GAFA, Geom. funct. anal.
Vol. 15 (2005) 1284 – 1318
1016-443X/05/061284-35
DOI 10.1007/s00039-005-0541-8
ONLINE FIRST: December 2005

GAFA Geometric And Functional Analysis

SHARP CONCENTRATION OF RANDOM POLYTOPES

V.H. Vu

Abstract. We prove that key functionals (such as the volume and the number of vertices) of a random polytope is strongly concentrated, using a martingale method. As applications, we derive new estimates for high moments and the speed of convergence of these functionals.

1 Introduction

Let K be a convex set with volume one in $\mathbb{R}^d$ and μ the uniform distribution concentrated on K. Choose n random points in $\mathbb{R}^d$ independently according to μ. The convex hull of these points, denoted by K_n, is called a *random polytope*. The study of the key functionals (such as the volume, the number of vertices, etc.) of K_n, started by Efron [Ef] and Rényi and Sulanke [RéS], is a popular topic in convex geometry. Interesting in its own right, this study is also stimulated by applications in theoretical computer science, where randomness plays an essential role (see, e.g. [PS], [E], [Bo]). Here is a partial list of functionals which have been considered [WW]:

- The volume of K_n, $\mathrm{Vol}(K_n)$.
- The number of vertices of K_n, $N(K_n)$.
- The intrinsic volumes of K_n.
- The number of faces of dimension i, $f_i(K_n)$(of course $f_0(K_n)=N(K_n)$).
- The mean width of K_n.
- The distance between K_n and K.

For each of these key functionals, one knows, approximately or up to a constant factor, its expectation. Usually, more accurate estimates can be obtained if one possesses extra information about the boundary of K. The two most important special cases are the following:

- K is smooth, i.e. the boundary $\mathrm{bd}(K)$ is twice differentiable with positive curvature bounded away from zero and infinity.
- K is a polytope.

Research supported in part by an NSF CAREER award and by an Alfred P. Sloan fellowship.

Let us quote a typical result about expectations (see [WW, Theorem 1.4]).

Theorem 1.1. *Let K be a smooth convex body. There is a constant $c(K)$ such that*

$$\mathbf{E}\big(\mathrm{Vol}(K_n)\big) = 1 - \big(c(K) + o(1)\big)n^{-2/(d+1)}.$$

It is simply impossible to mention all results of this type, so we refer the reader to the survey papers by Gruber [G] and Weil and Wieacker [WW], both of which contain an extensive list of references.

Let Y_n be a functional of K_n. Once the expectation of Y_n is determined, the main task is to study the distribution of Y_n around it. The ultimate goal would be to estimate the tail probability

$$\mathbf{P}\big(|Y_n - \mathbf{E}(Y_n)| \geq T\big)\,. \tag{1}$$

A precise and complete answer to this question would, of course, tell everything one would like to know about Y_n. On the other hand, it seems very hard to achieve this goal. Here are several closely related questions which have been studied

- *Higher moments.* What is the variance of Y_n? What is the kth moment of Y_n, for any fixed k?
- *Limit and speed of convergence.* Is it true that with probability one
$$\lim_{n\to\infty} \frac{Y_n}{\mathbf{E}(Y_n)} = 1\,?$$
Assume that the answer is affirmative. The next question is to estimate the speed of convergence. One would like to find the fastest growing function $f(n)$ such that
$$\lim_{n\to\infty} \left|\frac{Y_n}{\mathbf{E}(Y_n)} - 1\right| f(n) = 0\,.$$
- *Central limit theorem.* Does Y_n satisfy a central limit theorem?

Each of these questions can be answered if we have can answer (1). However, even these questions are highly non-trivial and are far from being solved completely (for recent progresses, see [R1], [CG], [Hu]). Let us quote a paragraph from Weil and Wieacker's survey [WW, p. 1431], which, while written some years ago, still captures the general state of the topic

> *"We finally emphasize that the results described so far give mean values hence first-order information on random sets and point processes. This is due to the geometric nature of the underlying integral geometric results. There are also some less geometric methods to obtain higher-order informations or distributions, but generally the determination of variance, e.g. is a major open problem".*

One of the main obstacles to go beyond first moment is the fact that the integral geometric techniques, which are very effective in computing expectation, become much harder to apply for higher moments. A recent paper of Reitzner [R1], in which estimates on the variance are obtained, shows that one needs additional ideas and tools and also some non-trivial calculations to apply these techniques for the second moment.

The goal of our paper is to introduce a new method to obtain a bound on the tail probability in (1). This method has a probabilistic and combinatorial nature, and does not rely on the usual integral geometric techniques. Our results assert that the key functionals of K_n have very light tails, i.e. with very high probability they are close to their means. Using these results we will be able to get quite accurate estimates for the moments of any order and the speed of convergence. In a more recent paper [V3], we combined these concentration results together with a result of Reitzner [R2] to prove central limit theorems for several functionals.

One way to obtain sharp concentration results is to use high moment estimates (via Markov inequality or Chebyshev inequality). This is, as the reader would have already sensed, *not* the approach we are going to follow. In fact, our argument is almost orthogonal to this approach as we shall prove a sharp concentration result first and use it to compute moments afterwards.

Our main tool is the so-called *divide and conquer martingale* method, which has been developed recently in a sequence of papers [KV1,2], [V1] in probabilistic combinatorics. The heart of this method is a martingale inequality [KV1], [V1] (see Lemma 3.1 in section 3), which is a refinement of Azuma's inequality. We believe that this method would be useful for many other problems in statistical geometry. For this reason, we would like to emphasize methodology in this paper. In order to make the presentation cleaner and easier to follow, we are going to focus on only two functionals: the volume and the number of vertices of K_n. These are arguably the most important and also the easiest to visualize functionals. The treatment of other functionals will appear later.

Let us conclude this section by describing the structure of the paper. In section 2, we are going to discuss what kind of sharp concentration results one may expect and how one derives information about the high moments and the speed of convergence from such results. Next, we present our results. We will first present a general result for an arbitrary convex body K. We believe this result is sharp up to a logarithmic term. After this, we focus on the case when K is smooth and provide a stronger result where the extra logarithmic term is removed.

In the next section, section 3, we describe our main probabilistic tool, the divide and conquer martingale lemma, mentioned above. This lemma is a bit technical, so we are going to first present the well-known Azuma inequality and show why this classical inequality is not sufficient for the current study. This discussion is intended to give the reader a better understanding of our lemma, which refines Azuma's work.

Section 4 contains a few geometrical statements which we need for the proof of a concentration result concerning the volume of K_n when K is a general convex body. The critical notions in this part are ϵ-wet part and ϵ-floating body, taken from a paper of Bárány and Larman [BL]. The proof of the concentration result is presented in section 5. This is the simplest proof and we hope that it would provide the reader a clear idea about our approach. This part also contains a new geometrical ingredient, namely the use of Vapnik–Chervonenkis dimensions.

The next section, section 6, is devoted to the proof of a more accurate result concerning the volume of K_n for the special case when K is smooth. This is perhaps the most interesting and also the most technical part of the paper, where one needs to combine geometric arguments with probabilistic ones in a not entirely trivial way.

The next section, section 7, focuses on the number of vertices of K_n when K is smooth. The final section, section 8, is devoted to concluding remarks. We discuss a tantalizing conjecture about the monotonicity of the expectation of the number of vertices of K_n (as a function in n). A lemma proved in section 7 sheds some light on this problem. We will also briefly mention few other random models where our method can be applied.

NOTATION. In the whole paper, we assume that n is large, whenever needed. The asymptotic notation is used under the assumption that $n \to \infty$. Given non-negative functions $f(n)$ and $g(n)$, we write $f(n) = O(g(n))$ ($f(n) = \Omega(g(n))$) if there is a positive constant c, independent of n, such that $f(n) \leq Cg(n)$ ($f(n) \geq Cg(n)$) for all sufficiently large value of n. We write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. In this case, we say that $f(n)$ and $g(n)$ have the same order of magnitude. Finally $f(n) = o(g(n))$ if $f(n)/g(n)$ tends to zero as n tends to infinity.

In the paper, the hidden constants in O, Ω and Θ will depend on the fixed convex body K. Vol and $\mathbf{P}$ denote volume and probability, respectively. Consider a (measurable) subset S of K and a random point x

$$\mathrm{Vol}(S) = \mathbf{P}(x \in S)\,.$$

$\mathbf{E}$ denotes expectation. Let t_i, $i = 1, \dots, n$, be independent random variables and $Y = Y(t_1, \dots, t_n)$ be a random variable depending on $t_1, \dots, t_n$. $\mathbf{E}(Y|t_1, \dots, t_i)$ is the conditional expectation of Y conditioned on the first i variables.

2 Sharp Concentration Results and Consequences

The goal of the first two subsections is to give a motivation for our results and are highly recommended.

2.1 Exponential tails. Given a random variable Y. We would like to show that Y is strongly concentrated (or has a light tail distribution). *What is the best possible result we can hope for?*

If we cannot prove directly that Y has a very concrete distribution (such as normal or Poisson, for instance), then a good candidate is the following inequality:

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda \mathbf{Var}(Y)}\big) \leq c_1 \exp(-c_2\lambda)\,, \tag{2}$$

where c_1, c_2 are absolute constants. This inequality means that Y has an exponential (sub-gaussian) tail.

Inequality (2) is satisfactory for many practical purposes. In the next subsection, we are going to use it to derive estimates for the moments of Y and also to derive some information about the speed of convergence.

2.2 High moments and speed of convergence via concentration. Assume that one can prove the following tail inequality

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq c_1 \exp(-c_2\lambda)\,, \tag{3}$$

for all $\lambda > 0$, where V and c_1, c_2 are positive constants. Then one can conclude that for any fixed $k \geq 2$, the kth moment M_k of Y satisfies

$$M_k = O(V^{k/2})\,. \tag{4}$$

Thus, (2) would imply $M_k = O(\mathbf{Var}(Y)^{k/2})$.

The (rather standard) proof goes as follows. Write $\mu(t)$ for the probability $\mathbf{P}(|Y - \mathbf{E}(Y)| \geq t)$. The kth moment of Y can be expressed as

$$\begin{aligned} M_k &= \int_0^\infty t^k \partial \mathbf{P}\big(|Y - \mathbf{E}(Y)| < t\big) \\ &= -\int_0^\infty t^k \partial \mu(t) \end{aligned}$$

$$= \lim_{M\to\infty} - \int_0^M t^k \partial\mu(t)$$
$$= \lim_{M\to\infty} \left(\big(- t^k \mu(t)|_0^M \big) + \int_0^M k t^{k-1} \mu(t) \partial t \right).$$

By (3), $\mu(t)$ decreases exponentially. Thus,

$$\lim_{M\to\infty} \big(- t^k \mu(t)|_0^M \big) = 0\,.$$

Replace $t = \sqrt{\lambda V}$ and use (3) again, we have

$$\begin{aligned}\int_0^M k t^{k-1} \mu(t) \partial t &\le \int_0^\infty k t^{k-1} \mu(t) \partial t \\ &= \int_0^\infty k(\sqrt{\lambda V})^{k-1} \mathbf{P}\big(|Y - \mathbf{E}(Y)| \ge \sqrt{\lambda V}\big) \frac{\sqrt{V}}{2\sqrt{\lambda}} \partial\lambda \\ &\le \frac{k}{2} V^{k/2} \int_0^\infty \lambda^{\frac{k}{2}-1} c_1 \exp(-c_2\lambda) \partial\lambda \\ &= O(V^{k/2})\,,\end{aligned}$$

proving (4).

In order to derive a result about the speed of convergence, let us assume that Y_n satisfy

$$\mathbf{P}\big(|Y_n - \mathbf{E}(Y_n)| \ge \sqrt{\lambda V_n}\big) \le c_1 \exp(-c_2\lambda)\,, \tag{5}$$

for all sufficiently large n. Set $\lambda_n = \frac{2}{c_2} \ln n$. Let $\delta(n)$ be a sequence tending arbitrarily slowly to zero with n and define

$$f(n) = \frac{\delta(n)\mathbf{E}(Y_n)}{\sqrt{\lambda_n V_n}}\,. \tag{6}$$

We have

$$\begin{aligned}\mathbf{P}\left(\left| \frac{Y_n}{\mathbf{E}(Y_n)} - 1 \right| f(n) \ge \delta(n) \right) &= \mathbf{P}\big(|Y_n - \mathbf{E}(Y_n)| \ge \delta(n)\mathbf{E}(Y_n) f(n)^{-1}\big) \\ &\le \mathbf{P}\big(|Y_n - \mathbf{E}(Y_n)| \ge \sqrt{\lambda_n V_n}\big) \\ &\le c_1 \exp(-c_2\lambda_n) \\ &= c_1 n^{-2}.\end{aligned}$$

As the series n^{-2} converges, by the Borel–Cantelli lemma, we can conclude that almost surely

$$\left| \frac{Y_n}{\mathbf{E}(Y_n)} - 1 \right| f(n)$$

tends to zero with n.

2.3 Concentration of the volume, general $\boldsymbol{K}$. In this subsection, Y denotes the volume of K_n. We write Y_n instead of Y when talking about limit theorems. Our aim is to prove an inequality similar to (2). In general, our results have the following form

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq c_1 \exp(-c_2\lambda) + p_{NT}\,. \tag{7}$$

The quantity p_{NT} represents an error term (p_{NT} stands for the probability that some non-typical, briefly NT, event occurs). This error term is very small and is negligible in applications such as the ones presented in the previous subsection.

The critical issue is the magnitude of V. Ideally, we would like to have V comparable to the variance of Y. The problem is that at this moment one does not know (even up to a constant factor) the variance of any functional of K_n. However, by looking at available estimates in special cases, we have some ground to believe that our V (in (7)) is at most a logarithmic factor away from the actual variance of Y. For the case K is smooth, we are able to remove this extra logarithmic factor.

In order to state the result for a general K, we need to introduce a few notions. These notions are fairly well known in the literature (see [BL] or [WW]). For a convex body K and a half space H, we call the intersection $H \cap K$ an *ϵ-cap* if $\mathrm{Vol}(H \cap K) = \epsilon$. The union of all ϵ-caps is the *ϵ-wet part* of K, which we shall denote by F_ϵ. The complement $\overline{F_\epsilon}$ of the ϵ-wet part F_ϵ is called the *ϵ-floating body* of K. We denote by ρ_ϵ the volume of the ϵ-wet part.

In a nice paper, Bárány and Larman [BL] proved that for any convex body K

$$\mathrm{Vol}(\overline{K_n}) = \Theta(\rho_{1/n})\,,$$

where $\overline{K_n} = K \backslash K_n$. They have shown that for every sufficiently small ϵ ($\epsilon \leq \epsilon_0$, where ϵ_0 depends only on K)

$$\Omega\left(\epsilon\left(\ln \tfrac{1}{\epsilon}\right)^{d-1}\right) = \rho_\epsilon = O\big(\epsilon^{2/(d+1)}\big)\,. \tag{8}$$

The upper bound is attained when K is smooth and the lower bound is attained when K is a polytope. The lower bound was obtained independently by Dwyer [D] for simple polytopes.

Consider the floating body F_ϵ and a point $x \in \overline{F_\epsilon}$. We say that *$x$ sees y* (with respect to F_ϵ) if the segment xy does not intersect F_ϵ. Let $S_{x,\epsilon}$ denote the set of those y that x sees. Define

$$g(\epsilon) = \sup_{x \in \overline{F_\epsilon}} \mathrm{Vol}(S_{x,\epsilon})\,.$$

The set $S_{x,\epsilon}$ is the union of all ϵ-caps containing x.

Set $C = 3g(\epsilon)$, $V = 36ng(\epsilon)^2\rho_\epsilon$.

Theorem 2.1. *There are positive constants* $,\alpha, c$ *and* ϵ_0 *such that the following holds. For any* $\alpha \ln n/n < \epsilon \le \epsilon_0$ *and* $0 < \lambda \le V/4C^2 = n\rho_\epsilon$, *we have*

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \ge \sqrt{\lambda V}\big) \le 2\exp(-\lambda/4) + p_{NT}\,, \tag{9}$$

where $p_{NT} = \exp(-c\epsilon n)$.

The constants 3 and 36 in the definitions of C and V can be improved but we make no attempt to optimize them.

We would like to make a comment about the magnitude of V. Let us first consider the special cases when K is either smooth or a polytope.

K is smooth. It is well known (and not too hard to show) that for sufficiently small ϵ we have $g(\epsilon) = \Theta(\epsilon)$ and $\rho_\epsilon = \Theta(\epsilon^{2/(d+1)})$. Thus,

$$V = \Theta\big(n\epsilon^2\epsilon^{2/(d+1)}\big)\,.$$

Substituting $\epsilon = \Theta(\ln n/n)$, we have

$$V = \Theta\big(n^{-(d+3)/(d+1)}\ln^{(2d+4)/(d+1)} n\big)\,.$$

In [R1], Reitzner proved that the variance of the volume of K_n, when K is smooth, is $O(n^{-(d+3)/(d+1)})$. A matching lower bound for the case $d = 2$ has been claimed by Buchta (unpublished manuscript). A more recent result of Reitzner (not yet published) shows that $\Omega(n^{-(d+3)/(d+1)})$ is also the lower bound of the variance for all d.

If this is the case, our V from Theorem 2.1 would differ from the variance by a logarithmic factor $\ln^{(2d+4)/(d+1)} n$. In fact, this factor can be removed, as we will show in subsection 2.5.

K is a polytope. We are going to use the trivial estimate that $g(\epsilon) \le \rho_\epsilon$. It follows that $V = O(n\rho_\epsilon^3)$. Using (8), we have $\rho_\epsilon = O\big(\epsilon\big(\ln\frac{1}{\epsilon}\big)^{d-1}\big)$. Substituting $\epsilon = \Theta(\ln n/n)$, we have

$$V = O(n^{-2}\ln^{3d} n)\,.$$

Bárány (private conversation) recently mentioned that he could prove that the variance is around $n^{-2}\ln^{d-1} n$.

Let us derive a corollary of Theorem 2.1 for the special case when K is a polytope and $\epsilon = \alpha\ln n/n$ for some large constant α. In this case $p_{NT} = 2\exp(-c\alpha\ln n)$ and

$$\rho_\epsilon = \Theta\big(\epsilon\ln^{d-1}\tfrac{1}{\epsilon}\big) = \Theta(\ln^d n/n)\,.$$

and

$$\lambda_0 = n\rho_\epsilon = \Theta(\ln^d n)\,.$$

Corollary 2.2. *There is a positive constant c such that the following holds. For any sufficiently large constant α, there are positive constants β and c' such that for $0 < \lambda \le \beta \ln^d n$, we have*

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \ge \sqrt{c'\lambda n^{-2}\ln^{3d} n}\big) \le 2\exp(-\lambda/4) + p_{NT}\,, \qquad (10)$$

where $p_{NT} = \exp(-c\alpha \ln n)$.

Given the above consideration, we tend to believe that the estimate $V = O(ng(\epsilon)^2\rho_\epsilon)$ for $\epsilon = \Theta(\ln n/n)$ is at most a logarithmic factor away from the variance of Y.

2.4 High moments and speed of convergence. We can use the theorem above to estimate the moments of the volume of K_n, using the argument described in subsection 2.2.

Corollary 2.3. *For a general convex body K, the kth moment of the volume of K_n satisfies*

$$M_k = O(V^{k/2})\,, \qquad (11)$$

where

$$V = 36ng(\epsilon)^2\rho_\epsilon$$

and $\epsilon = \alpha \ln n/n$ for some sufficiently large constant α.

If K is a polytope, we have $V = O(n^{-2}\ln^{3d} n)$ as shown in the last subsection. This yields

Corollary 2.4. *Let K be a polytope. Then the kth moment of the volume of K_n satisfies*

$$M_k = O\big(n^{-k}\ln^{3kd/2} n\big)\,. \qquad (12)$$

Next, we derive a result about the speed of convergence.

Corollary 2.5. *There is a constant α such that the following holds. Almost surely,*

$$\lim_{n\to\infty}\left|\frac{Y_n}{\mathbf{E}(Y_n)} - 1\right| f(n) = 0$$

for

$$f(n) = \delta(n)n^{-1}g(\epsilon)^{-2}\rho_\epsilon^{-1}\ln^{-1/2} n$$

where $\epsilon = \alpha\ln n/n$ and $\delta(n)$ is a function tending to zero arbitrarily slowly.

For the case K is a polytope,

$$g(\epsilon)^{-2}\rho(\epsilon)^{-1} = \Omega\big(\rho(\epsilon)^{-3}\big) = \Omega\big(\epsilon^{-3}\ln^{3-3d}\tfrac{1}{\epsilon}\big)\,.$$

By substituting $\epsilon = \alpha \ln n/n$, it follows that

$$g(\epsilon)^{-2}\rho(\epsilon)^{-1} = \Omega(n^3\ln^{-3d} n)\,.$$

Thus, we can have

$$f(n) = \delta(n)n^2\ln^{-3d-1/2} n\,.$$

Corollary 2.6. *Let K be a polytope. Then almost surely,*

$$\lim_{n\to\infty}\left|\frac{Y_n}{\mathbf{E}(Y_n)}-1\right|\delta(n)n^2\ln^{-3d-1/2}n=0$$

for any function $\delta(n)$ tending to zero with n.

We do not formulate corollaries for the smooth case as we have better bounds in the next subsection. The proof of Corollary 2.3 follows the idea in subsection 2.2, but one needs a more careful analysis. The details are presented in the Appendix. Corollary 2.5 follows directly from the relevant the discussion in subsection 2.2. □

2.5 Better results for smooth K.

Theorem 2.7. *For any smooth convex body K, there are positive constants c,α,ϵ_0 such that the following holds. For any $0<\epsilon\le\epsilon_0$, $V\ge\alpha n^{-(d+3)/(d+1)}$, $C\ge\epsilon$ and $0<\lambda<V/4C^2$, we have*

$$\mathbf{P}\big(|Y-\mathbf{E}(Y)|\ge\sqrt{\lambda V}\big)\le 2\exp(-\lambda/4)+p_{NT}$$

where

$$p_{NT}=\exp(-c\epsilon n)+\exp\big(-cn^{(d-1)/(3d+5)}\big)\,.$$

The critical point here is that we can set $V=O(n^{-(d+3)/(d+1)})$, without the extra logarithmic term. As mentioned in subsection 2.3, it is believed that $n^{-(d+3)/(d+1)}$ is the right order of magnitude of the variance of the volume. Another important feature is that V does not depend on ϵ. Thus, we can set ϵ quite large in order to make p_{NT} very small. In Theorem 2.1 there is a trade-off between p_{NT} and V as both depend on ϵ.

By setting $\epsilon=n^{-(2d+6)/(3d+5)}$, the two exponents in p_{NT} are the same. So we can write $p_{NT}=\exp(-cn^{(d-1)/(3d+5)})$ (we can slightly change c to swallow the 2 factor). Set $V=\alpha n^{-(d+3)/(d+1)}$ and $C=\epsilon$. The bound on λ is

$$0<\lambda\le V/4C^2=\tfrac{\alpha}{4}n^{\frac{(d-1)(d+3)}{(d+1)(3d+5)}}\,.$$

We obtain the following corollary

Corollary 2.8. *For any smooth convex body K, there are positive constants c and α such that the following holds. For any $0<\lambda\le\frac{\alpha}{4}n^{\frac{(d-1)(d+3)}{(d+1)(3d+5)}}$*

$$\mathbf{P}\big(|Y-\mathbf{E}(Y)|\ge\sqrt{\lambda V}\big)\le 2\exp(-\lambda/4)+\exp\big(-cn^{(d-1)/(3d+5)}\big)\,.$$

Next, we present two corollaries about moments of Y and the speed of convergence of $Y_n/\mathbf{E}(Y_n)$.

Corollary 2.9. *Let K be a smooth convex body. Then for any fixed k, the kth moment of the volume of K_n satisfies*

$$M_k=O\big(n^{-\frac{k}{2}\frac{d+3}{d+1}}\big)\,.$$

Corollary 2.10. *Let K be a smooth convex body. Then almost surely,*
$$\lim_{n\to\infty}\left|\frac{Y_n}{\mathbf{E}(Y_n)}-1\right|\delta(n)n^{(d+3)/(d+1)}\ln^{-1/2} n = 0$$
for any function $\delta(n)$ tending to zero with n.

Remark. The case $k=2$ of Corollary 2.9 was proved earlier by Reitzner [R1], using a different method. In the same paper, he used this result to show that almost surely
$$\lim_{n\to\infty}\left|\frac{Y_n}{\mathbf{E}(Y_n)}-1\right| = 0\,.$$

2.6 Results for the number of vertices. We first consider the case when K is smooth. Let Z denote the number of vertices of K_n. It is well known that, via Efron,
$$\mathbf{E}(Z) = n\mathbf{E}(\overline{K_n}) = n\big(1-\mathbf{E}(Y)\big)\,.$$
It suggests that Z behaves like $n(1-Y)$. If this is the case, then we expect that $\mathbf{Var}(Z) = \Theta(n^2\mathbf{Var}(Y))$. Thus, if we want to formulate a theorem like Theorem 2.7 for Z, then we would have to multiply V by n^2. We can achieve this via the following theorem.

Theorem 2.11. *For any smooth convex body K, there are positive constants c,α,ϵ_0 such that the following holds. For any $0<\epsilon\le\epsilon_0$, $V\ge \alpha n^{(d-1)/(d+1)}$, $C\ge n\epsilon$ and $0<\lambda<V/4C^2$, we have*
$$\mathbf{P}\big(|Z-\mathbf{E}(Z)|\ge\sqrt{\lambda V}\big)\le 2\exp(-\lambda/4)+p_{NT}$$
where
$$p_{NT} = \exp(-c\epsilon n)+\exp\big(-cn^{(d-1)/(3d+5)}\big)\,.$$

Corollary 2.12. *Let K be a smooth convex body. For any fixed k, the kth moment of Z is $O(n^{\frac{k}{2}\frac{d-1}{d+1}})$.*

3 Divide and Conquer Martingale

Our approach to the claimed sharp concentration results is via martingale inequalities. The most well-known martingale inequality is perhaps that of Azuma (see [Az] or [S]). While this inequality is not sufficiently strong for our purposes, we, however, choose to present it first for pedagogical reasons, as it would serve for a better understanding of our real tool, Lemma 3.1.

3.1 Azuma's inequality. Let t_i, $i=1,\dots,n$, be independent random points from a probability space $\mathbf{\Omega}$, and $Y=Y(t_1,\dots,t_n)$ be a function with real values. For $t=(t_1,\dots,t_n)$, define
$$C_i(t) = \big|\mathbf{E}(Y|t_1,\dots,t_{i-1},t_i)-\mathbf{E}(Y|t_1,\dots,t_{i-1})\big|\,.$$

Notice that we define C_i as a function of t, but it only depends on the first i coordinates of t.

Azuma's inequality. We have

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq T\big) \leq 2\exp\left(-\frac{T^2}{2\sum_{i=1}^n \|C_i\|_\infty^2}\right). \tag{13}$$

Azuma's inequality and its variants play an important role in several areas of mathematics such as convex analysis and combinatorics (e.g. see [Mi], [M], [S] and [AS] for many fundamental applications).

Let us try to apply (13) for the volume of a random polytope in a ball in three dimensions. In order to have a result similar to (2) or Theorem 2.7, we would like to have

$$T = \sqrt{\lambda V}$$

where V is of order $n^{-(d+3)/(d+1)}$ (well, here $d = 3$, but this actual value does not play any role). On the other hand, $\sum_{i=1}^n \|C_i\|_\infty^2$ turns out to be too large. Let us consider C_n. Assume that $t_1, \dots, t_{n-1}$ form a regular $n-1$-gon on an equator of K. It is trivial that the volume of the convex hull of $t_1, \dots, t_n$ is maximum when t_n is the pole P above this equator. Moreover $\mathbf{E}(Y|t_1, \dots, t_n)$ is the average volume of $\mathrm{Conv}(t_1, \dots, t_n)$ when t_n is taken randomly in K. One can show that the difference between the maximum volume $\mathrm{Vol}(\mathrm{Conv}(t_1, \dots, t_{n-1}, P))$ and the average volume $\mathbf{E}(Y|t_1, \dots, t_n)$ is a positive constant. It follows that

$$\|C_n\|_\infty = \Omega(1)\,.$$

Therefore,

$$\sum_{i=1}^n \|C_i\|_\infty^2 \geq \|C_n\|_\infty^2 = \Omega(1)\,.$$

So the right-hand side in (13) is upper bounded by

$$2\exp\big(-\lambda V/\Omega(1)\big) = 2\exp\big(-c\lambda n^{-(d+3)/(d+1)}\big)\,,$$

for some positive constant c. Since $2\exp(-c\lambda n^{-(d+3)/(d+1)})$ is larger than one if $\lambda = o(n^{(d+3)/(d+1)})$, this yields nothing unless λ is very large, i.e. $\lambda = \Omega(n^{(d+3)/(d+1)})$. In the case $\lambda = \Omega(n^{(d+3)/(d+1)})$, $T = \Omega(1)$ and the best one can get is a result of the following form

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq a\big) \leq 2\exp(-b)\,, \tag{14}$$

where a is a constant and b is a constant depending on a. This is an easy result (keep in mind that Y, the volume of K_n, is at most 1) and one can get a much stronger one through Lemma 4.2 (this lemma immediately implies that with probability $1 - o(1)$, the volume Y is $1 - o(1)$).

To make matters worse, let us point out that we cannot get even (14). The point is that the estimate

$$\sum_{i=1}^{n} \|C_i\|_\infty^2 \geq \|C_n\|_\infty^2 = \Omega(1)$$

is very generous. The truth is that for m relatively close to n, $\|C_m\|_\infty$ is also lower bounded by a constant. So one has

$$\sum_{i=1}^{n} \|C_i\|_\infty^2 = \omega(n)$$

for some function $\omega(n)$ tending to infinity.

3.2 Divide and conquer Martingale. Recall the definition the ith martingale difference

$$C_i(t) = \left|\mathbf{E}(Y|t_1, \dots, t_{i-1}, t_i) - \mathbf{E}(Y|t_1, \dots, t_{i-1})\right|.$$

We now want to view this quantity as a function in both t and t_i. For this purpose, we write $\mathbf{E}(Y|t_1, \dots, t_{i-1}, x)$ instead of $\mathbf{E}(Y|t_1, \dots, t_{i-1}, t_i)$, given $t_i = x$. This notation emphasizes the dependence on x.

Define

$$C_i(x,t) = \left|\mathbf{E}(Y|t_1, \dots, t_{i-1}, x) - \mathbf{E}(Y|t_1, \dots, t_{i-1})\right|.$$

Furthermore, set

$$V_i(t) = \int C_i(x,t)^2 \partial x\,,$$

$$V(t) = \sum_{i=1}^{n} V_i(t)\,,$$

$$C_i(t) = \sup_x C_i(x,t).$$

and

$$C(t) = \max_{i=1}^{n} C_i(t) = \sup_{i,x} C_i(x,t)\,.$$

The critical point here is that the definition of $V_i(t)$ is based on the l_2 norm, rather than the l_∞ norm.

Our key tool is Lemma 3.1 below. This lemma has a bit of history. It is a (fairly straightforward) generalization of Lemma 3.1 in [V1], which, in turn, is the continuous version of a concentration result about binary random variables by Kim and Vu in [KV1] (Theorem 3.1 of [KV1]). This theorem, on the other hand, is a refinement of several earlier results (see [V1] or [KV1] for discussion).

Lemma 3.1. *For any positive λ, C and V satisfying $\lambda \leq V/4C^2$, we have*

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq 2\exp\left(-\tfrac{\lambda}{4}\right) + \mathbf{P}\big(V(t) \geq V \ \text{ or } \ C(t) \geq C\big)\,. \quad (15)$$

It is useful to compare (27) with (7). The term p_{NT} in (7) corresponds to the probability $\mathbf{P}(V(t) \geq V \text{ or } C(t) \geq C)$ in (27). Bounding this term is the main technical part of the proofs.

Lemma 3.1 has many interesting applications in various areas. For instance, it played a key role in the development of concentration results for polynomials [KV1], [V1,2]. We would like to refer the reader to [V1] (especially section 3) for more details.

The proof of Lemma 3.1 is basically the same as the proof of Lemma 3.1 of [V1], which follows the scheme described in [KV1]. As this proof is independent of the rest of the paper, we choose to present it in the appendix.

3.3 One more lemma. Let Ω' and Ω'' be probability spaces and Ω''' their product. Let A be an event in Ω''' which occurs with probability at least $1-\delta'$, for some $0 < \delta' < 1$.

Lemma 3.2. *For any* $1 > \delta > \delta'$

$$\mathbf{P}_{\Omega'}\big(\mathbf{P}_{\Omega''}(A|x) \leq 1-\delta\big) \leq \delta'/\delta\,,$$

where x is a random point in Ω' and $\mathbf{P}_{\Omega'}$ and $\mathbf{P}_{\Omega''}$ are the probabilities according to Ω' and Ω'', respectively.

Proof of Lemma 3.2. Recall that $\mathbf{P}_{\Omega'''}(A) \geq 1-\delta'$. On the other hand,

$$\mathbf{P}_{\Omega'''}(A) = \int_{\Omega'} \mathbf{P}(A|x)\partial x \leq 1 - \delta \mathbf{P}_{\Omega'}\big(\mathbf{P}_{\Omega''}(A|x) \leq 1-\delta\big)\,.$$

The claim follows immediately. □

4 VC-Dimensions and Geometrical Lemmas

Let us start with an important definition. Let L be a finite collection of points. For a point x, define

$$\Delta_{x,L} = \mathrm{Vol}\big(\mathrm{Conv}(L \cup x)\big) - \mathrm{Vol}\big(\mathrm{Conv}(L)\big)\,.$$

It is trivial that if x belongs to $\mathrm{Conv}(L)$, then $\Delta_{x,L}$ is zero.

The following lemma is immediate from the definition of $g(\epsilon)$ in section 2.3.

Lemma 4.1. *Let L be a set whose convex hull contains the floating body F_ϵ. Then for any x*

$$\Delta_{x,L} \leq g(\epsilon)\,.$$

Next, we present a lemma which asserts that with very high probability the random polytope K_n contains the floating body F_ϵ, given that ϵ is sufficiently large. In this lemma, we consider the general set up where the random points are chosen with respect to an arbitrary distribution in $\mathbb{R}^d$. The definition of caps and floating bodies is modified accordingly. We say that a half space H is an ϵ-cap if the volume of H is ϵ. The ϵ-wet part is still the union of all ϵ-caps and the ϵ-floating body is the complement of the ϵ-wet part.

LEMMA 4.2. *There are positive constants c and c' such that the following holds for every sufficiently large n. For any $\epsilon \geq c' \ln n/n$, the probability that K_n does not contain F_ϵ is at most* $\exp(-c\epsilon n)$.

This lemma was proved by Bárány and Dalla [BD] for the case when the distribution is uniform in a convex body K. The reason we present the lemma in the above general form is that it might be useful for other cases (for example when the points are chosen from $\mathbb{R}^d$ with respect to the normal distribution). Furthermore, our proof, which relies on the notion of VC-dimensions, is different from the proof in [BD], which relies on the notion of Macbeath regions.

Proof of Lemma 4.2. We first recall the definition of VC-(Vapnik–Chervonenkis) dimension. Let X be a set and $\mathcal{F}$ be a family of subset of X. For a subset $X' \subset X$, the restriction of $\mathcal{F}$ on X' is

$$\mathcal{F}|_{X'} = \{S \cap X' \mid S \in \mathcal{F}\}.$$

A subset $A \subset X$ is *shattered* by $\mathcal{F}$ if each subset of A can be obtained as the intersection of some $S \in \mathcal{F}$ with A, i.e. if $\mathcal{F}|_A = 2^A$. The VC-dimension of $\mathcal{F}$, denoted by $\dim_{VC}(\mathcal{F})$, is the supremum of the sizes of all finite shattered finite subsets of X. The following fact is well known (see, e.g. [M, Lemma 10.3.1]).

LEMMA 4.3. *The VC-dimension of the family of half spaces in $\mathbb{R}^d$ is $d+1$.*

Given a set X with a probability measure μ and a family $\mathcal{F}$ of measurable subsets of X, a subset N of X is an ϵ-net of $\mathcal{F}$ if $N \cap S$ is not empty for any $S \in \mathcal{F}$ with $\mu(S) \geq \epsilon$. The following is a celebrated result of Haussler and Welzl [HW].

Theorem 4.4. *There is a constant c such that the following holds. If X is a set with a probability measure μ and $\mathcal{F}$ is a family of measurable subsets of X such that $\dim_{VC}(\mathcal{F}) = d$, then $\mathcal{F}$ has an ϵ-net of size $cd\epsilon^{-1} \ln 1/\epsilon$.*

The proof of this theorem uses a probabilistic argument. One takes a random set of the desired size $cd\epsilon^{-1} \ln 1/\epsilon$ and shows that with positive probability this set hits all sets of measure at least ϵ in $\mathcal{F}$. A closer look at the proof reveals that the probability of failure is quite small (see, for instance, [HW] or [M, p. 239-241]).

Corollary 4.5. *There are constants c and α such that the following holds. If X is a set with a probability measure μ and $\mathcal{F}$ is a family of measurable subsets of X such that $\dim_{VC}(\mathcal{F}) = d$, then a random set of size $N = cd\epsilon^{-1}\ln 1/\epsilon$ fails to hit all sets of measure at least ϵ in $\mathcal{F}$ with probability at most*

$$\alpha^d\big(\epsilon^{c/4}\ln\tfrac{1}{\epsilon}\big)^d.$$

In our case $\mathcal{F}$ is the family of half spaces in $\mathbb{R}^d$ with VC-dimension $d+1$. Let S be a set of m random points. Notice that

$$\mathbf{P}(F_\epsilon \not\subset K_m) \le \mathbf{P}(S \text{ does not hit all } \epsilon-\text{caps}).$$

Since d is a constant and ϵ is sufficiently small, we have

$$\alpha^d\big(\epsilon^{c/4}\ln\tfrac{1}{\epsilon}\big)^d \le \exp\big(-\beta\ln\tfrac{1}{\epsilon}\big)$$

for some positive constant β. Set $l = n/N$. We assume, without loss of generality, that l is an integer. The assumption that $\epsilon \ge c'\ln n/n$ is used here to guarantee that l is at least one. We sample n random points in l rounds, in each round we sample N points. (Notice that the fact that a point may be sampled more than once works in our favor.) The probability that the collection of n points fails to hit all ϵ-caps is at most

$$\exp\bigg(-\beta\ln\frac{1}{\epsilon}\bigg)^l = \exp\bigg(-\beta l\ln\frac{1}{\epsilon}\bigg) = \exp\bigg(-\Omega\bigg(\frac{n}{\epsilon^{-1}\ln\frac{1}{\epsilon}}\ln\frac{1}{\epsilon}\bigg)\bigg) = \exp\big(-\Omega(\epsilon n)\big)$$

concluding the proof. Here the hidden constant in Ω depends on d, but not on ϵ and n. □

5 Proof of Theorem 2.1

Set $V' = n^{-1}V = 36g(\epsilon)^2\rho_\epsilon$. Notice that since $g(\epsilon) \ge \epsilon \ge \alpha\ln n/n$, where α is a sufficiently large constant, we can assume that $n\exp(-c\epsilon n) \le n^{-4} < g(\epsilon)^2\rho_\epsilon$.

Recall that $p_{NT} = \exp(-c\epsilon n)$, again by setting α large we can assume that $p_{NT} \ge n\exp\big(-\frac{c}{2}\epsilon n\big)$. Thus by adjusting c we can replace p_{NT} by a new (more convenient) error term

$$n\exp(-c\epsilon n).$$

It suffices to prove that

$$\mathbf{P}\big(C(t) \ge C \text{ or } V(t) \ge V\big) \le n\exp(-c\epsilon n)$$

for some positive constant c.

We are going to prove the following claim.

Claim 5.1. *There is a positive constant c such that, for any $1 \le i \le n$,*

$$\mathbf{P}\big(C_i(t) \ge C \text{ or } V_i(t) \ge V'\big) \le \exp(-c\epsilon n).$$

Claim 5.1 and the trivial union bound imply that

$$\mathbf{P}\big(C(t) \geq C \text{ or } V(t) \geq V\big) \leq n\exp(-c\epsilon n)\,.$$

Using Lemma 3.1, we can conclude that for any $0 < \lambda \leq V/4C^2$,

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq 2\exp(-\lambda/4) + n\exp(-c\epsilon n)\,, \tag{16}$$

proving Theorem 2.1. It remains to verify the claim and that is what we do next. In the following, we denote by $\Omega_{(j)}$ and $\Omega^{\langle j\rangle}$ the product spaces spanned by $\{t_1,\dots,t_j\}$ and $\{t_j,\dots,t_n\}$, respectively.

Proof of Claim 5.1. To start, notice that by the triangle inequality

$$\begin{aligned} C_i(x,t) &= \big|\mathbf{E}(Y|t_1,\dots,t_{i-1},x) - \mathbf{E}(Y|t_1,\dots,t_{i-1})\big| \\ &\leq \mathbf{E}_{x'}\big|\mathbf{E}(Y|t_1,\dots,t_{i-1},x) - \mathbf{E}(Y|t_1,\dots,t_{i-1},x')\big|\,, \end{aligned}$$

where $\mathbf{E}_{x'}$ denotes the expectation over a random point x'. Let us fix (arbitrarily) $t_1,\dots,t_{i-1}$ and x. Let L be the union of $\{t_1,\dots,t_{i-1}\}$ and the random set $\{t_{i+1},\dots,t_n\}$. Since

$$\mathrm{Vol}\big(\mathrm{Conv}(L\cup x)\big) = \mathrm{Vol}\big(\mathrm{Conv}(L)\big) + \Delta_{x,L}\,,$$

we have

$$\mathbf{E}(Y|t_1,\dots,t_{i-1},x) = \mathbf{E}\big(\mathrm{Vol}(\mathrm{Conv}(L))|t_1,\dots,t_{i-1}\big) + \mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1})\,.$$

Now comes a key inequality

$$\mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1}) \leq \mathbf{P}\big(F_\epsilon \not\subset \mathrm{Conv}(L)\;\big|\; t_1,\dots,t_{i-1}\big) + g(\epsilon)\mathbf{I}_{x\in\overline{F_\epsilon}}\,. \tag{17}$$

Inequality (17) follows from the observations below:

- $\Delta_{x,L}$ is at most 1.
- If $\mathrm{Conv}(L)$ contains F_ϵ, then by Lemma 4.1 $\Delta_{x,L}$ is zero if $x \in \mathrm{Conv}(L)$ and at most $g(\epsilon)$ otherwise.

Set $\delta = n^{-4}$. We say that the set $\{t_1,\dots,t_{i-1}\}$ is *typical* if

$$\mathbf{P}_{\Omega^{\langle i+1\rangle}}\big(F_\epsilon \subset \mathrm{Conv}(L)\;\big|\; t_1,\dots,t_{i-1}\big) \geq 1-\delta\,.$$

The rest of the proof has two steps. In the first step, we show that if $\{t_1,\dots,t_{i-1}\}$ is typical then $C_i(t) \leq C$ and $V_i(t) \leq V'$. In the second step, we bound the probability that $\{t_1,\dots,t_{i-1}\}$ is not typical.

First step. Assume that $\{t_1,\dots,t_{i-1}\}$ is typical. We first bound $C_i(x,t)$. Observe that

$$\begin{aligned} C_i(x,t) &\leq \mathbf{E}_{x'}\big|\mathbf{E}(Y|t_1,\dots,t_{i-1},x) - \mathbf{E}(Y|t_1,\dots,t_{i-1},x')\big| \\ &\leq \mathbf{E}_{x'}\big|\mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1}) - \mathbf{E}(\Delta(x',L)|t_1,\dots,t_{i-1})\big| \\ &\leq \mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1}) + \mathbf{E}_{x'}\mathbf{E}(\Delta_{x',L}|t_1,\dots,t_{i-1}) \\ &\leq g(\epsilon)\mathbf{I}_{x\in\overline{F_\epsilon}} + g(\epsilon)\mathbf{P}(x'\in\overline{F_\epsilon}) + 2n^{-4} \quad \text{(using (17))} \\ &= g(\epsilon)(\mathbf{I}_{x\in\overline{F_\epsilon}} + \rho_\epsilon) + 2n^{-4} \\ &\leq 2g(\epsilon) + 2n^{-4} \leq 3g(\epsilon) = C\,. \end{aligned}$$

In the last inequality we used the fact that for $\epsilon = \Omega(\ln n/n)$, $g(\epsilon) \geq \epsilon \gg n^{-4}$. It follows that

$$C_i(t) = \max_x C_i(x,t) \leq C\,.$$

To simplify the calculation concerning $V_i(t)$, observe that

$$C_i(x,t)^2 \leq \big(g(\epsilon)(\mathbf{I}_{x\in\overline{F_\epsilon}} + \rho_\epsilon) + 2n^{-4}\big)^2 \leq g(\epsilon)^2(\mathbf{I}_{x\in\overline{F_\epsilon}} + \rho_\epsilon)^2 + n^{-4}\,,$$

since both $g(\epsilon)(\mathbf{I}_{x\in\overline{F_\epsilon}} + \rho_\epsilon)$ and $2n^{-4}$ are at most $1/8$ given ϵ is sufficiently small. Thus, we have

$$\begin{aligned}
V_i(t) &= \int C_i(x,t)^2 \partial x \\
&\leq n^{-4} + \int g(\epsilon)^2(\mathbf{I}_{x\in\overline{F_\epsilon}} + \rho_\epsilon)^2 \partial x \\
&= n^{-4} + \int g(\epsilon)^2\big(\mathbf{I}_{x\in\overline{F_\epsilon}} + 2\mathbf{I}_{x\in\overline{F_\epsilon}}\rho_\epsilon + \rho_\epsilon^2\big)\partial x \\
&= n^{-4} + g(\epsilon)^2\rho_\epsilon^2 + g(\epsilon)^2(1+2\rho_\epsilon)\int \mathbf{I}_{x\in\overline{F_\epsilon}}\partial x \\
&= n^{-4} + g(\epsilon)^2\rho_\epsilon^2 + g(\epsilon)^2(1+2\rho_\epsilon)\rho_\epsilon \\
&= n^{-4} + g(\epsilon)^2(\rho_\epsilon + 3\rho_\epsilon^2) \\
&\leq n^{-4} + 4g(\epsilon)^2\rho_\epsilon \leq 5g(\epsilon)^2\rho_\epsilon < V'\,.
\end{aligned}$$

In the last inequality we used the fact that for $\epsilon = \Omega(\ln n/n)$, $g(\epsilon)^2\rho_\epsilon \geq \epsilon^3 \gg n^{-4}$.

Second step. In this step, we bound the probability that $\{t_1,\dots,t_{i-1}\}$ is not typical. Consider a set L of $n-1$ random points $\{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_n\}$. Lemma 4.2 yields

$$\mathbf{P}\big(F_\epsilon \not\subset \mathrm{Conv}(L)\big) \leq \exp(-c_0\epsilon n)\,, \tag{18}$$

for some positive constant c_0 depending only on K. Applying Lemma 3.2 with $\Omega' = \Omega_{(i-1)}$, $\Omega'' = \Omega^{\langle i+1\rangle}$, $\delta' = \exp(-c\epsilon n)$ and $\delta = n^{-4}$, we have

$$\begin{aligned}
&\mathbf{P}_{\Omega_{(i-1)}}\big(\{t_1,\dots,t_{i-1}\} \text{ is not typical}\big) \\
&\quad= \mathbf{P}_{\Omega_{(i-1)}}\big(\mathbf{P}(F_\epsilon \not\subset \mathrm{Conv}(L)|t_1,\dots,t_{i-1}) \leq 1-\delta\big) \\
&\quad\leq \delta'/\delta \\
&\quad= n^4\exp(-c_0\epsilon n) \\
&\quad\leq \exp(-c\epsilon n)
\end{aligned}$$

for $c = c_0/2$, given $c_0\epsilon n \geq 8\ln n$. This latest condition can be satisfied by setting the constant α in the lower bound of ϵ to be sufficiently large. Our proof is completed. □

6 Proof of Theorem 2.7

In this section K is a fixed, smooth, convex body. All constants defined may depend on K.

The leading idea in the proof is the following. Instead of using (17) to bound the conditional expectation $\mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1})$, we try to bound it by a better quantity which depends on the location of x, given that $x \in \overline{F_\epsilon}$. This would lead to a better estimate of $V_i(t)$ as we can integrate over x.

This idea is, however, hard to visualize quantitatively. But it motivates the following approach, which actually works. We are going to show that (with very high probability), the set where $\Delta_{x,L}$ is large has very small measure. This way, we gain enough improvement so that when we integrate over x and estimate $V_i(t)$, we will be able to remove the extra logarithmic term.

The proof has four parts. In the first part, we describe a few geometrical lemmas. Some of these lemmas do rely on the fact that K is smooth and are not true otherwise (we will specify which ones). Next, we use these lemmas to prove the main lemma. In the third part, we derive a corollary of Theorem 2.1. Finally, the proof of the theorem comes in the fourth part.

6.1 Geometrical lemmas. In the following, we assume that ϵ is sufficiently small, whenever needed.

LEMMA 6.1. *There are positive constants c_1, c_2 such that any ϵ-cap C of K has diameter at least $c_1\epsilon^{1/(d+1)}$ and at most $c_2\epsilon^{1/(d+1)}$.*

LEMMA 6.2. *There is a positive constant c_3 such that the following holds. Let x be a point on the boundary of K and $D(x,\epsilon)$ the set of all points on the boundary which are of distance at most ϵ to x. Then the convex hull of $D(x,\epsilon)$ has volume at most $c_3\epsilon^{d+1}$.*

These two lemmas follow from the general principle that a smooth body looks locally like a ball (after a proper affine transformation).

The following lemma is an easy consequence.

LEMMA 6.3. *There is a positive constant c_4 such that the following holds. Let C be an ϵ-cap of K. The union of all ϵ-caps intersecting C has volume at most $c_4\epsilon$.*

Proof of Lemma 6.3. Let x be a point in $C \cap \mathrm{bd}(K)$, where $\mathrm{bd}(K)$ denotes the boundary of K. By Lemma 6.1, the distance from x to any point in the intersection of $\mathrm{bd}(K)$ and the union in question is $O(\epsilon^{1/(d+1)})$ and we can conclude the proof by applying Lemma 6.2. □

Let L be a finite set of points. If a cap does not intersect L, we say that it avoids L. We use $E_{\delta,L}$ to denote the union of all δ-caps that avoid L. By definition, $E_{\delta,L}$ does not contain any point from L.

LEMMA 6.4. *There is a positive constant c_5 such that the following holds. For any set L, the set $E_{\delta,L}$ contains at least $\lfloor c_5\delta^{-1}\mathrm{Vol}(E_\delta)\rfloor$ pairwise disjoint δ-caps.*

Proof of Lemma 6.4. Without loss of generality, let us assume that $E_{\delta,L}$ contains at least one δ-cap. Let $C_1,\dots,C_m$ be a maximal system of pairwise non-intersecting δ-caps in $E_{\delta,L}$. By the maximality, we have

$$E_{\delta,L} = \cup_{i=1}^m \mathcal{C}_i\,,$$

where $\mathcal{C}_i$ is the union of all δ-caps in $E_{\delta,L}$ intersecting C_i. By Lemma 6.3, the volume of $\mathcal{C}_i$ is $O(\delta)$. It follows that

$$\mathrm{Vol}(E_\delta) \le \sum_{i=1}^m \mathrm{Vol}(\mathcal{C}_i) = O(m\delta)\,,$$

which implies $m = \Omega(\delta^{-1}\mathrm{Vol}(E_\delta))$, as claimed. □

LEMMA 6.5. *There are positive constants c_6 and c_7 such that the following holds. There is a system of $m \le c_6(\epsilon^{-1}\rho_\epsilon)$ pairwise disjoint $(c_7\epsilon)$-caps $C_1,\dots,C_m$ such that any ϵ-cap contains at least one of the C_i's.*

LEMMA 6.6. *There are constants c_8 and c_9 such that the following holds. There is a system of $(c_9\epsilon)$-caps $C_1,\dots,C_m$ where $m \le c_8(\epsilon^{-1}\rho_\epsilon)$ and any cap of volume at most ϵ is contained in at least one of the C_i's.*

REMARK. Both Lemma 6.5 and Lemma 6.6 can be proved using Lemmas 6.1 and 6.2. We omit the routine proofs. Moreover, it was shown by Bárány that Lemma 6.6 holds for all convex bodies [B].

Assume that x is a vertex of $\mathrm{Conv}(L\cup x)$ and the caps determined by the hyperplanes defined by the facets containing x have volume at most ϵ. Then by Lemma 6.1 the diameter of each cap is $O(\epsilon^{1/(d+1)})$. Moreover, since x itself belongs to an ϵ-cap, there is a point x_0 on the boundary of K which has distance $O(\epsilon^{1/(d+1)})$ to x. As every cap in concern contains x, the distance from x_0 to any point in these caps is also $O(\epsilon^{1/(d+1)})$. Thus, by Lemma 6.2, the convex hull of these caps has volume $O(\epsilon)$. On the other hand, this convex hull contains $\mathrm{Conv}(L\cup x)\backslash\mathrm{Conv}(L)$. So

$$\Delta_{x,L} = \mathrm{Vol}\big(\mathrm{Conv}(L\cup x)\backslash\mathrm{Conv}(L)\big) = O(\epsilon)\,.$$

This results in the following lemma.

LEMMA 6.7. *There is a positive constant c_{10} such that the following holds. Let L be a finite set of points and assume that $\Delta_{x,L}$ is at least ϵ. Then one of the caps determined by the facets containing x has volume at least $c_{10}\epsilon$ and so $x \in E_{c_{10}\epsilon,L}$.*

6.2 The main lemma. Consider a finite set L of points in K. A point x is *δ-large with respect to L* if $\Delta_{x,L} \geq \delta$. Define

$$X_{\delta,L} = \{x \mid x \text{ is } \delta\text{-large with respect to } L\}.$$

Let L be a set of n random points. The main lemma of our proof asserts that with high probability, there are few points which are very large with respect to L.

LEMMA 6.8. *There are positive constants c, c', c'' and c''' such that the following holds. For any $\delta > c'n^{-1}$ and any $T \geq \max\{c''\delta, \rho_\delta \exp(-c'''\delta n)\}$, we have*

$$\mathbf{P}\big(\mathrm{Vol}(X_{\delta,L}) \geq T\big) \leq \exp(-cnT).$$

Proof of Lemma 6.8. Lemma 6.7 implies that

$$\mathrm{Vol}(X_{\delta,L}) \leq \mathrm{Vol}(E_{c_{10}\delta,L}).$$

We are going to bound the probability that $\mathrm{Vol}(E_{c_{10}\delta,L})$ is at least T. Let c_{11} be a sufficiently small positive constant and fix a pairwise disjoint $(c_{11}\delta)$-cap system $C_1, \dots, C_m$ such that every $c_{10}\delta$ cap contains at least one of the C_i, using Lemma 6.5.

Assume that the volume of $E_{c_{10}\delta,L}$ is at least T. Then using Lemma 6.4, we have a system $A_1, \dots, A_l$ of disjoint $c_{10}\delta$-caps where $l = \Omega(\delta^{-1}T)$ and every A_i avoids L. The assumption that T/δ is sufficiently large ($T \geq c''\delta$) guarantees that l is at least one. Each A_i contains one of the caps C_j (given that c_{11} is sufficiently small). So, we can conclude that the system $C_1, \dots, C_m$ contains a subsystem of l elements each of which avoids L.

As L contains n random points, the probability that a fixed $(c_{11}\delta)$-cap avoids L is at most

$$(1 - c_{11}\delta)^n \leq \exp(-c_{11}\delta n).$$

By the union bound, the probability that the system $C_1, \dots, C_m$ contains a subsystem of l elements avoiding L is at most

$$\begin{aligned}\binom{m}{l}\exp(-c_{11}\delta n)^l &= \binom{m}{l}\exp(-c_{11}\delta nl)\\ &\leq \Big(\frac{em}{l}\Big)^l \exp(-c_{11}\delta nl)\\ &= \exp\Big(\Big(-c_{11}\delta n + \ln\frac{em}{l}\Big)l\Big).\end{aligned}$$

We have $m = O(\delta^{-1}\rho_\delta)$ and $l = \Omega(\delta^{-1}T)$. Thus,

$$\ln \tfrac{em}{l} \leq \ln(c_{12}\rho_\delta T^{-1}),$$

for some positive constant c_{12}. It follows that there are positive constants c' and c''' such that if $T \geq \rho_\delta \exp(-c'''\delta n)$ and $\delta \geq c' n^{-1}$ then

$$\ln(c_{12}\rho_\delta T^{-1}) \leq \tfrac{1}{2}c_{11}\delta n.$$

(We need the assumption $\delta \geq c' n^{-1}$ to ignore the constant c_{12}.) Choosing c' and c''' properly, we have

$$\exp\big(\big(-c_{11}\delta n + \ln \tfrac{em}{l}\big) l\big) \leq \exp\big(-\tfrac{c_{11}}{2}\delta n l\big) \leq \exp(-cnT),$$

for some positive constant c, using the fact that $l = \Omega(\delta^{-1}T)$. This concludes the proof. □

6.3 A corollary of Theorem 2.1. We also need the following corollary of Theorem 2.1.

COROLLARY 6.9. *Let K be a smooth convex body and Y be the volume of K_n. There are positive constants c and c' such that the following holds*

$$\mathbf{P}(Y \leq 1 - c'\rho_{1/n}) \leq \exp\big(-cn^{(d-1)/(3d+5)}\big).$$

Proof of Corollary 6.9. We will apply Theorem 2.1 with $V = c_0 n^{-\alpha}$, $\epsilon = n^{-\beta}$ and $\lambda = c_0 n^{\gamma}$ for some positive constants $c_0, \alpha, \beta, \gamma$. Recall (from subsection 2.3) that

$$\rho_{1/n} = \Theta\big(n^{-2/(d+1)}\big)$$

and

$$\mathbf{E}(Y) = 1 - \Theta(\rho_{1/n}).$$

By choosing α and β such that $-\alpha + \beta = -4/(d+1)$, we have $\sqrt{\lambda V} = \Theta(\rho_{1/n})$. Thus, for a sufficiently large constant c'

$$\mathbf{P}(Y \leq 1 - c'\rho_{1/n}) \leq \mathbf{P}\big(Y \leq \mathbf{E}(Y) - \sqrt{\lambda V}\big).$$

On the other hand, due to Theorem 2.1

$$\mathbf{P}\big(Y \leq \mathbf{E}(Y) - \sqrt{\lambda V}\big) \leq \mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq 2\exp(-\lambda/4) + \exp\big(-\Omega(\epsilon n)\big).$$

Next, we are going to optimize the right-hand side. We would like to choose λ as large as possible and ϵ as small of possible, with respect to the following constraints

- $-\alpha + \gamma = -4/(d+1)$;
- $V = c_0 n^{-\alpha} \geq 36 n g(\epsilon)^2 \rho_\epsilon = \Theta(n\epsilon^2\rho(\epsilon)) = \Theta(n^{1-(2+\frac{2}{d+1})\beta})$;
- $\lambda = c_0 n^{\gamma} \leq \lambda_0 = n\rho_\epsilon = n^{1-\frac{2}{d+1}\beta}$.

In the second constraint, we used the fact that $g(\epsilon) = \Theta(\epsilon)$. We can set c_0 to be a sufficiently small positive constant. The optimal choices for α, β, γ are

$$\alpha = \frac{d^2 + 12d + 19}{(d+1)(3d+5)}, \quad \beta = \frac{2d+6}{3d+5}, \quad \gamma = \frac{d-1}{3d+5}.$$

With this setting, both λ and ϵn are of order $n^{(d-1)/(3d+5)}$. Thus,

$$2\exp(-\lambda/4) + \exp(-\Omega(\epsilon n)) \le \exp\big(-cn^{(d-1)/(3d+5)}\big)$$

for some properly chosen positive constant c, concluding the proof. □

6.4 Proof of the theorem. Recall that

$$p_{NT} = \exp(-c\epsilon n) + \exp\big(-cn^{(d-1)/(3d+5)}\big)\,.$$

Arguing as in the beginning of the proof of Theorem 2.1, we can replace this by a more convenient error term

$$p_{NT} = n\exp(-c\epsilon n) + n\exp\big(-cn^{(d-1)/(3d+5)}\big)\,.$$

Set $p_N^{[1]}T = \exp(-c\epsilon n)$ and $p_{NT}^{[2]} = \exp(-cn^{(d-1)/(3d+5)})$ for some sufficiently small c. It suffices to prove

$$\mathbf{P}\big(C(t) \ge C\big) \le np_{NT}^{[1]} \tag{19}$$

and

$$\mathbf{P}\big(V(t) \ge V\big) \le np_{NT}^{[2]}\,. \tag{20}$$

The proof of (19) is exactly as before. In the following, we are going to prove (20), using Lemma 6.8 and Corollary 6.9.

Claim 6.10. *For any* $1 \le i \le n$,

$$\mathbf{P}\big(V_i(t) \ge n^{-1}V\big) \le p_{NT}^{[2]}\,.$$

Inequality (20) follows immediately from the above claim and the union bound.

Proof of Claim 6.10. Set $\epsilon_0 = n^{-(2d+6)/(3d+5)}$. Since K is smooth, there is a constant a depending on K such that $g(\epsilon_0) \le a\epsilon_0$.

We define $\delta_0 = c_0 n^{-1}$, $T_0 = c_0\rho_{1/n}$, $\delta_j = 2^j\delta_0$, $T_j = (j+1)^{-2}4^{-j}T_0$, where c_0 is a positive constant to be determined, for all $j = 1, 2\dots$. Let j_0 be the first integer such that $\delta_{j_0} \ge a\epsilon_0$, it is easy to verify that for any $j \le j_0$, the assumption of Lemma 6.8 is satisfied by the pair (δ_j, T_j). Thus, in this range of j, one can apply Lemma 6.8 for (δ_j, T_j).

Recall that (see the proof of Claim 5.1)

$$C_i(x,t) \le \mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1}) + \mathbf{E}_{x'}\mathbf{E}(\Delta_{x',L}|t_1,\dots,t_{i-1})\,.$$

Thus,

$$\begin{aligned} V_i(t) &= \int_K C_i(x,t)^2\partial x \\ &\le \int_K \big(\mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1}) + \mathbf{E}_{x'}\mathbf{E}(\Delta_{x',L}|t_1,\dots,t_{i-1})\big)^2\partial x \\ &\le 4\int_K \mathbf{E}(\Delta_{x,L}|t_1,\dots,t_{i-1})^2\partial x \end{aligned}$$

by Holder's inequality. We say that the set $L = \{t_1,\dots,t_{i-1},t_{i+1},\dots,t_n\}$ is *nice* if the following hold:

- $\mathrm{Conv}(L)$ contains the floating body F_{ϵ_0};
- $\mathrm{Vol}(\mathrm{Conv}(L)) \geq 1 - c_1\rho_{1/n}$ for some sufficiently large constant c_1;
- $\mathrm{Vol}(X_{\delta_j,L}) \leq T_j$ for all $j = 0, 1, 2, \dots j_0$.

Following the proof of Theorem 2.7, we say that the set $\{t_1, \dots, t_{i-1}\}$ is typical if

$$\mathbf{P}_{\Omega^{\langle i+1\rangle}}(L \text{ is not nice}|t_1, \dots, t_{i-1}) \leq n^{-4}.$$

Similarly to the proof of Theorem 2.7, we are going to conclude in two steps. In the first step, we show that if $\{t_1, \dots, t_{i-1}\}$ is typical, then $V_i(t) \leq n^{-1}V$. In the second step, we bound the probability that $\{t_1, \dots, t_{i-1}\}$ is not typical.

First step. Assume that $\{t_1, \dots, t_{i-1}\}$ is typical. To estimate $\int_K \mathbf{E}(\Delta_{x,L}|t_1, \dots, t_{i-1})^2 \partial x$, observe that

$$\int_K \mathbf{E}(\Delta_{x,L}|t_1, \dots, t_{i-1})^2 \partial x = \int_{\Omega^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x \right) \partial t^{\langle i+1\rangle}$$

where $L = \{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n\}$. Let $\Omega_1^{\langle i+1\rangle}$ be the set of those $(t_{i+1}, \dots, t_n)$ such that the set L is nice and $\Omega_2^{\langle i+1\rangle}$ be the rest.

$$\begin{aligned}\int_{\Omega^{\langle i+1\rangle}} &\left(\int_K \Delta_{x,L}^2 \partial x \right) \partial t^{\langle i+1\rangle} \\ &= \int_{\Omega_1^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x \right) \partial t^{\langle i+1\rangle} + \int_{\Omega_2^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x \right) \partial t^{\langle i+1\rangle}.\end{aligned}$$

Since $\Delta_{x,L}$ is at most one, $\int_K \Delta_{x,L}^2 \partial x$ is at most one and

$$\int_{\Omega_2^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x \right) \partial t^{\langle i+1\rangle} \leq \mathbf{P}_{\Omega^{\langle i+1\rangle}}(L \text{ not nice}|t_1, \dots, t_{i-1}) \leq n^{-4}$$

since $\{t_1, \dots, t_{i-1}\}$ is typical. Furthermore, $\Delta_{x,L}$ is zero if $x \in \mathrm{Conv}(L)$. Thus, for a fixed vector $t^{\langle i+1\rangle}$ such that L is nice, the collection of those x's where $\Delta_{x,L}$ is not zero has measure at most $1 - \mathrm{Vol}(\mathrm{Conv}(L)) \leq c_1\rho_{1/n}$ and

$$\begin{aligned}\int_K \Delta_{x,L}^2 \partial x &\leq \delta_0^2 c_1 \rho_{1/n} + \sum_{j=0}^{\infty} \delta_{j+1}^2 \mathrm{Vol}(X_{\delta_j,L}) \\ &\leq c_1\delta_0^2\rho_{1/n} + \sum_{j=0}^{j_0} \delta_{j+1}^2 T_j \\ &\leq c_1\delta_0^2\rho_{1/n} + \sum_{j=0}^{j_0} 4^{j+1}\delta_0^2 \times (j+1)^{-2} 4^{-j} T_0 \\ &\leq c_1\delta_0^2\rho_{1/n} + 4\tfrac{\pi^2}{6}\delta_0^2 T_0\end{aligned}$$

$$= O(n^{-2}\rho_{1/n})\,.$$

In the second inequality, we can change $\sum_0^\infty$ to $\sum_0^{j_0}$ because for $j > j_0$, δ_j is larger than $g(\epsilon_0)$ so $X_{\delta_j,L}$ is empty. (Notice that the floating body F_{ϵ_0} is contained in the convex hull of L. Thus there is no point x such that $\Delta_{x,L}$ is larger than $g(\epsilon_0)$.) Therefore,

$$\int_{\Omega_1^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x\right)\partial t^{\langle i+1\rangle} = \int_{\Omega^{\langle i+1\rangle}} O(n^{-2}\rho_{1/n})\partial t^{\langle i+1\rangle} = O(n^{-2}\rho_{1/n})\,.$$

It follows that

$$\int_{\Omega^{\langle i+1\rangle}} \left(\int_K \Delta_{x,L}^2 \partial x\right)\partial t^{\langle i+1\rangle} = O(n^{-2}\rho_{1/n}) + n^{-4} \le c_2 n^{-2}\rho_{1/n}$$

for some constant c_2.

Second Step. We bound the probability that $\{t_1,\dots,t_{i-1}\}$ is not typical. Again we use Lemma 3.2 in a way similar to the proof of Theorem 2.7.

Lemma 4.2, Corollary 6.9 and Lemma 6.8 imply that the probability that a set L of $n-1$ random points is not nice is at most

$$\exp\big(-\Omega(n\epsilon_0)\big) + \exp\big(-\Omega\big(n^{(d-1)/(3d+5)}\big)\big) + \sum_{j=0}^{j_0} \exp\big(-\Omega(nT_j)\big)\,. \quad (21)$$

By the definition of ϵ_0, $n\epsilon_0 = n^{(d-1)/(3d+5)}$. Moreover, by the definition of j_0, δ_i and T_i,

$$nT_j \ge nT_{j_0} \gg n^{(d-1)/(3d+5)}$$

for any $0 \le j \le j_0$. Thus, the term $\exp(-\Omega(n^{(d-1)/(3d+5)}))$ in the above sum is dominating. So we can conclude that the probability that L is not nice is at most

$$\exp(-\Omega\big(n^{(d-1)/(3d+5)}\big))\,. \quad (22)$$

So by Lemma 3.2, the probability that $\{t_1,\dots,t_{i-1}\}$ is not typical is at most

$$n^4 \exp\big(-\Omega\big(n^{(d-1)/(3d+5)}\big)\big) = \exp\big(-\Omega\big(n^{(d-1)/(3d+5)}\big)\big)$$

concluding the proof. □

7 Proof of Theorem 2.11

Let L be a set of points. For a point x, define

$$\Gamma_{x,L} = \big|N(\mathrm{Conv}(L\cup x)) - N(\mathrm{Conv}(L))\big|\,.$$

$\Gamma_{x,L}$ will play the role of $\Delta_{x,L}$ in the proofs of Theorems 2.1 and 2.7. It is trivial that if x belongs to $\mathrm{Conv}(L)$, then $\Gamma_{x,L}$ is zero. Notice that $N(\mathrm{Conv}(L\cup x)) - N(\mathrm{Conv}(L))$ can be both positive and negative, so the absolute value sign is necessary.

For $k \geq 2$, we say that a point x is *k-wide* with respect to L if $\Gamma_{x,L} = k$. Define

$$U_{k,L} = \{x | x \text{ is } k - \text{wide } L\}.$$

The quantity $U_{k,L}$ is an analogue of $X_{\delta,L}$. The following lemma is a variant of Lemma 6.8.

LEMMA 7.1. *There are positive constants c, c', c'' and c''' such that the following holds. For any $k \geq c''$ and $T \geq c'' \max\{k/c'''n, \rho_{k/c'''n} \exp(-c'k)\}$ we have*

$$\mathbf{P}\big(\mathrm{Vol}(U_{k,L}) \geq T\big) \leq \exp(-cnT)$$

where L is a set of n random points.

Proof of Lemma 7.1. Set $\delta = k/c'''n$. For a point $x \notin \mathrm{Conv}(L)$, if $x \notin E_{\delta,L}$, then every cap containing x and disjoint from L has volume less than δ. Then by arguing as in Lemma 6.7, there is a cap of volume $O(\delta)$ containing $\mathrm{Conv}(L \cup x) \backslash \mathrm{Conv}(L)$.

Let $U_{k,\delta,L}$ be the set of all x in $U_{k,L}$ which are not in $E_{\delta,L}$. Clearly

$$U_{k,L} \subset E_{\delta,L} \cup U_{k,\delta,L}.$$

Therefore

$$\mathbf{P}\big(\mathrm{Vol}(U_{k,L}) \geq T\big) \leq \mathbf{P}\big(\mathrm{Vol}(E_{\delta,L}) \geq T/2\big) + \mathbf{P}\big(\mathrm{Vol}(U_{k,\delta,L}) \geq T/2\big). \quad (23)$$

We can bound $\mathbf{P}(\mathrm{Vol}(E_{\delta,L}) \geq T/2)$ using Lemma 6.8. (By setting $c'' \geq 2$ the assumption of this lemma is satisfied.) This gives

$$\mathbf{P}(E_{\delta,L}) \geq T/2 \leq \exp(-\Omega(nT)).$$

To bound $\mathbf{P}(\mathrm{Vol}(U_{k,\delta,L}) \geq T/2)$, let $U'_{k,\epsilon,L}$ be the union of all caps of volume at most ϵ and contains at least k points from L. By the argument in the first paragraph of the proof, there is some constant c_0 such that

$$U_{k,\delta,L} \subset U'_{k,c_0\delta,L}. \quad (24)$$

Arguing as in the proof of Lemma 6.4 (using the maximal disjoint subsystem trick), we can find a set $C_1, \dots, C_l$ of disjoint $(c\delta)$-caps in $U'_{k,c_0\delta,L}$, whose union has volume at least $c_1 \mathrm{Vol}(U'_{k,c\delta,L})$, for some constant c_1. (We need to set c'' sufficiently large here to guarantee that l is at least one.) It is essential to keep in mind that

$$l = \Omega\big(\delta^{-1} \mathrm{Vol}(U'_{k,c\delta,L})\big). \quad (25)$$

Using the covering lemma (Lemma 6.6), there is a constant c_2 and a fixed system of $(c_2\delta)$-caps $D_1, \dots, D_m$ such that each C_i is in some D_j. Assume, for convenience, that $C_i \subset D_i$ for $i = 1, \dots, l$. Next, we use the maximal disjoint subsystem trick again to get a subsystem of $l' = c_3 l$

pairwise disjoint elements from $D_1, \dots, D_l$. We have proved the following. The fixed system $D_1, \dots, D_m$ contains

$$l' = \Omega\big(\delta^{-1}\mathrm{Vol}(U'_{k,c\delta,L})\big)\,. \tag{26}$$

pairwise disjoint elements, each of which contains at least k points from L.

The probability that a fixed δ-cap D contains at least k points from a random set of n points is

$$\sum_{i=k}^{n}\binom{n}{i}\delta^i(1-\delta)^{n-i} \le \sum_{i=k}^{n}\binom{n}{i}\delta^i \le \sum_{i=k}^{n}\left(\frac{en}{i}\right)^i\delta^i.$$

In the second we use the well-known inequality $\binom{n}{i} \le (\frac{en}{i})^i$. Assume that $en\delta < k/2$. Thus, the terms in the sum decrease by a factor at least 2, thus the probability is bounded from above by

$$2\left(\tfrac{en}{ik}\right)^k\delta^k \le \left(\tfrac{3\delta n}{k}\right)^k$$

provided that $k \ge 3$.

Let $D_1, \dots, D_{l'}$ be pairwise disjoint $(c_2\delta)$-caps. By setting c''' sufficiently large, we have $k/2 \ge enc_2\delta$, so we can apply the above calculation. The probability that D_j contains at least k points is at most $(3c_2\delta n/k)^k$. These events are negatively correlated, so the probability that each of the D_j contains at least k point is at most

$$\left(\left(\frac{3c_2\delta n}{k}\right)^k\right)^{l'} = \exp\left(-\Omega(kl'\ln\frac{k}{3c_2\delta n}\right) = \exp\big(-\Omega(kl')\big)\,.$$

Recall that $l' = \Omega(\delta^{-1}\mathrm{Vol}(U'_{k,c\delta,L})$. It follows that

$$\begin{aligned}\mathbf{P}\big(\mathrm{Vol}(U_{k,\delta,L}) \ge T/2\big) &\le \mathbf{P}\big(\mathrm{Vol}(U'_{k,\delta,L}) \ge T/2\big)\\ &\le \exp\big(-\Omega(k\delta^{-1}T)\big)\\ &= \exp\big(-\Omega(nT)\big)\,,\end{aligned}$$

proving Lemma 7.1.

To prove Theorem 2.11, there are two issues. The bound for $C(t)$ and the bound for $V(t)$. To handle $C(t)$ notice that first one can show with very high probability $\Delta_{x,L}$ is at most $g(\epsilon) \le c_0\epsilon$. Next, each cap of size $c_0\epsilon$ is contained in some member of C_i of a fixed collection $C_1, \dots, C_m$ of caps of size $c_1\epsilon$, where $m = O(\rho_{ep}\epsilon^{-1})$. The probability that C_i contains more than $2nc_1\epsilon$ points (twice the expected number) is $\exp(-\Omega(n\epsilon))$. Thus by choosing $\epsilon \gg n^{-1}\ln n$, we can conclude that $\mathbf{P}(C(t) \ge \epsilon) = \exp(-\Omega(\epsilon n))$. To handle $V(t)$, one basically needs to rework the proof of Theorem 2.7

with $\Gamma_{x,L}$ playing the role of $\Delta_{x,L}$ and $U_{k,L}$ playing the role of $X_{\delta,L}$. We omit the details, which are technical, but routine.

8 Concluding Remarks

8.1 A corollary of Lemma 7.1. A tantalizing problem concerning the number of vertices of K_n is the monotonicity of expectations. As before, Z_n denotes the number of vertices of K_n.

QUESTION. Is it true that $\mathbf{E}(Z_{n+1}) \geq \mathbf{E}(Z_n)$?

Lemma 7.1 sheds some light on this question. Let L be a collection of n random points and x be an extra random point. The problem here is that adding x to L can decrease the number of vertices of the convex hull. This happens if x is outside $\mathrm{Conv}(L)$ and sees many facets of L. In this case, adding x we have a new vertex in x but may lose more than one original vertices.

We have a fairly good idea what the probability that x is outside $\mathrm{Conv}(L)$ is. This probability is exactly the volume of $\overline{K_n}$. In order to show that $\mathbf{E}(Z_{n+1}) \geq \mathbf{E}(Z_n)$, one would like to prove that typically the probability that x see many facets is small compared to $\mathrm{Vol}(\overline{K_n})$.

We know that with high probability, $\mathrm{Vol}(\overline{K_n}) = \Theta(\rho_{1/n})$. Lemma 7.1 implies the following corollary, which asserts that the probability that x sees k vertices from $\mathrm{Conv}(L)$ is exponentially small in k, conditioned on x lying outside $\mathrm{Conv}(L)$. This bound the probability that one loses $k-1$ vertices by adding x to L.

COROLLARY 8.1. *There are positive constants c and c' such that the following holds. Let L be a collection of n random points and x be an extra random point. For any $k \geq 1$, the probability that x is k-wide with respect to L is at most $c'\rho_{1/n}\exp(-ck)$.*

COROLLARY 8.2. *There are positive constants c and c' such that the following holds. Let L be a collection of n random points and x be an extra random point. For any $k \geq 2$, the probability that*

$$N\big(\mathrm{Conv}(L)\big) - N\big(\mathrm{Conv}(L \cup x)\big) \geq k-1$$

is at most $c'\rho_{1/n}\exp(-ck)$.

Since $\int_0^\infty t^l \exp(-ct)\partial t$ converges for any fixed l, we have

COROLLARY 8.3. *Let x be a random point and $V_n(x)$ be the number of vertices of K_n seen from x. Then for any fixed integer $l \geq 1$, there is a* *constant c_l such that*

$$\mathbf{E}\big(V_n(x)^l\big) \le c_l \rho_{1/n} = O\big(n^{-2/(d+1)}\big)\,.$$

It x see at most k vertices, then it sees at most $O(k^{d+1})$ facets. Therefore,

COROLLARY 8.4. *Let x be a random point and $F_n(x)$ be the number of facets of K_n seen from x. Then for any fixed integer $l \ge 1$, there is a constant c_l such that*

$$\mathbf{E}\big(F_n(x)^l\big) \le c_l \rho_{1/n} = O\big(n^{-2/(d+1)}\big)\,.$$

8.2 Other random models. Our method can be applied for other models of random polytopes as well. For instance, it is not too hard to derive results for the model of random polytopes with points chosen from the normal distribution in $\mathbb{R}^d$. Details will appear elsewhere.

In all models we know of, the points $t_1, \dots, t_n$ are drawn from the same distribution. On the other hand, our key tool, Lemma 3.1, is still valid when these points are drawn from different distribution (see Lemma 9.1 in the Appendix). This may lead to a new direction of research.

9 Appendix

9.1 Proof of Lemma 3.1. We shall prove the lemma in an even more general setting, where we allow the random variables $t_1, \dots, t_n$ to be chosen from different probability spaces. To this end, let $\mathbf{\Omega_1}, \dots, \mathbf{\Omega_n}$ be probability spaces and $\mathbf{\Omega}^*$ be their product (with the natural product measure); $t = (t_1, \dots, t_n)$ denote a point from $\mathbf{\Omega}^*$, where $t_i \in \mathbf{\Omega_i}$ are the coordinates of t. Let $Y = Y(t)$ be a real value function on $\mathbf{\Omega}^*$. We will assume that measurability is not an issue in what follows.

As in section 3, define

$$C_i(x,t) = \big|\mathbf{E}(Y|t_1, \dots, t_{i-1}, x) - \mathbf{E}(Y|t_1, \dots, t_{i-1})\big|\,.$$

Furthermore, set

$$V_i(t) = \int_{\mathbf{\Omega_i}} C_i(x,t)^2 \partial x\,,$$

$$V(t) = \sum_{i=1}^n V_i(t)\,,$$

$$C_i(t) = \sup_{x \in \mathbf{\Omega_i}} C_i(x,t)\,,$$

and

$$C(t) = \max_{i=1}^{n} C_i(t) = \sup_{i,x\in\mathbf{\Omega_i}} C_i(x,t)\,.$$

LEMMA 9.1. *For any positive λ, C and V satisfying $\lambda \le V/4C^2$, we have*

$$\mathbf{P}\big(|Y-\mathbf{E}(Y)| \ge \sqrt{\lambda V}\big) \le 2\exp\left(-\tfrac{\lambda}{4}\right) + \mathbf{P}\big(V(t)\ge V \text{ or } C(t)\ge C\big)\,. \quad (27)$$

Lemma 3.1 is the special case when $\mathbf{\Omega_1} = \cdots = \mathbf{\Omega_n} = \mathbf{\Omega}$. To prove Lemma 9.1, we first need the following lemma.

LEMMA 9.2. *Let X be a real value function on $\mathbf{\Omega}^*$. Assume that $V, C\lambda$ are positive number satisfying $C_X \le C$, $V_X \le V$ and $\lambda \le 4V/C^2$. Then*

$$\mathbf{P}\big(|X-\mathbf{E}(X)| \ge \sqrt{\lambda V}\big) \le 2e^{-\lambda/4}\,.$$

Proof of Lemma 9.2. We can assume, without loss of generality that $\mathbf{E}(X)=0$. So it suffices to prove that

$$\mathbf{P}\big(|X| \ge \sqrt{\lambda V}\big) \le 2e^{-\lambda/4}. \quad (28)$$

LEMMA 9.3. *Let Z be a function form $\mathbf{\Omega}^*$ to $\mathbb{R}$ with mean 0. If $u \le 1/C_Z$, then*

$$\mathbf{E}(e^{uZ}) \le e^{u^2V_Z}.$$

To see that Lemma 9.3 implies (28), set $u=\sqrt{\lambda/4V}$. Since $\lambda \le 4V/C^2$, $u \le 1/C \le 1/C_X$. Lemma 9.3 yields

$$\mathbf{E}(e^{uX}) \le e^{u^2V_X} \le e^{u^2V}.$$

By Markov's inequality

$$\begin{aligned}\mathbf{P}\big(X \ge \sqrt{\lambda V}\big) &= \mathbf{P}(e^{uX} \ge e^{u\sqrt{\lambda V}})\\ &= \mathbf{P}(e^{uX} > e^{\lambda/2})\\ &\le e^{u^2V-\lambda/2} = e^{-\lambda/4}\,.\end{aligned}$$

Inequality (28) follows by symmetry.

Proof of Lemma 9.3. First we need the following simple statement.

PROPOSITION 9.4. *Let $\mathbf{\Omega}$ be an arbitrary probability space. Suppose that $f(x)$ is a measurable function on $\mathbf{\Omega}$ which has absolute value at most 1 and mean 0 (with respect to μ). Then*

$$\int_{\mathbf{\Omega}} e^{f(x)}\partial x \le e^{\int_{\mathbf{\Omega}} f^2(x)\partial x}.$$

Equality holds if and only if $f \equiv 0$.

Proof of Proposition 9.4. Assume that f is not identically zero. By Taylor's series expansion

$$\int_{\mathbf{\Omega}} e^{f(x)}\partial x = 1 + \int_{\mathbf{\Omega}} f(x)\partial x + \frac{1}{2!}\int_{\mathbf{\Omega}} f^2(x)\partial x + \ldots$$

Since $\int_{\mathbf{\Omega}} f(x)\partial x = 0$ and $|f(x)| \leq 1$, the right-hand side is at most

$$1 + \sum_{i=2}^{\infty} \frac{1}{i!} \int_{\mathbf{\Omega}} f^2(x)\partial x < 1 + \int_{\mathbf{\Omega}} f^2(x)\partial x < e^{\int_{\mathbf{\Omega}} f^2(x)\partial x}. \qquad \square$$

REMARK. As $c_0 = \sum_{i=2}^{\infty} 1/i! < 1$, we can improve the proposition by replacing $\int_{\mathbf{\Omega}} f^2(x)\partial x$ in the exponent by $c_0 \int_{\mathbf{\Omega}} f^2(x)\partial x$. This may lead to a constant better than $1/4$ in the exponent of the bound in Lemma 9.2. However, from the practical point of view, this does not make a big difference and we do not try to optimize the constants in this paper.

The proof of Lemma 9.3 uses induction on n. If $n = 1$, then for all $x \in \mathbf{\Omega_1}$, $C_1(x,t) = |Z(x)|$ and $V(x) = \int_{\mathbf{\Omega_1}} Z^2(x)\partial x$, where the measure is generalized by the unique variable $t_1 = x$. This yields that $|uZ(x)| \leq uC_Z \leq 1$ for all x and the statement of the lemma follows from Proposition 9.4 by substituting $f = uZ$.

Now consider a generic n. First notice $C_1(x,t) = |\mathbf{E}(Z|t_1 = x)|$ does not depend on t. Consequently, $V_1(t) = \int_{\mathbf{\Omega_1}} C_1^2(x,t)\partial x$ is a constant and we can set $V_1 = V_1(t)$. Let $\mathbf{\Omega'}$ be the product of the spaces $\mathbf{\Omega_2}, \dots, \mathbf{\Omega_n}$. For each $x \in \mathbf{\Omega_1}$, consider the following function from $\mathbf{\Omega'}$ to $\mathbb{R}$:

$$Z_x(t') = Z(x, t_2, \dots, t_n) - \mathbf{E}(Z|t_1 = x)\,.$$

By definition, $Z_x(t')$ has mean 0, for any $x \in \mathbf{\Omega_1}$. Moreover, $V_{Z_x} \leq V_Z - V_1$. By the induction hypothesis

$$\mathbf{E}_{\mathbf{\Omega_1}}(e^{uZ_x}) \leq e^{u^2(V_{Z_x})} \leq e^{u^2(V_Z - V_1)}.$$

On the other hand, by Fubini's theorem

$$\begin{aligned}
\mathbf{E}_{\mathbf{\Omega}^*}(e^{uZ}) &= \mathbf{E}_{\mathbf{\Omega'}}\left(\int_{\mathbf{\Omega_1}} e^{uZ_x} e^{u\mathbf{E}(Z|t_1=x)}\partial x\right) \\
&= \int_{\mathbf{\Omega_1}} e^{u\mathbf{E}(Z|t_1=x)}(\mathbf{E}_{\mathbf{\Omega'}} e^{uZ_x})\partial x \\
&\leq \int_{\mathbf{\Omega_1}} e^{u\mathbf{E}(Z|t_1=x)} e^{u^2(V_Z - V_1)}\partial x \\
&= e^{u^2(V_Z - V_1)} \int_{\mathbf{\Omega_1}} e^{u\mathbf{E}(Z|t_1=x)}\partial x\,.
\end{aligned}$$

Set $f(x) = u\mathbf{E}(Z|t_1 = x)$. By the assumption on u, $f(x)$ satisfies the conditions of Proposition 9.4, so

$$\int_{\mathbf{\Omega_1}} e^{u\mathbf{E}(Z|t_1=x)}\partial x = \int_{\mathbf{\Omega_1}} e^{f(x)}\partial x \leq e^{\int_{\mathbf{\Omega_1}} f^2(x)\partial x} = e^{u^2 V_1}\,,$$

completing the proof. $\square$

Now we prove Lemma 9.1 with the help of Lemma 9.2.

We say that a vector t is bad if either $C(t) \geq C$ or $V(t) \geq V$. Let B be the collection of bad vectors,

$$\mathbf{B} = \{t | C(t) \geq C \text{ or } V(t) \geq V\}.$$

For any $t \in \mathbf{B}$, let $i(t)$ be the smallest index i (between 1 and n) such that either $C_i(x,t) \geq C$ for some x or $\sum_{j=1}^{i} V_j(t) \geq V$. Let $\mathbf{B}_t = \{z \in \mathbf{\Omega}^* \mid \mathbf{z_i} = \mathbf{t_i}$ for all $i < i(t)\}$. It is clear that

- $\mathbf{B}_t \subset \mathbf{B}$;
- For any $t, t' \in \mathbf{B}$, $\mathbf{B}_t$ and $\mathbf{B}_{t'}$ are either identical or disjoint.

Define a function Y' as follows:

- $Y'(z) = Y(z)$ if $z \notin \mathbf{B}$;
- $Y'(z) = \mathbf{E}_{\mathbf{B}_t}(Y)$ if $z \in \mathbf{B}_t \subset \mathbf{B}$.

In other words, if z is not bad, then we keep the value of $Y(z)$. If z is bad, then there is a unique set $\mathbf{B}_t$ containing z and we change $Y(z)$ to the expectation of Y over this set. It is immediate that the expectations of Y and Y' are the same. Moreover,

$$C_{Y'} \leq C$$
$$V_{Y'} \leq V$$
$$\mathbf{P}(Y \neq Y') \leq \mathbf{P}(\mathbf{B}).$$

We can conclude the proof by applying Lemma 9.2 to Y'. □

9.2 Proof of Corollary 2.3. Notice that for $\lambda > \lambda_0$ it is trivial that

$$\begin{aligned}\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) &\leq \mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda_0 V}\big) \\ &\leq 2\exp(-\lambda_0/4) + 2n\exp(-c\epsilon n).\end{aligned}$$

It follows that for any $\lambda > 0$,

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq \sqrt{\lambda V}\big) \leq 2\exp(-\lambda/4) + 2\exp(-\lambda_0/4) + 2n\exp(-c\epsilon n). \tag{29}$$

Now we are going to adapt the argument shown in subsection 2.2 to estimate M_k, the kth moment of Y. To start, we have

$$M_k = \int_0^\infty t^k \partial\mathbf{P}\big(|Y - \mathbf{E}(Y)| < t\big).$$

Notice that in this case Y, which is the volume of K_n, is at most one. Thus, we only need to integrate from 0 to 1. Therefore, with $\mu(t) = \mathbf{P}(|Y - \mathbf{E}(Y)| \geq t)$, we have

$$\begin{aligned}M_k &= \int_0^1 t^k \partial\mathbf{P}\big(|Y - \mathbf{E}(Y)| < t\big) \\ &= -\int_0^1 t^k \partial\mu(t)\end{aligned}$$

$$= \left(\left(-t^k\mu(t)|_0^1 \right) + \int_0^1 kt^{k-1}\mu(t)\partial t \right)$$
$$= \int_0^1 kt^{k-1}\mu(t)\partial t \,.$$

Replacing $t = \sqrt{\lambda V}$, we obtain

$$\int_0^1 kt^{k-1}\mu(t)\partial t \le \int_0^{1/V} k\big(\sqrt{\lambda V}\big)^{k-1}\mathbf{P}\big(|Y-\mathbf{E}(Y)| \ge \sqrt{\lambda V}\big)\frac{\sqrt{V}}{2\sqrt{\lambda}}\partial\lambda$$
$$\text{by(29)} \quad \le \frac{k}{2}V^{k/2}\int_0^{1/V} \lambda^{\frac{k}{2}-1}2\big((\exp(-\lambda/4)+\exp(-\lambda_0)+n\exp(-c\epsilon n)\big)\partial\lambda$$
$$= kV^{k/2}\int_0^{1/V} \lambda^{\frac{k}{2}-1}\big((\exp(-\lambda/4)+\exp(-\lambda_0)+n\exp(-c\epsilon n)\big)\partial\lambda \,.$$

We integrate each term in the sum separately. First of all,

$$\int_0^{1/V} \lambda^{\frac{k}{2}-1}\exp(-\lambda/4)\partial\lambda \le \int_0^{\infty} \lambda^{\frac{k}{2}-1}\exp(-\lambda/4)\partial\lambda = c_k \,, \qquad (30)$$

where c_k depends only on k.

Recall that $\epsilon = \alpha \ln n/n$. We are going to show that if the constant α is sufficiently large, then the contribution of the other two terms is negligible. In this estimate, the upper limit $1/V$ does play a role. By definition

$$V = 36ng(\epsilon)^2\rho_\epsilon \ge 36n\epsilon^3 \ge n^{-2} \,,$$

as $\rho_\epsilon \ge \epsilon$ and $\epsilon = \alpha\frac{\ln n}{n} \ge n^{-1}$. Moreover, by Theorem 2.1,

$$\lambda_0 = n\rho_\epsilon \ge n\epsilon = \alpha \ln n \,.$$

It follows that

$$\int_0^{1/V} \lambda^{\frac{k}{2}-1}\exp(-\lambda_0)\partial\lambda \le \int_0^{1/V} \lambda^{\frac{k}{2}-1}n^{-\alpha}\partial\lambda$$
$$= \tfrac{2}{k}(1/V)^{k/2}n^{-\alpha}$$
$$\le \tfrac{2}{k}n^k n^{-\alpha}.$$

So if we have $\alpha > k$ then

$$\int_0^{1/V} \lambda^{\frac{k}{2}-1}\exp(-\lambda_0)\partial\lambda = o(1) \,. \qquad (31)$$

Now we focus on the last term

$$\int_0^{1/V} \lambda^{\frac{k}{2}-1}n\exp(-c\epsilon n)\partial\lambda \le \tfrac{2}{k}n^{k+1}n^{-c\alpha}.$$

If we have $\alpha > (k+1)/c$ then

$$\int_0^{1/V} \lambda^{\frac{k}{2}-1}\exp(-c\epsilon n)\partial\lambda = o(1) \,. \qquad (32)$$

From (30)–(32) we can deduce that for any fixed k, there is a constant c_k (defined as in (30) such that

$$M_k \le \big(c_k + o(1)\big)kV^{k/2} = O(V^{k/2})\,, \tag{33}$$

where $\epsilon = \alpha \ln n/n$ for some sufficiently large constant α and

$$V = 36ng(\epsilon)^2\rho_\epsilon\,. \qquad \square$$

Acknowledgement. We would like to thank I. Bárány for bringing the subject to our attention and many enlightening conversations.

References

[AS] N. Alon, J. Spencer, The Probabilistic Method (2nd Eidition), Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[Az] K. Azuma, Weighted sums of certain dependent random variables, Tohoku Math. J. (2) 19 (1967), 357–367.

[B] I. Bárány, Intrinsic volumes and f-vectors of random polytopes, Math. Ann. 285:4 (1989), 671–699.

[BD] I. Bárány, L. Dalla, Few points to generate a random polytope, Mathematika 44:2 (1997), 325–331.

[BL] I. Bárány, D. Larman, Convex bodies, economic cap coverings, random polytopes, Mathematika 35:2 (1988), 274–291.

[Bo] K.H. Borgwardt, The simplex Method: A Probabilistic Analysis, Algorithms and Combinatorics: Study and Research Texts, 1. Springer-Verlag, Berlin, 1987.

[CG] A.J. Cabo, P. Groeneboom, Limit theorems for functionals of convex hulls, Probab. Theory Related Fields 100:1 (1994), 31–55.

[D] R. Dwyer, On the convex hull of random points in a polytope, J. Appl. Probab. 25:4 (1988), 688–699.

[E] H. Edelsbrunner, Algorithms in Combinatorial Geometry, Springer-Verlag, Heidelberg, Germany, 1987.

[Ef] B. Efron, The convex hull of a random set of points, Biometrika 52 (1965), 331–343.

[G] P. Gruber, Comparisons of best and random approximation of convex bodies by polytopes, in "Stochastic Geometry, Convex Bodies and Empirical Measures," II International Conference (Agrigento, 1996), Rend. Circ. Mat. Palermo (2) Suppl. 50 (1997), 189–216.

[HW] D. Haussler, E. Welzl, ϵ-nets and simplex range queries, Discrete Comput. Geom. 2:2 (1987), 127–151.

[Hu] I. Hueter, Limit theorems for the convex hull of random points in higher dimensions, Trans. Amer. Math. Soc. 351:11 (1999), 4337–4363.

[KV1] J.H. Kim, V.H. Vu, Concentration of multi-variate polynomials and its applications, Combinatorica 20:3 (2000), 417–434.

[KV2] J.H. Kim, V.H. Vu, Divide and conquer martingales and the number of triangles in a random graph, Random Structures Algorithms 24:2 (2004), 166–174.

[M] J. Matousek, Lectures on Discrete Geometry, Graduate Texts in Mathematics 212. Springer-Verlag, New York, 2002.

[Mi] V. Milman, A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies, Funkcional. Anal. i Priložen. 5:4 (1971),28–37.

[PS] F. Preparata, I. Shamos, Computational Geometry. An introduction, Texts and Monographs in Computer Science, Springer-Verlag, New York, 1985.

[R1] M. Reitzner, Random polytopes and the Efron–Stein jackknife inequality, Ann. Probab. 31:4 (2003), 2136–2166.

[R2] M. Reitzner, Central Limit Theorems for Random Polytopes, Probability and Related Fields, to appear.

[RéS] A. Rényi, R. Sulanke, Über die konvexe Hülle von n zufällig gewälten Punkten, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 2 (1963), 75–84.

[S] M. Steele, Probability theory and combinatorial optimization, CBMS-NSF Regional Conference Series in Applied Mathematics, 69, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[V1] V.H. Vu, Concentration of non-Lipschitz functions and applications, Random Structures and Algorithms 20:3 (2002), 262–316.

[V2] V.H. Vu, On the concentration of multivariate polynomials with small expectation, Random Structures Algorithms 16:4 (2000), 344–363.

[V3] V.H. Vu, Central limit theorems for random polytopes in a smooth convex set, Advances in Mathematics, to appear.

[WW] W. Weil, J. Wieacker, Stochastic geometry, Handbook of Convex Geometry, Vol. A, B, 1391–1438, North-Holland, Amsterdam, 1993.

Van H. Vu, Department of Mathematics, UCSD, La Jolla, CA 92093-0112, USA
http://www.math.ucsd.edu/~ vanvu vanvu@ucsd.edu
and
Department of Mathematics, Rutgers, NJ 08854, USA

Received: September 2004
Revision: November 2004
Accepted: November 2004