HAL open science

Simplicial Homology of Random Configurations

Laurent Decreusefond, Eduardo Ferraz, Hugues Randriambololona, Anaïs Vergne

To cite this version:

Laurent Decreusefond, Eduardo Ferraz, Hugues Randriambololona, Anaïs Vergne. Simplicial Homology of Random Configurations. 2011. hal-00578955v2

HAL Id: hal-00578955

https://imt.hal.science/hal-00578955v2

Preprint submitted on 12 Jul 2011 (v2), last revised 4 Jul 2013 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SIMPLICIAL HOMOLOGY OF RANDOM CONFIGURATIONS

L. DECREUSEFOND, E. FERRAZ, H. RANDRIAMBOLOLONA, AND A. VERGNE

Abstract. Given a Poisson process on a d-dimensional torus, its random geometric simplicial complex is the complex whose vertices are the points of the Poisson process and simplices are given by the Čech complex associated to the coverage of each point. By means of Malliavin calculus, we compute explicitly the nth order moment of the number of k-simplices. The two first order moments of this quantity allow us to find the mean and the variance of the Euler caracteristic. Also, we show that the number of any connected geometric simplicial complex converges to the Gaussian law when the intensity of the Poisson point process tends to infinity. We use a concentration inequality to find bounds for the for the distribution of the Betti number of first order and the Euler characteristic in such simplicial complex.

1. Motivation

As technology goes on [13, 14, 17], one can expect a wide expansion of the so-called *sensor networks*. Such networks represent the next evolutionary step in building, utilities, industrial, home, agriculture, defense and many other contexts [4]. These networks are built upon a multitude of small and cheap sensors which are devices with limited transmission capabilities and power. Each sensor monitors a region around itself by measuring some environmental quantities (e.g., temperature, humidity), detecting intrusion, etc, and broadcasts its collected information to other sensors or to a central node. Two questions are of paramount importance: can information be shared among the whole network, is the whole region totally monitored?

Several researches have recently been dedicated to this problem considering a variety of situations. One can distinguish three main scenarios: those where it is possible to choose the position of each sensor, those where sensors are arbitrarily deployed in the target region with the control of a central station and those where the sensor locations are random in a decentralized system. In many cases, placing the sensors is impossible or implies a high cost. Sometimes this impossibility comes from the fact that the cost of placing each sensor is too large and sometimes the network has an inherent random behavior (like in the ad-hoc case, where users move). In addition, this policy cannot take into account the configuration of the network in the case of failure of some sensor. The drawback of the second scenario is a higher cost of sensors, since each one has to communicate with the central station. Besides, the central station itself increases the cost of the whole system. Moreover, if sensors are supposed to know their positions, an absolute positioning system has to be included in each sensor, making their hardware even more complex and then more expensive. It is thus important to investigate the third scenario: randomly located sensors, no central station. Actually, if we can predict some characteristics of the topology of a random network, the number of sensors (or, as well, the power supply of them) can be *a priori* determined such that a given network may operate

2010 *Mathematics Subject Classification.* 60G55,60H07,55U10.

Key words and phrases. Čech complex, Concentration inequality, Homology, point processes, Rips-Vietoris complex.

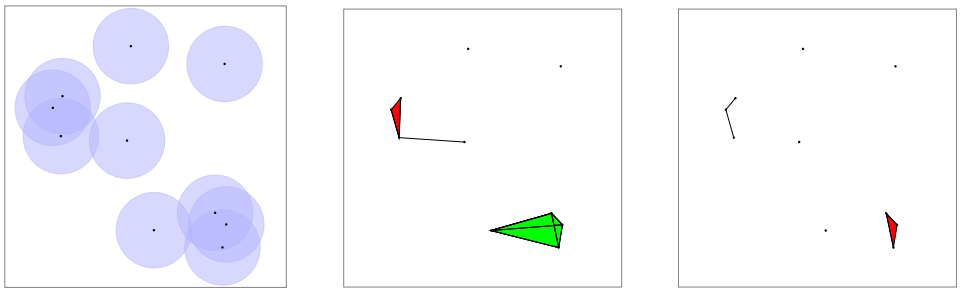

FIGURE 1. From left to right. a) Sensors and their coverage; b) simplicial complex representation when sensors are monitoring the region; c) simplicial complex representation when sensors are communicating among them.

with high probability. For instance, we can choose the mean number of sensors such that, if they are randomly deployed, there is more than 99% of probability the network to be completely connected.

Usually, sensors are deployed in the plane or in the ambient space, thus mathematically speaking, one has to deal with configurations in $\mathbf{R}^2$, $\mathbf{R}^3$ or a manifold. The recent works of Ghrist and his collaborators [8, 6] show how, in any dimension, 0algebraic topology can be used to compute the coverage of a given configuration of sensors. Trying to pursue their work for random settings, we completely solved the problem in one dimension in [7] by basic methods inspired by queuing theory, without using the forthcoming sophisticated tools of algebraic topology. Since we cannot order points in $\mathbf{R}^d$, it is not possible to generalize the results obtained in this earlier work to higher dimension.

A simplicial complex is a generalization of a graph: while we represent a graph with points and edges, a simplicial complex can be represented by points, edges, filled triangles, filled tetrahedrons and so on. It is usual to intepret ϵ as the radius of monitoring but a different interpretation can be given if the sensors are communicating among them. In this case, we suppose that sensors have a power supply allowing them to transmit theirs ID's and, at the same time, sensors have receivers which can identify the transmitted ID's of other sensors above a threshold power. The sensors, knowing mutually the ID's of the close neighbors, are considered connected, creating an information network. The problem remains analogous as the previous one, except that we substitute the coverage radius ϵ by a communication one of $\epsilon/2$. We can see examples of simplicial complexes representations given by sensors communicating among them or monitoring a region in Fig. 1.

In this work we consider that sensors are the atoms of a Poisson point process. Instead of the Euclidean norm, we use the maximum norm along this paper. We consider this for three reasons: this norm represents a superior and an inferior limits for the euclidean norm (we can inscribe and circumscribe a circle with two squares); due to the random interactions with the environment (causing shadowing and fading), even the euclidean norm cannot capture with precision the real behavior of this kind of sensor networks, so we choose the norm that allows us to simplify the calculations; using the maximum norm, the Čech complex becomes equal to the Rips-Vietoris complexes, which allows us to apply directly some results of the algebraic topology. Finally, we assume that sensors lie over d-torus with sides a, $\mathbb{T}^d_a$. This choice was motivated by three factors: it avoids the side effects, it helps to determine weather or not a sensor network in the d-box is completely covered

and if ϵ is small compared to a, the calculations for all parameters in the d-torus are a good approximation for the $[0, a]^d$ box.

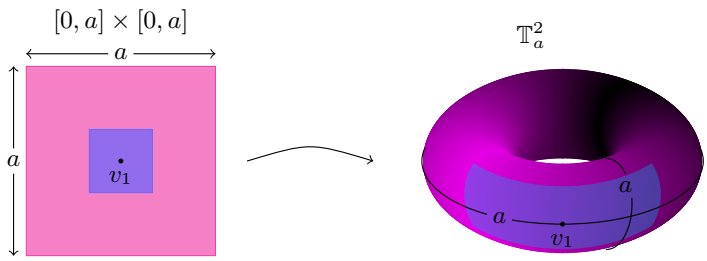

FIGURE 2. Illustration of the coverage of a point and the region where points can lie, in the 2 dimensional case.

When we were writing this paper, we were informed about [12]. Though there are certain similarities between our work and that of Kahle, we would like to point out the differences. In [12], the number of points is fixed and the positions are i.i.d. random variables in the plane. It does not differ very much from a Poisson point process where the number of points is random. However, for our initial motivation, Poisson process fits better since due to failures or movements, we don't know at each time the number of sensors. Moreover, Kahle is only concerned with asymptotic regimes of the mean value of the Betti numbers. We do give exact formulas for any moment of the number of simplices. Using the Morse inequality, we can then deduce the asymptotic regimes for the Betti numbers as in [12]. By using Malliavin calculus, we can evaluate the speed of convergence in the CLT and give a concentration inequality to bound the distribution tail of the first Betti number.

The paper is organized in the following way: Sections 2 and 3 are primers on algebraic topology and Malliavin calculus. In Section 4, the average number of simplices and the mean of the Euler characteristics are computed. Section 5 applies the Malliavin calculus in order to find the explicity expression of second order moments of the number of k-simplices and the Euler characteristic. Using the same solving strategy of the preview section, Section 6 finds the expression for the nth order moment of the number of simplices. Section 7 is devoted to asymptotic analysis: We prove a central limit theorem valid for any finite simplex and we use concentration inequality to estimate the distribution queue of the number of simplices of any order.

2. Algebraic Topology

For further reading on topology, see [10, 1, 15]. Graphs can be generalized to more generic topological objects known as simplicial complexes. While graphs model binary relations, simplicial complexes represent higher order relations. Given a set of points V, a k-simplex is an unordered subset $\{v_0, v_1, \cdots, v_k\}$ where $v_i \in V$ and $v_i \neq v_j$ for all $i \neq j$. The faces of the k-simplex $\{v_0, v_1, \cdots, v_k\}$ are defined as all the $(k-1)$-simplices of the form $\{v_0, \cdots, v_{j-1}, v_{j+1}, \cdots, v_k\}$ with $0 \leq j \leq k$. A simplicial complex is a collection of simplices which is closed with respect to the

inclusion of faces, i.e., if $\{v_0, v_1, \cdots, v_k\}$ is a k-simplex then all its faces are in the set of $(k-1)$-simplices.

Given $\mathcal{U} = (U_v, v \in \mathfrak{T})$ a collection of open sets, the Čech complex of $\mathcal{U}$ denoted by $\mathcal{C}(\mathcal{U})$, is the abstract simplicial complex whose k-simplices correspond to $(k+1)$-tuples of distinct elements of $\mathcal{U}$ that have non empty intersection, so $\{v_0, v_1, \cdots, v_k\}$ is a k-simplex if and only if $\bigcap_{i=0}^{k} U_{v_k} \neq \emptyset$.

One can define an orientation for a simplicial complex by defining an order on vertices. A change in the orientation corresponds to a change in the sign of the coefficient as

$$[v_0, \cdots, v_i, \cdots, v_j, \cdots, v_k] = -[v_0, \cdots, v_j, \cdots, v_i, \cdots, v_k].$$

For each integer k, $C_k(X)$ is the vector space spanned by the set of oriented k-simplices of X. The boundary map ∂_k is defined to be the linear transformations $\partial_k : C_k \to C_{k-1}$ which acts on basis elements $[v_0, \cdots, v_k]$ via

$$\partial_k[v_0, \cdots, v_k] = \sum_{i=0}^{k} (-1)^k [v_0, \cdots, v_{i-1}, v_{i+1}, \cdots, v_k].$$

Examples of such operations are given in Fig. 3.

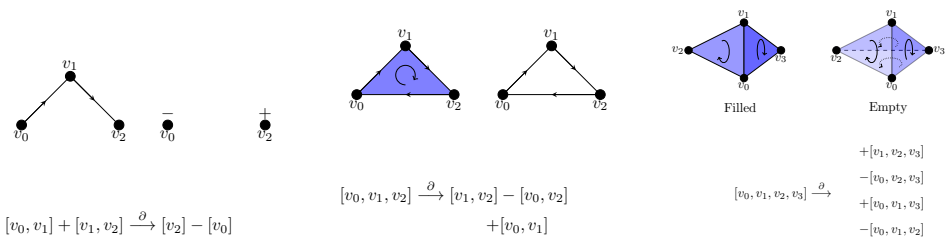

FIGURE 3. Examples of boundary maps. From left to right. An application over 1-simplices. Over a 2-simplex. Over a 3-simplex, turning a filled tetrahedron to an empty one.

This map gives rise to a chain complex: a sequence of vector spaces and linear transformations

$$\cdots \xrightarrow{\partial_{k+2}} C_{k+1}(X) \xrightarrow{\partial_{k+1}} C_k(X) \xrightarrow{\partial_k} C_{k-1}(X) \cdots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X).$$

A standard result then asserts that for any integer k,

$$\partial_k \circ \partial_{k+1} = 0.$$

If one defines

$$Z_k = \ker \partial_k \text{ and } B_k = \mathrm{im}\partial_{k+1},$$

this induces that $B_k \subset Z_k$.

The k-dimensional homology of X, denoted $H_k(X)$ is the quotient vector space,

$$H_k(X) = \frac{Z_k(X)}{B_k(X)}.$$

and the k-th Betti number of X is its dimension:

$$\beta_k = \dim H_k = \dim Z_k - \dim B_k.$$

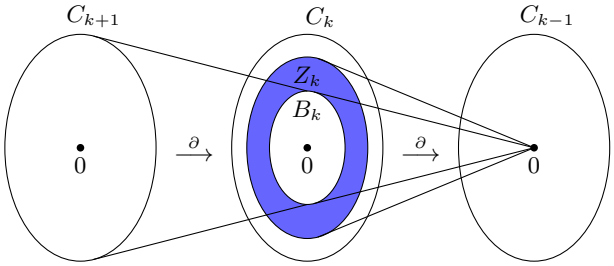

FIGURE 4. A chain complex showing the sets C_k, Z_k and B_k.

Let s_k be the number of k-simplices in a simplicial complex X. The well known topological invariant named Euler characteristic for X, denoted by $\chi(X)$, is an integer defined by:

$$\chi(X) = \sum_{i=0}^{\infty} (-1)^i \beta_i.$$

A well known theorem states that this is also given by:

$$\chi(X) = \sum_{i=0}^{\infty} (-1)^i s_i.$$

The simplicial complexes we consider are of a special type. They can be considered as a generalization of geometric random graphs.

Definition 1. *Given ω a finite set of points on the torus. For $\epsilon > 0$, we define $\mathcal{U}_\epsilon(\omega) = \{B_{d_\infty}(v, \epsilon), v \in \omega\}$ and $\mathcal{C}_\epsilon(\omega) = \mathcal{C}(\mathcal{U}_\epsilon(\omega))$, where $B_{d_\infty}(x, r) = \{y \in \mathbb{T}_d^a, \|x - y\|_\infty < r\}$.*

Theorem 1. *Suppose $\epsilon < a/4$. Then $\mathcal{C}_\epsilon(\omega)$ has the same homotopy type as $\mathcal{U}_\epsilon(\omega)$. In particular they have the same Betti numbers.*

Proof. This will follow from the so-called nerve lemma of Leray, as stated in [19, Theorem 7.26] or [3, Theorem 10.7]. One only needs to check that any non-empty intersection of sets $B_{d_\infty}(v, \epsilon)$ is contractible.

Consider such a non-empty intersection, and let x be a point contained in it. Then, since $\epsilon < a/4$, the ball $B_{d_\infty}(x, 4\epsilon)$ can be identified with a cube in the Euclidean space. Then each $B_{d_\infty}(v, \epsilon)$ containing x is contained in $B_{d_\infty}(x, 4\epsilon)$, hence also becomes a cube with this identification, hence convex. Then the intersection of these convex sets is convex, hence contractible. □

Definition 2. *Let ω be a finite set of points in $\mathbb{T}_d^a$. For any $\epsilon > 0$, the Rips-Vietoris complex of ω, $\mathcal{R}_\epsilon(\omega)$, is the abstract complex whose k-simplices correspond to unordered $(k+1)$-tuples of points in ω which are pairwise within distance less than ϵ of each other.*

Lemma 2. *For the torus $\mathbb{T}_a^d$ equipped with the product distance d_∞, $\mathcal{R}_\epsilon(\omega)$ has the homotopy type of the Čech complex $\mathcal{C}_{2\epsilon}(\omega)$*

The proof is given in [8] in a slightly different context, but it is easy to check that it works here as well. It must be pointed out that Čech and Rips-Vietoris simplicial complexes can be defined similarly for any distance on $\mathbb{T}_a^d$ but it is only for the product distance that the homotopy type of both complexes coincides.

By Lemma 2, k points are forming a $(k-1)$-simplex whenever they are two-by-two closer than 2ϵ from each other. We define along the paper $h(v_1, \cdots, v_k)$

as

$$\begin{aligned} h(v_1, \cdots, v_k) &= h_k(v_1, \cdots, v_k) \\ &= \prod_{1 \le i < j \le k} \mathbf{1}_{[\|v_i - v_j\| < 2\epsilon]}, \end{aligned}$$

which determines if a set of k distinct ordered points generates a $(k-1)$-simplex.

Proposition 3. *Let $\omega \in \mathbb{T}_a^d$ be a set of points, generating the simplicial complex $\mathcal{C}_\epsilon(\omega)$. Then, if $i > d$, $\beta_i(\omega) = 0$.*

Proof. By Theorem 1, $\mathcal{C}_\epsilon(\omega)$ has the same homology as $\mathcal{U}_\epsilon(\omega)$. But $\mathcal{U}_\epsilon(\omega)$ is an open manifold of dimension d, so its Betti numbers $\beta_i(\omega)$ vanish for $i > d$, see for example [9, Theorem 22.24]. □

Proposition 4. *Let $\omega \in \mathbb{T}_a^d$ be a set of points, generating the simplicial complex $\mathcal{C}_\epsilon(\omega)$. There are only two possible values for the d-th Betti number of $\mathcal{C}_\epsilon(\omega)$:*

i) $\beta_d = 0$, or
ii) $\beta_d = 1$.

If the second holds, then we also have $\chi(\mathcal{C}_\epsilon(\omega)) = 0$.

Proof. By Theorem 1, $\mathcal{C}_\epsilon(\omega)$ has the same homology as $\mathcal{U}_\epsilon(\omega)$. Now, $\mathcal{U}_\epsilon(\omega)$ is an open submanifold of the torus, so there are only two possibilities:

i) $\mathcal{U}_\epsilon(\omega)$ is a strict open submanifold, hence non-compact
ii) $\mathcal{U}_\epsilon(\omega) = \mathbb{T}_a^d$.

In the first case, $\beta_d(\omega) = 0$ by [9, Corollary 22.25]. In the second case $\mathcal{C}_\epsilon(\omega)$ has same homology as the torus, hence $\beta_d(\omega) = 1$ and $\chi(\omega) = 0$. □

2.1. **Application to sensor networks.** We now interpret the topological properties of simplicial complexes in terms of connectivity and coverage. In terms of coverage in a network, a 0-simplex represents a single sensor and the existence of a k-simplex means that the $(k+1)$ points of this simplex are covering the convex hull containing those points. We can see in Figure 1, examples of some simplices and their interpretation in terms of sensor networks.

In a very intuitive fashion, the number of k-simplices itself shows some tendency in the network: if in two networks with identical number of sensors, one of them has more 1-simplices than the other, this first one has a tendency to be more connected; by the same reason, if a network has more 2-simplices than another one, the region on the first case tends to be more strongly covered.

In a more sophisticated way, Theorem 1 formalizes that, in order to determine coverage of sensors, it suffices interpret them as Čech complexes. Unfortunately, a moment of thought shows that constructing the Čech complex cannot be done by pairwise only communications between sensors. Thus, the only complex that can be computed this way is the Rips-Vietoris complex.

An interpretation to Euler characteristic is given by Proposition 4, where we see that $\chi = 0$ is a necessary condition to have a complete coverage of the torus, and $\beta_d = 1$ is a necessary and sufficient condition. This could in turn translate into conditions for coverage in $[0, a]^d$ when considered as embedded in Euclidean space (i.e., not as a torus), but then one needs to be careful about border effects. For example, one can say that $\beta_d = 1$ is a sufficient condition for coverage of $[\epsilon, a - \epsilon]^d$.

3. Stochastic Model

3.1. **Poisson point process.** To characterize the randomness of the system, we consider that the set of points is represented by a Poisson point process ω with

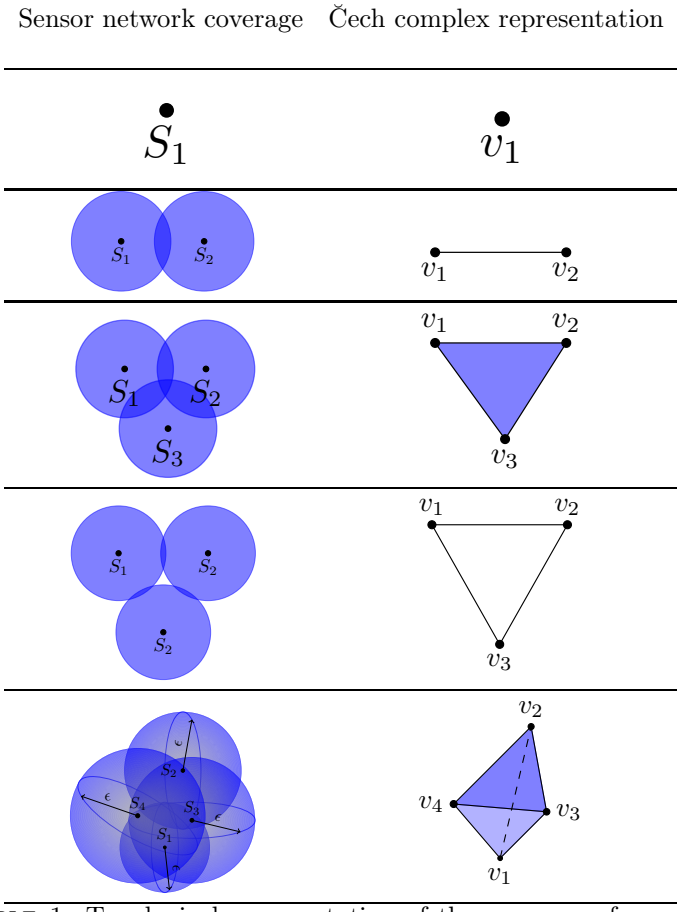

TABLE 1. Topological representation of the coverage of a sensor network. Each node v represents a sensor. From top to bottom, the highest order simplex is a vertex, an edge, a triangle, three edges, a tetrahedron.

intensity λ in a Polish space Y. The space of configurations on Y, is the set of locally finite simple point measures (cf [18]):

$$\Omega^Y = \left\{ \omega = \sum_{k=0}^{n} \delta(x_k) \ : \ (x_k)_{k=0}^{k=n} \subset Y, \ n \in \mathbb{N} \cup \{\infty\} \right\},$$

where $\delta(x)$ denotes the Dirac measure at $x \in Y$. Simple measure means that $\omega(\{x\}) \leq 1$ for any $x \in Y$. Locally finite means that $\omega(K) < \infty$ for any compact K of Y. It is often convenient to identify an element ω of Ω^Y with the set corresponding to its support, i.e., $\sum_{k=0}^{n} \delta(x_k)$ is identified with the unordered set $\{x_1, \cdots, x_n\}$. For $A \in \mathcal{B}(Y)$, we have $\delta(x_k)(A) = \mathbf{1}_{[x_k \in A]}$, so

$$\omega(A) = \sum_{x_k \in \omega} \mathbf{1}_{[x_k \in A]} = \int_A \mathrm{d}\omega(x),$$

counts the number of atoms in A. The configuration space Ω^Y is endowed with the vague topology and its associated σ-algebra denoted by $\mathcal{F}^Y$. Since ω is a Poisson point process of intensity λ:

i) For any A, $\omega(A)$ is a random variable of parameter $\lambda S(A)$, i.e.,

$$\mathbf{P}(\omega(A)=k)=e^{-\lambda S(A)}\frac{(\lambda S(A))^k}{k!}.$$

ii) For $A'\in\mathcal{B}(Y)$, for any disjoints A, A', the random variables $\omega(A)$ and $\omega(A')$ are independent.

Along this paper, we refer $\mathbf{E}_\lambda\left[F(\omega)\right]$ as the mean of some function F depending on ω given that the intensity of this process is λ and $\mathbf{P}_\lambda[\omega\in Y]=\mathbf{E}_\lambda\left[\mathbf{1}_{[\omega\in Y]}\right]$. The definitions of $\mathbf{V}_\lambda\left[F(\omega)\right]$ and $\mathrm{Cov}_\lambda\left[F(\omega),G(\omega)\right]$ are straightforward. Define $\Delta_n=\{(x_1,\cdots,x_n)\in Y^n\mid x_i\neq x_j,\forall i\neq j\}$. Let $f(x_1,\cdots,x_n)$ be a measurable function and let $F(\omega)$ be a random variable given by

$$F(\omega)=\sum_{\substack{x_i\in\omega\cap A,1\leq i\leq n\\ x_i\neq x_j\ \ if\ \ i\neq j}}f(x_1,\,\cdots\,,\,x_n)=\int_{A\cap\Delta_n}f(x_1,\cdots,x_n)\ \mathrm{d}\omega(x_1)\cdots\ \mathrm{d}\omega(x_n),$$

A well known property of the Poisson point processes [5] states that

$$\mathbf{E}_\lambda\left[F(\omega)\right]=\int_A f(x_1,\,\cdots\,,\,x_n)\ \mathrm{d}\lambda(x_1)\cdots\ \mathrm{d}\lambda(x_n).$$

A real function $f\ :\ Y^n\to\mathbb{R}$ is called symmetric if

$$f(x_{\sigma(1)},\,\cdots\,,\,x_{\sigma(n)})=f(x_1,\,\cdots\,,\,x_n)$$

for all permutations σ of $\mathfrak{S}_n$.The space of symmetric square integrable random variables is denoted by $L^2(\lambda)^{\circ n}$. For $f\in L^2(\lambda)^{\circ n}$, the multiple Poisson stochastic integral $I_n(f_n)$ is then defined as

$$I_n(f_n)(\omega)=\int_{\Delta_n}f_n(x_1,\cdots,x_n)(\ \mathrm{d}\omega(x_1)-\ \mathrm{d}\lambda(x_1))\cdots(\ \mathrm{d}\omega(x_n)-\ \mathrm{d}\lambda(x_n)).$$

If $f_n\in L^2(\lambda)^{\circ n}$ and $g_m\in L^2(\lambda)^{\circ m}$, the isometry formula

$$\mathbf{E}_\lambda\left[I_n(f_n)I_m(g_m)\right]=n!\mathbf{1}_{[m=n]}\langle f_n,g_m\rangle_{L^2(\lambda)^{\circ n}} \tag{1}$$

holds true (see [18]). Furthermore, we have:

Theorem 5. *Every random variable $F\in L^2(\Omega^Y,\mathbf{P})$ admits a (unique) Wiener-Poisson decomposition of the type*

$$F=\mathbf{E}_\lambda\left[F\right]+\sum_{n=1}^{\infty}I_n(f_n),$$

where the series converges in $L^2(\mathbf{P})$ and, for each $n\geq 1$, the kernel f_n is an element of $L^2(\lambda)^{\circ n}$. Moreover, we have the isometry

$$\|F-\mathbf{E}_\lambda\left[F\right]\|^2_{L^2(\lambda)^{\circ n}}=\sum_{n=1}^{\infty}n!\|f_n\|^2_{L^2(\mathbb{R}_+)^{\circ n}}. \tag{2}$$

For $f_n\in L^2(\lambda)^{\circ n}$ and $g_m\in L^2(\lambda)^{\circ m}$, we define $f_n\otimes_k^l g_m$, $0\leq l\leq k$, to be the function:

$$(y_{l+1},\cdots,y_n,x_{k+1},\cdots,x_m)\longmapsto \tag{3}$$
$$\int_{Y^l}f_n(y_1,\cdots,y_n)g_m(y_1,\cdots,y_k,x_{k+1},\cdots,x_m)\ \mathrm{d}\lambda(y_1)\ldots\ \mathrm{d}\lambda(y_l).$$

We denote by $f_n\circ_k^l g_m$ the symmetrization in $n+m-k-l$ variables of $f_n\otimes_k^l g_m$, $0\leq l\leq k$. This leads us to the next proposition, shown in [18]:

Proposition 6. *For $f_n \in L^2(\lambda)^{\circ n}$ and $g_m \in L^2(\lambda)^{\circ m}$, we have*

$$I_n(f_n)I_m(g_m) = \sum_{s=0}^{2(n\wedge m)} I_{n+m-s}(h_{n,m,s}),$$

where

$$h_{n,m,s} = \sum_{s\leq 2i\leq 2(s\wedge n\wedge m)} i!\binom{n}{i}\binom{m}{i}\binom{i}{s-i} f_n \circ_i^{s-i} g_m$$

belongs to $L^2(\lambda)^{\circ n+m-s}$, $0 \leq s \leq 2(m\wedge n)$.

In what follows, given $f \in L^2(\lambda)^{\circ q}$ $(q \geq 2)$ and $t \in Y$, we denote by $f(*,t)$ the function on Y^{q-1} given by $(x_1, \cdots, x_{q-1}) \longmapsto f(x_1, \cdots, x_{q-1}, t)$.

Definition 3. *Let $\operatorname{Dom} D$ be the the set of random variables $F \in L^2(P)$ admitting a chaotic decomposition such that*

$$\sum_{n=1}^{\infty} qq!\|f_n\|^2 < \infty.$$

Let D be defined by

$$D \ : \ \operatorname{Dom} D \to L^2(\Omega^Y \times Y, P\times\lambda)$$
$$F = \mathbf{E}_\lambda[F] + \sum_{n\geq 1} I_n(f_n) \longmapsto D_tF = \sum_{n\geq 1} nI_{n-1}(f_n(*,t)).$$

It is known, cf. [11]*, that we also have*

$$D_tF(\omega) = F(\omega\cup\{t\}) - F(\omega),\ dP\times dt \text{ a.e..}$$

Definition 4. *The Ornstein-Uhlenbeck operator L is given by*

$$LF = -\sum_{n=1}^{\infty} nI_n(f_n),$$

whenever $F \in \operatorname{Dom} L$, given by those $F \in L^2P$ such that their chaotic expansion verifies

$$\sum_{n=1}^{\infty} q^2q!\|f_n\|^2 < \infty.$$

Note that $\mathbf{E}_\lambda[LF] = 0$, by definition and (1)*.*

Definition 5. *For $F \in L^2(\mathbf{P})$ such that $\mathbf{E}_\lambda[F] = 0$, we may define L^{-1} by*

$$L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

Combining Stein's method and Malliavin calculus yields the following theorem, see [16]:

Theorem 7. *Let $F \in \operatorname{Dom} D$ be such that $\mathbf{E}_\lambda[F] = 0$ and $Var(F) = 1$. Then,*

$$d_W(F, \mathcal{N}(0,1)) \leq \mathbf{E}_\lambda\left[\left|1 - \int_Y [D_tF \times D_tL^{-1}F]\ d\lambda(t)\right|\right] + \int_Y \mathbf{E}_\lambda\left[|D_tF|^2\left|D_tL^{-1}F\right|\right]\ d\lambda(t).$$

Another result from the Malliavin calculus used in this work is the following one, quoted from [18]:

Theorem 8. *Let $F \in \operatorname{Dom} D$ be such that $DF \leq K$, a.s., for some $K \geq 0$ and $\|DF\|_{L^\infty(\Omega, L^2(Y))} < \infty$. Then*

$$\mathbf{P}(F - \mathbf{E}_\lambda[F] \geq x) \leq \exp\left(-\frac{x}{2K}\log\left(1 + \frac{xK}{\|DF\|_{L^\infty(\Omega,L^2(Y))}}\right)\right). \tag{4}$$

Proposition 9. *Let X a compact subset of $\mathbf{R}^d$ and consider the map $\tau : X \to Y$ as $x_i = ky_i$ for $x_i \in X$, $y_i \in Y$ and k a positive real constant. Denote by $\tau_*\omega$ the image measure of ω by τ, i.e., $\tau_* : \Omega^X \to \Omega^Y$ maps*

$$\omega = \sum_{i=1}^{\infty} \delta(x_i) \quad \text{to} \quad \tau_* \sum_{i=1}^{\infty} \delta(kx_i).$$

The application $\tau_* : \Omega^X \to \Omega^Y$ maps the Poisson measure λ on Ω^X to the Poisson measure $\lambda_\tau = \lambda/k^d$ on Ω^Y. Moreover, if ϵ_τ is the distance in Y such that two points will be connected, the homology of the two simplicial complexes $\mathcal{C}_\epsilon(\omega)_{\omega \in \mathbb{T}^d_{[a]}}$ and $\mathcal{C}_{\epsilon_\tau}(\tau_*\omega)_{\tau_*\omega \in \mathbb{T}^d_{[ak]}}$ are the same for any k if $\lambda_\tau = \lambda/k^d$ and $\epsilon_\tau = k\epsilon$.

Proof. A slightly changing on Propositions 6.1.7 and 6.1.8 of [18] is enough to show that τ_* maps the Poisson measure λ on Ω^X to the Poisson measure $\lambda_\tau = \lambda/k^d$ on Ω^Y. Then, it suffices to realize that for $x_i \in X$ and for $y_i \in Y$:

$$\begin{aligned} h(x_1, \cdots, x_k) &= \prod_{1 \leq i < j \leq k} \mathbf{1}_{[\|x_i - x_j\| < 2\epsilon]} \\ &= \prod_{1 \leq i < j \leq k} \mathbf{1}_{[\|kx_i - kx_j\| < 2k\epsilon]}, \end{aligned}$$

hence

$$h(y_1, \cdots, y_k) = \prod_{1 \leq i < j \leq k} \mathbf{1}_{[\|y_i - y_j\| < 2\epsilon_\tau]},$$

which concludes the proof. □

4. First order moments

Consider that a Poisson point process ω generates a Čech complex $\mathcal{C}_\epsilon(\omega)$, and, even though the number of k-simplices, the Betti's number and the Euler characteristic are functions of $\mathcal{C}_\epsilon(\omega)$, we denote them, respectively, $N_{k+1}(\mathcal{C}_\epsilon(\omega)) = N_{k+1}$, $\beta_k(\mathcal{C}_\epsilon(\omega)) = \beta_k$ and $\chi(\mathcal{C}_\epsilon(\omega)) = \chi$. In this section, we evaluate the mean of the number of $k-1$-simplices, $\mathbf{E}_\lambda[N_k]$ and the mean Euler characteristic, $\mathbf{E}_\lambda[\chi]$.

Theorem 10. *Let $\epsilon \leq a/6$. Then, the mean number of $(k-1)$-simplices $N_k(\mathcal{C}_\epsilon(\omega))$ is given by*

$$\mathbf{E}_\lambda[N_k] = \frac{\lambda^k (ak(2\epsilon)^{k-1})^d}{k!}.$$

Proof. If $(u_{i,1}, \cdots, u_{i,d})$ represents the coordinates of a point v_i, we can separate the indicator function as follows:

$$\mathbf{1}_{[\|v_i - v_j\| < 2\epsilon]} = \prod_{l=1}^{d} \mathbf{1}_{[\{|u_{i,l} - u_{j,l}| < 2\epsilon\} \cup \{|u_{i,l} - u_{j,l}| > a - 2\epsilon\}]}.$$

The number of $(k-1)$-simplices can be counted by the expression:

$$N_k = \frac{1}{k!} \sum_{\substack{v_1, \cdots, v_k \in \omega \\ v_i \neq v_j \ \text{if} \ i \neq j}} h(v_1, \cdots, v_k).$$

Since ω is a Poisson point process of intensity λ, for a borel sets A_i, i integer we have:

$$\mathbf{E}_\lambda\left[\sum_{\substack{v_1,\cdots v_k\in\omega\\ v_i\neq v_j \ \ if \ \ i\neq j}} h(v_1,\cdots,v_k)\right] = \lambda^k\int_{A_1}\cdots\int_{A_k} h(v_1,\cdots,v_k)\ \mathrm{d}v_1\ldots\ \mathrm{d}v_k.$$

Taking $\frac{1}{k!}h = f$, $A_i = \mathbb{T}_a^d$ and defining

$$\{|x_i - x_j| < 2\epsilon\}\cup\{|x_i - x_j| > a - 2\epsilon\} = d_\epsilon(x_i, x_j),$$

we have:

$$\begin{aligned}
\mathbf{E}_\lambda\left[N_k\right] &= \frac{\lambda^k}{k!}\int_{\mathbb{T}_a^d}\cdots\int_{\mathbb{T}_a^d} h(v_1,\cdots,v_k)\ \mathrm{d}v_1\ldots\ \mathrm{d}v_k\\
&= \frac{\lambda^k}{k!}\prod_{l=1}^{d}\int_0^a\cdots\int_0^a\prod_{1\le i<j\le k}\mathbf{1}_{[d_\epsilon(u_{i,l},u_{j,l})]}\ \mathrm{d}u_{1,l}\ldots\ \mathrm{d}u_{k,l}\\
&= \frac{\lambda^k}{k!}\left(\int_0^a\ldots\int_0^a\prod_{1\le i<j\le k}\mathbf{1}_{[d_\epsilon(x_i,x_j)]}\ \mathrm{d}x_1\ldots\ \mathrm{d}x_k\right)^d \qquad (5)\\
&= \frac{\lambda^k}{k!}\Bigg(\int_0^a\underbrace{\int_0^a\mathbf{1}_{[d_\epsilon(x_k,x_{k-1})]}\cdots\int_0^a}_{m\text{ integrals}}\prod_{i=k-m+1}^{k}\mathbf{1}_{[d_\epsilon(x_i,x_{k-m})]}\cdots\\
&\qquad\qquad\int_0^a\prod_{i=2}^{k}\mathbf{1}_{[d_\epsilon(x_i,x_1)]}\ \mathrm{d}x_1\ldots\ \mathrm{d}x_k\Bigg)^d. \qquad (6)
\end{aligned}$$

Since $6\epsilon\le a$, the integration region is convex (see Fig. 5).

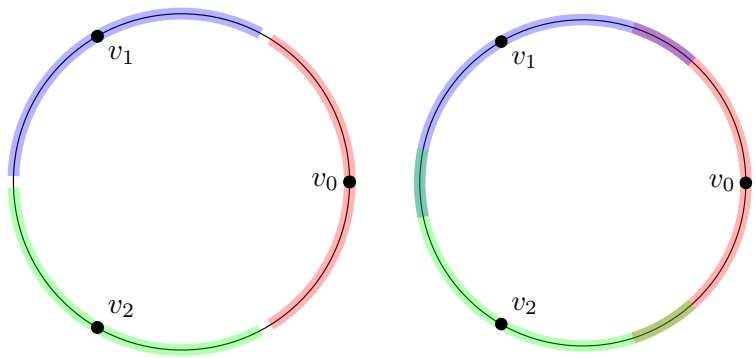

FIGURE 5. a) Maximum cover in $\mathbb{T}_a$ and $\epsilon = a/6$. The red region shows the cover of a point v_0, the blue region is the cover of v_1 and the green region is the cover of v_2. b) Maximum cover in the same conditions of a) when $\epsilon = a/5$. In this case, we the three covers intersect each other pairwise, but there is no intersection of the three covers.

Then, we can rewrite the integral in Eq. (6) as

$$\int_0^a \int_0^a \mathbf{1}_{[d_\epsilon(x_k,x_{k-1})]} \cdots \int_0^a \prod_{i=2}^{k} \mathbf{1}_{[d_\epsilon(x_i,x_1)]} \, \mathrm{d}x_1 \ldots \, \mathrm{d}x_k = \int_0^a \int_{x_k-2\epsilon}^{x_k+2\epsilon} \int_{\max(x_k,x_{k-1})-2\epsilon}^{\min(x_k,x_{k-1})+2\epsilon} \cdots \int_{\max(x_k,\cdots,x_2)-2\epsilon}^{\min(x_k,\cdots,x_2)+2\epsilon} \mathrm{d}x_1 \ldots \, \mathrm{d}x_k. \tag{7}$$

Then, consider a subset of the integration region $[0,a]^d$ of Eq. (7), defined as $A_{1,2,\cdots,k}$, such that $x_1 \geq x_1 \geq \ldots \geq x_k$. In this case, we can write the integral over $A_{1,2,\cdots,k}$ as:

$$\int_0^a \int_{x_k-2\epsilon}^{x_k+2\epsilon} \cdots \int_{\max(x_k,\cdots,x_2)-2\epsilon}^{\min(x_k,\cdots,x_2)+2\epsilon} \mathbf{1}_{[x_i \geq x_j \text{ if } i \leq j]} \, \mathrm{d}x_1 \ldots \, \mathrm{d}x_k = \int_0^a \int_{x_k}^{x_k+2\epsilon} \int_{x_{k-1}}^{x_k+2\epsilon} \cdots \int_{x_2}^{x_k+2\epsilon} \mathrm{d}x_1 \ldots \, \mathrm{d}x_k.$$

For $\sigma \in \mathfrak{S}_k$, we denote by A_σ the set $A_{\sigma(1),\cdots,\sigma(k)}$. Then,

$$\bigcup_{\sigma \in \mathfrak{S}_k} A_\sigma = [0,a]^d.$$

Moreover, since the function $h(x_1, \cdots, x_k)$ is symmetric, we can exchange the integration variables in the integral of Eq. (5) without changing its result. As a consequence, if $\sigma \in \mathfrak{S}_k$,

$$\int_{A_\sigma} \cdots \int \prod_{1 \leq i < j \leq k} \mathbf{1}_{[d_\epsilon(x_i,x_j)]} \, \mathrm{d}x_1 \ldots \, \mathrm{d}x_k = \int_{A_\sigma} \cdots \int \prod_{1 \leq i < j \leq k} \mathbf{1}_{[d_\epsilon(x_i,x_j)]} \, \mathrm{d}x_{i_1} \ldots \, \mathrm{d}x_{i_k}.$$

Thus, we have

$$\begin{aligned}
\int_0^a \cdots \int_0^a \prod_{1 \leq i < j \leq k} \mathbf{1}_{[d_\epsilon(x_i,x_j)]} \, \mathrm{d}x_1 \ldots \, \mathrm{d}x_k &= \sum_{\sigma \in \mathfrak{S}_k} \int_{A_\sigma} \cdots \int \prod_{1 \leq i < j \leq k} \mathbf{1}_{[d_\epsilon(x_i,x_j)]} \, \mathrm{d}x_{i_1} \ldots \, \mathrm{d}x_{i_k} \\
&= k! \int_{A_{\mathrm{Id}}} \cdots \int \prod_{1 \leq i < j \leq k} \mathbf{1}_{[d_\epsilon(x_i,x_j)]} \, \mathrm{d}x_1 \ldots \, \mathrm{d}x_k \\
&= k! \int_0^a \int_{x_k}^{x_k+2\epsilon} \int_{x_{k-1}}^{x_k+2\epsilon} \cdots \int_{x_2}^{x_k+2\epsilon} \mathrm{d}x_1 \ldots \, \mathrm{d}x_k.
\end{aligned}$$

Then, by the change of variables $y_{k-i} = (x_{k-i} - x_k)/2\epsilon$ for $i = \overline{1,k}$, we get:

$$k! \int_0^a \int_{x_k}^{x_k+2\epsilon} \int_{x_{k-1}}^{x_k+2\epsilon} \cdots \int_{x_2}^{x_k+2\epsilon} \mathrm{d}x_1 \ldots \, \mathrm{d}x_k = (2\epsilon)^{k-1} k! \int_0^a \int_0^1 \int_{y_{k-1}}^1 \cdots \int_{y_2}^1 \mathrm{d}y_1 \ldots \, \mathrm{d}y_{k-1} \, \mathrm{d}x_k.$$

The integral in the right-handed term is evaluated substituting all y_i for $y_i' - 1$, so

$$(8) \quad (2\epsilon)^{k-1}k! \int_0^a \int_{-1}^0 \int_{y_{k-1}'}^0 \dots \int_{y_2'}^0 \mathrm{d}y_1' \dots \mathrm{d}y_{k-1}' \, \mathrm{d}x_k = (2\epsilon)^{k-1}k! a \frac{1}{(k-1)!} = a(2\epsilon)^{k-1}k.$$

Finally, plug Eq. (8) into Eq. (6) to obtain:

$$\mathbf{E}_\lambda [N_k] = \frac{\lambda^k (ak(2\epsilon)^{k-1})^d}{k!},$$

and thus the proof is complete. □

Remark. The possibility of writing $\mathbf{E}_\lambda [N_k]$ as Eq. (5) is due the fact that we use the maximum norm. This simplifies the calculations since we can treat each component individually. However, considering the Euclidean norm it is still possible to find a closed-form expression for $\mathbf{E}_\lambda [N_2]$ and $\mathbf{E}_\lambda [N_3]$ when we consider the Rips-Vietoris complex in $\mathbb{T}_a^2$. These expressions, obtained after solving some integrals, are given by:

$$\begin{aligned} \mathbf{E}_\lambda [N_2] &= \frac{\pi(a\lambda\epsilon)^2}{2}, \\ \mathbf{E}_\lambda [N_3] &= \pi \left(\pi - \frac{3\sqrt{3}}{4} \right) \frac{\lambda^3 a^2 \epsilon^4}{6}. \end{aligned}$$

Consider now the Bell's polynomial $B_d(x)$, defined as (see [2])

$$B_n(x) = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} x^k,$$

where n is an positive integer and $\left\{ {n \atop k} \right\}$ is the Stirling number of the second kind. An equivalent definition of B_n can be:

$$B_n(x) = e^{-x} \sum_{k=0}^{\infty} \frac{x^k k^d}{k!}.$$

Theorem 11. *Let $\epsilon \leq a/6$. The mean of the Euler characteristic mean of the simplicial complex $\mathcal{C}_\epsilon(\omega)$ is given by*

$$\mathbf{E}_\lambda [\chi] = \left(\frac{a}{2\epsilon} \right)^d e^{-\lambda(2\epsilon)^d} (-B_d(-\lambda(2\epsilon)^d)).$$

Proof. Since

$$N_k \leq \frac{1}{k!} \prod_{j=0}^{k-1} (N_1 - j) \leq \frac{N_1^k}{k!},$$

then

$$\sum_{k=1}^{\infty} N_k \leq \sum_{k=1}^{\infty} \frac{N_1^k}{k!} = e^{N_1}.$$

As $\mathbf{E}_\lambda\left[e^{N_1}\right] < \infty$, we have $\mathbf{E}_\lambda\left[-\sum_{k=1}^{\infty}(-1)^k N_k\right] = -\sum_{k=1}^{\infty}(-1)^k \mathbf{E}_\lambda\left[N_k\right]$ and

$$\begin{aligned}
\mathbf{E}_\lambda\left[\chi\right] &= \mathbf{E}_\lambda\left[-\sum_{k=1}^{\infty}(-1)^k N_k\right] \\
&= -\sum_{k=1}^{\infty}(-1)^k \mathbf{E}_\lambda\left[N_k\right] \\
&= -\sum_{k=1}^{\infty}(-1)^k \frac{\lambda^k (ak(2\epsilon)^{k-1})^d}{k!} \\
&= \frac{a^d}{-(2\epsilon)^d}\sum_{k=0}^{\infty}\frac{(-\lambda(2\epsilon)^d)^{k+1}(k+1)^d}{(k+1)!} \\
&= \frac{a^d e^{-\lambda(2\epsilon)^d}}{-(2\epsilon)^d} e^{\lambda(2\epsilon)^d}\sum_{k=0}^{\infty}\frac{(-\lambda(2\epsilon)^d)^k k^d}{k!} \\
&= \left(\frac{a}{2\epsilon}\right)^d e^{-\lambda(2\epsilon)^d}(-B_d(-\lambda(2\epsilon)^d)).
\end{aligned}$$

The proof is thus complete. □

If we take $d = 1$, $d = 2$ and $d = 3$, we obtain:

$$\begin{aligned}
\mathbf{E}_\lambda\left[\chi\right](\mathcal{C}_\epsilon(\omega))_{\omega\in\mathbb{T}_{[a]}} &= a\lambda e^{-\lambda 2\epsilon}, \\
\mathbf{E}_\lambda\left[\chi\right](\mathcal{C}_\epsilon(\omega))_{\omega\in\mathbb{T}^2_{[a^2]}} &= a^2\lambda e^{-\lambda(2\epsilon)^2}\left(1-\lambda(2\epsilon)^2\right), \\
\mathbf{E}_\lambda\left[\chi\right](\mathcal{C}_\epsilon(\omega))_{\omega\in\mathbb{T}^3_{[a^3]}} &= a^3\lambda e^{-\lambda(2\epsilon)^3}\left(1-3\lambda(2\epsilon)^3+(\lambda(2\epsilon)^3)^2\right).
\end{aligned}$$

Remark. For c a positive real, $\mathbf{E}_\lambda\left[\chi\right]$ is invariant under the transformation $\lambda' = \lambda/c$, $\epsilon' = c\epsilon$ and $a' = ca$. Taking $c = 1/2\epsilon$, we obtain:

$$\mathbf{E}_\lambda\left[\chi\right] = a'^d e^{-\lambda'}(-B_d(-\lambda')).$$

Hence, the mean depends actually only on a' and λ'.

The following result is well known.

Lemma 12. *If $B_d(x)$ is the Bell's polynomial and for $d \geq 1$, the following relations are valid:*

$$\begin{aligned}
\frac{d}{dx}B_d(x) &= \frac{B_{d+1}(x)}{x} - B_d(x), \\
\frac{d}{dx}(e^x B_d(x)) &= \frac{e^x}{x}B_{d+1}(x).
\end{aligned}$$

According to these relations, it is routine to prove the following theorem.

Theorem 13. *The function $(\lambda \mapsto \mathbf{E}_\lambda\left[\chi\right])$ has exactly d non-negative real roots. Moreover, between each consecutive roots and after the last one, there is exactly one critical point.*

We can see by the expression of $\mathbf{E}_\lambda\left[\chi\right]$ that this quantity tends to 0 as λ tends to infinity. This convergence is due the fact that the Euler characteristic of the Čech complex of the cover tends to the Euler characteristic of the d-Torus where the points are deployed. This is shown in the following theorem.

Theorem 14. *The Betti numbers of $\mathcal{C}(\mathcal{U}_\epsilon)$ converge in probability to the Betti number of the torus as λ goes to infinity:*

$$\mathbf{P}_\lambda\left(\bigcap_{i=0}^{d}\left(\beta_i(\mathcal{C}_\epsilon) = \beta_i(\mathbb{T}^d_{[a]})\right)\right) \xrightarrow{\lambda\to\infty} 1.$$

Proof. Let $\eta < \epsilon/2$, by compacity of the torus, there exists a finite collection of balls $\mathfrak{B}$ of radius η covering $\mathbb{T}^d_{[a]}$. Since $\eta < \epsilon/2$, if x belongs to some ball $B \in \mathfrak{B}$ then $B \subset B(x, \epsilon)$, hence

$$\bigcap_{B\in\mathfrak{B}} (\omega(B) \neq 0) \subset \left(\mathcal{U}_\epsilon(\omega) = \mathbb{T}^d_{[a]}\right).$$

Thus,

$$\begin{aligned}\mathbf{P}_\lambda\left(\mathcal{U}_\epsilon(\omega) \neq \mathbb{T}^d_{[a]}\right) &\leq \mathbf{P}_\lambda\left(\bigcup_{B\in\mathfrak{B}} (\omega(B) = 0)\right)\\ &\leq K\exp(-\lambda(2\eta)^d) \xrightarrow{\lambda\to\infty} 0.\end{aligned}$$

Moreover, by the nerve lemma

$$\left(\mathcal{U}_\epsilon(\omega) = \mathbb{T}^d_{[a]}\right) \subset \bigcap_{i=0}^{d} \left(\beta_i(\mathcal{C}_\epsilon) = \beta_i(\mathbb{T}^d_{[a]})\right),$$

and the result follows. □

The following lemma is straightforward.

Lemma 15. *Let k_1, k_2 and k_3 be real positive constants and $f : \mathbf{R}^2_+ \to \mathbf{R}$ defined as*

$$f(x, y) = \exp\left(-\frac{k_1 - x}{2k_2}\log\left(1 + \frac{(k_1 - x)k_2}{k_3 y}\right)\right).$$

Then, for $k_1 - x > 0$, the function is strictly increasing with respect to x and with respect to y.

Theorem 16. *For $y > \lambda a^d$, we have*

$$\mathbf{P}_\lambda(\beta_0 \geq y) \leq \exp\left(-\frac{y - \lambda a^d}{2}\log\left(1 + \frac{y - \lambda a^d}{(2^d - 1)^2\lambda}\right)\right).$$

Proof. To apply Theorem 8, we need to evaluate $\max(D\beta_0)$ and $\|D\beta_0\|_{L^\infty(\Omega,L^2(Y))}$. Since there are more points than connected components, $\mathbf{E}_\lambda[\beta_0] \leq \mathbf{E}_\lambda[N_0] = \lambda a^d$. According to the definition of D, $\max(D\beta_0)$ is the maximum variation of β_0 induced by the addition of an arbitrary point. If this point is at a distance smaller than ϵ from ω, then $D\beta_0 \leq 0$, otherwise, $D\beta_0 = 1$, so $\max(D\beta_0) = 1$. Besides, this added point can join at most two connected components in each dimension, so in d dimensions it can join at most 2^d connected component, which means that $D\beta_0$ ranges from $-(2^d - 1)$ to 1, and then

$$\|D\beta_0\|_{L^\infty(\Omega,L^2(Y))} \leq \lambda \max|D\beta_0|^2 = \lambda(2^d - 1)^2.$$

Using Lemma 15 and Theorem 8, we get:

$$\mathbf{P}_\lambda(\beta_0 \geq y) \leq \exp\left(-\frac{y - \lambda a^d}{2}\log\left(1 + \frac{y - \lambda a^d}{(2^d - 1)^2\lambda}\right)\right),$$

for $y > \lambda a^d \geq \mathbf{E}_\lambda[\beta_0]$. □

5. Second order moments

We use all the definitions of the previous section.

Lemma 17. *We can rewrite N_k as*

$$N_k = \frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}\lambda^{k-i} I_i\left(\int_{(\mathbb{T}^d_a)^i} h(x_1, \ldots, x_k)\ dx_1 \ldots\ dx_{k-i}\right).$$

Proof. We have that

$$\int_{\Delta_k} h(x_1,\cdots,x_k)(\ \mathrm{d}\omega(x_1)-\lambda\ \mathrm{d}x_1)\dots(\ \mathrm{d}\omega(x_i)-\lambda\ \mathrm{d}x_i)\lambda\ \mathrm{d}x_{i+1}\dots\lambda\ \mathrm{d}x_k$$
$$=\sum_{j=0}^{i}(-1)^j\binom{i}{j}\int_{\Delta_k} h(x_1,\cdots,x_k)\ \mathrm{d}\omega(x_1)\dots\ \mathrm{d}\omega(x_j)\lambda\ \mathrm{d}x_{j+1}\dots\lambda\ \mathrm{d}x_k.$$

Thus, after some algebrism with the binomial factors, we have

$$\frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}\sum_{j=0}^{i}(-1)^j\binom{i}{j}\int_{\Delta_k} h(x_1,\cdots,x_k)\ \mathrm{d}\omega(x_1)\dots\ \mathrm{d}\omega(x_j)\lambda\ \mathrm{d}x_{j+1}\dots\lambda\ \mathrm{d}x_k$$
$$=\frac{1}{k!}\int_{\Delta_k} h(x_1,\cdots,x_k)\ \mathrm{d}\omega(x_1)\dots\ \mathrm{d}\omega(x_k)=N_k,$$

concluding the proof. □

Definition 6. *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two simplices with common vertices. For $L\in\mathcal{P}(\{1,2\})$, let us denote m_L the number of vertices belonging exactly to the list L of simplices.*

Then $M=m_{12}+m_1+m_2$ is the total number of vertices and $\mathcal{J}_2$ represents the integral on these two simplices:

$$\mathcal{J}_2(m_{12},m_1,m_2)=\int_{\Delta_{m_{12}+m_1}}\int_{\Delta_{m_{12}+m_2}} h_{m_{12}+m_1}h_{m_{12}+m_2}\ dx_1\dots\ dx_M.$$

with $x_1,\cdots,x_M$ being the M vertices.

Lemma 18. *For $d=1$ and $\epsilon=1/2$, we have*

$$\mathcal{J}_2(m_{12},m_1,m_2)=m_{12}+m_1+m_2+\frac{2m_1m_2}{m_{12}+1}. \tag{9}$$

Proof. Let us split the integration domain of $\mathcal{J}_2$ in two domains S_1 and S_2 corresponding to the cases:

(1) All the vertices are connected with each other, thus there is only one simplex. The integral on S_1 is simply the number of points in the simplex: $M=m_{12}+m_1+m_2$.
(2) There are at least two vertices at distance $d>1$, which leads to two simplices. By symetry we can choose to order the m_L vertices for each $L\in\mathcal{P}(\{1,2\})$ from lowest to greatest or the opposite and choose which simplex is on which side of the axis. Thus we have the integral on S_2 equal to $2m_{12}!m_1!m_2!\mathcal{A}$, with $\mathcal{A}$ an integral whose calculation is detailed below.

We choose to enumerate the vertices of the simplexes such that:

- $x_1,\cdots,x_{m_{12}}$ are the m_{12} common vertices.
- $x_{m_{12}+1},\cdots,x_{m_{12}+m_1}$ are the m_1 vertices of only $\mathcal{C}_1$.
- $x_{m_{12}+m_1+1},\cdots,x_M$ are the m_2 vertices of only $\mathcal{C}_2$.

Without loss of generality we can choose the origin to be x_1. The vertices are now order as described in Fig. 6:

$$0\le x_{m_{12}}\le x_{m_{12}-1}\le\cdots\le x_2\le 1,$$
$$-1\le x_2-1\le x_{m_{12}+1}\le x_{m_{12}+2}\le\cdots\le x_{m_{12}+m_1}\le x_{m_{12}+1}+1,$$
$$x_{m_{12}+1}\le x_{m_{12}+m_1+1}-1\le x_M\le x_{M-1}\le\cdots\le x_{m_{12}+m_1+1}\le 1,$$

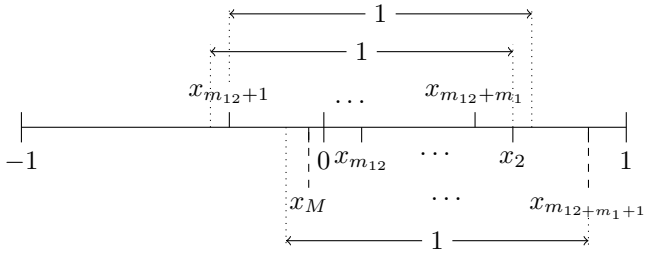

FIGURE 6. Example of relative positions of the points

Let us denote $J_a(f)(x) = \int_a^x f(u)\ \mathrm{d}u$ then we write the composition $J_a^{(2)}(f)(x) = \int_a^x \int_a^u f(v)\ \mathrm{d}v\ \mathrm{d}u$. We also denote $m = m_{12} + 1$ and $n = m_{12} + m_1 + 1$, then we have:

$$\mathcal{A} = \int_0^1 J_0^{(m_{12}-2)}(\mathbf{1})(x_2) \int_{x_2-1}^0 -J_{x_m+1}^{(m_1-1)}(\mathbf{1})(x_m) \int_{x_m+1}^1 J_{x_n-1}^{(m_2-1)}(\mathbf{1})(x_n)\ \mathrm{d}x_n\, \mathrm{d}x_m\, \mathrm{d}x_2.$$

We easily find that:

$$\begin{aligned} J_0^{(m_{12}-2)}(\mathbf{1})(x_2) &= \frac{x_2^{m_{12}-2}}{(m_{12}-2)!}, \\ -J_{x_m+1}^{(m_1-1)}(\mathbf{1})(x_m) &= \frac{1}{(m_1-1)!}, \\ J_{x_n-1}^{(m_2-1)}(\mathbf{1})(x_n) &= \frac{1}{(m_2-1)!}. \end{aligned}$$

Thus we have:

$$\begin{aligned} A &= \frac{1}{(m_{12}-2)!(m_1-1)!(m_2-1)!} \int_0^1 x_2^{m_{12}-2} \int_{x_2-1}^0 -x_m\ \mathrm{d}x_m\ \mathrm{d}x_2 \\ &= \frac{1}{(m_1-1)!(m_2-1)!(m_{12}+1)!}, \end{aligned}$$

concluding the proof. □

Theorem 19. *Let $\epsilon \leq a/6$. Then, the covariance between the number of $(k-1)$-simplices, N_k, and the number of $(l-1)$-simplices, N_l, for $l \leq k$ is given by*

$$\begin{aligned} &Cov_\lambda\left[N_k, N_l\right] \\ &= \sum_{i=0}^{l-1} \frac{1}{i!(k-l+i)!(l-i)!} (\lambda(2\epsilon)^d)^{k+i} \left(\frac{a}{2\epsilon}\right)^d \left(k+i+2\frac{i(k-l+i)}{l-i+1}\right)^d. \end{aligned} \tag{10}$$

Proof. We want to evaluate $\mathbf{E}_\lambda\left[(N_k - \mathbf{E}_\lambda\left[N_k\right])(N_l - \mathbf{E}_\lambda\left[N_l\right])\right]$. By Lemma 17, this can be written as

$$\mathbf{E}_\lambda\left[\frac{1}{k!}\sum_{i=1}^k \binom{k}{i} \lambda^{k-i} I_i\left(f_i^k\right) \frac{1}{l!}\sum_{i=1}^l \binom{l}{i} \lambda^{l-i} I_i\left(f_i^l\right)\right],$$

where

$$f_j^n = \int_{(\mathbb{T}_a^d)^j} h(v_1, \cdots, v_n)\ \mathrm{d}v_1 \ldots\ \mathrm{d}v_{n-j}.$$

Using the isometry formula, given by Eq. (1), we have

$$\begin{aligned}
\mathrm{Cov}_\lambda\left[N_k, N_l\right] &= \frac{1}{k!l!}\sum_{i=1}^{l}\binom{k}{i}\binom{l}{i}\lambda^{k+l-2i}\mathbf{E}_\lambda\left[I_i\left(f_i^k\right)I_i\left(f_i^l\right)\right] \\
&= \frac{1}{k!l!}\sum_{i=1}^{l}\binom{k}{i}\binom{l}{i}\lambda^{k+l-2i}i!\langle f_i^k f_i^l\rangle_{L^2(\lambda)^{\circ i}} \\
&= \sum_{i=0}^{l-1}\frac{1}{i!(k-l+i)!(l-i)!}\lambda^{k-l+2i}\langle f_{l-i}^k f_{l-i}^l\rangle_{L^2(\lambda)^{\circ(l-i)}}. \qquad (11)
\end{aligned}$$

Hence, we are reduced to compute

$$\begin{aligned}
\langle f_j^k f_j^l\rangle_{L^2(\lambda)^{\circ(j)}} = \int_{(\mathbb{T}_a^d)^j}\Bigg(&\int_{(\mathbb{T}_a^d)^{l-j}} h(v_1,\cdots,v_l)\,\mathrm{d}v_{j+1}\ldots\,\mathrm{d}v_l \\
&\int_{(\mathbb{T}_a^d)^{k-j}} h(v_1,\cdots,v_k)\,\mathrm{d}v_{j+1}\ldots\,\mathrm{d}v_k\Bigg)\lambda\,\mathrm{d}v_1\ldots\lambda\,\mathrm{d}v_j.
\end{aligned}$$

Since $a > \epsilon/6$, we have

$$\begin{aligned}
\langle f_j^k f_j^l\rangle_{L^2(\lambda)^{\circ(j)}} = &\int_{[0,a]^d}\lambda\,\mathrm{d}v_1\int_{([0,a]^d)^{k-1}} h(0,v_2,\cdots,v_k) \\
&\times h(0,v_2,\cdots,v_j,v_1',\cdots,v_{l-j}')\,\mathrm{d}v_{l-j}'\ldots\,\mathrm{d}v_1'\,\mathrm{d}v_k\ldots\,\mathrm{d}v_{j+1}\lambda\,\mathrm{d}v_j\ldots\lambda\,\mathrm{d}v_2.
\end{aligned}$$

Moreover, if $v_i = (u_{i,1},\cdots,u_{i,d})$ and $v_i' = (u_{i,1}',\cdots,u_{i,d}')$ and we proceed to the following substitutions:

$$\begin{aligned}
u_{i,1} &= 2\epsilon x_i \text{ if } 2\le i\le j, \\
u_{i,1} &= 2\epsilon y_{k-j} \text{ if } j+1\le i\le k, \\
u_{i,1}' &= 2\epsilon z_i \text{ if } 1\le i\le l-j,
\end{aligned}$$

This results in a Jacobian $(2\epsilon)^{k+l-2i-1}$ and we recognize the integral to be exactely $\mathcal{J}_2(j,k-j,l-j)$ as defined in Definition 6. Thus, we have:

$$\langle f_j^k f_j^l\rangle_{L^2(\lambda)^{\circ(j)}} = \lambda^i a^d(2\epsilon)^{k+l-2i-1}\left(\mathcal{J}_2(j,k-j,l-j)\right)^d.$$

Finally, using Eq. (9) and Eq. (11) gives the result. □

Remark. We remark that the possibility of writing $\mathrm{Var}(N_k)$ as Eq. (10) is due the fact that we use the maximum norm. This simplifies the calculations since we can treat each component individually. However, considering the Euclidean norm it is still possible to find analytically a closed-form expression for $\mathrm{Var}(N_k)$, but its calculation involves nasty integrals and a generic term cannot be found. When we consider the Rips-Vietoris complex in $\mathbb{T}_a^2$, the variance of the number of 1-simplices and 2-simplices are given by:

$$\mathbf{V}_\lambda\left[N_2\right] = \left(\frac{a}{2\epsilon}\right)^2\left(\frac{\pi}{2}(4\lambda\epsilon^2)^2 + \pi^2(4\lambda\epsilon^2)^3\right),$$

and

$$\begin{aligned}
\mathbf{V}_\lambda\left[N_3\right] = \left(\frac{a}{2\epsilon}\right)^2\Bigg((4\lambda\epsilon)^3\frac{\pi}{6}\left(\pi-\frac{3\sqrt{3}}{4}\right) + (4\lambda\epsilon^2)^4\pi\left(\frac{\pi^2}{2}-\frac{5}{12}-\frac{\pi\sqrt{3}}{2}\right) \\
+(4\lambda\epsilon^2)^5\frac{\pi^2}{4}\left(\pi-\frac{3\sqrt{3}}{4}\right)^2\Bigg).
\end{aligned}$$

Since we have an expression for the variance of the number of k-simplices, it is possible to calculate one for the Euler characteristic.

Theorem 20. *Let $\epsilon \leq a/6$. Then, the variance of the Euler characteristic in a d torus is:*

$$\mathbf{V}_\lambda[\chi] = \left(\frac{a}{2\epsilon}\right)^d \sum_{n=1}^{\infty} c_n^d (\lambda(2\epsilon)^d)^n,$$

where

$$c_n^d = \sum_{j=\lceil (n+1)/2 \rceil}^{n} \left[2 \sum_{i=n-j+1}^{j} \frac{(-1)^{i+j}}{(n-j)!(n-i)!(i+j-n)!} \left(n + \frac{2(n-i)(n-j)}{1+i+j-n} \right)^d \right.$$
$$\left. - \frac{1}{(n-j)!^2(2j-n)!} \left(n + \frac{2(n-j)^2}{1+2j-n} \right)^d \right].$$

Proof. The variance of χ is given by:

$$\begin{aligned}
\mathbf{V}_\lambda[\chi] &= \mathbf{E}_\lambda\left[(\chi - \mathbf{E}_\lambda[\chi])^2\right] = \mathbf{E}_\lambda\left[\left(\sum_{k=1}^{\infty}(-1)^k N_k - \sum_{k=1}^{\infty}(-1)^k \mathbf{E}_\lambda[N_k]\right)^2\right] \\
&= \mathbf{E}_\lambda\left[\left(\sum_{k=1}^{\infty}(-1)^k (N_k - \mathbf{E}_\lambda[N_k]\right)^2\right] \\
&= \mathbf{E}_\lambda\left[\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(-1)^{i+j}(N_i - \mathbf{E}_\lambda[N_i])(N_j - \mathbf{E}_\lambda[N_j])\right].
\end{aligned}$$

We remark that $N_i \leq \frac{N_1^i}{i!}$, so there is a constant c such that

$$\mathbf{E}_\lambda\left[\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|(N_i - \mathbf{E}_\lambda[N_i])(N_j - \mathbf{E}_\lambda[N_j])|\right] \leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left|\frac{N_1^i}{i!}\frac{N_1^j}{j!}\right| \leq c\mathbf{E}_\lambda[e^{n_1}]^2 < \infty.$$

Thus the alternating serie converges absolutely allowing us to exchange the mean with the sums and we can write

$$\mathbf{V}_\lambda[\chi] = \sum_{i=1}^{\infty}(-1)^i \sum_{j=1}^{\infty}(-1)^j \mathrm{Cov}_\lambda[N_i, N_j].$$

The result follows by Eq. (10) and some tedious but straightforward algebra. □

Lemma 21. *Let n be a positive integer, then*

$$\sum_{j=1}^{n}\binom{n}{j}\left(\binom{j-1}{n-j-1} - \binom{j-1}{n-j}\right) = (-1)^n.$$

Proof. We first simplify the expression:

$$\sum_{j=1}^{n}\binom{n}{j}\left(\binom{j-1}{n-j-1} - \binom{j-1}{n-j}\right) = \sum_{j=1}^{n}\frac{2n-3j}{j}\binom{n}{j}\binom{j}{n-j},$$

Then, applying hypergeometric functions, we solve the sum:

$$\sum_{j=1}^{n}\frac{2n-3j}{j}\binom{n}{j}\binom{j}{n-j} = (-1)^n.$$

□

Theorem 22. *In one dimension, the expression of the variance of the Euler characteristic is:*

$$\mathbf{V}_\lambda[\chi] = a\left(\lambda e^{-2\lambda\epsilon} - 4\lambda^2\epsilon e^{-4\lambda\epsilon}\right).$$

Proof. If $d = 1$, according to Theorem 20:

$$\mathbf{V}_\lambda[\chi] = \frac{a}{2\epsilon}\sum_{n=1}^{\infty} c_n^1 (2\lambda\epsilon)^n, \tag{12}$$

and we define

$$\alpha_n = \sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\left[2\sum_{i=n-j+1}^{j}\frac{(-1)^{i+j}n}{(n-j)!(n-i)!(i+j-n)!} - \frac{n}{(n-j)!^2(2j-n)!}\right].$$

and $\beta_n = c_n^1 - \alpha$. It is well known that

$$\sum_{i=0}^{2j-n}(-1)^i\binom{j}{i} = (-1)^{2j-n-1}\binom{j-1}{2j-n},$$

using Stiffel's relation, we obtain:

$$\begin{aligned}
\alpha_n &= (-1)^n\frac{n}{n!}\sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\left[\binom{n}{j}2\sum_{i=0}^{2j-n}(-1)^i\binom{j}{i} + 2(-1)^n\binom{n}{j}\right] \\
&= \frac{1}{(n-1)!}\sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\left[2\binom{n}{j}\binom{j-1}{n-j-1} - \binom{n}{j}\binom{j}{n-j} - 2(-1)^n\binom{n}{j}\right] \\
&= \frac{1}{(n-1)!}\sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\left[\binom{n}{j}\left(\binom{j-1}{n-j} - \binom{j-1}{n-j-1}\right) - 2(-1)^n\binom{n}{j}\right]. \qquad (13)
\end{aligned}$$

The identity $\binom{n}{j} = \binom{n}{n-j}$ allows us to write that

$$\sum_{j=\lceil(n+1)/2\rceil}^{n}(-2(-1)^n)\binom{n}{j} = \sum_{j=0}^{n}\binom{n}{j} = 2^n, \quad n \text{ odd},$$

$$\sum_{j=\lceil(n+1)/2\rceil}^{n}(-2(-1)^n)\binom{n}{j} = \binom{n}{n/2} + \sum_{j=0}^{n} -\binom{n}{j} = -2^n + \binom{n}{n/2}, \quad n \text{ even}.$$

Since $\binom{j-1}{n-j} = 0$ for $j < \lceil\frac{n+1}{2}\rceil$, we have

$$\sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\binom{n}{j}\left(\binom{j-1}{n-j} - \binom{j-1}{n-j-1}\right) = \sum_{j=1}^{n}\binom{n}{j}\left(\binom{j-1}{n-j} - \binom{j-1}{n-j-1}\right)$$

for n odd and

$$\begin{aligned}
&\sum_{j=\lceil\frac{n+1}{2}\rceil}^{n}\binom{n}{j}\left(\binom{j-1}{n-j} - \binom{j-1}{n-j-1}\right) \\
&\qquad = -\binom{n}{n/2} + \sum_{j=1}^{n}\binom{n}{j}\left(\binom{j-1}{n-j} - \binom{j-1}{n-j-1}\right).
\end{aligned}$$

for n even. According to Lemma 21, we get:

$$\begin{aligned}
\sum_{j=\lceil (n+1)/2\rceil}^{n} \binom{n}{j}\left[\binom{j-1}{n-j-1}-\binom{j-1}{n-j}\right] &= -1, \quad n \text{ odd},\\
\sum_{j=\lceil (n+1)/2\rceil}^{n} \binom{n}{j}\left[\binom{j-1}{n-j-1}-\binom{j-1}{n-j}\right] &= 1-\binom{n}{n/2}, \quad n \text{ even}.
\end{aligned}$$

Then, we substitute these two last expressions in Eq. (13) to obtain

$$\alpha_n = (-1)^n \frac{(1-2^n)\mathbf{1}_{[n\geq 1]}}{(n-1)!},$$

and thus

$$\sum_{i=0}^{\infty} \alpha_n x^n = -xe^{-x} + 2xe^{-2x}.$$

Proceeding along the same line, β_n is given by

$$\begin{aligned}
\beta_n &= \sum_{j=\lceil \frac{n+1}{2}\rceil}^{n} \left[2 \sum_{i=n-j+1}^{j} \frac{(-1)^{i+j} 2(n-i)(n-j)}{(n-j)!(n-i)!(i+j-n+1)!} \right.\\
&\qquad\qquad \left. - \frac{2(n-j)^2}{(n-j)!^2(2j-n+1)!} \right]\\
&= (-1)^n \left(\frac{(-2+2^n)\mathbf{1}_{[n\geq 1]}}{(n-1)!} - \frac{2\mathbf{1}_{[i\geq 2]}}{(i-2)!} \right),
\end{aligned}$$

and again we can simplify the power serie $\sum_{i=0}^{\infty} \beta_n x^n$:

$$\sum_{i=0}^{\infty} \beta_n x^n = 2xe^{-x} - 2(x+x^2)e^{-2x}.$$

Then, substituting α_n and β_n in Eq. (12) yields the result. □

Theorem 23. *We have $D\chi \leq 2$ and $\|D\chi\|_{L^\infty(\Omega, L^2(\mathbb{T}_a^d))} < \infty$ and*

$$\mathbf{P}(\chi - \bar{\chi} \geq x) \leq \exp\left(-\frac{x}{4} \log\left(1 + \frac{2x}{\mathbf{V}_\lambda[\chi]} \right) \right).$$

Proof. In two dimensions, the Euler characteristic is:

$$\chi = \beta_0 - \beta_1 + \beta_2.$$

Therefore we can bound $D\chi$ by the variation of $\beta_0 - \beta_1$ added to the variation of β_2 when we add a vertex to a simplicial complex.

If we add a vertex on the torus, either the vertex is isolated or not. In the first case, it forms a new connected component incrementing β_0 by 1, and the number of holes that is β_1 is the same. Otherwise, as there is no new connected component, β_0 is the same, but the new vertex can at most fill a hole incrementing β_1 by 1. Therefore, the variation of $\beta_0 - \beta_1$ is at most 1.

Now, let us look at the variation of β_2 when we add a vertex to a simplicial complex. According to Proposition 3 is at most 1, showing that $D\chi \leq 2$. Then, we use Eq. (4) to complete the proof. □

6. Nth order moments

For this section, without loss of generality, using Proposition 9, we can choose $k = 1/2\epsilon$, so $\lambda_\tau = \lambda(2\epsilon)^d$, $\epsilon_\tau = 1/2$ and $ak = a/2\epsilon$.

We are interested in the central moment, so we introduce the following notation for the centralized number of $(k-1)$-simplices: $\tilde{N_k} = N_k - \bar{N}_k$.

Finally, let us denote that $\binom{i}{j} = 0$ as soon as $i \leq 0$ or $j \leq 0$ or $i - j \leq 0$ for i and j integers.

Definition 7. *We extend the Definition 6 used in the second order moments calculations.*

Let $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ be three simplices with common vertices. For $L \in \mathcal{P}(\{1,2,3\})$, let us denote m_L the number of vertices belonging exactly to the list L of simplices.

Then $M = m_{123} + m_{12} + m_{13} + m_{23} + m_1 + m_2 + m_3$ is the total number of vertices and $\mathcal{J}_3$ represents the integral on these three simplices:

$$\mathcal{J}_3 = \int_{\Delta_{p_1}} \int_{\Delta_{p_2}} \int_{\Delta_{p_3}} h_{p_1} h_{p_2} h_{p_3} \ dx_1 \dots \ dx_M.$$

with p_i being the number of vertices of simplex $\mathcal{C}_i$ for $i = 1, \cdots, 3$, for instance $p_1 = m_{123} + m_{12} + m_{13} + m_1$, and $x_1, \cdots, x_M$ being the M vertices.

Definition 8. *We denote $\mathcal{J}_3(i,j,s,t)$ the integral defined above such that*

- $m_{123} = 2t - i - j + s \vee 0$
- $m_{12} = i + j - s - t \vee 0$
- $m_{13} = i - t \vee 0$
- $m_{23} = j - t \vee 0$
- $m_1 = k - i \vee 0$
- $m_2 = k - j \vee 0$
- $m_3 = k - s \vee 0.$

Theorem 24. *The third moment of the number of $(k-1)$-simplices is given by:*

$$\mathbf{E}_\lambda \left[\tilde{N_k}^3 \right] = \sum_{i,j,s,t} \lambda^{3k-i-j} t! \binom{k}{i} \binom{k}{j} \binom{k}{s} \binom{i}{t} \binom{j}{t} \binom{t}{i+j-s-t} \mathcal{J}_3(i,j,s,t),$$

with $s \geq |i-j|$.

Proof. The chaos decomposition of the number of $(k-1)$-simplices is as shown in lemma 17:

$$\tilde{N_k} = I_1(f_1) + \cdots + I_k(f_k) = \sum_{i=1}^{k} I_i(f_i),$$

with

$$f_i(x_1, \cdots, x_i) = \binom{k}{i} \int h(x_1, \cdots, x_k) \lambda^{k-i} \ \mathrm{d}x_k \dots \ \mathrm{d}x_{i+1},$$

And

$$I_i(f_i) = \int f_i(\ \mathrm{d}\omega(x_1) - \ \mathrm{d}\lambda(x_1)) \dots (\ \mathrm{d}\omega(x_i) - \ \mathrm{d}\lambda(x_i)).$$

Then, we define $g_{i,j,i+j-s} = \sum_{t=\lceil \frac{i+j-s}{2} \rceil}^{i+j-s\wedge i\wedge j} t!\binom{i}{t}\binom{j}{t}\binom{t}{i+j-s-t} f_i \circ_t^{u-t} f_j$ and using the chaos expansion (cf Proposition 6):

$$\begin{aligned}
\tilde{N_k}^3 &= (I_1(f_1) + \cdots + I_k(f_k))^3 \\
&= \left(\sum_{i=1}^{k}\sum_{j=1}^{k} I_i(f_i)I_j(f_j)\right)(I_1(f_1) + \cdots + I_k(f_k)) \\
&= \sum_{i,j=1}^{k}\sum_{s=|i-j|}^{i+j} I_s(g_{i,j,i+j-s})(I_1(f_1) + \cdots + I_k(f_k)) \\
&= \sum_{i,j,l=1}^{k}\sum_{s=|i-j|}^{i+j} I_s(g_{i,j,i+j-s})I_l(f_l).
\end{aligned}$$

When taking the expectation of $\tilde{N_k}$, we use the isometry formula in Eq. (1). Denoting $u = i + j - s$, we obtain:

$$\begin{aligned}
\mathbf{E}_\lambda\left[\tilde{N_k}^3\right] &= \mathbf{E}_\lambda\left[\sum_{i,j=1}^{k}\sum_{s=|i-j|\vee 1}^{i+j\wedge k} I_s(g_{i,j,u})I_s(f_s)\right] \\
&= \sum_{i,j=1}^{k}\sum_{s=|i-j|\vee 1}^{i+j\wedge k}\int g_{i,j,u}f_s\lambda^s \,\mathrm{d}x_1\ldots\,\mathrm{d}x_s \\
&= \sum_{i,j=1}^{k}\sum_{s=|i-j|\vee 1}^{i+j\wedge k}\sum_{t=\lceil\frac{u}{2}\rceil}^{u\wedge i\wedge j}\lambda^s t!\binom{i}{t}\binom{j}{t}\binom{t}{u-t}\int(f_i\circ_t^{u-t} f_j)f_s\,\mathrm{d}x_1\ldots\,\mathrm{d}x_s.
\end{aligned}$$

Then we recognize the integral defined in Definition 8:

$$\mathbf{E}_\lambda\left[\tilde{N_k}^3\right] = \sum_{i,j,=1}^{k}\sum_{s=|i-j|\vee 1}^{i+j\wedge k}\sum_{t=\lceil\frac{u}{2}\rceil}^{u\wedge i\wedge j}\lambda^{3k-i-j}t!\binom{k}{i}\binom{k}{j}\binom{k}{s}\binom{i}{t}\binom{j}{t}\binom{t}{u-t}\mathcal{J}_3(i,j,s,t).$$

Finally, relaxing the boundaries on the sums conclude the proof. □

Definition 9. *Let $\mathcal{C}_1,\cdots,\mathcal{C}_n$ be n simplices with common vertices. For $L \in \mathcal{P}(\{1,\cdots,n\})$, let us denote m_L the number of vertices belonging exactly to the list L of simplices.*

Then $M = \sum_{L\in\mathcal{P}(\{1,\cdots,n\})} m_L$ is the total number of vertices and $\mathcal{J}_n$ represents the integral on these n simplices:

$$\mathcal{J}_n = \int_{\Delta_{p_1}}\cdots\int_{\Delta_{p_n}} h_{p_1}\ldots h_{p_n}\ dx_1\ldots\ dx_M.$$

with p_i being the number of vertices of simplex $\mathcal{C}_i$ for $i = 1,\cdots,n$, and $x_1,\cdots,x_M$ being the M vertices.

Theorem 25. *The expression of the nth power of the number of $(k-1)$-simplices is given by:*

$$\tilde{N_k^n} = \sum_{i_1,\cdots,i_n}\sum_{s_1,\cdots s_{n-2}}\sum_{t_1,\cdots t_{n-2}}\left(\prod_{j=1}^{n-2} t_j!\binom{m_{j,1}}{t_j}\binom{m_{j,2}}{t_j}\binom{t_j}{u_j-t_j}\right) I_a(\circ_{j\in A}f_{i_j})I_b(\circ_{j\in\bar{A}}f_{i_j}). \tag{14}$$

With for $j \in \{1,\cdots,n-2\}$:

- $1 \leq i_1, \cdots, i_n \leq k$,
- $s_j \geq |m_{j,1} - m_{j,2}|$,
- $m_{j,1} = i_{2j-1}$ *if* $1 \leq j \leq \lfloor \frac{n}{2} \rfloor$ *and* $s_{2(j-\lfloor \frac{n}{2} \rfloor)-1}$ *otherwise,*
- $m_{j,2} = i_{2j}$ *if* $1 \leq j \leq \lfloor \frac{n}{2} \rfloor$ *and* $s_{2(j-\lfloor \frac{n}{2} \rfloor)}$ *otherwise,*
- $u_j = m_{j,1} + m_{j,2} - s_j$,
- $A \subset \{1, \cdots, n\}$,
- *If* n *is even, then* $a = s_{n-3}$ *and* $b = s_{n-2}$,
- *If* n *is odd, then* $a = s_{n-2}$ *and* $b = i_n$.

Proof. The decomposition of the centralized number of $(k-1)$-simplices is:

$$\tilde{N_k} = I_1(f_1) + \cdots + I_k(f_k) = \sum_{i=1}^{k} I_i(f_i).$$

Now, we raise $\tilde{N_k}$ to the nth power:

$$\tilde{N_k}^n = \left(\sum_{i=1}^{k} I_i(f_i) \right)^n.$$

First, we consider the case where n is even, we can group the factors by 2:

$$\tilde{N}_k^n = \left(\sum_{i_1=1}^{k} I_{i_1}(f_{i_1}) \sum_{i_2=1}^{k} I_{i_2}(f_{i_2}) \right) \cdots \left(\sum_{i_{n-1}=1}^{k} I_{i_{n-1}}(f_{i_{n-1}}) \sum_{i_n=1}^{k} I_{i_n}(f_{i_n}) \right).$$

We then use the chaos expansion of Proposition 6:

$$I_i(f_i) I_j(f_j) = \sum_{s=0}^{2(i \wedge j)} I_{i+j-s} \left(\sum_{s \leq 2t \leq 2(s \wedge i \wedge j)} t! \binom{i}{t} \binom{j}{t} \binom{t}{s-t} f_i \circ_t^{s-t} f_j \right)$$

$$= \sum_{s=|i-j|}^{i+j} I_s \left(\sum_{i+j-s \leq 2t \leq 2(i+j-s) \wedge i \wedge j)} t! \binom{i}{t} \binom{j}{t} \binom{t}{i+j-s-t} f_i \circ_t^{i+j-s-t} f_j \right).$$

Let us denote $g_s = t! \binom{i}{t} \binom{j}{t} \binom{t}{i+j-s-t} f_i \circ_t^{i+j-s-t} f_j$, so we can re-write, relaxing the boundaries on the sums:

$$I_i(f_i) I_j(f_j) = \sum_{s \geq |i-j|} \sum_{t} I_s(g_s).$$

Thus, we have:

$$\tilde{N}_k^n = \sum_{i_1, i_2 = 1}^{k} \sum_{s_1 \geq |i_1 - i_2|} \sum_{t_1} I_{s_1}(g_{s_1}) \cdots \sum_{i_{n-1}, i_n = 1}^{k} \sum_{s_{n/2} \geq |i_{n-1} - i_n|} \sum_{t_{n/2}} I_{s_{n/2}}(g_{s_{n/2}}).$$

We go on grouping terms by 2 until we only have a product of 2 chaos left: First we made $n/2$ chaos expansions, leading to $n/2$ sums with indexes s_j, $j = 1, \cdots, n/2$. To reduce the number of chaos to 2, we have to make other chaos expansions. For $j \geq \frac{n}{2} + 1$, the sum indexed by s_j represents the expansion of the chaos indexed $s_{2(j-\frac{n}{2})-1}$ and $s_{2(j-\frac{n}{2})-1}$. We have 2 chaos remaining when $j = 2(j - \frac{n}{2}) + 2$, i.e. when $j = n - 2$.

Moreover, there are as much sums indexed with t_j as with s_j, that is $n-2$. Thus we can write:

$$\tilde{N}_k^n = \sum_{i_1, \cdots, i_n = 1}^{k} \sum_{s_1, \cdots s_{n-2}} \sum_{t_1, \cdots t_{n-2}} I_{s_{n-3}}(\phi_{s_{n-3}}) I_{s_{n-2}}(\phi_{s_{n-2}}),$$

With $s_j \geq |m_{j,1} - m_{j,2}|$ for $j \in \{1, \cdots, n-2\}$ if we denote:

- $m_{j,1} = i_{2j-1}$ if $1 \leq j \leq \frac{n}{2}$ and $s_{2(j-\frac{n}{2})-1}$ otherwise,
- $m_{j,2} = i_{2j}$ if $1 \leq j \leq \frac{n}{2}$ and $s_{2(j-\frac{n}{2})}$ otherwise.

Then, denoting $u_j = m_{j,1} + m_{j,2} - s_j$ and A the subset of $\{1, \cdots, n\}$ such that if $j \in A$ then the chaos i_j is expanded in the chaos s_{n-3}, we have:

$$I_{s_{n-3}}(\phi_{s_{n-3}}) I_{s_{n-2}}(\phi_{s_{n-2}}) = \left(\prod_{j=1}^{n-2} t_j! \binom{m_{j,1}}{t_j} \binom{m_{j,2}}{t_j} \binom{t_j}{u_j - t_j} \right) I_{s_{n-3}}(\circ_{j \in A} f_{i_j}) I_{s_{n-2}}(\circ_{j \in \bar{A}} f_{i_j}).$$

The notation $\circ_{j \in A} f_{i_j}$ represents the product defined in Eq. (3) of the functions f_{i_j} for $j \in A$, but whom variables depend on all the $i_1, \cdots, i_n, s_1, \cdots, s_{n-2}$, and $t_1, \cdots, t_{n-2}$.

Now, if n is odd, we consider $n-1$ which is even, therefore we have:

$$\begin{aligned} \tilde{N}_k^n &= \sum_{i_1, \cdots, i_{n-1}=1}^{k} \sum_{s_1, \cdots, s_{n-3}} \sum_{t_1, \cdots, t_{n-3}} I_{s_{n-4}}(\phi_{s_{n-4}}) I_{s_{n-3}}(\phi_{s_{n-3}}) \sum_{i_n=1}^{k} I_{i_n}(f_{i_n}) \\ &= \sum_{i_1, \cdots, i_n=1}^{k} \sum_{s_1, \cdots, s_{n-2}} \sum_{t_1, \cdots, t_{n-2}} I_{s_{n-2}}(\phi_{s_{n-2}}) I_{i_n}(f_{i_n}), \end{aligned}$$

with $s_j \geq |m_{j,1} - m_{j,2}|$ for $j \in \{1, \cdots, n-2\}$ using the same notations for $n-1$ instead of n:

- $m_{j,1} = i_{2j-1}$ if $1 \leq j \leq \frac{n-1}{2}$ and $s_{2(j-\frac{n-1}{2})-1}$ otherwise,
- $m_{j,2} = i_{2j}$ if $1 \leq j \leq \frac{n-1}{2}$ and $s_{2(j-\frac{n-1}{2})}$ otherwise.

And with $u_j = m_{j,1} + m_{j,2} - s_j$,

$$I_{s_{n-2}}(\phi_{s_{n-2}}) = \left(\prod_{j=1}^{n-2} t_j! \binom{m_{j,1}}{t_j} \binom{m_{j,2}}{t_j} \binom{t_j}{u_j - t_j} \right) I_{s_{n-2}}(\circ_{j \in \{1, \cdots, n-1\}} f_{i_j}),$$

concluding the proof. □

Theorem 26. *The expression of the nth moment of the number of $(k-1)$-simplices is given by:*

$$\mathbf{E}_\lambda \left[\tilde{N_k}^n \right] = \sum_{i_1, \cdots, i_n} \sum_{s_1, \cdots, s_{n-3}} \sum_{t_1, \cdots, t_{n-2}} \lambda^{nk+c} \left(\prod_{j=1}^{n} \lambda^{-i_j} \binom{k}{i_j} \right) \left(\prod_{j=1}^{n-2} t_j! \binom{m_{j,1}}{t_j} \binom{m_{j,2}}{t_j} \binom{t_j}{u_j - t_j} \right) \mathcal{J}_n(i_1, \cdots, i_n, s_1, \cdots, s_{n-3}, t_1, \cdots, t_{n-2}).$$

With for $j \in \{1, \cdots, n-2\}$:

- *if $j \leq n-3$, $s_j \geq |m_{j,1} - m_{j,2}|$,*
- *$m_{j,1} = i_{2j-1}$ if $1 \leq j \leq \lfloor \frac{n}{2} \rfloor$ and $s_{2(j-\lfloor \frac{n}{2} \rfloor)-1}$ otherwise,*
- *$m_{j,2} = i_{2j}$ if $1 \leq j \leq \lfloor \frac{n}{2} \rfloor$ and $s_{2(j-\lfloor \frac{n}{2} \rfloor)}$ otherwise,*
- *$m_{j,3} = s_j$ if $1 \leq j \leq n-3$ and s_{n-3} otherwise,*
- *$u_j = m_{j,1} + m_{j,2} - m_{j,3}$,*
- *If n is even, then $c = s_{n-3}$ and $s_{n-3} \geq |m_{n-2,1} - m_{n-2,2}| \vee |m_{n-3,1} - m_{n-3,2}|$,*
- *If n is odd, then $c = i_n$ and $i_n \geq |m_{n-2,1} - m_{n-2,2}|$.*

Proof. The expression of the nth power of the number of $(k-1)$-simplices is given in Eq. (14):

$$\tilde{N}_k^n = \sum_{i_1,\cdots,i_n=1}^{k} \sum_{s_1,\cdots,s_{n-2}} \sum_{t_1,\cdots,t_{n-2}} \left(\prod_{j=1}^{n-2} t_j! \binom{m_{j,1}}{t_j} \binom{m_{j,2}}{t_j} \binom{t_j}{u_j - t_j} \right) I_a(\circ_{j\in A} f_{i_j}) I_b(\circ_{j\in \bar{A}} f_{i_j}).$$

If n is even, we have:

$$\tilde{N}_k^n = \sum_{i_1,\cdots,i_n=1}^{k} \sum_{s_1,\cdots,s_{n-2}} \sum_{t_1,\cdots,t_{n-2}} \left(\prod_{j=1}^{n-2} t_j! \binom{m_{j,1}}{t_j} \binom{m_{j,2}}{t_j} \binom{t_j}{u_j - t_j} \right) I_{s_{n-3}}(\circ_{j\in A} f_{i_j}) I_{s_{n-2}}(\circ_{j\in \bar{A}} f_{i_j}).$$

So let us focus on the only part of the equation which is likely to change when we take the expected value, that we will denote:

$$K = \sum_{s_{n-3}} \sum_{s_{n-2}} I_{s_{n-3}}(\circ_{j\in A} f_{i_j}) I_{s_{n-2}}(\circ_{j\in \bar{A}} f_{i_j}).$$

We then use the property of Eq. (1) and recognize the integral from Definition 9:

$$\begin{aligned} \mathbf{E}_\lambda[K] &= \sum_{s_{n-3}} \left(\prod_{j=1}^{n} \lambda^{k-i_j} \binom{k}{i_j} \right) \lambda^{s_{n-3}} \mathcal{J}_n(i_1,\cdots,i_n,s_1,\cdots,s_{n-3},t_1,\cdots,t_{n-2}) \\ &= \sum_{s_{n-3}} \lambda^{nk+s_{n-3}} \left(\prod_{j=1}^{n} \lambda^{-i_j} \binom{k}{i_j} \right) \mathcal{J}_n(i_1,\cdots,i_n,s_1,\cdots,s_{n-3},t_1,\cdots,t_{n-2}), \end{aligned}$$

with $s_{n-3} \geq |m_{n-2,1} - m_{n-2,2}| \vee |m_{n-3,1} - m_{n-3,2}|$.

Then for n odd we directly write:

$$\begin{aligned} K' &= \sum_{i_n} \sum_{s_{n-2}} I_{i_n}(\circ_{j\in I} f_{i_j}) I_{s_{n-2}}(\circ_{j\in \bar{I}} f_{i_j}), \\ \mathbf{E}_\lambda[K'] &= \sum_{i_n} \lambda^{nk+i_n} \left(\prod_{j=1}^{n} \lambda^{-i_j} \binom{k}{i_j} \right) \mathcal{J}_n(i_1,\cdots,i_n,s_1,\cdots,s_{n-3},t_1,\cdots,t_{n-2}), \end{aligned}$$

with $i_n \in \{|m_{n-2,1} - m_{n-2,2}| \vee 1, k\}$.

The binomials with the i_j allow us to relax the boundaries on the sums on i_j, concluding the proof. □

7. Convergence

Let Γ be an arbitrary connected simplicial complex containing n points and X be a random simplicial complex in a compact set B generated by the Poisson point process ω. The number of occurrences of Γ in X is denoted as G_Γ. In what follows, we denote by $\mathbf{V}_\lambda[X]$ the variance of the random variable X under the probability distribution $\mathbf{P}_\lambda$. In this section, c is a constant which may vary from line to line and may depend on everything but λ.

Lemma 27. *The random variable G_Γ has a chaos representation given by:*

$$G_\Gamma = \sum_{i=0}^{n} I_i(f_i),$$

where f_i is a bounded symetric function.

Proof. Let k be the number of 1-simplices of the connected simplicial complex G_Γ and J_k be the set of the k couples of indices such that if $(i,j) \in J_k$, then the points x_i and x_j form an 1-simplex. Let

$$\tilde{h}(x_1, \cdots, x_n) = \frac{1}{c_\Gamma} \prod_{(i,j)\in J_k} \mathbf{1}_{[\|x_i - x_j\| \le \epsilon]},$$

where c_Γ is the number of permutations of $\{x_1, \cdots, x_n\}$ such that

$$\tilde{h}(x_1, \cdots, x_n) = \tilde{h}(x_{\sigma(1)}, \cdots, x_{\sigma(n)}),$$

and let $f(x_1, \cdots, x_n)$ be the symetrization of $\tilde{h}(x_1, \cdots, x_n)$. This means that

$$G_\Gamma = \sum_{\substack{x_1,\cdots,x_n \in \omega \\ x_i \neq x_j i f i \neq j}} f(x_1, \cdots, x_n) = \int_{\Delta_n} f(x_1, \cdots, x_n)\, \mathrm{d}\omega(x_1) \cdots\, \mathrm{d}\omega(x_n).$$

Using the binomial expansion and some algebra, we obtain

$$G_\Gamma = \sum_{i=0}^{n} \int_{\Delta_i} \left(\binom{n}{i} \int_{\Delta_{n-i}} f(x_1, \cdots, x_n) \lambda\, \mathrm{d}x_1 \ldots \lambda\, \mathrm{d}x_{n-i} \right)$$
$$(\,\mathrm{d}\omega(x_{n-i+1}) - \lambda\, \mathrm{d}x_{n-i+1}) \ldots (\,\mathrm{d}\omega(x_n) - \lambda\, \mathrm{d}x_n).$$

We define, for any $i \in \{1, \cdots, n\}$,

$$f_i(x_{i+1}, \cdots, x_n) = \binom{n}{i} \lambda^{n-i} \int_{B^{n-i}} f(x_1, \cdots, x_n)\, \mathrm{d}x_1 \ldots\, \mathrm{d}x_{n-i}$$

and to conclude the proof, we note that, since B is a compact set and h is bounded, f_i is bounded. □

Lemma 28. *For any Γ connected simplicial complex containing n points, for λ large enough,*

$$\mathbf{E}_\lambda [G_\Gamma] \le c\lambda^n \text{ and } \mathbf{V}_\lambda [G_\Gamma] \le P_\Gamma^{2n-1}(\lambda).$$

where $P_\Gamma^{2n-1}(\lambda)$ is a polynomial on λ of degree $2n-1$ depending on Γ.

Proof. Using 27 and the chaos properties, we obtain

$$\mathbf{E}_\lambda [G_\Gamma] = \lambda^n \int_{\Delta_n} f(x_1, \cdots, x_n)\, \mathrm{d}x_1 \ldots\, \mathrm{d}x_n \le c\lambda^n,$$

since f is bounded. Furthermore,

$$\begin{aligned}
\mathbf{V}_\lambda [G_\Gamma] &= \sum_{i=1}^{n} i! \|f_i\|^2_{L^2(B,\lambda)} \\
&= \sum_{i=1}^{n} i! \int_{\Delta_i} \left(\lambda^{n-i} \binom{n}{i} \int_{B^{n-i}} f(x_1, \cdots, x_n)\, \mathrm{d}x_1 \ldots\, \mathrm{d}x_{n-i} \right)^2 \lambda\, \mathrm{d}x_1 \ldots \lambda\, \mathrm{d}x_i \\
&= \sum_{i=1}^{n} i!\, \lambda^{2n-i} \int_{\Delta_i} \left(\binom{n}{i} \int_{B^{n-i}} f(x_1, \cdots, x_n)\, \mathrm{d}x_1 \ldots\, \mathrm{d}x_{n-i} \right)^2 \mathrm{d}x_1 \ldots\, \mathrm{d}x_i.
\end{aligned}$$

and since f is bounded, $\mathbf{V}_\lambda [G_\Gamma]$ is a polynomial of degree $2n-1$. □

Lemma 29. *For λ large enough, if $k > 1$,*

$$\mathbf{E}_\lambda \left[(I_k(1))^2 \right] < c\lambda^{2k-1},$$

and $\mathbf{E}_\lambda \left[(I_k(1))^2 \right]$ is constant if $k = 0$.

Proof. The proof is trivial for the case $k = 0$. If $k \geq 1$, for $i \leq k$ we have

$$\int_{\Delta_i} \mathrm{d}\omega(x_1) \dots \mathrm{d}\omega(x_i) = \prod_{j=0}^{i-1} (\omega(B) - j),$$

so we can rewrite $I_k(1)$ as follows:

$$I_k(1) = \sum_{i=0}^{k} \left[\binom{k}{i} (-\lambda S(B))^i \prod_{j=0}^{k-i-1} (\omega(B) - j) \right].$$

Thus, $\mathbf{E}_\lambda \left[(I_k(1))^2 \right]$ can be written as

$$\mathbf{E}_\lambda \left[(I_k(1))^2 \right] = \mathbf{E}_\lambda \left[(\omega(B) - \lambda S(B))^{2k} + \sum_{2 \leq i+j \leq 2k-1} c_{i,j} \omega(B)^i (\lambda S(B))^j \right],$$

where the $c_{i,j}$ are integer constants.

If we differentiate the k-th central moment

$$\mathbf{E}[(N - \lambda')^k] = \sum_{i=0}^{\infty} (r - \lambda')^k e^{-\lambda'} \frac{(\lambda')^i}{i!}$$

of a random variable N distributed as Poisson with mean λ', with respect to λ' we find the following recurrence:

$$\mathbf{E}[(N - \lambda')^{k+1}] = \lambda' \left(\frac{d\mathbf{E}[(N - \lambda')^k]}{d\lambda'} + k\mathbf{E}[(N - \lambda')^{k-1}] \right).$$

Hence, using induction we can show that $\mathbf{E}[(N - \lambda')^k]$ is a polynomial on λ with maximum degree $\lfloor k/2 \rfloor$, for $k > 1$. Since $\mathbf{E}_\lambda \left[\omega(B)^i \right]$ is the Bell Polynomial of degree i on λ, it follows straightforwardly that the polynomial

$$\mathbf{E}_\lambda \left[\sum_{2 \leq i+j \leq 2k-1} c_{i,j} \omega(B)^i (\lambda S(B))^j \right]$$

has degree at most $2k - 1$, and the proof is thus complete. □

Definition 10. *Let f_i, g_j and h_k be, respectively, functions of i-th, j-th and k-th chaos of the Wiener-Poisson decomposition of some square integrable function of ω. For $0 \leq s \leq 2(n \wedge m)$, we define*

$$f_i \star_s g_j = \sum_{s \leq 2n \leq 2(s \wedge i \wedge j)} n! \binom{i}{n} \binom{j}{n} f_i \circ_n^{s-n} g_j.$$

For $0 \leq r \leq 2((i + j - s) \wedge k)$, we abuse of the notation to write

$$h_k \star_r (f_n \star_s g_m) = h_k \star_r f_n \star_s g_m.$$

Lemma 30. *If $|f_n(x_1, \cdots, x_n)|$ is bounded by a positive real c, then*

$$\mathbf{E}_\lambda \left[I_n(f_n)^2 \right] \leq c^2 \mathbf{E}_\lambda \left[I_n(1)^2 \right].$$

Proof. We use the isometry formula given by Eq. (2), so

$$\begin{aligned} \mathbf{E}_\lambda \left[I_n(f_n)^2 \right] &= n! \|f_n\|_{L^2(\lambda)^{\circ n}} \\ &= n! \int_{B^n} f_n^2(x_1, \cdots, x_n) \lambda \, \mathrm{d}x_1 \dots \lambda \, \mathrm{d}x_n \\ &\leq n! \int_{B^n} c^2 \lambda \, \mathrm{d}x_1 \dots \lambda \, \mathrm{d}x_n \\ &= c^2 \mathbf{E}_\lambda \left[I_n(1)^2 \right], \end{aligned}$$

and the proof is complete. □

Theorem 31. *Let* $F = \frac{G_\Gamma - \mathbf{E}_\lambda[G_\Gamma]}{\sqrt{\mathbf{V}_\lambda[G_\Gamma]}}$*, then, for* λ *large enough,*

$$\int_B \mathbf{E}_\lambda \left[|D_t F|^2 |D_t L^{-1} F|\right] \lambda \ dt \leq \frac{c}{\lambda^{1/2}}.$$

Proof. Provided that G_Γ has n points, Lemma 27 shows that $G_\Gamma = \sum_{i=0}^n I_n(f_n)$, so

$$D_t F = \frac{1}{\sqrt{\mathbf{V}_\lambda[G_\Gamma]}} \sum_{i=1}^n i I_{i-1}(f_i(*,t)),$$

$$D_t L^{-1} F = \frac{1}{\sqrt{\mathbf{V}_\lambda[G_\Gamma]}} \sum_{i=1}^n I_{i-1}(f_i(*,t)).$$

Let us define

$$g_{i-1} = \frac{f_i(*,t)}{\lambda^{n-i}}.$$

According to Eq. (7), we note that g_i does not depend on λ. Using the triangular inequality, we have

$$|D_t F|^2 |D_t L^{-1} F| \quad \leq \quad \sum_{i,j,k=0}^{n-1} \frac{\lambda^{3n-3-i-j-k}(i+1)(j+1)}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} |I_i(g_i) I_j(g_j) I_k(g_k)|.$$

Then, we apply twice the chaos expansion and use again the triangular inequality to obtain:

$$|D_t F|^2 |D_t L^{-1} F| \leq \sum_{i,j,k=0}^{n-1} \sum_{s=0}^{2(i\wedge j)} \sum_{r=0}^{2((i+j-s)\wedge k)} \frac{\lambda^{3n-3-i-j-k}(i+1)(j+1)}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} \times \\ |I_{i+j+k-s-r}(g_i \star_r g_j \star_s g_k)|,$$

Since f_i is bounded, g_i is bounded as so $g_i \star_r g_j \star_s g_k$ for i,j,k,r,s in the range of their indexes above. We define

$$c(i,j,k,r,s) = \sup\{g_i \star_r g_j \star_s g_k\}(i+1)(j+1),$$

and we use Jensen's inequality and Lemma 30 to write

$$\begin{aligned} &\mathbf{E}_\lambda \left[|D_t F|^2 |D_t L^{-1} F|\right] \\ &\leq \sum_{i,j,k=0}^{n-1} \sum_{s=0}^{2(i\wedge j)} \sum_{r=0}^{2((i+j-s)\wedge k)} \frac{\lambda^{3n-3-i-j-k}}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} \mathbf{E}_\lambda \left[(I_{i+j+k-s-r}(g_i \star_r g_j \star_s g_k))^2\right]^{\frac{1}{2}} \\ &\leq \sum_{i,j,k=0}^{n-1} \sum_{s=0}^{2(i\wedge j)} \sum_{r=0}^{2((i+j-s)\wedge k)} \frac{\lambda^{3n-3-i-j-k} c(i,j,k,r,s)}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} \mathbf{E}_\lambda \left[(I_{i+j+k-s-r}(1))^2\right]^{\frac{1}{2}}. \end{aligned}$$

Using Lemmas 28 and 29 we obtain:

$$\begin{aligned} \int_B \mathbf{E}_\lambda \left[|D_t F|^2 |D_t L^{-1} F|\right] \lambda \ \mathrm{d}t \quad &\leq \quad \sum_{i,j,k=0}^{n-1} \sum_{s=0}^{2(i\wedge j)} \sum_{r=0}^{2((i+j-s)\wedge k)} c(i,j,k,r,s) \times \\ &\quad \frac{\lambda^{3n-3-i-j-k}}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} \mathbf{E}_\lambda \left[(I_{i+j+k-s-r}(1))^2\right]^{1/2} \int_B \lambda \ \mathrm{d}t \\ &\leq \quad \frac{c\lambda^{3n-3}}{\mathbf{V}_\lambda[G_\Gamma]^{\frac{3}{2}}} \int_B \lambda \ \mathrm{d}t \leq \frac{c}{\lambda^{1/2}}, \end{aligned}$$

concluding the proof. □

Theorem 32. *Let*

$$F = \frac{G_\Gamma - \mathbf{E}_\lambda\left[G_\Gamma\right]}{\sqrt{Var(G_\Gamma)}}.$$

Then, when λ is large enough

$$\mathbf{E}_\lambda\left[\left|1 - \langle DF, DL^{-1}F\rangle_{L^2(\lambda)}\right|\right] \leq \frac{c}{\lambda^{1/2}},$$

for some constant c.

Proof. The expressions of D_tF and $D_tL^{-1}F$ are given by

$$D_tF = \frac{1}{\sqrt{\mathbf{V}_\lambda\left[G_\Gamma\right]}}\sum_{i=1}^n iI_{i-1}(f_i(*,t)),$$

and

$$D_tL^{-1}Z = \frac{1}{\sqrt{\mathbf{V}_\lambda\left[G_\Gamma\right]}}\sum_{i=1}^n I_{i-1}(f_i(*,t)).$$

The inner product $\langle D_tL^{-1}F, D_tF\rangle_{L^2(\lambda)}$ is expressed by:

$$\langle D_tL^{-1}F, D_tF\rangle_{L^2(\lambda)} = \frac{1}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\int_B \sum_{i,j=1}^n iI_{i-1}(f_i(*,t))I_{j-1}(f_j(*,t))\lambda\ \mathrm{d}t.$$

Then,

$$\begin{aligned}\langle D_tL^{-1}Z, D_tZ\rangle_{L^2(\lambda)} &= \frac{1}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\sum_{i,j=1}^n i\int_B I_{i-1}(f_i(*,t))I_{j-1}(f_j(*,t))\lambda\ \mathrm{d}t\\ &= \frac{1}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\int_B I_0(f_1(t)^2\lambda\ \mathrm{d}t\\ &+ \frac{1}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\sum_{\substack{i,j=1\\(i,j)\neq(1,1)}}^n i\int_B I_{i-1}(f_i(*,t))I_{j-1}(f_j(*,t))\lambda\ \mathrm{d}t.\end{aligned}$$

Defining g_{i-1} as in Theorem 31 and using the chaos expansion, we get:

$$\begin{aligned}\langle D_tL^{-1}F, D_tF\rangle_{L^2(\lambda)} &= \frac{\|f_1\|^2_{L^2(\lambda)}}{\mathbf{V}_\lambda\left[G_\Gamma\right]} + \frac{1}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\sum_{i=2}^n i(i-1)!\int_B \|f_n(*,t)\|^2_{L^2(\lambda)}\lambda\ \mathrm{d}t\\ &+ \sum_{i=1}^{n-1}(i+1)\frac{\lambda^{2n-2i-2}}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\int_B\sum_{s=0}^{2(i-1)} I_{2i-s}(g_i\star_s g_i)\lambda\ \mathrm{d}t\\ &+ \sum_{i,j=0}^{n-1}(i+1)\frac{\lambda^{2n-i-j-2}}{\mathbf{V}_\lambda\left[G_\Gamma\right]}\int_B\sum_{s=0}^{2(i\wedge j)} I_{i+j-s}(g_i\star_s g_j)\lambda\ \mathrm{d}t.\end{aligned}$$

Since

$$\begin{aligned}\int_B \|f_i(*,t)\|^2_{L^2(\lambda)}\lambda\ \mathrm{d}t &= \int_B\left(\int_{B^{i-1}} f_i^2(t_1,\cdots,t_{i-1},t)\lambda\ \mathrm{d}t_1\ldots\lambda\ \mathrm{d}t_{i-1}\right)\lambda\ \mathrm{d}t\\ &= \int_{B^i} f_i^2(t_1,\cdots,t_{i-1},t)\lambda\ \mathrm{d}t_1\ldots\lambda\ \mathrm{d}t_{i-1}\lambda\ \mathrm{d}t\\ &= \|f_i\|_{L^2(\lambda)},\end{aligned}$$

and given the isometry formula

$$\begin{aligned}\mathbf{V}_\lambda\left[G_\Gamma\right] &= \|G_\Gamma\|_{L^2\Omega} - \mathbf{E}_\lambda\left[G_\Gamma\right]^2\\ &= \sum_{i=0} n!\|f_i\|^2_{L^2(B^i)} - \|f_0\|^2 = \sum_{i=1} n!\|f_i\|^2_{L^2(\lambda)},\end{aligned}$$

we have

$$\frac{\|f_1\|^2_{L^2(\lambda)}}{\mathbf{V}_\lambda[G_\Gamma]} + \frac{1}{\mathbf{V}_\lambda[G_\Gamma]}\sum_{i=2}^{n} i(i-1)! \int_B \|f_i(*,t)\|^2_{L^2(\lambda)}\,\lambda\,\mathrm{d}t = 1.$$

Hence

$$\langle D_t L^{-1} Z, D_t Z\rangle_{L^2(\lambda)} = 1 + \sum_{\substack{i,j=0\\(i,j)\neq(1,1)}}^{n-1} \sum_{\substack{s=0\\ s\neq 2i \text{ if } i=j}}^{2(i\wedge j)} (i+1)\frac{\lambda^{2n-i-j-2}}{\mathbf{V}_\lambda[G_\Gamma]}\int_B I_{i+j-s}(g_i\star_s g_j)\lambda\,\mathrm{d}t.$$

Let $c(i,j,s)$ be defined as

$$c(i,j,s) = \sup\{g_i\star_s g_j\}(i+1).$$

Then, we use the triangular inequality, Jensen's inequality and Lemma 30 to obtain:

$$\begin{aligned}
&\mathbf{E}_\lambda\left[|1-\langle DF, DL^{-1}F\rangle_{L^2(\lambda)}|\right]\\
&\quad\le \mathbf{E}_\lambda\left[\sum_{\substack{i,j=0\\(i,j)\neq(1,1)}}^{n-1}\sum_{\substack{s=0\\ s\neq 2i \text{ if } i=j}}^{2(i\wedge j)} (i+1)\frac{\lambda^{2n-i-j-2}}{\mathbf{V}_\lambda[G_\Gamma]}\int_B |I_{i+j-s}(g_i\star_s g_j)|\,\lambda\,\mathrm{d}t\right]\\
&\quad\le \sum_{\substack{i,j=0\\(i,j)\neq(1,1)}}^{n-1}\sum_{\substack{s=0\\ s\neq 2i \text{ if } i=j}}^{2(i\wedge j)} \frac{(i+1)\lambda^{2n-i-j-2}}{\mathbf{V}_\lambda[G_\Gamma]}\int_B \mathbf{E}_\lambda\left[(I_{i+j-s}(g_i\star_s g_j))^2\right]^{\frac{1}{2}}\lambda\,\mathrm{d}t\\
&\quad\le \sum_{\substack{i,j=0\\(i,j)\neq(1,1)}}^{n-1}\sum_{\substack{s=0\\ s\neq 2i \text{ if } i=j}}^{2(i\wedge j)} c(i,j,s)\frac{\lambda^{2n-i-j-2}}{\mathbf{V}_\lambda[G_\Gamma]}\mathbf{E}_\lambda\left[(I_{i+j-s}(1))^2\right]^{1/2}\int_B \lambda\,\mathrm{d}t.
\end{aligned}$$

Finally, using Lemmas 29 and 28, there is a constant c such that:

$$\mathbf{E}_\lambda\left[|1-\langle DF, DL^{-1}F\rangle_{L^2(\lambda)}|\right] \le \frac{c}{\lambda^{1/2}}$$

for λ large enough. □

Theorem 33. *There exists a constant c such that, for λ large enough, the Wasserstein distance between $F = \frac{G_\Gamma - \mathbf{E}_\lambda[G_\Gamma]}{\sqrt{Var(G_\Gamma)}}$ and $\mathcal{N}(0,1)$ is given by:*

$$d_W(F,\mathcal{N}(0,1)) \le \frac{c}{\lambda^{1/2}}.$$

Proof. The proof comes straightforwardly from Theorem 3.1 as stated in [16]:

$$d_W(F,\mathcal{N}(0,1)) \le \mathbf{E}_\lambda\left[|1-\langle DF, DL^{-1}F\rangle_{L^2(\lambda)}|\right] + \int_B \mathbf{E}_\lambda\left[|D_tF|^2|D_tL^{-1}F|\right]\lambda\,\mathrm{d}t,$$

which we can use since f_n is bounded, so $F\in \operatorname{Dom} D$. We use theorems 31 and 32 in the first and second terms, respectively.

□

8. Discussion and Conclusion

Through this work, we successfully obtained the moments of the number of high order interactions, named simplices, and the two first order moments of the Euler characteristic. The assumptions are quite reasonable and are justified for real systems. The randomness of the positions reflect the lack of capacity of choosing the position where they are deployed, while the randomness of the number of sensors comprises various random factors that can happens with a sensor that suppose to be in the target region (for instance, the sensor can fall outside the region or can shut off due the lack of energy). The fact that they fall over a torus, can be justified

if we consider that we can warrant a total coverage in the box $[0,a]^d$ if the Euler characteristic is zero in the Torus, which would not be possible to find if we consider that points fall directly over $[0,a]^d$. Besides, the behavior in the torus or in the box is close from each other when we consider small cover of each point. We could criticize the use of the maximum norm, which is not physically realistic, but this norm allows us to find closed form formulas results and can represent bounds for the Euclidean one. If $\mathbf{E}_\lambda\left[N'_k\right]$ and $\mathbf{E}_\lambda\left[\chi'\right]$ represent, respectively, the mean number of k-simplices and the mean of the Euler characteristic, using the Euclidean norm we obtain:

$$\frac{\lambda^{k+1}a^d}{(k+1)!}\frac{(k+1)^d\left((2\epsilon)^k\right)^d}{\sqrt{d}^{d(k)}} \leq \mathbf{E}_\lambda\left[N'_k\right] \leq \frac{\lambda^{k+1}a^d}{(k+1)!}(k+1)^d\left((2\epsilon)^k\right)^d$$

$$\left(\frac{a\sqrt{d}}{2\epsilon}\right)^d e^{-\lambda(2\epsilon)^d}\left(-B_d\left(-\lambda\frac{(2\epsilon)^d}{\sqrt{d}^d}\right)\right) \leq \mathbf{E}_\lambda\left[\chi'\right] \leq \left(\frac{a}{2\epsilon}\right)^d e^{-\lambda(2\epsilon)^d}(-B_d(-\lambda(2\epsilon)^d))$$

From the properties obtained of the Euler characteristic mean and based in some simulations, we conjecture that, in a random simplicial complex as defined in this work, there is always two main kinds of holes in this complex, β_i and β_{i+1}. So, for instance, consider figure 8 where points are placed in 5 dimensions. When λ is small, in average, the components are isolated from the others, so $\beta_0 > 0$ and we do not have other kinds of holes, so β_0 is the dominating Betti number. If λ is increased (which means, in average, increase the number of points), the components connect with each other, decreasing β_0 and we begin to form the first cycles, so β_0 and β_1 are the dominating Betti numbers. Increasing λ even more, the complex becomes completely connected and we have a large number of cycles. For λ even larger, those cycles begin to vanish and we have the first voids. Then, we follow this reasoning until all the region is completely covered and $\chi = 0$.

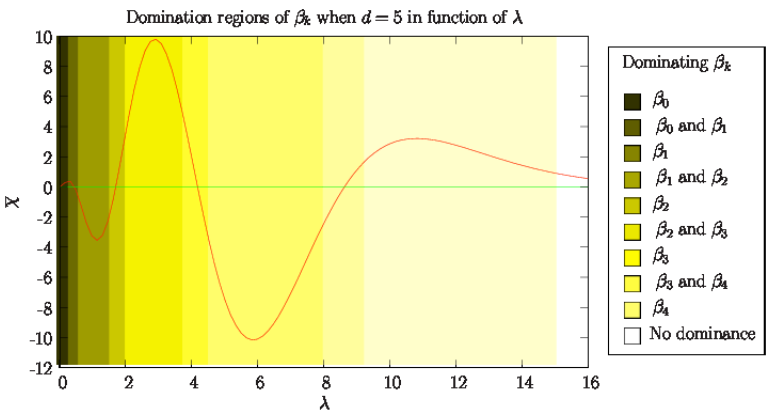

FIGURE 7. Behavior of χ with respect to λ.

References

[1] M. A. Armstrong, *Basic topology*, Springer-Verlag, 1997.

[2] E. T. Bell, *Exponential polynomials*, Ann, Math. (1934), no. 35, 258–277.

[3] A. Björner, *Topological methods*, (1995), 1819–1872.

[4] C.-Y. Chong and S. P. Kumar, *Sensor networks: Evolution, opportunities, and challenges*, Proceedings of IEEE, vol. 91, august 2003, pp. 1247–1256.

[5] D. Daley and D. Vere Jones, *An introduction to the theory of point processes*, Springer, 2002.

[6] V. de Silva and R. Ghrist, *Coordinate-free coverage in sensor networks with controlled boundaries via homology*, International Journal of Robotics Research **25** (2006).

[7] L. Decreusefond and E. Ferraz, *On the poisson random geometric graph: the one dimensional case*, 2010.

[8] R. Ghrist, *Coverage and hole-detection in sensor networks via homology*, Fouth International Conference on Information Processing in Sensor Networks (IPSN'05), UCLA, 2005, pp. 254–260.

[9] M. J. Greenberg and J. R. Harper, *Algebraic topology: a first course*, Addison-Wesley, New York, NY, 1981.

[10] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002.

[11] Y. Ito, *Generalized poisson functionals*, Probab. Theory Related Fields (1988), no. 77, 1–28.

[12] M. Kahle, *Random geometric complexes*, Discrete Comput. Geom. **45** (2011), no. 3, 553–573.

[13] J.M. Kahn, R.H. Katz, and K.S.J. Pister, *Mobile networking for smart dust*, Intl. Conf. on Mobile Computing and Networking (Seattle, WA), august 1999.

[14] F.L. Lewis, *Wireless sensor networks*, ch. 2, John Wiley, New York, 2004.

[15] J. Munkres, *Elements of algebraic topology*, Addison Wesley, 1993.

[16] G. Peccati, J.L. Solé, M.S. Taqqu, and F. Utzet, *Stein's method and normal approximation of poisson functionals*, Annals of Probability **38** (2010), no. 2, 443–478.

[17] G.J. Pottie and W.J. Kaiser, *Wireless integrated network sensors*, Communications of the ACM **43** (2000), 51–58.

[18] N. Privault, *Stochastic analysis in discrete and continuous settings*, Springer, 2009.

[19] J.J. Rotman, *An introduction to algebraic topology (graduate texts in mathematics)*, Springer, July 1998.

E-mail address: `laurent.decreusefond@telecom-paristech.fr`

E-mail address: `eduardo.ferraz@telecom-paristech.fr`

E-mail address: `hugues.randriam@telecom-paristech.fr`

E-mail address: `anais.vergne@telecom-paristech.fr`

Institut Telecom, Telecom ParisTech, CNRS LTCI, Paris, France