NEW KOLMOGOROV BOUNDS FOR FUNCTIONALS OF BINOMIAL POINT PROCESSES

Raphaël Lachièze-Rey, Giovanni Peccati

NEW KOLMOGOROV BOUNDS FOR FUNCTIONALS OF BINOMIAL POINT PROCESSES[1]

by Raphaël Lachièze-Rey[2] and Giovanni Peccati[3]

Abstract: We obtain explicit Berry-Esseen bounds in the Kolmogorov distance for the normal approximation of non-linear functionals of vectors of independent random variables. Our results are based on the use of Stein's method and of random difference operators, and generalise the bounds recently obtained by Chatterjee (2008), concerning normal approximations in the Wasserstein distance. In order to obtain lower bounds for variances, we also revisit the classical Hoeffding decompositions, for which we provide a new proof and a new representation. Several applications are discussed in detail: in particular, new Berry-Esseen bounds are obtained for set approximations with random tessellations, as well as for functionals of covering processes.
Key words: Berry-Esseen Bounds; Binomial Processes; Covering Processes; Random Tessellations; Stochastic Geometry; Stein's method.
2010 MSC: 60F05, 60D05

1 Introduction

1.1 Overview

Let $X = (X_1, ..., X_n)$ be a collection of independent random variables, defined on some probability space $(\Omega, \mathscr{F}, \mathbf{P})$ and taking values in some Polish space $(E, \mathscr{E})$; let $f : E^n \to \mathbb{R}$ be a measurable function such that $f(X)$ is square-integrable. The aim of the present paper is to deduce a new class of explicit upper bounds for the *Kolmogorov distance* $d_K(f(X), N)$, between the distribution of $f(X)$ and that of a Gaussian random variable $N \sim \mathscr{N}(m, \sigma^2)$ such that $m = \mathbf{E}f(X)$ and $\sigma^2 = \mathbf{Var}f(X)$. Recall that $d_K(f(X), N)$ is defined as:

$$d_K(f(X), N) = \sup_{t \in \mathbb{R}} |\mathbf{P}[f(X) \leqslant t] - \mathbf{P}[N \leqslant t]| .$$

The problem of obtaining explicit estimates on the distance between the distributions of $f(X)$ and N has been recently dealt with in the paper [4], where the author was able to apply a standard version of Stein's method (see e.g. [17]) in order to deduce effective upper bounds on the *Wasserstein distance*

$$d_W(f(X), N) = \sup_{h} |\mathbf{E}[h(f(X))] - \mathbf{E}[h(N)]| ,$$

where the supremum runs over 1-Lipschitz functions, by using a class of difference operators that we shall explicitly describe in Section 2.1 below (see e.g. [5, 14, 21] for some relevant applications of these bounds).

[1]This research has been supported by the grant F1R-MTH-PUL-12PAMP (PAMPAS) at Luxembourg University

[2]Laboratoire MAP5 Université Paris Descartes, Sorbonne Paris Cité, Paris. Email: raphael.lachieze-rey@parisdescartes.fr.

[3]Unité de Recherche en Mathématiques, Université du Luxembourg, Luxembourg. Email: giovanni.peccati@gmail.com.

It is a well known fact that upper bounds on $d_W(f(X), N)$ also yield a (typically suboptimal) bound on $d_K(f(X), N)$ via the standard relation $d_K(f(X), N) \leqslant 2\sqrt{d_W(f(X), N)}$. The challenge we are setting ourselves in the present paper is to deduce upper bounds on $d_K(f(X), N)$ that are *potentially of the same order* as the bounds on $d_W(f(X), N)$ that can be deduced from [4]. Our main abstract findings appear in the statement of Theorem 4.2 below. In order to prove our main bounds, we shall exploit some novel estimates on the solution of the Stein's equations associated with the Kolmogorov distance, that are strongly inspired by computations developed in [7, 26] in the framework of normal approximations for functionals of Poisson random measures.

Another important contribution of the present work (ses Section 2.2) is a novel representation (in terms of difference operators) of the kernels determining the *Hoeffding decomposition* (see e.g. [13, 22, 29], as well as [28, Chapter 5]) of a random variable of the type $f(X)$. This new representation is put into use for deducing effective lower bounds on $\mathbf{Var} f(X)$.

As demonstrated in the sections to follow, we are mainly interested by geometric applications and, in particular, by the normal approximation of geometric functionals whose dependency structure can be assessed by using second order difference operators. One of the applications developed in detail in Section 6.1 is that of *Voronoi set approximations*, where a given set K is estimated by the union of Voronoi cells. Remarkably, our bounds allow one to deduce normal approximation bounds for the volume approximation of sets K having a highly non-regular boundary. The present paper is associated with the work [18], where it is proved that, for a large class of sets with self-similar boundary of dimension $s > d - 1$, the variance of the volume approximation is asymptotically of the same order as $n^{-2+s/d}$ and the Kolmogorov distance between the volume approximation and the normal law is smaller than some multiple of $n^{-s/2d}$ multiplied by a logarithmic term. It turns out that the crucial feature for a set to be well behaved with respect to Voronoi approximation is its density at the boundary, which is mathematically independent of its fractal dimension (see [18] for an in-depth discussion of these phenomena). For illustrative purposes, we will also present an application of our methods to covering processes (re-obtaining the results of [11] in a slightly more general framework, see Section 6.2 below), as well as to some models already studied in [4] and [21].

In the recent reference [10], Gloria and Nolen have effectively used Theorem 4.2 below for deducing Berry-Esseen bounds in the Kolmogorov distance for the effective conductance on the discrete torus.

1.2 Plan

Section 2 contains our main results concerning decompositions of random variables. Section 3 deals with some estimates associated with Stein's method, and Section 4 contains our main abstract findings. Section 5 focusses on estimates based on second order difference operators. Finally, several applications are developed in Section 6.

From now on, every random object is defined on an adequate common probability space $(\Omega, \mathscr{F}, \mathbf{P})$, with $\mathbf{E}$ denoting expectation with respect to $\mathbf{P}$.

2 Decomposing random variables

2.1 Some difference operators

Let $(E, \mathscr{E})$ be a Polish space endowed with its Borel σ-field. Given two vectors $y = (y_1, ..., y_n) \in E^n$ and $y' = (y'_1, ..., y'_n) \in E^n$, for every $C \subseteq [n] := \{1, ..., n\}$ and every measurable function $f : E^n \to \mathbb{R}$, we denote by $f^C(y, y')$ the quantity that is obtained from $f(y)$ by replacing y_i with y'_i whenever $i \in C$. For instance, if $n = 4$ and $C = \{1, 4\}$, then

$$f^C(y, y') = f(y'_1, y_2, y_3, y'_4)$$

and

$$f^C(y', y) = f(y_1, y'_2, y'_3, y_4).$$

Given $C \subseteq [n]$, we introduce the operator

$$\Delta_C f(y, y') = f(y) - f^C(y, y').$$

When $C = \{j\}$ (to simplify the notation), we shall often write $f^{\{j\}} = f^j$ and $\Delta_{\{j\}} = \Delta_j$, for $j = 1, ..., n$, in such a way that

$$\Delta_{\{j\}} f(y, y') = \Delta_j f(y, y') = f(y) - f^j(y, y') = f(y) - f(y_1, ..., y_{j-1}, y'_j, y_{j+1}, ..., y_n),$$

and

$$\Delta_{\{j\}} f(y', y) = \Delta_j f(y', y) = f(y') - f^j(y', y) = f(y') - f(y'_1, ..., y'_{j-1}, y_j, y'_{j+1}, ..., y'_n).$$

We can canonically iterate the operator Δ_j as follows: for every $k \geq 2$ and every choice of distinct indices $1 \leqslant i_1 < \cdots < i_k \leqslant n$, the quantity $\Delta_{i_1} \cdots \Delta_{i_k} f(y, y')$, is defined as

$$\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(y, y') - (\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(y, y'))_{i_k},$$

where $(\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(y, y'))_{i_k}$ is obtained by replacing y_{i_k} with y'_{i_k} inside the argument of

$$\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(y, y').$$

Note that the operator $\Delta_{i_1} \cdots \Delta_{i_k}$ defined in this way is invariant with respect to permutations of the indices $i_1, ..., i_k$. For instance, if $n = 2$,

$$\begin{aligned} \Delta_1 \Delta_2 f(y, y') &= \Delta_2 \Delta_1 f(y, y') \\ &= f(y'_1, y'_2) - f(y'_1, y_2) - f(y_1, y'_2) + f(y_1, y_2). \end{aligned}$$

The notation introduced above also extends to random variables: if $X = (X_1, ..., X_n)$ and $X' = (X'_1, ..., X'_n)$ are two random vectors with values in E^n, then we write

$$\Delta_C f(X, X') := f(X) - f^C(X, X'), \quad C \subseteq [n],$$

and define $\Delta_{i_1} \cdots \Delta_{i_k} f(X, X')$, $1 \leqslant i_1 < \cdots < i_k \leqslant n$, exactly as above. The definitions of $\Delta_C f(X', X)$ and $\Delta_{i_1} \cdots \Delta_{i_k} f(X', X)$ are given analogously. Now assume that $E[|f(X)|] < \infty$. Our aim in this section is to discuss two representations of the quantity $f(X) - E[f(X)]$, that are based on the use of the difference operators Δ_j. The first one is a reformulation of the classical *Hoeffding decomposition* for functions of independent random variables (see e.g. [13, 22, 29], as well as [28, Chapter 5]). The second one comes from [4] (see also [5, Chapter 7]) and will play an important role in the derivation of our main estimates.

2.2 A new look at Hoeffding decompositions

Throughout this section, for every fixed integer $n \geq 1$ we write $X = (X_1, ..., X_n)$ to indicate a vector of independent random variables with values in the Polish space E, and let $X' = (X'_1, ..., X'_n)$ be an independent copy of X. If $f : E^n \to \mathbb{R}$ is a measurable function such that $\mathbf{E}[f(X)^2] < \infty$, then the classical theory of Hoeffding decompositions for functions of independent random variables (see e.g. [15, 29]) implies that $f(X)$ admits a unique decomposition of the type

$$f(X) = \mathbf{E}[f(X)] + \sum_{k=1}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \varphi_{i_1,...,i_k}(X_{i_1}, ..., X_{i_k}), \tag{2.1}$$

where the square-integrable kernels $\varphi_{i_1,...,i_k}$ verify the degeneracy condition

$$\mathbf{E}[\varphi_{i_1,...,i_k}(X_{i_1}, ..., X_{i_k}) \,|\, X_{j_1}, ..., X_{j_a}] = 0,$$

for any strict subset $\{j_1, ..., j_a\}$ of $\{i_1, ..., i_k\}$. The derivation of (2.1) is customarily based on some implicit recursive application of the inclusion-exclusion principle, and the kernels $\varphi_{i_1,...,i_k}$ can be represented as linear combinations of conditional expectations. As abundantly illustrated in the above-mentioned references, a representation such as (2.1) is extremely useful for analysing the variance of a wide range of random variables (in particular, U-statistics). Our aim in the present section is to point out a very compact way of writing the decomposition (2.1), that is based on the use of the operators Δ_j introduced above. Albeit not surprising, such an approach towards Hoeffding decompositions seems to be new and of independent interest, and will be quite useful in the present paper for explicitly deriving lower bounds on variances. Our starting point is the following statement, where we make use of the notation introduced in Section 2.1.

Lemma 2.1. *For every $f : E^n \to \mathbb{R}$*

$$f(y) - f(y') = \sum_{k=1}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (-1)^k \Delta_{i_1} \cdots \Delta_{i_k} f(y', y). \tag{2.2}$$

Proof. The key observation is that, for every $k \geq 1$ and every $B = \{i_1, ..., i_k\}$,

$$\Delta_{i_1} \cdots \Delta_{i_k} f(y', y) = \sum_{A \subseteq B} (-1)^{|A|} f^A(y', y),$$

a relation that can be easily proved by recursion. By virtue of this fact, one can now rewrite the right-hand side of (2.2) as

$$\sum_{A \subseteq [n]} \psi(A) \times Z(A), \tag{2.3}$$

where $\psi(A) := f^A(y', y)$ and $Z(A) := \sum_{B: B \neq \emptyset, A \subseteq B} (-1)^{|B \setminus A|}$. Standard combinatorial considerations yield that $Z([n]) = 1$, $Z(\emptyset) = -1$ and $Z(A) = 0$, for every non-empty strict subset of $[n]$. This implies that (2.3) is indeed equal to $\psi([n]) - \psi(\emptyset)$, and the desired conclusion follows at once. □

Now fix an integer n, as well as n-dimensional vectors X and X' as above (in particular, X' is an independent copy of X): the following statement provides an alternate description of the Hoeffding decomposition of $f(X)$ in terms of the difference operators defined above.

Theorem 2.2 (Hoeffding decompositions)**.** *Let $f : E^n \to \mathbb{R}$ be such that $E[f(X)^2] < \infty$. One has the following representation for $f(X)$:*

$$f(X) = \mathbf{E}[f(X)] + \sum_{k=1}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (-1)^k \mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_k} f(X', X) | X\right]. \qquad (2.4)$$

Formula (2.4) *coincides with the Hoeffding decomposition* (2.1) *of $f(X)$: in particular, one has that, for any choice of $i_1, ..., i_k$, $\mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_k} f(X', X) | X\right] = \varphi_{i_1,...,i_k}(X_{i_1}, ..., X_{i_k})$, and consequently*

$$\mathbf{E}\Big\{\mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_k} f(X', X) | X\right] \times \mathbf{E}\left[\Delta_{j_1} \cdots \Delta_{j_l} f(X', X) | X\right]\Big\} = 0, \qquad (2.5)$$

whenever $\{i_1, ..., i_k\} \neq \{j_1, ..., j_l\}$.

Proof. By Lemma 2.1,

$$f(X) = f(X') + \sum_{k=1}^{n} \sum_{1 \leqslant i_1 < \cdots i_k \leqslant n} (-1)^k \Delta_{i_1} \cdots \Delta_{i_k} f(X', X),$$

and (2.4) follows at once by taking conditional expectations with respect to X on both sides. To prove (2.5), it suffices to show the following stronger result: for every $1 \leqslant i_1 < \ldots < i_k \leqslant n$ (all k indices different),

$$\mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_k} f(X', X) | X_{i_1}, \ldots, X_{i_{k-1}}\right] = 0.$$

This is a consequence of the following fact: the random variable $\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X)$ is a function of $X_{i_1}, \ldots, X_{i_{k-1}}$ and of X'. By independence, it follows that

$$\mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X) | X_{i_1}, \ldots, X_{i_{k-1}}\right] = \mathbf{E}\left[(\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X))_{i_k} | X_{i_1}, \ldots, X_{i_{k-1}}\right]$$

where the random variable $(\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X))_{i_k}$ has been obtained from $\Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X)$ by replacing X'_{i_k} with X_{i_k}. Since (as already observed)

$$\Delta_{i_k} \Delta_{i_1} \cdots \Delta_{i_{k-1}} f(X', X) = \Delta_{i_1} \cdots \Delta_{i_k} f(X', X),$$

we deduce immediately the desired conclusion. □

The next statement is a direct consequence of (2.4)–(2.5).

Corollary 2.3. *Let $f(X)$ be as in the statement of Theorem 2.2. Then, the variance of $f(X)$ can be expanded as follows:*

$$\mathbf{Var}(f(X)) = \sum_{k=1}^{n} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{E}\left[\left(\mathbf{E}\left[\Delta_{i_1} \cdots \Delta_{i_k} f(X', X) | X\right]\right)^2\right]. \qquad (2.6)$$

As a first application of (2.6), we present a useful lower bound for variances.

Corollary 2.4. *Let $f(X)$ be as in the statement of Theorem 2.2. Then, one has the lower bound*

$$\mathbf{Var}(f(X)) \geq \sum_{i=1}^{n} \mathbf{E}\left[\left(\mathbf{E}\left[\Delta_i f(X', X)|X\right]\right)^2\right]$$

In particular, if $X = (X_1, \ldots, X_n)$ is a collection of n i.i.d. random variables with common distribution equal to μ, and $f : E^n \to \mathbb{R}$ is a symmetric mapping such that $E[f(X)^2] < \infty$, then

$$\mathbf{Var}(f(X)) \geq n \int_E \left(\mathbf{E}[f(X) - f(x, X_2, \ldots, X_n)]\right)^2 \, \mu(dx).$$

Remark 2.5. The estimates in Corollary 2.4 should be compared with the classical *Efron-Stein inequality* (see e.g. [1, Chapter 3]), stating that

$$\mathbf{Var}(f(X)) \leqslant \frac{1}{2} \sum_{i=1}^{n} \mathbf{E}\left[\Delta_i f(X, X')^2\right],$$

which, in the case where the X_i are i.i.d. and f is symmetric, becomes

$$\mathbf{Var}(f(X)) \leqslant \frac{n}{2} \int_E \mathbf{E}[(f(X) - f(x, X_2, \ldots, X_n))^2] \, \mu(dx).$$

For instance, if $f(X) = X_1 + \cdots + X_n$ is a sum of real-valued independent and square-integrable random variables, then the Efron-Stein upper bounds coincides with the lower bound in Corollary 2.4, that is:

$$\sum_{i=1}^{n} \mathbf{E}\left[\left(\mathbf{E}\left[\Delta_i f(X', X)|X\right]\right)^2\right] = \frac{1}{2} \sum_{i=1}^{n} \mathbf{E}\left[\Delta_i f(X, X')^2\right] = \sum_{i=1}^{n} \mathbf{Var}(X_i).$$

Heuristically, in the general case where the X_i are i.i.d. and f is symmetric, it seems that, in order for the Efron-Stein upper bound and the lower bound of Corollary 2.4 to have the same magnitude, it is necessary that the functional $f(X)$ is not *homogeneous*, meaning that the law of $f(X) - f(x, X_2, \ldots, X_n)$ depends on x. Examples of such a behaviour will be described in Section 6.1, where we will deal with Voronoi approximations.

2.3 Another subset-based interpolation

Let $n \geq 1$, let $f : E^n \to \mathbb{R}$, and let $y, y' \in E^n$. In [4], the following formula is pointed out:

$$f(y) - f(y') = \sum_{A \subsetneq [n]} \frac{1}{\binom{n}{|A|}(n - |A|)} \sum_{j \notin A} \Delta_j f(y^A, y'), \tag{2.7}$$

where the vector y^A has been obtained from y by replacing y_i with y_i' whenever $i \in A$, in such way that, with our notation, $\Delta_j f(y^A, y') = f(y^A) - f(y^{A \cup \{j\}}) = f^A(y, y') - f^{A \cup \{j\}}(y, y')$.

Now consider a vector $X = (X_1, ..., X_n)$, with independent components and with values in E^n, and let X' be an independent copy of X. For every $A \subseteq [n]$, we define $X^A = (X_1^A, ..., X_n^A)$ according to the above convention, that is:

$$X_i^A = \begin{cases} X_i \text{ if } i \notin A \\ X_i' \text{ otherwise.} \end{cases}$$

The following statement is a direct consequence of (2.7).

Proposition 2.6 (See [4], Lemma 2.3)**.** *For every $f, g : A^n \to \mathbb{R}$ such that $E[f(X)^2], E[g(X)^2] < \infty$,*

$$\mathbf{Cov}(f(X), g(X)) = \frac{1}{2} \sum_{A \subsetneq [n]} \frac{1}{\binom{n}{|A|}(n - |A|)} \sum_{j \notin A} \mathbf{E}[\Delta_j g(X, X')\Delta_j f(X^A, X')]. \tag{2.8}$$

To simplify the notation, we shall sometimes write

$$\frac{1}{\binom{n}{|A|}(n - |A|)} := \kappa_{n,A}.$$

Observe that, for every j, $\sum_{A \subsetneq [n]: j \notin A} \kappa_{n,A} = 1$

Remark 2.7. As demonstrated in [5, Lemmas 7.8-7.10], the identity (2.8) can also be used to deduce effective lower bounds on variances. Such lower bounds seem to have a different nature from the ones that can be proved by means of Hoeffding decompositions.

3 Stein's method and a new approximate Taylor expansion

Let U and V be two real-valued random variables. The *Kolmogorov distance* between the distributions of U and V is given by

$$d_K(U, V) = \sup_{t \in \mathbb{R}} |\mathbf{P}(U \leqslant t) - \mathbf{P}(V \leqslant t)|.$$

As anticipated in the Introduction, our aim in this paper is to provide upper bounds for quantities of the type $d_K(W, N)$, where $W = f(X)$ and N is a standard Gaussian random variable, that are based on the use of Stein's method. The following statement gathers together some classical facts concerning Stein's equations and their solutions (see Points (a)–(e) below), together with a new important approximate Taylor expansion for solutions of Stein's equations, that we partially extrapolated from reference [7] (see Point (f) below), generalising previous findings from [26]; see also [2, Theorem 2].

Proposition 3.1. *Let $N \sim \mathcal{N}(0, 1)$ be a centred Gaussian random variable with variance 1 and, for every $t \in \mathbb{R}$, consider the Stein's equation*

$$g'(w) - wg(w) = \mathbf{1}_{w \leqslant t} - \mathbf{P}(N \leqslant t), \tag{3.1}$$

where $w \in \mathbb{R}$. Then, for every real t, there exists a function $g_t : \mathbb{R} \to \mathbb{R} : w \mapsto g_t(w)$ with the following properties:

(a) *g_t is continuous at every point $w \in \mathbb{R}$, and infinitely differentiable at every $w \neq t$;*

(b) *g_t satisfies the relation (3.1), for every $w \neq t$;*

(c) *$0 < g_t \leqslant c := \frac{\sqrt{2\pi}}{4}$;*

(d) *for every* $u, v, w \in \mathbb{R}$,

$$|(w+u)g_t(w+u) - (w+v)g_t(w+v)| \leqslant \left(|w| + \frac{\sqrt{2\pi}}{4}\right)(|u|+|v|)\,; \tag{3.2}$$

(e) *adopting the convention*

$$g_t'(t) := tg_t(t) + 1 - \mathbf{P}(N \leqslant t), \tag{3.3}$$

one has that $|g_t'(w)| \leqslant 1$, *for every real* w.

(f) *using again the convention* (3.3), *for all* $w, h \in \mathbb{R}$ *one has that*

$$|g_t(w+h) - g_t(w) - g_t'(w)h| \leqslant \frac{|h|^2}{2}\left(|w| + \frac{\sqrt{2\pi}}{4}\right) + |h|(\mathbf{1}_{[w,w+h)}(t) + \mathbf{1}_{[w+h,w)}(t)) \tag{3.4}$$

$$= \frac{|h|^2}{2}\left(|w| + \frac{\sqrt{2\pi}}{4}\right) + h\left(\mathbf{1}_{[w,w+h)}(t) - \mathbf{1}_{[w+h,w)}(t)\right). \tag{3.5}$$

Proof. The proofs of Points (a)–(e) are classical, and can be found e.g. in [17, Lemma 2.3]. We will prove (f) by following the same line of reasoning adopted in [7, Proof of Theorem 3.1]. Fix $t \in \mathbb{R}$, recall the convention (3.3) and observe that, for every $w, h \in \mathbb{R}$, we can write

$$g_t(w+h) - g_t(w) - hg_t'(w) = \int_0^h \left(g_t'(w+u) - g'(w)\right) du.$$

Since g_t solves the Stein's equation (3.1) for every real w, we have that, for all $w, h \in \mathbb{R}$,

$$g_t(w+h) - g_t(w) - hg_t'(w)$$
$$= \int_0^h ((w+u)g_t(w+u) - wg_t(w))\, du + \int_0^h \left(\mathbf{1}_{\{w+u\leqslant t\}} - \mathbf{1}_{\{w\leqslant t\}}\right) du := I_1 + I_2.$$

It follows that, by the triangle inequality,

$$\left|g_t(w+h) - g_t(w) - hg_t'(x)\right| \leqslant |I_1| + |I_2|. \tag{3.6}$$

Using (3.2), we have

$$|I_1| \leqslant \int_0^h \left(|w| + \frac{\sqrt{2\pi}}{4}\right)|u|du = \frac{h^2}{2}\left(|w| + \frac{\sqrt{2\pi}}{4}\right). \tag{3.7}$$

Furthermore, observe that

$$\begin{aligned}
|I_2| &= \mathbf{1}_{\{h<0\}}\left|\int_0^h \left(\mathbf{1}_{\{w+u\leqslant t\}} - \mathbf{1}_{\{w\leqslant t\}}\right) du\right| + \mathbf{1}_{\{h\geq 0\}}\left|\int_0^h \left(\mathbf{1}_{\{w+u\leqslant t\}} - \mathbf{1}_{\{w\leqslant t\}}\right) du\right| \\
&= \mathbf{1}_{\{h<0\}}\left|-\int_h^0 \mathbf{1}_{\{w+u\leqslant t<w\}}du\right| + \mathbf{1}_{\{h\geq 0\}}\left|-\int_0^h \mathbf{1}_{\{w\leqslant t<w+u\}}du\right| \\
&= \mathbf{1}_{\{h<0\}}\int_h^0 \mathbf{1}_{\{w+u\leqslant t<w\}}du + \mathbf{1}_{\{h\geq 0\}}\int_0^h \mathbf{1}_{\{w\leqslant t<w+u\}}du.
\end{aligned}$$

Bounding u by h in both integrals provides the following upper bound:

$$\begin{aligned} |I_2| &\leqslant \mathbf{1}_{\{h<0\}}(-h)\mathbf{1}_{[w+h,w)}(t) + \mathbf{1}_{\{h\geq 0\}}h\mathbf{1}_{[w,w+h)}(t) \\ &\leqslant h\left(\mathbf{1}_{[w,w+h)}(t) - \mathbf{1}_{[w+h,w)}(t)\right) = |h|\left(\mathbf{1}_{[w,w+h)}(t) + \mathbf{1}_{[w+h,w)}(t)\right). \end{aligned} \tag{3.8}$$

Applying the estimates (3.7) and (3.8) to (3.6) concludes the proof. □

An immediate consequence of Proposition 3.1 is that for $N \sim \mathcal{N}(0,1)$ and for every real-valued random variable W, one has that

$$d_K(W,N) = \sup_{t\in\mathbb{R}} |\mathbf{E}g_t'(W) - Wg_t(W)| \tag{3.9}$$

(observe in particular that convention (3.3) defines unambiguously the quantity $g_t'(x)$ for every $t, x \in \mathbb{R}$) .

4 New Berry-Esseen bounds in the Kolmogorov distance

Let $n \geq 1$ be an integer, and consider a a vector $X = (X_1, ..., X_n)$ of independent random variables with values in the Polish space E. Let $X' = (X_1', \ldots, X')$ be an independent copy of X. Consider a function $f : E^n \to \mathbb{R}$ such that $W := f(X)$ is a centred and square-integrable random variable. We shall adopt the same notation introduced in Sections 2.1, 2.2, 2.3 and 3. For every $A \subsetneq [n]$, we write

$$T_A = \sum_{j\notin A} \Delta_j f(X,X')\Delta_j f(X^A,X')$$
$$T_A' = \sum_{j\notin A} \Delta_j f(X,X')|\Delta_j f(X^A,X')|$$

and

$$T = \frac{1}{2}\sum_{A\subsetneq[n]} \kappa_{n,A}T_A,$$
$$T' = \frac{1}{2}\sum_{A\subsetneq[n]} \kappa_{n,A}T_A'.$$

Observe that each T_A' is a sum of symmetric random variables in such way that $0 = \mathbf{E}[T'] = \mathbf{E}[T_A']$, $A \subsetneq [n]$.

Remark 4.1. An immediate application of (2.8) implies that $\mathbf{Var}(f(X)) = \mathbf{E}[T]$. We stress that the random variables T_A and T already appear in [4], in the context of normal approximations in the Wasserstein distance. Our use of the class of random objects $\{T', T_A' : A \subsetneq [n]\}$ for deducing bounds in the Kolmogorov distance is new.

The next statement is the main abstract finding of the paper.

Theorem 4.2. *Let the assumptions and notation of the present section prevail, let $N \sim \mathcal{N}(0,1)$, and assume that $\mathbf{E}W = 0$ and $\mathbf{E}W^2 = \sigma^2 \in (0,\infty)$. Then,*

$$\begin{aligned} d_K(\sigma^{-1}W, N) &\leqslant \frac{1}{\sigma^2}\sqrt{\mathbf{Var}(\mathbf{E}(T|X))} + \frac{1}{\sigma^2}\sqrt{\mathbf{Var}(\mathbf{E}\left(T'|X\right))} \\ &\quad + \frac{1}{4\sigma^4}\mathbf{E}\sum_{j,A,j\notin A}\kappa_{n,A}|f(X)|\left|\Delta_j f(X,X')^2\Delta_j f(X^A,X')\right| \\ &\quad + \frac{\sqrt{2\pi}}{16\sigma^3}\sum_{j=1}^n \mathbf{E}|\Delta_j f(X,X')|^3 \end{aligned} \tag{4.1}$$

$$\begin{aligned} &\leqslant \frac{1}{\sigma^2}\sqrt{\mathbf{Var}(\mathbf{E}(T|X))} + \frac{1}{\sigma^2}\sqrt{\mathbf{Var}(\mathbf{E}\left(T'|X\right))} \\ &\quad + \frac{1}{4\sigma^3}\sum_{j=1}^n\sqrt{\mathbf{E}|\Delta_j f(X,X')|^6} + \frac{\sqrt{2\pi}}{16\sigma^3}\sum_{j=1}^n \mathbf{E}|\Delta_j f(X,X')|^3. \end{aligned} \tag{4.2}$$

Proof. By homogeneity, we can assume that $\sigma = 1$, without loss of generality. By virtue of (3.9), the Kolmogorov distance between W and N is the supremum over $t \in [0,1]$ of

$$|\mathbf{E}g_t'(W) - Wg_t(W)| \leqslant \mathbf{E}|g_t'(W) - g_t'(W)T| + |\mathbf{E}(g_t(W)W - g_t'(W)T)|, \tag{4.3}$$

where the derivative $g_t'(w)$ is defined for every real w, thanks to the convention (3.3). Since W is $\sigma(X)$-measurable, $|g_t'| \leqslant 1$ and $\mathbf{E}T = \mathbf{E}W^2 = 1$, one infers that

$$\mathbf{E}|g_t'(W) - g_t'(W)T| \leqslant \mathbf{E}[|g_t'(W) \times \mathbf{E}[T-1 \mid X]|] \leqslant \mathbf{E}|\mathbf{E}[T-1 \mid X]| \leqslant \sqrt{\mathbf{Var}(\mathbf{E}(T|X))}.$$

Our aim is now to show that the quantity $|\mathbf{E}(g_t(W)W - g_t'(W)T)|$ is bounded by the last three summands on the right-hand side of (4.1) (with $\sigma = 1$). Reasoning as in [4], the relation (2.8) applied to $\mathbf{E}g_t(W)W$ and the definition of T yield

$$\begin{aligned} |\mathbf{E}g_t(W)W - g_t'(W)T| &= \left|\frac{1}{2}\sum_{A\subsetneq[n]}\kappa_{n,A}\sum_{j\notin A}\mathbf{E}(R_{A,j} - \tilde{R}_{A,j})\right| \\ &\leqslant \frac{1}{2}\sum_{A\subsetneq[n]}\kappa_{n,A}\sum_{j\notin A}\mathbf{E}|R_{A,j} - \tilde{R}_{A,j}|, \end{aligned}$$

with

$$\begin{aligned} R_{A,j} &= \Delta_j((g_t\circ f)(X))\Delta_j f(X^A), \\ \tilde{R}_{A,j} &= g_t'(f(X))\Delta_j f(X)\Delta_j f(X^A), \end{aligned}$$

where, here and for the rest of the proof, we use the simplified notation $\Delta_j f(X^A) = \Delta_j f(X^A, X')$, $\Delta_j f(X) = \Delta_j f(X,X')$, and so on. We have

$$\mathbf{E}|R_{A,j} - \tilde{R}_{A,j}| = \mathbf{E}\big[|g_t(f(X) - \Delta_j f(X)) - g_t(f(X)) - g_t'(f(X))(-\Delta_j f(X))| \times |\Delta_j f(X^A)|\big].$$

Now we use (3.5) with $w = f(X), h = -\Delta_j f(X)$, together with the fact that

$$h\left(\mathbf{1}_{[w,w+h)}(t) - \mathbf{1}_{[w+h,w)}(t)\right) = -h(\mathbf{1}_{\{w>t\}} - \mathbf{1}_{\{w+h>t\}})$$

to deduce that

$$|\mathbf{E}[g_t(W)W - g_t'(W)T]| \leqslant \frac{1}{2}\mathbf{E}\sum_{j,A,j\notin A}\kappa_{n,A}\Big\{\big(|f(X)| + \sqrt{2\pi}/4\big)\frac{|\Delta_j f(X)|^2|\Delta_j f(X^A)|}{2} + \Delta_j\left(\mathbf{1}_{f(X)>t}\right)\Delta_j f(X)\left|\Delta_j f(X^A)\right|\Big\}. \tag{4.4}$$

Using the independence of X and X', one proves immediately that, for $j \notin A$,

$$\mathbf{E}\Delta_j\left(\mathbf{1}_{f(X)>t}\right)\Delta_j f(X)\left|\Delta_j f(X^A)\right| = 2\mathbf{E}\mathbf{1}_{f(X)>t}\Delta_j f(X)\left|\Delta_j f(X^A)\right|,$$

from which it follows that the right-hand side of (4.4) is bounded by

$$\begin{aligned}&\frac{1}{4}\mathbf{E}\sum_{j,A,j\notin A}\kappa_{n,A}\left(|f(X)| + \frac{\sqrt{2\pi}}{4}\right)\left|\Delta_j f(X)^2\Delta_j f(X^A)\right| + \left|\mathbf{E}\left[\mathbf{1}_{f(X)>t}\times T'\right]\right|\\ &\leqslant \frac{1}{4}\mathbf{E}\sum_{j,A,j\notin A}\kappa_{n,A}\left(|f(X)| + \frac{\sqrt{2\pi}}{4}\right)\left|\Delta_j f(X)^2\Delta_j f(X^A)\right| + \sqrt{\mathbf{Var}(\mathbf{E}(T'\,|\,X))},\end{aligned}$$

where we have applied the Cauchy-Schwartz inequality, together with the fact that indicator functions are bounded by 1. The bound (4.1) is obtained by using the Hölder inequality in order to deduce that, for all j, A,

$$\mathbf{E}|\Delta_j f(X)|^2|\Delta_j f(X^A)| \leqslant \mathbf{E}|\Delta_j f(X)|^3,$$

and (4.2) follows by

$$\begin{aligned}\mathbf{E}|f(X)||\Delta_j f(X)|^2|\Delta_j f(X^A)| &\leqslant \sqrt{\mathbf{E}f(X)^2}\sqrt{\mathbf{E}\Delta_j f(X)^4\Delta_j f(X^A)^2}\\ &\leqslant \sqrt{(\mathbf{E}\Delta_j f(X)^{4(3/2)})^{2/3}(\mathbf{E}\Delta_j f(X^A))^{2(3)})^{1/3}} \leqslant (\mathbf{E}\Delta_j f(X)^6)^{1/2},\end{aligned}$$

where we have used the fact that X and X^A have the same distribution. □

Remark 4.3. Recall that the *Wasserstein distance* between the laws of two real-valued random variables U, V is defined as

$$d_W(U,V) := \sup_h |\mathbf{E}[h(U)] - \mathbf{E}[h(V)]|,$$

where the supremum runs over all 1-Lipschitz functions $h : \mathbb{R} \to \mathbb{R}$. In [4, Theorem 2.2], one can find the following bound: under the assumptions of Theorem 4.2,

$$d_W(W,N) \leqslant \frac{1}{\sigma^2}\sqrt{\mathbf{Var}(\mathbf{E}(T|X))} + \frac{1}{2\sigma^3}\sum_{j=1}^n \mathbf{E}|\Delta_j f(X,X')|^3. \tag{4.5}$$

Example 4.4. Consider a vector $X = (X_1, ..., X_n)$ of i.i.d. random variables with mean zero and variance 1, and assume that $\mathbf{E}|X_1|^4 < \infty$. Define $W = f(X) = n^{-1/2}(X_1 + \cdots + X_n)$. It is easily seen that, in this case, for every $j \notin A$, $\Delta_j f(X^A, X') = n^{-1/2}(X_j - X'_j)$, in such a way that

$$T = \frac{1}{2n}\sum_{j=1}^{n}(X_j - X'_j)^2 \quad \text{and} \quad T' = \frac{1}{2n}\sum_{j=1}^{n}\operatorname{sign}(X_j - X'_j)(X_j - X'_j)^2.$$

We also have, denoting $\hat{X}^j$ the vector X after removing X_j,

$$\begin{aligned}\mathbf{E}|f(X)\Delta_j f(X)^2\Delta_j f(X^A)| &\leqslant \mathbf{E}|f(X) - f(\hat{X}^j)||\Delta_j f(X)^2\Delta_j f(X^A)| + \mathbf{E}|f(\hat{X}^j)|\mathbf{E}|\Delta_j f(X)^2||\Delta_j f(X^A)| \\ &\leqslant \mathbf{E}n^{-2}|X_j||X_j - X'_j|^2|X_j - X'_j| + \mathbf{E}|f(\hat{X}^j)|\mathbf{E}n^{-3/2}|X_j - X'_j|^2|X_j - X'_j| \\ &\leqslant 8(n^{-2}\mathbf{E}X_j^4 + n^{-3/2}\mathbf{E}X_j^3).\end{aligned}$$

(note that the bound (4.2) can be used instead, whenever $\mathbf{E}X_1^6 < \infty$). An elementary application of (4.1) yields therefore that there exists a finite constant $C > 0$, independent of n, such that

$$d_K(W, N) \leqslant \frac{C}{\sqrt{n}},$$

providing a rate of convergence that is consistent with the usual Berry-Esseen estimates. One should notice that the estimate (4.5) yields the similar bound $d_W(W, N) \leqslant C/\sqrt{n}$.

5 Symmetric functions and geometric applications

In this section we adapt our results to random structures with local dependence, in a spirit close to [4, Section 2.3] – see Remark 5.4 below. Our principal focus will be on measurable and symmetric real-valued mappings f on E^n: we recall that $f : E^n \to \mathbb{R}$ is said to be *symmetric* if

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

for any permutation σ of $[n]$ and vector $x \in E^n$.

In the following, X and X' denote two independent sets of n i.i.d. random variables with common generic distribution μ. We will use the following short-hand notation: for any random vector Z of dimension n, and for every $1 \leqslant i \neq j \leqslant n$,

$$\Delta_i f(Z) := \Delta_i f(Z, X'),\ \Delta_{i,j} f(Z) := \Delta_i\Delta_j f(Z, X'),$$

where the notation is the same as in Section 2.1; we also adopt the additional convention that $\Delta_{i,i} = \Delta_i$. Now let $\tilde{X}$ be a further independent copy of X. We shall use the following terminology: a vector $Z = (Z_1, ..., Z_n)$ is a *recombination* of $\{X, X', \tilde{X}\}$, if $Z_i \in \{X_i, X'_i, \tilde{X}_i\}$ for every $1 \leqslant i \leqslant n$.

The next statement provides a bound for the normal approximation of geometric functionals that is amenable to geometric analysis, and can be heuristically regarded as the binomial counterpart to the second order Poincaré inequalities on the Poisson space (in the Kolmogorov distance), proved in [19].

Theorem 5.1. *Let $f : E^n \to \mathbb{R}$ be a symmetric measurable functional such that $W = f(X)$ is centred, and $\sigma^2 = \mathbf{Var}(W) < \infty$. Let N be a centred Gaussian random variable with variance 1. Define*

$$\begin{aligned}
B_n(f) &:= \sup_{(Y,Z,Z')} \mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0\}}\Delta_1 f(Z)^2\Delta_2 f(Z')^2\right],\\
B'_n(f) &:= \sup_{(Y,Y',Z,Z')} \mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0,\Delta_{1,3}f(Y')\neq 0\}}\Delta_2 f(Z)^2\Delta_3 f(Z')^2\right],
\end{aligned}$$

where the suprema run over all vectors Y, Y', Z, Z' that are recombinations of $\{X, X', \tilde{X}\}$. Then,

$$\begin{aligned}
d_K(\sigma^{-1}W, N) \leqslant \Bigg[&\frac{4\sqrt{2}n^{1/2}}{\sigma^2}\left(\sqrt{nB_n(f)} + \sqrt{n^2 B'_n(f)} + \sqrt{\mathbf{E}\Delta_1 f(X)^4}\right) \qquad (5.1)\\
&+\frac{n}{4\sigma^4}\sup_{A\subseteq[n]}\mathbf{E}|f(X)\Delta_1 f(X^A)^3| + \left(\frac{\sqrt{2\pi}}{16\sigma^3}n\mathbf{E}|\Delta_1 f(X)^3|\right)\Bigg].
\end{aligned}$$

Remark 5.2. We shall often use the following bounds, following at once from the Cauchy-Schwartz inequality,

$$\begin{aligned}
B'_n(f) &\leqslant \sup_{(Y,Y',Z,Z')}\sqrt{\mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0,\Delta_{1,3}f(Y')\neq 0\}}\Delta_2 f(Z)^4\right]\mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0,\Delta_{1,3}f(Y')\neq 0\}}\Delta_3 f(Z')^4\right]}\\
&\leqslant \sup_{(Y,Y',Z)}\mathbf{E}\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0,\Delta_{1,3}f(Y')\neq 0\}}\Delta_2 f(Z)^4 \qquad (5.2)
\end{aligned}$$

and

$$B_n(f) \leqslant \sup_{(Y,Z)}\mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0\}}\Delta_1 f(Z)^4\right]. \qquad (5.3)$$

In the framework of the applications developed in this paper, such estimates simplify some computations and do not worsen the associated rates of convergence.

In the applications developed below, we will often consider functions f that are obtained as restrictions to E^n of general real-valued mappings on the set $\cup_{n\geq 1}E^n$, corresponding to the class of all finite ordered point configurations (with possible repetitions). Now fix $f : \cup_{n\geq 1}E^n \to \mathbb{R}$ and, for every $n \geq 1$ and every $x = (x_1, ..., x_n) \in E^n$, introduce the notation $\hat{x}^i$ to indicate the element of E^{n-1} obtained by deleting the ith coordinate of x, that is: $\hat{x}^i = (x_1, ..., x_{i-1}, x_i, ..., x_n)$. Analogously, write $\hat{x}^{ij} \in E^{n-2}$ to denote the vector obtained from x by removing its i-th and j-th coordinates. We write

$$\begin{aligned}
D_i f(x) &= f(x) - f(\hat{x}^i),\\
D_{i,j} f(X) &= f(x) - f(\hat{x}^i) - f(\hat{x}^j) + f(\hat{x}^{ij}) = D_{j,i}f(x).
\end{aligned}$$

Proposition 5.3. *Let f be a functional defined on $\cup_{k\leqslant n}E^k$ such that its restriction to E^n satisfies the hypotheses of Theorem 5.1. Then we have*

$$B'_n(f) \leqslant 2^8 \sup_{(Y,Y',Z,Z')} \mathbf{E}\left[\mathbf{1}_{\{D_{1,2}f(Y)\neq 0\}}\mathbf{1}_{\{D_{1,3}f(Y')\neq 0\}}D_2f(Z)^2D_2f(Z')^2\right]$$
$$B_n(f) \leqslant 2^6 \sup_{(Y,Z,Z')} \mathbf{E}\left[\mathbf{1}_{\{D_{1,2}f(Y)\neq 0\}}D_1f(Z)^2D_2f(Z')^2\right].$$

Proof. First observe that

$$|\Delta_j f(X)| \leqslant |D_jf(X)| + |D_jf(X^j)| \tag{5.4}$$
$$\Delta_{i,j}f(X) = D_{i,j}f(X) - D_{i,j}f(X^i) - D_{i,j}f(X^j) + D_{i,j}f(X^{\{i,j\}}). \tag{5.5}$$

Let Y, Y', Z, Z' be recombinations of $\{X, X', \tilde{X}\}$. Using the bounds above, there are recombinations $Y^{(i)}, Y'^{,(i)}, i = 1, \dots, 4$ and $Z^{(l)}, Z'^{,(l)}, l = 1, 2$, such that

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{1}_{\{\Delta_{1,2}f(Y)\neq 0,\Delta_{1,3}f(Y')\neq 0\}}\Delta_2f(Z)^2\Delta_3f(Z')^2\right]\\
&\leqslant \mathbf{E}\left[\sum_{i=1}^{4}\mathbf{1}_{\{D_{1,2}f(Y^{(i)})\neq 0\}}\sum_{j=1}^{4}\mathbf{1}_{\{D_{1,3}f(Y'^{,(j)})\neq 0\}}\sum_{l,m=1}^{2}4D_2f(Z^{(l)})^2D_3(Z'^{,(m)})^2\right]\\
&\leqslant 256 \sup_{(Y,Y',Z,Z')}\mathbf{E}\left[\mathbf{1}_{\{D_{1,2}f(Y)\neq 0\}}\mathbf{1}_{\{D_{1,3}f(Y')\neq 0\}}D_2f(Z)^2D_3f(Z')^2\right],
\end{aligned}$$

which gives the bound on $B'_n(f)$. The bound on $B_n(f)$ is obtained analogously. □

Remark 5.4. Our framework is more restrictive than that of [4, Theorem 2.5], where it is not assumed that f is symmetric, but rather that its dependency graph is symmetric, meaning that the relation $\Delta_{i,j}f(X) = 0$ is equivalent to $\Delta_{\sigma(i),\sigma(j)}f(X^\sigma) = 0$ for any $i \neq j$ and every permutation σ of $\{1, ..., n\}$, where $X_i^\sigma := X_{\sigma(i)}$. One should notice that this subtlety is not exploited in most applications of [4] – see e.g. [21]. Under our symmetry assumption, a bound analogous to the main estimate in [4, Theorem 2.5] can be retrieved from (5.1) by using the bounds

$$\begin{aligned}
&\sqrt{\mathbf{E}\Delta_jf(X)^4} + \sqrt{nB_n(f)} + \sqrt{n^2B'_n(f)}\\
&\leqslant 3\sqrt{\mathbf{E}\Delta_jf(X)^4 + nB_n(f) + n^2B'_n(f)}\\
&\leqslant 3\sqrt{8\sum_{j,k=1}^{n}\sup_{(Y,Y',Z,Z')}\mathbf{E}1_{\{\Delta_{1,j}f(Y)\neq 0\}}1_{\{\Delta_{1,k}f(Y')\neq 0\}}\Delta_jf(Z)^2\Delta_kf(Z')^2}\\
&\leqslant 6\sqrt{2}\sqrt{\sum_{j,k=1}^{n}\sup_{(Y,Y',Z)} n^{-2}\mathbf{E}(\sup_{j=1}^{n}|\Delta_jf(Z)|)^4\delta_1(Y)\delta_1(Y')}\\
&\leqslant 6\sqrt{2}(\mathbf{E}M(X)^8)^{1/4}(\mathbf{E}\delta_1(X)^4)^{1/4}
\end{aligned}$$

where $M(X) = \sup_i|\Delta_if(X)|$ and $\delta_1(X) = \#\{j : \Delta_{1,j}f(X) \neq 0\}$. One should notice that the additional term involving quantities of the type $\mathbf{E}|f(X)\Delta_1f(X)^2\Delta_1f(X^A)|$ appears in our bounds because we are dealing with the Kolmogorov distance: in general, we shall control this term by using the rough estimate $\mathbf{E}|f(X)\Delta_1f(X)^2\Delta_1f(X^A)| \leqslant \sigma\sqrt{\mathbf{E}\Delta_jf(X)^6}$, that one can e.g. deduce by applying twice the Cauchy-Schwartz inequality – see Section 6 for more details.

Proof of Theorem 5.1. Assume without loss of generality that $\sigma = 1$. Our estimate follows by appropriately bounding each of the four summands appearing on the right-hand side of (4.1). We have for $A \subseteq [n], 1 \leqslant j \leqslant n$, by Hölder inequality,

$$\begin{aligned}\mathbf{E}|f(X)\Delta_j f(X)^2\Delta_j f(X^A)| &= \mathbf{E}|f(X)^{2/3}\Delta_j f(X)^2||\Delta_j f(X)^{1/3}\Delta_j f(X^A)| \\ &\leqslant \left(\mathbf{E}|f(X)\Delta_j f(X)^3|\right)^{2/3}\left(\mathbf{E}|f(X)\Delta_j f(X^A)^3|\right)^{1/3} \\ &\leqslant \sup_{A\subseteq[n]} \mathbf{E}|f(X)\Delta_j f(X^A)^3|,\end{aligned}$$

because $\Delta_j f(X) = \Delta_j f(X^{\emptyset})$. The two last terms on the right-hand side of (4.1) are therefore bounded by the last two terms in (5.1), in view of the symmetry of f and of the relation $\sum_{A\subsetneq[n]:1\notin A}\kappa_{n,A} = 1$. To control the first two summands in (4.1), we first bound the square root of the variance of a random variable of the type $U := \frac{1}{2}\sum_{A\subsetneq[n]}\kappa_{n,A}U_A$, for a general family of square-integrable random variables $U_A(X,X'), A \subsetneq [n]$. Using e.g. [4, Lemma 4.4], we infer that

$$\sqrt{\mathbf{Var}\left(\mathbf{E}\left(U|X\right)\right)} \leqslant \frac{1}{2}\sum_{A\subsetneq[n]}\kappa_{n,A}\sqrt{\mathbf{Var}\mathbf{E}\left(U_A|X\right)} \leqslant \frac{1}{2}\sum_{A\subsetneq[n]}\kappa_{n,A}\sqrt{\mathbf{E}\left(\mathbf{Var}(U_A|X')\right)}. \tag{5.6}$$

This inequality will be used both for $U_A = T_A$ and $U_A = T'_A$. Let us now bound each summand separately. Fix $A \subseteq [n]$. Introduce the substitution operator based on $\tilde{X} = (\tilde{X}_i)_{1\leqslant i\leqslant n}$

$$\tilde{S}_i(X) = (X_1, \ldots, \tilde{X}_i, \ldots, X_n).$$

Recall that, by the Efron-Stein's inequality, for any square-integrable functional $Z(X_1, \ldots, X_n)$,

$$\mathbf{Var}(Z) \leqslant \frac{1}{2}\sum_{i=1}^n \mathbf{E}(\tilde{\Delta}_i Z(X))^2$$

where

$$(\tilde{\Delta}_i Z)(X) := Z(\tilde{S}_i(X)) - Z(X)$$

is clearly centred. Applying this to $Z(X) = U_A(X,X')$ for fixed X',

$$\mathbf{Var}(U_A|X') \leqslant \frac{1}{2}\sum_{i=1}^n \mathbf{E}\left[\left(\tilde{\Delta}_i U_A(X,X')\right)^2|X'\right].$$

From this relation, we therefore infer that

$$\sqrt{\mathbf{Var}\left(\mathbf{E}\left(U|X\right)\right)} \leqslant \frac{1}{\sqrt{8}}\sum_{A\subsetneq[n]}\kappa_{n,A}\sqrt{\sum_{i=1}^n \mathbf{E}\left(\tilde{\Delta}_i U_A\right)^2}.$$

Now recall that $U_A = T_A$ or $U_A = T'_A$, i.e. $U_A = \sum_{j\notin A}\Delta_j f(X)g(\Delta_j f(X^A))$, where either g is the identity or $g(\cdot) = |\cdot|$. Expanding the square yields

$$\sum_{i=1}^n \mathbf{E}\left(\tilde{\Delta}_i U_A\right)^2 = \sum_{i=1}^n\sum_{j,k\notin A}\mathbf{E}|\tilde{\Delta}_i(\Delta_j f(X)g(\Delta_j f(X^A)))||\tilde{\Delta}_i(\Delta_k f(X)g(\Delta_k f(X^A)))|. \tag{5.7}$$

Now fix $1 \leqslant i \leqslant n$, write $\tilde{X}^i = \tilde{S}_i(X)$ and observe that for $j \notin A$,

$$\tilde{\Delta}_i(\Delta_j f(X) g(\Delta_j f(X^A))) = \tilde{\Delta}_i(\Delta_j f(X)) g(\Delta_j f(X^A)) + \Delta_j f(\tilde{X}^i) \tilde{\Delta}_i(g(\Delta_j f(X^A))). \tag{5.8}$$

We note immediately that, in the case $i = j$, using $|\tilde{\Delta}_i g(V(X))| \leqslant |\tilde{\Delta}_i(V(X))|$ and $\tilde{\Delta}_i(\Delta_i(V(X))) = \tilde{\Delta}_i(V(X))$ for any random variable $V(X)$, the right-hand side of (5.8) is bounded by the simpler expression

$$|\tilde{\Delta}_i f(X) \Delta_i f(X^A)| + |\Delta_i f(\tilde{X}^i) \tilde{\Delta}_i f(X^A)| \leqslant \frac{1}{2} \left[\tilde{\Delta}_i f(X)^2 + \Delta_i f(X^A)^2 + \Delta_i f(\tilde{X}^i)^2 + \tilde{\Delta}_i f(X^A)^2 \right]. \tag{5.9}$$

Now let us examine each summand appearing in (5.7) separately. If $i \notin A$ and $i = j = k$, using (5.9), the summand is smaller than

$$\frac{1}{4} \mathbf{E} \left[\tilde{\Delta}_i f(X)^2 + \Delta_i f(X^A)^2 + \Delta_i f(\tilde{X}^i)^2 + \tilde{\Delta}_i f(X^A)^2 \right]^2 \leqslant 4 \mathbf{E} \Delta_1 f(X)^4.$$

In the case where i, j, k are pairwise distinct, introduce the vector $\bar{X}$ by

$$\begin{cases} \bar{X}_i & = \tilde{X}_i \\ \bar{X}_l & = X'_l \text{ if } l \neq i, \end{cases}$$

and, for $x \in E^n$ and some mapping ψ on E^n, define, for $1 \leqslant l \leqslant n$,

$$\bar{\Delta}_l \varphi(x) = \psi(x) - \psi(x_1, \ldots, x_{l-1}, X_l, x_{l+1}, \ldots, x_n).$$

Then, the corresponding summands are bounded by

$$4 \sup_{(Y, Y', Z, Z')} \mathbf{E} \left| \bar{\Delta}_i(\bar{\Delta}_j f(Y)) \bar{\Delta}_j f(Y') \bar{\Delta}_i(\bar{\Delta}_k f(Z)) \bar{\Delta}_k f(Z') \right|.$$

Using $\bar{X} \overset{(d)}{=} X'$ and the fact that if Y is a recombination, switching the roles of $\tilde{X}_i$ and X'_i in Y still yields a recombination of $\{X, X', \tilde{X}\}$, the previous expression is bounded by

$$\begin{aligned} &= 4 \sup_{(Y, Y', Z, Z')} \mathbf{E} \left| \Delta_i(\Delta_j f(Y)) \Delta_j f(Y') \Delta_i(\Delta_k f(Z)) \Delta_k f(Z') \right| \\ &\leqslant 4 \sup_{(Y, Y', Z, Z')} \mathbf{E} \mathbf{1}_{\{\Delta_{i,j} f(Y) \neq 0\}} (|\Delta_j f(Y)| + |\Delta_j f(Y^i)|) |\Delta_j f(Y')| \times \\ &\qquad \times \mathbf{1}_{\{\Delta_{i,k} f(Z) \neq 0\}} (|\Delta_k f(Z)| + |\Delta_k f(Z^i)|) |\Delta_k f(Z')| \\ &\leqslant 16 B'_n(f), \end{aligned}$$

where we have used Cauchy-Schwarz inequality. The case $i \neq j = k$ is treated with the same vector $\bar{X}$ and operators $\bar{\Delta}_l$. Using similar computations and Cauchy-Schwarz inequality, we have the upper bound

$$\begin{aligned}
&4 \sup_{(Y,Y',Z,Z')} \mathbf{E}\bar{\Delta}_i(\bar{\Delta}_j f(Y))\bar{\Delta}_j f(Y')\bar{\Delta}_i(\bar{\Delta}_j f(Z))\bar{\Delta}_j f(Z') \\
&\leqslant 4 \sup_{(Y,Y')} \left[\mathbf{E}\bar{\Delta}_i(\bar{\Delta}_j f(Y))^2\bar{\Delta}_j f(Y')^2\right] \\
&= 4 \sup_{(Y,Y')} \left[\mathbf{E}\Delta_j(\Delta_i f(Y))^2\Delta_j f(Y')^2\right] \\
&\leqslant 4 \sup_{(Y,Y')} \mathbf{E}\mathbf{1}_{\{\Delta_{i,j}f(Y)\neq 0\}}(|\Delta_i f(Y)| + |\Delta_i f(Y^j)|)^2\Delta_j f(Y')^2 \\
&\leqslant 16 \sup_{(Y,Y',Z)} \mathbf{E}\mathbf{1}_{\{\Delta_{i,j}f(Y)\neq 0\}}\Delta_i f(Z)^2\Delta_j f(Y')^2 \\
&\leqslant 16 B_n(f),
\end{aligned}$$

where the suprema run over recombinations Y, Y', Z, Z' of $\{X, X', \tilde{X}\}$. Finally, if $i = j \neq k$, the corresponding summands on the right-hand side of (5.7) are bounded by

$$\begin{aligned}
4 \sup_{(Y,Y',Z)} \mathbf{E}&\left|\bar{\Delta}_i f(Y)^2\bar{\Delta}_i(\bar{\Delta}_k f(Y'))\bar{\Delta}_k f(Z)\right| \\
&\leqslant 4 \sup_{(Y,Y',Z)} \mathbf{E}\mathbf{1}_{\{\Delta_{i,k}f(Y')\neq 0\}}(|\Delta_k f(Y)| + |\Delta_k f(Y^i)|)\Delta_i f(Y)^2|\Delta_k f(Z)| \\
&\leqslant 8B_n(f).
\end{aligned}$$

This yields

$$\begin{aligned}
\sum_{i=1}^n \mathbf{E}\left(\tilde{\Delta}_i U_A\right)^2 &\leqslant 16n \sum_{j,k\notin A} \left[\mathbf{1}_{\{j=k=1\}}\mathbf{E}\Delta_1 f(X)^4 + (\mathbf{1}_{\{k\neq j=1\}} + \mathbf{1}_{\{k=j\neq 1\}})B_n(f) + \mathbf{1}_{\{k\neq j\neq 1\}}B_n'(f)\right] \\
&\leqslant 16n\left(\mathbf{1}_{\{1\notin A\}}\Delta_1 f(X)^4 + 2(n-|A|)B_n(f) + (n-|A|)^2 B_n'(f)\right),
\end{aligned}$$

and using the inequality $\sqrt{x+y} \leqslant \sqrt{x} + \sqrt{y}$ $(x, y \geq 0)$ we deduce that

$$\sqrt{\sum_{i=1}^n \mathbf{E}\left(\tilde{\Delta}_i U_A\right)^2} \leqslant \sqrt{16n}\left(\mathbf{1}_{\{1\notin A\}}\sqrt{\mathbf{E}\Delta_1 f(X)^4} + \sqrt{2B_n(f)}\sqrt{n-|A|} + \sqrt{B_n'(f)}(n-|A|)\right).$$

Finally,

$$\begin{aligned}
&\sqrt{\mathbf{Var}(\mathbf{E}(U|X))} \leqslant \\
&\sqrt{8n}\left(\sqrt{\mathbf{E}\Delta_1 f(X)^4}\sum_{A\subsetneq[n]:1\notin A}\kappa_{n,A} + \sqrt{B_n(f)}\sum_{A\subsetneq[n]}\kappa_{n,A}\sqrt{n-|A|} + \sqrt{B_n'(f)}\sum_{A\subsetneq[n]}\kappa_{n,A}(n-|A|)\right)
\end{aligned}$$

and the result follows by evaluating the three sums over $A \subsetneq [n]$ in the last expression.

□

6 Applications

6.1 Set approximation with random tessellations

Let K be a compact subset of $\mathbb{R}^d$ with positive volume, and let $X = (X_i)$ be a locally finite collection of points. Assume the only information available about K is given by the values of the indicator function $1_{\{x \in K\}}, x \in X$. Then, the *Voronoi reconstruction*, or *Voronoi approximation*, of K based on X is defined as

$$K^X = \{y \in \mathbb{R}^d : \text{ the closest point from } y \text{ in } X \text{ lies in } K\}.$$

This chapter is devoted to the study of the error committed when one approximates the volume of $K \subseteq [0,1]^d$ with that of K^X, when X is a random input consisting in n i.i.d points in $[0,1]^d$.

The underlying structure in this approximation scheme is the Voronoi tessellation based on X. For $x \in [0,1]^d$, denote by $V(x; X)$ the Voronoi cell with nucleus x among X, i.e. the convex set formed by points $y \in [0,1]^d$ such that $\|y - x\| \leqslant \|y - x'\|$ for any point $x' \in (X, x)$, where in all this section $(X, x) := X \cup \{x\}$, and we extend the set notation $\in$ to ordered collections of points in an obvious way. The volume approximation described above is denoted

$$\varphi(X) = \text{Vol}(K^X) = \sum_i 1_{\{X_i \in K\}} \text{Vol}(V(X_i; X)).$$

Along the same lines, one can also approximate the perimeter of K via the relation $\varphi_{\text{Per}}(X) = \text{Vol}(K^X \Delta K)$ where Δ denotes the symmetric difference of sets.

This set approximation can serve in image reconstruction and estimation: it has first been introduced by Einmahl and Khmaladze [8] as a discriminating statistic in the two-sample problem. These authors proved a strong law of large numbers in dimension 1. Heveling and Reitzner [12] proved that if K is convex and compact and $X = X'$ is a homogeneous Poisson process with intensity n, $\mathbf{E}\varphi(X') = \text{Vol}(K)$, and $\mathbf{Var}(\varphi(X')) \leqslant cn^{-1-1/d}S(K)$ where c is an explicit constant and $S(K)$ is the surface area of K. They also established that $\mathbf{E}\varphi_{\text{Per}}(X') = c'n^{-1/d}S(K)(1 + O(n^{-1/d}))$ and $\mathbf{Var}(\varphi_{\text{Per}}(X')) \leqslant c'n^{-1-1/d}S(K)$. Reitzner, Spodarev and Zaporozhets [23] extended these results to sets with finite variational perimeter, and also gave upper bounds for $\mathbf{E}|\varphi(X')^q - \text{Vol}(K)^q|$ for $q \geq 1$. Schulte [27] proved a similar lower bound for the variance, i.e. $CS(K)n^{-1-1/d} \leqslant \mathbf{Var}(\varphi(X'))$ with K a convex body and C a universal constant, and the corresponding CLT

$$d_W\left(\frac{\varphi(X') - \mathbf{E}\varphi(X')}{\sqrt{\mathbf{Var}(\varphi(X'))}}, N\right) \to 0.$$

Yukich [31] then gave an upper bound on the speed of convergence in Kolmogorov distance.

For Binomial input, Penrose proved that for measurable K and X consisting in n iid variables with density $\kappa(x) > 0$ on $[0,1]^d$,

$$\mathbf{E}\varphi(X) \to \text{Vol}(K), \tag{6.1}$$

without assumption on K, not even the negligibility of its boundary. Yukich [31] managed to extend to a non-Poissonized setting the estimates on the variance magnitude as well as the central limit theorem for the Volume approximation. See also [3] for a result involving the Hausdorff distance.

In this section, we consider a binomial input $X = (X_1, \ldots, X_n)$, where the X_i are n iid variables uniformly distributed on $[0,1]^d$. We give asymptotic upper bounds for the moments of $\varphi(X) - \mathbf{E}\varphi(X)$, as well as a central limit theorem with rates of convergence in the Kolmogorov distance, that is new in the literature. Note that, in the words of Heveling and Reitzner [12],"the general problem whether K^X approximates K for complicated sets seems to be difficult", and many applications of set approximation are concerned with the detection or approximation of sets with an irregular boundary, see for instance [6] or the survey [16, Chap. 11]. Our results also hold for large classes of irregular sets, with a possibly fractal boundary. The regularity of the boundary of K will be assessed in terms of the following quantities. Call below *Lebesgue-boundary* of K, written ∂K, the class of points x such that for all $\varepsilon > 0$, $\text{Vol}(B(x,\varepsilon) \cap K) > 0$ and $\text{Vol}(B(x,\varepsilon) \cap K^c) > 0$. Let $\beta > 0$. Denote by $d(x, A)$ the Euclidean distance from a point $x \in \mathbb{R}^d$ to a subset $A \subseteq \mathbb{R}^d$. Define

$$\begin{aligned}
\partial K^r &= \{x : d(x, \partial K) \leqslant r\} \\
\partial K^r_+ &= K^c \cap \partial K^r \\
\gamma(K, r) &= \int_{\partial K^r_+} \left(\frac{\text{Vol}(B(x, \beta r) \cap K)}{r^d} \right)^2 dx.
\end{aligned}$$

K is said to satisfy the *weak rolling ball condition* if

$$\gamma(K) := \liminf_{r>0} \text{Vol}(\partial K^r)^{-1}(\gamma(K,r) + \gamma(K^c, r)) > 0. \tag{6.2}$$

This assumption somehow implies that either K or K^c occupies a constant positive proportion of space as one zooms in on a typical point close to ∂K, at least in a non-negligible region of $[0,1]^d$. It is related to a weak form of the *rolling ball condition* used in set estimation (see for instance condition (a) of Theorem 1 in [6], the definition of standard sets in [25], Remark 4 in [27], or the survey [16, Chap. 11] and references therein), where for each $x \in \partial K$ a ball of radius βr touching x should lie in $\partial(K^c)^r_+$ or ∂K^r_+. In our weaker form of the condition, the ball is somehow allowed to be deformed to fit in the parallel body. It certainly allows sets which boundary is smooth in a certain sense, and does not discard a priori fractal sets. It is proved in [18] that a class of fractal sets including for instance the 2-dimensional Von Koch flake and antiflake satisfy the condition, as well as the hypotheses of the following theorem with $\alpha = 2 - s$, $s = \log(4)/\log(3)$ being the fractal dimension of the boundary.

Theorem 6.1. *Let $K \subset [0,1]^d$ such that*

$$Vol(\partial K^r) \leqslant S_+(K) r^\alpha \tag{6.3}$$

for some $S_+(K), \alpha > 0$. Then for $n, q \geq 1$,

$$\mathbf{E}|\varphi(X) - \mathbf{E}\varphi(X)|^q \leqslant S_+(K) C_{d,q,\alpha} n^{-q/2 - \alpha/d}, \tag{6.4}$$

for some $C_{d,q,\alpha} > 0$ explicit in the proof. If furthermore K satisfies the weak rolling ball condition (6.2) and

$$Vol(\partial K^r) \geq S_-(K) r^\alpha \tag{6.5}$$

for some $S_-(K) > 0$, then for n sufficiently large

$$C_d^- S_-(K)\gamma(K) \leqslant \frac{\mathbf{Var}(\varphi(K,X))}{n^{-1-\alpha/d}} \leqslant C_d^+ S_+(K) C_{d,2,\alpha},$$

for some $C_d^-, C_d^+ > 0$, and for every $\varepsilon > 0$, there is $c_\varepsilon > 0$ not depending on n such that

$$d_K\left(\frac{\varphi(X) - \mathbf{E}\varphi(X)}{\sqrt{\mathbf{Var}(\varphi(X))}}, N\right) \leqslant c_\varepsilon n^{-1/2+\alpha/2d} \log(n)^{3+\alpha/d+\varepsilon},$$

for $n \geq 1$.

Remarks 1.

1. The previous theorem also applies to smooth sets. Blashke's theorem (see for instance [30, Theorem 1]), yields that any $\mathcal{C}^1$ manifold K with Lipschitz normal admits inside and outside rolling balls in the traditional sense, and satisfies in particular our weak rolling ball condition. Furthermore, such a set and its complement have positive reach, which proves by Steiner formula that the upper and lower bounds (6.3), (6.5) are satisfied, see the pioneering work of Federer [9]. The result might still hold if the boundary is only piecewise regular, see for instance Remark 4 in [27].
2. If (6.2) is not satisfied, we can still get a lower bound on the variance (and therefore a rate of convergence), but its magnitude will not match that of the upper bound, see Lemma 6.8. It might be difficult for such a set to get a clear estimate of the variance. See also the counterexample in [18].
3. The constant β in the rolling ball condition is left at our choice. The larger β, the easier it is for K to verify the condition.
4. Conditions (6.3) and (6.5) imply that K has Minkowski dimension equal to $d - \alpha$, and furthermore that K has lower and upper Minkowski content (see for instance [18]). Self similar sets satisfy these hypotheses, and are treated in [18], as well as some examples, such as the Von Koch flake, that also satisfies the weak rolling ball condition. We provide as well the example of a set K with lower and upper Minkowski content for $\alpha = 1/2$ that does not satisfy the rolling ball condition. Simulations indicate that for this example the variance is indeed negligible with respect to $n^{-1-\alpha/d}$, but it is still possible to get a rate of convergence for Kolmogorov distance to the normal law.
5. The uniformity of the distribution of the X_i's does not have a crucial importance, apart from easing certain geometric estimates. The results should hold, up to constants, if the common distribution of the X_i's is only assumed to have a density bounded from below by some constant $\kappa > 0$ on the domain ∂K^r, for some $r > 0$.
6. The Berry-Essen bounds is derived from (5.1). It turns out that each of the terms on the right hand side of (5.1) contributes with the same power of n, heuristically indicating that this power is likely to be optimal.

The proof of the theorem is decomposed into several independent results. The variance lower bound is established in the specific framework of Voronoi volume approximation. The Kolmogorov distance and moments upper bounds are potentially valid in a more general framework.

Theorem 6.2. *Define $\sigma^2 = \mathbf{Var}(\varphi(X))$. Assume that $Vol(\partial K^r) \leqslant S_+(K) r^\alpha$ for some $S_+(K), \alpha > 0$. Then (6.4) holds, and for every $\varepsilon > 0$ there is a constant c_ε not depending on n such that for $n \geq 1$,*

$$d_K\left(\sigma^{-1}(\varphi(X) - \mathbf{E}\varphi(X)), N\right) \leqslant c_\varepsilon \left(\sigma^{-2} n^{-3/2-\alpha/2d} + \sigma^{-3} n^{-2-\alpha/d} + \sigma^{-4} n^{-3-\alpha/d}\right) \log(n)^{3+\alpha/2d+\varepsilon} \tag{6.6}$$

where N is a standard Gaussian variable.

Say that two points $x, y \in [0,1]^d$ are *Voronoi neighbours* among a point set X if $V(x;X) \cap V(y;X) \neq \emptyset$. More generally, denote $d_V(x,y;X)$ the Voronoi distance between x and y, i.e. the minimal integer $k \geq 1$ such that we can form a path $x_0 = x; x_1 \in X, \ldots, x_{k-1} \in X, x_k = y$ where x_i and x_{i+1} are Voronoi neighbours. Denote $v(x,y;X) = \mathrm{Vol}(V(x,(X,y)) \cap V(y,X))$ the volume that the cell $V(y,X)$ loses when x is added to X. We have the explicit expression, for $x \notin X$,

$$\varphi(X,x) - \varphi(X) = 1_{\{x \in K\}} \sum_{y \in X \cap K^c} v(x,y;X) - 1_{\{x \in K^c\}} \sum_{y \in X \cap K} v(x,y;X). \tag{6.7}$$

Since $v(x,y;X) = 0$ if x and y are not Voronoi neighbours in (X,x,y), the concatenation of X with x and y, the following properties hold.

Proposition 6.3. *Let $X = (X_i)_{1 \leqslant i \leqslant n}$ be a finite collection of points.*

(i) *For $1 \leqslant i \leqslant n$ such that $X_i \in K$ (resp. K^c), if every Voronoi neighbour of X_i among X is also in K (resp. K^c), then $D_i\varphi(X) = 0$.*

(ii) *For every point X_j at Voronoi distance > 2 from some $X_i \in X$, $D_{i,j}\varphi(X) = 0$.*

Remark 6.4. These properties mean somehow that φ is of range 2 with respect to the Voronoi tessellation. An analogue of Theorem 6.2 should hold for any functional with finite range, such as the perimeter approximation induced by φ_{Per}. On the other hand, the variance lower bound derived in this section is specific to the volume approximation.

We define for $x \in \mathbb{R}^d, X = (X_i)$ a finite collection of points, $k \geq 1$,

$$R_k(x;X) = \sup\{\|y - x\| : y \in V(X_i;X), d_V(x, X_i;X) \leqslant k\}$$

the distance to the furthest point in the cell of a k-th order Voronoi neighbour, with $R(x;X) := R_0(x;X)$. If x does not have k-th order neighbours, we put the convention $R_k(x;X) = \mathrm{diam}([0,1]^d) = \sqrt{d}$. We have obviously

$$\mathrm{Vol}(V(x;X)) \leqslant \kappa_d R(x;X)^d, x \in \mathbb{R}^d, \tag{6.8}$$

where κ_d is the volume of the unit sphere in $\mathbb{R}^d$.

Proof of Theorem 6.2. We will use Theorem 5.1 with the functional $f(X) = \varphi(X) - \mathbf{E}\varphi(X)$. Let us start with a crucial bound.

Lemma 6.5. *Assume that (6.3) holds. Define for some $k \geq 0$, the random variable*

$$U_k = 1_{\{d(X_1,\partial K)\leqslant R_k(X_1;X)\}} R_k(X_1;X)^d.$$

Then for some $c_{d,qd+\alpha,k} > 0$,

$$\mathbf{E}U_k^q \leqslant S_+(K) c_{d,qd+\alpha,k} n^{-q-\alpha/d},\ n \geq 1, q \geq 1.$$

Proof. Under this form, it is problematic to give a sharp upper bound because the law of $R_k(X_1;X)$ depends on the position of X_1 within $[0,1]^d$. To inject some stationarity in the problem, we will bound $R_k(X_1;X) = R_k(X_1;\hat{X}^1)$ by introducing a closely related quantity $\overline{R}_k(X_1;\hat{X}^1)$ whose conditional law with respect to $\hat{X}^1$ is independent of the value of X_1. To this end, introduce the process

$$X' = \bigcup_{m\in\mathbb{Z}^d} (X+m),$$

which law is invariant under translations. Remark that given any $t \in \mathbb{R}^d, X'$ has a.s. exactly n points in $[t,t+1]^d$. For $x \in \mathbb{R}^d$, call

$$\mathcal{C}_x = \{[x-t,x-t+1]^d; t \in [0,1]^d\} = \{[y,y+1]^d : y \in \mathbb{R}^d, x \in [y,y+1]^d\},$$

the family of translates of $[0,1]^d$ that contain x. Then by stationarity of X', the law $\mu_{k,n}$ of

$$\overline{R_k}(x,X) := \sup_{C\in\mathcal{C}_x} R_k(x,X'\cap C)$$

does not depend on x (and it is indeed only a function of x and X). Also, for $x \in [0,1]^d$, $[0,1]^d \in \mathcal{C}_x$, whence $R_k(x,X) \leqslant \overline{R_k}(x,X)$. This yields

$$\mathbf{E}U_k^q \leqslant \int_{[0,1]^d} dx 1_{\{d(x;\partial K)\leqslant \overline{R}_k(x;\hat{X}^1)\}} \overline{R}_k(x;\hat{X}^1)^{qd} \tag{6.9}$$

$$\leqslant \int_{\mathbb{R}_+\times[0,1]^d} 1_{\{d(x,\partial K)\leqslant r\}} r^{qd} \mu_{k,n-1}(dr)dx \tag{6.10}$$

$$\leqslant S_+(K)\mathbf{E}\overline{R}_k(0;\hat{X}^1)^{qd+\alpha} \quad \text{using (6.3).} \tag{6.11}$$

Let us now bound the probability of the event $\overline{R_k}(0,X) \geqslant r$, for some $r \geqslant 0$. If this event is realised, there is a k-th order Voronoi neighbour $z \in X'$ of 0 and a point y in the Voronoi cell of z such that $\|y\| \geqslant r$. There is therefore a sequence of points $x_1 = 0, x_2 \in X', \dots, x_k = z, x_{k+1} = y$ such that for $i < k$, x_i and x_{i+1} are Voronoi neighbours. Since the midpoint z_i of x_i and x_{i+1} has x_i and x_{i+1} as closest neighbours in $(X',0)$, the open ball $B^o(z_i, \|x_i - x_{i+1}\|/2)$ has an empty intersection with X'. Since z is the point of X' closest to y, $B^o((z+y)/2, \|z-y\|/2)\cap X = \emptyset$ also. We therefore have k (possibly empty) open balls $B_1,\dots,B_k$, with respective radii $r_i, i=1,\dots,k$, such that $[x_i,x_{i+1}]$ is a diameter of B_i, and such that X' has a point in none of them. Since $\|y\| \geqslant r$, the radius of at least one of these balls is larger than $r/2k$. Define

$$i_0 := \min\{1 \leqslant i \leqslant k : r_i \geqslant r/2k\}.$$

We have by the triangular inequality $\|x_{i_0}\| \leqslant i_0 r/2k \leqslant r/2$, and the ball $B(x_{i_0}, r/2k)$ is empty of points of X' and is contained in $[-r,r]^d$. It is easy to find $\gamma_d > 0$ such that at least one of the cubes $[g, g+\gamma_d r]^d, g \in \gamma_d r\mathbb{Z}^d \cap [-r,r]^d$ is contained in every ball with radius $r/2$ contained in $[-r,r]^d$. This yields

$$\begin{aligned}\mathbf{P}(\overline{R_k}(0,X) \geqslant r) &\leqslant \mathbf{P}(\exists g \in \gamma_d r\mathbb{Z}^d \cap [-r,r]^d : X' \cap [g, g+\gamma_d r]^d = \emptyset) \\ &\leqslant \#(\gamma_d'\mathbb{Z}^d \cap [-1,1]^d)\mathbf{P}([0, 0+\gamma_d r]^d \cap X' = \emptyset).\end{aligned}$$

Since $\#[0, 0+\gamma_d r]^d \cap X' \geqslant n$ for $r \geqslant \gamma_d^{-1}$ and $X' \cap [0, 0+\gamma_d r] = X \cap [0, 0+\gamma_d r]$ for $r \leqslant \gamma_d^{-1}$, we finally have

$$\mathbf{P}(\overline{R_k}(0,X) \geqslant r) \leqslant 2^d \gamma_d^{-d}(1-\gamma_d^d r^d)^n \leqslant 2^d \gamma_d^{-d} \exp(-n\gamma_d^d r^d).$$

It then follows that for $u > 0$,

$$\begin{aligned}\mathbf{E}\overline{R_k}(0,\hat{X}^1)^u &= \int_0^\infty \mathbf{P}(\overline{R_k}(0,\hat{X}^1) \geqslant r^{1/u})dr \leqslant 2^d\gamma_d^{-d} \int_0^\infty \exp(-(n-1)\gamma_d^d r^{d/u})dr \\ &\leqslant 2^d\gamma_d^{-d}(n-1)^{-u/d} \int_0^\infty \exp(-\gamma_d^d r^{d/u})dr.\end{aligned}$$

The conclusion follows by reporting this in (6.9). □

Proposition 6.3 and (6.8) yield for $q \geq 1$

$$|\mathbf{E}D_1 f(X)^q| \leqslant \kappa_d^q \mathbf{E}U_1^{qd}.$$

Lemma 6.5 implies, for $q \geq 1$,

$$\mathbf{E}|D_1 f(X)|^q \leqslant c_{d,qd+\alpha}\kappa_d^q S_+(K) n^{-q-\alpha/d}, \tag{6.12}$$

therefore the second term of the right-hand side of (6.6) follows immediately from the last estimate in (5.1). We now state Rhee-Talagrand's inequality [24], which then immediately yields (6.4).

Lemma 6.6 (Rhee-Talagrand's inequality)**.** *Let $\psi(X)$ be a symmetric measurable functional with finite q-th moment . Then for $q \geq 1$*

$$\mathbf{E}|\psi(X) - \mathbf{E}\psi(X)|^q \leqslant n^{q/2} c_q \mathbf{E}D_1|\psi(X)|^q$$

with $c_q = 2^q(18\sqrt{q}q')^{q'}$, where $1/q + 1/q' = 1$. For $q = 2$, Stein-Efron's inequality yields the better constant $c_2 = 1/2$.

Let us bound the two first terms of (5.1). We need for that to control the maximum radius of Voronoi cells over X. We first introduce the event on the circumscribed radii of the Voronoi spheres,

$$\Omega_n(X) = \left(\max_{1 \leqslant j \leqslant n} (R(X_j; X)) \leqslant n^{-1/d}\rho_n \right)$$

where $\rho_n = \log(n)^{1/d+\varepsilon'}$ for ε' sufficiently small. We have the following lemma, proved later for the sake of readability.

Lemma 6.7. *For all $\eta > 0$, $n^\eta \mathbf{P}(\Omega_n(X)^c) \to 0$ as $n \to \infty$.*

To bound the first term of (5.1), let Y, Y', Z be recombinations of $\{X, X', \tilde{X}\}$. Introduce the event $\Omega := \Omega_n(Y) \cap \Omega_n(Y') \cap \Omega_n(Z) \cap \Omega_n(Z')$ which satisfies $\mathbf{P}(\Omega^c) \leqslant 4\mathbf{P}(\Omega_n(X)^c)$. Recall the fact that $D_{ij}f(X)$ can only be non-zero if X_j is at Voronoi distance $\leqslant 2$ from X_i, and that $D_j f(X)$ can only be non-zero if X_j has a Voronoi neighbour which cell touches ∂K. In the notation of (5.1), we have

$$\begin{aligned}
\mathbf{E}1_{\{D_{1,2}\varphi(Y)\neq 0\}} D_1\varphi(Z)^4 &\leqslant \mathbf{E}1_\Omega 1_{\{D_{1,2}\varphi(Y)\neq 0\}} D_1\varphi(Z)^4 + \mathbf{P}(\Omega^c) \\
&\leqslant \kappa_d^4 n^{-4}\rho_n^{4d}\mathbf{E}[1_{\{d(Y_1,\partial K)\leqslant 2n^{-1/d}\rho_n\}}\mathbf{E}[1_{\{\|Y_1-Y_2\|\leqslant 2n^{-1/d}\rho_n\}}|Y_1]] + \mathbf{P}(\Omega^c) \\
&\leqslant \kappa_d^5 n^{-4}\rho_n^{4d}2^d n^{-1}\rho_n^d \mathbf{P}(d(Y_1,\partial K)\leqslant 2n^{-1/d}\rho_n) + \mathbf{P}(\Omega^c) \\
&\leqslant C_{1,2} n^{-5-\alpha/d}\rho_n^{5d+\alpha}
\end{aligned}$$

for some $C_{1,2} \geq 0$, whence Proposition 5.3 and (5.3) yield $nB_n(f) \leqslant C' n^{-4-\alpha/d}\rho_n^{5d+\alpha}$ for some $C' > 0$. With a similar computation,

$$\begin{aligned}
&\mathbf{E}\mathbf{1}_{\{\Omega\}}\mathbf{1}_{\{D_{1,2}\varphi(Y)\neq 0, D_{1,3}\varphi(Y')\neq 0\}} D_2\varphi(Z)^4 \\
&\quad\leqslant \kappa_d^4 n^{-4}\rho_n^{4d}\mathbf{P}(\|Y_1-Y_2\|\leqslant 2n^{-1/d}\rho_n, \|Y_1'-Y_3'\|\leqslant 2n^{-1/d}\rho_n, d(Y_1,\partial K)\leqslant 2n^{-1/d}\rho_n) + \mathbf{P}(\Omega^c) \\
&\quad\leqslant C_{2,3} n^{-6-\alpha/d}\rho_n^{6d+\alpha},
\end{aligned}$$

from where $n^2 B_n'(f) \leqslant C'' n^{-4-\alpha/d}\rho_n^{6d+\alpha}$ for some $C'' > 0$. Therefore the first term of (5.1) is bounded by

$$\sigma^{-2}\sqrt{n}(n^{-2-\alpha/2d})\log(n)^{3+\alpha/2d+d\varepsilon'/2}$$

up to a constant, which yields the first term of (6.6). It remains to bound the term

$$\mathbf{E}|f(X)||D_j f(X^A)|^3$$

from (5.1). Recall that under $\Omega_n(X^A)$, all Voronoi cells volumes, and therefore all $|D_j f(X^A)|$,$1 \leqslant j \leqslant n$, are bounded by $\kappa_d n^{-1}\rho_n^d$, and also, $D_j f(X^A) = 0$ if X_j and X_j' are at distance more than $2n^{-1/d}\rho_n$ from $K's$ boundary. We have

$$\begin{aligned}
\mathbf{E}|f(X)D_j f(X^A)|^3 &\leqslant \mathbf{E}\left(|f(X)||D_j f(X^A)|^3 1_{\Omega_n}(X^A)\right) + \mathbf{P}(\Omega_n(X)^c) \\
&\leqslant cn^{-3}\rho_n^{3d}\mathbf{E}\left[|f(X)|1_{\{X_j \text{ or } X_j'\in\partial K^{2n^{-1/d}\rho_n}\}}\right] + \mathbf{P}(\Omega_n(X)^c) \\
&\leqslant cn^{-3}\rho_n^{3d}\mathbf{E}\left(\left(|f(\hat{X}^j)| + |D_j f(X)|\right)1_{\left\{X_j \text{ or } X_j'\in\partial K^{2n^{-1/d}\rho_n}\right\}}\right) + \mathbf{P}(\Omega_n(X)^c).
\end{aligned}$$

We have

$$\mathbf{E}|D_j f(X)| \leqslant c' n^{-1-\alpha/d}$$

by (6.12), while the other term is bounded by independence by

$$\begin{aligned}
\mathbf{E}|f(\hat{X}^j)|1_{\left\{X_j \text{ or } X_j'\in\partial K^{2n^{-1/d}\log(n)}\right\}} &\leqslant 2\mathbf{E}|f(\hat{X}^j)|\mathbf{P}\left(X_j\in\partial K^{2n^{-1/d}\rho_n}\right) \\
&\leqslant c''\sigma n^{-\alpha/d}\rho_n^\alpha.
\end{aligned}$$

Finally, for some $C > 0$,

$$\mathbf{E}|f(X)D_j f(X^A)|^3 \leqslant Cn^{-3-\alpha/d}\log(n)^{3+\varepsilon/2}(\sigma\log(n)^{\alpha/d+\varepsilon/2} + n^{-1}),$$

which gives the desired bound.

□

Proof of Lemma 6.7. We can find a constant $\gamma_d > 0$ such that the intersection with $[0,1]^d$ of every ball centred in $[0,1]^d$ of radius $r \leqslant 1$ contains a cube $g + [0, \gamma_d r]^d$ for some $g \in \gamma_d r\mathbb{Z}^d$. If $\max_{1\leqslant j\leqslant n} R(X_j; X) > n^{-1/d}\rho_n$, then two Voronoi neighbours X_i, X_j are at distance more than $n^{-1/d}\rho_n$ from one another, and the open ball with diameter $[X_i, X_j]$ does not contain points of X, by the construction of the Voronoi tessellation. It follows that a cube $g + [0, \gamma_d n^{-1/d}\rho_n]^d \subseteq [0,1]^d$ is empty of points of X, for some $g \in \gamma_d n^{-1/d}\rho_n\mathbb{Z}^d$, and this event happens with a probability bounded by

$$\begin{aligned}(\gamma_d n^{-1/d}\rho_n)^{-d}\mathbf{P}([0,\gamma_d n^{-1/d}\rho_n]^d \cap X = \emptyset) &\leqslant \gamma_d^{-d} n\rho_n^{-d}(1-\gamma_d^d n^{-1}\rho_n^d)^n \\ &\leqslant \gamma_d^{-d} n\rho_n^{-d}\exp(n\log(1-\gamma_d^d n^{-1}\rho_n^d)) \\ &\leqslant \gamma_d^{-d} n\rho_n^{-d}\exp(-\gamma_d^d\log(n)^{1+d\varepsilon'}),\end{aligned}$$

which proves the result. □

Proof of Theorem 6.1. It only remains to prove the lower bound on the variance in (6.5). Lemma 2.4 states that the variance is larger than $n\|h\|^2_{L^2([0,1]^d,\ell)}$, where

$$h(x) = \mathbf{E}\varphi(\hat{X}^1, x) - \mathbf{E}\varphi(X), \quad x \in [0,1]^d.$$

We decompose h as follows:

$$\begin{aligned}h(x) &= (\mathbf{E}\varphi(\hat{X}^1, x) - \varphi(\hat{X}^1) - (\mathbf{E}\varphi(X) - \varphi(\hat{X}^1)), x \in [0,1]^d \\ &=: h_1(x) - h_2. \end{aligned} \tag{6.13}$$

Voronoi volume approximation is not homogeneous in the sense that points falling close to K's boundary have more influence than other points of X_n. The following lemma shows that this inhomogeneity makes h_1 the dominant term in the previous decomposition.

Lemma 6.8. *Let K be a measurable subset of $[0,1]^d$, define h_1 as in (6.13). Then we have*

$$\int_{[0,1]^d} h_1(x)^2 dx \geq C_d(\gamma(K, n^{-1/d}) + \gamma(K^c, n^{-1/d}))n^{-2}$$

for some $C_d > 0$.

Let us first conclude the proof of Theorem 6.1. If the weak rolling ball condition is satisfied along with (6.5), it yields

$$\int_{[0,1]^d} h_1(x)^2 dx \geq C_d S_-(K)\gamma(K)(n^{-1/d})^\alpha n^{-2}.$$

According to Lemma 6.5, $h_2 = O(n^{-1-\alpha/d})$, which is indeed negligible with respect to $\|h_1\|_{L^2} \geq C_{d,K}n^{-1-\alpha/2d}$. □

Proof of Lemma 6.8. It follows from (6.7) that for $x \in K^c$

$$|\varphi(x, \hat{X}^1) - \varphi(\hat{X}^1)| = \sum_{j=2}^{n} 1_{\{X_j \in K\}} v(x, X_j; \hat{X}^1),$$

where we notice that the summand distribution does not depend on j. Then

$$\begin{aligned}
|h_1(x)| &\geq 1\left(x \in \partial K_+^{n^{-1/d}}\right)(n-1)\mathbf{E}1_{\{X_2 \in K\}} v(x, X_2; \hat{X}^1) \\
&\geq 1\left(x \in \partial K_+^{n^{-1/d}}\right)(n-1)\mathbf{E}\int_{y \in K} v(x, y; \hat{X}^{1,2})dy \\
&\geq 1\left(x \in \partial K_+^{n^{-1/d}}\right)(n-1)\mathrm{Vol}(B(x, \beta n^{-1/d}) \cap K) \inf_{y: \|y-x\| \leqslant \beta n^{-1/d}} \mathbf{E}v(x, y; \hat{X}^{1,2}).
\end{aligned}$$

If for some $y \in [0,1]^d, \varepsilon > 0$, no point of $\hat{X}^{1,2} := (X_i)_{i \neq 1,2}$ falls in $B(y, 6\varepsilon)$, then $B(y, 3\varepsilon) \subset V(y, \hat{X}^{1,2})$. If furthermore $x \in [0,1]^d$ lies at distance less than ε from y, then with $z = x + \varepsilon\|x - y\|^{-1}(x-y)$,

$$B(z, \varepsilon) \subset V(x, (\hat{X}^{1,2}, y)) \subset B(y, 3\varepsilon) \subset V(y; \hat{X}^{1,2}),$$

and therefore $v(x, y; \hat{X}^{1,2}) \geq \kappa_d \varepsilon^d$. We finally have

$$\inf_{y: \|y-x\| \leqslant \beta n^{-1/d}} \mathbf{E}v(x, y; \hat{X}^{1,2}) \geq \kappa_d \beta^d n^{-1} \mathbf{P}(\hat{X}^{1,2} \cap B(y, 6\beta n^{-1/d}) = \emptyset) \geq c_d' n^{-1}$$

for some $c_d' > 0$. With a completely similar result for $x \in K$, we have for some $c_d'' > 0$

$$\int_W h_1(x)^2 dx \geq c_d'' \left(\int_{\partial K_+^{n^{-1/d}}} \mathrm{Vol}(B(x, \beta n^{-1/d}) \cap K)^2 dx + \int_{\partial K_-^{n^{-1/d}}} \mathrm{Vol}(B(x, \beta n^{-1/d}) \cap K^c)^2 dx \right).$$

Remark 6.9. All three terms of (5.1) give in the case of Theorem 6.1 a bound of order $n^{-1/2+\alpha/2d}\log(n)^q$ for some $q > 0$. In these conditions it seems hard to reach a Berry-Essen bound negligible with a better magnitude than $n^{-1/2+\alpha/2d}$, but removing the log is an open problem.

□

6.2 Covering processes

Let $(\mathcal{K}, \mathscr{K})$ be the space of compact subsets of $\mathbb{R}^d$, endowed with the hit-and-miss topology and a Borel probability measure ν. Let E_n be a cube of volume n, and $C_1, \dots, C_n$ iid uniform variables in E_n, called the *germs*. Let n iid compact sets $K_1, \dots, K_n$ be distributed as ν, called the *grains*, and define the *germ-grain process* $X_i = C_i + K_i$, $i = 1, \dots, n$. An important feature of the model regarding Gaussian approximation is the radius

$$R_i := \sup\{\|x\| : x \in K_i\}, 1 \leqslant i \leqslant n.$$

We consider the random closed set formed by the union of the grains translated by the germs

$$F_n = (\cup_{k=1}^n X_k) \cap E_n.$$

We are interested in the volume of C_n covered by F_n

$$f_V(X_1,\dots,X_n)=\mathrm{Vol}(F_n),$$

the number of isolated grains

$$f_I(X_1,\dots,X_n)=\#\{k : X_k\cap X_j\cap E_n=\emptyset, k\neq j\},$$

and their centred versions with unit variance $\tilde f_V, \tilde f_I$. The functional f_V denotes the total volume of the germ-grain process, and $n^{-1}f_V(X_1,\dots,X_n)$ can serve as an estimator for the *fraction volume*, i.e. the portion of the space occupied by the boolean model $\cup_k X_k$, and therefore be used in estimating the parameters of ν (see [20] for insights on the boolean model statistics).

Kolmogorov Berry-Essen bounds in $n^{-1/2}$ for binomial input for f_V or f_I have only been obtained very recently in [11] with balls with deterministic identical radii (with the possibility to extend the method to a random radius), using size-biased couplings. Chatterjee [4] obtained similar bounds in Wasserstein distance. We present here the first such bounds in the unbounded random grain context. Furthermore, the computations are quite straightforward and the method is generalisable to similar local functionals of the boolean model, such as the perimeter, or other Minkowski functionals. The use of the bound (5.1) is crucial to have a decay in $n^{-1/2}$ in the context of random grains. The variance is a straightforward computation of integral geometry, it is a consequence of for instance [16, Th. 4.4] that under the conditions of the theorem below, we have $cn\leqslant \mathbf{Var} f(X_1,\dots,X_n)\leqslant Cn$ for some $c,C>0$, for $f=f_V$ or $f=f_I$.

Theorem 6.10. *Assume that* $\mathbf{E}R_1^{5d}<\infty$. *Let N be a standard Gaussian variable. Then we have for some $C>0$,*

$$d_K(\tilde f_V(X_1,\dots,X_n),N)\leqslant Cn^{-1/2}.$$

If $\mathbf{E}R_1^{8d}<\infty$, *for some $C'>0$,*

$$d_K(\tilde f_I(X_1,\dots,X_n),N)\leqslant C'n^{-1/2}.$$

Proof. Let first $f=f_V$. Given a n-tuple $x=(x_1,\dots,x_n)\in\mathcal{K}^n$, we have $D_{i,j}f(x)=0$ as soon as $\mathrm{Vol}(x_i\cap x_j)=0$, which gives us a sufficient condition. Let us estimate the right hand side of (5.1). Introduce independent copies $X',\tilde X$ of X, and for U a random compact set among those families, denote by $c(U)$, $r(U), K(U)$ its centre, radius, and grain, so that

$$\{c(X_i),c(X_i'),c(\tilde X_i),K(X_i),K(X_i'),K(\tilde X_i),1\leqslant i\leqslant n\}$$

is a family of independent variables. Let us write $V_i=\mathrm{Vol}(X_i), V_i'=\mathrm{Vol}(X_i')$. We have $|D_1f_V(X)|\leqslant V_1$, and since the volume has a finite moment of order 5,

$$\sup_{n\geq 1}\mathbf{E}|D_1f(X)|^3<\infty,\ \sup_{n\geq 1}\mathbf{E}|D_1f(X)|^4<\infty.$$

We also have for $A\subseteq[n]$

$$\begin{aligned}\mathbf{E}|f(X)||D_jf(X^A)|^3&\leqslant \mathbf{E}|f(X^{\hat j})D_jf(X^A)|^3+\mathbf{E}|D_jf(X)D_jf(X^A)^3|\\&\leqslant \mathbf{E}|f(X^{\hat j})|(V_j^3+(V_j')^3)+\mathbf{E}D_jf(X)^4\\&\leqslant \mathbf{E}|f(X^{\hat j})|2\mathbf{E}V_j^3+\mathbf{E}V_j(X)^4,\end{aligned}$$

whence

$$\sigma^{-4}n\mathbf{E}|f(X)D_jf(X^A)^3| \leqslant Cn^{-1/2}$$

for some $C>0$.

To estimate $B_n(f), B'_n(f)$, we use Proposition 5.3, (5.2), and (5.3). Fix Y,Y',Z recombinations of $\{X,X',\tilde{X}\}$, we have

$$\begin{aligned}\mathbf{E}\mathbf{1}_{\{D_{1,2}f(Y)\neq 0\}}D_1f(Z)^4 &\leqslant \mathbf{E}\mathbf{1}_{\{Y_2\cap Y_1\neq\emptyset\}}\mathrm{Vol}(Z_1)^4\\ &\leqslant \mathbf{E}\left[\kappa_d^4 r(Z_1)^{4d}\mathbf{P}(c(Y_2)\in B(c(Y_1),r(Y_1)+r(Y_2))|Y_1,Z_1,r(Y_2))\right]\\ &\leqslant n^{-1}\kappa_d^5\mathbf{E}r(Z_1)^{4d}(r(Y_1)+r(Y_2))^d\end{aligned}$$

whence $\sup_n nB_n(f)<\infty$ since $\mathbf{E}R_1^{5d}<\infty$.

Then,

$$\begin{aligned}&\mathbf{E}\mathbf{1}_{\{D_{1,2}f(Y)\neq 0, D_{1,3}f(Y')\neq 0\}}D_2f(Z)^4\\ &\quad\leqslant \mathbf{E}\left[\mathrm{Vol}(Z_2)^4\mathbf{1}_{\{D_{12}f(Y)\neq 0\}}\mathbf{P}(c(Y'_3)\in B(c(Y'_1),r(Y'_1)+r(Y'_3))|Z_2,Y_1,Y_2,Y'_1,r(Y'_3))\right]\\ &\quad\leqslant n^{-1}\kappa_d^5\mathbf{E}\left[r(Z_2)^4(r(Y'_1)+r(Y'_3))^d\mathbf{P}(c(Y_2)\in B(c(Y_1),r(Y_1)+r(Y_2))|Z_2,Y_1,Y'_1,Y'_3,r(Y_2))\right]\\ &\quad\leqslant n^{-2}\kappa_d^6\mathbf{E}r(Z_2)^4(r(Y'_1)+r(Y'_3))^d(r(Y_1)+r(Y_2))^d.\end{aligned}$$

Using the definition of recombinations, the variables Y'_1, Z_2, Y'_3 are pairwise independent, and the expectation above is finite because of $\mathbf{E}r(X_1)^{5d}<\infty$. We indeed have $\sup_n n^2B'_n(f)<\infty$, which concludes the proof for the Kolmogorov bound on $\tilde{f}_V$.

Dealing with $f=f_I$ is slightly more complicated. Introduce $d_{i,j}(X)$ the distance between i and j in the germ-grain process X, defined as the smallest number q such that there is a chain $i_1=i,\dots,i_q=j$ such that $X_{i_k}\cap X_{i_{k+1}}\neq\emptyset$. Call $B_i^p(X)$ the number of points at distance $\leqslant p$ from the point i for the distance $d_{\cdot,\cdot}(X)$. For some $1\leqslant i,j\leqslant n$, the value of the functional

$$\mathbf{1}_{\{X_j \text{ is isolated}\}} := \mathbf{1}_{\{X_j\cap X_k\cap E_n=\emptyset, k\neq j\}}$$

can be affected by the removal of X_i only if $X_i\cap X_j\neq\emptyset$, therefore, for $1\leqslant i\leqslant n$,

$$|D_if_I(X)|\leqslant \#B_i^1(X),$$

whence,

$$\mathbf{E}|D_1f_I(X)|^q\leqslant \mathbf{E}\#B_i^1(X)^q, q\leqslant 1. \tag{6.14}$$

We will estimate this bound later. With the same notation than for the functional f_V, let us now deal with $B_n(f), B'_n(f)$. Remark that $D_{i,j}f_I(X)=0$ if $d_{i,j}(X)>2$. We have

$$B_n(f)\leqslant \sup_{(Y,Z)}\mathbf{E}\mathbf{1}_{\{2\in B_1^2(Y)\}}\#B_1^1(Z)^4$$

and

$$\mathbf{1}_{\{2\in B_1^2(Y)\}}\leqslant\sum_k\mathbf{1}_{\{X_1\cap X_k\neq\emptyset, X_2\cap X_k\neq\emptyset\}}.$$

To simplify notation, remark that for Y, Z recombinations of $\{X, X', \tilde{X}\}$, $\#B_1^p(Y) \leqslant \#B_1^p(T)$, where T is the concatenation of Y and Z and is in fact composed of m iid variables distributed as X_1, where $n \leqslant m \leqslant 2n$. We then have

$$B_n(f) \leqslant \sup_{n \leqslant m \leqslant 2n} \mathbf{E} \sum_{k=1}^{m} \mathbf{1}_{\{T_1 \cap T_k \neq \emptyset, T_k \cap T_2 \neq \emptyset\}} \sum_{1 \leqslant k_1, k_2, k_3, k_4 \leqslant m} \mathbf{1}_{\{T_{k_i} \cap T_1 \neq \emptyset, i=1,\ldots,4\}}, \tag{6.15}$$

and the supremum is reached for $m = 2n$. We have similarly, with $m = 3n$,

$$B'_n(f) \leqslant \mathbf{E} \sum_{k=1}^{m} \mathbf{1}_{\{T_1 \cap T_k \neq \emptyset, T_2 \cap T_k \neq \emptyset\}} \sum_{k'=1}^{m} \mathbf{1}_{\{T_1 \cap T_{k'} \neq \emptyset, T_3 \cap T_{k'} \neq \emptyset\}} \sum_{\mathbf{k}=(k_1,k_2,k_3,k_4) \in [m]^4} \mathbf{1}_{\{T_1 \cap T_{k_i} \neq \emptyset\}}. \tag{6.16}$$

To estimate (6.14)-(6.16), it is useful to introduce some more notation. Call graph on $[n]$ the finite data of distinct edges $t = \{\{i_1, j_1\}, \ldots, \{i_q, j_q\}\}$. For such a graph, introduce the probability

$$p(t) = \mathbf{P}(T_{i_1} \cap T_{j_1} \neq \emptyset, \ldots, T_{i_q} \cap T_{j_q} \neq \emptyset).$$

Say that this graph is a tree when it is connected and has no cycles. Let us prove that for every tree t with q distinct vertices,

$$p(t) \leqslant (d\kappa_d n^{-1})^{q-1} \mathbf{E} r(T_1)^{(q-1)d}. \tag{6.17}$$

Let t be such a tree, and let an arbitrary vertex i_0 of t, designated to be the root of t. Call $\mathcal{G}_k(t), k \geq 1$, the members of the k-th generation, noticing that there can not be more than q generations, i.e. $\mathcal{G}_k(t) = \emptyset$ for $k > q$. Call $\mathcal{G}_k^-(t) = \cup_{j<k} \mathcal{G}_j(t), \mathcal{G}_k^+(t) = \mathcal{G}_k(t) \setminus \mathcal{G}_k^-(t)$, and call $\mathcal{G}_k^{k+1}(t)$ the collection of all pairs (i,j) such that $i \in \mathcal{G}_k(t), j \in \mathcal{G}_{k+1}(t), \{i,j\} \in t$. We have

$$\begin{aligned}
p(t) \leqslant \mathbf{E} \Big[& \mathbf{1}_{\{T_i \cap T_j \neq \emptyset; \{i,j\} \in t; i,j \in \mathcal{G}_q^-(t)\}} \\
& \mathbf{P}\Big(c(T_j) \in B(c(T_i), r(T_i) + r(T_j)); (i,j) \in \mathcal{G}_{q-1}^q(t) \,\Big|\, c(T_i), i \in \mathcal{G}_q^-(t); r(T_i), i \in [m] \Big) \Big] \\
\leqslant \mathbf{E} & \left[\mathbf{1}_{\{T_i \cap T_j \neq \emptyset; \{i,j\} \in t, i,j \in \mathcal{G}_q^-(t)\}} \prod_{(i,j) \in \mathcal{G}_{q-1}^q(t)} n^{-1} \kappa_d (r(T_i) + r(T_j))^d \right] \\
\leqslant (\kappa_d & n^{-1})^{\#\mathcal{G}_{q-1}^q(t)} \mathbf{E} \left[\mathbf{1}_{\{T_i \cap T_j \neq \emptyset; \{i,j\} \in t, i,j \in \mathcal{G}_q^-(t)\}} \prod_{\{i,j\} \in t: i,j \in \mathcal{G}_q^+(t)} (r(T_i) + r(T_i))^d \right].
\end{aligned}$$

Applying this procedure inductively back until the 1-st generation , that is the root i_0 of the tree, yields

$$p(t) \leqslant (\kappa_d n^{-1})^{\sum_{k \geq 1} \#\mathcal{G}_k^{k+1}(t)} \mathbf{E} \left[\prod_{(i,j) \in \cup_k \mathcal{G}_k^{k+1}(t)} (r(T_i) + r(T_j))^d \right].$$

Now, $\cup_{k\geq 1}\mathcal{G}_k^{k+1}(t)$, contains all the $q-1$ edges of t, whence

$$p(t) \leqslant \kappa_d^{q-1} n^{-(q-1)} \mathbf{E} \prod_{\{i,j\}\in t} (r(T_i)+r(T_j))^d \leqslant (d\kappa_d n^{-1})^{q-1} \mathbf{E} r(T_1)^{(q-1)d},$$

by using Cauchy-Schwarz inequality, whence (6.17) follows.

We have

$$\mathbf{E}|D_1 f_I(X)|^6 \leqslant \sum_{\mathbf{k}=(k_1,\dots,k_6)\in[m]^6} p(\{1,k_i\}, i=1,\dots,6) \leqslant Cn^{-5}$$

for some $C>0$, by using $\mathbf{E}r(X_1)^{5d}<\infty$, which treats all the terms of (5.1) except the ones containing $B_n(f)$ and $B_n'(f)$.

We call, for $u_1,\dots,u_q$ distinct integers, $l\geq 0, p\geq 4$,

$$[m]^p_{u_1,\dots,u_q;l} = \{\mathbf{k}=(k_1,\dots,k_p)\in[m]^p \ : \ \#\{u_1,\dots,u_q,k_1,\dots,k_p\}\} = q+l.$$

We can easily prove that there are constants C_l not depending on m such that

$$\#[m]^p_{u_1,\dots,u_q;l} \leqslant C_l n^l. \tag{6.18}$$

We have, for T with $2n$ iid components, using (6.15),

$$\begin{aligned} B_n(f) &\leqslant \sum_{k=1}^n \sum_{\mathbf{k}=(k_i)\in[2n]^4} p(\{1,k\},\{2,k\},\{1,k_i\}; i=1,\dots,4) \\ &\leqslant \sum_{l=0}^5 \sum_{\mathbf{k}\in[m]^5_{1,2;l}} p(\{1,k_1\},\{2,k_1\},\{1,k_i\}; i=2,\dots,5). \end{aligned}$$

For $\mathbf{k}\in[m]^5_{1,2;l}$, one can easily extract a tree with $l+1$ edges from $\{\{1,k_1\},\{2,k_1\},\{1,k_i\}; i=2,\dots,5\}$, whence (6.17) yields

$$B_n(f) \leqslant C\sum_{l=0}^5 \sum_{\mathbf{k}\in[m]^5_{1,2;l}} n^{-l-1} \leqslant C' n^{-1},$$

using also (6.18). This gives $\sup_n nB_n(f)<\infty$. Similar computations yield

$$\begin{aligned} B_n'(f) &\leqslant \mathbf{E}\sum_k \mathbf{1}_{\{T_1\cap T_k\neq\emptyset, T_2\cap T_k\neq\emptyset\}} \sum_{k'} \mathbf{1}_{\{T_1\cap T_{k'}\neq\emptyset, T_3\cap T_{k'}\neq\emptyset\}} \sum_{\mathbf{k}=(k_1,k_2,k_3,k_4)\in[m]^4} \mathbf{1}_{\{T_1\cap T_{k_i}\neq\emptyset\}} \\ &\leqslant \sum_{\mathbf{k}=(k_i)\in[m]^6} p(\{1,k_1\},\{2,k_1\},\{1,k_2\},\{3,k_2\},\{1,k_i\}, i=3,\dots,6) \\ &= \sum_{l=0}^6 \sum_{\mathbf{k}=(k_i)\in[m]^6_{1,2,3;l}} p(\{2,k_1\},\{3,k_2\},\{1,k_i\}, i=1,\dots,6) \end{aligned}$$

and for $\mathbf{k} \in [m]^6_{1,2,3;l}$ one can extract a tree with $l+2$ edges from $\{\{2,k_1\},\{3,k_2\},\{1,k_i\}; i = 1...6\}$, whence

$$B'_n(f) \leqslant \sum_{l=0}^{6} \sum_{\mathbf{k}\in[m]^6_{1,2,3;l}} (\kappa_d d n^{-1})^{l+2} \leqslant C n^{-2},$$

which concludes the proof.

□

6.3 Further applications

It is proved in [4] that, in the notation of Theorem 4.2 and for $\sigma = 1$,

$$d_W(W,N) \leqslant \delta_1 + \delta_2 \tag{6.19}$$

$$\delta_1 := \sqrt{\mathbf{Var}(\mathbf{E}(T|X))} \tag{6.20}$$

$$\delta_2 := 2c \sum_{j=1}^{n} \mathbf{E}|\Delta_j f(X)|^3 \tag{6.21}$$

where d_W is the 1-Wasserstein distance. This bound has been successfully applied in [4], [5], and [21] to several normal approximation problems. Without fully developing the details, we indicate here how we can obtain similar bounds in the Kolmogorov's distance by using the techniques developed in this paper. Assuming that $\sigma = 1$, the new terms in (4.2) with respect to (6.19) are

$$\delta'_1 = \sqrt{\mathbf{Var}(\mathbf{E}(T'|X))}$$
$$\delta'_2 = 6 \sum_{j=1}^{n} \sqrt{\mathbf{E}|D_j f(X)|^6}.$$

The term δ'_1 is very close in its expression to δ_1. In the examples developed below, it is indeed possible to apply the bound already derived for δ_1 to δ'_1. The term δ'_2 has to be dealt with separately, it is in general more straightforward. Remark that δ'_2 can be replaced by the bound $\delta''_2 = \sup_A \sum_{j=1}^{n} \mathbf{E}|f(X) D_j f(X^A)^3|$ from (4.1), which can give a better convergence rate or less restrictive hypotheses, but it requires a specific analysis and we do not develop it below.

Nearest neighbours statistics. Let $k \geq 1, i \geq 1$, let $\psi : (\mathbb{R}^d)^k \to \mathbb{R}$ be a measurable function and let

$$f(x_1,\ldots,x_n) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi(x_i^{(1)},\ldots,x_i^{(k)})$$

where the $x_i^{(j)}$ are the k nearest neighbours of x_i among $(x_1,\ldots,x_n)$ for the Euclidean distance, ordered by increasing distance to x_i, with an arbitrary tie breaking rule. Given n i.i.d random variables $X_1,\ldots,X_n$ in $\mathbb{R}^d$, in [4] Chatterjee obtains estimates on the Wasserstein distance between $f(X)$ and the normal law under the assumptions that for $i \neq j$, $\|X_i - X_j\|$ is a continuous random variable. He obtains the bounds, for $p \geq 8$,

$$\delta_1 \leqslant C_d \frac{k^4 \gamma_p^2}{\sigma^2 n^{(p-8)/2p}},$$
$$\delta_2 \leqslant C_d \frac{k^3 \gamma_p^3}{\sigma^3 n^{(p-6)/2p}},$$

where $\gamma_p := (\mathbf{E}|\psi(X_1,\ldots,X_n)|^p)^{1/p}, C_d > 0$. These bounds are obtained through [4, Theorem 2.5], which is similar to Theorem 5.1, where our bound on δ_1' is already smaller or equal to the bound on δ_1 from [4, Theorem 2.5], up to a constant, see Remark 5.4. Therefore we have $\delta_1' \leqslant C\delta_1$. In order to obtain an explicit bound on the Kolmogorov distance, it therefore only remains to bound δ_2'. In [4] it is shown that $\mathbf{E}\sup_{j=1}^n |\Delta_j f(X)|^p \leqslant (n^2+n)n^{-p/2}\gamma_p^p$ from where the bounds

$$\delta_1' \leqslant C_{k,d} n^{1/2} \left(\mathbf{E} \sup_{j=1}^n |\Delta_j f(X)|^p \right)^{2/p} \leqslant C_{k,d} n^{4/p} n^{1/2} n^{-1} \gamma_p^2 = C_{k,d} \frac{\gamma_p^2}{n^{(p-8)/2p}}$$
$$\delta_2 \leqslant C_{k,d} n \left(\mathbf{E} \sup_{j=1}^n |\Delta_j f(X)|^p \right)^{3/p} \leqslant C_{k,d} \frac{\gamma_p^3}{n^{1/2-6/p}}$$
$$\delta_2' \leqslant C_{k,d} n \left(\mathbf{E} \sup_{j=1}^n |\Delta_j f(X)|^p \right)^{3/p} \leqslant \delta_2.$$

easily follow. We observe that in [4] a more general situation is actually considered : for each i, a different functional ψ_i is applied to $(x_i^{(1)},\ldots,x_i^{(k)})$ in the definition of f. However, all the explicit examples developed in such reference are purely geometric, in the sense that this subtlety is not exploited, and the functional $f(X)$ is symmetric. These examples includes the *average distance to the nearest neighbour*, the *degree count in the nearest-neighbour graph,* and the*Levina-Bickel statistic with parameter* k, which is defined by

$$f(x_1,\ldots,x_l) = \frac{1}{n}\sum_{i=1}^n \left(\frac{1}{k-1}\sum_{j=1}^{k-1} \log\left(\frac{\|x_i - x_i^{(k)}\|}{\|x_i - x_i^{(j)}\|} \right) \right).$$

Flux through a random conductor. In [21], Nolen considers the solution of an elliptic partial differential equation with a stationary random conductivity coefficient $a(x)$ over the torus $[0,L)^d, L>0$. The random function $a(x)$ depends on the local contributions of a set of i.i.d variables $Z = (Z_1,\ldots,Z_k)$ indexed by $\mathbb{Z}^d \cap [0,L)^d$. He derives a bound on the Wasserstein distance between the normal law and the average flux $\Gamma(Z)$ of the solution. He obtains the bounds

$$\delta_1 \leqslant CL^{-3d/2}\sigma^{-2}\log(L)\left(\mathbf{E}\Phi_0^{8q}\right)^{1/2q}, \tag{6.22}$$
$$\delta_2 \leqslant C\sigma^{-3}L^{-2d}\mathbf{E}\Phi_0^6, \tag{6.23}$$

where σ^2 is the variance and Φ_0 is an integral related to the gradient of the solution over $[0,1)^d$ (see [21] for details).

Our method allows one to extend this result to the Kolmogorov distance, under slightly stronger assumptions. Gloria and Nolen [10] have also used Theorem 4.2 for a Kolmogorov Berry-Essen bound with a discretised version of the problem. Once again, the simple inequality $||a| - |b|| \leqslant |a-b|, a, b \in \mathbb{R}$, yields that the upper bound on $\mathbf{Var}(T(Z, Z')|Z')$ derived in [21, (2.25)-(2.27)] and then used in (4.53) can be used in an exact similar fashion to bound $\mathbf{Var}(T'(Z, Z')|Z')$ where T' is defined as in our Theorem 4.2. This yields that δ_1' satisfies the same bound as δ_1, up to a constant. Then, [21, Lemma 4.1] provides the estimate

$$\mathbf{E}|\Delta_j \Gamma(Z)|^q \leqslant C_q L^{-qd} \mathbf{E}|\Phi_0(Z)|^{2q}$$

which readily yields the first term of 6.22, and the bound on the Kolmogorov distance

$$\delta_1 + \delta_2 + \delta_1' + \delta_2' \leqslant C(\delta_1 + L^{-2d}\sqrt{\mathbf{E}|\Phi_0|^{12}}).$$

Note that the new condition $\mathbf{E}|\Phi_0|^{12} < \infty$ might be weakened if one uses (4.1) instead of (4.2), as it is done in the proof of Theorem 6.1.

References

[1] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities.* Oxford, 2013.

[2] S. Bourguin and G. Peccati. *Stochastic analysis for Poisson point processes: Malliavin calculus, Wiener-Ito chaos expansions and stochastic geometry*, chapter Stein and Chen–Stein methods and Malliavin calculus on the Poisson space. Springer, in preparation.

[3] P. Calka and N. Chenavier. Extreme values for characteristic radii of a Poisson-Voronoi tessellation. *Extremes*, 17:359–385, 2014.

[4] S. Chatterjee. A new method of normal approximation. *Ann. Probab.*, 36(4):1584–1610, 2008.

[5] S. Chatterjee. *Superconcentration and related topics.* Springer, 2013.

[6] A. Cuevas, R. Fraiman, and A. Rodriguez-Casal. A nonparametric approach to the estimation of lengths and surface areas. *Ann. Stat.*, 35(3):1031–1051, 2007.

[7] P. Eichelsbacher and C. Thaele. New Berry-Essen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.*, (102), 2014. preprint.

[8] J. H. J. Einmahl and E. V. Khmaladze. The two-sample problem in Rm and measure-valued martingales. *Lecture Notes-Monograph Series*, pages 434–463, 2001.

[9] H. Federer. Curvature measures. *Trans. Am. Math. Soc.*, 93:418–491, 1959.

[10] A. Gloria and J. Nolen. A quantitative central limit theorem for the effective conductance on the discrete torus. arXiv:1410.5734, 2014.

[11] L. Goldstein and M. Penrose. Normal approximation for coverage models over binomial point processes. *Ann. Appl. Probab.*, 20(2):696–721, 2010.

[12] M. Heveling and M. Reitzner. Poisson-voronoi approximation. *Ann. Appl. Probab.*, 19(2):719–736, 2009.

[13] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, 19(3):293–325, 1948.

[14] C. Houdré and Ü. Islak. A central limit theorem for the length of the longest common subsequence in random words. arXiv:1408.1559, 2014.

[15] S. Karlin and Y. Rinott. Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates. *Ann. Stat.*, 10(2):485–501, 1982.

[16] W. S. Kendall and I. Molchanov. *New perspectives in stochastic geometry.* Oxford University Press, 2010.

[17] L. Goldstein L. H. Y. Chen and Q.-M. Shao. *Normal Approximation by Stein's Method.* Springer-Verlag, 2011.

[18] R. Lachièze-Rey and S. Vega. Boundary density and Voronoi approximation of irregular sets. arXiv:1501.04724, 2015.

[19] G. Last, G. Peccati, and M. Schulte. Normal approximation on poisson spaces: Mehler's formula, second order poincaré inequalities and stabilization. arXiv:1401.7568, 2014.

[20] I. Molchanov. *Statistics of the Boolean Model for Practitioners and Mathematicians.* Wiley, 1997.

[21] J. Nolen. Normal approximation for the net flux through a random conductor. arXiv:1406.2186, 2014.

[22] G. Peccati. Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations. *Ann. Prob.*, 32(3A):1796–1829, 2004.

[23] M. Reitzner, Y. Spodarev, and D. Zaporozhets. Set reconstruction by Voronoi cells. *Adv. Appl. Probab.*, 44(4):938–953, 2012.

[24] W. T. Rhee and M. Talagrand. Martingale inequalities and the jackknife estimate of variance. *Stat. Probab. Lett.*, 4:5–6, 1986.

[25] A. Rodriguez-Casal. Set estimation under convexity-type assumptions. *Ann. Inst. H. Poincaré Prob. Stat.*, 43:763–774, 2007.

[26] M. Schulte. A central limit theorem for the poisson-voronoi approximation. *Adv. Appl. Math.*, 49(3-5):285–306, 2012.

[27] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *J. Theor. Prob.*, september 2014.

[28] R. J. Serfling. *Approximation theorems of mathematical statistics.* Wiley, 1980.

[29] R. Vitale. Covariances of symmetric statistics. *J. Multiv. Anal.*, 41:14–26, 1992.

[30] G. Walther. On a generalization of Blaschke's rolling theorem and the smoothing of surfaces. *Math. Meth. Appl. Sci.*, 22:301–316, 1999.

[31] J. E. Yukich. Surface order scaling in stochastic geometry. *Ann. Appl. Probab.*, 25(1):177–210, 2015.