 1, 2, \dots, J$, $\Phi_{\Lambda_M + c}$ is a (b_j, d_j)-marker in $\Lambda_M + c$ for PH_q for any $c \in I_j$. Here, $(b_j, d_j) \in A$, $j = 1, 2, \dots, J$ are all distinct. From (4.7), we have

$$\xi_q(\Phi_{\Lambda_L})(A) \ge \sum_{j=1}^J \xi_q(\Phi_{\Lambda_L})\bigl(\{(b_j, d_j)\}\bigr) \ge \sum_{j=1}^J |I_j| = k,$$

which implies (4.8).

For a closed set $A \subset \Delta$, from (4.8), we obtain

$$\begin{aligned}
\nu_q(A) &\ge \limsup_{L\to\infty} \frac{1}{L^N} \mathbb{E}\bigl[\xi_q(\Phi_{\Lambda_L})\bigr](A) \\
&\ge \limsup_{L\to\infty} \frac{1}{L^N} \sum_{i=1}^{m^N} \mathbb{P}(\Phi_{\Lambda_M + c_i} \text{ is an } A\text{-marker in } \Lambda_M + c_i \text{ for } \mathrm{PH}_q) \\
&= \frac{1}{M^N} \mathbb{P}(\Phi_{\Lambda_M} \text{ is an } A\text{-marker in } \Lambda_M \text{ for } \mathrm{PH}_q).
\end{aligned}$$

This completes the proof. □

PROOF OF THEOREM 4.7. If $(b,d) \in S_q$, then for every $\varepsilon > 0$, there exists $M = M_\varepsilon \in \mathbb{N}$ such that

$$\mathbb{P}\bigl(\Phi_{\Lambda_M} \text{ is a } B_\varepsilon\bigl((b,d)\bigr)\text{-marker in } \Lambda_M \text{ for } \mathrm{PH}_q\bigr) > 0.$$

From (4.6), we see that $\nu_q(\bar{B}_\varepsilon((b,d))) > 0$ for any $\varepsilon > 0$, which implies $(b,d) \in \operatorname{supp} \nu_q$. Therefore, $S_q \subset \operatorname{supp} \nu_q$. □

For a bounded set $\Lambda \subset \mathbf{R}^N$, the restriction $\mathfrak{N}(\Lambda)$ of $\mathfrak{N}(\mathbf{R}^N)$ on Λ can be identified with $\bigcup_{k=0}^{\infty} \Lambda^k/\sim$, where $\sim$ is the equivalence relation induced by permutations on coordinates. Let Π be the probability distribution of homogeneous Poisson point process with unit intensity. It is clear that the local densities, which are sometimes called Janossy densities, of the restriction of Π on Λ are given by

$$\Pi|_\Lambda(dx_1 \dots dx_n) = \begin{cases} e^{-|\Lambda|}\, dx_1\, dx_2 \cdots dx_k & \text{on } \Lambda^k, \\ e^{-|\Lambda|} & \text{on } \Lambda^0 = \{\varnothing\}. \end{cases}$$

In other words, for a bounded measurable (local) function $f : \mathfrak{N}(\Lambda) \to \mathbf{R}$,

$$\mathbb{E}_\Pi[f] = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\Lambda^n} f(x_1, \dots, x_n) \Pi|_\Lambda(dx_1 \cdots dx_n).$$

For a probability measure Θ on $\mathfrak{N}(\mathbf{R}^N)$, if $\Theta|_\Lambda$ is absolutely continuous with respect to $\Pi|_\Lambda$ for a bounded set Λ, then $\Theta|_\Lambda$ is absolutely continuous with respect to the Lebesgue measure on each Λ^k for every k; thus the Radon–Nikodym density $d\Theta|_\Lambda/d\Pi|_\Lambda$ is defined a.e. on Λ^k for every k.

PROOF OF THEOREM 1.9. Assume that $(b, d) \in R_q$ and it is realizable by $\{y_1, \dots, y_m\}$. From continuity of persistence diagram in Lemma 4.4, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\xi_q(\{z_1, \dots, z_m\})(B_\varepsilon(\{(b, d)\})) \geq 1$ for any $(z_1, \dots, z_m) \in B_\delta(y_1) \times \cdots \times B_\delta(y_m)$ and the balls $\{B_\delta(y_i)\}_{i=1}^m$ are disjoint. From Example 4.6(i), there exists $M \in \mathbb{N}$ such that any $\{z_1, \dots, z_m\}$ is a $B_\varepsilon((b, d))$-marker in Λ_M. Hence, we see that

$$\begin{aligned} &\mathbb{P}\big(\Phi_{\Lambda_M} \text{ is a } B_\varepsilon\big((b, d)\big)\text{-marker in } \Lambda_M \text{ for } \mathrm{PH}_q\big) \\ &\quad \geq \Theta|_{\Lambda_M}\left(\bigcap_{i=1}^m \{\Phi(B_\delta(y_i) = 1)\} \cap \left\{\Phi\left(\Lambda_M \Big\backslash \bigcup_{i=1}^m B_\delta(y_i)\right) = 0\right\}\right) \\ &\quad = e^{-|\Lambda_M|} \int_{B_\delta(y_1) \times \cdots \times B_\delta(y_m)} f_{\Lambda_M}(z_1, \dots, z_m)\, dz_1 \cdots dz_m \\ &\quad > 0, \end{aligned}$$

where $f_{\Lambda_M} = d\Theta|_{\Lambda_M}/d\Pi|_{\Lambda_M}$. Hence, $R_q \subset S_q \subset \operatorname{supp} \nu_q$ by Theorem 4.7. Since $\operatorname{supp} \nu_q \subset \overline{R_q}$ as mentioned after Example 1.8, we conclude that $\operatorname{supp} \nu_q = \overline{R_q}$. □

Point processes are often specified by the local conditional distributions given a configuration outside, that is, there exists a measurable function $q_\Lambda : \mathfrak{N}(\Lambda) \times \mathfrak{N}(\Lambda^c)$, called a specification, for each bounded Borel set $\Lambda \in \mathscr{B}(\mathbf{R}^N)$ such that for every bounded measurable function f on $\mathfrak{N}(\mathbf{R}^N)$

$$\mathbb{E}_\Theta[f | \mathcal{F}_{\Lambda^c}](\xi) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Lambda^k} f(\mathbf{x} \cup \xi_{\Lambda^c}) q_\Lambda(\mathbf{x}, \xi_{\Lambda^c})\, d\mathbf{x} \qquad \Theta\text{-a.e.}\,\xi,$$

where $\mathcal{F}_{\Lambda^c}$ is the σ-field generated by the mappings $\xi \mapsto \Phi_\xi(A)$, $A \in \mathscr{B}(\Lambda^c)$ and $\mathbf{x} \cup \xi_{\Lambda^c}$ denotes $\delta_{x_1} + \cdots + \delta_{x_k} + \xi_{\Lambda^c}$ if $\mathbf{x} = (x_1, x_2, \dots, x_k)$. In this case, the local density of Θ is given by

$$\frac{d\Theta|_\Lambda}{d\Pi|_\Lambda}(\mathbf{x}) = e^{|\Lambda|} \cdot \mathbb{E}_\Theta\big[q_\Lambda(\mathbf{x}, \xi_{\Lambda^c})\big] \qquad \text{for } \mathbf{x} \in \bigcup_{k=0}^{\infty} \Lambda^k. \tag{4.9}$$

EXAMPLE 4.9. (1) The DLR equations due to Dobrushin–Lanford–Ruelle provide local conditional distributions of a Gibbs point process. In this formalism, a measurable function $U_\Lambda : \mathfrak{N}(\Lambda) \times \mathfrak{N}(\Lambda^c) \to (-\infty, \infty]$ is understood as the conditional energy of particles given a configuration outside of Λ, if it satisfies

$$q_\Lambda(\mathbf{x}, \xi_{\Lambda^c}) = Z_\Lambda(\xi)^{-1} e^{-U_\Lambda(\mathbf{x}, \xi_{\Lambda^c})},$$

where

$$Z_\Lambda(\xi) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Lambda^k} e^{-U_\Lambda(\mathbf{x}, \xi_{\Lambda^c})}\, d\mathbf{x}.$$

If there exists $B \geq 0$ such that $-Bk \leq U(\mathbf{x}, \xi_{\Lambda^c}) < \infty$ for any $k \geq 1$, $\mathbf{x} \in \Lambda^k$ and Θ-a.e. ξ, then it is easy to see that the local density satisfies the positivity condition in Theorem 1.9. If U is of hard-core type, that is, $U(\mathbf{x}, \xi_{\Lambda^c}) = \infty$ for all $\mathbf{x}$ in some open set almost surely, then the positivity condition fails.

(2) For the Ginibre point process and the zeros of the Gaussian entire function given in Example 1.10, Ghosh-Peres [13] showed that both processes exhibit the so-called "rigidity" meaning that for a bounded Borel set Λ there exists a nonnegative integer $N(\xi_{\Lambda^c}) \in \{0, 1, 2, \dots\}$ which is measurable with respect to $\mathcal{F}_{\Lambda^c}$ such that $q(\cdot, \xi_{\Lambda^c})$ is supported on $\Lambda^{N(\xi_{\Lambda^c})}$. Roughly speaking, the number of points inside Λ is determined from a given configuration outside of Λ. Ghosh [12], moreover, showed that there exist positive constants $m(\xi_{\Lambda^c})$ and $M(\xi_{\Lambda^c})$ such that almost surely

$$m(\xi_{\Lambda^c})\big|\Delta(\mathbf{x})\big|^2 \leq q_\Lambda(\mathbf{x}, \xi_{\Lambda^c}) \leq M(\xi_{\Lambda^c})\big|\Delta(\mathbf{x})\big|^2 \qquad \text{for a.e. } \mathbf{x} \in \Lambda^{N(\xi_{\Lambda^c})},$$

where $\Delta(\mathbf{x}) = \prod_{1 \leq i < j \leq k}(x_j - x_i)$ is the Vandermonde determinant. From this inequality and (4.9), we have, for any $k \geq 0$,

$$\frac{d\Theta|_\Lambda}{d\Pi|_\Lambda}(\mathbf{x}) \geq e^{|\Lambda|}\mathbb{E}_\Theta\big[m(\xi_{\Lambda^c}); N(\xi_{\Lambda^c}) = k\big] \cdot \big|\Delta(\mathbf{x})\big|^2 \qquad \text{for } \mathbf{x} \in \Lambda^k.$$

Since $\Theta(\xi(\Lambda) = k) > 0$ and $m(\xi_{\Lambda^c})$ is positive, the left-hand side is positive almost everywhere on $\bigcup_{k=0}^{\infty} \Lambda^k$.

We remark that the Ginibre point process is an important example of determinantal point processes. Determinantal (resp., permanental) point processes provides an important class of point processes that are negatively (resp., positively) correlated (cf. [20, 36]). In both cases, the local density can be expressed in terms of the so-called correlation kernel so that for a given kernel one can basically check whether the positivity condition is satisfied or not.

5. Central limit theorem for persistent Betti numbers. In this section, let $\Phi = \mathcal{P}$ be a homogeneous Poisson point process with unit intensity, and we prove Theorem 1.12. The idea is to apply a result in [31] which shows a central limit theorem for a certain class of functionals defined on Poisson point processes.

We here summarize necessary properties for functionals to achieve the central limit theorem. First of all, let us consider a sequence $\{W_n\}$ of Borel subsets in $\mathbf{R}^N$ satisfying the following conditions:

(A1) $|W_n| = n$ for all $n \in \mathbb{N}$;
(A2) $\bigcup_{n\ge 1} \bigcap_{m\ge n} W_m = \mathbf{R}^N$;
(A3) $\lim_{n\to\infty} |(\partial W_n)^{(r)}|/n = 0$ for all $r > 0$;
(A4) there exists a constant $\gamma > 0$ such that $\mathrm{diam}(W_n) \le \gamma n^{\gamma}$.

Given such a sequence, let $\mathcal{W} = \mathcal{W}(\{W_n\})$ be the collection of all subsets A in $\mathbf{R}^N$ of the form $A = W_n + x$ for some W_n in the sequence and some point $x \in \mathbf{R}^N$.

Let H be a real-valued functional defined on $\mathscr{F}(\mathbf{R}^N)$. The functional H is said to be *translation invariant* if it satisfies $H(\mathcal{X} + y) = H(\mathcal{X})$ for any $\mathcal{X} \in \mathscr{F}(\mathbf{R}^N)$ and $y \in \mathbf{R}^N$. Let D_0 be the add one cost function

$$D_0 H(\mathcal{X}) = H\big(\mathcal{X} \cup \{0\}\big) - H(\mathcal{X}), \qquad \mathcal{X} \in \mathscr{F}(\mathbf{R}^N),$$

which is the increment in H caused by inserting a point at the origin. The functional H is *weakly stabilizing* on $\mathcal{W}$ if there exists a random variable $D(\infty)$ such that $D_0 H(\mathcal{P}_{A_n}) \xrightarrow{\text{a.s.}} D(\infty)$ as $n \to \infty$ for any sequence $\{A_n \in \mathcal{W}\}_{n\ge 1}$ tending to $\mathbf{R}^N$. The *Poisson bounded moment condition* on $\mathcal{W}$ is given by

$$\sup_{0\in A\in\mathcal{W}} \mathbb{E}\big[\big(D_0 H(\mathcal{P}_A)\big)^4\big] < \infty.$$

Then we restate Theorem 3.1 in [31] in the following form.

LEMMA 5.1 ([31], Theorem 3.1). *Let H be a real-valued functional defined on $\mathscr{F}(\mathbf{R}^N)$. Assume that H is translation invariant and weakly stabilizing on $\mathcal{W}$, and satisfies the Poisson bounded moment condition. Then there exists a constant $\sigma^2 \in [0, \infty)$ such that $n^{-1}\,\mathrm{Var}[H(\mathcal{P}_{W_n})] \to \sigma^2$ and*

$$\frac{H(\mathcal{P}_{W_n}) - \mathbb{E}[H(\mathcal{P}_{W_n})]}{n^{1/2}} \xrightarrow{d} \mathcal{N}\big(0, \sigma^2\big) \qquad \text{as } n \to \infty.$$

By using Lemma 5.1, we prove the following theorem.

THEOREM 5.2. *Let $\Phi = \mathcal{P}$ be a homogeneous Poisson point process with unit intensity. Assume that the sequence $\{W_n\}$ satisfies* (A1)–(A4). *Then for any $0 \le r \le s < \infty$,*

$$\frac{\beta_q^{r,s}(\mathbb{K}(\mathcal{P}_{W_n})) - \mathbb{E}[\beta_q^{r,s}(\mathbb{K}(\mathcal{P}_{W_n}))]}{n^{1/2}} \xrightarrow{d} \mathcal{N}\big(0, \sigma_{r,s}^2\big) \qquad \text{as } n \to \infty.$$

In particular, Theorem 1.12 is derived from this theorem by taking $W_n = \Lambda_{L_n}$ with $L_n = n^{1/N}$.

For the proof of Theorem 5.2, the essential part is to show the weak stabilization of the persistent Betti number $\beta_q^{r,s}(\mathbb{K}(\cdot))$ as a functional on $\mathscr{F}(\mathbf{R}^N)$, on which we focus below.

We remark that, for almost surely, the Poisson point process $\mathcal{P}$ consists of infinite points in $\mathbf{R}^N$ which do not have accumulation points. In view of this property, we first show a stabilization of persistent Betti numbers in the following deterministic setting.

LEMMA 5.3. *Let P be a set of points in $\mathbf{R}^N$ without accumulation points. Then, for each fixed $r \le s$, there exist constants D_∞ and $R > 0$ such that*

$$D_0\beta_q^{r,s}\big(\mathbb{K}(P_{\bar{B}_a(0)})\big) = D_\infty$$

for all $a \ge R$.

PROOF. Let $P' = P \cup \{0\}$. Let $K_{r,a} = K(P_{\bar{B}_a(0)}, r)$ be the simplicial complex defined on $P_{\bar{B}_a(0)}$ with parameter r, and similarly let $K'_{r,a} = K(P'_{\bar{B}_a(0)}, r)$.

From the definition (2.3), $D_0\beta_q^{r,s}(\mathbb{K}(P_{\bar{B}_a(0)}))$ can be expressed as

$$\begin{aligned} &D_0\beta_q^{r,s}\big(\mathbb{K}(P_{\bar{B}_a(0)})\big)\\ &\quad= \dim \frac{Z_q(K'_{r,a})}{Z_q(K'_{r,a}) \cap B_q(K'_{s,a})} - \dim \frac{Z_q(K_{r,a})}{Z_q(K_{r,a}) \cap B_q(K_{s,a})}\\ &\quad= \big(\dim Z_q\big(K'_{r,a}\big) - \dim Z_q(K_{r,a})\big)\\ &\qquad - \big(\dim Z_q\big(K'_{r,a}\big) \cap B_q\big(K'_{s,a}\big) - \dim Z_q(K_{r,a}) \cap B_q(K_{s,a})\big). \end{aligned}$$

Hence, it suffices to show the stabilization with respect to a for $\dim Z_q(K_{r,a})$ and $\dim(Z_q(K_{r,a}) \cap B_q(K_{s,a}))$ separately.

Let us study $\dim Z_q(K_{r,a})$. Since the dimension takes nonnegative integer values, we show the bounded and the nondecreasing properties. First of all, note that $K_{r,a} \subset K'_{r,a}$, and hence $Z_q(K_{r,a}) \subset Z_q(K'_{r,a})$. Let us express $K'_{r,a}$ as a disjoint union $K'_{r,a} = K_{r,a} \sqcup K^0_{r,a}$, where $K^0_{r,a}$ is the set of simplices having the point 0, and let $K^0_{r,a,q} = \{\sigma \in (K'_{r,a})_q : 0 \in \sigma\}$.

Let $\partial_{q,a}$ and $\partial'_{q,a}$ be the qth boundary maps on $K_{r,a}$ and $K'_{r,a}$, respectively. Then we can obtain the following block matrix form

$$\partial'_{q,a} = \begin{bmatrix} M_{1,\rho} & \mathbf{0} \\ M_{2,\rho} & \partial_{q,a} \end{bmatrix}, \tag{5.1}$$

where the first columns and rows are arranged by the simplices in $K^0_{r,a,q}$ and $K^0_{r,a,q-1}$, and the second columns and rows correspond to the simplices in $K_{r,a}$.

Recall that any simplex $\sigma \in K(P,r)$ containing the point 0 is included in $\bar{B}_{\rho(r)}(0)$. Hence, the set $K^0_{r,a,q}$ becomes independent of a for $a \geq \rho(r)$, which we denote by $K^0_{r,*,q}$. From this observation and Lemma 2.9 applied to the matrix form (5.1), we have

$$\dim Z_q(K'_{r,a}) - \dim Z_q(K_{r,a}) \leq |K^0_{r,a,q}| = |K^0_{r,*,q}|,$$

which gives the boundedness.

In order to show the nondecreasing property, let us consider a homomorphism defined by

$$f: \frac{Z_q(K'_{r,a_1})}{Z_q(K_{r,a_1})} \ni [c] \quad \longmapsto \quad [c] \in \frac{Z_q(K'_{r,a_2})}{Z_q(K_{r,a_2})}$$

for $a_1 \leq a_2$. This map is well defined because $Z_q(K_{r,a_1}) \subset Z_q(K_{r,a_2})$ and $Z_q(K'_{r,a_1}) \subset Z_q(K'_{r,a_2})$ hold. Suppose that $f[c] = 0$. Then the cycle $c \in Z_q(K'_{r,a_1})$ is in $Z_q(K_{r,a_2})$. It means that the q-simplices consisting of c do not contain the point 0, and hence $c \in Z_q(K_{r,a_1})$. This shows that the map f is injective. From this observation, we have the inequality

$$\dim Z_q(K'_{r,a_1})/Z_q(K_{r,a_1}) \leq \dim Z_q(K'_{r,a_2})/Z_q(K_{r,a_2}),$$

which leads to the desired nondecreasing property. This completes the proof of the stabilization of $\dim Z_q(K_{r,a})$.

Let us study the stabilization of $\dim(Z_q(K_{r,a}) \cap B_q(K_{s,a}))$. The strategy is basically the same as above. It follows from Lemma 2.8 that

$$\dim \frac{Z_q(K'_{r,a}) \cap B_q(K'_{s,a})}{Z_q(K_{r,a}) \cap B_q(K_{s,a})} \leq \dim \frac{Z_q(K'_{r,a})}{Z_q(K_{r,a})} + \dim \frac{B_q(K'_{s,a})}{B_q(K_{s,a})}.$$

Then, from the same reasoning used in $\dim Z_q(K_{r,a})$, we have the stabilization $|K^0_{s,a,q+1}| = |K^0_{s,*,q+1}|$ for large a. Hence, we have the boundedness

$$\dim Z_q(K'_{r,a}) \cap B_q(K'_{s,a}) - \dim Z_q(K_{r,a}) \cap B_q(K_{s,a}) \leq |K^0_{r,*,q}| + |K^0_{s,*,q+1}|.$$

Similarly, for sufficiently large $a_1 \leq a_2$, we can show the injectivity of the map

$$f: \frac{Z_q(K'_{r,a_1}) \cap B_q(K'_{s,a_1})}{Z_q(K_{r,a_1}) \cap B_q(K_{s,a_1})} \longrightarrow \frac{Z_q(K'_{r,a_2}) \cap B_q(K'_{s,a_2})}{Z_q(K_{r,a_2}) \cap B_q(K_{s,a_2})}, \qquad f[c] = [c],$$

from which the nondecreasing property follows. This completes the proof of the lemma. □

PROPOSITION 5.4. *The functional $\beta_q^{r,s}(\mathbb{K}(\cdot))$ is weakly stabilizing.*

PROOF. Let $R > 0$ be chosen as in Lemma 5.3 and let $\{A_n \in \mathcal{W}\}_{n \geq 1}$ be a sequence tending to $\mathbf{R}^N$. Then there exists $n_0 \in \mathbb{N}$ such that $B_R(0) \subset A_n$ for all $n \geq n_0$.

For $n \ge n_0$, let us set $L_{r,n} = K(\mathcal{P}_{A_n}, r)$. Then, since A_n is bounded, there exists $a > R$ such that

$$B_R(0) \subset A_n \subset B_a(0).$$

Then, as in the same way used for showing the injectivity in the proof of Lemma 5.3, we can show

$$\frac{Z_q(K'_{r,R})}{Z_q(K_{r,R})} \subset \frac{Z_q(L'_{r,n})}{Z_q(L_{r,n})} \subset \frac{Z_q(K'_{r,a})}{Z_q(K_{r,a})},$$

where $K_{r,a} = K(\mathcal{P}_{\bar{B}_a(0)}, r)$ as before. Since the dimensions of $Z_q(K'_{r,R})/Z_q(K_{r,R})$ and $Z_q(K'_{r,a})/Z_q(K_{r,a})$ are equal, for all $n \ge n_0$

$$\dim Z_q(K'_{r,R}) - \dim Z_q(K_{r,R}) = \dim Z_q(L'_{r,n}) - \dim Z_q(L_{r,n}).$$

We can also show that $\dim Z_q(L'_{r,n}) \cap B_q(L'_{s,n}) - \dim Z_q(L_{r,n}) \cap B_q(L_{s,n})$ is invariant for $n \ge n_0$ in a similar manner. This completes the proof. □

PROOF OF THEOREM 5.2. For fixed $r \le s$, we regard the persistent Betti number $\beta_q^{r,s}(\mathbb{K}(\cdot))$ as a functional on $\mathscr{F}(\mathbf{R}^N)$, and check the three conditions stated in Lemma 5.1. First, the translation invariance is obvious, because κ is translation invariant. Next, let us consider the Poisson bounded moment condition on $\mathcal{W}$. We note the following estimate:

$$\begin{aligned} |D_0\beta_q^{r,s}(\mathbb{K}(\mathcal{P}_A))| &= |\beta_q^{r,s}(\mathbb{K}(\mathcal{P}_A \cup \{0\})) - \beta_q^{r,s}(\mathbb{K}(\mathcal{P}_A))| \\ &\le \sum_{j=q,q+1} |K_j(\mathcal{P}_A \cup \{0\}, s) \setminus K_j(\mathcal{P}_A, s)| \\ &\le \sum_{j=q,q+1} F_j(\mathcal{P}_{\bar{B}_{\rho(s)}(0)}, s). \end{aligned}$$

Here, the second inequality follows from Lemma 2.11. Then the fourth moment is uniformly bounded because of the finiteness of moments of the Poisson point process on $\bar{B}_{\rho(s)}(0)$. We showed the weak stabilization in Proposition 5.4. The proof of Theorem 5.2 is now complete. □

6. Conclusions. In this paper, we studied a convergence of persistence diagrams and persistent Betti numbers for stationary point processes, and a central limit theorem of persistent Betti numbers for homogeneous Poisson point process. Several important problems are still yet to be solved:

1. We showed the existence of limiting persistence diagram for simplicial complexes built over stationary ergodic point processes. Such convergence results can be expected for more general random simplicial/cell complexes studied in [17, 18]. It would also be important to investigate the rate of convergence from the statistical and computational point of view.

2. Attractiveness/repulsiveness of point processes are reflected on persistence diagrams (see Figure 1). For example, the mass of the limiting persistence diagram ν_q for negatively correlated point process seems to become more concentrated than that for positively correlated point process.

3. The moments of the limiting persistence diagram, $\int_\Delta |y-x|^n \nu_q(dx\,dy)$, should be studied. Other properties of limiting persistence diagrams such as continuity, absolute continuity/singularity, comparison, etc. should also be investigated thoroughly for practical purposes (cf. [16, 26]).

4. The central limit theorem for persistent Betti numbers (even for usual Betti numbers) is only proved for Poisson point processes. It could be extended to more general stationary point processes. We also expect that a scaled persistence diagram converges to a Gaussian field on Δ.

APPENDIX A: CONVERGENCE-DETERMINING CLASS FOR VAGUE CONVERGENCE

We provide a sufficient condition for a class of $\mathscr{B}$-sets to be a convergence determining class for vague convergence. We use the same notation as in Section 3. Assume that a class $\mathscr{A} \subset \mathscr{B}$ is closed under finite intersections. Let us define

$$\mathcal{R}(\mathscr{A}) = \left\{ \bigcup_{\text{finite}} A_i : A_i \in \mathscr{A} \right\}.$$

Then $\mathcal{R}(\mathscr{A})$ is closed under both finite intersections and finite unions. Furthermore, if $\mu_n(A) \to \mu(A)$ for all $A \in \mathscr{A}$, then so does for all $A \in \mathcal{R}(\mathscr{A})$, because

$$\mu\left(\bigcup_{i=1}^m A_i\right) = \sum_i \mu(A_i) - \sum_{i \neq j} \mu(A_i \cap A_j) + \cdots + (-1)^{m-1} \mu\left(\bigcap_{i=1}^m A_i\right).$$

LEMMA A.1. *Assume that a class $\mathscr{A}$ is closed under finite intersections, and that*:

(i) *each open set $G \in \mathscr{B}$ is a countable union of $\mathcal{R}(\mathscr{A})$-sets, and*
(ii) *each closed set $F \in \mathscr{B}$ is a countable intersection of $\mathcal{R}(\mathscr{A})$-sets.*

If $\mu_n(A) \to \mu(A)$ for all $A \in \mathscr{A}$, then μ_n converges vaguely to μ. In particular, the class $\mathscr{A}$ is a convergence-determining class for μ provided that $\mathscr{A} \subset \mathscr{B}_\mu$.

PROOF. Let $G \in \mathscr{B}$ be an open set. By assumption, there are sets $A_i \in \mathcal{R}(\mathscr{A})$ such that

$$G = \bigcup_{i=1}^\infty A_i.$$

Given $\varepsilon > 0$, choose an m such that

$$\mu\left(\bigcup_{i=1}^{m} A_i\right) > \mu(G) - \varepsilon.$$

Then we have

$$\mu(G) - \varepsilon < \mu\left(\bigcup_{i=1}^{m} A_i\right) = \lim_{n\to\infty} \mu_n\left(\bigcup_{i=1}^{m} A_i\right) \le \liminf_{n\to\infty} \mu_n(G).$$

Since ε is arbitrary, we get

$$\mu(G) \le \liminf_{n\to\infty} \mu_n(G).$$

Now for a closed set $F \in \mathscr{B}$, take $A_i \in \mathcal{R}(\mathscr{A})$ such that

$$F = \bigcap_{i=1}^{\infty} A_i.$$

Since $A_i \in \mathscr{B}$, for given $\varepsilon > 0$, we can choose m large enough such that

$$\mu(F) + \varepsilon > \mu\left(\bigcap_{i=1}^{m} A_i\right).$$

Then it follows from $\bigcap_{i=1}^{m} A_i \in \mathcal{R}(\mathscr{A})$ that

$$\mu(F) + \varepsilon > \mu\left(\bigcap_{i=1}^{m} A_i\right) = \lim_{n\to\infty} \mu_n\left(\bigcap_{i=1}^{m} A_i\right) \ge \limsup_{n\to\infty} \mu_n(F).$$

Letting $\varepsilon \to 0$, we get

$$\mu(F) \ge \limsup_{n\to\infty} \mu_n(F).$$

Therefore, the conclusion follows from Lemma 3.1. □

For given $\mathscr{A}$, let $\mathscr{A}_{x,\varepsilon}$ be the class of $\mathscr{A}$-sets satisfying $x \in A^\circ \subset A \subset B_\varepsilon(x)$, where A° is the interior of A. Let $\partial\mathscr{A}_{x,\varepsilon}$ be the class of their boundaries, that is, $\partial\mathscr{A}_{x,\varepsilon} = \{\partial A : A \in \mathscr{A}_{x,\varepsilon}\}$.

The following theorem gives a sufficient condition for a class $\mathscr{A}$ to be a convergence-determining class for vague convergence of Radon measures (see Theorem 2.4 in [1] for an analogous result on weak convergence of probability measures).

THEOREM A.2. *Suppose that $\mathscr{A}$ is closed under finite intersections and, for each $x \in S$ and $\varepsilon > 0$, $\partial\mathscr{A}_{x,\varepsilon}$ contains either $\varnothing$ or uncountably many disjoint sets. Then $\mathscr{A}$ is a convergence-determining class. Moreover, for any measure $\mu \in \mathfrak{M}$, $\mathscr{A}$ contains a countable convergence-determining class for μ.*

PROOF. Fix an arbitrary $\mu \in \mathfrak{M}$, and let $\mathscr{A}_\mu = \mathscr{A} \cap \mathscr{B}_\mu$ be the class of μ-continuity sets in $\mathscr{A}$. Since

$$\partial(A \cap B) \subset (\partial A) \cup (\partial B),$$

$\mathscr{A}_\mu$ is again closed under finite intersections.

Let $G \in \mathscr{B}$ be an open set. For $x \in G$, choose $\varepsilon > 0$ such that $B_\varepsilon(x) \subset G$. By the assumption, if $\partial \mathscr{A}_{x,\varepsilon}$ does not contain $\varnothing$, then it must contain uncountably many disjoint sets. Hence, in either case, $\partial \mathscr{A}_{x,\varepsilon}$ contains a set ∂A_x of μ-measure 0, or $A_x \in \mathscr{A}_\mu$. Therefore, G can be written as

$$G = \bigcup_{x \in G} A_x^\circ = \bigcup_{x \in G} A_x.$$

Since S is a separable metric space, there is a countable subcollection $\{A_{x_i}^\circ\}$ of $\{A_x^\circ : x \in G\}$ which covers G, namely,

$$G = \bigcup_{i=1}^\infty A_{x_i}^\circ.$$

Let $\{G_i\}_{i=1}^\infty$ be a countable basis of S. For each i, we have just shown that there are countable sets $\{A_{i,j}\}_{j=1}^\infty \subset \mathscr{A}_\mu$ such that

$$G_i = \bigcup_{j=1}^\infty A_{i,j}^\circ = \bigcup_{j=1}^\infty A_{i,j}.$$

Set

$$\mathscr{A}'_\mu = \Big\{ \bigcap_{\text{finite}} A_{i,j} \Big\}.$$

Then $\mathscr{A}'_\mu \subset \mathscr{A}_\mu$ is countable and closed under finite intersections. The remaining task is to show that $\mathscr{A}'_\mu$ satisfies the two conditions in Lemma A.1. The condition for open sets is clear from the construction of $\mathscr{A}'_\mu$.

Next, let $F \in \mathscr{B}$ be a closed (thus compact) set. For each $\varepsilon > 0$, let

$$F^{(\varepsilon)} = \Big\{ x \in S : d(x, F) = \inf_{y \in F} \rho(x, y) \le \varepsilon \Big\}.$$

Then $F = \bigcap_{p=1}^\infty F^{(\frac{1}{p})}$. We claim that, for each $\varepsilon > 0$, there exist $m = m(\varepsilon)$ and a collection of sets $\{C_k\}_{k=1}^m \subset \mathscr{A}'_\mu$ such that

$$F \subset \bigcup_{k=1}^m C_k \subset F^{(\varepsilon)}.$$

Indeed, for each $x \in F$, there is a pair (i_x, j_x) such that $x \in A_{i_x,j_x}^\circ \subset A_{i_x,j_x} \subset G_{i_x} \subset B_\varepsilon(x)$. Let $C_x = A_{i_x,j_x}$. Then

$$F \subset \bigcup_{x \in F} C_x^\circ.$$

Since F is compact, there is a finite collection $\{C^\circ_{x_k}\}_{k=1}^m$ such that

$$F \subset \bigcup_{k=1}^{m} C^\circ_{x_k}.$$

Finally, note that $C^\circ_{x_k} \subset C_{x_k} \subset F^{(\varepsilon)}$, we have

$$F \subset \bigcup_{k=1}^{m} C_{x_k} \subset F^{(\varepsilon)}.$$

Therefore, the condition for closed sets in Lemma A.1 is satisfied, which completes the proof of Theorem A.2. □

COROLLARY A.3. *The class*

$$\mathscr{A} = \{(r_1, r_2] \times (s_1, s_2], [0, r_2] \times (s_1, s_2] \subset \Delta : 0 \le r_1 \le r_2 \le s_1 \le s_2 \le \infty\}$$

satisfies the conditions of Proposition 3.4, namely, for any measure μ, it contains a countable convergence determining class for μ.

PROOF. It suffices to check the conditions in Theorem A.2. It is clear that $\mathscr{A}$ is closed under finite intersection and $\partial \mathscr{A}_{x,\varepsilon}$ contains uncountably many disjoint sets for any $x \in \Delta$ and $\varepsilon > 0$. □

APPENDIX B: SIMPLICIAL COMPLEX AND HOMOLOGY

B.1. Simplicial complex. We first introduce a combinatorial object called simplicial complex. Let $P = \{1, \dots, n\}$ be a finite set (not necessary to be points in a metric space). A *simplicial complex* with the vertex set P is defined by a collection K of nonempty subsets in P satisfying the following properties:

(i) $\{i\} \in K$ for $i = 1, \dots, n$, and
(ii) if $\sigma \in K$ and $\emptyset \ne \tau \subset \sigma$, then $\tau \in K$.

Each subset σ with $q + 1$ vertices is called a q-simplex. We denote the set of q-simplices by K_q. A subcollection $T \subset K$ which also becomes a simplicial complex is called a subcomplex of K.

EXAMPLE B.1. Figure 5 shows two polyhedra of simplicial complexes:

$$K = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\},$$
$$T = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

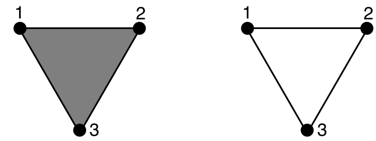

FIG. 5. *The polyhedra of the simplicial complexes K (left) and T (right).*

B.2. Homology. The procedure to define homology is summarized as follows:

1. Given a simplicial complex K, build a chain complex $C_*(K)$. This is an algebraization of K characterizing the boundary.
2. Define homology by quotienting out certain subspaces in $C_*(K)$ characterized by the boundary.

We begin with the procedure 1 by assigning orientations on simplices. When we deal with a q-simplex $\sigma = \{i_0, \dots, i_q\}$ as an ordered set, there are $(q+1)!$ orderings on σ. For $q > 0$, we define an equivalence relation $i_{j_0}, \dots, i_{j_q} \sim i_{\ell_0}, \dots, i_{\ell_q}$ on two orderings of σ such that they are mapped to each other by even permutations. By definition, two equivalence classes exist, and each of them is called an oriented simplex. An oriented simplex is denoted by $\langle i_{j_0}, \dots, i_{j_q}\rangle$, and its opposite orientation is expressed by adding the minus $-\langle i_{j_0}, \dots, i_{j_q}\rangle$. We write $\langle\sigma\rangle = \langle i_{j_0}, \dots, i_{j_q}\rangle$ for the equivalence class including $i_{j_0} < \cdots < i_{j_q}$. For $q = 0$, we suppose that we have only one orientation for each vertex.

Let $\mathbf{F}$ be a field. We construct a $\mathbf{F}$-vector space $C_q(K)$ as

$$C_q(K) = \mathrm{Span}_{\mathbf{F}}\{\langle\sigma\rangle \mid \sigma \in K_q\}$$

for $K_q \neq \varnothing$ and $C_q(K) = 0$ for $K_q = \varnothing$. Here, $\mathrm{Span}_{\mathbf{F}}(A)$ for a set A is a vector space over $\mathbf{F}$ such that the elements of A formally form a basis of the vector space. Furthermore, we define a linear map called the *boundary map* $\partial_q : C_q(K) \to C_{q-1}(K)$ by the linear extension of

$$\partial_q \langle i_0, \dots, i_q\rangle = \sum_{\ell=0}^{q} (-1)^\ell \langle i_0, \dots, \widehat{i_\ell}, \dots, i_q\rangle, \tag{B.1}$$

where $\widehat{i_\ell}$ means the removal of the vertex i_ℓ. We can regard the linear map ∂_q as algebraically capturing the $(q-1)$-dimensional boundary of a q-dimensional object.

For example, the image of the 2-simplex $\langle\sigma\rangle = \langle 1, 2, 3\rangle$ is given by $\partial_2\langle\sigma\rangle = \langle 2, 3\rangle - \langle 1, 3\rangle + \langle 1, 2\rangle$, which is the boundary of σ (see Figure 5).

In practice, by arranging some orderings of the oriented q- and $(q-1)$-simplices, we can represent the boundary map as a matrix

$$M_q = (M_{\sigma,\tau})_{\sigma \in K_{q-1}, \tau \in K_q}$$

with the entry $M_{\sigma,\tau} = 0, \pm 1$ given by the coefficient in (B.1). For the simplicial complex K in Example B.1, the matrix representations M_1 and M_2 of the boundary maps are given by

$$M_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \qquad M_1 = \begin{bmatrix} -1 & 0 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \tag{B.2}$$

Here, the 1-simplices (resp. 0-simplices) are ordered by $\langle 1, 2\rangle$, $\langle 2, 3\rangle$, $\langle 1, 3\rangle$ (resp., $\langle 1\rangle$, $\langle 2\rangle$, $\langle 3\rangle$).

We call a sequence of the vector spaces and linear maps

$$\cdots \longrightarrow C_{q+1}(K) \xrightarrow{\partial_{q+1}} C_q(K) \xrightarrow{\partial_q} C_{q-1}(K) \longrightarrow \cdots$$

the *chain complex* $C_*(K)$ of K. As an easy exercise, we can show $\partial_q \circ \partial_{q+1} = 0$ for every q. Hence, the subspaces $Z_q(K) = \ker \partial_q$ and $B_q(K) = \operatorname{im} \partial_{q+1}$ satisfy $B_q(K) \subset Z_q(K)$. Then the *qth (simplicial) homology* is defined by taking the quotient space

$$H_q(K) = Z_q(K)/B_q(K).$$

Intuitively, the dimension of $H_q(K)$ counts the number of q-dimensional holes in K and each generator of the vector space $H_q(K)$ corresponds to these holes. We remark that the homology as a vector space is independent of the orientations of simplices.

For a subcomplex T of K, the inclusion map $\iota \colon T \hookrightarrow K$ naturally induces a linear map in homology $\iota_q \colon H_q(T) \to H_q(K)$. Namely, an element $[c] \in H_q(T)$ is mapped to $[c] \in H_q(K)$, where the equivalence class $[c]$ is taken in each vector space.

For example, the simplicial complex K in Example B.1 has

$$Z_1(K) = \mathrm{Span}_{\mathbf{F}}[1\ 1\ -1]^T = B_1(K)$$

from (B.2). Hence, $H_1(K) = 0$, meaning that there are no 1-dimensional hole (ring) in K. On the other hand, since $Z_1(T) = Z_1(K)$ and $B_1(T) = 0$, we have $H_1(T) \simeq \mathbf{F}$, meaning that T consists of one ring. Hence, the induced linear map $\iota_1 \colon H_1(T) \to H_1(K)$ means that the ring in T disappears in K under $T \hookrightarrow K$.

APPENDIX C: CONTINUITY OF PERSISTENCE DIAGRAMS OF κ-COMPLEXES

We give a stability result for persistence diagrams of κ-filtrations which extends the stability result obtained in [6]. The notation used here follows the paper [6]. We first recall the definition of the Hausdorff distance and the bottleneck distance. The Hausdorff distance d_H on $\mathscr{F}(\mathbf{R}^N)$ for $\sigma, \sigma' \in \mathscr{F}(\mathbf{R}^N)$ is given by

$$d_H(\sigma, \sigma') = \max\Big\{\max_{x\in\sigma} \inf_{x'\in\sigma'} \|x - x'\|, \max_{x'\in\sigma'} \inf_{x\in\sigma} \|x - x'\|\Big\}.$$

We define the ℓ_∞-metric on Δ by $d_\infty((b_1, d_1), (b_2, d_2)) = \max(|b_1 - b_2|, |d_1 - d_2|)$, where $\infty - \infty = 0$. For $(b, d) \in \Delta$, we define $d_\infty((b, d), \partial\Delta) = (d - b)/2$. For finite multisets X and Y in Δ, a partial matching between X and Y is a subset $M \subset X \times Y$ such that for every $x \in X$ there is at most one $y \in Y$ such that $(x, y) \in M$ and for every $y \in Y$ there is at most one $x \in X$ such that $(x, y) \in M$. An $x \in X$ (resp., $y \in Y$) is unmatched if there is no $y \in Y$ (resp., $x \in X$) such that $(x, y) \in M$. We say that a partial matching M is δ-matching if $d_\infty(x, y) \le \delta$ for every $(x, y) \in M$, $d_\infty(x, \partial\Delta) \le \delta$ if $x \in X$ is unmatched, and $d_\infty(y, \partial\Delta) \le \delta$ if $y \in Y$ is unmatched.

The bottleneck distance is defined as follows:

$$d_B(X, Y) := \inf\{\delta > 0 : \text{there exists a } \delta\text{-matching between } X \text{ and } Y\}.$$

For $\Xi, \Xi' \in \mathscr{F}(\mathbf{R}^N)$ and $\kappa, \kappa' : \mathscr{F}(\mathbf{R}^N) \to [0, \infty]$, we define two complexes

$$\mathbb{K}_\kappa(\Xi) = \{K_\kappa(\Xi, t)\}_{t\ge0}, \qquad \mathbb{K}_{\kappa'}(\Xi') = \{K_{\kappa'}(\Xi', t)\}_{t\ge0}.$$

Let C be a *correspondence* between Ξ and Ξ', that is, $C \subset \Xi \times \Xi'$ such that $p_1(C) = \Xi$ and $p_2(C) = \Xi'$, where p_i is the projection onto the ith coordinate for $i = 1, 2$. We define the transpose C^T of C, which is also a correspondence, by

$$C^T := \{(x', x) \in \Xi' \times \Xi : (x, x') \in C\}.$$

A correspondence C defines a map from $\mathscr{F}(\Xi)$ to $\mathscr{F}(\Xi')$ as

$$C(\sigma) = \{x' \in \Xi' : (x, x') \in C, x \in \sigma\}.$$

The distortion of C is defined as

$$\operatorname{dis}(C) := \max\Big\{\sup_{\sigma\subset\Xi} |\kappa(\sigma) - \kappa'(C(\sigma))|, \sup_{\sigma'\subset\Xi'} |\kappa(C^T(\sigma')) - \kappa'(\sigma')|\Big\}.$$

LEMMA C.1. *If* $\operatorname{dis}(C) \le \varepsilon$, *then* $H_q(\mathbb{K}_\kappa(\Xi))$ *and* $H_q(\mathbb{K}_{\kappa'}(\Xi'))$ *are* ε-*interleaving*.

PROOF. Assume that $\sigma \in K_\kappa(\Xi, t)$ and $\kappa(\sigma) \le t$. Then it follows from $|\kappa(\sigma) - \kappa'(C(\sigma))| \le \varepsilon$ that

$$\kappa'(\sigma') \le \kappa'(C(\sigma)) \le \kappa(\sigma) + \varepsilon \le t + \varepsilon \qquad \text{for any } \sigma' \subset C(\sigma),$$

which implies $\sigma' \in K_{\kappa'}(\Xi', t + \varepsilon)$, and hence C is ε-simplicial from $\mathbb{K}_\kappa(\Xi)$ to $\mathbb{K}_{\kappa'}(\Xi')$. Symmetrically, C^T is also ε-simplicial. Therefore, the conclusion follows from Proposition 4.2 in [6]. □

Let us define

$$S((\kappa, \Xi), (\kappa', \Xi')) := \sup_{\substack{\sigma\subset\Xi, \sigma'\subset\Xi' \\ d_H(\sigma,\sigma')\le d_H(\Xi,\Xi')}} |\kappa(\sigma) - \kappa'(\sigma')|. \tag{C.1}$$

We remark that $S((\kappa, \Xi), (\kappa', \Xi')) = \|\kappa - \kappa'\|_\infty$ if $\Xi = \Xi'$.

LEMMA C.2. *Let C denote the correspondence defined by $C = \{(x, x') \in \Xi \times \Xi' : \|x - x'\| \leq d_H(\Xi, \Xi')\}$. Then*

$$\mathrm{dis}(C) \leq S\big((\kappa, \Xi), (\kappa', \Xi')\big).$$

PROOF. We easily see that

$$\sup_{\sigma \subset \Xi} d_H\big(\sigma, C(\sigma)\big) \leq d_H\big(\Xi, \Xi'\big) \quad \text{and} \quad \sup_{\sigma' \subset \Xi'} d_H\big(C^T(\sigma'), \sigma'\big) \leq d_H\big(\Xi, \Xi'\big),$$

which implies the assertion. □

For $D_q(\kappa, \Xi) = D_q(\mathbb{K}_\kappa(\Xi))$ and $D_q(\kappa', \Xi') = D_q(\mathbb{K}_{\kappa'}(\Xi'))$, we obtain the following continuity result.

THEOREM C.3.

$$d_B\big(D_q(\kappa, \Xi), D_q(\kappa', \Xi')\big) \leq S\big((\kappa, \Xi), (\kappa', \Xi')\big). \tag{C.2}$$

PROOF. It follows from Lemma C.2 and Lemma C.1 that $H_q(\mathbb{K}_\kappa(\Xi))$ and $H_q(\mathbb{K}_{\kappa'}(\Xi'))$ are $S((\kappa, \Xi), (\kappa', \Xi'))$-interleaving. Therefore, we obtain (C.2) from [5]. □

COROLLARY C.4. *Suppose that κ is Lipschitz continuous with respect to d_H, that is, there exists a constant $\gamma > 0$ such that*

$$\big|\kappa(\sigma) - \kappa(\sigma')\big| \leq \gamma d_H(\sigma, \sigma') \qquad \textit{for } \sigma, \sigma' \in \mathscr{F}(\mathbf{R}^N).$$

Then

$$d_B\big(D_q(\kappa, \Xi), D_q(\kappa, \Xi')\big) \leq \gamma d_H\big(\Xi, \Xi'\big). \tag{C.3}$$

PROOF. From the assumption and (C.1), we see that

$$S\big((\kappa, \Xi), (\kappa, \Xi')\big) \leq \gamma d_H\big(\Xi, \Xi'\big).$$

Therefore, (C.3) follows. □

Acknowledgment. The authors would like to thank Shu Kanazawa, Ippei Obayashi and Kenkichi Tsunoda for useful discussions. The authors would also like to thank the associate editor and the referee for many valuable comments which really help to improve this paper.

REFERENCES

[1] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749

[2] BŁASZCZYSZYN, B. and YOGESHWARAN, D. (2015). Clustering comparison of point processes, with applications to random geometric models. In *Stochastic Geometry*, *Spatial Statistics and Random Fields*. *Lecture Notes in Math*. **2120** 31–71. Springer, Cham. MR3330572

[3] BOBROWSKI, O. and KAHLE, M. (2014). Topology of random simplicial complexes: A survey. Topology in Statistical Inference, the Proceedings of Symposia in Applied Mathematics. To appear. Available at arXiv:1409.4734.

[4] CARLSSON, G. (2009). Topology and data. *Bull*. *Amer. Math*. *Soc*. (*N.S.*) **46** 255–308. MR2476414

[5] CHAZAL, F., DE SILVA, V., GLISSE, M. and OUDOT, S. (2016). *The Structure and Stability of Persistence Modules*. *SpringerBriefs in Mathematics*. Springer, Cham. MR3524869

[6] CHAZAL, F., DE SILVA, V. and OUDOT, S. (2014). Persistence stability for geometric complexes. *Geom*. *Dedicata* **173** 193–214. MR3275299

[7] CHIU, S. N., STOYAN, D., KENDALL, W. S. and MECKE, J. (2013). *Stochastic Geometry and Its Applications*, 3rd ed. *Wiley Series in Probability and Statistics*. Wiley, Chichester. MR3236788

[8] COHEN-STEINER, D., EDELSBRUNNER, H. and HARER, J. (2007). Stability of persistence diagrams. *Discrete Comput*. *Geom*. **37** 103–120. MR2279866

[9] EDELSBRUNNER, H. and HARER, J. L. (2010). *Computational Topology*. Amer. Math. Soc., Providence, RI. MR2572029

[10] EDELSBRUNNER, H., LETSCHER, D. and ZOMORODIAN, A. (2002). Topological persistence and simplification. *Discrete Comput*. *Geom*. **28** 511–533. MR1949898

[11] GAMEIRO, M., HIRAOKA, Y., IZUMI, S., KRAMAR, M., MISCHAIKOW, K. and NANDA, V. (2015). A topological measurement of protein compressibility. *Jpn*. *J*. *Ind*. *Appl*. *Math*. **32** 1–17. MR3318898

[12] GHOSH, S. (2012). Rigidity and Tolerance in Gaussian zeroes and Ginibre eigenvalues: Quantitative estimates. Preprint. Available at arXiv:1211.3506.

[13] GHOSH, S. and PERES, Y. (2017). Rigidity and tolerance in point processes: Gaussian zeros and Ginibre eigenvalues. *Duke Math*. *J*. **166** 1789–1858. MR3679882

[14] GRESCHONIG, G. and SCHMIDT, K. (2000). Ergodic decomposition of quasi-invariant probability measures. *Colloq*. *Math*. **84/85** 495–514. MR1784210

[15] HATCHER, A. (2002). *Algebraic Topology*. Cambridge Univ. Press, Cambridge. MR1867354

[16] HIRAOKA, Y., NAKAMURA, T., HIRATA, A., ESCOLAR, E. G., MATSUE, K. and NISHIURA, Y. (2016). Hierarchical structures of amorphous solids characterized by persistent homology. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **113** 7035–7040.

[17] HIRAOKA, Y. and SHIRAI, T. (2016). Tutte polynomials and random-cluster models in Bernoulli cell complexes. In *Stochastic Analysis on Large Scale Interacting Systems*. *RIMS Kôkyûroku Bessatsu*, *B*59 289–304. Res. Inst. Math. Sci. (RIMS), Kyoto. MR3675939

[18] HIRAOKA, Y. and SHIRAI, T. (2017). Minimum spanning acycle and lifetime of persistent homology in the Linial–Meshulam process. *Random Structures Algorithms* **51** 315–340. MR3683365

[19] HIRAOKA, Y. and TSUNODA, K. (2016). Limit theorems for random cubical homology. Preprint. Available at arXiv:1612.08485.

[20] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. and VIRÁG, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. *University Lecture Series* **51**. Amer. Math. Soc., Providence, RI. MR2552864

[21] KACZYNSKI, T., MISCHAIKOW, K. and MROZEK, M. (2004). *Computational Homology*. *Applied Mathematical Sciences* **157**. Springer, New York. MR2028588

[22] KAHLE, M. (2011). Random geometric complexes. *Discrete Comput*. *Geom*. **45** 553–573. MR2770552

[23] KAHLE, M. (2014). Topology of random simplicial complexes: A survey. In *Algebraic Topology*: *Applications and New Directions*. *Contemp*. *Math*. **620** 201–221. Amer. Math. Soc., Providence, RI. MR3290093

[24] KALLENBERG, O. (1986). *Random Measures*, 4th ed. Akademie-Verlag, Berlin. MR0854102

[25] KARAMCHANDANI, N., MANJUNATH, D., YOGESHWARAN, D. and IYER, S. K. (2006). Evolving random geometric graph models for mobile wireless networks. In 2006 *4th International Symposium on Modeling and Optimization in Mobile*, *Ad Hoc and Wireless Networks* 1–7. IEEE.

[26] KUSANO, G., FUKUMIZU, K. and HIRAOKA, Y. (2016). Persistence weighted Gaussian kernel for topological data analysis. In *Proceedings of the* 33*rd International Conference on Machine Learning* 2004–2013.

[27] MIYOSHI, N. and SHIRAI, T. (2014). A cellular network model with Ginibre configured base stations. *Adv. in Appl*. *Probab*. **46** 832–845. MR3254344

[28] NAKAMURA, T., HIRAOKA, Y., HIRATA, A., ESCOLAR, E. G. and NISHIURA, Y. (2015). Persistent homology and many-body atomic structure for medium-range order in the glass. *Nanotechnol*. **26** 304001.

[29] OWADA, T. and ADLER, R. J. (2017). Limit theorems for point processes under geometric constraints (and topological crackle). *Ann*. *Probab*. **45** 2004–2055. DOI:10.1214/16-AOP1106. MR3650420

[30] PENROSE, M. (2003). *Random Geometric Graphs*. *Oxford Studies in Probability* **5**. Oxford Univ. Press, Oxford. MR1986198

[31] PENROSE, M. D. and YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann*. *Appl*. *Probab*. **11** 1005–1041. MR1878288

[32] PUGH, C. and SHUB, M. (1971). Ergodic elements of ergodic actions. *Compos*. *Math*. **23** 115–122. MR0283174

[33] REED, M. and SIMON, B. (1978). *Methods of Modern Mathematical Physics*. *IV*. *Analysis of Operators*. Academic Press, New York. MR0493421

[34] RILEY, S., EAMES, K., ISHAM, V., MOLLISON, D. and TRAPMAN, P. (2015). Five challenges for spatial epidemic models. *Epidemics* **10** 68–71.

[35] SAADATFAR, M., TAKEUCHI, H., ROBINS, V., FRANCOIS, N. and HIRAOKA, Y. (2017). Pore configuration landscape of granular crystallization. *Nat*. *Commun*. **8** 15082.

[36] SHIRAI, T. and TAKAHASHI, Y. (2003). Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *J*. *Funct*. *Anal*. **205** 414–463. DOI:10.1016/S0022-1236(03)00171-X. MR2018415

[37] TRINH, K. D. (2017). A remark on the convergence of Betti numbers in the thermodynamic regime. *Pac*. *J*. *Math*. *Ind*. **9** Art. 4, 7. DOI:10.1186/s40736-017-0029-0. MR3620264

[38] YOGESHWARAN, D. and ADLER, R. J. (2015). On the topology of random complexes built over stationary point processes. *Ann*. *Appl*. *Probab*. **25** 3338–3380. MR3404638

[39] YOGESHWARAN, D., SUBAG, E. and ADLER, R. J. (2017). Random geometric complexes in the thermodynamic regime. *Probab*. *Theory Related Fields* **167** 107–142. MR3602843

[40] ZOMORODIAN, A. and CARLSSON, G. (2005). Computing persistent homology. *Discrete Comput*. *Geom*. **33** 249–274. MR2121296

Y. HIRAOKA
ADVANCED INSTITUTE FOR MATERIALS RESEARCH
TOHOKU UNIVERSITY
2-1-1, KATAHIRA, AOBA-KU
SENDAI, 980-8577
JAPAN
E-MAIL: hiraoka@tohoku.ac.jp

T. SHIRAI
K. D. TRINH
INSTITUTE OF MATHEMATICS FOR INDUSTRY
KYUSHU UNIVERSITY
744, MOTOOKA, NISHI-KU
FUKUOKA, 819-0395
JAPAN
E-MAIL: shirai@imi.kyushu-u.ac.jp
trinh@imi.kyushu-u.ac.jp