Adv. Appl. Prob. **53,** 57–80 (2021)
*doi:*10.1017/apr.2020.46

FUNCTIONAL LIMIT THEOREMS FOR THE EULER CHARACTERISTIC PROCESS IN THE CRITICAL REGIME

ANDREW M. THOMAS* AND
TAKASHI OWADA,* *Purdue University*

Abstract

This study presents functional limit theorems for the Euler characteristic of Vietoris–Rips complexes. The points are drawn from a nonhomogeneous Poisson process on $\mathbb{R}^d$, and the connectivity radius governing the formation of simplices is taken as a function of the time parameter t, which allows us to treat the Euler characteristic as a stochastic process. The setting in which this takes place is that of the critical regime, in which the simplicial complexes are highly connected and have nontrivial topology. We establish two 'functional-level' limit theorems, a strong law of large numbers and a central limit theorem, for the appropriately normalized Euler characteristic process.

Keywords: Functional central limit theorem; functional strong law of large numbers; Euler characteristic; geometric simplicial complex

2010 Mathematics Subject Classification: Primary 60F17
Secondary 55U10; 60C05; 60D05

1. Introduction

The Euler characteristic is one of the oldest and simplest topological summaries. It is at once local and global, combinatorial and topological, owing to its representation as either the alternating sum of Betti numbers of a topological space, or the alternating sum of simplices in its triangulation. Beyond its theoretical beauty, the Euler characteristic has recently made its way into the field of applied mathematics, notably topological data analysis (TDA). For instance, the Euler characteristic of sublevel (or superlevel) sets of random fields has found broad applications [1, 8]. In TDA, the technique of capturing the dynamic evolution of topology is generally studied in *persistent homology*—see [7] for a good introduction. Persistent homology originated in computational topology [10] and has received much attention as a useful machinery for exploring the manner in which topological holes appear and/or disappear in a filtered topological space. The primary objective of the current study is to associate the Euler characteristic with some filtered topological space by treating it as a stochastic process in the time parameter t.

Due to recent rapid development of TDA in conjunction with probability theory, there has been a growing interest in the study of random geometric complexes. We focus on the *Vietoris–Rips complex* [15, 16, 19], because of its ease of application, especially in

Received 1 October 2019; revision received 19 June 2020.
* Postal address: Department of Statistics, Purdue University, West Lafayette, IN 47907, USA. Email address: thoma186@purdue.edu

computational topology; though much research has also been done on the *Čech complex* [4, 6, 9, 11, 15, 16, 20, 25], and on the notion of generalizing both types of complex [13]. An elegant survey of progress in these areas can be found in [5]. These studies are mostly concerned with the asymptotic behavior of topological invariants such as the Euler characteristic and Betti numbers. Among them, [9] derived a concentration inequality for the Euler characteristic built over a Čech complex on a d-dimensional torus, as well as its asymptotic mean and variance; and [14] established a multivariate central limit theorem for the intrinsic volumes, including the Euler characteristic. Furthermore, [23] proved ergodic theorems for the Euler characteristic over a stationary and ergodic point process.

Most of the studies cited in the last paragraph start with either an independent and identically distributed (i.i.d.) random sample $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ or a Poisson point process $\mathcal{P}_n = \{X_1, \ldots, X_{N_n}\}$, where N_n is a Poisson random variable with mean n, independent of $(X_i)_i$. Subsequently, we will consider a simple Boolean model of the union of balls centered around $\mathcal{X}_n$ or $\mathcal{P}_n$ with a sequence of non-random radii $s_n \to 0$, $n \to \infty$. Then the behavior of topological invariants based on the Boolean model can be split up into several distinct regimes. When $ns_n^d \to 0$ as $n \to \infty$, we have what is called the *sparse* (or subcritical) regime, in which there occur many small connected components. If $ns_n^d \to \infty$ as $n \to \infty$, we have the *dense* (or supercritical) regime, which is characterized by a large connected component with few topological holes as a result of a slower decay rate of s_n. The intermediate case where ns_n^d converges to a positive and finite constant is called the *critical* regime, in which the stochastic features of a geometric complex are less assured, and arguably more interesting, owing to the emergence of highly connected components with nontrivial topologies. The present study focuses exclusively on the critical regime. This is because the behaviors of the Euler characteristic in other regimes, e.g. the sparse and dense regimes, are essentially trivial. For example, in the dense regime, the Euler characteristic is asymptotic to 1 (see [4]).

Within the context of geometric complexes—such as the Čech and Vietoris–Rips complexes—few attempts have been made thus far at deriving limit theorems on the functional level for topological invariants (for some exceptions, see [3, 19, 20]). From the viewpoint of persistent homology, such functional information is crucial for the understanding of topological invariants in a filtered topological space. With this in mind, the current study proceeds to establish functional limit theorems for the Euler characteristic defined as a stochastic process. More specifically, we shall prove a functional strong law of large numbers and a functional central limit theorem in the space $D[0, \infty)$ of right-continuous functions with left limits. Our results are the first functional limit theorems in the literature for a topological invariant under the critical regime that have neither time/radius restrictions nor restrictions on the number/size of components in the underlying simplicial complex. The primary benefit in our results lies in information obtainable about topological changes as the time parameter t varies. For example, if we let $\chi_n(t)$ be the Euler characteristic considered as a stochastic process, then as consequences of our main theorems, one can capture the limiting behavior of various useful functions of the Euler characteristic process via the continuous mapping theorem. We elaborate on these at the end of Section 3. Other potential applications can be found in Chapter 14 of [2] and in [24].

In Section 2 we discuss all the topological background necessary for the paper. In Section 3 we discuss our main results: the functional strong law of large numbers and functional central limit theorems for the Euler characteristic process in the critical regime. All of the proofs in the paper are collected in Section 4.

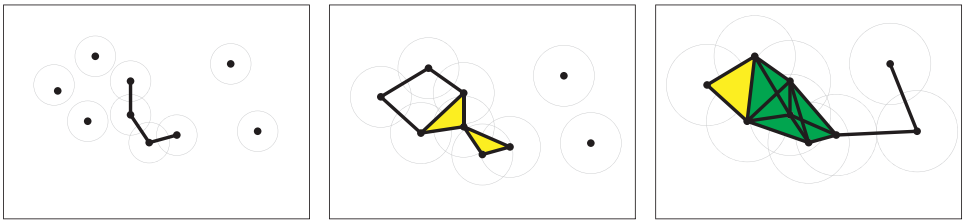

FIGURE 1. A family of Vietoris–Rips complexes.

2. Preliminaries

2.1. Topology

The main concept in the present paper is the Euler characteristic. Before introducing it we begin with the notions of a *simplex* and an *(abstract) simplicial complex*. Let $\mathbb{N}$, $\mathbb{N}_0$ be the positive and nonnegative integers respectively, and let $B(x, r)$ be the closed ball centered at x with radius $r \geq 0$.

Definition 2.1. Let $\mathcal{X}$ be a finite set. An abstract simplicial complex $\mathcal{K}$ is a collection of nonempty subsets of $\mathcal{X}$ which satisfy the following conditions:

1. All singleton subsets of $\mathcal{X}$ are in $\mathcal{K}$.
2. If $\sigma \in \mathcal{K}$ and $\tau \subset \sigma$, then $\tau \in \mathcal{K}$.

If $\sigma \in \mathcal{K}$ and $|\sigma| = k + 1$, with $k \in \mathbb{N}_0$, then σ is said to have dimension k and is called a *k-simplex* in $\mathcal{K}$. The dimension of $\mathcal{K}$ is the dimension of the largest simplex in $\mathcal{K}$.

It can be shown (cf. [10]) that every abstract simplicial complex $\mathcal{K}$ of dimension d can be embedded into $\mathbb{R}^{2d+1}$. The image of such an embedding, denoted geom($\mathcal{K}$), is called the *geometric realization* of $\mathcal{K}$. A topological space Y is said to be *triangulable* if there exists a simplicial complex $\mathcal{K}$ together with a homeomorphism between Y and geom($\mathcal{K}$). We now define the Euler characteristic.

Definition 2.2. Take $\mathcal{K}$ to be a simplicial complex and let $S_k(\mathcal{K})$ be the number of k-simplices in $\mathcal{K}$. Then the Euler characteristic of $\mathcal{K}$ is defined as

$$\chi(\mathcal{K}) := \sum_{k=0}^{\infty} (-1)^k S_k(\mathcal{K}).$$

If Y is a triangulable topological space with an associated simplicial complex $\mathcal{K}$, then we have $\chi(Y) = \chi(\mathcal{K})$, and $\chi(Y)$ is independent of the triangulation (see Theorem 2.44 in [12]). Therefore, the Euler characteristic is a topological invariant (and in fact a homotopy invariant).

Our setting for this study is always in $\mathbb{R}^d$, so we may take $\mathcal{X}$, $\mathcal{Y}$ to be arbitrary finite subsets of $\mathbb{R}^d$. To conclude this section, we will now define the Vietoris–Rips complex: the aforementioned simplicial complex that allows us to get a topological, as well as combinatorial, structure from our data $\mathcal{X}$. A family of Vietoris–Rips complexes $(\mathcal{R}(\mathcal{X}, t),\ t \geq 0)$ for points in $\mathbb{R}^2$ can be seen in Figure 1; yellow represents a 2-simplex and green represents a 3-simplex, which cannot be embedded in $\mathbb{R}^2$.

Definition 2.3. Let $\mathcal{X} = \{x_1, \dots, x_n\}$ be a finite subset of $\mathbb{R}^d$ and $t \geq 0$. The Vietoris–Rips complex $\mathcal{R}(\mathcal{X}, t)$ is the (abstract) simplicial complex with the following properties:

1. All singleton subsets of $\mathcal{X}$ are in $\mathcal{R}(\mathcal{X}, t)$.
2. A k-simplex $\sigma = \{x_{i_0}, \dots, x_{i_k}\}$ is in $\mathcal{R}(\mathcal{X}, t)$ if

$$B(x_{i_j}, t) \cap B(x_{i_\ell}, t) \neq \emptyset$$

for all $0 \leq j < \ell \leq k$.

2.2. Tools

Throughout, we let $\mathcal{P}_n$ denote a Poisson point process on $\mathbb{R}^d$ with intensity measure $n \int_A f(x)\mathrm{d}x$, where A is a Borel subset of $\mathbb{R}^d$, and f is a probability density function. Writing m for Lebesgue measure on $\mathbb{R}^d$, we assume that f is bounded almost everywhere, i.e., $\|f\|_\infty := \inf\{a \in \mathbb{R} : m(f^{-1}(a, \infty)) = 0\} < \infty$.

For two finite subsets $\mathcal{Y} \subset \mathcal{X}$ of $\mathbb{R}^d$ with $|\mathcal{Y}| = k + 1$, and $t \geq 0$, we define

$$h_t^k(\mathcal{Y}) := \mathbf{1}\{\mathcal{Y} \text{ forms a } k\text{-simplex in } \mathcal{R}(\mathcal{X}, t)\} = \prod_{x,y \in \mathcal{Y},\, x \neq y} \mathbf{1}\{B(x, t) \cap B(y, t) \neq \emptyset\}. \tag{2.1}$$

In the below we present obvious but highly useful properties of this indicator function. First, it is translation- and scale-invariant: for any $c > 0$, $x \in \mathbb{R}^d$, and $y_0, \dots, y_k \in \mathbb{R}^d$,

$$h_t^k(cy_0 + x, \dots, cy_k + x) = h_{t/c}^k(y_0, \dots, y_k).$$

Furthermore, for any fixed $y_i \in \mathbb{R}^d$, $i = 0, \dots, k$, it is nondecreasing in t, i.e.,

$$h_s^k(y_0, \dots, y_k) \leq h_t^k(y_0, \dots, y_k), \quad 0 \leq s \leq t. \tag{2.2}$$

Using (2.1), we can define k-simplex counts by $S_k(\mathcal{X}, t) := \sum_{\mathcal{Y} \subset \mathcal{X}} h_t^k(\mathcal{Y})$. As declared in the introduction, we shall exclusively focus on the critical regime, so that $ns_n^d \to 1$, $n \to \infty$. Finally, in order to formulate the Euler characteristic as a stochastic process, let $r_n(t) := s_n t$ and define

$$\chi_n(t) := \sum_{k=0}^{\infty} (-1)^k S_k(\mathcal{P}_n, r_n(t)) = \sum_{k=0}^{\infty} (-1)^k \sum_{\mathcal{Y} \subset \mathcal{P}_n} h_{r_n(t)}^k(\mathcal{Y}), \quad t \geq 0. \tag{2.3}$$

Notice that (2.3) is almost surely (a.s.) a finite sum, because the cardinality of $\mathcal{P}_n$, denoted by $|\mathcal{P}_n|$, is finite a.s., and $S_k(\mathcal{P}_n, r_n(t)) \equiv 0$ for all $k \geq |\mathcal{P}_n|$. Furthermore, for a Borel subset A of $\mathbb{R}^d$, define a restriction of the Euler characteristic to A by

$$\chi_{n,A}(t) := \sum_{k=0}^{\infty} (-1)^k \sum_{\mathcal{Y} \subset \mathcal{P}_n} h_{r_n(t)}^k(\mathcal{Y}) \mathbf{1}\{\mathrm{LMP}(\mathcal{Y}) \in A\}, \tag{2.4}$$

where $\mathrm{LMP}(\mathcal{Y})$ represents the leftmost point of $\mathcal{Y}$, i.e., the least point with respect to lexicographic order in $\mathbb{R}^d$. This restriction is useful for proving finite-dimensional convergence in the case when A is bounded. When A is bounded we get a finite number of random variables for the dependency graph, so that we may use Stein's method for normal approximation. See Section 4.3 for more details. Clearly, $\chi_{n,\mathbb{R}^d}(t) = \chi_n(t)$.

3. Main results

The first contribution of the present paper is the functional strong law of large numbers (FSLLN) for χ_n in the space $D[0, \infty)$ of right-continuous functions with left limits. More precisely, almost sure convergence of χ_n/n to the limiting mean will be established in terms of the uniform metric. Our proof techniques rely on the Borel–Cantelli lemma to prove a strong law of large numbers for each fixed t, and we then extend this to the functional case. As for the method of proofs in other studies, [22] and [25] have established concentration inequalities that can lead to the desired (static) strong law of large numbers. Although these concentration inequalities can yield sharper bounds, a downside is that extra conditions need to be put on an underlying density f. For example f must have bounded support. For this reason, we have adopted a different approach via the Borel–Cantelli lemma, by which one can prove $n^{-1}\big(\chi_n(t) - \mathbb{E}[\chi_n(t)]\big) \to 0$ a.s. by showing that the sum of the fourth moments is convergent. The relevant article taking an approach similar to ours is [11].

The second contribution of this paper is to show the weak convergence of the process

$$\bar{\chi}_n(t) := n^{-1/2}\big(\chi_n(t) - \mathbb{E}[\chi_n(t)]\big), \quad t \geq 0,$$

with respect to the Skorokhod J_1-topology. Proving finite-dimensional weak convergence of $\bar{\chi}_n$ in conjunction with its tightness will allow us to obtain the desired convergence in $D[0, \infty)$. Finite-dimensional convergence will be established via the Cramér–Wold device and Stein's method, as in Theorem 2.4 in [22], by adhering closely to the proof of Theorem 3.9 in the same source. In addition, the tightness will be proven via Theorem 13.5 in [2]. These functional limit theorems enable us to capture dynamic features of topological changes in $D[0, \infty)$. The proofs for all results in this section are postponed to Section 4.

In order to obtain a clear picture of our limit theorems, it will be beneficial to start with some results on asymptotic moments of χ_n. Define for $k_1, k_2 \in \mathbb{N}_0$, $t, s \geq 0$, and a Borel subset A of $\mathbb{R}^d$,

$$\Psi_{k_1,k_2,A}(t, s) := \sum_{j=1}^{(k_1 \wedge k_2)+1} \psi_{j,k_1,k_2,A}(t, s),$$

where $k_1 \wedge k_2 = \min\{k_1, k_2\}$, and

$$\begin{aligned}\psi_{j,k_1,k_2,A}(t, s) := {} & \frac{\int_A f(x)^{k_1+k_2+2-j}\mathrm{d}x}{j!(k_1+1-j)!(k_2+1-j)!} \\ & \times \int_{(\mathbb{R}^d)^{k_1+k_2+1-j}} h_t^{k_1}(0, y_1, \ldots, y_{k_1}) \\ & \times h_s^{k_2}(0, y_1, \ldots, y_{j-1}, y_{k_1+1}, \ldots, y_{k_1+k_2+1-j})\mathrm{d}\mathbf{y}.\end{aligned}$$

Here we set $h_t^k(0, y_1, \ldots, y_k) = 1$ if $k = 0$, so that $\Psi_{0,0,A}(t, s) = \psi_{1,0,0,A}(t, s) = \int_A f(x)dx$. In the sequel, we write $\Psi_{k_1,k_2}(t, s) := \Psi_{k_1,k_2,\mathbb{R}^d}(t, s)$ with $\psi_{j,k_1,k_2}(t, s) := \psi_{j,k_1,k_2,\mathbb{R}^d}(t, s)$.

Proposition 3.1. *For $t, s \geq 0$ and $A \subset \mathbb{R}^d$ open with $m(\partial A) = 0$, we have*

$$n^{-1}\mathbb{E}[\chi_{n,A}(t)] \to \sum_{k=0}^{\infty} (-1)^k \psi_{k+1,k,k,A}(t, t), \qquad n \to \infty, \tag{3.1}$$

$$n^{-1}\mathrm{Cov}\big(\chi_{n,A}(t),\, \chi_{n,A}(s)\big) \to \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}(-1)^{k_1+k_2}\Psi_{k_1,k_2,A}(t,s), \qquad n\to\infty, \tag{3.2}$$

so that both of the right-hand sides are convergent for every such $A \subset \mathbb{R}^d$.

We can now introduce the FSLLN for the process χ_n.

Theorem 3.1. (FSLLN for χ_n.) *As* $n\to\infty$,

$$\frac{\chi_n(t)}{n} \to \sum_{k=0}^{\infty}(-1)^k\psi_{k+1,k,k}(t,t) \ \text{ a.s. in } D[0,\infty),$$

where $D[0,\infty)$ *is equipped with the uniform topology.*

Before stating our functional central limit theorem (FCLT) for χ_n, let us define its limiting process. First define $(\mathcal{H}_k,\ k\in\mathbb{N}_0)$ as a family of zero-mean Gaussian processes on a generic probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with intra-process covariance

$$\mathbb{E}[\mathcal{H}_k(t)\mathcal{H}_k(s)] = \Psi_{k,k}(t,s), \tag{3.3}$$

and inter-process convariance

$$\mathbb{E}[\mathcal{H}_{k_1}(t)\mathcal{H}_{k_2}(s)] = \Psi_{k_1,k_2}(t,s), \tag{3.4}$$

for all $k, k_1, k_2 \in \mathbb{N}_0$ with $k_1 \neq k_2$ and $t, s \geq 0$. In the proof of Proposition 3.1, the functions $\Psi_{k_1,k_2}(t,s)$ naturally appear in the covariance calculation of χ_n, which in turn implies that the covariance functions in (3.3) and (3.4) are well-defined. With this notation in mind, we now define the limiting Gaussian process for $\bar{\chi}_n$ as

$$\mathcal{H}(t) := \sum_{k=0}^{\infty}(-1)^k\mathcal{H}_k(t), \qquad t\geq 0, \tag{3.5}$$

so that

$$\mathbb{E}[\mathcal{H}(t)\mathcal{H}(s)] = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}(-1)^{k_1+k_2}\Psi_{k_1,k_2}(t,s), \quad t,s\geq 0. \tag{3.6}$$

Once again, Proposition 3.1 implies that the right-hand side of (3.6) can define the covariance functions of a limiting Gaussian process, since it is obtained as a (scaled) limit of the covariance functions of χ_n. In particular, since (3.6) is convergent, for every $t\geq 0$, $\mathcal{H}(t)$ is definable in the $L^2(\Omega)$ sense. Note that the Euler characteristic in (2.3) and the process (3.5) exhibit similar structure, in the sense that $S_k\big(\mathcal{P}_n, r_n(t)\big)$ in (2.3) and $\mathcal{H}_k(t)$ both correspond to the spatial distribution of k-simplices.

Now, we proceed to stating the FCLT for χ_n.

Theorem 3.2. (FCLT for χ_n.) *As* $n\to\infty$,

$$\bar{\chi}_n \Rightarrow \mathcal{H} \ \text{ in } D[0,\infty),$$

where $D[0,\infty)$ *is equipped with the Skorokhod* J_1*-topology. Furthermore, for every* $0 < T < \infty$, *we have that* $\big(\mathcal{H}(t),\ 0\leq t\leq T\big)$ *has a continuous version with Hölder continuous sample paths of any exponent* $\gamma \in [0, 1/2)$.

Remark 3.1. The results of Theorem 3.1 and Theorem 3.2 also hold for the Čech complex, in the case of the latter theorem only up to finite-dimensional weak convergence of $\bar{\chi}_n$. The definition of a k-simplex of the Čech complex requires a non-empty intersection of 'multiple' closed balls. This makes it more difficult to establish the required tightness for the Čech complex. Specifically, obtaining bounds as in Lemma 4.2 seems much harder. If one were able to establish such a nice bound, the rest of the argument for tightness would essentially be the same as the Vietoris–Rips case.

Example 3.1. Consider a map $x \mapsto \sup_{0\le t\le 1} |x(t)|$ from $D[0, 1]$ to $\mathbb{R}_+$. This map is continuous on $C[0, 1]$, the space of continuous functions on $[0, 1]$. Since the limits in Theorems 3.1 and 3.2 are both continuous, we get that as $n \to \infty$,

$$n^{-1} \sup_{0\le t\le 1} |\chi_n(t)| \to \sup_{0\le t\le 1} \Big|\sum_{k=0}^{\infty} (-1)^k \psi_{k+1,k,k}(t, t)\Big| \quad a.s.,$$

$$n^{-1/2} \sup_{0\le t\le 1} \big|\chi_n(t) - \mathbb{E}[\chi_n(t)]\big| \Rightarrow \sup_{0\le t\le 1} |\mathcal{H}(t)|.$$

In particular, the latter claims that the supremum of a mean-centered Euler characteristic process can be approximated by $n^{1/2} \sup_{0\le t\le 1} |\mathcal{H}(t)|$ for large enough n.

4. Proofs

We first deal with moment asymptotics of χ_n, in Section 4.1. In Section 4.2 we prove the FSLLN in Theorem 3.1. Subsequently, we establish Theorem 3, the proof of which is divided into two parts, with the first part devoted to finite-dimensional weak convergence, and the second to tightness. The proofs frequently refer to *Palm theory for Poisson processes* for computing the moments of various Poisson functionals. A brief citation is given in Lemma 5.1 of the appendix. Finally we verify Hölder continuity of the limiting Gaussian process $\mathcal{H}$, adhering closely to what is established for subgraph counting processes in Proposition 4.2 of [18].

For simplicity of description, we assume throughout this section that $ns_n^d = 1$. However, generalizing it to $ns_n^d \to 1$, $n \to \infty$, is straightforward. In the following, we write $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$ for $a, b \in \mathbb{R}$.

4.1 Proof of moment asymptotics

Without loss of generality, the proof of Proposition 3.1 only handles the case when $A = \mathbb{R}^d$. Throughout this section, let $\mathcal{Y}$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$ denote collections of i.i.d. random points with density f. We begin with the following lemma.

Lemma 4.1.

(*i*) *For $t \ge 0$ we have, as $n \to \infty$,*

$$\frac{n^k}{(k+1)!}\, \mathbb{E}\left[h^k_{r_n(t)}(\mathcal{Y})\right] \to \psi_{k+1,k,k}(t, t).$$

(*ii*) *For all $n \in \mathbb{N}$,*

$$n^k \mathbb{E}\left[h^k_{r_n(t)}(\mathcal{Y})\right] \le (a_t)^k,$$

where

$$a_t := (2t)^d \theta_d \|f\|_\infty \tag{4.1}$$

with $\theta_d = m\big(B(0, 1)\big)$, i.e., the volume of the unit ball in $\mathbb{R}^d$.

(*iii*) *For* $1 \le j \le (k_1 \wedge k_2) + 1$, $k_1, k_2 \in \mathbb{N}_0$, *and* $t, s \ge 0$,

$$\frac{n^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!}\,\mathbb{E}\big[h^{k_1}_{r_n(t)}(\mathcal{Y}_1)h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = j\big\}\big] \to \psi_{j,k_1,k_2}(t,s)$$

as $n \to \infty$.

(*iv*) *For all* $n \in \mathbb{N}$,

$$n^{k_1+k_2+1-j}\mathbb{E}\big[h^{k_1}_{r_n(t)}(\mathcal{Y}_1)h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = j\big\}\big] \le (a_{t\vee s})^{k_1+k_2+1-j}.$$

Proof. We shall prove (*iii*) and (*iv*) only, since (*i*) and (*ii*) can be established by a similar and simpler argument. With the change of variables $x_1 = x$ and $x_i = x + s_n y_{i-1}$, $i = 1, \dots, k_1 + k_2 + 2 - j$, the left-hand side of (*iii*) equals

$$\begin{aligned}
&\frac{n^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!}\int_{(\mathbb{R}^d)^{k_1+k_2+2-j}} h^{k_1}_{r_n(t)}(x_1,\dots,x_{k_1+1}) \\
&\qquad\qquad \times h^{k_2}_{r_n(s)}(x_1,\dots,x_j,x_{k_1+2},\dots,x_{k_1+k_2+2-j}) \prod_{i=1}^{k_1+k_2+2-j} f(x_i)\mathrm{d}\mathbf{x} \\
&= \frac{(ns_n^d)^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!}\int_{\mathbb{R}^d}\int_{(\mathbb{R}^d)^{k_1+k_2+1-j}} h^{k_1}_t(0,y_1\dots,y_{k_1}) \qquad (4.2)\\
&\quad \times h^{k_2}_s(0,y_1,\dots,y_{j-1},y_{k_1+1},\dots,y_{k_1+k_2+1-j})f(x) \prod_{i=1}^{k_1+k_2+1-j} f(x+s_n y_i)\mathrm{d}\mathbf{y}\,\mathrm{d}x.
\end{aligned}$$

Recall that $ns_n^d = 1$ and note that $\prod_{i=1}^{k_1+k_2+1-j} f(x+s_n y_i) \to f(x)^{k_1+k_2+1-j}$, $n \to \infty$, holds under the integral sign because of the Lebesgue differentiation theorem. Thus, (4.2) converges to $\psi_{j,k_1,k_2}(t,s)$ as $n \to \infty$.

Now let us turn to proving (*iv*). Without loss of generality, we may assume $s \le t$. After the same change of variables as in (*iii*), the left-hand side of (*iv*) is bounded by

$$\begin{aligned}
&\big(\|f\|_\infty\big)^{k_1+k_2+1-j}\int_{(\mathbb{R}^d)^{k_1+k_2+1-j}} h^{k_1}_t(0,y_1\dots,y_{k_1}) \\
&\qquad\qquad \times h^{k_2}_s(0,y_1,\dots,y_{j-1},y_{k_1+1},\dots,y_{k_1+k_2+1-j})\mathrm{d}\mathbf{y}. \qquad (4.3)
\end{aligned}$$

By the definition of the indicators $h^{k_1}_t$, $h^{k_2}_s$, each of the y_i in (4.3) must have distance at most $2t$ from the origin. Therefore, (4.3) can be bounded by

$$\big(\|f\|_\infty\big)^{k_1+k_2+1-j} m\big(B(0,2t)\big)^{k_1+k_2+1-j} = (a_t)^{k_1+k_2+1-j}. \qquad \square$$

Proof of Proposition 3.1. We prove only (3.2), as the proof techniques for (3.1) are very similar to (3.2). Specifically, we shall make use of (*ii*), (*iii*), and (*iv*) of Lemma 4.1. We start by writing

$$\begin{aligned}
n^{-1}\mathrm{Cov}\big(\chi_n(t),\chi_n(s)\big) &= n^{-1}\mathbb{E}\Bigg[\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}(-1)^{k_1+k_2}S_{k_1}\big(\mathcal{P}_n,r_n(t)\big)S_{k_2}\big(\mathcal{P}_n,r_n(s)\big)\Bigg] \qquad (4.4)\\
&\quad - n^{-1}\mathbb{E}\Bigg[\sum_{k=0}^{\infty}(-1)^k S_k\big(\mathcal{P}_n,r_n(t)\big)\Bigg]\mathbb{E}\Bigg[\sum_{k=0}^{\infty}(-1)^k S_k\big(\mathcal{P}_n,r_n(s)\big)\Bigg].
\end{aligned}$$

Next, Palm theory for Poisson processes, Lemma 5.1(*ii*), along with the bounds given in Parts (*ii*) and (*iv*) of Lemma 4.1, yields that

$$\begin{aligned}
&\mathbb{E}\Big[S_{k_1}\big(\mathcal{P}_n, r_n(t)\big)S_{k_2}\big(\mathcal{P}_n, r_n(s)\big)\Big] \\
&= \sum_{j=0}^{(k_1\wedge k_2)+1} \mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n} h^{k_1}_{r_n(t)}(\mathcal{Y}_1)h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\,\mathbf{1}\big\{|\mathcal{Y}_1\cap\mathcal{Y}_2|=j\big\}\Big] \\
&= \frac{n^{k_1+k_2+2}}{(k_1+1)!(k_2+1)!}\,\mathbb{E}\big[h^{k_1}_{r_n(t)}(\mathcal{Y}_1)\big]\mathbb{E}\big[h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\big] \\
&\quad + \sum_{j=1}^{(k_1\wedge k_2)+1} \frac{n^{k_1+k_2+2-j}}{j!(k_1+1-j)!(k_2+1-j)!}\,\mathbb{E}\Big[h^{k_1}_{r_n(t)}(\mathcal{Y}_1)h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\mathbf{1}\big\{|\mathcal{Y}_1\cap\mathcal{Y}_2|=j\big\}\Big] \\
&\le \frac{n^2(a_t)^{k_1}(a_s)^{k_2}}{(k_1+1)!(k_2+1)!} + \sum_{j=1}^{(k_1\wedge k_2)+1} \frac{n(a_{t\vee s})^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!}.
\end{aligned}$$

Here it is straightforward to see that

$$\sum_{k=0}^{\infty}\frac{(a_t)^k}{(k+1)!} < e^{a_t} < \infty, \quad \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{j=1}^{(k_1\wedge k_2)+1}\frac{(a_{t\vee s})^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!} < 2e^{3a_{t\vee s}} < \infty.$$

So Fubini's theorem is applicable to the first term in (4.4). Repeating the same argument for the second term of (4.4), one can get

$$\begin{aligned}
n^{-1}\mathrm{Cov}\big(\chi_n(t), \chi_n(s)\big) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}(-1)^{k_1+k_2}\sum_{j=1}^{(k_1\wedge k_2)+1}\frac{n^{k_1+k_2+1-j}}{j!(k_1+1-j)!(k_2+1-j)!} \\
\times\,\mathbb{E}\Big[h^{k_1}_{r_n(t)}(\mathcal{Y}_1)h^{k_2}_{r_n(s)}(\mathcal{Y}_2)\mathbf{1}\big\{|\mathcal{Y}_1\cap\mathcal{Y}_2|=j\big\}\Big].
\end{aligned}$$

By virtue of Lemma 4.1(*iii*)–(*iv*), the dominated convergence theorem implies that the last expression converges to $\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}(-1)^{k_1+k_2}\Psi_{k_1,k_2}(t,s)$ as required. □

4.2. Proof of FSLLN

To prove the FSLLN, we first establish a result which allows us to extend a 'pointwise' strong law for a fixed t into a functional one, if the processes are nondecreasing and there is a deterministic and continuous limit. We again would like to emphasize that our approach in this section gives an improvement from the viewpoint of assumptions on the density f. Unlike the existing results, such as those of [25], ours do not require f to have compact and convex support.

Proposition 4.1. *Let $(X_n,\ n\in\mathbb{N})$ be a sequence of random elements in $D[0,\infty)$ with nondecreasing sample paths. Suppose $\lambda:[0,\infty)\to\mathbb{R}$ is a continuous and nondecreasing function. If we have*

$$X_n(t)\to\lambda(t), \qquad n\to\infty, \quad a.s. \tag{4.5}$$

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

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

for every $0 \leq T < \infty$. Hence, it holds that $X_n \to \lambda$ a.s. in $D[0, \infty)$ endowed with the uniform topology.

Proof. Fix $0 \leq T < \infty$. Note that λ is uniformly continuous on $[0, T]$. Given $\epsilon > 0$, choose $k = k(\epsilon) \in \mathbb{N}$ such that for all $s, t \in [0, T]$,

$$|s - t| \leq 1/k \quad \text{implies} \quad \big|\lambda(s) - \lambda(t)\big| < \epsilon. \tag{4.6}$$

Since $X_n(t)$ and $\lambda(t)$ are both nondecreasing in t, we see that

$$\begin{aligned}
\sup_{t\in[0,T]} \big|X_n(t) - \lambda(t)\big| &= \max_{1\leq i\leq k} \sup_{t\in[(i-1)T/k,\, iT/k]} \big|X_n(t) - \lambda(t)\big| \\
&\leq \max_{1\leq i\leq k} \left\{ \Big(X_n(iT/k) - \lambda((i-1)T/k)\Big) \vee \Big(\lambda(iT/k) - X_n((i-1)T/k)\Big) \right\} \\
&\leq \max_{1\leq i\leq k} \left\{ \Big(X_n(iT/k) - \lambda(iT/k)\Big) \vee \Big(\lambda((i-1)T/k) - X_n((i-1)T/k)\Big) \right\} + \epsilon \\
&\leq \max_{0\leq i\leq k} \big|X_n(iT/k) - \lambda(iT/k)\big| + \epsilon,
\end{aligned}$$

where the second inequality follows from (4.6). By the SLLN in (4.5), the last expression tends to ϵ a.s. as $n \to \infty$. Since ϵ is arbitrary, this completes the proof. □

Proof of Theorem 3.1. Since (2.3) is a.s. represented as a sum of finitely many terms, it can be split into two parts:

$$\chi_n(t) = \sum_{k=0}^{\infty} S_{2k}\big(\mathcal{P}_n, r_n(t)\big) - \sum_{k=0}^{\infty} S_{2k+1}\big(\mathcal{P}_n, r_n(t)\big) =: \chi_n^{(1)}(t) - \chi_n^{(2)}(t) \quad \text{a.s.}$$

Denoting by $K(t)$ the limit of (3.1) with $A = \mathbb{R}^d$, we decompose it in a way similar to the above:

$$K(t) = \sum_{k=0}^{\infty} \psi_{2k+1,2k,2k}(t, t) - \sum_{k=0}^{\infty} \psi_{2k+2,2k+1,2k+1}(t, t) =: K^{(1)}(t) - K^{(2)}(t).$$

Our final goal is to prove that for every $0 < T < \infty$,

$$\sup_{0\leq t\leq T} \left| \frac{\chi_n(t)}{n} - K(t) \right| \to 0, \qquad n \to \infty, \quad \text{a.s.},$$

which is clearly implied by

$$\sup_{0\leq t\leq T} \left| \frac{\chi_n^{(i)}(t)}{n} - K^{(i)}(t) \right| \to 0, \qquad n \to \infty, \quad \text{a.s.}$$

for each $i = 1, 2$. As $\chi_n^{(i)}(t)/n$ and $K^{(i)}(t)$ satisfy the conditions of Proposition 4.1, it suffices to show that

$$\frac{\chi_n^{(i)}(t)}{n} \to K^{(i)}(t), \quad n \to \infty, \quad \text{a.s.}$$

for every $t \geq 0$. We will prove only the case $i = 1$, and henceforth omit the superscript (1) from $\chi_n^{(1)}(t)$ and $K^{(1)}(t)$. It then suffices to show that

$$n^{-1}|\chi_n(t) - \mathbb{E}[\chi_n(t)]| \to 0, \qquad n \to \infty, \quad \text{a.s.}, \tag{4.7}$$

and

$$\left|n^{-1}\mathbb{E}[\chi_n(t)] - K(t)\right| \to 0, \qquad n \to \infty. \tag{4.8}$$

First we will deal with (4.8). It follows from the customary change of variables as in the proof of Lemma 4.1 that

$$\begin{aligned}
&\left|n^{-1}\mathbb{E}[\chi_n(t)] - K(t)\right| \\
&\quad = \left| \sum_{k=1}^{\infty} \frac{1}{(2k+1)!} \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{2k}} h_t^{2k}(0, y_1, \ldots, y_{2k}) \right. \\
&\qquad\qquad \left. \times f(x)\Big(\prod_{i=1}^{2k} f(x + s_n y_i) - f(x)^{2k} \Big) \mathbf{dy}\, \mathrm{d}x \right| \\
&\quad \leq \sum_{k=1}^{\infty} \frac{1}{(2k+1)!} \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{2k}} h_t^{2k}(0, y_1, \ldots, y_{2k}) f(x) \left| \prod_{i=1}^{2k} f(x + s_n y_i) - f(x)^{2k} \right| \mathbf{dy}\, \mathrm{d}x.
\end{aligned}$$

Similarly to the proof of Lemma 4.1 Part (*ii*) or (*iv*), one can show that the last term above is bounded by $2\sum_{k=1}^{\infty} (a_t)^{2k}/(2k+1)! < \infty$ (where a_t is defined in (4.1)). Thus, the dominated convergence theorem concludes the proof of (4.8).

Now, let us turn our attention to (4.7). From the Borel–Cantelli lemma it suffices to show that, for every $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\Big(\left|\chi_n(t) - \mathbb{E}[\chi_n(t)]\right| > \epsilon n \Big) < \infty.$$

By Markov's inequality, the left-hand side above is bounded by

$$\frac{1}{\epsilon^4} \sum_{n=1}^{\infty} \frac{1}{n^4} \mathbb{E}\Big[\big(\chi_n(t) - \mathbb{E}[\chi_n(t)]\big)^4 \Big].$$

Since $\sum_n n^{-2} < \infty$, we only need to show that

$$\limsup_{n\to\infty} \frac{1}{n^2} \mathbb{E}\Big[\big(\chi_n(t) - \mathbb{E}[\chi_n(t)]\big)^4 \Big] < \infty. \tag{4.9}$$

Applying Fubini's theorem as in the proof of Proposition 3.1, along with Hölder's inequality, we get that

$$\begin{aligned}
&\frac{1}{n^2}\, \mathbb{E}\Big[\big(\chi_n(t) - \mathbb{E}[\chi_n(t)]\big)^4 \Big] \\
&= \frac{1}{n^2} \sum_{(k_1,\ldots,k_4)\in\mathbb{N}^4} \mathbb{E}\Big[\prod_{i=1}^{4} \Big(S_{2k_i}\big(\mathcal{P}_n, r_n(t)\big) - E\big[S_{2k_i}\big(\mathcal{P}_n, r_n(t)\big)\big] \Big) \Big] \\
&\leq \Big[\sum_{k=1}^{\infty} \Big\{ \frac{1}{n^2} \mathbb{E}\Big[\Big(S_{2k}\big(\mathcal{P}_n, r_n(t)\big) - \mathbb{E}\big[S_{2k}\big(\mathcal{P}_n, r_n(t)\big)\big] \Big)^4 \Big] \Big\}^{1/4} \Big]^4.
\end{aligned}$$

Now, (4.9) can be obtained if we show that

$$\sum_{k=1}^{\infty} \left\{ \limsup_{n\to\infty} \frac{1}{n^2}\, \mathbb{E}\Big[\Big(S_{2k}\big(\mathcal{P}_n, r_n(t)\big) - \mathbb{E}\big[S_{2k}\big(\mathcal{P}_n, r_n(t)\big)\big]\Big)^4\Big] \right\}^{1/4} < \infty. \tag{4.10}$$

From this point on, let us introduce the shorthand $S_{2k} \coloneqq S_{2k}\big(\mathcal{P}_n, r_n(t)\big)$. In order to find an appropriate upper bound for (4.10), using the binomial expansion we write

$$\mathbb{E}\big[\big(S_{2k} - \mathbb{E}[S_{2k}]\big)^4\big] = \sum_{\ell=0}^{4} \binom{4}{\ell} (-1)^{\ell}\, \mathbb{E}[S_{2k}^{\ell}]\big(\mathbb{E}[S_{2k}]\big)^{4-\ell}. \tag{4.11}$$

For every $\ell \in \{0, \dots, 4\}$, one can write $\mathbb{E}[S_{2k}^{\ell}]\big(\mathbb{E}[S_{2k}]\big)^{4-\ell}$ as

$$\mathbb{E}\Bigg[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n^{(1)}} \sum_{\mathcal{Y}_2 \subset \mathcal{P}_n^{(2)}} \sum_{\mathcal{Y}_3 \subset \mathcal{P}_n^{(3)}} \sum_{\mathcal{Y}_4 \subset \mathcal{P}_n^{(4)}} \prod_{i=1}^{4} h_{r_n(t)}^{2k}(\mathcal{Y}_i) \Bigg], \tag{4.12}$$

where for every $i, j \in \{1, \dots, 4\}$, we have either that $\mathcal{P}_n^{(i)} = \mathcal{P}_n^{(j)}$ or that $\mathcal{P}_n^{(i)}$ is an independent copy of $\mathcal{P}_n^{(j)}$. If $|\mathcal{Y}_1 \cup \cdots \cup \mathcal{Y}_4| = 8k + 4$, i.e., $\mathcal{Y}_1, \dots, \mathcal{Y}_4$ do not have any common elements, then Palm theory (Lemma 5.1) shows that (4.12) is equal to $\big(\mathbb{E}[S_{2k}]\big)^4$, which grows at the rate of $O(n^4)$ (see Lemma 4.1(*i*)). In this case, the total contribution to (4.11) disappears, because

$$\sum_{\ell=0}^{4} \binom{4}{\ell} (-1)^{\ell} \big(\mathbb{E}[S_{2k}]\big)^4 = 0.$$

Suppose next that $|\mathcal{Y}_1 \cup \cdots \cup \mathcal{Y}_4| = 8k + 3$; that is, there is exactly one common element between $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for some $i \neq j$, with no other overlappings. Then (4.12) is equal to

$$\mathbb{E}\Bigg[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n} \sum_{\mathcal{Y}_2 \subset \mathcal{P}_n} h_{r_n(t)}^{2k}(\mathcal{Y}_1) h_{r_n(t)}^{2k}(\mathcal{Y}_2)\, \mathbf{1}\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = 1\} \Bigg] \big(\mathbb{E}[S_{2k}]\big)^2.$$

Although the growth rate of the above term is $O(n^3)$ (see Lemma 4.1 Parts (*i*) and (*iii*)), an overall contribution to (4.11) is again canceled. This is because

$$\begin{aligned} &\left\{ \binom{4}{2}(-1)^2 + \binom{4}{3}(-1)^3 \binom{3}{2} + \binom{4}{4}(-1)^4 \binom{4}{2} \right\} \\ &\quad \times \mathbb{E}\Bigg[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n} \sum_{\mathcal{Y}_2 \subset \mathcal{P}_n} h_{r_n(t)}^{2k}(\mathcal{Y}_1) h_{r_n(t)}^{2k}(\mathcal{Y}_2)\, \mathbf{1}\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = 1\} \Bigg] \big(\mathbb{E}[S_{2k}]\big)^2 = 0. \end{aligned}$$

By the above discussion, we only need to consider the case where there are at least two common elements within $\mathcal{Y}_1, \dots, \mathcal{Y}_4$. Among many such cases, let us deal with a specific term,

$$\begin{aligned} n^{-2}\mathbb{E}\Bigg[&\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n} \sum_{\mathcal{Y}_2 \subset \mathcal{P}_n} \sum_{\mathcal{Y}_3 \subset \mathcal{P}_n} \sum_{\mathcal{Y}_4 \subset \mathcal{P}_n} \prod_{i=1}^{4} h_{r_n(t)}^{2k}(\mathcal{Y}_i) \\ &\times \mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = \ell_1,\ |\mathcal{Y}_3 \cap \mathcal{Y}_4| = \ell_2,\ \big|(\mathcal{Y}_1 \cup \mathcal{Y}_2) \cap (\mathcal{Y}_3 \cup \mathcal{Y}_4)\big| = 0\big\} \Bigg], \end{aligned} \tag{4.13}$$

where $\ell_1, \ell_2 \in \{1, \dots, 2k+1\}$. Palm theory allows us to write (4.13) as

$$\prod_{i=1}^{2} \frac{n^{4k+1-\ell_i}}{\ell_i!\big((2k+1-\ell_i)!\big)^2}\, \mathbb{E}\Big[h^{2k}_{r_n(t)}(\mathcal{Y}_1) h^{2k}_{r_n(t)}(\mathcal{Y}_2)\, \mathbf{1}\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = \ell_i\}\Big]. \tag{4.14}$$

By Lemma 4.1(*iv*) and the fact that $\ell!(2k+1-\ell)! \geq k!$ for any $\ell \in \{1, \dots, 2k+1\}$, one can bound (4.14) by

$$\prod_{i=1}^{2} \frac{(a_t)^{4k+1-\ell_i}}{\ell_i!\big((2k+1-\ell_i)!\big)^2} \leq \frac{(a_t)^{8k+2-\ell_1-\ell_2}}{k!}.$$

Now, the ratio test shows that

$$\sum_{k=1}^{\infty} \left\{ \frac{(a_t)^{8k+2-\ell_1-\ell_2}}{k!} \right\}^{1/4} < \infty$$

as desired. Notice that all the cases except (4.13) can be handled in a very similar way, and so (4.10) follows.

4.3. Proof of finite-dimensional convergence in Theorem 3.2

Proof of finite-dimensional convergence in Theorem 3.2. Throughout the proof, C^* denotes a generic positive constant that potentially varies across and within the lines. Recall (2.4), and define $\bar{\chi}_{n,A}(t)$ analogously to $\bar{\chi}_n(t)$ by mean-centering and scaling by $n^{-1/2}$. We first consider the case where A is an open and bounded subset of $\mathbb{R}^d$ with $m(\partial A) = 0$.

From the viewpoint of the Cramér–Wold device, one needs to establish weak convergence of $\sum_{i=1}^{m} a_i \bar{\chi}_n(t_i)$ for every $0 < t_1 < \cdots < t_m$, $m \in \mathbb{N}$, and $a_i \in \mathbb{R}$, $i = 1, \dots, m$. Our proof exploits Stein's normal approximation method in Theorem 2.4 of [22]. Let $(Q_{j,n},\ j \geq 1)$ be an enumeration of disjoint subsets of $\mathbb{R}^d$ congruent to $(0, r_n(t_m)]^d$, such that $\mathbb{R}^d = \bigcup_{j=1}^{\infty} Q_{j,n}$. Let $H_n = \{j \in \mathbb{N} : Q_{j,n} \cap A \neq \emptyset\}$. Define

$$\xi_{j,n} := \sum_{k=0}^{\infty} (-1)^k \sum_{\mathcal{Y} \subset \mathcal{P}_n} \sum_{i=1}^{m} a_i h^k_{r_n(t_i)}(\mathcal{Y}) \mathbf{1}\big\{\mathrm{LMP}(\mathcal{Y}) \in A \cap Q_{j,n}\big\},$$

and also

$$\bar{\xi}_{j,n} := \frac{\xi_{j,n} - \mathbb{E}[\xi_{j,n}]}{\sqrt{\mathrm{Var}\big(\sum_{i=1}^{m} a_i \chi_{n,A}(t_i)\big)}}.$$

Then, we have $\sum_{i=1}^{m} a_i \chi_{n,A}(t_i) = \sum_{j \in H_n} \xi_{j,n}$.

Now, we define H_n to be the vertex set of a *dependency graph* (see Section 2.1 of [22] for the formal definition) for the random variables $(\bar{\xi}_{j,n},\ j \in H_n)$ by setting $j \sim j'$ if and only if the condition

$$\inf\big\{\|x - y\| : x \in Q_{j,n},\ y \in Q'_{j,n}\big\} \leq 4 r_n(t_m)$$

is satisfied. This is because $\xi_{j,n}$ and $\xi_{j',n}$ become independent whenever $j \sim j'$ fails to hold. Now we must ensure that the other conditions of Theorem 2.4 in [22] are satisfied with respect to the dependency graph $(H_n, \sim)$. First, $\sum_{j \in H_n} \bar{\xi}_{j,n}$ is a zero-mean random variable with unit variance. We know that $|H_n| = O(s_n^{-d})$ as A is bounded. Furthermore, the maximum degree of

any vertex of H_n is uniformly bounded by a positive and finite constant. Let Z denote a standard normal random variable. Then the aforementioned theorem implies that

$$\Big|\mathbb{P}\Big(\sum_{j\in H_n}\bar{\xi}_{j,n}\le x\Big)-\mathbb{P}(Z\le x)\Big|\le C^*\left(\sqrt{s_n^{-d}\max_j\mathbb{E}\big[|\bar{\xi}_{j,n}|^3\big]}+\sqrt{s_n^{-d}\max_j\mathbb{E}\big[|\bar{\xi}_{j,n}|^4\big]}\right)$$
$$\le C^*\left(\sqrt{s_n^{-d}n^{-3/2}\max_j\mathbb{E}\big[|\xi_{j,n}-\mathbb{E}[\xi_{j,n}]|^3\big]}+\sqrt{s_n^{-d}n^{-2}\max_j\mathbb{E}\big[|\xi_{j,n}-\mathbb{E}[\xi_{j,n}]|^4\big]}\right), \quad (4.15)$$

where the second inequality follows from Proposition 3.1, which claims that $\mathrm{Var}\big(\sum_{i=1}^m a_i\chi_{n,A}(t_i)\big)$ is asymptotically equal to n up to multiplicative constants. Minkowski's inequality implies that

$$\Big(\mathbb{E}\big[|\xi_{j,n}-\mathbb{E}[\xi_{j,n}]|^p\big]\Big)^{1/p}\le\big(\mathbb{E}\big[|\xi_{j,n}|^p\big]\big)^{1/p}+\mathbb{E}\big[|\xi_{j,n}|\big].$$

Recall that for fixed $\mathcal{Y}\subset\mathbb{R}^d$, $h_t^k(\mathcal{Y})$ is nondecreasing in t. Then, we have that

$$|\xi_{j,n}|\le\sum_{k=0}^{\infty}\sum_{\mathcal{Y}\subset\mathcal{P}_n}\sum_{i=1}^{m}|a_i|h^k_{r_n(t_i)}(\mathcal{Y})\mathbf{1}\big\{\mathrm{LMP}(\mathcal{Y})\in A\cap Q_{j,n}\big\}$$
$$\le C^*\sum_{k=0}^{\infty}\sum_{\mathcal{Y}\subset\mathcal{P}_n}h^k_{r_n(t_m)}(\mathcal{Y})\mathbf{1}\big\{\mathrm{LMP}(\mathcal{Y})\in A\cap Q_{j,n}\big\}$$
$$\le C^*\sum_{k=0}^{\infty}\binom{\mathcal{P}_n\big(\mathrm{Tube}(Q_{j,n},\,2r_n(t_m))\big)}{k+1}$$
$$\le C^*\cdot 2^{\mathcal{P}_n(\mathrm{Tube}(Q_{j,n},\,2r_n(t_m)))},$$

where

$$\mathrm{Tube}\big(Q_{j,n},\,2r_n(t_m)\big)=\big\{x\in\mathbb{R}^d:\inf_{y\in Q_{j,n}}\|x-y\|\le 2r_n(t_m)\big\}.$$

By the assumption $ns_n^d=1$, one can easily show that $\mathcal{P}_n\big(\mathrm{Tube}(Q_{j,n},\,2r_n(t_m))\big)$ is stochastically dominated by a Poisson random variable with positive and finite parameter, which does not depend on j and n. Denote such a Poisson random variable by Y. Then, for $p=3,4$,

$$\max_j\mathbb{E}\Big[\big|\xi_{j,n}-\mathbb{E}[\xi_{j,n}]\big|^p\Big]\le C^*\Big[\big(\mathbb{E}[2^{pY}]\big)^{1/p}+\mathbb{E}(2^Y)\Big]<\infty.$$

Referring back to (4.15) and noting $ns_n^d=1$, we can see that

$$\Big|\mathbb{P}\Big(\sum_{j\in H_n}\bar{\xi}_{j,n}\le x\Big)-\mathbb{P}(Z\le x)\Big|\le C^*\Big(\sqrt{s_n^{-d}n^{-3/2}}+\sqrt{s_n^{-d}n^{-2}}\Big)=O(n^{-1/4})\to 0,\qquad n\to\infty,$$

which implies that $\sum_{j\in H_n}\bar{\xi}_{j,n}\Rightarrow\mathcal{N}(0,\,1)$ as $n\to\infty$; equivalently,

$$\sum_{i=1}^{m}a_i\bar{\chi}_{n,A}(t_i)\Rightarrow\mathcal{N}(0,\,\Sigma_A),\qquad n\to\infty,$$

where

$$\Sigma_A := \sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j \sum_{k_i=0}^{\infty}\sum_{k_j=0}^{\infty} (-1)^{k_i+k_j}\Psi_{k_i,k_j,A}(t_i, t_j).$$

Subsequently we claim that

$$\sum_{i=1}^{m} a_i \bar{\chi}_n(t_i) \Rightarrow \mathcal{N}(0, \Sigma_{\mathbb{R}^d}), \qquad n \to \infty,$$

which completes the proof. To show this, take $A_K = (-K, K)^d$ for $K > 0$. It then suffices to verify that

$$\mathcal{N}(0, \Sigma_{A_K}) \Rightarrow \mathcal{N}(0, \Sigma_{\mathbb{R}^d}), \qquad K \to \infty,$$

and for each $t \geq 0$ and $\epsilon > 0$,

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{P}\Big(\big|\bar{\chi}_n(t) - \bar{\chi}_{n,A_K}(t)\big| > \epsilon\Big) = 0.$$

The former condition is obvious from the fact that $\Sigma_{A_K} \to \Sigma_{\mathbb{R}^d}$ as $K \to \infty$. The latter is also a direct consequence of Proposition 3.1, together with Chebyshev's inequality and the fact that $\chi_n(t) - \chi_{n,A_K}(t) = \chi_{n,\mathbb{R}^d\setminus A_K}(t)$. □

4.4. Proof of tightness in Theorem 3.2

Before we begin the proof, we add a few more useful properties of h_t^k. For $0 \leq s < t < \infty$, we define

$$h_{t,s}^k(\mathcal{Y}) = h_t^k(\mathcal{Y}) - h_s^k(\mathcal{Y}), \qquad \mathcal{Y} = (y_0, \ldots, y_k) \in (\mathbb{R}^d)^{k+1}.$$

Lemma 4.2.

(*i*) *For any* $0 \leq s \leq t \leq T < \infty$,

$$\int_{(\mathbb{R}^d)^k} h_{t,s}^k(0, y_1, \ldots, y_k)\mathrm{d}\mathbf{y} \leq C_{d,k,T}(t^d - s^d),$$

where $C_{d,k,T} = k^2(2^d\theta_d)^k T^{d(k-1)}$.

(*ii*) *Let* $j \in \{1, \ldots, (k_1 \wedge k_2) + 1\}$ *and suppose that* $\mathbf{y}_0 \in (\mathbb{R}^d)^{j-1}$, $\mathbf{y}_1 \in (\mathbb{R}^d)^{k_1+1-j}$, *and* $\mathbf{y}_2 \in (\mathbb{R}^d)^{k_2+1-j}$. *Then, for* $0 \leq t_1 \leq s \leq t_2 \leq T < \infty$,

$$\int_{(\mathbb{R}^d)^{k_1+k_2+1-j}} h_{s,t_1}^{k_1}(0, \mathbf{y}_0, \mathbf{y}_1) h_{t_2,s}^{k_2}(0, \mathbf{y}_0, \mathbf{y}_2)\mathrm{d}\mathbf{y}_0\mathrm{d}\mathbf{y}_1\mathrm{d}\mathbf{y}_2$$
$$\leq 36(k_1k_2)^6((2T)^d\theta_d)^{2(k_1+k_2)}\big(t_2^d - t_1^d\big)^2.$$

Proof. We note that for any $0 \leq s < t$ with $y_0 \equiv 0$,

$$h_{t,s}^k(0, y_1, \ldots, y_k) = \mathbf{1}\big\{2s < \max_{0\leq i<j\leq k} \|y_i - y_j\| \leq 2t\big\}$$
$$\leq \prod_{i=1}^{k} \mathbf{1}\big\{y_i \in B(0, 2T)\big\}\Bigg(\sum_{i=1}^{k} \mathbf{1}\big\{2s < \|y_i\| \leq 2t\big\} + \sum_{1\leq i<j\leq k} \mathbf{1}\big\{2s < \|y_i - y_j\| \leq 2t\big\}\Bigg).$$

For each $i = 1, \dots, k$, let $\mathbf{y}^{(i)}$ be the tuple $(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_k) \in (\mathbb{R}^d)^{k-1}$ with the ith coordinate omitted. Then

$$\begin{aligned}\int_{(\mathbb{R}^d)^k} h_{t,s}^k(0, y_1, \dots, y_k)\mathrm{d}\mathbf{y} &\le \sum_{i=1}^k \int_{B(0,2T)^{k-1}} \int_{\mathbb{R}^d} \mathbf{1}\{2s < \|y_i\| \le 2t\}\mathrm{d}y_i \, \mathrm{d}\mathbf{y}^{(i)} \\ &\quad + \sum_{1\le i<j\le k} \int_{B(0,2T)^{k-1}} \int_{\mathbb{R}^d} \mathbf{1}\{2s < \|y_i - y_j\| \le 2t\}\mathrm{d}y_i \, \mathrm{d}\mathbf{y}^{(i)}. \\ &= \left(k + \binom{k}{2}\right) m\big(B(0, 2T)\big)^{k-1}\big[m\big(B(0, 2t)\big) - m\big(B(0, 2s)\big)\big] \\ &\le C_{d,k,T}(t^d - s^d)\end{aligned}$$

as required.

Part (*ii*) is essentially the same as Lemma 7.1 in [18], so the proof is skipped. □

Proof of tightness in Theorem 3.2. To show tightness, it suffices to use Theorem 13.5 from [2], which requires that for every $0 < T < \infty$, there exist a $C > 0$ such that

$$\mathbb{E}\big[|\bar{\chi}_n(t_2) - \bar{\chi}_n(s)|^2 |\bar{\chi}_n(s) - \bar{\chi}_n(t_1)|^2\big] \le C(t_2^d - t_1^d)^2 \tag{4.16}$$

for all $0 \le t_1 \le s \le t_2 \le T$ and $n \in \mathbb{N}$. To demonstrate (4.16), we will give an abridged proof—tightness will be similarly established for analogous processes seen in [18, 21]. Let us begin with some helpful notation, namely,

$$\begin{aligned}h_{n,t,s}^k(\mathcal{Y}) &:= h_{r_n(t),r_n(s)}^k(\mathcal{Y}) = h_{r_n(t)}^k(\mathcal{Y}) - h_{r_n(s)}^k(\mathcal{Y}), \\ \zeta_{n,t,s}^k &:= S_k\big(\mathcal{P}_n, r_n(t)\big) - S_k\big(\mathcal{P}_n, r_n(s)\big) = \sum_{\mathcal{Y}\subset\mathcal{P}_n} h_{n,t,s}^k(\mathcal{Y}).\end{aligned}$$

By the same argument as in the proof of Proposition 3.1, one can apply Fubini's theorem to obtain

$$\begin{aligned}&\mathbb{E}\big[|\bar{\chi}_n(t_2) - \bar{\chi}_n(s)|^2 |\bar{\chi}_n(s) - \bar{\chi}_n(t_1)|^2\big] \\ &\quad = \frac{1}{n^2} \sum_{(k_1,k_2,k_3,k_4)\in\mathbb{N}_0^4} (-1)^{k_1+k_2+k_3+k_4} \mathbb{E}\Big[\big(\zeta_{n,t_2,s}^{k_1} - \mathbb{E}[\zeta_{n,t_2,s}^{k_1}]\big)\big(\zeta_{n,t_2,s}^{k_2} - \mathbb{E}[\zeta_{n,t_2,s}^{k_2}]\big) \\ &\qquad\qquad \times \big(\zeta_{n,s,t_1}^{k_3} - \mathbb{E}[\zeta_{n,s,t_1}^{k_3}]\big)\big(\zeta_{n,s,t_1}^{k_4} - \mathbb{E}[\zeta_{n,s,t_1}^{k_4}]\big)\Big].\end{aligned} \tag{4.17}$$

Our objective now is to find a suitable bound for

$$\mathbb{E}\Big[\big(\zeta_{n,t_2,s}^{k_1} - \mathbb{E}[\zeta_{n,t_2,s}^{k_1}]\big)\big(\zeta_{n,t_2,s}^{k_2} - \mathbb{E}[\zeta_{n,t_2,s}^{k_2}]\big)\big(\zeta_{n,s,t_1}^{k_3} - \mathbb{E}[\zeta_{n,s,t_1}^{k_3}]\big)\big(\zeta_{n,s,t_1}^{k_4} - \mathbb{E}[\zeta_{n,s,t_1}^{k_4}]\big)\Big]. \tag{4.18}$$

To this end, let us refine the notation once more by setting $\xi_1 := \zeta_{n,t_2,s}^{k_1}$, $\xi_2 := \zeta_{n,t_2,s}^{k_2}$, $\xi_3 := \zeta_{n,s,t_1}^{k_3}$, and $\xi_4 := \zeta_{n,s,t_1}^{k_4}$. Furthermore, let $h_1 := h_{n,t_2,s}^{k_1}$, $h_2 := h_{n,t_2,s}^{k_2}$, $h_3 := h_{n,s,t_1}^{k_3}$, and $h_4 := h_{n,s,t_1}^{k_4}$. Define $[n] := \{1, 2, \dots, n\}$, and for any $\sigma \subset [4]$ let $\xi_\sigma = \prod_{i\in\sigma} \xi_i$, where we set $\xi_\emptyset = 1$ by convention. Then we can express (4.18) quite simply as

$$\sum_{\sigma\subset[4]} (-1)^{|\sigma|}\, \mathbb{E}[\xi_\sigma] \prod_{i\in[4]\setminus\sigma} \mathbb{E}[\xi_i]. \tag{4.19}$$

For $\sigma \subset [4]$ with $\sigma \neq \emptyset$, and finite subsets $\mathcal{Y}_j \subset \mathbb{R}^d$, $j \in \sigma$, we define $\mathcal{Y}_\sigma := \bigcup_{j\in\sigma} \mathcal{Y}_j$. Given a subset $\tau \subset \sigma \subset [4]$, we also define

$$\mathcal{I}_{\tau,\sigma}(\mathcal{Y}_\sigma) := \prod_{j\in\tau} \mathbf{1}\{\text{there exists } p \in \tau \setminus \{j\} \text{ such that } \mathcal{Y}_j \cap \mathcal{Y}_p \neq \emptyset\} \times \prod_{j\in\sigma\setminus\tau} \mathbf{1}\{\mathcal{Y}_j \cap \mathcal{Y}_q = \emptyset \text{ for all } q \in \sigma \setminus \{j\}\}.$$

Note that $\mathcal{I}_{\tau,\sigma}(\mathcal{Y}_\sigma) = 0$ whenever $|\tau| = 1$, and

$$\sum_{\tau\subset\sigma} \mathcal{I}_{\tau,\sigma}(\mathcal{Y}_\sigma) = 1. \tag{4.20}$$

Furthermore, if $\tau = \sigma$, we write $\mathcal{I}_\sigma(\cdot) := \mathcal{I}_{\sigma,\sigma}(\cdot)$. It follows from (4.20) and the Palm theory in the appendix that, for each non-empty $\sigma \subset [4]$,

$$\begin{aligned}
\mathbb{E}[\xi_\sigma] &= \mathbb{E}\Big[\sum_{\mathcal{Y}_j\subset\mathcal{P}_n,\, j\in\sigma} \prod_{i\in\sigma} h_i(\mathcal{Y}_i)\Big] \\
&= \sum_{\tau\subset\sigma} \mathbb{E}\Big[\sum_{\mathcal{Y}_j\subset\mathcal{P}_n,\, j\in\sigma} \mathcal{I}_{\tau,\sigma}(\mathcal{Y}_\sigma) \prod_{i\in\sigma} h_i(\mathcal{Y}_i)\Big] \\
&= \sum_{\tau\subset\sigma} \mathbb{E}\Big[\sum_{\mathcal{Y}_j\subset\mathcal{P}_n,\, j\in\tau} \mathcal{I}_{\tau}(\mathcal{Y}_\tau) \prod_{i\in\tau} h_i(\mathcal{Y}_i)\Big] \prod_{i\in\sigma\setminus\tau} \mathbb{E}[\xi_i].
\end{aligned}$$

Hence, (4.19) is equal to

$$\begin{aligned}
&\sum_{\sigma\subset[4]} \sum_{\tau\subset\sigma} (-1)^{|\sigma|}\, \mathbb{E}\Big[\sum_{\mathcal{Y}_j\subset\mathcal{P}_n,\, j\in\tau} \mathcal{I}_{\tau}(\mathcal{Y}_\tau) \prod_{i\in\tau} h_i(\mathcal{Y}_i)\Big] \prod_{i\in\sigma\setminus\tau} \mathbb{E}[\xi_i] \prod_{i\in[4]\setminus\sigma} \mathbb{E}[\xi_i] \\
&= \sum_{\tau\subset[4]} \mathbb{E}\Big[\sum_{\mathcal{Y}_j\subset\mathcal{P}_n,\, j\in\tau} \mathcal{I}_{\tau}(\mathcal{Y}_\tau) \prod_{i\in\tau} h_i(\mathcal{Y}_i)\Big] \prod_{i\in[4]\setminus\tau} \mathbb{E}[\xi_i] \sum_{\tau\subset\sigma\subset[4]} (-1)^{|\sigma|} \\
&= \mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n} \sum_{\mathcal{Y}_2\subset\mathcal{P}_n} \sum_{\mathcal{Y}_3\subset\mathcal{P}_n} \sum_{\mathcal{Y}_4\subset\mathcal{P}_n} \mathcal{I}_{[4]}(\mathcal{Y}_{[4]}) \prod_{i=1}^4 h_i(\mathcal{Y}_i)\Big],
\end{aligned}$$

where the last line follows from the fact that

$$\sum_{\tau\subset\sigma\subset[4]} (-1)^{|\sigma|} = \binom{4-|\tau|}{0}(-1)^{|\tau|} + \cdots + \binom{4-|\tau|}{4-|\tau|}(-1)^4 = 0$$

unless $\tau = [4]$. Substituting this back into (4.17) and taking the absolute value of $(-1)^{k_1+k_2+k_3+k_4}$, we get

$$\begin{aligned}
&\mathbb{E}\big[|\bar\chi_n(t_2) - \bar\chi_n(s)|^2 |\bar\chi_n(s) - \bar\chi_n(t_1)|^2\big] \\
&\leq \sum_{(k_1,k_2,k_3,k_4)\in\mathbb{N}_0^4} \frac{1}{n^2} \mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n} \sum_{\mathcal{Y}_2\subset\mathcal{P}_n} \sum_{\mathcal{Y}_3\subset\mathcal{P}_n} \sum_{\mathcal{Y}_4\subset\mathcal{P}_n} \mathcal{I}_{[4]}(\mathcal{Y}_{[4]}) \prod_{i=1}^4 h_i(\mathcal{Y}_i)\Big].
\end{aligned}$$

Now, it suffices to show that the right-hand side above is less than $C\big(t_2^d - t_1^d\big)^2$ for some $C > 0$. We can break the above summand into four distinct cases:

(I) $b_{12} = |\mathcal{Y}_1 \cap \mathcal{Y}_2| > 0$, $b_{34} = |\mathcal{Y}_3 \cap \mathcal{Y}_4| > 0$, with all other pairwise intersections empty.

(II) $b_{13} = |\mathcal{Y}_1 \cap \mathcal{Y}_3| > 0$, $b_{24} = |\mathcal{Y}_2 \cap \mathcal{Y}_4| > 0$, with all other pairwise intersections empty.

(III) $b_{14} = |\mathcal{Y}_1 \cap \mathcal{Y}_4| > 0$, $b_{23} = |\mathcal{Y}_2 \cap \mathcal{Y}_3| > 0$, with all other pairwise intersections empty.

(IV) For each i, there exists a $j \neq i$ such that $\mathcal{Y}_i \cap \mathcal{Y}_j \neq \emptyset$, but **(I)**–**(III)** do not hold.

We prove appropriate upper bounds for Cases **(I)** and **(IV)**, and the other two cases follow from the proof for **(I)**. The Palm theory in the appendix implies that

$$\begin{aligned}
&\frac{1}{n^2}\mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n}\sum_{\mathcal{Y}_3\subset\mathcal{P}_n}\sum_{\mathcal{Y}_4\subset\mathcal{P}_n}\prod_{i=1}^{4} h_i(\mathcal{Y}_i) \\
&\qquad \times \mathbf{1}\big\{|\mathcal{Y}_1\cap\mathcal{Y}_2| = b_{12},\, |\mathcal{Y}_3\cap\mathcal{Y}_4| = b_{34},\, |\mathcal{Y}_i\cap\mathcal{Y}_j| = 0 \text{ for other } (i,j)\big\}\Big] \qquad (4.21)\\
&= \frac{1}{n^2}\mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n} h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)\mathbf{1}\big\{|\mathcal{Y}_1\cap\mathcal{Y}_2| = b_{12}\big\}\Big] \\
&\qquad \times \mathbb{E}\Big[\sum_{\mathcal{Y}_3\subset\mathcal{P}_n}\sum_{\mathcal{Y}_4\subset\mathcal{P}_n} h_3(\mathcal{Y}_3)h_4(\mathcal{Y}_4)\mathbf{1}\big\{|\mathcal{Y}_3\cap\mathcal{Y}_4| = b_{34}\big\}\Big] \\
&= \frac{n^{k_1+k_2+1-b_{12}}}{b_{12}!(k_1+1-b_{12})!(k_2+1-b_{12})!} \\
&\qquad \times \mathbb{E}\big[h_1(X_1,\dots,X_{k_1+1})h_2(X_1,\dots,X_{b_{12}},X_{k_1+2},\dots,X_{k_1+k_2+2-b_{12}})\big] \\
&\times \frac{n^{k_3+k_4+1-b_{34}}}{b_{34}!(k_3+1-b_{34})!(k_4+1-b_{34})!} \\
&\qquad \times \mathbb{E}\big[h_3(X_1,\dots,X_{k_3+1})h_4(X_1,\dots,X_{b_{34}},X_{k_3+2},\dots,X_{k_3+k_4+2-b_{34}})\big].
\end{aligned}$$

In the remainder of the proof, assume for ease of description that $(2T)^d\theta_d > 1$, $\|f\|_\infty > 1$, and $T > 1$. Moreover, assume without loss of generality that $k_1 \geq k_2$ and $k_3 \geq k_4$. Using trivial bounds and the customary changes of variable (i.e., $x_1 = x$ and $x_i = x + s_n y_{i-1}$ for $i = 2,\dots,k_1+k_2+2-b_{12}$), applying Lemma 4.2($i$), and recalling that $a_T = (2T)^d\theta_d\|f\|_\infty$, we see that

$$\begin{aligned}
&n^{k_1+k_2+1-b_{12}}\mathbb{E}[h_1(X_1,\dots,X_{k_1+1})h_2(X_1,\dots,X_{b_{12}},X_{k_1+2},\dots,X_{k_1+k_2+2-b_{12}})] \\
&\leq (\|f\|_\infty)^{k_1+k_2+1-b_{12}}\int_{(\mathbb{R}^d)^{k_2+1-b_{12}}}\int_{(\mathbb{R}^d)^{k_1+1-b_{12}}}\int_{(\mathbb{R}^d)^{b_{12}-1}} h_{t_2,s}^{k_1}(0,\mathbf{y}_0,\mathbf{y}_1) \\
&\qquad\qquad \times h_{t_2,s}^{k_2}(0,\mathbf{y}_0,\mathbf{y}_2)\mathrm{d}\mathbf{y}_0\mathrm{d}\mathbf{y}_1\mathrm{d}\mathbf{y}_2 \\
&\leq (\|f\|_\infty)^{k_1+k_2}\big((2T)^d\theta_d\big)^{k_2+1-b_{12}}\int_{(\mathbb{R}^d)^{k_1+1-b_{12}}}\int_{(\mathbb{R}^d)^{b_{12}-1}} h_{t_2,s}^{k_1}(0,\mathbf{y}_0,\mathbf{y}_1)\mathrm{d}\mathbf{y}_0\mathrm{d}\mathbf{y}_1 \\
&\leq (\|f\|_\infty)^{k_1+k_2}\big((2T)^d\theta_d\big)^{k_2+1-b_{12}}C_{d,k_1,T}(t_2^d - s^d) \\
&\leq k_1^2(a_T)^{k_1+k_2}(t_2^d - s^d).
\end{aligned}$$

Hence (4.21) is bounded by

$$\frac{(a_T)^{k_1+k_2}k_1^2}{b_{12}!(k_1+1-b_{12})!(k_2+1-b_{12})!}(t_2^d-s^d)\frac{(a_T)^{k_3+k_4}k_3^2}{b_{34}!(k_3+1-b_{34})!(k_4+1-b_{34})!}(s^d-t_1^d)$$
$$\le \frac{(a_T)^{k_1+k_2+k_3+k_4}k_1^2k_3^2}{b_{12}!(k_1+1-b_{12})!(k_2+1-b_{12})!b_{34}!(k_3+1-b_{34})!(k_4+1-b_{34})!}(t_2^d-t_1^d)^2.$$

Finally we see that

$$\sum_{\substack{k_1\ge k_2,\, k_3\ge k_4,\\ 1\le b_{12}\le k_2+1,\\ 1\le b_{34}\le k_4+1}} \frac{(a_T)^{k_1+k_2+k_3+k_4}k_1^2k_3^2}{b_{12}!(k_1+1-b_{12})!(k_2+1-b_{12})!b_{34}!(k_3+1-b_{34})!(k_4+1-b_{34})!}<\infty,$$

since

$$\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{k_1}\sum_{\ell=1}^{k_2+1}\frac{(a_T)^{k_1+k_2}k_1^2}{\ell!(k_1+1-\ell)!(k_2+1-\ell)!}=\sum_{\ell=1}^{\infty}\sum_{k_1=\ell-1}^{\infty}\frac{(a_T)^{k_1}k_1^2}{\ell!(k_1+1-\ell)!}\sum_{k_2=\ell-1}^{k_1}\frac{(a_T)^{k_2}}{(k_2+1-\ell)!}$$
$$\le e^{a_T}\sum_{\ell=1}^{\infty}\frac{(a_T)^{\ell-1}}{\ell!}\sum_{k_1=\ell-1}^{\infty}\frac{(a_T)^{k_1}k_1^2}{(k_1+1-\ell)!}<\infty. \tag{4.22}$$

Now, for Cases **(I)**–**(III)**, we have an upper bound of the form $C(t_2^d-t_1^d)^2$, as desired.

Thus we need only demonstrate the same for Case (**IV**). In addition to the notation b_{ij}, $1\le i<j\le 4$, introduced above, define for $\mathcal{Y}_i\in(\mathbb{R}^d)^{k_i+1}$, $k_i\in\mathbb{N}_0$, $i=1,\dots,4$,

$$b_{ijk}:=|\mathcal{Y}_i\cap\mathcal{Y}_j\cap\mathcal{Y}_k|,\qquad 1\le i<j<k\le 4,$$
$$b_{1234}:=|\mathcal{Y}_1\cap\mathcal{Y}_2\cap\mathcal{Y}_3\cap\mathcal{Y}_4|,$$

and

$$b:=b_{12}+b_{13}+b_{14}+b_{23}+b_{24}+b_{34}-b_{123}-b_{124}-b_{134}-b_{234}+b_{1234}, \tag{4.23}$$

so that $|\mathcal{Y}_1\cup\mathcal{Y}_2\cup\mathcal{Y}_3\cup\mathcal{Y}_4|=k_1+k_2+k_3+k_4+4-b$ with $b\ge 3$. Let $\mathcal{B}$ be the collection of $\mathbf{b}=(b_{12},\dots,b_{1234})\in\mathbb{N}_0^{11}$ satisfying the conditions in Case (**IV**). For a non-empty $\sigma\subset[4]$ and $\mathcal{Y}_i\in(\mathbb{R}^d)^{k_i+1}$, $i=1,\dots,4$, let

$$j_\sigma:=\Big|\bigcap_{i\in\sigma}\Big(\mathcal{Y}_i\setminus\bigcup_{j\in[4]\setminus\sigma}\mathcal{Y}_j\Big)\Big|. \tag{4.24}$$

In particular, the j_σ are functions of $\mathbf{b}$ such that $\sum_{\sigma\subset[4],\,\sigma\ne\emptyset}j_\sigma=|\mathcal{Y}_1\cup\mathcal{Y}_2\cup\mathcal{Y}_3\cup\mathcal{Y}_4|$. The Palm theory in the appendix yields

$$\frac{1}{n^2}\,\mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n}\sum_{\mathcal{Y}_3\subset\mathcal{P}_n}\sum_{\mathcal{Y}_4\subset\mathcal{P}_n}\prod_{i=1}^{4}h_i(\mathcal{Y}_i)\,\mathbf{1}\{\text{case (IV) holds}\}\Big]$$
$$=\sum_{\mathbf{b}\in\mathcal{B}}\frac{1}{n^2}\,\mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n}\sum_{\mathcal{Y}_3\subset\mathcal{P}_n}\sum_{\mathcal{Y}_4\subset\mathcal{P}_n}\prod_{i=1}^{4}h_i(\mathcal{Y}_i)\,\mathbf{1}\{|\mathcal{Y}_1\cap\mathcal{Y}_2|=b_{12},\,|\mathcal{Y}_1\cap\mathcal{Y}_3|=b_{13},$$

$$\ldots, |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big]$$

$$= \sum_{\mathbf{b}\in\mathcal{B}} \frac{n^{k_1+k_2+k_3+k_4+2-b}}{\prod_{\sigma\subset[4],\,\sigma\neq\emptyset} j_\sigma!}\, \mathbb{E}\Big[\prod_{i=1}^{4} h_i(\mathcal{Y}_i)\, \mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = b_{12},\, |\mathcal{Y}_1 \cap \mathcal{Y}_3| = b_{13},$$

$$\ldots, |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big].$$

Under the conditions in Case (**IV**), at least one of the b_{ij} is nonzero, so we may assume without loss of generality that $b_{13} > 0$. Then we have

$$n^{k_1+k_2+k_3+k_4+2-b}\mathbb{E}\Big[\prod_{i=1}^{4} h_i(\mathcal{Y}_i)\, \mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = b_{12}, \ldots, |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big]$$

$$= n^{k_1+k_2+k_3+k_4+2-b} \int_{(\mathbb{R}^d)^{k_1+k_2+k_3+k_4+4-b}} h_1(\mathbf{x}_0, \mathbf{x}_1) h_3(\mathbf{x}_0, \mathbf{x}_3) h_2(\mathbf{x}_2) h_4(\mathbf{x}_4)$$

$$\times \prod_{x\in\bigcup_{i=0}^{4}\mathbf{x}_i} f(x)\mathrm{d}\,(\mathbf{x}_0 \cup \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_4),$$

where $\mathbf{x}_0$ is a collection of elements in $\mathbb{R}^d$ with $|\mathbf{x}_0| = b_{13} > 0$. In other words, $\mathbf{x}_0 \in (\mathbb{R}^d)^{b_{13}}$, so that $\mathbf{x}_1 \in (\mathbb{R}^d)^{k_1+1-b_{13}}$ and $\mathbf{x}_3 \in (\mathbb{R}^d)^{k_3+1-b_{13}}$ with $\mathbf{x}_1 \cap \mathbf{x}_3 = \emptyset$. Moreover, $\mathbf{x}_2 \in (\mathbb{R}^d)^{k_2+1}$ and $\mathbf{x}_4 \in (\mathbb{R}^d)^{k_4+1}$, such that if $\mathbf{x}_2 \cap \mathbf{x}_4 = \emptyset$, then $\mathbf{x}_i \cap (\mathbf{x}_0 \cup \mathbf{x}_1 \cup \mathbf{x}_3) \neq \emptyset$ for $i = 2, 4$, and if $\mathbf{x}_2 \cap \mathbf{x}_4 \neq \emptyset$, then $(\mathbf{x}_2 \cup \mathbf{x}_4) \cap (\mathbf{x}_0 \cup \mathbf{x}_1 \cup \mathbf{x}_3) \neq \emptyset$.

Now, let us perform the change of variables $\mathbf{x}_i = x\mathbf{1} + s_n\mathbf{y}_i$ for $i = 0, \ldots, 4$, where $\mathbf{1}$ is a vector with all entries 1, and the first element of $\mathbf{y}_0$ is taken to be 0. In addition to this, we apply the translation and scale invariance of the h_i to get

$$n^{k_1+k_2+k_3+k_4+2-b}\mathbb{E}\Big[\prod_{i=1}^{4} h_i(\mathcal{Y}_i)\, \mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = b_{12}, \ldots, |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big]$$

$$= n^{k_1+k_2+k_3+k_4+2-b} s_n^{d(k_1+k_2+k_3+k_4+3-b)} \int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^{k_1+k_2+k_3+k_4+3-b}} h_{t_2,s}^{k_1}(\mathbf{y}_0, \mathbf{y}_1) h_{s,t_1}^{k_3}(\mathbf{y}_0, \mathbf{y}_3)$$

$$\times h_{t_2,s}^{k_2}(\mathbf{y}_2) h_{s,t_1}^{k_4}(\mathbf{y}_4) \prod_{y\in\bigcup_{i=0}^{4}\mathbf{y}_i} f(x + s_n y)\mathrm{d}\big((\mathbf{y}_0 \cup \cdots \cup \mathbf{y}_4) \setminus \{0\}\big)\mathrm{d}x.$$

Using $ns_n^d = 1$, together with the trivial bounds $h_{t_2,s}^{k_2}(\mathbf{y}_2) \le h_T^{k_2}(\mathbf{y}_2)$, $h_{s,t_1}^{k_4}(\mathbf{y}_4) \le h_T^{k_4}(\mathbf{y}_4)$, and $f(x + s_n y) \le \|f\|_\infty$, one can bound the last expression by

$$\|f\|_\infty^{k_1+k_2+k_3+k_4+3-b} \int_{(\mathbb{R}^d)^{k_1+k_2+k_3+k_4+3-b}} h_{t_2,s}^{k_1}(\mathbf{y}_0, \mathbf{y}_1) h_{s,t_1}^{k_3}(\mathbf{y}_0, \mathbf{y}_3)$$

$$\times h_T^{k_2}(\mathbf{y}_2) h_T^{k_4}(\mathbf{y}_4)\mathrm{d}\big((\mathbf{y}_0 \cup \cdots \cup \mathbf{y}_4) \setminus \{0\}\big)$$

$$= \|f\|_\infty^{k_1+k_2+k_3+k_4+3-b} \int_{(\mathbb{R}^d)^{k_1+k_3+1-b_{13}}} h_{t_2,s}^{k_1}(\mathbf{y}_0, \mathbf{y}_1) h_{s,t_1}^{k_3}(\mathbf{y}_0, \mathbf{y}_3)$$

$$\times \Big\{ \int_{(\mathbb{R}^d)^{k_2+k_4+2-b+b_{13}}} h_T^{k_2}(\mathbf{y}_2) h_T^{k_4}(\mathbf{y}_4)\mathrm{d}\big((\mathbf{y}_2 \cup \mathbf{y}_4) \setminus (\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_3)\big) \Big\}\mathrm{d}\big(\mathbf{y}_0 \setminus \{0\}\big)\mathrm{d}\mathbf{y}_1\mathrm{d}\mathbf{y}_3. \tag{4.25}$$

Suppose $h_T^{k_2}(\mathbf{y}_2)h_T^{k_4}(\mathbf{y}_4) = 1$, so that

$$\mathbf{y}_2 \cap \mathbf{y}_4 \neq \emptyset, \qquad \mathbf{y}_2 \cap (\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_3) = \emptyset, \qquad \mathbf{y}_4 \cap (\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_3) \neq \emptyset. \tag{4.26}$$

Then there exists $y' \in \mathbf{y}_4 \cap (\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_3)$ such that all points in $\mathbf{y}_2$ are at distance at most $4T$ from y'. Since y' itself lies within distance $2T$ from the origin (recall that the first element of $\mathbf{y}_0$ is 0), we conclude that all points in $\mathbf{y}_2 \cap \mathbf{y}_4$ are at distance at most $6T$ from the origin. As $b_{13} \le k_1 + k_3 + 1$ and $b \ge 3$, we have

$$\begin{aligned} &\int_{(\mathbb{R}^d)^{k_2+k_4+2-b+b_{13}}} h_T^{k_2}(\mathbf{y}_2)h_T^{k_4}(\mathbf{y}_4)\mathrm{d}\big((\mathbf{y}_2 \cup \mathbf{y}_4) \setminus (\mathbf{y}_0 \cup \mathbf{y}_1 \cup \mathbf{y}_3)\big) \\ &\quad \le m\big(B(0, 6T)\big)^{k_2+k_4+2-b+b_{13}} = \big((6T)^d\theta_d\big)^{k_2+k_4+2-b+b_{13}} \le \big((6T)^d\theta_d\big)^{k_1+k_2+k_3+k_4}. \end{aligned} \tag{4.27}$$

If $\mathbf{y}_2$ and $\mathbf{y}_4$ do not satisfy (4.26), it is still easy to check (4.27).

Applying (4.27), along with Lemma 4.2(*ii*), one can bound (4.25) by

$$\begin{aligned} &\|f\|_\infty^{k_1+k_2+k_3+k_4+3-b}\big((6T)^d\theta_d\big)^{k_1+k_2+k_3+k_4} \times 36(k_1k_3)^6\big((2T)^d\theta_d\big)^{2(k_1+k_3)}\big(t_2^d - t_1^d\big)^2 \\ &\quad \le 36(k_1k_2k_3k_4)^6\big((6T)^d\theta_d\|f\|_\infty\big)^{3(k_1+k_2+k_3+k_4)}\big(t_2^d - t_1^d\big)^2. \end{aligned}$$

Thus, we conclude that

$$\begin{aligned} &\frac{1}{n^2}\,\mathbb{E}\Big[\sum_{\mathcal{Y}_1\subset\mathcal{P}_n}\sum_{\mathcal{Y}_2\subset\mathcal{P}_n}\sum_{\mathcal{Y}_3\subset\mathcal{P}_n}\sum_{\mathcal{Y}_4\subset\mathcal{P}_n}\prod_{i=1}^4 h_i(\mathcal{Y}_i)\,\mathbf{1}\{\text{case } \textbf{(IV)} \text{ holds}\}\Big] \\ &\quad \le 36(k_1k_2k_3k_4)^6\big((6T)^d\theta_d\|f\|_\infty\big)^{3(k_1+k_2+k_3+k_4)}\sum_{\mathbf{b}\in\mathcal{B}}\frac{1}{\prod_{\sigma\subset[4],\,\sigma\neq\emptyset} j_\sigma!}\,\big(t_2^d - t_1^d\big)^2. \end{aligned}$$

To complete the proof, we need to show that

$$\sum_{k_1\le k_2\le k_3\le k_4}(k_1k_2k_3k_4)^6\big((6T)^d\theta_d\|f\|_\infty\big)^{3(k_1+k_2+k_3+k_4)}\sum_{\mathbf{b}\in\mathcal{B}}\frac{1}{\prod_{\sigma\subset[4],\,\sigma\neq\emptyset} j_\sigma!} < \infty.$$

As seen in the calculation at (4.22), the term $(k_1k_2k_3k_4)^6\big((6T)^d\theta_d\|f\|_\infty\big)^{3(k_1+k_2+k_3+k_4)}$ is negligible, while proving

$$\sum_{k_1\le k_2\le k_3\le k_4}\sum_{\mathbf{b}\in\mathcal{B}}\frac{1}{\prod_{\sigma\subset[4],\,\sigma\neq\emptyset} j_\sigma!} < \infty$$

is straightforward.

4.5 Proof of Hölder continuity of $\mathcal{H}$

Proof of Hölder continuity in Theorem 3.2. Since $\mathcal{H}(t) - \mathcal{H}(s)$ has a normal distribution for $0 \le s < t < \infty$, we have for every $m \in \mathbb{N}$ that

$$\mathbb{E}\Big[\big(\mathcal{H}(t) - \mathcal{H}(s)\big)^{2m}\Big] = \prod_{i=1}^m (2i-1)\Big(\mathbb{E}\Big[\big(\mathcal{H}(t) - \mathcal{H}(s)\big)^2\Big]\Big)^m.$$

Proposition 3.1 ensures that $\left(\sum_{k=0}^{M}(-1)^k\mathcal{H}_k(t),\ M\in\mathbb{N}_0\right)$ constitutes a Cauchy sequence in $L^2(\Omega)$. Therefore we have

$$\mathbb{E}\Big[\big(\mathcal{H}(t)-\mathcal{H}(s)\big)^2\Big]=\lim_{M\to\infty}\mathbb{E}\Bigg[\bigg(\sum_{k=0}^{M}(-1)^k\big(\mathcal{H}_k(t)-\mathcal{H}_k(s)\big)\bigg)^2\Bigg]$$
$$\le\Bigg[\sum_{k=0}^{\infty}\Big\{\mathbb{E}\Big[\big(\mathcal{H}_k(t)-\mathcal{H}_k(s)\big)^2\Big]\Big\}^{1/2}\Bigg]^2,$$

where the second line is due to the Cauchy–Schwarz inequality. We see at once that

$$\mathbb{E}\Big[\big(\mathcal{H}_k(t)-\mathcal{H}_k(s)\big)^2\Big]=\Psi_{k,k}(t,t)-2\Psi_{k,k}(t,s)+\Psi_{k,k}(s,s)$$
$$\le\Psi_{k,k}(t,t)-\Psi_{k,k}(t,s)=\sum_{j=1}^{k+1}\big(\psi_{j,k,k}(t,t)-\psi_{j,k,k}(t,s)\big)\tag{4.28}$$

by monotonicity due to (2.2) and symmetry of $\Psi_{k,k}(\cdot,\cdot)$ in its arguments. Now, we note that

$$\psi_{j,k,k}(t,t)-\psi_{j,k,k}(t,s)=\frac{\int_{\mathbb{R}^d}f(x)^{2k+2-j}\mathrm{d}x}{j!((k+1-j)!)^2}$$
$$\times\int_{(\mathbb{R}^d)^{k+1-j}}\int_{(\mathbb{R}^d)^{k+1-j}}\int_{(\mathbb{R}^d)^{j-1}}h_t^k(0,\mathbf{y}_0,\mathbf{y}_1)h_{t,s}^k(0,\mathbf{y}_0,\mathbf{y}_2)\mathrm{d}\mathbf{y}_0\mathrm{d}\mathbf{y}_1\mathrm{d}\mathbf{y}_2.$$

Applying a bound $h_t^k(0,\mathbf{y}_0,\mathbf{y}_1)\le\prod_{y\in\mathbf{y}_1}\mathbf{1}\{\|y\|\le 2T\}$ and integrating out $\mathbf{y}_1$, as well as using Lemma 4.2(i), we get

$$\psi_{j,k,k}(t,t)-\psi_{j,k,k}(t,s)\le\frac{k^2}{T^dj!\big((k+1-j)!\big)^2}(a_T)^{2k+1-j}\big(t^d-s^d\big)$$
$$\le\frac{dk^2}{Tj!\big((k+1-j)!\big)^2}(a_T)^{2k+1-j}(t-s),$$

where a_T is given in (4.1). Substituting this back into (4.28), we obtain

$$\mathbb{E}\Big[\big(\mathcal{H}_k(t)-\mathcal{H}_k(s)\big)^2\Big]\le\frac{dk^2}{T}\sum_{j=1}^{k+1}\frac{(a_T)^{2k+1-j}}{j!\big((k+1-j)!\big)^2}(t-s)$$
$$\le\frac{dk^2}{T(k+1)!a_T}\big(a_T(1+a_T)\big)^{k+1}(t-s).$$

Therefore, we conclude that

$$\mathbb{E}\Big[\big(\mathcal{H}(t)-\mathcal{H}(s)\big)^{2m}\Big]\le\prod_{i=1}^{m}(2i-1)\Big(\frac{d}{Ta_T}\Big)^m\bigg(\sum_{k=0}^{\infty}\frac{k\big(a_T(1+a_T)\big)^{(k+1)/2}}{\sqrt{(k+1)!}}\bigg)^{2m}(t-s)^m.$$

One can easily check that the infinite sum on the right-hand side converges via the ratio test. As a result, we can apply the Kolmogorov continuity theorem [17]. This implies that there exists a continuous version of $(\mathcal{H}_k(t),\ 0\le t\le T)$ with Hölder continuous sample paths on $[0,T]$ with any exponent $\gamma\in[0,(m-1)/2m)$. As m is arbitrary, we can conclude by letting $m\to\infty$.

5. Appendix

We briefly cite the necessary Palm theory for Poisson processes, for easy reference for the proofs in Section 4.

Lemma 5.1. (Lemma 8.1 in [18] and Theorems 1.6–1.7 in [22].) *Suppose $\mathcal{P}_n$ is a Poisson point process on $\mathbb{R}^d$ with intensity nf. Further, for every $k_i \in \mathbb{N}_0$, $i = 1, \dots, 4$, let $h_i(\mathcal{Y})$ be a real-valued measurable bounded function defined for $\mathcal{Y} \in (\mathbb{R}^d)^{k_i+1}$. By a slight abuse of notation we let $\mathcal{Y}_i$ be collections of $k_i + 1$ i.i.d. points with density f on the right-hand side of each equation below. We have the following results:*

(*i*)
$$\mathbb{E}\Big[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n} h_1(\mathcal{Y}_1)\Big] = \frac{n^{k_1+1}}{(k_1+1)!}\,\mathbb{E}[h_1(\mathcal{Y}_1)].$$

(*ii*) For every $\ell \in \{0, \dots, (k_1 \wedge k_2) + 1\}$,
$$\mathbb{E}\Big[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_2 \subset \mathcal{P}_n} h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = \ell\big\}\Big]$$
$$= \frac{n^{k_1+k_2+2-\ell}}{\ell!(k_1+1-\ell)!(k_2+1-\ell)!}\,\mathbb{E}\big[h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = \ell\big\}\big].$$

(*iii*) For every $\mathbf{b} = (b_{12}, b_{13}, \dots, b_{1234}) \in \mathbb{N}_0^{11}$, we have
$$\mathbb{E}\Big[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_2 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_3 \subset \mathcal{P}_n} h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)h_3(\mathcal{Y}_3)$$
$$\times\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = b_{12},\ |\mathcal{Y}_1 \cap \mathcal{Y}_3| = b_{13},\ |\mathcal{Y}_2 \cap \mathcal{Y}_3| = b_{23},\ |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3| = b_{123}\big\}\Big]$$
$$= \frac{n^{k_1+k_2+k_3+3-b_{12}-b_{13}-b_{23}+b_{123}}}{\prod_{\sigma \subset [3],\,\sigma \neq \emptyset} j_\sigma!}\,\mathbb{E}\big[h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)h_3(\mathcal{Y}_3)$$
$$\times\,\mathbf{1}\big\{|\mathcal{Y}_1 \cap \mathcal{Y}_2| = b_{12},\ |\mathcal{Y}_1 \cap \mathcal{Y}_3| = b_{13},\ |\mathcal{Y}_2 \cap \mathcal{Y}_3| = b_{23},\ |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3| = b_{123}\big\}\Big],$$
where
$$j_\sigma = \Big|\bigcap_{i \in \sigma}\Big(\mathcal{Y}_i \setminus \bigcup_{j \in [3]\setminus\sigma} \mathcal{Y}_j\Big)\Big|.$$

(*iv*) Furthermore, we have
$$\mathbb{E}\Big[\sum_{\mathcal{Y}_1 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_2 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_3 \subset \mathcal{P}_n}\sum_{\mathcal{Y}_4 \subset \mathcal{P}_n} h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)h_3(\mathcal{Y}_3)h_4(\mathcal{Y}_4)$$
$$\times\,\mathbf{1}\big\{|\mathcal{Y}_i \cap \mathcal{Y}_j| = b_{ij},\ 1 \le i < j \le 4,\ |\mathcal{Y}_i \cap \mathcal{Y}_j \cap \mathcal{Y}_k| = b_{ijk},\ 1 \le i < j < k \le 4,$$
$$|\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big]$$
$$= \frac{n^{k_1+k_2+k_3+k_4+4-b}}{\prod_{\sigma \subset [4],\,\sigma \neq \emptyset} j_\sigma!}\,\mathbb{E}\Big[h_1(\mathcal{Y}_1)h_2(\mathcal{Y}_2)h_3(\mathcal{Y}_3)h_4(\mathcal{Y}_4)\,\mathbf{1}\big\{|\mathcal{Y}_i \cap \mathcal{Y}_j| = b_{ij},\ 1 \le i < j \le 4,$$
$$|\mathcal{Y}_i \cap \mathcal{Y}_j \cap \mathcal{Y}_k| = b_{ijk},\ 1 \le i < j < k \le 4,\ |\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3 \cap \mathcal{Y}_4| = b_{1234}\big\}\Big],$$
where b and j_σ are defined in (4.23) and (4.24), respectively.

Acknowledgements

The authors cordially thank the two anonymous referees and the anonymous editor for their insightful and in-depth comments. Their suggestions have allowed us to improve the clarity and flow of the paper. T. O.'s research is partially supported by the National Science Foundation's Division of Mathematical Sciences Grant #1811428.

References

[1] ADLER, R. J. (2008). Some new random field tools for spatial analysis. *Stoch. Environm. Res. Risk Assessment* **22**, 809.
[2] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
[3] BISCIO, C. A. N., CHENAVIER, N., HIRSCH, C. AND SVANE, A. M. (2020). Testing goodness of fit for point processes via topological data analysis. *Electron. J. Statist.* **14**, 1024–1074.
[4] BOBROWSKI, O. AND ADLER, R. J. (2014). Distance functions, critical points, and the topology of random Čech complexes. *Homol. Homotopy Appl.* **16**, 311–344.
[5] BOBROWSKI, O. AND KAHLE, M. (2018). Topology of random geometric complexes: a survey. *J. Appl. Comput. Topol.* **1**, 331–364.
[6] BOBROWSKI, O. AND MUKHERJEE, S. (2015). The topology of probability distributions on manifolds. *Prob. Theory Relat. Fields* **161**, 651–686.
[7] CARLSSON, G. (2009). Topology and data. *Bull. Amer. Math. Soc.* **46**, 255–308.
[8] CRAWFORD, L., MONOD, A., CHEN, A. X., MUKHERJEE, S. AND RABADÁN, R. (2019). Predicting clinical outcomes in glioblastoma: an application of topological and functional data analysis. *J. Amer. Statist. Assoc.* **115**, 1139–1150.
[9] DECREUSEFOND, L., FERRAZ, E., RANDRIAMBOLOLONA, H. AND VERGNE, A. (2014). Simplicial homology of random configurations. *Adv. Appl. Prob.* **46**, 325–347.
[10] EDELSBRUNNER, H. AND HARER, J. (2010). *Computational Topology: an Introduction*. American Mathematical Society, Providence, RI.
[11] GOEL, A., TRINH, K. D. AND TSUNODA, K. (2019). Strong law of large numbers for Betti numbers in the thermodynamic regime. *J. Statist. Phys.* **174**, 865–892.
[12] HATCHER, A. (2002). *Algebraic Topology*. Cambridge University Press.
[13] HIRAOKA, Y., SHIRAI, T. AND TRINH, K. D. (2018). Limit theorems for persistence diagrams. *Ann. Appl. Prob.* **28**, 2740–2780.
[14] HUG, D., LAST, G. AND SCHULTE, M. (2016). Second-order properties and central limit theorems for geometric functionals of Boolean models. *Ann. Prob.* **26**, 73–135.
[15] KAHLE, M. (2011). Random geometric complexes. *Discrete Computat. Geom.* **45**, 553–573.
[16] KAHLE, M. AND MECKES, E. (2013). Limit theorems for Betti numbers of random simplicial complexes. *Homol. Homotopy Appl.* **15**, 343–374.
[17] KARATZAS, I. AND SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer, New York.
[18] OWADA, T. (2017). Functional central limit theorem for subgraph counting processes. *Electron. J. Prob.* **22**, 38 pp.
[19] OWADA, T. (2019). Topological crackle of heavy-tailed moving average processes. *Stoch. Process. Appl.* **129**, 4965–4997.
[20] OWADA, T. AND THOMAS, A. M. (2020). Limit theorems for process-level Betti numbers for sparse and critical regimes. *Adv. Appl. Prob.* **52**, 1–31.
[21] PENROSE, M. D. (2000). Central limit theorems for k-nearest neighbour distances. *Stoch. Process. Appl.* **85**, 295–320.
[22] PENROSE, M. D. (2003). *Random Geometric Graphs*. Oxford University Press.
[23] SCHNEIDER, R. AND WEIL, W. (2008). *Stochastic and Integral Geometry*. Springer, New York.
[24] WHITT, W. (2002). *Stochastic-Process Limits: an Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer, New York.
[25] YOGESHWARAN, D., SUBAG, E. AND ADLER, R. J. (2017). Random geometric complexes in the thermodynamic regime. *Prob. Theory Relat. Fields* **167**, 107–142.