



ANNALES
HENRI LEBESGUE

ANNA GUSAKOVA

MATTHIAS REITZNER

CHRISTOPH THÄLE

VARIANCE EXPANSION AND BERRY-ESSEEN BOUND FOR THE NUMBER OF VERTICES OF A RANDOM POLYGON IN A POLYGON

ASYMPTOTIQUE DE LA VARIANCE ET
BORNES DE BERRY-ESSEEN POUR LE
NOMBRE DE SOMMETS D'UN POLYGONE
ALÉATOIRE DANS UN POLYGONE

Keywords: Berry–Esseen bound, central limit theorem, geometric probability, Poisson point process, random convex chain, random polygon, variance expansion.

2020 *Mathematics Subject Classification:* 52A22, 60D05.

DOI: <https://doi.org/10.5802/ahl.180>

(*) This work has been initiated during the virtual Hausdorff Trimester Program *The Interplay between High Dimensional Geometry and Probability*.

CT was supported by the German Research Foundation (DFG) via SPP 2265 *Random Geometric Systems*. AG was supported by the DFG under Germany's Excellence Strategy EXC 2044 – 390685587, *Mathematics Münster: Dynamics - Geometry - Structure*.

ABSTRACT. — Fix a container polygon P in the plane and consider the convex hull P_n of $n \geq 3$ independent and uniformly distributed in P random points. In the focus of this paper is the vertex number of the random polygon P_n . The precise variance expansion for the vertex number is determined up to the constant-order term, a result which can be considered as a second-order analogue of the classical expansion for the expectation of Rényi and Sulanke (1963). Moreover, a sharp Berry–Esseen bound is derived for the vertex number of the random polygon P_n , which is of the same order as one over the square-root of the variance. The latter is optimal and improves the earlier result of Bárány and Reitzner (2006) by removing the factor $(\log \log n)^{60}$ in the planar case. The main idea behind the proof of both results is a decomposition of the boundary of the random polygon P_n into random convex chains and a careful merging of the variance expansions and Berry–Esseen bounds for the vertex numbers of the individual chains. In the course of the proof, we derive similar results for the Poissonized model.

RÉSUMÉ. — Fixons un polygone conteneur P dans le plan et considérons l’enveloppe convexe P_n de $n \geq 3$ points aléatoires indépendants et uniformément distribués dans P . Cet article est consacré à l’étude du nombre de sommets du polygone aléatoire P_n . Nous déterminons l’asymptotique précise de sa variance jusqu’au terme constant, un résultat qui peut être considéré comme un analogue au second ordre du développement classique pour l’espérance de Rényi et Sulanke (1963). De plus, nous obtenons une borne de Berry–Esseen précise pour le nombre de sommets du polygone aléatoire P_n , du même ordre que l’inverse de la racine carrée de la variance. Cette borne est optimale et améliore le résultat précédent de Bárány et Reitzner (2006) en supprimant le facteur $(\log \log n)^{60}$ dans le cas planaire. L’idée principale dans la preuve des deux résultats est une décomposition du bord du polygone aléatoire P_n en chaînes convexes aléatoires et une fusion soigneuse des asymptotiques de la variance et des bornes de Berry–Esseen pour le nombre de sommets des chaînes individuelles. Au cours de la preuve, nous obtenons des résultats similaires pour le modèle poissonisé.

1. Introduction and result

Let $P \subset \mathbb{R}^2$ be a polygon in the plane with $\ell \geq 3$ vertices. We refer to P as a container in what follows. Let X_1, \dots, X_n be $n \geq 3$ independent random points, distributed uniformly in P . We denote by P_n the random convex hull $[X_1, \dots, X_n]$ of these points. It is a random polygon in the container polygon P , see Figure 1.1. In this article we are interested in the combinatorial structure of P_n , more precisely in the variance expansion and the fluctuations of the number $f_0(P_n)$ of vertices of P_n . Note that this quantity is the same as the number $f_1(P_n)$ of edges of P_n .

The random variable $f_0(P_n)$ has been intensively studied in the literature and has attracted a lot of interest in geometric probability as well as convex and integral geometry. For example, as $n \rightarrow \infty$, Rényi and Sulanke in their fundamental article [RS63] have established the asymptotics

$$(1.1) \quad \mathbb{E}f_0(P_n) = \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log \left(\frac{F_i}{\text{area}(P)} \right) + \frac{2\gamma\ell}{3} + o(1),$$

for the expected number of vertices of P_n , where $\gamma \approx 0.57721\dots$ is the Euler–Mascheroni constant, F_i , $i = 1, \dots, \ell$, are the areas of the triangles formed by three consecutive vertices of P (that is, $F_i = \text{area}([v_{i-1}, v_i, v_{i+1}])$ if v_1, \dots, v_ℓ are the vertices of P with the convention that $v_0 = v_\ell$ and $v_{\ell+1} = v_1$). The corresponding variance asymptotics

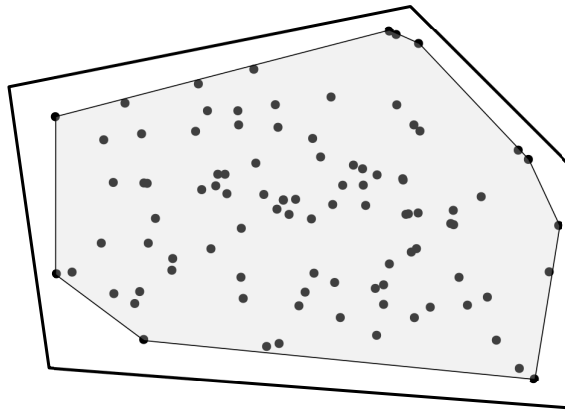


Figure 1.1. Random polygon in a polygon.

$$(1.2) \quad \text{Var } f_0(P_n) = \frac{10\ell}{27} \log n(1 + o(1)),$$

is due to Groeneboom [Gro88]. In the same paper, Groeneboom also proved a central limit theorem for $f_0(P_n)$, as $n \rightarrow \infty$. Denoting by $\Phi(\cdot)$ the distribution function of a standard Gaussian random variable, Bárány and Reitzner [BR06] have obtained the following quantitative version of the central limit theorem:

$$(1.3) \quad \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{f_0(P_n) - \mathbb{E} f_0(P_n)}{\sqrt{\text{Var } f_0(P_n)}} \leq x \right) - \Phi(x) \right| \leq c \frac{(\log \log n)^{60}}{\sqrt{\log n}},$$

where $c > 0$ is some constant not depending on n . Note that in the published version [BR10b] of [BR06] the Poissonized model for P_n has been treated, which involves the additional randomization that the number of generating random points is Poisson distributed with mean n . The quantitative central limit theorem in [BR10b] for this Poisson random polytope P_η comes with an error term $c(\log \log n)^{60}/\sqrt{\log n}$ and thus contains an additional double logarithmic factor like (1.3). The Poisson random polytope P_η will also play a prominent role in our analysis. In Theorem 3.13 we will remove the double logarithmic factor in the error term of the Poissonized version, thus improving for dimension two the speed of convergence in the main result of [BR10b]. In the unpublished preprint [BR06], a de-Poissonization argument leads to the central limit theorem (1.3). A similar de-Poissonization argument will also be used in Section 3.6 to obtain Theorem 1.1.

The double logarithmic factor also appears in a central limit theorem of Pardon [Par11, Par12], which holds for general planar convex container. This is one of the few result in the literature, in which the considered class of container sets is not restricted to polygons or convex sets bounded by a sufficiently smooth closed curve. The purpose of the present article is to demonstrate how the double-logarithmic factor in (1.3) can be removed for polygons. Since the resulting rate of convergence is then of the same order as $(\text{Var } f_0(P_n))^{-1/2}$ and since $f_0(P_n)$ is an integer-valued random variable, an argument of Englund [Eng81, Section 6] shows that, up to numerical constants, our result is in fact optimal. Moreover, our technique allows us

at the same time to determine the precisely expansion of (1.2) up to the constant-order term, leading to a second-order analogue of the classical result (1.1) for the expectation of Rényi and Sulanke. Even further, our technique allows us to make more precise the $o(1)$ -term in (1.1).

THEOREM 1.1. — *For every planar polygon $P \subset \mathbb{R}^2$ and $n \geq 2$ we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{f_0(P_n) - \mathbb{E}f_0(P_n)}{\sqrt{\text{Var } f_0(P_n)}} \leq x \right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log n}},$$

where $c > 0$ is some constant independent of n , with

$$\mathbb{E}f_0(P_n) = \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log \left(\frac{F_i}{\text{area}(P)} \right) + \frac{2\gamma\ell}{3} + O \left(n^{-\frac{1}{4}} (\log n)^2 \right)$$

and

$$\text{Var } f_0(P_n) = \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log \left(\frac{F_i}{\text{area}(P)} \right) + \frac{(10\gamma - 2\pi^2)\ell}{27} + O \left(n^{-\frac{1}{4}} (\log n)^4 \right),$$

as $n \rightarrow \infty$.

Remark 1.2. — On the way of the proof of Theorem 1.1 we obtain in Theorem 3.13 the same statements also for the Poisson random polytope model previously considered in [BR10b], that is, under the additional randomization that the number of generating points follows a Poisson distribution with mean n .

Remark 1.3. — Although this paper focusses on the vertex number of the random polygon P_n , we briefly discuss a direct consequence of Theorem 1.1 on the area of P_n . First, the classical Efron identity [Efr65] connects the expected number of vertices of P_n with its area:

$$\frac{\mathbb{E}\text{area}(P_n)}{\text{area}(P)} = 1 - \frac{\mathbb{E}f_0(P_{n+1})}{n+1}.$$

Applying the expansion for $\mathbb{E}f_0(P_{n+1})$ in Theorem 1.1 we conclude that, as $n \rightarrow \infty$,

$$\begin{aligned} & \frac{\mathbb{E}\text{area}(P \setminus P_n)}{\text{area}(P)} \\ &= \frac{2\ell}{3} (\log n) n^{-1} + \left[\frac{2}{3} \sum_{i=1}^{\ell} \log \left(\frac{F_i}{\text{area}(P)} \right) + \frac{2\gamma\ell}{3} \right] n^{-1} + O \left(n^{-\frac{5}{4}} (\log n)^2 \right). \end{aligned}$$

Similarly, one can apply Buchta's identity [Buc05, Corollary 1], which connects the variance of the area of P_n with the first two moments of its vertex number:

$$\frac{\text{Var } \text{area}(P_n)}{\text{area}(P)^2} = \frac{\text{Var } f_0(P_{n+2}) + A_n - B_n}{(n+1)(n+2)}$$

with $A_n := (\mathbb{E}f_0(P_{n+2}))^2 - \frac{n+2}{n+1} (\mathbb{E}f_0(P_{n+1}))^2$ and $B_n := (2n+3)\mathbb{E}f_0(P_{n+2}) - 2(n+2)\mathbb{E}f_0(P_{n+1})$. In connection with the variance expansion in Theorem 1.1 this leads to

$$\begin{aligned} & \frac{\text{Var area}(P_n)}{\text{area}(P)^2} \\ &= \frac{28\ell}{27}(\log n)n^{-2} + \left[\frac{28}{27} \sum_{i=1}^{\ell} \log \left(\frac{F_i}{\text{area}(P)} \right) + \frac{(28\gamma - 2\pi^2)\ell}{27} \right] n^{-2} + O\left(n^{-\frac{9}{4}}(\log n)^4\right), \end{aligned}$$

as $n \rightarrow \infty$.

We would like to point out that if the container set P of the random polygon P_n has a C^2 -smooth boundary with everywhere positive curvature the first Berry–Esseen bound for the vertex number in [Rei05] also contained an additional logarithmic factor, and the presumably optimal Berry–Esseen bound has been found in [LRSY19, Theorem 3.5]. Moreover, the approach in [LRSY19] even allows to deal with higher dimensional random polytopes and with other geometric and combinatorial parameters. However, we would like to stress at this point that the transition from smooth container sets to polygons appears to be highly non-trivial. One reason for this fact is the observation that, in contrast to smooth containers, the geometry of a random polygon in a polygon is not locally determined in a sufficiently strong sense. For example, for an arbitrary boundary point x of the container set one can ask for the expected number of vertices of the random polygon P_n one can “see” from x . While in the smooth case this number stays bounded for large n , in case of a polygon it grows to infinity at a double logarithmic speed. These and similar geometric facts are the reason why already the proof of a sub-optimal Berry–Esseen bound like (1.3) is considerably more involved compared to its counterpart for smooth container sets. Another aspect to be mentioned in this context is the strong concentration of the number of vertices of a random polygon in a small neighbourhood around the corners of the container polygon. This concentration phenomenon, which apparently does not take place in smoothly bounded container sets, makes it much harder to approximate the random polygon by the so-called floating body associated with the container. More precisely, even the careful refinement from [BR10b] of this approach automatically leads to the double-logarithmic factor in (1.3).

We shall now briefly explain how we overcome these difficulties. In particular, this description makes it evident that our approach cannot be extended to deal with the vertex number, or more generally the number of faces of arbitrary dimension, of convex hulls of random points in polytopes of dimension more than two. The main ingredient we use in the proof of Theorem 1.1 is a decomposition of the boundary of the random polygon P_n into so-called random convex chains, where each chain corresponds to one of the corners of the container polygon, see Figure 3.1. In fact, the vertex number of P_n is the same as the sum of the number of vertices of these chains, which after Poissonization of the construction become independent random variables. By a suitable affine transformation, each chain can be transformed into the following standard form without changing its combinatorial structure (of course, the number m below depends in a suitable way on the size of the corresponding chain in P_n). Consider the triangle T with vertices $(0,0)$, $(0,1)$ and $(1,0)$ and let T_m be the convex hull of $(0,1)$ and $(1,0)$ together with $m \geq 1$ independent and uniformly distributed random points in T . The following Berry–Esseen bound for the

vertex number $f_0(T_m)$ of the random convex chain T_m has been obtained in [GT21, Corollary 9]:

$$(1.4) \quad \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{f_0(T_m) - \mathbb{E}f_0(T_m)}{\sqrt{\text{Var } f_0(T_m)}} \leq x \right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log m}},$$

for some absolute constant $c > 0$ and $m \geq 1$. We would like to mention at this occasion that this result is a consequence of an unexpected connection (which does not persist for polygons different from a triangle) between the random variables $f_0(T_m)$, the location of zeros of certain orthogonal polynomials related to probability generating function of $f_0(T_m)$ and a Berry–Esseen bound for sums of independent Bernoulli random variables. We remark that (1.4) is the main motivation behind Theorem 1.1 and the highly non-trivial transition from (1.4) to the Berry–Esseen bound for $f_0(P_n)$ in Theorem 1.1 is our main contribution. In particular, it involves a careful merge of the Berry–Esseen bounds of the individual convex chains based on Poissonization and conditioning arguments. A similar strategy will be applied in order to determine the asymptotic behaviour of the expectation and variance of $f_0(P_n)$ in Theorem, 1.1.

2. Preliminaries

In this paper, $[A]$ stands for the convex hull of a set $A \subseteq \mathbb{R}^2$, $\#A$ for its cardinality and $\text{aff}(A)$ for the affine hull of A . In addition, if A is a line segment, we denote its length by $|A|$. We write $B(x, r)$ for the closed disc centred at $x \in \mathbb{R}^2$ with radius $r > 0$. Given a set $A \subset \mathbb{R}^2$ and a point $x \in \mathbb{R}^2$ denote by $d(A, x) := \inf_{y \in A} \|y - x\|$ the distance between A and x . For two functions f and g we write $f = O(g)$ if $\limsup_{x \rightarrow \infty} |f(x)/g(x)| < \infty$ and $f = o(g)$ if $\lim_{x \rightarrow \infty} |f(x)/g(x)| = 0$. Given a line $L := \{(x, y) \in \mathbb{R}^2 : ux + vy = t\}$ we denote by

$$L^+ := \{(x, y) \in \mathbb{R}^2 : ux + vy \geq t\} \quad \text{and} \quad L^- := \{(x, y) \in \mathbb{R}^2 : ux + vy \leq t\}$$

the positive and negative half-planes into which L divides \mathbb{R}^2 , respectively. Given a convex body $K \subset \mathbb{R}^2$ we consider a function $v : K \mapsto \mathbb{R}$, defined as

$$v(z) := \min\{\text{area}(H \cap K) : H \text{ is a half-plane with } z \in H\}.$$

Then the floating body $K(v \geq \delta)$ with parameter $\delta > 0$ is a level set of the function v , namely $K(v \geq \delta) := \{z \in K : v(z) \geq \delta\}$. The wet part is the set $K(v < \delta) := K \setminus K(v \geq \delta)$. In case of the polygon P with $\text{area}(P) = 1$ the following formula for the area of the wet part (and an analogous formula for volume in arbitrary dimensions) was independently obtained in [Sch91] and [BB93]:

$$(2.1) \quad \text{area}(P(v < \delta)) = \frac{\ell}{4} \delta \log \frac{1}{\delta} (1 + o(1)),$$

where $o(1)$ is meant for $\delta \rightarrow 0$.

LEMMA 2.1. — Choose n independent uniform random points X_1, \dots, X_n in a container polygon P . Then there is a constant $b_0 > 0$ such that the event $\mathcal{A}_n := \{P(v \geq b_0 n^{-1} \log n) \subset P_n\}$ satisfies

$$(2.2) \quad \mathbb{P}(\overline{\mathcal{A}_n}) = O(n^{-6}).$$

We also have to deal with the Poissonized random polytope in the following. To define this model properly, let N be a Poisson random variable with mean $n \in \mathbb{R}_+$, and choose N independent uniform random points X_1, \dots, X_N in a polygon P , which are also independent of N . Then X_1, \dots, X_N is a homogeneous Poisson point process η in the polygon P with $\mathbb{E}(\#\eta) = n$. We denote its convex hull by $P_\eta = [\eta]$. The next lemma is a combination of [BR10b, Lemma 5.3] and the Poissonized version of [BD97, Theorem 1].

LEMMA 2.2. — There are constants $b_0, c_0 > 0$ independent of $n > 1$, such that the events

$$\mathcal{A}_n^\pi := \{P(v \geq b_0 n^{-1} \log n) \subset P_\eta\}$$

and

$$\mathcal{B}_n^\pi := \{\#\left(\eta \cap P(v < b_0 n^{-1} \log n)\right) \leq c_0 (\log n)^2\}$$

satisfy

$$\mathbb{P}(\overline{\mathcal{A}_n^\pi}) \leq \mathbb{P}(\overline{\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi}) = O(n^{-6}).$$

During the proofs of this paper we switch several times between the Poisson model P_η and the binomial model P_n of the random polygons we consider. For this the following two estimates will turn out to be helpful. The first is a slight extension of a result of Matsunawa [Mat82], the second one estimates the total variation distance between two Poisson random variables.

LEMMA 2.3. — Let $n \in \mathbb{N}$ and $p > 0$. Put $\binom{n}{m} = 0$ for $n < m$. Then

$$\sum_{m=0}^{\infty} m^k \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right| \leq \begin{cases} 2p & k=0, \\ 4np^2 & k=1, \\ 6np^2(1+np) & k=2. \end{cases}$$

Proof. — The inequality for $k=0$ is due to Matsunawa [Mat82] verifying a conjecture of Vervaat [Ver69]. The case $k=1$ follows from

$$\begin{aligned} \sum_{m=0}^{\infty} m \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right| &= np \sum_{m=1}^{\infty} \left| \frac{(np)^{m-1}}{(m-1)!} e^{-np} - \binom{n-1}{m-1} p^{m-1} (1-p)^{n-m} \right| \\ &= np \sum_{m=0}^{\infty} \left| \frac{(np)^m}{m!} e^{-np} - \binom{n-1}{m} p^m (1-p)^{n-m-1} \right| \\ &\leq 4np^2. \end{aligned}$$

The case $k = 2$ is a combination of case $k = 1$ with

$$\begin{aligned}
 & \sum_{m=0}^{\infty} m(m-1) \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right| \\
 & \leq (np)^2 \sum_{m=0}^{\infty} \left| \frac{(np)^m}{m!} e^{-np} - \frac{n-1}{n} \binom{n-2}{m} p^m (1-p)^{n-m-2} \right| \\
 & \leq (np)^2 \sum_{m=0}^{\infty} \left| \frac{(np)^m}{m!} e^{-np} - \binom{n-2}{m} p^m (1-p)^{n-m-2} \right| \\
 & \quad + np^2 \sum_{m=0}^{\infty} \left| \binom{n-2}{m} p^m (1-p)^{n-m-2} \right| \\
 & \leq 6p(np)^2 + np^2.
 \end{aligned}$$

This completes the argument. \square

LEMMA 2.4. — Let $t > 0$. Then for $\varepsilon \geq 0$ we have

$$\sum_{m=0}^{\infty} \left| \frac{(t+\varepsilon)^m}{m!} e^{-(t+\varepsilon)} - \frac{t^m}{m!} e^{-t} \right| \leq 2\varepsilon.$$

Proof. — The inequality $|ab - cd| \leq |a - c|b + c|b - d|$, valid for $a, b, c, d \geq 0$, shows that

$$\begin{aligned}
 \sum_{m=0}^{\infty} \left| \frac{(t+\varepsilon)^m}{m!} e^{-(t+\varepsilon)} - \frac{t^m}{m!} e^{-t} \right| & \leq \sum_{m=0}^{\infty} \frac{(t+\varepsilon)^m - t^m}{m!} e^{-(t+\varepsilon)} + \sum_{m=0}^{\infty} \frac{t^m}{m!} (e^{-t} - e^{-(t+\varepsilon)}) \\
 & = (e^{t+\varepsilon} - e^t) e^{-(t+\varepsilon)} + e^t (e^{-t} - e^{-(t+\varepsilon)}) \\
 & = 2(1 - e^{-\varepsilon}) \leq 2\varepsilon,
 \end{aligned}$$

and the proof is complete. \square

3. Proof of Theorem 1.1

The strategy of the proof of Theorem 1.1 consists of the following main steps:

Step 1: As a first step we randomize the model further by taking instead of a fixed number n of points, distributed uniformly in P , a random number N of points. More precisely, as the number of points we use a Poisson random variable with mean n so that the collection of random points becomes a homogeneous Poisson point process. This construction is known as Poissonization and it appears to be very helpful for our purposes. In fact, after Poissonization the number of points in disjoint regions become independent random variables.

Step 2: In order to prove the variance asymptotic and the Berry–Esseen bound for the Poisson model described in Step 1 we need to introduce an additional construction, which allows us to exclude some “bad” and “unlikely” events on which we do not have sufficient control on the geometric configuration.

Step 3: The third step is devoted to the proof of the variance asymptotic and the Berry–Esseen bound for the Poisson random convex chain. The corresponding results for the classical convex chain, proven in [GT21], are transformed using so-called “transfer lemma”, which has been used in similar situations before in the literature, see [BR06, BR10b, BV07, Rei05, Vu06].

Step 4: The fourth step is concerned with the proof of the variance asymptotic and the Berry–Esseen bound for the Poisson random polygon in P under the conditions introduced in Step 2. As already explained in the introduction, the idea of the proof is to decompose the boundary of the random polygon into independent random convex chains, each of which corresponds to one corner of container polygon. The result can now be obtained from the corresponding result for the Poissonized convex chain from Step 3.

Step 5: In this step we remove the constraints introduced in Step 2 and show that this does not change the quality of the variance asymptotic and the Berry–Esseen bound for the Poisson model.

Step 6: In the last step we deduce the variance asymptotic and the Berry–Esseen bound for the original model from the one for the Poisson model via de-Poissonization. The transition from the Poisson model to the binomial one is made using again the “transfer lemma”.

In order to simplify some of our arguments we can and will from now on assume that the container polygon P has area one. This is indeed possible as the vertex number of P_n does not change under rescaling of the container polygon P .

3.1. Step 1: Introducing the Poisson model

Let η be a homogeneous Poisson point process in the container polygon P with $\mathbb{E}(\#\eta) = n$. Formally, η can be constructed as follows. Let N be a Poisson random variable with mean n and, independently of N , let X_1, X_2, \dots be a sequence of independent and uniformly distributed random points in P . Then η can be defined as the random set of points $\{X_1, \dots, X_N\}$, which is interpreted as the empty set if $N = 0$, an event having probability e^{-n} . We will refer to η as a homogeneous Poisson point process with intensity n on P . The well-known multivariate Mecke formula is a useful tool to compute expectations of Poisson functionals, see [LP18, Theorem 4.1]. It says that

$$(3.1) \quad \mathbb{E} \sum_{(X_1, \dots, X_k) \in \eta_{\neq}^k} f(X_1, \dots, X_k, \eta) = n^k \mathbb{E} \int_{P^k} f(x_1, \dots, x_k, \eta \cup \{x_1, \dots, x_k\}) dx_1 \dots dx_k,$$

where η_{\neq}^k denotes the set of all k -tuples of distinct points of η , and f is a non-negative measurable function acting on a k -tuple of points and a locally-finite point configuration in \mathbb{R}^2 to which these points belong.

We consider now the random polygon P_η , which (we recall) was defined as a convex hull of the Poisson point process η , that is, $P_\eta = [\eta]$. There exists a clear connection

the length of the edge e_i , by α_i the angle of P at the vertex v_i and by F_i the area of the triangle $[v_{i-1}, v_i, v_{i+1}] = [e_i, e_{i+1}]$, $1 \leq i \leq \ell$. For every edge e_i , $1 \leq i \leq \ell$, if $\eta \neq \emptyset$ there is one point $Z_i = Z_i(\eta) := (x_i, y_i)$ from the Poisson point process η , such that either $L^+(Z_i) \cap \eta = \{Z_i\}$ or $L^-(Z_i) \cap \eta = \{Z_i\}$, where $L(Z_i)$ denotes the line parallel to e_i passing through Z_i , see Figure 3.1. If $\eta = \emptyset$ we set Z_i , $1 \leq i \leq \ell$ to be the vertex of P maximizing $d(y, \text{aff}(e_i))$, $y \in P$. In case if there are few such vertices we may take the lexicographically smallest one. Without loss of generality we assume that $L(Z_i)$ is parametrized in a way such that $L^-(Z_i) \cap \eta = \{Z_i\}$ and thus $\eta \subset L^+(Z_i)$. It might happen that for some or few $1 \leq i \leq \ell$ we have $Z_i = Z_{i+1}$ with $Z_{\ell+1} = Z_1$ or even $\eta = \emptyset$, but we will treat these situations separately and then, by a conditioning argument, exclude them from the forthcoming discussion.

Denote the point of intersection of $L(Z_i)$ and $L(Z_{i+1})$ by $V_i := L(Z_i) \cap L(Z_{i+1})$, see Figure 3.1, and put $\Delta_i := [Z_i, V_i, Z_{i+1}]$, $1 \leq i \leq \ell$, and $\Delta_i := Z_i$ if $Z_i = Z_{i+1}$. The length of the edges of the triangle Δ_i with vertex V_i are denoted by $\delta_{i,i} = \|Z_i - V_i\|$ and $\delta_{i,i+1} = \|V_i - Z_{i+1}\|$.

In the next sections we will need the first moments and covariances of the logarithmic area of Δ_i , and the probability that the triangle is not too small. We start with the following lemma.

LEMMA 3.1. — *There is a constant $c_P > 0$ depending only on P , such that for $1 \leq i \leq \ell$ we have*

$$\mathbb{P}(d(Z_i, \text{aff}(e_i)) \geq x) \leq e^{-c_P n x}.$$

Proof. — If $\eta = \emptyset$ then Z_i is the vertex of P maximizing the distance, and the inequality is trivial. Recall that ℓ_i denotes the length of e_i . If $d(Z_i, \text{aff}(e_i)) \geq x$ and m is some point in P with $d(m, \text{aff}(e_i)) = x$, then $L^-(Z_i) \cap P$ contains the triangle $[e_i, m]$ with base length ℓ_i and height x ,

$$\text{area}([e_i, m]) = \frac{1}{2} \ell_i x.$$

On the other hand, according to the definition of Z_i we have $\eta \subset L^+(Z_i)$ and, thus, the above triangle with base ℓ_i and height x does not contain points of η . Thus,

$$\mathbb{P}(d(Z_i, \text{aff}(e_i)) \geq x) \leq \mathbb{P}(\#(\eta \cap [e_i, m]) = 0) = e^{-n \frac{1}{2} \ell_i x}.$$

The result follows with $c_P := \min_i \frac{\ell_i}{2}$. \square

From now on we assume that $d(Z_i, \text{aff}(e_i)) \leq n^{-\frac{3}{4}}$, $Z_i \neq Z_{i+1}$, and $\delta_{i,j} \geq \tilde{c}_P n^{-\frac{1}{4}}$ with $\tilde{c}_P := \sqrt{2}(\min_{1 \leq j \leq \ell} \sin \alpha_j)^{-\frac{1}{2}}$ for $1 \leq i \leq \ell$ and $j = i, i+1$ (note that the latter condition already implies that $Z_i \neq Z_{i+1}$, which we have included only for convenience). We denote this event by \mathcal{E} . Observe that given \mathcal{E} we have

$$(3.2) \quad \text{area}(\Delta_i) \geq n^{-\frac{1}{2}}.$$

In a first step we show that \mathcal{E} happens with high probability.

COROLLARY 3.2. — *There are constants $\underline{c}_P, \bar{c}_P > 0$ only depending on P , such that*

$$\underline{c}_P n^{-1} \leq \mathbb{P}(\bar{\mathcal{E}}) \leq \bar{c}_P n^{-\frac{1}{4}}.$$

Proof. — First of all note that

$$(3.3) \quad \mathbb{P}(\bar{\mathcal{E}}) \leq \sum_{i=1}^{\ell} \mathbb{P}\left(d(Z_i, \text{aff}(e_i)) > n^{-\frac{3}{4}}\right) + \sum_{i=1}^{\ell} \mathbb{P}(Z_i \neq Z_{i+1}) \\ + \sum_{i=1}^{\ell} \sum_{j=i}^{i+1} \mathbb{P}\left(0 < \delta_{i,j} < \tilde{c}_P n^{-1/4}, d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}}, \forall k = 1, \dots, \ell\right).$$

The event that $d(Z_i, \text{aff}(e_i)) > n^{-\frac{3}{4}}$ occurs according to Lemma 3.1 with probability $\leq e^{-c_P n^{\frac{1}{4}}}$. The probability of the event that $V_i = Z_i = Z_{i+1}$ for some $i = 1, \dots, \ell$, again according to Lemma 3.1, can be estimated by

$$\sum_{i=1}^{\ell} \mathbb{P}(Z_i = Z_{i+1}) \\ \leq \sum_{i=1}^{\ell} \mathbb{P}\left(Z_i = Z_{i+1}, d(Z_i, \text{aff}(e_i)) \leq n^{-\frac{3}{4}}, d(Z_{i+1}, \text{aff}(e_{i+1})) \leq n^{-\frac{3}{4}}\right) + 2\ell e^{-c_P n^{\frac{1}{4}}} \\ \leq \sum_{i=1}^{\ell} \mathbb{P}\left(\#\left(\eta \cap B\left(v_i, (\sin(\alpha_i/2))^{-1} n^{-\frac{3}{4}}\right)\right) \geq 1\right) + 2\ell e^{-c_P n^{\frac{1}{4}}} \\ \leq \ell \left[1 - \exp\left(-2\pi \max_{1 \leq i \leq \ell} (\sin(\alpha_i/2))^{-2} n^{-\frac{1}{2}}\right)\right] + 2\ell e^{-c_P n^{\frac{1}{4}}} = O\left(n^{-\frac{1}{2}}\right).$$

Finally, we estimate the probabilities that $0 < \delta_{i,j} < \tilde{c}_P n^{-\frac{1}{4}}$ for $i = 1, \dots, \ell$ and $j = i, i+1$ using the multivariate Mecke formula (3.1). Denote by $L_i(x)$ the line through x parallel to e_i and by $L_i^+(x)$ the corresponding halfplane not containing the edge e_i . Taking for simplicity $i = 1, j = 2$ we have for $\delta_{1,2} = \delta_{1,2}(\eta)$ that

$$\mathbb{P}\left(0 < \delta_{1,2} < \tilde{c}_P n^{-\frac{1}{4}}, d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}}, \forall k = 1, \dots, \ell\right) \\ = \mathbb{E} \sum_{(X_1, X_2, X_3) \in \eta_{\neq}^3} \mathbb{1}\left(0 < \delta_{1,2}(\eta) < \tilde{c}_P n^{-\frac{1}{4}}\right) \mathbb{1}\left(\eta \subset L_1^+(X_1) \cap L_2^+(X_2) \cap L_3^+(X_3)\right) \\ \times \mathbb{1}\left(d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}} \quad \forall k = 1, \dots, \ell\right) \\ = n^3 \mathbb{E} \int_P \int_P \int_P \mathbb{1}\left(0 < \delta_{1,2}(\eta \cup \{x_1, x_2, x_3\}) < \tilde{c}_P n^{-\frac{1}{4}}\right) \\ \times \mathbb{1}\left(\eta \cup \{x_1, x_2, x_3\} \subset L_1^+(x_1) \cap L_2^+(x_2) \cap L_3^+(x_3)\right) \\ \times \mathbb{1}\left(d(Z_k(\eta \cup \{x_1, x_2, x_3\}), \text{aff}(e_k)) \leq n^{-\frac{3}{4}} \quad \forall k = 1, \dots, \ell\right) dx_1 dx_2 dx_3.$$

Let us remark that $(X_1, X_2, X_3) \in \eta_{\neq}^3$ in the second line automatically implies that $Z_1 = X_1$, $Z_2 = X_2$ and $Z_3 = X_3$ are almost surely distinct points. Also, the random variable $\delta_{1,2}(\eta)$ is determined by the random points X_1 and X_2 in the second and $\delta_{1,2}(\eta \cup \{x_1, x_2, x_3\})$ by the points x_1 and x_2 in the third line. Recall that V_1 and V_2 are the points on $L_2(x_2)$ whose positions are determined by x_1, x_2 and x_3 , see Figure 3.1 (in fact, the dependence of V_2 on the position of x_3 is the reason why we consider the three points x_1, x_2, x_3). Now, we fix V_1, V_2 and first integrate x_2 along

the segment $[V_1, V_2] \subset L_2(x_2) \cap P$. Without loss of generality assume that the edge e_2 is parallel to the one of the axes in \mathbb{R}^2 . In this case we have $x_2 = (x_{2,1}, x_{2,2})$ with (say) $x_{2,2} = d(Z_2, \text{aff}(e_2))$ and, thus, V_1, V_2 depend on x_1, x_3 and $x_{2,2}$ only. In these new coordinates we perform the integration with respect to $x_{2,1} \in [V_1, V_2]$. Because of

$$\int_{V_1}^{V_2} \mathbb{1}(\delta_{1,2} < \tilde{c}_P n^{-\frac{1}{4}}) dx_{2,1} \leq \tilde{c}_P n^{-\frac{1}{4}} = \tilde{c}_P n^{-\frac{1}{4}} \|V_1 - V_2\|^{-1} \int_{V_1}^{V_2} dx_{2,1}$$

and $\|V_1 - V_2\| = \ell_2(1 + O(n^{-\frac{3}{4}}))$, which holds under the assumption that $d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}}$ for all $k = 1, \dots, \ell$, we have

$$\begin{aligned} & \mathbb{P}(0 < \delta_{1,2} < \tilde{c}_P n^{-\frac{1}{4}}, d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}} \quad \forall k = 1, \dots, \ell) \\ & \leq \tilde{c}_P n^{-\frac{1}{4}} \ell_2^{-1} (1 + O(n^{-\frac{3}{4}})) n^3 \mathbb{E} \int_P \int_P \int_P \mathbb{1}(\eta \subset L_1^+(x_1) \cap L_2^+(x_2) \cap L_3^+(x_3)) \\ & \quad \times \mathbb{1}(d(Z_k(\eta \cup \{x_1, x_2, x_3\}), \text{aff}(e_k)) \leq n^{-\frac{3}{4}} \quad \forall k = 1, \dots, \ell) dx_1 dx_2 dx_3 \\ & = \tilde{c}_P n^{-\frac{1}{4}} \ell_2^{-1} (1 + O(n^{-\frac{3}{4}})) \mathbb{P}(0 < \delta_{1,2}, d(Z_k, \text{aff}(e_k)) \leq n^{-\frac{3}{4}} \quad \forall k = 1, \dots, \ell), \\ & \leq \tilde{c}_P n^{-\frac{1}{4}} \ell_2^{-1} (1 + O(n^{-\frac{3}{4}})) (1 - \ell e^{-c_P n^{\frac{1}{4}}}) = O(n^{-\frac{1}{4}}), \end{aligned}$$

where in the third line we have applied backwards the same argument, based on the multivariate Mecke formula, as above. Combining the three estimates with (3.3) yields the right hand side of the inequality.

On the other hand, there is some small $c_1 > 0$ depending on P such that the probability that $B(v_1, n^{-1}) \cap P$ contains precisely one point of η is

$$\begin{aligned} \mathbb{P}(\#(\eta \cap B(v_1, n^{-1})) = 1) &= n \text{area}(B(v_1, n^{-1}) \cap P) e^{-n \text{area}(B(v_1, n^{-1}) \cap P)} \\ &= c_1 n^{-1} e^{-c_1 n^{-1}} \end{aligned}$$

for n sufficiently large. The area of the parallel strips along the edges e_1 and e_2 of width n^{-1} without the disk $B(v_1, n^{-1})$ is upper bounded by $c_2 n^{-1}$ with some $c_2 > 0$ depending on P . Hence, the probability that this region contains no points of η is lower bounded by e^{-c_3} with $c_3 > 0$ depending on P . If these independent events occur then the single point in $B(v_1, n^{-1})$ is just $Z_1 = Z_2$, which proves the left hand side inequality, that is

$$\mathbb{P}(\bar{\mathcal{E}}) \geq \mathbb{P}(Z_1 = Z_2) \geq c_1 n^{-1} e^{-c_1 n^{-1} - c_3} \geq \underline{c}_P n^{-1}.$$

This completes the argument. \square

In the second step we investigate the moments and mixed moments of $\log \delta_{i,i}$ and $\log \delta_{i,i+1}$ under the condition \mathcal{E} .

LEMMA 3.3. — *Let $i, j, k, l \in \{1, \dots, \ell\}$ with $j = i, i + 1$ and $l = k, k + 1$. The first to logarithmic moments of $\delta_{i,j}$ satisfy*

$$(3.4) \quad \mathbb{E}(\log \delta_{i,j} | \mathcal{E}) = \log \ell_j - 1 + O(n^{-\frac{1}{4}} \log n)$$

and

$$(3.5) \quad \mathbb{E} \left((\log \delta_{i,j})^2 \mid \mathcal{E} \right) = (\log \ell_j - 1)^2 + 1 + O \left(n^{-\frac{1}{4}} (\log n)^2 \right).$$

For the mixed logarithmic moments we obtain

$$(3.6) \quad \mathbb{E} \left[(\log \delta_{i,i+1}) (\log \delta_{i+1,i+1}) \mid \mathcal{E} \right] = (\log \ell_{i+1} - 1)^2 + 1 - \frac{\pi^2}{6} + O \left(n^{-\frac{1}{4}} \log n \right)$$

and for $j \neq l$,

$$(3.7) \quad \mathbb{E}[(\log \delta_{i,j})(\log \delta_{k,l}) \mid \mathcal{E}] = (\log \ell_j - 1)(\log \ell_l - 1) + O \left(n^{-\frac{1}{4}} \log n \right).$$

Proof. — We assume the event \mathcal{E} and prove (3.4) e.g. for $\delta_{i,i+1}$ and $i = 1$. Our argument will be similar to that in the proof of Corollary 3.2 and we will also use the same notation introduced as there. In particular, recall that $L_i(x)$ is the line through x parallel to e_i and $L_i^+(x)$ the corresponding halfplane not containing the edge e_i . The multivariate Mecke formula (3.1) yields

$$\begin{aligned} \mathbb{E}(\log \delta_{1,2} \mathbb{1}(\mathcal{E})) &= \mathbb{E} \sum_{(X_1, X_2, X_3) \in \eta_{\neq}^3} \log \delta_{1,2}(\eta) \mathbb{1}(\eta \subset L_1^+(X_1) \cap L_2^+(X_2) \cap L_3^+(X_3), \mathcal{E}) \\ &= n^3 \mathbb{E} \int_P \int_P \int_P \log \delta_{1,2}(\eta \cup \{x_1, x_2, x_3\}) \\ &\quad \times \mathbb{1}(\eta \cup \{x_1, x_2, x_3\} \subset L_1^+(x_1) \cap L_2^+(x_2) \cap L_3^+(x_3)) \\ &\quad \times \mathbb{1}(\mathcal{E}(\eta \cup \{x_1, x_2, x_3\})) \, dx_2 dx_3 dx_1. \end{aligned}$$

As before, we assume without loss of generality that e_2 is parallel to of the axes of \mathbb{R}^2 and that $x_2 = (x_{2,1}, x_{2,2})$ with $x_{2,2} = d(Z_2, \text{aff}(e_2))$. Thus, $L_2(x_2)$, V_1 , V_2 depend on x_1, x_3 and $x_{2,2}$ only. We integrate $x_{2,1}$ on the line segment $[V_1, V_2]$ as in Figure 3.1. Moreover, note that given \mathcal{E} , we have $\|V_1 - V_2\| = \ell_2(1 + O(n^{-\frac{3}{4}}))$. Because $\int \log x \, dx = (\log x - 1)x$ we get

$$\begin{aligned} &\int_{V_1}^{V_2} \log \delta_{1,2}(\eta \cup \{x_1, x_2, x_3\}) \mathbb{1}(\mathcal{E}(\eta \cup \{x_1, x_2, x_3\})) \, dx_{2,1} \\ &= \int_{\tilde{c}_p n^{-\frac{1}{4}}}^{\|V_1 - V_2\| + \tilde{c}_p n^{-\frac{1}{4}}} \log x \, dx \\ &= (\log \ell_2 - 1 + O(n^{-\frac{1}{4}})) (\ell_2 + O(n^{-\frac{1}{4}})) - O(n^{-\frac{1}{4}} \log n) \\ &= (\log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n)) \int_{\tilde{c}_p n^{-\frac{1}{4}}}^{\|V_1 - V_2\| + \tilde{c}_p n^{-\frac{1}{4}}} dx \\ &= (\log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n)) \int_{V_1}^{V_2} \mathbb{1}(\mathcal{E}(\eta \cup \{x_1, x_2, x_3\})) \, dx_{2,1}. \end{aligned}$$

Thus, using again the multivariate Mecke formula (3.1),

$$\mathbb{E}(\log \delta_{1,2} \mathbb{1}(\mathcal{E})) = (\log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n))$$

$$\begin{aligned} & n^3 \mathbb{E} \int_P \int_P \int_P \mathbb{1}(\eta \cup \{x_1, x_2, x_3\} \subset L_1^+(x_1) \cap L_2^+(x_2) \cap L_3^+(x_3)) \\ & \quad \times \mathbb{1}(\mathcal{E}(\eta \cup \{x_1, x_2, x_3\})) \, dx_2 dx_3 dx_1 \\ & = (\log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n)) \mathbb{E} \mathbb{1}(\mathcal{E}), \end{aligned}$$

and

$$\mathbb{E}(\log \delta_{1,2} | \mathcal{E}) = \log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n).$$

In this way we obtain Equation (3.4). In the same way, this time using $\int (\log x)^2 dx = ((\log x)^2 - 2 \log x + 2)x$, one proves (3.5).

Analogously, using the identity $\int_0^a \log x \log(a-x) dx = a[(\log a - 1)^2 + 1 - \frac{\pi^2}{6}]$ and thus

$$\begin{aligned} \int_{\epsilon}^{a-\epsilon} \log x \log(a-x) dx &= \int_0^a \log x \log(a-x) dx - 2 \int_0^{\epsilon} (\log a + O(\epsilon)) \log x dx \\ &= a \left[(\log a - 1)^2 + 1 - \frac{\pi^2}{6} + O(\epsilon \log \epsilon) \right], \end{aligned}$$

we obtain for the expectation of the product of the two neighbouring log-distances

$$\mathbb{E}((\log \delta_{1,2})(\log \delta_{2,2}) | \mathcal{E}) = (\log \ell_2 - 1)^2 + 1 - \frac{\pi^2}{6} + O(n^{-\frac{1}{4}} \log n).$$

Considering the product of two distances not on the same line $L_i(\cdot)$, e.g.

$$\mathbb{E}(\log \delta_{1,1} \log \delta_{1,2} | \mathcal{E}) \quad \text{or} \quad \mathbb{E}(\log \delta_{1,2} \log \delta_{2,3} | \mathcal{E})$$

we rewrite this as a multiple integral using the multivariate Mecke formula once again, and integrate first with respect to $x_{2,1}$ to obtain

$$\begin{aligned} \mathbb{E}((\log \delta_{1,1})(\log \delta_{1,2}) | \mathcal{E}) &= (\log \ell_2 - 1 + O(n^{-\frac{1}{4}} \log n)) \mathbb{E}(\log \delta_{1,1} | \mathcal{E}) \\ &= (\log \ell_2 - 1)(\log \ell_1 - 1) + O(n^{-\frac{1}{4}} \log n), \end{aligned}$$

and similarly for all other cases. The proof is thus completed. \square

The conditional second moments can be written in a more concise way as the conditional variance and covariance of the involved quantities.

COROLLARY 3.4. — *For $i, j, k, l \in \{1, \dots, \ell\}$ with $j = i, i + 1$ it holds that*

$$(3.8) \quad \text{Var}(\log \delta_{i,j} | \mathcal{E}) = 1 + O(n^{-\frac{1}{4}} (\log n)^2)$$

and with $j = i, i + 1$, $l = k, k + 1$, and $(i, j) \neq (k, l)$,

$$(3.9) \quad \text{Cov}(\log \delta_{i,j}, \log \delta_{k,l} | \mathcal{E}) = \mathbb{1}(l = j) \left(1 - \frac{\pi^2}{6} \right) + O(n^{-\frac{1}{4}} \log n).$$

For $i \in \{1, \dots, \ell\}$, the logarithmic area of the triangle Δ_i equals

$$\log \text{area}(\Delta_i) = \log \frac{\sin \alpha_i}{2} + \log \delta_{i,i} + \log \delta_{i,i+1}.$$

Because the edges of Δ_i of length $\delta_{i,i}$ and $\delta_{i,i+1}$ are parallel to e_i and e_{i+1} , we see that

$$F_i = \frac{\sin \alpha_i}{2} \ell_i \ell_{i+1}, \quad 1 \leq i \leq \ell,$$

is the area of the triangle $[v_{i-1}, v_i, v_{i+1}] = [e_i, e_{i+1}]$. Lemma 3.3 yields immediately the conditional expectation, variance and covariances of $\log \text{area}(\Delta_i)$. For example,

$$\text{Cov} \left(\log \text{area}(\Delta_i), \log \text{area}(\Delta_k) \mid \mathcal{E} \right) = \sum_{j=i, i+1, \quad l=k, k+1} \text{Cov}(\log \delta_{i,j}, \log \delta_{k,l} \mid \mathcal{E}).$$

Combined with the formula for the conditional variance (3.8) and covariance (3.9) this yields the following result.

COROLLARY 3.5. — *For $i, k \in \{1, \dots, \ell\}$ one has that*

$$(3.10) \quad \mathbb{E}(\log \text{area}(\Delta_i) \mid \mathcal{E}) = \log F_i - 2 + O \left(n^{-\frac{1}{4}} \log n \right),$$

$$(3.11) \quad \text{Var}(\log \text{area}(\Delta_i) \mid \mathcal{E}) = 2 + O \left(n^{-\frac{1}{4}} (\log n)^2 \right),$$

and

$$(3.12) \quad \begin{aligned} \text{Cov}(\log \text{area}(\Delta_i), \log \text{area}(\Delta_k) \mid \mathcal{E}) \\ = \mathbb{1}(|i - k| = 1) \left(1 - \frac{\pi^2}{6} \right) + O \left(n^{-\frac{1}{4}} \log n \right), \end{aligned}$$

for $k \neq i$.

3.3. Step 3: Berry–Esseen bound for the Poisson random convex chain

Let T be the canonical triangle with vertices $(0, 1)$, $(0, 0)$ and $(1, 0)$, and χ be a homogeneous Poisson point process with $\mathbb{E}(\#\chi) = M > 1$ in T . The convex hull of the two vertices $(0, 1)$, $(1, 0)$ and the points of χ is denoted by

$$T_\chi := [\chi, (0, 1), (1, 0)].$$

Denote by $f_0(T_\chi)$ the number of vertices of T_χ . In order to obtain the Berry–Esseen bound for the Poissonized version of the random convex chain as in (1.4) we will use the following transfer lemma taken from [BR10b, Lemma 3.2]. The proof of this lemma can be found, for example, in [BV07] (see also the remark after [BR10b, Lemma 3.2]).

LEMMA 3.6. — *Given two sequences of random variables ξ_n and ξ'_n with means $\mu_n \in \mathbb{R}$ and $\mu'_n \in \mathbb{R}$, and variances $0 < \sigma_n^2 < \infty$ and $0 < \sigma_n'^2 < \infty$, respectively. Assume that there are sequences $\varepsilon_1(n)$, $\varepsilon_2(n)$, $\varepsilon_3(n)$ and $\varepsilon_4(n)$, all tending to zero as $n \rightarrow \infty$, such that*

- (i) $|\mu'_n - \mu_n| \leq \varepsilon_1(n) \sigma_n$,
- (ii) $|\sigma_n'^2 - \sigma_n^2| \leq \varepsilon_2(n) \sigma_n^2$,

- (iii) for every $x \in \mathbb{R}$, $|\mathbb{P}(\xi'_n \leq x) - \mathbb{P}(\xi_n \leq x)| \leq \varepsilon_3(n)$,
- (iv) for any $x \in \mathbb{R}$, $|\mathbb{P}(\frac{\xi_n - \mu'_n}{\sigma'_n} \leq x) - \Phi(x)| \leq \varepsilon_4(n)$.

Then there is a positive constant $C > 0$ such that

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\xi_n - \mu_n}{\sigma_n} \leq x \right) - \Phi(x) \right| \leq C \sum_{i=1}^4 \varepsilon_i(n).$$

In order to verify conditions (i) and (ii) of the previous lemma for our model we derive the following asymptotic formulas for the expectation and the variance of $f_0(T_\chi)$. For later purposes we also formulate our result for general homogeneous Poisson point processes.

LEMMA 3.7. — Consider a homogeneous Poisson point process χ in the canonical triangle T with $\mathbb{E}(\#\chi) = M > 1$. Then

$$(3.13) \quad \mathbb{E}f_0(T_\chi) = \frac{2}{3} \log M + \frac{2\gamma + 7}{3} + O(M^{-1/2}),$$

$$(3.14) \quad \text{Var } f_0(T_\chi) = \frac{10}{27} \log M + \frac{10\gamma + 2\pi^2 - 28}{27} + O(M^{-1/2}),$$

as $M \rightarrow \infty$.

Proof. — Denote by $H_n = \sum_{i=1}^n \frac{1}{i}$ the n^{th} harmonic sum and by $H_n^{(2)} = \sum_{i=1}^n \frac{1}{i^2}$ the n^{th} harmonic sum of second order. Set $H_0 = H_0^{(2)} := 0$ for convenience. It is well known that

$$(3.15) \quad H_n = \log n + \gamma + O(n^{-1}),$$

$$(3.16) \quad H_n^{(2)} = \frac{\pi^2}{6} + O(n^{-1}),$$

as $n \rightarrow \infty$. Recall that γ is the Euler–Mascheroni constant. Let T_k , $k \geq 1$ be the random convex chain, which is build based on the sample of k independent random points Y_1, \dots, Y_k , uniformly distributed inside the canonical triangle T . That is, $T_k := [Y_1, \dots, Y_k, (0, 1), (1, 0)]$. It is known from [Buc12, Corollary 1 and Corollary 2] that

$$(3.17) \quad \mathbb{E}f_0(T_k) = \frac{2}{3}H_k + \frac{7}{3},$$

$$(3.18) \quad \text{Var } f_0(T_k) = \frac{10}{27}H_k + \frac{4}{9}H_k^{(2)} - \frac{28}{27} + \frac{4}{9(k+1)};$$

note that the result in [Buc12] is stated for the quantity $N_k = f_0(T_k) - 2$. Let Y be a Poisson random variable with mean $M \in \mathbb{R}$, $M > 0$. Then

$$\mathbb{E}H_Y = \sum_{k=0}^{\infty} (H_k - H_{\lceil M \rceil}) \mathbb{P}(Y = k) + H_{\lceil M \rceil}.$$

Because of the trivial estimate

$$|H_Y - H_{\lceil M \rceil}| \leq \max \left\{ \frac{Y - \lceil M \rceil}{\lceil M \rceil}, \frac{\lceil M \rceil - Y}{Y + 1} \right\} \leq \left| \frac{Y - M}{M} \right| + \left| \frac{M - Y + 1}{Y + 1} \right|$$

we have

$$\begin{aligned} |\mathbb{E}H_Y - H_{\lceil M \rceil}| &\leq \mathbb{E} \frac{|Y - M|}{M} + \mathbb{E} \frac{|M - Y + 1|}{Y + 1} \\ &\leq \sqrt{\mathbb{E} \frac{(Y - M)^2}{M^2}} + \sqrt{\mathbb{E} \frac{(M - Y + 1)^2}{(Y + 1)^2}}. \end{aligned}$$

Since $\mathbb{E}Y = M$ and $\mathbb{E}Y^2 = M^2 + M$, we see that

$$\mathbb{E} \frac{(Y - M)^2}{M^2} = \frac{\text{Var}(Y)}{M^2} = M^{-1}.$$

Analogously, since $\mathbb{E}(Y + 1)^{-1} = M^{-1}(1 - e^{-M}) \geq M^{-1}(1 - M^{-1}) = (M - 1)M^{-2}$, and

$$\begin{aligned} \mathbb{E}(Y + 1)^{-2} &= \sum_{k=2}^{\infty} \frac{M^{k-2}}{k!} \frac{k}{k-1} e^{-M} \leq \sum_{k=2}^{\infty} \frac{M^{k-2}}{k!} \left(1 + \frac{3}{k+1}\right) e^{-M} \leq M^{-2} (1 + 3M^{-1}), \end{aligned}$$

we see that

$$\mathbb{E} \frac{(M - Y + 1)^2}{(Y + 1)^2} = \mathbb{E} \frac{(M + 2)^2}{(Y + 1)^2} - 2\mathbb{E} \frac{(M + 2)}{(Y + 1)} + 1 \leq 37M^{-1}.$$

Because of $\log \lceil M \rceil = \log M + O(M^{-1})$, equation (3.15) yields

$$(3.19) \quad \mathbb{E}H_Y = H_{\lceil M \rceil} + O(M^{-\frac{1}{2}}) = \log M + \gamma + O(M^{-\frac{1}{2}}).$$

Together with (3.17) this proves (3.13). Similarly, by the law of total variance we have

$$\begin{aligned} \text{Var } f_0(T_X) &= \mathbb{E}_Y \text{Var}(f_0(T_k) | Y = k) + \text{Var}_Y \mathbb{E}(f_0(T_k) | Y = k) \\ &= \mathbb{E} \left(\frac{10}{27} H_Y + \frac{4}{9} H_Y^{(2)} - \frac{28}{27} + \frac{4}{9(Y+1)} \right) + \text{Var} \left(\frac{2}{3} H_Y \right). \end{aligned}$$

By (3.19), and because $\mathbb{E}H_Y^{(2)} = \frac{\pi^2}{6} + O(\mathbb{E}(Y + 1)^{-1})$ by (3.16) with $\mathbb{E}(Y + 1)^{-1} = O(M^{-1})$, we obtain

$$\text{Var } f_0(T_X) = \left(\frac{10}{27} \log M + \frac{10\gamma + 2\pi^2 - 28}{27} + O(M^{-\frac{1}{2}}) \right) + \frac{4}{9} \text{Var } H_Y.$$

To prove (3.14) it remains to show that the variance of H_Y is bounded by a constant times $M^{-\frac{1}{2}}$. For this we use the Poincaré inequality for Poisson random variables, which says that

$$\text{Var } f(Y) \leq M \mathbb{E}(f(Y + 1) - f(Y))^2$$

for functions $f : \{0, 1, 2, \dots\} \rightarrow \mathbb{R}$ for which $\text{Var } f(Y) < \infty$. This inequality can be considered as a special case of the general Poincaré inequality [LP18, Theorem 18.7] for functionals of Poisson random measures. Applying this to $f(Y) = H_Y$ we conclude that

$$\text{Var } H_Y \leq M \mathbb{E} \frac{1}{(Y + 1)^2} \leq \sum_{k=0}^{\infty} 2 \frac{M^{k+1}}{(k + 2)!} e^{-M} \leq 2M^{-1},$$

which completes the argument. \square

Now we are prepared to prove a Berry–Esseen bound for the number of vertices of the Poisson random chain T_χ in the canonical triangle T .

LEMMA 3.8. — *Consider a homogeneous Poisson point process χ in the canonical triangle T with $\mathbb{E}(\#\chi) = M \geq 2$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{f_0(T_\chi) - \mathbb{E}f_0(T_\chi)}{\sqrt{\text{Var } f_0(T_\chi)}} \leq x \right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log M}},$$

for some absolute constant $c > 0$.

Proof. — Let T_k denotes the convex chain, build on random points X_1, \dots, X_k independently and uniformly distributed in T . We apply Lemma 3.6 with $\xi'_M := f_0(T_{\lceil M \rceil})$ and $\xi_M := f_0(T_\chi)$. The condition (i) with $\varepsilon_1(M) \leq c_1 M^{-\frac{1}{2}} / \log M$ for $c_1 > 0$ independent of M can be verified by formulas (3.13), (3.17) and (3.14). Analogously, condition (ii) with $\varepsilon_2(M) = c_2 M^{-\frac{1}{2}} / \log M$ for $c_2 > 0$ independent of M follows from (3.14) and (3.18). The condition (iv) with $\varepsilon_4(M) = c / \sqrt{\log \lceil M \rceil} = c_4 / \sqrt{\log M}$ for some constant $c_4 > 0$ independent of M follows immediately from (1.4).

For the verification of condition (iii) we use the convex floating body introduced in Section 2, and follow an approach already used somewhat implicitly in [Gro88] as well as [Rei05]. Recall that $\mathcal{A}_{\lceil M \rceil}$ is the event that the vertices of the convex hull of $\lceil M \rceil$ random points are contained in the wet part $T(v < b_0 \lceil M \rceil^{-1} \log \lceil M \rceil)$ of the triangle T , which clearly implies that all vertices of the convex chain $T_{\lceil M \rceil}$ are contained in this wet part. Slightly more general, for $M > 1$ define \mathcal{A}_M to be the event that the vertices of the convex hull of the $\lceil M \rceil$ random points are contained in the wet part $T(v < b_0 M^{-1} \log M)$ of the triangle T . Analogously, \mathcal{A}_M^π is the event that all vertices of the Poisson convex hull, and thus all vertices of the Poisson convex chain T_χ belong to $T(v < b_0 M^{-1} \log M)$. Because

$$T(v < \lceil M \rceil^{-1} \log \lceil M \rceil) \subset T(v < M \log M),$$

by Lemma 2.1 and Lemma 2.2 we have

$$\mathbb{P}(\overline{\mathcal{A}_M}) \leq \mathbb{P}(\overline{\mathcal{A}_{\lceil M \rceil}}) = O(M^{-6}) \quad \text{and} \quad \mathbb{P}(\overline{\mathcal{A}_M^\pi}) = O(M^{-6}).$$

Combining these estimates yields, for any $x \in \mathbb{R}$,

$$(3.20) \quad \left| \mathbb{P}(f_0(T_{\lceil M \rceil}) \leq x) - \mathbb{P}(f_0(T_\chi) \leq x) \right| \leq \left| \mathbb{P}(f_0(T_{\lceil M \rceil}) \leq x, \mathcal{A}_M) - \mathbb{P}(f_0(T_\chi) \leq x, \mathcal{A}_M^\pi) \right| + O(M^{-6}).$$

For T_χ we have that $\#(\chi \cap T(v < b_0 M^{-1} \log M))$ is Poisson distributed with mean Mp with

$$p := \frac{\text{area}(T(v < b_0 M^{-1} \log M))}{\text{area}(T)} = O(M^{-1}(\log M)^2)$$

by (2.1), and for $T_{\lceil M \rceil}$ the number of points in $T(v < b_0 M^{-1} \log M)$ is binomial distributed with mean $\lceil M \rceil p$. Denote by E_m the event that precisely m points of the Poisson or binomial process are in $T(v \geq b_0 M^{-1} \log M)$. Coupling both processes in

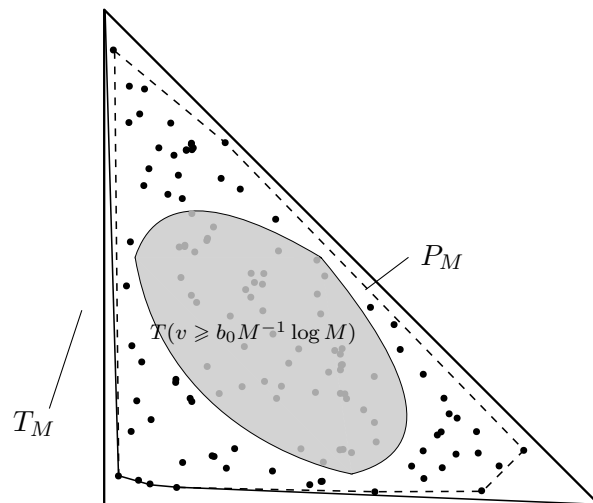


Figure 3.2. Illustration of the construction used in the proof of Lemma 3.8. The convex hull P_M (or P_χ) is indicated by the dashed segments, while the convex chain T_M (or T_χ) by a solid line. The floating body of T is drawn in grey.

the canonical way and using Lemma 2.4, Lemma 2.3 with $k = 0$, together with (3.20) yields

$$\begin{aligned}
 & \left| \mathbb{P} \left(f_0 \left(T_{\lceil M \rceil} \right) \leq x \right) - \mathbb{P} \left(f_0(T_\chi) \leq x \right) \right| \\
 & \leq \sum_{m=0}^{\infty} \mathbb{P} \left(f_0(T_M) \leq x, \mathcal{A}_M \mid E_m \right) \left| \frac{(Mp)^m}{m!} e^{-Mp} - \binom{\lceil M \rceil}{m} p^m (1-p)^{n-m} \right| + O(M^{-6}) \\
 & \leq \sum_{m=0}^{\infty} \left| \frac{(\lceil M \rceil p)^m}{m!} e^{-\lceil M \rceil p} - \frac{(Mp)^m}{m!} e^{-Mp} \right| \\
 & \quad + \sum_{m=0}^{\infty} \left| \frac{(\lceil M \rceil p)^m}{m!} e^{-\lceil M \rceil p} - \binom{\lceil M \rceil}{m} p^m (1-p)^{n-m} \right| + O(M^{-6}) \\
 & \leq 2p + 2p + O(M^{-6}) \\
 & = O(M^{-1}(\log M)^2).
 \end{aligned}$$

Thus condition (iii) in Lemma 3.6 holds with $\varepsilon_3(M) = c_3 M^{-1}(\log M)^2$ for some $c_3 > 0$ independent of M . An application of Lemma 3.6 finishes the proof of Lemma 3.8. \square

3.4. Step 4: Berry–Esseen bound for the Poisson model under condition \mathcal{E}

In the next step we consider the random variable $f_0(P_\eta)$, conditioned on the event \mathcal{E} we introduced in Step 2:

$$\xi := (f_0(P_\eta) | \mathcal{E}).$$

Let us also condition on the positions of the points $Z_1, \dots, Z_\ell \in \eta$ and introduce the random variable

$$\xi' := (f_0(P_\eta) \mid \mathcal{E}, Z_1, \dots, Z_\ell).$$

It should be mentioned that under \mathcal{E} all ℓ points Z_1, \dots, Z_ℓ are well defined and in fact distinct. Note that, conditionally on the positions of Z_1, \dots, Z_ℓ , the restriction of η to the interior of the polygon $L^+(Z_1) \cap \dots \cap L^+(Z_\ell)$ is again a Poisson point process with the same intensity. Indeed, this follows on the one hand from the fact that the ℓ -fold reduced Palm distribution of a stationary Poisson point process coincides with its distribution, see [Han82, Corollary 1]. Since after fixing Z_1, \dots, Z_ℓ , the triangles Δ_i , $1 \leq i \leq \ell$ are fixed and they have disjoint interiors, the restriction of η to the interior of Δ_i is again a Poisson point process whose intensity measure is n times the Lebesgue measure restricted to Δ_i , see [LP18, Theorem 5.2]. Hence, the random variable ξ' can be decomposed into the sum of ℓ independent random variables ξ'_i , $1 \leq i \leq \ell$, where each ξ'_i is defined as a number of vertices of the random convex chain, formed by the Poisson point process η restricted to the triangle Δ_i . More precisely, given an arbitrary triangle $\Delta \subset P$ with vertices v_1, v_2, v_3 and a Poisson point process η we define

$$T_\eta(\Delta, v_1, v_2) := [(\eta \cap \Delta), v_1, v_2].$$

Then we take

$$\xi'_i := f_0(T_\eta(\Delta_i, Z_i, Z_{i+1})) - 2,$$

where the -2 is coming from the fact that we exclude the two endpoints Z_i and Z_{i+1} of the convex chain. Now, consider for each $1 \leq i \leq \ell$ the affine transformation $A_i : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ which maps the triangle Δ_i with vertices Z_i, V_i and Z_{i+1} to the canonical triangle T with vertices $(0, 1)$, $(0, 0)$ and $(1, 0)$. Using the mapping property [LP18, Theorem 5.1] and restriction property [LP18, Theorem 5.2] of Poisson point processes we conclude that $\eta_i := A_i(\eta \cap \Delta_i)$ is a homogeneous Poisson point process on T with intensity $2n \text{ area}(\Delta_i)$.

Since the number of vertices is invariant under affine transformations we conclude that

$$\xi'_i \stackrel{d}{=} f_0(T_{\eta_i}) - 2.$$

In order to prove a Berry–Esseen bound for the random variable

$$\xi' = \sum_{i=1}^{\ell} \xi'_i + \ell,$$

where the additional summand $+\ell$ is coming from the fact that we excluded in the definition of ξ_1, \dots, ξ_ℓ the points Z_1, \dots, Z_ℓ , we will use the following lemma.

LEMMA 3.9. — *Let X_1, \dots, X_k be independent random variables with $\mu_i := \mathbb{E}X_i < \infty$, $\sigma_i := \sqrt{\text{Var } X_i} \in (0, \infty)$, $1 \leq i \leq k$ and let G_1, \dots, G_k be independent standard Gaussian random variables. Let $\varepsilon_i > 0$, $1 \leq i \leq k$ be such that*

$$(3.21) \quad \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{X_i - \mu_i}{\sigma_i} \leq x \right) - \mathbb{P}(G_i \leq x) \right| \leq \varepsilon_i, \quad 1 \leq i \leq k.$$

Then for $X := X_1 + \dots + X_k$ we have

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{X - \mathbb{E}X}{\sqrt{\text{Var } X}} \leq x \right) - \Phi(x) \right| \leq \sum_{i=1}^k \varepsilon_i.$$

Proof. — Let \mathcal{F} be the space of cumulative distribution functions, namely

$$\mathcal{F} := \left\{ F : \mathbb{R} \mapsto [0, 1] : F \text{ right-continuous, monotone increasing,} \right. \\ \left. F(-\infty) = 0, F(\infty) = 1 \right\},$$

where $F(\pm\infty)$ has to be interpreted as the appropriate limit. First of all let us recall that the classical Kolmogorov (or uniform) metric $d : \mathcal{F} \times \mathcal{F} \rightarrow [0, \infty)$ on the space \mathcal{F} is defined as

$$d(F_1, F_2) := \sup_{x \in \mathbb{R}} |F_1(x) - F_2(x)|.$$

Given two random variables X, Y with cumulative distribution functions F_X, F_Y , respectively, we write

$$d(X, Y) := d(F_X, F_Y) = \sup_{x \in \mathbb{R}} |\mathbb{P}(X \leq x) - \mathbb{P}(Y \leq x)|.$$

Using this notation and the fact that $d(aX + b, aY + b) = d(X, Y)$ for any $a > 0, b \in \mathbb{R}$ the conditions in (3.21) can be written in the form

$$(3.22) \quad d(X_i - \mu_i, \sigma_i G_i) \leq \varepsilon_i, \quad 1 \leq i \leq k.$$

Further, we apply the so-called semi-additivity property of the Kolmogorov metric, which says that for any independent Y_1, \dots, Y_k and any independent Y'_1, \dots, Y'_k one has that

$$d(Y_1 + \dots + Y_k, Y'_1 + \dots + Y'_k) \leq \sum_{i=1}^n d(Y_i, Y'_i),$$

see [Zol76, Section 2.3, Equation (1.2)]. Recalling that $X = X_1 + \dots + X_k$, and taking $Y_i = X_i - \mu_i$ and $Y'_i = \sigma_i G_i$ we conclude by (3.22) that

$$d(X - \mathbb{E}X, \sigma_1 G_1 + \dots + \sigma_k G_k) = d \left(\frac{X - \mathbb{E}X}{\sqrt{\text{Var } X}}, \frac{\sigma_1 G_1 + \dots + \sigma_k G_k}{(\sigma_1^2 + \dots + \sigma_k^2)^{1/2}} \right) \leq \sum_{i=1}^k \varepsilon_i.$$

Finally, we need to observe that $(\sigma_1 G_1 + \dots + \sigma_k G_k)/(\sigma_1^2 + \dots + \sigma_k^2)^{1/2}$ has the standard Gaussian distribution. This completes the proof of Lemma 3.9. \square

We apply Lemma 3.9 with $k = \ell$ to the random variables $X_1 := \xi'_1, \dots, X_\ell := \xi'_\ell$.

COROLLARY 3.10. — *There exists a constant $c > 0$ such that for any $n \geq 4$ we have that*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\xi' - \mathbb{E}\xi'}{\sqrt{\text{Var } \xi'}} \leq x \right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log n}}.$$

Proof. — Lemma 3.8 yields (3.21) for the random variables $X_i = \xi'_i$ with $\varepsilon_i = c(\log M)^{-1/2}$, $1 \leq i \leq \ell$. It is clear that for $X := \xi' - \ell = \sum_{i=1}^\ell \xi'_i$ we have $X - \mathbb{E}X = \xi' - \mathbb{E}\xi'$ and $\text{Var } X = \text{Var } \xi'$. Moreover, according to (3.2) we have $\text{area}(\Delta_i) \geq n^{-\frac{1}{2}}$ and

$$(3.23) \quad M = n \text{ area}(\Delta_i) \geq n^{\frac{1}{2}} \geq 2$$

given \mathcal{E} . Thus $\varepsilon_i \leq c(\frac{1}{2} \log n)^{-1/2}$ for all $1 \leq i \leq \ell$, and the proof of Corollary 3.10 is complete. \square

Note that the obtained bound is independent of the exact position of the points Z_1, \dots, Z_ℓ if we condition on the event \mathcal{E} . This already suggests that the same bound holds for the random variable $f_0(P_\eta)$ conditionally on \mathcal{E} only. Our next result ensures that this is indeed the case.

LEMMA 3.11. — *There exists a constant $c > 0$ such that for any $n \geq 2$ we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{(f_0(P_\eta)|\mathcal{E}) - \mu}{\sigma} \leq x \right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log n}}$$

with

$$\mu = \mathbb{E}(f_0(P_\eta) | \mathcal{E}) = \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log F_i + \frac{2\gamma\ell}{3} + O(n^{-\frac{1}{4}} \log n)$$

and

$$\sigma^2 = \text{Var}(f_0(P_\eta) | \mathcal{E}) = \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma - 2\pi^2)\ell}{27} + O(n^{-\frac{1}{4}} (\log n)^2).$$

Proof. — Let $\mathbf{Z} = (Z_1, \dots, Z_\ell)$. Recall that $\xi = (f_0(P_\eta)|\mathcal{E})$ and $\xi' = (f_0(P_\eta)|\mathcal{E}, \mathbf{Z})$. Moreover, we define $\mu := \mathbb{E}\xi$, $\mu' := \mathbb{E}\xi'$ and $\sigma^2 := \text{Var}\xi$, $\sigma'^2 := \text{Var}\xi'$ (in this proof we suppress the dependence on the parameter n in our notation). Using the representation $\xi' = \sum_{i=1}^{\ell} \xi'_i + \ell$ together with Lemma 3.7, thanks to the condition \mathcal{E} , we obtain

$$\begin{aligned} \mu' &= \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + \frac{(2\gamma + 4)\ell}{3} + O(n^{-\frac{1}{4}}), \\ \sigma'^2 &= \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + \frac{(10\gamma + 2\pi^2 - 28)\ell}{27} + O(n^{-\frac{1}{4}}). \end{aligned}$$

Note, that we used the estimate (3.23) for M here. Corollary 3.5 shows that

$$\begin{aligned} \mu &= \mathbb{E}(\mu' | \mathcal{E}) = \mathbb{E} \left(\frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + \frac{(2\gamma + 4)\ell}{3} + O(n^{-\frac{1}{4}}) \middle| \mathcal{E} \right) \\ &= \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log F_i + \frac{2\gamma\ell}{3} + O(n^{-\frac{1}{4}} \log n), \end{aligned}$$

which already coincides with the expectation (1.1) by Rényi and Sulanke. Analogously, the expected conditional variance is given by

$$\begin{aligned} \mathbb{E}(\sigma'^2 | \mathcal{E}) &= \mathbb{E} \left(\frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + \frac{(10\gamma + 2\pi^2 - 28)\ell}{27} + O(n^{-\frac{1}{4}}) \right) \\ &= \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma + 2\pi^2 - 48)\ell}{27} + O(n^{-\frac{1}{4}} \log n). \end{aligned}$$

And for the variance of the expectation we use Corollary 3.5 again,

$$\begin{aligned} \text{Var}(\mu'|\mathcal{E}) &= \frac{4}{9} \text{Var} \left(\sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + O(n^{-\frac{1}{4}}) \middle| \mathcal{E} \right) \frac{4}{9} \sum_{i=1}^{\ell} \text{Var}(\log \text{area}(\Delta_i) | \mathcal{E}) \\ &\quad + \frac{8}{9} \sum_{i=1}^{\ell} \text{Cov} \left(\log \text{area}(\Delta_i), \log \text{area}(\Delta_{i+1}) \middle| \mathcal{E} \right) + O(n^{-\frac{1}{4}} \log n) \\ &= \frac{8}{9} \ell \left(2 - \frac{\pi^2}{6} \right) + O(n^{-\frac{1}{4}} (\log n)^2). \end{aligned}$$

Hence, by the law of total variance,

$$\begin{aligned} \sigma^2 &= \mathbb{E}_{\mathbf{Z}} \left(\underbrace{\text{Var}(f_0(P_{\eta}) | \mathcal{E}, \mathbf{Z})}_{=\sigma'^2} \middle| \mathcal{E} \right) + \text{Var}_{\mathbf{Z}} \left(\underbrace{\mathbb{E}(f_0(P_{\eta}) | \mathcal{E}, \mathbf{Z})}_{=\mu'} \middle| \mathcal{E} \right) \\ &= \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma - 2\pi^2)\ell}{27} + O(n^{-\frac{1}{4}} (\log n)^2). \end{aligned}$$

This also shows that

$$(3.24) \quad \mu - \mu' = \frac{2}{3} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + O(1),$$

$$(3.25) \quad \sigma'^2 - \sigma^2 = \frac{10}{27} \sum_{i=1}^{\ell} \log \text{area}(\Delta_i) + O(1),$$

$$(3.26) \quad \sigma + \sigma' \geq \sigma = \left(\frac{10\ell}{27} \log n \right)^{\frac{1}{2}} + O(1).$$

Next, we observe that

$$\begin{aligned} (3.27) \quad & \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\xi - \mu}{\sigma} \leq x \right) - \Phi(x) \right| \\ &= \sup_{y \in \mathbb{R}} \left| \mathbb{P}(\xi \leq y) - \Phi \left(\frac{y - \mu}{\sigma} \right) \right| \\ &\leq \sup_{y \in \mathbb{R}} \left| \mathbb{E} \mathbb{P}(\xi' \leq y) - \mathbb{E} \Phi \left(\frac{y - \mu'}{\sigma'} \right) \right| + \sup_{y \in \mathbb{R}} \left| \mathbb{E} \Phi \left(\frac{y - \mu'}{\sigma'} \right) - \Phi \left(\frac{y - \mu}{\sigma} \right) \right| \\ &\leq \mathbb{E} \sup_{y \in \mathbb{R}} \left| \mathbb{P}(\xi' \leq y) - \Phi \left(\frac{y - \mu'}{\sigma'} \right) \right| + \mathbb{E} \sup_{y \in \mathbb{R}} \left| \Phi \left(\frac{y - \mu'}{\sigma'} \right) - \Phi \left(\frac{y - \mu}{\sigma} \right) \right|, \end{aligned}$$

where the expectation is taken with respect to the law of the random vector \mathbf{Z} . From Corollary 3.10 we have

$$\begin{aligned} (3.28) \quad & \mathbb{E} \sup_{y \in \mathbb{R}} \left| \mathbb{P}(\xi' \leq y) - \Phi \left(\frac{y - \mu'}{\sigma'} \right) \right| \\ &= \mathbb{E} \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\xi' - \mu'}{\sigma'} \leq x \right) - \Phi(x) \right| = O \left(\frac{1}{\sqrt{\log n}} \right). \end{aligned}$$

It remains to deal with the random variable

$$Y_{\mathbf{Z}} := \sup_{y \in \mathbb{R}} \left| \Phi \left(\frac{y - \mu'}{\sigma'} \right) - \Phi \left(\frac{y - \mu}{\sigma} \right) \right|.$$

Assume that the supremum is attained at y_0 . Then, with $\phi(t) := \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ the density of standard normal distribution,

$$(3.29) \quad Y_{\mathbf{Z}} = \left| \Phi \left(\frac{y_0 - \mu'}{\sigma'} \right) - \Phi \left(\frac{y_0 - \mu}{\sigma} \right) \right| \\ \leq \sup_{t \in \mathbb{R}} |\phi(t)| \cdot \left| \frac{y_0 - \mu'}{\sigma'} - \frac{y_0 - \mu}{\sigma} \right| \leq \left| \frac{y_0 - \mu'}{\sigma'} - \frac{y_0 - \mu}{\sigma} \right|,$$

where y_0 is such that

$$\frac{1}{\sigma'} \phi \left(\frac{y_0 - \mu'}{\sigma'} \right) = \frac{1}{\sigma} \phi \left(\frac{y_0 - \mu}{\sigma} \right).$$

By taking logarithms on both sides and putting all terms to the left hand side of the equality, we see that the last equation is equivalent to

$$y_0^2 (\sigma^2 - \sigma'^2) - 2y_0 (\mu' \sigma^2 - \mu \sigma'^2) + \mu'^2 \sigma^2 - \mu^2 \sigma'^2 - \sigma^2 \sigma'^2 \log \left(\frac{\sigma}{\sigma'} \right) = 0.$$

This quadratic equation has the following solutions:

$$y_0^{\pm} = \frac{\mu' \sigma^2 - \mu \sigma'^2 \pm \sigma \sigma' \sqrt{(\mu' - \mu)^2 + \log(\sigma/\sigma') |\sigma'^2 - \sigma^2|}}{\sigma^2 - \sigma'^2}.$$

Substituting this back into (3.29) leads to the bound

$$Y_{\mathbf{Z}} \leq \frac{|\mu' - \mu| + \sqrt{(\mu' - \mu)^2 + \log(\sigma/\sigma') |\sigma'^2 - \sigma^2|}}{\sigma + \sigma'} \\ \leq 2 \frac{|\mu' - \mu|}{\sigma + \sigma'} + \frac{\sqrt{|\log(\sigma/\sigma')| |\sigma'^2 - \sigma^2|}}{\sigma + \sigma'},$$

where we used the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for all $a, b > 0$. Observe that given \mathcal{E} we have $\text{area}(\Delta_i) \geq c n^{-\frac{1}{2}}$ and thus $\sigma'^2 \geq \frac{5\ell}{27} \log n + O(1)$. Hence, there exist constants $c_1, c_2 > 0$ such that $c_1 < \sigma/\sigma' < c_2$. Further, $\sqrt{|\sigma'^2 - \sigma^2|} \leq |\sigma'^2 - \sigma^2| + 1$. Thus, using (3.24), (3.25), and (3.26) we conclude that, for some constant $C_1 > 0$,

$$Y_{\mathbf{Z}} \leq C_1 \left(\sum_{i=1}^{\ell} \frac{|\log \text{area}(\Delta_i)|}{\sqrt{\log n}} + \frac{1}{\sqrt{\log n}} \right),$$

and Corollary 3.5 yields, for another constant $C_2 > 0$,

$$\mathbb{E} Y_{\mathbf{Z}} \leq C_2 \left(\frac{\mathbb{E} |\log \text{area}(\Delta_1)|}{\sqrt{\log n}} + \frac{1}{\sqrt{\log n}} \right) = \frac{O(1)}{\sqrt{\log n}}.$$

Together with (3.27) and (3.28) this completes the proof of Lemma 3.11. \square

3.5. Step 5: Removing the condition \mathcal{E}

In order to remove the remaining condition \mathcal{E} in Lemma 3.11 we use Lemma 3.6 again. We will apply this lemma to the random variables $\xi'_n = f_0(P_\eta|\mathcal{E})$ and $\xi_n = f_0(P_\eta)$. Note that condition (iv) then follows from Lemma 3.11. Checking the other conditions requires a more careful analysis.

LEMMA 3.12. — *The random variables $\xi_n := f_0(P_\eta)$ and $\xi'_n := f_0(P_\eta|\mathcal{E})$ satisfy conditions (i)–(iii) of Lemma 3.6 with $\varepsilon_1(n) = c_1 n^{-\frac{1}{4}}(\log n)^{\frac{3}{2}}$, $\varepsilon_2(n) = c_2 n^{-\frac{1}{4}}(\log n)^3$, $\varepsilon_3(n) = c_3 n^{-\frac{1}{4}}$, where $c_1, c_2, c_3 > 0$ are positive constants not depending on n .*

Proof. — The proof of this lemma will basically follow the lines of the proof of [BR10b, Lemma 8.2]. We will start by estimating $\mathbb{E}(f_0(P_\eta)^k|\mathcal{E})$ for $k = 1, 2$. For this we assume additionally that the event $\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi$ holds, that is

$$P(v \geq b_0 n^{-1} \log n) \subset P_\eta \quad \text{and} \quad \#(\eta \cap P(v < b_0 n^{-1} \log n)) \leq c_0 (\log n)^2,$$

and make use of Lemma 2.2. It follows that $f_0(P_\eta)^k \mathbb{1}(\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi) \leq c_0^k (\log n)^{2k}$. As a consequence,

$$\begin{aligned} \mathbb{E}(f_0(P_\eta)^k | \mathcal{E}) &= \mathbb{E}(f_0(P_\eta)^k \mathbb{1}(\overline{\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi}) | \mathcal{E}) + \mathbb{E}(f_0(P_\eta)^k \mathbb{1}(\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi) | \mathcal{E}) \\ &\leq \frac{\mathbb{E}((\#\eta)^k \mathbb{1}(\overline{\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi}) \mathbb{1}(\mathcal{E}))}{\mathbb{P}(\mathcal{E})} + C (\log n)^{2k} \\ &\leq \frac{(\mathbb{E}(\#\eta)^{2k})^{\frac{1}{2}} \mathbb{P}(\overline{\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi})^{\frac{1}{2}}}{\mathbb{P}(\mathcal{E})} + C (\log n)^{2k}, \end{aligned}$$

by Hölder's inequality and where $C > 0$ is some constant. Because for $m \geq 1$, $\mathbb{E}N^m \leq m^m(n^m + 1)$ for a Poisson random variable N with mean np ,

$$(\mathbb{E}(\#\eta)^{2k})^{\frac{1}{2}} = O(n^k),$$

and from Corollary 3.2 and Lemma 2.2 we see that

$$\frac{\mathbb{P}(\overline{\mathcal{A}_n^\pi \cap \mathcal{B}_n^\pi})^{\frac{1}{2}}}{\mathbb{P}(\mathcal{E})} = O(n^{-2}).$$

Thus, for $k = 1, 2$

$$(3.30) \quad \mathbb{E}(f_0(P_\eta)^k | \mathcal{E}) = O((\log n)^{2k}).$$

In order to verify conditions (i)–(iii) in Lemma 3.6 we will use the following simple inequality from [BR10b, Claim 8.3], which says that

$$|\mathbb{E}(\zeta) - \mathbb{E}(\zeta|A)| \leq (\mathbb{E}(\zeta|A) + \mathbb{E}(\zeta|\overline{A})) \mathbb{P}(\overline{A}),$$

for any non-negative random variable ζ and any event A .

For condition (i) we take $\zeta = \mathbb{E}f_0(P_\eta)$ and $A = \mathcal{E}$. By Lemma 3.11 we get $\mathbb{E}(f_0(P_\eta)|\mathcal{E}) = O(\log n)$. Using this and (3.30), and the estimate in Corollary 3.2 for $\mathbb{P}(\mathcal{E})$, we conclude that

$$(3.31) \quad |\mathbb{E}(f_0(P_\eta)) - \mathbb{E}(f_0(P_\eta)|\mathcal{E})| = O(n^{-\frac{1}{4}}(\log n)^2),$$

and, thus,

$$(3.32) \quad \mathbb{E}f_0(P_\eta) = \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log F_i + \frac{2\gamma\ell}{3} + O\left(n^{-\frac{1}{4}}(\log n)^2\right).$$

In the same way, for condition (ii) we take $\zeta = \mathbb{E}f_0(P_\eta)^2$ and $A = \mathcal{E}$. Then

$$\begin{aligned} |\mathrm{Var}(f_0(P_\eta)) - \mathrm{Var}(f_0(P_\eta)|\mathcal{E})| \\ \leq \left| \mathbb{E}f_0(P_\eta)^2 - \mathbb{E}(f_0(P_\eta)^2|\mathcal{E}) \right| + \left| (\mathbb{E}f_0(P_\eta))^2 - (\mathbb{E}(f_0(P_\eta)|\mathcal{E}))^2 \right|. \end{aligned}$$

For the first term we get from Lemma 3.11 the bound

$$\mathbb{E}(f_0(P_\eta)^2|\mathcal{E}) = \mathrm{Var}(f_0(P_\eta)|\mathcal{E}) + (\mathbb{E}(f_0(P_\eta)|\mathcal{E}))^2 = O((\log n)^2),$$

and combine this with (3.30) and Corollary 3.2, in order to obtain

$$\begin{aligned} \left| \mathbb{E}f_0(P_\eta)^2 - \mathbb{E}(f_0(P_\eta)^2|\mathcal{E}) \right| &\leq \left(\mathbb{E}(f_0(P_\eta)^2|\mathcal{E}) + \mathbb{E}(f_0(P_\eta)^2|\bar{\mathcal{E}}) \right) \mathbb{P}(\bar{\mathcal{E}}) \\ &= O\left(n^{-\frac{1}{4}}(\log n)^4\right). \end{aligned}$$

For the second term we note that again by Lemma 3.11 and (3.31) we get

$$\left| (\mathbb{E}f_0(P_\eta))^2 - (\mathbb{E}(f_0(P_\eta)|\mathcal{E}))^2 \right| = O\left(n^{-\frac{1}{4}}(\log n)^3\right).$$

Putting these estimates together we conclude that

$$|\mathrm{Var}(f_0(P_\eta)) - \mathrm{Var}(f_0(P_\eta)|\mathcal{E})| = O\left(n^{-\frac{1}{4}}(\log n)^4\right),$$

implying that

$$(3.33) \quad \mathrm{Var}(f_0(P_\eta)) = \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma - 2\pi^2)\ell}{27} + O\left(n^{-\frac{1}{4}}(\log n)^4\right).$$

This shows that $\varepsilon_2(n) = c_2 n^{-\frac{1}{4}}(\log n)^3$, where $c_2 > 0$ is some absolute constant. Analogously, by (3.31) we set $\varepsilon_1(n) = c_1 n^{-\frac{1}{4}}(\log n)^{\frac{3}{2}}$, where $c_1 > 0$ is some absolute constant.

Finally, taking $\zeta = \mathbb{1}(f_0(P_\eta) \leq x)$ and $A = \mathcal{E}$ we obtain from Corollary 3.2 that

$$|\mathbb{P}(f_0(P_\eta)|\mathcal{E} \leq x) - \mathbb{P}(f_0(P_\eta) \leq x)| \leq 2\mathbb{P}(\bar{\mathcal{E}}) = O\left(n^{-\frac{1}{4}}\right).$$

This completes the argument. \square

By Lemma 3.6 and Lemma 3.12 together with Lemma 3.11 we conclude the required Berry–Esseen bound for the Poisson model of random polygons. Expectation and variance have been obtained in (3.32) and (3.33).

THEOREM 3.13. — *Consider the Poisson random polygon P_η induced by a homogeneous Poisson point process η in a polygon P of unit area with $\mathbb{E}(\#\eta) = n$. Then, for any $n \geq 2$,*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(\frac{f_0(P_\eta) - \mathbb{E}f_0(P_\eta)}{\sqrt{\mathrm{Var} f_0(P_\eta)}} \geq x\right) - \Phi(x) \right| \leq \frac{c}{\sqrt{\log n}}$$

for some constant $c > 0$ independent of n , with

$$\mathbb{E}f_0(P_\eta) = \frac{2\ell}{3} \log n + \frac{2}{3} \sum_{i=1}^{\ell} \log F_i + \frac{2\gamma\ell}{3} + O\left(n^{-\frac{1}{4}}(\log n)^2\right)$$

and

$$\text{Var}(f_0(P_\eta)) = \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma - 2\pi^2)\ell}{27} + O\left(n^{-\frac{1}{4}}(\log n)^4\right).$$

3.6. Step 6: Going back to uniform model

This is the last step in the proof of Theorem 1.1, in which we apply Lemma 3.6 with $\xi'_n = f_0(P_\eta)$ and $\xi_n = f_0(P_n)$. Condition (iv) there holds due to Theorem 3.13 with $\varepsilon_4(n) = c_4/\sqrt{\log n}$ for some $c_4 > 0$. Condition (i) with $\varepsilon_1(n) = c_1/\sqrt{\log n}$, $c_1 > 0$, follows from (1.1) and Theorem 3.13.

Proof. — In order to check condition (ii) we need a more precise asymptotics for the variance of $f_0(P_n)$ compared to the one given by (1.2). In particular we need a result of the form $\text{Var}(f_0(P_n)) = \frac{10\ell}{27} \log n + O(1)$. First note that by the trivial estimate $f_0(P_\eta) \leq \#\eta$ and using Lemma 2.2 we have

$$\begin{aligned} \left| \mathbb{E}\left(f_0(P_\eta)^k \mathbb{1}(\mathcal{A}_n^\pi)\right) - \mathbb{E}f_0(P_\eta)^k \right| &= \left| \mathbb{E}\left(f_0(P_\eta)^k (\mathbb{1}(\mathcal{A}_n^\pi) - 1)\right) \right| \\ (3.34) \qquad \qquad \qquad &\leq \left(\mathbb{E}\left(f_0(P_\eta)^{2k}\right)^{\frac{1}{2}} (\mathbb{E}(\mathbb{1}(\mathcal{A}_n^\pi) - 1)^2)^{\frac{1}{2}} \right) \\ &= O\left(n^k \mathbb{P}(\overline{\mathcal{A}_n^\pi})^{\frac{1}{2}}\right) = O\left(n^{k-3}\right) \end{aligned}$$

for $k = 1, 2$, say. The definition of the variance gives

$$\begin{aligned} &|\text{Var}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \text{Var} f_0(P_\eta)| \\ &\leq \left| \mathbb{E}\left(f_0(P_\eta)^2 \mathbb{1}(\mathcal{A}_n^\pi)\right) - \mathbb{E}f_0(P_\eta)^2 \right| + \left| \left(\mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi))\right)^2 - (\mathbb{E}f_0(P_\eta))^2 \right|. \end{aligned}$$

We use now (3.34) to bound the first term by n^{-1} . Theorem 3.13 and (3.34) show, that

$$\begin{aligned} &\left| \left(\mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi))\right)^2 - (\mathbb{E}f_0(P_\eta))^2 \right| \\ &= |\mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \mathbb{E}f_0(P_\eta)| |\mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) + \mathbb{E}f_0(P_\eta)| \end{aligned}$$

is bounded (up to multiplicative constant) by $n^{-2} \log n = O(n^{-1})$. Hence

$$|\text{Var}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \text{Var} f_0(P_\eta)| = O(n^{-1}).$$

The same result holds for $\text{Var}(f_0(P_n) \mathbb{1}(\mathcal{A}_n)) - \text{Var} f_0(P_n)$ with an actually slightly simpler proof because $f_0(P_n) \leq n$ in which n is non-random. These estimates show that

$$\begin{aligned}
& |\mathrm{Var} f_0(P_\eta) - \mathrm{Var} f_0(P_n)| \\
& \leq \left| \mathrm{Var}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \mathrm{Var}(f_0(P_n) \mathbb{1}(\mathcal{A}_n)) \right| + O(n^{-1}) \\
& \leq \left| \mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi))^2 - \mathbb{E}(f_0(P_n) \mathbb{1}(\mathcal{A}_n))^2 \right| \\
& \quad + \left| \mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \mathbb{E}(f_0(P_n) \mathbb{1}(\mathcal{A}_n)) \right| (\mathbb{E} f_0(P_\eta) + \mathbb{E} f_0(P_n)) + O(n^{-1}).
\end{aligned}$$

For P_η we have that $\# \eta \cap P(v < b_0 n^{-1} \log n)$ is Poisson distributed with mean

$$np := n \operatorname{area} \left(P(v < b_0 n^{-1} \log n) \right) = O \left((\log n)^2 \right),$$

by (2.1), and for P_n the number of points in $P(v < b_0 n^{-1} \log n)$ is binomial distributed with mean p . Denote by E_m the event that precisely m points of the Poisson or binomial process are in $P(v < b_0 n^{-1} \log n)$. Coupling both processes in the canonical way yields

$$\begin{aligned}
& \left| \mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi))^k - \mathbb{E}(f_0(P_n) \mathbb{1}(\mathcal{A}_n))^k \right| \\
& = \sum_{m=0}^{\infty} \mathbb{E} \left((f_0(P_n) \mathbb{1}(\mathcal{A}_n))^k | E_m \right) \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right| \\
& \leq \sum_{m=0}^{\infty} m^k \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right|.
\end{aligned}$$

Lemma 2.3 and the fact that the expected number of vertices in both models is of order $\log n$ implies

$$|\mathrm{Var} f_0(P_\eta) - \mathrm{Var} f_0(P_n)| = O \left(n^{-1} (\log n)^5 \right).$$

Thus,

$$\mathrm{Var} f_0(P_n) = \frac{10\ell}{27} \log n + \frac{10}{27} \sum_{i=1}^{\ell} \log F_i + \frac{(10\gamma - 2\pi^2)\ell}{27} + O \left(n^{-\frac{1}{4}} (\log n)^4 \right),$$

which proves the variance expansion in Theorem (1.1) and shows that condition (ii) holds with $\varepsilon_2(n) = c_2 n^{-1} (\log n)^4$ for some $c_2 > 0$ independent of n .

The slightly better estimate for the expectation in Theorem 1.1 (in comparison to (1.1)) follows analogously from Theorem 3.13 and the estimate

$$\begin{aligned}
|\mathbb{E} f_0(P_\eta) - \mathbb{E} f_0(P_n)| & \leq \left| \mathbb{E}(f_0(P_\eta) \mathbb{1}(\mathcal{A}_n^\pi)) - \mathbb{E}(f_0(P_n) \mathbb{1}(\mathcal{A}_n)) \right| + O(n^{-1}) \\
& = O \left(n^{-1} (\log n)^4 \right).
\end{aligned}$$

Condition (iii) can be verified following the same approach, which was already used in Lemma 3.8. By Lemma 2.1, Lemma 2.2 and Lemma 2.3,

$$\begin{aligned}
 & \left| \mathbb{P}(f_0(P_n) \leq x) - \mathbb{P}(f_0(P_\eta) \leq x) \right| \\
 &= \left| \mathbb{P}(f_0(P_n) \leq x, \mathcal{A}_n) - \mathbb{P}(f_0(P_\eta) \leq x, \mathcal{A}_n^\pi) \right| + O(n^{-6}) \\
 &\leq \sum_{m=0}^{\infty} \left| \frac{(np)^m}{m!} e^{-np} - \binom{n}{m} p^m (1-p)^{n-m} \right| + O(n^{-6}) \\
 &\leq 2p + O(n^{-6}) \\
 &= O\left(n^{-1}(\log n)^2\right).
 \end{aligned}$$

Thus, (ii) holds with $\varepsilon_3(n) = c_3 n^{-1}(\log n)^2$ for some constant $c_3 > 0$ independent of n . Finally, by combining all estimates and using Lemma 3.6 we complete the proof of Theorem 1.1. \square

Acknowledgements

We would like to thank Sam Johnston (Bath) for an enlightening discussion on the content of Lemma 3.9. Moreover, we thank two anonymous referees for careful reading and for stimulating comments.

BIBLIOGRAPHY

- [BB93] Imre Bárány and Christian Buchta, *Random polytopes in a convex polytope, independence of shape, and concentration of vertices.*, Math. Ann. **297** (1993), no. 3, 467–498. ↑880
- [BD97] Imre Bárány and Leoni Dalla, *Few points to generate a random polytope*, Mathematika **44** (1997), no. 2, 325–331. ↑881
- [BR06] Imre Bárány and Matthias Reitzner, *Random polytopes*, 2006, Preprint available at <https://citeseerx.ist.psu.edu/doc/10.1.1.77.2284>. ↑877, 883
- [BR10a] ———, *On the variance of random polytopes*, Adv. Math. **225** (2010), no. 4, 1986–2001. ↑884
- [BR10b] ———, *Poisson polytopes*, Ann. Probab. **38** (2010), no. 4, 1507–1531. ↑877, 878, 879, 881, 883, 884, 890, 900
- [Buc05] Christian Buchta, *An identity relating moments of functionals of convex hulls*, Discrete Comput. Geom. **33** (2005), 125–142. ↑878
- [Buc12] ———, *On the boundary structure of the convex hull of random points*, Adv. Geom. **12** (2012), 179–190. ↑891
- [BV07] Imre Bárány and Van Vu, *Central limit theorems for Gaussian polytopes*, Ann. Probab. **35** (2007), 1593–1621. ↑883, 890
- [CY17] Pierre Calka and Joseph E. Yukich, *Variance asymptotics and scaling limits for random polytopes*, Adv. Math. **304** (2017), 1–55. ↑884
- [Efr65] Bradley Efron, *The convex hull of a random set of points*, Biometrika **52** (1965), 331–343. ↑878
- [Eng81] Gunnar Englund, *A Remainder Term Estimate for the Normal Approximation in Classical Occupancy*, Ann. Probab. **9** (1981), 684–692. ↑877

- [Gro88] Piet Groeneboom, *Limit theorems for convex hulls*, Probab. Theory Relat. Fields **79** (1988), no. 3, 327–368. ↑877, 893
- [GT21] Anna Gusakova and C. Thäle, *On Random Convex Chains, Orthogonal Polynomials, PF Sequences and Probabilistic limit Theorems*, Mathematika **67** (2021), no. 2, 434–446. ↑880, 883
- [Han82] Karl-Heinz Hanisch, *On inversion formulae for n -fold Palm distributions of point processes in LCS-spaces*, Math. Nachr. **106** (1982), 171–179. ↑895
- [LP18] Günter Last and Mathew Penrose, *Lectures on the Poisson Process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, 2018. ↑883, 892, 895
- [LRSY19] Raphaël Lachièze-Rey, Matthias Schulte, and Joseph E. Yukich, *Normal approximation for stabilizing functionals*, Ann. Appl. Probab. **29** (2019), no. 2, 931–993. ↑879
- [Mat82] Tadashi Matsunawa, *Some strong ε -equivalence of random variables*, Ann. Inst. Stat. Math. **34** (1982), no. 2, 209–224. ↑881
- [Par11] John Pardon, *Central limit theorems for random polygons in an arbitrary convex set*, Ann. Probab. **39** (2011), no. 3, 881–903. ↑877
- [Par12] ———, *Central limit theorems for uniform model random polygons*, J. Theor. Probab. **25** (2012), no. 3, 823–833. ↑877
- [Rei05] Matthias Reitzner, *Central limit theorems for random polytopes*, Probab. Theory Relat. Fields **133** (2005), 483–507. ↑879, 883, 893
- [RS63] Alfréd Rényi and Rolf Sulanke, *Über die konvexe Hülle von n zufällig gewählten Punkten*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **2** (1963), 75–84. ↑876
- [Sch91] Carsten Schütt, *The convex floating body and polyhedral approximation*, Isr. J. Math. **73** (1991), 65–77. ↑880
- [Ver69] Wim Vervaat, *Upper bounds for the distance in total variation between the binomial or negative binomial and the Poisson distribution*, Stat. Neerl. **23** (1969), 79–86. ↑881
- [Vu06] Van Vu, *Central limit theorems for random polytopes in a smooth convex set*, Adv. Math. **207** (2006), no. 1, 221–243. ↑883
- [Zol76] Vladimir M. Zolotarev, *Metric distances in spaces of random variables and their distributions*, Math. USSR, Sb. **30** (1976), no. 3, 373–401. ↑896

Manuscript received on 30th June 2022,
revised on 12th May 2023,
accepted on 18th June 2023.

Recommended by Editors S. Gouëzel and N. Privault.
Published under license CC BY 4.0.



eISSN: 2644–9463

This journal is a member of Centre Mersenne.



Anna GUSAKOVA
Fachbereich Mathematik,
Universität Münster, Germany
gusakova@uni-muenster.de

Matthias REITZNER
Institut für Mathematik,
Universität Osnabrück, Germany
matthias.reitzner@uni-osnabrueck.de

Christoph THÄLE
Fakultät für Mathematik,
Ruhr-Universität Bochum, Germany
christoph.thaele@rub.de