

Recurrence relationships for the mean number of faces and vertices for random convex hulls

Richard Cowan¹

Abstract: This paper studies the convex hull of n random points in \mathbb{R}^d . A recently-proved topological identity of the author is used in combination with identities of Efron and Buchta to find the expected number of vertices of the convex hull – yielding a new recurrence formula for all dimensions d . A recurrence for the expected number of facets and $(d-2)$ -faces is also found, this analysis building on a technique of Rényi and Sulanke. Other relationships for the expected count of i -faces ($1 \leq i < d$) are found when $d \leq 5$, by applying the Dehn–Sommerville identities. A general recurrence identity (see equation (3) below) for this expected count is conjectured.

Published version: This is the personal-archive version of the paper which is published in the journal *Discrete and Computational Geometry*, 2008. The original publication is available at www.springerlink.com. The Springer DOI is: 10.1007/s00454-008-9122-6.

Introduction

Consider the experiment where points $P_1, P_2, \dots, P_n, \dots$ are placed randomly and independently in \mathbb{R}^d ($d \geq 1$) according to a common probability law. This law is defined by the (induced) probability measure μ defined on \mathcal{B}_d , the Borel sets of \mathbb{R}^d . Let the convex hull of the first n points be denoted by H_n . The d -dimensional volume of H_n is given by V_n , whilst μ_n is defined as the probability content (that is, μ -measure) of H_n .

The convex hull H_n is a convex polytope. We denote the number of i -faces of H_n by $F_{n,i}$. When $i = 0$, however, we follow the practice of other authors and use the notation $N_n := F_{n,0}$, for the number of vertices. We also introduce $F_n := F_{n,d-1}$ as the number of facets.

Combining the power of three distinct ideas – the Efron/Buchta identity ([11], [5], [17]),

¹Richard Cowan, School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia. email: R.Cowan@maths.usyd.edu.au

an integral technique of Rényi/Sulanke ([15],[16]) modified to suit our purpose, and an identity involving μ_n that follows from recent work of the author [8] – we derive various new recurrence identities for $\mathbb{E}N_n$, $\mathbb{E}F_n$ and $\mathbb{E}F_{n,d-2}$. To be more precise, we note that these identities are only *partial recurrences* as there are parity conditions on them. We do not, however, use the word ‘partial’ in further descriptions.

Mostly we focus on the case where μ gives zero probability mass to every j -dimensional hyperplane (for all $j < d$) and, in particular, to any one-point set. This implies that the polytope H_n is almost surely *simplicial* with dimension equal to $\min(d, n - 1)$. We refer to this constraint on μ as the *hyperplane condition*; we note that it is weaker than absolute continuity with respect to Lebesgue measure as it allows μ to have positive mass on some types of Lebesgue-measure-zero sets. For example, when $d = 2$ we allow a measure which is positive on a curve no part of which is a line-segment.

In the hyperplane-conditioned case our results are as follows. For vertices, $\mathbb{E}N_n = n$ if $n \leq d + 1$, whilst

$$\mathbb{E}N_n = \frac{n}{2} + \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j}, \quad (n-d) \geq 3 \text{ and odd.} \quad (1)$$

For facets, $\mathbb{E}F_n = 0$ if $n < d$, $\mathbb{E}F_d \equiv \mathbb{E}F_{d,d-1} = 1$, $\mathbb{E}F_{d+1} \equiv \mathbb{E}F_{d+1,d-1} = d + 1$ and

$$\mathbb{E}F_n = \frac{1}{2} \binom{n}{d} + \frac{1}{2} \sum_{j=1}^{n-d} (-1)^{j-1} \binom{n}{j} \mathbb{E}F_{n-j}, \quad (n-d) \geq 3 \text{ and odd.} \quad (2)$$

Employing the Dehn–Sommerville identities ([10], [18] and [12], p145) in combination with (1) and (2), we find a recurrence for $(d-2)$ -faces, and a complete list of recurrences for i -faces when $i < d \leq 5$.

All of our new results have a common form, namely, for $0 \leq i < d$,

$$\begin{aligned} \mathbb{E}F_{n,i} &= \frac{1}{2} \binom{n}{i+1} + \frac{1}{2} \sum_{j=1}^{n-i-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}F_{n-j, i} & (n-d) \geq 3 \text{ and odd} \\ &= \binom{n}{i+1} & n \leq d+1. \end{aligned} \quad (3)$$

This suggests to us that (3) may have general validity in the case where μ satisfies the *hyperplane condition*, even though we only prove it here for all d when i equals 0, $d - 2$ or $d - 1$ (and additionally for $i = 1$ when $d = 4, 5$ and $i = 2$ when $d = 5$).

The paper concludes with a brief discussion of the difficulties that arise when μ does not satisfy the hyperplane condition. The discussion suggests that this condition is the most appropriate for the current study.

Some detail which has been omitted from the published version is available on the author's website (<http://www.maths.usyd.edu.au/u/richardc>).

The identities of Efron, Buchta and Cowan

Assuming that μ is an absolutely continuous measure (the words ‘with respect to Lebesgue measure’ will be implicitly assumed for the rest of this paper), Efron [11] showed that $\mathbb{E}\mu_n = 1 - \mathbb{E}N_{n+1}/(n+1)$. In the important special case where μ is uniformly distributed on a bounded convex $K \subset \mathbb{R}^d$, being zero outside K , then μ_n equals V_n divided by $|K|$, the volume of K . Efron's result is better known in this restricted context, where it becomes $\mathbb{E}V_n/|K| = 1 - \mathbb{E}N_{n+1}/(n+1)$.

Developing Efron's idea, Buchta [5] showed that, provided $\mu = 0$ for all single points in \mathbb{R}^d ,

$$\mathbb{E}\mu_n^k = \mathbb{E} \prod_{i=1}^k \left(1 - \frac{N_{n+i}}{n+i}\right), \quad n \geq 1. \quad (4)$$

We call this the Efron/Buchta identity. It is a powerful formula linking moments of μ_n to moments of the number of vertices of convex hulls formed by random points. We note that Buchta's contribution weakened the condition on μ to one which is implied by our *hyperplane condition*.

Cowan [8] (in his Theorem 2 and Corollary 1) has recently shown that, for any absolutely continuous measure ν defined on $(\mathbb{R}^d, \mathcal{B}_d)$,

$$\nu_n := \nu(H_n) = \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \bar{\nu}_{n-j}^{(n)}, \quad n \geq 2, \quad (n-d) \text{ even}, \quad (5)$$

where $\bar{\nu}_j^{(n)}$ equals the average ν -measure of all $\binom{n}{j}$ j -hulls of P_1, P_2, \dots, P_n . Note that this identity does not involve expectations; it is true for all realisations P_1, P_2, \dots, P_n whose convex hull has dimension d . We show in Appendix A that (5) remains true for any measure ν satisfying the *hyperplane condition* (which is weaker than *absolute continuity*).

In particular, it holds when $\nu = \mu$ if μ satisfies the condition – and also, as noted in [8], if ν equals the absolutely-continuous volume measure V . So we can write the following mean-

value identities involving $\mu_n := \mu(\mathbf{H}_n)$ and V_n , by applying expectations and using the obvious identities $\mathbb{E}\bar{\mu}_{n-j}^{(n)} = \mathbb{E}\mu_{n-j}$ and $\mathbb{E}\bar{V}_{n-j}^{(n)} = \mathbb{E}V_{n-j}$.

$$\mathbb{E}\mu_n = \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}\mu_{n-j}; \quad (6)$$

$$\mathbb{E}V_n = \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}V_{n-j}. \quad (7)$$

Identity (6) is new, while (7) has been known since 1990 (see [4]).

The summation ranges in (5)–(7) can be written with fewer terms, using only the range $j \leq n - d - 1$, because the other terms are zero. We resist this shortening, however, for the convenience of formulae derived in the sequel. To obtain the first of these formulae, we now combine the new identity (6) with the Efron/Buchta identity (4) when $k = 1$.

A recurrence involving $\mathbb{E}N_n$ for general d , when μ satisfies the hyperplane condition

The hyperplane condition on μ implies that both (4) and the $(\nu = \mu)$ -version of (5) hold. It also tells us immediately that $\mu_j = 0$ for $j \leq d$ and $\mathbb{E}N_n = n$ for $n \leq d + 1$.

We can state from (6), and the equation (4) with $k = 1$, that

$$\begin{aligned} 1 - \frac{\mathbb{E}N_{n+1}}{n+1} &= \mathbb{E}\mu_n = \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}\mu_{n-j} \\ &= \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \left(1 - \frac{\mathbb{E}N_{n-j+1}}{n-j+1}\right), \end{aligned}$$

when $(n - d)$ is even and $n \geq 2$. Rearranging, we have

$$\begin{aligned} 1 - \frac{\mathbb{E}N_{n+1}}{n+1} &= \frac{1}{2} \left(1 + (-1)^n - (1 - 1)^n - \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \frac{\mathbb{E}N_{n-j+1}}{n-j+1}\right). \\ \therefore \mathbb{E}N_{n+1} &= \frac{n+1}{2} \left(1 - (-1)^n + \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \frac{\mathbb{E}N_{n-j+1}}{n-j+1}\right) \end{aligned}$$

when $(n - d)$ is even and $n \geq 2$. This is perhaps written better as

$$\mathbb{E}N_n = \frac{n}{2} \left(1 - (-1)^{n-1} + \sum_{j=1}^{n-2} (-1)^{j-1} \binom{n-1}{j} \frac{\mathbb{E}N_{n-j}}{n-j}\right) \quad (8)$$

when $(n - d)$ is odd and $n \geq 3$. The form (1) follows easily from (8); when d is odd (implying n is even), (8) becomes

$$\mathbb{E}N_n = n + \frac{1}{2} \sum_{j=1}^{n-2} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j} = n - \frac{n}{2} \mathbb{E}N_1 + \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j},$$

whilst when d is even (implying n is odd), (8) becomes

$$\mathbb{E}N_n = \frac{1}{2} \sum_{j=1}^{n-2} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j} = \frac{n}{2} \mathbb{E}N_1 + \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j}.$$

Thus (1) is valid, since $N_1 = 1$. Note that equation (1) behaves appropriately when $d = 1$, where it is known from elementary considerations that $\mathbb{E}N_1 = 1$ and $\mathbb{E}N_j = 2$ for $j \geq 2$.

Example 1: Let the points be placed randomly inside a convex quadrilateral in \mathbb{R}^2 according to the uniform distribution. The quadrilateral has vertices labelled ABCD and its shape is characterised by two numbers a and b . These may be chosen in many ways, but we use the style in Cowan and Chiu [9]: the diagonal AC is divided by the other diagonal BD into two segments of ratio $a : 1$, with BD in turn being divided in the ratio $b : 1$. These authors show that $\mathbb{E}N_5 = \frac{25}{6} + \frac{10}{9}\theta$ and $\mathbb{E}N_4 = \frac{11}{3} + \frac{4}{9}\theta$, where $\theta := ab/(1 + a^2)(1 + b^2)$. By applying (1), we can check these values. When $n = 5$, $n - d = 3$, an odd number; so

$$\begin{aligned} \mathbb{E}N_5 &= \frac{5}{2} + \frac{1}{2} \left[\binom{5}{1} \mathbb{E}N_4 - \binom{5}{2} \mathbb{E}N_3 + \binom{5}{3} \mathbb{E}N_2 - \binom{5}{4} \mathbb{E}N_1 \right] \\ &= \frac{5}{2} + \frac{1}{2} \left[5 \left(\frac{11}{3} + \frac{4}{9}\theta \right) - 10 \times 3 + 10 \times 2 - 5 \times 1 \right] = \frac{25}{6} + \frac{10}{9}\theta. \end{aligned}$$

$\mathbb{E}N_6$ is not known at this stage, but if it is found, then we shall obtain $\mathbb{E}N_7$ immediately from (1). This illustrates the general principle when $d = 2$: find the result for the even cases, perhaps with some difficulty, and get the odd cases for free.

Example 2: Let the points be uniformly distributed inside the unit ball in \mathbb{R}^3 . Kingman [13] showed, when $n = 5$, that $\mathbb{E}N_5 = 4 + \frac{134}{143}$ and we know trivially that $\mathbb{E}N_j = j$ for $j \leq 4$. So, we can compute $\mathbb{E}N_6$.

$$\begin{aligned} \mathbb{E}N_6 &= 3 + \frac{1}{2} \left[\binom{6}{1} \mathbb{E}N_5 - \binom{6}{2} \mathbb{E}N_4 + \binom{6}{3} \mathbb{E}N_3 - \binom{6}{4} \mathbb{E}N_2 + \binom{6}{5} \mathbb{E}N_1 \right] \\ &= 3 + \frac{1}{2} \left[6 \left(4 + \frac{134}{143} \right) - 15 \times 4 + 20 \times 3 - 15 \times 2 + 6 \times 1 \right] = 5 + \frac{116}{143}. \end{aligned}$$

This agrees with a result of Miles [14] calculated by other methods. Using our new identity (1), we get this even- n answer for free!

Remark 1: For the case where μ is assumed to have the uniform distribution within a bounded domain K , μ is proportional to V (on sets $\subset K$). In this case, Buchta and Reitzner [6] have combined the identity (7) – or more precisely an algebraically equivalent version due to Affentranger and Badertscher ([1], [2])– with the Efron identity. Our theory in this section has much greater generality.

Some ideas from Rényi and Sulanke

We now find a recurrence relationship for the expected facet-number, F_n , when μ satisfies the *hyperplane condition*. Initially let $d = 2$ and $n > 2$ (whereby $N_n = F_n$). We commence with an approach introduced in the ‘uniform in convex K ’ case by Rényi and Sulanke ([15],[16]). Following these authors but with modification to suit our purpose, we argue that the line segment P_1P_2 is a *side* (i.e. a 1-face) of the polygon H_n if and only if all $n - 2$ other points lie *either* in the half-plane to the left of the line through P_1P_2 *or* in the half-plane to the right. Line direction is defined arbitrarily as going from P_1 to P_2 . The probability of this is, for $n > 2$,

$$\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} (m^{n-2} + (1-m)^{n-2}) \mu(dP_1) \mu(dP_2),$$

where m is defined as the μ -measure of the left half-plane. Thus, by considering all pairs of points,

$$\mathbb{E}(N_n) = \mathbb{E}(F_n) = \binom{n}{2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} (m^{n-2} + (1-m)^{n-2}) \mu(dP_1) \mu(dP_2), \quad n > 2.$$

Apart from our use of a more general μ , our argument differs from Rényi and Sulanke (and some other researchers who have exploited their idea, for example Affentranger [1] and Buchta [3]) in another important way. They defined m as the *smaller* μ -measure of the left and right half-planes. This was an essential choice for the asymptotic agenda of Rényi and Sulanke, but our choice has another use. Because any positioning of P_1 and P_2 is as likely as the same configuration with subscripts interchanged, one can write, for $n > 2$,

$$\mathbb{E}(N_n) = \mathbb{E}(F_n) = 2 \binom{n}{2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} m^{n-2} \mu(dP_1) \mu(dP_2) = 2 \binom{n}{2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} (1-m)^{n-2} \mu(dP_1) \mu(dP_2).$$

Such an identity, and the double representations based on symmetric half-spaces, extend to d -dimensions in an obvious way. For $n > d$, the extended form is:

$$\mathbb{E}(F_n) = 2 \binom{n}{d} \int \dots \int m^{n-d} \mu(dP_1) \dots \mu(dP_d) = 2 \binom{n}{d} \int \dots \int (1-m)^{n-d} \mu(dP_1) \dots \mu(dP_d). \quad (9)$$

Remark 2: We note that this double representation has also been used by Buchta and Reitzner [7] in their study of random points inside a 3-dimensional tetrahedron. Our analysis, shown in the next section, and that of these two authors, take different pathways from this common form, equation (9).

General- d recurrence formulae for $\mathbb{E}(F_n)$ and $\mathbb{E}(F_{n,d-2})$

We exploit the double representation in (9), with an analysis that leads to a family of new identities. For $n > d$,

$$\begin{aligned}
\mathbb{E}(F_n) &= 2 \binom{n}{d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} (1-m)^{n-d} \mu(dP_1) \mu(dP_2) \cdots \mu(dP_d) \\
&= 2 \binom{n}{d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \sum_{j=0}^{n-d} \binom{n-d}{j} (-1)^j m^j \mu(dP_1) \mu(dP_2) \cdots \mu(dP_d) \\
&= 2 \binom{n}{d} \sum_{j=0}^{n-d} \binom{n-d}{j} (-1)^j \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} m^j \mu(dP_1) \mu(dP_2) \cdots \mu(dP_d) \\
&= \binom{n}{d} \sum_{j=0}^{n-d} \frac{\binom{n-d}{j}}{\binom{d+j}{d}} (-1)^j \left[2 \binom{d+j}{d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} m^j \mu(dP_1) \mu(dP_2) \cdots \mu(dP_d) \right] \\
&= 2 \binom{n}{d} + \sum_{j=1}^{n-d} (-1)^j \binom{n}{d+j} \mathbb{E}(F_{d+j}). \tag{10}
\end{aligned}$$

From (10), $\mathbb{E}F_{d+1} = 2(d+1) - \mathbb{E}F_{d+1}$; so $\mathbb{E}F_{d+1} = d+1$. When $n \geq d+2$, (10) progresses (by substituting $m := d+j$) to

$$\mathbb{E}(F_n) = (-1)^{n-d} \mathbb{E}(F_n) + 2 \binom{n}{d} + \sum_{m=d+1}^{n-1} (-1)^{m-d} \binom{n}{m} \mathbb{E}(F_m).$$

Therefore if $n-d$ is odd (and ≥ 3) this yields results agreeing with (2).

$$\begin{aligned}
\mathbb{E}(F_n) &= \binom{n}{d} + \frac{1}{2} \sum_{m=d+1}^{n-1} (-1)^{m-d} \binom{n}{m} \mathbb{E}(F_m) \\
&= \binom{n}{d} + \frac{1}{2} \sum_{j=1}^{n-d-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}F_{n-j} \quad \text{where } j := n-m \\
&= \frac{1}{2} \binom{n}{d} + \frac{1}{2} \sum_{j=1}^{n-d} (-1)^{j-1} \binom{n}{j} \mathbb{E}F_{n-j} \quad \text{because } \mathbb{E}F_d = 1.
\end{aligned} \tag{11}$$

Since H_n is a simplicial d -polytope when $n > d$, we can relate $F_{n,d-2}$ to $F_{n,d-1}$ (which we know as F_n). Each facet of H_n is a $(d-1)$ -simplex, with each of its d $(d-2)$ -faces pasted onto

a $(d - 2)$ -face of another facet. This means that $dF_{n,d-1} = 2F_{n,d-2}$. Thus $F_{n,d-2} = \frac{1}{2}d F_n$, for $n > d$. We note that this simple counting identity is one of the Dehn-Sommerville equations applied to the dual polytope of H_n .

So from (11), we have the following recurrence for $(n - d) \geq 3$ and odd.

$$\begin{aligned}\mathbb{E}F_{n,d-2} &= \frac{d}{2} \mathbb{E}F_n = \frac{d}{2} \binom{n}{d} + \frac{d}{4} \sum_{j=1}^{n-d-1} (-1)^{j-1} \binom{n}{j} \frac{2}{d} \mathbb{E}F_{n-j, d-2} \\ &= \frac{1}{2} \binom{n}{d-1} + \frac{1}{2} \sum_{j=1}^{n-d+1} (-1)^{j-1} \binom{n}{j} \mathbb{E}F_{n-j, d-2},\end{aligned}$$

using $\mathbb{E}F_{d-1,d-2} = 1$ and $\mathbb{E}F_{d,d-2} = d$.

It is easy to show that the results of this section, derived by the methods which exploit the Rényi/Sulanke idea, are consistent with those for $\mathbb{E}N_n$ derived by use of the Cowan and Efron/Buchta identities. When $d = 2$, one uses the identity $N_n = F_n$, $n > 2$. Likewise when $d = 3$, one employs the identities for $n > 3$: $F_{n,1} = \frac{3}{2}F_n$; $N_n = 2 + F_{n,1} - F_n = 2 + \frac{1}{2}F_n$.

When $d = 2, 3$: agreement with the earlier results for $\mathbb{E}N_n$

In this section, we show that these results derived by the methods which exploit the Rényi/Sulanke idea are consistent with those for $\mathbb{E}N_n$ derived by use of the Cowan and Efron/Buchta identities, when $d = 2, 3$.

For $d = 2$, $N_n = F_n$ when $n \geq 3$; so, when n is odd and ≥ 5 , a reduction formula for $\mathbb{E}N_n$ is provided by (11).

$$\begin{aligned}\mathbb{E}N_n &= \binom{n}{2} + \frac{1}{2} \sum_{j=1}^{n-3} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j} \\ &= \frac{1}{2} \binom{n}{1} + \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \mathbb{E}N_{n-j}, \quad \text{because } \mathbb{E}N_2 = 2 \text{ and } \mathbb{E}N_1 = 1.\end{aligned}$$

This latter form, which agrees with our earlier result (1) derived for all d , is also valid for $n = 3$.

If $d = 3$, we see that $F_{n,1} = \frac{3}{2}F_n$, $n \geq 4$. From Euler's identity, $N_n = 2 + F_{n,1} - F_n = 2 + \frac{1}{2}F_n$

when $n \geq 4$. Applying this to (2), gives, for $n \geq 6$ and even:

$$\begin{aligned}
\mathbb{E}N_n &= 2 + \frac{1}{2}\mathbb{E}F_n = 2 + \frac{1}{2}\left[\frac{1}{2}\binom{n}{3} + \frac{1}{2}\sum_{j=1}^{n-3}(-1)^{j-1}\binom{n}{j}\mathbb{E}F_{n-j}\right] \\
&= 2 + \frac{1}{2}\left[\frac{1}{2}\binom{n}{3} + \frac{1}{2}\binom{n}{3}\mathbb{E}F_3 + \frac{1}{2}\sum_{j=1}^{n-4}(-1)^{j-1}\binom{n}{j}2(\mathbb{E}N_{n-j} - 2)\right] \\
&= 2 + \frac{1}{2}\binom{n}{3} + \frac{1}{2}\sum_{j=1}^{n-4}(-1)^{j-1}\binom{n}{j}\mathbb{E}N_{n-j} - \sum_{j=1}^{n-4}(-1)^{j-1}\binom{n}{j} \quad (\text{using } \mathbb{E}F_3 = 1) \\
&= 2 + \frac{1}{2}\binom{n}{3} + \frac{1}{2}\sum_{j=1}^{n-1}(-1)^{j-1}\binom{n}{j}\mathbb{E}N_{n-j} - \binom{n}{1}\frac{\mathbb{E}N_1}{2} + \binom{n}{2}\frac{\mathbb{E}N_2}{2} - \binom{n}{3}\frac{\mathbb{E}N_3}{2} \\
&\quad - \left[\sum_{j=0}^n(-1)^{j-1}\binom{n}{j} + \binom{n}{0} + \binom{n}{n} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3}\right] \\
&= \frac{1}{2}\binom{n}{1} + \frac{1}{2}\sum_{j=1}^{n-1}(-1)^{j-1}\binom{n}{j}\mathbb{E}N_{n-j} \quad \text{because } \mathbb{E}N_3 = 3, \mathbb{E}N_2 = 2 \text{ and } \mathbb{E}N_1 = 1.
\end{aligned}$$

This result, which is also valid for $n = 4$, agrees with (1).

Other cases when $d = 4, 5$

In this section, we deal with some new cases where $d = 4, 5$ and show that the general form given in (3) holds. Firstly let $d = 4$. The Euler identity states that $F_{n,1} = F_{n,0} + F_{n,2} - F_{n,3}$ when $n \geq 5$. This becomes $F_{n,1} = F_{n,0} + F_{n,3} = N_n + F_n$, using the *simplicial* identity, $F_{n,2} = 2F_{n,3}$. So, using (1) and (2), we can write for $n \geq 7$ and odd:

$$\begin{aligned}
\mathbb{E}F_{n,1} &= \mathbb{E}N_n + \mathbb{E}F_n = \frac{1}{2}\left[\binom{n}{1} + \binom{n}{4} + \sum_{j=1}^{n-1}(-1)^{j-1}\binom{n}{j}\mathbb{E}N_{n-j} + \sum_{j=1}^{n-4}(-1)^{j-1}\binom{n}{j}\mathbb{E}F_{n-j}\right] \\
&= \frac{1}{2}\left[\binom{n}{1} + \binom{n}{4} + \sum_{j=1}^{n-5}(-1)^{j-1}\binom{n}{j}\mathbb{E}F_{n-j,1} + \binom{n}{4}\mathbb{E}F_4 \right. \\
&\quad \left. + \binom{n}{4}\mathbb{E}N_4 - \binom{n}{3}\mathbb{E}N_3 + \binom{n}{2}\mathbb{E}N_2 - \binom{n}{1}\mathbb{E}N_1\right] \\
&= \frac{1}{2}\left[2\binom{n}{2} - 3\binom{n}{3} + 6\binom{n}{4} + \sum_{j=1}^{n-2}(-1)^{j-1}\binom{n}{j}\mathbb{E}F_{n-j,1} \right. \\
&\quad \left. - \binom{n}{4}\mathbb{E}F_{4,1} + \binom{n}{3}\mathbb{E}F_{3,1} - \binom{n}{2}\mathbb{E}F_{2,1}\right] \\
&= \frac{1}{2}\binom{n}{2} + \frac{1}{2}\sum_{j=1}^{n-2}(-1)^{j-1}\binom{n}{j}\mathbb{E}F_{n-j,1},
\end{aligned}$$

using $\mathbb{E}F_{4,1} = 6$, $\mathbb{E}F_{3,1} = 3$ and $\mathbb{E}F_{2,1} = 1$. This formula, which also holds for $n = 5$, conforms to (3).

The Euler and ‘simplicial’ identities in the paragraph above can also be viewed as the Dehn-Sommerville system of equations. These equations apply to simple convex polytopes, but can also be used for the dual of any simple polytope – which will be a simplicial polytope. In this theory, our polytopes are simplicial, so Dehn-Sommerville equations can be applied.

When $d = 5$, there are three Dehn-Sommerville equations and these are sufficient to allow expression of $F_{n,1}$, $F_{n,2}$ and $F_{n,3}$ in terms of N_n and F_n . For $n \geq 6$,

$$F_{n,1} = 3N_n + \frac{1}{2}F_n - 6;$$

$$F_{n,2} = 2N_n + 2F_n - 4;$$

$$F_{n,3} = \frac{5}{2}F_n.$$

It can now be shown, using the method we have employed for $d = 4$, that recurrence relationships conforming to (3) hold for $\mathbb{E}F_{n,1}$ and $\mathbb{E}F_{n,2}$.

For $d > 5$, the Dehn-Sommerville equations do not provide enough information for a complete set of recurrence relationships. But we might hope, given the agreement that has held in so many cases, that (3) still holds for all i when $d > 5$. We leave this as an interesting conjecture, or at least, a plausible speculation.

When μ does not satisfy the hyperplane condition

In this section, we report that a theory similar to that developed above is not possible if we weaken the hyperplane condition.

Firstly, the Efron/Buchta identity does not hold if the weakened form of the condition allows single-point sets to have positive probability. This identity can, however, be modified, to the form (when $k = 1$) of $\mathbb{E}\mu_n = 1 - \mathbb{E}N_{n+1}^{(1)}/(n+1)$. Here $N_n^{(1)}$ equals the number of vertices containing only one of the n random points (higher-order vertices being possible). (See the Appendix B for notes on this generalised Efron/Buchta equation.) So, in principle the weakening of the hyperplane condition might still allow a recurrence relationship in $\mathbb{E}N_n^{(1)}$ to be developed.

Unfortunately, however, we cannot exploit Cowan’s recurrence in ν_n (or μ_n) in these weak-

ened circumstances. The reason, explained in Appendix A, is the potential for H_n to have dimension less than d — and in particular for the dimension (which becomes a random variable when μ is not *hyperplane conditioned*) to take both odd and even values.

Relaxing the hyperplane condition also renders the Dehn-Sommerville equations inapplicable, as H_n is not longer a simplicial polytope. Finally, we note that our arguments based on ideas of Rényi and Sulanke are also invalid if μ has positive values on hyperplanes.

So, for many reasons, we have been unable to extend our theory to the case of more general μ . It seems that the hyperplane condition provides the most natural and best framework for our theory.

Second moments: potential for research

We conclude our paper with one further comment. Because the recurrence relationship (5) is for *all* point configurations, we can find the variance of μ_n by simply taking the variance of the right-hand side of (5). Thus

$$\begin{aligned}\text{Var}(\mu_n) &= \frac{1}{4} \text{Var} \left(\sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \bar{\mu}_{n-j}^{(n)} \right) \\ &= \frac{1}{4} \sum_{j=1}^{n-1} \binom{n}{j}^2 \text{Var}(\bar{\mu}_{n-j}^{(n)}) + \frac{1}{2} \sum_{j=1}^{n-2} \sum_{i=j+1}^{n-1} (-1)^{i+j} \binom{n}{j} \binom{n}{i} \text{Cov}(\bar{\mu}_{n-j}^{(n)}, \bar{\mu}_{n-i}^{(n)}),\end{aligned}$$

when $(n-d)$ is even and $n \geq 2$. This can then be combined with the ($k=2$)-version of the Efron-Buchta identity:

$$\text{Var}(\mu_n) = \mathbb{E} \prod_{i=1}^2 \left(1 - \frac{N_{n+2}}{n+i} \right) - \left(1 - \frac{\mathbb{E} N_{n+1}}{n+1} \right)^2.$$

We see that covariance issues severely complicate this line of argument. In principle, these complications can be resolved and further research (beyond the scope of this paper) is in progress.

APPENDIX A: Discussion of the identity (5)

Cowan's Theorem 2 [8], stated below in slightly variant form, is a topological result for arbitrarily placed points.

Cowan's Topological Theorem: $P_1, P_2, \dots, P_n, n \geq d + 1$, are points in \mathbb{R}^d whose convex hull has dimension d . For any reference point $P \in \mathbf{H}_n$, define $c_j(P)$ as the number of sub-collections of j points taken from $\{P_1, P_2, \dots, P_n\}$ whose convex hull contains P . Then,

$$\begin{aligned} \Psi(P) := c_1(P) - c_2(P) + \dots + (-1)^{n-1}c_n(P) &= (-1)^d && \text{for almost all } P \in \overset{\circ}{\mathbf{H}}_n, \\ &= 0 && P \in \partial\mathbf{H}_n, \end{aligned} \quad (12)$$

where $\partial\mathbf{H}_n$ and $\overset{\circ}{\mathbf{H}}_n$ denote the boundary and interior respectively of \mathbf{H}_n .

Remark 3: Cowan conjectures that the words ‘almost all’ in his Topological Theorem can be replaced by ‘all’. In [8], he proves this conjecture when $d \leq 2$.

Remark 4: He also notes ([8], Remark 4) that the set of positions for $P \in \overset{\circ}{\mathbf{H}}_n$ not encompassed within his proof of the Topological Theorem is of dimension $(d - 2)$. Denote this set of *exceptional* positions for P by \mathcal{E} . We make the important observation here, based on the proofs and remarks in [8], that \mathcal{E} is a union of subsets of hyperplanes – each hyperplane being of dimension $\leq d - 2$.

Remark 5: If the n points placed in \mathbb{R}^d have convex hull of dimension h less than d , we reduce d in the Theorem so that it equals h (and identify the hyperplane containing \mathbf{H}_n with \mathbb{R}^h). We call this a *reduced-dimension version* of the Theorem.

Proof of (5) when ν satisfies the hyperplane condition. In [8], this result was proved when ν was any absolutely continuous measure on $(\mathbb{R}^d, \mathcal{B}_d)$ – or on $(\mathbb{R}^h, \mathcal{B}_h)$ if a reduced-dimension version is required – but we now prove it under the weaker hyperplane condition. The argument, similar to that in [8] (but not marred by the author’s typographical omission of $\binom{n}{j}$ in formulae (5) and (7) of that paper and a substantive inaccuracy within his Corollary 1) runs as follows.

For $P \in \mathbb{R}^d$ and $\mathbf{H} \subset \mathbb{R}^d$ let $I_{\mathbf{H}}(\cdot)$ be defined as the indicator function of the domain \mathbf{H} , namely $I_{\mathbf{H}}(P) = 1$ if $P \in \mathbf{H}$, being zero otherwise. Clearly,

$$\int_{\mathbb{R}^d} I_{\mathbf{H}}(P) \nu(dP) = \nu(\mathbf{H}).$$

The entity c_j in Cowan’s Topological Theorem is the sum of indicator functions of the j -subset convex hulls. So

$$\int_{\mathbb{R}^d} c_j(P) \nu(dP) = \binom{n}{j} \bar{\nu}_j^{(n)}.$$

Therefore, from integration of (12), we have the following. With $n \geq d + 1$,

$$\int_{\mathbb{R}^d} \Psi(P) \nu(dP) = \int_{\mathring{\mathbf{H}}_n} (-1)^d \nu(dP). \quad (13)$$

This is true because $\nu(\mathcal{E}) = 0$ (see Remark 4 above). Therefore

$$\binom{n}{1} \bar{\nu}_1^{(n)} - \binom{n}{2} \bar{\nu}_2^{(n)} + \dots + (-1)^{n-1} \binom{n}{n} \bar{\nu}_n^{(n)} = (-1)^d \nu(\mathring{\mathbf{H}}_n).$$

We note that $\bar{\nu}_n^{(n)} = \nu(\mathbf{H}_n)$ and, because $\partial\mathbf{H}_n$ comprises facets which lie in hyperplanes, $\nu(\partial\mathbf{H}_n) = 0$. So $\nu(\mathring{\mathbf{H}}_n) = \nu(\mathbf{H}_n)$ and

$$\begin{aligned} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \bar{\nu}_j^{(n)} &= [(-1)^d - (-1)^{n-1}] \nu(\mathbf{H}_n) \\ &= 2(-1)^n \nu(\mathbf{H}_n), & \text{when } (n-d) \text{ is even} & \quad (14) \end{aligned}$$

$$= 0 \quad \text{when } (n-d) \text{ is odd.} \quad (15)$$

Division throughout (14) by $(-1)^n$ and rearrangement of the summation yields (5).

Consider now the random experiment with $n \geq d + 1$. If realisations yield an \mathbf{H}_n of dimension d with probability 1, as is the case when μ satisfies the hyperplane condition, then (14) holds for almost all outcomes P_1, P_2, \dots, P_n . Thus (5) holds. In particular it holds when $\nu = \mu$, if μ is *hyperplane conditioned*. When $n < d + 1$, (5) holds trivially since both sides equal zero when ν is *hyperplane conditioned*.

For more general ν . Contrary to an inaccurate statement in the author's previous paper ([8], Corollary 1), the equation (13) above will only be true if $\nu(\mathcal{E}) = 0$, and when $d > 2$ this may not be the case for general ν . This qualification was overlooked in [8]. If the conjecture in Remark 3 is resolved positively, of course, then $\mathcal{E} = \emptyset$ and $\nu(\mathcal{E}) = 0$; so (13) will hold for all ν , including those which do not satisfy the hyperplane condition.

If $\nu(\mathcal{E}) = 0$, the arguments following (13) progress (as in [8]) to yield a generalisation of (14) and (15), namely:

$$\sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \bar{\nu}_{n-j}^{(n)} = 2\nu(\mathring{\mathbf{H}}_n) + \nu(\partial\mathbf{H}_n) = 2\nu(\mathbf{H}_n) - \nu(\partial\mathbf{H}_n) \quad \text{when } (n-d) \text{ is even} \quad (16)$$

$$= \nu(\partial\mathbf{H}_n) \quad \text{when } (n-d) \text{ is odd.} \quad (17)$$

Equation (16) implies the following generalisation of (5):

$$\nu_n := \nu(\mathbf{H}_n) = \frac{1}{2} \sum_{j=1}^{n-1} (-1)^{j-1} \binom{n}{j} \bar{\nu}_{n-j}^{(n)} + \frac{1}{2} \nu(\partial\mathbf{H}_n), \quad n \geq 2, (n-d) \text{ even,} \quad (18)$$

An example illustrates the situation.

Example 3: Consider the random experiment where $d = 2$, n is a fixed even number and μ has as support the sides of a given convex polygon K . Clearly μ does not satisfy the hyperplane condition, but, because $d = 2$, $\mathcal{E} = \emptyset$. Thus $\nu(\mathcal{E}) = 0$ and the $(\nu = \mu)$ -version of (18) holds for all realisations P_1, P_2, \dots, P_n whose convex hull H_n has dimension 2 (where we note that $n - d$ is even). It is easy to calculate $\mu(\partial H_n)$ as the μ -measure of those sides of H_n which lie within sides of K .

The reader will note that, for some realisations in this example, H_n will have dimension 1; all points will by chance lie on one side of K . This does not stop us using the Topological Theorem; we can still use it in a reduced-dimension version with $h := \dim(H_n) = 1$ (see Remark 5). But $n - h$ is odd in these cases, so (18) is not usable – and the opposite-parity equation (17) is not helpful as it contains no information about $\nu(H_n)$. This means that no overall formula can be obtained for μ_n , nor for $\mathbb{E}\mu_n$.

We note that similar difficulties, resulting from mixed odd-even parity issues, arise in all cases where μ does not satisfy the hyperplane condition.

APPENDIX B: The Efron-Buchta identity for general μ and small k .

It is not difficult, in principle, to modify Buchta's argument (in [5]) to cover the case where μ may give positive weight to single points. Expressions would then be written in terms of the random variables $N_n^{(r)}$, defined as the number of vertices of H_n which have multiplicity r .

When $k = 1$, the modification is a trivial exercise, yielding

$$\mathbb{E}\mu_n = 1 - \frac{\mathbb{E}N_{n+1}^{(1)}}{n+1},$$

and the $k = 2, 3$ cases are also relatively easy:

$$\begin{aligned} \mathbb{E}\mu_n^2 &= \mathbb{E} \prod_{i=1}^2 \left(1 - \frac{N_{n+2}^{(1)}}{n+i}\right) - \frac{2\mathbb{E}N_{n+2}^{(2)}}{(n+2)(n+1)}; \\ \mathbb{E}\mu_n^3 &= \mathbb{E} \prod_{i=1}^3 \left(1 - \frac{N_{n+3}^{(1)}}{n+i}\right) - \frac{6(\mathbb{E}N_{n+3}^{(3)} + \mathbb{E}(N_{n+3}^{(2)}(n+1 - N_{n+3}^{(1)})))}{(n+3)(n+2)(n+1)}. \end{aligned}$$

Expressions for $k > 3$ are more complicated and efforts to date at finding a general form have not been successful.

References

- [1] F. Affentranger, Generalization of a formula of C. Buchta about the convex hull of random points. *Elem. Math.* **43**, (1988), 39–45, 151–152.
- [2] E. Badertscher, An explicit formula about the convex hull of random points. *Elem. Math.* **44**, (1989) 104–106.
- [3] C. Buchta, Über die konvexe Hülle von Zufallspunkten in Eibereichen. *Elem. Math.* **38**, (1983), 153–156.
- [4] C. Buchta, Distribution-independent properties of the convex hull of random points. *J. Theor. Probab.* **3**, (1990), 387–393.
- [5] C. Buchta, An identity relating moments of functions of convex hulls. *Discr. Comput. Geom.* **33**, (2005), 125–142.
- [6] C. Buchta and M. Reitzner, Equiaffine inner parallel curves of a plane convex body and the convex hulls of randomly chosen points. *Probab. Theory Relat. Fields* **108**, (1997), 385–415.
- [7] C. Buchta and M. Reitzner, The convex hull of random points in a tetrahedron: Solution of Blaschke’s problem and more general results. *J. reine angew. Math.* **536**, (2001), 1–29.
- [8] R. Cowan, Identities linking volumes of convex hulls. *Adv. Appl. Prob.* **39**, (2007), 630–644.
- [9] R. Cowan and S. N. Chiu, Extension of Deltheil’s study on random points in a convex quadrilateral. *Adv. Appl. Prob.* **37**, (2005), 857–858.
- [10] M. Dehn, Die Eulersche Formel in Zusammenhang mit dem Inhalt in der nicht-Euklidischen Geometrie, *Math. Ann.* **61**, (1905), 561–586.
- [11] B. Efron, The convex hull of a random set of points. *Biometrika* **52**, (1965), 331–343.
- [12] B. Grünbaum, *Convex Polytopes*. (2nd Edition, prepared by Kaibel, V., Klee, V. and Ziegler, G. M.), Springer, Berlin. 2003.
- [13] J. F. C. Kingman, Random secants of a convex body. *J. Appl. Prob.* **6**, (1969), 660–672.
- [14] R. E. Miles, Isotropic random simplices. *Adv. Appl. Prob.* **3**, (1971), 353–382.
- [15] A. Rényi and R. Sulanke, Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahr. Verw. Geb.* **2**, (1963), 75–84.

- [16] A. Rényi and R. Sulanke, Über die konvexe Hülle von n zufällig gewählten Punkten. II *Z. Wahr. Verw. Geb.* **3**, (1963), 138–147.
- [17] R. Schneider and W. Weil, *Stochastic and Integral Geometry*. Springer, Berlin. 2008.
- [18] D. Sommerville, The relations connecting the angle sums and volume of a polytope in space of n dimensions, *Proc. Roy. Soc. London Ser. A* **115**, (1927), 103-119