arXiv:1701.07130v1 [math.AC] 25 Jan 2017

Random Monomial Ideals[1]

Jesús A. De Loera, Sonja Petrović, Lily Silverstein, Despina Stasi, Dane Wilburne

Abstract: Inspired by the study of random graphs and simplicial complexes, and motivated by the need to understand average behavior of ideals, we propose and study probabilistic models of random monomial ideals. We prove theorems about the probability distributions, expectations and thresholds for events involving monomial ideals with given Hilbert function, Krull dimension, first graded Betti numbers, and present several experimentally-backed conjectures about regularity, projective dimension, strong genericity, and Cohen-Macaulayness of random monomial ideals.

1. Introduction

Randomness has long been used to study polynomials. A commutative algebraist's interest in randomness stems from the desire to understand "average" or "typical" behavior of ideals and rings. A natural approach to the problem is to define a probability distribution on a set of ideals or varieties, which, in turn, induces a distribution on algebraic invariants or properties of interest. In such a formal setup, questions of expected (typical) or unlikely (rare, non-generic) behavior can be stated formally using probability. Let us consider some examples. Already in the 1930's, Littlewood and Offord [32] studied the expected number of real roots of a random algebraic equation defined by random coefficients. The investigations on random varieties, defined by random coefficients on a fixed Newton polytope support, have generated a lot of work now considered classical (see, e.g., [26, 29, 39] and the references therein). A probabilistic analysis of algorithms called *smooth analysis* has been used in algebraic geometry, see [5, 9]. Roughly speaking, smooth analysis measures the expected performance of an algorithm under slight random perturbations of worst-case inputs. Another more recent example of a notion of algebraic randomness appears in [15, 16], where they consider the distribution of Betti numbers generated randomly using concepts from the Boij-Söderberg theory [18]. Still, there are many other examples of algebraic topics in which probabilistic analysis plays a useful role (see e.g., [13] and the references therein). Our paper introduces a new probabilistic model on monomial ideals inside a polynomial ring.

Why work on a probabilistic model for monomial ideals? There are at least three good reasons: First, monomial ideals are the simplest of ideals and play a fundamental role in commutative algebra, most notably as Gröbner degenerations of general ideals, capturing the range of values shown in all ideals for several invariants (see [12, 17]). Second, monomial ideals provide a strong link to algebraic combinatorics (see [23, 33, 38]). Third, monomial ideals naturally generalize graphs, hypergraphs, and simplicial complexes; and, since the seminal paper of Erdös and Rényi [19], probabilistic methods have been successfully applied to study those objects (see, e.g., [1, 7, 27] and the references therein). Thus, our work is an extension of both classical probabilistic combinatorics and the new trends in stochastic topology.

Our goal is to provide a formal probabilistic setup for constructing and understanding distributions of monomial ideals and the induced distributions on their invariants (degree, dimension, Hilbert

[1]This work is supported by the NSF collaborative grants DMS-1522662 to Illinois Institute of Technology and DMS-1522158 to the Univ. of California, Davis.

function, regularity, etc.). To this end, we define below a simple probability distribution on the set of monomial ideals. Drawing a random monomial ideal from this distribution allows us to study the average behavior of algebraic invariants. While we are interested in more general probability distributions on ideals, some of which we describe in Section 5, we begin with the most basic model one can consider to generate random monomial ideals; inspired by classical work, we call this family the *Erdős-Rényi-type model*, or the *ER-type model for random monomial ideals*.

The ER-type model for random monomial ideals. Let k be a field and let $S = k[x_1, \ldots, x_n]$ be the polynomial ring in n indeterminates.

We construct a random monomial ideal in S by producing a random set of generators as follows: Given an integer D and a parameter $p = p(n, D)$, $0 \leq p \leq 1$, include *with probability* p each non-constant monomial of total degree at most D in n variables in a generating set of the ideal. Note that the boundary values $p = 0$ and $p = 1$ make the model trivial so we will not consider those; effectively one may thus define the model parameter space to be $(0, 1)$. In other words, starting with $B = \emptyset$, each monomial in S of degree at most D is added to the set B with equal probability p. The resulting random monomial ideal is then simply $I = (B)$. When $B = \emptyset$, we simply let $I = (0)$.

Henceforth, we will denote by $\mathcal{B}(n, D, p)$ the resulting Erdős-Rényi-type distribution on the sets of monomials. Since sets of monomials are now joint random variables, a random set of monomials $\mathfrak{B}$ drawn from the distribution $\mathcal{B}(n, D, p)$ will be denoted by the standard notation for distributions $\mathfrak{B} \sim \mathcal{B}(n, D, p)$. Note that if $B \subset S$ is any fixed set of monomials of degree at most D each and $\mathfrak{B} \sim \mathcal{B}(n, D, p)$, then

$$P(\mathfrak{B} = B) = p^{|B|}(1-p)^{\binom{D+n}{D}-|B|-1}.$$

In turn, the distribution $\mathcal{B}(n, D, p)$ induces a distribution $\mathcal{I}(n, D, p)$ on ideals. We will use the notation $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ to indicate that $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $\mathfrak{I} = (\mathfrak{B})$ is a random monomial ideal generated by the ER-type model.

Before we state our results, let us establish some probabilistic notation that helps highlight our point of view. Readers already familiar with random structures and probability background may skip this part, but those interested in further details are directed to many excellent texts on random graphs and the probabilistic method including [1, 7, 25].

Given a random variable X, we denote its *expected value* by $\mathbb{E}[X]$, its *variance* by $\mathrm{Var}[X]$, and its *covariance* with random variable Y as $\mathrm{Cov}[X, Y]$. We will often use in our proofs four well-known facts: the *linearity* property of expectation, i.e., $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, the *first moment method* (a special case of Markov's inequality), which states that $\mathrm{P}(X \neq 0) \leq \mathbb{E}[X]$, and the *second moment method* (a special case of Chebyshev's inequality), which states that $\mathrm{P}(X = 0) \leq \mathrm{Var}[X]/(\mathbb{E}[X])^2$ for X a non-negative integer-valued random variable. Finally, an *indicator random variable* $\mathbf{1}_A$ for an event A is a random variable such that $\mathbf{1}_A = 1$, if event A occurs, and $\mathbf{1}_A = 0$ otherwise. Indicator random variables behave particularly nicely with regard to taking expectations: $\mathbb{E}[\mathbf{1}_A] = P(A)$ and $\mathrm{Var}[\mathbf{1}_A] = P(A)(1-P(A))$ for any event A. Following convention, we abbreviate by saying a property holds *a.a.s*, to mean that a property holds *asymptotically almost surely* if, over a sequence of sets, the probability of having the property converges to 1.

An important point is that $\mathcal{I}(n, D, p)$ is a *parametric* probability distributions on the set of monomial ideals, because it clearly depends on the value of the probability parameter p. A key concern of our paper is to investigate how invariants evolve as p changes. To express some of our results, we will need to use asymptotic analysis of probabilistic events: For functions $f, g : \mathbb{N} \to \mathbb{R}$, we write $f(n) = \mathrm{o}(g(n))$, and $g(n) = \omega(f(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. We also write $f(n) \sim g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 1$, and $f(n) \asymp g(n)$ if there exist positive constants n_0, c_1, c_2 such that $c_1 g(n) \leq$

$f(n) \leq c_2 g(n)$ when $n \geq n_0$. When a sequence of probabilistic events is given by $f(n)$ for $n \in \mathbb{N}$, we say that f happens *asymptotically almost surely*, abbreviated *a.a.s.*, if $\lim_{n\to\infty} \mathrm{P}\left(f(n)\right) = 1$.

In analogy to graph-theoretic properties, we define an *(monomial) ideal-theoretic property* Q to be the set of all (monomial) ideals that have property Q. A property Q is *monotone increasing* if for any monomial ideals I and J such that $I \subseteq J$, $I \in Q$ implies $J \in Q$ as well. We will see that several algebraic invariants on monomial ideals are monotone.

Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. A function $f : \mathbb{N} \to \mathbb{R}$ is a *threshold function* for a monotone property Q if for any $p : \mathbb{N} \to [0, 1]$:

$$\lim_{D\to\infty} \mathrm{P}\left(\mathfrak{I} \in Q\right) = \begin{cases} 0, & \text{if } p = \mathrm{o}\left(f(D)\right), \\ 1, & \text{if } p = \omega\left(f(D)\right). \end{cases}$$

A similar definition holds for the case when f is a function of n and $n \to \infty$. A threshold function gives a zero/one law that specifies when a certain behavior appears.

Our results. Our results, summarized in the list (A)-(E) below, describe the kind of monomial ideals generated by the ER-type model, in the sense that we can get a handle on both the average and the extreme behavior of the random ideals and how various ranges of the probability parameter p control those properties. The list also gives an outline of the paper.

(A) Hilbert functions and the distribution of monomial ideals. In Section 2, we begin our investigation of random monomial ideals by showing that $\mathcal{I}(n, D, p)$ is not a uniform distribution on all monomial ideals, but instead, the probability of choosing a particular monomial ideal under the ER-type model is completely determined by its Hilbert function and the first total Betti number of its quotient ring.

For an ideal I, denote by $\beta_{i,j} = \beta_{i,j}(S/I)$ the (i, j)-th graded Betti number of S/I, that is, the number of syzygies of degree j at step i of the minimal free resolution. We will collect the *first graded Betti numbers*, counting the minimal generators of I, in a vector $\hat{\beta}_1 = (\beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,reg_0})$, and denote by $\beta_1 = \beta_{1,1} + \beta_{1,2} + \cdots + \beta_{1,reg_0}$ the *first total Betti number* of S/I. We will denote by $h_I(\cdot)$, or simply h_I, the Hilbert function of the ideal I. We can now state the following foundational result:

Theorem 1.1. *Let $I \subseteq S$ be a fixed monomial ideal generated in degree at most D. The probability that the random monomial ideal $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ equals I is*

$$P(\mathfrak{I} = I) = p^{\beta_1(S/I)}(1-p)^{-1+\sum_{d=1}^{D} h_I(d)}.$$

In this way, two monomial ideals with the same Hilbert function and the same number of minimal generators have the same probability of occurring. Then, the following natural question arises: How many monomial ideals in n variables with generators of degree less than or equal to D and with a given Hilbert function are there? It is well-known that one can compute the Hilbert function from the graded Betti numbers (see e.g., [38]). Thus it is no surprise that we can state a combinatorial lemma to count monomial ideals in terms of their graded Betti numbers. If we denote by $NMon(n, D, h, \hat{\beta}_1)$ the number of possible monomial ideals in n variables, generated in degree no more than D, Hilbert function h, and first graded Betti numbers $\hat{\beta}_1 = (\beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,reg_0})$, then $NMon(n, D, h, \hat{\beta}_1)$ is equal to the number of $0-1$ vertices of a certain convex polytope (Lemma 2.3).

Based on Theorem 1.1 and Lemma 2.3, we can provide a closed formula for the induced distribution on Hilbert functions under the ER-type model, that is, the probability that the ER-type model places on any given Hilbert function. This is presented in Theorem 2.5.

(B) The Krull dimension and random monomial ideals. Section 3 is dedicated to a first fundamental ring invariant: the Krull dimension $\dim S/\mathfrak{I}$. By encoding the Krull dimension as the transversal number of a certain hypergraph, we can show (see Theorem 3.1 for the precise statement) that the probability that $\dim(S/\mathfrak{I})$ is equal to t for $0 \leq t \leq n$, for a random monomial ideal $\mathfrak{I} \sim \mathcal{I}(n, D, p)$, is given by a polynomial in p of degree $\sum_{i=1}^{t+1} \binom{D}{i}\binom{n}{i}$. This formula is exponentially large but we make it explicit in some interesting values of t (Theorem 3.2); which in particular give a complete description of the case for two- and three-variable polynomials.

Turning to asymptotic behavior, we prove that the Krull dimension of $(S/\mathfrak{I})$ can be controlled by bounding the asymptotic growth of the probability parameter $p = p(D)$ as $D \to \infty$. The evolution of the Krull dimension in terms of p is illustrated in Figure 1. The result is obtained by combining the family of threshold results from Theorem 3.4 and is stated in the following corollary:

Corollary 1.2. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. For any $0 \leq t < n$, if the parameter $p = p(D)$ is such that $p = \omega\left(D^{-(t+1)}\right)$ and $p = \mathrm{o}\left(D^{-t}\right)$, then $\dim(S/\mathfrak{I}) = t$ asymptotically almost surely.*

It is very useful to consider the evolution – as the probability p increases from 0 to 1 – of the random monomial ideal from the ER-type model and its random generating set $\mathfrak{B}$. For very small values of p, $\mathfrak{B}$ is all but guaranteed to be empty, and the random monomial ideal is asymptotically almost surely the zero ideal. As p increases, the random monomial ideal evolves into a more complex ideal generated in increasingly smaller degrees and support. Simultaneously, as the density of $\mathfrak{B}$ continues to increase with p, smaller-degree generators appear; these divide increasingly larger numbers of monomials and the random ideal starts to become less complex as its minimal generators begin to have smaller and smaller support, causing the Krull dimension to drop. Finally, the random ideal becomes 0-dimensional and continues to evolve towards the maximal ideal.

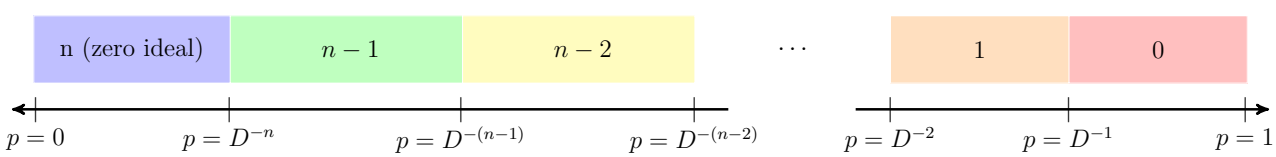

FIG 1. *Evolution of the Krull dimension of $S/\mathfrak{I}$, where $\mathfrak{I}$ is the random monomial ideal $\mathfrak{I} \sim \mathcal{I}(n, D, p)$, as the probability parameter p changes.*

(C) First Betti numbers and random monomial ideals. The ER-type model induces a distribution on the Betti numbers of the coordinate rings of random monomial ideals. To study this distribution, we first ask: *what are the first Betti numbers generated under the model?* A natural way to 'understand' a distribution is to compute its expected value, i.e., the average first Betti number. In Theorem 4.1 we establish the asymptotic behavior for the expected number of minimal generators in a random monomial ideal

We also establish and quantify threshold behavior of first graded Betti numbers of random monomial ideals $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ in two regimes, as n or D go to infinity, in Theorem 4.4. As in the Krull dimension case, we combine the threshold results to obtain the following corollary:

Corollary 1.3. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ and n be fixed. For any $d = d(D)$, such that $1 < d \leq D$ and $\lim_{D\to\infty} d(D) = \infty$, if the parameter $p = p(D)$ is such that $p = \omega\left(d^{-n}\right)$ and $p = \mathrm{o}\left((d-1)^{-n}\right)$, then $\operatorname{initdeg}(\mathfrak{I}) = d$ asymptotically almost surely. Furthermore, the same result holds in the case when D is fixed and $n \to \infty$, with bounds $p(n) = \omega\left(n^{-d}\right)$ and $p(n) = \mathrm{o}\left(n^{-d+1}\right)$.*

While Corollary 1.3 explains how the parameter p controls the smallest non-zero first Betti number, it gives no information regarding the largest such number; the latter, of course, leads the complexity of the first Betti numbers. To that end, we study the degree complexity $\mathrm{reg}_0(\mathfrak{I})$, introduced by Bayer and Mumford in [4] as the maximal degree of any reduced Gröbner basis of I; for monomial ideals, this is simply the highest degree of a minimal generator. In Theorem 4.7 we show that p can be specified so that, even though B contains many large degree monomials, none are minimal.

Intuitively, one can think of the above results in following terms: when the growth of p is bounded *above* by that of $1/D$, Corollary 1.3 provides the minimum degree of a minimal generator of $\mathfrak{I}$. When the growth of p is bounded *below* by that of $1/D$, Theorem 4.7 provides the maximum degree of a minimal generator of $\mathfrak{I}$. The evolution of the minimum degree of a minimal generator for the case when n is fixed and D tends to infinity is illustrated in Figure 2:

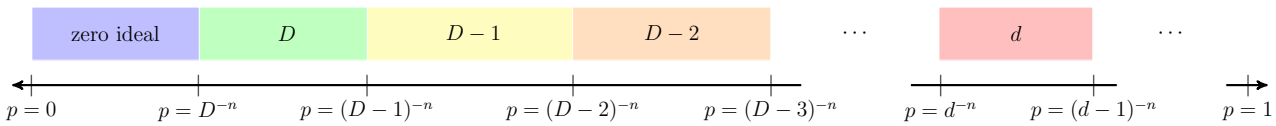

FIG 2. *Evolution of the initial degree of $S/\mathfrak{I}$, where $\mathfrak{I}$ is the random monomial ideal $\mathfrak{I} \sim \mathcal{I}(n, D, p)$, as the model parameter p changes. The initial degree, that is, the smallest integer k such that $\beta_{1,j} = 0$ for $j < k$ and $\beta_{1,k} \neq 0$, is the value displayed in each box. d is a function of D such that $d \to \infty$ as $D \to \infty$.*

(D) Other probabilistic models for generating monomial ideals. In Section 5 we define more general models for random monomial ideals. This is useful to study distributions on ideals that place varying probabilities on monomials depending on a property of interest. For example, one may wish to change the likelihood of monomials according to their total degree. To that end, we define a *general multiparameter model* for random monomials, show how the ER-type model is a special case of it, and introduce the graded model. In Theorem 5.2 we show there is a choice of probability parameters in the general model for random monomial ideals that recovers, as a special case, the multiparameter model for random simplicial complexes of [11], which itself generalizes various models for random graphs and clique complexes presented in [27, 31].

(E) Experiments & conjectures. Section 6 contains a summary of computer simulations of additional algebraic properties of random monomial ideals using the ER-type model. For each triple of selected values of model parameters p, n, and D, we generate 1000 monomial ideals $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ and compute their various properties. Our simulation results are summarized in Tables 3, 4, 5, 6, 7, and 8. The experiments suggest several trends and conjectures, which we state explicitly.

2. Hilbert functions and the Erdős-Rényi distribution on monomial ideals

The *Hilbert function*, and the Hilbert polynomial that it determines, are important tools in the classification of ideals via the Hilbert scheme [17]. Of course, monomial ideals are key to this enterprise: given an arbitrary homogeneous ideal I and a monomial order on $k[x_1, \ldots, x_n]$, the monomial ideal generated by the leading terms of the polynomials in I has the same Hilbert function as I. Therefore the study of all Hilbert functions reduces to the study of Hilbert functions of monomial ideals. Moreover, the Hilbert function $h_I(\cdot)$ of a monomial ideal $I \subseteq S$ has a useful combinatorial meaning: the value of $h_I(d)$, $d \leq D$, counts the number of standard monomials - monomials of degree exactly d that are *not* contained in I.

In this section we explore the Erdős-Rényi distribution of monomial ideals and the induced distribution on Hilbert functions. It turns out they are intimately related.

2.1. Distribution of monomial ideals

Our first theorem precisely describes the distribution $\mathcal{I}(n, D, p)$ on ideals specified by the ER-type model. Recall that Theorem 1.1 states that the Hilbert function and first total Betti number determine the probability of any given ideal in n variables generated in degree at most D.

Proof of Theorem 1.1. Fix $I \subseteq S$ generated in degree at most D and let M be the unique minimal set of generators of I. Then, $\mathfrak{I} = I$ if and only if $\mathfrak{B} \supseteq M$ and no monomial $\mathbf{x}^\alpha$ such that $\mathbf{x}^\alpha \notin I$ is in $\mathfrak{B}$. Let A_1 denote the event that each of the $\beta_1(S/I)$ elements of M is in $\mathfrak{B}$ and let A_2 denote the event that no monomial $\mathbf{x}^\alpha$ such that $\mathbf{x}^\alpha \notin I$ is in $\mathfrak{B}$. Then, the event $\mathfrak{I} = I$ is equivalent to the event $A_1 \cap A_2$. Since the events A_1 and A_2 are independent, $P(\mathfrak{I} = I) = P(A_1 \cap A_2) = P(A_1)P(A_2)$. Observe that $P(A_1) = p^{\beta_1(S/I)}$, since each of the $\beta_1(S/I)$ elements of M is chosen to be in $\mathfrak{B}$ independently with probability p and $P(A_2) = (1-p)^{-1+\sum_{d=1}^{D} h_I(d)}$, since there are exactly $-1 + \sum_{d=1}^{D} h_I(d)$ monomials of degree at most D not contained in I and each of them is excluded from $\mathfrak{B}$ independently with probability $1 - p$. □

Consider a special case of this theorem in the following example.

Example 2.1 (Principal random monomial ideals in 2 variables)**.** Fix $n = 2$ and $D > 0$. In this example, we calculate the probability of observing the principal random monomial ideal $(\mathfrak{B}) = (x^\alpha y^\gamma) \subseteq k[x, y]$, $\alpha + \gamma \leq D$. We observe the ideal $(x^\alpha y^\gamma)$ exactly when $x^\alpha y^\gamma \in \mathfrak{B}$ *and* $x^{\alpha'} y^{\gamma'} \notin \mathfrak{B}$ for every monomial $x^{\alpha'} y^{\gamma'}$ with $\alpha' < \alpha \leq D$ or $\gamma' < \gamma \leq D$. Thus, we must count the lattice points below the "staircase" whose only corner is at $(\alpha, \gamma) \in \mathbb{Z}^2_{\geq 0}$. A simple counting argument shows that there are exactly $\frac{\gamma}{2}(2D - \gamma + 3) + \frac{\alpha}{2}(2D - \alpha + 3) - \alpha\gamma - 1$ lattice points. Hence,

$$\mathrm{P}\left(\mathfrak{I} = (x^\alpha y^\gamma)\right) = p(1-p)^{\frac{\gamma}{2}(2D-\gamma+3)+\frac{\alpha}{2}(2D-\alpha+3)-\alpha\gamma-1}.$$

Note that the right hand side of this expression is maximized at $\alpha = 1, \gamma = 0$ or $\alpha = 0, \gamma = 1$, so that among the principal ideals, the most likely to appear under the ER-type model are (x) and (y).

A corollary of Theorem 1.1 follows directly from the well-known formulas for Stanley-Reisner rings:

Corollary 2.2. *If I is a square-free monomial ideal, then the probability of I under the ER-type model is determined by the number of minimal non-faces and faces of the associated simplicial complex.*

2.2. Distribution of Hilbert functions

Since the Erdős-Rényi-type model specifies a distribution on monomial ideals, it provides a formal probabilistic procedure to generate random Hilbert functions, in turn providing a very natural question: *what is the induced probability of observing a particular Hilbert function? In other words, what is the most likely Hilbert function under the ER-type model? If we randomly generate a monomial ideal, are all Hilbert functions equally likely?*

Not all possible non-negative vectors are first graded Betti vectors, but Lemma 2.3 below gives an algorithmic way to determine, using polyhedral geometry, whether a particular potential set of first graded Betti numbers can occur for a monomial ideal in n variables. We note that Onn and Sturmfels introduced other polytopes useful in the study of initial (monomial) ideals of a zero-dimensional ideal for n generic points in affine d-dimensional space: the *staircase polytope* and the simpler *corner cut polyhedron* [34].

Lemma 2.3. *Let α, γ denote exponent vectors of monomials with n variables and total degree no more than D. Denote by $NMon(n, D, h, \hat{\beta}_1)$ the number of possible monomial ideals in n variables, with generating monomials of degree no more than D, given Hilbert function h, and specific first graded Betti numbers vector $\hat{\beta}_1 = (\beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,reg_0})$. Then $NMon(n, D, h, \hat{\beta}_1)$ is equal to the number of $0-1$ solutions for the system of equations and inequalities with $\binom{n+D}{D} - 1$ variables, indexed by possible monomials, and the following constraints:*

$$\sum_{deg(\alpha)=L} x_\alpha = \binom{n+L}{L} - h(L), \quad \forall L = 1, \ldots D$$
$$\sum_{deg(\alpha)=\beta_{i,1}} x_\alpha = \beta_{1,i}, \quad \forall i = 1 \ldots reg_0$$
$$x_\alpha \leq x_\gamma, \quad \forall\, \alpha|\gamma,\, deg(\alpha) + 1 = deg(\gamma).$$

Thus $NMon(n, D, h, \hat{j})$ equals the number of $0-1$ vertices of a convex polytope. The monomial ideals in n variables with generators of degree less than or equal to D and with given Hilbert function can be easily expressed as a sum of terms $NMon(n, D, h, \hat{j})$ running over possible first graded Betti numbers vectors.

Proof of Lemma 2.3. The variables are indicator variables recording when a monomial x^α is included or not included. The first type of equations forces the values of the Hilbert function at degree L to be satisfied. The second type of equations forces the exact number of monomials generators of given degree i be as dictated by the value of $\beta_{1,i}$. The inequalities $x_\alpha \leq x_\gamma$, for all $\alpha|\gamma$, ensure the set of chosen monomials is closed under divisibility and thus forms an ideal. The non-redundant subset of these satisfying $deg(\alpha) + 1 = deg(\gamma)$ suffices to cut out the polyhedron. □

Since we are interested in computing the probability of having a particular Hilbert function, we will need to understand which values of first graded Betti numbers occur for ideals with a particular Hilbert function. There can be many vectors of first graded Betti numbers for a given Hilbert function, but it turns out that there are upper bounds for each potential graded Betti number. In 1993, in independent work, Bigatti [6] and Hulett [24] proved that the largest values for the graded Betti numbers, among all rings of the form a polynomial ring modulo a homogeneous ideal, are attained by the so-called *lex-segment ideal*. Their proof holds in coefficient fields of characteristic zero, but in 1996 Pardue [35] extended the theorem to an arbitrary field. For completeness, recall that a monomial ideal I is said to be lex-segment if $x^a \in I$ whenever $x^a >_{lex} x^b$ and $x^b \in I$. If I is lex-segment, then the degree-j component $[I]_j$ of the ideal I consists of the $\dim_k [I]_j$ lexicographically largest monomials in $[S]_j$. It follows that for any Hilbert function h of a homogenous ideal, there exists a unique lex-segment ideal L such that $h_L = h$ (that is, $h_L(r) = h(r)$ for all $r \geq 0$). In addition, the information is explicit, as there are simple combinatorial formulas for the Betti numbers of lex-segment ideals. The Bigatti-Hulett-Pardue theorem, stated below, shows that lex-segment ideals give the necessary bounds to prove of Theorem 2.5.

Lemma 2.4 (Bigatti-Hulett-Pardue)**.** *Let I be an ideal of the polynomial ring and let L be the lex-segment ideal such that $h_I = h_L$. Then, the corresponding graded Betti numbers satisfy the inequality $\beta_{i,j}(I) \leq \beta_{i,j}(L)$.*

Theorem 2.5. *Let h be a Hilbert function for an ideal of the polynomial ring in n variables. Denote by $\beta^L_{1,j}$ the lex-segment ideal bounds on the first graded Betti numbers. Denote by $NMon(n, D, h, \hat{j})$*

the number of distinct monomial ideals in n variables, with generators of degree less than or equal to D, Hilbert function h, and vector of first graded Betti numbers $\hat{j}$. Then the probability of generating a random monomial ideal $\mathfrak{I}$ from the ER-type model with a given Hilbert function h is expressed by the following formula

$$P(h_{\mathfrak{I}} = h) = \sum_{j_1=0}^{\beta^L_{1,1}} \sum_{j_2=0}^{\beta^L_{1,2}} \cdots \sum_{j_{reg_0}=0}^{\beta^L_{1,reg_0}} NMon(n, D, h, \hat{j}) \cdot p^{||\hat{j}||_1} (1-p)^{-1+\sum_{d=1}^{D} h(d)}, \tag{2.1}$$

where $\hat{j}$ denotes the vector $\hat{j} = (j_1, j_2, \ldots, j_{reg_0})$ and $||\hat{j}||_1 = j_1 + \cdots + j_{reg_0}$.

Proof of Theorem 2.5. To derive the above formula for the probability of randomly generating a monomial ideal with a particular Hilbert function h using the ER model, we decompose the random event into probability events that are combinatorially easy to count. The probability of generating random monomial ideals with a given Hilbert function h is the sum of disjoint events of choosing a possible such monomial ideal with a certain vector of first graded Betti numbers $\hat{j}$; that is:

$$P(h_{\mathfrak{I}} = h) = \sum_{j_1=0}^{\beta^L_{1,1}} \sum_{j_2=0}^{\beta^L_{1,2}} \cdots \sum_{j_{reg_0}=0}^{\beta^L_{1,reg_0}} P(h_{\mathfrak{I}} = h \text{ and } \beta_{1,i}(S/\mathfrak{I}) = j_i, i = 1, \ldots, reg_0). \tag{2.2}$$

Further, the event $\{h_{\mathfrak{I}} = h \text{ and } \beta_{1,i}(S/\mathfrak{I}) = j_i\}$ is the disjoint union of choosing each possible monomial ideal I with the given Hilbert function h and graded Betti numbers $\hat{j}$, thus:

$$P(h_{\mathfrak{I}} = h \text{ and } \beta_{1,i}(S/\mathfrak{I}) = j_i) = \sum_{I \,:\, h_I = h \text{ and } \beta_{1,i}(I) = j_i} P(\mathfrak{I} = I). \tag{2.3}$$

Notice that each of the probabilities $P(\mathfrak{I} = I)$ is provided by Theorem 1.1 and, in particular, only depends on the Hilbert function and the first *total* Betti number of I. Therefore:

$$\begin{aligned} P(h_{\mathfrak{I}} = h \text{ and } \beta_{1,i}(S/\mathfrak{I}) = j_i) &= \sum_{I \,:\, h_I = h \text{ and } \beta_{1,i}(I) = j_i} p^{||\hat{j}||_1} (1-p)^{-1+\sum_{d=1}^{D} h(d)} \\ &= NMon(n, D, h, \hat{j}) \cdot p^{||\hat{j}||_1} (1-p)^{-1+\sum_{d=1}^{D} h(d)}. \end{aligned} \tag{2.4}$$

All that is left to do is to calculate the numbers $NMon(n, D, h, \hat{j})$. For this we rely on the distinct decompositions of the vector of first graded Betti number β_1 into its possible component values and a polyhedral interpretation of $NMon(n, D, h, \beta_1)$ where the Betti number β_1 is decomposed as $\beta_1 = (\beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,reg_0})$, reg_0 being the degree complexity of the ideal in question.

It is well-known that each $\beta_{1,i}$ is bounded below by 0 and above by the corresponding Betti number for the lex-segment ideals of Macaulay and Bigatti-Hulett-Pardue et al (see [6, p. 35]). Essentially, because of the bounds, the search for all possible sets of graded Betti numbers for a particular Hilbert function is a bounded problem. On the other hand, we saw in Lemma 2.3 one can express the number of monomial ideals with parameters n, D, h and vector of first graded Betti numbers $(\beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,reg_0})$ as the number of vertices of a convex polytope presented by the equations and inequalities in Lemma 2.3. □

Note that, from the point of view of computer science, it is useful to know that the Hilbert function h can be specified with one polynomial (the Hilbert polynomial) as well as a list of finitely many values that do not fit the polynomial.

3. Krull dimension

In this section, we study the Krull dimension $\dim S/\mathfrak{I}$ of a random monomial ideal under the model $\mathcal{I}(n, D, p)$. First, in Section 3.2, we give an explicit formula to calculate the probability of $\mathfrak{I}$ having a given Krull dimension. Second, in Section 3.3, we establish threshold results characterizing the asymptotic behavior of the Krull dimension and we provides ranges of values of p that, asymptotically, control the Krull dimension of the random monomial ideal.

3.1. Hypergraph Transversals and Krull dimension

Let us recall some necessary definitions. A *hypergraph* H is a pair $H = (V, E)$ such that V is a finite vertex set and E is a collection of non-empty subsets of V, called *edges*; we allow edges of size one. A *clutter* is a hypergraph where no edge is contained in any other (see, e.g., [10, Chapter 1]). A *transversal* of H, also called a *hitting set* or a *vertex cover*, is a set $T \subseteq V$ such that no edge in E has a nonempty intersection with T. The *transversal number* of H, also known as its *vertex cover number* and denoted $c(H)$, is the minimum cardinality of any transversal of H.

Given a collection B of monomials in $k[x_1, \ldots, x_n]$, the *support hypergraph* of B is the hypergraph $G(B)$ with vertex set $V = \{x_1, \ldots, x_n\}$, and edge set $E = \{\mathrm{supp}(x^\alpha) \mid x^\alpha \in B\}$, where $\mathrm{supp}(x^\alpha)$ denotes the support of the monomial x^α, i.e., $x_i \in \mathrm{supp}(x^\alpha)$, if and only if $x_i \mid x^\alpha$. What we require here is that for $I = (B)$, the Krull dimension of S/I is determined by the transversal number of $G(B)$ through the formula

$$\dim S/I + c\left(G(B)\right) = n. \tag{3.1}$$

This formula is known to hold for square-free monomial ideals minimally generated by B (see e.g., [20]). Since the support hypergraph of B is equal to that of its maximal square-free subset and $c\left(G(B)\right)$ is invariant under the choice of generating set B of I, the equation, obviously, also holds for general ideals.

3.2. Probability distribution of Krull dimension

The relationship between Krull dimension and hypergraph transversals leads to a complete characterization of the probability of producing a monomial ideal with a particular fixed Krull dimension in the ER-type model (Theorem 3.1). Explicitly computing the values given by Theorem 3.1 requires an exhaustive combinatorial enumeration, which we demonstrate for several special cases in Theorem 3.2. Nevertheless, where the size of the problem makes enumeration prohibitive, Theorem 3.1 gives that $\mathrm{P}\left(\dim S/\mathfrak{I} = t\right)$ is always a polynomial in p of a specific degree, and thus can be approximated by numerically evaluating $\mathrm{P}\left(\dim S/\mathfrak{I} = t\right)$ for a sufficient number of p values, and interpolating.

Theorem 3.1. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. For any integer t, $0 \le t \le n$, the probability that $S/\mathfrak{I}$ has Krull dimension t is given by a polynomial in p of degree $\sum_{i=1}^{t+1} \binom{D}{i}\binom{n}{i}$. More precisely,*

$$\mathrm{P}\left(\dim S/\mathfrak{I} = t\right) = \sum_{C \in \mathcal{C}_{n-t}} \prod_{\sigma \in E(C)} 1 - (1-p)^{\binom{D}{|\sigma|}} \prod_{\substack{\sigma' \subset \{x_1, \ldots, x_n\} \\ \sigma \not\subseteq \sigma' \forall \sigma \in E(C)}} (1-p)^{\binom{D}{|\sigma'|}},$$

where $\mathcal{C}_{n-t}$ is the set of all clutters on $\{x_1, \ldots, x_n\}$ with transversal number $n - t$.

Proof. Let $\mathfrak{B}$ be the random generating set that gave rise to $\mathfrak{I}$. By Equation (3.1), the probability that $S/\mathfrak{I}$ has Krull dimension t is equal to the probability that $c(G(\mathfrak{B})) = n - t$. Let $G_{\min}(\mathfrak{B})$ be the hypergraph obtained from $G(\mathfrak{B})$ by deleting all edges of $G(\mathfrak{B})$ that are strict supersets of other edges in $G(\mathfrak{B})$: in other words, $G_{\min}(\mathfrak{B})$ contains all edges σ of $G(\mathfrak{B})$ for which there is no σ' in $G(\mathfrak{B})$ such that $\sigma' \subsetneq \sigma$.

Also let $\mathcal{C}_{n-t}$ denote the set of clutters on $\{x_1, \ldots, x_n\}$ with transversal number $n - t$. Then $c(G(\mathfrak{B})) = n-t$ if and only if $G_{\min}(\mathfrak{B}) = C$ for some $C \in \mathcal{C}_{n-t}$. As these are disjoint events, we have

$$\mathrm{P}\left(\dim S/\mathfrak{I} = t\right) = \sum_{C \in \mathcal{C}_{n-t}} \mathrm{P}\left(G_{\min}(\mathfrak{B}) = C\right). \tag{3.2}$$

Let $\sigma \subset \{x_1, \ldots, x_n\}$. There are $\binom{D}{|\sigma|}$ monomials supported on σ, so

$$\mathrm{P}\left(\sigma \in E(G(\mathfrak{B}))\right) = 1 - (1-p)^{\binom{D}{|\sigma|}}, \text{ and } \mathrm{P}\left(\sigma \notin E(G(\mathfrak{B}))\right) = (1-p)^{\binom{D}{|\sigma|}}.$$

Note that both expressions are polynomials in p of degree exactly $\binom{D}{|\sigma|}$. Now $G_{\min}(\mathfrak{B}) = C$ if only if:

1. every edge σ of C is an edge of $G(\mathfrak{B})$, and
2. every $\sigma' \subset \{x_1, \ldots, x_n\}$ such that σ' does not contain any $\sigma \in E(C)$ is *not* an edge of $G(\mathfrak{B})$.

Note that condition 2 is the contrapositive of: any edge of $G(\mathfrak{B})$ which is not an edge of C is not minimal. Since edges are included in $G(\mathfrak{B})$ independently, the above are also independent events and the probability of their intersection is

$$\begin{aligned}
\mathrm{P}\left(G_{\min}(\mathfrak{B}) = C\right) &= \prod_{\sigma \in E(C)} \mathrm{P}\left(\sigma \in E(G(\mathfrak{B}))\right) \prod_{\substack{\sigma' \subset \{x_1,\ldots,x_n\} \\ \sigma \not\subseteq \sigma' \forall \sigma \in E(C)}} \mathrm{P}\left(\sigma' \notin E(G(\mathfrak{B}))\right) \\
&= \prod_{\substack{\sigma \in E(C) \\ \text{s.t. } |\sigma| \leq t+1}} \mathrm{P}\left(\sigma \in E(G(\mathfrak{B}))\right) \prod_{\substack{\sigma' \subset \{x_1,\ldots,x_n\} \\ \text{s.t. } |\sigma'| \leq t \text{ and} \\ \sigma \not\subseteq \sigma' \forall \sigma \in E(C)}} \mathrm{P}\left(\sigma' \notin E(G(\mathfrak{B}))\right).
\end{aligned} \tag{3.3}$$

The last equality follows because for $C \in \mathcal{C}_{n-t}$, if σ is an edge of C then $|\sigma| \leq t+1$, and also if σ' is a subset of $\{x_1, \ldots, x_n\}$ satisfying $\sigma \not\subseteq \sigma'$ for all $\sigma \in E(C)$, then $|\sigma'| \leq t$. To show the first statement, suppose $\sigma \in E(C)$ with $|\sigma| > t+1$. Since C is a clutter, no proper subset of σ is an edge of C, so for every $\sigma' \in E(C)$, $\sigma' \neq \sigma$, σ' contains at least one vertex not in σ. Hence the set $T = \{x_1, \ldots, x_n\} \backslash \sigma$ intersects every edge of C except σ. By taking the union of T with any one vertex in σ, we create a transversal of C of cardinality $|T| + 1 = n - |\sigma| + 1 < n - t$, contradicting $c(C) = n - t$.

For the second statement, suppose $|\sigma'| > t$. By assumption no edge of C is a subset of σ', so every edge of C contains at least one vertex in the set $T = \{x_1, \ldots, x_n\} \backslash \sigma'$. Hence T is a transversal of C with $|T| = n - |\sigma'| < n - t$, again a contradiction.

No subset of $\{x_1, \ldots, x_n\}$ can appear in both index sets of (3.3), and each subset that does appear contributes a polynomial in p of degree $\binom{D}{i}$. It follows that (3.3) is a polynomial in p of degree no greater than $\sum_{i=1}^{t+1} \binom{D}{i}\binom{n}{i}$, and hence so is $\mathrm{P}\left(\dim S/\mathfrak{I} = t\right)$ as (3.2) is a sum of such polynomials.

To prove this bound is in fact the precise degree of the polynomial, we show there is a particular clutter for which (3.2) has degree exactly $\sum_{i=1}^{t+1} \binom{D}{i}\binom{n}{i}$, and furthermore, that for every other clutter the expression in (3.3) is of strictly lower degree. Hence the sum over clutters in (3.2) has no cancellation of this leading term.

Consider the hypergraph K_n^{t+1} that contains all $\binom{n}{t+1}$ edges of cardinality $t+1$ and no other edges. Then $K_n^{t+1} \in \mathcal{C}_{n-t}$ and

$$\begin{aligned}\mathrm{P}\left(G_{\min}(\mathfrak{B}) = K_n^{t+1}\right) &= \prod_{\substack{\sigma\in\{x_1,\dots,x_n\}\\ |\sigma|=t+1}} \mathrm{P}\left(\sigma \in E(G(\mathfrak{B}))\right) \prod_{\substack{\sigma\in\{x_1,\dots,x_n\}\\ |\sigma|\le t}} \mathrm{P}\left(\sigma \notin E(G(\mathfrak{B}))\right)\\ &= \left(1-(1-p)^{\binom{D}{t+1}}\right)^{\binom{n}{t+1}}(1-p)^{Dn+\binom{D}{2}\binom{n}{2}+\cdots+\binom{D}{t}\binom{n}{t}}\end{aligned}$$

which has the correct degree. On the other hand, if $C \in \mathcal{C}_{n-t}$ and $C \neq K_n^{t+1}$, then at least one edge σ of C is not an edge of K_n^{t+1}; hence $|\sigma| \le t$. All subsets properly containing σ are neither edges of C, nor do they satisfy condition 2 above, hence these subsets are not indexed by either product in 3.3. In particular there are positively many subsets of cardinality $t+1$ which do not contribute factors to $\mathrm{P}\left(G_{\min}(\mathfrak{B}) = C\right)$. □

To use the main formula of Theorem 3.1, one must be able to enumerate all clutters on n vertices with transversal number $n-t$. When t is very small or very close to n this is tractable, as we see in the following theorem.

Theorem 3.2. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. Then,*

a) $\mathrm{P}\left(\dim S/\mathfrak{I} = 0\right) = \left(1-(1-p)^D\right)^n$.

b) $\mathrm{P}\left(\dim S/\mathfrak{I} = 1\right) = \sum_{j=0}^{n-1}\binom{n}{j}(1-(1-p)^D)^j(1-p)^{D(n-j)}\left(1-(1-p)^{\binom{D}{2}}\right)^{\binom{n-j}{2}}$.

c) $\mathrm{P}\left(\dim S/\mathfrak{I} = n-1\right) = -(1-p)^{\binom{n+D}{n}-1} + \sum_{j=1}^{n}(-1)^{j-1}\binom{n}{j}(1-p)^{\binom{n+D}{n}-1-\binom{n+D-j}{n}}$.

d) $\mathrm{P}\left(\dim S/\mathfrak{I} = n\right) = (1-p)^{\binom{n+D}{n}-1}$.

For $n \le 4$, the listed formulas gives the complete probability distribution of Krull dimension induced by $\mathcal{I}(n, D, p)$, for any integer D.

Proof. Proof of Part (a) For $\mathfrak{B} \sim \mathcal{B}(n, D, p)$, $\mathfrak{I} = (\mathfrak{B})$, $S/\mathfrak{I}$ is zero dimensional if and only if $G_{\min}(\mathfrak{B}) \in \mathcal{C}_n$. There is a single clutter on n vertices with transversal number n: the one with edge set $\{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$. Hence by Theorem 3.1,

$$\mathrm{P}\left(\dim S/\mathfrak{I} = 0\right) = \prod_{i=1}^{n} 1-(1-p)^{\binom{D}{|\{x_i\}|}} = \left(1-(1-p)^D\right)^n.$$

Proof of Part (b): $S/\mathfrak{I}$ is one-dimensional if and only if $G_{\min}(\mathfrak{B}) \in \mathcal{C}_{n-1}$. We wish to describe $\mathcal{C}_{n-1}$. Suppose C is a clutter on n vertices and exactly j of the vertices are contained in a 1-edge. Then $j \neq n$ else C would be the clutter from part a, so let $0 \le j \le n-1$, and denote by V the set of these j vertices. Then V is a subset of any transversal of C. Let $W = \{x_1, \dots, x_n\}\backslash V$, then it can be shown that $c(C) = n-1$ if and only $E(C) = \{\{x_i\} \mid x_i \in V\} \cup \{\{x_i, x_k\} \mid x_i, x_k \in W, x_i \neq x_k\}$. Hence

$$\begin{aligned}\mathrm{P}\left(G_{\min}(\mathfrak{B}) = C\right) &= \prod_{x_i\in V}\mathrm{P}\left(\{x_i\} \in E(G(\mathfrak{B}))\right) \prod_{x_i,x_k\in W}\mathrm{P}\left(\{x_i,x_k\} \in E(G(\mathfrak{B}))\right) \prod_{x_i\in W}\mathrm{P}\left(\{x_i\} \notin E(G(\mathfrak{B}))\right)\\ &= (1-(1-p)^D)^j\left(1-(1-p)^{\binom{D}{2}}\right)^{\binom{n-j}{2}}(1-p)^{D(n-j)}.\end{aligned}$$

The expression for $\mathrm{P}\left(\dim S/\mathfrak{I} = 1\right)$ is obtained by summing over all $\binom{n}{j}$ ways of selecting the j 1-edges, for each $0 \le j \le n-1$.

Proof of Part (c): For the case of $(n-1)$–dimensionality, Theorem 3.1 requires us to consider clutters with transversal number 1: clutters where some vertex appears in all the edges. However, for this case we can give a simpler argument by looking at the monomials in $\mathfrak{B}$ directly. Now the condition equivalent to $(n-1)$–dimensionality is that there is some x_i that divides every monomial in $\mathfrak{B}$.

Fix an i, then there are $\binom{n+D}{D} - 1 - \binom{n+D-1}{D-1}$ monomials that x_i does *not* divide. If F_i is the event that x_i divides every monomial in $\mathfrak{B}$, then $\mathrm{P}(F_i) = (1-p)^{\binom{n+D}{D}-1-\binom{n+D-1}{D-1}}$. To get an expression for $(n-1)$-dimensionality, we need to take the union over all F_i, which we can do using an inclusion-exclusion formula considering the events that two variables divide every monomial, three variables divide every monomial, etc. Finally, we subtract the probability that $\mathfrak{B}$ is empty.

Proof of Part (d): Since only the zero ideal has Krull dimension n, this occurs if and only if $\mathfrak{B}$ is empty, which has probability $(1-p)^{\binom{n+D}{D}-1}$. □

3.3. Threshold and asymptotic results

Next we study the asymptotic behavior of $\dim(S/\mathfrak{I})$. In particular, we derive ranges of values of p that control the Krull dimension as illustrated in Figure 1. As a special case we obtain a threshold result for the random monomial ideal being zero-dimensional.

Lemma 3.3. *Let n, t be fixed integers, with $0 \le t \le n-1$. If $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $p(D) = \mathrm{o}\left(D^{-t}\right)$, then a.a.s. $\mathfrak{B}$ will contain no monomials of support size t or less as D tends to infinity.*

Proof. By the first moment method, the probability that $\mathfrak{B}$ contains some monomial of support at most t is bounded above by the expected number of such monomials. As the number of monomials in n variables with support of size at most t is strictly less than $\binom{n}{t}\binom{D+t}{t}$, the expectation is bounded above by the quantity $p\binom{n}{t}\binom{D+t}{t}$. This quantity tends to zero when $p(D) = \mathrm{o}\left(D^{-t}\right)$ and n and t are constants, thus establishing the lemma. □

Theorem 3.4. *Let n, t be integers and $p = p(D)$ with $1 \le t \le n$ and $0 \le p \le 1$. Suppose that $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. Then D^{-t} is a threshold function for the property that $\dim(S/\mathfrak{I}) \le t-1$. In other words,*

$$\lim_{D\to\infty} \mathrm{P}\left(\dim(S/\mathfrak{I}) \le t-1\right) = \begin{cases} 0, & \text{if } p = \mathrm{o}\left(D^{-t}\right) \\ 1, & \text{if } p = \omega\left(D^{-t}\right) \end{cases}.$$

Proof. Let $p = \mathrm{o}\left(D^{-t}\right)$. Consider the support hypergraph of $\mathfrak{B}$, $G(\mathfrak{B})$. If every monomial in $\mathfrak{B}$ has support of size $t+1$ or more, then any $(n-t)$-vertex set of $G(\mathfrak{B})$ is a transversal, giving $c(G(\mathfrak{B})) \le n-t$. Thus Equation 3.1 and Lemma 3.3 give that $\dim(S/\mathfrak{I}) \le t-1$ holds with probability tending to 0.

Now, let $p = \omega\left(D^{-t}\right)$. Consider $\sigma \subset \{x_1, \ldots, x_n\}$ with $|\sigma| = t$ and let X_σ be the random variable that records the number of monomials in $\mathfrak{B}$ with support exactly the set σ. There are exactly $\binom{D}{t}$ such monomials of degree at most D: each monomial can be divided by all variables in σ and the remainder is a monomial in (some or all of) the variables of σ of degree no greater than $D-t$.

Thus, $\mathbb{E}[X_\sigma] = \binom{D}{t}p$ tends to infinity with D, when $p = \omega\left(D^{-t}\right)$. Further, $\mathrm{Var}[X_\sigma] \le \mathbb{E}[X_\sigma]$, as X_σ is a sum of independent indicator random variables, one for each monomial with support exactly σ. Hence we can apply the second moment method to conclude that as D tends to infinity,

$$\mathrm{P}(X_\sigma = 0) \le \frac{\mathrm{Var}[X_\sigma]}{\mathbb{E}[X_\sigma]^2} \le \frac{1}{\binom{D}{t}p} \to 0.$$

In other words, a.a.s. $\mathfrak{B}$ will contain a monomial with support σ. Since σ was an arbitrary t-subset of the variables $\{x_1, \ldots, x_n\}$, it follows that the same holds for every choice of σ.

The probability that any of the events $X_\sigma = 0$ occurs is bounded above by the sum of their probabilities and tends to zero, as there are a finite number of sets σ. Then, with probability tending to 1, $G(\mathfrak{B})$ contains all size-t edges, which implies that, for every vertex set of size $n-t$ or less, $G(\mathfrak{B})$ contains an edge of size t that is disjoint from it. Thus the transversal number of the hypergraph is at least $n-t+1$, and, by Formula 3.1, $\dim(S/\mathfrak{I}) = n - c(G(\mathfrak{B})) \leq t-1$ a.a.s. □

The smallest choice of t-values in Theorem 3.4, $t = 1$, gives a threshold function for the event that $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ has zero Krull dimension:

Corollary 3.5. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. The function $1/D$ is a threshold function for the 0-dimensionality of $S/\mathfrak{I}$.*

Now we move to obtain Corollary 1.2 announced in the Introduction. Theorem 3.4 establishes a threshold result for each choice of constant t. But, if both $p = \omega\left(D^{-(t+1)}\right)$ and $p = \mathrm{o}\left(D^{-t}\right)$ hold, then the theorem gives that events $\dim(S/\mathfrak{I}) > t$ and $\dim(S/\mathfrak{I}) < t$ each hold with probability tending to 0. Therefore the probability of their union, in other words the probability that $\dim(S/\mathfrak{I})$ is either strictly larger than t or strictly smaller than t, also tends to 0. By combining the threshold results applied to two consecutive constants t and $t+1$ in this way, we have established the Corollary.

4. Minimal generators of random monomial ideals from the ER-type model

A way to measure the complexity of a monomial ideal is to study the number of minimal generators. As before, we will denote by $\beta_1 := \beta_1(S/I)$ the first total Betti number of the quotient ring of the ideal I, that is, the total number of minimal generators of I. Subsection 4.1 addresses this invariant and provides a formula for the asymptotic average value of β_1. A more refined invariant is, of course, the set of first *graded Betti numbers* $\beta_{1,d}(S/I)$, recording the number of minimal generators of I of degree d. Subsection 4.2 provides threshold results that explain how the asymptotic growth of the parameter value p influences the appearance of minimal generators of given degree. Subsection 4.3 provides an analogous result for the *degree complexity* $\mathrm{reg}_0(I)$ of a monomial ideal I, that is, the maximum degree of a minimal generator.

4.1. The first total Betti number

Consider the random variable $\beta_1(S/\mathfrak{I})$ for $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ in the regime $D \to \infty$. Before stating the results of this section, we introduce a concept from multiplicative number theory. The *order r divisor function*, denoted by $\tau_r(k)$, is a function that records the number of ordered factorizations of an integer k into exactly r parts (e.g., $\tau_3(4) = 6$).

Theorem 4.1. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ and let $\beta_1(S/\mathfrak{I})$ be the random variable denoting the number of minimal generators of the random monomial ideal $\mathfrak{I}$. Then,*

$$\lim_{D\to\infty} \mathbb{E}\left[\beta_1(S/\mathfrak{I})\right] = p \sum_{k=0}^{\infty} \tau_n(k+2)(1-p)^k. \tag{3}$$

Proof. For each $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$, let $\mathbf{1}_\alpha$ be the indicator random variable for the event that x^α is a minimal generator of $\mathfrak{I}$. Excluding the constant monomial 1 and x^α itself, there are

$\prod_{i=1}^n(\alpha_i+1)-2$ monomials in S that divide x^α. Therefore, $\mathbb{E}[\mathbf{1}_\alpha] = (1-p)^{\prod_{i=1}^n(\alpha_i+1)-2}$ and letting $D\to\infty$ shows that

$$\lim_{D\to\infty}\mathbb{E}[\beta_1(S/\mathfrak{I})] = \sum_{0\neq\alpha\in\mathbb{Z}^n_{\geq 0}}\mathbb{E}[\mathbf{1}_\alpha] \;=\; p\sum_{0\neq\alpha\in\mathbb{Z}^n_{\geq 0}}(1-p)^{\prod_{i=1}^n(\alpha_i+1)-2}.$$

Since $\prod_{i=1}^n(\alpha_i+1)-2$ is also the number of lattice points contained in the box $\prod_{i=1}^n[0,\alpha_i]$, excluding the origin and α, it follows that every integer $k\geq 0$ appears as an exponent of q in the infinite sum above exactly as many times as $k+2$ can be factored into a product of n ordered factors, so that the above sum may be rewritten as $\sum_{k=0}^\infty \tau_n(k+2)(1-p)^k$. □

When $n=2$, $\tau_2(k)$ is simply the usual divisor function which counts the number of divisors of an integer k. The ordinary generating function of $\tau_2(k)$ has the form of a *Lambert series* (see [14]), which leads to the following result in the two-variable case.

Corollary 4.2. *Let $\mathfrak{I}\sim\mathcal{I}(2,D,p)$. Then,*

$$\lim_{D\to\infty}\mathbb{E}[\beta_1(S/\mathfrak{I})] = \frac{p}{(1-p)^2}\left[\sum_{k=1}^\infty\frac{(1-p)^k}{1-(1-p)^k}\right]-\frac{p}{1-p}.$$

Proof. This follows from the Lambert series identity $\sum_{k=1}^\infty\frac{(1-p)^k}{1-(1-p)^k}=\sum_{k=1}^\infty\tau_2(k)(1-p)^k$ in (3). □

Corollary 4.2 provides an expression for the asymptotic expected number of minimal generators of a random monomial ideal in two variables that can be easily evaluated for any fixed value of p. For example, if $\mathfrak{I}\sim\mathcal{I}(2,D,10^{-5})$, then asymptotically (i.e., for very large values of D) one expects $\mathfrak{I}$ to have about 12 generators on average. On the other hand, when $n\geq 3$, the right-hand side of (3) is difficult to compute exactly, but one can obtain bounds from the fact that $n\leq\tau_n(k)\leq k^n$, which are valid for every $k\geq 2$. Hence

$$n\leq\lim_{D\to\infty}\mathbb{E}[\beta_1(S/\mathfrak{I})]\leq\frac{p}{(1-p)^2}\mathrm{Li}_{-n}(1-p), \tag{4}$$

where $\mathrm{Li}_s(x)$ denotes the *polylogarithm function of order* s defined by $\mathrm{Li}_s(x)=\sum_{k=1}^\infty\frac{x^k}{k^s}$. A polylogarithm of negative integer order can be expressed as a rational function of its argument:

$$\mathrm{Li}_{-n}(1-p)=\sum_{j=0}^n j!S(n+1,j+1)\left(\frac{1-p}{p}\right)^{j+1},$$

where $S(n+1,k+1)$ denotes a *Stirling number of the second kind* (see [30]). The upper bound in (4) is thus of order $\mathrm{O}\left(p^{-n}\right)$.

4.2. The first graded Betti numbers

Here we study the behavior of the random variables $\beta_{1,k}(S/\mathfrak{I})$ for $\mathfrak{I}\sim\mathcal{I}(n,D,p)$. Specifically, we show how the choice of p in the ER-type model controls the minimum degree of a generator of $\mathfrak{I}$, by controlling the degrees of the monomials in the random set $\mathfrak{B}$.

Lemma 4.3. *Let $\mathfrak{B}\sim\mathcal{B}(n,D,p)$, $\mathfrak{I}=\langle\mathfrak{B}\rangle$ and let*

$$\mathfrak{B}_d := \{x^\alpha\in\mathfrak{B} : |\alpha|\leq d\}.$$

a) Let d, D be fixed, $d \leq D$, and $p = p(n)$. Then $f(n) = n^{-d}$ is a threshold function for the property that $|\mathfrak{B}_d| = 0$.

b) Let n be fixed, $d = d(D) \leq D$ be such that $\lim_{D\to\infty} d(D) \to \infty$, and $p = p(D)$. Then $f(d) = d^{-n}$ is a threshold function for the property that $|\mathfrak{B}_d| = 0$.

Proof. For part a), we must show that when $p(n) = \mathrm{o}\left(n^{-d}\right)$, then $|\mathfrak{B}_d| = 0$ with probability tending to 1, and when $p(n) = \omega\left(n^{-D}\right)$, then $|\mathfrak{B}_d| > 0$ with probability tending to 1.

For d fixed, the number of monomials of degree at most d in n variables is $\binom{n+d}{d} \asymp n^d$ (by considering the usual asymptotic bounds for binomial coefficients). For any $\alpha \in \mathbb{Z}^n_{\geq 0}$, with $0 < ||\alpha||_1 \leq d$, let $\mathbf{1}_\alpha$ be the indicator random variable for the event that $x^\alpha \in \mathfrak{B}$, so that $|\mathfrak{B}_d| = \sum_\alpha \mathbf{1}_\alpha$, and therefore $\mathbb{E}\left[|\mathfrak{B}_d|\right] = \sum_\alpha \mathbb{E}\left[\mathbf{1}_\alpha\right] = p \cdot \left(\binom{n+d}{d} - 1\right)$. By the first moment method:

$$\mathrm{P}\left(|\mathfrak{B}_d| > 0\right) \leq \mathbb{E}\left[|\mathfrak{B}_d|\right] = p \cdot \left(\binom{n+d}{d} - 1\right) \leq pe^d n^d.$$

If $p(n) = \mathrm{o}\left(n^{-d}\right)$, then $pe^d n^d \to 0$ as $n \to \infty$ and hence $\lim_{n\to\infty} \mathrm{P}\left(|\mathfrak{B}_d| > 0\right) = 0$.

Now suppose $p(n) = \omega\left(n^{-d}\right)$. In this case $\mathbb{E}\left[|\mathfrak{B}_d|\right] \to \infty$, since:

$$\mathbb{E}\left[|\mathfrak{B}_d|\right] = p \cdot \left(\binom{n+d}{d} - 1\right) \geq \left(\frac{2}{d}\right)^d \cdot pn^d.$$

The variance of $|\mathfrak{B}_d|$ is calculated as follows:

$$\mathrm{Var}\left[|\mathfrak{B}_d|\right] = \sum_\alpha \mathrm{Var}\left[\mathbf{1}_\alpha\right] + \sum_{\alpha\neq\alpha'} \mathrm{Cov}\left[\mathbf{1}_\alpha, \mathbf{1}_{\alpha'}\right] = \left(\binom{n+d}{d} - 1\right) \cdot p \cdot (1-p) = \mathbb{E}\left[|\mathfrak{B}_d|\right] \cdot (1-p),$$

where we have again used that $\mathrm{Var}\left[\mathbf{1}_\alpha\right] = \mathbb{E}\left[\mathbf{1}_\alpha^2\right] - \mathbb{E}\left[\mathbf{1}_\alpha\right]^2 = p - p^2$ and for $\alpha \neq \alpha'$, $\mathbf{1}_\alpha$ and $\mathbf{1}_{\alpha'}$ are independent and thus have 0 covariance. Then by the second moment method and the fact that $\mathbb{E}\left[|\mathfrak{B}_d|\right] \to \infty$ when $p = \omega\left(n^{-d}\right)$:

$$\mathrm{P}\left(|\mathfrak{B}_d| = 0\right) \leq \frac{\mathrm{Var}\left[|\mathfrak{B}_d|\right]}{\left(\mathbb{E}\left[|\mathfrak{B}_d|\right]\right)^2} = \frac{1-p}{\mathbb{E}\left[|\mathfrak{B}_d|\right]} \leq \frac{1}{\mathbb{E}\left[|\mathfrak{B}_d|\right]} \to 0 \text{ as } n \to \infty.$$

Part b) follows in a similar fashion, once one notices that, for fixed n, $\binom{n+d}{d} = \binom{n+d}{n} \asymp d^n$ as d tends to infinity. □

From Lemma 4.3 we immediately obtain a threshold result for the first graded Betti numbers:

Theorem 4.4. *Let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$.*

a) Let $d \leq D$ be fixed and $p = p(n)$. The function $f(n) = n^{-d}$ is a threshold function for the property that $\mathrm{initdeg}(\mathfrak{I}) > d$.

b) Let n be fixed, $d = d(D) \leq D$ be such that $\lim_{D\to\infty} d(D) = \infty$, and $p = p(D)$. Then $f(d) = d^{-n}$ is a threshold function for the property that $\mathrm{initdeg}(\mathfrak{I}) < d$.

Proof. By definition, $\mathrm{initdeg}(\mathfrak{I}) > d$ if and only if $\sum_{k=1}^d \beta_{1,k}(S/\mathfrak{I}) = 0$. Equivalently, $\mathrm{initdeg}(\mathfrak{I}) > d$ if and only if $\mathfrak{B} \neq \emptyset$ and $\mathfrak{B}$ contains no monomials of degree k for each $1 \leq k \leq d$ and therefore

$$P\left(\sum_{k=1}^d \beta_{1,k}(S/\mathfrak{I}) = 0\right) = P\left(|\mathfrak{B}_d| = 0\right).$$

Both a) and b) now follow directly from Lemma 4.3. □

The case $d = D$ of Theorem 4.4 deserves special mention, as it provides a threshold function for the property that the random ideal generated by the ER-type model is the zero ideal.

Corollary 4.5. *Let $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $\mathfrak{I} = \langle \mathfrak{B} \rangle$.*

a) Let D be fixed and $p = p(n)$. Then $f(n) = n^{-D}$ is a threshold function for the property that $\mathfrak{I} = (0)$.
b) Let n be fixed and $p = p(D)$. Then $f(D) = D^{-n}$ is a threshold function for the property that $\mathfrak{I} = (0)$.

On the other hand, in the critical region of this threshold function, both the expected number of monomials in $\mathfrak{B}$ and the variance of the number of monomials in $\mathfrak{B}$ is constant.

Proposition 4.6. *Let $\mathfrak{B} \sim \mathcal{B}(n, D, p)$.*

a) For any n, D, if $p = 1/\left(\binom{n+D}{D} - 1\right)$ then $\mathbb{E}\left[|\mathfrak{B}|\right] = 1$ and $\mathrm{Var}\left[|\mathfrak{B}|\right] = 1$.
b) For D fixed, if $p = p(n) \asymp n^{-D}$ then $k_1 \leq \lim_{n\to\infty} \mathbb{E}\left[|\mathfrak{B}|\right] \leq k_2$ and $c_1 \leq \lim_{n\to\infty} \mathrm{Var}\left[|\mathfrak{B}|\right] \leq c_2$, for some constants k_1, k_2, c_1, c_2.
c) For n fixed, if $p = p(D) \asymp D^{-n}$ then $k_1 \leq \lim_{D\to\infty} \mathbb{E}\left[|\mathfrak{B}|\right] \leq k_2$ and $c_1 \leq \lim_{D\to\infty} \mathrm{Var}\left[|\mathfrak{B}|\right] \leq c_2$, for some constants k_1, k_2, c_1, c_2.

Proof. Since $\mathfrak{B} = \mathfrak{B}_D$, each of these claims follows directly from the calculated values of $\mathbb{E}\left[|\mathfrak{B}_D|\right]$ and $\mathrm{Var}\left[|\mathfrak{B}_D|\right]$ from the proof of Lemma 4.3. □

By combining threshold functions from Theorem 4.4 we can obtain Corollary 1.3 announced in the Introduction: the result specifies ranges of p for which $\mathrm{initdeg}(\mathfrak{I}) = d$ asymptotically almost surely. The argument is identical to the one used in Section 3.3 to obtain Corollary 1.2 from Theorem 3.4.

4.3. Degree complexity

In the previous subsection, we saw how the choice of p influences the smallest degree of a minimal generator of a random monomial ideal. In this subsection, we ask the complementary question: what can one say about the *largest* degree of a minimal generator of a random monomial ideal? The following theorem establishes an asymptotic bound for the degree complexity $\mathrm{reg}_0(\mathfrak{I})$ of $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ for certain choices of probability parameter p.

Theorem 4.7. *Let n be fixed, $\mathfrak{I} \sim \mathcal{I}(n, D, p)$, $p = p(D)$ and let $r = r(D)$ be a function tending to infinity as $D \to \infty$. If $p = \omega\left(\frac{1}{r}\right)$, then $\mathrm{reg}_0(\mathfrak{I}) \leq nr$ a.a.s.*

Proof. By the proof of Theorem 3.5, for each variable x_i, a.a.s. $\mathfrak{B}$ will contain a monomial of the form x_i^j where $j \leq r$. If $j_1, \ldots, j_n$ are integers such that $1 \leq j_i \leq r$ for each $1 \leq i \leq n$, then the ideal $(x_1^{j_1}, \ldots, x_n^{j_n})$ contains all monomials of degree $\geq rn$ in S, since if $\deg(x^\alpha) \geq r$, by the pigeonhole principle at least one $\alpha_i \geq \lfloor \frac{r}{n} \rfloor$. Hence, with probability tending to 1, $\mathfrak{I}$ will contain every monomial in S of degree nr and thus every minimal generator has degree at most nr. □

To illustrate this result, suppose that $p = \omega\left(1/\log D\right)$ and let $\mathfrak{I} \sim \mathcal{I}(n, D, p)$. Then, by Theorem 4.7, as D tends to infinity, one should expect $\mathrm{reg}_0(\mathfrak{I})$ to be at most $n \log D$. One should keep in mind that $\mathfrak{I}$ never contains generators of degree larger than D, and so the bound given in this theorem is not optimal for certain choices of r.

5. Other probabilistic models

As mentioned in the Introduction, the study of 'typical' ideals from a family of interest may require not only the ER-type model, but also more general models of random monomial ideals. To that end, we define the most general probabilistic model on sets of monomials which, a priori, provides no structure, but it does provide a framework within which other models can be recovered.

A general model. Fix a degree bound $D > 0$. To *each* monomial $x^\alpha \in k[x_1, \ldots, x_n]$ with $0 < \deg(x^\alpha) \leq D$, the *general model* for random monomials assigns an arbitrary probability $0 \leq p_\alpha \leq 1$ of selecting the monomial x^α:

$$\mathrm{P}\left(x^\alpha\right) = p_\alpha. \tag{5.1}$$

Hence the general model has many parameters, namely $\{p_\alpha : \alpha \in \mathbb{N}^n \setminus \{0\}, |\alpha| \leq D\}$. It is clear that that, for a fixed n and fixed degree bound D, the ER-type model is a special case of the general model, where $p_\alpha = p(n, D)$, for all $\alpha \in \mathbb{N}^n$ such that $0 < |\alpha| \leq D$, does not depend on α.

Of course, there are many other interesting models one can consider. We define another natural extension of ER-type model.

The graded model. Fix a degree bound $D > 0$. The *graded model* for random monomials places the following probability on each monomial $x^\alpha \in k[x_1, \ldots, x_n]$ with $0 < \deg(x^\alpha) \leq D$:

$$p_\alpha = f(n, D, |\alpha|), \tag{5.2}$$

where $0 \leq p_\alpha \leq 1$. Note that, given $\deg(x^\alpha) = |\alpha|$, the probability of each monomial is otherwise constant in α. The graded model is a D-parameter family of probability distributions on random sets of monomials that induces a distribution on monomial ideals in the same natural way as the ER-type model. Namely, the probability model selects random monomials and they are added to a generating set with probability according to the model, thus producing a random monomial ideal in $k[x_1, \ldots, x_n]$ whose generators are of degree at most D. More precisely, let $\mathbf{p}(n, D) = (p_1(n, D), \ldots, p_D(n, D))$ be such that $0 \leq p_i(n, D) < 1$ for all $1 \leq i \leq D$. Initialize $B = \emptyset$ and for each monomial x^α of degree i, $1 \leq i \leq D$, add x^α to B with probability $p_i(n, D)$. The resulting random monomial ideal is $I = (B)$ (again with the convention that if $B = \emptyset$, then we set $I = (0)$). Denote by $\mathcal{G}(n, D, \mathbf{p})$ the resulting induced distribution on monomial ideals. As before, since ideals are now random variables, we will write $\mathfrak{I} \sim \mathcal{G}(n, D, \mathbf{p})$ for a random monomial ideal generated using the graded model.

The analogue of Theorem 1.1 for $\mathcal{G}(n, D, \mathbf{p})$, which has an almost identical proof, is as follows:

Theorem 5.1. *Fix n, D, and the graded model parameters $\mathbf{p}(n, D) = (p_1, \ldots, p_D)$. For any fixed monomial ideal $I \subseteq S$, random monomial ideals from the graded model distribution $\mathfrak{I} \sim \mathcal{G}(n, D, \mathbf{p})$ satisfy the following:*

$$P(\mathfrak{I} = I) = \prod_{d=1}^{D} p_d^{\beta_{1,d}(S/I)} (1 - p_d)^{h_I(d)},$$

where $\beta_{1,d}(S/I)$ is the number of degree-d minimal generators of I (that is, the first graded Betti number of S/I), and $h_I(d)$ is its Hilbert function.

5.1. Random monomial ideals generalize random simplicial complexes

An *abstract simplicial complex* on the vertex set $[n]$ is a collection of subsets (called *faces*) of $[n]$ closed under the operation of taking subsets. The *dimension* $\dim(\sigma)$ of a face $\sigma \subseteq [n]$ is $|\sigma| - 1$. If Y is an

abstract simplicial complex on $[n]$, then the *dimension* $\dim(Y)$ of Y is $\max\{\dim(\sigma) : \sigma \text{ is a face of } Y\}$ and the *Stanley-Reisner ideal* I_Y in the polynomial ring $k[x_1, \ldots, x_n]$ is the square-free monomial ideal generated by the monomials corresponding to the non-faces τ of Y: $I_Y := \langle \tau : \tau \subseteq [n], \tau \notin Y \rangle$. For example, if $n = 2$ and $Y = \{\emptyset, \{1\}, \{2\}\}$, then $I_Y = \langle x_1 x_2 \rangle \subset k[x_1, x_2]$. A fundamental result in combinatorial commutative algebra is that the Stanley-Reisner correspondence constitutes a bijection between abstract simplicial complexes on $[n]$ and the square-free monomial ideals in $k[x_1, \ldots, x_n]$. See, for example, [33, Theorem 1.7].

In [11], Costa and Farber introduced a model for generating abstract simplicial complexes at random. Their model, which we call the *Costa-Farber model*, is a probability distribution on $\Delta_n^{(r)}$, the set of all abstract simplicial complexes on $[n]$ of dimension at most r. In it, one selects a vector $\tilde{\mathbf{p}} = (\tilde{p}_0, \ldots, \tilde{p}_r, 0, \ldots, 0)$ of probabilities such that $0 \leq \tilde{p}_l \leq 1$ for all $0 \leq l \leq r$. Then, one retains each of the n vertices with probability $\tilde{p}_0$, and for each pair ij of remaining vertices, one forms the edge ij with probability $\tilde{p}_1$, and for each triple of vertices ijk that are pairwise joined by an edge, one forms the 2 dimensional face ijk with probability $\tilde{p}_2$, and so on.

One can study the combinatorial and topological properties of the random simplicial complexes generated by the Costa-Farber model. In view of the Stanley-Reisner correspondence, this model can be seen as a model for generating random square-free monomial ideals. Thus, the general model for generating random monomial ideals can be viewed as a generalization of the Costa-Farber model, which in turn has been proved to generalize many other models of random combinatorial objects, for example, the Erdős-Rényi model for random graphs and Kahle's model for random clique complexes [27]. See [11, Section 2.3] for details.

The relationship between our models for random monomial ideals and the Costa-Farber model can be made precise in the following way: there exists a choice of parameters p_α in (5.1) such that the resulting distribution on square-free monomial ideals in $S = k[x_1, \ldots, x_n]$ is precisely the distribution on the abstract simplicial complexes on $[n]$ under the Costa-Farber model.

Theorem 5.2. *Let $\tilde{\mathbf{p}} = (\tilde{p}_0, \tilde{p}_1, \ldots, \tilde{p}_r, 0, \ldots, 0)$ denote the n-vector of probabilities in the Costa-Farber model for random simplicial complexes. Let $Y \subset \Delta_n^{(r)}$ be a simplicial complex on $[n]$ of dimension at most r and let I_Y be the Stanley-Reisner ideal corresponding to Y. Fix $D = r + 1$ and specify the following probabilities p_α, where $\alpha \in \mathbb{N}^n$ and $0 < ||\alpha||_1 \leq r + 1$, for the general monomial model (5.1):*

$$p_\alpha = \begin{cases} 1 - \tilde{p}_{\deg(x^\alpha)-1}, & \text{if } 0 \neq \alpha \in \{0,1\}^n, \\ 0, & \text{otherwise}. \end{cases} \tag{5.3}$$

Then, $P_{CF}(Y) = P(I_Y)$, where the former is probability under the Costa-Farber model and the latter under the distribution on random monomial ideals induced by the general model (5.1).

In other words, the specification of probabilities in Theorem 5.2 recovers the Costa-Farber model on random simplicial complexes as a sub-model of the model (5.1) on random monomial ideals. Note that this specification of probabilities can be considered as an instance of the graded model with support restricted to the vertices of the unit hypercube.

Proof. From [11, Equation (1)], the following probability holds under the Costa-Farber model:

$$P_{CF}(Y) = \prod_{i=0}^{r} \tilde{p}_i^{f_i(Y)} (1 - \tilde{p}_i)^{e_i(Y)},$$

where $f_i(Y)$ denotes the number of i-dimensional faces of Y and $e_i(Y)$ denotes the number of i-dimensional *minimal* non-faces of Y (i.e., the number of i-dimensional non-faces of Y that do not

strictly contain another non-face). The minimal non-faces of Y correspond exactly to the minimal generators of the Stanley-Reisner ideal I_Y. Thus, I_Y has exactly $e_i(Y)$ minimal generators of degree $i+1$, that is, $e_i(Y) = \beta_{1,i+1}(S/I_Y)$. Each i-dimensional face of Y corresponds to a degree $i+1$ standard square-free monomial of I_Y, hence we have a Hilbert function value$h_{I_Y}(i+1) = f_i(Y)$. Next, note that the specification of probabilities in (5.3) depends only on the degree of each monomial, so that if $\deg(x^\alpha) = \deg(x^{\alpha'})$, then $p_\alpha = p_{\alpha'}$. Hence, for each $1 \leq j \leq r+1$, we denote by p_j the probability assigned to the degree j monomials in (5.3), so $p_j = 1 - \tilde{p}_{j-1}$. We now apply Theorem 5.1 to conclude that

$$P(I_Y) = \prod_{j=1}^{r+1} p_j^{\beta_{1,j}} (1-p_j)^{h_{I_Y}(j)} = \prod_{i=0}^{r} (1-\tilde{p}_i)^{e_i(Y)} \tilde{p}_i^{f_i(Y)} = P_{CF}(Y), \text{ as desired.}$$

□

Simplicial complex Y	Ideal I_Y	Faces of Y	Non-faces of Y	$P_{\mathrm{CF}}(Y)$	$P(I_Y)$
Void	$k[x_1, x_2]$	none	$\emptyset,\ x_1,\ x_2,\ x_1x_2$	0	0
$\emptyset$	(x_1, x_2)	$\emptyset$	$x_1,\ x_2,\ x_1x_2$	$(1-\tilde{p}_0)^2$	$p_{(1,0)}p_{(0,1)}$
1	(x_2)	$\emptyset,\ x_1$	$x_2,\ x_1x_2$	$\tilde{p}_0(1-\tilde{p}_0)$	$p_{(0,1)}(1-p_{(1,0)})$
2	(x_1)	$\emptyset,\ x_2$	$x_1,\ x_1x_2$	$\tilde{p}_0(1-\tilde{p}_0)$	$p_{(1,0)}(1-p_{(0,1)})$
1 2	(x_1x_2)	$\emptyset,\ x_1,\ x_2$	x_1x_2	$\tilde{p}_0^2(1-\tilde{p}_1)$	$(1-p_{(1,0)})(1-p_{(0,1)})p_{(1,1)}$
1 — 2	(0)	$\emptyset,\ x_1,\ x_2,\ x_1x_2$	none	$\tilde{p}_0^2\tilde{p}_1$	$(1-p_{(1,0)})(1-p_{(0,1)})(1-p_{(1,1)})$

Table 1

A table illustrating the correspondence between random monomial ideals and the Costa-Farber model

Table 5.1 illustrates Theorem 5.2 in the case $n = 2$ and $r = 1$. The Costa-Farber model generates a random simplicial complex on $\{1, 2\}$ by specifying a probability vector $\tilde{\mathbf{p}} = (\tilde{p}_0, \tilde{p}_1)$ and starting with two vertices x_1 and x_2. We retain each vertex independently with probability $\tilde{p}_0$. In the event that we retain both x_1 and x_2, we connect them with probability $\tilde{p}_1$. By Theorem 5.2, to recover this model from Equation (5.1) we set $D = 2$ and $p_{(2,0)} = p_{(0,2)} = 0$, $p_{(1,0)} = p_{(0,1)} = 1 - \tilde{p}_0$, and $p_{(1,1)} = 1 - \tilde{p}_1$. The probability distribution of the set of all simplicial complexes Y on $\{1, 2\}$ and, equivalently, all square-free monomial ideals $I_Y \subseteq k[x_1, x_2]$ appear in the table.

To conclude, note that there are other models that do not sample individual monomials, but instead sample ensembles of monomials all at the same time. Examples of these are the methods to generate random integer partitions and Ferrers diagrams [36] and random lattice polytopes [3].

6. Experiments & conjectures

This section collects experimental work with two purposes in mind: illustrating the theorems we proved in this paper, and also motivating further conjectures about the ER-type model. These experiments were performed using `Macaulay2` [22] and `SageMath` [37], including the `MPFR` [21] and `NumPy` [2] modules. Graphics were created using `SageMath`.

We ran several sets of experiments using the ER-type model. All experiments considered varying number of variables n, varying maximum degree D, and varying probability p. For each choice of a triple (n, D, p), we generated a sample of $N = 1000$ monomial ideals. We then computed some

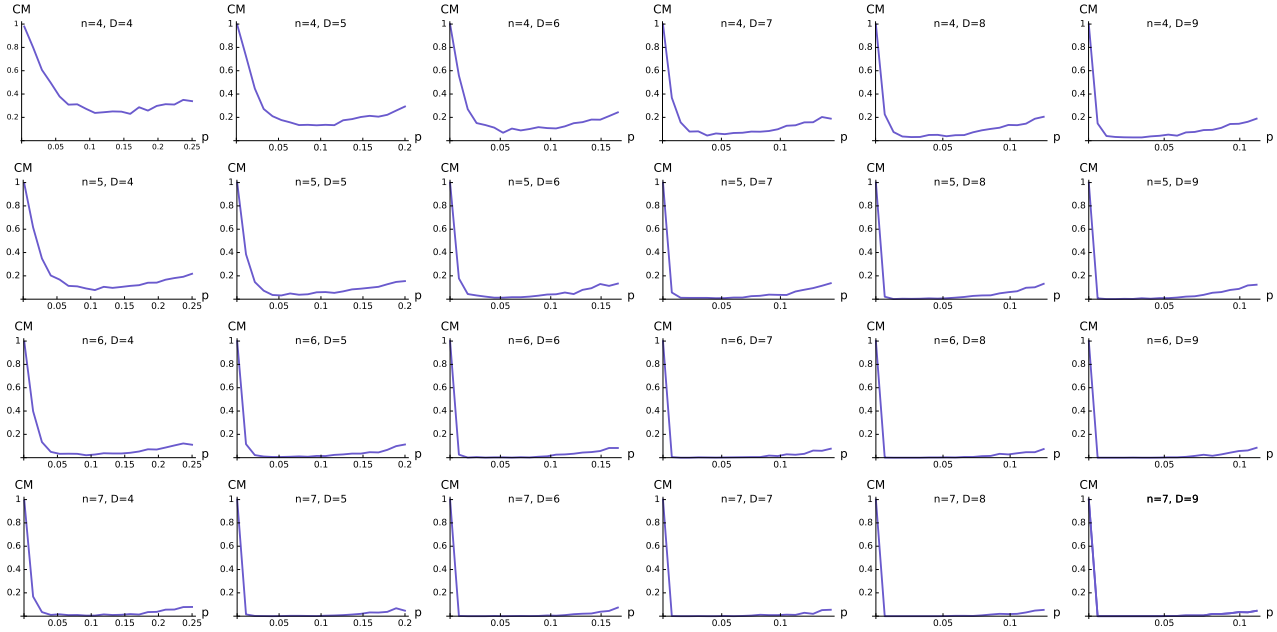

FIG 3. *Frequency of random monomial ideals $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ for which $S/\mathfrak{I}$ is Cohen-Macaulay for $(n, D) = (4, 4)$ through $(n, D) = (7, 9)$, as the model parameter p takes values between D^{-n} and D^{-1}. The sample size for each fixed value of p is* 1000.

algebraic properties and tabulated the results. The figures in this section summarize the results and support the conjectures we state.

Cohen-Macaulayness. Recall that an ideal has an (arithmetically) *Cohen-Macaulay* quotient ring if its depth equals its Krull dimension. As is well-known, Cohen-Macaulayness is a very special property from which one derives various special results; see [8] for a standard reference.

Figure 3 shows the percent of ideals in the random samples whose quotient rings are arithmetically Cohen-Macaulay for $n = 4, \ldots, 7$, $D = 4, \ldots, 9$. We see that, as n and D get larger, very few Cohen-Macaulay ideals are generated, suggesting that Cohen-Macaulayness is a "rare" property in a meaningful sense. We speculate that the appearances of Cohen-Macaulay ideals are largely or entirely due to the appearance of the zero ideal and of zero-dimensional ideals, both of which are trivially Cohen-Macaulay. For the smallest n and D values, the zero ideal appears with observable frequency for all p values, as do zero-dimensional ideals. As p decreases toward the zero ideal threshold or increases toward the zero-dimensional threshold, these probabilities grow, resulting in a "U" shape. (Note that probabilities continue to increase to 1 for larger p beyond the domain of these plots.) However, as n and D grow, the thresholds for the zero ideal and for zero-dimensionality become pronounced, and even as p approaches either threshold the frequency of Cohen-Macaulayness stays low. When p is so low that only the zero ideal is generated, Cohen-Macaulayness is of course observed with probability 1, but as soon as nontrivial ideals are generated, the frequency plummets to near zero. These experimental results suggest the following conjecture:

Conjecture 1. *For large n and D, the only Cohen-Macaulay ideals generated by the ER-type model are the trivial cases (the zero ideal or zero-dimensional ideals), with probability approaching* 1. *In particular, using the threshold functions of Theorems 3.4 and 4.4, we conjecture that as n goes to infinity, the probability that $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ is Cohen-Macaulay will go to zero for $p = p(D)$ satisfying*

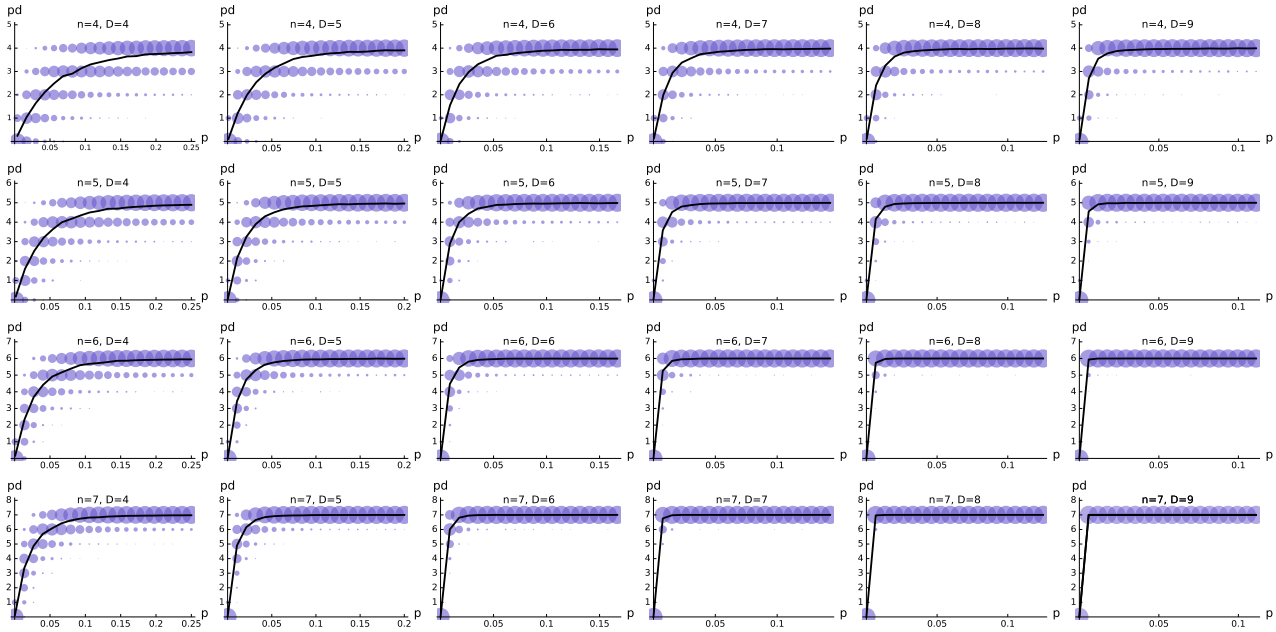

FIG 4. *Projective dimension of $S/\mathfrak{I}$ for random monomial ideals $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ for $(n, D) = (4, 4)$ through $(n, D) = (7, 9)$. The values of the model parameter p are 20 evenly spaced values between D^{-n} and D^{-1}, and the sample size for each value of p is* 1000*. For each p, circle sizes indicate the proportion of quotient rings with that projective dimension, while the black curve indicates the mean projective dimension.*

$p(D) = \mathrm{o}\left(D^{-1}\right)$ *and* $p(D) = \omega\left(D^{-n}\right)$.

Projective Dimension. Further exploring the complexity of minimal free resolutions, and how much the ranges of Betti numbers vary, we investigate the *projective dimension* of $S/\mathfrak{I}$; i.e., the length of the minimal free resolution of $S/\mathfrak{I}$.

By the Hilbert syzygy theorem, the projective dimension is at most n. We see experimentally that for large n and D, the projective dimension of S modulo any non-zero random ideal tends toward this upper bound. (Since $\mathrm{pd}(S/\mathfrak{I}) = 0$ if and only if $\mathfrak{I} = (0)$, the projective dimension will always concentrate at 0 below the $p = D^{-n}$ threshold for the zero ideal.)

Conjecture 2. *As n and D increase, the projective dimension of the quotient ring of any non-zero ideal in the ER-type model is equal to n with probability approaching 1.*

Note that Conjecture 2 implies Conjecture 1. By the Auslander-Buchsbaum formula ([8], see also [38]), $\mathrm{pd}(S/\mathfrak{I}) = n - \mathrm{depth}(S/\mathfrak{I})$ for all $\mathfrak{I}$. Hence Conjecture 2 implies $\mathrm{depth}(S/\mathfrak{I}) \to 0$, which implies that the quotient ring of a nonzero ideal will be Cohen-Macaulay if and only if it is zero-dimensional.

Strong genericity. A monomial ideal $I \subset k[x_1, \ldots, x_n]$ is said to be *strongly generic* if no two minimal generators agree on a non-zero exponent of the same variable. For example, $\left(x_1x_2, x_1^2x_3\right)$ is a strongly generic monomial ideal in $k[x_1, x_2, x_3]$, but (x_1x_2, x_1x_3) is not. Strong genericity is interesting because the minimal free resolution of $S/\mathfrak{I}$ has a combinatorial interpretation using the *Scarf polyhedral complex* (see [33, Chapter 6] for details.), when $\mathfrak{I}$ is strongly generic.

The zero ideal, with no generators, is trivially strongly generic, as is any principal ideal. At first it might seem that increasing p, and thus increasing the expected number of monomials included in $\mathfrak{I}$, would always lower the frequency of strong genericity. However at the other extreme of $p = 1 - \epsilon$,

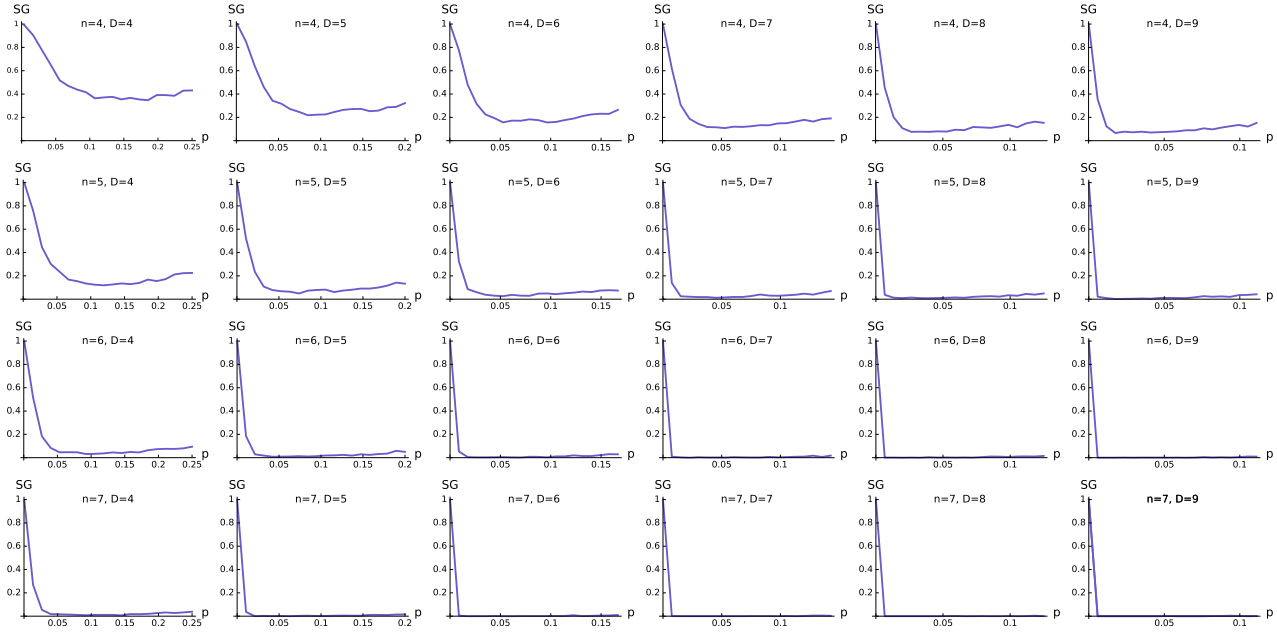

FIG 5. *Frequency of strong genericity in the ER-type model, for* $(n, D) = (4, 4)$ *through* $(n, D) = (7, 9)$*, as the model parameter* p *takes values between* D^{-n} *and* D^{-1}*. The sample size for each fixed value of* p *is* 1000*.*

the maximal ideal $(x_1, x_2, \ldots, x_n)$ occurs frequently, and this ideal is strongly generic. Between these extremal cases, we observe experimentally that very few random monomial ideals are strongly generic, and suspect a connection between this behavior and our results in Section 4 about the number of minimal generators.

Conjecture 3. *As* n *and* D *increase, there will be a lower threshold function* $l(n, D)$ *and an upper threshold function* $u(n, D)$*, such that the probability of being strongly generic will go to zero for* $p = \mathrm{o}\left(u(n, D)\right)$ *and* $p = \omega\left(l(n, D)\right)$*, while the probability of being strongly generic will go to one for* $p = \mathrm{o}\left(l(n, D)\right)$ *as well as for* $p = \omega\left(u(n, D)\right)$*.*

Note that because the plots in Figure 5 display p values between D^{-n} and D^{-1}, the behavior as p approaches 1 is not visible.

We now turn to several properties of monomial ideals for which we experimentally observed noticeable patterns, but for which we do not have explicit conjectures to present.

Castelnuovo-Mumford regularity. The *Castelnuovo-Mumford regularity* (or simply regularity) of an ideal is a measure of the variation and spread of the the degrees of the generators at each of the modules in a minimal free resolution of the ideal. To compute the regularity requires computing the resolutions and the number of rows in the `Macaulay2`-formatted Betti diagram. Of course, the degree complexity studied in Section 4 is similar but less refined measure of complexity, as it counts the number of entries in the first column of the Betti diagram only.

The simulations on the range of values realized for regularity of random monomial ideals, which do not take into account any zero ideals generated (since their regularity is $-\infty$), show an interesting trend. Namely, already for the small values of n and D depicted in Figure 6, there is quite a range of values obtained under the model. As n grows these values concentrate more tightly around their mean, while seeming to follow a binomial distribution (discretized version of a normal distribution).

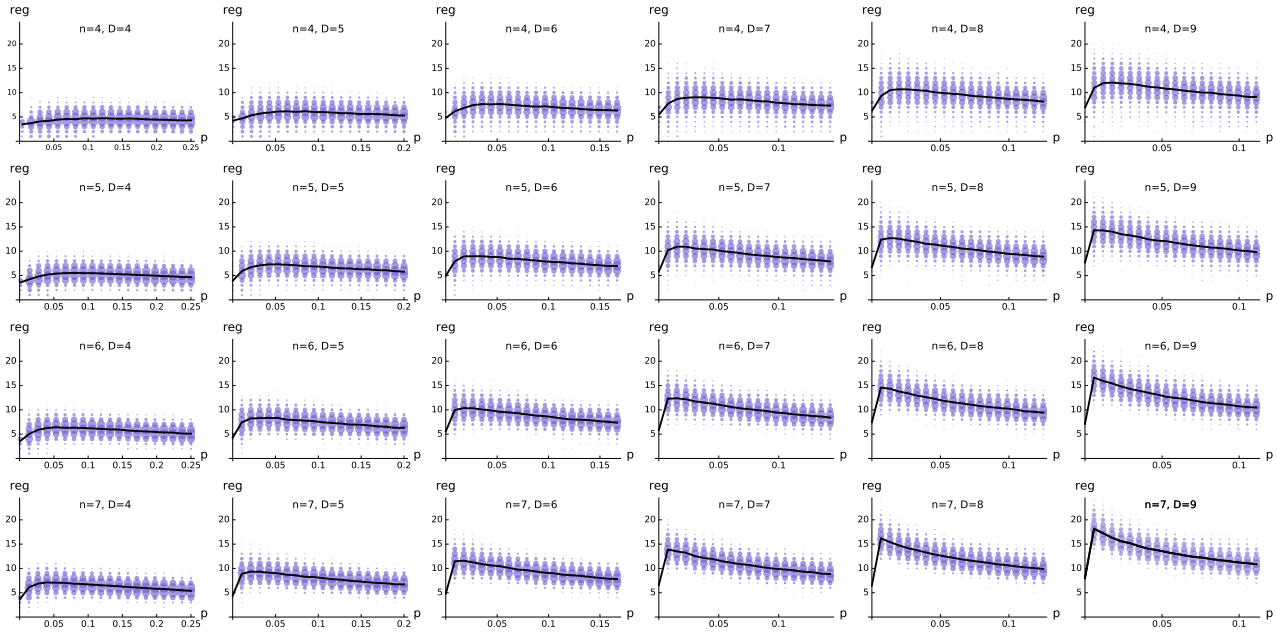

FIG 6. *Castelnuovo-Mumford regularity of random monomial ideals* $\mathfrak{I} \sim \mathcal{I}(n, D, p)$ *for* $(n, D) = (4, 4)$ *through* $(n, D) = (7, 9)$. *Note that* n *increases across rows and* D *increases down columns, and that the values of the model parameter* p *are 20 evenly spaced values between* D^{-n} *and* D^{-1}. *The sample size for each fixed value of* p *is* 1000. *Circle sizes indicate the relative frequencies of occurrence of random monomial ideals of corresponding regularity, and the mean regularity for each* p *value is indicated by the black curves.*

Simplicial homology of associated simplicial complexes. One motivation for studying random monomial ideals is that they provide a way to generate *random simplicial complexes* different from earlier work, where the authors randomly generate k-dimensional faces of complexes for some fixed integer k (see [11, 27, 31]). Instead, by Stanley-Reisner duality, the indices in a monomial we generate are the elements of a *non-face* of a simplicial complex.

There are two natural ways to randomly generate sets of square-free monomials and our experiments considered both of them. First, as radicals of random monomial ideals drawn from the ER-type model and, second, directly as random square-free monomial ideals drawn from the general model, as described in Theorem 5.2, that places zero probability on non-square-free-monomials and probability p on square-free monomials. In our experiments, given an ideal in n variables, we computed the $\mathbb{Z}_2$-homology H_i, for i from 0 to $n-1$, of the associated simplicial complex. Figure 7 displays the homological properties of the random simplicial complexes obtained via the Stanley-Reisner correspondence from the first approach. Figure 8 displays the homological properties of the random simplicial complexes obtained via the second method, directly generating square-free ideals, avoiding taking the radical.

The patterns of appearance and disappearance of higher homology groups, visible in both figures, are familiar in the world of random topology; see for example [28]. For each i there is a lower threshold below which H_i always vanishes, and an upper threshold above which H_i also vanishes, while between these two thresholds we see a somewhat normal-looking curve. Since our model parameter p controls how many faces are *removed* from a complex, higher dimensional objects are associated with lower values of p: a reversal of the typical behavior in prior random topological models. Another interesting pattern in both sets of experiments (unfortunately not evident in the plots) is that there was an overall tendency for each random simplicial complex to have no more than one non-trivial homology group.

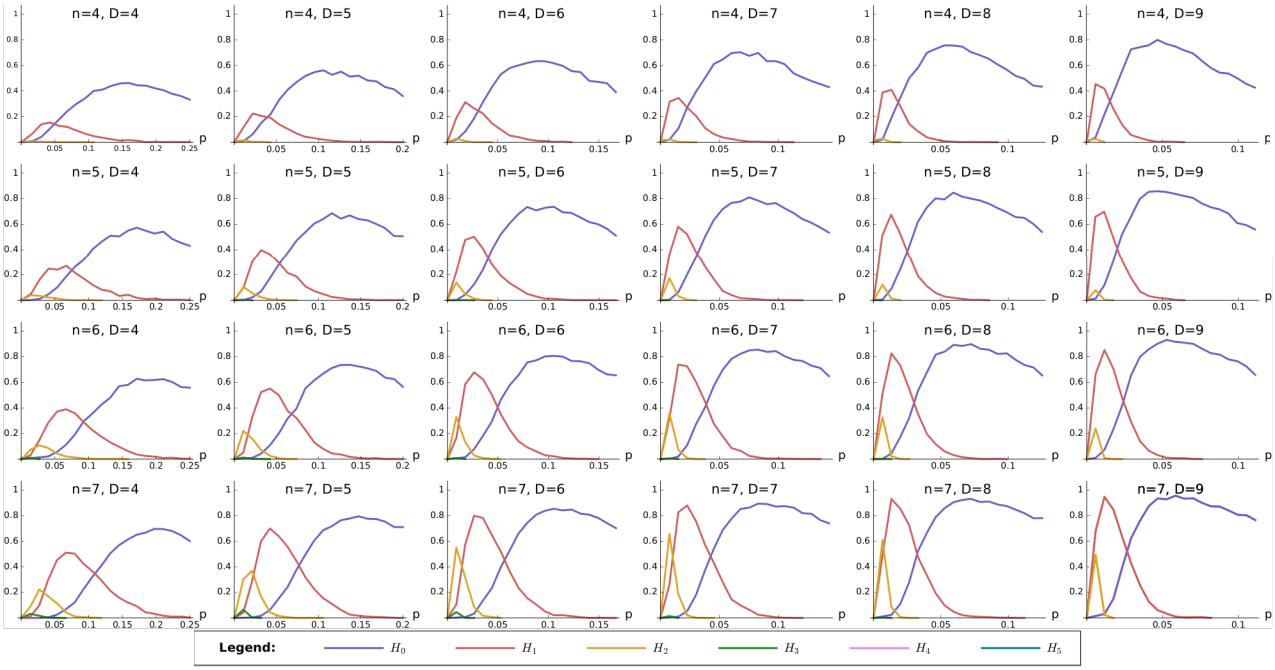

FIG 7. *Non-trivial homology groups for the random simplicial complexes whose nonfaces are radicals of random monomial ideals in the* ***ER-type*** *model setting. Each subplot in Figure 7 represents a fixed pair of* (n, D) *values, with* n *increasing down columns, and* D *increasing across rows. For each* n, D *and* p *value we generated* 1000 *ideals* $\mathrm{rad}(\mathfrak{I} \sim \mathcal{I}(n, D, p))$. *Each subplot registers the frequency at which the second kind of random simplicial complex exhibited* $H_i \neq \{0\}$ *over* $\mathbb{Z}_2$. *The values of* p *in each subplot are chosen in the interval* $[D^{-n}, D^{-1}]$. *We note that* $\mathfrak{I}$ *tends to the maximal ideal as* $p \to 1$, *so every curve will tend to* 0.

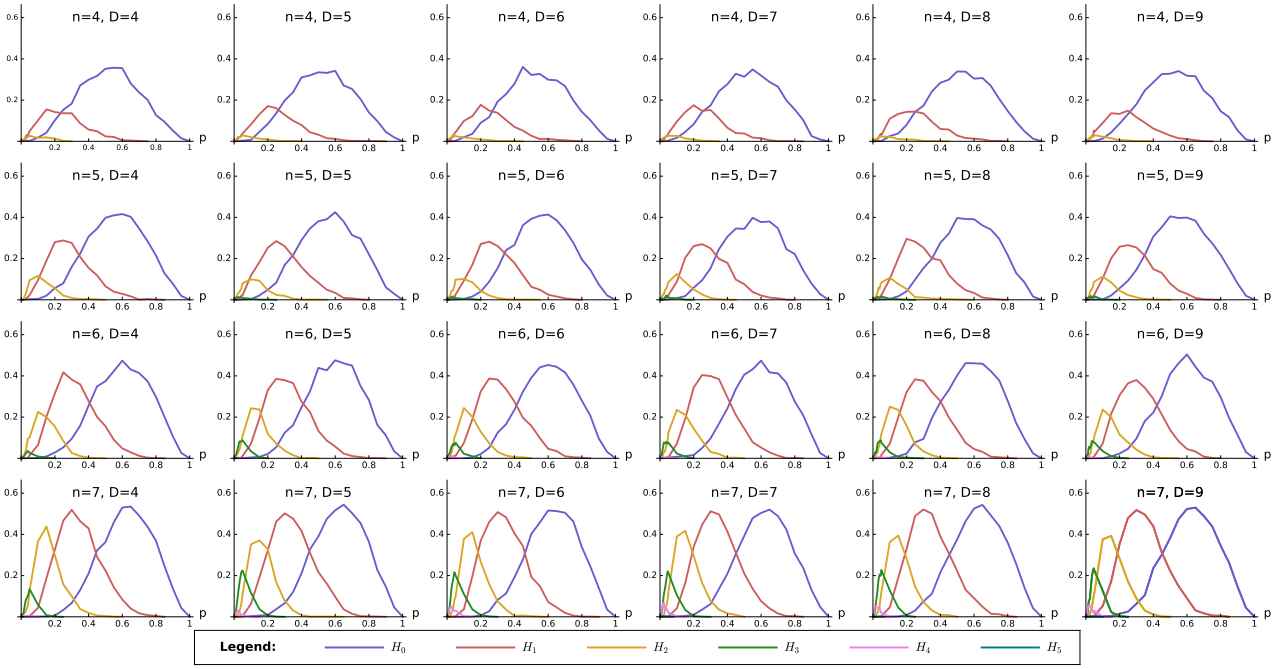

FIG 8. *Non-trivial homology groups for the random simplicial complexes whose non-faces are random square-free monomial ideals. Each subplot is for a pair* (n, D) *values, with* n *increasing down columns, and* D *increasing across rows. With a sample size of 1000 monomial ideals for each* n, D *and* p, *each subplot shows the frequency at which the second kind of random simplicial complex exhibited* $H_i \neq \{0\}$ *over* $\mathbb{Z}_2$. *The* p *values for this set of experiments are* $\{0.05n \mid 0 \leq n \leq 20\}$. *The subplots register the frequency at which these randomly generated simplicial complexes exhibited* $H_i \neq \{0\}$.

That is, if for a particular n, D and p the experiments found that H_0 was nontrivial with frequency x, and H_1 was nontrivial with frequency y, one might expect to see random simplicial complexes with H_0 and H_1 *simultaneously* nontrivial with frequency xy. This would be the case if these events were statistically uncorrelated. However we consistently observed frequencies much lower than this, i.e. a negative correlation between nonzero homology at i and nonzero homology at j, for every $i \neq j$.

7. Acknowledgements

We are grateful for the comments, references and suggestions from Eric Babson, Boris Bukh, Seth Sullivant, Agnes Szanto, Ezra Miller, and Sayan Mukherjee. We are also grateful to REU student Arina Ushakova who did some initial experiments for this project.

References

[1] Alon, N., and Spencer, J. H. *The Probabilistic Method*, 4th ed. Wiley, 2016.

[2] Ascher, D., Dubois, P. F., Hinsen, K., Hugunin, J., and Oliphant, T. *Numerical Python*, ucrl-ma-128569 ed. Lawrence Livermore National Laboratory, Livermore, CA, 1999.

[3] Bárány, I., and Matousek, J. The randomized integer convex hull. *Discrete & Computational Geometry 33*, 1 (2005), 3–25.

[4] Bayer, D., and Mumford, D. What can be computed in algebraic geometry? In *Computational algebraic geometry and commutative algebra* (Apr. 1992), University Press, pp. 1–48.

[5] Beltrán, C., and Pardo, L. M. Smale's 17th problem: average polynomial time to compute affine and projective solutions. *Journal of the AMS 22*, 2 (2009.), 363–385.

[6] Bigatti, A. M. Upper bounds for the Betti numbers of a given Hilbert function. *Commutative Algebra 21* (1993), 2317–2334.

[7] Bollobás, B. *Random graphs*, 2nd ed. Cambridge University Press, 2001.

[8] Bruns, W., and Herzog, J. *Cohen-Macaulay rings*. Cambridge University Press, 1998.

[9] Bürgisser, P., and Cucker, F. On a problem posed by Steve Smale. *Annals of Mathematics 174*, 3 (2011), 1785–1836.

[10] Cornuéjols, G. *Combinatorial Optimization: Packing and Covering*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, 2001.

[11] Costa, A., and Farber, M. Random simplicial complexes. In *Configuration spaces - Geometry, Topology, and Representation Theory*, INdAM Series. Springer, To appear.

[12] Cox, D., Little, J. B., and O'Shea, D. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2007.

[13] De Loera, J. A., Petrović, S., and Stasi, D. Random sampling in computational algebra: Helly numbers and violator spaces. *Journal of Symbolic Computation 77* (2016), 1–15.

[14] Dilcher, K. Some q-series identities related to divisor functions. *Discrete mathematics 145*, 1 (1995), 83–93.

[15] Ein, L., Erman, D., and Lazarsfeld, R. Asymptotics of random Betti tables. *Journal für die reine und angewandte Mathematik (Crelle's Journal)*. (To appear).

[16] Ein, L., and Lazarsfeld, R. Asymptotic syzygies of algebraic varieties. *Inventiones mathematicae 190*, 3 (2012), 603–646.

[17] Eisenbud, D. *Commutative algebra with a view toward algebraic geometry*, vol. 150 of *Graduate Texts in Mathematics*. Springer, 1995.

[18] Eisenbud, D., and Schreyer, F.-O. Betti numbers of graded modules and cohomology of vector bundles. *Journal of the American Mathematical Society 22*, 3 (2009), 859–888.