

Betti numbers in the random connection model for higher-dimensional simplicial complexes and the Boolean model

Dominik Pabst  a,b,*

^a Institute of Stochastics, Karlsruhe Institute of Technology, Karlsruhe, 76131, Germany

^b Institute of Theoretical Physics, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, 91058, Germany

ARTICLE INFO

2020 MSC:

60F05

60D05

05C80

60B99

Keywords:

Random connection model

Betti numbers

Central limit theorems

Boolean model

Simplicial complexes

ABSTRACT

Random simplicial complexes, as generalizations of random graphs, have become increasingly popular in the literature in recent years. In this paper, we consider a new model for a random simplicial complex that was introduced in [1], which generalizes the Random Connection Model (RCM) in a natural way and includes several models used in the literature as special cases. We focus on the marked stationary case with vertices in $\mathbb{R}^d \times \mathbb{A}$, where the mark space $\mathbb{A}$ is an arbitrary Borel space. We will derive a central limit theorem (CLT) for an abstract class of functionals and show that many of the typical functionals considered in the study of simplicial complexes, such as Betti numbers, fall into this class. As an important special case, we obtain a CLT for Betti numbers in the Boolean model.

1. Introduction

Random graphs provide an effective mathematical framework for modeling and analyzing complex systems built on pairwise interactions. Applications of random graphs span a wide range of disciplines, including the spread of diseases in epidemiology, communication and search algorithms in computer science, the structure of social and economic networks, and critical phenomena in statistical physics. While graph models effectively capture pairwise interactions, their extension into higher dimensions through random simplicial complexes enables a deeper exploration of group dynamics and multi-way relationships. In recent years, random simplicial complexes have gained significant attention in mathematics, ranging from combinatorial models, such as extensions of the classic Erdős-Rényi graph [4,5], to applications in topological data analysis (see for example [6]), where the Vietoris-Rips and the Čech complex are used to study the structure of spatial data and extract topological features.

Betti numbers are among the most fundamental quantities studied in simplicial complexes, providing essential information about the topological structure of a space, such as its connectedness, loops, and voids. In topological data analysis, the Betti numbers play a crucial role in extracting topological features from data, offering a powerful way to quantify the underlying shape of datasets. CLTs for the Betti numbers of the Čech and the Vietoris-Rips complex are, for example, provided in [7,8]. So-called soft random simplicial complexes, generalizations of the Čech and the Vietoris-Rips complex with additional randomness, have been introduced in [9]. Asymptotic results for Betti numbers of soft random simplicial complexes can be found in [10].

Formally, a *simplicial complex* is a collection of non-empty finite subsets, known as simplices, of a vertex set V , with the condition that for each simplex, all of its faces (non-empty subsets) are also included. A simplex containing $n + 1$ vertices is called an n -simplex,

* correspondence to: Institute of Theoretical Physics, Friedrich Alexander University Erlangen-Nuremberg, Erlangen, 91058, Germany.

E-mail address: dominik.pabst@fau.de

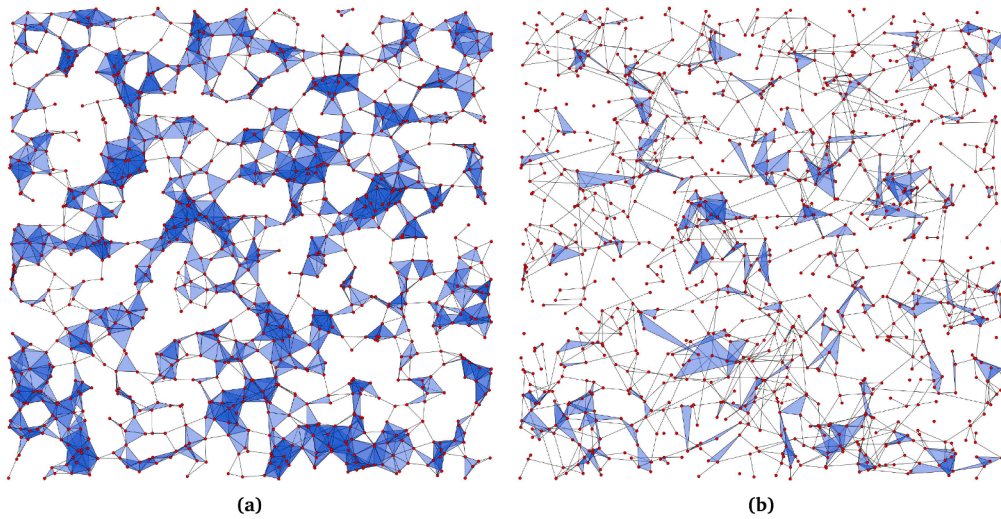


Fig. 1. (a) A realization of the model with the connection functions from (1) and (b) a realization of the model with the connection functions from (2) based on the same realization of a Poisson process.

and n is also referred to as the dimension of the simplex. The dimension of a simplicial complex is defined as the supremum of the dimensions of its simplices and can be infinite. In this sense, graphs can be understood as 1-dimensional simplicial complexes. The j -skeleton of a simplicial complex refers to the subcomplex formed by restricting the complex to simplices of dimension at most j .

In this paper we present a model extending the *Random Connection Model (RCM)* to a random simplicial complex, as first proposed in [1]. To this end, we fix a natural number α , which determines the maximal dimension of the simplicial complex. In the special case $\alpha = 1$, the model reduces to a random graph and coincides with the classical RCM, which has been studied as a random graph model for several decades (see, e.g., [2,11–13]) since it was introduced by Penrose [14] in 1991. In particular, all results obtained in this paper apply to random graphs as a special case of the present framework. Although the model can be defined and studied in a more general setting (as in [1]), the methods employed in this paper rely on certain properties of Euclidean space. Nevertheless, it turns out that we are not limited to $\mathbb{R}^d$ alone; we can, in fact, consider the product space $\mathbb{R}^d \times \mathbb{A}$, where $(\mathbb{A}, \mathcal{T})$ is an arbitrary Borel space equipped with a probability measure Θ .

The vertex set of the random simplicial complex will be the set of points of a Poisson process Φ on $\mathbb{R}^d \times \mathbb{A}$ with intensity measure $\gamma \lambda_d \otimes \Theta$, where $\gamma > 0$ is the intensity of the model and λ_d denotes the Lebesgue measure on $\mathbb{R}^d$. For all $j \in \{1, \dots, \alpha\}$ we fix a measurable, symmetric function $\varphi_j : (\mathbb{R}^d \times \mathbb{A})^{j+1} \rightarrow [0, 1]$ which we assume to be additionally translation-invariant in the sense of (7). We call the functions $\varphi_1, \dots, \varphi_\alpha$ the connection functions of the model and define the random simplicial complex Δ in the following way.

- (0) The vertex set of Δ is the set of points of Φ .
- (1) For each pair of points $x, y \in \Phi$ add the edge $\{x, y\}$ independently with probability $\varphi_1(x, y)$ to the complex Δ .
- (2) For each triple $x, y, z \in \Phi$, all of whose edges have been added to the complex, add the triangle $\{x, y, z\}$ independently with probability $\varphi_2(x, y, z)$ to the complex Δ .
- $\vdots$
- (α) For $\alpha + 1$ points $x_{i_0}, \dots, x_{i_\alpha} \in \Phi$, all of whose subsimplices have been added to the complex, add the simplex $\{x_{i_0}, \dots, x_{i_\alpha}\}$ independently with probability $\varphi_\alpha(x_{i_0}, \dots, x_{i_\alpha})$ to the complex Δ .

We refer to the introduced framework as the *marked stationary* case, which is consistent with the terminology used in the literature for the RCM. We begin by illustrating the model through two examples. As two-dimensional simplicial complexes are especially illustrative, we choose $\alpha = 2$. In our first example model, we choose the connection functions

$$\varphi_1(x, y) = \mathbb{1}\{\|x - y\| \leq r\}, \quad \varphi_2(x, y, z) \equiv p \quad (1)$$

for $r > 0$ and $p \in [0, 1]$. The model is built in two steps: first, a geometric graph with parameter r is formed; second, each potential triangle is added with fixed probability p . For the second example, we consider the connection functions

$$\varphi_1(x, y) = \exp(-\theta_1 \|x - y\|), \quad \varphi_2(x, y, z) = \exp(-\theta_2 V(x, y, z)) \quad (2)$$

for $\theta_1, \theta_2 > 0$, where $V(x, y, z)$ denotes the (two-dimensional) volume of the triangle spanned by x, y, z . Fig. 1 shows realizations of both models in dimension $d = 2$, within the unit square $[0, 1]^2$. Both are based on the same realization of a Poisson process, using $r = p = \frac{1}{2}$ for (a), and $\theta_1 = 4, \theta_2 = \frac{1}{100}$ for (b).

The connection functions can also be chosen such that Δ coincides with the α -skeleton of the Čech complex, the Vietoris-Rips complex, or the complexes considered in [9]. However, for functionals for which there exists $j \in \mathbb{N}$ such that they depend only on the j -skeleton of a simplicial complex, this does not impose any restriction, since one can simply choose $\alpha \geq j$. This applies to all concrete

example functionals considered in this paper, except for the Euler characteristic. For example, the p th Betti number of a simplicial complex is already determined by its $(p + 1)$ -skeleton. In [Section 6.2](#), we will see that the CLT can nevertheless be applied to the Euler characteristic of both the Čech and the Vietoris-Rips complex.

Our main result is a CLT that applies to various functionals, including the *Betti numbers*. The proof builds on the methods developed in [\[2\]](#), where the RCM is studied as a random graph in the unmarked stationary setting. Note that the unmarked stationary case discussed in [\[2\]](#) arises as a special case of the marked framework presented here by taking the mark space $\mathbb{A}$ to consist of a single element. Consequently, the CLT we prove is novel in two aspects: it applies to simplicial complexes with dimensions greater than one and also addresses the marked case, even specifically for graphs. In [\[2\]](#), the CLT is applied to the Betti numbers of the clique complex of the RCM. The clique complex of a graph is the complex formed by adding all simplices of dimension greater than 1 whose edges are already present in the graph. To the best of our knowledge, the result in [\[2\]](#) represents the pioneering CLT addressing higher Betti numbers (excluding the zeroth Betti number) within the context of the RCM. This result can be recovered as a special case of the present framework by choosing all connection functions except for φ_1 to be constant and equal to 1 in the unmarked case. In contrast, our results apply directly to the first Betti number β_1 of the underlying random graph (corresponding to the case $\alpha = 1$). In particular, this yields a CLT for β_1 of the random graph, which is new even in the unmarked case. A CLT for the number of components, i.e., the zeroth Betti number, can be found in [\[13\]](#) (Theorem 9.1). The paper [\[15\]](#) studies the asymptotic behavior of various functionals in the RCM on a general space. These include the number of vertices of a fixed degree, as well as the number of induced subgraphs or connected components of a given order.

An important application of our result refers to the *Boolean model*, one of the flagship models of stochastic geometry. In [\[16\]](#), a CLT (Theorem 9.1) for a large class of functionals in the Boolean model is presented. However, this does not include the Betti numbers, as the functionals considered there are additive. Indeed, as far as we know, CLTs for Betti numbers in the Boolean model have been known only in a very restrictive special case, specifically when balls with identical radii are used as grains. This special case corresponds to the Čech complex, and a corresponding result can be found, for example, in [\[8\]](#) (Theorem 4.7). For more details on the Boolean model, we refer to [Section 8](#). In this paper, results and methods derived from the PhD thesis [\[17\]](#) are presented.

The main results of the paper are stated in [Section 2](#). After introducing the required preliminaries in [Section 3](#) and providing a precise definition of the model in [Section 4](#), [Section 5](#) establishes and proves the CLT for weakly stabilizing functionals. [Section 6](#) studies the class of weakly stabilizing functionals. We show that many typical functionals arising in the context of graphs and simplicial complexes are weakly stabilizing under suitable integrability assumptions. [Section 7](#) addresses the question of when the asymptotic variance in the CLT is strictly positive. Finally, [Section 8](#) illustrates how to derive a CLT for Betti numbers within the Boolean model.

2. Main results

Our main result is a CLT for so-called weakly stabilizing functionals according to the upcoming definition. We denote by Δ the random simplicial complex described in the introduction. A precise definition of the model is given in [Section 4](#). For a measurable set $W \subset \mathbb{R}^d$, let Δ_W denote the restriction of Δ to simplices with vertices in $W \times \mathbb{A}$. A key tool for deriving a CLT and defining weakly stabilizing functionals is a suitable difference operator, which measures the influence of adding an additional vertex to the model.

To define the difference operator, we work with an augmented Poisson process Ψ on $\mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$, in which each point carries additional auxiliary marks taking values in $[0, 1] \times \mathbb{M}$. Here $\mathbb{M}$ denotes a countable product of copies of the unit interval. The auxiliary marks in $[0, 1] \times \mathbb{M}$ are used to determine which simplices are included in the complex, but they are not part of the vertex labels of the resulting simplicial complex. The full construction can be found in [Section 4.1](#). For $(x, a, t, u) \in \mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$, let $\Delta^{(x,a,t,u)}$ denote the simplicial complex obtained from the augmented point process $\Psi + \delta_{(x,a,t,u)}$. In particular, the vertices of $\Delta^{(x,a,t,u)}$ lie in $\mathbb{R}^d \times \mathbb{A}$. For a measurable, bounded set $W \subseteq \mathbb{R}^d$ and a real-valued function f defined for finite simplicial complexes, we define the difference operator

$$\Lambda_{(x,a,t,u)} f(\Delta_W) := f(\Delta_W^{(x,a,t,u)}) - f(\Delta_{W \setminus \{x\}}^{(x,a,t,u)}). \quad (3)$$

By construction, the complex $\Delta_{W \setminus \{x\}}^{(x,a,t,u)}$ is obtained from $\Delta_W^{(x,a,t,u)}$ by removing the additional vertex (x, a) and is therefore a subcomplex. The operator $\Lambda_{(x,a,t,u)} f$ thus quantifies the effect of adding the vertex (x, a) on the value of the functional f . In the main results below, this operator is evaluated at a typical added point of the form $(0, V, 1, U)$, where $V \sim \Theta$ and $U \sim \mathbb{Q}$ are independent random elements, in particular independent of Ψ . Here Θ denotes the mark distribution on $\mathbb{A}$, and $\mathbb{Q}$ is a countable product of the uniform distribution on $[0, 1]$. Due to the translation-invariant situation, the Euclidean component of the added point does not matter, and we choose the origin 0 as usual.

Furthermore, for a simplicial complex K with vertices in $\mathbb{R}^d \times \mathbb{A}$, we denote by $K + t$, $t \in \mathbb{R}^d$, the simplicial complex obtained from K by translating all Euclidean components of vertices of K by t . That is, each simplex $\{(x_0, a_0), \dots, (x_j, a_j)\}$, $j \in \mathbb{N}_0$, is replaced by the translated simplex $\{(x_0 + t, a_0), \dots, (x_j + t, a_j)\}$. Moreover, let $\mathcal{W}$ be the set of all (half-open) cubes in $\mathbb{R}^d$. Finally, when we refer to a sequence of cubes $(W_n)_{n \in \mathbb{N}}$ tending to $\mathbb{R}^d$, briefly denoted by $W_n \rightarrow \mathbb{R}^d$, we mean that for every bounded set $B \subset \mathbb{R}^d$, there exists an $N \in \mathbb{N}$ such that $B \subseteq W_n$ holds for all $n \geq N$.

Definition 2.1 (Weakly stabilizing functionals). Let f be a measurable real-valued function defined for finite simplicial complexes with vertices in $\mathbb{R}^d \times \mathbb{A}$.

- (i) The function f is called translation invariant if $f(K) = f(K + t)$ holds for all $t \in \mathbb{R}^d$ and all finite simplicial complexes K with vertices in $\mathbb{R}^d \times \mathbb{A}$.

- (ii) The function f is called weakly stabilizing if it is translation invariant and there exists a real-valued random variable Z such that for all sequences of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, the convergence

$$\Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_{W_n}) \xrightarrow{\mathbb{P}} Z \quad (4)$$

holds as $n \rightarrow \infty$. We call Z the random variable associated with f .

- (iii) We say f satisfies the moment condition if there is some $\varepsilon > 0$ with

$$\sup_{\substack{W \in \mathcal{W}, \\ \mathbf{0} \in W}} \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W) \right|^{2+\varepsilon} \right] < \infty. \quad (5)$$

Observe that for a weakly stabilizing functional f , the random variable Z associated to f is almost everywhere uniquely determined, which justifies the usage of the definite article. If f also satisfies the moment condition (5), the convergence in (4) even holds in L^2 . Before we address concrete example functionals in Section 6, let us first note that the class of weakly stabilizing functionals satisfying the moment condition is non-empty, as all constant functionals belong to this class. Moreover, the class is closed under taking linear combinations. The moment condition for such a linear combination follows by applying the Minkowski inequality.

The following lemma establishes the existence of an asymptotic variance and forms a crucial step in proving a CLT for weakly stabilizing functionals. It also provides a criterion for the asymptotic variance to be strictly positive, which serves as the starting point for a later analysis.

Lemma 2.2. *Let f be a weakly stabilizing functional that satisfies the moment condition (5). There exists $\sigma^2 \in [0, \infty)$ such that for every sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, the convergence*

$$\frac{\mathbb{V}(f(\Delta_{W_n}))}{|W_n|} \rightarrow \sigma^2 \quad \text{as } n \rightarrow \infty$$

holds. If the random variable Z associated with f from Definition 2.1 satisfies $\mathbb{P}(Z \neq 0) > 0$, then $\sigma^2 > 0$.

An explicit representation of σ^2 is given in Lemma 5.4. The main result is the following CLT for weakly stabilizing functionals. A multivariate version is given in Section 5.

Theorem 2.3 (CLT for weakly stabilizing functionals). *Let f be a weakly stabilizing functional satisfying the moment condition (5). Then for any sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, we have*

$$\frac{f(\Delta_{W_n}) - \mathbb{E}[f(\Delta_{W_n})]}{\sqrt{|W_n|}} \xrightarrow{d} N_{\sigma^2} \quad \text{as } n \rightarrow \infty,$$

where σ^2 is the limit from Lemma 2.2. If $\mathbb{P}(Z \neq 0) > 0$ for the random variable Z associated with f , then the asymptotic variance σ^2 is positive.

A natural question is whether one can formulate concrete conditions on the function f and the connection functions that ensure positivity of the asymptotic variance. To this end, we will derive a corollary that provides a condition under which the asymptotic variance is indeed strictly positive for weakly additive functionals. Here we call a functional f weakly additive if, for a finite simplicial complex K with connected components $Z_1, \dots, Z_m$, the equality

$$f(K) = \sum_{i=1}^m f(Z_i)$$

always holds. This property was formulated for the Betti numbers in Lemma 3.1 (iii). For a finite set of points $(x_1, a_1), \dots, (x_k, a_k) \in \mathbb{R}^d \times \mathbb{A}$, we denote by $\Delta((x_1, a_1), \dots, (x_k, a_k))$ the random simplicial complex we obtain by applying the construction of Δ (in terms of the multi-staged random experiment described in the introduction) to the vertex set $\{(x_1, a_1), \dots, (x_k, a_k)\}$.

Corollary 2.4. *Let f be a function defined on finite simplicial complexes, which is weakly additive, weakly stabilizing, invariant under isomorphism and satisfies the moment condition (5). Assume that there exists a finite connected simplicial complex K on $r+1$ vertices, $r \in \mathbb{N}_0$, with $\dim(K) \leq \alpha$, such that $|f(K) - f(L)| > 0$ for every induced subcomplex L of K on r vertices. Additionally, suppose that one of the following two conditions is fulfilled.*

- (i) *The connection functions are such that $\varphi_j \in (0, 1)$ almost everywhere for each $j \leq \min\{\alpha, \dim(K) + 1\}$.*
(ii) *The simplicial complex K satisfies*

$$\int_{(\mathbb{R}^d)^r} \int_{\mathbb{A}^{r+1}} \mathbb{P}(\Delta((\mathbf{0}, a_0), (x_1, a_1), \dots, (x_r, a_r)) \cong K) \Theta^{r+1}(da) dx > 0,$$

where we abbreviate $da = d(a_0, \dots, a_r)$ and $dx = d(x_1, \dots, x_r)$.

Then the asymptotic variance σ^2 from Theorem 2.3 applied to f is strictly positive.

3. Preliminaries

This section provides some mathematical background used in this paper, beginning with the notion of a Borel space. A Borel space $(\mathbb{A}, \mathcal{T})$ is a measurable space such that there is a bijective measurable map $\phi : \mathbb{A} \rightarrow U$ into some measurable subset U of the unit interval $[0, 1]$ with measurable inverse. We assume that all random elements appearing in this work live on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

3.1. Poisson processes

To introduce Poisson processes on $\mathbb{R}^d \times \mathbb{A}$ let $\mathbb{N}(\mathbb{R}^d \times \mathbb{A})$ be the set of all locally finite counting measures on $\mathbb{R}^d \times \mathbb{A}$. Here a counting measure η on $\mathbb{R}^d \times \mathbb{A}$ is called locally finite, when $\eta(B \times \mathbb{A}) < \infty$ holds for all bounded Borel sets B . We equip $\mathbb{N}(\mathbb{R}^d \times \mathbb{A})$ with the smallest σ -field such that the mappings $\mathbb{N}(\mathbb{R}^d \times \mathbb{A}) \rightarrow [0, \infty]$, $\eta \mapsto \eta(B)$ are measurable for all measurable sets $B \subseteq \mathbb{R}^d \times \mathbb{A}$. A random element Φ in $\mathbb{N}(\mathbb{R}^d \times \mathbb{A})$ is called a Poisson process on $\mathbb{R}^d \times \mathbb{A}$ with intensity measure $\gamma \lambda_d \otimes \Theta$, if the following two properties are satisfied.

- (a) For every measurable set $B \subseteq \mathbb{R}^d \times \mathbb{A}$ the random variable $\Phi(B)$ follows a Poisson distribution with parameter $(\gamma \lambda_d \otimes \Theta)(B)$.
- (b) For pairwise disjoint measurable sets $B_1, \dots, B_m \subseteq \mathbb{R}^d \times \mathbb{A}$, $m \in \mathbb{N}$, the random variables $\Phi(B_1), \dots, \Phi(B_m)$ are stochastically independent.

Theorem 3.6 in [18] ensures that a Poisson process in this setting always exists and its distribution is uniquely determined by γ and Θ (Proposition 3.2 in [18]). Throughout this paper Φ will always be a Poisson process on $\mathbb{R}^d \times \mathbb{A}$ with intensity measure $\gamma \lambda_d \otimes \Theta$. For a measurable set $B \subseteq \mathbb{R}^d$ we will denote by Φ_B the restriction of Φ to $B \times \mathbb{A}$.

3.2. Simplicial complexes and Betti numbers

A simplicial complex is a family of nonempty, finite subsets of a vertex set, which is closed under taking subsets. An element σ of a simplicial complex is called a j -simplex if it has $j + 1$ elements and j is called the dimension of the simplex. We will denote a j -simplex with vertices $v_0, \dots, v_j$ by $[v_0, \dots, v_j]$. Two simplicial complexes K, L with vertex sets V_K, V_L are called isomorphic (denoted by $K \cong L$) if there is a bijective map $f : V_K \rightarrow V_L$ having the property

$$\sigma \in K \iff f(\sigma) \in L.$$

For a vertex v , that is a 0-simplex, of a simplicial complex K and $n \in \mathbb{N}$, we call

$$\deg_n(v, K) := \|\{\sigma \in K \mid v \in \sigma, \dim(\sigma) = n\}\| \quad (6)$$

the n -simplex degree of the vertex v in K . The j -skeleton of a simplicial complex is its restriction

$$S_j(K) := \{\sigma \in K \mid \dim(\sigma) \leq j\}$$

to simplices of dimension not bigger than j and is again a simplicial complex. Furthermore we denote by $f_i(K)$ the number of i -simplices of a simplicial complex K . Throughout this paper we will only consider finite simplicial complexes that are complexes containing only a finite number of simplices. While we will derive a CLT for a broad range of functionals of simplicial complexes, the Betti numbers are of particular interest to us. The Betti numbers are intricate quantities, and their definition involves several preparatory steps. For this purpose, we fix a finite simplicial complex K and denote the field containing only the two elements 0 and 1 by $\mathbb{Z}_2$. For $p \in \mathbb{N}_0$, let $\sigma_1, \dots, \sigma_k$ be the p -simplices of K . We define the p th chain group $C_p(K)$ of K as the vector space over $\mathbb{Z}_2$, for which $\{\sigma_1, \dots, \sigma_k\}$ is a basis, that is

$$C_p(K) := \left\{ \sum_{i=1}^k a_i \sigma_i \mid a_i \in \mathbb{Z}_2 \right\}.$$

Due to the choice of $\mathbb{Z}_2$ as the space of coefficients (which is common in applications), $C_p(K)$ can be identified with the power set of $\{\sigma_1, \dots, \sigma_k\}$. For $p \in \mathbb{N}$ define a linear map $\partial_p : C_p(K) \rightarrow C_{p-1}(K)$ by linear extension of

$$\partial_p[v_0, \dots, v_p] := \sum_{j=0}^p [v_0, \dots, v_{j-1}, v_{j+1}, \dots, v_p].$$

With the additional definitions $C_{-1}(K) := \{0\}$ and $\partial_0 \equiv 0$ we obtain a so-called chain complex

$$\cdots \longrightarrow C_{p+1}(K) \xrightarrow{\partial_{p+1}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K) \xrightarrow{\partial_{p-1}} \cdots \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} \{0\},$$

satisfying $\partial_p \circ \partial_{p+1} \equiv 0$. The last condition is equivalent to

$$B_p(K) := \text{image}(\partial_{p+1}) \subseteq \text{kernel}(\partial_p) =: Z_p(K).$$

The p th Betti number $\beta_p(K)$ of K is defined by

$$\beta_p(K) := \dim(Z_p(K)) - \dim(B_p(K)),$$

which is the dimension of the quotient vector space $Z_p(K)/B_p(K)$. We summarize the most important properties of the Betti numbers in the following lemma.

Lemma 3.1. *Let $p \in \mathbb{N}_0$ and K be a finite simplicial complex whose connected components are given by $Z_1, \dots, Z_m$. Furthermore, let $L \subseteq K$ be a subcomplex of K . Then the following statements hold.*

- (i) $\beta_p(K) = 0$ if $p > \dim(K)$.
- (ii) $\beta_p(K) = \beta_p(S_{p+1}(K))$.
- (iii) $\beta_p(K) = \sum_{i=1}^m \beta_p(Z_i)$.
- (iv) $|\beta_p(K) - \beta_p(L)| \leq (f_p(K) - f_p(L)) + (f_{p+1}(K) - f_{p+1}(L))$.

The first two properties follow immediately from the definition of the Betti numbers. A proof of the third one, which sometimes is called weakly additive, can be found in [19] (Proposition 2.6). Property (iv) of Lemma 3.1 is crucial to obtain a CLT for the Betti numbers, and is proved, for example, in [8] (Lemma 2.2).

4. The RCM for higher-dimensional simplicial complexes

In this section, we provide a precise definition of the random simplicial complex Δ . We also establish variance results related to the difference operator introduced in Section 2.

4.1. Definition of the model

First of all, we fix some natural number $\alpha \in \mathbb{N}$, which will be the maximum dimension of the simplicial complex Δ . Furthermore, for each $j \in \{1, \dots, \alpha\}$, we choose a measurable, symmetric and translation-invariant function $\varphi_j : (\mathbb{R}^d \times \mathbb{A})^{j+1} \rightarrow [0, 1]$, where the latter means that

$$\varphi_j((x_0 + t, a_0), \dots, (x_j + t, a_j)) = \varphi_j((x_0, a_0), \dots, (x_j, a_j)) \quad (7)$$

holds for all $x_0, \dots, x_j, t \in \mathbb{R}^d$ and $a_0, \dots, a_j \in \mathbb{A}$. We call the functions $\varphi_1, \dots, \varphi_\alpha$ the connection functions of the model. For the definition of Δ we need an even bigger space than $\mathbb{R}^d \times \mathbb{A}$, where the additional randomness to decide which simplices are part of the complex comes from. We write

$$\mathbb{M}^{(j)} := [0, 1]^{\mathbb{N}^j} = \{(u_z)_{z \in \mathbb{N}^j} \mid u_z \in [0, 1] \text{ for all } z \in \mathbb{N}^j\}$$

for the space of j -times indexed sequences with members in $[0, 1]$, equipped with the product σ -field of the Borel σ -fields on $[0, 1]$. We consider the space $\mathbb{M} := M^{(2)} \times \dots \times M^{(2\alpha)}$ together with the product σ -field and the probability distribution $\mathbb{Q} := \otimes_{j=1}^\alpha \otimes_{z \in \mathbb{N}^{2j}} \mathcal{U}([0, 1])$, where $\mathcal{U}([0, 1])$ is the uniform distribution on $[0, 1]$. A random element U in $\mathbb{M}$ with distribution $\mathbb{Q}$ is therefore a countable sequence of i.i.d. uniformly distributed random variables. For $u \in \mathbb{M}$, $j \in \{1, \dots, \alpha\}$ and $z = (k_1, l_1, \dots, k_j, l_j) \in \mathbb{N}^{2j}$ we write $u_{k_1, l_1, \dots, k_j, l_j}$ for the component of u associated with z .

Now let Ψ be a Poisson process on $\mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$ with intensity measure $\gamma \lambda_d \otimes \Theta \otimes \mathcal{U}([0, 1]) \otimes \mathbb{Q}$ and Φ its projection on $\mathbb{R}^d \times \mathbb{A}$, which is a Poisson process with intensity measure $\gamma \lambda_d \otimes \Theta$. Observe that, since a countable product of Borel spaces is again a Borel space, Ψ is again a Poisson process on a product of $\mathbb{R}^d$ with some Borel space. For a point $(x, a, t, u) \in \Psi$, the coordinate $t \in [0, 1]$ is used to impose an ordering on the points of the process, while the components of $u \in \mathbb{M}$ will be used to decide which simplices are contained in Δ .

We define the vertex set of Δ as the set of points of Φ . To determine which higher-dimensional simplices belong to the complex Δ , we associate to each potential simplex σ a coordinate u_σ taken from one of the marks $u \in \mathbb{M}$ of the points of Ψ . For this purpose, we introduce a suitable ordering of the points of Ψ . We assign to each point in Ψ a unique pair coordinates $(k, l) \in \mathbb{N}_0^2$ as follows. Consider the partition $(Q_i)_{i \in \mathbb{N}_0}$ of $\mathbb{R}^d$ into cubes $Q_i := [0, 1]^d + z_i$, where $z_0, z_1, \dots$ is a fixed enumeration of $\mathbb{Z}^d$ with $z_0 = \mathbf{0}$ (origin). For a point $(x, a, t, u) \in \Psi$, the index k is determined uniquely by the condition $x \in Q_k$, and l is such that the point is the l th smallest point of Ψ in Q_k with respect to the ordering defined by

$$(x, a, t, u) < (y, b, s, v) \quad \text{if } t < s. \quad (8)$$

With probability one, each point of Ψ is assigned a unique pair of coordinates $(k, l) \in \mathbb{N}_0^2$, which identifies the point within the process.

Fix points $(x_0, a_0, t_0, u^{(0)}), \dots, (x_j, a_j, t_j, u^{(j)}) \in \Psi$, $j \in \{1, \dots, \alpha\}$, with $t_0 < \dots < t_j$ and consider the potential simplex $\sigma := [(x_0, a_0), \dots, (x_j, a_j)]$ of Δ . To this simplex, we assign the coordinate

$$u(\sigma) := u_{k_0, l_0, \dots, k_{j-1}, l_{j-1}},$$

where (k_i, l_i) denotes the coordinates of the point $(x_i, a_i, t_i, u^{(i)})$ within Ψ . We define the rule

$$\sigma \in \Delta : \iff u(\rho) \leq \varphi_{|\rho|-1}(x_\rho) \text{ for all } \rho \subseteq \sigma \text{ with } |\rho| \geq 2,$$

where $\varphi_{|\rho|-1}(x_\rho)$ is the value of $\varphi_{|\rho|-1}$, when inserting the elements of ρ . This decision rule uniquely determines the simplicial complex Δ . Observe that translating the Euclidean components of all points of Ψ by the same vector while keeping their marks fixed results in another simplicial complex (with the same distribution), which, on the level of realizations, need not be isomorphic to the original one. For instance, two points that were connected by an edge before the shift may no longer be connected afterwards.

For $W \subseteq \mathbb{R}^d$, we write Δ_W for the restriction of Δ to simplices with vertices in $W \times \mathbb{A}$, and Ψ_W for the restriction of Ψ to $W \times \mathbb{A} \times [0, 1] \times \mathbb{M}$. Since the connection functions and the distribution of the Poisson process Ψ are translation invariant with respect to the Euclidean component, it follows that for every $t \in \mathbb{R}^d$, the random variables $f(\Delta_W)$ and $f(\Delta_{W+t})$ have the same distribution whenever f is translation invariant in the sense of Definition 2.1.

This construction has another important property, which we briefly discuss and which also motivates its particular form. Whenever W is a union of cubes from the decomposition $(Q_i)_{i \in \mathbb{N}_0}$, the subcomplex Δ_W is fully determined by Ψ_W , as this holds for the coordinates of the points of Ψ_W within Ψ . Consequently, for two disjoint sets W_1, W_2 , which are unions of cubes from the decomposition $(Q_i)_{i \in \mathbb{N}_0}$, the subcomplexes $\Delta_{W_1}, \Delta_{W_2}$ are independent since the processes Ψ_{W_1}, Ψ_{W_2} are independent. This property will be crucial for the derivation of a CLT.

The idea of introducing an additional $[0, 1]$ -component in the mark space to order the points of the process already appears in [13, 15], among other constructions. We refer in particular to the approach developed in Section 5 of [13]. That approach also relies on a decomposition of the underlying space and an implicit introduction of coordinates for the points, analogous to those used here. However, in [13] additional assumptions on the decomposition are imposed which are not required in the present setting. This framework was subsequently adapted in [2] to the case of $\mathbb{R}^d$ using a decomposition into cubes. In all of these works, the objects under consideration are random graphs, which lead to a simpler overall construction and notation compared to the random simplicial complex studied here. In the special case $\alpha = 1$, the present construction reduces to a random graph.

4.2. Difference operator and variance results

In what follows, we denote by $\Delta^{(x,a,t,u)}$, $(x, a, t, u) \in \mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$, the simplicial complex constructed from the point process $\Psi + \delta_{(x,a,t,u)}$ using the previously introduced procedure from Section 4.1. Observe that for $t = 1$, the complex $\Delta^{(x,a,1,u)}$ contains Δ as a subcomplex almost surely, since the added point does not change the coordinates of the other points. Since this no longer holds for general $t \in [0, 1]$, we will work with the difference operator introduced in Section 2 and recall its definition. For a measurable, bounded set $W \subseteq \mathbb{R}^d$ and a real-valued function f defined for finite simplicial complexes, let

$$\Lambda_{(x,a,t,u)} f(\Delta_W) := f(\Delta_W^{(x,a,t,u)}) - f(\Delta_{W \setminus \{x\}}^{(x,a,t,u)}). \quad (9)$$

The complex $\Delta_{W \setminus \{x\}}^{(x,a,t,u)}$ is a subcomplex of $\Delta_W^{(x,a,t,u)}$ since it is obtained from the latter by removing the vertex (x, a) together with all simplices containing it. This is the crucial property of this operator, ensuring that it depends only on the resulting complex and the location of the added point and not on how the ordering is affected. Therefore, it is more suitable for our analysis than the classical difference operator for Poisson functionals

$$D_{(x,a,t,u)} f(\Delta_W) := f(\Delta_W^{(x,a,t,u)}) - f(\Delta_W)$$

from the literature, where one compares the functional before and after adding a point. Here the second complex is not a subcomplex of the first one, because the added point changes the ordering and thus the assignment of the i.i.d. uniformly distributed random variables to the simplices.

The idea of defining such a well-behaved operator as in (9) was already used in [13] and later in [2]. Observe that in principle one could define something like the “space of simplicial complexes” on which f lives, but for this paper this is not necessary and one can think of quantities like $f(\Delta_W)$ or $\Lambda_{(x,a,t,u)} f(\Delta_W)$ as functions of Ψ .

Now consider independent typical marks $V \sim \Theta$ and $U \sim \mathbb{Q}$, also independent of Ψ . Due to the translation-invariant framework introduced in Section 4.1, it follows that for translation-invariant functionals f (cf. Definition 2.1),

$$\Lambda_{(x,V,t,U)} f(\Delta_W) \stackrel{d}{=} \Lambda_{(0,V,t,U)} f(\Delta_{W-x}).$$

Furthermore, the mark t only affects the ordering of the points. Hence, choosing a different value for t may change the simplicial complex on the level of realizations but not its distribution, i.e.

$$\Lambda_{(0,V,t,U)} f(\Delta_W) \stackrel{d}{=} \Lambda_{(0,V,t',U)} f(\Delta_W) \quad \text{for all } t, t' \in [0, 1].$$

A crucial step in deriving CLTs for Poisson functionals is controlling their variance. To this end, the theory of Poisson processes provides variance formulas that utilize classical difference operators. However, it has been shown that these can be transferred to operators like those from (9). In the following, we present two such results that will be applied in later sections. The first one is a generalization of Theorem 5.1 in [13]. To state it, we denote by Ψ_t the restriction of Ψ to $\mathbb{R}^d \times \mathbb{A} \times [0, t) \times \mathbb{M}$.

Lemma 4.1. *Let $W \subseteq \mathbb{R}^d$ be measurable and bounded and f a function defined for finite simplicial complexes with vertices in $\mathbb{R}^d \times \mathbb{A}$, such that $\mathbb{E}[f(\Delta_W)^2] < \infty$. Then*

$$\mathbb{V}(f(\Delta_W)) = \gamma \int_0^1 \int_W \int_{\mathbb{A}} \int_{\mathbb{M}} \mathbb{E} \left[\mathbb{E} \left[\Lambda_{(x,a,t,u)} f(\Delta_W) \mid \Psi_t \right]^2 \right] \mathbb{Q}(du) \Theta(da) dx dt.$$

Proof. Since W is bounded, it intersects only finitely many cubes of the decomposition $(Q_i)_{i \in \mathbb{N}_0}$ used for the construction of Δ . Let Q_W denote the union of those cubes intersecting W . From the construction of the simplicial complex, it follows that the difference operator $D_{(x,a,t,u)} f(\Delta_W)$ vanishes whenever $x \notin Q_W$. Therefore, Theorem B.1 in [13] shows

$$\mathbb{V}(f(\Delta_W)) = \gamma \int_0^1 \int_{Q_W} \int_{\mathbb{A}} \int_{\mathbb{M}} \mathbb{E} \left[\mathbb{E} \left[D_{(x,a,t,u)} f(\Delta_W) \mid \Psi_t \right]^2 \right] \mathbb{Q}(du) \Theta(da) dx dt.$$

By construction, $\Delta_W = \Delta_{W \setminus \{x\}}^{(x,a,t,u)}$, and hence $\mathbb{E} \left[f(\Delta_{W \setminus \{x\}}^{(x,a,t,u)})^2 \right] < \infty$. Observe that inserting (x, a, t, u) can only affect simplices containing at least one point (y, b, s, v) with $s > t$. Since the presence of each simplex is determined by the corresponding mark in $\mathbb{M}$ of its largest point with respect to the ordering (8), the complexes Δ_W and $\Delta_{W \setminus \{x\}}^{(x,a,t,u)}$ have the same conditional distribution given Ψ_t . Consequently,

$$\mathbb{E} \left[f(\Delta_W) \mid \Psi_t \right] = \mathbb{E} \left[f(\Delta_{W \setminus \{x\}}^{(x,a,t,u)}) \mid \Psi_t \right] \quad \mathbb{P}\text{-almost surely.} \quad (10)$$

Applying the Mecke formula to the Poisson process Ψ and the function

$$g((x, a, t, u), \Psi) := \mathbb{1}\{x \in Q_W\} |f(\Delta_W)|,$$

where Δ_W is regarded as a function of Ψ , yields

$$\begin{aligned} \gamma \int_0^1 \int_{Q_W} \int_{\mathbb{A}} \int_{\mathbb{M}} \mathbb{E} \left[|f(\Delta_W^{(x,a,t,u)})| \right] \mathbb{Q}(du) \Theta(da) dx dt \\ = \mathbb{E} \left[|f(\Delta_W)| \Psi(Q_W \times \mathbb{A} \times [0, 1] \times \mathbb{M}) \right] < \infty. \end{aligned}$$

The finiteness follows from the Cauchy-Schwarz inequality. Therefore, for almost all $(x, a, t, u) \in Q_W \times \mathbb{A} \times [0, 1] \times \mathbb{M}$, the expectation $\mathbb{E} \left[|f(\Delta_W^{(x,a,t,u)})| \right]$ is finite, and together with (10) we obtain

$$\begin{aligned} \mathbb{E} \left[D_{(x,a,t,u)} f(\Delta_W) \mid \Psi_t \right] &= \mathbb{E} \left[f(\Delta_W^{(x,a,t,u)}) \mid \Psi_t \right] - \mathbb{E} \left[f(\Delta_W) \mid \Psi_t \right] \\ &= \mathbb{E} \left[f(\Delta_W^{(x,a,t,u)}) \mid \Psi_t \right] - \mathbb{E} \left[f(\Delta_{W \setminus \{x\}}^{(x,a,t,u)}) \mid \Psi_t \right] \\ &= \mathbb{E} \left[\Lambda_{(x,a,t,u)} f(\Delta_W) \mid \Psi_t \right] \end{aligned}$$

almost surely. Since $\Lambda_{(x,a,t,u)} f(\Delta_W)$ vanishes whenever $x \notin W$, the statement follows. $\square$

With the help of Lemma 4.1 we can deduce a Poincaré type inequality (compare Theorem 5.2 in [13]).

Lemma 4.2. *Let $W \subseteq \mathbb{R}^d$ be measurable and bounded and f a function defined for finite simplicial complexes with vertices in $\mathbb{R}^d \times \mathbb{A}$, such that $\mathbb{E} [f(\Delta_W)^2] < \infty$. Then for all $t \in [0, 1]$ it is true that*

$$\mathbb{V}(f(\Delta_W)) \leq \gamma \int_{\mathbb{R}^d} \int_{\mathbb{A}} \int_{\mathbb{M}} \mathbb{E} \left[(\Lambda_{(x,a,t,u)} f(\Delta_W))^2 \right] \mathbb{Q}(du) \Theta(da) dx.$$

Proof. Applying Jensen's inequality for conditional expectations to the integral representation of $\mathbb{V}(f(\Delta_W))$ from Lemma 4.1 yields the desired inequality by observing that the distribution of $\Lambda_{(x,a,t,u)} f(\Delta_W)$ does not depend on t . $\square$

5. Proof of the CLT and auxiliary results

The goal of this section is to prove the CLT (Theorem 2.3). The approach is based on the derivation of Theorem 2.4 in [2], whose proof technique is, in turn, based on that of Theorem 3.1 in [3]. In [2], however, the difference operator used there is constructed slightly differently from the operator in (9), as the underlying Poisson process is first restricted to W , and then the simplicial complex is constructed from this restriction. Theorem 2.4 in [2] is a CLT for weakly stabilizing functionals in the RCM as a random graph in the unmarked stationary case. The proof is based on a general result for normal approximation of random variables, which can be found as Lemma 2.2 in [3] and will be stated here again for completeness. For that, denote by N_{σ^2} , $\sigma^2 \in [0, \infty)$, a normally distributed random variable with mean 0 and variance σ^2 .

Lemma 5.1 (Lemma 2.2 in [3]). *Let $(Y_n)_{n \in \mathbb{N}}$ and $(X_{n,L})_{n,L \in \mathbb{N}}$ be sequences of random variables with mean 0 satisfying the following two properties.*

(i) *For all $L \in \mathbb{N}$ it is true that*

$$\mathbb{V}(X_{n,L}) \rightarrow \sigma_L^2 \in [0, \infty) \quad \text{and} \quad X_{n,L} \xrightarrow{d} N_{\sigma_L^2} \quad \text{as } n \rightarrow \infty.$$

(ii) *We have $\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{V}(Y_n - X_{n,L}) = 0$.*

Then the limit $\sigma^2 := \lim_{L \rightarrow \infty} \sigma_L^2$ exists and

$$\mathbb{V}(Y_n) \rightarrow \sigma^2 \quad \text{and} \quad Y_n \xrightarrow{d} N_{\sigma^2} \quad \text{as } n \rightarrow \infty$$

holds.

Before proving the CLT, we first analyze the class of weakly stabilizing functionals (Definition 2.1). Recall that $V \sim \Theta$ and $U \sim \mathbb{Q}$ are independent random elements, and in particular independent of Ψ . Throughout this section, we will consider the difference operator $\Lambda_{(0,V,1,U)} f(\Delta_W)$, defined in (9), where W is a cube and f a weakly stabilizing functional according to Definition 2.1. From now on, we use the notation $|W| := \lambda_d(W)$ for a measurable set $W \subseteq \mathbb{R}^d$.

Remark 5.2. Let f be a weakly stabilizing functional satisfying the moment condition (5). Then $\mathbb{E}[f(\Delta_W)^2] < \infty$ holds for all $W \in \mathcal{W}$.

Remark 5.2 can be shown by approximating the function f with bounded functions and applying the Poincaré-type inequality from Lemma 4.2. A proof can be found in [17] (Corollary 5.4). The next proposition is analogous to the result in Lemma 2.7 in [2], but we use a different argument.

Proposition 5.3. Let f be a weakly stabilizing functional. Then for each $t \in [0, 1]$ there is a real-valued random variable Z_t such that for all sequences of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ the convergence

$$\Lambda_{(0,V,t,U)} f(\Delta_{W_n}) \xrightarrow{\mathbb{P}} Z_t$$

holds as $n \rightarrow \infty$.

Proof. Let $(W_n)_{n \in \mathbb{N}}$ be a sequence of cubes with $W_n \rightarrow \mathbb{R}^d$ and $t \in [0, 1]$. As a consequence of Lemma 5.6 in [20] we get

$$\mathbb{E} \left[\min \{ |\Lambda_{(0,V,1,U)} f(\Delta_{W_m}) - \Lambda_{(0,V,1,U)} f(\Delta_{W_n})|, 1 \} \right] \rightarrow 0 \quad \text{for } m, n \rightarrow \infty.$$

Since for every choice of $n, m \in \mathbb{N}$ the equality in distribution

$$\left(\Lambda_{(0,V,1,U)} f(\Delta_{W_n}), \Lambda_{(0,V,1,U)} f(\Delta_{W_m}) \right) \stackrel{d}{=} \left(\Lambda_{(0,V,t,U)} f(\Delta_{W_n}), \Lambda_{(0,V,t,U)} f(\Delta_{W_m}) \right)$$

holds, Lemma 5.6 in [20] implies that the sequence $(\Lambda_{(0,V,t,U)} f(\Delta_{W_n}))_{n \in \mathbb{N}}$ converges in probability. Since this is true for every such sequence of cubes, one can show in precisely the same manner as in the proof of Lemma 2.13 in [2], that the limit variables according to different sequences of cubes have to be the same $\mathbb{P}$ -almost surely. $\square$

We proceed with the proof of Lemma 2.2. The argument follows the same basic structure as the proof of Lemma 2.8 in [2], though certain details differ, among other reasons, because the difference operator used in [2] is constructed somewhat differently. For this reason, and to enhance clarity, we present the proof here, including details omitted in [2]. We begin with an auxiliary result that provides an explicit representation of the asymptotic variance.

We decompose the probability space $\Omega = \Omega_1 \times \Omega_2$, where $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$ is the probability space on which the Poisson process Ψ is defined, and $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ is the probability space on which the marks V, U are defined. We denote the expectation with respect to $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$ by $\tilde{\mathbb{E}}$. Finally, let Ψ_t be the restriction of Ψ to $\mathbb{R}^d \times \mathbb{A} \times [0, t] \times \mathbb{M}$.

Lemma 5.4. Let f be a weakly stabilizing functional satisfying the moment condition (5). For every sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, it is true that

$$\frac{\mathbb{V}(f(\Delta_{W_n}))}{|W_n|} \rightarrow \gamma \int_0^1 \mathbb{E}[\tilde{\mathbb{E}}[Z_t | \Psi_t]^2] dt < \infty \quad \text{as } n \rightarrow \infty,$$

where the random variables $Z_t, t \in [0, 1]$, are the random variables appearing in Proposition 5.3.

Proof. We fix $t \in [0, 1]$ and a sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. The moment condition (5), which also applies to the operator $\Lambda_{(0,V,t,U)}$, provides the uniform integrability of the sequence $(|\Lambda_{(0,V,t,U)} f(\Delta_{W_n})|^2)_{n \in \mathbb{N}}$. From Theorem 5.12 in [20], Proposition 5.3, and the moment condition (5), we obtain

$$\mathbb{E} \left[\left| \Lambda_{(0,V,t,U)} f(\Delta_{W_n}) - Z_t \right|^2 \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad (11)$$

and also $\mathbb{E}[Z_t^2] < \infty$. Jensen's inequality for conditional expectations, combined with the dominated convergence theorem, yields

$$\begin{aligned} \int_0^1 \mathbb{E} \left[\tilde{\mathbb{E}} \left[\Lambda_{(0,V,t,U)} f(\Delta_{W_n}) - Z_t \mid \Psi_t \right]^2 \right] dt \\ \leq \int_0^1 \mathbb{E} \left[\left| \Lambda_{(0,V,t,U)} f(\Delta_{W_n}) - Z_t \right|^2 \right] dt \rightarrow 0 \end{aligned}$$

and therefore

$$h(W_n) := \int_0^1 \mathbb{E} \left[\tilde{\mathbb{E}} \left[\Lambda_{(0,V,t,U)} f(\Delta_{W_n}) \mid \Psi_t \right]^2 \right] dt \rightarrow \int_0^1 \mathbb{E}[\tilde{\mathbb{E}}[Z_t | \Psi_t]^2] dt.$$

The integral on the right-hand side is finite and will be denoted by c in the following. Since this convergence holds for every sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, we can find for every $\varepsilon > 0$ a radius $r_\varepsilon > 0$ such that

$$\sup_{\substack{W \in \mathcal{W} \\ B(0, r_\varepsilon) \subseteq W}} |h(W) - c| < \varepsilon. \quad (12)$$

Lemma 4.1 provides for the functional $f(\Delta_{W_n})$

$$\mathbb{V}(f(\Delta_{W_n})) = \gamma \int_{W_n} \int_0^1 \mathbb{E} \left[\tilde{\mathbb{E}} \left[\Lambda_{(x,V,t,U)} f(\Delta_{W_n}) \mid \Psi_t \right]^2 \right] dt dx$$

$$= \gamma \int_{W_n} h(W_n - x) dx.$$

The last equality follows from

$$\mathbb{E} \left[\Lambda_{(\mathbf{0}, V, t, U)} f(\Delta_{W_n - x}) \mid \Psi_t \right] \stackrel{d}{=} \mathbb{E} \left[\Lambda_{(x, V, t, U)} f(\Delta_{W_n}) \mid \Psi_t \right]$$

which is a consequence of the translation-invariant framework, where the distribution of Ψ , the connection functions, and the functional f are invariant under translations with respect to the Euclidean component. For $\varepsilon > 0$, we choose $r_\varepsilon > 0$ according to (12) and obtain

$$\begin{aligned} \frac{\mathbb{V}(f(\Delta_{W_n}))}{|W_n|} &= \frac{\gamma}{|W_n|} \int_{W_n} \mathbb{1}\{d(x, \partial W_n) > r_\varepsilon\} h(W_n - x) dx \\ &\quad + \frac{\gamma}{|W_n|} \int_{W_n} \mathbb{1}\{d(x, \partial W_n) \leq r_\varepsilon\} h(W_n - x) dx. \end{aligned} \quad (13)$$

Jensen's inequality and the moment condition (5) yield, for $x \in W_n$,

$$\begin{aligned} h(W_n - x) &\leq \int_0^1 \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, t, U)} f(\Delta_{W_n - x}) \right|^2 \right] dt \\ &= \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_{W_n - x}) \right|^2 \right] \leq \sup_{\substack{W \in \mathcal{W}, \\ \mathbf{0} \in W}} \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W) \right|^2 \right] < \infty, \end{aligned}$$

where we used that the distribution of $\Lambda_{(\mathbf{0}, V, t, U)} f(\Delta_{W_n - x})$ does not depend on $t \in [0, 1]$. Hence, the second integral in (13) converges to zero as $n \rightarrow \infty$ (compare (38) in [1] for more details). On the other hand, using (12), it follows that

$$\begin{aligned} &\left| \frac{1}{|W_n|} \int_{W_n} \mathbb{1}\{d(x, \partial W_n) > r_\varepsilon\} h(W_n - x) dx - c \right| \\ &\leq \frac{1}{|W_n|} \int_{W_n} \mathbb{1}\{d(x, \partial W_n) > r_\varepsilon\} |h(W_n - x) - c| dx \\ &\quad + \frac{c}{|W_n|} \left| \{x \in W_n \mid d(x, \partial W_n) \leq r_\varepsilon\} \right| \\ &\leq \varepsilon + \frac{c}{|W_n|} \left| \{x \in W_n \mid d(x, \partial W_n) \leq r_\varepsilon\} \right|. \end{aligned}$$

As the second term also asymptotically vanishes as $n \rightarrow \infty$ and $\varepsilon > 0$ was arbitrary, the desired convergence follows. $\square$

Lemma 5.5. Assume that f is a weakly stabilizing functional for which the moment condition (5) is fulfilled. Then the following statements hold.

(i) For every cube $W \in \mathcal{W}$, we have in L^2

$$\mathbb{E} \left[\Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W) \mid \Psi_t \right] \rightarrow \Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W) \quad \text{in } L^2 \text{ as } t \rightarrow 1,$$

(ii) The following convergence holds in L^2

$$\mathbb{E} [Z \mid \Psi_t] \rightarrow Z \quad \text{as } t \rightarrow 1.$$

Before proceeding with the proof, we introduce the following notation. We denote by $T(\eta)$ the simplicial complex constructed from a counting measure η on $\mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$ according to Section 4.1, without formally introducing a space of simplicial complexes. Recall that this construction is based on a decomposition $(Q_i)_{i \in \mathbb{N}_0}$ of $\mathbb{R}^d$ into cubes and $Q_0 = [0, 1]^d$ is the cube containing the origin $\mathbf{0}$.

Proof. We first prove (i). Let ξ_t be the restriction of the Poisson process Ψ to $\mathbb{R}^d \times \mathbb{A} \times [t, 1] \times \mathbb{M}$, and denote the expectation with respect to ξ_t by $\mathbb{E}_t[\cdot]$. Furthermore, write

$$\Lambda_{(\mathbf{0}, V, 1, U)} f(T(\eta)_W) := f\left(T(\eta + \delta_{(\mathbf{0}, V, 1, U)})_W\right) - f\left(T(\eta + \delta_{(\mathbf{0}, V, 1, U)})_{W \setminus \{\mathbf{0}\}}\right),$$

for a counting measure η on $\mathbb{R}^d \times \mathbb{A} \times [0, 1] \times \mathbb{M}$ and let A_t be the event that ξ_t has no points in $(W \cup Q_0) \times \mathbb{A} \times [t, 1] \times \mathbb{M}$. Observe that $\Lambda_{(\mathbf{0}, V, 1, U)} f(T(\Psi)_W) = \Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W)$ and that A_t is independent of Ψ_t . On A_t , we have

$$T(\Psi_t + \xi_t + \delta_{(\mathbf{0}, V, 1, U)})_W = T(\Psi_t + \delta_{(\mathbf{0}, V, 1, U)})_W \quad \mathbb{P}\text{-almost surely}$$

and hence

$$\Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_W) = \Lambda_{(\mathbf{0}, V, 1, U)} f(T(\Psi_t)_W) \quad \mathbb{P}\text{-almost surely.} \quad (14)$$

The conditional expectation appearing in (i) can be expressed as

$$\mathbb{P}(A_t) \Lambda_{(\mathbf{0}, V, 1, U)} f(T(\Psi_t)_W) + \mathbb{E}_t \left[\Lambda_{(\mathbf{0}, V, 1, U)} f(T(\Psi_t + \xi_t)_W) \mathbb{1}_{A_t^c} \right],$$

and thus, we obtain

$$\begin{aligned} & \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_W) | \Psi_t] - \Lambda_{(0,V,1,U)}f(T(\Psi_t)_W) \\ &= (\mathbb{P}(A_t) - 1) \Lambda_{(0,V,1,U)}f(T(\Psi_t)_W) + \mathbb{E}_t[\Lambda_{(0,V,1,U)}f(T(\Psi_t + \xi_t)_W) \mathbb{1}_{A_t^c}]. \end{aligned} \quad (15)$$

Due to (14) and $\mathbb{P}(A_t) \rightarrow 1$ as $t \rightarrow 1$, the convergence

$$\Lambda_{(0,V,1,U)}f(T(\Psi_t)_W) \xrightarrow{\mathbb{P}} \Lambda_{(0,V,1,U)}f(\Delta_W) \quad \text{as } t \rightarrow 1$$

holds and therefore, the first term in (15) converges to 0 in probability. Applying Hölder's inequality (to the expectation with respect to ξ_t), we can also deduce that the second term in (15) converges to 0 in probability. Thus, the convergence in (i) holds in probability. Together with the moment condition (5), this implies convergence in L^2 .

We now turn to (ii). First, let a sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ be given. Using Minkowski's inequality and the moment condition (5), we have

$$\sup_{n \in \mathbb{N}} \left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] - \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) \right\|_{L^2} < \infty. \quad (16)$$

Additionally, Jensen's inequality for conditional expectations yields

$$\left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] - \tilde{\mathbb{E}}[Z | \Psi_t] \right\|_{L^2} \leq \left\| \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) - Z \right\|_{L^2} \rightarrow 0$$

as $n \rightarrow \infty$. This gives us

$$\begin{aligned} \left\| \tilde{\mathbb{E}}[Z | \Psi_t] - Z \right\|_{L^2} &\leq \limsup_{n \rightarrow \infty} \left\| \tilde{\mathbb{E}}[Z | \Psi_t] - \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] \right\|_{L^2} \\ &\quad + \limsup_{n \rightarrow \infty} \left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] - \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) \right\|_{L^2} \\ &\quad + \limsup_{n \rightarrow \infty} \left\| \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) - Z \right\|_{L^2} \\ &= \limsup_{n \rightarrow \infty} \left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] - \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) \right\|_{L^2}. \end{aligned}$$

Since the last limit superior is finite by (16), for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_n}) | \Psi_t] - \Lambda_{(0,V,1,U)}f(\Delta_{W_n}) \right\|_{L^2} \\ \leq \left\| \tilde{\mathbb{E}}[\Lambda_{(0,V,1,U)}f(\Delta_{W_N}) | \Psi_t] - \Lambda_{(0,V,1,U)}f(\Delta_{W_N}) \right\|_{L^2} + \varepsilon. \end{aligned}$$

From part (i), it follows that

$$\limsup_{t \rightarrow 1} \left\| \tilde{\mathbb{E}}[Z | \Psi_t] - Z \right\|_{L^2} \leq \varepsilon.$$

Since $\varepsilon > 0$ was chosen arbitrarily, the assertion follows. $\square$

With Lemmas 5.4 and 5.5 at hand, we now turn to the proof of Lemma 2.2.

Proof of Lemma 2.2. The desired convergence of the variance follows directly from Lemma 5.4 by choosing

$$\sigma^2 := \gamma \int_0^1 \mathbb{E}[\tilde{\mathbb{E}}[Z_t | \Psi_t]^2] dt.$$

It remains to show that $\sigma^2 > 0$ whenever $\mathbb{P}(Z \neq 0) > 0$. To this end, we demonstrate the continuity of the integrand in the definition of σ^2 at $t = 1$, that is

$$\mathbb{E}[\tilde{\mathbb{E}}[Z_t | \Psi_t]^2] \rightarrow \mathbb{E}[Z^2] \quad \text{as } t \rightarrow 1. \quad (17)$$

Since $\mathbb{E}[Z^2] > 0$ whenever $\mathbb{P}(Z \neq 0) > 0$, the convergence in (17) implies the positivity of σ^2 . To conclude the proof, it remains to prove (17). For this purpose, consider the event B_t that Ψ has no points in $Q_0 \times \mathbb{A} \times [t, 1] \times \mathbb{M}$. By the construction of the simplicial complex, on the event B_t , we have

$$T(\Psi + \delta_{(0,V,t,U)})_W = T(\Psi + \delta_{(0,V,1,U)})_W \quad \mathbb{P}\text{-almost surely}$$

for every cube $W \in \mathcal{W}$ and therefore, in particular,

$$\Lambda_{(0,V,t,U)}f(\Delta_W) = \Lambda_{(0,V,1,U)}f(\Delta_W) \quad \mathbb{P}\text{-almost surely}.$$

Since the probability of the event B_t converges to 1 as $t \rightarrow 1$, we obtain

$$\Lambda_{(0,V,t,U)}f(\Delta_W) \xrightarrow{\mathbb{P}} \Lambda_{(0,V,1,U)}f(\Delta_W) \quad \text{as } t \rightarrow 1 \quad (18)$$

for every cube $W \in \mathcal{W}$. The moment condition in fact yields convergence in L^2 . Now consider a sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. The triangle inequality implies for each $n \in \mathbb{N}$ that

$$\begin{aligned} \|Z_t - Z\|_{L^2} &\leq \|Z_t - \Lambda_{(0,V,t,U)} f(\Delta_{W_n})\|_{L^2} \\ &\quad + \|\Lambda_{(0,V,t,U)} f(\Delta_{W_n}) - \Lambda_{(0,V,1,U)} f(\Delta_{W_n})\|_{L^2} + \|\Lambda_{(0,V,1,U)} f(\Delta_{W_n}) - Z\|_{L^2}. \end{aligned}$$

The first and last term on the right-hand side vanish, due to (11), as $n \rightarrow \infty$. Using that the convergence in (18) also holds in L^2 , we thus obtain the convergence $Z_t \rightarrow Z$ as $t \rightarrow 1$. Together with Jensen's inequality for conditional expectations, this yields

$$\mathbb{E}[\tilde{\mathbb{E}}[Z_t - Z | \Psi_t]^2] \rightarrow 0 \quad \text{as } t \rightarrow 1.$$

Together with part (ii) of Lemma 5.5, the convergence (17) follows and the proof is completed. $\square$

Using Lemmas 2.2 and 5.1, we prove the CLT for weakly stabilizing functionals stated in Theorem 2.3.

Proof of Theorem 2.3. Fix a sequence $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. The distributions of $f(\Delta_{W+t})$ and $f(\Delta_W)$ are the same for all cubes $W \in \mathcal{W}$ and $t \in \mathbb{R}^d$ due to the translation-invariant framework (f , the connection functions, and the distribution of Ψ are translation invariant with respect to the Euclidean component). Therefore, we may assume that all cubes of the sequence $(W_n)_{n \in \mathbb{N}}$ are centered at the origin. Hence, there exists a sequence $(m_n)_{n \in \mathbb{N}}$ of positive numbers with $m_n \rightarrow \infty$ such that

$$W_n = \left[-\frac{\sqrt[4]{m_n}}{2}, \frac{\sqrt[4]{m_n}}{2} \right]^d,$$

that is, m_n is the volume of W_n . Now fix $L \in \mathbb{N}$ and decompose the space into cubes of side length L , i.e., consider the decomposition $([0, L]^d + L \cdot z_i)_{i \in \mathbb{N}_0}$ with the previously chosen enumeration $z_0, z_1, \dots$ of the lattice $\mathbb{Z}^d$ with $z_0 = \mathbf{0}$. Let $l_n \in \mathbb{N}_0$ and $(G_i)_{i \in \mathbb{N}}$ be an enumeration of the cubes in this decomposition, such that $G_i \subseteq W_n$ holds precisely for $i \leq l_n$. Since L is a natural number, each cube G_i is a union of cubes from the decomposition $(Q_i)_{i \in \mathbb{N}_0}$. We define a double sequence, which will play the role of the double sequence in Lemma 5.1, by

$$X_{n,L} := \frac{1}{\sqrt{m_n}} \sum_{i=1}^{l_n} f(\Delta_{G_i}) - \mathbb{E}[f(\Delta_{G_i})].$$

Recall that, by the construction of Δ introduced in Section 4.1, the subcomplex Δ_{G_i} is almost surely completely determined by the process Ψ_{G_i} and the processes Ψ_{G_i} , $i \in \mathbb{N}$, are independent. Therefore, from the latter representation, it follows immediately that $X_{n,L}$ is a sum of i.i.d. random variables. Moreover, Remark 5.2 yields that these random variables have finite variance. It follows that

$$\mathbb{V}(X_{n,L}) = \frac{l_n}{m_n} \mathbb{V}(f(\Delta_{G_1})) \rightarrow \frac{1}{L^d} \mathbb{V}(f(\Delta_{G_1})) =: \sigma_L^2 \in [0, \infty)$$

as $n \rightarrow \infty$. The convergence $\frac{l_n}{m_n} \rightarrow \frac{1}{L^d}$ follows from the second convergence in (9.15) in [21]. Note that by Lemma 2.2, we have $\sigma^2 = \lim_{L \rightarrow \infty} \sigma_L^2$. Thus, the sequence of random variables $(X_{n,L})$ satisfies a CLT as $n \rightarrow \infty$ (cf. e.g., Theorem 27.1 in [22]), i.e., it holds

$$X_{n,L} \xrightarrow{d} N_{\sigma_L^2} \quad \text{as } n \rightarrow \infty.$$

Thus, the double sequence $(X_{n,L})_{n,L \in \mathbb{N}}$ satisfies the first condition of Lemma 5.1. To prove the second condition, let $Y_n := \frac{1}{\sqrt{m_n}} (f(\Delta_{W_n}) - \mathbb{E}[f(\Delta_{W_n})])$. Then, by the Poincaré inequality (Lemma 4.2), we have

$$\begin{aligned} \mathbb{V}(Y_n - X_{n,L}) &\leq \frac{\gamma}{m_n} \int_{W_n} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \sum_{i=1}^{l_n} \mathbb{1}_{\{x \in G_i\}} \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right)^2 \right] dx \\ &= \frac{\gamma}{m_n} \int_{W_n \setminus (\cup_{i=1}^{l_n} G_i)} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) \right)^2 \right] dx \\ &\quad + \frac{\gamma}{m_n} \sum_{i=1}^{l_n} \int_{G_i} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right)^2 \right] dx. \end{aligned} \quad (19)$$

The expectation in the first summand of (19) is uniformly bounded in n due to the moment condition (5). From the first equation in (9.15) in [21], it follows that the first summand asymptotically vanishes as $n \rightarrow \infty$. The convergence (11) from the proof of Lemma 2.2 provides for $t = 1$

$$\mathbb{E} \left[\left| \Lambda_{(0,V,1,U)} f(\Delta_{W_n}) - Z \right|^2 \right] \rightarrow 0.$$

This convergence holds for any sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. Therefore, for any $\varepsilon > 0$, there exists a radius $r_\varepsilon > 0$ such that it is true that

$$\mathbb{E} \left[\left| \Lambda_{(0,V,1,U)} f(\Delta_W) - \Lambda_{(0,V,1,U)} f(\Delta_V) \right|^2 \right] < \varepsilon \quad (20)$$

for all $W, V \in \mathcal{W}$ with $B(\mathbf{0}, r_\varepsilon) \subset W \cap V$. For a fixed $\varepsilon > 0$, we define for a cube $W \in \mathcal{W}$

$$\begin{aligned} \text{int}_\varepsilon(W) &:= \{x \in W \mid d(x, \partial W) \geq r_\varepsilon\}, \\ \partial_\varepsilon(W) &:= \{x \in W \mid d(x, \partial W) < r_\varepsilon\} = W \setminus \text{int}_\varepsilon(W). \end{aligned}$$

For $L > 2r_\varepsilon$, we have $|\partial_\varepsilon G_i| = L^d - (L - 2r_\varepsilon)^d$, and for sufficiently large L , we get $|\partial_\varepsilon G_i| \leq 2r_\varepsilon d L^{d-1}$. Due to the distributional equality

$$(\Lambda_{(0,V,1,U)} f(\Delta_W), \Lambda_{(0,V,1,U)} f(\Delta_V)) \stackrel{d}{=} (\Lambda_{(x,V,1,U)} f(\Delta_{W+x}), \Lambda_{(x,V,1,U)} f(\Delta_{V+x}))$$

for all cubes $W, V \in \mathcal{W}$ and $x \in \mathbb{R}^d$, it follows from (20)

$$\mathbb{E} \left[\left| \Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right|^2 \right] < \varepsilon \quad \text{for all } x \in \text{int}_\varepsilon(G_i), i \leq l_n.$$

The moment condition (5) ensures the existence of a constant $\rho > 0$ such that

$$\mathbb{E} \left[\left| \Lambda_{(x,V,1,U)} f(\Delta_W) - \Lambda_{(x,V,1,U)} f(\Delta_V) \right|^2 \right] \leq \rho \quad \text{for all } W, V \in \mathcal{W}, x \in \mathbb{R}^d.$$

Thus, for the second summand in (19), using the previous three estimates for sufficiently large L , we obtain

$$\begin{aligned} & \frac{\gamma}{m_n} \sum_{i=1}^{l_n} \int_{G_i} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right)^2 \right] dx \\ & \leq \frac{\gamma}{m_n} \sum_{i=1}^{l_n} \left(\int_{\text{int}_\varepsilon(G_i)} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right)^2 \right] dx \right. \\ & \quad \left. + \int_{\partial_\varepsilon(G_i)} \mathbb{E} \left[\left(\Lambda_{(x,V,1,U)} f(\Delta_{W_n}) - \Lambda_{(x,V,1,U)} f(\Delta_{G_i}) \right)^2 \right] dx \right) \\ & \leq \gamma \frac{l_n}{m_n} \left(\varepsilon (L - 2r_\varepsilon)^d + 2\rho r_\varepsilon d L^{d-1} \right) \\ & \leq \gamma \varepsilon + 2\gamma \rho r_\varepsilon d L^{-1}, \end{aligned}$$

where we used the trivial estimates $l_n L^d \leq m_n$ and $(L - 2r_\varepsilon)^d \leq L^d$ for the last inequality. We conclude that

$$\limsup_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{V}(Y_n - X_{n,L}) \leq \gamma \varepsilon.$$

Since ε was arbitrary, condition (ii) in Lemma 5.1 follows, and thus the desired distributional convergence holds. Lemma 2.2 provides the additional result on the positivity of the asymptotic variance σ^2 . $\square$

Finally, from Theorem 2.3, we also directly obtain a multivariate CLT for weakly stabilizing functionals.

Corollary 5.6 (Multivariate CLT for weakly stabilizing functionals). *Let $m \in \mathbb{N}$ and $g_1, \dots, g_m$ be weakly stabilizing functionals that satisfy the moment condition (5). According to Proposition 5.3, there exist random variables $Z_t^{(i)}$, $i \in \{1, \dots, m\}$, $t \in [0, 1]$, such that*

$$\Lambda_{(0,V,t,U)} g_i(\Delta_{W_n}) \xrightarrow{\mathbb{P}} Z_t^{(i)} \quad \text{as } n \rightarrow \infty$$

for all sequences of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. Then, for all $i, j \in \{1, \dots, m\}$ and any sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$,

$$\lim_{n \rightarrow \infty} \frac{\text{Cov}(g_i(\Delta_{W_n}), g_j(\Delta_{W_n}))}{|W_n|} = \gamma \int_0^1 \mathbb{E}[Z_t^{(i)} | \Psi_t] \mathbb{E}[Z_t^{(j)} | \Psi_t] dt =: \sigma_{i,j} \quad (21)$$

and the covariance matrix $\Sigma := (\sigma_{i,j})_{1 \leq i, j \leq m}$ is positive semi-definite. Let N_Σ be an m -dimensional centered normally distributed random vector with covariance matrix Σ . Then

$$\frac{1}{\sqrt{|W_n|}} \left(g_1(\Delta_{W_n}) - \mathbb{E}[g_1(\Delta_{W_n})], \dots, g_m(\Delta_{W_n}) - \mathbb{E}[g_m(\Delta_{W_n})] \right) \xrightarrow{d} N_\Sigma$$

as $n \rightarrow \infty$.

Proof. The claim follows by applying Theorem 2.3, since a linear combination of weakly stabilizing functionals that satisfy the moment condition (5) is itself weakly stabilizing and fulfills the moment condition. For further details, refer to the proof of Corollary 2.12 in [2]. $\square$

6. Betti numbers and other weakly stabilizing functionals

After deriving a CLT for an abstract class of functionals in the previous section, we will now focus on specific example functionals in this section. We will show that many functionals from the literature on random simplicial complexes or random graphs, such as Betti numbers or the Euler characteristic, satisfy the conditions of Theorem 2.3 under certain integrability conditions. To demonstrate that these functionals are weakly stabilizing, we will use a characterization of weak stabilization that simplifies the definition in two aspects. Using this characterization, we can even establish almost sure convergence of the difference operator $\Lambda_{(0,V,1,U)}$. Throughout

this section, when referring to the added vertex, we write $\mathbf{0}$ instead of $(\mathbf{0}, V)$. However, before we turn to example functionals, we will examine the edge degree $\deg_1(\mathbf{0}, \Delta^0)$ of the origin in $\Delta^0 := \Delta^{(\mathbf{0}, V, 1, U)}$. This will be the key quantity for verifying the moment condition for the functionals, as it will allow us to control the difference operator.

It is well known that the edge degree of a deterministic vertex added to the RCM, even on more general spaces, follows a Poisson distribution. More precisely, adding a deterministic vertex $(\mathbf{0}, a)$ for some fixed $a \in \mathbb{A}$ would lead to a Poisson distribution of the edge degree of the added vertex with parameter

$$\pi(a) := \gamma \int_{\mathbb{R}^d} \int_{\mathbb{A}} \varphi_1((\mathbf{0}, a), (y, b)) \Theta(db) dy. \quad (22)$$

Therefore, by choosing a random mark V of the added vertex, which is independent of the remaining part of the model, we gain a mixed Poisson distribution with mixing parameter $\pi(V)$. A mixed Poisson distribution can be considered as the result of a two-stage random experiment, where in the first sub-experiment the random parameter is drawn, and in the second sub-experiment a Poisson-distributed random variable is generated, with the outcome of the first sub-experiment serving as the parameter. Mixed Poisson distributions also appear in generalized random graphs (see, for example, Section 6.3 in [23]). The precise definition and some stochastic properties of a mixed Poisson distribution can be found in [24], where the crucial statement for our work is Proposition 1. This states, among other things, that the s th moment of a mixed Poisson distribution exists if the s th moment of its mixing parameter exists, i.e.

$$\mathbb{E}[\pi(V)^s] < \infty \implies \mathbb{E}[\deg_1(\mathbf{0}, \Delta^0)^s] < \infty, \quad s \in \mathbb{N}. \quad (23)$$

We will use this statement to verify the moment condition (5) for the considered functionals under suitable moment assumptions on the mixing parameter $\pi(V)$. Note that in the unmarked case, the mixing parameter is constant and therefore all its moments exist, as long as this value is assumed to be finite, which leads to the typical integrability condition. Next, we will formulate the aforementioned characterization of weak stabilization.

Proposition 6.1 (Characterization of weak stabilization). *For a translation-invariant functional f defined on finite simplicial complexes with vertices in $\mathbb{R}^d \times \mathbb{A}$ the following statements are equivalent.*

(i) f is weakly stabilizing.

(ii) The operator $\Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_{W_n})$ converges for all increasing sequences of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ in probability as $n \rightarrow \infty$.

Proof. The argumentation works in exactly the same way like the argumentation in the proof of Lemma 2.13 in [2]. $\square$

If we want to show that a certain functional f is weakly stabilizing, this is simplified in two ways by Proposition 6.1. On the one hand, we only need to consider increasing sequences $(W_n)_{n \in \mathbb{N}}$ of cubes, which in turn provides us with an increasing sequence of simplicial complexes $(\Delta_{W_n}^0)_{n \in \mathbb{N}}$. On the other hand, for each such sequence, it is only necessary to prove the existence of the limit of $\Lambda_{(\mathbf{0}, V, 1, U)} f(\Delta_{W_n})$, and not that these limits agree for all sequences.

6.1. Betti numbers

In this subsection, we aim to prove, using Proposition 6.1, that the Betti numbers defined in Section 3.2 satisfy the conditions of Theorem 2.3 whenever an integrability assumption for the mixing parameter $\pi(V)$ is satisfied.

Proposition 6.2. *Let $p \in \mathbb{N}_0$ such that $\mathbb{E}[\pi(V)^{3(p+1)}] < \infty$ holds. Then the p th Betti number β_p is a weakly stabilizing functional satisfying the moment condition (5).*

Proof. First of all, β_p is clearly translation invariant since it depends only on the combinatorial structure of a simplicial complex. To show that the convergence (4) holds, we fix an increasing sequence $(W_n)_{n \in \mathbb{N}}$ of cubes with $W_n \rightarrow \mathbb{R}^d$ and a realization $\omega \in \Omega$ such that $\deg_1(\mathbf{0}, \Delta^0)$ is finite, which holds with probability one. Then $(\Delta_{W_n}^0)_{n \in \mathbb{N}}$ is an increasing sequence of finite simplicial complexes and Lemma 3.8 in [2] shows that the limit

$$Z := \lim_{n \rightarrow \infty} \Lambda_{(\mathbf{0}, V, 1, U)} \beta_p(\Delta_{W_n}) = \lim_{n \rightarrow \infty} \beta_p(\Delta_{W_n}^0) - \beta_p(\Delta_{W_n \setminus \{\mathbf{0}\}}^0)$$

exists. Hence, Proposition 6.1 implies that β_p is weakly stabilizing. To verify the moment condition (5), we apply Lemma 3.1 (iv) with $K = \Delta_{W_n}^0$, $L = \Delta_{W_n \setminus \{\mathbf{0}\}}^0$ and obtain

$$\begin{aligned} \sup_{\substack{W \in \mathcal{W}, \\ \mathbf{0} \in W}} \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, 1, U)} \beta_p(\Delta_W) \right|^3 \right] &\leq \sup_{\substack{W \in \mathcal{W}, \\ \mathbf{0} \in W}} \mathbb{E} \left[\left| \deg_p(\mathbf{0}, \Delta_W^0) + \deg_{p+1}(\mathbf{0}, \Delta_W^0) \right|^3 \right] \\ &\leq \mathbb{E} \left[\left| \deg_p(\mathbf{0}, \Delta^0) + \deg_{p+1}(\mathbf{0}, \Delta^0) \right|^3 \right] \\ &\leq \mathbb{E} \left[\left| \binom{\deg_1(\mathbf{0}, \Delta^0)}{p} + \binom{\deg_1(\mathbf{0}, \Delta^0)}{p+1} \right|^3 \right] \\ &\leq \mathbb{E} \left[\left| \deg_1(\mathbf{0}, \Delta^0)^p + \deg_1(\mathbf{0}, \Delta^0)^{p+1} \right|^3 \right] < \infty. \end{aligned}$$

For the third inequality, we used the fact that every p -simplex that contains the origin can be identified with a p -element set of neighbors of the origin. Due to (23) and the assumptions, the upper bound is finite. $\square$

For $\alpha = 1$ and $p = 1$, [Proposition 6.2](#) yields a CLT for the first Betti number of the random graph, which has not been established so far, even in the unmarked case. Finally, we aim to provide a law of large numbers for the Betti numbers. For this, we require the integrability condition

$$\text{ess sup}_{a \in \mathbb{A}} \int_{\mathbb{R}^d} \int_{\mathbb{A}} \varphi_1((0, a), (y, b)) \Theta(db) dy < \infty. \quad (24)$$

Note that (24) implies that the mixing parameter $\pi(V)$ is bounded $\mathbb{P}$ -almost surely.

Corollary 6.3. *Suppose that condition (24) holds. Then there exists a constant $b_p \geq 0$ for $p \in \mathbb{N}_0$, such that for all sequences of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$, we have*

$$\frac{\beta_p(\Delta_{W_n})}{|W_n|} \xrightarrow{\mathbb{P}} b_p.$$

Proof. We fix a sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. First, the convergence of $\mathbb{E}[\beta_p(\Delta_{W_n})] |W_n|^{-1}$ follows in the same way as in the proof of Theorem 3.6 in [2], and we set b_p as the limit of this sequence. In the proof of Theorem 3.6 in [2], properties (iii), (iv) from [Lemma 3.1](#) of the Betti numbers were used, as well as the existence of the limit

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}[f_i(\Delta_{W_n})]}{|W_n|}$$

for each $i \in \{0, \dots, \alpha\}$, where $f_i(K)$ denotes the number of i -simplices of a simplicial complex K . The latter is shown for the present situation in Section 7 of [1], where condition (24) is used. Since, according to [Lemma 2.2](#) and [Proposition 6.2](#), the limit of $\mathbb{V}(\beta_p(\Delta_{W_n})) |W_n|^{-1}$ also exists, the convergence in probability follows directly from Chebyshev's inequality. The assumption from [Proposition 6.2](#) is satisfied, as (24) implies the almost sure boundedness of $\pi(V)$. $\square$

6.2. Euler characteristic

Having established that the Betti numbers fall into the class of functionals to which [Theorem 2.3](#) is applicable, in this section we will turn to other example functionals. Besides the Betti numbers, the Euler characteristic is one of the most popular topological quantities when studying random simplicial complexes (see, for example, [9,25]). Recall that $f_i(K)$ denotes the number of i -simplices of a simplicial complex K . For a finite simplicial complex K , the Euler characteristic is defined by

$$\chi(K) := \sum_{i=0}^{\dim(K)} (-1)^i f_i(K).$$

Like the Betti numbers, the Euler characteristic can be defined on much more general topological spaces. The CLT in [16] for the Boolean model can be applied, for example, to the Euler characteristic, which is found there as the zeroth intrinsic volume.

Corollary 6.4. *Assume $\mathbb{E}[\pi(V)^{3\alpha}] < \infty$. Then the Euler characteristic χ is weakly stabilizing and the moment condition (5) holds for χ .*

Proof. To apply [Proposition 6.1](#), we fix an increasing sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. The difference operator from (9) for the Euler characteristic satisfies

$$\Lambda_{(0,V,1,U)} \chi(\Delta_{W_n}) = \sum_{i=0}^{\alpha} (-1)^i \deg_i(0, \Delta_{W_n}^0). \quad (25)$$

Since $\mathbb{E}[\pi(V)] < \infty$ implies $\deg_1(0, \Delta^0) < \infty$, and therefore $\deg_i(0, \Delta^0) < \infty$ for all $i \leq \alpha$, almost surely, the right-hand side of (25) converges almost surely as $n \rightarrow \infty$. Hence, according to [Proposition 6.1](#), χ is weakly stabilizing. Furthermore, (25) yields

$$\begin{aligned} \sup_{\substack{W \in \mathcal{W}, \\ 0 \in W}} \mathbb{E} \left[\left| \Lambda_{(0,V,1,U)} \chi(\Delta_W) \right|^3 \right] &\leq \sum_{i,j,m=0}^{\alpha} \mathbb{E} [\deg_i(0, \Delta^0) \deg_j(0, \Delta^0) \deg_m(0, \Delta^0)] \\ &\leq \sum_{i,j,m=0}^{\alpha} \mathbb{E} [\deg_1(0, \Delta^0)^{i+j+m}]. \end{aligned}$$

Here, for the second inequality, we used $\deg_k(0, \Delta^0) \leq \deg_1(0, \Delta^0)^k$ (compare the proof of [Proposition 6.2](#)). Due to the condition $\mathbb{E}[\pi(V)^{3\alpha}] < \infty$ and (23), all moments appearing in the last expression exist, and the upper bound is therefore finite. $\square$

Since the Euler characteristic can also be expressed as a linear combination of Betti numbers according to the Euler-Poincaré formula

$$\chi(K) = \sum_{p=0}^{\dim(K)} (-1)^p \beta_p(K),$$

it could be directly inferred from [Proposition 6.2](#) that the Euler characteristic is weakly stabilizing and satisfies the moment condition (5). However, this would lead to the stronger condition $\mathbb{E}[\pi(V)^{3(\alpha+1)}] < \infty$. This is because, in the proof of [Proposition 6.2](#), the p th

Betti number is estimated from above using f_p and f_{p+1} . For $p = \alpha$, however, the term f_{p+1} does not appear, which leads to a weaker moment condition.

As mentioned in the introduction, even when using [Corollary 6.4](#), the CLT cannot be applied to the Euler characteristic of the Čech or Vietoris-Rips complex, since their Euler characteristics do not coincide with those of their α -skeletons. As both complexes are uniquely determined by their vertex sets, we may define their Euler characteristics as functions of finite point sets, which in turn yield functionals on finite simplicial complexes through their vertex set. We briefly introduce both complexes and omit the mark space, as it is not needed here and would only complicate the notation. The Čech complex with parameter $r > 0$ on a vertex set X contains a simplex consisting of the points $x_0, \dots, x_n \in X$ if and only if

$$\bigcap_{i=0}^n B(x_i, r) \neq \emptyset,$$

where $B(x, r)$ denotes the ball of radius r centered at x . For the Vietoris-Rips complex, the corresponding condition is

$$\|x_i - x_j\| \leq r \quad \text{for all } i \neq j.$$

For a finite set $X \subset \mathbb{R}^d$, we define $H_C^r(X)$ and $H_{VR}^r(X)$ as the Euler characteristics of the Čech and Vietoris-Rips complexes with parameter $r > 0$ on the vertex set X , respectively. In this way, we obtain functionals on finite simplicial complexes through

$$\chi_C^r(K) := H_C^r(\text{Vert}(K)), \quad \chi_{VR}^r(K) := H_{VR}^r(\text{Vert}(K)),$$

where $\text{Vert}(K)$ denotes the vertex set of K . It follows immediately that χ_C^r and χ_{VR}^r are translation invariant. Since these definitions only depend on the vertex set of the simplicial complex, the choice of α and the connection functions are irrelevant.

Corollary 6.5. *The functionals χ_C^r and χ_{VR}^r are weakly stabilizing and satisfy the moment condition (5).*

Proof. Fix an increasing sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ and observe that

$$\Lambda_{(\mathbf{0}, V, 1, U)} \chi_C^r(\Delta_{W_n}) = \sum_{i=0}^{\infty} (-1)^i \deg_i^C(\mathbf{0}, \Delta_{W_n}^0), \quad (26)$$

where $\deg_i^C(\mathbf{0}, \Delta_{W_n}^0)$ denotes the i -simplex degree of $\mathbf{0}$ in the Čech complex with parameter $r > 0$ on the vertex set $\text{Vert}(\Delta_{W_n}) \cup \{\mathbf{0}\}$. Since in the Čech complex the vertex $\mathbf{0}$ can only be connected to points at distance at most $2r$, the operator in (26) becomes constant as $n \rightarrow \infty$ once $B(\mathbf{0}, 2r) \subset W_n$. Therefore, χ_C^r is weakly stabilizing, and the same argument shows that χ_{VR}^r is weakly stabilizing. To verify the moment condition, let N be the number of vertices of Δ in $B(\mathbf{0}, 2r)$. Thus, we obtain for χ_C^r

$$\begin{aligned} \sup_{\substack{W \in \mathcal{W}, \\ \mathbf{0} \in W}} \mathbb{E} \left[\left| \Lambda_{(\mathbf{0}, V, 1, U)} \chi_C^r(\Delta_W) \right|^3 \right] &\leq \mathbb{E} \left[\left| \sum_{i=0}^{\infty} \deg_i^C(\mathbf{0}, \Delta) \right|^3 \right] \leq \mathbb{E} \left[\left| \sum_{i=0}^{\infty} \binom{N}{i} \right|^3 \right] \\ &\leq \mathbb{E}[8^N] = \exp(7\gamma |B(\mathbf{0}, 2r)|). \end{aligned}$$

The last equality can be computed directly using that N follows a Poisson distribution with parameter $\gamma |B(\mathbf{0}, 2r)|$. The same bound applies to the difference operator of χ_{VR}^r , and therefore both functionals satisfy the moment condition (5). $\square$

6.3. Subcomplex counts and further example functionals

The Euler characteristic considered in the previous section is a linear combination of the functionals f_i (number of i -simplices), i.e., quantities that count specific substructures, in this case simplices of a certain dimension. As further example functionals, we want to consider functionals that count certain (possibly even more complex) substructures within simplicial complexes.

For this purpose, let L be a connected, finite simplicial complex, and let g_L be the functional that assigns to a simplicial complex the number of induced subcomplexes isomorphic to L . An induced subcomplex of a simplicial complex K is the restriction of this complex to simplices with vertices in a given subset I of the vertex set of K , denoted by $K[I]$. Writing $\text{Vert}(K)$ for the vertex set of a simplicial complex K , the functional g_L can be written as

$$g_L(K) := \sum_{I \subseteq \text{Vert}(K)} \mathbb{1}\{K[I] \cong L\}.$$

Since $K[I] \cong L$ can only occur if $|I|$ equals the number of vertices of L , the sum could be restricted to subsets I of this size. However, as this is not necessary and would make the formulas more cumbersome, we will always sum over all subsets in the following. We recover the functional f_i as a special case when we choose L as a simplicial complex consisting only of an i -simplex and all its subsimplices. Furthermore, we denote by $h_L(K)$ the number of connected components of a simplicial complex K isomorphic to L .

Functionals of this type can often be found in the literature. For example, counts of components of a certain form (i.e., isomorphic to a fixed chosen graph) were already considered in [13], while the CLT in [2] is applied to isomorphic subgraph counts, among other things. In [15], functionals counting induced connected subgraphs or connected components of a given size are studied.

Proposition 6.6. *Let L be a connected simplicial complex on $r + 1$ vertices for some $r \in \mathbb{N}_0$.*

(i) *If the integrability condition (24) is satisfied, then g_L is a weakly stabilizing functional that fulfills the moment condition (5).*

(ii) In the case that $\mathbb{E}[\pi(V)^3] < \infty$, h_L is a weakly stabilizing functional that meets the moment condition (5).

Proof. For the proof, let us fix an increasing sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. Furthermore, we write $\Delta^0[\mathbf{0}, I]$ for the restriction of Δ^0 to the set $I \cup \{\mathbf{0}\}$, i.e., to simplices with vertices in this set. Considering the functional g_L , we first note that in the case $r = 0$, the functional coincides with the number of vertices f_0 . In this case, the difference operator is identically equal to 1 once $\mathbf{0} \in W_n$, and the statement is therefore obvious. In the following, let $r \in \mathbb{N}$. The operator $\Lambda_{(\mathbf{0}, V, 1, U)} g_L(\Delta_{W_n})$ counts all induced subcomplexes of $\Delta_{W_n}^0$ isomorphic to L that contain the origin and is thus non-negative. Concretely, the operator satisfies

$$\begin{aligned} \Lambda_{(\mathbf{0}, V, 1, U)} g_L(\Delta_{W_n}) &= \sum_{I \subseteq \text{Vert}(\Delta_{W_n})} \mathbb{1}\{\Delta_{W_n}^0[\mathbf{0}, I] \cong L\} \\ &\xrightarrow{a.s.} \sum_{I \subseteq \text{Vert}(\Delta)} \mathbb{1}\{\Delta^0[\mathbf{0}, I] \cong L\} =: Z_L \quad \text{as } n \rightarrow \infty. \end{aligned}$$

An application of Mecke's formula (Theorem 4.4 in [18]) shows that the expected value of Z_L is equal to

$$\frac{\gamma^r}{r!} \int_{(\mathbb{R}^d)^r} \int_{\mathbb{A}^{r+1}} \mathbb{P}(\Delta((\mathbf{0}, a_0), (x_1, a_1), \dots, (x_r, a_r)) \cong L) \Theta^{r+1}(da) dx.$$

where $da = d(a_0, \dots, a_r)$ and $dx = d(x_1, \dots, x_r)$. We can bound the integrand by $(r+1)! f_L((\mathbf{0}, a_0), (x_1, a_1), \dots, (x_r, a_r))$, where f_L denotes the simplex function of L according to Definition 4.1 in [1]. Therefore, Corollary 7.2 in [1] implies that the expectation of Z_L is finite, where condition (24) is applied. Hence, Z_L is almost everywhere finite and Proposition 6.1 implies that g_L is weakly stabilizing. To prove the validity of the moment condition for the functional g_L , we show $\mathbb{E}[Z_L^3] < \infty$. Initially, we have

$$Z_L^3 = \sum_{I, J, M \subseteq \text{Vert}(\Delta)} \mathbb{1}\{\Delta^0[\mathbf{0}, I], \Delta^0[\mathbf{0}, J], \Delta^0[\mathbf{0}, M] \cong L\}.$$

Since L is connected, we can bound the indicator function appearing here by the indicator function of the event that Δ^0 restricted to $I \cup J \cup M \cup \{\mathbf{0}\}$ is connected. By distinguishing cases based on the cardinality of this set, we can find constants $C(s)$ such that

$$Z_L^3 \leq \sum_{s=r}^{3r} \sum_{I \subseteq \text{Vert}(\Delta), |I|=s} C(s) \mathbb{1}\{\Delta^0[\mathbf{0}, I] \text{ is connected}\}.$$

Clearly a simplicial complex on $s+1$ vertices is connected if and only if it is isomorphic to a connected simplicial complex with vertex set $\{1, \dots, s+1\}$ and there are only finitely many of them. Denoting the set of such simplicial complexes by B_s , we obtain

$$Z_L^3 \leq \sum_{s=r}^{3r} \sum_{I \subseteq \text{Vert}(\Delta), |I|=s} \sum_{K \in B_s} C(s) \mathbb{1}\{\Delta^0[\mathbf{0}, I] \cong K\} = \sum_{s=r}^{3r} \sum_{K \in B_s} C(s) Z_K.$$

Following the same argument used previously for Z_L , we find that the expected values of the random variables Z_K are also finite, proving statement (i).

To describe the difference operator of the functional h_L for the second part of the proposition, let $Z(v, K)$ denote the connected component of a simplicial complex K that contains the vertex $v \in K$. Moreover, for a subset $I \subseteq \text{Vert}(K)$, we denote by $K[I] \stackrel{cc}{\cong} L$ the event that the restriction of K to I is a connected component of K isomorphic to L . Finally, we write $I \leftrightarrow v$ if there is an edge between a vertex of I and v . With this notation, we derive

$$\begin{aligned} \Lambda_{(\mathbf{0}, V, 1, U)} h_L(\Delta_{W_n}) &= \mathbb{1}\{Z(\mathbf{0}, \Delta_{W_n}^0) \cong L\} - \sum_{I \subseteq \text{Vert}(\Delta_{W_n})} \mathbb{1}\{\Delta_{W_n}^0[I] \stackrel{cc}{\cong} L, I \leftrightarrow \mathbf{0}\} \\ &\xrightarrow{a.s.} \mathbb{1}\{Z(\mathbf{0}, \Delta^0) \cong L\} - \sum_{I \subseteq \text{Vert}(\Delta)} \mathbb{1}\{\Delta[I] \stackrel{cc}{\cong} L, I \leftrightarrow \mathbf{0}\} \end{aligned}$$

as $n \rightarrow \infty$. To understand the almost sure finiteness of this limit variable, we observe that the sum involved is bounded above by $\deg_1(\mathbf{0}, \Delta^0)$. Thus, together with the assumption and (23), we also have

$$\mathbb{E}[|\Lambda_{(\mathbf{0}, V, 1, U)} h_L(\Delta_W)|^3] \leq \mathbb{E}[(1 + \deg_1(\mathbf{0}, \Delta^0))^3] < \infty.$$

□

As the final example functional, we consider, for $m \in \mathbb{N}, l \in \mathbb{N}_0$, the number of vertices with m -simplex degree equal to l , denoted by d_m^l . For a finite simplicial complex K , define

$$d_m^l(K) := \sum_{v \in \text{Vert}(K)} \mathbb{1}\{\deg_m(v, K) = l\}$$

with the definition (6) of a simplex degree. This functional for $m = 1$ appears in the literature for random graphs and is examined, for example, in Section 8.4 of [23] for the Preferential Attachment Model. In the RCM, vertex degree distributions are studied in [26], for example.

Corollary 6.7. Let $m \in \mathbb{N}, l \in \mathbb{N}_0$, and assume that $\mathbb{E}[\pi(V)^3] < \infty$. Then d_m^l is a weakly stabilizing functional that satisfies the moment condition (5).

Proof. We again fix an increasing sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ to apply [Proposition 6.1](#). The operator $\Lambda_{(0,V,1,U)} d_m^l(\Delta_{W_n})$ can be written as

$$\begin{aligned} & \mathbb{1}\{\deg_m(\mathbf{0}, \Delta_{W_n}^0) = l\} \\ & + \sum_{v \in \text{Vert}(\Delta_{W_n})} \mathbb{1}\{\deg_m(v, \Delta_{W_n}^0) = l\} - \mathbb{1}\{\deg_m(v, \Delta_{W_n \setminus \{0\}}^0) = l\} \\ & \xrightarrow{a.s.} \mathbb{1}\{\deg_m(\mathbf{0}, \Delta^0) = l\} \\ & + \sum_{v \in \text{Vert}(\Delta)} \mathbb{1}\{\deg_m(v, \Delta^0) = l\} - \mathbb{1}\{\deg_m(v, \Delta_{\mathbb{R}^d \setminus \{0\}}^0) = l\}. \end{aligned}$$

The limit variable is almost surely finite, as the absolute value of the sum involved can be bounded above by $\deg_1(\mathbf{0}, \Delta^0)$. With this reasoning, we also have $|\Lambda_{(0,V,1,U)} d_m^l(\Delta_W)| \leq 1 + \deg_1(\mathbf{0}, \Delta^0)$, from which the desired statement follows together with the assumption and [\(23\)](#). $\square$

Finally, we point out that, although only isomorphism-invariant example functionals were considered in this section, [Theorem 2.3](#) also permits more general functionals.

7. On the positivity of the asymptotic variance

For the functionals introduced in the previous section, [Corollary 2.4](#) gives a criterion for the strict positivity of the asymptotic variance in the CLT ([Theorem 2.3](#)). We now turn to its proof, which is based on the final part of [Theorem 2.3](#).

Proof of Corollary 2.4. We will prove the result for the case in which the second statement holds, as the first statement implies this. This follows because the function in the integral from (ii) is almost everywhere strictly positive in this case. Now, let K be as specified in statement (ii) and $A \subseteq \Omega$ the event that the connected component containing the origin in Δ^0 is isomorphic to K . Using Mecke's formula ([Theorem 4.4](#) in [\[18\]](#)), we find that the expected number of connected components in Δ^0 isomorphic to K that contain the origin, and consequently $\mathbb{P}(A)$, is positive whenever the integral from (ii) is positive.

Now let $\omega \in A$ and $W \in \mathcal{W}$ be a sufficiently large cube such that all the vertices of the connected component of the origin in Δ^0 are contained in W . Since this connected component is isomorphic to K , it is in particular finite, and we can always find such a cube. By restricting the connected component of the origin to all vertices except the origin, we obtain a complex that is isomorphic to an induced subcomplex L of K on r vertices. Due to the last required property of K and the weak additivity of f , it follows that

$$|\Lambda_{(0,V,1,U)} f(\Delta_W)| = |f(K) - f(L)| > 0.$$

Therefore, the random variable Z associated with f (from [Definition 2.1](#)) is not zero on A , and in particular it follows that $\mathbb{P}(Z \neq 0) > 0$, which, according to [Theorem 2.3](#), enforces the positivity of the asymptotic variance σ^2 . $\square$

For the application of [Corollary 2.4](#) to a functional f with the required properties, the first question that arises is whether there exists a connected simplicial complex K with $|f(K) - f(L)| > 0$ for all induced subcomplexes with one fewer vertex, for which the conditions (i),(ii) could be verified in the case $\dim(K) \leq \alpha$. If this is the case, then the asymptotic variance from [Theorem 2.3](#) is always positive according to [Corollary 2.4](#), provided that the connection functions take values only in $(0, 1)$. All functionals from [Section 6](#) are isomorphism-invariant and weakly additive. We want to provide examples for the choice of the complex K from [Corollary 2.4](#), starting with the Betti numbers. For the zeroth Betti number, a simplicial complex consisting of a single vertex can always be chosen. A minimal example (in terms of the number of vertices) for a simplicial complex with a p -dimensional hole for $p \in \mathbb{N}$, i.e., with a positive p th Betti number, is given by the family of all proper subsets of a set with $p + 2$ elements, for example by

$$K_p = \{\sigma \subseteq \{1, \dots, p+2\} : 0 < |\sigma| < p+2\}. \quad (27)$$

Since every simplicial complex with a positive p th Betti number has at least $p + 2$ vertices, $\beta_p(L) = 0$ holds for all induced subcomplexes L of K_p on $p + 1$ vertices. In case $\varphi_{p+1} \equiv 1$, however, the two conditions from [Corollary 2.4](#) are not fulfilled for this simplicial complex. In this situation, we construct a simplicial complex L_p for which Condition (ii) from [Corollary 2.4](#) can be verified. On the vertex set $\{1, \dots, p+3\}$, we consider the p -simplices

$$\sigma_j := \{1, \dots, p+2\} \setminus \{j\}, \quad \rho_j := \{1, \dots, p+1, p+3\} \setminus \{j\}$$

for $j \in \{1, \dots, p+1\}$. Let L_p be the simplicial complex consisting of all simplices σ_j, ρ_j for $j \in \{1, \dots, p+1\}$ and all their subsimplices, i.e.

$$L_p = \{\tau \subseteq \{1, \dots, p+3\} : \exists j \in \{1, \dots, p+1\} \text{ with } \emptyset \neq \tau \subseteq \sigma_j \text{ or } \emptyset \neq \tau \subseteq \rho_j\}.$$

For the p -dimensional simplicial complex L_p , the properties $\beta_p(L_p) = 1$ and $\beta_p(M) = 0$ for all induced subcomplexes M of L_p on $p + 2$ vertices can be verified. Additionally, it does not contain the boundary of a $(p + 1)$ -simplex, meaning all its subsimplices without the $(p + 1)$ -simplex itself, which allows it to be applied when $\varphi_{p+1} \equiv 1$. Another possible choice for this situation is the simplicial complex O_p from [Definition 3.3](#) in [\[27\]](#). Since O_p , however, contains $2p + 2$ vertices and thus, in case $p \geq 2$, more than the simplicial complex L_p , the complex L_p may be better suited for verifying Condition (ii) from [Corollary 2.4](#). In the case $p = 1$, both complexes are isomorphic.

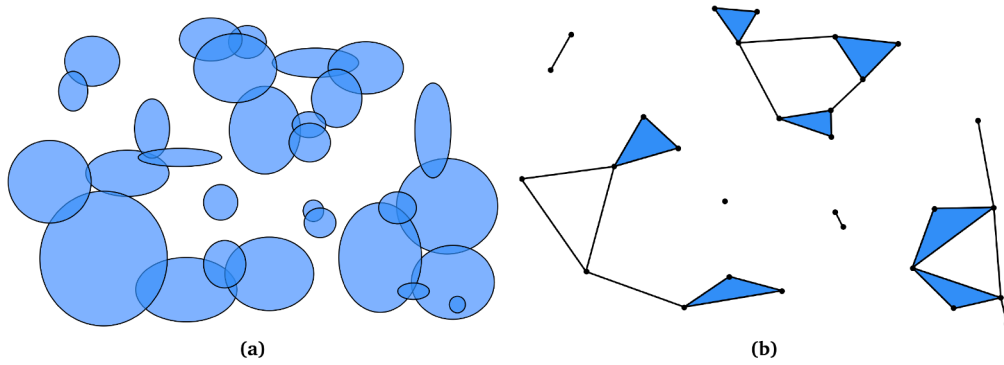


Fig. 2. (a) A union of convex sets and (b) the corresponding homotopy-equivalent nerve.

Considering the Euler characteristic χ , the complex K_1 from (27) can be used since $\chi(K_1) = 0$ and $\chi(M) = 1$ holds for all induced subcomplexes M of K_1 on 2 vertices. For the functionals g_L, h_L , the simplicial complex L satisfies the desired property. To find a suitable simplicial complex for d_m^l (the number of vertices with m -simplex degree l), we ask whether there exists a connected simplicial complex K with $\deg_m(v, K) = l$ for all vertices v of K . Such a simplicial complex would satisfy the desired property. The existence of such a simplicial complex with dimension m is equivalent to the existence of a $1-(v, m+1, l)$ -design (in the sense of Definition 12.1 in [28]) for some $v \in \mathbb{N}$. Such a design exists according to Theorem 12.3 in [28] whenever we can find a $b \in \mathbb{N}$ with

$$m+1 < v, \quad b \leq \binom{v}{m+1}, \quad b(m+1) = vl.$$

This scenario occurs with $v = k(m+1)$ and $b = kl$, provided $k \geq 2$ is sufficiently large. If the associated simplicial complex is not connected, each connected component satisfies the desired property.

8. Application to the Boolean model

Finally, in this section, we want to unravel the connection to the Boolean model that was already hinted at in the introduction, thereby obtaining a CLT for Betti numbers in the Boolean model. To this end, we choose as the marking space the set $\mathbb{A} = \mathcal{K}^d$ of all non-empty, compact and convex subsets of $\mathbb{R}^d$, equipped with the σ -field induced by the Hausdorff metric. According to the Theorems A.19 and A.26 in [18], this is a Borel space. As before, let Θ be a probability measure on $\mathcal{K}^d$. Additionally, we choose the connection functions

$$\varphi_j((x_0, K_0), \dots, (x_j, K_j)) = \mathbb{1} \left\{ \bigcap_{i=0}^j (x_i + K_i) \neq \emptyset \right\}, \quad j \in \{1, \dots, \alpha\}, \quad (28)$$

which are obviously translation invariant in the sense of (7). Since these connection functions only take values in $\{0, 1\}$, we can forgo the additional marking space $\mathbb{M}$ and thus construct the random simplicial complex Δ directly from the Poisson process Φ on $\mathbb{R}^d \times \mathcal{K}^d$. To establish the connection to the Boolean model, we first want to introduce it. In stochastic geometry, the Boolean model (in $\mathbb{R}^d$) corresponding to the Poisson process Φ with intensity measure $\gamma \lambda_d \otimes \Theta$ refers to the random closed set

$$Z := \bigcup_{(x, K) \in \Phi} (x + K)$$

and the individual union sets are called the grains of the model. The probability measure Θ is called the grain distribution and γ the intensity of the Boolean model. For a bounded, measurable set $W \subset \mathbb{R}^d$, define

$$Z^W := \bigcup_{(x, K) \in \Phi_W} (x + K),$$

which is a finite union of convex sets. Be aware of the difference compared to the restriction of Z to W . A variant of the Nerve Theorem (p. 76 in [29]) shows that the set Z^W is homotopy equivalent to the nerve of the individual union sets. The nerve of a finite collection A of nonempty sets is the simplicial complex defined by

$$N(A) := \{\emptyset \neq \sigma \subseteq A : \bigcap_{B \in \sigma} B \neq \emptyset\}.$$

Fig. 2 shows a union of 29 convex sets and the corresponding nerve N with $\beta_0(N) = 6$, $\beta_1(N) = 3$ and $\beta_p(N) = 0$ for all $p \geq 2$.

Note that Δ_W is the α -skeleton of the nerve of the grains $x + K$ with $x \in W$. Consequently, it follows that for a functional f defined on arbitrary topological spaces, which is invariant under homotopy equivalence and depends only on their α -skeletons for simplicial complexes, the equality $f(Z^W) = f(\Delta_W)$ holds almost surely. This is true for the p th Betti number β_p whenever $p \leq \alpha - 1$ due to Lemma 3.1 (ii). Therefore, for $p \leq \alpha - 1$, applying Theorem 2.3 to β_p yields a CLT for $\beta_p(Z^W)$. Since $\alpha \in \mathbb{N}$ can be chosen arbitrarily, the condition $p \leq \alpha - 1$ imposes no restriction at all.

Finally, we take a look at the integrability condition of [Proposition 6.2](#) in the Boolean model. For the k th moment of the mixing parameter $\pi(V)$ from (22), $V \sim \Theta$, we obtain by applying Jensen's inequality

$$\begin{aligned} \mathbb{E}[\pi(V)^k] &= \int_{\mathcal{K}^d} \left(\gamma \int_{\mathbb{R}^d} \int_{\mathcal{K}^d} \mathbb{1}\{K \cap (L + y) \neq \emptyset\} \Theta(dL) dy \right)^k \Theta(dK) \\ &= \int_{\mathcal{K}^d} \left(\gamma \int_{\mathcal{K}^d} |K - L| \Theta(dL) \right)^k \Theta(dK) \\ &\leq \gamma^k \int_{\mathcal{K}^d} \int_{\mathcal{K}^d} |K - L|^k \Theta(dL) \Theta(dK), \end{aligned} \quad (29)$$

where $A - B := \{a - b \mid a \in A, b \in B\}$ denotes the Minkowski difference of two sets $A, B \subseteq \mathbb{R}^d$. [Proposition 6.2](#) states that [Theorem 2.3](#) is applicable for the p th Betti number, whenever the integral in (29) is finite for $k = 3(p + 1)$. We summarize the discussion of this section in the following corollary.

Corollary 8.1. *Let $p \in \mathbb{N}_0$ and consider a Boolean model Z with grain distribution Θ and intensity γ satisfying*

$$\int_{\mathcal{K}^d} \int_{\mathcal{K}^d} |K - L|^{3(p+1)} \Theta(dL) \Theta(dK) < \infty. \quad (30)$$

Then there exists $\sigma^2 \geq 0$ such that for any sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$ it is true that

$$\frac{\beta_p(Z^{W_n}) - \mathbb{E}[\beta_p(Z^{W_n})]}{\sqrt{|W_n|}} \xrightarrow{d} N_{\sigma^2} \quad \text{as } n \rightarrow \infty.$$

Proof. Choose $\alpha = p + 1$, the mark space $\mathbb{A} = \mathcal{K}^d$ and the connection functions from (28). The condition (30) ensures that $\mathbb{E}[\pi(V)^{3(p+1)}] < \infty$ as demonstrated in (29). Hence, [Proposition 6.2](#) shows that the p th Betti number β_p satisfies the requirements of the CLT. [Theorem 2.3](#) therefore yields

$$\frac{\beta_p(\Delta_{W_n}) - \mathbb{E}[\beta_p(\Delta_{W_n})]}{\sqrt{|W_n|}} \xrightarrow{d} N_{\sigma^2} \quad \text{as } n \rightarrow \infty,$$

for any sequence of cubes $(W_n)_{n \in \mathbb{N}}$ with $W_n \rightarrow \mathbb{R}^d$. Since the Nerve Theorem (see, e.g., p. 76 in [29]) states that $\beta_p(Z^{W_n}) = \beta_p(\Delta_{W_n})$ almost surely, the assertion follows. $\square$

The following remark provides a criterion under which condition (30) is always satisfied for every $p \in \mathbb{N}_0$.

Remark 8.2. Assume that the Boolean model has bounded grains, meaning that there exists an $R > 0$ with

$$\Theta(\{K \in \mathcal{K}^d : K \subseteq B(\mathbf{0}, R)\}) = 1,$$

where $B(x, r)$ denotes the closed ball with center x and radius r . In this situation, we obtain almost surely

$$\begin{aligned} \pi(V) &= \gamma \int_{\mathbb{R}^d} \int_{\mathcal{K}^d} \mathbb{1}\{V \cap (y + K) \neq \emptyset\} \Theta(dK) dy \\ &\leq \gamma \int_{\mathbb{R}^d} \int_{\mathcal{K}^d} \mathbb{1}\{B(\mathbf{0}, R) \cap B(y, R) \neq \emptyset\} \Theta(dK) dy \\ &= \gamma(2R)^d |B(\mathbf{0}, 1)|, \end{aligned}$$

which means that the mixing parameter $\pi(V)$ is almost surely bounded and therefore, all its moments exist.

Acknowledgements

The results of this paper stem from the author's PhD thesis [17]. The author is grateful to Daniel Hug, the supervisor of the PhD thesis, for the support and insightful feedback throughout the research. This work was partially supported by the [Deutsche Forschungsgemeinschaft](#) (DFG, German Research Foundation) through the SPP 2265, under grant numbers [HU 1874/5-1](#) and [ME 1361/16-1](#).

References

- [1] D. Pabst, Central limit theorems for the Euler characteristic in the random connection model for higher-dimensional simplicial complexes, 2025, [arXiv:2506.11918](#).
- [2] V.H. Can, K.D. Trinh, Random connection models in the thermodynamic regime: central limit theorems for add-one cost stabilizing functionals, *Electron. J. Probab.* 27 (2022) 1–40.
- [3] K.D. Trinh, On central limit theorems in stochastic geometry for add-one cost stabilizing functionals, *Electron. Commun. Probab.* 24 (76) (2019) 1–15.
- [4] A. Costa, M. Farber, Random simplicial complexes, in: F. Callegaro, F. Cohen, C.D. Concini, E.M. Feichtner, G. Gaiffi, M. Salvetti (Eds.), *Configuration Spaces: Geometry, Topology and Representation Theory*, vol. 14 Springer, 2016, pp. 129–153.
- [5] M. Kahle, Topology of random clique complexes, *Discrete Math.* 309 (6) (2009) 1658–1671.
- [6] C.A.N. Biscio, N. Chenavier, C. Hirsch, A.M. Svane, Testing goodness of fit for point processes via topological data analysis, *Electron. J. Stat.* 14 (1) (2020) 1024–1074.
- [7] M. Kahle, E. Meckes, Limit theorems for Betti numbers of random simplicial complexes, *Homol. Homotopy Appl.* 15 (1) (2013) 343–374.

- [8] D. Yogeshwaran, E. Subag, R.J. Adler, Random geometric complexes in the thermodynamic regime, *Probab. Theory Relat. Fields* 167 (2017) 107–142.
- [9] J.D. Candela, Central limit theorems for soft random simplicial complexes, 2025, [arXiv:2311.1065](#).
- [10] J.D. Candela, Soft random simplicial complexes, 2025, [arXiv:2311.13034](#).
- [11] A. Caicedo, M. Dickson, Critical exponents for marked random connection models, *Electron. J. Probab.* 29 (151) (2024) 1–57.
- [12] M. Dickson, M. Heydenreich, The triangle condition for the marked random connection model, 2022, [arXiv:2210.07727](#).
- [13] G. Last, F. Nestmann, M. Schulte, The random connection model and functions of edge-marked Poisson processes: second order properties and normal approximation, *Ann. Appl. Probab.* 31 (1) (2021) 128–168.
- [14] M. Penrose, On a continuum percolation model, *Adv. Appl. Probab.* 23 (3) (1991) 536–556.
- [15] M. Penrose, Inhomogeneous random graphs, isolated vertices, and poisson approximation, *J. Appl. Probab.* 55 (1) (2018) 112–136.
- [16] D. Hug, G. Last, M. Schulte, Second-order properties and central limit theorems for geometric functionals of Boolean models, *Ann. Appl. Probab.* 26 (1) (2016) 73–135.
- [17] D. Pabst, Das Random Connection Model Für höherdimensionale Simplicialkomplexe, Ph.D. thesis, Karlsruhe Institute of Technology, 2024.
- [18] G. Last, M. Penrose, *Lectures on the Poisson Process*, Cambridge University Press, 2018.
- [19] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.
- [20] O. Kallenberg, *Foundations of Modern Probability*, Vol. 99 of Probability Theory and Stochastic Modelling, Springer, third edition, 2021.
- [21] R. Schneider, W. Weil, *Stochastic and Integral Geometry*, Probability and Its Applications, Springer, 2008.
- [22] P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, 3rd edition, 1995.
- [23] R. van der Hofstad, *Random Graphs and Complex Networks*, Vol. 1 of Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2016.
- [24] M. Kuba, A. Panholzer, On moment sequences and mixed Poisson distributions, *Probab. Surv.* 13 (2016) 89–155.
- [25] A.M. Thomas, T. Owada, Functional limit theorems for the Euler characteristic process in the critical regime, *Adv. Appl. Probab.* 53 (2021) 57–80.
- [26] S.K. Iyer, The random connection model: connectivity, edge lengths, and degree distributions, *Random Struct. Algorithms* 52 (2018) 283–300.
- [27] M. Kahle, Random geometric complexes, *Discrete Comput. Geom.* 45 (3) (2011) 553–573.
- [28] S. Jukna, *Extremal Combinatorics*, Vol. 173 of Texts in Theoretical Computer Science, Springer, second edition, 2011.
- [29] H. Edelsbrunner, *A Short Course in Computational Geometry and Topology*, Springer Briefs in Applied Sciences and Technology, Springer International Publishing AG, 1st edition, 2014.