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

by

Irene Hueter
Purdue University

Technical Report #93-39

Department of Statistics
Purdue University

July 1993
Limit Theorems for the Convex Hull of Random Points in Higher Dimensions

Irene Hueter*
Department of Statistics
Purdue University

July 1993

1 Introduction

Let $Z_1, Z_2, \ldots, Z_n$ be $n$ i.i.d. random vectors, each with absolutely continuous distribution $F$. Let $N_n$ be the number of vertices of the convex hull of $\{Z_1, Z_2, \ldots, Z_n\}$. What is the limiting distribution of $N_n$ as $n$ grows to $\infty$?

Although numerous papers investigate the behaviour of random convex hulls, we do not know the asymptotic distribution nor the second moment of any convex hull functional in the higher dimensional space. In this context, even in two dimensions there are only a few results (see [6],[1] and [7]). 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 combinatorial 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 two cases, namely, for a set of $n$ points 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 [7] the same approach was generalized in various directions. Among others, bivariate exponentially-tailed,

*Research supported by Swiss National Science Foundation
rotationally invariant distributions, especially, the normal distribution, were studied 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, which includes the normal distribution, whose tails decay at an exponential rate, in order to establish a central limit theorem for $N_n$ and to investigate the variance of $N_n$. Two cases must be distinguished. In the first case, the order of magnitude of $Var[N_n]$ is obtained exactly, in the second case, only an upper bound for the order of $Var[N_n]$ is obtained. The rotational invariance is of key importance to our analysis.

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

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

in the plane, and twenty years later, Dwyer [3] showed

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

in $d$ dimensions, for sufficiently large $n$. 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 1.1** 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 (10) below are satisfied. Denote by $\kappa_d = 2\pi^{d/2}/\Gamma(\frac{d}{2})$ the surface area of the unit $d$-ball.

(i) If $L(n)\varepsilon(n)^{1/2} \not\to \infty$, as $n \to \infty$, then

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

for some positive finite constant $c_1$, where

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

(ii) If $L(n)\varepsilon(n)^{1/2} \to \infty$, as $n \to \infty$, then

$$\frac{N_n - E[N_n]}{(Var[N_n])^{1/2}} \xrightarrow{L} \mathcal{N}(0,1),$$
where for some positive finite constant $c_2$
\[
E[N_n] \leq c_2 L(n)^2 \varepsilon(n)^{-(d-3)/2}
\]
\[
\text{Var}[N_n] \sim \mathcal{O}(L(n)^3 \varepsilon(n)^{-(d-4)/2}).
\]

In view of Dwyer’s result, in the second case of the above theorem our upper bound for the order of $E[N_n]$ is not sharp by the condition $L(n)\varepsilon(n)^{1/2} \to \infty$. We suspect that the correct order of the variance $\text{Var}[N_n]$ is still the same as that for $E[N_n]$, namely $\mathcal{O}(\varepsilon(n)^{-(d-1)/2})$. However, our method is far from being precise in this case. 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 $E[N_n]$ can be established.

**Theorem 1.2** 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 $n \to \infty$,
\[
\frac{N_n - c_3 (\ln n)^{(d-1)/2}}{(c_4 (\ln n)^{(d-1)/2})^{1/2}} \xrightarrow{L} \mathcal{N}(0, 1),
\]
where $0 < c_4 < \infty$ and
\[
0 < c_3 \leq 2\sqrt{d-1} (2\pi)^{(d-1)/2} \Gamma(d/2).
\]

Actually, Raynaud [9] gave an estimate for the first moment of the number of facets $F_n$, which implies an upper bound for $E[N_n]$, namely,
\[
E[N_n] \leq E[F_n] \sim 2^d d^{-1/2} (d-1)^{-1} \pi^{(d-1)/2} (\ln n)^{(d-1)/2}.
\]

Our upper bound is an improvement for large values of the dimension $d$. The inequality $E[N_n] \leq E[F_n]$ 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 take the exact value for $E[N_n]$.

The step from two to $d$ ($d > 2$) dimensions is not that straightforward. A key role is played by the jump measure for the moment computations. However, the derivation of the jump measure of the process of the vertices is a delicate issue, which reflects the fact that, in more than two dimensions, the natural order of the parameter space of the process running through the convex hull vertices is lost. Therefore, the bulk of this paper is devoted to define a way to follow a typical path on the surface of the convex hull such that all the vertices can be recorded.
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 try to mimic the two-dimensional situation. In Section 5, we shall study the behaviour of the local jump measure from a vertex point and somehow relate the moments of $N_n$ with the jump measure by looking at some (super)martingales. Finally, the last section contains the proofs of our 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 an important role in the estimation of the moments of $N_n$. Let $Z_1, Z_2, \ldots, Z_n$ be $n$ i.i.d. random vectors with common absolutely continuous distribution $F$. Denote the tail probability of the radial component by $F_R(x) = P(\|Z_1\| > x)$. We think of distributions with "exponential tails" as those distributions whose $F_R(x)$ satisfies

\[ x = L \left( \frac{1}{F_R(x)} \right), \]  

(5)

for a monotonically increasing function $L$ that varies slowly at infinity, i.e. as $x \to \infty$, for each $\lambda > 0$,

\[ \lim_{x \to \infty} L(\lambda x)/L(x) = 1. \]  

(6)

Since $F_R(0) = 1$ and $F_R(\infty) = 0$, $L(1) = 0$ and $L(\infty) = \infty$ hold. A slowly varying function $L(x)$ can be expressed in the form

\[ L(x) = a(x) \exp\left\{ \int_1^x \varepsilon(t)/t \, dt \right\} \]

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, we have $L(s) = \sqrt{2 \ln s}$ for the $d$-dimensional normal distribution. If we put $s = 1/F_R(x)$, we have

\[ x = L(s) = \exp\left\{ \int_1^s \varepsilon(t)/t \, dt \right\}. \]  

(7)

Let us further define

\[ 0 < \nu(u) = \varepsilon(L^{-1}(u)) = \varepsilon(1/F_R(u)) \to 0, \quad u \to \infty. \]  

(8)
Then the elementary correspondence
\[
\frac{dx}{ds} = \frac{\varepsilon(s)}{s} \iff \frac{ds}{s} = \frac{dx}{\nu(x) x}
\]
implies the useful representation
\[
F_R(x) = \exp(\ln F_R(x)) = \exp\left\{ - \int_0^x du / (\nu(u) u) \right\}. \tag{9}
\]

Proceeding along the lines of Carnal [2] and Dwyer [3], we impose the following smoothness conditions on \(\nu\) and thus, on \(L\), that are satisfied by most distributions in question, and forces that the function \(\varepsilon\) be slowly varying, too,
\[
\begin{align*}
\nu(x) & \text{ is monotone (decreasing) for large } x, \tag{10} \\
x \cdot \nu'(x) \cdot \ln(\nu(x)) & = o(1) \text{ as } x \to \infty, \\
\nu(x) \cdot \ln(x) & = o(1) \text{ as } x \to \infty.
\end{align*}
\]

For instance, the standard normal distribution with density function \((2\pi)^{-d/2} \exp(-\|x\|^2/2)\) has \(\nu(r) = r^{-2}\). 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, \(\varepsilon(n) \neq O(L(n)^{-2})\). Take the example \(F_R(r) = c_0 \exp(-r^k)\) for \(k > 0\), thus, for large \(r\), \(L(r) \sim (\ln r)^{1/k}\) and \(\varepsilon(r) \sim 1/(k \ln r)\). That means, for \(k > 2\), \(\varepsilon(n)^{-1/2} > L(n)\). Roughly speaking, the smaller the function \(\nu(\cdot)\) around the value \(L(n)\), the thinner is the tail of the distribution.

## 3 Vertex Process

It is very natural to investigate 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. Also, it will be convenient to have available some more notation and to define certain special regions related to the set of vertex points of the convex hull.

As the number of convex hull vertices is invariant under affine transformations, we may as well shift the whole sample by a certain amount. More precisely, it will be advantageous to think of our distribution to be centered at \((0, 0, \ldots, r_2)\), where \(r_2\) is the radius of a ball that contains a sample of \(n\) points with law \(F\) with a probability tending to 1 (Indeed, we will make this “shift assumption” for the rest of the paper). Let us first introduce the process which runs through the vertices of the convex hull.
Definition 3.1 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 of 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. 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. As the vector $\mathbf{a}$ runs through all $\mathbb{R}^{d-1}$, roughly one half of all convex hull vertices is counted. However, the process $W_n$ is close to be 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 $n \to \infty$. An easy coupling argument between a Poisson and a Binomial distribution will establish an upper bound for this distance. We try to find a Poisson point process whose extreme sample points look pretty much like the extremes of the original sample $Z_1, Z_2, \ldots, Z_n$.

For this purpose, we may first 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

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

and

$$P(Z_1 \in A_n^*) \longrightarrow 0.$$  \hspace{1cm} (11)

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

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

where $r_1$ and $r_2$ are chosen as follows:

$$r_0 = L(n) \hspace{1cm} (13)$$

$$r_1 = r_0 - \varepsilon_n/2$$

$$r_2 = r_0 + \varepsilon_n/2,$$

and $\varepsilon_n$ is such that

$$nF_R(r_1) = \gamma_1 \hspace{1cm} (14)$$

$$nF_R(r_2) = 1/\gamma_2$$
with \( \gamma_1, \gamma_2 \to \infty \). Note that there is still a lot of freedom for the choice of \( \varepsilon_n \). Then it is easily verified that the region \( A^*_n \) satisfies (11) and (12). Clearly, \( \varepsilon_n \to 0 \). This tells us 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.

Now let \( \eta_n \) denote the sample point process of size \( n \) and let \( \xi_n \) denote the Poisson point process on \( \mathbb{R}^d \) with intensity measure \( n \int \! dF \), which is an inhomogeneous Poisson point process.

**Lemma 3.1** There exist processes \( \tilde{\eta}_n \) and \( \tilde{\xi}_n \) defined on the same probability space such that

\[
\tilde{\eta}_n \overset{d}{=} \eta_n|_{A^*_n}, \quad \tilde{\xi}_n \overset{d}{=} \xi_n|_{A^*_n} \tag{15}
\]

and

\[
P(\tilde{\eta}_n \neq \tilde{\xi}_n) \leq 2P(Z_1 \in A^*_n). \tag{16}
\]

("\( \overset{d}{=} \)" stands for "restriction to".) We do not show the proof here, which is standard. It employs a comparison 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). Remark on the side that there actually are rotationally invariant distributions such that no such region \( A^*_n \) exists, and therefore, Poisson approximation is not possible, i.e. the algebraic-tailed distributions. Curiously, along with them goes an \( E[N_n] \) that is constant for large \( n \) (see Carnal [2], Dwyer [3], and also Aldous et al. [1]).

Now we are ready to define the vertex process \( \{W(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\} \) of a realization of the Poisson point process \( \xi_n \). Then, by the independence of all events of a Poisson process defined on disjoint sets, the process \( \{W(\mathbf{a}) : \mathbf{a} \in \mathbb{R}^{d-1}\} \) is Markovian. Note that \( W \) still depends on \( n \), although we drop the index. The requirement \( P(Z_1 \in A^*_n) \to 0 \) and Lemma 3.1 guarantee 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. For the normal distribution, this bound is of the order \( n^{-1} \).

### 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 runned “straightly”, the two-dimensional situation with one-dimen-
sional 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, only bounded. In fact, the farther away from the path the vertices are, which is tanta-
mount to the less convex hull vertices there are, the less accurate the upperbound estimate for the jump measure, and therefore, for the moments of $N_n$, will be.

On the other hand, the requirement to visit each neighbourhood exactly once, is not such a big 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 containing a straight path in order to reach conclusions about $N_n$. Of course, the path need run close enough to each point. That is meant by “visiting a neighbourhood”.

Hence, the first step will be to define such a typical small region for the local investigations, and the second step to give a new definition of 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 close to a paraboloid which is easier to deal with than with the boundary of the annulus $A_n^*$. We are going to show that in variational distance the vertex process defined by using the local paraboloid approximation is sufficiently close to the vertex process $W$.

To this end, write

$$X' = \sqrt{2r_2} \varepsilon_n$$  \hspace{1cm} (17)

and define

$$S_n = \{(u_1, u_2, \ldots, u_d) \in A_n^* : \sum_{i=1}^{d-1} u_i^2 \leq (X')^2\}.$$  \hspace{1cm} (18)

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

$$v = \left(\sum_{i=1}^{d-1} u_i^2\right)/2r_2 \approx r_2 - (r_2 - \sum_{i=1}^{d-1} u_i^2)^{1/2}$$
as local approximation to the outer boundary \( \{(u_1, \ldots, u_d) \in \mathbb{R}^d : \sum_{i=1}^{d-1} u_i^2 + (r_2 - u_d)^2 = r_2^2 \} \) of \( A_n^* \). Further define

\[
B_n = \{(u_1, \ldots, u_d) \in \mathbb{R}^d : (\sum_{i=1}^{d-1} u_i^2)/2r_2 \leq u_d, \sum_{i=1}^{d-1} u_i^2 \leq (X')^2 \}. \tag{19}
\]

Let \( \xi_n|_{S_n} \) and \( \xi_n|_{B_n} \) be the restrictions of the Poisson point process with intensity \( n \int dF \) to \( S_n \) and \( B_n \), respectively. Let \( W^S \) and \( W^B \) be the vertex processes defined as in Definition 3.1, but now by the sample replaced by a realization of \( \xi_n|_{S_n} \) and \( \xi_n|_{B_n} \), respectively.

**Lemma 4.1** Let the point processes \( \delta^S_n \) and \( \delta^B_n \) be defined by

\[
\delta^S_n = \{ W^S(a_1, \ldots, a_{d-1}) : 2r_2 (\sum_{i=1}^{d-1} a_i^2)^{1/2} \leq \sqrt{2r_2 \varepsilon_n} \}, \tag{20}
\]

\[
\delta^B_n = \{ W^B(a_1, \ldots, a_{d-1}) : 2r_2 (\sum_{i=1}^{d-1} a_i^2)^{1/2} \leq \sqrt{2r_2 \varepsilon_n} \}. \tag{21}
\]

Then, as \( n \to \infty \),

\[
P(\delta^S_n \neq \delta^B_n) = O(nF_R(r_2)(2\varepsilon_n/r_2)^{(d-1)/2}). \tag{22}
\]

**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 \((0,0,\ldots,r_2)\). Therefore,

\[
P((S_n \setminus B_n) \cup (B_n \setminus S_n)) \leq F_R(r_2)(2r_2\varepsilon_n)^{(d-1)/2}/r_2^{(d-1)/2}, \tag{23}
\]

and, by (14),

\[
nP((S_n \setminus B_n) \cup (B_n \setminus S_n)) \leq (1/\gamma_2)(2\varepsilon_n/r_2)^{(d-1)/2},
\]

tends to zero because \( \gamma_2 \to \infty \), as \( n \to \infty \). \( \square \)

Later on \( \varepsilon_n \) will be chosen such that \( P(W(0) \notin B_n) \to 0 \).

**Note.** In the previous proof, we rely on the following argument. For small \((a_1, \ldots, a_{d-1})\), i.e. for \( a = (a_1, \ldots, a_{d-1}) \) such that \( \|a\| \leq X'/r_2 \), the slope parameter \( a_1 \) of \( W(a_1, \ldots, a_{d-1}) = (y_1, \ldots, y_d) \) can be looked at as the angle between the vectors.
(0, . . . , 0, r_2) and ( . . . , 0, y_i, 0 . . . , 0, r_2). Consequently, in a neighbourhood of the origin in \( \mathbb{R}^d \), for each coordinate \( y_i \) (1 \leq i \leq d - 1) of the process \( W \), \( y_i/r_2 \) grows proportionally to the corresponding slope parameter \( a_i \). We will rely on this "linearity argument" when defining the vertex process \( W^\pi \) below (see Definition 4.1), and later on, when determining the range of the parameter space that brings relevant contribution to \{ \( W_n( a) : a \in \mathbb{R}^{d-1} \) \}.

For all that follows, we restrict attention to the region \( A_n^* \cap B_n \). Now there obviously are thousands of ways to arrange in order the points of the process \{ \( W( a) : a \in \mathbb{R}^{d-1} \) \}. Beware that the reduction to the one-dimensional parameter space brings along the loss of information contained in \( (d - 2) \) parameters. When trying to determine the jump measure, we are faced with the problem that it makes a big and crucial difference whether one of the thrown away coordinates (or even more than one) is relatively small or big, especially whenever the first coordinate is small. Also, this makes clear that we must content with an upper bound. Fortunately, if we are careful enough, we can rule out the situation where any of the remaining \( (d - 2) \) parameter values is bigger than the first, and therefore, we will obtain a reasonable upper bound for the jump measure.

Our construction roughly runs as follows. We will project the points \{ \( W( a) \in B_n : a \in \mathbb{R}^{d-1} \) \} to the hyperplane \( \mathbb{R}^{d-1} \times \{0\} \) to eliminate curvature, and then, scan through the projected region in a specified direction by an object whose sides all have curvature zero. Each time one of the projected vertex points is hit by this object, we decide to see the point. This happens precisely once for each projected point.

The projection of \( B_n \) onto the hyperplane \( \mathbb{R}^{d-1} \times \{0\} \) is homeomorphic to a \( (d - 1) \)-ball \( B_n^\pi \) of radius \( X' \) centered at the origin and preserves distances. This can be seen from the fact that the major contribution of the distance between two vertex points stems from the distance computed only from the first \( (d - 1) \) coordinates and by looking at the polar coordinates of the \( d \)-sphere of radius \( r_2 \) centered at the origin:

\[
\begin{align*}
x_1 &= r_2 \sin \theta_1 \cdot \ldots \cdot \sin \theta_{d-1} \\
x_2 &= r_2 \sin \theta_1 \cdot \ldots \cdot \sin \theta_{d-2} \cos \theta_{d-1} \\
&\quad \vdots \\
x_j &= r_2 \sin \theta_1 \cdot \ldots \cdot \sin \theta_{d-j} \cos \theta_{d-j+1} \\
&\quad \vdots \\
x_{d-1} &= r_2 \sin \theta_1 \cdot \sin \theta_{d-j} \cos \theta_{d-j+1} \\
x_d &= r_2 \cos \theta_1,
\end{align*}
\]
where $0 < \theta_j \leq \pi \ (2 \leq j \leq d - 1)$ and $-\pi < \theta_1 \leq \pi$. For sufficiently small $|\theta_1|$, say, $\leq X'/r_2$, we have $\sin \theta_1 \approx \theta_1$ and $\cos \theta_1 \approx 1$, the error term being of order no larger than $\theta_1^2 \leq (X'/r_2)^2$. Consequently, $\{(x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1} : \theta_1 = \theta_*, \theta_* \leq X'/r_2, \ 0 < \theta_j \leq \pi, \forall 2 \leq j \leq d - 1\}$ is equal to the $(d - 1)$-sphere of radius $r_2\theta_1$ centered at the origin, up to error terms no larger than $(X'/r_2)^2$ and $\{(x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1} : \theta_1 \leq X'/r_2, \ 0 < \theta_j \leq \pi, \forall 2 \leq j \leq d - 1\}$ is approximately equal to the ball $B_n^*$ in $\mathbb{R}^{d-1}$ of radius $r_2(X'/r_2)$ centered at the origin.

In short, the above "projection argument" tells us that the change of each of the coordinates under the projection is "negligible" under our point of view. Henceforth, we can project all points of the process $W$ that lie in $B_n$ to $\mathbb{R}^{d-1} \times \{0\}$ such that its local jump behaviour is preserved. Moreover, the projected vertices are approximately uniformly distributed on $B_n^*$ by the rotational invariance. In particular, there is no preferred direction for the points to fall. We may choose an arbitrary direction and, while walking through $B_n^*$ diametrically, count the number of jumps of the process $W$ in $B_n^*$.

The last order of business is to define in which order we see the projected vertices in $B_n^*$. Near at hand would be to scan through $B_n^*$ by a hyperplane in $\mathbb{R}^{d-1}$, which is not far from the geometrical object we have in mind.

Define the following half-closed convex sets in $\mathbb{R}^{d-1}$

$$C(0) = \{(u_1, u_2, \ldots, u_{d-1}) \in \mathbb{R}^{d-1} : 0 \leq u_1, 0 \leq u_k \leq u_1, \forall 2 \leq k \leq d - 1\}$$

$$C^*(0) = \{(u_1, u_2, \ldots, u_{d-1}) \in \mathbb{R}^{d-1} : u_1 = v_1, u_k = \pm v_k, 2 \leq k \leq d - 1, \text{ where } (v_1, v_2, \ldots, v_{d-1}) \in C(0)\}$$

$$\tilde{C}(0) = C(0) \cup C^*(0)$$

$$\tilde{C}(a) = \tilde{C}(0) + (a, 0, \ldots, 0)$$

(24)

For a subset $A \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$, let $\text{dist}(x, A) = \min\{\|x - y\|, y \in A\}$, and let $\tilde{C}(a)$ denote the complement of $\tilde{C}(a)$ in $\mathbb{R}^{d-1}$. After this rather long passage of motivation, we finally are in a position to give a suitable version of a local vertex process.

**Definition 4.1** Let $\tilde{Z}_1, \tilde{Z}_2, \ldots$, denote a realization of the Poisson point process $\xi_{[B_n]}$, and, for each $1 \leq i$, let $Z_i^* = \pi_{\mathbb{R}^{d-1} \times \{0\}}(\tilde{Z}_i)$ be the projection of $\tilde{Z}_i$ onto $\mathbb{R}^{d-1} \times \{0\}$. Furthermore, let $\tilde{Z}_i = Z_i^* \times V_i, \forall i$. Finally, for each $c \in \mathbb{R}$, let $V(c)$ be the point $Z_k^*$ of $Z_1^*, Z_2^*,$ such that $V(c) \subset \tilde{C}(c)$ and $\text{dist}(V(c), \tilde{C}(c))$ is minimal. If there are several of such points, then we take the one with the smallest first coordinate. This happens with probability zero for fixed $c$.

Then, for each $c \in \mathbb{R}$, define the "local vertex process" $W(c) = (V(c), V_k)$. 

11
(The superscript \( \pi \) is supposed to indicate that we locally project the vertices that are close to the path.) It is clear that the process \( \{W^\pi(a) : a \in \mathcal{R}\} \) is Markovian and right-continuous. Our “path” is sausage-like and has positive \((d-1)\)-dimensional volume measure. Remark that the values of the original vertex process are not changed. It is easy to see that we count the right number of jumps in \(B_n^a\) when scanning through \(B_n^a\) by \(\tilde{C}(a)\), i.e. when \(a\) runs through \([-X'/r_2, X'/r_2]\), after having enlarged the region \(B_n^a\) by the neglected region \(\tilde{C}(0)^c \cup B_n^a\) appropriately.

**Note.** The above definition of the vertex process \(W^\pi\) is understood to make sense only for a sufficiently small slope parameter. Otherwise, with a probability approaching 1, the convex hull vertices will come from \(A_n^* \setminus B_n\).

## 5 Jump Measure

This section is concerned with the derivation of an estimate for the local jump measure of the vertex process \(W^\pi\) (and whence, of \(W\)) and with establishing two expressions in terms of supermartingales that will provide upper bounds for the first two moments of the number of points in \(B_n\). Throughout this section, we will always employ the paraboloid approximation and assume, if not stated otherwise, that the time parameter \(a \in \mathcal{R}\) 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 our path. Then the jump measure mimics a \((d-1)\)-dimensional measure. 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. We only need keep track of the length of the path to compute the volume of the tube.

Denote the common density of the vectors \(Z_1, Z_2, \ldots\) by \(f(\cdot)\). For each \(a > 0\) and \(b > 0\) with \(a < b\), define the \(\sigma\)-algebra

\[
\mathcal{F}_{a,b} = \sigma\{W^\pi(c) : a \leq c \leq b\}.
\]

(25)

Define the sector

\[
S_0 = \{(u_1, u_2, \ldots, u_d) \in \mathcal{R}^d : u_d \geq (u_1^2 + u_2^2 + \ldots + u_{d-1}^2)/2r_2\},
\]

and let \(C_0 = C_0(S_0)\) be the set of continuous functions \(g : S_0 \to \mathcal{R}\) with compact support contained in \(S_0\). Furthermore, for each \(a_1 > 0\) and \((x_1, x_2, \ldots, x_{d-1}, y) \in S_0\), let the linear operator \(L_{a_1} : C_0 \to C_0\) be defined by

\[
[L_{a_1}g](x_1, x_2, \ldots, y) = n(2X')^{(d-2)} \sqrt{d-1} \int_0^{X_0-x_1} u
\]

(26)
\[ f(x_1 + u, x_2, \ldots, x_{d-1}, r_2 - y - \sqrt{d-1}a_1 u) \]
\[ \cdot [g(x_1 + u, x_2, \ldots, x_{d-1}, y + \sqrt{d-1}a_1 u) - g(x_1, x_2, \ldots, x_{d-1}, y)]du, \]

where \( X_0 \) is the "bigger" intersection of the approximating paraboloid \( v = (u_1^2 + u_2^2 + \ldots + u_{d-1}^2)/2r_2 \) with the line of slope \((a_1, 0, \ldots, 0)\) through \((x_1, x_2, \ldots, y)\).

**Proposition 5.1** For each \( g \in C_0 \) and \( b_1 > 0 \), the process

\[ \{Y_g(b_2) = g(W^\pi(b_2)) - \int_{b_1}^{b_2} [L_c g](W^\pi(c))dc : b_2 \geq b_1\} \]

is a supermartingale with respect to the filtration \( \{\mathcal{F}_{b_1,b_2} : b_2 \geq b_1\} \).

**Proof.** It suffices to show that for each \( a_1 > 0 \) and \( w = (x_1, x_2, \ldots, x_{d-1}, y) \in B_n \),

\[ \lim_{h \to 0} \frac{1}{h} E\{g(W^\pi(a_1 + h)) - g(w)|W^\pi(a_1) = w\} \leq [L_{a_1} g](w). \]  \hspace{1cm} (27)

Since relation (27) is crucial for our conclusions about the jump behaviour of the vertex process, the arguments will be developed in detail. For each \((x_1, x_2, \ldots, y) \in B_n\), define

\[ A_h = \{(u_1, \ldots, u_d) \in \mathbb{R}^d : au_1 + (y - ax_1) \leq u_d \leq (a + h)u_1 + (y - ax_1), \]
\[ |u_j| \leq X', 2 \leq j \leq d - 1\}. \]

The probability of finding more than one point of the Poisson point process \( \xi_n \) in \( A_h \) is \( o(h) \). Hence, for \( g \in C_0 \),

\[ E\{g(W^\pi(a_1 + h)) - g(w)|W^\pi(a_1) = w\} \]
\[ = n \int_{A_h} f(v)\{g(v) - g(w)\}dv + o(h) \]
\[ \leq nh\sqrt{d-1}\int_{[-X',X]^d-2} \int_0^\infty u_1 \]
\[ \cdot f(x_1 + u_1, x_2 + u_2, \ldots, x_{d-1} + u_{d-1}, r_2 - y - \sum_{i=1}^{d-1} a_i u_i) \]
\[ \cdot \{g(W^\pi(a_1 + h)) - g(w)\}du_1 du_2 \ldots \cdot du_{d-1} + o(h) \]
\[ \leq nh\sqrt{d-1}\int_{[-X',X]^d-2} \int_0^\infty u_1 \]

13
\[
\begin{aligned}
&= nh(2X')^{d-2}\sqrt{d-1} \int_0^\infty u \\
&\quad \cdot f(x_1 + u, x_2, \ldots, x_{d-1}, r_2 - y - \sqrt{d-1}a_1u) \\
&\quad \cdot \{g(W^n(a_1 + h)) - g(w)\} du + o(h)
\end{aligned}
\]

The reasoning in the above chain of equalities and inequalities runs as follows. In deriving the above conditional expectation, we need know, first, the path that we follow, and second, the value of the density \( f \) at the new point \( W^n(a_1 + h) \). Remember that when defining the process \( W^n \), we are thinking of following a path along the \( x_1 \)-axis and of seeing the convex hull vertices in the neighbourhood \( B^n \) of the path in a certain unambiguously defined order, which is determined by the shape of \( \hat{C}(\cdot) \), i.e. at time \( a_1 \), we see \( W^n(a_1) \). By the rotational invariance, we can choose any path, in particular, the path along the first coordinate axis.

Instead of moving to the new point (which would be a deviation of our path), we project \( W^n(a_1 + h) = (V_1, V_2, \ldots, V_{d-1}, V_d) \) to \( (V_1, x_2, \ldots, x_{d-1}, V_d) \), where \( V_d \) is such that \( f(V_1, V_2, \ldots, V_{d-1}, r_2 - V_d) = f(V_1, x_2, \ldots, x_{d-1}, r_2 - V_d) \). \( hV_d - y \) is the vertical distance at the point \( W^n(a_1 + h) \) between the hyperplanes \( H(a_1 + h, 0, \ldots, 0) \) and \( H(a_1, 0, \ldots, 0) \) through \((x_1, x_2, \ldots, y)\). Although we do not know the vector \((a_2, \ldots, a_{d-1})\), because we reduced the parameter space to the one-dimensional parameter space of the process \( W^n \), we have \( a_h \leq a_1 \) (see Definition 4.1), and therefore, we can find an estimate for \( V_d \) in the following way. Assume that \( V_d = y + \sum_{i=1}^{d-1} a_iu_i \). Observe that we are looking for an upper bound for the density \( f \), and thus, for an upper bound for \( \sum_{i=1}^{d-1} a_iu_i \) for each \( a = (a_1, a_2, \ldots, a_{d-1}) \) such that \( \sum_{i=1}^{d-1} a_i^2 \leq (X'/r_2)^2 \) and for each \( u = (u_1, u_2, \ldots, u_{d-1}) \) such that \( \sum_{i=1}^{d-1} u_i^2 \leq (X')^2 \). Now by Definition 4.1, \( \sum_{i=1}^{d-1} a_iu_i \leq a_1 \sum_{i=1}^{d-1} u_i \), and

\[
\sup_{(u_1, \ldots, u_{d-1})} \sum_{i=1}^{d-1} u_i \leq \sqrt{d-1} \max_{1 \leq i \leq d-1} u_i.
\]
Since $u_k$ grows proportionally to $a_k$, $\max_i u_i = u_1$. Consequently, we have $\sum_{i=1}^{d-1} a_i u_i \leq \sqrt{d-1} a_1 u_1$. Finally, to obtain the first inequality, we also notice that, when turning the supporting hyperplane $H(a_1, \ldots, 0)$ infinitesimally by changing the slope from $a_1$ to $a_1 + \epsilon$ and jumping by $u_1$, we increase the $d$-th coordinate by less than $h u_1 \sqrt{d - 1}$; to get the second, it as well remains to be seen that in order that the norm of the vector $(a_1 u_1, x_2, \ldots, x_{d-1}, r_2 - y - \sum_{i=1}^{d-1} a_i u_i)$ be small, $\sum_{i=1}^{d-1} a_i u_i$ must be positive and large, which implies that $\sum_{i=1}^{d-1} u_i (2x_i + u_i)$ is nonnegative. Moreover, for the last equality, we observe that the integration of the first coordinate above $X_0$ is negligible by Lemma 4.1. Since $g$ has compact support, the remainder term is $o(h)$, uniformly in $a_1$, for $|a_1| \leq X'/r_2$ 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 our Proposition.

We now define the random counting measure $\eta(a, b; \cdot)$ by

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

(28) 

for $b > a > 0$ and $B \in B^d$ (the family of Borel sets in $\mathbb{R}^d$). Clearly, for $0 < a < b \leq X'/r_2$, the identity

$$\eta(a, b; B) = \sum_{\begin{substack}{w(c, \ldots, 0) \neq w(c-, \ldots, 0), a \leq c \leq b} \end{substack}} 1_B (W(c, 0, \ldots, 0) - W(c-, 0, \ldots, 0)),$$  

(29) 

holds since the reformulation of the vertex process in Definition 4.1 does not affect the values of the vertex process $W$. Moreover, let $\eta_a$ denote the number of jumps of $W^\pi$ in the time interval $[0, a]$:

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

(30) 

Denote the jump measure of the process $W^\pi$ in $w \in \mathbb{R}^d$ at time $a > 0$ by

$$M^\pi(W^\pi(a), w; B)$$  

(31) 

for $B \in B^d$, which can be interpreted as the probability that the process $W^\pi$ jumps from $w$ to $w + B$ at time $a$. Unfortunately, as explained before, we do not know anything about $M^\pi(W^\pi(a), \cdot; \cdot)$. However, the infinitesimal generator of $W^\pi$ is dominated by the infinitesimal generator $L_a$ of an unknown jump process $W_*$ with known jump measure $M(\cdot; \cdot; \cdot)$. The process $W_*$ is uniquely determined by its jump measure given
in (32) below. The operator $L_a$ can be written with respect to the jump measure $M(\cdot; \cdot)$ in the following way

$$[L_ag](w) = \int_{\mathbb{R}^d} [g(w + z) - g(w)] M(w; dz).$$

(32)

To state the Proposition and Lemma below, we need introduce

$$\eta_*(a, b; B) = \sum_{\substack{w_*(c) \neq w_*(c-) \in \mathcal{B}^d \cap B \leq b}} 1_B(W_*(c) - W_*(c-)),

(33)$$

for $b > a > 0$ and $B \in \mathcal{B}^d$, where $W_*$ is the process with jump measure defined in (32), correspondingly,

$$\eta_* = \eta_*(0, a; \mathbb{R}^d),

(34)$$

and the $\sigma$-algebra

$$\mathcal{F}_{a, b}^* = \sigma\{W_*(c) : a \leq c \leq b\}.

(35)$$

The next result, due to Stroock ([11], Theorem 1.3), is important to our analysis.

**Proposition 5.2 (Stroock)** For each bounded Borel measurable function $g : \mathbb{R}^d \to \mathbb{R}$ that vanishes in a neighbourhood of the origin, for each $a > 0$ and $\theta \in \mathbb{R}^d$, the process

$$\exp \left\{ <i\theta, W_*(b) - W_*(a)> + \int_{\mathbb{R}^d} g(z) \eta_*(a, b; dz) \right. \left. - \int_a^b dc \int_{\mathbb{R}^d} \{e^{i\theta, z} - \frac{1}{2} M(W_*(c); dz) \} \right\}, \quad b \geq a$$

is a martingale with respect to the filtration $\{\mathcal{F}_{a, b}^* : b \geq a\}$, where $W_*$ is the process with jump measure defined in (32), $< x, y >$ denotes the inner product on $\mathbb{R}^d$, and where $i = \sqrt{-1}$.

We do not reproduce Stroock's proof here. The uniform boundedness of the random measure $M(W_*(\cdot; \cdot))$ is used, which follows by virtue of Proposition 5.1, by (26) and (32). The above martingale is helpful at tackling the asymptotic moments of certain convex hull functionals. As a consequence of Proposition 5.2, we find
Lemma 5.3 For $b > 0$

\[ \{ \eta_{*b} - \int_0^b M(W_*(c); \mathbb{R}^d) \, dc : b > 0 \} \] (36)

and

\[ \{ \eta_{*b}^2 - \int_0^b (2\eta_{*} + 1) M(W_*(c); \mathbb{R}^d) \, dc : b > 0 \} \] (37)

are $\mathcal{F}_{0,b}^*$-martingales.

Proof. Write $J(\lambda)$ for the martingale in Proposition 5.2 when $\theta = 0$, $a = 0$ and $g \equiv \lambda$. Then once differencing $J(\lambda)$ with respect to $\lambda$ and setting $\lambda = 0$ lead to the first assertion (36), by (30).

The second claim is a consequence of the properties of the Poisson point process (and can also be found in [6], Lemma 2.6). It is convenient to define the quantity $\eta_{*a,b} = \eta_{*b} - \eta_{*a}$, for $a \leq b$. We have

\[
E \left\{ \eta_{*b+h}^2 - \eta_{*b}^2 \mid \mathcal{F}_{0,b}^* \right\} = E \left\{ \eta_{*b+h}^2 + 2 \eta_{*b} \eta_{*b+h} \mid \mathcal{F}_{0,b}^* \right\} = E \left\{ (1 + 2\eta_{*b}) (h M(W_*(b); \mathbb{R}^d) + R_h) \mid \mathcal{F}_{0,b}^* \right\},
\]

where we use the relation $E[\eta_{*b+h} \mid \mathcal{F}_{0,b}] \sim h M(W_*(b); \mathbb{R}^d)$ (see (36)), 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 $0 < h < 1$,

\[
h M(W_*(b); \mathbb{R}^d) \exp \{-h M(W_*(b); \mathbb{R}^d) \} \leq E \{ \eta_{*b+h}^2 \mid W_*(b) \} \leq h M(W_*(b); \mathbb{R}^d) + h^2 [M(W_*(b); \mathbb{R}^d)]^2.
\]

An easy exercise now finishes the proof. \qed

Corollary 5.4 For $b > 0$

\[ \{ \eta_{b} - \int_0^b M(W_*(c); \mathbb{R}^d) \, dc : b > 0 \} \] (39)

and

\[ \{ \eta_{b}^2 - \int_0^b (2\eta_{*} + 1) M(W_*(c); \mathbb{R}^d) \, dc : b > 0 \} \] (40)

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

Proof. Since the integrals in (39) and (40), respectively, essentially depend on the jump measure, in light of Lemma 5.1 and (32), we can conclude that, for each $a > 0$, $w \in B_n$ and $B \in B^d,$

\[ M_\tau(W^*(a), w; B) \leq M(W_*(a), w; B). \]

Whence (39) and (40) follow.
Before bringing this section to an end, we wish to transform the local jump measure \( M(W_*(a) : \cdot) \) such that it is independent of the time parameter \( a \), i.e. that it is the jump measure of a stationary local process. Assume \( W_*(a) = (X_1(a), X_2(a), \ldots, X_d(a)) \). Let us transform the process \( \{W_*(a)\} \) into the process \( \{T_*(a) = (R_1(a), R_2(a), \ldots, R_d(a))\} \) by the following substitution

\[
\begin{align*}
R_1(a) &= X_1(a) - ar_2\sqrt{d-1} \quad (41) \\
R_k(a) &= X_k(a) \quad (2 \leq k \leq d-1) \\
R_d(a) &= X_d(a) - a\sqrt{d-1}X_1(a) + a^2r_2(d-1)/2.
\end{align*}
\]

Let \( K_d(y) \) be the fraction of the surface area of a unit \( d \)-sphere cut off by a hyperplane at distance \( y > 0 \) and

\[
G_2(x) = \int_x^\infty K_2(x/y)|dF_R(y)|. \quad (42)
\]

Following Dwyer ([3], Section 1 and 3) and setting \( d = 2 \), we obtain under the conditions in (10), for sufficiently large \( x \),

\[
G_2(x) \sim (2\pi)^{-1/2}\sqrt{\nu(x)}F_R(x), \quad (43)
\]

\[
G_2'(x) = \frac{\partial}{\partial x}G_2(x) \sim -(2\pi)^{-1/2}F_R(x)/(x\sqrt{\nu(x)}). \quad (44)
\]

Then the local infinitesimal generator becomes, for each continuously differentiable real-valued function \( g \) with compact support, defined on \( S_0 \),

\[
\lim_{h \to 0} \frac{1}{h} E\{g(T_*(a + h)) - g(x_1, x_2, \ldots, x_d) \mid T_*(a) = (x_1, x_2, \ldots, x_d)\} \quad (45)
\]

\[
\leq -n(2X')(d-2)\sqrt{d-1} \int_0^{\sqrt{2r_2x_d-x_1}} u \\
\cdot G_2'(\| (x_1 + u, x_2, \ldots, r_2 - x_d) \|) \\
\cdot [g(x_1 + u, x_2, \ldots, x_d) - g(x_1, x_2, \ldots, x_d)]du \\
- r_2\sqrt{d-1} \frac{\partial}{\partial x_1}g(x_1, \ldots, x_d) \\
- x_1\sqrt{d-1} \frac{\partial}{\partial x_d}g(x_1, \ldots, x_d).
\]
The process $T_*$ 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 $(0,0,\ldots,r_2)$ while the coordinate system is fixed.

Lemma 5.5 As $n \to \infty$,
\begin{align}
E[\eta_b] & \leq b \cdot E[M(T_*(0); \mathbb{R}^d)], \tag{46} \\
E[\eta_b^2] & \leq E[\eta_b] + 2 \int_0^b da \int_0^a E[M^*(T_*(0); \mathbb{R}^d) M(T_*(a-c); \mathbb{R}^d)] \, dc, \tag{47}
\end{align}
where $T_*$ is the process transformed above from $W_*$ described in (32) and $M^*(\cdot; \cdot)$ is the "backward jump measure" of $T_*$. \hfill \Box

Proof. In view of Lemma 5.3 and (45) with $g \equiv 1$ it remains to verify the second claim (47). Consider the time reversed process \{(-R_1(a-c), R_2(a-c), \ldots, R_d(a-c)) : c \geq 0\}. The process \{(-R_1(c), R_2(c), \ldots, R_d(c)) : c \leq 0\} is recovered from the process \{T_*(c) = (R_1(c), R_2(c), \ldots, 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 $T_*$. Let $M^*(\cdot; \cdot)$ denote the jump measure of the time reversed process. Now (47) follows from the stationarity of $T_*$ and Corollary 5.4. \hfill \Box

As yet, we have not made use of the special form of the density $G_0^*$. Note that another choice of the function $g$ in Proposition 5.2 would lead to results concerning other convex hull functionals like obtained in Hueter [7] for the perimeter and the area of the planar convex hull.

6 Proofs of Theorem 1.1 and 1.2

For the proof of the central limit theorem for $N_n$, two lemmata will be needed. Assume that we already know how to "walk over the whole surface" of the convex hull, i.e. 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" as will become clear shortly. Since the number of jumps within each element of the partition are identical random
variables by the rotational invariance, if the elements have equal size, we can build a stationary sequence of random variables. Again by the spherical symmetry of the distribution there must exist a (local) transformation of \( W^x \) into another local jump process which is stationary. Call this process \( T^x \). We shall show that the process \( \{T^x(a) : a \in \mathcal{R}\} \) is strongly mixing and the dependence decreases exponentially fast to zero with increasing "distance", for sufficiently large \( n \).

In showing that the whole 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, i.e. for all elements visited by \( \{T^x(a) : a \in \mathcal{R}\} \). Then, in terms of mixing coefficients, each of the remaining elements is equivalent to one element already furnished with a mixing coefficient. It is easily seen that the summation of all mixing coefficients, raised to a certain power, converges (This is one of the conditions that need be satisfied for the CLT).

**Lemma 6.1** Let \( \mathcal{F}_0 = \sigma \{T^x(c) : c \leq 0\} \), \( \mathcal{F}_a^+ = \sigma \{T^x(c) : c \geq a\} \), and let \( A \in \mathcal{F}_0 \) and \( B \in \mathcal{F}_a^+ \). Then, for some constants \( 0 < c', c'' < \infty \),

\[
|P(A \cap B) - P(A)P(B)| \leq \tau_n(a) \tag{48}
\]

where

\[
\tau_n(a) \leq 4 \exp \left\{ nG'_2(r_0 + \varepsilon_n/2 - a^2 c' r_2/16) \cdot (2X')^{d-2} a^3 r_2^2 c'' \right\} \xrightarrow{n \to \infty} 0.
\]

**Proof.** Let \( T^x(a) = (S_1(a), S_2(a), \ldots, S_d(a)) \), for each \( a > 0 \). Assume that at time \( a > 0 \) 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 (48) 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 \( \xi_n \).

In fact, the point \( T^x(a) \) 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 \( \xi_n \) in the region \( K_a \) between the paraboloid and the hyperplane, namely, in the region

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

20
Consequently, if for any \( a > 0 \) \( T^*(0) \) and \( T^*(a) \) 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 \( \{T^*(a) : a \in \mathbb{R} \} \),

\[
P(A \cap B|E) = P(A|S_d(0) < a^2 \gamma^2 r_2/8) \, P(B|S_d(a) < a^2 \gamma^2 r_2/8).
\]

Now, the bigger the distance of the point \( T^*(a) \) to the \( x_1 \)-axis the smaller the dependence from the point \( T^*(0) \). Thus, we do not decrease the dependence when choosing \( T^*(a) \) on the \( x_1 \)-axis from all possible values of \( T^*(a) \) in accordance to Definition 4.1. Suppose so. Then \( P(S_d(a) \geq a^2 \gamma^2 r_2/8) = P(S_d(0) \geq a^2 \gamma^2 r_2/8) \) by the stationarity of \( T^* \). Therefore, by applying some rather crude estimates, we get

\[
|P(A \cap B) - P(A)P(B)| \leq P(E^c) + 2P(S_d(0) \geq a^2 \gamma^2 r_2/8) \leq 4P(S_d(0) \geq a^2 \gamma^2 r_2/8),
\]

where \( E^c \) denotes the complementary event to \( E \). However, \( \{ S_d(0) \geq a^2 \gamma^2 r_2/8 \} \) if and only if the region \( K_0 \) contains no points of the Poisson point process \( \xi_n \). We are going to give a lower bound for the probability \( P(K_0) \), by using the intermediate value theorem and the fact that the tail of the distribution decreases exponentially fast.

Let \( \lambda_d \) denote the Lebesgue measure in \( \mathbb{R}^d \). Then, for some \( 0 < \delta < 1/2 \),

\[
P(K_0) = -G^*_d(r_0 + \epsilon_n/2 - (1 - \delta)a^2 \gamma^2 r_2/8) \, \lambda_d(K_0)
\]

\[
\geq -G^*_d(r_0 + \epsilon_n/2 - a^2 \gamma^2 r_2/16) \, \lambda_d(K_0)
\]

\[
= -G^*_d(r_0 + \epsilon_n/2 - a^2 \gamma^2 r_2/16) (2X')^{d-2} a^3 \gamma^3 r_2^2 c^{''''}
\]

for a constant \( 0 < c^{''''} < \infty \). Hence,

\[
P(S_d(0) \geq a^2 \gamma^2 r_2/8) = \exp\{-nP(K_0)\} \leq \exp\{nG^*_d(r_0 + \epsilon_n/2 - a^2 \gamma^2 r_2/16) (2X')^{d-2} a^3 \gamma^3 r_2^2 c^{''''}\}.
\]

Since by (44), for large \( x \), \( G^*_d(x) \sim -(2\pi)^{-1/2}F_d(x)/(x\sqrt{\nu(x)}) \) and \( \nu(u) \to 0 \), as \( u \to \infty \), the right-hand side of the last inequality decreases to zero exponentially fast in \( a \) iff \( a > 2(2\epsilon_n/r_2)^{1/2}/\gamma \).

Recall that \( \eta_a \) denotes the number of jumps of the process \( \{W^*(c) : 0 < c \leq a\} \), and thus, of \( \{T^*(c) : 0 < c \leq a\} \) (see (30)).
Lemma 6.2 For each $a > 0$ the moment generating function

$$M(\lambda) = E \exp\{\lambda \eta_a\}$$

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

Proof. The proof makes an easy exercise. We give a sketch here. With probability one there exists some $0 < b_n < \infty$ such that $P(\{T^\ast(a) : a > b_n\} = \emptyset) \rightarrow 1$ as $n \rightarrow \infty$ because $F$ is absolutely continuous. $\eta_a$ can not be bigger than $c'a/b_n$ times the number of points of the Poisson point process $\xi_n$ in $A_n^\ast$, where $c'$ is a positive finite constant. This number of points has a Poisson distribution with parameter $nP(A_n)\cdot c'a/b_n$. Furthermore, $P(A_n) \leq F_R(r_1)$. Hence, $nP(A_n) \leq \gamma_1$, where $\gamma_1$ is defined in (14) ($\gamma_1$ is a function of $n$ of order $o(n)$). Consequently, we have

$$E \exp\{\lambda \eta_a\} \leq \exp\{\exp\{\lambda^\gamma_1 - 1\} \gamma_1 ac'/b_n\}$$

for every $0 < \lambda < \infty$. \hfill \square

Proposition 6.3 Let $\eta_b$ be the number of jumps of the process $\{T^\ast(c) : 0 \leq c \leq b\}$. Assume $Z_1, Z_2, \ldots$ is a sequence of i.i.d. random vectors, each with distribution $F$, where $F$ is an exponentially-tailed, spherically symmetric distribution such that the smoothness conditions in (10) hold.

(i) Then as $n \rightarrow \infty$, for sufficiently small $b > 0$,

$$E(\eta_b) \sim b \tilde{c}_1,$$

and

$$\text{Var}(\eta_b) \sim b \tilde{c}_1 + b^2 \tilde{c}_2,$$

where $\tilde{c}_1 \leq 2^{d-3}\sqrt{d-1}(X')^d/(\sqrt{2\pi} \nu(r_0)r_0)$ and $\tilde{c}_2 \leq c_2(d-1) \nu(r_0)r_0^{d-1}$ and $c_2$ is some positive finite constant.

(ii) $\eta_b$ converges, appropriately normalized, to a standard normally distributed random variable, i.e.

$$(\eta_b - \tilde{c}_1 b)/((\tilde{c}_1 b + \tilde{c}_2 b^2)^{1/2} \rightarrow \mathcal{N}(0,1),$$

where $\tilde{c}_1$ and $\tilde{c}_2$ are as in (i) above and $\mathcal{N}(0,1)$ denotes the standard normal distribution.
Proof. ad(i) Remember $G_2'(x) \sim -(2\pi)^{-1/2}F_R(x)/(x\sqrt{\nu(x)})$. Let $T_*(0) = (x_1, \ldots, x_d)$. From (45) we derive

$$M(T_*(0); \mathbb{R}^d) = -n(2X')^{(d-2)}\sqrt{d-1} \int_0^{\sqrt{2\pi x_d-x_1}} u \cdot G_2'([x_1+u, x_2, \ldots, r_2-x_2])du$$

$$\leq -n2^{d-3}(X')^d\sqrt{d-1}G_2'(r_0)$$

$$\sim n2^{d-3}(X')^d\sqrt{d-1} F_R(r_0)/(\sqrt{2\pi}r_0\sqrt{\nu(r_0)})$$

where we take advantage of the fact that $G_2'(u) \sim G_2'(r_0)$ for $u \in (r_1, r_2)$, as $n \to \infty$, which is a consequence of the slow variation of the functions $L$ and $\nu$, and thus, of $L(\cdot)\sqrt{\nu(L(\cdot))}$, and the fact that integration above $X'$, and therefore, outside the region $A_{n*}$ is negligible. Recall that $nF_R(r_0) = 1$. Now apply Lemma 5.5.

Moreover, also by Lemma 5.5 and the intermediate value theorem, for any sufficiently small $b > 0$,

$$E[\eta_b^2] \leq E[\eta_b] + b^2E[M^*(T_*(0); \mathbb{R}^d) M(T_*(a'); \mathbb{R}^d)]$$

for some $a' \in (0, b)$. It is clear that both measures $M^*(T_*(0); \mathbb{R}^d)$ and $M(T_*(a'); \mathbb{R}^d)$ are of order $O((X')^d/(r_0\sqrt{\nu(r_0)}))$ by the same arguments again. Since $X' \neq o(1)$ and the order $O(1/(r_0\sqrt{\nu(r_0)}))$ is the order of a lower bound for $M(T_*(0); \mathbb{R}^d)$ (for otherwise it is not hard to see that $E[N_n] = o(\varepsilon(n)^{-(d-1)/2})$ must hold, which is a contradiction to the known results), the upper bound for the variance in (i) follows.

ad(ii) Partition the surface of the $d$-sphere of radius, say, $r_0$ into “equal” elements, i.e. such that each element $\tilde{S}_i$ is a congruent copy of a specific element $\tilde{S}_0$ of the partition (Each element is a homeomorphic image of a nonempty bounded subset in $\mathbb{R}^{d-1}$). Let $P_i$ be the smallest “cone”, with its top vertex at the center of the sphere that contains $\tilde{S}_i$, and let $M_i$ be the number of points of the process $T^\pi$ in $P_i$. Then the $M_i$ are identically distributed random variables by the rotational invariance, and the distribution of $M_i$ only depends on the size of $\tilde{S}_0$. Suppose that $\tilde{S}_0$ has surface area equal to the volume area of $[0,1] \times [-X', X']^{d-2}$.

Furthermore, assume that the $M_i$ are indexed such that our path crosses them in strictly increasing order, i.e. first $M_1$, then $M_2$, $M_3$ and so forth. The sequence $M_1, M_2, \ldots$ satisfies the mixing condition of Lemma 6.1 for $A \in \sigma(M_1, M_2, \ldots, M_j)$ and $B \in \sigma(M_{j+1}: l > m)$ with mixing coefficient $\tau_n(h(m))$ for some positive real-valued function $h$. Note that $EM_t^k$ is finite for each $k \geq 1$ by Lemma 6.2. Thus, the conditions for a central limit theorem for the sequence $M_1, M_2, \ldots$ are satisfied (see e.g. Ibragimov and Linnik [8], Theorem 18.5.3). As $\eta_k = \sum_{i=1}^k M_i$ and $\eta_k - \eta_l$ only
depends on the difference \( k - l \) again by the rotational invariance, (i) now proves the limit in (54). \(\square\)

**Proof of Theorem 1.1.** The number \( N_n \) of convex hull vertices corresponds to the number of points in the set \( \{ W_n(a) : \| a \| \leq r_0 c_d \} \), where \( c_d \) is such that the \((d - 1)\)-ball of radius \( r_0 c_d \) has volume area equal to \( r_0^{d-1} \). We will show that with high probability the number of points in the set \( \{ T^n(a) : a \in [-\kappa_d r_0 (2X'\sqrt{\nu(r_0)})^{-(d-2)}/2, \kappa_d r_0 (2X'\sqrt{\nu(r_0)})^{-(d-2)}/2} \} \) is equal to the number of points in the set \( \{ W_n(a) : \| a \| \leq r_0 c_d \} \).

For the next part of the proof some more notation is called for. Let \( B_* \) be the \((d - 1)\)-ball of radius \( r_0 c_d \) centered at the origin. Define

\[
m_* = \kappa_d r_0 (2X'\sqrt{\nu(r_0)})^{-(d-2)}
\]

and the interval \( I_* = [-m_*/2, m_*/2] \). Remember \( \eta_n \) to be the sample point process, and \( \xi_n \) the Poisson point process on \( \mathbb{R}^d \) with intensity measure \( n \int dF \). In Lemma 3.1 also the point processes \( \tilde{\eta}_n \) and \( \tilde{\xi}_n \) were considered. Let \( W'_n, \tilde{W}_n, \) and \( W \) be the vertex processes based on the processes \( \eta_n, \tilde{\eta}_n \) and \( \tilde{\xi}_n \), respectively. Next define the following counters

\[
\begin{align*}
N_n &= \# \{ W_n(a) : a \in B_* \} \\
\tilde{N}_n &= \# \{ \tilde{W}_n(a) : a \in B_* \} \\
N'_n &= \# \{ W'_n(a) : a \in B_* \} \\
\tilde{L}_n &= \# \{ W(a) : a \in B_* \} \\
L^n \pi &= \# \{ T^n(a) : a \in I_* \}.
\end{align*}
\]

Since the processes \( W^n \) and \( W \) have the same local jump measure as described, for instance, by the relation (29), \( L^n \pi \) has the same distribution as \( \tilde{L}_n \) whenever the range of their parameter space corresponds to each other. This argument will be given shortly. By Lemma 3.1, \( \lim_{n \to \infty} P(\tilde{N}_n \neq \tilde{L}_n) = 0 \). Furthermore, \( N'_n \) and \( \tilde{N}_n \) have the same distribution and by (11) \( \lim_{n \to \infty} P(N'_n \neq N_n) = 0 \). Therefore,

\[
\lim_{n \to \infty} P(L^n \pi \neq N_n) = 0.
\]

Hence, it suffices to show that

\[
(L^n \pi - \alpha_n)/\beta_n \xrightarrow{L} \mathcal{N}(0,1),
\]

for suitably chosen norming constants \( \alpha_n \) and \( \beta_n \) because then

\[
(N_n - \alpha_n)/\beta_n \xrightarrow{L} \mathcal{N}(0,1)
\]

24
follows.

The verification of the limiting law for $L_n^x$, i.e. of a central limit theorem for the stationary mixing sequence $M_1, M_2, \ldots$, uses a "blocking system". Since this technique is quite standard and a careful writing down of every argument would bring along even more additional notation than was necessary in the preceding paragraph of the proof, this step is left to the reader. Roughly, the idea is to cover the parameter space $B_\ast$ by strips along a regular square grid such that the number of points of the process $W$ falling on those strips are asymptotically negligible compared to the total number of points of the process $\{W(a): a \in B_\ast\}$. However, we wish to assure that the paraboloid approximation does no harm to our computations. By Lemma 4.1, it is easily verified that the error is of order no larger than

$$\mathcal{O}(nF_R(r_2)),$$

which is equal to $\mathcal{O}(1/\gamma_2)$ with $\gamma_2 \to \infty$ by (14).

In view of Proposition 6.3, in order to complete the proof we must count the number of "rectangles" in $B_\ast$ implied by the partition $\{\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_{k_n}\}$ of the surface of the convex hull with appropriate $k_n$. On each rectangle such that the length of the side parallel to the $x_1$-axis is equal to $b$, the expectation of the number of points of $W$ is equal to $\tilde{c} \cdot b$ and the variance is equal to $\tilde{c}_1 b + \tilde{c}_2 b^2$, where $\tilde{c}_1$ and $\tilde{c}_2$ are given in Proposition 6.3. The $b$ that we need choose corresponds to $X'$. The "width" $w_n$ of the path the walker follows is determined such that the expected number of points of $W$ on the path intersected with a $d$-cube of side length $w_n$ is bounded away from zero, as $n \to \infty$. That $w_n = (2X')r_0\sqrt{\nu(r_0)}$ can be seen in the following way.

There are two cases to distinguish between. Either it must be the case that the expected jump measure $E[M(T_\ast(0); \mathbb{R}^d)]$ is independent of the dimension, namely, if $X' = \mathcal{O}(1)$, then we may resort to the relation $2\pi r_0/E[N_n] = 2\pi r_0/(2\sqrt{\pi}\nu(r_0)^{-1/2}) = r_0\sqrt{\pi\nu(r_0)}$, which is valid in two dimensions (see [2]), in order to determine the correct order for the side length $w_n$, i.e. $w_n = c_0 r_0\sqrt{\nu(r_0)}$ for some positive finite constant $c_0$, or it must be the case that the order of $E[M(T_\ast(0); \mathbb{R}^d)]$ depends on the dimension $d$. $X' = \mathcal{O}(1)$ corresponds to the case $r_0\sqrt{\nu(r_0)} \neq \infty$, because then on an interval of sufficient large, but constant length, with high probability, the number of convex hull vertices is bounded away from zero, as $n \to \infty$, whereas in the second case, $r_0\sqrt{\nu(r_0)} \to \infty$, which means that, for sure, $X' \to \infty$ in order that, with high probability, there is at least one point of the Poisson point process $\xi_n$ in the region where the paraboloid approximation is applied. Consequently, in the second case, since with each dimension the expected jump measure increases by a factor $X'$, the side length $w_n = 2X'r_0\sqrt{\nu(r_0)}$ is of the right order by again using the relation

25
For either case, we have found \( w_n = 2X' r_0 \sqrt{\nu(r_0)} \) (Set \( c_0 = 2X' \) in the first case).

It remains to find the length of the path. As the number of convex hull vertices in the elements \( P_i \), defined in the proof of Proposition 6.3, only depends on the size of \( \tilde{S}_i \), but not on its shape, we patch together the homeomorphic preimages of all the \( \tilde{S}_i \) such that we wind up with a large rectangle in \( \mathbb{R}^{d-1} \) that has one side length equal to \( r_0 \kappa_d \) and the \( (d-2) \) remaining sides, each of length equal to \( r_0 \). Consequently, the path has length equal to \( m_* \). Hence,

\[
L_n = \eta_{m_*}.
\]

(55)

Now either \( X' = \mathcal{O}(1) \) is large enough because on each interval of length \( l_n \) such that \( l_n \rightarrow \infty \) there are infinitely many, in fact, \( \mathcal{O}(l_n/(r_0 \sqrt{\nu(r_0)})) \) many expected convex hull vertices if \( r_0 \sqrt{\nu(r_0)} \neq \infty \) or \( X' = \mathcal{O}(r_0 \sqrt{\nu(r_0)}) \) suffices because on each interval of length of order \( \mathcal{O}(l_n \sqrt{\nu(r_0)}) \) such that \( l_n \rightarrow \infty \) there are \( \mathcal{O}(l_n) \), thus, infinitely many expected convex hull vertices, whenever the width of the path \( w_n \) is of the right order. Note that \( X' \) corresponds to the value of \( b \) in Proposition 6.3 that we need choose. For the variance of \( N_n \), an additional factor \( X' \) must be taken into account because in (53) the quadratic term is dominating the linear term whenever \( X' \rightarrow \infty \). This concludes our proof.

\[\square\]

**Proof of Theorem 1.2.** The verification runs along the same line used in the proof of Theorem 1.1 with the pleasant difference that an explicit upper bound for the jump measure \( M(T_*(\cdot); \mathbb{R}^d) \) can be computed. Let \( T_*(0) = (x_1, x_2, \ldots, x_d) \). From (45) together with \( G'_2(x) \sim (2\pi)^{-1/2} \exp(-\|x\|^2/2) \), for sufficiently large \( x \), it is easily deduced that

\[
M((x_1, \ldots, x_d); \mathbb{R}^d) \leq n \sqrt{d-1} (2X')^{d-2} \exp(-\|r_2 - x_d\|^2/2) \\
\cdot \int_{0}^{\sqrt{2r_2x_d - x_1}} \frac{(2\pi)^{-1/2} u \exp(-\|x_1 + u\|^2/2) du}{(2X')^{d-2}} \\
\sim n \sqrt{d-1} (2X')^{d-2} \exp(-\|r_2 - x_d\|^2/2)/\sqrt{\pi}.
\]

Rather tedious calculations involving the densities of the vertex process (which can be found in [7]) yield

\[
E[M(T_*(0); \mathbb{R}^d)] \leq \sqrt{d-1}(2X')^{d-2}(\pi)^{-1/2}\]

Note that \( r_0 = L(n) = \nu(r_0)^{-1/2} = \sqrt{2 \ln n} \). Consequently,
\[ E[N_n] \leq \kappa_d r_0(2X' \sqrt{n(r_0)})^{-(d-2)} E[M(T_*(0); \mathbb{R}^d)] \]
\[ \sim \sqrt{d-1} \kappa_d (2 \ln n)^{(d-1)/2}/\sqrt{\pi}. \]

References


