INTRINSIC VOLUMES OF RANDOM CUBICAL COMPLEXES

MICHAEL WERMAN AND MATTHEW L. WRIGHT

ABSTRACT. Intrinsic volumes, which generalize both Euler characteristic and Lebesgue volume, are important properties of d-dimensional sets. A random cubical complex is a union of unit cubes, each with vertices on a regular cubic lattice, constructed according to some probability model. We analyze and give exact polynomial formulae, dependent on a probability, for the expected value and variance of the intrinsic volumes of several models of random cubical complexes. We then prove a central limit theorem for these intrinsic volumes. For our primary model, we also prove an interleaving theorem for the zeros of the expected-value polynomials. The intrinsic volumes of cubical complexes are useful for understanding the shape of random d-dimensional sets and for characterizing noise in applications.

1. INTRODUCTION

In the literature about random combinatorial objects, a number of recent papers analyze topological properties of various random structures. In particular, Kahle studied the Betti numbers of simplicial complexes obtained from random points in $\mathbb{R}^d$ [9, 10], and Kahle and Meckes gave limit theorems for such Betti numbers [11]. Linial and Meshulam studied homology groups of random 2-dimensional complexes [14], while Meshulam and Wallach generalized to higher-dimensional complexes [15]. Cohen et. al. obtained results about collapsibility and topological embeddings of the Linial-Meshulam complexes [4].

Among these papers on random complexes, topological properties (e.g. Euler characteristic, homology, Betti numbers) are prevalent, but the intrinsic volumes are scarcely to be found. The intrinsic volumes[1] generalize both Euler characteristic and Lebesgue volume, providing information about not only the topology but also the geometry of sets. In contrast to the Betti numbers, the intrinsic volumes are additive (local), which makes them easier to compute in some cases. The intrinsic volumes are useful in applications such as imaging and stereology, and can be computed quickly for low-dimensional binary images in pixel (or voxel) form [6, 8, 13, 18, 21]. Furthermore, the intrinsic volumes arise in Adler and Taylor's work on random fields [1] (see also [23]), and various applications of the Euler characteristic to image classification can be found in [16].

This paper analyzes the intrinsic volumes of several models of random cubical complexes. A random cubical complex is a union of unit cubes with vertices on a regular d-dimensional cubic lattice. These complexes are relevant for digital images and discretized domains in which data is represented on a cubical lattice of possibly high dimension. This work is motivated in part by the desire to understand noise in digital images. The construction of a random cubical complex depends on the particular model and a probability. We give polynomial formulae, dependent on probability p, for the expected values and variances of the intrinsic volumes of these complexes. We analyze these polynomials and, for our primary model, prove an interleaving theorem about their

[1]Intrinsic volumes known by various names, including *Hadwiger measures*, *Minkowski functionals*, and *quermass-integrale*. These concepts are equivalent up to normalization.

zeros. Moreover, we prove that as the size of the complexes tend toward infinity, the distribution of a particular intrinsic volume approaches a normal distribution.

We first provide background about the intrinsic volumes in Section 2. Section 3 describes rigorously our models of random cubical complexes. We then analyze our primary model, the *voxel model*, giving expected values and variances in Sections 4 and 5, respectively. Section 6 contains a central limit theorem for the intrinsic volumes. In Section 7, we turn to other models of random cubical complexes.

2. Background: Intrinsic Volumes

The intrinsic volumes are *valuations*.[2] A **valuation** is a real-valued function on a class of sets that provides a notion of *size* for each set and satisfies the additive property. The **additive property** says that if v is a valuation, then $v(X) + v(Y) = v(X \cap Y) + v(X \cup Y)$ for sets on which v is defined. The **intrinsic volumes** are valuations, related to Euler characteristic and Lebesgue volume, normalized to be *intrinsic* to sets and independent of the ambient dimension in which sets may be embedded. We give a brief introduction to the intrinsic volumes here; the interested reader may pursue the references for more details.

Defined on various classes of sets in Euclidean space $\mathbb{R}^d$, the intrinsic volumes are denoted $\mu_0, \mu_1, \ldots, \mu_d$. Intuitively, μ_k assigns a real number to each set that provides a notion of the k-dimensional size of the set [20]. The 0-dimensional valuation μ_0 is the Euler characteristic, the only topological invariant among the intrinsic volumes. The intrinsic volume μ_1 gives a notion of the *length* of a set; μ_2 gives a notion of *area*, etc. For example, if X is a compact, convex 3-dimensional set, then $\mu_2(X)$ is one-half the surface area of X. For a d-dimensional set, μ_d gives d-dimensional Lebesgue volume. The intrinsic volumes are invariant with respect to rigid motions (translations and rotations) of sets. Furthermore, the classic Hadwiger Theorem says that the intrinsic volumes form a basis for the vector space of all rigid-motion invariant valuations that are continuous on convex sets with respect to the Hausdorff metric.

The intrinsic volumes are commonly defined on the class of compact convex sets in $\mathbb{R}^d$, but it is possible to be much more general. In their text, Klain and Rota approach intrinsic volumes via valuations on a lattice of subsets of $\mathbb{R}^d$ [12]. Another common approach defines the intrinsic volumes on compact convex sets via the Steiner tube formula [19]. Hadwiger's formula can be used to define intrinsic volumes for all sets in an o-minimal structure, including sets that are neither closed nor convex [3, 22]. We do not need the full generality of these approaches in this paper, since it suffices for our present purposes to define the intrinsic volumes for finite unions of cubes. The intrinsic volumes of a closed unit cube, given in Definition 1, agree with the intrinsic volumes as defined in any of the above sources.

Definition 1. A **closed i-cube** shall refer to any translate of the unit cube $[0,1]^i$, for any non-negative integer i. If X is a closed i-cube, then the **intrinsic volume** μ_k of X is $\mu_k(X) = \binom{i}{k}$. Note that if $k > i$, then $\mu_k(X) = 0$.

When the intersection $X \cap Y$ of two closed cubes X and Y (of possibly different dimension) is also a closed cube, then the intrinsic volume of the union $\mu_k(X \cup Y)$ follows from Definition 1 and the additive property. Thus, intrinsic volumes are well-defined for a class of (non-convex) sets obtained from finite unions of closed cubes.

Additivity also induces intrinsic volumes on the interiors of cubes. An **open i-cube** is any translate of $(0,1)^i$, for any positive integer i. We may refer to a point as *either* an open or a closed 0-cube. The following proposition gives intrinsic volumes of open cubes, which we introduce as an

---

[2]Valuations are not *measures*, though the concepts are similar. In particular, the intrinsic volumes may take on negative values and are not monotonic.

auxiliary notion that will simplify many later calculations. Furthermore, this proposition gives a special case of the intrinsic volumes on definable sets found in [3].

Proposition 2. *If $X = (0,1)^i$ is an open i-cube, then the intrinsic volume μ_k of X is*

$$\mu_k(X) = (-1)^{i-k} \binom{i}{k}. \tag{1}$$

Proof. Let X be an open i-cube. Then X has $\binom{i}{j}2^{i-j}$ faces of dimension j, for $j = 0, 1, \ldots, i$. Inclusion-exclusion (i.e. the additive property) lets us write $\mu_k(X)$ as an alternating sum over all closed faces of X. Since μ_k is zero for any closed face of dimension less than k, we have:

$$\mu_k(X) = \sum_{j=k}^{i} (-1)^{i-j} \binom{i}{j} 2^{i-j} \mu_k\left([0,1]^j\right) = \sum_{j=k}^{i} (-2)^{i-j} \binom{i}{j}\binom{j}{k}.$$

We then use the binomial coefficient identity $\binom{i}{j}\binom{j}{k} = \binom{i}{k}\binom{i-k}{i-j}$ to obtain

$$\mu_k(X) = \binom{i}{k} \sum_{j=k}^{i} \binom{i-k}{i-j}(-2)^{i-j} = \binom{i}{k} \sum_{t=0}^{i-k} \binom{i-k}{t}(-2)^t = \binom{i}{k}(-1)^{i-k}$$

as desired. $\qquad\square$

The intrinsic volumes of open cubes will appear frequently in sequel, so we assign them special notation. Let $\mu_{k,i}$ refer to the intrinsic volume μ_k of an open unit i-cube. That is,

$$\mu_{k,i} = (-1)^{i-k} \binom{i}{k}. \tag{2}$$

Proposition 2 implies that the Euler characteristic of an open i-cube X is $\mu_0(X) = (-1)^i$. We note that μ_0 is sometimes called the *combinatorial* Euler characteristic[3] and is found in the literature [3, 5, 22].

The cubical complexes that we study in this paper may be decomposed into disjoint unions of open cubes. This perspective simplifies calculations because additivity of the intrinsic volumes is especially easy to work with over *disjoint* unions. Throughout the paper, we will consider a cubical complex to be a disjoint union of open cubes, even if the complex itself is topologically closed. For a closed cubical complex, we could obtain the same results by considering a non-disjoint union of closed cubes, but this would require more bookkeeping via the inclusion-exclusion principle. Our approach shifts this bookkeeping into the definition of intrinsic volumes of open cubes, while also permitting us to work with cubical complexes that are not necessarily closed (which we do in Section 7).

3. Random Cubical Complexes

We construct random geometric objects that are subsets of a regular array of d-dimensional unit cubes. Let $n \in \{2, 3, 4, \ldots, \}$, and consider a regular lattice of d-dimensional unit cubes, with n cubes in each direction, for a total of n^d unit d-cubes. Let $\mathcal{L}$ be this lattice of unit cubes. For simplicity, so as not to bother with the border of $\mathcal{L}$, we suppose that opposite $(d-1)$-dimensional faces of $\mathcal{L}$ are identified, so that $\mathcal{L}$ has the topology of a d-dimensional torus. The number of i-dimensional faces of any d-cube is $\binom{d}{i}2^{d-i}$. Furthermore, any i-cube is a face of 2^{d-i} d-cubes.

---

[3]The combinatorial Euler characteristic is not a homotopy invariant, but on compact sets it agrees with the *topological* Euler characteristic (which has value 1 on any contractible set).

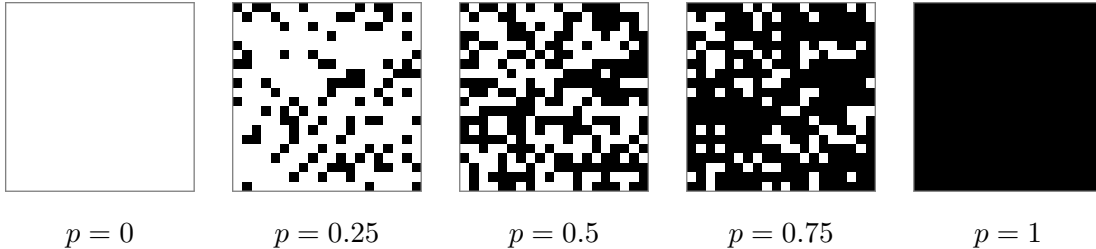|				
$p = 0$	$p = 0.25$	$p = 0.5$	$p = 0.75$	$p = 1$

FIGURE 1. Example of the two-dimensional *voxel model* of random cubical complexes ($d = 2$, $n = 20$). The black set is the cubical complex, which is regarded as a subset of the torus.

Thus, the number of i-cubes in $\mathcal{L}$, which we denote N_i, is the number of d-cubes, times the number of i-cubes per d-cube, divided by the number of d-cubes of which each i-cube is a face:

$$(3) \qquad N_i = \frac{n^d \binom{d}{i} 2^{d-i}}{2^{d-i}} = \binom{d}{i} n^d.$$

We randomly select a subset C of $\mathcal{L}$, such that C is a union of open unit cubes of various dimensions. Such a subset C we shall call a **random cubical complex**. In general, there is no requirement that C itself be open or closed. We are interested in understanding the intrinsic volumes $\mu_k(C)$, for all $k = 0, 1, \ldots, d$. Of course, the intrinsic volumes $\mu_k(C)$ depend on how we select the unit cubes that comprise C. Each selection method yields interesting connections among the $\mu_k(C)$. We now describe the three models of cubical complexes that we will examine.

Our primary selection method we call the **voxel model**. In this model, we select the closed unit d-cubes of $\mathcal{L}$ independently with equal probability. Specifically, the unit d-cubes of $\mathcal{L}$ are included in C independently with probability p, and a unit i-cube of $\mathcal{L}$, for $i < d$, is included in C exactly when it is a face of an included d-cube. In this case, C is a closed set, consisting of a union of closed d-cubes that possibly overlap on some of their faces. Figure 1 illustrates the voxel model in two dimensions for several values of p. We find that the expected value and variance of $\mu_k(C)$ are given by polynomials in $q = 1 - p$ of degree 2^{d-k}. We show various identities between these polynomials and prove an interleaving property of the roots of the expected-value polynomials.

The voxel model is the most natural for various applications, but we also consider other possible models. A related model is the **plaquette model** which includes all cubes of dimension less than d, while including the d-cubes independently with probability p. This model is related to the plaquette model of percolation [2, 7], and is analogous to the models of random simplicial complexes studied by Linial and Meshulam [14].

Additionally, the following two models allow for randomness to occur in all dimensions. The **closed faces model** allows lower-dimensional closed cubes to be included. Specifically, this model selects each open i-cube independently with probability p, and then forms a cubical complex consisting of the closure of all selected cubes. The **independent faces model** includes each open i-cube independently with probability depending on i, and the resulting cubical complex might not be closed. For an interesting model, we choose to include open i cubes with probability p^i. These models may be useful for applications in which features of different dimensions occur simultaneously.

We emphasize that regarding our random cubical complexes as subsets of the d-torus allows us to ignore boundary effects. Furthermore, as $n \to \infty$, the distribution of intrinsic volumes of random cubical complexes living in $\mathbb{R}^d$ converges to that of complexes (of the same model) living in the d-torus. That is, boundary effects become negligible as $n \to \infty$. Thus, since we are primarily interested in behavior as $n \to \infty$, the torus setting is natural and simplifies computation.
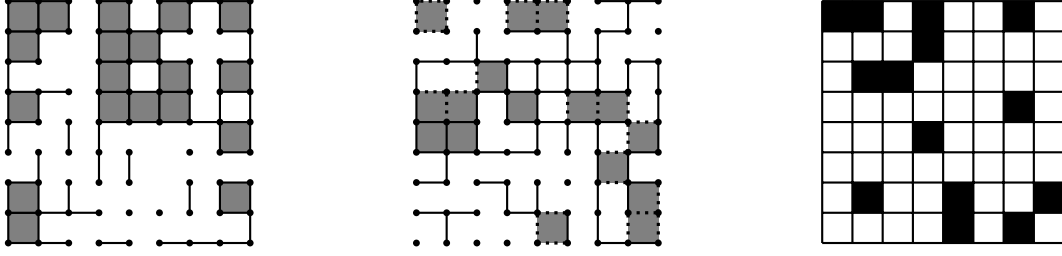
FIGURE 2. The *closed faces model* (left) selects each open i-cube independently with probability p, and then forms a cubical complex consisting of the closure of all selected cubes. In the illustration, $d = 2$, $n = 8$, and $p = 0.25$. The *independent faces model* (center) includes each open i-cube independently with probability p^i. The diagram shows $d = 2$, $n = 8$, and $p = 0.5$. The *plaquette model* (right) includes top-dimensional cubes with probability p and includes all cubes of lower dimension; here $d = 2$, $n = 8$, and $p = 0.25$.

4. VOXEL MODEL: EXPECTED VALUE

We now compute the expected value of $\mu_k(C)$ for the voxel model. Recall that this model selects d-cubes from $\mathcal{L}$ independently with probability p, and that C is the union of the closure of all selected cubes. Since $\mathcal{L}$ is a disjoint union of open cubes, it suffices to sum the expected value of $\mu_k(X)$ over all open cubes X (of all dimensions) in $\mathcal{L}$. For any open i-cube X, let $\xi_{X,k}$ be the random variable that indicates the contribution of X to $\mu_k(C)$. That is, $\xi_{X,k} = \mu_{k,i} = (-1)^{i-k}\binom{i}{k}$ if X is included in C, and $\xi_{X,k} = 0$ otherwise. When the subscript k is clear, we will write ξ_X instead of $\xi_{X,k}$. Writing the expected value as a sum over all open cubes, indexed by dimension, we obtain

$$(4) \qquad \mathbb{E}(\mu_k(C)) = \sum_{\text{open } X \in \mathcal{L}} \mathbb{E}(\xi_X) = \sum_{i=k}^{d} N_i P_i \mu_{k,i},$$

where P_i is the probability that any particular open i-cube is included in C.

Since any i-cube is a face of 2^{d-i} d-cubes, the probability that any particular i-cube is *not* included in C is $(1-p)^{2^{d-i}}$. The probability that any particular i-cube *is* included in C is then $1 - (1-p)^{2^{d-i}}$. Let $q = 1 - p$, and we have

$$P_i = 1 - q^{2^{d-i}}.$$

Recalling the count N_i of i-cubes in $\mathcal{L}$ from equation (3) and the value of $\mu_{k,i}$ from equation (2), we obtain

$$(5) \qquad \mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d} \binom{d}{i} n^d \left(1 - q^{2^{d-i}}\right)(-1)^{i-k}\binom{i}{k}.$$

In equation (5), n only appears as a factor n^d, the (d-dimensional) volume of $\mathcal{L}$. We normalize by volume to remove this factor. The expected value of $\mu_k(C)$ per unit volume is a polynomial in q, which we denote $E_{d,k}(q)$. That is,

$$(6) \qquad E_{d,k}(q) = \frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d}(-1)^{i-k}\binom{d}{i}\binom{i}{k}\left(1 - q^{2^{d-i}}\right).$$

The following theorem expresses our computation in a simpler form.

5

Theorem 3. *The expected value of $\mu_k(C)$, normalized by volume, is a polynomial in q of degree 2^{d-k}, given by the following formula*

(7)
$$E_{d,k}(q) = \frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \begin{cases} \sum_{i=k}^{d}(-1)^{i-k+1}\binom{d}{i}\binom{i}{k}q^{2^{d-i}} & \text{if } k < d \\ 1 - q = p & \text{if } k = d. \end{cases}$$

Proof. Use the following identity to simplify equation (6):

(8)
$$\sum_{i=k}^{d}(-1)^{i-k}\binom{d}{i}\binom{i}{k} = \begin{cases} 0 & \text{if } k \in \{0, 1, \ldots, d-1\} \\ 1 & \text{if } k = d. \end{cases}$$

$\square$

The expected value polynomials $E_{d,k}(q)$ appear in Table 1 for small d and k. We seem to have a family of polynomials indexed by two parameters, d and k. However, in the next section we show that, up to constant multiples, we have a family of polynomials indexed by the single parameter $d - k$, and the roots of these polynomials interleave.

	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$
$k = 0$	$1 - q$	$q - q^2$	$-q + 2q^2 - q^4$	$q - 3q^2 + 3q^4 - q^8$	$-q + 4q^2 - 6q^4 + 4q^8 - q^{16}$
$k = 1$		$1 - q$	$2(q - q^2)$	$3(-q + 2q^2 - q^4)$	$4(q - 3q^2 + 3q^4 - q^8)$
$k = 2$			$1 - q$	$3(q - q^2)$	$6(-q + 2q^2 - q^4)$
$k = 3$				$1 - q$	$4(q - q^2)$
$k = 4$					$1 - q$

TABLE 1. Expected value polynomials $E_{d,k}(q)$, written in factored form to illustrate Proposition 4.

4.1. **Expected Value Polynomials.** The polynomial $E_{d,k}(q)$ is a constant multiple of the polynomial $E_{d-k,0}(q)$. Thus, if we understand the sequence of polynomials $E_{d,0}$ that give the expected Euler characteristics for $d = 1, 2, \ldots$, then we understand the expected intrinsic volumes $E_{d,k}$ for any d and k. We can regard the quantity $d - k$ as the *codimension* of the valuation μ_k on a d-dimensional complex, and the roots of the polynomial $E_{d,k}(q)$ depend only on this codimension. Specifically, we have the following proposition.

Proposition 4. *The expected value polynomials satisfy*

(9)
$$E_{d,k}(q) = \binom{d}{k}E_{d-k,0}(q).$$

Proof. Use the binomial coefficient identity $\binom{d}{i}\binom{i}{k} = \binom{d}{k}\binom{d-k}{d-i}$ to rewrite equation (6) as

$$E_{d,k}(q) = \binom{d}{k}\sum_{i=k}^{d}(-1)^{i-k}\binom{d-k}{d-i}\left(1 - q^{2^{d-i}}\right).$$

Now let $j = i - k$ to re-index the sum, obtaining

$$E_{d,k}(q) = \binom{d}{k}\sum_{j=0}^{d-k}(-1)^j\binom{d-k}{j}\left(1 - q^{2^{d-k-j}}\right) = \binom{d}{k}E_{d-k,0}(q)$$
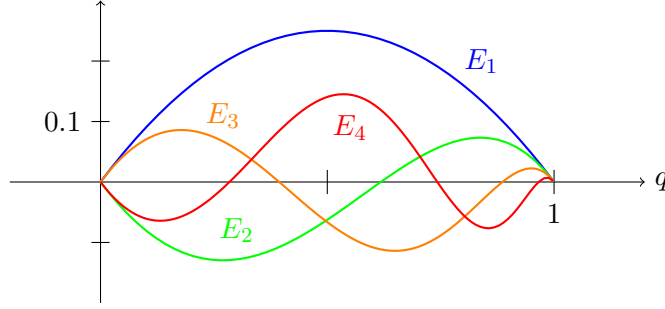
as desired.

$\square$

FIGURE 3. Graphs of the polynomials $E_d(q)$ that give the expected Euler characteristic for the voxel model random cubical complexes, where d is the dimension of the complex.

To simplify notation, let $E_d(q) = E_{d,0}(q)$. These polynomials, which give the expected Euler characteristic per unit volume, are particularly nice. The graphs of the first four of these polynomials are shown in Figure 3.

One may ask why these polynomials (other than E_1) are not symmetric about $q = 1/2$. To answer this question, let C_p be a d-dimensional cubical complex that includes unit d-cubes with probability p, and similarly for C_{1-p}. Let C_p^c denote the complement of C_p (in the d-torus), and observe that $\mu_0(C_p) = -\mu_0(C_p^c)$ by additivity. While the expected number of unit d-cubes in the complex C_{1-p} is the same as that in C_p^c, the complex C_{1-p} is a closed set, but C_p^c is open. Two diagonally-adjacent closed cubes form a connected set, while two diagonally-adjacent open cubes do not. Thus, we do not expect $\mu_0(C_p)$ to equal $\mu_0(C_{1-p})$, even in absolute value. For example, if $d = 2$ and $p = q = 1/2$, the complex C_p is likely to have more holes than connected components, since diagonally-adjacent included squares are connected, but diagonally-adjacent holes are not.

We now examine the polynomial family E_d, with the goal of understanding their roots on the interval $[0, 1]$. For $d > 0$, we observe from equation (7) that

$$(10) \qquad E_d(q) = \sum_{i=0}^{d} (-1)^{i+1} \binom{d}{i} q^{2^{d-i}}.$$

Thus, the expected Euler characteristic polynomials have a form similar to $(1 - q)^d$ and satisfy the following recurrence.

Lemma 5. *The expected Euler characteristic polynomials satisfy the recurrence*

$$(11) \qquad E_d(q) = E_{d-1}(q^2) - E_{d-1}(q).$$

Proof. Apply the identity $\binom{d}{i} = \binom{d-1}{i-1} + \binom{d-1}{i}$ to equation (10) and simplify. $\qquad \square$

Remark 6. The polynomial $E_2(q)$ has a root at $q = \frac{\sqrt{5}-1}{2}$, which is the reciprocal of the golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$. In other words, if $q = \frac{1}{\varphi}$ is the probability of *excluding* each pixel from the 2-dimensional voxel model, then the expected Euler characteristic of the random cubical complex is zero.

4.2. **Interleaving of Roots.** In Figure 3, we observe that the roots of E_{d+1} interleave the roots of E_d in the open interval $(0, 1)$. We now prove that this interleaving of roots occurs for all d.

Theorem 7. *The polynomial $E_d(q)$ has $d + 1$ roots in the closed interval $[0, 1]$, including roots at the endpoints $q = 0$ and $q = 1$. Moreover, the roots of E_d in the open interval $(0, 1)$ interleave the roots of E_{d+1} in that interval.*

7

Proof. Note that $E_d(q)$ is a polynomial with $d+1$ nonzero coefficients that alternate in sign when arranged by degree. Thus, Descartes' Rule of Signs says that E_d has at most d positive roots. Furthermore, $E_d(q)$ certainly has a root at $q = 0$. Denote the roots of E_d as follows: let $q_{d,0} = 0$, and let $0 < q_{d,1} < q_{d,2}, \dots$ denote the positive roots of $E_d(q)$.

We will prove by induction that E_d has exactly $d+1$ roots in the interval $[0,1]$, including the roots $q_{d,0} = 0$ and $q_{d,d} = 1$, and that they satisfy

$$\sqrt{q_{d,i}} < q_{d,i+1} \tag{12}$$

for all $i = 0, 1, \dots, d-1$.

Base case: The roots of $E_1(q)$ are $q_{1,0} = 0$ and $q_{1,1} = 1$. Thus, E_1 has exactly two roots in the interval $[0,1]$, and these roots satisfy inequality (12).

Induction: Suppose E_d has exactly $d+1$ roots in the interval $[0,1]$, and these roots satisfy inequality (12).

For any $i \in \{0, 1, \dots, d-1\}$, continuity of E_d implies that E_d does not change sign on the interval $(q_{d,i}, q_{d,i+1})$. Without loss of generality, assume that E_d is positive on this interval. We exhibit $r, s \in (q_{d,i}, q_{d,i+1})$ such that $E_{d+1}(r) < 0$ and $E_{d+1}(s) > 0$.

If $i > 0$, then $q_{d,i} > 0$, so inequality (12) implies $\sqrt{q_{d,i}} \in (q_{d,i}, q_{d,i+1})$. Then $E_d\left(\sqrt{q_{d,i}}\right) > 0$, and by Lemma 5 we have

$$E_{d+1}\left(\sqrt{q_{d,i}}\right) = E_d(q_{d,i}) - E_d\left(\sqrt{q_{d,i}}\right) = -E_d\left(\sqrt{q_{d,i}}\right) < 0.$$

In this case, let $r = \sqrt{q_{d,i}}$. (See Figure 4.) Similarly, for the $i = 0$ case, since E_d is continuous and non-constant there exists $\epsilon \in (0, q_{d,1})$ such that $E_d(\epsilon^2) < E_d(\epsilon)$. Thus, $E_{d+1}(\epsilon) = E_d(\epsilon^2) - E_d(\epsilon) < 0$, so we let $r = \epsilon$.

Likewise, if $i < d-1$, then $q_{d,i+1} < 1$, so $q_{d,i+1}^2 \in (q_{d,i}, q_{d,i+1})$ by inequality (12). Then $E_d(q_{d,i+1}^2) > 0$, which implies

$$E_{d+1}\left(q_{d,i+1}\right) = E_d\left(q_{d,i+1}^2\right) - E_d\left(q_{d,i+1}\right) = E_d\left(q_{d,i+1}^2\right) > 0,$$

so let $s = q_{d,i+1}$. For the $i = d-1$ case, there exists $\delta \in (q_{d,d-1}, 1)$ such that $E_d(\delta) > E_d(\sqrt{\delta})$. This implies $E_{d+1}(\sqrt{\delta}) = E_d(\delta) - E_d(\sqrt{\delta}) > 0$, so $s = \delta$.

Since $E_{d+1}(r) < 0$ and $E_{d+1}(s) > 0$, continuity implies that it must have at least one root in the interval $(r, s) \subset (q_{d,i}, q_{d,i+1})$. There are d such intervals, so E_{d+1} has at least d roots in the interval $(0, 1)$.

Now $E_{d+1}(1) = 0$, so $q = 1$ is also a root of E_{d+1}. Thus, E_{d+1} has at least $d+1$ positive roots. Since it cannot have any more than $d+1$ positive roots, E_{d+1} has *exactly* $d+1$ positive roots. Since $q = 0$ is also a root, E_{d+1} has exactly $d+2$ roots in the interval $[0,1]$, including roots at 0 and 1.
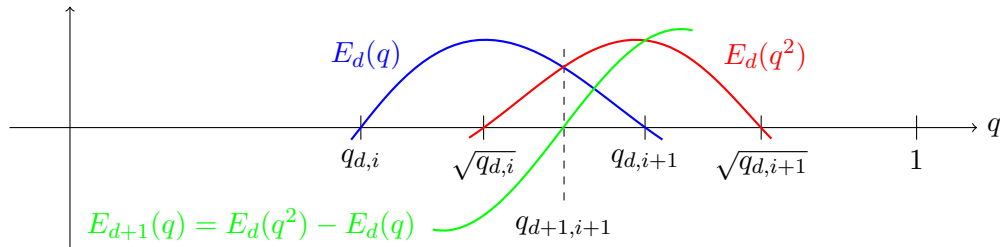

FIGURE 4. Polynomial E_{d+1} has a root between any two roots of E_d.

For $i > 0$ we have $q_{d+1,i} < q_{d,i}$, and so $\sqrt{q_{d+1,i}} < \sqrt{q_{d,i}}$. Also, for $i < d$, we have $\sqrt{q_{d,i}} < q_{d+1,i+1}$. Thus $\sqrt{q_{d+1,i}} < q_{d+1,i+1}$ for any i, which proves that the roots of E_{d+1} in $[0,1]$ satisfy inequality (12).

Lastly, the inequalities in the above paragraph imply that for $i = 1, 2, \ldots, d-1$, we have the inequality $q_{d+1,i} < q_{d,i} < q_{d+1,i+1}$. Thus, the roots of E_d interleave the roots of E_{d+1} on $(0,1)$. $\square$

5. VOXEL MODEL: VARIANCE

We now compute the variance of $\mu_k(C)$ for the voxel model. Let $V_{d,k}(q)$ denote this variance, normalized by volume. This normalized variance is a polynomial in q, as given in the following theorem.

Theorem 8. *For $n > 2$, the variance of $\mu_k(C)$, normalized by volume, is a polynomial in q of degree 2^{d-k}, as follows:*

$$(13) \quad V_{d,k}(q) = \frac{1}{n^d} \mathrm{Var}(\mu_k(C)) = \sum_{i=0}^{d} \sum_{j=0}^{d} \sum_{s=0}^{d} (-1)^{i+j} \binom{i}{k} \binom{j}{k} \binom{d}{s} N_{i,j,s} q^{2^{d-i}+2^{d-j}} \left(q^{-2^{d-s}} - 1 \right),$$

where

$$(14) \qquad\qquad N_{i,j,s} = \sum_{\ell=0}^{s} (-1)^{s-\ell} \binom{s}{\ell} \binom{\ell}{i} \binom{\ell}{j} 2^{s+\ell-i-j}.$$

5.1. Proof of Theorem 8.
As before, we let $\xi_{X,k}$ be the random variable that indicates the contribution of an open i-cube X to $\mu_k(C)$. That is, $\xi_{X,k} = \mu_{k,i} = (-1)^{i-k} \binom{i}{k}$ if X is included in C, and $\xi_{X,k} = 0$ otherwise, and we often omit the subscript k. We recall that $\mu_k(C) = \sum_X \xi_X$, where the sum is over all open cubes X of all dimensions in $\mathcal{L}$. We express the variance of $\mu_k(C)$ in terms of the indicator variables ξ_X and analyze the contributions of different pairs of cubes.

Proof. We can express the variance of $\mu_k(C)$ as follows:

$$\mathrm{Var}(\mu_k(C)) = \mathbb{E}\left((\mu_k(C))^2 \right) - (\mathbb{E}(\mu_k(C)))^2$$

$$= \mathbb{E}\left(\left(\sum_X \xi_X \right)^2 \right) - \left(\mathbb{E}\left(\sum_X \xi_X \right) \right)^2$$

$$= \sum_{X,Y} \mathbb{E}(\xi_X \xi_Y) - \sum_{X,Y} \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y),$$

where $\sum_{X,Y}$ means that the sum is over all pairs of open cubes X and Y (of same or different dimensions, including the cases where X and Y are the same open cube) in $\mathcal{L}$. Thus, we have:

$$(15) \qquad\qquad \mathrm{Var}(\mu_k(C)) = \sum_{X,Y} \left(\mathbb{E}(\xi_X \xi_Y) - \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y) \right).$$

We say that X and Y are **close** if they are faces of some common d-cube in $\mathcal{L}$; otherwise, we say that X and Y are **far**. If X and Y are close, then ξ_X and ξ_Y are dependent. However, if X and Y are far, then ξ_X and ξ_Y are independent, because the set of d-cubes whose selection would cause X to be included in C is disjoint from the set of d-cubes whose selection would cause Y to be included.

Let i be the dimension of open cube X and let j be the dimension of open cube Y. If X and Y are far, then independence implies that

$$\mathbb{E}(\xi_X \xi_Y) = \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y) = P_i \mu_{k,i} P_j \mu_{k,j}.$$

9

Thus, far pairs of cubes make no contribution to the variance as expressed in equation (15), and it suffices to sum over close cubes. That is,

$$\text{Var}(\mu_k(C)) = \sum_{\text{close } X,Y} \left(\mathbb{E}(\xi_X \xi_Y) - \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y) \right). \tag{16}$$

In order to systematically sum over all close pairs of cubes, we observe that if X and Y are close, then there is a unique *smallest* cube S of which X and Y are faces (the assumption that $n > 2$ is essential here). We call S the **common cube** of X and Y, and we let s be the dimension of S. Let $N_{i,j,s}$ be the number of pairs X, Y consisting of an i-cube and a j-cube that are faces of any particular common cube of dimension s. Let $P_{i,j,s}$ be the probability that both X and Y are included in C when X and Y have a common cube of dimension s. We can then rewrite the sum in equation (16) as:

$$\text{Var}(\mu_k(C)) = \sum_{i=0}^{d} \sum_{j=0}^{d} \sum_{s=0}^{d} N_s N_{i,j,s} \left(\mathbb{E}(\xi_X \xi_Y) - \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y) \right)$$

$$= \sum_{i=0}^{d} \sum_{j=0}^{d} \sum_{s=0}^{d} N_s N_{i,j,s} \left(P_{i,j,s} \mu_{k,i} \mu_{k,j} - P_i \mu_{k,i} P_j \mu_{k,j} \right)$$

$$= \sum_{i=0}^{d} \sum_{j=0}^{d} \sum_{s=0}^{d} \mu_{k,i} \mu_{k,j} N_s N_{i,j,s} \left(P_{i,j,s} - P_i P_j \right) \tag{17}$$

(In fact, we don't really need the indices of summation to start at 0; it suffices to start at $i = k$, $j = k$, and $s = \max(i,j)$, but there is no harm in starting each index at 0 because all unnecessary terms contain binomial coefficients which evaluate to zero.) It remains to compute $N_{i,j,s}$ and $P_{i,j,s}$.

We use inclusion-exclusion to compute $N_{i,j,s}$. An s-cube has $\binom{s}{i}2^{s-i}$ faces of dimension i. Thus, the number of pairs X and Y, of dimensions i and j, that are faces of any s-cube S, is $\binom{s}{i}2^{s-i}\binom{s}{j}2^{s-j}$. However, the previous formula counts pairs X and Y that are faces of some common cube smaller than S. Inclusion-exclusion gives the number of pairs of cubes that are faces of a common cube S and *not* faces of any smaller common cube:

$$N_{i,j,s} = \sum_{\ell=0}^{s} (-1)^{s-\ell} \binom{s}{\ell} 2^{s-\ell} \binom{\ell}{i} 2^{\ell-i} \binom{\ell}{j} 2^{\ell-j}$$

$$= \sum_{\ell=0}^{s} (-1)^{s-\ell} \binom{s}{\ell} \binom{\ell}{i} \binom{\ell}{j} 2^{s+\ell-i-j}. \tag{18}$$

Now we compute $P_{i,j,s}$. Suppose that S is the common cube of X and Y. Then X and Y are included in C if either of the following two events occur:

    (a) cube S is included, or
    (b) cube S is not included, but other d-cubes containing X and Y are included.

Informally, let $\mathbb{P}(S)$ denote the probability that cube S is included in C, and let $\mathbb{P}(\neg S)$ be the probability that S is not included in C. Then:

$$P_{i,j,s} = \mathbb{P}(S) + \mathbb{P}(\neg S) \cdot \mathbb{P}(X \mid \neg S) \cdot \mathbb{P}(Y \mid \neg S) \tag{19}$$

Recall that $P_s = 1 - q^{2^{d-s}}$ is the probability of including an s-cube. The probability of including X without S is the probability of excluding S and including another d-cube of which X is a face. There are $2^{d-i} - 2^{d-s}$ such d-cubes; thus

$$\mathbb{P}(X \mid \neg S) = 1 - q^{2^{d-i}-2^{d-s}}.$$

Therefore equation (19) becomes

$$(20) \qquad P_{i,j,s} = \left(1 - q^{2^{d-s}}\right) + q^{2^{d-s}} \left(1 - q^{2^{d-i}-2^{d-s}}\right) \left(1 - q^{2^{d-j}-2^{d-s}}\right).$$

Routine algebra shows that

$$P_{i,j,s} - P_i P_j = q^{2^{d-i}+2^{d-j}} \left(q^{-2^{d-s}} - 1\right),$$

and the proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.2. **Variance Polynomials.** The variance polynomials are more complicated than the expected value polynomials previously discussed. Table 2 lists the variance polynomials $V_{d,k}(q)$ for d and k from 0 to 3. While the expected value polynomials have nonzero coefficients only for terms of the form q^{2^i}, the variance polynomials have many more nonzero coefficients. (Curiously, the coefficient of q^{13} in $V_{3,0}$ is zero.) Furthermore, these coefficients are not unimodal, nor do they alternate in sign in any clear way. Figure 5 displays the graphs of the variance polynomials for the Euler characteristic for low-dimensional complexes, and Figure 6 graphs the variance polynomials for all intrinsic volumes in dimension $d = 3$.

We observe interesting patterns among the variance polynomials in Table 2 when we look along diagonals of the form $V_{d,d-i}$; that is, when we consider intrinsic volumes of codimension i. Most obviously, along the $k = d$ diagonal (codimension 0), we find $V_{d,d}(q) = q - q^2$. This is because the

	$d = 0$	$d = 1$	$d = 2$	$d = 3$
$k = 0$	$q - q^2$	$q - 4q^2 + 6q^3 - 3q^4$	$q - 7q^2 + 20q^3 - 13q^4$ $-24q^5 + 28q^6 + 4q^7 - 9q^8$	$q - 10q^2 + 42q^3 - 30q^4 -$ $120q^5 + 156q^6 + 60q^7 -$ $102q^8 + 64q^9 - 120q^{10} -$ $48q^{11} + 114q^{12} + 12q^{14} +$ $8q^{15} - 27q^{16}$
$k = 1$		$q - q^2$	$4q - 18q^2 + 28q^3 - 14q^4$	$9q - 69q^2 + 192q^3 - 105q^4 -$ $264q^5 + 276q^6 + 60q^7 - 99q^8$
$k = 2$			$q - q^2$	$9q - 42q^2 + 66q^3 - 33q^4$
$k = 3$				$q - q^2$

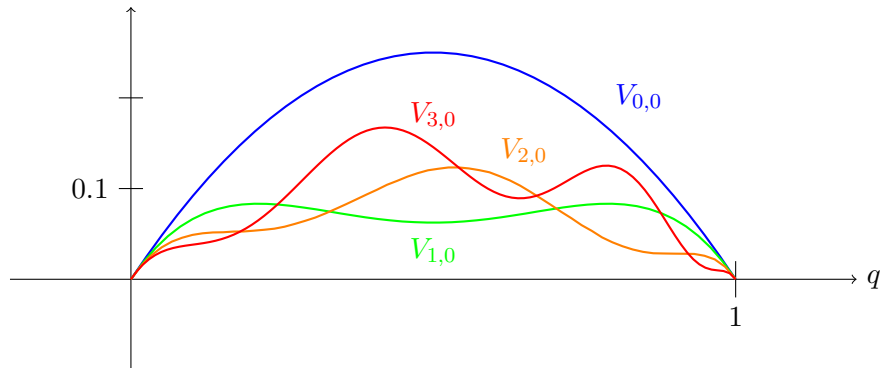TABLE 2. Variance polynomials $V_{d,k}(q)$

FIGURE 5. Graphs of the polynomials $V_{d,0}(q)$ that give the variance of the Euler characteristic for random cubical complexes of dimensions 0 to 3.
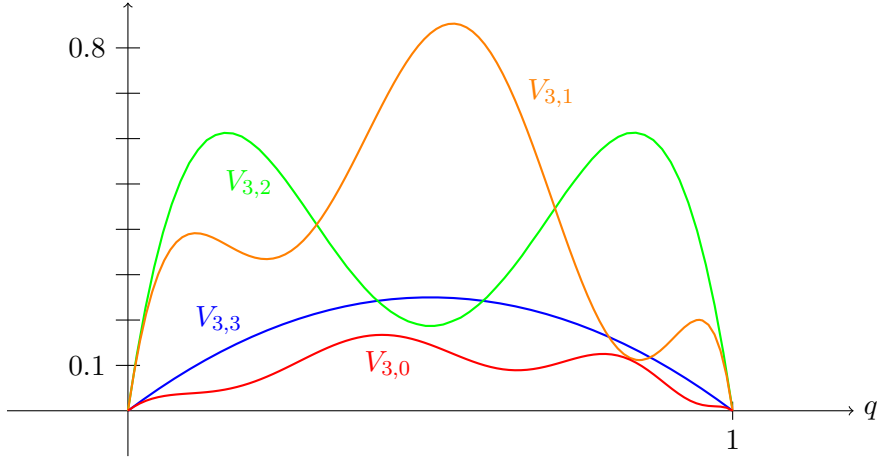
FIGURE 6. Graphs of the polynomials $V_{3,k}(q)$ that give the variance of the intrinsic volumes for random cubical complexes of dimension 3.
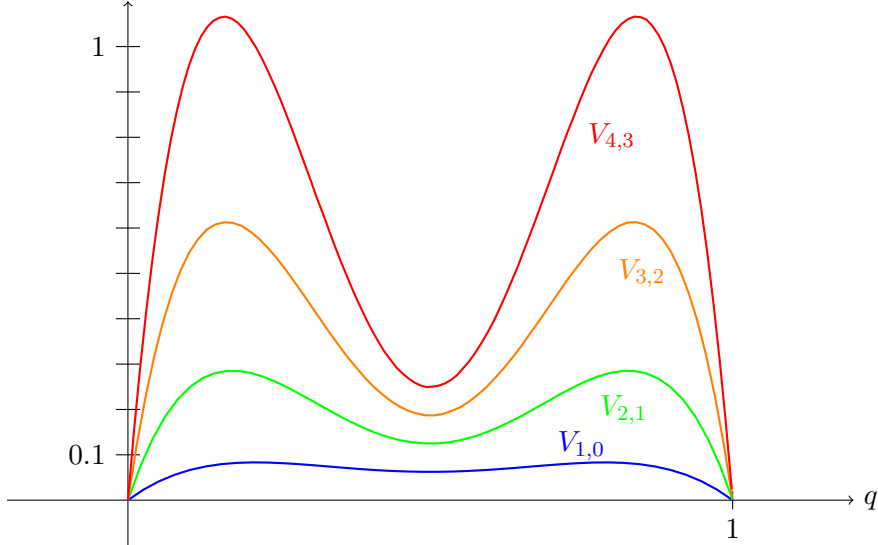

FIGURE 7. Graphs of the polynomials $V_{d,d-1}(q)$ that give the variance of the intrinsic volume μ_{d-1} for random cubical complexes of dimensions $d = 1$ to 4.

intrinsic volume μ_d is actually the *volume* for d-dimensional cubes. Thus, $\mu_d(C)$ depends only on the number of d-cubes selected, and the lower-dimensional faces introduce no dependencies.

Along the codimension-1 diagonal, where $k = d - 1$, we find polynomials of degree 4. Figure 7 illustrates that the graphs of these polynomials all have the same general shape: they are symmetric about $q = 1/2$, with local minima at $q = 1/2$. At first glance, these polynomials may appear to be multiples of each other, but this is not so. In fact, the critical points at which $V_{d,d-1}$ attains local maxima gradually spread outward and approach $\frac{1}{2} \pm \frac{\sqrt{2}}{4}$ as $d \to \infty$. The following proposition gives explicitly $V_{d,d-1}$ for any d.

Proposition 9. *When $k = d - 1$, the variance polynomials satisfy*

$$(21) \qquad V_{d,d-1}(q) = d^2 q - (5d^2 - d)q^2 + (8d^2 - 2d)q^3 - (4d^2 - d)q^4.$$

12

Each such polynomial has a local minimum at $q = 1/2$, and $V_{d,d-1}$ has local maxima at $q = \frac{1}{2} \pm \frac{1}{8d-2}\sqrt{(4d-1)(2d-1)}$.

Proof. Equation (21) can be obtained from simplifying equation (13) when $k = d-1$. The statement about local extrema is then elementary calculus. □

Presumably, similar patterns exist along other diagonals of Table 2. For example, the variance polynomials for $k = d - 2$ also have the same general shape, but are harder to analyze than in the $d - 1$ case. It would also be nice to find a recurrence that connects variance polynomials within a row or column of Table 2.

6. Central Limit Theorem

Having computed the mean and variance of the intrinsic volumes for the voxel model of random cubical complexes, we now prove a central limit theorem. Specifically, we show that for fixed d and q, the distribution of $\mu_k(C)$ tends toward a normal distribution as n tends toward infinity. We rely on an application of Stein's method to sums of random variables with small dependence neighborhoods, detailed in the survey article by Ross [17].

We have seen that $\mu_k(C) = \sum_X \xi_X$ and, in our variance computation, noted that the ξ_X are only locally dependent. That is, ξ_X and ξ_Y are dependent only if cubes X and Y are faces of some common d-cube. Each i-cube is a face of 2^{d-i} d-cubes, and each d-cube has 3^d total faces of all dimensions (including itself). Thus, for any particular ξ_X, the number of ξ_Y that are dependent with ξ_X is not greater than $2^{d-i}3^d \leq 6^d$. Therefore, we say that the size of the dependence neighborhood of ξ_X is bounded by 6^d. We emphasize that this bound is independent of n, which allows us to prove the following theorem.

Theorem 10. *For the d-dimensional voxel model and fixed $q \in (0,1)$, as $n \to \infty$ the random variable $\mu_k(C)$ converges in distribution to a normal random variable with mean $E_{d,k}(q)$ and variance $V_{d,k}(q)$.*

Proof. Let $\sigma^2 = \text{Var}(\mu_k(C)) = n^d V_{d,k}(q)$. Define random variable W as follows:

$$W = \frac{1}{\sigma} \sum_X (\xi_X - \mathbb{E}(\xi_X)) = \frac{1}{\sigma} (\mu_k(C) - \mathbb{E}(\mu_k(C))).$$

Thus W is the standardization (mean zero, unit variance) of the variable $\mu_k(C)$. Let Z be a standard normal random variable.

Since the ξ_X have finite fourth moments and dependency neighborhoods of maximum size 6^d, we can apply Theorem 3.5 in the survey article by Ross [17] to obtain

$$(22) \qquad d_W(W, Z) \leq \frac{6^{2d}}{\sigma^3} \sum_X \mathbb{E}\left|\xi_X - \mathbb{E}(\xi_X)\right|^3 + \frac{\sqrt{26} \cdot 6^{3d/2}}{\sqrt{\pi}\sigma^2} \sqrt{\sum_X \mathbb{E}\left(\xi_X - \mathbb{E}(\xi_X)\right)^4},$$

where $d_W(W, Z)$ is the Wasserstein distance between the distributions of W and Z.

We can also bound the expected value $\mathbb{E}\left|\xi_X - \mathbb{E}(\xi_X)\right|^a$ independently of n. For $a \in \{3, 4\}$, a simple calculation shows that $\mathbb{E}\left|\xi_X - \mathbb{E}(\xi_X)\right|^a \leq 2\binom{d}{k}^a$. The lattice $\mathcal{L}$ contains a total of $(2n)^d$ unit cubes (of all dimensions), so we have:

$$\sum_X \mathbb{E}\left|\xi_X - \mathbb{E}(\xi_X)\right|^a \leq (2n)^d \cdot 2\binom{d}{k}^a \in O(n^d).$$

Since $\sigma^2 \in O(n^d)$, both terms in the right side of inequality (22) are $O(n^{-d/2})$, which means that the right side of the inequality converges to zero as $n \to \infty$. Convergence in Wasserstein metric implies convergence of distribution, so the proof is complete. □

7. Other Models

7.1. Closed Faces.

The closed faces model is similar to the voxel model, but it allows closed faces of dimension less than d to be included, regardless of which d-cubes are included. It is constructed by the following process. Each open i-cube (for all $i = 0, 1, \ldots, d$) is selected independently with probability p. Then C is the *closure* of the union of all selected cubes. (Refer to Figure 2 left.)

To analyze this model, we first determine the probability that any open i-cube X is included in C. Observe that X is a face of 3^{d-i} cubes of dimensions $i, i+1, \ldots, d$ (here we consider X to be a face of itself). If *any* of these 3^{d-i} cubes are selected, then X will be included in C. Since each open cube is selected independently with probability p, the probability that X is included in C is $P_i = 1 - q^{3^{d-i}}$, where as before, $q = 1 - p$.

Proposition 11. *For the closed faces model, the expected value of $\mu_k(C)$, normalized by volume, is*

$$(23) \qquad \frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \begin{cases} \sum_{i=k}^{d}(-1)^{i-k+1}\binom{d}{i}\binom{i}{k}q^{3^{d-i}} & \text{if } k < d \\ 1 - q = p & \text{if } k = d. \end{cases}$$

The normalized variance is similar to that given in Theorem 8 except that $P_i = 1 - q^{3^{d-i}}$ and

$$(24) \qquad P_{i,j,s} = \left(1 - q^{3^{d-s}}\right) + q^{3^{d-s}}\left(1 - q^{3^{d-i}-3^{d-s}}\right)\left(1 - q^{3^{d-j}-3^{d-s}}\right).$$

Proof. The expected value can be written $\mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d} N_i P_i \mu_{k,i}$, as in equation (4). From the previous discussion, the probability that any open i-cube is included in closed faces model is $P_i = 1 - q^{3^{d-i}}$. The quantities N_i and $\mu_{k,i}$ are unchanged, so we have

$$\frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d}(-1)^{i-k}\binom{d}{i}\binom{i}{k}\left(1 - q^{3^{d-i}}\right).$$

We apply identity (8) to obtain the formula in equation (23).

For the variance, equation (13) holds, with probabilities P_i and $P_{i,j,s}$ specific to the closed faces model. We found P_i for this model above. To find $P_{i,j,s}$, suppose i-cube X is a face of s-cube S. There are 3^{d-i} open cubes whose selection causes X to be included in C, and 3^{d-s} of these also cause S to be included. Thus, there are $3^{d-i} - 3^{d-s}$ open cubes whose selection causes X, but not S, to be included. In the notation of equation (19), the probability of including X but not S is

$$\mathbb{P}(X \mid \neg S) = 1 - q^{3^{d-i}-3^{d-s}}.$$

Therefore, $P_{i,j,s}$, given by equation (19), becomes

$$P_{i,j,s} = \left(1 - q^{3^{d-s}}\right) + q^{3^{d-s}}\left(1 - q^{3^{d-i}-3^{d-s}}\right)\left(1 - q^{3^{d-j}-3^{d-s}}\right)$$

as in equation (24). $\qquad \square$

The expected value polynomials for the closed faces model have a very similar form to those for the voxel model, with q raised to powers of 3 instead of powers of 2. Similar to Proposition 4, the expected value of μ_k for the d-dimensional closed faces model is a multiple of the expected value of μ_0 for the $(d-k)$-dimensional closed faces model. Furthermore, the expected value polynomials for the closed faces model satisfy a recurrence like that in Lemma 5 and their roots, which depend only on the codimension $d - k$, interleave just as in the voxel model. Likewise, the variance polynomials for the closed faces model are similar to, but of higher degree than, those for the voxel model. The distribution of $\mu_k(C)$ for the closed faces model also tends toward a normal distribution as $n \to \infty$, yielding a central limit theorem similar to that in Section 6.

The patterns observed in the voxel and closed faces model suggest that there may be a family of models of random complexes in dimension d such that for any positive integer m, there is a model in which the probability of including an open i-cube is $1 - q^{m^{d-i}}$. For the voxel model, $m = 2$, and for the closed faces model, $m = 3$. Specific geometric models corresponding to larger m remain to be found.

7.2. Independent Faces. Suppose that each open i-cube in $\mathcal{L}$ is included in C independently with probability P_i. We emphasize that in this model, if a high-dimensional cube X is included in C, the faces of X are *not* automatically included in C. In other words, C is not necessarily a closed set. (Refer to Figure 2 center.)

In general, we could choose any set of probabilities $P_0, \ldots, P_d$. If we choose $P_i = p$ for all i, then, because there are no dependencies among the faces, we obtain expected value polynomials that are linear in p. A more interesting model results if $P_i = p^i$. This implies that lower-dimensional faces have *greater* probability of being included than do higher-dimensional faces, and *all* vertices (0-cubes) are included in C. We compute the expected value and variance of $\mu_k(C)$ for this model.

Proposition 12. *If each open i-cube is included in C independently with probability p^i, then the expected value and variance of $\mu_k(C)$, normalized by volume, are:*

$$(25) \qquad \frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \binom{d}{k}p^k(1-p)^{d-k}$$

$$(26) \qquad \frac{1}{n^d}\mathrm{Var}(\mu_k(C)) = \sum_{i=k}^{d} \binom{d}{i}\binom{i}{k}^2 \left(p^i - p^{2i}\right).$$

Proof. The expected value can be written $\mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d} N_i P_i \mu_{k,i}$ (see equation (4)), which becomes

$$\frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \sum_{i=k}^{d}(-1)^{i-k}\binom{d}{i}\binom{i}{k}p^i.$$

We apply the identity $\binom{d}{i}\binom{i}{k} = \binom{d}{k}\binom{d-k}{d-i}$ and re-index the sum to obtain equation (25).

Since *all* open cubes are included independently, the variance can be expressed as a sum over all open cubes in $\mathcal{L}$:

$$\mathrm{Var}(\mu_k(C)) = \sum_{\text{open } X} \mathrm{Var}(\xi_X).$$

The variance of ξ_X for any open i-cube X is $\binom{i}{k}^2(p^i - p^{2i})$. Thus,

$$\mathrm{Var}(\mu_k(C)) = \sum_{i=k}^{d}\binom{d}{i}n^d\binom{i}{k}^2(p^i - p^{2i}),$$

and we divide by n^d to obtain equation (26). $\qquad\square$

The intrinsic volumes for the independent faces model satisfy a central limit theorem similar to that of the voxel model in Section 6. However, for the independent faces model, we observe that the expected value polynomial is always of degree d, and the variance polynomial is always of degree $2d$, regardless of k. Thus, the expected value of $\mu_k(C)$ cannot be reduced to a multiple of $\mu_0(C)$ as in the voxel and closed faces models (e.g. equation (9)).

We note that unlike the voxel model, the expected values of the independent faces model are never negative. Furthermore, both the expected value and variance polynomials are unimodal in the independent faces model.

7.3. Plaquette Model. In this model, all $(d-1)$-cubes in $\mathcal{L}$ are included in C, and d-cubes are included independently with probability p. (Refer to Figure 2 right.) The normalized expected values and variances of $\mathbb{E}(\mu_k(C))$ are linear and quadratic, respectively, in p.

Proposition 13. *The expected value and variance of $\mu_k(C)$, normalized by volume, are:*

(27)
$$\frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \begin{cases} (-1)^{d-k}\binom{d}{k}(p-1) & \text{if } k < d \\ p & \text{if } k = d \end{cases}$$

(28)
$$\frac{1}{n^d}\mathrm{Var}(\mu_k(C)) = \binom{d}{k}^2(p-p^2).$$

Proof. As is now familiar, the expected value can be written $\mathbb{E}(\mu_k(C)) = \sum_{i=k}^d N_i P_i \mu_{k,i}$ (equation (4)). However, now $P_i = 1$ if $i < d$ since open cubes of dimension less than d are always included in C, and $P_d = p$. Thus,

$$\frac{1}{n^d}\mathbb{E}(\mu_k(C)) = \left(\sum_{i=k}^{d-1}(-1)^{i-k}\binom{d}{i}\binom{i}{k}\right) + (-1)^{d-k}\binom{d}{k}p$$

We apply identity (8) to obtain equation (27).

The variance can be expressed as a sum over all pairs of cubes, as in equation (15):

$$\mathrm{Var}(\mu_k(C)) = \sum_{X,Y}\left(\mathbb{E}(\xi_X\xi_y) - \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y)\right).$$

In this model, $\mathbb{E}(\xi_X\xi_y) = \mathbb{E}(\xi_X)\mathbb{E}(\xi_Y)$ for all pairs of cubes *except* when X and Y are the *same* d-dimensional cube. If X is a d-cube, then $\mathbb{E}(\xi_X^2) - (\mathbb{E}(\xi_X))^2 = \binom{d}{k}^2(p-p^2)$. Since there are n^d d-cubes, the variance reduces to

$$\mathrm{Var}(\mu_k(C)) = n^d\binom{d}{k}^2(p-p^2),$$

and we divide by n^d to obtain equation (28). $\qquad\square$

We notice that the expected values for the plaquette model satisfy a similar relationship to those from the voxel model. If $E_{d,k}$ is the normalized expected value for $\mu_k(C)$ in the d-dimensional plaquette model, then

$$E_{d,k} = \binom{d}{k}E_{d-k,0}$$

just as in equation (9). However, in the plaquette model, the variance polynomials are all multiples of each other, unlike the variance polynomials in the voxel model. The plaquette model also yields a central limit theorem similar to that in Section 6.

8. Further Work

Besides the intrinsic volumes, it is desirable to find the expected homology of random cubical complexes. However, computing the expected Betti numbers for these complexes seems difficult. We conjecture that, for most values of p and large n, one Betti number dominates all others, and each Betti number gives way to the next as p approaches the zeros of the expected Euler characteristic function. Such is the situation observed in other studies of the topology of random structures [1, 9, 10].

9. Acknowledgements

References

[1] Robert J. Adler and Jonathan E. Taylor, *Random Fields and Geometry*, Springer, 2007.

[2] Michael Aizenman, Chayes, J. T., Chayes, L., Frhlich, J., and Russo, L., "On a Sharp Transition from Area Law to a Perimeter Law in a System of Random Surfaces," *Communications in Mathematical Physics*, vol. 92, March 1983, p. 19–69.

[3] Yuliy Baryshnikov, Robert Ghrist, and Matthew Wright, "Hadwiger's Theorem for Definable Functions," *Advances in Mathematics*, vol. 245, October 2013, p. 573–586.

[4] Daniel Cohen, Armindo Costa, Michael Farber, and Thomas Kappeler, "Topology of Random 2-Complexes," *Discrete and Computational Geometry*, vol. 47, issue 1, January 2012, p. 117–149.

[5] Justin Curry, Robert Ghrist, and Michael Robinson, "Euler Calculus and its Applications to Signals and Sensing," *Proceedings of Symposia in Applied Mathematics*, AMS, 2012.

[6] Stephen B. Gray, "Local properties of binary images in two dimensions," *IEEE Transactions on Computers*, vol. 20, issue 5, May 1971, p. 551–561.

[7] Geoffrey R. Grimmett and Alexander E. Holroyd, "Plaquettes, Spheres, and Entanglement," *Electronic Journal of Probability*, vol. 15 (2010), p. 1415–1428.

[8] Ralph Guderlei, Simone Klenk, Johannes Mayer, Volker Schmidt, and Evgeny Spodarev, "Algorithms for the computation of the Minkowski functionals of deterministic and random polyconvex sets," *Image and Vision Computing*, vol 25, no. 4 (2007), p. 464–474.

[9] Matthew Kahle, "Random geometric complexes", *Discrete and Computational Geometry*, vol. 45, issue 3, April 2011, p. 553–573.

[10] Matthew Kahle, "Topology of random simplicial complexes: a survey," *AMS Contemporary Volumes in Mathematics*, 2014.

[11] Matthew Kahle and Elizabeth Meckes, "Limit theorems for Betti numbers of random simplicial complexes," *Homology, Homotopy, and Applications*, vol. 15, no. 1 (2013), p. 343–374.

[12] Daniel A. Klain and Gian-Carlo Rota, *Introduction to Geometric Probability*, Cambridge University Press, 1997.

[13] Xiaoxing Li, Paulo R.S. Mendonça, Rahul Bhotika, "Texture Analysis Using Minkowski Functionals," *Medical Imaging 2012: Image Processing*, Proceedings of SPIE, vol. 8314.

[14] Nathan Linial and Roy Meshulam, "Homological Connectivity of Random 2-Complexes," *Combinatorica*, vol. 26 (2006), p. 475–487.

[15] Roy Meshulam and N. Wallach, "Homological Connectivity of Random k-Dimensional Complexes," *Random Structures & Algorithms*, vol. 34, no. 3 (2000), p. 408–417.

[16] Eitan Richardson and Michael Werman. "Efficient classification using the Euler characteristic," *Pattern Recognition Letters*, vol. 49 (2014), p. 99–106.

[17] Nathan Ross, "Fundamentals of Stein's Method," *Probability Surveys*, vol. 8, 2011, p. 210–293.

[18] Katja Schladitz, Joachim Ohser, and Werner Nagel, "Measurement of intrinsic volumes of sets observed on lattices," *Discrete Geometry for Computer Imagery*, 2006, p. 247–258.

[19] Rolf Schneider and Wolfgang Weil, *Stochastic and Integral Geometry*, Springer, 2009.

[20] Stephen H. Schanuel, "What is the Length of a Potato?" *Lecture Notes in Mathematics*, Springer, 1986, p. 118–126.

[21] Anne Marie Svane, "Estimation of Intrinsic Volumes from Digital Grey-scale Images," *Journal of Mathematical Imaging and Vision*, vol. 49, no. 2 (2014), p. 352–376.

[22] Lou van den Dries, *Tame Topology and O-Minimal Structures*, Cambridge University Press, 1998.

[23] Matthew L. Wright, "Hadwiger Integration of Random Fields," *Topological Methods in Nonlinear Analysis*, to appear (2014).

The Hebrew University of Jerusalem, Israel
E-mail address: `werman@cs.huji.ac.il`

Institute for Mathematics and its Applications, University of Minnesota, USA
E-mail address: `mlwright@ima.umn.edu`