



**HAL**  
open science

## Simplicial Homology of Random Configurations

Laurent Decreasefond, Eduardo Ferraz, Hugues Randriambololona, Anaïs Vergne

► **To cite this version:**

Laurent Decreasefond, Eduardo Ferraz, Hugues Randriambololona, Anaïs Vergne. Simplicial Homology of Random Configurations. *Advances in Applied Probability*, 2014, 46 (2), pp.1-23. hal-00578955v4

**HAL Id: hal-00578955**

**<https://imt.hal.science/hal-00578955v4>**

Submitted on 4 Jul 2013

**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 three first order moments of the number of  $k$ -simplices, and provide a way to compute higher order moments. Then, we derive the mean and the variance of the Euler characteristic. Using the Stein method, we estimate the speed of convergence of the number of occurrences of any connected subcomplex converges towards the Gaussian law when the intensity of the Poisson point process tends to infinity. We use a concentration inequality for Poisson processes to find bounds for the tail distribution of the Betti number of first order and the Euler characteristic in such simplicial complexes.

## 1. MOTIVATION

Algebraic topology is the domain of mathematics in which the topological properties of a set are analyzed through algebraic tools. Initially developed for the classification of manifolds, it is by now heavily used in image processing and geometric data analysis. More recently, applications to sensor networks were developed in [7, 11]. Imagine that we are given a bounded domain in the plane and sensors which can detect intruders within a fixed distance. The so-called coverage problem consists in determining whether the domain is fully covered, i.e. whether there is any part of the domain in which an intrusion can occur without being detected. The mathematical set which is to be analyzed here is the union of the balls centered on each sensor. If this set has no hole then the domain is fully covered. It turns out that algebraic topology provides a computationally effective procedure to determine whether this property holds. In view of the rapid development of the

---

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

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

L. Decreusefond was partially supported by ANR Masterie.

The authors would like to thank the anonymous referee for his/her thorough reading which helped us to improve the presentation of the current article.

technology of sensor networks [17, 18, 24], which are small and cheap devices with limited capacity of autonomy and communications, devoted to measure some local physical quantity (temperature, humidity, intrusion, etc.), this kind of question is likely to become recurrent.

The coverage problem, via homology techniques, for a set of sensors was first addressed in the papers [7, 11]. The method consists in calculating from the geometric data, a combinatorial object known as a simplicial complex which is a list of points, edges, triangles, tetrahedron, etc. satisfying some compatibility conditions: all the faces of a  $k$ -simplex ( $k = 0$  means points,  $k = 1$  means edges, etc.) of the complex must belong to the set of  $(k - 1)$ -simplices of the complex. Then, an algebraic structure on these lists and linear maps, known as boundary operators, are constructed. Some of the topological properties (like connectivity and coverage) are given by the so-called Betti numbers which mathematically speaking are dimension of some quotient vector spaces (see below). Another key parameter is the alternated sum of the Betti numbers, known as Euler characteristic which gives some information on the global topology of the studied set. Persistence homology [5, 10] is both a way to compute the Betti numbers avoiding a (frequent) combinatorial explosion and a way to detect the robustness of the topological properties of a set with respect to some parameter: For instance, in the intrusion detection setting, how the connectivity of the covering domain is altered by variations of the detection distance.

When points (i.e. sensors) are randomly located in the ambient space, may it be  $\mathbb{R}^d$  or a manifold, it is natural to ask about the statistical properties of the Betti numbers and the Euler characteristic. We completely solved the problem in one dimension (see [8]) by basic methods inspired by queuing theory, without using the forthcoming sophisticated tools of algebraic topology. Since we cannot order points in  $\mathbb{R}^d$ , it is not possible to generalize the results obtained in this earlier work to higher dimension.

A very few papers deal with the properties of random simplicial complexes. In [15] and [22], for Binomial point processes whose number of points are going to infinity, the asymptotic regimes of the mean value of the Betti numbers and simplices numbers are investigated. In [16], these results are refined by providing Poisson and Gaussian approximations of the Betti numbers in asymptotic regimes.

As will be apparent below, for the Rips-Vietoris simplicial complex, the number of  $k$ -simplices boils down to the number of  $(k + 1)$ -cliques of the underlying graph.

As we mainly analyze this kind of complex, our work has thus strong links with the pioneering work of Penrose [22] and with [3, 23] as well. In [3], for Poisson input, the limiting regimes of the number of  $k$ -simplices on a square of size  $a$ , are investigated through limit theorems of U-statistics. The size of the square is growing to infinity with a constant mean number of points per unit of surface. In [23], there is an extension of the latter result to non-linear manifolds and non-uniform distributions of the points. In the above cited papers, sophisticated combinatorial arguments are at the root of the arguments of the convergence theorems. We here replace this line of thought by a functional analytic approach which transfers the difficulty to the computation of a (possibly involved) deterministic integral. By doing so, we can, in principle, obtain a CLT for the number of occurrences of any connected sub-complex and not only for cliques.

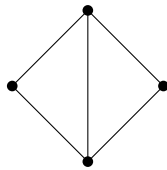


FIGURE 1. A sub-complex which is not a clique

One of our main contributions are exact formulas for the first three moments of the number of simplices for both the Poisson and the Binomial processes at the price of working on a square bounded domain, which we embed into a torus in order to avoid side effects. The rationale behind this simplification is that when the size of the covering balls is small compared to the size of the square, the topology of the two sets (the union of balls in the square and the union of balls in the corresponding torus) must be similar. Our method could be generalized to compute the moments of any order but the computations become more and more tricky as the order increases. We also investigate the properties of the moments of the Euler characteristic. Moreover, by using Malliavin calculus, we go further than the previously cited works since we can evaluate the speed of convergence in the CLT. We also give a concentration inequality to bound the distribution tail of the first Betti number.

During the final preparation of this paper, we learned that such an approach was used independently in [26] for the analysis of  $U$ -statistics (functionals of a fixed

chaos in our vocabulary) of Poisson processes. Both approaches rely on the ideas which appeared in [9, 20].

Our method goes as follows: We write the numbers of  $k$ -simplices (i.e. points, edges, triangles, tetrahedron, etc) as iterated integrals with respect to the underlying Poisson process. Then, the computation of the means is reduced to the computation of deterministic iterated integrals thanks to Campbell formula. By using the definition of the Euler characteristic as an alternating sum of the numbers of simplices, we find its expectation. The point is that even if the summing index goes to infinity, the expectation of  $\chi$  depends only on the  $d$ -th power of the intensity of the Poisson process where  $d$  is the dimension of the underlying space. By de-poissonization, we obtain the exact values of the mean number of simplices of any order and then the mean value of the Euler characteristic for Binomial processes. Using the multiplication formula of iterated integrals, one can reproduce the same line of thought for higher order moments to the price of an increased complexity in the computations. We obtain closed form formulas for the variance of the number of  $k$ -simplices and of the Euler characteristic and a series expansion for third order moments. Our method is a priori suitable for any higher order moments but the computations become much involved. Using Stein's method mixed with Malliavin calculus, we generalize the results of [22] by proving a precise (i.e. with speed of convergence) CLT for sub-complexes count. As expected, the speed of convergence is of the order of  $\lambda^{-1/2}$ .

The paper is organized as follows: 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 characteristic are computed. This is sufficient to bound the tail distribution of  $\beta_0$  using concentration inequality. Section 5 applies the Malliavin calculus in order to find the explicit expression of second order moments of the number of  $k$ -simplices and the Euler characteristic. Using the same strategy, in Section 6, we find the expression for the third order moment of the number of simplices. In Section 7, we prove a central limit theorem for the number of occurrences of a finite simplex into a Poisson random geometric complex.

## 2. ALGEBRAIC TOPOLOGY

For further reading on algebraic topology, see [1, 13, 19]. Graphs can be generalized to more generic combinatorial objects known as simplicial complexes. While graphs model binary relations, simplicial complexes represent higher order relations.

Given a set of vertices  $V$ , a  $k$ -simplex is an unordered subset  $\{v_0, v_1, \dots, v_k\}$  where  $v_i \in V$  and  $v_i \neq v_j$  for all  $i \neq j$ . The faces of the  $k$ -simplex  $\{v_0, v_1, \dots, v_k\}$  are defined as all the  $(k - 1)$ -simplices of the form  $\{v_0, \dots, v_{j-1}, v_{j+1}, \dots, v_k\}$  with  $0 \leq j \leq k$ . A simplicial complex  $\mathcal{C}$  is a collection of simplices which is closed with respect to the inclusion of faces, i.e. if  $\{v_0, v_1, \dots, v_k\}$  is a  $k$ -simplex then all its faces are in the set of  $(k - 1)$ -simplices.

One can define an orientation on simplices by defining an order on vertices and with the convention that:

$$[v_0, \dots, v_i, \dots, v_j, \dots, v_k] = -[v_0, \dots, v_j, \dots, v_i, \dots, v_k],$$

for  $0 \leq i, j \leq k$ .

For each integer  $k$ ,  $\mathcal{C}_k$  is the vector space spanned by the set of oriented  $k$ -simplices of  $\mathcal{V}$ . For any integer  $k$ , the boundary map  $\partial_k$  is the linear transformation  $\partial_k : \mathcal{C}_k \rightarrow \mathcal{C}_{k-1}$  which acts on basis elements  $[v_0, \dots, v_k]$  as

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

and  $\partial_0$  is the null function. Examples of such operations are given in Table 1.

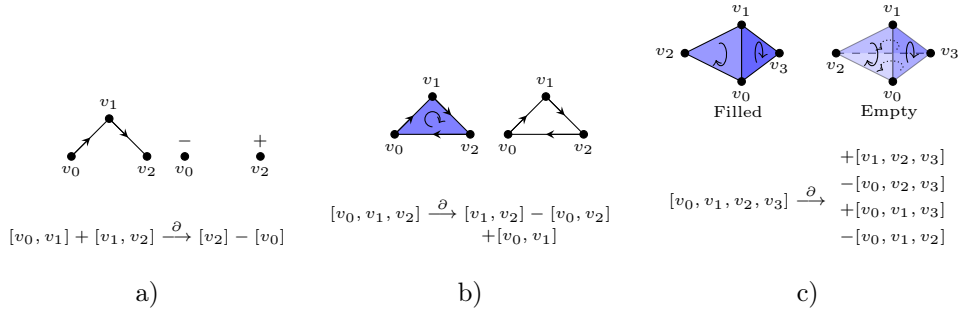


TABLE 1. Examples of boundary maps. From left to right. An application over 1-simplices. Over a 2-simplex. Over a 3-simplex, turning a filled tetrahedron to an empty one.

These maps give rise to a chain complex: a sequence of vector spaces and linear transformations:

$$\dots \xrightarrow{\partial_{k+2}} \mathcal{C}_{k+1} \xrightarrow{\partial_{k+1}} \mathcal{C}_k \xrightarrow{\partial_k} \dots \xrightarrow{\partial_2} \mathcal{C}_1 \xrightarrow{\partial_1} \mathcal{C}_0.$$

A standard result then asserts that for any integer  $k$ ,  $\partial_k \circ \partial_{k+1} = 0$ . If one defines  $Z_k = \ker \partial_k$  and  $B_k = \text{im } \partial_{k+1}$ , this induces that  $B_k \subset Z_k$ .

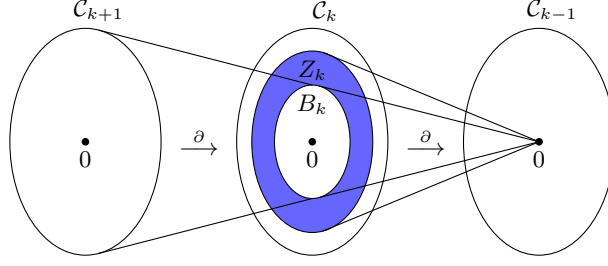


FIGURE 2. A chain complex showing the sets  $\mathcal{C}_k$ ,  $Z_k$  and  $B_k$ .

The  $k$ -th homology group of  $\mathcal{C}$ , denoted by  $H_k$ , is the quotient vector space,  $H_k = Z_k/B_k$  and the  $k$ -th Betti number of  $\mathcal{C}$  is its dimension:  $\beta_k = \dim H_k = \dim Z_k - \dim B_k$ .

The simplicial complexes we consider are of a special type since they are built on topological rules.

**Definition 1.** Given  $\mathcal{U} = (U_v, v \in \omega)$  a collection of open sets of some topological space  $X$ , 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  $\{v_0, v_1, \dots, v_k\}$  is a  $k$ -simplex if and only if  $\bigcap_{i=0}^k U_{v_i} \neq \emptyset$ .

For the applications we have in mind, the set  $U_v$  is meant to be the covered zone by the sensor located at  $v$ . For the sake of tractability, it is supposed to be a ball centered at  $v$  with a fixed radius. As said earlier, in order to avoid side effects, we work on the  $d$ -dimensional torus of length  $a$ , denoted by  $\mathbb{T}_a^d$  and to simplify the computations, we consider the  $l^\infty$  distance. Namely, the torus is defined as the quotient of the action of the group of translations  $a\mathbb{Z}^d$  on  $\mathbb{R}^d$ , i.e.  $\mathbb{T}_a^d = \mathbb{R}^d/a\mathbb{Z}^d$ . The space  $X = [0, a)^d$  can be embedded in  $\mathbb{T}_a^d$  as a fundamental domain of this action. If we equip  $X$  with the distance

$$\rho_d(x, y) = \inf_{k \in \mathbb{Z}^d} \|x - y + ka\|_\infty$$

where  $\|x\|_\infty$  is the  $l^\infty$ -norm in  $\mathbb{R}^d$ , then the embedding of  $X$  into the torus is a bijective isometry. One can thus identify these two spaces and use the representation which is the most convenient according to the situation.

**Definition 2.** Given  $\omega$  a finite set of points on the torus  $\mathbb{T}_a^d$ . For  $\epsilon > 0$ , we define  $\mathcal{U}_\epsilon(\omega) = \{B_{\rho_d}(v, \epsilon), v \in \omega\}$  and  $\mathcal{C}_\epsilon(\omega) = \mathcal{C}(\mathcal{U}_\epsilon(\omega))$ , where  $B_{\rho_d}(x, \epsilon) = \{y \in \mathbb{T}_a^d, \rho_d(x, y) < \epsilon\}$ .

The following result is known for  $\mathbb{R}^d$ , there is a slight modification of the proof for the torus.

**Theorem 1.** *Suppose  $\epsilon < a/4$ . Then  $\mathcal{C}_\epsilon(\omega)$  has the same homology vector spaces 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 [27, Theorem 7.26] or [4, Theorem 10.7]. One only needs to check that any non-empty intersection of sets  $B_{\rho_d}(v, \epsilon)$  is contractible.

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

For any finite sets of points  $\omega$  of the torus, according to the geometrical definition of the Čech complex, the Betti numbers have a geometrical meaning:  $\beta_0(\omega)$  (with obvious notations) is the number of connected components and for  $k \geq 1$ ,  $\beta_k(\omega)$  is the number of  $k$ -dimensional holes of  $\mathcal{U}_\epsilon(\omega)$ . For  $k = 0$  and  $k = 1$ , an intuitive explanation can be given. By definition,  $\beta_0(\omega)$  is the number of points minus the number of independent edges. Each time there exists a cycle with  $n$  points, we can remove an edge without altering  $\beta_0(\omega)$  since there are  $n - 1$  independent edges in such a cycle. Doing this repeatedly, one can reduce the original graph to as many linear chains of edges as there are connected components. A linear chain of edges which contains  $n$  points has  $n - 1$  edges, hence a  $\beta_0(\omega)$  equal to 1. Thus  $\beta_0(\omega)$  counts the number of connected components. As to  $\beta_1(\omega)$ , we remark that  $\ker \partial_1$  is composed by the cycles and that  $B_1$  is the set of linear combinations of edges forming triangles, hence  $\beta_1(\omega)$  is the number of cycles which are not triangles, hence represents the number of “coverage” holes. The well known topological invariant named Euler characteristic for  $\mathcal{U}_\epsilon(\omega)$ , denoted by  $\chi(\omega)$ , is an integer defined by:

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

Let  $s_k(\omega)$  be the number of  $k$ -simplices in the simplicial complex  $\mathcal{C}_\epsilon(\omega)$ . A well known theorem states that this is also given by:

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



**Definition 3.** Let  $\omega$  be a finite set of points in  $\mathbb{T}_a^d$ . For any  $\epsilon > 0$ , the Rips-Vietoris complex of  $\omega$ ,  $\mathcal{R}_\epsilon(\omega)$ , is the abstract simplicial complex whose  $k$ -simplices correspond to unordered  $(k + 1)$ -tuples of points in  $\omega$  which are pairwise within distance less than  $\epsilon$  of each other.

**Lemma 2.** For the torus  $\mathbb{T}_a^d$  equipped with the product distance  $\rho_d$ , the Rips-Vietoris complex  $\mathcal{R}_{2\epsilon}(\omega)$  has the same Betti numbers as the Čech complex  $\mathcal{C}_\epsilon(\omega)$ .

The proof is given in [11] 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 homology vector spaces of both complexes coincide.

**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 [12, Theorem 22.24].  $\square$

**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)  $\beta_d(\omega) = 0$ , or
- ii)  $\beta_d(\omega) = 1$ .

If the latter condition holds, then we also have  $\chi(\omega) = 0$ .

*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 sub-manifold of the torus, so there are only two possibilities:

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

In the first case,  $\beta_d(\omega) = 0$  by [12, Corollary 22.25]. In the second case  $\mathcal{C}_\epsilon(\omega)$  has same homology as the torus, hence  $\beta_d(\omega) = 1$  and  $\chi(\omega) = 0$ .  $\square$

### 3. POISSON POINT PROCESS AND MALLIAVIN CALCULUS

The space of configurations on  $X = [0, a)^d$ , is the set of locally finite simple point measures (see [6, 25] for details):

$$\Omega^X = \left\{ \omega = \sum_{k=1}^n \delta(x_k) : (x_k)_{k=1}^{k=n} \subset X, n \in \mathbb{N} \cup \{\infty\} \right\},$$

where  $\delta(x)$  denotes the Dirac measure at  $x \in X$ . It is often convenient to identify an element  $\omega$  of  $\Omega^X$  with the set corresponding to its support, i.e.  $\sum_{k=1}^n \delta(x_k)$  is identified with the unordered set  $\{x_1, \dots, x_n\}$ . For  $A \in \mathcal{B}(X)$ , we have  $\delta(x)(A) = \mathbf{1}_A(x)$ , so that

$$\omega(A) = \sum_{x \in \omega} \mathbf{1}_A(x),$$

counts the number of atoms in  $A$ . Simple measure means that  $\omega(\{x\}) \leq 1$  for any  $x \in X$ . Locally finite means that  $\omega(K) < \infty$  for any compact  $K$  of  $X$ . The configuration space  $\Omega^X$  is endowed with the vague topology and its associated  $\sigma$ -algebra denoted by  $\mathcal{F}^X$ . To characterize the randomness of the system, we consider that the set of points is represented by a Poisson point process  $\omega$  with intensity measure  $d\Lambda(x) = \lambda dx$  in  $X$ . The parameter  $\lambda$  is called the intensity of the Poisson process. Since  $\omega$  is a Poisson point process of intensity measure  $\Lambda$ :

i) For any compact  $A$ ,  $\omega(A)$  is a random variable of parameter  $\Lambda(A)$ :

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

ii) For any disjoint sets  $A, A' \in \mathcal{B}(X)$ , the random variables  $\omega(A)$  and  $\omega(A')$  are independent.

Along this paper, we refer  $\mathbf{E}_\Lambda [F]$  as the mean of some function  $F$  depending on  $\omega$  given that the intensity measure of this process is  $\Lambda$ . The notations  $\text{Var}_\Lambda [F]$  and  $\text{Cov}_\Lambda [F, G]$  are defined accordingly. As said above, a configuration  $\omega$  can be viewed as a measure on  $X$ . It also induces a measure on any  $X^n$ , called the factorial measure associated to  $\omega$  of order  $n$ , defined by

$$\omega^{(n)}(C) = \sum_{\substack{(x_1, \dots, x_n) \in \omega \\ x_i \neq x_j}} \mathbf{1}_C(x_1, \dots, x_n),$$

for any  $C \in X^n$ , with the convention that  $\omega^{(n)}$  is the null measure if  $\omega$  has less than  $n$  points. Let  $f \in L^1(\Lambda^{\otimes n})$  and let  $F$  be a random variable given by

$$F(\omega) = \sum_{\substack{x_i \in \omega \\ x_i \neq x_j}} f(x_1, \dots, x_n) = \int f(x_1, \dots, x_n) d\omega^{(n)}(x_1, \dots, x_n).$$

The Campbell-Mecke formula for Poisson point processes states that

$$\mathbf{E}_\Lambda [F] = \int_{X^n} f(x_1, \dots, x_n) d\Lambda(x_1) \dots d\Lambda(x_n).$$

In view of this result, it is natural to introduce the compensated factorial measures defined by :

$$d\omega_\Lambda^{(1)}(x) = d\omega(x) - d\Lambda(x)$$

and for  $n \geq 2$ , for any  $f \in L^1(\Lambda^{\otimes n})$ ,

$$\begin{aligned} & \int f(x_1, \dots, x_n) d\omega_\Lambda^{(n)}(x_1, \dots, x_n) \\ &= \int \left( \int f(x_1, \dots, x_n) (d(\omega - \sum_{j=1}^{n-1} \delta(x_j))(x_n) - d\Lambda(x_n)) \right) \\ & \hspace{15em} d\omega_\Lambda^{(n-1)}(x_1, \dots, x_{n-1}). \end{aligned}$$

A real-valued function  $f : X^n \rightarrow \mathbb{R}$  is called symmetric if

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

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

$$I_n(f_n)(\omega) = \int f_n(x_1, \dots, x_n) d\omega_\Lambda^{(n)}(x_1, \dots, x_n).$$

It is known that  $I_n(f_n) \in L^2(\Omega^X, \mathbf{P})$ . Moreover, if  $f_n \in L^2(X, \Lambda)^{\circ n}$  and  $g_m \in L^2(X, \Lambda)^{\circ m}$ , the isometry formula

$$(1) \quad \mathbf{E}_\Lambda [I_n(f_n)I_m(g_m)] = n! \mathbf{1}_m(n) \langle f_n, g_m \rangle_{L^2(X, \Lambda)^{\circ n}}$$

holds true. Furthermore, we have:

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

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

where the series converges in  $L^2(\Omega^X, \mathbf{P})$  and, for each  $n \geq 1$ , the kernel  $f_n$  is an element of  $L^2(X, \Lambda)^{\circ n}$ . Moreover, by definition  $\text{Var}_\Lambda [F] = \|F - \mathbf{E}_\Lambda [F]\|_{L^2(\Omega^X, \mathbf{P})}^2$  then we have the isometry

$$(2) \quad \text{Var}_\Lambda [F] = \sum_{n=1}^{\infty} n! \|f_n\|_{L^2(X, \Lambda)^{\circ n}}^2.$$

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

$$(3) \quad (y_{l+1}, \dots, y_n, x_{k+1}, \dots, x_m) \mapsto \int_{X^l} f_n(y_1, \dots, y_n) g_m(y_1, \dots, y_k, x_{k+1}, \dots, x_m) d\Lambda(y_1) \dots d\Lambda(y_l).$$

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 (see [25] for a proof):

**Proposition 6.** For  $f_n \in L^2(X, \Lambda)^{\circ n}$  and  $g_m \in L^2(X, \Lambda)^{\circ m}$ , 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  $L^2(X, \Lambda)^{\circ n+m-s}$ ,  $0 \leq s \leq 2(m \wedge n)$ .

In what follows, given  $f \in L^2(X, \Lambda)^{\circ q}$  ( $q \geq 2$ ) and  $t \in X$ , we denote by  $f(*, x)$  the function on  $X^{q-1}$  given by  $(x_1, \dots, x_{q-1}) \mapsto f(x_1, \dots, x_{q-1}, x)$ .

**Definition 4.** Let  $\text{Dom } D$  be the set of random variables  $F \in L^2(\Omega^X, \mathbf{P})$  admitting a chaotic decomposition such that

$$\sum_{n=1}^{\infty} n! \|f_n\|_{L^2(X, \Lambda)^{\circ n}}^2 < \infty.$$

Let  $D$  be defined by

$$D : \text{Dom } D \rightarrow L^2(\Omega^X \times X, \mathbf{P} \otimes \Lambda)$$

$$F = \mathbf{E}_{\Lambda}[F] + \sum_{n \geq 1} I_n(f_n) \mapsto D_x F = \sum_{n \geq 1} n I_{n-1}(f_n(*, x)).$$

It is known, cf. [14], that we also have

$$D_x F(\omega) = F(\omega \cup \{x\}) - F(\omega), \quad \mathbf{P} \otimes \Lambda - a.e.$$

**Definition 5.** The Ornstein-Uhlenbeck operator  $L$  is given by

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

whenever  $F \in \text{Dom } L$ , given by those  $F \in L^2(\Omega^X, \mathbf{P})$  such that their chaos expansion verifies

$$\sum_{n=1}^{\infty} n^2 \|f_n\|_{L^2(X, \Lambda)^{\circ n}}^2 < \infty.$$

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

**Definition 6.** For  $F \in L^2(\Omega^X, \mathbf{P})$  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 [21]:

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

$$d_W(F, \mathcal{N}(0, 1)) \leq \mathbf{E}_\Lambda \left[ \left| 1 + \int_X [D_x F \times D_x L^{-1} F] d\Lambda(x) \right| \right] \\ + \int_X \mathbf{E}_\Lambda \left[ |D_x F|^2 |D_x L^{-1} F| \right] d\Lambda(x).$$

Another result from the Malliavin calculus used in this work is the following one, quoted from [25]:

**Theorem 8.** *Let  $F \in \text{Dom } D$  be such that  $DF \leq K$ , a.s., for some  $K \geq 0$  and we denote*

$$\|DF\|_{L^\infty(L^2(X, \Lambda), \mathbf{P})} := \sup_\omega \int_X |D_x F(\omega)|^2 d\Lambda(x) < \infty.$$

Then

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

#### 4. FIRST ORDER MOMENTS

Let  $\omega$  denote a generic realization of a Poisson point process on the torus  $\mathbb{T}_a^d$  and  $\mathcal{C}_\epsilon(\omega)$  the associated Čech complex with  $\epsilon < a/4$ . A Poisson process in  $\mathbb{R}^d$  of intensity  $\lambda$  dilated by a factor  $\alpha$  is a Poisson process of intensity  $\lambda\alpha^{-d}$ . Hence, statistically, the homological properties of a Poisson process of intensity  $\lambda$ , inside a torus of length  $a$  with ball sizes  $\epsilon$  are the same as that of a Poisson process of intensity  $\lambda\alpha^{-d}$ , inside a torus of length  $\alpha a$  with ball sizes  $\alpha\epsilon$ . Thus there are only two degrees of freedom among  $\lambda$ ,  $a$ , and  $\epsilon$ . For instance, we can set  $a = 1$  and the general results are obtained by a multiplication of magnitude  $a^d$ . Strictly speaking, Betti numbers, Euler characteristic and number of  $k$ -simplices are functions of  $\mathcal{C}_\epsilon(\omega)$  but we will skip this dependence for the sake of notations. We also define  $N_k$  as the number of  $(k - 1)$ -simplices.

In this section, we evaluate the mean of the number of  $(k - 1)$ -simplices  $\mathbf{E}_\Lambda [N_k]$  and the mean of the Euler characteristic  $\mathbf{E}_\Lambda [\chi]$ . We introduce some notations. Let

$$\Delta_k^{(d)} = \{(v_1, \dots, v_k) \in ([0, a)^d)^k, v_i \neq v_j, \forall i \neq j\}.$$

For any integer  $k$ , we define  $\varphi_k^{(d)}$  as:

$$\varphi_k^{(d)} : ([0, a)^d)^k \longrightarrow \{0, 1\} \\ (v_1, \dots, v_k) \longmapsto \begin{cases} \prod_{1 \leq i < j \leq k} \mathbf{1}_{[0, 2\epsilon)}(\rho_d(v_i, v_j)) & \text{if } (v_1, \dots, v_k) \in \Delta_k^{(d)}, \\ 0 & \text{otherwise.} \end{cases}$$

In words, this means that  $\varphi_k^{(d)}(v_1, \dots, v_k) = 1$  if  $[v_1, \dots, v_k]$  is a  $(k-1)$ -simplex and 0 otherwise.

**Theorem 9.** *The mean number of  $(k-1)$ -simplices  $N_k$  is given by*

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

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

$$N_k(\omega) = \frac{1}{k!} \int \varphi_k^{(d)}(v_1, \dots, v_k) d\omega^{(k)}(v_1, \dots, v_k).$$

According to the Campbell-Mecke formula and since the max-distance can be tensorized, we have:

$$\begin{aligned} \mathbf{E}_\Lambda [N_k] &= \frac{\lambda^k}{k!} \int_{X^k} \varphi_k^{(d)}(v_1, \dots, v_k) dv_1 \dots dv_k \\ &= \frac{\lambda^k}{k!} \left( \int_{[0,a]^k} \varphi_k^{(1)}(x_i, x_j) dx_1 \dots dx_k \right)^d. \end{aligned}$$

A moment of thought reveals that for any  $(x_1, \dots, x_k) \in \Delta_k^{(1)}$ , since  $\epsilon < a/4 < a/2$ , there exists a unique index  $i$  such that for all  $j \in \{1, \dots, k\} \setminus \{i\}$ , one and only one of the two following conditions holds:

$$x_i < x_j < x_i + 2\epsilon \text{ or } x_i < x_j + a < x_i + 2\epsilon.$$

Let  $\zeta(x_1, \dots, x_k)$  denote this index  $i$ . Hence, by invariance by translation of the Lebesgue measure,

$$\begin{aligned} &\int_{\zeta^{-1}(1)} \varphi_k^{(1)}(x_i, x_j) dx_1 \dots dx_k \\ &= (k-1)! \int_0^a dx_1 \int_{[x_1, x_1+2\epsilon]^{k-1}} \prod_{j=2}^{k-1} \mathbf{1}_{x_j < x_{j+1}} dx_2 \dots dx_k = a(2\epsilon)^{k-1}. \end{aligned}$$

The very same identity holds for any integral on the set  $\zeta^{-1}(i)$  for any  $i \in \{1, \dots, k\}$  hence

$$\int_{[0,a]^k} \varphi_k^{(1)}(x_i, x_j) dx_1 \dots dx_k = ka(2\epsilon)^{k-1}.$$

The proof is thus complete.  $\square$

By deppoissonization, we can estimate the mean number of  $k$ -simplices for a Binomial process: a process with  $n$  points uniformly distributed over the torus.

**Corollary 10.** *The mean number  $(k-1)$ -simplices  $N_k$  given  $N_1 = n$  is*

$$\mathbf{E}_\Lambda [N_k | N_1 = n] = \binom{n}{k} k^d \left( \frac{2\epsilon}{a} \right)^{d(k-1)}.$$

*Proof.* According to Theorem 9, we have:

$$\frac{\lambda a^d (\lambda (2\epsilon)^d)^{k-1} k^d}{k!} = \sum_{n=0}^{\infty} \mathbf{E}_{\Lambda} [N_k | N_1 = n] e^{-\lambda a^d} \frac{(\lambda a^d)^n}{n!}.$$

The principle of dePoissonization is then to invert the transform  $\Theta$  defined by:

$$\begin{aligned} \Theta : \mathbb{R}^{\mathbb{N}} &\longrightarrow \mathbb{R}[\lambda] \\ (\alpha_n, n \geq 0) &\longmapsto \sum_{n \geq 0} \alpha_n e^{-\lambda} \frac{\lambda^n}{n!}. \end{aligned}$$

We have that  $(\lambda a^d)^k = \sum_{n \geq k} \frac{n!}{(n-k)!} \frac{(\lambda a^d)^n}{n!} e^{-\lambda a^d}$ . The result follows.  $\square$

*Remark.* Considering the maximum norm simplifies the calculations. However, even for 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  $\mathbb{T}_a^2$ . We are limited to small orders because no formula seems to be known for the area of the intersection of  $k$  balls in general position. For  $k = 2$  and  $3$ , the expectations are given by the following formulas:

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

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

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

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

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

These polynomials appear rather surprisingly in the computations of the mean value of the Euler characteristic.

**Theorem 11.** *The mean of the Euler characteristic 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

$$N_k \leq \frac{1}{k!} \prod_{j=0}^{k-1} (N_1 - j) \leq \frac{N_1^k}{k!}, \text{ then } \sum_{k=1}^{\infty} N_k \leq \sum_{k=1}^{\infty} \frac{N_1^k}{k!} \leq e^{N_1}.$$

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

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

The proof is thus complete.  $\square$

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

$$\begin{aligned} \mathbf{E}_{\Lambda} [\chi] &= a\lambda e^{-\lambda 2\epsilon}, \text{ for } d = 1; \\ \mathbf{E}_{\Lambda} [\chi] &= a^2 \lambda e^{-\lambda(2\epsilon)^2} (1 - \lambda(2\epsilon)^2), \text{ for } d = 2; \\ \mathbf{E}_{\Lambda} [\chi] &= a^3 \lambda e^{-\lambda(2\epsilon)^3} (1 - 3\lambda(2\epsilon)^3 + (\lambda(2\epsilon)^3)^2), \text{ for } d = 3. \end{aligned}$$

The next corollary is an immediate consequence of Corollary 10, obtained again by depossionization.

**Corollary 12.** *The expectation of  $\chi$  for a binomial point process with  $n$  points is given by:*

$$\mathbf{E}_{\Lambda} [\chi | N_1 = n] = \sum_{k=0}^n (-1)^k \binom{n}{k} k^d \left(\frac{2\epsilon}{a}\right)^{d(k-1)}.$$

So far, we have not say a word about Betti numbers. It turns out that the preceding computations lead to a bound of the tail of  $\beta_0$ , the number of connected components.

**Theorem 13.** *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.*  $\beta_0$  is the number of connected components. Since there are more points than connected components,  $\mathbf{E}_{\Lambda} [\beta_0] \leq \mathbf{E}_{\Lambda} [N_1] = \lambda a^d$ . According to the definition of  $D$ ,  $\sup_{x \in X} D_x \beta_0$  is the maximum variation of  $\beta_0$  induced by the addition of an arbitrary point. If the point  $x$  is at a distance smaller than  $\epsilon$  from  $\omega$ , then  $D_x \beta_0 \leq 0$ , otherwise,  $D_x \beta_0 = 1$ , so  $D_x \beta_0 \leq 1$  for any  $x \in X$ . 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

$$\|D\beta_0\|_{L^\infty(L^2(X,\Lambda),\mathbf{P})} \leq \sup_{\omega} \int_X |D_x \beta_0|^2 d\Lambda(x) \leq \lambda a^d (2^d - 1)^2.$$

Since the function  $f$  defined by

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

is strictly increasing with respect to  $x$  and  $y$  for  $k_1 > x$ , it follows from Theorem 8 that:

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

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

## 5. SECOND ORDER MOMENTS

We now deal with the computations of the second order moments. The proofs rely on the chaos decomposition of the number of simplices (see Lemma 14) and the multiplication formula for iterated integrals (see Proposition 6). The computations are rather technical and postponed to Appendix A. We make the following convention: For any integer  $k$ ,

$$\int_{X^0} \varphi_k^{(d)}(v_1, \dots, v_k) dv_1 \dots dv_k = \varphi_k^{(d)}(v_1, \dots, v_k).$$

**Lemma 14.** *We can rewrite  $N_k$  as*

$$N_k = \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} \lambda^{k-i} I_i \left( \int_{X^{k-i}} \varphi_k^{(d)}(v_1, \dots, v_k) dv_1 \dots dv_{k-i} \right).$$

*Proof.* For  $k = 1$ , the result is immediate with the convention made above. Once we have seen that

$$\begin{aligned} & \int \varphi_k^{(d)}(v_1, \dots, v_k) d\omega^{(k)}(v_1, \dots, v_k) \\ &= \int \left( \int \varphi_k^{(d)}(v_1, \dots, v_k) (d(\omega - \sum_{j=1}^{k-1} \delta(v_j))(v_k) - d\Lambda(v_k)) \right) d\omega^{(k-1)}(v_1, \dots, v_{k-1}) \\ & \quad + \int \left( \int_X \varphi_k^{(d)}(v_1, \dots, v_k) d\Lambda(v_k) \right) d\omega^{(k-1)}(v_1, \dots, v_{k-1}), \end{aligned}$$

the result follows by induction.  $\square$

**Theorem 15.** *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*

$$\text{Cov}_\Lambda [N_k, N_l] = \sum_{i=1}^l \frac{\lambda a^d (\lambda (2\epsilon)^d)^{k+l-i-1}}{i!(k-i)!(l-i)!} \left( k+l-i+2 \frac{(k-i)(l-i)}{i+1} \right)^d.$$

*Remark.* As for the first moment it is still possible to find, considering the Euclidean norm, a closed-form expression for  $\text{Var}_\Lambda [N_k]$ . We did not find a general expression for any dimension. However, 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:

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

and

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

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

**Theorem 16.** *The variance of the Euler characteristic is:*

$$\text{Var}_\Lambda [\chi] = \lambda a^d \sum_{n=1}^{\infty} c_n^d (\lambda (2\epsilon)^d)^{n-1},$$

where

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

**Theorem 17.** *In one dimension, the expression of the variance of the Euler characteristic is:*

$$\text{Var}_\Lambda [\chi] = a (\lambda e^{-2\lambda\epsilon} - 4\lambda^2 \epsilon e^{-4\lambda\epsilon}).$$

**Theorem 18.** *If  $d = 2$ , we have  $D\chi \leq 2$  and thus*

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

*Proof.* In two dimensions, according to Proposition 3, the Euler characteristic is given by:

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

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 increasing  $\beta_0$  by 1, and the number of holes, i.e.  $\beta_1$ , remains the same. Otherwise, as there is no new connected component,  $\beta_0$  is the same, but the new vertex can at most fill one hole, increasing  $\beta_1$  by 1. Therefore, the variation of  $\beta_0 - \beta_1$  is at most 1.

Furthermore, when we add a vertex to a simplicial complex, we know from Proposition 4 that  $D\beta_2 \leq 1$  hence  $D\chi \leq 2$ . Then, we use Eq. (4) to complete the proof.  $\square$

## 6. THIRD ORDER MOMENTS

Higher order moments can be computed in a similar way but the computations become trickier as the order increases. We here restrict our computations to the third order moments to illustrate the general procedure. The proof is given in Appendix B.

We are interested in the central moment, so we introduce the following notation for the centralized number of  $(k-1)$ -simplices:  $\widetilde{N}_k = N_k - \mathbf{E}_\Lambda [N_k]$ .

**Theorem 19.** *The third central moment of the number of  $(k-1)$ -simplices is given by:*

$$\mathbf{E}_\Lambda [\widetilde{N}_k^3] = \sum_{i,j,s,t} \frac{\lambda^{3k-i-j-t}}{(k!)^3} \binom{k}{i} \binom{k}{j} \binom{k}{s} \binom{i}{t} \binom{j}{t} \binom{t}{i+j-s-t} \mathcal{J}_3(k, i, j, s, t),$$

with  $s \geq |i-j|$ , and  $\mathcal{J}_3(k, i, j, s, t)$  is an integral depending on  $k, i, j, s$  and  $t$ , defined below in (9).

## 7. CONVERGENCE

Before going further, we must answer a natural question: Do we retrieve the torus homology when the intensity of the Poisson process goes to infinity, so that the number of points becomes arbitrary large? The answer is positive as shows the next theorem.

**Theorem 20.** *The Betti numbers of  $\mathcal{C}_\epsilon(\omega)$  converge in probability to the Betti numbers of the torus as  $\lambda$  goes to infinity:*

$$\mathbf{P}_\Lambda \left( \bigcap_{i=0}^d \left( \beta_i(\omega) = \binom{d}{i} \right) \right) \xrightarrow{\lambda \rightarrow \infty} 1,$$

where  $\binom{d}{i}$  is the  $i$ -th Betti number of the  $d$ -dimensional torus, see [13].

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

$$\bigcap_{B \in \mathfrak{B}} (\omega(B) \neq 0) \subset (\mathcal{U}_\epsilon(\omega) = \mathbb{T}_a^d).$$

Thus,

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

Moreover, by the nerve lemma

$$(\mathcal{U}_\epsilon(\omega) = \mathbb{T}_a^d) \subset \bigcap_{i=0}^d \left( \beta_i(\omega) = \binom{d}{i} \right),$$

and the result follows.  $\square$

Let  $\Gamma$  be an arbitrary connected simplicial complex of  $n$  vertices. The number of occurrences of  $\Gamma$  in  $C_\epsilon(\omega)$  is denoted as  $G_\Gamma(\omega)$ . It must be noted that with our construction of the simplicial complex, a complex  $\Gamma$  appears in  $C_\epsilon(\omega)$  as soon as its edges are in  $C_\epsilon(\omega)$ . The set of edges of  $\Gamma$ , denoted by  $J_\Gamma$  is a subset of  $\{1, \dots, n\} \times \{1, \dots, n\}$ . Let us define the following function on the vertices of  $\Gamma$ :

$$\widetilde{h}^\Gamma(v_1, \dots, v_n) = \frac{1}{c_\Gamma} \prod_{(i,j) \in J_\Gamma} \mathbf{1}_{\rho_d(v_i, v_j) < \epsilon},$$

where  $c_\Gamma$  is the number of permutations  $\sigma$  of  $\{v_1, \dots, v_n\}$  such that

$$\widetilde{h}^\Gamma(v_1, \dots, v_n) = \widetilde{h}^\Gamma(v_{\sigma(1)}, \dots, v_{\sigma(n)}),$$

and let  $f^\Gamma$  be the symmetrization of  $\widetilde{h}^\Gamma$ . Then, we have:

$$(5) \quad G_\Gamma(\omega) = \int f^\Gamma(v_1, \dots, v_n) d\omega^{(n)}(v_1, \dots, v_n).$$

**Lemma 21.** *The random variable  $G_\Gamma$  has a chaos representation given by:*

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

where  $f_i^\Gamma$  is the bounded symmetric function defined as

$$(6) \quad f_i^\Gamma(v_{i+1}, \dots, v_n) = \binom{n}{i} \lambda^{n-i} \int_{X^{n-i}} f^\Gamma(v_1, \dots, v_n) dv_1 \dots dv_{n-i}.$$

*Proof.* From (5), using the binomial expansion and some algebra, we obtain

$$\begin{aligned} G_\Gamma(\omega) &= \sum_{i=0}^n \int \left( \binom{n}{i} \int_{X^{n-i}} f^\Gamma(v_1, \dots, v_n) \lambda \, dv_1 \dots \lambda \, dv_{n-i} \right) d\omega_\Lambda^{(i)}(v_{n-i+1}, \dots, v_n). \end{aligned}$$

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

$$f_i^\Gamma(v_{i+1}, \dots, v_n) = \binom{n}{i} \lambda^{n-i} \int_{X^{n-i}} f^\Gamma(v_1, \dots, v_n) \, dv_1 \dots \, dv_{n-i}.$$

To conclude the proof, we note that, since  $X$  is a bounded set and  $h^\Gamma$  is bounded,  $f_i^\Gamma$  is bounded.  $\square$

**Theorem 22.** *There exists  $c > 0$  such that for any  $\lambda \geq 1$ ,*

$$d_W \left( \frac{G_\Gamma - \mathbf{E}_\Lambda[G_\Gamma]}{\sqrt{\text{Var}_\Lambda[G_\Gamma]}}, \mathcal{N}(0, 1) \right) \leq \frac{c}{\lambda^{1/2}}.$$

*Proof.* Let  $F = \frac{G_\Gamma - \mathbf{E}_\Lambda[G_\Gamma]}{\sqrt{\text{Var}_\Lambda[G_\Gamma]}}$ . Provided that  $\Gamma$  has  $n$  vertices, according to Lemma 21, we have the following identities:

$$\begin{aligned} D_t F &= \frac{1}{\sqrt{\text{Var}_\Lambda[G_\Gamma]}} \sum_{i=1}^n i I_{i-1}(f_i^\Gamma(*, t)), \\ -D_t L^{-1} F &= \frac{1}{\sqrt{\text{Var}_\Lambda[G_\Gamma]}} \sum_{i=1}^n I_{i-1}(f_i^\Gamma(*, t)), \\ \text{Var}_\Lambda[G_\Gamma] &= \sum_{i=1}^n i! \|f_i^\Gamma\|_{L^2(X, \Lambda)^{\otimes i}}^2. \end{aligned}$$

Hence,  $\text{Var}_\Lambda[G_\Gamma]$  is a polynomial of degree  $2n - 1$  with respect to  $\lambda$ . From Proposition 6, it is tedious but straightforward to see that  $\langle DL^{-1}F, DF \rangle_{L^2(X, \Lambda)}$  is a polynomial of degree  $2n - 2$  with random coefficients depending on the integrals over  $X^i$  of the  $f_i^\Gamma$ . According to Lemma 21, these coefficients are bounded almost-surely. Hence there exists a constant  $c > 0$  such that

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

The same kind of computations shows that

$$\int_X \mathbf{E}_\Lambda \left[ |D_x F|^2 |D_x L^{-1} F| \right] \lambda \, dx \leq c \lambda^{-1/2}.$$

Then, the result follows from Theorem 7.  $\square$

## REFERENCES

- [1] M. A. Armstrong, *Basic topology*, Springer-Verlag, 1997.
- [2] E. T. Bell, *Exponential polynomials*, Ann. Math. (1934), no. 35, 258–277.
- [3] Rabi N. Bhattacharya and Jayanta K. Ghosh, *A class of U-statistics and asymptotic normality of the number of k-clusters*, J. Multivariate Anal. **43** (1992), no. 2, 300–330.
- [4] A. Björner, *Topological methods*, (1995), 1819–1872.
- [5] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of persistence diagrams*, Discrete Comput. Geom. **37** (2007), no. 1, 103–120.
- [6] D. Daley and D. Vere Jones, *An introduction to the theory of point processes*, Springer, 2002.
- [7] V. de Silva and R. Ghrist, *Coordinate-free coverage in sensor networks with controlled boundaries via homology*, International Journal of Robotics Research **25** (2006).
- [8] L. Decreusefond and E. Ferraz, *On the one dimensional Poisson random geometric graph*, Journal of Probability and Statistics (2011).
- [9] L. Decreusefond, A. Joulin, and N. Savy, *Upper bounds on Rubinstein distances on configuration spaces and applications*, Communications on stochastic analysis **4** (2010), no. 3, 377–399.
- [10] H. Edelsbrunner, D. Letscher, and A. Zomorodian, *Topological persistence and simplification*, Discrete Comput. Geom. **28** (2002), no. 4, 511–533, Discrete and computational geometry and graph drawing (Columbia, SC, 2001).
- [11] R. Ghrist, *Coverage and hole-detection in sensor networks via homology*, Fourth International Conference on Information Processing in Sensor Networks (IPSN'05), UCLA, 2005, pp. 254–260.
- [12] M. J. Greenberg and J. R. Harper, *Algebraic topology: a first course*, Addison-Wesley, New York, NY, 1981.
- [13] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002.
- [14] Y. Ito, *Generalized poisson functionals*, Probab. Theory Related Fields (1988), no. 77, 1–28.
- [15] M. Kahle, *Random geometric complexes*, Discrete Comput. Geom. **45** (2011), no. 3, 553–573.
- [16] M. Kahle and E. Meckes, *Limit theorems for Betti numbers of random simplicial complexes*, (2010), 29.
- [17] 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.
- [18] F.L. Lewis, *Wireless sensor networks*, ch. 2, John Wiley, New York, 2004.
- [19] J. Munkres, *Elements of algebraic topology*, Addison Wesley, 1993.
- [20] I. Nourdin and G. Peccati, *Normal approximations with malliavin calculus: From stein's method to universality*, Cambridge University Press, 2012.
- [21] 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.
- [22] M. Penrose, *Random geometric graphs*, Oxford Studies in Probability, vol. 5, Oxford University Press, Oxford, 2003.
- [23] M. D. Penrose and J. E. Yukich, *Limit theory for point processes in manifolds*, (2011).
- [24] G.J. Pottie and W.J. Kaiser, *Wireless integrated network sensors*, Communications of the ACM **43** (2000), 51–58.

- [25] N. Privault, *Stochastic analysis in discrete and continuous settings with normal martingales*, Lecture Notes in Mathematics, vol. 1982, Springer-Verlag, Berlin, 2009.
- [26] M. Reitzner and M. Schülte, *Central limit theorems for U-statistics of poisson point processes*, Annals of Probability, to appear (2013).
- [27] J.J. Rotman, *An introduction to algebraic topology (graduate texts in mathematics)*, Springer, July 1998.

## APPENDIX A. PROOFS OF THE SECOND ORDER MOMENTS

A.1. **Proof of Theorem 15.** By Lemma 14, we can rewrite the covariance between  $N_k$  and  $N_l$  with  $l \leq k$ :

$$\begin{aligned} \text{Cov}_\Lambda [N_k, N_l] &= \mathbf{E}_\Lambda [(N_k - \mathbf{E}_\Lambda [N_k])(N_l - \mathbf{E}_\Lambda [N_l])] \\ &= \mathbf{E}_\Lambda \left[ \frac{1}{k!} \sum_{i=1}^k \binom{k}{i} \lambda^{k-i} I_i(f_i^k) \frac{1}{l!} \sum_{j=1}^l \binom{l}{j} \lambda^{l-j} I_j(f_j^l) \right], \end{aligned}$$

where

$$f_i^k(v_{k-i+1}, \dots, v_k) = \int_{X^{k-i}} \varphi_k^{(d)}(v_1, \dots, v_k) dv_1 \dots dv_{k-i}.$$

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

$$\begin{aligned} \text{Cov}_\Lambda [N_k, N_l] &= \frac{1}{k!l!} \sum_{i=1}^l \binom{k}{i} \binom{l}{i} \lambda^{k+l-2i} \mathbf{E}_\Lambda [I_i(f_i^k) I_i(f_i^l)] \\ &= \frac{1}{k!l!} \sum_{i=1}^l \binom{k}{i} \binom{l}{i} \lambda^{k+l-2i} i! \langle f_i^k, f_i^l \rangle_{L^2(X, \Lambda)^{\circ i}} \\ &= \sum_{i=1}^l \frac{1}{i!(k-i)!(l-i)!} \lambda^{k+l-2i} \langle f_i^k, f_i^l \rangle_{L^2(X, \Lambda)^{\circ i}} \end{aligned}$$

Hence, we are reduced to compute

$$\begin{aligned} \langle f_i^k, f_i^l \rangle_{L^2(X, \Lambda)^{\circ i}} &= \int_{X^i} \left( \int_{X^{l-i}} \varphi_l^{(d)}(v_1, \dots, v_l) dv_{i+1} \dots dv_l \right) \\ &\quad \times \left( \int_{X^{k-i}} \varphi_k^{(d)}(v_1, \dots, v_k) dv_{i+1} \dots dv_k \right) \lambda dv_1 \dots \lambda dv_i. \end{aligned}$$

Let us denote  $\mathcal{J}_2(m_1, m_2, m_{12})$  the integral on two simplices of respectively  $m_1 + m_{12}$  and  $m_2 + m_{12}$  vertices with  $m_{12} > 0$  common vertices:

$$\begin{aligned} \mathcal{J}_2(m_1, m_2, m_{12}) &= \\ &\int_{X^M} \varphi_{m_1+m_{12}}^{(d)}(v_1, \dots, v_{m_1+m_{12}}) \varphi_{m_2+m_{12}}^{(d)}(v_{m_1+1}, \dots, v_M) dv_1 \dots dv_M, \end{aligned}$$

with  $M = m_1 + m_2 + m_{12}$ . Then we can rewrite:

$$\langle f_i^k, f_i^l \rangle_{L^2(X, \Lambda)^{\circ i}} = \lambda^i \mathcal{J}_2(l-i, k-i, i),$$

and it then remains to compute  $\mathcal{J}_2(m_1, m_2, m_{12})$ .

First, thanks to the tensorization property of the max-distance, we can write:

$$\mathcal{J}_2(m_1, m_2, m_{12}) = \left( \int_{[0, a]^M} \varphi_{m_1+m_{12}}^{(1)}(x_1, \dots, x_{m_1+m_{12}}) \varphi_{m_2+m_{12}}^{(1)}(x_{m_1+1}, \dots, x_M) \, dx_1 \dots dx_M \right)^d$$

Let us split the integration domain of  $\mathcal{J}_2$  into two parts:

- $A_1 = \{(x_1, \dots, x_M) \in \Delta_M^{(1)}, \varphi_M^{(1)}(x_1, \dots, x_M) = 1\}$ , we recognize the integral calculated in the proof of Theorem 9:

$$\begin{aligned} \int_{A_1} \varphi_{m_1+m_{12}}^{(1)}(x_1, \dots, x_{m_1+m_{12}}) \varphi_{m_2+m_{12}}^{(1)}(x_{m_1+1}, \dots, x_M) \, dx_1 \dots dx_M \\ = M(2\epsilon)^{M-1}a. \end{aligned}$$

- $A_2 = \{(x_1, \dots, x_M) \in \Delta_M^{(1)}, \varphi_M^{(1)}(x_1, \dots, x_M) \neq 1\}$ .

As in the proof of Theorem 9, we denote  $\zeta(x_1, \dots, x_M)$  the index  $i$  such that  $x_i < x_j < x_i + 2\epsilon$  or  $x_i < x_j + a < x_i + 2\epsilon$ , which exists since  $\epsilon < a/4$  and  $m_{12} > 0$ . By symmetry, we can reduce the analysis to the situation where  $\zeta(x_1, \dots, x_M) = 1$  and  $x_1$  pertains to the first simplex of  $m_1 + m_{12}$ . We then order the three sets of vertices such that:

$$x_1 < \dots < x_{m_1}, \quad x_{m_1+1} < \dots < x_{m_1+m_{12}}, \quad \text{and} \quad x_{m_1+m_{12}+1} < \dots < x_M.$$

Since  $(x_1, \dots, x_M)$  belongs to  $A_2$ , we have  $x_M - x_1 > 2\epsilon$ .

Let us denote  $J_a(f)(x) = \int_x^a f(u) \, du$  and by induction

$$J_a^{(m)}(f)(x) = \int_x^a J_a^{(m-1)}(f)(u) \, du.$$

Then we have by invariance by translation of the Lebesgue measure,

$$\begin{aligned} \int_{A_2} \varphi_{m_1+m_{12}}^{(1)}(x_1, \dots, x_{m_1+m_{12}}) \varphi_{m_2+m_{12}}^{(1)}(x_{m_1+1}, \dots, x_M) \, dx_1 \dots dx_M \\ = 2m_1!m_2!m_{12}! \int_0^a J_{x_1+2\epsilon}^{(m_1-1)}(\mathbf{1})(x_1) \int_{x_1+2\epsilon}^{x_1+4\epsilon} (-J_{x_M-2\epsilon}^{(m_2-1)}(\mathbf{1})(x_M)) J_{x_1+2\epsilon}^{(m_{12})}(\mathbf{1})(x_M-2\epsilon) \, dx_M \, dx_1. \end{aligned}$$

We easily find that:

$$\begin{aligned} J_{x_1+2\epsilon}^{(m_1-1)}(\mathbf{1})(x_1) &= \frac{(2\epsilon)^{m_1-1}}{(m_1-1)!} \\ -J_{x_M-2\epsilon}^{(m_2-1)}(\mathbf{1})(x_M) &= \frac{(2\epsilon)^{m_2-1}}{(m_2-1)!}, \\ J_{x_1+2\epsilon}^{(m_{12})}(\mathbf{1})(x_M-2\epsilon) &= \frac{(x_1 - x_M + 4\epsilon)^{m_{12}}}{m_{12}!}. \end{aligned}$$



Thus we have:

$$\begin{aligned} \int_{A_2} \varphi_{m_1+m_{12}}^{(1)}(x_1, \dots, x_{m_1+m_{12}}) \varphi_{m_2+m_{12}}^{(1)}(x_{m_1+1}, \dots, x_M) dx_1 \dots dx_M \\ = \frac{2m_1m_2}{m_{12}+1} (2\epsilon)^{M-1} a. \end{aligned}$$

Then,

$$\mathcal{J}_2(m_1, m_2, m_{12}) = (m_1 + m_2 + m_{12} + \frac{2m_1m_2}{m_{12}+1})^d a^d (2\epsilon)^{(m_1+m_2+m_{12}-1)d}$$

concluding the proof.

**A.2. Proof of Theorem 16.** The variance of  $\chi$  is given by:

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

We remark that  $N_i \leq N_1^i/i!$ , thus

$$\begin{aligned} &\mathbf{E}_\Lambda \left[ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |(N_i - \mathbf{E}_\Lambda [N_i]) (N_j - \mathbf{E}_\Lambda [N_j])| \right] \\ &\leq \mathbf{E}_\Lambda \left[ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} N_i N_j + \mathbf{E}_\Lambda [N_i] \mathbf{E}_\Lambda [N_j] + N_i \mathbf{E}_\Lambda [N_j] + N_j \mathbf{E}_\Lambda [N_i] \right] \\ &\leq \mathbf{E}_\Lambda \left[ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{N_1^{i+j} + \mathbf{E}_\Lambda [N_1^i] \mathbf{E}_\Lambda [N_1^j] + N_1^i \mathbf{E}_\Lambda [N_1^j] + N_1^j \mathbf{E}_\Lambda [N_1^i]}{i!j!} \right] \\ &\leq \mathbf{E}_\Lambda \left[ e^{2N_1} + e^{2\mathbf{E}_\Lambda [N_1]} + 2e^{N_1 + \mathbf{E}_\Lambda [N_1]} \right] \\ &< \infty. \end{aligned}$$

Thus, we can write

$$\text{Var}_\Lambda [\chi] = \sum_{i=1}^{\infty} (-1)^i \sum_{j=1}^{\infty} (-1)^j \text{Cov}_\Lambda [N_i, N_j].$$

The result follows by Theorem 15.

A.3. **Proof of Theorem 17.** If  $d = 1$ , according to Theorem 16:

$$(7) \quad \text{Var}_\Lambda [\chi] = \frac{a}{2\epsilon} \sum_{n=1}^{\infty} c_n^1 (2\lambda\epsilon)^n.$$

Moreover, 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  $\beta_n = c_n^1 - \alpha_n$ . 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:

$$\begin{aligned} \alpha_n &= (-1)^n \frac{n}{n!} \sum_{j=\lceil \frac{n+1}{2} \rceil}^n \left[ \binom{n}{j} 2 \sum_{i=0}^{2j-n} (-1)^i \binom{j}{i} + 2(-1)^n \binom{n}{j} \right] \\ &= \frac{1}{(n-1)!} \sum_{j=\lceil \frac{n+1}{2} \rceil}^n \left[ 2 \binom{n}{j} \binom{j-1}{n-j-1} - \binom{n}{j} \binom{j}{n-j} - 2(-1)^n \binom{n}{j} \right] \\ (8) \quad &= \frac{1}{(n-1)!} \sum_{j=\lceil \frac{n+1}{2} \rceil}^n \left[ \binom{n}{j} \left( \binom{j-1}{n-j} - \binom{j-1}{n-j-1} \right) - 2(-1)^n \binom{n}{j} \right]. \end{aligned}$$

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

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

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

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

By known formulas on hypergeometric functions, we have that:

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

Then, we substitute these last two expressions in Eq. (8) to obtain

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

and thus

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

Proceeding along the same line,  $\beta_n$  is given by

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

and again we can simplify the power series  $\sum_{n=0}^{\infty} \beta_n x^n$  as

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

Then, substituting  $\alpha_n$  and  $\beta_n$  in Eq. (7) yields the result.

## APPENDIX B. PROOF OF THE THIRD ORDER MOMENT

**B.1. Proof of Theorem 19.** From Lemma 14, we know that the chaos decomposition of the number of  $(k-1)$ -simplices is given by

$$\widetilde{N}_k = \sum_{i=1}^k I_i(h_i),$$

with

$$h_i(v_1, \dots, v_i) = \frac{1}{k!} \binom{k}{i} \lambda^{k-i} \int_{X^{k-i}} \varphi_k^{(d)}(v_1, \dots, v_k) dv_{i+1} \dots dv_k,$$

and

$$I_i(h_i) = \int_{X^i} h_i(v_1, \dots, v_i) d\omega_{\Lambda}^{(i)}(v_1, \dots, v_i).$$

Then, we define denoting  $u = i + j - s$ ,

$$g_{i,j,s,t} = t! \binom{i}{t} \binom{j}{t} \binom{t}{u-t} h_i \circ_t^{u-t} h_j$$

and using the chaos expansion (cf Proposition 6), we have

$$\begin{aligned} \widetilde{N}_k^3 &= \left( \sum_{i=1}^k I_i(h_i) \right)^3 \\ &= \left( \sum_{i=1}^k \sum_{j=1}^k I_i(h_i) I_j(h_j) \right) \left( \sum_{l=1}^k I_l(h_l) \right) \\ &= \sum_{i,j,l=1}^k \sum_{s=|i-j|}^{i+j} \sum_{t=\lceil \frac{s}{2} \rceil}^{u \wedge i \wedge j} I_s(g_{i,j,s,t}) I_l(h_l). \end{aligned}$$

According to (1), we obtain:

$$\begin{aligned}
\mathbf{E}_\Lambda \left[ \widetilde{N}_k^3 \right] &= \mathbf{E}_\Lambda \left[ \sum_{i,j=1}^k \sum_{s=|i-j|\vee 1}^{i+j\wedge k} \sum_{t=\lceil \frac{u}{2} \rceil}^{u\wedge i\wedge j} I_s(g_{i,j,s,t}) I_s(h_s) \right] \\
&= \sum_{i,j=1}^k \sum_{s=|i-j|\vee 1}^{i+j\wedge k} \sum_{t=\lceil \frac{u}{2} \rceil}^{u\wedge i\wedge j} \int_{X^s} g_{i,j,s,t} h_s \lambda^s \, dv_1 \dots \, dv_s \\
&= \sum_{i,j=1}^k \sum_{s=|i-j|\vee 1}^{i+j\wedge k} \sum_{t=\lceil \frac{u}{2} \rceil}^{u\wedge i\wedge j} \lambda^s t! \binom{i}{t} \binom{j}{t} \binom{t}{u-t} \int_{X^s} (h_i \circ_t^{u-t} h_j) h_s \, dv_1 \dots \, dv_s.
\end{aligned}$$

We denote  $\mathcal{J}_3(k, i, j, s, t)$  the following integral:

$$\begin{aligned}
(9) \quad \mathcal{J}_3(k, i, j, s, t) &= \int_{X^{3k-t-s}} \varphi_k^{(d)}(v_1, \dots, v_k) \varphi_k^{(d)}(v_1, \dots, v_t, v_{k+1}, \dots, v_{2k-t}) \\
&\quad \varphi_k^{(d)}(v_1, \dots, v_{2t-i-j+s}, v_{t+1}, \dots, v_i, v_{2k-j+1}, \dots, v_{3k-t-s}) \, dv_1 \dots \, dv_{3k-t-s},
\end{aligned}$$

for  $i, j, s$  and  $t$  bounded as in the previous sums. We recognize the integral on three  $(k-1)$ -simplices with  $u-t$ ,  $i-t$ , and  $j-t$  common vertices to only two of them, and  $2t-u$  common vertices to the three of them. Then, we can write:

$$\begin{aligned}
\mathbf{E}_\Lambda \left[ \widetilde{N}_k^3 \right] &= \sum_{i,j=1}^k \sum_{s=|i-j|\vee 1}^{i+j\wedge k} \sum_{t=\lceil \frac{u}{2} \rceil}^{u\wedge i\wedge j} \frac{\lambda^{3k-i-j} t!}{(k!)^3} \binom{k}{i} \binom{k}{j} \binom{k}{s} \binom{i}{t} \binom{j}{t} \binom{t}{u-t} \\
&\quad \mathcal{J}_3(k, i, j, s, t).
\end{aligned}$$

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

*E-mail address:* [laurent.decreasefond@telecom-paristech.fr](mailto:laurent.decreasefond@telecom-paristech.fr)

*E-mail address:* [eduardo.ferraz@telecom-paristech.fr](mailto:eduardo.ferraz@telecom-paristech.fr)

*E-mail address:* [hugues.randriam@telecom-paristech.fr](mailto:hugues.randriam@telecom-paristech.fr)

*E-mail address:* [anais.vergne@telecom-paristech.fr](mailto:anais.vergne@telecom-paristech.fr)

INSTITUT TELECOM, TELECOM PARISTECH, CNRS LTCl, 46, RUE BARRAULT, PARIS - 75634, FRANCE