Electronic Journal of Statistics
Vol. 14 (2020) 1024–1074
ISSN: 1935-7524
https://doi.org/10.1214/20-EJS1683

Testing goodness of fit for point processes via topological data analysis

Christophe A. N. Biscio*

Aalborg University, Department of Mathematical Sciences
Skjernvej 4, 9220 Aalborg Ø, Denmark
e-mail: christophe@math.aau.dk

Nicolas Chenavier†

Université Littoral Côte d'Opale
EA 2797, LMPA, 50 rue Ferdinand Buisson, 62228 Calais, France
e-mail: nicolas.chenavier@univ-littoral.fr

Christian Hirsch*

University of Groningen, Bernoulli Institute
Nijenborgh 9, 68161 Groningen, The Netherlands
e-mail: c.p.hirsch@rug.nl

Anne Marie Svane*

Aalborg University, Department of Mathematical Sciences
Skjernvej 4, 9220 Aalborg Ø, Denmark
e-mail: annemarie@math.aau.dk

Abstract: We introduce tests for the goodness of fit of point patterns via methods from topological data analysis. More precisely, the persistent Betti numbers give rise to a bivariate functional summary statistic for observed point patterns that is asymptotically Gaussian in large observation windows. We analyze the power of tests derived from this statistic on simulated point patterns and compare its performance with global envelope tests. Finally, we apply the tests to a point pattern from an application context in neuroscience. As the main methodological contribution, we derive sufficient conditions for a functional central limit theorem on bounded persistent Betti numbers of point processes with exponential decay of correlations.

MSC 2010 subject classifications: Primary 60D05; Secondary 55N20; 60F17.
Keywords and phrases: Point processes, goodness-of-fit tests, central limit theorem, topological data analysis, persistent Betti number.

Received June 2019.

*CB, CH and AS are supported by The Danish Council for Independent Research — Natural Sciences, grant DFF – 7014-00074 *Statistics for point processes in space and beyond*, and by the *Centre for Stochastic Geometry and Advanced Bioimaging*, funded by grant 8721 from the Villum Foundation.
†NC was partially supported by the French ANR grant ASPAG (ANR-17-CE40- 0017).

Contents

1. Introduction

Topological data analysis (TDA) provides insights into a variety of datasets by capturing some of their most salient properties via refined topological features. Since the mathematical field of topology specializes in describing invariants of objects independently of the choice of a precise metric, these features are robust against small perturbations or different embeddings of the object [12, 13]. Among the most classical topological invariants are the Betti numbers. Loosely speaking, they capture the number of k-dimensional holes of the investigated structure. TDA refines this idea substantially by constructing filtrations and

tracing when topological features appear and disappear. In point pattern analysis, simplicial complexes are built so that they are topologically equivalent to a union of disks with the same radius and centered at the data points, see the first three panels of Figure 1. As the radius increases, a sequence of simplicial complexes is then defined. Examples of such complexes are the basic Čech complex or the more elaborate α-complex, which is based on the Delaunay triangulation, see [19]. In that framework, 1-dimensional features correspond to loops in the simplicial complexes while 0-dimensional features correspond to connected components. When moving up in the filtration, additional edges appear and at some point create new loops. On the other hand, more and more triangles also appear, thereby causing completely filled loops to disappear. Usually, the filtration is indexed by time, and we refer to the appearance and disappearance of features as births and deaths. We refer the reader to [19] for a detailed presentation of these concepts. The *persistence diagram* visualizes the time points when the features are born and die, see the bottom-right panel in Figure 1. Persistent Betti numbers count the number of events in upper-left blocks of the persistence diagram and are also illustrated in the figure.

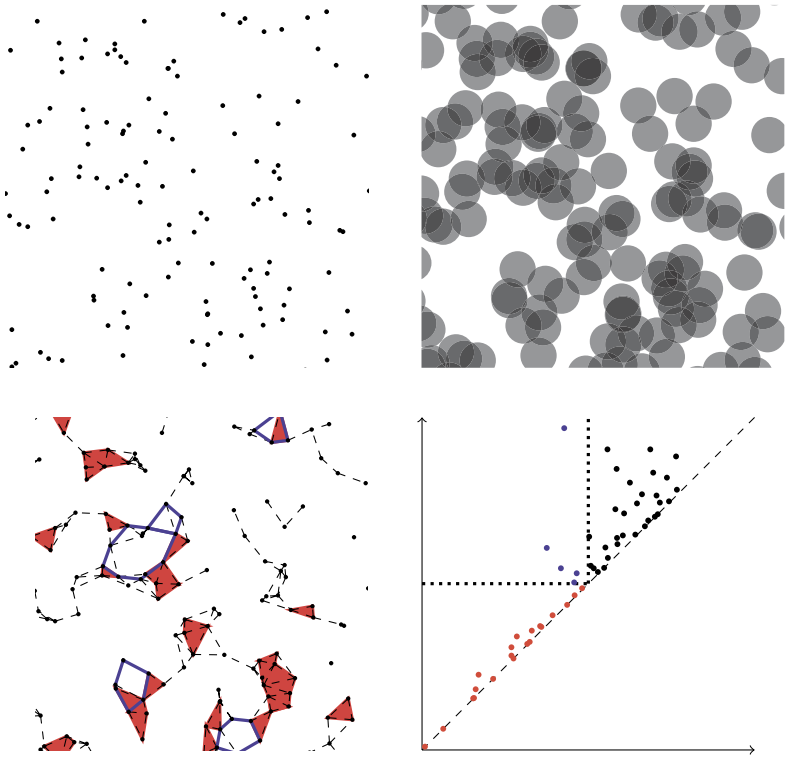

FIG 1. *Top: Realization of Poisson point process (left) and union of disks centered at the points of the process (right). Bottom: Alpha-complex corresponding to the union of disks with alive (blue) and dead (red) loops marked (left). Associated persistence diagram (right).*

In this paper, we leverage persistent Betti numbers to derive goodness-of-fit tests for planar point processes. Here, the abstract general definition of persistent Betti numbers gives way to a clear geometric intuition induced by a picture of growing disks centered at the points of the pattern and all having radius r, corresponding to the index of the filtration. Features of dimension 0 correspond to connected components in the union of disks, interpreted as point clusters, whereas boundaries of the complement set can be considered as the loops forming the 1-dimensional features. Since the notion of clusters in the sense of connected components lies at the heart of persistent Betti numbers in degree 0, they become highly attractive as a tool to detect clustering in point patterns. Our tests are based on a novel functional central limit theorem (CLT) for the persistent Betti numbers in large domains in $\mathbb{R}^2$ and outperform in certain cases tests based on Ripley's K-function (see e.g. Table 2). We think that investigating Betti numbers in higher dimension should also provide more efficient tests. The present work embeds into two active streams of current research.

First, now that TDA has become widely adopted, the community is vigorously working towards putting the approach on a firm statistical foundation paving the way for hypothesis testing. On the one hand, this encompasses large-sample Monte Carlo tests when working on a fixed domain [7, 10, 14]. Although these tests are highly flexible, the test statistics under the null hypothesis must be re-computed each time when testing observations in a different window. In large domains, this becomes time-consuming. On the other hand, there has been substantial progress towards establishing CLTs in large domains for functionals related to persistent Betti numbers [41, 42, 31, 37, 26]. However, these results are restricted to the null hypothesis of complete spatial randomness – i.e., the Poisson point process – and establish asymptotic Gaussianity on a multivariate, but not on a functional level. Our proof of a functional CLT is based on recently developed stabilization techniques for point processes with exponential decay of correlations [9]. As explained in the final section of [11], the main technical step towards a functional CLT are bounds on the cumulants.

Second, the introduction of global rank envelope tests has led to a novel surge of research activity in goodness-of-fit tests for point processes [36]. One of the reasons for their popularity is that they rely on functional summary statistics rather than scalar quantities. Thus, they reveal a substantially more fine-grained picture of the underlying point pattern. In the overwhelming majority of cases, variants of the K-function are used as a functional summary statistic, thereby essentially capturing the relative density of point pairs at different distances. Here, the persistent Betti numbers offer an opportunity to augment the basic second-order information by more refined characteristics of the data. Still, even for classical summary statistics, rigorous limit theorems in large domains remain scarce. For instance, a functional central limit theorem of the estimated K-function is proven in detail only for the Poisson point process in [23] and an extension to α-determinantal point processes is outlined in [24].

The rest of the manuscript is organized as follows. First, in Section 2, we introduce the concepts of M-bounded persistence diagrams and M-bounded persistent Betti numbers. Next, in Section 3, we state the two main results of the

paper, a CLT for the M-bounded persistence diagram and a functional CLT for the M-bounded persistent Betti numbers. In Section 4, we provide specific examples of point processes satisfying the conditions of the main results. Sections 5 and 6 explore TDA-based tests for simulated and real datasets, respectively. Finally, Section 7 summarizes the findings and points to possible avenues of future research. The proofs of the main results are deferred to Sections 8 and 9 of the appendix.

2. M-bounded persistent Betti numbers

For a locally finite point set $\mathcal{X} \subset \mathbb{R}^2$, persistent Betti numbers provide refined measures for the amount of clusters and voids on varying length scales. More precisely, we let

$$U_r(\mathcal{X}) = \bigcup_{x \in \mathcal{X}} B_r(x). \tag{1}$$

denote the union of closed disks of radius $r \geq 0$ centered at points in $\mathcal{X}$. A 0-dimensional topological feature is a connected component of this union, corresponding to a cluster of points in $\mathcal{X}$, while a 1-dimensional feature can be thought of as a bounded connected component of the background space, often identified with its boundary loop, and describes a vacant area in the plane. As the disks grow, new features arise and vanish; we say that they are born and die again. The persistent Betti numbers quantify this evolution of clusters and loops. Henceforth, we consider the persistence diagram only until a fixed deterministic radius $r_{\mathsf{f}} \geq 0$.

As r approaches the critical radius for continuum percolation, long-range phenomena emerge [33]. Thus, determining whether two points are connected could require exploring large regions in space. While useful quantitative bounds on cluster sizes are known for Poisson point processes [1], for more general classes of point processes the picture remains opaque and research is currently at a very early stage [28, 8]. Recently, a central limit theorem for persistent Betti numbers has been established in the Poisson setting [31, 26], but for general point processes the long-range interactions pose a formidable obstacle towards proving a fully-fledged functional CLT.

From a more practical point of view, these long-range dependencies are of less concern. Although large features can carry interesting information, we expect that spatially bounded topological features already provide a versatile tool for the statistical analysis of both simulated point patterns and real datasets, even when focusing only on features of a bounded size. For that purpose, we concentrate on features whose spatial diameter does not exceed a large deterministic threshold M.

To define these *M-bounded features*, we introduce the Gilbert graph $G_r(\mathcal{X})$ on the vertex set $\mathcal{X}$. The Gilbert graph $G_r(\mathcal{X})$ has for vertices the points in $\mathcal{X}$ and two points are connected by an edge if the distance between them is at most $2r$ or, equivalently, if the two disks of radius r centered at the points intersect.

2.1. M-bounded clusters

The 0-dimensional M-bounded features alive at time $r > 0$ are the connected components of $G_r(\mathcal{X})$ with diameter at most M. Starting at $r = 0$, all points belong to separate connected components that merge into larger clusters when r increases. We thus say that all components are *born at time 0*.

To define the death time of a component, let $\mathcal{C}_r(x)$ denote the connected component of $x \in \mathcal{X}$ in $G_r(\mathcal{X})$. The components of $x, y \in \mathcal{X}$ meet at time

$$R(x, y) = \inf\{r > 0 : \mathcal{C}_r(x) = \mathcal{C}_r(y)\}.$$

Then, the *death time* of $x \in \mathcal{X}$ is the smallest $R(x, y)$ such that the spatial diameter of $\mathcal{C}_r(x)$ exceeds M or such that P_x is lexicographically smaller than P_y, where P_x, P_y are the points of $\mathcal{C}_r(x) \cap \mathcal{X}$ and $\mathcal{C}_r(y) \cap \mathcal{X}$ whose associated disks meet at time $R(x, y)$. This ordering determines which component dies when two of them meet. See Figure 2.1a for an illustration.

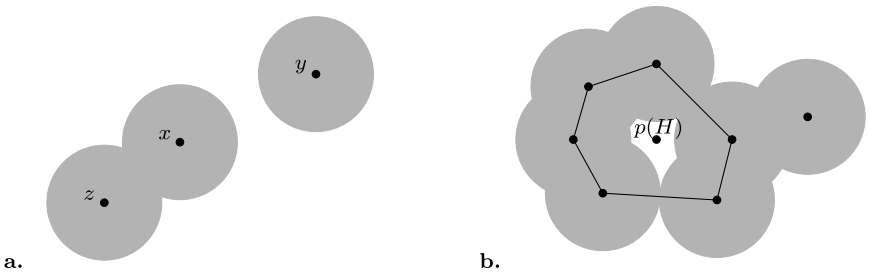

FIG 2. *a. The point z is already dead, while x is still alive, because z is lexicographically smaller than x. The point x will die when the balls around x and y meet because x is the lexicographically smaller of the two. b. A hole is formed when the two lower balls meet. The size of the hole is the diameter of the polygon formed by the centers of balls touching the hole at this time point. If this is smaller than M, we say that the hole is born. Otherwise, we wait until two balls merge to split the hole in two smaller pieces and recompute the size. The hole is identified with the point $p(H)$ which is covered last by the balls.*

2.2. M-bounded loops

Next, we introduce 1-dimensional features. At time $r > 0$, these correspond to holes, i.e., bounded connected components in the vacant phase $V_r(\mathcal{X}) = \mathbb{R}^2 \setminus U_r(\mathcal{X})$. In contrast to the clusters, there are no holes at time 0, so that both birth and death times must be specified. Moreover, it needs to be defined how holes are related for different radii r.

The *death time* of a hole H_s in $V_s(\mathcal{X})$ is the first time $r > s$ when the hole is completely covered by disks, i.e., $H_s \subseteq U_r(\mathcal{X})$. We identify a hole H with the point $p(H)$ that is covered last. Thus, holes H_s in $V_s(\mathcal{X})$ and H_r in $V_r(\mathcal{X})$, are identified if $p(H_r) = p(H_s)$.

New holes in $V_r(\mathcal{X})$ can only be formed when two balls merge, which corresponds to including a new edge in $G_r(\mathcal{X})$. A new hole can appear in two ways: either a finite component is separated from the infinite component, or an existing hole is split in two. In both cases, we define the size of the newly created piece(s) as follows: Let $x_1, \dots, x_k \in \mathcal{X}$ be the points in $\mathcal{X}$ such that the disks of radius r around the points intersect the boundary of the hole H in $V_r(\mathcal{X})$. Then, *the size of* H is the diameter of the set $\{x_1, \dots, x_k\}$. The size remains unchanged until the next time the hole is split into smaller pieces. Then the size is recomputed for both new pieces. This definition ensures that the size decreases when the balls grow and only changes when a new edge is added to $G_r(\mathcal{X})$.

The *birth time* of a hole H is the minimal s such that there is a hole H_s in $V_s(\mathcal{X})$ with $p(H) = p(H_s)$ and size less than M. By an M-bounded loop, we mean a loop with size lower than M. Figure 2.1b illustrates this definition.

2.3. The persistence diagram

We now adapt the definition of the persistence diagram in [26] to only include M-bounded features. That is, we define the *qth M-bounded persistence diagram*, $q \in \{0, 1\}$, as the empirical measure

$$\mathsf{PD}^{M,q}(\mathcal{X}) = \sum_{i \in I^{M,q}(\mathcal{X})} \delta_{(B_i^M, D_i^M)}, \tag{2}$$

where $I^{M,q}(\mathcal{X})$ is an index set over all M-bounded q-dimensional features that die before time r_{f} and B_i^M, D_i^M are the birth and death times of the ith feature. Then, the *qth M-bounded persistent Betti numbers*

$$\beta_{b,d}^{M,q}(\mathcal{X}) = \mathsf{PD}^{M,q}(\mathcal{X})([0, b] \times [d, r_{\mathsf{f}}])$$

are the number of M-bounded features born before time $b \geq 0$ and dead after time $d \leq r_{\mathsf{f}}$. When $q = 0$, all features are born at time 0, so that only death times are relevant. Hence, we write $\beta_d^{M,0}$ instead of the more verbose $\beta_{b,d}^{M,0}$.

3. Main results

Henceforth, $\mathcal{P}$ denotes a simple stationary point process in $\mathbb{R}^2$ with intensity $\rho > 0$. We think of $\mathcal{P}$ as a random variable taking values in the space of locally finite subsets $\mathcal{N}$ of $\mathbb{R}^2$ endowed with the smallest σ-algebra $\mathfrak{N}$ such that the number of points in any given Borel set becomes measurable. Throughout the manuscript, we assume that the factorial moment measures exist and are absolutely continuous. In particular, writing $\boldsymbol{x} = (x_1, \dots, x_p) \in \mathbb{R}^{2p}$, the *pth factorial moment density* $\rho^{(p)}$ is determined via the identity

$$\mathbb{E}\Big[\prod_{i \leq p} \mathcal{P}(A_i)\Big] = \int_{A_1 \times \cdots \times A_p} \rho^{(p)}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{3}$$

for any pairwise disjoint bounded Borel sets $A_1, \dots, A_p \subset \mathbb{R}^2$, where $\mathcal{P}(A_i)$ denotes the number of points of $\mathcal{P}$ in A_i. Moreover, as we rely on the framework of [9], we also require that $\mathcal{P}$ exhibits *exponential decay of correlations*. Loosely speaking this expresses an approximate factorization of the factorial moment densities and is made precise in Section 4 below. Many of the most prominent examples of point processes appearing in spatial statistics exhibit exponential decay of correlations [9, Section 2.2].

Our first main result is a CLT for the persistence diagram built on the restriction $\mathcal{P}_n = \mathcal{P} \cap W_n$ of the point process $\mathcal{P}$ to a large observation window $W_n = [-\sqrt{n}/2, \sqrt{n}/2]^2$. With a slight abuse of notation, we write $\mathcal{P} \cup \boldsymbol{x} = \mathcal{P} \cup \{x_1, \dots, x_p\}$. To prove the CLT, we impose an additional condition concerning moments under the *reduced p-point Palm distribution* $\mathbb{P}^!_{\boldsymbol{x}}$. We recall that this distribution is determined via

$$\mathbb{E}\Big[\sum_{(X_1,\dots,X_p)\in\mathcal{P}^p_{\neq}} f(X_1,\dots,X_p;\mathcal{P})\Big] = \int_{\mathbb{R}^{2p}} \mathbb{E}^!_{\boldsymbol{x}}[f(\boldsymbol{x};\mathcal{P}\cup\boldsymbol{x})]\rho^{(p)}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{4}$$

for any bounded measurable $f : \mathbb{R}^{2p} \times \mathcal{N} \to \mathbb{R}$, where $\mathcal{P}^p_{\neq}$ denotes p-tuples of pairwise distinct points in $\mathcal{P}$. In the following, $\mathbb{P}_{\boldsymbol{x}}$ denotes the unreduced Palm measure characterized via $\mathbb{E}_{\boldsymbol{x}}[f(\mathcal{P})] = \mathbb{E}^!_{\boldsymbol{x}}[f(\mathcal{P}\cup\boldsymbol{x})]$ for any non-negative measurable $f : \mathcal{N} \to [0,\infty)$. Then, we impose the following moment condition.

(M) For every $p \geq 1$

$$\sup_{\substack{l\leq p\\ \boldsymbol{x}\in\mathbb{R}^{2l}}} \mathbb{E}^!_{\boldsymbol{x}}[\mathcal{P}(W_1)^p] < \infty.$$

To state the CLT for the persistence diagram precisely, we let

$$\langle f, \mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle = \int_{[0,r_{\mathsf{f}}]^2} f(b,d)\mathsf{PD}^{M,q}(\mathcal{P}_n)(\mathrm{d}b,\mathrm{d}d) = \sum_{i\in I^{M,q}(\mathcal{P}_n)} f\big(B_i^M, D_i^M\big)$$

denote the integral of a bounded measurable function $f : [0, r_{\mathsf{f}}]^2 \to \mathbb{R}$ with respect to the measure $\mathsf{PD}^{M,q}(\mathcal{P}_n)$.

We first recall the definition of exponential decay of correlations from [9]. To this end, we define the *separation distance* between $\boldsymbol{x} = \{x_1, \dots, x_p\} \subset \mathbb{R}^2$ and $\boldsymbol{x}' = \{x_{p+1}, \dots, x_{p+q}\} \subset \mathbb{R}^2$ as in [9, Formula (1.3)] via

$$\mathsf{dist}(\boldsymbol{x}, \boldsymbol{x}') = \inf_{\substack{i\leq p\\ j\leq q}} |x_i - x_{p+j}|. \tag{5}$$

Definition 3.1. Let $\mathcal{P}$ be a stationary point process in $\mathbb{R}^2$, such that the k-point correlation function $\rho^{(k)}$ exists for all $k \geq 1$. Then, $\mathcal{P}$ exhibits *exponential decay of correlations* if there exist $a < 1$, $\phi : [0,\infty) \to [0,\infty)$ such that

1. $\lim_{t\to\infty} t^n\phi(t) = 0$ for all $n \geq 1$,
2. $\liminf_{t\to\infty} \log\phi(t)/t^b < 0$ for some $b > 0$,

3.
$$|\rho^{(p+q)}(\boldsymbol{x}\cup\boldsymbol{x}')-\rho^{(p)}(\boldsymbol{x})\rho^{(q)}(\boldsymbol{x}')|\le(p+q)^{a(p+q)}\phi(\mathsf{dist}(\boldsymbol{x},\boldsymbol{x}'))$$
for any $\boldsymbol{x}=\{x_1,\dots,x_p\},\boldsymbol{x}'=\{x_{p+1},\dots,x_{p+q}\}\subset\mathbb{R}^2$.

Theorem 3.2 (CLT for persistence diagrams)**.** *Let $M>0$, $q\in\{0,1\}$ and $f:[0,r_{\mathsf{f}}]^2\to\mathbb{R}$ be a bounded measurable function. Assume that $\mathcal{P}$ exhibits exponential decay of correlations and satisfies condition* **(M)**. *Furthermore, assume that $\liminf_{n\to\infty}\mathsf{Var}(\langle f,\mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle)n^{-\nu}=\infty$ for some $\nu>0$. Then,*
$$\frac{\langle f,\mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle-\mathbb{E}[\langle f,\mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle]}{\sqrt{\mathsf{Var}(\langle f,\mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle)}}$$
converges in distribution to a standard normal random variable as $n\to\infty$.

In order to derive a functional CLT for the persistent Betti numbers, we add a further constraint on $\mathcal{P}$, which is needed to establish a lower bound on the variance via a conditioning argument in the vein of [40, Lemma 4.3]. For this purpose, we consider a random measure Λ, which is jointly stationary with $\mathcal{P}$ and which we think of as capturing additional useful information on the dependence structure of $\mathcal{P}$. For instance, if $\mathcal{P}$ is a Cox point process, we choose Λ to be the random intensity measure. If $\mathcal{P}$ is a Poisson cluster process, then Λ would describe the cluster centers. If the dependence structure is exceptionally simple, it is also possible to take $\Lambda=0$. The idea of using additional information is motivated from conditioning on the spatially refined information coming from the clan-of-ancestors construction in Gibbsian point processes [40].

The point process $\mathcal{P}$ is *conditionally m-dependent* if $\mathcal{P}\cap A$ and $\mathcal{P}\cap A'$ are conditionally independent given $\sigma(\Lambda,\mathcal{P}\cap A'')$ for any bounded Borel sets $A,A',A''\subset\mathbb{R}^2$ such that the distance between A and A' is larger than some $m>0$. Here, $\sigma(\Lambda,\mathcal{P}\cap A'')$ denote the σ-algebra generated by Λ and $\mathcal{P}\cap A''$.

Finally, we impose an absolute continuity-type assumption on the Poisson point process in a fixed box with respect to $\mathcal{P}$ when conditioned on Λ and the outside points. More precisely, we demand that there exists $r_{\mathsf{AC}}>6M\vee 3r_{\mathsf{f}}$ with the following property, where $\mathcal{Q}$ denotes a homogeneous Poisson point process in the window $W_{r_{\mathsf{AC}}^2}$.

(AC) Let $E_1,E_2\in\mathfrak{N}$ be such that $\min_{i\in\{1,2\}}\mathbb{P}(\mathcal{Q}\in E_i)>0$. Then,
$$\mathbb{E}\big[\min_{i\in\{1,2\}}\mathbb{P}\big(\mathcal{P}_{r_{\mathsf{AC}}^2}\in E_i\,|\,\Lambda,\mathcal{P}\setminus W_{r_{\mathsf{AC}}^2}\big)\big]>0.$$

Although **(AC)** appears technical, Section 4 illustrates that it is tractable for many commonly used point processes.

Since the persistent Betti numbers exhibit jumps at the birth- and death times of features, we work in the Skorokhod topology [6, Section 14]. One of the main difficulties of this paper is that the functionals are not simple functionals depending on points or pairs of points but are more complex.

Theorem 3.3 (Functional CLT for persistent Betti numbers)**.** *Let $M > 0$ and $\mathcal{P}$ be a conditionally m-dependent point process with exponential decay of correlations and satisfying conditions* **(M)** *and* **(AC)**. *Then, the following convergence statements hold true.*

q=0. *The one-dimensional process*

$$\big\{n^{-1/2}\big(\beta_d^{M,0}(\mathcal{P}_n) - \mathbb{E}[\beta_d^{M,0}(\mathcal{P}_n)]\big)\big\}_{d\le r_{\mathsf{f}}}$$

converges weakly in Skorokhod topology to a centered Gaussian process.

q=1. *The two-dimensional process*

$$\big\{n^{-1/2}\big(\beta_{b,d}^{M,1}(\mathcal{P}_n) - \mathbb{E}[\beta_{b,d}^{M,1}(\mathcal{P}_n)]\big)\big\}_{b,d\le r_{\mathsf{f}}}$$

converges weakly in Skorokhod topology to a centered Gaussian process.

Additionally, proceeding as in [9, Theorem 1.12] we obtain convergence of the rescaled variances and also an expression for the covariance structure of the limiting Gaussian process in Theorem 3.3. To be more precise, let $f, g : [0, r_{\mathsf{f}}]^2 \to \mathbb{R}$ be bounded measurable functions. Then,

$$\lim_{n\to\infty} \frac{1}{n}\mathsf{Cov}(\langle f, \mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle, \langle g, \mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle) = \sigma_{f,g,q}^2,$$

where

$$\sigma_{f,g,q}^2 := \mathbb{E}_o\big[\xi_{f,q}(o,\mathcal{P})\xi_{g,q}(o,\mathcal{P})\big]\rho + \int_{\mathbb{R}^2} a_{f,g,q}(x)\mathrm{d}x. \tag{6}$$

Here,

$$a_{f,g,q}(x) = \mathbb{E}_{o,x}\Big[\xi_{f,q}(o,\mathcal{P})\xi_{g,q}(x,\mathcal{P})\Big]\rho^{(2)}(o,x) - \mathbb{E}_o[\xi_{f,q}(o,\mathcal{P})]\mathbb{E}_o[\xi_{g,q}(o,\mathcal{P})]\rho^2$$

and $\xi_{f,q}, \xi_{g,q}$ denote TDA-related scores, whose precise definition is given in identity (12) in Section 8.

Furthermore, it would be attractive to replace **(AC)** by the assumption that there exists $\nu > 0$ such that $n^{-\nu}\mathsf{Var}[\beta_{b,d}^{M,q}(\mathcal{P}_n)] \to \sigma^2(M,b,d)$ for all $b \le d \le r_{\mathsf{f}}$, and then establish the functional CLT under the scaling $n^{\nu/2}$. Indeed, this would open the door to studying point processes whose variance grows sub-volume order. However, the refined variance computations that are needed for the tightness argument in Section 9.2 are currently tied to having $\nu = 1$. It is an exciting avenue of further research to think about alternative approaches with the potential to work for more general $\nu > 0$.

As an application of Theorem 3.3, we obtain a functional CLT for the following two characteristics, which are modified variants of the *accumulated persistence function* from [7]:

$$\mathsf{APF}_r^{M,0}(\mathcal{P}_n) = \sum_{i\in I^{M,0}(\mathcal{P}_n)} D_i^M \mathbb{1}\{D_i^M \le r\}$$

and

$$\mathsf{APF}_r^{M,1}(\mathcal{P}_n) = \sum_{i\in I^{M,1}(\mathcal{P}_n)} (D_i^M - B_i^M)\mathbb{1}\{B_i^M \le r\}.$$

Corollary 3.4 (Functional CLT for the APF)**.** *Let $M > 0$ and $\mathcal{P}$ be as in Theorem 3.3. Then, both $\{n^{-1/2}(\mathsf{APF}_r^{M,0}(\mathcal{P}_n) - \mathbb{E}[\mathsf{APF}_r^{M,0}(\mathcal{P}_n)])\}_{r\le r_{\mathrm{f}}}$ and $\{n^{-1/2}(\mathsf{APF}_r^{M,1}(\mathcal{P}_n) - \mathbb{E}[\mathsf{APF}_r^{M,1}(\mathcal{P}_n)])\}_{r\le r_{\mathrm{f}}}$ converge to centered Gaussian processes.*

Moreover, we can again describe the covariance structure of the limit in Corollary 3.4. We provide the details only for $\mathsf{APF}_r^{M,1}$, since the expressions for $\mathsf{APF}_r^{M,0}$ are similar. Then,

$$\lim_{n\to\infty} \frac{1}{n}\mathsf{Cov}(\mathsf{APF}_r^{M,1}, \mathsf{APF}_{r'}^{M,1}) = \sigma^2_{(d-b)\mathbb{1}\{b\le r\},(d-b)\mathbb{1}\{b\le r'\},1},$$

where the right-hand side is defined in (6).

4. Examples of point processes

In this section, we give examples of point processes which satisfy the assumptions of our main theorems. More precisely, we show that log-Gaussian Cox processes with compactly supported covariance functions and Matérn cluster processes both satisfy the conditions of Theorems 3.2 and 3.3. We also show under which conditions determinantal point processes, and in particular the Ginibre point process, exhibit exponential decay of correlations and verify condition **(M)**. However, checking **(AC)** for determinantal point processes appears to be very challenging and will not be addressed in this paper.

Conversely, we do not expect that hard-core point processes satisfy the functional central limit theorem in the generality of Theorem 3.3. Indeed, hard-core conditions put a strict lower bound on the death time of clusters and the birth time of loops. We believe that suitable repulsive point processes, where the hard-core conditions only need to be imposed with a certain probability can be embedded in the framework of Theorem 3.3.

Note that in most cases, including the examples below, the theoretical Betti numbers are not known explicitly.

4.1. Log-Gaussian Cox process

Let $Y = \{Y(x)\}_{x\in\mathbb{R}^2}$ be a stationary Gaussian process with mean $\mu \in \mathbb{R}$ and covariance function $c(x, x') = c(x-x')$. Then, the random measure on $\mathbb{R}^2$ defined as $\mathbf{\Lambda}(B) = \int_B \exp(Y(x))\mathrm{d}x$, for any Borel subset $B \subset \mathbb{R}^2$ has moments of any order. Let $\mathcal{P}$ be a Cox process with random intensity measure $\mathbf{\Lambda}$, referred to as a *Log-Gaussian Cox process*. By [16, Equation (7)], the factorial moment densities of $\mathcal{P}$ are given by

$$\rho^{(j)}(u_1, \dots, u_j) = \exp\left(j\mu + \frac{jc(0)}{2}\right) \prod_{1\le i<i'\le j} \exp(c(u_i - u_{i'})).$$

To apply Theorems 3.2 and 3.3, we assume that c is bounded and of compact support, which ensures that $\mathcal{P}$ exhibits exponential decay of correlation.

We show below that condition (**M**) is satisfied. Let $\boldsymbol{x} = (x_1, \dots, x_l) \in \mathbb{R}^{2l}$. According to [17, Theorem 1], the Log-Gaussian Cox process $\mathcal{P}$ under the reduced Palm version is also a Log-Gaussian Cox process $\mathcal{P}_{\boldsymbol{x}}$ with underlying Gaussian process $Y_{\boldsymbol{x}}(x) = Y(x) + \sum_{i \le l} c(x, x_i)$. According to [18, Equation (5.4.5)],

$$\mathbb{E}^!_{\boldsymbol{x}}[\mathcal{P}(W_1)^p] = \mathbb{E}[\mathcal{P}_{\boldsymbol{x}}(W_1)^p] = \sum_{1 \le j \le p} \Delta_{j,l,p} \int_{W_1^j} \rho^{(j)}_{\boldsymbol{x}}(u_1, \dots, u_j) \mathrm{d}u_1 \cdots \mathrm{d}u_j$$

for suitable coefficients $\Delta_{j,l,p} \in \mathbb{R}$, where $\rho^{(j)}_{\boldsymbol{x}}(u_1, \dots, u_j)$ denotes the jth factorial moment density with respect to $\mathcal{P}_{\boldsymbol{x}}$. Therefore, it is enough to prove that

$$\sup_{\boldsymbol{x} \in \mathbb{R}^{2l}} \int_{W_1^j} \rho^{(j)}_{\boldsymbol{x}}(u_1, \dots, u_j) \mathrm{d}u_1 \cdots \mathrm{d}u_j < \infty,$$

for all $j, l \ge 1$. Now, Equation (8) in [16] gives that

$$\rho^{(j)}_{\boldsymbol{x}}(u_1, \dots, u_j) = \exp\Big(j\mu + \frac{jc(0)}{2} + \sum_{\substack{1 \le i \le j \\ 1 \le k \le l}} c(u_i, x_k)\Big) \prod_{1 \le i < i' \le j} \exp(c(u_i - u_{i'})),$$

where the right-hand side is bounded as μ and c are bounded independently of $\boldsymbol{x}$. This verifies condition **(M)**.

Since conditionally on $\boldsymbol{\Lambda}$, the point process $\mathcal{P}$ is a Poisson point process, the conditional m-dependence property holds with $\Lambda = \boldsymbol{\Lambda}$.

It remains to verify condition (**AC**). By [35, Equation (6.2)], conditionally on Λ, the distribution of the point process $\mathcal{P}_{r_{\mathsf{AC}}^2}$ admits the density with respect to a homogeneous Poisson point process $\mathcal{Q}$ with intensity 1 in $W_{r_{\mathsf{AC}}^2}$ given by

$$f_{\boldsymbol{\Lambda}}(\phi) = \exp(|W_{r_{\mathsf{AC}}^2}| - \boldsymbol{\Lambda}(W_{r_{\mathsf{AC}}^2})) \prod_{x \in \phi} \exp(Y(x)),$$

where $\phi \in \mathfrak{N}$. In particular, $f_{\boldsymbol{\Lambda}}(\phi)$ is strictly positive for all ϕ. Therefore, if E_1, E_2 are two events such that $\min_{i \in \{1,2\}} \mathbb{P}(\mathcal{Q} \in E_i) > 0$, then $\mathbb{P}(\mathcal{P}_{r_{\mathsf{AC}}^2} \in E_i \,|\, \Lambda = \boldsymbol{\Lambda}) > 0$. This verifies condition (**AC**).

4.2. Matérn cluster process

Let η be a homogeneous Poisson point process in $\mathbb{R}^2$ with intensity $\gamma > 0$. Given a realization of η, we define a family of independent point processes $(\Phi_x)_{x \in \eta}$, where Φ_x, $x \in \eta$, is a homogeneous Poisson point process with intensity 1 in the disk $B_R(x)$ of radius $R > 0$ centered at $x \in \mathbb{R}^2$. The point process $\mathcal{P} = \bigcup_{x \in \eta} \Phi_x$ is referred to as a *Matérn cluster process*. Since $\mathcal{P}$ is $2R$-dependent, it exhibits exponential decay of correlations.

Next, we verify condition **(M)**. For this purpose, we deduce from [17, Section 5.3.2] that a Matérn cluster process is a Cox process whose random intensity measure $\boldsymbol{\Lambda}$ has as density the random field $(\lambda(x))_{x \in \mathbb{R}^2}$ given by

$$\lambda(x) = \gamma \eta(B_R(x)).$$

Now, let $\boldsymbol{x} = (x_1, \dots, x_l) \in \mathbb{R}^{2l}$ and $p \geq 1$ be fixed. From [17, Equations (19) and (20)] we obtain that

$$\mathbb{E}^!_{\boldsymbol{x}}[\mathcal{P}(W_1)^p] = \frac{1}{\mathbb{E}\big[\prod_{i\leq l}\lambda(x_i)\big]} \cdot \mathbb{E}\big[\mathcal{P}(W_1)^p \prod_{i\leq l}\lambda(x_i)\big].$$

Since $\eta(B_R(x))$ is increasing in η for every $x \in \mathbb{R}^2$ in the sense of [32], the Harris-FKG inequality [32, Theorem 20.4] gives that

$$\mathbb{E}\big[\prod_{i\leq l}\lambda(x_i)\big] \geq \prod_{i\leq l}\mathbb{E}[\lambda(x_i)] = (\gamma\pi R^2)^l,$$

where we used that $\lambda(x_i) = \eta(B_R(x_i))$ is a Poisson random variable with parameter πR^2. In order to bound $\mathbb{E}\Big[\mathcal{P}(W_1)^p \prod_{i\leq l}\lambda(x_i)\Big]$, we first apply the Hölder inequality and stationarity, to arrive at

$$\mathbb{E}\Big[\mathcal{P}(W_1)^p \prod_{i\leq l}\lambda(x_i)\Big] \leq \mathbb{E}\big[\mathcal{P}(W_1)^{p(l+1)}\big]^{1/(l+1)}\mathbb{E}[\lambda(o)^{l+1}]^{l/(l+1)}.$$

First, $\mathbb{E}[\lambda(o)^{l+1}] = \mathbb{E}[\eta(B_R(o))^{l+1}]$ is finite since η is a Poisson point process. For the remaining part, we note that $\mathcal{P}(W_1) \leq \sum_{y\in\eta\cap(W_1\oplus B_R(o))} \#\Phi_x$, where $W_1 \oplus B_R(o) = \{x+y : x \in W_1,\, y \in B_R(o)\}$ denotes the Minkowski sum. Hence,

$$\begin{aligned}\mathbb{E}\big[\mathcal{P}(W_1)^{p(l+1)}\big] &\leq \mathbb{E}\Big[\Big(\sum_{x\in\eta\cap(W_1\oplus B_R(o))} \#\Phi_x\Big)^{p(l+1)}\Big] \\ &\leq \mathbb{E}\Big[\prod_{x\in\eta\cap(W_1\oplus B_R(o))} e^{p(l+1)\#\Phi_x}\Big] \\ &= \exp\big(\gamma|W_1 \oplus B_R(o)|(\mathbb{E}[e^{p(l+1)\#\Phi_0}] - 1)\big),\end{aligned}$$

where Φ_0 is a homogeneous Poisson point process of intensity 1 in the disk $B_R(o)$ [32, Theorem 3.9]. Again, since $\#\Phi_0$ is a Poisson random variable with parameter πR^2, the latter expression is finite. Taking the supremum over all $\boldsymbol{x}$ and all $l \leq p$, this verifies condition **(M)**. The point process $\mathcal{P}$ is also conditionally m-dependent, by taking $m = 2R$ and $\Lambda = \eta$.

It remains to prove (**AC**). By [35, Equation (6.2)], conditional on $\Lambda = \eta$, the distribution of $\mathcal{P}_{r^2_{\mathsf{AC}}}$ admits the density

$$f_\eta(\phi) = \gamma \exp(|W_{r^2_{\mathsf{AC}}}| - \mathbf{\Lambda}(W_{r^2_{\mathsf{AC}}})) \prod_{x\in\phi}\eta(B_R(x))$$

with respect to the distribution of a homogeneous Poisson point process. Now, consider the event

$$\mathcal{E} = \{W_{r^2_{\mathsf{AC}}} \subset \eta \oplus B_R(o)\},$$

the density f_η is positive. Therefore, if E_1, E_2 are such that $\min_{i\in\{1,2\}}\mathbb{P}(\mathcal{Q} \in E_i) > 0$, then almost surely

$$\min_{i\in\{1,2\}}\mathbb{P}(\mathcal{P}_{r^2_{\mathsf{AC}}} \in E_i|\eta)\mathbf{1}_{\mathcal{E}}(\eta) > 0.$$

Since $\mathcal{E}$ occurs with positive probability, this proves condition **(AC)**.

4.3. Determinantal point process

As mentioned in [9, Section 2.2.2], any determinantal point process with kernel K verifying for some function ϕ as in Definition 3.1,

$$|K(z_1, z_2)| \leq \phi(|z_1 - z_2|)$$

with $z_1, z_2 \in \mathbb{C}$, exhibits exponential decay. According to [22, Theorem 2], for $\boldsymbol{x} \in \mathbb{R}^{2l}$ we have $\mathbb{E}^!_{\boldsymbol{x}}[\mathcal{P}(W_1)^p] \leq \mathbb{E}[\mathcal{P}(W_1)^p]$, where the right-hand side is finite by [27, Lemma 4.2.6]. Hence, we obtain an upper bound for $\mathbb{E}^!_{\boldsymbol{x}}[\mathcal{P}(W_1)^p]$, which is independent of $\boldsymbol{x}$, thereby verifying condition **(M)**. In particular, the *Ginibre* point process, which is a determinantal point process with kernel

$$K(z_1, z_2) = \exp(z_1\overline{z_2}) \exp\left(-\frac{|z_1|^2 + |z_2|^2}{2}\right),$$

with $z_1, z_2 \in \mathbb{C}$, exhibits exponential decay and verifies condition **(M)**. To apply Theorem 3.2, it still remains to check that the variance condition is satisfied.

5. Simulation study

We elucidate in a simulation study, how cluster- and loop-based test statistics derived from Theorem 3.3 can detect deviations from complete spatial randomness (CSR) and how effective they are in comparison to classical goodness-of-fit tests. The simulations are carried out with `spatstat` and `TDA` [21, 2].

For the entire simulation study, the null model $\mathsf{Poi}(2)$ is a Poisson point process with intensity 2 in a 10×10 observation window. Moreover, we ignore the constraint of M-boundedness in the simulations. Although the proof of Theorem 3.3 relies on the M-boundedness, the simulation study illustrates that it is not critical to impose this condition.

5.1. Deviation tests

As a first step, we derive scalar cluster- and loop-based test statistics.

5.1.1. Definition of test statistics

As a test statistic based on clusters, we use the integral over the number of cluster deaths in a finite time interval $[0, r_{\mathsf{c}}]$ with $r_{\mathsf{c}} \leq r_{\mathsf{f}}$, i.e.,

$$\int_0^{r_{\mathsf{c}}} \mathsf{PD}^0(\mathcal{P}_n)([0, d])\mathrm{d}d.$$

After subtracting the mean, this test statistic becomes reminiscent of the classical Cramér-von-Mises statistic except that we do not consider squared deviations. Although squaring would make it easier to detect two-sided deviations,

it would also require knowledge of quantiles of the square integral of a centered Gaussian process. Albeit possible, this incurs substantial computational expenses. Hence, our simpler alternative has the appeal that as an integral of a Gaussian process, the test statistic is asymptotically normal and therefore characterized by its mean and variance.

However, this statistic requires that the intensity is known in advance, whereas in practice we need to estimate it from data. To arrive at an intensity-adapted version, we note that by scaling and change of variable, we have for every $a > 0$,

$$\int_0^{r_c/a} \mathsf{PD}^0(\mathcal{Q}_n)([0,d])\mathrm{d}d = \frac{1}{a}\int_0^{r_c} \mathsf{PD}^0((a\mathcal{P})_{a^2 n})([0,d])\mathrm{d}d.$$

Moreover, if $\mathcal{P} = \mathcal{Q}$ is a Poisson point process with intensity λ and if we choose $a = \sqrt{\lambda}$, then $\mathcal{Q}^* := a\mathcal{Q}$ becomes a Poisson point process with unit intensity. After dividing both sides by an, the above relation specializes to

$$\frac{1}{\sqrt{\lambda}n}\int_0^{r_c/\sqrt{\lambda}} \mathsf{PD}^0(\mathcal{Q}_n)([0,d])\mathrm{d}d = \frac{1}{\lambda n}\int_0^{r_c} \mathsf{PD}^0(\mathcal{Q}^*_{\lambda n})([0,d])\mathrm{d}d.$$

This computation motivates the intensity-adapted test statistic

$$T_{\mathsf{C}} = \frac{1}{\sqrt{\lambda}|W_n|}\int_0^{r_c/\sqrt{\lambda}} \mathsf{PD}^0(\mathcal{P}_n)([0,d])\mathrm{d}d. \tag{7}$$

Note that in practice, the intensity λ above is replaced by $\hat{\lambda}$, the standard estimator of the intensity. Of course, one could proceed without this rescaling, but in the simulation study, we found this statistic to have a better power.

As a test statistic based on loops, we use the accumulated persistence function, which aggregates the life times of all loops with birth times in a time interval $[0, r_{\mathsf{L}}]$ with $r_{\mathsf{L}} \le r_{\mathsf{f}}$. That is,

$$\int_{[0,r_{\mathsf{L}}]\times[0,\infty)} (d-b)\mathsf{PD}^1(\mathcal{P}_n)(\mathrm{d}b,\mathrm{d}d).$$

By Corollary 3.4, after centering and rescaling, this statistic converges in the large-volume limit to a normal random variable. As in the cluster-based test, we adapt to the intensity to obtain

$$T_{\mathsf{L}} = \frac{1}{\sqrt{\lambda}|W_n|}\int_{[0,r_{\mathsf{L}}/\sqrt{\lambda}]\times[0,\infty)} (d-b)\mathsf{PD}^1(\mathcal{P}_n)(\mathrm{d}b,\mathrm{d}d). \tag{8}$$

The statistics T_{C} and T_{L} are specific possibilities to define scalar characteristics from the persistence diagram. Depending on the application context other choices, such as APF^0 instead of T_{C} could be useful. However, in the simulation study below, we found the weighting by life times of clusters to be detrimental.

Finally, we compare the TDA-based statistics to a quantity derived from the classical Ripley L-function. More precisely, we let

$$T_{\mathsf{Rip}} = \int_0^{r_{\mathsf{Rip}}} \hat{L}(r)\mathrm{d}r \tag{9}$$

denote the integral of the estimated Ripley's L-function until a radius $r_{\mathsf{Rip}} > 0$. We note that there is no need for a normalization as $\hat{L}$ already accounts for an estimated intensity. By existing FCLTs for Ripley's K-function, we expect that T_{Rip} is also asymptotically normal [23].

5.1.2. Exploratory analysis

As alternatives to the Poisson null hypothesis, we consider the attractive Matérn cluster and the repulsive Strauss and Ginibre processes. More precisely, the Matérn cluster process $\mathsf{MatC}(2, 0.5, 1)$ features a Poisson parent process with intensity 2 and generates a $\mathsf{Poi}(1)$ number of offspring uniformly in a disk of radius 0.1 around each parent. The Strauss process $\mathsf{Str}(4, 0.6, 0.5)$ has interaction parameter 0.6 and interaction radius 0.5. The intensity parameter 4 was tuned so as to match approximately the intensity of the null model. Additionally, we include the Baddeley-Silverman process that is known for its complex higher-order interactions [3]. Figure 3 shows realizations of the null model and the alternatives.

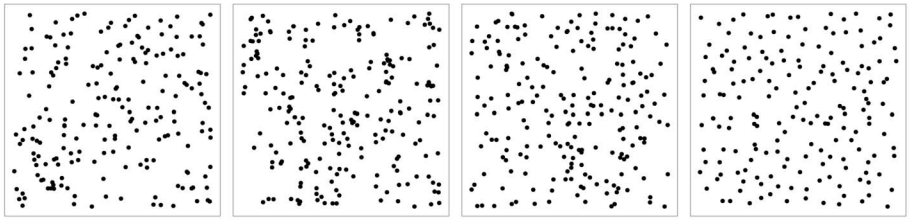

FIG 3. *From left to right: Samples from the* $\mathsf{Poi}(2)$ *null model, the* $\mathsf{MatC}(2, 0.5, 1)$ *process, the* $\mathsf{Str}(4, 0.6, 0.5)$ *process and the Baddeley-Silverman process.*

When considering persistence diagrams, we expect loosely speaking that more regular point patterns can lead to loops with shorter lifetimes and more clustered point patterns lead to longer lifetimes. Indeed, in a regular point patterns the points are at similar distances and we see only few exceptionally large holes. On the other hand, in clustered point patterns the loops connecting different cluster centers become much larger than the typical distance between points and therefore live for a substantial amount of time.

For the alternatives introduced above, Figure 4 illustrates the persistence diagram. From the cluster-based diagrams, it becomes apparent that in comparison to the null model, in the Matérn cluster process, there is a pronounced peak of deaths at early times, whereas this happens very rarely in the Strauss process. When analyzing loops, we see that loops with long life times appear earlier in the null model than in the Matérn cluster process. Conversely, while some loops with substantial life time emerge at later times in the null model, there are very few such cases in the Strauss model. Due to the complex higher-order interaction of the Baddeley-Silverman process, its behavior is difficult to

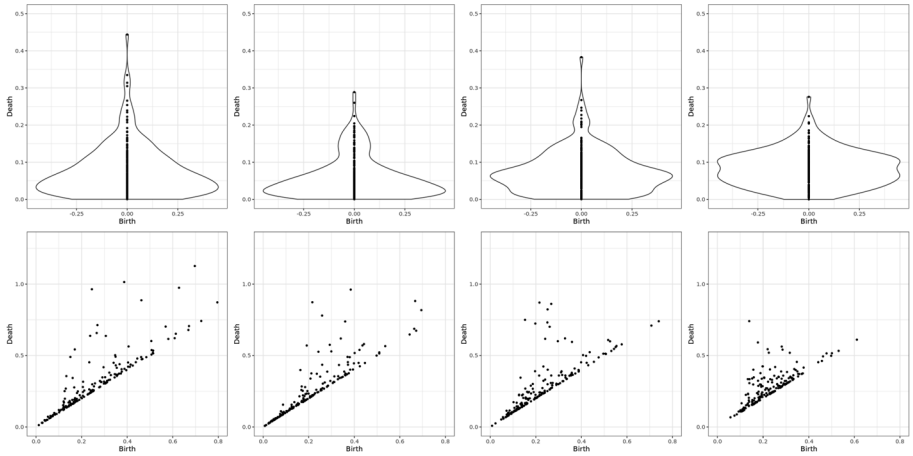

FIG 4. *Persistence diagrams for cluster-based features with density plots (top) and loop-based features (bottom) for the* $\mathsf{Poi}(2)$ *null model, the* $\mathsf{MatC}(2, 0.5, 1)$ *process, the* $\mathsf{Str}(4, 0.6, 0.5)$ *process and the Baddeley-Silverman process (from left to right).*

predict in advance. However, the samples in Figure 4 show that its topological characteristics are closer to those of a repulsive than a attractive point pattern.

These observations are not only true for the specific examples of the Matérn cluster and Strauss process, but extend more generally to attractive and repulsive point patterns. For different types of perturbed lattices, this philosophy is vividly illustrated in [41, Section 1].

5.1.3. Mean and variance under the null model

Now, we determine the mean and variance of T_{C} and T_{L} under the null model with $r_{\mathsf{f}} = 1.5$. For this purpose, we compute the number of cluster deaths and accumulated loop life times for 10,000 independent draws of the null model. Here, we normalize by $\sqrt{\lambda}|W_n|$ as in (7).

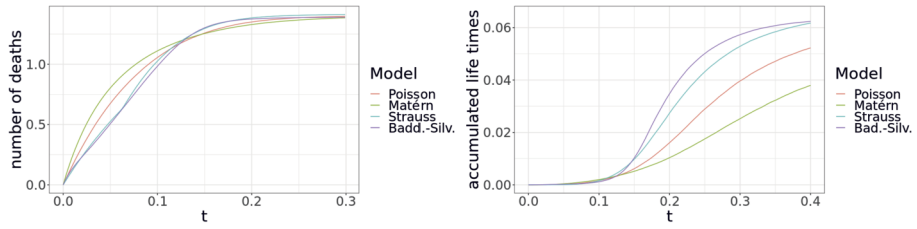

FIG 5. *Mean normalized number of cluster deaths (top) and accumulated loop life times (bottom) for the null model (red) and the alternatives (green, blue and pink) based on 10,000 realizations.*

Comparing the mean curves for the number of cluster deaths in the null model with those of the alternatives matches up nicely with the intuition about attraction and repulsion. For late times, they all approach a common value, namely the expected number of points in the observation window. However, Figure 5 shows that for the Matérn model, the slope is far steeper for early times, caused by merging of components of points within a cluster. In contrast, for the Strauss process the increase is at first much less pronounced than in the Poisson model, thereby reflecting the repulsive nature of the pair potential. The Baddeley-Silverman process is surprisingly similar to the Strauss model.

For the loops, a radically different picture emerges. Here, the curve for the Strauss process lies above the accumulated loop life times of the null model. The Strauss model spawns substantially more loops than the Poisson model, although most of them live for a shorter period. Still, taken together these competing effects lead to a net increase of the accumulated loop life times in the Strauss model. A similar picture also emerges for the Baddeley-Silverman process.

Finally, Table 1 shows the times needed to compute the test statistics for the 10,000 realizations on an AWS (Amazon Web Service) c5.9xlarge instance.

TABLE 1
Time needed for 10,000 evaluations of the test statistics.

T_{C}	T_{L}	T_{Rip}
4.13s	3.5s	3.4s

5.1.4. Type I and II errors

By Theorem 3.3, the statistics T_{C} and T_{L} are asymptotically normal, so that knowing the mean and variance allows us to construct a deviation test whose nominal confidence level is asymptotically exact. For the loops, we can choose the entire relevant time range, so that $r_{\mathsf{L}} = 0.5$. For the cluster features, this choice would be unreasonable, as for late times, we simply obtain the number of points in the observation window, which is not discriminative. Hence, we set $r_{\mathsf{C}} = 0.1$. We stress that in situations with no a priori knowledge of a good choice of r_{C}, the test power can degrade substantially.

To analyze the type I and II errors, we draw 1,000 realizations from the null model and from the alternatives, respectively. Next, Table 2 shows the rejection rates of this test setup. Under the null model the rejection rates are close to the nominal 5%-level, thereby illustrating that already for moderately large point patterns, the approximation by the Gaussian limit is accurate.

Using the mean and standard deviation from the null model, we now compute the test powers for the alternatives. The mean and deviations are $\mathbb{E}[T_{\mathsf{C}}] = 0.62$, $\mathsf{sd}(T_{\mathsf{C}}) = 0.047$, $\mathbb{E}[T_{\mathsf{L}}] = 0.028$ and $\mathsf{sd}(T_{\mathsf{L}}) = 0.0035$. Since Theorem 3.3 is designed for the type I error and not the type II error, it makes sense to study the Strauss process even though it may be very difficult to verify condition **(M)**. Already an inspection of Figure 3 reveals that the alternatives differ visibly from

the null hypothesis. This is reflected in the rejection rates when using the L-function based statistics T_{Rip}. More precisely, in the Matérn-cluster and Strauss process alternatives the null hypothesis is rejected in 92.6% and 86.3% of the cases. Using the TDA-based statistics T_{C} and T_{L} alone yields smaller rejection rates. However, Table 2 illustrates that the linear combination $T_{\mathsf{C}} - 10T_{\mathsf{L}}$ leads to comparable rejection rates as the L-function. This result based on ad-hoc coefficients illustrates how worthwhile it can be for future work to think about more conceptual approaches for choosing an appropriate linear functional.

Although in general, there may be no clear-cut reason why certain deviation tests work better for specific types of point processes, sometimes more can be said. For instance, the pair correlation function of the Baddeley-Silverman process on infinite sampling windows coincides with that of a Poisson point process. Consequently, the L-function exhibits a low power as it fails to take into account higher-order interactions. In this setting, all three TDA-based statistic outperform the L-function statistics. We have not included the Ginibre alternative in the table as all four tests reject at a perfect rate of 100%.

TABLE 2
Rejection rates for the test statistics under the CSR null model and the alternatives.

	Poi	MatC	Str	Badd. − Silv.
T_{C}	4.8%	55.7%	52.0%	65.6%
T_{L}	4.5%	63.0%	54.5%	84.7%
$T_{\mathsf{C}} - 10T_{\mathsf{L}}$	4.5%	87.9%	89.2%	88.1%
T_{Rip}	5.0%	92.6%	86.3%	59.3%

5.1.5. Null hypothesis of clustering

Section 5.1.4 discusses how to test the null hypothesis of complete spatial randomness via cluster- and loop-based statistics. In the present section, we illustrate at the hand of a Matérn cluster process that these test statistics also allow for testing clustering or regularity of the point patterns. In other words, the null hypothesis is now a Matérn cluster process $\mathsf{MatC}(2, 0.5, 1)$.

Similarly to the Poisson null model, in practice the intensity of the Matérn cluster process is not known but must be estimated from data. However, for the Matérn process the issue is more severe, since it is described by three parameters that all need to be estimated. For this purpose, we resort to the minimum contrast method [35, Chapter 10].

In addition, now that the null model does not depend any longer on a single parameter, we cannot apply a simple rescaling to arrive at an intensity-adapted version of the estimators. Therefore, we need to resort to nested Monte Carlo simulations. That is, once we estimated the parameters of the Matérn process, we determine the mean and variance of the test statistics based on $n_{\mathsf{nest}} = 100$ independent samples. Due to the complexity of the parameter estimation, we should expect substantial effects on the empirical results.

Table 3 summarizes the results of the simulation study of 1,000 simulation runs when using complete spatial randomness as alternative to the null hypoth-

esis of clustering. First, all tests are conservative as the rejection rates under the null hypothesis are far below the nominal 5%-level. This is particularly pronounced for T_{C} and T_{Rip} where the null-hypothesis is never rejected. Regarding the type II error with CSR alternative, although the linearly combined TDA-based test statistic $T_{\mathsf{C}} - 10T_{\mathsf{L}}$ is outperformed by T_{Rip}, it still exhibits a reasonable rejection rate of 71.2%.

TABLE 3
Rejection rates for the test statistics under the clustered null model and the alternatives.

	MatC	Poi
T_{C}	0.0%	39.1%
T_{L}	2.6%	51.1%
$T_{\mathsf{C}} - 10T_{\mathsf{L}}$	2.8%	71.2%
T_{Rip}	0.0%	82.7%

5.2. Envelope tests

Leveraging Theorem 3.3 shows that the deviation statistics T_{C} and T_{L} are asymptotically normal. Using a simulation-based estimate for the asymptotic mean and variance under the null model allowed us to construct a deviation test whose confidence level is asymptotically precise.

Recently, global envelope tests have gained widespread popularity, because they are both powerful and provide graphical insights as to why a null hypothesis is rejected [36]. The global envelope tests are fundamentally Monte Carlo-based tests and therefore do not relate directly to the large-volume CLT. However, they also rely on a functional summary statistic as input. Most of the applications in spatial statistics use a distance-based second-order functional such as Ripley's L-function. In this section, we compare such classical choices with cluster- and loop-based statistics.

We follow the simulation set-up of Section 5.1.4. That is, the null hypothesis is complete spatial randomness, and we study Matérn-cluster, Strauss and Baddeley-Silverman process as alternatives. Since the Ginibre process is already perfectly set apart from complete spatial randomness by the deviation tests, we do not consider it here further.

For the convenience of the reader, we outline the basic framework of global envelope tests and refer to [36] for details. Practitioners in spatial statistics value highly functional summary statistics of the form $T = \{T(r)\}_{r\geq 0}$ that can describe properties of the process at different scales. Popular examples for T include Ripley's L-function and its variants. The idea behind envelopes is to provide a Monte-Carlo based tool for assessing how well a given point pattern conforms with a null hypothesis.

More precisely, let $T_1 = \{T_1(r)\}_{r\geq 0}$ denote the curve associated with an observed point pattern and $T_2, \ldots, T_{s+1}$ result from s independent samples of the null model. Then, for each fixed $r > 0$ and $k \geq 1$, we define

$$T^k_{\mathsf{low}}(r) = \mathsf{min}^k_{1\leq i\leq s+1} T_i(r)$$
$$T^k_{\mathsf{upp}}(r) = \mathsf{max}^k_{1\leq i\leq s+1} T_i(r)$$

as the kth smallest and kth largest value of the family $\{T_1(r), T_2(r), \dots, T_{s+1}(r)\}$. Then, we refer to the curves $T^k_{\mathsf{low}} = \{T^k_{\mathsf{low}}(r)\}_{r\geq 0}$ and $T^k_{\mathsf{upp}} = \{T^k_{\mathsf{upp}}(r)\}_{r\geq 0}$ as lower and upper envelopes.

If we fix k first, then the observed curve $\{T_1(r)\}_{r\geq 0}$ leaving the envelope at some index r is often taken as an indication that the observation does not conform with the null hypothesis. However, as the order statistics in the definition of T^k_{low} and T^k_{upp} are computed separately for each index, we would naïvely only obtain an exact significance test if we considered only a single index, which has been fixed in advance.

In order to arrive at a test with exact significance level accounting for several values of r, the global rank envelope test orders the entire curves $T_1, \dots, T_{s+1}$ according to their *extreme rank measures* $R_1, \dots, R_{s+1}$ given by

$$R_i := \max\{k : T^k_{\mathsf{low}}(r) \leq T_i(r) \leq T^k_{\mathsf{upp}}(r) \text{ for all } r\}.$$

In words, R_i measures the centrality of $T_i(r)$ by determining the largest k such that kth envelopes contain T_i.

To turn these definitions into a test at a significance level α, we let

$$k_\alpha := \max\{k : \#\{i : R_i < k\} \leq (s+1)\alpha\}$$

denote the largest integer k such that the number of curves with rank at most k does not exceed $(s+1)\alpha$. Then, the envelope test rejects the null hypothesis if the observed statistic $T_1(r)$ falls outside the interval $[T^{k_\alpha}_{\mathsf{low}}, T^{k_\alpha}_{\mathsf{upp}}]$ for some $r \geq 0$.

5.2.1. Power analysis

Next, we analyze the power of the envelope test for the null-hypothesis and alternatives described in Section 5.1.4. Before presenting the results, we explain in detail the choice of parameters in the envelope tests in [36]. First, we choose a significance level of $\alpha = 5\%$. Second, we follow the suggestion in [36, Section 8] and generate $s = 2,499$ realizations of the null model. The most delicate parameter involves the interval from which to select the r-values. To be consistent with the simulation study of the deviation tests, we take the same values for r_{C}, r_{L} and r_{Rip}. Although other choices may also be of interest, we refer to the discussion in [36, Section 10] revealing that the envelope tests are highly robust with respect to changing the interval.

Then, we perform the global envelope test from [36] with four summary statistics: $T^{\mathsf{env}}_{\mathsf{C}}$, $T^{\mathsf{env}}_{\mathsf{L}}$, $T^{\mathsf{env}}_{\mathsf{C}} - 10T^{\mathsf{env}}_{\mathsf{L}}$, $T^{\mathsf{env}}_{\mathsf{Rip}}$. These test statistics are defined as the analogs in Section 5.1, except that we do not compute the averaging integral.

The rejection rates from Table 4 illustrate that while the L-based test benefits from the functional statistics, this is not the case for the TDA-based tests. Nevertheless, we still see that for the complex Baddeley-Silverman process the linearly combined statistics $T^{\mathsf{env}}_{\mathsf{C}} - 10T^{\mathsf{env}}_{\mathsf{L}}$ achieves an impressive rate of 92.8% and outperforms the L-based envelope test with rejection rate of 58.5%.

TABLE 4
Rejection rates for the envelope test under the CSR null model and the alternatives.

	Poi	MatC	Str	Badd. − Silv.
$T_{\mathsf{C}}^{\mathsf{env}}$	6.1%	69.0%	46.6%	64.9%
$T_{\mathsf{L}}^{\mathsf{env}}$	4.3%	41.9%	35.3%	69.8%
$T_{\mathsf{C}}^{\mathsf{env}} - 10T_{\mathsf{L}}^{\mathsf{env}}$	5.9%	81.6%	80.0%	92.4%
$T_{\mathsf{Rip}}^{\mathsf{env}}$	4.6%	92.6%	97.6%	73.1%

6. Analysis of the minicolumn dataset

In this section, we explore to what extent the deviation tests from Section 5 provide insights when dealing with real data. For this purpose, we analyze the minicolumn dataset provided by scientists at the Centre for Stochastic Geometry and Advanced Bioimaging.

As it should serve only to illustrate the application of Theorem 3.3, the present analysis is very limited in scope, and we refer to [15] for a far more encompassing study. For instance, that work considers two datasets and investigates 3D data together with marks for the directions attached to the neurons.

6.1. Exploratory analysis

The minicolumn dataset consists of 634 points emerging as two-dimensional projections of a three-dimensional point pattern of neurons. As neurons are believed to arrange in vertical columns, the projections are expected to exhibit clustering, see [34, 39]. The projections are taken along z-axis, since neuroscientists expect an arrangement in vertical columns. A visual inspection of the point pattern in Figure 6 supports this hypothesis.

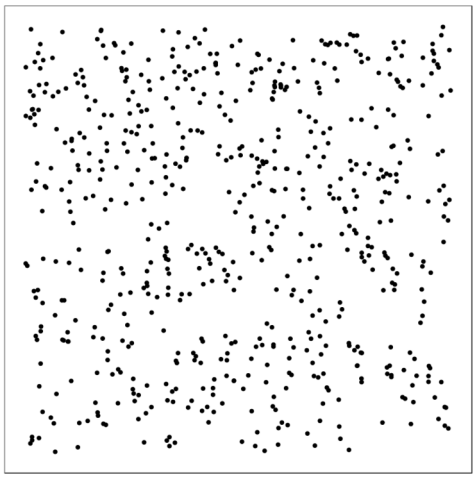

FIG 6. *Projected minicolumn point pattern*

As a first step, we explore whether the purported clustering already manifests in the persistence diagram. Comparing the loop-based persistence diagram of the minicolumn data with the persistence diagram of a homogeneous Poisson point process in Figure 7 shows that loops with substantial life times tend to be born later in the minicolumn model. This suggests clustering since loops formed by points within a cluster typically disappear rapidly.

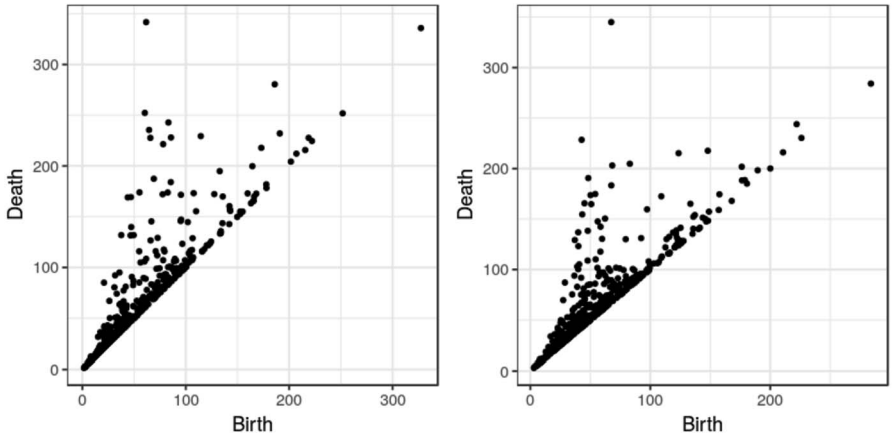

FIG 7. *Persistence diagram for the minicolumn data (left) and a homogeneous Poisson point process with the same intensity (right)*

Now, we explore whether the impressions from the persistence diagrams are reflected in the summary statistics from Section 5. When comparing in Figure 8 (left) the number of cluster death at different points in time, we note that until time 35, the curve for the observed data runs a bit above the curve for the null model. This provides already a first indication towards clustering. Next, we proceed to the loop-based features. As shown in Figure 8 (right), the curve for the observed pattern runs substantially below the one of the null model. This reflects a property that we have seen already in the persistence diagram: loops with substantial life time tend to be born earlier in the null model, thereby leading to a steeper increase of the accumulated life times.

6.2. Test for complete spatial randomness

Under the impression of the previous visualizations, we now test the minicolumn pattern against the null model. As in Section 5, we deduce from Theorem 3.3 that the statistics are asymptotically normal under the null model, so that we only need to determine means and variances.

When given a specific dataset, a subtle issue concerns the choice of the integration interval. The simplest option would be to take the whole intervals shown in Figure 8. For instance, for the loop-based features, this means $r_{\mathsf{L}} = r_{\mathsf{f}} = 120$. However, for the cluster-based features the choice of the interval

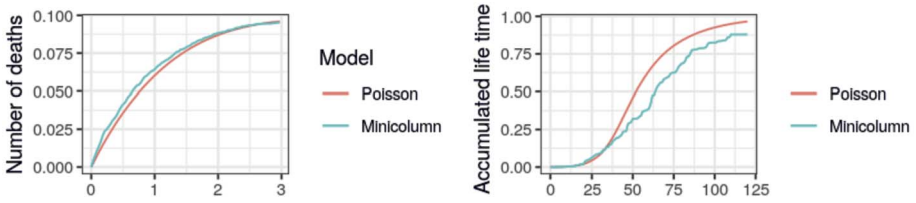

FIG 8. *Number of cluster deaths (left) and accumulated loop life times (right) for the Poisson null model (red) and the minicolumn dataset (green).*

is less clear, since taking the whole interval is not discriminatory. Therefore, we choose $r_{\mathsf{C}} = 1$.

As a general strategy, we propose that visualizations of summary statistics such as the ones in Figure 8 should guide the choice of r_{C} and r_{L}. If there are r-regions where the plot for the data differs substantially from that of the null model, these are good regions for selecting candidates for r_{C} and r_{L}. Since this is an ad hoc procedure, we strongly advise to study how sensitive the results are with respect to different choices of r_{C} and r_{L}.

Choosing $r_{\mathsf{C}} = 1$ and $r_{\mathsf{L}} = 120$, both the cluster-based and the loop-based test reject the null-hypothesis at the 5%-level, since the corresponding p-values are 0.18% and 0.031%. The tests are fairly robust with respect to the choice of the integration bound. More precisely, when changing the integration domains for the cluster- and loop-based tests to $[0, 2]$ and $[0, 70]$, then the null hypothesis is still rejected with p-values 4.55% and 1.4%. However, if we change the intervals to $[0, 2.5]$ and $[0, 60]$, then the p-value decrease to 10.7% and 5.1% so that the null-hypothesis is no longer rejected.

7. Discussion

In this paper, we elucidated how to apply tools from TDA to derive goodness-of-fit tests for planar point patterns. For this purpose, we derived sufficient conditions for a large-domain functional CLT for the M-bounded persistent Betti numbers on point processes exhibiting exponential decay of correlations. Following the framework developed in [9], the main difficulty arose from a detailed analysis of geometric configurations when bounding higher-order cumulants.

A simulation study revealed that the asymptotic Gaussianity is already accurate for patterns consisting of a few hundred data points. Additionally, as functional summary statistics, the persistent Betti numbers can also be used in the context of global envelope tests. Here, our finding is that TDA-based statistics can provide helpful additional information for point patterns with complex interactions.

Finally, we applied the TDA-based tests on a point pattern from a neuroscientific dataset. As conjectured from the application context, the functional summary statistics indicate a clustering of points and the tests reject the Pois-

son null-model. However, the analysis also reveals a certain sensitivity to the range of birth times considered in the statistics.

In future work, we plan to extend the present analysis to dimensions larger than 2. On a technical level, the definition of higher-dimensional features requires a deeper understanding of persistent homology groups. Additionally, when thinking of broader application scenarios, a further step is to extend the testing framework from mere point patterns to random closed sets involving a richer geometric structure.

Acknowledgments

We thank the two referees for their exceptionally detailed reports. In particular, they suggested to use intensity-adapted test statistics and to use a combination of cluster- and loop-based statistics. We thank J. Møller and A. D. Christoffersen for valuable discussions on the minicolumn dataset and for providing references. We also thank J. Yukich for inspiring discussions and helpful remarks. Finally, we thank the Centre for Stochastic Geometry and Advanced Bioimaging for collecting and sharing the data.

References

[1] D. Ahlberg, V. Tassion, and A. Teixeira. Sharpness of the phase transition for continuum percolation in $\mathbb{R}^2$. *Probab. Theory Related Fields*, 172(1):525–581, 2018. MR3851838

[2] A. Baddeley and R. Turner. spatstat: An R package for analyzing spatial point patterns. *J. Stat. Softw.*, 12(6):1–42, 2005.

[3] A. J. Baddeley and B. W. Silverman. A cautionary example on the use of second-order methods for analyzing point patterns. *Biometrics*, 40(4):1089–1093, 1984. MR0786181

[4] Y. Baryshnikov and J. E. Yukich. Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.*, 15(1A):213–253, 2005. MR2115042

[5] P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.*, 42:1656–1670, 1971. MR0383482

[6] P. Billingsley. *Convergence of Probability Measures*. J. Wiley & Sons, New York, second edition, 1999. MR1700749

[7] C. A. N. Biscio and J. Møller. The accumulated persistence function, a new useful functional summary statistic for topological data analysis, with a view to brain artery trees and spatial point process applications. *J. Comput. Graph. Statist.*, 28(3):671–681, 2019. MR4007749

[8] B. Błaszczyszyn and D. Yogeshwaran. Clustering and percolation of point processes. *Electron. J. Probab.*, 18:1–20, 2013. MR3091718

[9] B. Błaszczyszyn, D. Yogeshwaran, and J. E. Yukich. Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Probab.*, 47(2):835–895, 2019. MR3916936

[10] P. Bubenik. Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.*, 16:77–102, 2015. MR3317230

[11] P. Calka, T. Schreiber, and J. E. Yukich. Brownian limits, local limits and variance asymptotics for convex hulls in the ball. *Ann. Probab.*, 41(1):50–108, 2013. MR3059193

[12] G. Carlsson. Topology and data. *Bull. Amer. Math. Soc.*, 46(2):255–308, 2009. MR2476414

[13] F. Chazal and M. Bertrand. High-dimensional topological data analysis. In C. D. Toth, J. O'Rourke, and J. E. Goodman, editors, *Handbook of Discrete and Computational Geometry*. CRC, Boca Raton, third edition, 2017. MR3793131

[14] F. Chazal, B. T. Fasy, F. Lecci, A. Rinaldo, and L. Wasserman. Stochastic convergence of persistence landscapes and silhouettes. *J. Comput. Geom.*, 6(2):140–161, 2015. MR3323391

[15] A. D. Christoffersen, J. Møller, and H. S. Christensen. Modelling columnarity of pyramidal cells in the human cerebral cortex. *arXiv preprint arXiv:1908.05065*, 2019.

[16] J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. Palm distributions for log Gaussian Cox processes. *Scand. J. Stat.*, 44(1):192–203, 2017. MR3619701

[17] J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. A tutorial on Palm distributions for spatial point processes. *Int. Stat. Rev.*, 85(3):404–420, 2017. MR3723609

[18] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York, second edition, 2003. MR1950431

[19] H. Edelsbrunner and J. Harer. *Computational Topology*. American Mathematical Society, Providence, RI, 2010. MR2572029

[20] P. Eichelsbacher, M. Raič, and T. Schreiber. Moderate deviations for stabilizing functionals in geometric probability. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(1):89–128, 2015. MR3300965

[21] B. T. Fasy, J. Kim, F. Lecci, and C. Maria. Introduction to the R package TDA. *arXiv preprint arXiv:1411.1830*, 2014.

[22] A. Goldman. The Palm measure and the Voronoi tessellation for the Ginibre process. *Ann. Appl. Probab.*, 20(1):90–128, 2010. MR2582643

[23] L. Heinrich. Gaussian limits of empirical multiparameter K-functions of homogeneous Poisson processes and tests for complete spatial randomness. *Lith. Math. J.*, 55(1):72–90, 2015. MR3323283

[24] L. Heinrich. On the strong Brillinger-mixing property of α-determinantal point processes and some applications. *Appl. Math.*, 61(4):443–461, 2016. MR3532253

[25] L. Heinrich and V. Schmidt. Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. in Appl. Probab.*, 17(4):709–730, 1985. MR0809427

[26] Y. Hiraoka, T. Shirai, and K. D. Trinh. Limit theorems for persistence diagrams. *Ann. Appl. Probab.*, 28(5):2740–2780, 2018. MR3847972

[27] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian*

Analytic Functions and Determinantal Point Processes. American Mathematical Society, Providence, 2009. MR2552864

[28] S. Jansen. Continuum percolation for Gibbsian point processes with attractive interactions. *Electron. J. Probab.*, 21:No. 47, 22, 2016. MR3539641

[29] J. L. Jensen and H. R. Künsch. On asymptotic normality of pseudo likelihood estimates for pairwise interaction processes. *Ann. Inst. Statist. Math.*, 46(3):475–486, 1994. MR1309718

[30] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, second edition, 2002. MR1876169

[31] J. T. N. Krebs and W. Polonik. On the asymptotic normality of persistent Betti numbers. *arXiv preprint arXiv:1903.03280*, 2019.

[32] G. Last and M. Penrose. *Lectures on the Poisson process*. Cambridge University Press, Cambridge, 2018. MR3791470

[33] R. Meester and R. Roy. *Continuum Percolation*. Cambridge University Press, Cambridge, 1996. MR1409145

[34] J. Møller, F. Safavimanesh, and J. G. Rasmussen. The cylindrical K-function and Poisson line cluster point processes. *Biometrika*, 103(4):937–954, 2016. MR3620449

[35] J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. CRC, Boca Raton, 2004. MR2004226

[36] M. Myllymäki, T. Mrkvička, P. Grabarnik, H. Seijo, and U. Hahn. Global envelope tests for spatial processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 79(2):381–404, 2017. MR3611751

[37] T. Owada and A. Thomas. Limit theorems for process-level Betti numbers for sparse, critical, and Poisson regimes. *Adv. in Appl. Probab.*, 2020, to appear.

[38] G. Peccati and M. S. Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams*. Springer, Milan, 2011. MR2791919

[39] A. H. Rafati, F. Safavimanesh, K.-A. Dorph-Petersen, J. G. Rasmussen, J. Møller, and J. R. Nyengaard. Detection and spatial characterization of minicolumnarity in the human cerebral cortex. *Journal of Microscopy*, 261(1):115–126, 2016.

[40] A. Xia and J. E. Yukich. Normal approximation for statistics of Gibbsian input in geometric probability. *Adv. in Appl. Probab.*, 47(4):934–972, 2015. MR3433291

[41] D. Yogeshwaran and R. J. Adler. On the topology of random complexes built over stationary point processes. *Ann. Appl. Probab.*, 25(6):3338–3380, 2015. MR3404638

[42] D. Yogeshwaran, E. Subag, and R. J. Adler. Random geometric complexes in the thermodynamic regime. *Probab. Theory Related Fields*, 167(1-2):107–142, 2017. MR3602843

8. Proof of Theorem 3.2

The main tool to prove Theorem 3.2 is the general CLT [9, Theorem 1.14]. To make the paper self-contained, we state this theorem and recall the framework of [9].

Framework and CLT of [9] Let $\xi : \mathbb{R}^2 \times \mathcal{N} \to \mathbb{R}$ be a function referred to as a *score function*. Given $p \geq 1$, we say that the pair $(\xi, \mathcal{P})$ satisfies the *p-moment condition* if

$$\sup_{\substack{n \geq 1 \\ l \leq p,\, \boldsymbol{x} \in \mathbb{R}^{2l}}} \mathbb{E}_{\boldsymbol{x}}[|\xi(x_1, \mathcal{P}_n)|^p] < \infty. \tag{10}$$

We recall that, given $\mathcal{X} \in \mathcal{N}$ and $x \in \mathcal{X}$, the *radius of stabilization* $R^{\xi}(x, \mathcal{X})$ is defined as the smallest $r \in \mathbb{Z}_{\geq 0}$ such that

$$\xi(x, \mathcal{X} \cap B_r(x)) = \xi(x, (\mathcal{X} \cap B_r(x)) \cup (A \cap B_r^c(x)))$$

for all $A \in \mathcal{N}$. If no such finite r exists, we set $R^{\xi}(x, \mathcal{X}) = \infty$.

The main CLT of [9, Theorem 1.14] deals with rather general score functions in the sense that the radius of stabilization is assumed exponentially stabilizing on the input $\mathcal{P}$. Since we work with M-bounded features, we state the theorem only for bounded radii of stabilization.

Theorem 8.1 (CLT for score functions)**.** *Let $(\xi, \mathcal{P})$ be such that the following properties hold:*

- *the p-moment condition (10) holds for all $p > 1$;*
- *the point process $\mathcal{P}$ exhibits exponential decay of correlation as in Definition 3.1;*
- *the radius of stabilization is bounded, i.e., $\sup_{x \in \mathbb{R}^2} \sup_{\mathcal{X} \in \mathcal{N}} R^{\xi}(x, \mathcal{X}) < \infty$;*
- *the score function ξ satisfies a power growth condition, namely there exists $c \geq 1$ such that for all $r > 0$, $n \geq 1$ and $\mathcal{X} \in \mathcal{N}$,*

$$|\xi(x, \mathcal{X} \cap B_r(x))| \mathbb{1}\{\#(\mathcal{X} \cap B_r(x)) = n\} \leq c^n (1 \vee r^n). \tag{11}$$

Let f be a bounded measurable functions on W_1, and let

$$\mu_n^{\xi}(f) = \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n) f(n^{-1/d} x).$$

Assume that $\liminf_{n \to \infty} \mathsf{Var}(\mu_n^{\xi}(f)) n^{-\nu} > 0$ for some $\nu > 0$. Then,

$$\frac{\mu_n^{\xi}(f) - \mathbb{E}[\mu_n^{\xi}(f)]}{\sqrt{\mathsf{Var}(\mu_n^{\xi}(f))}}$$

converges in distribution to a standard normal random variable as $n \to \infty$.

Score functions in our context To be in the framework of Theorem 8.1, we need to express the quantity $\langle f, \mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle = \sum_{i\in I^{M,q}(\mathcal{P}_n)} f\big(B_i^M, D_i^M\big)$ in the form $\sum_{x\in\mathcal{P}_n} \xi(x,\mathcal{P}_n)$ for a suitable score function $\xi(x,\mathcal{P}_n)$.

In other words, we need to transform the indexing over features into an indexing over the points of the point process $\mathcal{P}_n$. We achieve this goal by assigning to each feature a point $x \in \mathcal{P}_n$ that either kills or gives birth to this feature, depending on whether $q = 0$ or $q = 1$.

First, the death of a cluster at time $r > 0$ is always caused by the merging of two points $x, x' \in \mathcal{P}_n$ at distance $2r$. Indeed, when the size of a component has a jump, this can only appear by attaching to another component. If $\mathcal{C}_r(x)$ dies by this merging, we say that x' *kills* $\mathcal{C}_r(x)$. This ensures that if two components both die when they merge, their deaths are caused by different points.

Similarly, if $q = 1$, then the birth of a hole at time $r > 0$ is caused by two points $x, x' \in \mathcal{P}_n$ at distance $2r$ whose connection creates a new hole. If only one M-bounded feature is born at time r, we choose the lexicographic minimum of x and x' and say that it *gives birth* to this hole. However, if a large hole is split into two M-bounded pieces, it can happen that two M-bounded pieces H, H' are born at the same time. In this situation, we assign one M-bounded piece to each of x and x'. Hence, we define the score functions as

$$\begin{aligned}\xi_0(x,\mathcal{P}_n)=\xi_{f,0}(x,\mathcal{P}_n) &= \sum_{i\in I^{M,0}(\mathcal{P}_n)} \mathbb{1}\{x \text{ kills the } i\text{th cluster}\} f(0, D_i^M),\\ \xi_1(x,\mathcal{P}_n)=\xi_{f,1}(x,\mathcal{P}_n) &= \sum_{i\in I^{M,1}(\mathcal{P}_n)} \mathbb{1}\{x \text{ gives birth to the } i\text{th hole}\} f(B_i^M, D_i^M).\end{aligned} \tag{12}$$

Notice that the main difficulty to deal with the score functions ξ_0, ξ_1 is that they do not only depend on points or pairs of points. Definition (12) translates the desired CLT for $\langle f, \mathsf{PD}^{M,q}(\mathcal{P}_n)\rangle$ into the framework of Theorem 8.1. Let $\mathcal{X} \in \mathcal{N}$. It remains to verify the conditions stated therein.

Proof of Theorem 3.2. As mentioned above, it is sufficient to check the assumptions of Theorem 8.1. The exponential decay of correlations is satisfied since it is one of our standing assumptions on the point process $\mathcal{P}$. Moreover, the radius of stabilization is bounded since we work with M-bounded features.

To check the power-growth condition, we note that in the worst case x can be responsible for the death of all other points of $\mathcal{X}$. Similarly, it can give birth to at most $\mathcal{X}(W_r(x)) - 1$ holes. Hence,

$$\xi_q(x, \mathcal{X} \cap W_r(x))\mathbb{1}\{\#(\mathcal{X} \cap W_r(x)) = n\} \le |f|_\infty (n-1) \le (1 + |f|_\infty)^n.$$

This proves that (11) holds.

It remains to verify the p-moment condition. We explain in detail how this is achieved if $q = 0$, noting that the case $q = 1$ can be deduced after minor modifications. If $x \in \mathcal{P}$ is responsible for the death of a component at time r, then there exists $x' \in \mathcal{P}_n$ at distance $2r$ from x. Since each ball grows for time at most r_{f}, we see that

$$|\xi_0(x, \mathcal{P}_n)| \le |f|_\infty \mathcal{P}_n(B_{2r_{\mathsf{f}}}(x)).$$

Leveraging stationarity and applying condition **(M)** concludes the proof. □

9. Proofs of Theorem 3.3 and Corollary 3.4

In the following, we assume $q = 1$, since the proofs for $q = 0$ are similar but easier. Hence, to simplify notation, we write $\beta_{b,d}(\mathcal{P}_n)$ for $\beta^{M,1}_{b,d}(\mathcal{P}_n)$.

Proof of Corollary 3.4. Note that if $(X(s))_{s\le r_\mathsf{f}}$ is a Gaussian process, then the process $(\int_0^r X(s)\mathrm{d}s)_{r\le r_\mathsf{f}}$ is also Gaussian. The plan is to start from Theorem 3.3 and then apply the continuous mapping theorem [30, Theorem 4.27]. To this end, we show that $\{\mathsf{APF}^{M,1}_r(\mathcal{P}_n)\}_{r\le r_\mathsf{f}}$ is a continuous functional of the persistent Betti numbers $\{\beta_{b,d}(\mathcal{P}_n)\}_{b,d\le r_\mathsf{f}}$. We assert that

$$\mathsf{APF}^{M,1}_r(\mathcal{P}_n) = \int_0^r \beta_{b,0}(\mathcal{P}_n)\mathrm{d}b + \int_0^{r_\mathsf{f}} \beta_{r,t}(\mathcal{P}_n)\mathrm{d}t - r\beta_{r,0}(\mathcal{P}_n). \tag{13}$$

The remainder of the proof proceeds in two steps. First, we verify identity (13). Second, we show that the right-hand side is continuous in β with respect to the Skorokhod topology.

To prove identity (13), linearity allows us to reduce the claim to the case where the persistence diagram consists of a single δ-measure at a point (B_0, D_0) for some $D_0 > B_0 > 0$. If $B_0 > r$, then both sides vanish. If $B_0 \le r$, then $\beta_{b,0} = \mathbb{1}\{b \ge B_0\}$ and $\beta_{r,t} = \mathbb{1}\{t \le D_0\}$, so that the right-hand side of (13) gives the asserted

$$(r - B_0) + D_0 - r = (D_0 - B_0).$$

Let $\beta \in D([0, r_\mathsf{f}]^2, \mathbb{R})$, where $D([0, r_\mathsf{f}]^2, \mathbb{R})$ is the Skorokhod space of càdlàg functions from $[0, r_\mathsf{f}]^2$ to $\mathbb{R}$. For any $r \ge 0$ put

$$\Phi_r(\beta) = \int_0^r \beta_{b,0}\mathrm{d}b + \int_0^{r_\mathsf{f}} \beta_{r,t}\mathrm{d}t - r\beta_{r,0}.$$

According to (13), it is sufficient to prove that the function $\Phi_r : D([0, r_\mathsf{f}]^2, \mathbb{R}) \to D([0, r_\mathsf{f}], \mathbb{R})$, $\beta \mapsto (\Phi_r(\beta))_{r\le r_\mathsf{f}}$ is continuous with respect to the Skorokhod topology. We prove this for the first integral. The arguments for the second are similar. Let $\beta' : [0, r_\mathsf{f}]^2 \to \mathbb{R}$ be càdlàg and $\lambda : [0, r_\mathsf{f}] \to [0, r_\mathsf{f}]$ be an increasing continuous bijection. Then,

$$\begin{aligned}\Big|\int_0^{\lambda(r)} \beta_{b,0}\mathrm{d}b - \int_0^r \beta'_{b,0}\mathrm{d}b\Big| &\le |\lambda(r) - r||\beta_{\cdot,0}|_\infty + \int_0^r |\beta_{b,0} - \beta'_{b,0}|\mathrm{d}b \\ &\le |\lambda(r) - r||\beta_{\cdot,0}|_\infty + \int_0^r |\beta_{\lambda(b),0} - \beta'_{b,0}|\mathrm{d}b + \int_0^r |\beta_{\lambda(b),0} - \beta_{b,0}|\mathrm{d}b.\end{aligned}$$

If β' approaches β in the Skorokhod metric, then by definition of this metric, we can choose λ such that the first two expressions become arbitrarily small. Moreover, since β itself is càdlàg, it follows that also the third expression tends to 0. □

The proof of Theorem 3.3 decomposes into two steps: lower and upper variance bounds and an upper bound on fourth-order cumulants. In what follows, we write

$$\beta(E,\mathcal{P}_n)=\beta_{b_+,d_+}(\mathcal{P}_n)+\beta_{b_-,d_-}(\mathcal{P}_n)-\beta_{b_+,d_-}(\mathcal{P}_n)-\beta_{b_-,d_+}(\mathcal{P}_n)$$

for the increment of $\beta_{b,d}$ in the block $E=(b_-,b_+]\times(d_-,d_+]$ with $b_-<b_+$ and $d_-<d_+$. Notice that this is minus the measure $\mathsf{PD}^{M,q}(\mathcal{P}_n)$ from (2) evaluated at the block E. Moreover, $\beta(E,\mathcal{P}_n)$ is the number of holes with birth time before b_- and death time between d_- and d_+ minus the number of holes with birth time before b_+ and death time between d_- and d_+. Following [5], two blocks $E,E'\subset[0,r_{\mathsf{f}}]^2$ are *neighboring* if they share a common side.

Proposition 9.1 (Variance lower bound)**.** *Let $\mathcal{P}$ be a conditionally m-dependent point process with exponential decay of correlations and satisfies condition* **(AC)**. *Moreover, let $a_1,\dots,a_k\neq 0$ and $E_1,\dots,E_k\subset[0,r_{\mathsf{f}}]^2$ be pairwise disjoint blocks such that each E_i contains some $(b,d)\in[0,r_{\mathsf{f}}]^2$ with $d>b$. Then,*

$$\liminf_{n\to\infty}\frac{1}{n}\mathsf{Var}\Big(\sum_{i\le k}a_i\beta(E_i,\mathcal{P}_n)\Big)>0.$$

Proposition 9.2 (Variance upper bound)**.** *Let $\mathcal{P}$ be a conditionally m-dependent point process with exponential decay of correlations and satisfies condition* **(M)**. *Then, there exist $n_0\ge 1$ and $\varepsilon_0,C_0>0$ such that*

$$\frac{1}{n}\mathsf{Var}\big(\beta(E,\mathcal{P}_n)\big)\le C_0|E|^{1/2+\varepsilon_0}$$

holds for all $n\ge n_0$ and blocks $E\subset[0,r_{\mathsf{f}}]^2$.

Now, the *kth cumulant* c^k of $k\ge 1$ real random variables $Y_1,\dots,Y_k$ equals

$$c^k(Y_1,\dots,Y_k)=\sum_{\{T_1,\dots,T_p\}\preceq\{1,\dots,k\}}(-1)^{p-1}(p-1)!\,\mathbb{E}\Big[\prod_{i\in T_1}Y_i\Big]\cdots\mathbb{E}\Big[\prod_{i\in T_p}Y_i\Big],$$

provided that all appearing moments are well-defined [38, Proposition 3.2.1]. Here, the sum ranges over all partitions $\{T_1,\dots,T_p\}$ of the set $\{1,\dots,k\}$.

Proposition 9.3 (Cumulant bound)**.** *Let $\mathcal{P}$ be a conditionally m-dependent point process with exponential decay of correlations satisfying conditions* **(AC)** *and* **(M)**. *Then, there exist $n_0'\ge 1$ and $\varepsilon_0',C_0'>0$ such that*

$$\frac{1}{n}c^4\big(\beta(E,\mathcal{P}_n),\beta(E,\mathcal{P}_n),\beta(E',\mathcal{P}_n),\beta(E',\mathcal{P}_n)\big)\le C_0'|E|^{1/2+\varepsilon_0'}|E'|^{1/2+\varepsilon_0'}$$

holds for all $n\ge n_0'$ and neighboring blocks $E,E'\subset[0,r_{\mathsf{f}}]^2$.

We postpone the proofs of Propositions 9.1–9.3 to Sections 9.1–9.3, respectively. To deduce Theorem 3.3 from these two central auxiliary results, we write

$$\overline{\beta}_{b,d}(\mathcal{P}_n)=\beta_{b,d}(\mathcal{P}_n)-\mathbb{E}[\beta_{b,d}(\mathcal{P}_n)]$$

for the centered persistent Betti numbers.

Proof of Theorem 3.3. Let $a'_1, \dots, a'_{k'} \neq 0$ and $(b_1, d_1), \dots, (b_{k'}, d_{k'}) \in [0, r_{\mathsf{f}}]^2$ be pairwise distinct, and put

$$X_n = \sum_{i \leq k'} a'_i \beta_{b'_i, d'_i}(\mathcal{P}_n).$$

Then, after suitable regrouping of terms, we can express X_n in the form

$$X_n = \sum_{i \leq k} a_i \beta(E_i, \mathcal{P}_n).$$

as in Proposition 9.1. Now, combining Proposition 9.2 with Theorem 3.2 and the variance asymptotics [9, Theorem 1.12] shows that the centered and rescaled random variable $n^{-1/2}(X_n - \mathbb{E}[X_n])$ converges in distribution to a Gaussian. Hence, the Cramér-Wold device yields convergence of the finite-dimensional distributions of $n^{-1/2}\overline{\beta}_{b,d}(\mathcal{P}_n)$.

Next, [38, Proposition 3.2.1] gives the general cumulant identity

$$\begin{aligned}\mathbb{E}[X^2Y^2] &= c^4(X, X, Y, Y) + \mathsf{Var}(X)\mathsf{Var}(Y) + 2\mathsf{Cov}(X, Y)^2 \\ &\leq c^4(X, X, Y, Y) + 3\mathsf{Var}(X)\mathsf{Var}(Y)\end{aligned}$$

for centered random variables X, Y. Hence, by Propositions 9.2 and 9.3,

$$\mathbb{E}\big[n^{-2}\overline{\beta}(E, \mathcal{P}_n)^2\overline{\beta}(E', \mathcal{P}_n)^2\big] \leq (C'_0/n + 3C_0^2)|E|^{1/2+\varepsilon''_0}|E'|^{1/2+\varepsilon''_0},$$

for some $\varepsilon''_0 > 0$. In particular, the process $\big\{n^{-1/2}\overline{\beta}_{b,d}(\mathcal{P}_n)\big\}_{b,d \leq r_{\mathsf{f}}}$ is tight in Skorokhod topology [25, Lemma 3]. In this context, we note that condition (8.4) of [25, Lemma 3] follows from the variance upper bound derived in Proposition 9.2 and that similar as in (2.18) [23], we have replaced the equality in (8.5) of [25, Lemma 3] by an inequality. Combining this property with the convergence of finite-dimensional distributions yields the asserted weak convergence. □

9.1. Proof of Proposition 9.1

To show the variance lower bound, we adapt a conditioning argument that has already been successfully applied in the setting of Gibbsian point processes [29, 40]. More precisely, we subdivide the window W_n into blocks of a fixed size and use the law of conditional variance to obtain a lower bound in the order of the number of blocks.

Associate with the jth feature H_j in $\mathsf{PD}^{M,1}(\mathcal{P}_n)$ a center point $y_j \in W_n$, for instance by taking the point $p(H_j)$ as defined in Section 2.2. Then,

$$\nu_n = \sum_{i \leq k} a_i \sum_{j \in I^{M,1}(\mathcal{P}_n)} \mathbb{1}\{(B_j^M, D_j^M) \in E_i\}\delta_{y_j}$$

defines a signed measure of total mass $\nu_n(W_n) = -\sum_{i \leq k} a_i \beta(E_i, \mathcal{P}_n)$.

In the vein of [40], the key towards proving a lower bound on the variance is the following non-degeneracy property, where r_{AC} is introduced in Section 3.

Lemma 9.4 (Non-degeneracy)**.** *It holds that*

$$\inf_{n \geq t \geq r_{\mathsf{AC}}^2} \mathbb{E}\big[\mathsf{Var}\big(\nu_n(W_t) | \, \Lambda, \mathcal{P} \setminus W_{r_{\mathsf{AC}}^2}\big)\big] > 0.$$

Before proving Lemma 9.4, we explain how it implies Proposition 9.1. In essence, the proof follows along the lines of [40, Lemma 4.3]. Nevertheless, since the details of the conditioning argument differ a bit from the corresponding picture for Gibbs processes, we explain how to adapt the main steps from [40, Lemma 4.3] in the present setting.

Proof of Proposition 9.1. The idea of proof is to consider a family of well-separated blocks in W_n. Then, we leverage the conditional m-dependence of the point process and the M-boundedness of the features to decompose the variance of their contributions as the sum of the variances. More precisely, we apply the assumption of conditional m-dependence with the conditioning set

$$A'' = \mathbb{R}^2 \setminus \bigcup_{z \in \mathbb{Z}^2} (6\rho z + W_{\rho^2})$$

chosen as the complement of the union of well-separated blocks of side length $\rho = m \vee r_{\mathsf{AC}}$. Then, the law of total variance yields the lower bound

$$\mathsf{Var}(\nu_n(W_n)) \geq \mathbb{E}\big[\mathsf{Var}\big(\nu_n(W_n) \,|\, \Lambda, \mathcal{P} \cap A''\big)\big].$$

Moreover, since $\rho > M$ the statistics $\nu_n((A'')^-)$ in the smaller domain

$$(A'')^- = \mathbb{R}^2 \setminus \bigcup_{z \in \mathbb{Z}^2} (6\rho z + W_{9\rho^2})$$

is measurable with respect to $\mathcal{P} \cap A''$. We obtain that

$$\mathbb{E}\big[\mathsf{Var}\big(\nu_n(W_n) | \Lambda, \mathcal{P} \cap A''\big)\big] = \mathbb{E}\big[\mathsf{Var}\big(\nu_n(\mathbb{R}^2 \setminus (A'')^-) | \Lambda, \mathcal{P} \cap A''\big)\big]$$

because $\nu_n((A'')^-)$ is $\mathcal{P} \cap A''$ measurable. Thanks to the conditional m-dependence, we have

$$\begin{aligned}\mathbb{E}\big[\mathsf{Var}\big(\nu_n(W_n) | \Lambda, \mathcal{P} \cap A''\big)\big] &= \sum_{\substack{z \in \mathbb{Z}^2 \\ 6\rho z + W_{9\rho^2} \subset W_n}} \mathbb{E}\big[\mathsf{Var}\big(\nu_n(6\rho z + W_{9\rho^2}) \,|\, \Lambda, \mathcal{P} \cap A''\big)\big] \\ &\geq \sum_{\substack{z \in \mathbb{Z}^2 \\ 6\rho z + W_{9\rho^2} \subset W_n}} \mathbb{E}\big[\mathsf{Var}\big(\nu_n(6\rho z + W_{9\rho^2}) \,|\, \Lambda, \mathcal{P} \setminus (6\rho z + W_{\rho^2})\big)\big].\end{aligned}$$

Now, the number of 6ρ-blocks contained in W_n is of order n, and we conclude by noting that Lemma 9.4 and $\rho > r_{\mathsf{AC}}$ imply that each of the contributions is bounded away from 0. □

To verify non-degeneracy, we rely on the techniques introduced in [40]. In particular, we make use of [40, Lemma 2.3], which we restate below to render the presentation self-contained.

Lemma 9.5. *Let Y be a real random variable and A_1, A_2 be Borel sets of $\mathbb{R}$. Then,*

$$\mathsf{Var}(Y) \geq \frac{1}{4} \min_{i\in\{1,2\}} \mathbb{P}(Y \in A_i) \inf_{x_1\in A_1, x_2\in A_2} |x_1 - x_2|^2.$$

Proof of Lemma 9.4. Write

$$F_1 = \{\mathcal{P} \cap W_{r^2_{\mathsf{AC}}/9} = \emptyset\} \quad \text{and} \quad F' = \{\mathcal{P} \cap (W_{r^2_{\mathsf{AC}}} \setminus W_{r^2_{\mathsf{AC}}/9}) = \emptyset\}$$

for the events that there are no points in $W_{r^2_{\mathsf{AC}}/9}$ and $W_{r^2_{\mathsf{AC}}} \setminus W_{r^2_{\mathsf{AC}}/9}$, respectively. Next, let

$$F_2 = \{-\beta(E_1, \mathcal{P}_{r^2_{\mathsf{AC}}/9}) = 1\} \cap \{\beta(E_2, \mathcal{P}_{r^2_{\mathsf{AC}}/9}) = \cdots = \beta(E_k, \mathcal{P}_{r^2_{\mathsf{AC}}/9}) = 0\}$$

denote the event that all but the first of the considered persistent Betti numbers vanish. Now, let $I_0(\mathcal{P}_n)$ denote the indices of all features entirely contained in $\mathbb{R}^2 \setminus W_{r^2_{\mathsf{AC}}}$ and put

$$Y = \sum_{i\leq k} a_i \#\{j \in I^{M,1}(\mathcal{P}_n) \setminus I_0(\mathcal{P}_n) : (B^M_j, D^M_j) \in E_i \text{ and } y_j \in W_t\}.$$

Then, assuming $a_1 > 0$, by Lemma 9.5 with $A_1 = \{0\}$ and $A_2 = [a_1, \infty)$,

$$\begin{aligned}\mathbb{E}\big[\mathsf{Var}\big(\nu_n(W_t)\,|\,\Lambda, \mathcal{P} \setminus W_{r^2_{\mathsf{AC}}}\big)\big] &= \mathbb{E}\big[\mathsf{Var}\big(Y\,|\,\Lambda, \mathcal{P} \setminus W_{r^2_{\mathsf{AC}}}\big)\big] \\ &\geq \frac{a_1^2}{4}\mathbb{E}\big[\min_{i\in\{1,2\}} \mathbb{P}(F' \cap F_i | \Lambda, \mathcal{P} \setminus W_{r^2_{\mathsf{AC}}})\big],\end{aligned}$$

and it remains to show that the right-hand side is non-zero.

Since $E_1, \ldots, E_k$ are pairwise disjoint and contain points above the diagonal, [26, Example 1.8] shows that under the homogeneous Poisson point process the event $F' \cap F_2$ has positive probability. Also $F' \cap F_1$ is of positive probability. Hence, an application of condition **(AC)** concludes the proof. □

9.2. Proof of Proposition 9.2

For a block $E = (b_-, b_+] \times (d_-, d_+] \subset [0, r_{\mathsf{f}}]^2$, we let ξ_E denote the score function associated with $\beta(E, \mathcal{P}_n)$. That is,

$$\xi_E(x, \mathcal{P}_n) = \#\{(B^M_i, D^M_i) \in E : x \text{ gives birth to the } i\text{th hole}\}$$

is the number of holes born by x with birth and death times in E. Note that if x gives birth to the ith hole, then it gets in contact with another point at time $B^M_i \in (b_-, b_+]$. In particular, $\mathcal{P}$ contains a point in the annulus $A_{2b_-, 2b_+}(x) = B_{2b_+}(x) \setminus B_{2b_-}(x)$.

Moreover, if the ith hole dies at time $D^M_i \in (d_-, d_+]$, then a previously vacant component is covered completely, which is caused by three disks centered at points in $\mathcal{P}$ meeting at a single point in the plane. The three center points

of the disks must form a triangle with no obtuse angle. Otherwise, two of the disks would meet for the first time in the interior of the third and hence no connected component in the background was covered by the merging. This could be interpreted as a feature that is born and dies at the same time, but we chose to exclude such features in our definition of 1-features.

Henceforth, let $B_d^{\pm}(x,y) \subset \mathbb{R}^2$ denote the two disks of radius $d > 0$ whose boundary passes through $x, y \in \mathbb{R}^2$. If $|x-y|/2 > d$, we let $B_d^{\pm}(x,y)$ be empty. The points in $B_d^+(x,y) \cup B_d^-(x,y)$ are exactly the points z such that the time when the boundaries of the three disks around x, y, and z meet in one point is at most d. For $d_+ > d_- \geq 0$, we let

$$D_{d_-,d_+}(x,y) = \big(B_{d_+}^+(x,y) \cup B_{d_+}^-(x,y)\big) \backslash \big(B_{d_- \vee a}^+(x,y) \cup B_{d_- \vee a}^-(x,y)\big),$$

where $a = |x-y|/2$. This set consists of all points z such that the boundaries of the three disks around x, y and z meet at time r with $d_- < r \leq d_+$. Some $z \in D_{d_-,d_+}(x,y)$ may still form a triangle having an obtuse angle with x and y, that is, the disks around x, y, and z already met earlier in an interior point of one of the disks. However, all z that can cause the death of a hole in E together with x and y must be contained in $D_{d_-,d_+}(x,y)$.

Now,

$$\xi_E(x, \mathcal{P}_n) \leq \mathcal{P}(B_M(x)) \mathbb{1}_{E_x}, \tag{14}$$

where E_x denotes the event that for some $\mathcal{P}' \subset \mathcal{P}$ with $x \in \mathcal{P}'$ the event $E_{x,\mathsf{b}}(\mathcal{P}') \cap E_{x,\mathsf{d}}(\mathcal{P}')$ occurs, where

$$\begin{aligned}
E_{x,\mathsf{b}}(\mathcal{P}') = \big\{&x \text{ creates an } M\text{-bounded hole in } \{U_r(\mathcal{P}')\}_{r \geq 0} \text{ with birth and} \\
&\text{death time in } E \text{ by connecting to some } x_1 \in \mathcal{P}' \cap A_{2b_-,2b_+}(x)\big\} \\
E_{x,\mathsf{d}}(\mathcal{P}') = \big\{&\exists y_1, y_2 \in \mathcal{P}' \cap B_M(x) \text{ such that } y_1, y_2, y_3 \text{ kill an } M\text{-bounded hole} \\
&\text{in } \{U_r(\mathcal{P}')\}_{r \geq 0} \text{ with birth and death time in } E \text{ for some} \\
&y_3 \in \mathcal{P}' \cap D_{d_-,d_+}(y_1, y_2)\big\}.
\end{aligned}$$

Here, we say that $y_1, y_2, y_3 \in \mathcal{P}'$ kill the hole H if the disks around the points meet for the first time at $p(H)$. In particular, any three points can kill at most one hole.

Similarly, for a block $E' = (b'_-, b'_+] \times (d'_-, d'_+] \subset [0, r_{\mathsf{f}}]^2$,

$$\xi_E(x, \mathcal{P}_n) \xi_{E'}(x', \mathcal{P}_n) \leq \mathcal{P}(B_M(x)) \mathcal{P}(B_M(x')) \mathbb{1}_{E''_{x,x'}},$$

where we let $E''_{x,x'}$ denote the event that for some $\mathcal{P}' \subset \mathcal{P}$ with $x, x' \in \mathcal{P}'$ the event

$$E_{x,\mathsf{b}}(\mathcal{P}') \cap E'_{x',\mathsf{b}}(\mathcal{P}') \cap E_{x,\mathsf{d}}(\mathcal{P}') \cap E'_{x',\mathsf{d}}(\mathcal{P}')$$

occurs.

Using this notation, the proof of the variance upper bound is now based on the following pivotal geometric moment bound. In the following, $\mathbb{P}_{\boldsymbol{x}}$ denotes the unreduced Palm measure characterized via

We recall from (3) that $\rho^{(p)}$ denotes the pth factorial moment density. In the following, we adhere to the convention $\int_{B^0} f(x) \mathrm{d}z = f(x)$.

Lemma 9.6 (Moment bound)**.** *Let $\mathcal{P}$ be a stationary point process having fast decay of correlations and satisfying condition* **(M)**. *Let $p \geq 0$ and $K_0 > 0$. Then, there exist $\varepsilon > 0$ and $C_{\mathsf{g}} > 0$ such that for all $n > 0$ and any ball $B \subset \mathbb{R}^2$ of radius $K > K_0$,*

1. $$\frac{1}{|B|^{p+1}} \int_{B^p} \mathbb{P}_{o,\boldsymbol{z}}(E_o)\rho^{(p+1)}(o,\boldsymbol{z})\mathrm{d}\boldsymbol{z} \leq C_{\mathsf{g}}|E|^{1/2+\varepsilon}$$
holds for all blocks $E \subset [0, r_{\mathsf{f}}]^2$.
2. $$\frac{1}{|B|^{p+1}} \int_{B^p} \mathbb{P}_{o,\boldsymbol{z}}(E''_{o,o})\rho^{(p+1)}(o,\boldsymbol{z})\mathrm{d}\boldsymbol{z} \leq C_{\mathsf{g}}|E|^{1/2+\varepsilon}|E'|^{1/2+\varepsilon}$$
holds for all neighboring blocks $E, E' \subset [0, r_{\mathsf{f}}]^2$, and
3. $$\frac{1}{|B|^{p+2}} \int_{B^{p+1}} \mathbb{P}_{o,z',\boldsymbol{z}}(E''_{o,z'})\rho^{(p+2)}(o,z',\boldsymbol{z})\mathrm{d}(z',\boldsymbol{z}) \leq C_{\mathsf{g}}|E|^{1/2+\varepsilon}|E'|^{1/2+\varepsilon}$$
holds for all neighboring blocks $E, E' \subset [0, r_{\mathsf{f}}]^2$.

The proof of Lemma 9.6 relies on a delicate geometric analysis that we defer to Section 9.4. We now prove Proposition 9.2. As in [9, Equation (1.6)], for $\boldsymbol{x} = (x_1, \ldots, x_p) \in \mathbb{R}^{2p}$ and $k_1, \ldots, k_p \geq 0$, we introduce the *mixed ξ_E-moments*

$$m_n^{(k_1,\ldots,k_p)}(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{x}}[\xi_E(x_1, \mathcal{P}_n)^{k_1} \cdots \xi_E(x_p, \mathcal{P}_n)^{k_p}]\rho^{(p)}(\boldsymbol{x}). \tag{15}$$

In the rest of the manuscript, we freely use that exponential decay of correlations implies boundedness of the factorial moment densities [9, Inequality (1.11)].

Proof of Proposition 9.2. To lighten notation, we write ξ instead of ξ_E. To give the paper more pleasant to read, we have not attempted to optimize the exponents occurring in the course of this proof. Proceeding as in [9, Equation (4.1)], the refined Campbell-Mecke formula [9, Equation (1.9)] gives that $\mathsf{Var}\big(\beta(E, \mathcal{P}_n)\big)$ equals

$$\int_{W_n} m_n^{(2)}(x)\mathrm{d}x + \int_{W_n \times W_n} \big(m_n^{(1,1)}(x,y) - m_n^{(1)}(x)m_n^{(1)}(y)\big)\mathrm{d}(x,y). \tag{16}$$

We derive bounds for the two summands separately.

By stationarity, (14) and Hölder's inequality, the first expression is at most

$$n\mathbb{E}_o[\mathcal{P}(B_M(o))^{16}]^{1/8}\mathbb{P}_o(E_o)^{7/8}\rho = n\big(\mathbb{E}_o[\mathcal{P}(B_M(o))^{16}]\rho\big)^{1/8}\big(\mathbb{P}_o(E_o)\rho\big)^{7/8}. \tag{17}$$

Hence, Lemma 9.6(1) with $p = 0$ yields the asserted upper bound.

To deal with the double integral in (16), we recall that ξ is a local score function and that $\mathcal{P}$ exhibits exponential decay of correlations. Hence, as in [9, Equation (3.26)], the factorial moment measure expansion shows that

$$\big|m_n^{(1,1)}(x,y) - m_n^{(1)}(x)m_n^{(1)}(y)\big| \leq c\phi(|x-y|/2)$$

for some $c > 0$. In particular, choosing a cut-off $K = |E|^{-1/128}$, we see that

$$\sup_{x \in W_n} \int_{W_n \setminus B_K(x)} \left| m_n^{(1,1)}(x,y) - m_n^{(1)}(x) m_n^{(1)}(y) \right| \mathrm{d}y \leq C|E|$$

holds for a suitable $C > 0$ and it suffices to derive an upper bound for

$$\int_{B_K(x)} m_n^{(1,1)}(x,y) + m_n^{(1)}(x) m_n^{(1)}(y) \mathrm{d}y.$$

For the second summand, we can argue similarly as in (17), so that it remains to bound the integral involving $m_n^{(1,1)}(x,y)$. Here, we set $z = y - x$, note that $\rho^{(2)}(x,y) = \rho^{(2)}(o,z)$ and combine (14) with Hölder's inequality to arrive at

$$m_n^{(1,1)}(x,y) \leq \mathbb{E}_{o,z}[\mathcal{P}(B_M(o))^{16}]^{1/16} \mathbb{E}_{o,z}[\mathcal{P}(B_M(z))^{16}]^{1/16} \mathbb{P}_{o,z}(E_o)^{7/8} \rho^{(2)}(o,z).$$

We bound $\mathbb{E}_{o,z}[\mathcal{P}(B_M(z))^{16}]$ thanks to condition **(M)**. Finally, by Jensen's inequality applied to the uniform distribution on $B_K(o)$,

$$\frac{1}{|B_K(o)|} \int_{B_K(o)} (\mathbb{P}_{o,z}(E_o) \rho^{(2)}(o,z))^{7/8} \mathrm{d}z \leq \Big(\frac{1}{|B_K(o)|} \int_{B_K(o)} \mathbb{P}_{o,z}(E_o) \rho^{(2)}(o,z) \mathrm{d}z \Big)^{7/8}, \tag{18}$$

so that applying Lemma 9.6(1) with $p = 1$ shows that the right-hand side is of order at most $(|E|^{3/4}|B_K(o)|)^{7/8} = |E|^{7 \cdot 47/512}$, thereby concluding the proof. □

9.3. *Proof of Proposition 9.3*

To prove Proposition 9.3, we take up the idea suggested in [25, Theorem 8] and [11, Theorem 8.1] and express c^4 in terms of cumulant measures induced by the functional of interest. A slight technical nuisance in the present setting comes from dealing with a product of two different functionals – one associated with the block E and the other with E' – whereas the semi-cluster measure machinery from [9, Section 4.3] relies on a single score function. However, this artificial difficulty can be overcome by formally attaching $\{1,2\}$-valued marks to $\mathcal{P}_n$. Taking up the notation from [20], we let $\breve{\mathbb{R}}^2 = \mathbb{R}^2 \times \{1,2\}$ and $\breve{\mathcal{P}}_n = \mathcal{P}_n \times \{1,2\}$ denote the correspondingly marked space and point process. Writing $E'' = (E, E')$, we define an augmented score function $\xi_{E''}$, where points with mark 1 are evaluated with the first score function and points with mark 2 are evaluated with the second score function. In other words,

$$\xi_{E''}((x,\tau), \breve{\mathcal{P}}_n) = \begin{cases} \xi_E(x, \mathcal{P}_n) & \text{if } \tau = 1, \\ \xi_{E'}(x, \mathcal{P}_n) & \text{if } \tau = 2. \end{cases}$$

We take the concise proof of Proposition 9.2 as a blueprint for the strategy of the more involved setting laid out in Proposition 9.3. In particular, we need

to address two main steps: bounds for mixed moments and a reduction of the integral to the diagonal.

In order to reduce to the diagonal, we decompose the cumulant measure into semi-cluster measures as in [4, Section 5.1] and [20, Section 3.2]. For the convenience of the reader, we reproduce the basic definitions. First, the *kth moment measure* $M^k(\mu_n)$ is given as

$$\langle \boldsymbol{f}, M^k(\mu_n)\rangle = \int \boldsymbol{f}(\breve{\boldsymbol{x}})M^k(\mu_n)(\mathrm{d}\breve{\boldsymbol{x}}) = \mathbb{E}[\langle f_1, \mu_n\rangle \cdots \langle f_k, \mu_n\rangle],$$

where $\boldsymbol{f} = f_1 \otimes \cdots \otimes f_k$ is non-negative and measurable with each f_i defined on $\breve{\mathbb{R}}^2$, and

$$\mu_n = \mu_{E'',n} = n^{-1} \sum_{\breve{x}\in\breve{\mathcal{P}}_n} \xi_{E''}(\breve{x}, \breve{\mathcal{P}}_n)\delta_{\breve{x}}$$

denotes the empirical measure associated with $\xi_{E''}$ and $\breve{\mathcal{P}}_n$. In terms of mixed ξ-moments, with $\breve{\boldsymbol{x}}_{T_i}$ the projection of $\breve{\boldsymbol{x}}$ to the coordinates in T_i, we write

$$\mathrm{d}M^k = \sum_{\{T_1,\ldots,T_p\}\preceq\{1,\ldots,k\}} m_n^{(T_1,\ldots,T_p)}\mathrm{d}\breve{\boldsymbol{x}}_{T_1}\cdots\mathrm{d}\breve{\boldsymbol{x}}_{T_p}, \tag{19}$$

where $\mathrm{d}\breve{\boldsymbol{x}}_{T_i}$ are the singular differentials determined via

$$\int_{\breve{\mathbb{R}}^{2|T_i|}} f(\breve{\boldsymbol{x}}_{T_i})\mathrm{d}\breve{\boldsymbol{x}}_{T_i} = \int_{\breve{\mathbb{R}}^2} f(\breve{x},\ldots,\breve{x})\mathrm{d}\breve{x}$$

where $f : \breve{\mathbb{R}}^{2|T_i|} \to [0,\infty)$ is any non-negative measurable function [20, Section 3.1]. As in (15), for $T_1,\ldots,T_p \preceq \{1,\ldots,k\}$, the mixed $\xi_{E''}$-moments are given as

$$m_n^{(T_1,\ldots,T_p)}(\breve{\boldsymbol{x}}) = \mathbb{E}_{\boldsymbol{x}}[\xi_{E''}(\breve{\boldsymbol{x}}_{T_1},\mathcal{P}_n)^{|T_1|}\cdots\xi_{E''}(\breve{\boldsymbol{x}}_{T_p},\mathcal{P}_n)^{|T_p|}]\rho^{(p)}(\boldsymbol{x}),$$

for every $\breve{\boldsymbol{x}} = ((x_1,\tau_1),\ldots,(x_k,\tau_k)) \in \breve{\mathbb{R}}^{2k}$.

Similarly, the *kth cumulant measure* $c_n^k = c^k(\mu_n)$ equals

$$\langle \boldsymbol{f}, c_n^k\rangle = c^k(\langle f_1,\mu_n\rangle,\ldots,\langle f_k,\mu_n\rangle),$$

so that

$$c_n^k = \sum_{\{T_1,\ldots,T_p\}\preceq\{1,\ldots,k\}} (-1)^{p-1}(p-1)!\, M^{T_1}\cdots M^{T_p}, \tag{20}$$

where

$$\mathrm{d}M^{T_i} = \sum_{\{T_1',\ldots,T_{p'}'\}\preceq T_i} m_n^{(T_1',\ldots,T_{p'}')}\mathrm{d}\breve{\boldsymbol{x}}_{T_1'}\cdots\mathrm{d}\breve{\boldsymbol{x}}_{T_{p'}'}$$

denotes the moment measure with coordinates in T_i.

Next, the space $\breve{W}_n^4$ decomposes into a union of subsets according to which coordinate is most distant from the diagonal [20, Lemma 3.1]. More precisely, write

$$D(\breve{\boldsymbol{x}}) = \max_{\{S,T\} \preceq \{1,2,3,4\}} \mathsf{dist}(\breve{\boldsymbol{x}}_S, \breve{\boldsymbol{x}}_T)$$

for the maximal separation of $\breve{\boldsymbol{x}}_S$ and $\breve{\boldsymbol{x}}_T$, where $\mathsf{dist}(\breve{\boldsymbol{x}}_S, \breve{\boldsymbol{x}}_T) = \mathsf{dist}(\boldsymbol{x}_S, \boldsymbol{x}_T)$. Then, put

$$\sigma(S,T) = \big\{\breve{\boldsymbol{x}} = (\breve{\boldsymbol{x}}_S, \breve{\boldsymbol{x}}_T) \in \breve{W}_n^4 : \, D(\breve{\boldsymbol{x}}) = \mathsf{dist}(\breve{\boldsymbol{x}}_S, \breve{\boldsymbol{x}}_T)\big\} \setminus \Delta.$$

Here, the marks are ignored for the diagonal $\Delta \subset \breve{W}_n^4$. We also put $W_n^{(1,2)} = (W_n \times \{1\})^2 \times (W_n \times \{2\})^2$.

Lemma 9.7 (Off-diagonal bounds)**.** *Let S, T denote a non-trivial partition of $\{1,2,3,4\}$. Then, there exist $n_{S,T} \geq 1$ and $\varepsilon_{S,T}, C_{S,T} > 0$ such that*

$$\frac{1}{n}\big|c_n^4(\sigma(S,T) \cap W_n^{(1,2)})\big| \leq C_{S,T}|E|^{1/2+\varepsilon_{S,T}}|E'|^{1/2+\varepsilon_{S,T}}$$

holds for all $n \geq n_{S,T}$ and neighboring blocks $E, E' \subset [0, r_{\mathsf{f}}]^2$.

Before proving Lemma 9.7, we elucidate how to deduce Proposition 9.3.

Proof of Proposition 9.3. First, integration over the cumulant measure decomposes into a diagonal and an off-diagonal part [20, Equation (3.28)]. That is,

$$\frac{1}{n}\langle \boldsymbol{f}, c_n^4 \rangle = \frac{1}{n}\int_{\Delta} \boldsymbol{f} \mathrm{d}c_n^4 + \frac{1}{n}\sum_{S,T}\int_{\sigma(S,T)} \boldsymbol{f} \mathrm{d}c_n^4.$$

where $\boldsymbol{f} = \mathbb{1}_{W_n^{(1,2)}}$ is the indicator function of the domain $W_n^{(1,2)}$ and the sum is over all non-trivial partitions S, T. By Lemma 9.7, the off-diagonal contributions in this decomposition are bounded above by $\sum_{S,T} C_{S,T}|E|^{1/2+\varepsilon_{S,T}}|E'|^{1/2+\varepsilon_{S,T}}$.

Next, when integrating over the diagonal, we leverage that in the decomposition (20), only $p = 1$ contributes [20, Lemma 3.1]. Hence,

$$\begin{aligned}\int_{\Delta} \boldsymbol{f} \mathrm{d}c_n^4 &= \int_{W_n} \mathbb{E}_x[\xi_E(x, \mathcal{P}_n)^2 \xi_{E'}(x, \mathcal{P}_n)^2] \rho \mathrm{d}x \\ &\leq n\mathbb{E}_o[\mathcal{P}(B_M)^{2/\varepsilon}]^{\varepsilon} \mathbb{P}_o(E''_{o,o})^{1-\varepsilon},\end{aligned}$$

so that applying Lemma 9.6(2) with $p = 1$ and noting the convention preceding that result concludes the proof. □

To prove Lemma 9.7, we decompose the cumulant measures into *semi-cluster measures* [4, Lemma 5.1]. More precisely, as in [4, 20], any two disjoint non-empty subsets $S', T' \preceq \{1,2,3,4\}$, induce a *cluster measure*

$$U^{S',T'}(A \times B) = M^{S' \cup T'}(A \times B) - M^{S'}(A) M^{T'}(B).$$

Now, c_n^4 decomposes into semi-cluster measures

$$c_n^4 = \sum_{\{S',T',T_1,\dots,T_p\} \preceq \{1,2,3,4\}} a_{S',T',T_1,\dots,T_p} U^{S',T'} M^{T_1} \cdots M^{T_p}, \tag{21}$$

for coefficients $a_{S',T',T_1,\dots,T_p} \in \mathbb{R}$, where the sum runs over all partitions such that S' and T' are non-empty subsets of S and T, respectively [20, Lemma 3.2].

Equipped with these ingredients, we now prove Lemma 9.7. Since the basic structure of the proof parallels that of Proposition 9.2, we only provide details for the steps that are substantially different.

Proof of Lemma 9.7. Putting $D_K = \{\breve{\boldsymbol{x}} \in W_n^{(1,2)} \cap \sigma(S,T) : D(\breve{\boldsymbol{x}}) > K\}$ for $K \geq 1$, we first derive an upper bound for

$$\Big| \int_{D_K} \mathrm{d}U^{S',T'} \mathrm{d}M^{T_1} \cdots \mathrm{d}M^{T_p} \Big| = \Big| \int_{D_K} (\mathrm{d}M^{S' \cup T'} - \mathrm{d}M^{S'} \mathrm{d}M^{T'}) \mathrm{d}M^{T_1} \cdots \mathrm{d}M^{T_p} \Big|.$$

For this purpose, we decompose the moment measures $\mathrm{d}M^{S' \cup T'}$, $\mathrm{d}M^{S'}$ and $\mathrm{d}M^{T'}$ according to (19). Hence, we need bounds for the absolute value of differences of mixed ξ-moments of the form

$$\begin{aligned} \Big| & m_n^{(S_1'',\dots,S_{p''}'',T_1'',\dots,T_{r''}'')}(\breve{\boldsymbol{x}}_{S_1''}, \dots, \breve{\boldsymbol{x}}_{S_{p''}''}, \breve{\boldsymbol{x}}_{T_1''}, \dots, \breve{\boldsymbol{x}}_{T_{r''}''}) \\ & - m_n^{(S_1'',\dots,S_{p''}'')}(\breve{\boldsymbol{x}}_{S_1''}, \dots, \breve{\boldsymbol{x}}_{S_{p''}''}) m_n^{(T_1'',\dots,T_{r''}'')}(\breve{\boldsymbol{x}}_{T_1''}, \dots, \breve{\boldsymbol{x}}_{T_{r''}''}) \Big|, \end{aligned} \tag{22}$$

where $\{S_1'', \dots, S_{p''}''\}$ and $\{T_1'', \dots, T_{r''}''\}$ are partitions of S' and T', respectively. Since we are working on the set $\sigma(S,T)$, as in the proof of Proposition 9.2, the fast decay of ξ-correlations bounds (22) by $c\phi(D(\breve{\boldsymbol{x}}_{S' \cup T'})/2)$ for a suitable $c > 0$.

Next, as in [20, Section 3.1] the singular differentials occurring in the expansion (19) of the moment measure M^k can be grouped into a single object. More precisely, we write $\tilde{\mathrm{d}}\breve{\boldsymbol{x}}$ for the measure that equals $\mathrm{d}\breve{\boldsymbol{x}}_{T_1} \cdots \mathrm{d}\breve{\boldsymbol{x}}_{T_p}$ on the subset of $\breve{\mathbb{R}}^{2k}$ consisting of all $\breve{\boldsymbol{x}} = (\breve{x}_1, \dots, \breve{x}_k)$ such that $\breve{x}_i = \breve{x}_j$ if $i, j \in T_r$ for some $r \leq p$ and $\breve{x}_i \neq \breve{x}_j$ otherwise.

In the setting of the present proof, we note that the bounds on the mixed moments from (22) only involve coordinates with indices in the set $S' \cup T'$. Hence, we need to consider also singular differentials only with respect to these coordinates, i.e., integrate with respect to $\tilde{\mathrm{d}}\breve{\boldsymbol{x}}_{S' \cup T'}$. In particular, we arrive at the bound

$$\Big| \int_{D_K} \mathrm{d}U^{S',T'} \mathrm{d}M^{T_1} \cdots \mathrm{d}M^{T_p} \Big| \leq c \int_{D_K} \phi(D(\breve{\boldsymbol{x}}_{S' \cup T'})/2) \tilde{\mathrm{d}}\breve{\boldsymbol{x}}_{S' \cup T'} \mathrm{d}M^{T_1} \cdots \mathrm{d}M^{T_p}. \tag{23}$$

Now, setting $K = |E|^{-\varepsilon/128} |E'|^{-\varepsilon/128}$, the exponential decay assumption on the function ϕ gives control on one integral over the window, while the integrals with respect to the remaining variables are controlled by the volume of balls. Then, a repeated application of Hölder's inequality provides suitable bounds on the moment measures such that

$$\frac{1}{n} \int_{D_K} \phi(D(\breve{\boldsymbol{x}}_{S' \cup T'})/2) \tilde{\mathrm{d}}\breve{\boldsymbol{x}}_{S' \cup T'} \mathrm{d}M^{T_1} \cdots \mathrm{d}M^{T_p} \leq C|E||E'|$$

holds for some $C > 0$. Hence, it suffices to provide upper bounds for

$$\frac{1}{n}\int_{\{\check{\boldsymbol{x}}\in W_n^{(1,2)}\colon D(\check{\boldsymbol{x}})\le K\}} \mathrm{d}M^{T_1'}\cdots\mathrm{d}M^{T_{p'}'},$$

where $\{T_1', \dots, T_{p'}'\}$ is an arbitrary partition of $\{1,2,3,4\}$. We explain how to proceed for $p' = 1$, noting that for $p' > 1$ the arguments are similar but easier.

We claim that for some $C' > 0$,

$$\frac{1}{n}\int_{W_n\times\{1\}}\int_{B_K(x_1)\times\{1\}}\int_{(B_K(x_1)\times\{2\})^2} \mathrm{d}M^{\{1,2,3,4\}} \le C'|E|^{1/2+\varepsilon/8}|E'|^{1/2+\varepsilon/8}. \tag{24}$$

To prove this claim, decompose $M^{\{1,2,3,4\}}$ according to (19) and let $\{T_1'', \dots, T_{p''}''\}$ be an arbitrary partition of $\{1,2,3,4\}$. As in the proof of Proposition 9.2, a repeated use of Hölder's inequality shows that on $W_n^{(1,2)}$, the mixed moments of the form

$$m_n^{(T_1'',\dots,T_{p''}'')}(\check{x}_1,\dots,\check{x}_4)$$

are bounded above by $c'\big(\mathbb{P}_{\boldsymbol{x}}(E''_{x_1,x_i})\rho^{(p'')}(\boldsymbol{x})\big)^{1-\varepsilon}$ for a suitable $c' > 0$ and some $i \le 4$. At this point, we may proceed similarly as in (18) by invoking Lemmas 9.6(2) and 9.6(3). As an illustration consider the setting where $p'' = 4$ and $i = 2$. Then, we set $z' = x_2 - x_1$, $z_3 = x_3 - x_1$ and $z_4 = x_4 - x_1$. We combine Jensen's inequality with Lemma 9.6(3) to show that

$$\begin{aligned}
&\frac{1}{|B_K|^3}\int_{B_K^3}\big(\mathbb{P}_{o,z',z_3,z_4}(E''_{o,z'})\rho^{(4)}(o,z',z_3,z_4)\big)^{1-\varepsilon}\mathrm{d}z'\mathrm{d}z_3\mathrm{d}z_4\\
&\quad\le\Big(\int_{B_K^3}\frac{1}{|B_K|^3}\mathbb{P}_{o,z',z_3,z_4}(E''_{o,z'})\rho^{(4)}(o,z',z_3,z_4)\mathrm{d}z'\mathrm{d}z_3\mathrm{d}z_4\Big)^{1-\varepsilon}\\
&\quad\le C_{\mathbf{g}}^{1-\varepsilon}|B_K||E|^{1/2+\varepsilon/4}|E'|^{1/2+\varepsilon/4}.
\end{aligned}$$

Hence, inserting the definition of K concludes the proof. □

9.4. *Proof of Lemma 9.6*

We now turn to the proof of Lemma 9.6. The proof is based on the following four lemmas that are used to bound the probability with which certain point configurations occur. Throughout we use the notation

$$\begin{aligned}
E &= (b_-, b_+]\times(d_-, d_+],\\
E' &= (b'_-, b'_+]\times(d'_-, d'_+],\\
\delta_b &= b_+ - b_-, \quad \delta_d = d_+ - d_-,\\
\delta_{b'} &= b'_+ - b'_-, \quad \delta_{d'} = d'_+ - d'_-.
\end{aligned}$$

The proofs make use of the inequalities

$$|\sqrt{x} - \sqrt{y}| \le \sqrt{|x-y|} \tag{25}$$

$$|\arcsin(x) - \arcsin(y)| \leq C_0\sqrt{|x-y|}, \tag{26}$$

where $C_0 > 0$ is some constant. Moreover, we repeatedly use that the volume of an annulus is given by

$$|A_{b_-,b_+}(o)| = b_+^2 - b_-^2 \leq 2b_+\delta_b.$$

Lemma 9.8. *Let $x, y \in \mathbb{R}^2$ and $a = |x-y|/2$. There is a constant $C > 0$ such that for all $0 \leq a \leq d_+ \leq r_{\mathsf{f}}$,*

$$\begin{aligned}|D_{d_-,d_+}(x,y)| &= 2d_+^2\Big(\pi - \arcsin\big(\tfrac{a}{d_+}\big) + \tfrac{a}{d_+}\sqrt{1-\big(\tfrac{a}{d_+}\big)^2}\Big)\\ &\quad - 2(d_- \vee a)^2\Big(\pi - \arcsin\big(\tfrac{a}{d_-\vee a}\big) + \tfrac{a}{d_-\vee a}\sqrt{1-\big(\tfrac{a}{d_-\vee a}\big)^2}\Big)\\ &\leq Cd_+\delta_d^{1/2}.\end{aligned}$$

Proof. Recall that

$$D_{d_-,d_+}(x,y) = (B^+_{d_+}(x,y) \cup B^-_{d_+}(x,y))\backslash(B^+_{d_-\vee a}(x,y) \cup B^-_{d_-\vee a}(x,y)).$$

The line through x and y cuts the disk $B^+_d(x,y)$ into two parts. The area of the larger part is given by

$$d^2\Big(\pi - \arcsin(\tfrac{a}{d}) + \tfrac{a}{d}\sqrt{1-(\tfrac{a}{d})^2}\Big).$$

$D_{d_-,d_+}(x,y)$ is the union of two such sets of radius d_+ from which we remove two sets of the same type with radius $d_- \vee a$ from the interior. This yields the formula for the area.

The inequality follows from

$$\begin{aligned}d_+^2 - (d_- \vee a)^2 &\leq 2d_+\delta_d,\\ a(d_+ - d_- \vee a)\sqrt{1-\big(\tfrac{a}{d_-\vee a}\big)^2} &\leq d_+\delta_d,\end{aligned}$$

and, using (25) and (26),

$$\begin{aligned}&d_+^2\Big(\arcsin\big(\tfrac{a}{d_-\vee a}\big) - \arcsin\big(\tfrac{a}{d_+}\big) + \tfrac{a}{d_+}\Big(\sqrt{1-\big(\tfrac{a}{d_+}\big)^2} - \sqrt{1-\big(\tfrac{a}{d_-\vee a}\big)^2}\Big)\Big)\\ &\quad\leq d_+^2\Big(C_0\sqrt{\tfrac{a}{d_-\vee a} - \tfrac{a}{d_+}} + \tfrac{a}{d_+}\sqrt{(\tfrac{a}{d_-\vee a})^2 - (\tfrac{a}{d_+})^2}\Big)\\ &\quad\leq C_1 d_+^{3/2}\delta_d^{1/2}.\end{aligned}$$

□

Lemma 9.9. *Let $0 \leq b_- < b_+ \leq r_{\mathsf{f}}$ and $0 \leq d_- < d_+ \leq r_{\mathsf{f}}$ and let B_M be a disk of radius M. Then, there is a constant $C > 0$ such that*

$$\int_{B_M^3} \mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|y_1-y_2|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3)\mathrm{d}y_3\mathrm{d}y_2\mathrm{d}y_1 \leq C|B_M|d_+^2(\delta_b \vee \delta_d)^{\frac{1}{2}}\delta_b \wedge \delta_d.$$

Proof. Integration with respect to y_3 yields:

$$\int_{B_M^3} \mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|y_1-y_2|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3)\mathrm{d}y_3\mathrm{d}y_2\mathrm{d}y_1 \le \int_{B_M^2} \mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|y_1-y_2|}{2}\Big)|D_{d_-,d_+}(y_1,y_2)|\mathrm{d}y_2\mathrm{d}y_1.$$

When $\delta_b \le \delta_d$, the claim follows directly from Lemma 9.8. Otherwise, letting $a = |y_1 - y_2|/2$, we split the integral in two terms according to whether $a < d_-$ or $a \ge d_-$. Applying Lemma 9.8 yields the bound

$$C_1|B_M|\Bigg(\int_{b_-\wedge d_-}^{b_+\wedge d_-} a\Big(d_+^2\Big(\pi - \arcsin\big(\tfrac{a}{d_+}\big) + \tfrac{a}{d_+}\sqrt{1-\big(\tfrac{a}{d_+}\big)^2}\Big) - d_-^2\Big(\pi - \arcsin\big(\tfrac{a}{d_-}\big) + \tfrac{a}{d_-}\sqrt{1-\big(\tfrac{a}{d_-}\big)^2}\Big)\Big)\mathrm{d}a \tag{27}$$

$$+\int_{b_-\vee d_-}^{b_+\wedge d_+} a\Big(d_+^2\Big(\pi - \arcsin(\tfrac{a}{d_+}) + \tfrac{a}{d_+}\sqrt{1-(\tfrac{a}{d_+})^2}\Big) - a^2\tfrac{\pi}{2}\Big)\mathrm{d}a\Bigg). \tag{28}$$

To bound (27), we apply the mean value theorem and perform the integration to obtain the bound

$$\begin{aligned} &C_1|B_M|d_+^3\int_{b_-\wedge d_-}^{b_+\wedge d_-}\Big(\tfrac{1}{\sqrt{1-\left(\frac{a}{d_-}\right)^2}}\big(\tfrac{a}{d_-}-\tfrac{a}{d_+}\big) + \tfrac{a^2}{d_+d_-}\tfrac{1}{\sqrt{1-\left(\frac{a}{d_-}\right)^2}}\big(\tfrac{a}{d_-}-\tfrac{a}{d_+}\big)\Big)\mathrm{d}a \\ &\le 2C_1|B_M|d_+^2\delta_d\int_{b_-\wedge d_-}^{b_+\wedge d_-}\tfrac{a}{\sqrt{d_-^2-a^2}}\mathrm{d}a \\ &= 2C_1|B_M|d_+^2\delta_d\Big(\sqrt{d_-^2-(b_-\wedge d_-)^2}-\sqrt{d_-^2-(b_+\wedge d_-)^2}\Big) \\ &\le 4C_1|B_M|r_{\mathsf{f}}d_+^2\delta_d\delta_b^{1/2}. \end{aligned}$$

To bound (28), we bound the integrand using Lemma 9.8 and note that

$$|b_+\wedge d_+ - b_-\vee d_-| \le \delta_d\wedge\delta_b.$$

This proves the claim when $\delta_d \le \delta_b$. □

Lemma 9.10. *Let B_M be a disk of radius $M > 0$. There is a constant $C > 0$ such that for all $b_-, b_+, b'_-, b'_+, d_-, d_+ \in [0, r_{\mathsf{f}}]$ with $d_- < d_+$ and either $b_- < b_+ = b'_- < b'_+$ or $b_- = b'_-$ and $b_+ = b'_+$,*

$$\begin{aligned} &\int_{B_M^4}\mathbb{1}_{(b_-,b_+]\times(b'_-,b'_+]\times(0,b_+\vee b'_+]}\big(\tfrac{|x_1-x_2|}{2},\tfrac{|x_1-x_3|}{2},\tfrac{|x_2-x_3|}{2}\big) \\ &\qquad\times\mathbb{1}_{D_{d_-,d_+}(x_2,x_3)}(y_1)\mathrm{d}y_1\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &\le C|B_M|\delta_b\wedge\delta_{b'}(\delta_b\vee\delta'_b)^{3/4}\delta_d^{3/4}. \end{aligned}$$

Proof. We may assume $d_+ > 3\delta_d \vee 8\sqrt{r_{\mathfrak{f}}(\delta_b + \delta_{b'})}$. Indeed, if $d_+ \leq 3\delta_d$, we can show the claim by first integrating with respect to y_1, then using that by Lemma 9.9,

$$|D_{d_-,d_+}(x_2, x_3)| \leq C_1 d_+^2 \leq 9C_1\delta_d^2,$$

and finally integrating with respect to x_2 and x_3 to provide a factor $|B_M|\delta_b\delta_{b'}$. If $d_+ \leq 8\sqrt{r_{\mathfrak{f}}(\delta_b + \delta_{b'})}$, we first integrate with respect to x_1, which yields the area of $A_{2b_-,2b_+}(x_2) \cap A_{2b'_-,2b'_+}(x_3)$. This is bounded by $C_2\delta_b \wedge \delta_{b'}$, and by Lemma 9.9 the remaining integral is bounded by

$$C_3|B_M|d_+^2\delta_d \leq 64C_3|B_M|r_{\mathfrak{f}}(\delta_b + \delta_{b'})\delta_d \leq 128C_3|B_M|r_{\mathfrak{f}}(\delta_b \vee \delta_{b'})\delta_d.$$

Let $a = |x_2 - x_3|/2$. We write the integral as a sum of three terms corresponding to whether I: $a < d_+/4$, II: $d_+/4 \leq a < b_- \wedge b'_-$, or III: $b_- \wedge b'_- \leq a \leq b_+ \vee b'_+$.

Term I: We first integrate with respect to y_1. Since

$$\tfrac{a}{d_-} \leq \tfrac{d_+}{4d_-} = \tfrac{d_- + \delta_d}{4d_-} \leq \tfrac{3}{4},$$

the mean value theorem applied to the formula in Lemma 9.8 implies that $|D_{d_-,d_+}(x_2, x_3)| \leq C_4\delta_d$. We then integrate with respect to x_2 and x_3 to obtain the bound $C_5|B_M|\delta_b\delta_{b'}\delta_d$.

Term II: When $d_+/4 \leq a \leq b_- \wedge b'_-$, we first integrate with respect to x_1 to obtain the area of $A_{2b_-,2b_+}(x_2) \cap A_{2b'_-,2b'_+}(x_3)$. To bound term II, we need to explicitly compute this area. For this, we first compute the area $A_a(b_1, b_2)$ of the intersection $B_{2b_1}(x_2) \cap B_{2b_2}(x_3)$ where $b_1, b_2 \in \{b_+, b_-, b'_+, b'_-\}$. By the assumption on d_+,

$$a^2 \geq d_+^2/16 \geq 4r_{\mathfrak{f}}(\delta_b + \delta_{b'}) \geq 2(b_1^2 - b_2^2). \tag{29}$$

This ensures that the line containing the two points where the boundaries of the disks $B_{2b_1}(x_2)$ and $B_{2b_2}(x_3)$ meet separates x_2 and x_3. The area of $B_{2b_1}(x_2) \cap B_{2b_2}(x_3)$ is

$$\begin{aligned}\mathcal{A}_a(b_1, b_2) = 4\Bigg(& b_1^2 \arccos\left(\tfrac{a^2+b_1^2-b_2^2}{2ab_1}\right) + b_2^2 \arccos\left(\tfrac{a^2+b_2^2-b_1^2}{2ab_2}\right) \\ & - b_1 \tfrac{a^2+b_1^2-b_2^2}{2a}\left(1 - \left(\tfrac{a^2+b_1^2-b_2^2}{2ab_1}\right)^2\right)^{1/2} \\ & - b_2 \tfrac{a^2+b_2^2-b_1^2}{2a}\left(1 - \left(\tfrac{a^2+b_2^2-b_1^2}{2ab_2}\right)^2\right)^{1/2}\Bigg).\end{aligned}$$

The area of $A_{2b_-,2b_+}(x_2) \cap A_{2b'_-,2b'_+}(x_3)$ is given by

$$\begin{aligned}&\mathcal{A}_a(b_+, b'_+) + \mathcal{A}_a(b_-, b'_-) - \mathcal{A}_a(b_+, b'_-) - \mathcal{A}_a(b'_+, b_-) \\ &\quad = \int_{b_-}^{b_+}\int_{b'_-}^{b'_+} \tfrac{\partial^2}{\partial b_1 \partial b_2}\mathcal{A}_a(b_1, b_2)\mathrm{d}b_1\mathrm{d}b_2.\end{aligned} \tag{30}$$

It is a straightforward computation to see that $\frac{\partial^2}{\partial b_1 \partial b_2}\mathcal{A}_a(b_1, b_2)$ is uniformly bounded by C_6/d_+^2 on the set of $a, b_1, b_2 \leq r_{\mathfrak{f}}$ satisfying (29) and $d_+/4 \leq a \leq b_1 \wedge b_2$. In particular, (29) guarantees that

$$\frac{a^2 + b_1^2 - b_2^2}{2ab_1} \leq \frac{3a}{4b_1} \leq \frac{3}{4},$$

such that arccos and $x \mapsto \sqrt{1-x^2}$ have bounded derivatives for the relevant values of x. It follows that (30) is bounded by $C_7\delta_b\delta_{b'}/d_+^2$. The remaining integral is of order $|B_M|d_+^2\delta_d$ by Lemma 9.9, which yields the appropriate bound.

Term III: In this case, we first integrate with respect to x_1 providing a factor $\delta_b \wedge \delta_{b'}$. The remaining integral is bounded using Lemma 9.9. □

The fourth lemma allows us to analyze which point configurations can cause the birth and death of M-bounded features. To state it, we recall the α-complex associated with a locally finite point set $\mathcal{X} \subseteq \mathbb{R}^2$, see e.g. [19, Sec. III.4] for details. It is built from the Delaunay triangulation, which is a triangulation of the plane with vertex set $\mathcal{X}$. For $r > 0$, $\alpha_r(\mathcal{X})$ is the union of all edges in the Delaunay triangulation with length at most $2r$ and all triangles such that the three balls of radius r centered at its vertices cover the triangle. Then $\alpha_r(\mathcal{X}) \subseteq U_r(\mathcal{X})$ and the inclusion is a homotopy equivalence, i.e. it preserves the topology.

Lemma 9.11. *Let $\mathcal{X} \subseteq \mathbb{R}^2$ be locally finite.*

(i) Each connected component of $\mathbb{R}^2 \backslash \alpha_r(\mathcal{X})$ contains at most one M-bounded connected component of $\mathbb{R}^2 \backslash U_r(\mathcal{X})$.
(ii) If an M-bounded loop is born at time b because two balls centered at x_1, x_2 meet, then there is an edge of length $2b$ joining x_1, x_2 in the α-complex.
(iii) If an M-bounded feature dies at time d because exactly three balls centered at points y_1, y_2, y_3 meet, then y_1, y_2, y_3 form a triangle with no obtuse angle in the α-complex.

Proof. The analogous statements hold for unbounded loops by the homotopy equivalence between the α-complex and the union of balls. (i) follows because any M-bounded loop is also an unbounded loop. An M-bounded feature is either born the same way as the corresponding unbounded component or when two balls meet to split off a component. In both cases, some unbounded loop is born by the merging, and hence an edge is added to the α-complex. This shows (ii). When an M-bounded loop dies, so does the corresponding unbounded loop, hence (iii) is clear. □

We are now ready to prove Lemma 9.6.

Proof of Lemma 9.6. **Proof of (1).** Stationarity and Equation (4) yield

$$\begin{aligned}&\int_{B^p} \mathbb{P}_{o,\boldsymbol{z}}(E_o)\rho^{(p+1)}(o,\boldsymbol{z})\mathrm{d}\boldsymbol{z}\\&\quad=\int_{[0,1]^2}\int_{(B+x)^p}\mathbb{P}_{x,\boldsymbol{z}}(E_x)\rho^{(p+1)}(x,\boldsymbol{z})\mathrm{d}\boldsymbol{z}\mathrm{d}x\\&\quad=\mathbb{E}\Bigg[\sum_{(x,\boldsymbol{z})\in\mathcal{P}^{p+1}_{\neq}}\mathbb{1}_{[0,1]^2}(x)\mathbb{1}_{(B+x)^p}(\boldsymbol{z})\mathbb{1}_{E_x}\Bigg].\end{aligned}\tag{31}$$

In the following, we let $\boldsymbol{y}=(y_1,y_2,y_3)$, and

$$g(x_1,x_2,\boldsymbol{y})=\mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|x_1-x_2|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3)$$

for simplicity. By definition of E_x, (31) is bounded by

$$\begin{aligned}&\mathbb{E}\Bigg[\sum_{x_1\in\mathcal{P}}\mathcal{P}(B+x_1)^p\mathbb{1}_{[0,1]^2}(x_1)\sum_{x_2\in\mathcal{P}}\sum_{\boldsymbol{y}\in\mathcal{P}^3_{\neq}}\mathbb{1}_{B_M(x_1)^3}(\boldsymbol{y})g(x_1,x_2,\boldsymbol{y})\Bigg]\\&\quad=\mathbb{E}\Bigg[\sum_{(x_1,x_2,\boldsymbol{y})\in\mathcal{P}^5_{\neq}}\mathcal{P}(B+x_1)^p\mathbb{1}_{[0,1]^2}(x_1)\mathbb{1}_{B_M(x_1)^3}(\boldsymbol{y})g(x_1,x_2,\boldsymbol{y})\Bigg]\\&\qquad+3\mathbb{E}\Bigg[\sum_{(x_1,\boldsymbol{y})\in\mathcal{P}^4_{\neq}}\mathcal{P}(B+y_1)^p\mathbb{1}_{[0,1]^2}(y_1)\mathbb{1}_{B_M(y_1)^2}(y_2,y_3)g(x_1,y_1,\boldsymbol{y})\Bigg]\\&\qquad+3\mathbb{E}\Bigg[\sum_{(x_1,\boldsymbol{y})\in\mathcal{P}^4_{\neq}}\mathcal{P}(B+x_1)^p\mathbb{1}_{[0,1]^2}(x_1)\mathbb{1}_{B_M(x_1)^3}(\boldsymbol{y})g(x_1,y_1,\boldsymbol{y})\Bigg]\\&\qquad+6\mathbb{E}\Bigg[\sum_{\boldsymbol{y}\in\mathcal{P}^3_{\neq}}\mathcal{P}(B+y_2)^p\mathbb{1}_{[0,1]^2}(y_2)\mathbb{1}_{B_M(y_2)^2}(y_1,y_3)g(y_1,y_2,\boldsymbol{y})\Bigg].\end{aligned}\tag{32}$$

Here, we have used that $g(x_1,x_2,\boldsymbol{y})$ is symmetric in x_1 and x_2 and in y_1, y_2, and y_3. Applying (4) again, we may bound the last term in (32) by

$$6\int_{B^3_{M+2}}\mathbb{E}_{\boldsymbol{y}}[\mathcal{P}(B+y_1)^p]g(y_1,y_2,\boldsymbol{y})\rho^{(3)}(\boldsymbol{y})\mathrm{d}\boldsymbol{y},\tag{33}$$

since $b_+\le M$. The remaining terms are treated similarly. Now choose a covering $B+x_1\subseteq\bigcup_{i\le\ell}W_1^{(i)}$, where each $W_1^{(i)}$ is a translation of W_1 and such that $\ell\le C_1|B|$ for some C_1 independent of K (for instance using that $B_K\subseteq W_{4\lceil K\rceil^2}$). Then, by the moment condition **(M)** for $\boldsymbol{x}=(x_1,\dots,x_k)$,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}}\big[\mathcal{P}(B_K+x_1)^p\big]&\le\sup_{\boldsymbol{x}\in\mathbb{R}^{2k}}\mathbb{E}_{\boldsymbol{x}}\Big[\mathcal{P}\Big(\bigcup_{i=1}^{\ell}W_1^{(i)}\Big)^p\Big]\\&\le\ell^p\sum_{i\le\ell}\sup_{\boldsymbol{x}\in\mathbb{R}^{2k}}\mathbb{E}_{\boldsymbol{x}}\big[\mathcal{P}\big(W_1^{(i)}\big)^p\big]\end{aligned}$$

$$\leq \ell^p \sum_{i\leq \ell} \sup_{\boldsymbol{x}\in\mathbb{R}^{2k}} \mathbb{E}^!_{\boldsymbol{x}}\big[\big(\mathcal{P}\big(W_1^{(i)}\big)+k\big)^{p\vee k}\big]$$
$$\leq C_2\ell^{p+1}\Big(\sup_{\boldsymbol{x}\in\mathbb{R}^{2k}} \mathbb{E}^!_{\boldsymbol{x}}\big[\mathcal{P}(W_1)^{p\vee k}\big]+k^{p\vee k}\Big)$$
$$\leq C_3|B|^{p+1}.$$

We apply this in (33) together with Lemma 9.9. Since each $\rho^{(k)}$ is bounded according to the assumption of fast decay of correlations, we obtain the bound $C_4|B|^{p+1}|E|^{1/2+\varepsilon}$.

Proof of (2). In the following, we use the notation

$$g'(x_1,x_2,\boldsymbol{y}) = \mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_1-x_2|}{2}\Big)\mathbb{1}_{D_{d'_-,d'_+}(y_1,y_2)}(y_3).$$

Note that since the blocks E and E' are neighboring, the features in E and E' are different. Putting $\boldsymbol{x}=(x_1,x_2,x_3)$, we now expand as in (31)

$$\begin{aligned}
&\int_{B^p} \mathbb{P}_{o,\boldsymbol{z}}(E''_{o,o})\rho^{(p+1)}(o,\boldsymbol{z})\mathrm{d}\boldsymbol{z}\\
&\quad\leq \mathbb{E}\Bigg[\sum_{\substack{x_1\in\mathcal{P}\cap[0,1]^2\\(x_2,x_3)\in\mathcal{P}^2_{\neq}}}\ \sum_{\substack{\boldsymbol{y},\boldsymbol{y}'\in\mathcal{P}^3_{\neq}\cap B_M(x_1)^3\\\boldsymbol{y}\neq\boldsymbol{y}'}} \mathcal{P}(B+x_1)^p g(x_1,x_2,\boldsymbol{y})g'(x_1,x_3,\boldsymbol{y}')\mathbb{1}_A(\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}')\Bigg]\\
&\quad\leq \mathbb{E}\Bigg[\sum_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}'\in\mathcal{P}^3_{\neq}\cap B^3_{M+2}\\\boldsymbol{y}\neq\boldsymbol{y}'}} \mathcal{P}(B+x_1)^p g(x_1,x_2,\boldsymbol{y})g'(x_1,x_3,\boldsymbol{y}')\mathbb{1}_A(\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}')\Bigg].
\end{aligned}\tag{34}$$

The condition $x_2\neq x_3$ comes from the fact that x_1 can give birth to at most one feature when connecting to another point, and since E and E' are neighboring, x_2 and x_3 correspond to different features. Similarly, $\boldsymbol{y}'\neq\boldsymbol{y}$ comes from the fact that a triangle can kill at most one feature.

The event A excludes certain point configurations that are not possible. If the triangles formed by $\boldsymbol{y}$ and $\boldsymbol{y}'$ share an edge, and the vertices of this edge coincide with x_2 and x_3, then $|x_2-x_3|>2(b_+\vee b'_+)$ is not allowed. Indeed, it follows from Lemma 9.11 that the triangles correspond to the same feature in the α-complex until x_2 and x_3 are joined. Thus, this must happen before both triangles are born, that is, at the latest at time $b_+\vee b'_+$. Moreover, if the two triangles share an edge, then the two points in $\boldsymbol{y},\boldsymbol{y}'$ not lying on this edge cannot be equal to x_1 and x_2 or to x_1 and x_3, as this would lead to crossing edges in the α-complex by Lemma 9.11 (since the triangles formed by $\boldsymbol{y},\boldsymbol{y}'$ cannot have any obtuse angles).

We now write the sum in (34) as a sum where each term is a sum over $\mathcal{P}^k_{\neq}$, $4\leq k\leq 9$, as in (32). Each such term comes from grouping $\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}'$ into sets of equal points. Consider for illustration the term corresponding to the situation $x_2=y'_1, x_1=y_2=y'_2, x_3=y_3=y'_3$. The sum is handled as in the proof of

Lemma 9.6(1) by applying (4) and bounding the involved Palm means. For this special point configuration, it is sufficient to bound $\mathbb{1}_A$ by 1.

$$\frac{1}{|B|^{p+1}}\mathbb{E}\Big[\sum_{(\boldsymbol{x},y_1)\in\mathcal{P}^4_{\neq}\cap B^4_{M+2}} \mathcal{P}(B+x_1)^p g(x_1,x_2,y_1,x_1,x_3)g'(x_1,x_2,\boldsymbol{x})\Big]$$
$$\leq C_5\int_{B^4_{M+2}} g(x_1,x_2,y_1,x_1,x_3)g'(x_1,x_3,\boldsymbol{x})\mathrm{d}y_1\mathrm{d}\boldsymbol{x}.$$

Now, we apply the Hölder inequality with $\frac{1}{q_1}+\frac{1}{q_2}=1$ to obtain the bound

$$C_5\Big[\int_{B^4_{M+2}} \mathbb{1}_{(b_-,b_+]\times(b'_-,b'_+]}\Big(\tfrac{|x_1-x_2|}{2},\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(x_1,x_3)}(y_1)\mathrm{d}y_1\mathrm{d}\boldsymbol{x}\Big]^{\frac{1}{q_1}}$$
$$\times\Big[\int_{B^4_{M+2}} \mathbb{1}_{D_{d_-,d_+}(x_1,x_3)}(y_1)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d'_-,d'_+}(x_1,x_2)}(x_3)\mathrm{d}y_1\mathrm{d}\boldsymbol{x}\Big]^{\frac{1}{q_2}}. \tag{35}$$

In the first integral, we first integrate with respect to x_2 and then apply Lemma 9.9, while in the second integral we first integrate with respect to y_1 and use the bound in Lemma 9.8 and then apply Lemma 9.9 again. Next we use that E and E' are neighboring blocks so that either $\delta_b=\delta_{b'}$ or $\delta_d=\delta_{d'}$.

When $\delta_b=\delta_{b'}$, we get the bound

$$C_6\big(\delta_b(\delta_{b'}\delta_d)^{\frac{3}{4}}\big)^{\frac{1}{q_1}}\big(\delta_d^{\frac{1}{2}}(\delta_{b'}\delta_{d'})^{\frac{3}{4}}\big)^{\frac{1}{q_2}} = C_6\delta_b^{\frac{3}{4}+\frac{1}{q_1}}\delta_d^{\frac{3}{4}\cdot\frac{1}{q_1}+\frac{1}{2}\cdot\frac{1}{q_2}}\delta_{d'}^{\frac{3}{4}\cdot\frac{1}{q_2}}, \tag{36}$$

so we take $1/q_1>1/4$ and $1/q_2>2/3$.

When $\delta_d=\delta_{d'}$, we use Lemma 9.9 to get the bound

$$C_7\big(\delta_b(\delta_{b'}\delta_d)^{\frac{3}{4}}\big)^{\frac{1}{q_1}}\big(\delta_d^{\frac{1}{2}}\delta_{b'}^{\frac{1}{2}}\delta_{d'}\big)^{\frac{1}{q_2}} = C_7\delta_b^{\frac{1}{q_1}}\delta_{b'}^{\frac{3}{4}\cdot\frac{1}{q_1}+\frac{1}{2}\cdot\frac{1}{q_2}}\delta_d^{\frac{3}{4}\cdot\frac{1}{q_1}+\frac{3}{2}\cdot\frac{1}{q_2}}, \tag{37}$$

so we take $1/q_1>1/2$ and $1/q_2>1/3$.

For a general term, note that there are at least four different points among $\boldsymbol{y},\boldsymbol{y}'$, so one of them, say y_1, cannot be equal to any of $\boldsymbol{x}$. We consider two cases:

I y_1 is not among y'_1,y'_2,y'_3.
II $y_1=y'_1$, $y_2=y'_2$, and $y_3=x_2$ and $y'_3=x_3$.

Since we no longer keep track of which edge kills which triangle, all possible point configurations allowed by A fall into one of the above cases after possibly renaming the variables.

In particular, if $y_1=y'_1$ and the points y_2,y_3,y'_2,y'_3 are all different, one of them cannot be any of x_1,x_2,x_3, and we could have taken this as y_1 and be in Case I. If $y_1=y'_1$, $y_2=y'_2$ and, say, y_3 is not any of x_1, x_2, x_3, we could have chosen y_3 as y_1 and be in Case I.

We further divide the Case I configurations allowed by A into the following two sub-cases that have to be treated separately:

Ia. x_3 is not any of y_2, y_3.
Ib. $x_2 = y_2 = y_2'$, $x_3 = y_3 = y_3'$, $|x_2 - x_3|/2 \le b_+ \vee b'_+$.

Again, after renaming the variables, we are always in one of the two sub-cases.

Case Ia: We apply the Hölder inequality to

$$\begin{aligned}
&\mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|x_1-x_2|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d'_-,d'_+}(y_1',y_2')}(y_3') \\
&\quad = \mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|x_1-x_2|}{2}\Big)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3) \\
&\qquad \times \mathbb{1}_{D_{d_-,d_+}(y_1,y_2)}(y_3)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d'_-,d'_+}(y_1',y_2')}(y_3').
\end{aligned} \tag{38}$$

The first factor is integrated with respect to x_3 and the remaining integral is bounded using Lemma 9.9. The second factor is first integrated wrt. y_1, the result is bounded using Lemma 9.8, and the remaining integral is bounded using Lemma 9.9. The rest of the argument proceeds as in the special case treated above.

Case Ib: The claim follows by applying the Hölder inequality to (38) and arguing as in Case Ia using Lemma 9.10 to bound the first integral.

Case II: We apply the Hölder inequality exactly as in (38) and argue as in Case Ia, except that the second integral is first integrated with respect to y_3 rather than y_1.

Proof of (3). As in (31), we find

$$\begin{aligned}
&\int_{B^{p+1}} \mathbb{P}_{o,z',\boldsymbol{z}}(E''_{o,z'})\rho^{(p+2)}(o,z',\boldsymbol{z})\mathrm{d}z'\mathrm{d}\boldsymbol{z} \\
&\le \mathbb{E}\Big[\sum_{(x,z')\in\mathcal{P}^2_{n\neq}} \mathcal{P}(B+x)^p\mathbb{1}_{[0,1]^2}(x)\mathbb{1}_{B+x}(z')\mathbb{1}_{E''_{x,z'}}\Big] \\
&\le \mathbb{E}\Big[\sum_{(x_1,z')\in\mathcal{P}^2_{\neq}}\sum_{(x_2,x_2')\in\mathcal{P}^2}\sum_{\boldsymbol{y}\in\mathcal{P}^3_{\neq}\cap B_M(x_1)^3}\sum_{\substack{\boldsymbol{y}'\in\mathcal{P}^3_{\neq}\cap B_M(z')^3\\ \boldsymbol{y}\neq\boldsymbol{y}'}} \mathcal{P}(B+x_1)^p\mathbb{1}_{[0,1]^2}(x_1) \\
&\qquad \times \mathbb{1}_{B+x_1}(z')g(x_1,x_2,\boldsymbol{y})g'(z',x_2',\boldsymbol{y}')\mathbb{1}_{\tilde{A}}(x_1,z',x_2,x_2',\boldsymbol{y},\boldsymbol{y}')\Big].
\end{aligned}$$

The set $\tilde{A}$ consists of tuples of points $(x_1,x_2,x_3,x_4,\boldsymbol{y},\boldsymbol{y}') \in \mathbb{R}^{20}$ and, similar to A, it excludes certain configurations of the points $(x_1,x_2,x_3,x_4,\boldsymbol{y},\boldsymbol{y}')$ that are not allowed by Lemma 9.11. If the triangles formed by $\boldsymbol{y}$ and $\boldsymbol{y}'$ share an edge, then the length of this edge must be at most $2(b_+ \vee b'_+)$. Moreover, if the two triangles share an edge, then the two points in $\boldsymbol{y},\boldsymbol{y}'$ not lying on this edge cannot be equal to x_1 and x_3 or to x_2 and x_4.

The contribution from the cases where two of the points x_1, x_2, x_2', z' are identical is bounded by

$$\mathbb{E}\Big[\sum_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{y}'\in\mathcal{P}^3_{\neq}\cap B^3_{2M+2}\\ \boldsymbol{y}'\neq\boldsymbol{y}}} \mathcal{P}(B+x_1)^p g(x_1,x_2,\boldsymbol{y})g'(x_1,x_3,\boldsymbol{y}')\mathbb{1}_{\tilde{A}}(x_1,x_2,x_1,x_3,\boldsymbol{y},\boldsymbol{y}')\Big],$$

which is handled exactly as in the proof of Lemma 9.6(2). Thus, it remains to treat the terms where x_1, x_2, x_2', z' are all different. Therefore, if we put $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$, we must bound

$$\mathbb{E}\Bigg[\sum_{\substack{\boldsymbol{x}\in\mathcal{P}_{\neq}^4 \\ x_1\in[0,1]^2}} \sum_{\boldsymbol{y}\in\mathcal{P}_{\neq}^3\cap B_M(x_1)^3} \sum_{\substack{\boldsymbol{y}'\in\mathcal{P}_{\neq}^3\cap B_M(x_2)^3 \\ \boldsymbol{y}\neq\boldsymbol{y}'}} \mathcal{P}(B+x_1)^p \mathbb{1}_{B+x_1}(x_2) \times g(x_1, x_3, \boldsymbol{y}) g'(x_2, x_4, \boldsymbol{y}') \mathbb{1}_{\tilde{A}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{y}') \Bigg].$$

The rest of the proof proceeds as the proof of Lemma 9.6(2) by suitable applications of the Hölder inequality. We divide into two cases according to whether all points in $\boldsymbol{y}$, $\boldsymbol{y}'$ are one of $\boldsymbol{x}$ or not. After renaming the variables, we may assume

I $y_1 = y_1' = x_1$, $y_2 = y_2' = x_2$, $y_3 = x_3$, and $y_3' = x_4$, or
II y_1 is not any of $\boldsymbol{x}$.

Notice that in Case I we exclude the case $y_1 = y_1' = x_1$, $y_2 = y_2' = x_3$, $y_3 = x_2$, and $y_3' = x_4$ because it was excluded by definition of $\tilde{A}$. After renaming variables, Case II is divided into

IIa y_1 is not any of $\boldsymbol{x}$ or $\boldsymbol{y}'$, and x_1 is not any of y_2, y_3.
IIb $y_1 = y_1'$ and y_1 is not any of $\boldsymbol{x}$, $y_2 = y_2' \neq x_3$, $y_3 = x_1$.
IIc $y_1 = y_1'$, $y_2 = x_2$, $y_3 = x_4$, $y_2' = x_1$, $y_3' = x_3$.
IId $y_1 = y_1'$, $y_2 = x_1$, $y_3 = x_2$, $y_2' = x_3$, $y_3' = x_4$.

In Case IIa, y_1 is not one of $\boldsymbol{y}'$, while in Case IIb, IIc, and IId it is. Case IIb corresponds to the situation in which the triangles formed by $\boldsymbol{y}, \boldsymbol{y}'$ share an edge, while in Case IIc and IId they share only one vertex. In Case IIc, each triangle contains one of the edges joining x_1 to x_3 and x_2 to x_4, while in Case IId they do not.

Case I: When $\delta_b = \delta_{b'}$, we first write

$$\begin{aligned} &\mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}}(y_1,y_2)(y_3)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_2-x_4|}{2}\Big)\mathbb{1}_{D_{d'_-,d'_+}}(y_1',y_2')(y_3') \\ &\quad = \mathbb{1}_{(b_-,b_+]}\Big(\tfrac{|x_1-x_3|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}}(y_1,y_2)(y_3)\mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_2-x_4|}{2}\Big) \qquad (39) \\ &\qquad \times \mathbb{1}_{(b'_-,b'_+]}\Big(\tfrac{|x_2-x_4|}{2}\Big)\mathbb{1}_{D_{d_-,d_+}}(y_1,y_2)(y_3)\mathbb{1}_{D_{d'_-,d'_+}}(y_1',y_2')(y_3'). \qquad (40) \end{aligned}$$

We then apply the Hölder inequality. Integrating first with respect to x_4 and then y_2 in (39) and integrating with respect to y_3 first in (40) yields a bound of order

$$(\delta_{b'}(\delta_b\delta_d)^{\frac{3}{4}})^{\frac{1}{q_1}}(\delta_d(\delta_{b'}\delta_{d'})^{\frac{3}{4}})^{\frac{1}{q_2}}.$$

This is the same as (36) since $\delta_b = \delta_{b'}$. When $\delta_d = \delta_{d'}$, we replace $\mathbb{1}_{D_{d_-,d_+}}(y_1,y_2)(y_3)$ by $\mathbb{1}_{D_{d'_-,d'_+}}(y_1',y_2')(y_3')$ in (39), to obtain a bound of order

$$(\delta_b(\delta_{b'}\delta_{d'})^{\frac{3}{4}})^{\frac{1}{q_1}}(\delta_d^{\frac{1}{2}}\delta_{b'}^{\frac{1}{2}}\delta_{d'})^{\frac{1}{q_2}},$$

which reduces to the same form as (37).

Case IIa: We apply the Hölder inequality to (39)–(40) and integrate first with respect to x_1 and then y_1 in (39) and with respect to y_1 first in (40). The remaining argument proceeds as in the proof of Lemma 9.6(2) Ia.

Case IIb: We apply the Hölder inequality to (39)–(40) and integrate first with respect to x_3 and then y_1 in (39) and with respect to y_3 first in (40) and argue as in the proof of Lemma 9.6(2) Ia.

Case IIc: In (39), we first integrate with respect to x_1. In (40), we first integrate with respect to y_2' and y_3' to obtain a factor $\delta_{d'}$. Then we integrate with respect to y_1 and x_2 and apply Lemma 9.9 to obtain a factor $\delta_d^{1/2}\delta_{b'}$. The resulting bounds are stricter than (36) and (37).

Case IId: Here we integrate (39) with respect to x_3 first and then y_1 while (40) is integrated first with respect to y_2 and then y_1.

In all cases treated above, a minor difference to (35) is that the integration domains are slightly more complicated due to the indicator $\mathbb{1}_{B+x_1}(x_2)$. However, it contributes at most a factor $C_7|B|$ to the bound, and this cancels when we divide by $|B|^{p+2}$. □