Modern Stochastics: Theory and Applications 10 (1) (2023) 1–18
https://doi.org/10.15559/22-VMSTA214

A limit theorem for persistence diagrams of random filtered complexes built over marked point processes

Tomoyuki Shirai[a], **Kiyotaka Suzaki**[b,*]

[a] *Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*
[b] *Mathematical Science Education Center, Headquarters for Admissions and Education, Kumamoto University, 2-40-1 Kurokami, Chuo-ku,Kumamoto, 860-8555, Japan*

shirai@imi.kyushu-u.ac.jp (T. Shirai), k-suzaki@kumamoto-u.ac.jp (K. Suzaki)

Received: 4 February 2022, Revised: 9 August 2022, Accepted: 29 August 2022,
Published online: 13 September 2022

Abstract Random filtered complexes built over marked point processes on Euclidean spaces are considered. Examples of these filtered complexes include a filtration of Čech complexes of a family of sets with various sizes, growths, and shapes. The law of large numbers for persistence diagrams is established as the size of the convex window observing a marked point process tends to infinity.

Keywords Marked point process, persistence diagram, persistent Betti number, random topology

2010 MSC 60K35, 60B10, 55N20

1 Introduction

Much attention has been paid to topological data analysis (TDA) over the last few decades and persistent homology has been playing a central role as one of the most important tools in TDA. Persistent homology measures persistence of topological feature, in particular, appearance and disapperance of homology generators in each

*Corresponding author.

dimension and enables us to view data sets in multi-resolutional way. There are several aspects to be discussed in the theory of persistent homology, among which we focus on the random aspect. Data sets to be analyzed are often represented as binomial processes if each data point is regarded as a sample from a certain probability distribution and as stationary point processes if data points are considered as part of a huge object. There have been many works on the topology of binomial processes from the viewpoint of manifold learning [11, 3, 4]. In the setting of stationary point processes, Yogeshwaran and Adler [19] discussed the topology of random complexes built over stationary point processes in the Euclidean space and showed the strong law of large numbers for Betti numbers of such random complexes. In the same setting, Hiraoka, Shirai and Trinh [12] proved the strong law of large numbers for persistence diagrams, which comprise all information about persistence Betti numbers, and also discussed the positivity of its limiting persistence diagram. In the present paper, we extend the framework to deal with random filtered complexes built over stationary *marked* point processes in order to include more natural examples such as *weighted* complexes ([2], [5], [13], and references therein).

Given data as a finite point configuration Ξ in $\mathbb{R}^d$, we consider the union of closed balls $\cup_{x\in\Xi}\overline{B}_t(x)$ of radius $t \geq 0$ centered at each data point $x \in \Xi$, which we denoted by $X(\Xi, t)$. We are interested in how the q-dimensional homology classes of $X(\Xi, t)$ behave as t grows. By the so-called Nerve theorem, it is well known that $X(\Xi, t)$ is homotopy equivalent to the Čech complex $C(\Xi, t)$, which is defined as a simplicial complex over points in Ξ consisting of q-simplices $\sigma = \{x_0, x_1, \dots, x_q\}$ for which $\cap_{i=0}^{q}\overline{B}_t(x_i) \neq \emptyset$. We thus obtain a filtration of simplicial complexes $C(\Xi) = \{C(\Xi, t)\}_{t\geq 0}$ from Ξ. The qth persistent homology of the filtration $C(\Xi)$ gives more topological information of data than the homologies of snapshots of $C(\Xi, t)$ (cf. [9] and [21]).

When we look at an atomic configuration, it is natural to consider the influence of atomic radii. In the usual setting, as explained above, we start from a finite set of points in Ξ and attach balls of radius t to construct the Čech complex, however, taking atomic radii into account, it would be natural to start from a finite set of balls with initial radii rather than a finite set of points. If points are considered to have different shapes, it would be better to attach a different shape of $V_t(x_i)$ to the ith point instead of a ball $B_t(x_i)$ depending on the shape of the ith point. In many applications, each point often has some extra information and so we would like to incorporate it in our framework. For this purpose, in the present paper, we introduce a filtration of simplicial complexes built over finite sets on $\mathbb{R}^d$ with *marks* in a complete separable metric space $\mathbb{M}$. Here by marks we mean additional information of data and some information at each point that can be expressed as a mark by taking $\mathbb{M}$ appropriately.

Now we introduce some notations to state our main theorems. We say that a nonempty finite subset Ξ of $\mathbb{R}^d \times \mathbb{M}$ is a simple marked point set if $\#(\Xi \cap \pi^{-1}\{x\}) \leq 1$ holds for any $x \in \mathbb{R}^d$, where $\#A$ is the cardinality of a set A and $\pi : \mathbb{R}^d \times \mathbb{M} \to \mathbb{R}^d$ is the natural projection. For a simple marked point set Ξ, by forgetting marks by π, we obtain a simple point set $\Xi_g = \pi(\Xi) \subset \mathbb{R}^d$ as the ground point set of Ξ. For a given simple marked point set Ξ, we define a filtration of simplicial complexes $\mathbb{K}(\Xi) = \{K(\Xi, t)\}_{t\geq 0}$ with the vertex sets in Ξ_g by assigning the birth time $\kappa(\sigma)$ for each simplex $\sigma_g = \pi(\sigma) \subset \Xi_g$, that is, $K(\Xi, t) = \{\sigma_g \subset \Xi_g : \kappa(\sigma) \leq t\}$, where κ is a function defined on the nonempty finite subsets of $\mathbb{R}^d \times \mathbb{M}$ with some appropriate conditions (see Section 2.1). We call $\mathbb{K}(\Xi)$ the *κ-filtered complex* built over a simple marked point set Ξ. This is a marked version of κ-complex (resp. filtration) introduced in [12] as a generalization of Čech and Vietoris–Rips complex (resp. filtration). For example, if we consider the case where $\mathbb{M}$ is the closed interval $[0, R]$ and

$$\kappa(\sigma) = \inf_{w \in \mathbb{R}^d} \max_{(x,r) \in \sigma} (\|x - w\| - r)^+ \text{ for finite } \sigma \subset \mathbb{R}^d \times \mathbb{M},$$

then $\mathbb{K}(\Xi, t)$ is the Čech complex of the family of closed balls $\{\overline{B}_{t+r}(x)\}_{(x,r)\in\Xi}$, which is the case where we start from balls with various initial radii (Example 2.1). Thus our framework enables us to consider a filtration of Čech complexes of a family of balls with various sizes naturally. Later, we give several examples of κ-filtered complexes, which include a filtration of Čech complexes of a family of sets with various growth speeds (Example 2.2) and various shapes (Example 2.3). These examples are often called weighted complexes.

For a κ-filtered complex $\mathbb{K}(\Xi)$, its qth *persistence diagram*

$$D_q(\mathbb{K}(\Xi)) = \{(b_i, d_i) \in \Delta : i = 1, 2, \dots, n_q\}$$

is defined by a multiset on $\Delta = \{(x, y) \in [0, \infty] \times [0, \infty] : x < y\}$ determined by the decomposition of the persistent homology (see Section 2.2), which is an expression of the qth persistent homology. Each (b_i, d_i) means that a qth homology class appears at $t = b_i$, persists for $b_i \le t < d_i$, and disappears at $t = d_i$ in $\mathbb{K}(\Xi)$. In this paper, the persistence diagram $D_q(\mathbb{K}(\Xi))$ is treated as the counting measure

$$\xi_q(\mathbb{K}(\Xi)) = \sum_{(b,d)\in D_q(\mathbb{K}(\Xi))} \delta_{(b,d)},$$

where $\delta_{(x,y)}$ denotes the Dirac measure at $(x, y) \in \Delta$. Let Φ be a marked point process on $\mathbb{R}^d$ with marks in $\mathbb{M}$. It is a point process on $\mathbb{R}^d \times \mathbb{M}$ such that the ground process $\Phi_g(\cdot) = \Phi(\cdot \times \mathbb{M})$ is a simple point process on $\mathbb{R}^d$. We assume that Φ_g has all finite moments, that is, $\mathbb{E}[\Phi_g(A)^p] < +\infty$ for any bounded Borel set A in $\mathbb{R}^d$ and $p \ge 1$. The restricted marked point process Φ on $A \times \mathbb{M}$ is denoted by Φ_A. We discuss the strong law of large numbers for persistence diagrams of a random κ-filtered complex built over a marked point process, or more precisely, the asymptotic behavior of $\xi_q(\mathbb{K}(\Phi_{A_n}))$ (ξ_{q,A_n} for short) of the κ-filtered complex $\mathbb{K}(\Phi_{A_n}) = \{K(\Phi_{A_n}, t)\}_{t\ge 0}$ as the size of window A_n tends to infinity, where $\{A_n\}_{n\in\mathcal{N}}$ is an increasing net of bounded convex sets in $\mathbb{R}^d$ with $\sup\{r > 0 : A_n \text{ contains a ball of radius } r\} \to \infty$ as $n \to \infty$. Such a net is called a *convex averaging net* in $\mathbb{R}^d$. It is a generalized version of a convex averaging sequence considered in [8], for example. The main purpose of this paper is to show the following.

Theorem. *Let Φ be a stationary ergodic marked point process and suppose its ground process Φ_g has all finite moments. Then for any nonnegative integer q, there exists a Radon measure ν_q on Δ such that for any convex averaging net $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$ in $\mathbb{R}^d$,*

$$\frac{1}{|A_n|}\xi_{q,A_n} \xrightarrow{v} \nu_q \quad \text{a.s. as } n \to \infty$$

where $|A|$ is the d-dimensional Lebesgue measure of A and $\xrightarrow{v}$ denotes the vague convergence of measures on Δ.

New feature of this theorem is two-fold: marks and averaging nets. The same limit theorem as above is first established in [12, Theorem 1.5] for persistence diagrams in the case of stationary ergodic point processes (without marks) on $\mathbb{R}^d$ and $\{A_n\}_{n\in\mathcal{N}}$ being the rectangles $\{[-L/2, L/2)^d\}_{L>0}$. Marked point processes are often useful from the application point of view (cf. [1, 6]) so that this extension greatly expanded the scope of application in TDA. The limit theorem along convex averaging sequences can also be found in a recent article [17] when the underlying filtered complexes are basically Čech complexes. Our theorem is also an extension of [17] to the case of the class of κ-complexes, which includes Čech complexes as a special example. We also remark that the papers [19] and [20] discuss the limiting behavior of Betti numbers of random Čech complexes built over stationary point processes.

The paper is organized as follows. We give the statement of our results after introducing some notation and fundamental facts in Section 2. Some examples of marked point processes and κ-filtered complexes are also presented in this section. In Section 3, we show the law of large numbers for persistent Betti numbers (Theorem 2.7) to prove the main theorem (Theorem 2.6).

2 Preliminaries and results

2.1 κ-filtered complexes

For a topological space S, let $\mathcal{F}(S)$ be the collection of all finite nonempty subsets in S. Given a function f on $\mathcal{F}(S)$, there exists a permutation invariant function f_k on S^k such that $f_k(s_1, s_2, \ldots, s_k) = f(\{s_1, s_2, \ldots, s_k\})$ for any positive integer k. We say a function f on $\mathcal{F}(S)$ is measurable if the permutation invariant functions $\{f_k\}$ are Borel measurable. In this paper, we extend the κ-filtration for (unmarked) point processes introduced in [12] to that for marked ones. Let $\mathbb{M}$ be a complete separable metric space, which stands for the set of marks, and $\kappa : \mathcal{F}(\mathbb{R}^d \times \mathbb{M}) \to [0, \infty)$ a measurable function satisfying the following:

(K1) $\kappa(A) \le \kappa(B)$ if $A \subset B$.

(K2) κ is invariant under the translations acting on the first component only, i.e.,

$$\kappa(T_a(A)) = \kappa(A)$$

for any $a \in \mathbb{R}^d$ and for any $A \in \mathcal{F}(\mathbb{R}^d \times \mathbb{M})$, where $T_a : (x, m) \mapsto (x + a, m)$.

(K3) There exists an increasing function $\rho : [0, \infty) \to [0, \infty)$ such that

$$\|x - y\| \le \rho(\kappa(\{(x, m), (y, n)\}))$$

for all $(x, m), (y, n) \in \mathbb{R}^d \times \mathbb{M}$, where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^d$.

Let $\pi : \mathbb{R}^d \times \mathbb{M} \to \mathbb{R}^d$ be the projection with respect to the first component. We say $\Xi \in \mathcal{F}(\mathbb{R}^d \times \mathbb{M})$ is a simple marked point set if for any $x \in \mathbb{R}^d$, $\#(\Xi \cap \pi^{-1}\{x\}) \le 1$

holds, where $\#A$ is the number of elememts in A. For any simple marked point set Ξ, we write as $\Xi_g = \pi(\Xi)$. The projection π naturally induces the bijection

$$\mathcal{F}(\Xi) \ni \sigma \mapsto \sigma_g \in \mathcal{F}(\Xi_g).$$

Once a simple marked point set Ξ is fixed, each subset

$$\sigma = \{(x_0, m_0), (x_1, m_1), \ldots, (x_q, m_q)\}$$

of Ξ can be regarded as a finite point configuration $\sigma_g = \{x_0, x_1, \ldots, x_q\}$ in $\mathbb{R}^d$ with marks $\{m_0, m_1, \ldots, m_q\}$ in $\mathbb{M}$. By the definition of simple marked point sets we see that each $x \in \Xi_g$ has a unique mark $m \in \mathbb{M}$ with $(x, m) \in \Xi$.

Given a simple marked point set Ξ, we construct a filtration

$$\mathbb{K}(\Xi) = \{K(\Xi, t)\}_{t \geq 0} \tag{1}$$

of simplicial complexes from the simple marked point set Ξ and the function κ by

$$K(\Xi, t) = \{\sigma_g \subset \Xi_g \ : \ \kappa(\sigma) \leq t\},$$

i.e., $\kappa(\sigma)$ is the birth time of a simplex σ_g in the filtration $\mathbb{K}(\Xi)$. Note that whether or not a q-simplex $\sigma_g = \{x_0, x_1, \ldots, x_q\} \subset \Xi_g$ belongs to $K(\Xi, t)$ depends not only on the q-simplex itself but also on the marked set

$$\sigma = \{(x_0, m_0), (x_1, m_1), \ldots, (x_q, m_q)\} \subset \Xi.$$

We call $\mathbb{K}(\Xi) = \{K(\Xi, t)\}_{t \geq 0}$ the κ-filtered complex built over Ξ. We also note that the conditions (K1) and (K3) of κ yield the following diameter bound

$$\operatorname{diam} \sigma_g \leq \rho(t)$$

for any simplex $\sigma \in K(\Xi, t)$. Indeed, for any $\sigma \in K(\Xi, t)$ and any $x, y \in \sigma_g$ we take $m, n \in \mathbb{M}$ with $(x, m), (y, n) \in \sigma$, then it is easy to see that

$$\|x - y\| \leq \rho(\kappa(\{(x, m), (y, n)\})) \leq \rho(\kappa(\sigma)) \leq \rho(t).$$

Example 2.1 (Čech and Vietoris–Rips filtered complex with various sizes)**.** For a fixed $R > 0$, let $\mathbb{M}$ be the closed interval $[0, R]$. Fundamental examples of κ on $\mathcal{F}(\mathbb{R}^d \times \mathbb{M})$ are

$$\kappa_C(\sigma) = \inf_{w \in \mathbb{R}^d} \max_{(x,r) \in \sigma} (\|x - w\| - r)^+$$

$$\text{and} \quad \kappa_R(\sigma) = \max_{(x_1,r_1),(x_2,r_2) \in \sigma} \frac{(\|x_1 - x_2\| - r_1 - r_2)^+}{2},$$

where $a^+ = \max\{a, 0\}$ for $a \in \mathbb{R}$. It is easy to see that they satisfy (K1), (K2), and (K3) with $\rho(t) = 2t + 2R$. We denote the corresponding κ-filtered complexes built over a simple marked point set Ξ by $\mathbb{C}(\Xi) = \{C(\Xi, t)\}_{t \geq 0}$ and $\mathbb{R}(\Xi) = \{R(\Xi, t)\}_{t \geq 0}$, respectively. For $\sigma \in \mathcal{F}(\Xi)$, we see that

$$\kappa_C(\sigma) \leq t \Leftrightarrow \bigcap_{(x,r) \in \sigma} \overline{B}_{t+r}(x) \neq \emptyset,$$

$$\kappa_R(\sigma) \leq t \Leftrightarrow \overline{B}_{t+r_1}(x_1) \cap \overline{B}_{t+r_2}(x_2) \neq \emptyset \quad \text{for any } (x_1, r_1), (x_2, r_2) \in \sigma,$$

where $\overline{B}_r(x) = \{y \in \mathbb{R}^d : \|y - x\| \le r\}$ is the closure of the open ball $B_r(x)$ of radius r centered at x. Hence $C(\Xi, t)$ and $R(\Xi, t)$ are the so-called Čech complex and Vietoris–Rips complex of the family of balls $\{\overline{B}_{t+r}(x)\}_{(x,r)\in\Xi}$.

Example 2.2 (Čech and Vietoris–Rips filtered complex with various growth speeds)**.** Let $\mathbb{M}$ be a finite family $\{\boldsymbol{r}_i(\cdot)\}_{i\in I}$ of right continuous, strictly increasing functions on $[0, \infty)$. We define functions on $\mathcal{F}(\mathbb{R}^d \times \mathbb{M})$ by

$$\kappa_C(\sigma) = \inf_{w\in\mathbb{R}^d} \max_{(x,\boldsymbol{r})\in\sigma} \boldsymbol{r}^{-1}(\|x - w\|)$$

$$\text{and} \quad \kappa_R(\sigma) = \max_{(x_1,\boldsymbol{r}_1),(x_2,\boldsymbol{r}_2)\in\sigma} (\boldsymbol{r}_1 + \boldsymbol{r}_2)^{-1}(\|x_1 - x_2\|),$$

where $\boldsymbol{r}^{-1}(t) = \inf\{s \ge 0 : \boldsymbol{r}(s) \ge t\}$. One can show that

$$\kappa_C(\sigma) \le t \Leftrightarrow \bigcap_{(x,\boldsymbol{r})\in\sigma} \overline{B}_{\boldsymbol{r}(t)}(x) \ne \emptyset,$$

$$\kappa_R(\sigma) \le t \Leftrightarrow \overline{B}_{\boldsymbol{r}_1(t)}(x_1) \cap \overline{B}_{\boldsymbol{r}_2(t)}(x_2) \ne \emptyset \text{ for any } (x_1, \boldsymbol{r}_1), (x_2, \boldsymbol{r}_2) \in \sigma$$

in the same way as in Example 2.1 above. In this case, (K3) is satisfied with $\rho(t) = 2\max_{i\in I} \boldsymbol{r}_i(t)$. The corresponding κ-filtered complexes are the Čech complexes and Vietoris–Rips complexes of the family of balls $\{\overline{B}_{\boldsymbol{r}(t)}(x)\}_{(x,\boldsymbol{r})\in\Xi}$ for a simple marked point set $\Xi \in \mathcal{F}(\mathbb{R}^d \times \mathbb{M})$.

Example 2.3 (Čech filtered complex with various shapes)**.** Let $\mathbb{M}$ be a finite family $\{C_i\}_{i\in I}$ of bounded convex sets in $\mathbb{R}^d$ satisfying that $0 \in \operatorname{int} C_i$ for every $i \in I$, where $\operatorname{int} C$ is the interior of C. We put $f_C(z) = \inf\{s \ge 0 : z \in sC\}$ for a convex set C and $z \in \mathbb{R}^d$. Consider the function on $\mathcal{F}(\mathbb{R}^d \times \mathbb{M})$ defined by

$$\kappa(\sigma) = \inf_{w\in\mathbb{R}^d} \max_{(x,C)\in\sigma} f_C(w - x).$$

This satisfies (K3) with $\rho(t) = 2t \max_{i\in I} \operatorname{diam} C_i$. For any simple marked point set $\Xi \in \mathcal{F}(\mathbb{R}^d \times \mathbb{M})$, it is easy to see that the corresponding $K(\Xi, t)$ is the Čech complex of the family of sets $\{t\overline{C} + x\}_{(x,C)\in\Xi}$.

2.2 *Persistent homologies and persistence diagrams*

In what follows, we fix a function κ satisfying the conditions (K1)–(K3) in Section 2.1. Now we give a brief introduction of persistent homology, persistence diagrams, and persistent Betti numbers for the κ-filtered complex $\mathbb{K}(\Xi)$. Let $\mathbb{F}$ be a field. Given a nonnegative integer q and $t \ge 0$, we denote by $H_q(K(\Xi, t))$ the qth homology group of the simplicial complex $K(\Xi, t)$ with coefficients in $\mathbb{F}$. For $r \le s$, the inclusion $K(\Xi, r) \hookrightarrow K(\Xi, s)$ induces the linear map $\iota_r^s : H_q(K(\Xi, r)) \to H_q(K(\Xi, s))$. We put $H_q(\mathbb{K}(\Xi)) = (H_q(\{\mathbb{K}(\Xi, t)\}_{t\ge 0}, \{\iota_r^s\}_{s\ge r\ge 0})$ and call it the qth persistent homology (or persistence module) of $K(\Xi)$. It is well known that there exist a unique nonnegative integer n_q and $b_i, d_i \in [0, \infty]$ with $b_i < d_i$, $i = 1, 2, \ldots, n_q$, such that the qth persistent homology $H_q(\mathbb{K}(\Xi))$ has a decomposition property

$$H_q(\mathbb{K}(\Xi)) \simeq \bigoplus_{i=1}^{n_q} I(b_i, d_i), \tag{2}$$

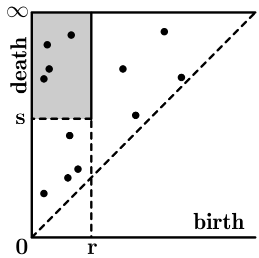

Fig. 1. $\beta_q^{r,s}$ counts the number of birth–death pairs in the gray region

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

$$U_r = \begin{cases} \mathbb{F} & b_i \le r < d_i, \\ 0 & \text{otherwise}, \end{cases}$$

and the identity map $f_r^s = \mathrm{id}_{\mathbb{F}}$ for $b_i \le r \le s < d_i$. Each $I(b_i, d_i)$ in (2) describes that a topological feature (qth homology class) appears at $t = b_i$, persists for $b_i \le t < d_i$, and disappears at $t = d_i$ in $\mathbb{K}(\Xi)$. We call the pair (b_i, d_i) its birth–death pair. The qth persistence diagram of $\mathbb{K}(\Xi)$ is defined by a multiset

$$D_q(\mathbb{K}(\Xi)) = \{(b_i, d_i) \in \Delta : i = 1, 2, \ldots, n_q\},$$

where $\Delta = \{(x, y) \in [0, \infty] \times [0, \infty] : x < y\}$. Let $m_{b,d}$ be the multiplicity of the point $(b, d) \in D_q(\Xi)$ and $\xi_q(\mathbb{K}(\Xi))$ the counting measure on Δ given by

$$\xi_q(\mathbb{K}(\Xi)) = \sum_{(b,d)} m_{b,d}\delta_{(b,d)},$$

where $\delta_{(x,y)}$ is the Dirac measure at $(x, y) \in \Delta$. We identify the persistence diagram $D_q(\mathbb{K}(\Xi))$ with the counting measure $\xi_q(\mathbb{K}(\Xi))$. The qth (r, s)-*persistent Betti number* is also defined by

$$\beta_q^{r,s}(\mathbb{K}(\Xi)) = \dim \frac{Z_q(K(\Xi, r))}{Z_q(K(\Xi, r)) \cap B_q(K(\Xi, s))},$$

where $Z_q(K(\Xi, r))$ and $B_q(K(\Xi, r))$ are the qth cycle group and boundary group of $K(\Xi, r)$, respectively. It is easy to see that this number is equal to the rank of $\iota_r^s : H_q(K(\Xi, r)) \to H_q(K(\Xi, s))$. By definition of the persistent Betti number, we have

$$\beta_q^{r,s}(\mathbb{K}(\Xi)) = \sum_{b \le r, s < d} m_{b,d} = \xi_q(\mathbb{K}(\Xi))([0, r] \times (s, \infty]). \tag{3}$$

Therefore the persistence Betti number $\beta_q^{r,s}$ counts the number of birth–death pairs in the persistence diagram $D_q(\mathbb{K}(\Xi))$ located in the gray region of Figure 1. Details for these facts can be found in [9], [12] and [21], for example.

2.3 Marked point processes

Now we consider marked point processes. Let X be a complete separable metric space and $\mathcal{B}(X)$ the Borel σ-field on X. A Borel measure μ on X is boundedly finite if $\mu(A) < \infty$ for every bounded Borel set A. We say that a sequence $\{\mu_n\}$ of boundedly finite measures on X converges to a boundedly finite measure μ on X in the $w^{\#}$-topology if

$$\int_X f\, d\mu_n \to \int_X f\, d\mu \text{ as } n \to \infty \tag{4}$$

for all bounded continuous functions f on X vanishing outside a bounded set. We denote by $\mathcal{M}_X^{\#}$ the totality of boundedly finite measures on $\mathcal{B}(X)$. $\mathcal{M}_X^{\#}$ is a complete separable metric space under the $w^{\#}$-topology. The corresponding σ-field $\mathcal{B}(\mathcal{M}_X^{\#})$ coincides with the smallest σ-field with respect to which the mappings $\mu \mapsto \mu(A)$ are measurable for all $A \in \mathcal{B}(X)$. If X is a locally compact Hausdorff space with countable base, we can take a metric so that X is complete and every bounded subset of X is relatively compact. Then a Borel measure is boundedly finite if and only if it is a Radon measure and $w^{\#}$-convergence coincides with vague convergence. We recall that a Radon measure is a measure on X taking finite values on compact sets and a sequence $\{\mu_n\}$ of Radon measures on X converges to a Radon measure μ on X vaguely (or in the vague topology) if (4) holds for each continuous function f on X vanishing outside a compact set. In this case, we write $\mu_n \xrightarrow{v} \mu$. Let $\mathcal{N}_X^{\#}$ be the totality of boundedly finite integer-valued measures. We call a measure in $\mathcal{N}_X^{\#}$ a counting measure for short. For a counting measure μ on X, there exist sequences of positive integers $\{k_i\}$ and points $\{x_i\}$ in X with at most finitely many x_i in any bounded Borel set such that

$$\mu = \sum_i k_i \delta_{x_i}.$$

Note that $\mathcal{N}_X^{\#}$ is a closed subset of $\mathcal{M}_X^{\#}$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An $(\mathcal{M}_X^{\#}, \mathcal{B}(\mathcal{M}_X^{\#}))$ (resp., $(\mathcal{N}_X^{\#}, \mathcal{B}(\mathcal{N}_X^{\#}))$-valued random variable ξ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a random measure (resp., point process) on X. A point process ξ is typically identified with the random point configuration of its atoms. The expectation measure (or mean measure) of ξ is defined so that $M(A) = \mathbb{E}[\xi(A)]$ for any $A \in \mathcal{B}(X)$. It is often denoted by $\mathbb{E}[\xi]$. We say that a point process ξ is simple if

$$\mathbb{P}(\xi(\{x\}) = 0 \text{ or } 1 \text{ for any } x \in X) = 1.$$

A marked point process on $\mathbb{R}^d$ with marks in $\mathbb{M}$ is a point process Φ on $\mathbb{R}^d \times \mathbb{M}$ whose marginal point process $\Phi_g(\cdot) = \Phi(\cdot \times \mathbb{M})$ on $\mathbb{R}^d$ is a simple point process on $\mathbb{R}^d$. The point process Φ_g is called the ground process of Φ. We say that the ground process Φ_g has all finite moments if $\mathbb{E}[\Phi_g(A)^p] < +\infty$ for every bounded $A \in \mathcal{B}(\mathbb{R}^d)$ and every $p \geq 1$. The translations $\{T_a\}_{a\in\mathbb{R}^d}$ on $\mathbb{R}^d \times \mathbb{M}$ induce the translations $\{T_{a*}\}_{a\in\mathbb{R}^d}$ on $\mathcal{N}^{\#}_{\mathbb{R}^d\times\mathbb{M}}$ defined by

$$(T_{a*}\mu)(A) = \mu(T_a^{-1}A)$$

for $a \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{M})$. A marked point process is called stationary if its probability distribution is translation invariant. A stationary marked point process Φ

is called ergodic if every member B of $\mathcal{B}(\mathcal{N}^{\#}_{\mathbb{R}^d \times \mathbb{M}})$ with $\mathbb{P} \circ \Phi^{-1}(T_{a*}B \Delta B) = 0$ for all $a \in \mathbb{R}^d$ satisfies $\mathbb{P} \circ \Phi^{-1}(B) = 0$ or 1.

Example 2.4 (point process with i.i.d. marks)**.** Let Φ_g be a point process on $\mathbb{R}^d$ and $\{X_i\}$ a measurable enumeration of Φ_g, that is, $\{X_i\}$ is a sequence of $\mathbb{R}^d$-valued random variables so that $\Phi_g = \sum_i \delta_{X_i}$ a.s. We take an i.i.d. sequence of $\mathbb{M}$-valued random variables $\{M_i\}$ such that $\Phi = \{X_i\}$ and $\{M_i\}$ are independent. A marked point process on $\mathbb{R}^d$ with marks in $\mathbb{M}$ is defined by

$$\Phi = \sum_i \delta_{(X_i, M_i)}.$$

If the point process Φ is stationary (and ergodic), then so is Φ.

Example 2.5. Let Φ_g be a simple stationary (ergodic) point process on $\mathbb{R}^d$ and $\{X_i\}$ a measurable enumeration of Φ_g. For a fixed $R > 0$ and for each i, we define a $\{0, 1\}$-valued random variable M_i by

$$M_i = \begin{cases} 1, & \text{if there exists } j \neq i \text{ such that } |X_i - X_j| \leq R, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The point process on $\mathbb{R}^d \times \{0, 1\}$ defined by

$$\Phi = \sum_i \delta_{(X_i, M_i)}$$

is a marked point process. In general, for measurable maps $M_i : \mathcal{N}^{\#}_{\mathbb{R}^d} \to \mathbb{M}$ $(i \geq 1)$, the point process on $\mathbb{R}^d \times \mathbb{M}$ defined by

$$\Phi = \sum_i \delta_{(X_i, M_i(\Phi_g))}$$

is a stationary marked point process.

Incidentally, for M_i $(i \geq 1)$ in (5), the point process

$$\Phi_I = \sum_{i : M_i = 0} \delta_{X_i}$$

on $\mathbb{R}^d$ is called a Matérn type I construction of Φ (see [14]).

Other examples and basic facts for marked point processes are available in [7] and [8].

2.4 Main theorems

In order to state the main results we introduce the notion of convex averaging nets in $\mathbb{R}^d$. Let $(\mathcal{N}, \leq)$ be a linearly ordered set. A family $\mathcal{A} = \{A_n\}_{n \in \mathcal{N}}$ of bounded Borel sets in $\mathbb{R}^d$ is a convex averaging net if

(i) A_n is convex for each $n \in \mathcal{N}$,

(ii) $A_n \subset A_m$ for $n \leq m$, and

(iii) $\sup_{n\in\mathcal{N}} r(A_n) = \infty$, where $r(A) = \sup\{r > 0 : A \text{ contains a ball of radius } r\}$.

Given a marked point process Φ and $A \in \mathcal{B}(\mathbb{R}^d)$, we denote the restricted marked point process Φ on $A \times \mathbb{M}$ by Φ_A, i.e., $\Phi_A(\cdot) = \Phi(\cdot \cap (A \times \mathbb{M}))$. Note that Φ_A can be regarded as a random simple marked point set for any bounded A. For any convex averaging net $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$, a random κ-filtered complex with parameter $n \in \mathcal{N}$ is defined by $\mathbb{K}(\Phi_{A_n}) = \{K(\Phi_{A_n}, t)\}_{t\geq 0}$. For the sake of simplicity, we often denote the corresponding qth persistence diagram $\xi_q(\mathbb{K}(\Phi_{A_n}))$ and qth (r, s)-persistent Betti number $\beta_q^{r,s}(\mathbb{K}(\Phi_{A_n}))$ by ξ_{q,A_n} and $\beta_{q,A_n}^{r,s}$, respectively.

Now we are in a position to state the main theorem.

Theorem 2.6. *Let Φ be a stationary marked point process and suppose its ground process Φ_g has all finite moments. Then for any nonnegative integer q, there exists a Radon measure ν_q on Δ such that for any convex averaging net $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$ in $\mathbb{R}^d$,*

$$\frac{1}{|A_n|}\mathbb{E}[\xi_{q,A_n}] \xrightarrow{v} \nu_q \quad as \quad n \to \infty,$$

where $|A|$ is the d-dimensional Lebesgue measure of A. Furthermore if Φ is ergodic, then

$$\frac{1}{|A_n|}\xi_{q,A_n} \xrightarrow{v} \nu_q \quad a.s.\ as\ n \to \infty$$

Theorem 2.6 can be proved by a general theory of the vague convergence for Radon measures and the following law of large numbers for persistent Betti numbers.

Theorem 2.7. *Let Φ be a stationary marked point process and suppose its ground process Φ_g has all finite moments. Then, for any $0 \leq r \leq s < \infty$ and nonnegative integer q, there exists a nonnegative number $\bar{\beta}_q^{r,s}$ such that for any convex averaging net $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$ in $\mathbb{R}^d$,*

$$\frac{1}{|A_n|}\mathbb{E}[\beta_{q,A_n}^{r,s}] \to \bar{\beta}_q^{r,s} \qquad as \quad n \to \infty.$$

Furthermore, if Φ is ergodic, then

$$\frac{1}{|A_n|}\beta_{q,A_n}^{r,s} \to \bar{\beta}_q^{r,s} \quad a.s.\ as\ n \to \infty$$

The proofs of Theorem 2.6 and Theorem 2.7 will be given in the next section. They are shown in the same way as [12, Theorem 1.5 and Theorem 1.11], in which such limit theorems for stationary (unmarked) point processes were proved.

3 Proof of Theorems 2.6 and 2.7

The aim of this section is to prove Theorem 2.6 and Theorem 2.7.

3.1 Convergence of persistent Betti numbers

Let M, h be positive numbers and $A \in \mathcal{B}(\mathbb{R}^d)$. We put

$$\begin{aligned}\Lambda_M &= [-M/2, M/2)^d,\\ \underline{A}^{(M)} &= \bigsqcup\{\Lambda_M + z \,:\, z \in M\mathbb{Z}^d \text{ and } (\Lambda_M + z) \subset A\},\\ \overline{A}^{(M)} &= \bigsqcup\{\Lambda_M + z \,:\, z \in M\mathbb{Z}^d \text{ and } (\Lambda_M + z) \cap A \neq \emptyset\},\\ \text{and } \partial A^h &= \{x \in \mathbb{R}^d \,:\, d(x, \partial A) \le h\},\end{aligned}$$

where $M\mathbb{Z}^d = \{Mz \,:\, z \in \mathbb{Z}^d\}$ and $d(x, \partial A) = \inf_{y \in \partial A} |x - y|$. Fundamental results treated in this paper for convex averaging nets are summarized in the next proposition.

Proposition 3.1. *Let $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$ be a convex averaging net in $\mathbb{R}^d$. Then for any $M > 0$ and $h > 0$, as $n \to \infty$,*

$$\frac{|\underline{A}_n^{(M)}|}{|A_n|} \to 1, \qquad \frac{|\overline{A}_n^{(M)}) \setminus \underline{A}_n^{(M)}|}{|A_n|} \to 0, \tag{6}$$

and

$$\frac{|\partial A_n^h|}{|A_n|} \to 0. \tag{7}$$

Proposition 3.1 is a special case of [15, Lemma 3.1]. For (6) and (7), see [10, Lemma 1] and [18, Lemma 2], respectively.

Next we need a version of the multi-dimensional ergodic theorem for stationary ergodic marked point processes.

Proposition 3.2. *Let Φ be a stationary ergodic marked point process and $Z \in L^p(\mathbb{P}\circ \Phi^{-1})$ for $1 \le p < +\infty$. If $\mathcal{A} = \{A_n\}_{n\in\mathcal{N}}$ is a convex averaging net, then for each $M > 0$*

$$\lim_{n\in\mathcal{N}} \frac{1}{|A_n|} \sum_{z\in M\mathbb{Z}^d\cap A_n} Z(T_{-z*}\Phi) = \frac{1}{M^d}\mathbb{E}[Z(\Phi)] \quad \text{a.s.}$$

$$\text{and} \quad \lim_{n\in\mathcal{N}} \frac{1}{|A_n|} \sum_{z\in M\mathbb{Z}^d\cap \underline{A}_n^{(M)}} Z(T_{-z*}\Phi) = \frac{1}{M^d}\mathbb{E}[Z(\Phi)] \quad \text{a.s.}$$

Proof. Take any $M > 0$. Applying [15, Theorem 3.7 and Corollary 3.10] to the probability space $(\mathcal{N}_X^\#, \mathcal{B}(\mathcal{N}_X^\#), \mathbb{P}\circ\Phi^{-1})$ and the translations $\{T_{z*}\}_{z\in M\mathbb{Z}^d}$, we have

$$\lim_{n\in\mathcal{N}} \frac{1}{|A_n|} \sum_{z\in M\mathbb{Z}^d\cap A_n} Z(T_{-z*}\Phi) = \frac{1}{M^d}\mathbb{E}[Z(\Phi)|\Phi^{-1}\mathcal{I}] \quad \text{a.s.}$$

$$\text{and} \quad \lim_{n\in\mathcal{N}} \frac{1}{|A_n|} \sum_{z\in M\mathbb{Z}^d\cap \underline{A}_n^{(M)}} Z(T_{-z*}\Phi) = \frac{1}{M^d}\mathbb{E}[Z(\Phi)|\Phi^{-1}\mathcal{I}] \quad \text{a.s.,}$$

where $\mathcal{I}$ is the invariant σ-field in $\mathcal{N}^\#_{\mathbb{R}^d\times\mathbb{M}}$ under the translations $\{T_{z*}\}_{z\in M\mathbb{Z}^d}$. Rotating $M\mathbb{Z}^d$ if necessary, we may assume that an element of $M\mathbb{Z}^d$ is ergodic (see [16, Theorem 1]). Therefore $\mathcal{I}$ is trivial, that is, for every $I \in \mathcal{I}$, $\mathbb{P}\circ\Phi^{-1}(I) = 0$ or 1. This implies that $\mathbb{E}[Z(\Phi)|\Phi^{-1}\mathcal{I}] = \mathbb{E}[Z(\Phi)]$ a.s. □

Let $S_q(\Xi, t)$ be the number of q-simplices in $\mathbb{K}(\Xi, t)$ for a simple marked point set Ξ. The following limit theorems for $S_q(\Phi, t)$ play important roles in the proof of Theorem 2.7.

Lemma 3.3. *Let Φ be a stationary ergodic marked point process and suppose its ground process Φ_g has all finite moments. Then for any nonnegative integer q, $t \geq 0$, $M > 0$, and convex averaging net $\mathcal{A} = \{A_n\}_{n \in \mathcal{N}}$,*

$$\lim_{n \in \mathcal{N}} \frac{1}{|A_n|} S_q(\Phi_{\underline{A}_n^{(M)}}, t) = \lim_{n \in \mathcal{N}} \frac{1}{|A_n|} S_q(\Phi_{\overline{A}_n^{(M)}}, t) = \lim_{n \in \mathcal{N}} \frac{1}{|A_n|} S_q(\Phi_{A_n}, t) \text{ a.s.}$$

Proof. The proof is similar to that of [20, Lemma 3.2]. Consider the function defined by

$$h_{q,t}^{(M)}(\Phi) = \frac{1}{q+1} \sum_{x \in \Phi_g \cap \Lambda_M} \#\{q\text{-simplices in } \mathbb{K}(\Phi, t) \text{ containing } x\}.$$

We recall that the conditions (K1), (K2), and $\operatorname{diam} \sigma_g \leq \rho(t)$ for every $\sigma_g \in \mathbb{K}(\Phi_{A_n}, t)$. Hence we obtain

$$\begin{aligned}\sum_{z \in M\mathbb{Z}^d \cap (A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M})} h_{q,t}^{(M)}(T_{-z*}\Phi) &\leq S_q(\Phi_{\underline{A}_n^{(M)}}, t) \leq S_q(\Phi_{A_n}, t) \\ &\leq S_q(\Phi_{\overline{A}_n^{(M)}}, t) \leq \sum_{z \in M\mathbb{Z}^d \cap (A_n \cup \partial A_n^{\sqrt{d}M})} h_{q,t}^{(M)}(T_{-z*}\Phi).\end{aligned}$$

Since the ground process Φ_g of Φ has all finite moments, we have $\mathbb{E}[h_{q,t}^{(M)}(\Phi)] \leq \mathbb{E}[\Phi_g(\Lambda_M \cup \partial \Lambda_M^{\rho(t)})^{q+1}] < +\infty$. If we notice the fact that $\{A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M}\}_{n \in \mathcal{N}}$ is also a convex averaging net, we see from Proposition 3.1 and Proposition 3.2 that

$$\begin{aligned}&\frac{1}{|A_n|} \sum_{z \in M\mathbb{Z}^d \cap (A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M})} h_{q,t}^{(M)}(T_{-z*}\Phi) \\ &= \frac{|A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M}|}{|A_n|} \cdot \frac{1}{|A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M}|} \sum_{z \in M\mathbb{Z}^d \cap (A_n \setminus \partial A_n^{\rho(t)+2\sqrt{d}M})} h_{q,t}^{(M)}(T_{-z*}\Phi) \\ &\to \frac{1}{M^d} \mathbb{E}[h_{q,t}^{(M)}(\Phi)] \text{ as } n \to \infty \text{ a.s.}\end{aligned}$$

We can similarly show that

$$\frac{1}{|A_n|} \sum_{z \in M\mathbb{Z}^d \cap (A_n \cup \partial A_n^{\sqrt{d}M})} h_{q,t}^{(M)}(T_{-z*}\Phi) \to \frac{1}{M^d} \mathbb{E}[h_{q,t}^{(M)}(\Phi)] \text{ as } n \to \infty \text{ a.s.}$$

Therefore we reach the desired result. □

Now we state a basic estimate on the persistent Betti numbers for nested filtered complexes $\mathbb{K}^{(1)} \subset \mathbb{K}^{(2)}$. The proof of the following lemma is given in [12, Lemma 2.11].

Lemma 3.4. *Let $\mathbb{K}^{(1)} = \{K_t^{(1)}\}_{t\geq 0}$ and $\mathbb{K}^{(2)} = \{K_t^{(2)}\}_{t\geq 0}$ be filtered complexes with $K_t^{(1)} \subset K_t^{(2)}$ for $t \geq 0$. Then*

$$|\beta_q^{r,s}(\mathbb{K}^{(1)}) - \beta_q^{r,s}(\mathbb{K}^{(2)})| \leq \sum_{j=q,q+1} \#K_{s,j}^{(2)} \setminus K_{s,j}^{(1)} + \#\{\sigma \in K_{s,j}^{(1)} \setminus K_{r,j}^{(1)} : t_\sigma^{(2)} \leq r\},$$

where $K_{s,j}^{(i)}$ is the set of j-simplices in $K_s^{(i)}$, and $t_\sigma^{(i)}$ is the birth time of σ in $\mathbb{K}^{(i)}$, $i = 1, 2$. In particular, if $t_\sigma^{(1)} = t_\sigma^{(2)}$ holds for any simplex σ in $\mathbb{K}^{(1)}$, then

$$|\beta_q^{r,s}(\mathbb{K}^{(1)}) - \beta_q^{r,s}(\mathbb{K}^{(2)})| \leq \sum_{j=q,q+1} \#K_{s,j}^{(2)} \setminus K_{s,j}^{(1)}.$$

Now we give the proof of Theorem 2.7.

Proof of Theorem 2.7. We first note that it can be proved that there exist $C_{q,t} \geq 0$ and $\bar{\beta}_q^{r,s} \geq 0$ such that

$$\mathbb{E}[S_q(\Phi_A, t)] \leq C_{q,t}|A| \text{ for bounded } A \in \mathcal{B}(\mathbb{R}^d) \tag{8}$$

and

$$\lim_{M\to\infty} \frac{1}{M^d}\mathbb{E}[\beta_{q,\Lambda_M}^{r,s}] = \bar{\beta}_q^{r,s} \tag{9}$$

in the same way as in [12, Theorem 1.11]. Take $0 \leq r \leq s < \infty$ and a nonnegative integer q and fix them. The set $\underline{A}_n^{(M)}$ is decomposed into rectangles

$$\underline{A}_n^{(M)} = \bigsqcup_{z\in M\mathbb{Z}^d \cap \underline{A}_n^{(M)}} (\Lambda_M + z).$$

We define a new filtered complex $\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}})$ by

$$\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}) = \bigsqcup_{z\in M\mathbb{Z}^d \cap \underline{A}_n^{(M)}} \mathbb{K}(\Phi_{\Lambda_M+z}).$$

From the second assertion in Lemma 3.4, we have

$$|\beta_{q,A_n}^{r,s} - \beta_q^{r,s}(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}))| \leq \sum_{j=q,q+1}\left(\sum_{z\in M\mathbb{Z}^d\cap\underline{A}_n^{(M)}} S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}+z}, s) + S_j(\Phi_{A_n\setminus\underline{A}_n^{(M)}}, s)\right). \tag{10}$$

Since Φ is stationary, κ satisfies the condition (K2), and it is easy to see that $|\underline{A}_n^{(M)}| = \#\{M\mathbb{Z}^d \cap \underline{A}_n^{(M)}\} \cdot M^d$, we have

$$\mathbb{E}[\beta_q^{r,s}(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}))] = \mathbb{E}\left[\sum_{z\in M\mathbb{Z}^d\cap\underline{A}_n^{(M)}} \beta_{q,\Lambda_M+z}^{r,s}\right] = \#\{M\mathbb{Z}^d \cap \underline{A}_n^{(M)}\} \cdot \mathbb{E}[\beta_{q,\Lambda_M}^{r,s}]$$
$$= |\underline{A}_n^{(M)}| \cdot \frac{1}{M^d}\mathbb{E}[\beta_{q,\Lambda_M}^{r,s}]. \tag{11}$$

In addition, we have

$$\mathbb{E}\left[\sum_{z\in M\mathbb{Z}^d\cap \underline{A}_n^{(M)}} S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}+z}, s)\right] = \#\{M\mathbb{Z}^d \cap \underline{A}_n^{(M)}\}\cdot \mathbb{E}[S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}}, s)] \\ \le |\underline{A}_n^{(M)}|\cdot \frac{C_{j,s}|\partial\Lambda_M^{2\rho(s)}|}{M^d} \tag{12}$$

and

$$\mathbb{E}[S_j(\Phi_{A_n\setminus \underline{A}_n^{(M)}}, s)] \le C_{j,s}|A_n \setminus \underline{A}_n^{(M)}| \tag{13}$$

from (8). Take $\varepsilon > 0$. We can find $M > 0$ such that

$$\left|\frac{1}{M^d}\mathbb{E}[\beta^{r,s}_{q,\Lambda_M}] - \bar{\beta}^{r,s}_q\right| < \varepsilon, \quad \sum_{j=q,q+1}\frac{C_{j,s}|\partial\Lambda_M^{2\rho(s)}|}{M^d} < \varepsilon. \tag{14}$$

By taking expectation on both sides of the inequality (10), we see that the estimates (11), (12), and (13) yield that

$$\begin{aligned}&\left|\frac{1}{|A_n|}\mathbb{E}[\beta^{r,s}_{q,A_n}] - \bar{\beta}^{r,s}_q\right| \\ &\le \frac{1}{|A_n|}\mathbb{E}\left[\left|\beta^{r,s}_{q,A_n} - \beta^{r,s}_q(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}))\right|\right] + \frac{|\underline{A}_n^{(M)}|}{|A_n|}\left|\frac{1}{M^d}\mathbb{E}[\beta^{r,s}_{q,\Lambda_M}] - \bar{\beta}^{r,s}_q\right| \\ &\quad + \bar{\beta}^{r,s}_q\left|\frac{|\underline{A}_n^{(M)}|}{|A_n|} - 1\right| \\ &\le \frac{|\underline{A}_n^{(M)}|}{|A_n|}\varepsilon + \sum_{j=q,q+1} C_{j,s}\frac{|A_n\setminus \underline{A}_n^{(M)}|}{|A_n|} + \frac{|\underline{A}_n^{(M)}|}{|A_n|}\varepsilon + \bar{\beta}^{r,s}_q\left|\frac{|\underline{A}_n^{(M)}|}{|A_n|} - 1\right|.\end{aligned}$$

Therefore we conclude that

$$\limsup_{n\in\mathcal{N}}\left|\frac{1}{|A_n|}\mathbb{E}[\beta^{r,s}_{q,A_n}] - \bar{\beta}^{r,s}_q\right| \le 2\varepsilon.$$

This implies the first assertion.

In order to prove the second assertion, we assume that Φ is ergodic. By virtue of the multi-dimensional ergodic theorem mentioned in Proposition 3.2, we see that

$$\begin{aligned}\frac{1}{|A_n|}\beta^{r,s}_q(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}})) &= \frac{1}{|A_n|}\sum_{z\in M\mathbb{Z}^d\cap \underline{A}_n^{(M)}}\beta^{r,s}_{q,\Lambda_M+z} \\ &= \frac{1}{|A_n|}\sum_{z\in M\mathbb{Z}^d\cap \underline{A}_n^{(M)}}\beta^{r,s}_q(\mathbb{K}((T_{-z*}\Phi))_{\Lambda_M}) \to \frac{1}{M^d}\mathbb{E}[\beta^{r,s}_{q,\Lambda_M}],\end{aligned} \tag{15}$$

and

$$\frac{1}{|A_n|}\sum_{z\in M\mathbb{Z}^d\cap\underline{A}_n^{(M)}} S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}+z}, s) = \frac{1}{|A_n|}\sum_{z\in M\mathbb{Z}^d\cap\underline{A}_n^{(M)}} S_j((T_{-z*}\Phi)_{\partial\Lambda_M^{2\rho(s)}}, s) \to \frac{1}{M^d}\mathbb{E}[S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}}, s)] \tag{16}$$

a.s. as $n\to\infty$ for any $M>0$. If we notice the fact that $S_j(\Phi_A, s) + S_j(\Phi_B, s) \le S_j(\Phi_{A\cup B}, s)$ holds for disjoint bounded $A, B\in\mathcal{B}(\mathbb{R}^d)$, we see from Lemma 3.3 that

$$\frac{1}{|A_n|}S_j(\Phi_{A_n\setminus\underline{A}_n^{(M)}}, s) \le \left|\frac{1}{|A_n|}S_j(\Phi_{A_n}, s) - \frac{1}{|A_n|}S_j(\Phi_{\underline{A}_n^{(M)}}, s)\right| \to 0 \tag{17}$$

a.s. as $n\to\infty$ for any $M>0$. Hence we can find $\Omega_0\in\mathcal{F}$ with $\mathbb{P}(\Omega_0)=1$ such that for any $\omega\in\Omega_0$ and positive integer M, the convergences (15), (16), and (17) hold as $n\to\infty$. Take any $\omega\in\Omega_0$ and $\varepsilon>0$. If we choose a positive integer M so that the inequalities (14) hold, we have

$$\begin{aligned}
&\left|\frac{1}{|A_n|}\beta_{q,A_n}^{r,s}(\omega) - \bar\beta_q^{r,s}\right| \\
&\le \frac{1}{|A_n|}\left|\beta_{q,A_n}^{r,s}(\omega) - \beta_q^{r,s}(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}(\omega)))\right| \\
&\quad + \left|\frac{1}{|A_n|}\beta_q^{r,s}(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}(\omega))) - \frac{1}{M^d}\mathbb{E}[\beta_{q,\Lambda_M}^{r,s}]\right| + \left|\frac{1}{M^d}\mathbb{E}[\beta_{q,\Lambda_M}^{r,s}] - \bar\beta_q^{r,s}\right| \\
&\le \sum_{j=q,q+1}\left(\frac{1}{|A_n|}\sum_{z\in M\mathbb{Z}^d\cap\underline{A}_n^{(M)}} S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}+z}(\omega), s) + \frac{1}{|A_n|}S_j(\Phi_{A_n\setminus\underline{A}_n^{(M)}}(\omega), s)\right) \\
&\quad + \left|\frac{1}{|A_n|}\beta_q^{r,s}(\mathbb{K}^\circ(\Phi_{\underline{A}_n^{(M)}}(\omega))) - \frac{1}{M^d}\mathbb{E}[\beta_{q,\Lambda_M}^{r,s}]\right| + \varepsilon.
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
\limsup_{n\in\mathcal{N}}\left|\frac{1}{|A_n|}\beta_{q,A_n}^{r,s}(\omega) - \bar\beta_q^{r,s}\right| &\le \sum_{j=q,q+1}\frac{1}{M^d}\mathbb{E}[S_j(\Phi_{\partial\Lambda_M^{2\rho(s)}}, s)] + \varepsilon \\
&\le \sum_{j=q,q+1}\frac{C_{j,s}|\partial\Lambda_M^{2\rho(s)}|}{M^d} + \varepsilon \\
&\le 2\varepsilon,
\end{aligned}$$

which implies that the second assertion is valid. The proof of Theorem 2.7 is now complete. ☐

3.2 Convergence of persistence diagrams

In this section we prove Theorem 2.6. To this end, we make use of similar arguments which can be found in the proof of the same kind of limit theorem for persistence diagrams built over stationary point process (Theorem 1.5 in [12]). Let X be a locally compact Hausdorff space with countable base and $\mathcal{C}$ the ring of all relatively compact sets in X. A class $\mathcal{C}' \subset \mathcal{C}$ is called a convergence-determining class (for vague convergence) if for any $\mu \in \mathcal{M}_X^{\#}$ and any sequence $\{\mu_n\} \subset \mathcal{M}_X^{\#}$, the condition

$$\mu_n(A) \to \mu(A) \text{ as } n \to \infty \text{ for all } A \in \mathcal{C}' \cap \mathcal{C}_\mu$$

implies the vague convergence $\mu_n \xrightarrow{v} \mu$, where $\mathcal{C}_\mu$ is the class of relatively compact continuity sets of μ, i.e., $\mathcal{C}_\mu = \{B \in \mathcal{C} : \mu(\partial B) = 0\}$. A class $\mathcal{C}'_\mu$ is called a convergence-determining class for $\mu \in \mathcal{M}_X^{\#}$ if for any sequence $\{\mu_n\} \subset \mathcal{M}_X^{\#}$, the condition

$$\mu_n(A) \to \mu(A) \text{ as } n \to \infty \text{ for all } A \in \mathcal{C}'_\mu$$

implies the vague convergence $\mu_n \xrightarrow{v} \mu$. We note that a class $\mathcal{C}'$ is a convergence-determining class if and only if for any $\mu \in \mathcal{M}_X^{\#}$, $\mathcal{C}' \cap \mathcal{C}_\mu$ is a convergence-determining class for μ. A convergence-determining class $\mathcal{C}'$ has the finite covering property if for any $B \in \mathcal{C}$, B is covered by a finite union of $\mathcal{C}'$-sets. The next lemma can be proved in the same way as Proposition 3.4 in [12].

Lemma 3.5. *Let X be a locally compact Hausdorff space with countable base and $\mathcal{C}'$ a convergence-determining class with finite covering property. Suppose that for every $\mu \in \mathcal{M}_X^{\#}$, $\mathcal{C}'$ contains a countable convergence-determining class for μ. Let $\{\xi_n\}$ be a net of random measures on X satisfying the following:*

(i) For every n, $\mathbb{E}[\xi_n] \in \mathcal{M}_X^{\#}$.

(ii) For every $A \in \mathcal{C}'$, there exists $c_A \geq 0$ such that $\mathbb{E}[\xi_n(A)] \to c_A$ as $n \to \infty$.

Then there exists a unique measure $\mu \in \mathcal{M}_X^{\#}$ such that $\mathbb{E}[\xi_n] \xrightarrow{v} \mu$ as $n \to \infty$ and $\mu(A) = c_A$. Furthermore, if $\xi(A) \to c_A$ almost surely as $n \to \infty$ for any $A \in \mathcal{C}'$, then $\xi_n \xrightarrow{v} \mu$ almost surely as $n \to \infty$.

An example of convergence-determining classes satisfying the conditions in Lemma 3.5 is the following.

Lemma 3.6 (Corollary A.3 in [12])**.** *The class*

$$\mathcal{C}' = \{(r_1, r_2] \times (s_1, s_2], [0, r_2] \times (s_1, s_2] \subset \Delta : 0 \leq r_1 \leq r_2 \leq s_1 \leq s_2 \leq \infty\}$$

is a convergence-determining class which satisfies the conditions in Lemma 3.5.

We finish with the proof of Theorem 2.6.

Proof of Theorem 2.6. Suppose that R is a rectangle of the form $(r_1, r_2] \times (s_1, s_2]$ or $[0, r_1] \times (s_1, s_2]$ in Δ. By virtue of Lemma 3.5 and Lemma 3.6, we have only to show that there exists $c_R \geq 0$ such that for any convex averaging net $\mathcal{A} = \{A_n\}_{n \in \mathcal{N}}$,

$$\frac{1}{|A_n|}\mathbb{E}[\xi_{q,A_n}(R)] \to c_R \quad \text{as } n \to \infty$$

and if Φ is ergodic, then

$$\frac{1}{|A_n|}\xi_{q,A_n}(R) \to c_R \quad \text{a.s. as } n \to \infty$$

It follows immediately from Theorem 2.7 and the fact that $\xi_{q,A_n}(R)$ is calculated as

$$\begin{aligned}\xi_{q,A_n}(R) &= \xi_{q,A_n}([0,r_2]\times(s_1,\infty)) - \xi_{q,A_n}([0,r_2]\times(s_2,\infty)) \\ &\quad + \xi_{q,A_n}([0,r_1]\times(s_2,\infty)) - \xi_{q,A_n}([0,r_1]\times(s_1,\infty)) \\ &= \beta^{r_2,s_1}_{q,A_n} - \beta^{r_2,s_2}_{q,A_n} + \beta^{r_1,s_2}_{q,A_n} - \beta^{r_1,s_1}_{q,A_n}\end{aligned}$$

for $R = (r_1, r_2] \times (s_1, s_2]$ and

$$\xi_{q,A_n}(R) = \beta^{r_1,s_1}_{q,A_n} - \beta^{r_1,s_2}_{q,A_n}$$

for $R = [0, r_1] \times (s_1, s_2]$. Thus we arrive at the desired result. □

Funding

This work was supported by JST CREST Grant Number JPMJCR15D3, Japan. The first named author (T.S.) was supported by JSPS KAKENHI Grant Numbers JP18H01124, JP20K20884, and JSPS Grant-in-Aid for Transformative Research Areas (A) JP22H05105. T.S. was also supported in part by JSPS KAKENHI Grant Numbers, JP20H00119 and JP21H04432.

References

[1] Baccelli, F., Blaszczyszyn, B.: Stochastic Geometry and Wireless Networks, Volume I – Theory. Found. Trends Netw. **3**(3-4), 249–449, (2009). Vol. 1, pp. 150. NoW Publishers. https://doi.org/10.1561/1300000006, https://hal.inria.fr/inria-00403039. Stochastic Geometry and Wireless Networks, Volume II – Applications; see http://hal.inria.fr/inria-00403040.

[2] Bell, G., Lawson, A., Martin, J., Rudzinski, J., Smyth, C.: Weighted persistent homology. Involve **12**(5), 823–837 (2019). MR3954298. https://doi.org/10.2140/involve.2019.12.823

[3] Bobrowski, O., Mukherjee, S.: The topology of probability distributions on manifolds. Probab. Theory Related Fields **161**(3–4), 651–686 (2015). MR3334278. https://doi.org/10.1007/s00440-014-0556-x

[4] Bobrowski, O., Oliveira, G.: Random Čech complexes on Riemannian manifolds. Random Structures Algorithms **54**(3), 373–412 (2019). MR3938773. https://doi.org/10.1002/rsa.20800

[5] Buchet, M., Chazal, F., Oudot, S.Y., Sheehy, D.R.: Efficient and robust persistent homology for measures. Comput. Geom. **58**, 70–96 (2016). MR3541079. https://doi.org/10.1016/j.comgeo.2016.07.001

[6] Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J.: Stochastic Geometry and Its Applications, 3rd edn. Wiley Series in Probability and Statistics, p. 544. John Wiley & Sons, Ltd., Chichester (2013). MR3236788. https://doi.org/10.1002/9781118658222

[7] Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods, 2nd edn. Probability and its Applications (New York), p. 469. Springer, (2003). MR1950431

[8] Daley, D.J., Vere-Jones, D.: An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure, 2nd edn. Probability and its Applications (New York), p. 573. Springer, (2008). MR2371524. https://doi.org/10.1007/978-0-387-49835-5

[9] Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. In: Discrete and computational geometry and graph drawing, Columbia, SC, 2001, vol. 28, pp. 511–533 (2002). MR1949898. https://doi.org/10.1007/s00454-002-2885-2

[10] Fritz, J.: Generalization of McMillan's theorem to random set functions. Studia Sci. Math. Hungar. **5**, 369–394 (1970). MR293956

[11] Goel, A., Trinh, K.D., Tsunoda, K.: Strong law of large numbers for Betti numbers in the thermodynamic regime. J. Stat. Phys. **174**(4), 865–892 (2019). MR3913900. https://doi.org/10.1007/s10955-018-2201-z

[12] Hiraoka, Y., Shirai, T., Trinh, K.D.: Limit theorems for persistence diagrams. Ann. Appl. Probab. **28**(5), 2740–2780 (2018). MR3847972. https://doi.org/10.1214/17-AAP1371

[13] Meng, Z., Anand, D.V., Lu, Y., Wu, J., Xia, K.: Weighted persistent homology for biomolecular data analysis (2020). https://doi.org/10.1038/s41598-019-55660-3

[14] Nguyen, T.V., Baccelli, F.: On the Generating Functionals of a Class of Random Packing Point Processes (2013). arXiv: 1311.4967. https://doi.org/10.48550/ARXIV.1311.4967

[15] Nguyen, X.-X., Zessin, H.: Ergodic theorems for spatial processes. Z. Wahrsch. Verw. Gebiete **48**(2), 133–158 (1979). MR534841. https://doi.org/10.1007/BF01886869

[16] Pugh, C., Shub, M.: Ergodic elements of ergodic actions. Compos. Math. **23**, 115–122 (1971). MR283174

[17] Spitz, D., Wienhard, A.: The self-similar evolution of stationary point processes via persistent homology (2020). arXiv: 2012.05751. https://doi.org/10.48550/ARXIV.2012.05751

[18] Xuan-Xanh, N., Zessin, H.: Punktprozesse mit Wechselwirkung. Z. Wahrsch. Verw. Gebiete **37**(2), 91–126 (1976/77). MR423601. https://doi.org/10.1007/BF00536775

[19] Yogeshwaran, D., Adler, R.J.: On the topology of random complexes built over stationary point processes. Ann. Appl. Probab. **25**(6), 3338–3380 (2015). MR3404638. https://doi.org/10.1214/14-AAP1075

[20] Yogeshwaran, D., Subag, E., Adler, R.J.: Random geometric complexes in the thermodynamic regime. Probab. Theory Related Fields **167**(1–2), 107–142 (2017). MR3602843. https://doi.org/10.1007/s00440-015-0678-9

[21] Zomorodian, A., Carlsson, G.: Computing persistent homology. Discrete Comput. Geom. **33**(2), 249–274 (2005). MR2121296. https://doi.org/10.1007/s00454-004-1146-y