TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 351, Number 11, Pages 4337–4363
S 0002-9947(99)02499-X
Article electronically published on July 21, 1999

LIMIT THEOREMS FOR THE CONVEX HULL OF RANDOM POINTS IN HIGHER DIMENSIONS

IRENE HUETER

Abstract. We give a central limit theorem for the number N_n of vertices of the convex hull of n independent and identically distributed random vectors, being sampled from a certain class of spherically symmetric distributions in $\mathbb{R}^d$ $(d > 1)$, that includes the normal family. Furthermore, we prove that, among these distributions, the variance of N_n exhibits the same order of magnitude as the expectation as $n \to \infty$. The main tools are Poisson approximation of the point process of vertices of the convex hull and (sub/super)-martingales.

1. Introduction

Let $Z_1, Z_2, \dots, Z_n$ be n i.i.d. random vectors in $\mathbb{R}^d$ $(d > 1)$, each with absolutely continuous distribution. Let N_n be the number of vertices of the convex hull of $\{Z_1, Z_2, \dots, Z_n\}$. What is the limiting distribution of N_n as n grows to ∞ ?

Although random convex hulls have attracted an enormous amount of attention in the mathematical literature, there are no results in the higher dimensional space $(d > 2)$ that concern the second moment or the asymptotic distribution of any convex hull functionals. The more difficult part is to derive the exact rate function in n associated with the second moments, whereas it is easy to provide expressions in terms of integrals. In this context of limit laws, even in two dimensions there are only a few results [6, 1, 8, 9]. The classical method, introduced by Rényi and Sulanke (1963) and further developed by Efron (1965), Carnal (1970), Raynaud (1970) and Dwyer (1991) to compute the first moment of convex hull functionals, is purely combinatorical and cannot capture the dependence structure between the multivariate extremes.

In a remarkable paper, Groeneboom (1988) found a powerful method to deal with the convex hull. The method makes extensive use of the one-dimensional process of consecutive vertices of the convex hull to prove a central limit theorem for N_n in the two cases where a set of n points is uniformly distributed on an r-polygon or on an ellipse. Precise results on the variance of N_n are obtained. The basic tools are Poisson approximation of the sample points near the boundary of the convex hull and some suitable martingales for the moment calculations. In [8] the same approach was generalized in various directions. Among others, bivariate

Received by the editors December 2, 1998 and, in revised form, January 22, 1999.
1991 *Mathematics Subject Classification.* Primary 52A22, 60D05.
Key words and phrases. Convex hull, Poisson point process, Markovian jump process, (sub/super)-martingales, central limit theorem, rotationally invariant distributions.

exponentially-tailed, rotationally invariant distributions, especially the normal distribution [9], were studied not only for their number of convex hull vertices, but also for the perimeter and the area of the convex hull.

The purpose of this paper is to develop Groeneboom's technique in higher dimensions for a class of spherically symmetric distributions, whose tails decay at an exponential rate that in a certain sense is not much slower than the one for the normal, in order to establish a central limit theorem for N_n and to investigate the *variance* of N_n. It will turn out that the variance has the same order as the expectation in the considered class. For the distributions left out in our study, it would be interesting to know whether the variance can differ in order from the expectation.

Let F_R be the tail probability of the radial component of Z_1, let L be a monotonely increasing, slowly varying function which satisfies $x = L(1/F_R(x))$ for sufficiently large x, and let the function $\varepsilon(\cdot)$ be given by $L(x) = \exp\{\int_1^x \varepsilon(t)/t\, dt\}$ (see Section 2). Carnal [2] proved

$$\mathbf{E}[N_n] \sim 2(\pi/\varepsilon(n))^{1/2}$$

in the plane, and twenty years later, Dwyer [3] showed, for $d \geq 2$ and sufficiently large n,

$$\mathbf{E}[N_n] \leq \sqrt{d}(8\pi d/(d-1))^{(d-1)/2}\, \varepsilon(n)^{-(d-1)/2}.$$

Here, "$a_n \sim b_n$" means that $a_n/b_n \to 1$ as $n \to \infty$. These asymptotical results tell us that the quicker the distribution tails off, the more convex hull vertices can be expected. The main results of this paper are the following.

Theorem 1. *Let N_n be the number of vertices of the convex hull of a sample of size n from a rotationally invariant, exponentially-tailed distribution such that the smoothness conditions in* (6) *below are satisfied and $L(n)\varepsilon(n)^{1/2} \not\to \infty$, as $n \to \infty$. Then there are two positive and finite constants c_1 and c_2 so that, as $n \to \infty$,*

$$\frac{N_n - c_1\, \varepsilon(n)^{-(d-1)/2}}{(\mathbf{Var}[N_n])^{1/2}} \xrightarrow{L} Z \sim \mathcal{N}(0,1), \tag{1}$$

where

$$\mathbf{Var}[N_n] \quad \sim \quad c_2 \varepsilon(n)^{-(d-1)/2}.$$

Here "$\xrightarrow{L}$" indicates convergence in distribution.

We wish to mention the special case of the normal distribution, where the necessary moment computations are more accessible, and thus, an upper bound for the asymptotic constant for $\mathbf{E}[N_n]$ can be established. Since N_n is affine invariant, the result is valid for any normal distribution.

Theorem 2. *Let N_n be the number of vertices of the convex hull of a sample of size n from any normal distribution in $\mathbb{R}^d$. Then, as $n \to \infty$,*

$$\frac{N_n - c_3(\ln n)^{(d-1)/2}}{\{c_4(\ln n)^{(d-1)/2}\}^{1/2}} \xrightarrow{L} Z \sim \mathcal{N}(0,1), \tag{2}$$

where $0 < c_3 \leq 2\sqrt{d-1}\,(2\pi)^{(d-1)/2}/\Gamma(d/2)$ and $0 < c_4 < \infty$.

Raynaud [12] gave an estimate for the first moment of the number of facets F_n, which implies an upper bound for $\mathbf{E}[N_n]$,

$$\mathbf{E}[N_n] \leq \mathbf{E}[F_n] \sim 2^d\, d^{-1/2}\,(d-1)^{-1}\,\pi^{(d-1)/2}\,(\ln n)^{(d-1)/2}.$$

Our upper bound is an improvement for large values of the dimension d. The inequality $\mathbf{E}[N_n] \leq \mathbf{E}[F_n]$ stems from the fact that, in higher dimensions, the convex hull is *simplicial* with probability one, but *not* simple. In other words, each vertex may be incident to more than d edges of the convex hull. Note that for $d = 2$ our bound as well as Raynaud's estimate takes the exact value for $\mathbf{E}[N_n]$.

The step from two to d ($d > 2$) dimensions is not straightforward. The jump measure plays a key role to our analysis of the moments. However, the derivation of the jump measure of the process of the vertices is a delicate issue, which reflects upon the fact that, in more than two dimensions, the natural order of the parameter space of the process visiting the convex hull vertices is lost. Therefore, the bulk of this paper is devoted to defining a way to follow a typical path on the surface of the convex hull such that "most" vertices are recorded and the essence of the structure of dependence is retained. Our strategy is to find some stationary jump process, to partition the surface of the convex hull into identical elements, and to count the number of jumps related to each element.

Since questions on random convex hulls, especially on Central Limit Theorems in higher dimensions, fall in the class of problems that are easy to state but rather hard to solve, and their solutions are of great interest in a variety of optimization problems, the following curious question deserves further investigation: Can the Stein-Chen method be applied and is it more effective ?

The remainder of this paper is organized as follows. Section 2 states some well-known properties of spherically symmetric distributions. Section 3 defines and discusses the process of vertices of the convex hull and the role of Poisson approximation in this context. In Section 4, we shall locally arrange the points of the process of vertices in order and compare the process of vertices to two jump processes depending on one parameter only. In Section 5, we shall study the behaviour of the local jump measure from a vertex point and derive some results that will eventually relate the moments of N_n with the jump measure. The final section contains the proofs of the main results.

2. Spherically symmetric distributions with exponential tails

In this section, we collect some well-known features of our distribution family that will play a prominent role in the estimation of the moments of N_n. Recall that $Z_1, Z_2, \ldots, Z_n$ are i.i.d. random vectors, each from an absolutely continuous distribution F. Denote the tail probability of the radial component by $F_R(x) = P(\|Z_1\| > x)$. Consider a certain class of "exponentially-tailed" distributions with sufficiently fast tails in the sense that their distribution $F_R(x)$ satisfies

$$x = L(\frac{1}{F_R(x)}) \tag{3}$$

for a monotonely increasing function L that *varies slowly* at infinity, e.g. for each $\lambda > 0$, $\lim_{x\to\infty} L(\lambda x)/L(x) = 1$. Since $F_R(0) = 1$ and $F_R(\infty) = 0$, we have $L(1) = 0$ and $L(\infty) = \infty$. A slowly varying function $L(x)$ can be expressed in the form

$$L(x) = a(x)\,\exp\{\int_1^x \varepsilon(t)/t\,dt\}$$

with $a \to a_0 \notin \{0, \infty\}$ and $\varepsilon \to 0$, as $x \to \infty$ (see i.e. Feller [5]). We shall assume $a \equiv 1$ for all that follows. For instance, the d-dimensional normal distribution has

$L(s) \sim \sqrt{2 \ln s}$. If we let $s = 1/F_R(x)$, we have

$$x = L(s) = \exp\{\int_1^s \varepsilon(t)/t \, dt\}. \tag{4}$$

Furthermore, we may define

$$0 < \nu(u) = \varepsilon(L^{-1}(u)) = \varepsilon(1/F_R(u)), \tag{5}$$

which $\to 0$, as $u \to \infty$. The elementary correspondence $dx/ds = x\,\varepsilon(s)/s \iff ds/s = dx/(\nu(x)\,x)$ implies the representation

$$F_R(x) = \exp(\ln F_R(x)) = \exp\{-\int_0^x du/(\nu(u)\,u)\}.$$

Proceeding along the lines of Carnal [2] and Dwyer [3], we impose the following *smoothness conditions* on ν and thus, on L, that are satisfied by most distributions in question, and force the function ε to be slowly varying:

$$\begin{aligned} &\text{(i)} && \nu(x) \text{ is monotone (decreasing) for large } x, \\ &\text{(ii)} && x \cdot \nu'(x) \cdot \ln(\nu(x)) = o(1) \text{ as } x \to \infty, \\ &\text{(iii)} && \nu(x) \cdot \ln(x) = o(1) \text{ as } x \to \infty. \end{aligned} \tag{6}$$

Examples. (a) The normal distribution with density function

$$(2\pi)^{-d/2} \exp(-\|x\|^2/2)$$

has $\nu(r) \sim r^{-2}$. It is worthwhile noting that in this case $\nu(L(n))^{1/2}\ L(n) = 1$. As we will see later on, the behaviour of $r\nu(r)^{1/2}$ for $r = L(n)$ will be of importance. In general, we find $\varepsilon(n) \neq \mathcal{O}(L(n)^{-2})$.

(b) Take the example $F_R(r) = c_0 \exp\{-r^k\}$ for $k > 0$. For large n, we obtain $L(n) \sim (\ln n)^{1/k}$ and $\varepsilon(n) \sim 1/(k \ln n)$. That means $L(n)\sqrt{\varepsilon(n)} \to 0$ iff $k > 2$, and $L(n)\sqrt{\varepsilon(n)} \to \infty$ iff $k < 2$. Thus, the case $k \geq 2$ is covered by Theorem 1.

These examples hint towards the fact that, roughly speaking, the smaller the function $\nu(\cdot)$ around the value $L(n)$, the thinner is the tail of the distribution, and thus, the more points are expected to span the convex hull.

3. Vertex process

It is natural to examine the jump process which visits precisely the vertices of the convex hull since these harbour the complete information about the convex hull. This process, however, is itself too complex to deal with and requires a number of simplifications, as we will see shortly. As the number of convex hull vertices is invariant under affine transformations, we may as well shift the whole sample. Indeed, for the rest of the paper, it will be advantageous to think of the distribution F being centered at $(0, 0, \ldots, r_2)$, where r_2 is the radius of a ball that contains a sample of n points from F with a probability tending to one. Let us introduce the process which runs through the vertices of the convex hull.

Definition 1. For each $\mathbf{a} = (a_1, a_2, \ldots, a_{d-1}) \in \mathbb{R}^{d-1}$ we define the "vertex process" $W_n(\mathbf{a}) = W_n(a_1, a_2, \ldots, a_{d-1}) = (X_1(\mathbf{a}), X_2(\mathbf{a}), \ldots, X_d(\mathbf{a}))$ as the point $(U_1, U_2, \ldots, U_d)$ of the sample such that $U_d - \sum_{i=1}^{d-1} a_i U_i$ is minimal. If there are several such points, then we take the one with the biggest first coordinate. This happens with probability zero for fixed $\mathbf{a}$.

Occasionally, the hyperplane just described in the definition will be called the "supporting hyperplane" of the convex hull and will be referred to as $H_{\mathbf{a}}$ in the sequel. As the vector $\mathbf{a}$ runs through all $\mathbb{R}^{d-1}$, roughly one half of all convex hull vertices is counted. The process $\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$ is a pure jump process, non-Markovian and has "right"-continuous paths. However, the process W_n is close to being Markovian in the sense that there is a process W endowed with the preferable *Markov property* such that the variational distance between W_n and W tends to zero, as $n \to \infty$. The question is "how can we find a Poisson point process such that the extremes of a realization look similar to the extremes of the original sample $Z_1, Z_2, \ldots, Z_n$?"

For this purpose, we determine the region A_n^* close to the boundary of the convex hull where "most" of the convex hull vertices fall. In other words, we are looking for a region $A_n^* \subset \mathbb{R}^d$ such that, as $n \to \infty$, the following two conditions hold:

$$\mathbf{P}(W_n(\mathbf{a}) \in A_n^*,\ \forall \mathbf{a} \in \mathbb{R}^{d-1}) \longrightarrow 1 \tag{7}$$

and

$$\mathbf{P}(Z_1 \in A_n^*) \longrightarrow 0. \tag{8}$$

By the rotational invariance of the distribution, the set A_n^* is an annulus

$$A_n^* = \{(u_1, u_2, \ldots, u_d) \in \mathbb{R}^d : r_1^2 \leq \sum_{i=1}^{d-1} u_i^2 + (r_2 - u_d)^2 \leq r_2^2\}, \tag{9}$$

with r_1 and r_2 chosen as

$$\begin{aligned} r_0 &= L(n), \\ r_1 &= r_0 - \varepsilon_n/2, \\ r_2 &= r_0 + \varepsilon_n/2, \end{aligned} \tag{10}$$

and ε_n is such that

$$\begin{aligned} nF_R(r_1) &= \gamma_1, \\ nF_R(r_2) &= 1/\gamma_2, \end{aligned} \tag{11}$$

where $\gamma_1 = \gamma_1(n)$ and $\gamma_2 = \gamma_2(n)$ are two functions of n that are dominated by the function n and both increase to ∞ as $n \to \infty$. Any such function will do for γ_2, whereas the function γ_1 is chosen sufficiently large by suitable tuning, as explained in the proof of Lemma 3 below. Hence, γ_2 is determined by γ_1 and the coupling equations in (10). If we set $x = r_0 = L(n)$ in (3), we recover the identities

$$\begin{aligned} nF_R(r_0) &= 1, \\ \varepsilon(n) &= \nu(r_0). \end{aligned} \tag{12}$$

Lemma 3. *The region A_n^* in* (9) *satisfies* (7) *and* (8).

Proof. First, $\mathbf{P}(Z_1 \in A_n^*) < F_R(r_1) = \gamma_1/n$, which tends to zero for every γ_1 that is dominated by the function n. Consider the two balls $B(r_1)$ and $B(r_2)$ of radii r_1 and r_2, centered at $(0, 0, \ldots, r_1)$ and $(0, 0, \ldots, r_2)$, respectively, where r_1 and r_2 are given in (10). Thus, $A_n^* = B(r_2) \setminus B(r_1)$. To verify that, with a probability that tends to 1, the ball $B(r_2)$ is large enough to contain $Z_1, Z_2, \ldots, Z_n$ is the easier part. In fact, the probability that all points are contained in the ball $B(r_2)$ is $\sim \exp\{-nF_R(r_2)\} = \exp\{-1/\gamma_2\}$, which tends to one, as $n \to \infty$, because $1/\gamma_2 \to 0$.

Now, we turn to show the trickier part, that there is a function $\gamma_1 = \gamma_1(n)$ that is dominated by the function n so that, with a probability that tends to 1, the vertices of the convex hull lie outside the ball $B(r_1)$. Suppose that, with probability p, at most $h(n)$ vertices span the convex hull of $Z_1, Z_2, \dots, Z_n$. Consider the event $E(r_1)$ that there are exactly $h(n)$ points outside $B(r_1)$ (and $n - h(n)$ points inside $B(r_1)$). The probability of $E(r_1)$ is given by

$$\mathbf{P}(E(r_1)) = [1 - F_R(r_1)]^{n-h(n)} F_R(r_1)^{h(n)} \binom{n}{h(n)}.$$

A sequence of straightforward calculus steps, Stirling's formula for large factorials, and $nF_R(r_1) = \gamma_1$ yield, for sufficiently large n,

$$\mathbf{P}(E(r_1)) \quad \sim \quad (2\pi)^{-1/2} \exp\{-\gamma_1 + h(n)[\gamma_1/n + \ln(\gamma_1 e/h(n))]\}.$$

The right-hand side of the last display is $\sim (2\pi)^{-1/2}$ if and only if $\gamma_1 \sim h(n)$, is positive if γ_1 is proportional to $h(n)$, and tends to zero otherwise. Thus, the same calculation, with $B(r_2)$ in place of $B(r_1)$, yields a probability of $E(r_2)$ that tends to zero. Substituting radius r in the previous event gives a probability $\mathbf{P}(E(r))$ that is a continuous function in the radius r that is nonincreasing above some positive critical value. Moreover, the probability that $Z_1, Z_2, \dots, Z_n$ are all contained in $B(r_0)$ is strictly positive. These together imply that there is a largest radius $r_* = r_*(h(n))$ so that the probability of $E(r_*)$ is nonzero. Let $nF_R(r_*) = \gamma_*$. We conclude that, for every $\gamma_1 > \gamma_*$, there is positive probability that at most $h(n)$ points lie outside $B(r_1)$; equivalently, there is positive conditional probability, given that there are at most $h(n)$ convex hull vertices, that the boundary of the convex hull lies outside $B(r_1)$. Since this argument is valid for every probability p, our statement follows. □

Note that there is a lot of freedom for the choice of ε_n, and thus, γ_1, and γ_2. (The reader should be aware that the function $\varepsilon(n)$ is *not* the same as the function ε_n before confusion arises.) In the proof of Corollary 11 together with the relation to X', as defined in (15), it will turn out that ε_n can be chosen such that $\varepsilon_n \to 0$ as $n \to \infty$. Hence, we can conclude that, as $n \to \infty$, with high probability, the boundary of the convex hull is contained between two spheres such that the difference between their radii shrinks to zero. Thus, the distribution of the random points indeed decays rather quickly, no slower than for the normal distribution.

Let η_n denote the sample point process of size n and let ξ_n denote the Poisson point process on $\mathbb{R}^d$ with intensity measure $n \int_\cdot dF$, which is an *inhomogeneous* Poisson point process.

Lemma 4. *There exist processes $\tilde{\eta}_n$ and $\tilde{\xi}_n$, defined on the same probability space, such that*

$$\tilde{\eta}_n \stackrel{L}{=} \eta_{n|_{A_n^*}}, \qquad \tilde{\xi}_n \stackrel{L}{=} \xi_{n|_{A_n^*}} \tag{13}$$

and

$$\mathbf{P}(\tilde{\eta}_n \neq \tilde{\xi}_n) \leq 2\mathbf{P}(Z_1 \in A_n^*) \to 0. \tag{14}$$

Here "$|$" stands for "restriction to", and "$\stackrel{L}{=}$" denotes equality in distribution. The proof is standard and omitted here. It employs a comparison via coupling between the Poisson distribution with parameter $n\int_{A_n^*} dF$ and the binomial distribution with parameters n and $n\int_{A_n^*} dF$ (see [6], Lemma 2.2 for details). We remark on the side that there exist rotationally invariant distributions such that no such region A_n^* exists, and therefore, Poisson approximation is not possible, for instance, the *algebraically tailed* distributions. Curiously, along with those goes an $E[N_n]$ that is constant for large n (see Carnal [2], Dwyer [3], and also Aldous et al. [1]).

We may define the vertex process $\{W(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$ of a realization of the Poisson point process ξ_n by applying Definition 1 to this realization of the Poisson point process. Notice that the process W still depends on n, although we drop the index to simplify matters and to distinguish W from W_n.

Markov property. It may be worthwhile to pause for a while and to discuss the Markov property to some extent. First, the original vertex process based on the sample lacks the Markov property because whether, at some time and point, the process W_n jumps to some point in a subset $B \subset \mathbb{R}^d$ depends on every sample point (but *not only* on the actual point) and on B. The process W, though, will exhibit the Markov property once its parameter space is reduced to a one-dimensional parameter space (described in Definition 2) as an immediate consequence of the stochastic independence of random counting measures of disjoint sets.

Note that the requirement $\mathbf{P}(Z_1 \in A_n^*) \to 0$ and Lemma 4 force that the variational distance between the vertex processes W_n and W tends to zero with increasing n. The upper bound for the variational distance depends on the distribution (being of order n^{-1} for the normal).

4. Local approximations of the vertex process

The purpose of this section is to find the "best" mapping of the $(d-1)$-dimensional parameter space of W onto a one-dimensional subspace of $\mathbb{R}$. Indeed, in order that the jump measure of the vertex process can be defined, the time range must exhibit an *order*. The basic idea is to follow a path that "visits each neighbourhood" of the surface of the convex hull exactly once and, in such a way that the path is as "straight" as possible. If the path ran on a line, the two-dimensional situation with one-dimensional time parameter would be perfectly imitated. However, since the convex hull vertices fall in lines with probability zero, the path will wobble slightly. The consequence thereof is that the jump measure cannot be determined precisely. We will find bounds. In fact, the farther away from the path the vertices are, which is tantamount to the less convex hull vertices there are, the less accurate the estimate will be for the jump measure, and thus, for the moments of N_n.

On the other hand, the requirement to visit each neighbourhood exactly once is not such big a difficulty as it might appear at first glance. Whenever there are not too few convex hull vertices (this would be the case if Poisson approximation fails, which already was excluded), by the rotational invariance, it suffices to study the convex hull only on a small region and to think of patching together all such small regions in order to reach conclusions about N_n.

Hence, the first step will be to define such a typical small region for the local investigations, and the second step will be to redefine the vertex process with one-dimensional parameter space. By the rotational invariance, in order to discuss the jump behaviour of the process W on a small region, we may as well look at the jump measure from a point of W chosen from a small sector $S_n \subset A_n^*$ around the origin. On a large scale, on S_n the boundary of A_n^* appears to be close to the graph of a paraboloid, which is easier to deal with than the boundary of the annulus A_n^*. Lemma 5 below will show that in variational distance the vertex process, defined by locally approximating A_n^* by a paraboloid, is sufficiently close to the vertex process W on A_n^*. To this end, write

$$X' = \sqrt{2r_2\varepsilon_n} \tag{15}$$

and define

$$S_n = \{(u_1, u_2, \dots, u_d) \in A_n^* : \sum_{i=1}^{d-1} u_i^2 \le (X')^2\}. \tag{16}$$

In the neighbourhood of the origin of the coordinate system, we will use the paraboloid

$$u_d = (\sum_{i=1}^{d-1} u_i^2)/2r_2 \approx r_2 - (r_2^2 - \sum_{i=1}^{d-1} u_i^2)^{1/2},$$

being a local approximation to the outer boundary $\{(u_1, \dots, u_d) \in \mathbb{R}^d : \sum_{i=1}^{d-1} u_i^2 + (r_2 - u_d)^2 = r_2^2\}$ of A_n^*. Additionally, define

$$B_n = \{(u_1, \dots, u_d) \in \mathbb{R}^d : (\sum_{i=1}^{d-1} u_i^2)/2r_2 \le u_d,\ \sum_{i=1}^{d-1} u_i^2 \le (X')^2\}. \tag{17}$$

Let $\xi_{n|_{S_n}}$ and $\xi_{n|_{B_n}}$ be the restrictions of the Poisson point process with intensity $n\int_\cdot dF$ to S_n and B_n, respectively. Let W^S and W^B be the vertex processes that are defined as in Definition 1, with the samples being replaced by realizations of $\xi_n|_{S_n}$ and $\xi_n|_{B_n}$, respectively.

Lemma 5. *Let the point processes δ_n^S and δ_n^B be defined by*

$$\begin{aligned} \delta_n^S &= \{W^S(\mathbf{a}) : 2r_2\|\mathbf{a}\| \le \sqrt{2r_2\varepsilon_n}\}, \\ \delta_n^B &= \{W^B(\mathbf{a}) : 2r_2\|\mathbf{a}\| \le \sqrt{2r_2\varepsilon_n}\}. \end{aligned} \tag{18}$$

Then, as $n \to \infty$,

$$\mathbf{P}(\delta_n^S \ne \delta_n^B) \le nF_R(r_2)(2\varepsilon_n/r_2)^{(d-1)/2}.$$

Proof. The probability that we lose any vertices of the convex hull while resorting to the paraboloid approximation tends to zero as $n \to \infty$. In fact, the paraboloid $u_d = (\sum_{i=1}^{d-1} u_i^2)/2r_2$ runs below the boundary of the d-ball of radius r_2 centered at $(0, 0, \dots, r_2)$. Therefore,

$$\mathbf{P}((S_n \setminus B_n) \cup (B_n \setminus S_n)) \le F_R(r_2)(2r_2\varepsilon_n)^{(d-1)/2}/r_2^{(d-1)/2}, \tag{19}$$

and by (11),

$$n\,\mathbf{P}((S_n \setminus B_n) \cup (B_n \setminus S_n)) \le (1/\gamma_2)(2\varepsilon_n/r_2)^{(d-1)/2}$$

tends to zero because $\gamma_2 \to \infty$. □

The scaling of the parameters in transformation (40) along with Lemma 5 will imply that the error, introduced by that approximation, for the *entire* boundary of A_n^*, is of order $1/\gamma_2 \to 0$.

Note. In the previous proof, we rely on the following argument. For small $(a_1, \dots, a_{d-1})$, i.e. for $\mathbf{a} = (a_1, \dots, a_{d-1})$ such that $\|\mathbf{a}\| \le X'/r_2$, the slope parameter a_i of $W(a_1, \dots, a_{d-1}) = (y_1, \dots, y_d)$ can be looked at as the angle between the vectors $(0, \dots, 0, r_2)$ and $(\dots, 0, y_i, 0, \dots, 0, r_2)$. Consequently, in a neighbourhood of the origin in $\mathbb{R}^d$, for each coordinate y_i $(1 \le i \le d-1)$ of the process W, the ratio y_i/r_2 grows proportionally to the corresponding slope parameter a_i. We will rely on this "linearity argument" when defining the vertex process W^π below (see Definition 2), and later on, when determining the range of the parameter space that brings relevant contribution to $\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$.

For all that follows, we restrict attention to W on the region $A_n^* \cap B_n$. Now there obviously are thousands of ways to arrange in order the points of the process $\{W(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$. Beware that the reduction to the one-dimensional parameter space brings along the loss of information that the $(d-2)$ ignored parameters contain. When trying to determine the jump measure, we are faced with the problem that it makes a crucial difference whether one of the thrown away coordinates (or more than one) is relatively small or big, especially whenever the first coordinate is small. This makes clear that we must content with some bounds. A suitable definition of ordering the vertices will rule out the situation where any of the remaining $(d-2)$ parameter values is bigger than the first one and allow for bounds of the jump measure.

To specify the order in which the convex hull vertices in $A_n^* \cap B_n$ are counted, we will rely on the following construction. Project the points $\{W(\mathbf{a}) \in B_n : \mathbf{a} \in \mathbb{R}^{d-1}\}$ to the hyperplane $\mathbb{R}^{d-1} \times \{0\}$ to eliminate curvature, and scan through the projected region in a specified direction by a geometric object whose sides all have curvature zero. (Observe that the direction of scanning does not matter.) Each time one of the projected vertex points is hit by this object, we decide to see the point and mark the current time on the (real) time axis. This happens precisely once for each projected point.

Hence, let B_n^π be the projection of $A_n^* \cap B_n$ to the hyperplane $\mathbb{R}^{d-1} \times \{0\}$. It is natural to scan through B_n^π by a cone-like object to collect the vertices projected onto B_n^π. So, we may define the following half-closed convex sets in $\mathbb{R}^{d-1}$:

$$C(0) = \{(u_1, u_2, \dots, u_{d-1}) \in \mathbb{R}^{d-1} : 0 \le u_1, 0 \le |u_k| \le u_1, \forall\, 2 \le k \le d-1\},$$
$$C(a) = C(0) + (a, 0, \dots, 0). \tag{20}$$

Denote the complement of $C(a)$ by $C^c(a)$. For $d = 3$, the set $C(a)$ corresponds to an infinite triangle with a top at $(0,0)$. Let $dist(\mathbf{x}, A) = \min\{\|\mathbf{x} - \mathbf{y}\| : \text{ with } \mathbf{y} \in A\}$ denote the distance between a subset $A \subset \mathbb{R}^d$ and $\mathbf{x} \in \mathbb{R}^d$. Furthermore, let $\{W(\mathbf{a}) \in B_n : \mathbf{a} \in \mathbb{R}^{d-1}\} = \{\tilde{Z}_1, \tilde{Z}_2, \tilde{Z}_3, \dots\}$ denote those points of the Poisson point process $\xi_{n|B_n}$ that are vertices of the convex hull of the entire realization of the Poisson point process $\xi_{n|B_n}$. By a continuity argument, it is seen that the points $\tilde{Z}_1, \tilde{Z}_2, \tilde{Z}_3, \dots$ can be thought of as being realized by "thinning" the Poisson point process $\xi_{n|B_n}$ by a collection of i.i.d. Bernoulli(p) random variables with appropriate $p \in (0,1)$ (such a p can be found due to the rotational invariance of the underlying distribution F), equivalently, as being realized by another Poisson point process $\zeta_{|B_n}$ such that the variables $\#\{W(\mathbf{a}) \in B_n : \mathbf{a} \in \mathbb{R}^{d-1}\}$ and $\#\{\tilde{Z}_i : i \ge 1\}$ have the same distribution. Thus, the Poisson point process $\zeta_{|B_n}$ generates the vertices of the convex hull of $\xi_{n|B_n}$.

After this passage of motivation, we are in a position to give a suitable version of a *local* vertex process, depending on only *one* parameter.

Definition 2. Let $\tilde{Z}_1, \tilde{Z}_2, \ldots$ denote a realization of a Poisson point process as described above. For each $1 \leq i \leq n$, let $Z_i^* = \pi_{\mathbb{R}^{d-1}\times\{0\}}(\tilde{Z}_i)$ be the projection of $\tilde{Z}_i$ onto $\mathbb{R}^{d-1} \times \{0\}$ and let $\tilde{Z}_i = Z_i^* \times V_i$. Finally, for each $c \in \mathbb{R}$, let $V^\pi(c)$ be the point Z_k^* of $Z_1^*, Z_2^*, \ldots$ such that $V^\pi(c) \subset C^c(c)$, and $dist(V^\pi(c), C(c))$ is minimal. (When $V^\pi(c)$ gets hit, $dist(V^\pi(c), C(c)) = 0$.) If there are several such points, take the one with the smallest first coordinate. This happens with probability zero for fixed c. Then call $W^\pi(c) = (V^\pi(c), V_k)$ the "local vertex process".

Note that in the previous definition the objects $C(a)$ do not hit any points outside B_n^π, and in view of the absolute continuity of F, each $C(a)$ hits two distinct points Z_j^* and Z_i^* *simultaneously* with probability zero. In fact, $C(a)$ detects a point $\tilde{Z}_i$ among the projected points if $\tilde{Z}_i$ lies in the union of the corresponding two hyperplanes in $\mathbb{R}^d$ that cross $\mathbb{R}^{d-1} \times \{0\}$ in $C(a)$. That the local vertex process $\{W^\pi(c) : c \in \mathbb{R}\}$ has the *Markov property* as a one-parameter process is an immediate consequence of the stochastic independence of events of any Poisson point process on independent regions, which can be seen as follows. First observe that all projected points are projections of a realization of a Poisson point process, and this projection is a one-to-one map, in fact, a homeomorphism. Because the curvature of the boundary of $A_n^* \cap B_n$ is negligibly small and the sets $C(a_1)$ are strictly nested (that is, $a_1 < a_2$ implies $C(a_2) \subset C(a_1)$), it follows that the position of the process at a future time point depends only on events of the Poisson point process on regions that are independent of all events in the past, and whence the stochastic independence of events of the Poisson process on independent regions applies. In particular, it is guaranteed that, given $W^\pi(a) = x$ and $W^\pi(a+h) = x$ for any $a \in \mathbb{R}$, $h > 0$ and some $x \in \mathbb{R}^d$, with probability one, for every $0 \leq s \leq h$, $W^\pi(a+s) = x$.

For the rest of the paper, this local Markov property will just be referred to as the Markov property. For $d = 2$, we scan along the first coordinate axis. In fact in this case, it is not necessary to project the points but a line is turned along the approximating parabola at the origin. In a small neighbourhood of the origin, for the distributions at hand, the two procedures are essentially equivalent in the sense that the error entailed by the projection is negligible. Observe that the values of the vertex points are unchanged but the parameter space was transformed. The process W^π is right-continuous. When walking along our path, we see the vertices in the d-dimensional region $A_n^* \cap B_n$ and follow each vertex point. Thus, the path is not straight. To scan B_n^π, the parameter a of $W^\pi(a)$ runs through $[-X'/r_2, X'/r_2)$ and the region B_n^π might be enlarged appropriately by the possibly neglected region $C^c(0) \cap B_n^\pi$. But for simplicity, we wish to think of B_n as a ball. The shape of B_n will turn out to be irrelevant later. Note that the above definition of the vertex process W^π is understood to make sense only for a sufficiently small slope parameter. Otherwise, with a probability approaching one, the convex hull vertices will come from $A_n^* \setminus B_n$.

5. Jump measure

We proceed to derive an estimate for the local jump measure of the vertex process W^π and to establish two pairs of expressions, one pair in terms of supermartingales that will provide upper bounds for the first two moments of the number of vertices in B_n, the other pair in terms of submartingales that will provide the corresponding lower bounds. Throughout this section, we will employ the paraboloid approximation and assume, if not stated otherwise, that the time parameter $a \in \mathbb{R}$ of the vertex process is small enough that $W^\pi(a)$ is in B_n with high probability. We shall count all convex hull vertices within a $(d-1)$-dimensional "tubular neighbourhood" that envelopes the path (its transversal intersections are cubes). By the rotational invariance of the distribution of the sample points, for the path we may choose any direction tangential to the surface of the d-ball of radius r_2 centered at $(0, \ldots, 0, r_2)$, in particular, the path that projects down to an interval of the x_1-axis.

For each $a > 0$ and $b > a$, define the σ-algebra

$$\mathcal{F}_{a,b} = \sigma\{W^\pi(c) : a \le c \le b\}$$

and the set

$$S_0 = \{(u_1, u_2, \ldots, u_d) \in \mathbb{R}^d : u_d \ge (u_1^2 + u_2^2 + \ldots + u_{d-1}^2)/2r_2\}. \tag{21}$$

Let f denote the joint density of $Z_1, Z_2, \ldots$. Let $C_f = C_f(S_0)$ be the set of continuous functions $g : S_0 \to \mathbb{R}$ with compact support contained in S_0 and with the same "set of increase" and "decrease" as f; in other words, $f(\mathbf{v}_1) \le f(\mathbf{v}_2)$ if and only if $g(\mathbf{v}_1) \le g(\mathbf{v}_2)$ for any pair $(\mathbf{v}_1, \mathbf{v}_2)$ (for instance, constant functions belong to C_f). Observe that $f(\mathbf{v}_1) = f(\mathbf{v}_2)$ if and only if $g(\mathbf{v}_1) = g(\mathbf{v}_2)$. We shall determine upper and lower bounds of the jump measure of W^π with respect to the class C_f by comparing the infinitesimal generator to that of two one-parameter processes, one having larger infinitesimal generator, the other having smaller infinitesimal generator, respectively. For each $a > 0$ and $(x_1, x_2, \ldots, x_{d-1}, y) \in S_0$, define the linear operators $L_a^l, L_a^u : C_f \to C_f$ by

$$\begin{aligned}
[L_a^l g](x_1, x_2, \ldots, y) = {} & n(2X')^{(d-2)} \int_0^{X_0 - x_1} u \\
& \cdot f(x_1 + u, x_2, \ldots, x_{d-1}, r_2 - y - au) \\
& \cdot [g(x_1 + u, x_2, \ldots, x_{d-1}, y + au) \\
& - g(x_1, x_2, \ldots, x_{d-1}, y)]\, du,
\end{aligned} \tag{22}$$

$$\begin{aligned}
[L_a^u g](x_1, x_2, \ldots, y) = {} & n(2X')^{(d-2)}\sqrt{d-1} \int_0^{X_0 - x_1} u \\
& \cdot f(x_1 + u, x_2, \ldots, x_{d-1}, r_2 - y - \sqrt{d-1}au) \\
& \cdot [g(x_1 + u, x_2, \ldots, x_{d-1}, y + \sqrt{d-1}au) \\
& - g(x_1, x_2, \ldots, x_{d-1}, y)]\, du,
\end{aligned} \tag{23}$$

where X_0 is the one intersection of the approximating paraboloid $v = \sum_{i=1}^{d-1} u_i^2/2r_2$ with the line of slope $(a, 0, \ldots, 0)$ through $(x_1, x_2, \ldots, y)$ that is larger in norm.

Proposition 6. *For each $g \in C_f$ and $b_1 > 0$, the process*

$$\{g(W^\pi(b_2)) - \int_{b_1}^{b_2} [L_c^u g](W^\pi(c))\, dc\}_{b_2 \ge b_1}$$

is a supermartingale with respect to the filtration $\{\mathcal{F}_{b_1,b_2} : b_2 \ge b_1\}$, and the process

$$\{g(W^{\pi}(b_2)) - \int_{b_1}^{b_2} [L^l_c g](W^{\pi}(c))\, dc\}_{b_2 \geq b_1}$$

is a submartingale with respect to the filtration $\{\mathcal{F}_{b_1,b_2} : b_2 \geq b_1\}$.

Proof. It suffices to show that, for each $a > 0$ and $\mathbf{w} = (x_1, x_2, \dots, x_{d-1}, y) \in B_n$, the following inequalities hold:

(24)

$$[L^l_{a_1} g](\mathbf{w}) \leq \lim_{h \to 0} \frac{1}{h} \mathbf{E}\{g(W^{\pi}(a_1 + h)) - g(\mathbf{w}) | W^{\pi}(a_1) = \mathbf{w}\} \leq [L^u_{a_1} g](\mathbf{w}).$$

Since relation (24) is important for the conclusions about the jump behaviour of the vertex process, the arguments will be developed in detail for the first statement, that involves the operator L^u_c. Our strategy is to compare the jump behaviour of $W^{\pi}(a_1)$ to that of a process moving *exactly* along the x_1-axis. (By the rotational invariance, any path may be selected, in particular, the one along an axis.) We need to specify (A) the path followed, and (B) the density f at the new point $W^{\pi}(a_1+h)$. Recall from the definition of W^{π} that the path moves within a tubular neighbourhood along the x_1-axis, and the convex hull vertices (of the Poisson point process) projected to B^{π}_n are detected in a certain order that is determined by the shape of $C(a)$; that is, at time a, the point $W^{\pi}(a)$ is seen. We shall derive bounds to (A) and (B).

For each $(x_1, x_2, \dots, y) \in B_n$, we may define the region

$$\begin{aligned} A_h \quad = \quad & \{(u_1, \dots, u_d) \in \mathbb{R}^d : a_1 u_1 + (y - a_1 x_1) \leq u_d \leq (a_1 + h) u_1 + (y - a_1 x_1), \\ & |u_j| \leq X',\, 2 \leq j \leq d-1\}, \end{aligned}$$

where the value of $X' = \sqrt{2 r_2 \varepsilon}$ will be chosen sufficiently large so that with high probability the jump length of the process W^{π} is less than X'. (For a more precise statement about X', see the proof of Corollary 11.) A property enjoyed by the Poisson point process implies that the probability of finding more than one point of the Poisson process ξ_n in A_h is $o(h)$.

Next, let $W^{\pi}(a_1) = (x_1, \dots, x_{d-1}, y)$ and assume that a jump occurs at time $a_1 + h$ so that $W^{\pi}(a_1 + h) = (V_1, V_2, \dots, V_d)$. Instead of moving to this new point, which might lie away from the path along the x_1-axis, choose a point $(V_1, x_2, \dots, x_{d-1}, \tilde{V}_d)$ so that $f(V_1, V_2, \dots, V_{d-1}, r_2 - V_d) = f(V_1, x_2, \dots, x_{d-1}, r_2 - \tilde{V}_d)$. To bound this value of f from above, we need to find an upper bound on $V_d = y + \sum_{i=1}^{d-1} a_i u_i$. The requirement $a_k \leq a_1$ in Definition 2 is crucial here. It follows that $u_k \leq u_1$ since the coordinates of the vertex process grow proportionally to the parameters a_k (e.g. see equations (40)). Due to the reduction of the parameter space to the one-dimensional parameter space (of the process W^{π}), the value of $(a_2, \dots, a_{d-1})$ is not known. Thus, to achieve an upper (or lower) bound for the value of f, we may maximize (or minimize) the sum $\sum_{i=1}^{d-1} a_i u_i$ over all $(a_1, a_2, \dots, a_{d-1})$ under the constraints $a_k \leq a_1$ for $k \neq 1$, $\sum_{i=1}^{d-1} u_i^2 \leq (X')^2$ and each $u_i \geq 0$. An elementary exercise shows that the maximum is achieved when all the terms $u_k a_k$ are the same, equivalently by $a_k \leq a_1$, equal to $u_1 a_1 / \sqrt{d-1}$. Another elementary consideration yields the lower bound. Hence,

$$a_1 u_1 \leq \sum_{i=1}^{d-1} a_i u_i \leq \sqrt{d-1}\, a_1 u_1. \tag{25}$$

Equivalently,

$$
\begin{aligned}
(26)\qquad & f(x_1+u_1, x_2+u_2, \dots, x_{d-1}+u_{d-1}, r_2-y-a_1u_1) \\
&\le f(x_1+u_1, x_2+u_2, \dots, x_{d-1}+u_{d-1}, r_2-y-\sum_{i=1}^{d-1} a_iu_i) \\
&\le f(x_1+u_1, x_2+u_2, \dots, x_{d-1}+u_{d-1}, r_2-y-\sqrt{d-1}a_1u_1).
\end{aligned}
$$

A similar argument implies that, when the supporting hyperplane $H_{(a_1,0,\dots,0)}$ is turned infinitesimally while the slope is changed from a_1 to a_1+h, and thus, the first coordinate jumps by u_1, the d-th coordinate is increased by no more than $hu_1\sqrt{d-1}$. Combining all of this leads to the following chain of equalities and inequalities, for every $g \in C_f$,

$$
\begin{aligned}
\mathbf{E}\{g(W^\pi(a_1 &+ h)) - g(\mathbf{w}) | W^\pi(a_1) = \mathbf{w}\} \\
&= n\int_{A_h} f(\mathbf{v})\{g(\mathbf{v})-g(\mathbf{w})\}d\mathbf{v} + o(h) \\
&\le nh\sqrt{d-1}\int_{[-X',X']^{d-2}}\int_0^\infty u_1 \\
&\qquad \cdot f(x_1+u_1, x_2+u_2, \dots, x_{d-1}+u_{d-1}, r_2-y-\sum_{i=1}^{d-1} a_iu_i) \\
&\qquad \cdot \{g(W^\pi(a_1+h)) - g(\mathbf{w})\}du_1du_2\cdot\dots\cdot du_{d-1} + o(h) \\
&\le nh\sqrt{d-1}\int_{[-X',X']^{d-2}}\int_0^\infty u_1 \\
&\qquad \cdot f(x_1+u_1, x_2, \dots, x_{d-1}, r_2-y-\sqrt{d-1}a_1u_1) \\
&\qquad \cdot \{g(x_1+u, x_2, \dots, x_{d-1}, y+\sqrt{d-1}a_1u) - g(\mathbf{w})\}du_1 \\
&\qquad\qquad \cdot\dots\cdot du_{d-1} + o(h) \\
&= nh(2X')^{d-2}\sqrt{d-1}\int_0^\infty u \\
&\qquad \cdot f(x_1+u, x_2, \dots, x_{d-1}, r_2-y-\sqrt{d-1}a_1u) \\
&\qquad \cdot \{g(x_1+u, x_2, \dots, x_{d-1}, y+\sqrt{d-1}a_1u) - g(\mathbf{w})\}du + o(h) \\
&= nh(2X')^{d-2}\sqrt{d-1}\int_0^{X_0-x_1} u \\
&\qquad \cdot f(x_1+u, x_2, \dots, x_{d-1}, r_2-y-\sqrt{d-1}a_1u) \\
&\qquad \cdot \{g(x_1+u, x_2, \dots, x_{d-1}, y+\sqrt{d-1}a_1u) - g(\mathbf{w})\}du + o(h).
\end{aligned}
$$

Observe that we used the fact that the functions f and g have the same sets of increase, by definition of C_f. In addition, observe that the contribution of the integration of the first coordinate above X_0 is negligible by Lemma 5: since g has compact support, the remainder term is $o(h)$, uniformly in a_1, for $|a_1| \le X'/r_2$ and $w \in S_0$. Furthermore, we used the uniform continuity of f and g on S_0 and the fact that the line of slope a_1 through w cuts the paraboloid in X_0. Now repeated conditioning arguments prove the first claim about the supermartingale.

The reasoning to verify the *submartingale* runs quite parallel, i.e. the reverse inequalities hold if the lower bounds are placed for the upper bounds above. This finishes the proof. □

Next, for $b > a > 0$ and $B \in \mathcal{B}^d$ (the family of Borel sets in $\mathbb{R}^d$), we may define the random counting measure $\eta(a, b\,;\,\cdot\,)$ by

$$\eta(a, b; B) = \sum_{\substack{W^\pi(c)\neq W^\pi(c-)\\ a\leq c\leq b}} 1_B(W^\pi(c) - W^\pi(c-)). \tag{27}$$

Since the reformulation of the vertex process in Definition 2 does not affect the values of the vertex process W, for $0 < a < b \leq X'/r_2$, the identity

$$\eta(a, b; B) = \sum_{\substack{W(c,0,\dots,0)\neq W(c-,0,\dots,0)\\ a\leq c\leq b}} 1_B(W(c, 0, \dots, 0) - W(c-, 0, \dots, 0)) \tag{28}$$

is a consequence. Moreover, let η_a denote the number of jumps of W^π in the time interval $[0, a]$,

$$\eta_a = \eta(0, a; \mathbb{R}^d). \tag{29}$$

Furthermore, for every $B \in \mathcal{B}^d$, denote the *jump measure* of the process W^π at the point $\mathbf{w} \in \mathbb{R}^d$ at time $a > 0$ by

$$M(W^\pi(a), \mathbf{w}; B), \tag{30}$$

which can be interpreted as the probability that the process W^π jumps from $\mathbf{w}$ to some point in $\mathbf{w} + B$ at time a. The infinitesimal generator of W^π is dominated by the infinitesimal generator L_a^u of the *unknown* jump process W^u with *known* jump measure $M(W^u(\cdot), \cdot\,;\,\cdot\,)$. The process W^u is uniquely determined by its jump measure given in (31) below. In fact, the operator L_a^u can be written with respect to the jump measure $M(W^u, \cdot; \cdot)$ in the following way:

$$[L_a^u g](\mathbf{w}) \;=\; \int_{\mathbb{R}^d} [g(\mathbf{w} + \mathbf{z}) - g(\mathbf{w})]\; M(\mathbf{w}; d\mathbf{z}). \tag{31}$$

Analogously, we define the jump process W^l as the process having jump measure $M(W^l(\cdot), \cdot; \cdot)$, expressed as the integrating measure in the integral expression for the operator L_a^l. Assume that W^u and W^l are Markovian. To state the proposition and lemma below, we need introduce the counting measures

$$\begin{aligned}
\eta^u(a, b; B) &= \sum_{\substack{W^u(c)\neq W^u(c-)\\ a\leq c\leq b}} 1_B(W^u(c) - W^u(c-))\,,\\
\eta^l(a, b; B) &= \sum_{\substack{W^l(c)\neq W^l(c-)\\ a\leq c\leq b}} 1_B(W^l(c) - W^l(c-))\,,
\end{aligned}$$

for $b > a > 0$ and $B \in \mathcal{B}^d$,

$$\begin{aligned}
\eta_a^u &= \eta^u(0, a; \mathbb{R}^d),\\
\eta_a^l &= \eta^l(0, a; \mathbb{R}^d),
\end{aligned} \tag{32}$$

and the σ-algebras

$$\begin{aligned}\mathcal{F}^u_{a,b} &= \sigma\{W^u(c) : a \le c \le b\},\\ \mathcal{F}^l_{a,b} &= \sigma\{W^l(c) : a \le c \le b\}.\end{aligned}$$

The next result, due to Stroock ([14], Theorem 1.3), is slightly adjusted for our purposes.

Proposition 7 (Stroock)**.** *For each bounded Borel measurable function $g : \mathbb{R}^d \to \mathbb{R}$ in C_f that vanishes in a neighbourhood of the origin, for each $a > 0$ and $\theta \in \mathbb{R}^d$, the process*

$$\exp\left\{\langle i\theta, W^u(b) - W^u(a)\rangle + \int_{\mathbb{R}^d} g(\mathbf{z})\,\eta^u(a,b\,;d\mathbf{z}) - \int_a^b dc \int_{\mathbb{R}^d} \{e^{\langle i\theta,\mathbf{z}\rangle + g(\mathbf{z})} - 1\}\, M(W^u(c);d\mathbf{z})\right\}_{b>a}$$

is a martingale with respect to the filtration $\{\mathcal{F}^u_{a,b} : b \ge a\}$, where W^u is the process with jump measure defined in (31), $\langle \mathbf{x},\mathbf{y}\rangle$ denotes the inner product on $\mathbb{R}^d$, and $i = \sqrt{-1}$. Similarly, the exponential is a martingale with respect to $\{\mathcal{F}^l_{a,b} : b \ge a\}$ if W^u and η^u are replaced by W^l and η^l.

The uniform boundedness of the random measures $M(W^u(\cdot),\cdot)$ and $M(W^l(\cdot),\cdot)$, required for this result to be valid, follows by virtue of Proposition 6, by (22) and (31). The above martingales will be helpful in tackling the asymptotic moments of the number of convex hull vertices. The following result is a consequence of Proposition 7.

Lemma 8. *For $b > 0$*

$$\{\eta^u_b - \int_0^b M(W^u(c);\,\mathbb{R}^d)\,dc\}_{b>0} \tag{33}$$

and

$$\{(\eta^u_b)^2 - \int_0^b (2\eta^u_c + 1)\; M(W^u(c);\,\mathbb{R}^d)\,dc\}_{b>0} \tag{34}$$

are $\mathcal{F}^u_{0,b}$-martingales. Similarly, the two expressions are $\mathcal{F}^l_{0,b}$-martingales if W^u, η^u_b and η^u_c, respectively, are replaced by W^l, η^l_b and η^l_c, respectively.

Proof. Write $J(\lambda)$ for the martingale in Proposition 7 when $\theta = \{\mathbf{0}\}$, $a = 0$ and $g \equiv \lambda$. Then once differentiating $J(\lambda)$ with respect to λ and setting $\lambda = 0$ lead to the first assertion (33) in view of (29).

The second claim is a consequence of the properties of the Poisson point process (and can also be found in [6], Lemma 2.6). Since the verification is short, we present the steps for $(\eta^u_b)^2$, say. It is convenient to define the quantity $\eta^u_{a,b} = \eta^u_b - \eta^u_a$, for $a \le b$. We have

$$\begin{aligned}\mathbf{E}\left\{(\eta^u_{b+h})^2 - (\eta^u_b)^2 \,|\, \mathcal{F}^u_{0,b}\right\} &= \mathbf{E}\left\{(\eta^u_{b,b+h})^2 + 2\,\eta^u_b\eta^u_{b,b+h} \,|\, \mathcal{F}^u_{0,b}\right\}\\ &= \mathbf{E}\{(1 + 2\eta^u_b)\,(h\,M(W^u(b);\mathbb{R}^d) + R_h) \,|\, \mathcal{F}^u_{0,b}\},\end{aligned} \tag{35}$$

where we use the relation $\mathbf{E}[\eta^u_{b,b+h}|\mathcal{F}^u_{0,b}] \sim h\,M(W^u(b);\mathbb{R}^d)$, derived in (33), and we have $|R_h| = o(h)$, for $b > 0$, as $h \downarrow 0$, by relying on the properties of the underlying

Poisson point process. In fact, for $0 < h < 1$,

$$\begin{aligned} & h\, M(W^u(b); \mathbb{R}^d)\, \exp\{-h\, M(W^u(b); \mathbb{R}^d)\} \\ & \quad \le \mathbf{E}\{(\eta^u_{b,b+h})^2 | W^u(b)\} \\ & \quad \le h\, M(W^u(b); \mathbb{R}^d) + h^2\, [M(W^u(b); \mathbb{R}^d)]^2. \end{aligned}$$

An easy exercise finishes the proof. □

Corollary 9. *For $b > 0$*

$$\{\eta_b - \int_0^b M(W^u(c);\, \mathbb{R}^d)\, dc\}_{b>0} \tag{36}$$

and

$$\{\eta_b^2 - \int_0^b (2\eta_c + 1)\, M(W^u(c);\, \mathbb{R}^d)\, dc\}_{b>0} \tag{37}$$

are $\mathcal{F}_{0,b}$-supermartingales, and

$$\{\eta_b - \int_0^b M(W^l(c);\, \mathbb{R}^d)\, dc\}_{b>0} \tag{38}$$

and

$$\{\eta_b^2 - \int_0^b (2\eta_c + 1)\, M(W^l(c);\, \mathbb{R}^d)\, dc\}_{b>0} \tag{39}$$

are $\mathcal{F}_{0,b}$-submartingales.

Proof. Since the integrals in (36) and (37), respectively, essentially depend on the jump measure, in light of Proposition 6 and (31), we can conclude that, for each $a > 0$, $\mathbf{w} \in B_n$ and $B \in \mathcal{B}^d$,

$$M(W^\pi(a), \mathbf{w}; B) \le M(W^u(a), \mathbf{w}; B),$$

whence (36) and (37) follow. Similar arguments imply (38) and (39). □

Before bringing this section to an end, we shall transform the processes W^u and W^l into *stationary* local processes T^u and T^l. Suppose $W^u(a) = (X_1(a),\, X_2(a),\, \ldots, X_d(a))$ and $W^l(a) = (Y_1(a), Y_2(a), \ldots, Y_d(a))$. Transform the process $\{W^u(a)\}$ into the process $\{T^u(a) = (R_1(a),\, R_2(a),\, \ldots,\, R_d(a))\}$ by the substitution

$$\begin{aligned} R_1(a) &= X_1(a) - a r_2 \sqrt{d-1}, \\ R_k(a) &= X_k(a), \qquad\qquad (2 \le k \le d-1) \\ R_d(a) &= X_d(a) - a\sqrt{d-1}\, X_1(a) + a^2 r_2 (d-1)/2, \end{aligned} \tag{40}$$

and similary, transform the process $\{W^l(a)\}$ into the process $\{T^l(a) = (S_1(a), S_2(a), \ldots, S_d(a))\}$,

$$\begin{aligned} S_1(a) &= Y_1(a) - a r_2, \\ S_k(a) &= Y_k(a), \qquad\qquad (2 \le k \le d-1) \\ S_d(a) &= Y_d(a) - a Y_1(a) + a^2 r_2/2. \end{aligned}$$

Let $Z^{(1)}$ denote the first coordinate of the random vector Z and denote $G_d(x) = \mathbf{P}(Z^{(1)} \geq x)$. Setting $d = 2$ and following Dwyer ([3], Section 1 and 3), we obtain under the conditions in (6), for sufficiently large x,

$$G_2(x) \sim (2\pi)^{-1/2}\sqrt{\nu(x)}F_R(x), \tag{41}$$

$$G_2'(x) = \frac{\partial}{\partial x}G_2(x) \sim -(2\pi)^{-1/2}F_R(x)/(x\sqrt{\nu(x)})\,. \tag{42}$$

Thus, the local infinitesimal generator of T^u becomes, for each continuously differentiable real-valued function $g \in C_f$ with compact support in S_0,

$$\begin{aligned}(43)\ \lim_{h\to 0}\frac{1}{h}\mathbf{E}\{g(T^u(a+h)) \quad &- \quad g(x_1, x_2, \dots, x_d) \mid T^u(a) = (x_1, x_2, \dots, x_d)\}\\ &= \quad -n(2X')^{(d-2)}\sqrt{d-1}\int_0^{\sqrt{2r_2x_d}-x_1} u\\ &\qquad \cdot G_2'(\|(x_1+u, x_2, \dots, r_2 - x_d)\|)\\ &\qquad \cdot [g(x_1+u, x_2, \dots, x_d) - g(x_1, x_2, \dots, x_d)]du\\ &\qquad - r_2\sqrt{d-1}\,\frac{\partial}{\partial x_1}g(x_1, \dots, x_d)\\ &\qquad - x_1\sqrt{d-1}\,\frac{\partial}{\partial x_d}g(x_1, \dots, x_d),\end{aligned}$$

and the jump measure in $(x_1, x_2, \dots, x_d)$ is

$$\begin{aligned}M(T^u(a); B) \quad &= \quad -n(2X')^{(d-2)}\sqrt{d-1}\int_0^{\sqrt{2r_2x_d}-x_1} u\\ &\qquad \cdot G_2'(\|(x_1+u, x_2, \dots, r_2 - x_d)\|)1_B(u, 0, \dots, 0)du.\end{aligned} \tag{44}$$

The process T^u jumps "horizontally". Its deterministic part runs through curves parallel to the paraboloid $v = \sum_{i=1}^{d-1} u_i^2/2r_2$. The transformation may be imagined as a rotation and a shift of the coordinate system, or perhaps simpler, just as a rotation of the convex hull around the point $(0, 0, \dots, r_2)$ while the coordinate system is fixed. In the same spirit, we derive the infinitesimal generator of T^l and find the very same expression above with the factor $\sqrt{d-1}$ removed everywhere. Since the infinitesimal generators of T^u and T^l do not depend on the time parameter a, the processes T^u and T^l are stationary. Now we can start reaping the fruits.

Lemma 10. *As* $n \to \infty$,

$$b\,\mathbf{E}[\,M(T^l(0); \mathbb{R}^d)\,] \quad \leq \quad \mathbf{E}[\eta_b\,] \leq b\,\mathbf{E}[\,M(T^u(0); \mathbb{R}^d)\,], \tag{45}$$

$$2I^l \quad \leq \quad \mathbf{E}[\eta_b^2\,] - \mathbf{E}[\eta_b\,] \leq 2I^u, \tag{46}$$

with

$$\begin{aligned}I^l \quad &= \quad \int_0^b da\int_0^a \mathbf{E}[\,M_*(T^l(0); \mathbb{R}^d)\,M(T^l(a-c); \mathbb{R}^d)\,]\,dc,\\ I^u \quad &= \quad \int_0^b da\int_0^a \mathbf{E}[\,M_*(T^u(0); \mathbb{R}^d)\,M(T^u(a-c); \mathbb{R}^d)\,]\,dc,\end{aligned} \tag{47}$$

where T^u *and* T^l*, respectively, are the processes derived above from* W^u *and* W^l*, respectively, as described in* (40) *and the* $M_*(\,\cdot\,;\,\cdot\,)$ *are the corresponding "backward jump measures" of* T^u *and* T^l*.*

Proof. It is enough to show one pair of inequalities since both pairs come with the same derivation. In view of Lemma 8 and (44), the second claim (46) remains to be verified. Consider the time reversed process $\{(-R_1(a-c), R_2(a-c), \dots, R_d(a-c)) : c \geq 0\}$. The process $\{(-R_1(c), R_2(c), \dots, R_d(c)) : c \leq 0\}$ is recovered from the process $\{T^u(c) = (R_1(c), R_2(c), \dots, R_d(c)) : c > 0\}$ by interchanging the sign of the first coordinate and by moving backwards in time. The new process has the same distribution as the original process T^u. Let $M_*(\cdot;\cdot)$ denote the jump measure of the time reversed process. Now (46) follows from the stationarity of T^u and Corollary 9. ∎

For the following result to hold, the observation $r_0\sqrt{\nu(r_0)} \not\to \infty$ is crucial. As yet, we have not made use of the special form of the density G_2'. Let T^π be a *stationary* process that has identical moments to the moments of W^π and has the Markov property. Such a process exists by the existence of the upper process T^u and the lower process T^l.

Corollary 11. *For each sufficiently small $a \geq 0$, as $n \to \infty$,*

$$\delta_1/(r_0\,(v(r_0))^{1/2}) \leq M(T^\pi(a);\mathbb{R}^d) \quad\leq\quad \delta_2/(r_0\,(v(r_0))^{1/2}), \tag{48}$$

for some constants $\delta_1 = \delta_1(d,F)$ and $\delta_2 = \delta_2(d,F)$, depending on F and d. Furthermore,

$$\begin{aligned} \mathbf{E}[M(T^\pi(a);\mathbb{R}^d)] &\sim \delta_3/(r_0\,(v(r_0))^{1/2}), \\ \mathbf{E}[M^*(T^\pi(0);\mathbb{R}^d)M(T^\pi(a);\mathbb{R}^d)] &\sim \delta_4/(r_0^2\,v(r_0)) \end{aligned} \tag{49}$$

for some $\delta_3 = \delta_3(d,F) \in (\delta_1,\delta_2)$ and some $\delta_4 = \delta_4(d,F)$.

Proof. Recall from (10) that $r_0 = L(n)$ and from (42) that

$$G_2'(x) \sim -(2\pi)^{-1/2}F_R(x)/(x\sqrt{\nu(x)}).$$

An easy calculation shows that, as $n \to \infty$, $G_2'(u) \sim G_2'(r_0)$ for every $u \in (r_1, r_2)$, being a consequence of the slow variation of the functions L and ν, and thus, of $L(n)(\nu(L(n)))^{1/2}$. From (12), we recover $nF_R(r_0) = 1$. Combining all of these gives

$$G_2'(r_0) \sim -(2\pi)^{-1/2}(nr_0\sqrt{\nu(r_0)})^{-1}.$$

Observe that the contribution of integration outside the region A_n^* is negligible, that is, for x_1 beyond X'. Consider $T^l(0) = (x_1, \dots, x_d)$ and $a \geq 0$ sufficiently small. Hence from Lemma 10 and (44), we derive the following inequalities:

$$\mathbf{E}[M(T^l(a);\mathbb{R}^d)] \quad\leq\quad \mathbf{E}[M(T^\pi(a);\mathbb{R}^d)] \leq \mathbf{E}[M(T^u(a);\mathbb{R}^d)],$$

where T^l is a stationary jump process with jump measure

$$\begin{aligned} M(T^l(a);\mathbb{R}^d) &= -n(2X')^{d-2}\int_0^{X'-x_1} u \\ &\qquad \cdot G_2'(\|(x_1+u, x_2, \dots, r_2 - x_d)\|)du \\ &\sim -n\,2^{d-3}(X')^d G_2'(r_0) \\ &= 2^{d-3}(X')^d/(\sqrt{2\pi}\,r_0\,\sqrt{v(r_0)}), \end{aligned}$$

and similarly, T^u is a stationary jump process with jump measure

$$M(T^u(a);\mathbb{R}^d) = \sqrt{d-1}M(T^l(a);\mathbb{R}^d).$$

It remains to be seen that we can choose X' between two positive finite constants. This is an immediate consequence of the fact that, for every $\varepsilon > 0$, with a probability increasing to one, the jumps of the process T^π are in $[(1-\varepsilon)s_n, (1+\varepsilon)s_n]$, where s_n denotes the expected jump length of the process T^π (equivalently, of W^π in a neighbourhood of the origin). In addition, s_n is bounded. This may be seen as follows. Consider a d-cube C_n of sides t_n so that the number M_n of realized points from the Poisson point process ξ_n in C_n is unbounded in n in expectation. The event that the jump length of T^π is too large or too small is related to how points cluster, equivalently, to having regions with too few Poisson points or to having regions with too many Poisson points. We may view M_n to be the sum of $r = r(n)$ i.i.d. Bernoulli random variables and apply the following large deviation estimate to the sum.

Lemma 12 (Hoeffding [7]). *Let $\{S_k\}_{k\geq 1}$ be the partial sums of a sequence of independent random variables with zero expectation. Let B denote a bound on the absolute values of the increments. Then for every $\varepsilon > 0$ and $k \geq 1$,*

$$\mathbf{P}(S_k \geq \varepsilon k) \geq e^{-c\varepsilon^2 k} \tag{50}$$

where $c = B^{-2}/2$.

This guarantees both directions, that is, there cannot be too many and not too few points on subsets of C_n, with the extreme events having a probability that decays to zero exponentially fast. In other words, the probability that the nearest neighbour-distance is not within positive and finite multiples of the expected nearest neighbour-distance tends to zero exponentially fast. Consequently, for every $\varepsilon > 0$, with high probability, the jump length of T^π lies in $[s_n(1-\varepsilon), s_n(1+\varepsilon)]$. Notice that up to a constant multiplier, all of T^l, T^u and T^π have the same expected jump lengths because of the bounds on the infinitesimal generators. If we choose X' equal to an appropriate constant, then $M(T^\pi(a); \mathbb{R}^d)$ is of order $1/(r_0\sqrt{v(r_0)})$. Since by assumption, $L(n)\sqrt{\varepsilon(n)} = r_0\sqrt{\nu(r_0)} \not\to \infty$ as $n \to \infty$, the jump measure $M(T^\pi(a); \mathbb{R}^d)$ is bounded away from zero; thus, the expected jump length, being proportional to $M(T^\pi(a); \mathbb{R}^d)^{-1}$, is bounded. The four claims follow. □

6. Proofs of Theorems 1 and 2

For the proof of the Central Limit Theorem for N_n, a series of lemmata will be needed to carefully handle the details. Assume that we already know how to visit each element of a partition of the surface exactly once. The path need not be connected, since in proving the CLT, we can deal with a sequence of random variables, indexed by a countable, possibly unordered set. It is sufficient to know the degree of dependence of each of the elements from one fixed element, which is measured by the *mixing coefficient.* Since the sequence of numbers of jumps within each element of the partition is a sequence of identical random variables by the rotational invariance, if the elements have equal size, we can build a stationary sequence of random variables. Recall the stationary process T^π. We will show that the process $\{T^\pi(a) : a \in \mathbb{R}\}$ is *strongly mixing* and the dependence decreases exponentially fast to zero with increasing "distance" for sufficiently large n.

In showing that the entire stationary sequence has the strong mixing property, the basic idea is to attach to each element a value that expresses the degree of dependence from a certain reference point. First, we define the mixing coefficient for each element along the path that we follow when increasing only the first coordinate of the slope parameter of the supporting hyperplane. Second, in terms of mixing coefficients, each of the remaining elements is equivalent to one element already furnished with a mixing coefficient. Because of the exponential decay of the mixing coefficients, it will be easily established that the summation of all mixing coefficients, raised to the power $2+\delta$ for each $\delta > 0$, converges.

Lemma 13. *Let $\mathcal{F}_0 = \sigma\{T^\pi(c) : c \leq 0\}$ and $\mathcal{F}_{a+} = \sigma\{T^\pi(c) : c \geq a\}$ for each $a > 0$, and let $A \in \mathcal{F}_0$ and $B \in \mathcal{F}_{a+}$. Then, for some constants $0 < c' < \infty$ and $1 \leq \gamma \leq \sqrt{d-1}$*

$$|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| \leq \tau_n(a) \tag{51}$$

where

$$\tau_n(a) \quad \leq \quad 4\exp\{nG_2'(r_0 + \varepsilon_n/2 - a^2\gamma^2 r_2/16)a^3 r_2^2 c'\},$$

which tends to zero exponentially fast iff $a > 2(2\varepsilon_n/r_2)^{1/2}/\gamma$.

Proof. Let $T^\pi(a) = (S_1(a), S_2(a), \ldots, S_d(a))$, for each $a \in \mathbb{R}$. Assume that at time $a \geq 0$ the boundary of the convex hull is approximated by the paraboloid $x_d = ((x_1 - r_2 a\gamma)^2 + (\sum_{i=2}^{d-1} x_i^2))/2r_2$ for some constant $1 \leq \gamma \leq \sqrt{d-1}$. We will show that the left-hand side of (51) is bounded above by four times the probability that a certain region between the approximating paraboloid at the origin and a hyperplane contains no points of the Poisson point process ξ_n.

In fact, the point $T^\pi(a)$ arises when the convex hull is supported by the hyperplane $\{(u_1, u_2, \ldots, u_d) \in \mathbb{R}^d : u_d = au_1 + (S_d(a) - aS_1(a))\}$ and there is no point of the Poisson point process ξ_n in the region K_a between the paraboloid and the hyperplane, namely, in the region

$$\begin{aligned} K_a \quad = \quad & \{(u_1, u_2, \ldots, u_d) \in \mathbb{R}^d : ((u_1 - r_2 a\gamma)^2 + (\sum_{i=2}^{d-1} u_i^2))/2r_2 \leq u_d \\ & \leq au_1 + (S_d(a) - aS_1(a)),\ |u_i| \leq X' \text{ for each } 2 \leq i \leq d-1\}, \end{aligned}$$

where X' is defined in (15) and has been specified to be constant (see the proof of Corollary 11). Consequently, if for any a the points $T^\pi(0)$ and $T^\pi(a)$ are such that the two regions K_0 and K_a are disjoint, the events $A \in \mathcal{F}_0$ and $B \in \mathcal{F}_{a+}$ are based on disjoint sets of the underlying Poisson point process, and thus, are independent of each other. More precisely, if $S_d(0)$ and $S_d(a)$ are smaller than $a^2\gamma^2 r_2/8$, then K_0 and K_a are disjoint. Let E denote the event $\{S_d(0) < a^2\gamma^2 r_2/8,\ S_d(a) < a^2\gamma^2 r_2/8\}$. Then, by the Markov property of the process $\{T^\pi(a) : a \in \mathbb{R}\}$,

$$\mathbf{P}(A \cap B|E) = \mathbf{P}(A|S_d(0) < a^2\gamma^2 r_2/8)\,\mathbf{P}(B|S_d(a) < a^2\gamma^2 r_2/8).$$

Now, the bigger the distance of the point $T^\pi(a)$ to the x_1-axis the smaller the dependence from the point $T^\pi(0)$. Thus, we do not decrease the dependence when choosing $T^\pi(a)$ on the x_1-axis among all possible values of $T^\pi(a)$. Suppose so. Then $\mathbf{P}(S_d(a) \geq a^2\gamma^2 r_2/8) = \mathbf{P}(S_d(0) \geq a^2\gamma^2 r_2/8)$ by the stationarity of T^π. Therefore, by applying some rather crude estimates, we get

$$\begin{aligned} |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| \quad & \leq \quad \mathbf{P}(E^c) + 2\mathbf{P}(S_d(0) \geq a^2\gamma^2 r_2/8) \\ & \leq \quad 4\mathbf{P}(S_d(0) \geq a^2\gamma^2 r_2/8), \end{aligned} \tag{52}$$

where E^c denotes the complementary event to E. However, we have $S_d(0) \geq a^2\gamma^2 r_2/8$ if and only if the region K_0 contains no points of the Poisson point process ξ_n. We are going to give a lower bound for the probability $\mathbf{P}(K_0)$, by using the intermediate value theorem and the fact that the tail of the distribution decreases exponentially fast: Let λ_d denote the Lebesgue measure in $\mathbb{R}^d$. Then, for some $0 < \delta < 1/2$,

$$\begin{aligned} \mathbf{P}(K_0) &= -G_2'(r_0 + \varepsilon_n/2 - (1-\delta)a^2\gamma^2 r_2/8)\,\lambda_d(K_0) \\ &\geq -G_2'(r_0 + \varepsilon_n/2 - a^2\gamma^2 r_2/16)\,\lambda_d(K_0) \\ &= -G_2'(r_0 + \varepsilon_n/2 - a^2\gamma^2 r_2/16)\,(2X')^{d-2}a^3\gamma^3 r_2^2 c'' \end{aligned} \tag{53}$$

for a constant $0 < c'' < \infty$. Hence,

$$\begin{aligned} \mathbf{P}(S_d(0) \geq a^2\gamma^2 r_2/8) &= \exp\{-nP(K_0)\} \\ &\leq \exp\{nG_2'(r_0 + \varepsilon_n/2 - a^2\gamma^2 r_2/16)\,(2X')^{d-2}a^3\gamma^3 r_2^2 c''\} \\ &\leq \exp\{nG_2'(r_0 + \varepsilon_n/2 - a^2\gamma^2 r_2/16)\,a^3\gamma^3 r_2^2 c'\}. \end{aligned}$$

Since by (42), for large x, we get $G_2'(x) \sim -(2\pi)^{-1/2}F_R(x)/(x\sqrt{\nu(x)})$ and $\nu(x) \to 0$, as $x \to \infty$, the right-hand side of the last inequality decreases to zero exponentially fast in a iff $a > 2(2\varepsilon_n/r_2)^{1/2}/\gamma$. □

Recall from (29) that η_a denotes the number of jumps of the process $\{W^\pi(c) : 0 < c \leq a\}$.

Lemma 14. *For each $a > 0$ the moment generating function*

$$\mathbf{E}\exp\{\lambda\eta_a\}$$

is finite for all values $0 < \lambda < \infty$.

Proof. The proof makes an easy exercise. We give a sketch here. An argument that parallels the reasoning for the correspondence of the parameter spaces of the process W_n and T^π in the proof of Lemma 17 implies that, after a definite time, the stationary process T^π has visited all of the surface of the convex hull. This happens after visiting k_n elements of the partition, equivalently, after traveling for time $w_n k_n$, where k_n and w_n are defined in Lemma 18. Therefore, there exists some $0 < b_n < \infty$ such that $\mathbf{P}(\{T^\pi(c) : c > b_n\} = \emptyset) \to 1$ as $n \to \infty$ because F is absolutely continuous. The variable η_a cannot be bigger than $c'a/b_n$ times the number of points of the Poisson point process ξ_n in A_n^*, where c' is a positive finite constant. This number of points has a Poisson distribution with parameter $n\mathbf{P}(A_n^*)c'a/b_n$. Furthermore, $\mathbf{P}(A_n^*) \leq F_R(r_1)$. Hence, $n\mathbf{P}(A_n^*) \leq \gamma_1$, where γ_1 is defined in (11). Consequently, we have

$$\mathbf{E}\exp\{\lambda\eta_a\} \quad \leq \quad \exp\{(e^\lambda - 1)\gamma_1 ac'/b_n\}$$

for every $0 < \lambda < \infty$. □

Proposition 15. *Let η_b be the number of jumps of the process $\{T^\pi(c) : 0 \leq c \leq b\}$. Assume that $Z_1, Z_2, \ldots$ is a sequence of i.i.d. random vectors, each from an exponentially-tailed, spherically symmetric distribution F such that the smoothness conditions in* (6) *hold and $L(n)\sqrt{\varepsilon(n)} \not\to \infty$. Then as $n \to \infty$ and for sufficiently small $b > 0$, the following hold:*

(i) *There exist some positive finite constants $\alpha_1 = \alpha_1(d, F)$ and $\alpha_2 = \alpha_2(d, F)$ such that*

$$\mathbf{E}(\eta_b) \quad \sim \quad b\,\tilde{c}_1, \tag{54}$$

$$\mathbf{Var}(\eta_b) \quad \sim \quad b\,\tilde{c}_1 + b^2\,\tilde{c}_2, \tag{55}$$

where

$$\alpha_1/(r_0\sqrt{v(r_0)}) \quad \le \quad \tilde c_1,\ \sqrt{\tilde c_2} \le \alpha_2/(r_0\sqrt{v(r_0)}).$$

(ii) *Suitably normalized, the random variable* η_b *converges to a standard normally distributed random variable, i.e.*

$$(\eta_b - \tilde c_1 b)/(\tilde c_1 b + \tilde c_2 b^2)^{1/2} \xrightarrow{L} Z \sim \mathcal{N}(0,1), \tag{56}$$

where $\tilde c_1$ *and* $\tilde c_2$ *are as in (i) above.*

Proof. The first claim (i) follows as an appeal to Lemma 10 and Corollary 11.

Consider the d-ball $B(r)$ of radius r (centered at $(0,\dots,r)$). By the rotational invariance of the underlying distribution and the properties of the Poisson point process, for every $r > 0$, the projections of the vertices of the convex hull of the Poisson point process onto the boundary $\partial B(r)$ of $B(r)$ are identically and uniformly distributed random variables. To prove (ii), it is natural to partition the surface $\partial B(r_0)$ of $B(r_0)$ into congruent copies $\tilde S_1, \tilde S_2, \dots, \tilde S_{k_n}$ for some positive integer $k_n = k(n)$, unboundedly increasing with n. Observe that it suffices that 'most' of them are congruent copies of a typical element $\tilde S_1$. We may think of the $\tilde S_i$ as $(d-1)$-cubes. Consider the number M_i of points of the process T^π that are orthogonally projected onto $\tilde S_i$. The random variables M_i are identically distributed, and by the rotational invariance, their distribution only depends on the surface area of $\tilde S_1$. The size of $\tilde S_1$, and thus, the number k_n, to be determined below in Lemma 18, will be chosen such that the M_i are asymptotically independent of each other and such that, with positive probability, the points of W^π, associated to $\tilde S_i$, are hit by T^π. It remains to be shown that the M_i satisfy a Central Limit Theorem.

Assume that the elements $\tilde S_i$ are indexed such that the path taken by T^π is connected and crosses them in strictly increasing order, i.e. first $\tilde S_1$, then $\tilde S_2$, $\tilde S_3$ and so forth. The sequence $M_1, M_2, \dots$ satisfies the mixing condition of Lemma 13 for $A \in \sigma\{M_1, M_2, \dots, M_j\}$ and $B \in \sigma\{M_{j+l} : l > m\}$ with mixing coefficient $\tau_n(h(m))$ for some positive real-valued function h that indicates the minimal distance of $\{\tilde S_1, \tilde S_2, \dots, \tilde S_j\}$ to $\{\tilde S_{j+m+1}, \tilde S_{j+m+2}, \dots\}$. (On this path, the distance from a fixed reference point may strictly increase for a while but jump down to a value near zero before strictly increasing again. So, we consider a block of $\tilde S_j$'s.) Note that the number of $\tilde S_j$ within small distance is bounded in n. On every $\tilde S_j$ that is not too close the mixing coefficient, given in (51), decays sufficiently fast. Therefore, we may introduce a blocking system in the parameter space, of big and small blocks, such that big blocks are asymptotically independent and the number of points of T^π on the small blocks are negligible. In fact, the small blocks correspond to $(d-1)$-stripes parallel to the coordinate axes and the big blocks correspond to $(d-1)$-cubes. Observe that, by Lemma 14, $\mathbf{E}M_1^k$ is finite for each $k \ge 1$. Consequently, the conditions for a Central Limit Theorem for the sequence $M_1, M_2, \dots$ apply (see e.g. Ibragimov and Linnik [10], Theorem 18.5.3). Write $\eta_k = \sum_1^{w(k)} M_i$ for some function $w(k)$. Since, by the rotational invariance, $\eta_k - \eta_l$ only depends on the difference $k - l$, recalling the results in (i) completes the proof of (56). □

Recall that η_n and ξ_n denote the sample point process (to be precise, one of its realizations) and the Poisson point process on $\mathbb{R}^d$ with intensity measure $n\int_. dF$, respectively. Furthermore, consider the restricted point processes $\tilde\eta_n$ and $\tilde\xi_n$, the point processes η_n and ξ_n, restricted to A_n^*. From (21), recall $S_0 = \{(u_1, u_2, \dots, u_d) \in \mathbb{R}^d : u_d \ge (u_1^2 + u_2^2 + \dots + u_{d-1}^2)/2r_2\}$.

Lemma 16. *For every continuous function f with compact support S_0 the variables $\int_{\mathbb{R}^d} f\,d\eta_n$, $\int_{\mathbb{R}^d} f\,d\tilde{\eta}_n$, and $\int_{\mathbb{R}^d} f\,d\xi_n$ have the same first two moments in the limit as $n \to \infty$. In addition, $\int_{\mathbb{R}^d} f\,d\xi_n$ and $\int_{\mathbb{R}^d} f\,d\eta_n$ have the same distribution as $n \to \infty$.*

Proof. By (7), for sufficiently large n, $\int_{\mathbb{R}^d} f\,d\xi_n = \int_{A_n^*} f\,d\xi_n + o(1)$. Similarly, for the integrating measure with respect to η_n. Since $A_n^* \subset S_0$, the same identity holds, with A_n^* replaced by S_0. By Lemma 4, the variables $\int_{A_n^*} f\,d\tilde{\xi}_n$ and $\int_{A_n^*} f\,d\xi_n$ have the same distribution, as do the variables $\int_{A_n^*} f\,d\tilde{\eta}_n$ and $\int_{A_n^*} f\,d\eta_n$. Hence by Lemma 4, $\int_{\mathbb{R}^d} f\,d\xi_n$ and $\int_{\mathbb{R}^d} f\,d\eta_n$ have the same distribution as $n \to \infty$. Moreover, as $n \to \infty$, the first two moments of each of these random variables converge to the corresponding first two moments, with each integrating measure being replaced by the corresponding integrating measure for the limiting point processes, e.g. ξ_n converges to ξ as $n \to \infty$. We conclude that the variables $\int_{S_0} f\,d\eta_n$, $\int_{S_0} f\,d\tilde{\eta}_n$, and $\int_{S_0} f\,d\xi_n$ have the same first two moments in the limit as $n \to \infty$, as desired. □

Note that, from the proof of Corollary 11 along with definition (15) of X', it follows that $\varepsilon_n \to 0$, as $n \to \infty$. Therefore, for sufficiently large n, the balls $B(r_0)$, $B(r_1)$, and $B(r_2)$ will have the same surface area. Therefore, it is irrelevant which of these balls we rely on. Consider a typical element $\tilde{S}_1$, say, of the partition of $\partial B(r_0)$, whose side length will be denoted by w_n. The function w_n of n will be such that, with positive probability, the orthogonal projection of $\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$ onto $\partial B(r_0)$ will have nonempty intersection with $\tilde{S}_1$. The cardinality of the partition is $k_n = k(n)$.

Let I_* be the interval $[-w_n k_n/2, w_n k_n/2)$. Define

$$\begin{aligned} L_n^\pi &= \#\{T^\pi(a) : a \in I_*\}, \\ N_n &= 2\#\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}. \end{aligned} \tag{57}$$

A key ingredient to this proof is the following result.

Lemma 17. *If $(L_n^\pi - \alpha(n))/\beta(n) \xrightarrow{L} Z \sim \mathcal{N}(0,1)$ for suitably chosen norming constants $\alpha(n)$ and $\beta(n)$, then $(N_n - \alpha(n))/\beta(n) \xrightarrow{L} Z \sim \mathcal{N}(0,1)$.*

Proof. In view of Lemma 16, it suffices to show that, as $n \to \infty$, with a probability increasing to one, twice the number of points in the set $\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}$ is equal to the number of points in the set $\{T^\pi(a) : a \in [-w_n k_n/2, w_n k_n/2)\}$ where w_n and k_n are given in Lemma 18 below.

We begin with verifying that associating the parameter space I_* to the process $T^\pi(a)$ corresponds to having $W_n(\mathbf{a}) \in A_n^*$, as an appeal to the analogy of lengths along the surface of the d-ball $B(r_0)$ of radius r_0 and the angles of the hyperplane, supporting the convex hull, with the coordinate axes, and thus, the analogy of lengths and surface areas.

The process $T^\pi(a)$ moves within a cubically tubular (cylindrical) neighbourhood of side lengths w_n along a (straight) line, where the parameter a indicates the length traveled since time 0. We need to determine the length of the shortest path that visits all of the surface $\partial B(r_0)$, intersected with the set A_n^*. It helps to visualize this surface as mapped onto a parallelepiped $R \subset \mathbb{R}^{d-1}$ of the same volume, which has one side length equal $r_0\kappa_d$ and each of the $(d-2)$ remaining sides of length equal to r_0, say. Recall the partition $\{\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_{k_n}\}$ from the proof of Proposition 15. Consider the number M_i of points of the process T^π that are orthogonally projected onto $\tilde{S}_i$. Since M_i depends only on the size of $\tilde{S}_i$, the $\tilde{S}_i$ may be patched together so that a parallelepiped R results. Thus, the length of the shortest path of $T^\pi(a)$ that covers all of R is $\kappa_d r_0 (r_0/w_n)^{(d-2)}$. Hence, let

$$w_n k_n = \kappa_d r_0 (r_0/w_n)^{(d-2)}.$$

This long path can be broken up into $\kappa_d (r_0/w_n)^{(d-1)}$ $(d-1)$-cubes, each one having each side length equal w_n. Hence,

$$k_n \sim \kappa_d (r_0/w_n)^{d-1}. \tag{58}$$

It remains to observe that the boundary of R which was not any boundary on the surface of the d-ball does not do any harm since, in Euclidean space, the number of elements $\tilde{S}_j$ mapped onto cubes with nonempty intersection with the boundary of R is dominated by the total number k_n.

Next, recall that η_n denotes the sample point process, ξ_n denotes the Poisson point process on $\mathbb{R}^d$ with intensity measure $n \int_\cdot dF$, and $\tilde{\eta}_n$ and $\tilde{\xi}_n$ denote the corresponding processes that have the same distribution as the processes $\eta_{n|_{A_n^*}}$ and $\xi_{n|_{A_n^*}}$ and in a common probability space as defined in Lemma 4. Let W_n', $\tilde{W}_n$, and W be the vertex processes based on the processes $\eta_{n|_{A_n^*}}$, $\tilde{\eta}_n$ and $\tilde{\xi}_n$, respectively. We may define the following counters

$$\begin{aligned} N_n &= 2\#\{W_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}, \\ \tilde{N}_n &= \#\{\tilde{W}_n(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}, \\ N_n' &= \#\{W_n'(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}, \\ \tilde{L}_n &= \#\{W(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\}, \\ L_n^\pi &= \#\{T^\pi(a) : a \in I_*\}, \end{aligned}$$

where $I_* = [-w_n k_n/2, w_n k_n/2)$ (thus, the letter N relates to sample points and the letter L to Poisson points). Since the processes W^π and W have the same local jump measures, described by relations (27) or (28), the process L_n^π has the same distribution as $\tilde{L}_n$ whenever the ranges of their parameter spaces correspond to each other. This argument is given above. By Lemma 4, $\lim_{n\to\infty} \mathbf{P}(\tilde{N}_n \neq \tilde{L}_n) = 0$ and the random variables N_n' and $\tilde{N}_n$ have the same distribution. Furthermore by (7), $\lim_{n\to\infty} \mathbf{P}(N_n' \neq N_n) = 0$. As a consequence,

$$\lim_{n\to\infty} \mathbf{P}(L_n^\pi \neq N_n) = 0.$$

This finishes the proof. □

Lemma 18. *For any two sequences* $\{\vartheta_1(n)\}_n$ *and* $\{\vartheta_2(n)\}_n$ *so that* $\vartheta_1(n), \vartheta_2(n) \to \infty$ *as* $n \to \infty$ *and every* $\varepsilon > 0$, *the functions* w_n *and* k_n *are given by*

$$w_n = \gamma(1 + \varepsilon\vartheta_1(n))\, r_0 (v(r_0))^{1/2} \tag{59}$$

and

$$k_n = \beta(1 + \varepsilon\vartheta_2(n))^{-(d-1)}\, (v(r_0))^{-(d-1)/2} \tag{60}$$

for some constants $0 < \beta, \gamma < \infty$.

Proof. Recall that the number k_n denotes the cardinality of the partition of the surface $\partial B(r_0)$ into elements that are congruent copies of $\tilde{S}_1$ and was expressed in (58). Keep in mind that we are not careful about *exact* constants, and hence, the choice of constants will be irrelevant. Observe that w_n needs to be chosen sufficiently large so that the process T^π hits some points with positive probability. By the inequalities in (48), the jump length of T^π is proportional to $r_0\sqrt{\nu(r_0)}$. Pick some $\varepsilon > 0$. In order that the $\tilde{S}_i$ be asymptotically independent of each other, given two sequences $\{\vartheta_1(n)\}_n$ and $\{\vartheta_2(n)\}_n$ that converge to ∞ arbitrarily slowly, let

$$w_n = \gamma(1+\varepsilon\vartheta_1(n))\, r_0(v(r_0))^{1/2}, \tag{61}$$

and thus,

$$k_n = \beta(1+\varepsilon\vartheta_2(n))^{-(d-1)}\,(v(r_0))^{-(d-1)/2} \tag{62}$$

for some constants $0 < \beta, \gamma < \infty$, as required. ☐

Observe that eventually ε can be chosen arbitrarily small.

Lemma 19. *Let $M_i = L_n^\pi|_{\tilde{S}_i}$ denote the number of points of the process T^π that orthogonally project to $\tilde{S}_i$. Then*

$$\begin{aligned} \mathbf{E}[M_1] &\sim \alpha_3, \\ \mathbf{Var}[M_1] &\sim \alpha_4 \end{aligned}$$

for some finite constants α_3, α_4. Furthermore, there are some positive finite constants c_1 and c_2 so that, as $n \to \infty$,

$$\begin{aligned} \mathbf{E}[L_n^\pi] &\sim c_1(v(r_0))^{-(d-1)/2}, \\ \mathbf{Var}[L_n^\pi] &\sim c_2(v(r_0))^{-(d-1)/2}. \end{aligned} \tag{63}$$

Proof. Let $\tilde{S}_1$ be a typical element of the partition of the boundary of $B(r_0)$. Consider $M_i = L_n^\pi|_{\tilde{S}_i}$, the number of points of the process T^π that orthogonally project to $\tilde{S}_i$. In view of Proposition 15 with $b = w_n$, the above choices of k_n and w_n, and by the asymptotic independence of the $\tilde{S}_i$, we have

$$\begin{aligned} \mathbf{E}L_n^\pi &= \sum_{i=1}^{k_n} \mathbf{E}[M_i] \\ &= k_n\, \mathbf{E}[M_1], \\ \mathbf{Var}L_n^\pi &\sim \sum_{i=1}^{k_n} \mathbf{Var}[M_i] \\ &= k_n\, \mathbf{Var}[M_1] \end{aligned}$$

for sufficiently large n. Recall the estimates of $\tilde{c}_1$ and $\sqrt{\tilde{c}_2}$ from Proposition 15, being proportional to $(r_0\sqrt{v(r_0)})^{-1}$. Furthermore from (54) and (55), we have

$$\begin{aligned} \mathbf{E}[M_1] &\sim b\,\tilde{c}_1 \sim w_n\,\tilde{c}_1, \\ \mathbf{Var}[M_1] &\sim w_n\tilde{c}_1 + w_n^2\tilde{c}_2. \end{aligned}$$

Since $\varepsilon > 0$ in (61) and (62) was arbitrary, the two claims immediately follow. ☐

Proof of Theorem 1. Throughout the proof, we are not careful about which (positive and finite) constants to choose since the final statements do not provide *exact* constants. Observe that, in the neighbourhood of the origin, up to some factor between 1 and $\sqrt{d-1}$, the processes W, W_n and W^π and the stationary processes T^u and T^l have their parameters increasing at the same speed (this obviously is not true for large values of the parameters). Furthermore, the number of different points of the processes W^π, W^u and W^l, respectively, have the same moments as the number of points T^π, T^u and T^l, respectively, by (43). By Lemmata 17 and 16, it suffices to establish that L_n^π, suitably normalized, converges to a normally distributed random variable. Note that, by Lemma 5, the total error injected by using the paraboloid approximation to the boundary of the set A_n^* is of the order $1/\gamma_2 \to 0$, as $n \to \infty$. The choice of w_n in Lemma 18 guarantees that the probability that, for a positive fraction of the k_n elements $\tilde{S}_j$, not all convex hull vertices in $\tilde{S}_j$ are hit by T^π tends to zero as $n \to \infty$. Hence, Proposition 15 along with the moments of L_n^π, provided in Lemma 19, gives the desired result. □

Proof of Theorem 2. The verification runs along the same lines used in the proof of Theorem 1 with the pleasant difference that explicit expressions of upper and lower bounds for the expected jump measure $M(T^\pi(0);\mathbb{R}^d)$ can be computed. Let $T^\pi(0) = (x_1, x_2, \dots, x_d)$. From (43) together with $G_2'(x) \sim (2\pi)^{-1/2}\exp(-\|x\|^2/2)$, for sufficiently large x, it is easily deduced that

$$\begin{aligned} M(T^\pi(0);\mathbb{R}^d) &\le n\sqrt{d-1}\,(2X')^{d-2}\exp(-\|(r_2 - x_d)\|^2/2) \\ &\quad \cdot \int_0^{\sqrt{2r_2x_d}-x_1} (2\pi)^{-1/2} u \exp(-\|(x_1+u)\|^2/2)du \\ &\sim n\sqrt{d-1}\,(2X')^{d-2}\exp(-\|(r_2 - x_d)\|^2/2)/\sqrt{\pi}. \end{aligned}$$

A rather long sequence of calculations involving the densities of the vertex process (see [9]) yields

$$\mathbf{E}[M(T^\pi(0);\mathbb{R}^d)] \le \sqrt{d-1}(2X')^{d-2}(\pi)^{-1/2}.$$

Now, $k_n = \kappa_d(\nu(r_0))^{-(d-1)/2}$ and it suffices to choose $2X' \le 1$. Note that $r_0 = L(n) = \nu(r_0)^{-1/2} = \sqrt{2\ln n}$. Consequently,

$$\begin{aligned} \mathbf{E}[N_n] &\le k_n\sqrt{d-1}\,\pi^{-1/2} \\ &= \kappa_d\sqrt{d-1}\,\pi^{-1/2}\,(2\ln n)^{(d-1)/2}. \end{aligned}$$

□

Acknowledgement

The author is grateful to an anonymous referee for valuable remarks that led to considerable improvement of the exposition.

References

[1] Aldous, D. J., Fristedt, J., Griffin, P. S. and Pruitt, W. E. (1991). *The number of extreme points in the convex hull of a random sample.* J. Appl. Prob. **28**, 287-304. MR **92j:**60022

[2] Carnal, H. (1970). *Die konvexe Hülle von n rotationssymmetrisch verteilten Punkten.* Z. Wahrscheinlichkeitstheorie verw. Geb. **15**, 168-176. MR **44:**3367

[3] Dwyer, R. A. (1991). *Convex hulls of samples from spherically symmetric distributions.* Discrete Appl. Math. **31**, 113-132. MR **92i:**60020

[4] Efron, B. (1965). *The convex hull of a random set of points.* Biometrika **52**, 331-343. MR **34:**6820

[5] Feller, W. (1971). *An Introduction to Probability Theory and its Applications, Vol. II.* Wiley, New York. MR **42:**5292

[6] Groeneboom, P. (1988). *Limit theorems for convex hulls.* Prob. Theory Rel. Fields **79**, 327-368. MR **89j:**60024

[7] Hoeffding, W. (1963). *Probability inequalities for sums of bounded random variables.* J. Amer. Statist. Assoc. **58, no. 301**, 13–30. MR **26:**1908

[8] Hueter, I. (1992). *The convex hull of n random points and its vertex process.* Doctoral Dissertation, University of Berne.
[9] Hueter, I. (1994). *The convex hull of a normal sample.* Adv. Appl. Prob. **26**, 855–875. MR **96g:**60018
[10] Ibragimov, I. A. and Linnik, Y. V. (1971). *Independent and stationary sequences of random variables.* Groningen: Wolters Noordhoff. MR **48:**1287
[11] Kallenberg, O. (1983). *Random Measures.* Akademie Verlag Berlin. MR **87g:**60048
[12] Raynaud, H. (1970). *Sur l'enveloppe convexe des nuages de points aléatoires dans* $\mathbb{R}^n$. J. Appl. Prob. **7**, 35-48. MR **41:**2736
[13] Rényi, A. and Sulanke, R. (1963). *Über die konvexe Hülle von n zufällig gewählten Punkten.* Z. Wahrscheinlichkeitstheorie verw. Geb. **2**, 75-84. MR **27:**6190
[14] Stroock, D. W. (1975). *Diffusion processes associated with Lévy generators.* Z. Wahrscheinlichkeitstheorie verw. Geb. **32**, 209-244. MR **55:**6587

Department of Mathematics, University of Florida, Gainesville, Florida 32611
E-mail address: hueter@math.ufl.edu