

Probability and Machine Learning in Combinatorial Commutative Algebra

By

LILY RACHEL SILVERSTEIN

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Professor Jesús De Loera, Chair

Professor Eric Babson

Professor Monica Vazirani

Committee in Charge

2019

We can only see a short distance ahead, but we can see plenty there that needs to be done.

—Alan Turing

My undertaking is not difficult, essentially... I should only have to be immortal to carry it out.

—Jorge Luis Borges

Dedicated to my parents, with all of my love.

Abstract

Many important computations in commutative algebra are known to be NP-hard. Despite their apparent intractability, these algebra problems—including computing the dimension of an algebraic variety, and computing the Hilbert series, projective dimension, and regularity of a homogeneous ideal—are indispensable in both applications and theoretical work. This dissertation advances our understanding of hard commutative algebra problems in several ways.

First, we introduce families of parameterized random models for monomial ideals, and derive the expected values and asymptotic behavior of important algebraic properties of the ideals and their varieties. The surprising result is that many theoretically intractable computations on monomial ideals are easily approximated by simple ratios among number of generators, number of variables, and degrees of generators. Though these approximations are not deterministic, they are guaranteed to hold asymptotically almost surely.

We derive threshold functions in the random models for Krull dimension, (strong) genericity, projective dimension, and Cohen-Macaulayness. In particular, we prove that in a rigorous sense, almost all monomial ideals have the maximum projective dimension possible according to the Hilbert Syzygy Theorem, and that almost no monomial ideals are Cohen-Macaulay. Furthermore, we derive specific parameter ranges in the models for which the minimal free resolution of a monomial ideal can be constructed combinatorially via the algebraic Scarf complex. We also give a threshold for every value of the Krull dimension.

Following recent advances in optimization and computer science, another chapter of this thesis demonstrates how machine learning can be used as a tool in computational commutative algebra. We use supervised machine learning to train a neural network to select the best algorithm to perform a Hilbert series computation, out of a portfolio of options, for each new instance of this problem. We also explore how accurately a neural network can predict NP-hard monomial ideal invariants such as dimension and projective dimension, using features of the ideals that are computable in polynomial time. We provide compelling

evidence that answers to these hard problems can be predicted for new instances based only on the historical data of previously seen examples.

Finally, we implement integer linear programming reformulations of computations on ideals, to take advantage of the sophisticated solving methods now available for this particular class of problems. We demonstrate significant timing improvements in computations such as dimension and degree, especially for large instances of these problems. We define new polytopes useful for enumerative problems in commutative algebra, including enumerating all monomial ideals with a particular Hilbert function, and enumerating the possible Betti tables for a particular Hilbert function.

Acknowledgments

I gratefully acknowledge various sources of funding during my graduate research: Fellowships from the UC Davis math department, generous support from my advisor, including under NSF grants DMS-1522158 and DMS-1522662, and student travel awards from SIAM, the AMS, and the UC Davis Graduate Student Association.

Thank you to the Institute of Computational and Experimental Research in Mathematics for funding me to participate in their Nonlinear Algebra semester in Fall 2018, and to Bernd Sturmfels and the Max Planck Institute for Mathematics in the Sciences for sponsoring a visit in Summer 2018. In both places I pushed forward my dissertation research, and was able to engage with an inspiring research community.

To my advisor, Jesús De Loera: working with you has been one of the greatest privileges of my life. Thank you for your incredible enthusiasm, dedication, and inspiration.

Thank you to Robert Krone, Serkan Hoşten, Sonja Petrović, Despina Stasi, Dane Wilburne, Jay Yang, and Zekai Zhao, for fruitful research collaborations that contributed to this dissertation.

Thank you to Eric Babson and Monica Vazirani for serving on my dissertation committee.

Thank you to Sarah Driver for invaluable assistance with the practical and administrative (and sometimes emotional) aspects of being a graduate student.

And thank you to Steven Glenn Jackson, who gave me my first copy of the bible [CLO07].

Contents

Abstract	iii
Acknowledgments	v
List of Figures	viii
List of Tables	x
Chapter 1. Introduction	1
1.1. Random graphs and simplicial complexes	7
1.2. Ideals and varieties	19
1.3. Resolutions of monomial ideals	27
1.4. Hilbert functions and series	41
1.5. Supervised machine learning	45
Chapter 2. Random models for monomial ideals	53
2.1. The Erdős-Rényi-type model	54
2.2. The graded model	55
2.3. The general model	56
2.4. Specialization to random simplicial complexes and graphs	57
Chapter 3. Krull dimension	59
3.1. Dimension as a vertex cover problem	59
3.2. Krull dimension probabilities in the ER-type model	63
3.3. Krull dimension thresholds in the ER-type model	67
3.4. Fast dimension computations using integer linear programming	69
Chapter 4. Projective dimension	73

4.1. Monomial ideals with large projective dimension	73
4.2. Most monomial ideals have large projective dimension	74
4.3. Cohen-Macaulayness	81
Chapter 5. Genericity and Scarf complexes in the graded model	86
5.1. A threshold for genericity	86
5.2. Almost a threshold for being Scarf	88
Chapter 6. Distributions of Hilbert functions	94
6.1. The De Loera polytope	95
6.2. Enumerating monomial ideals with specified Hilbert functions	98
6.3. Explicit Hilbert functions probabilities	101
6.4. Betti tables with the same Hilbert function	106
Chapter 7. Supervised machine learning in commutative algebra	110
7.1. Monomial ideal features and training data	111
7.2. Algorithm selection for Hilbert series computations	115
7.3. Fast predictions of algebraic invariants	121
Appendix A. Computational details	127
A.1. Code for discrete optimization techniques in commutative algebra	127
A.2. Code and data for machine learning on monomial ideals	128
Bibliography	129

List of Figures

1.1	Summary of thresholds for $\mathfrak{M} \sim \mathcal{M}(n, D, p)$	6
1.2	Erdős-Rényi random graphs	11
1.3	Connectedness of Erdős-Rényi random graphs	15
1.4	Several representations of a simplicial complex	17
1.5	First examples of varieties of polynomial ideals	20
1.6	Stanley-Reisner ideal of a simplicial complex	25
1.7	Taylor complex of a monomial ideal	36
1.8	Scarf complex of a monomial ideal	38
1.9	Staircase diagram of a monomial ideal	43
1.10	The anatomy of a neural network	49
1.11	Activation functions for neural networks	50
3.1	Support hypergraph of a monomial ideal	60
3.2	Reduced support hypergraph of a set of monomials.	62
3.3	Zero-dimensional threshold of ER-type model random monomial ideals	69
3.4	Computing dimension with integer programming	71
3.5	Computing dimension with integer programming II	72
4.1	Geometric interpretation of witnesses to large projective dimension	83
4.2	Pairs of witness lcm's with negative or zero covariance	84
4.3	Intersection types, color classes, the graph H , and the set V	85
5.1	Generic versus Scarf monomial ideals	93
7.1	What is a confusion matrix?	110

7.2	Machine learning dimension	123
7.3	Machine learning projective dimension	124
7.4	Machine learning regularity	126

List of Tables

6.1	Enumerating monomial ideals according to Hilbert function	99
6.2	List of monomial ideals with a common Hilbert function	102
6.3	Distribution of Hilbert functions in $\mathcal{I}(2, 5, p)$	103
6.4	Enumerating Betti tables with a particular Hilbert function	106
7.1	Monomial ideal training data	112
7.2	Feature transform for monomial ideals	116
7.3	Pivot rules in computing Hilbert series	117
7.4	How pivot rule choices affect the complexity of Hilbert series computations	118
7.5	Effectiveness of pivot rule choice for Hilbert series.	120
7.6	Cheap feature transform for monomial ideals	120

CHAPTER 1

Introduction

This is a dissertation about varieties of polynomial ideals, about graphs and simplicial complexes, about phase transitions and asymptotic thresholds, about minimal free resolutions and projective dimension, about integer programming and lattice polytopes, about pivot rules in a Hilbert series algorithm, and about supervised learning on artificial neural networks. Not only are there intimate relationships among the topics in this apparently miscellaneous list, there is also a single, fundamental theme underlying the entire collection. That theme, which permeates every page of this dissertation, is the pursuit of *faster, simpler methods for hard problems in computer algebra*.

To make the notion of a hard problem precise, we begin with some fundamental concepts from computational complexity theory. A *decision problem* in complexity theory is a class of instances, or specific inputs, on which a true-or-false statement can be evaluated. For example, the *subset sum problem* is: given a set of integers, is there a subset of them which sums to zero? A particular instance of this problem is: *Is there a subset of $\{1, -3, 8, -2, 4, -13, 5\}$ that sums to zero?*

The time complexity of an algorithm is measured with respect to the size of an instance. For instance, let n be the number of elements in a set of integers. Then the “brute-force” approach to solving the subset sum problem, which iterates over every possible subset, sums its elements, and then checks if the result is equal to zero, requires 2^n iterations in the worst case. Taking the cost of integer addition to be essentially negligible, the brute-force algorithm takes $\mathcal{O}(2^n)$ (“big oh” of 2^n) steps, which roughly means “no more than a constant multiple of 2^n ” (see Definition 1.1.12). This is an exponential algorithm since the complexity grows exponentially as n does.

As another example, matrix multiplication is the problem that takes two $n \times n$ matrices and computes their product. The standard algorithm taught in any intro linear algebra class uses n^3 multiplications and $n^3 - n^2$ additions; its time complexity is $\mathcal{O}(n^3)$. The complexity of this algorithm is polynomial in the input size. (There are faster algorithms than this one, by the way—see [Lan11].)

The matrix multiplication problem is not a decision problem, but we could state a decision version of it, for instance: is the product of two matrices equal to the zero matrix? If a problem admits a polynomial-time algorithm, then so does the decision version, since we can simply compute the answer in polynomial time, then check for equality.

The complexity class P is the class of all decision problems that admit a polynomial-time (in the input size) algorithm. Another important complexity class is NP, the class of all decision problems for which a *proposed* solution can be *verified* in polynomial time. The subset sum problem belongs to NP, because given a set A of n integers together with a candidate solution $B \subseteq A$, checking whether B sums to zero takes polynomial time.

Every problem in P is also in NP, since an algorithm which quickly solves a problem can also quickly check a solution. Amazingly, whether or not $P=NP$ remains an open question nearly 50 years after it was first formulated precisely [Coo71, GJ79]. There is no intuitive reason why the existence of polynomial-time verification of given solutions should imply a polynomial-time algorithm for *finding* a solution. On the other hand, all that is required for proving $P \neq NP$ is proving that even one problem in NP, like the subset sum problem, cannot admit a polynomial-time algorithm. Despite half a century of focused efforts from the most brilliant minds in computer science, this has never been done.

Deepening the mystery is the notion of NP-completeness. For two decision problems Q_1 and Q_2 , we say that Q_1 *polynomially transforms* to Q_2 if for any instance x of Q_1 , there is an algorithm which produces an instance y of Q_2 , in polynomial time in the size of x , such that the answer to x is yes if and only if the answer to y is yes. (See [PS98, Chapter 15].) A problem Q is *NP-hard* if every problem in NP polynomially transforms to Q . If Q is in NP *and* NP-hard, we say that Q is *NP-complete*. If a polynomial-time algorithm is ever found

for an NP-complete problem Q , this means that *every* other problem in NP can be solved in polynomial time, too. This, also, has never been done, despite 50 years of attempts to efficiently solve many famous problems known to be NP-complete. These include the subset sum problem, as well as:

- The *traveling salesperson* problem: Given the locations of n cities, and the pairwise distances between cities, what is the minimum length of a *tour*, a trip that begins and ends in the same city, and visits every other city exactly once? (Decision version: is there a trip of length $\leq K$?)
- The *minimum vertex cover* problem: Given a graph on n vertices, what is the minimum size of a *vertex cover*, a subset of vertices such that every edge of the graph contains at least one element in the subset? (Decision version: is there a vertex cover of size $\leq K$?)
- The *Boolean satisfiability* problem: Given n Boolean variables, and m logical clauses of the variables together with operations OR, AND, and NOT, is there an assignment of *true* and *false* values to the variables that makes every clause true?

There are many other interesting NP-complete problems; for these and further theory of algorithmic complexity, see [PS98, Kar72, GJ79].

The problems studied in this thesis, concerning computations on ideals in polynomial rings, are all at least as hard as NP-hard problems. One hard problem in commutative algebra is the *ideal membership problem*: given a polynomial $f \in \mathbb{k}[x_1, \dots, x_n]$, and an ideal $I = \langle f_1, \dots, f_r \rangle$, is $f \in I$? Mayr and Meyer famously proved in [MM82] that this problem is EXPSPACE-complete. The class EXPSPACE contains all problems that can be solved with exponential *space complexity*, and strictly contains both P and NP; thus the ideal membership problem is strictly harder than any NP-complete problem. One property of an ideal that relates to computational complexity is its *regularity* (see Definition 1.3.19). The regularity of an ideal I gives a bound on the degrees of the polynomials in a Gröbner basis of I [BS87b], a ubiquitous method for computations with multivariate polynomials. Unfortunately, [MM82] along with [BS88] shows that in the worst case, this *degree complexity* is

double exponential in n , the number of variables of the ring (i.e., an exponential function of an exponential function of n).

In practice, many polynomial computations are reduced to computations on monomial ideals; for instance, computing dimension, degree, and the Hilbert series of an ideal. The *initial ideal* of a polynomial ideal is a monomial ideal, which preserves many fundamental invariants of the original ideal, such as dimension and degree [CLO07]. Computations on an initial ideal provide bounds for other invariants of a polynomial ideal, such as projective dimension and regularity [HH11]. Monomial ideals are the simplest polynomial ideals, with varieties that are always unions of linear subspaces, yet they are general enough to capture the entire range of possible values for many invariants such as the Hilbert series of an ideal [Eis95].

Even for the apparently simpler case of monomial ideals, problems like computing the dimension of a variety are hard. In fact, the decision version of this problem—is the dimension of monomial ideal I no more than K ?—is NP-complete [BS92] (see Section 3.1). Other problems, like finding the Hilbert series of a monomial ideal, computing its projective dimension, or constructing a minimal free resolution, are *at least* as hard.

This dissertation is inspired by probabilistic and computational methods that have been successfully applied to hard problems in other fields of mathematics and computer science. Chapter 2 introduces new families of parameterized random models for monomial ideals, and in Chapters 3 to 6 we prove the expected values and asymptotics of important algebraic properties of random monomial ideals. The techniques in these chapters are similar to those of *probabilistic combinatorics*, notably the classic random graphs of Erdős-Rényi and Gilbert [ER59, Gil59], and more recent work on random simplicial complexes (e.g., [LM06, CF16, Kah09, BK18, BHK11]). One notable model, which first appeared in print in [DPS⁺19], is called the *ER-type model*, because of its resemblance to Erdős-Rényi random graphs:

DEFINITION (Definition 2.1.1). A random monomial ideal $I \sim \mathcal{I}(n, D, p)$ in the polynomial ring $S = \mathbb{k}[x_1, \dots, x_n]$ is produced by randomly selecting its generators independently,

with probability $p = p(n, D) \in (0, 1)$ each, from the set of all monomials in S of positive degree no more than D .

Random monomial ideals give insight into how invariants are distributed. The surprising result is that many theoretically intractable computations on monomial ideals are determined by simple ratios among number of generators, number of variables, and degrees of generators. Though these classifications are not deterministic, they are guaranteed to hold asymptotically almost surely, and give good approximations in small cases. An example is the dimension of ER-type model random monomial ideals:

THEOREM (Theorem 3.3.2). Fix n , so $p(n, D) = p(D)$, and let $I \sim \mathcal{I}(n, D, p)$. For $0 \leq t < n$, if $p = \omega(D^{-t-1})$ and $p = o(D^{-t})$, then $\dim(S/I) = t$ asymptotically almost surely as $D \rightarrow \infty$.

A second parameterized family of random monomial ideals that is important in this thesis is the *graded model*:

DEFINITION (Definition 2.2.3). A random monomial ideal $M \sim \mathcal{M}(n, D, p)$ in the polynomial ring $S = \mathbb{k}[x_1, \dots, x_n]$ is produced by randomly selecting its generators independently, with probability $p = p(n, D) \in (0, 1)$ each, from the set of all monomials in S of total degree *exactly* D .

When a monomial ideal is *generic* (see Section 1.3.6), there is an elegant combinatorial method for computing its minimal free resolution via the (*algebraic*) *Scarf complex* [BPS98, MS04] (see Section 1.3.5). For the graded model, we prove that $p(D) = D^{-n+3/2}$ is a threshold (see Section 1.1.2) for the genericity of a monomial ideal.

THEOREM (Theorem 5.1.1). Let $S = \mathbb{k}[x_1, \dots, x_n]$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, and $p = p(D)$. As $D \rightarrow \infty$, $p = D^{-n+3/2}$ is a threshold for \mathfrak{M} being (strongly) generic. In other words, if $p(D) = o(D^{-n+3/2})$ then \mathfrak{M} is (strongly) generic a.a.s., and if $p(D) = \omega(D^{-n+3/2})$ then \mathfrak{M} is (strongly) generic asymptotically almost surely.

This implies that the Scarf algorithm is correct asymptotically almost surely when $p = o(D^{-n+3/2})$. On the other hand, in Section 5.2 we use combinatorial methods for computing

Betti numbers, especially those developed in [Ale17b], to show that for $p = \omega(D^{-n+2-1/n})$, the Scarf complex of $\mathfrak{M} \sim \mathcal{M}(n, D, p)$ will almost surely be strictly smaller than the minimal free resolution. This and other thresholds for the graded model, many of which appear in [DHKS19], are summarized in Figure 1.1.

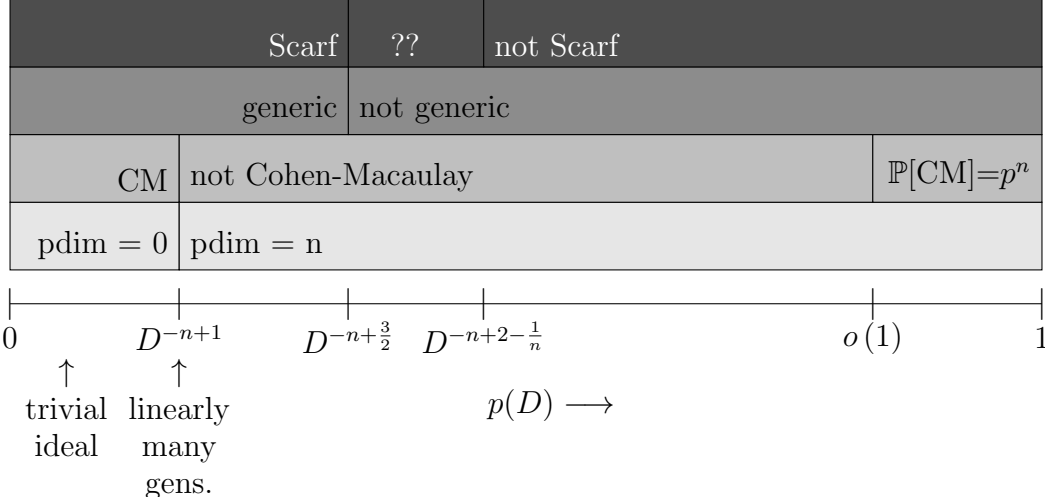


FIGURE 1.1. Typical homological properties and thresholds for random monomial ideals $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, when n is fixed and $D \rightarrow \infty$.

In Section 3.4 and Chapter 6, computations on ideals are reformulated as *integer linear programming* problems to take advantage of the sophisticated solving methods now available for this particular class of problems. Figures 3.4 and 3.5 illustrate the dramatic practical improvements gained by this method, which are currently being developed for incorporation into the Macaulay2 computer algebra system [GS].

In Chapter 7, we demonstrate how machine learning can be used as a tool in computational commutative algebra. by training a neural network to select the best algorithm to perform a Hilbert series computation (see Section 1.4.2), out of a portfolio of options, for each new instance of this problem. The premise of this research is that statistics on algorithm performance, gathered from sufficient samples on some distribution of algebraic inputs, can be used to learn “good choices” when running these algorithms on new inputs. The feasibility of learning algorithm choices has been demonstrated in other areas of math and

computer science. One example is learning optimal heuristics for branch-and-bound methods in integer and mixed-integer optimization [MALW14, KDN⁺17]. In [XHHLB08], Xu et al. used learning to design a “meta” SAT solver that selected, from a portfolio of different SAT solvers, the best choice for each instance of the problem. Researchers in combinatorial optimization have recently outlined some machine learning methodologies designed to apply generally to problems in their field (e.g., [BNVW17, DKZ⁺17, GR15]). The goal of this portion of the dissertation is to adapt these methodologies for computations in commutative algebra and algebraic geometry.

Also in Chapter 7, we explore how accurately a neural network (see Section 1.5.1) can predict NP-hard monomial ideal invariants such as dimension and projective dimension, using as input the same inexpensive features used for the algorithm selection problem. Examples of features (see Section 7.1) are number of minimal generators, average support size of generators, average proportion of generators that each variable divides, etc. In Section 7.3 we see compelling evidence that a neural network can predict the answers to these hard problems based only on a data set of randomly generated examples.

1.1. Random graphs and simplicial complexes

1.1.1. Probability basics. A *random variable* is a map $X : \Omega \rightarrow \mathbb{R}$ from a probability space (Ω, Σ, P) to the real numbers, \mathbb{R} . The *expected value* or *expectation* of X , written $\mathbb{E}[X]$, is defined by,

$$(1.1.1) \quad \mathbb{E}[X] = \int_{\Omega} X(\omega) \, dP(\omega).$$

It follows from linearity of the integral that expectation is *linear*, i.e. $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for constants a, b and random variables X, Y . The *variance* of X , written $\text{Var}[X]$, is defined by,

$$(1.1.2) \quad \text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

The *covariance* of two random variables, written $\text{Cov}[X, Y]$, is defined by

$$(1.1.3) \quad \text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

It follows that when X, Y are *independent*, $\text{Cov}[X, Y] = 0$. The converse is not true. When Ω is a discrete probability space, as will often be the case in the following chapters, the Lebesgue integral of Equation (1.1.1) takes the simpler form,

$$(1.1.4) \quad \mathbb{E}[X] = \sum_{\omega \in \Omega} \omega \mathbb{P}[X = \omega].$$

When X is a non-negative integer-valued random variable, it satisfies the following useful inequalities:

$$(1.1.5) \quad \mathbb{P}[X > 0] \leq \mathbb{E}[X], \text{ and}$$

$$(1.1.6) \quad \mathbb{P}[X = 0] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2}.$$

Equations (1.1.5) and (1.1.6) are sometimes called, respectively, the *first moment method* and *second moment method* [AS92], since the bounds are in terms of, respectively, expectation (the first moment) and variance (the second moment). Equation (1.1.5) can be thought of as a special case of Markov's inequality, and Equation (1.1.6) as a special case of Chebyshev's inequality. For these inequalities and other basics of measure-theoretic probability, we recommend [Bil95].

An *indicator random variable* $\mathbf{1}_A$ for an event A takes the value

$$(1.1.7) \quad \mathbf{1}_A = \begin{cases} 1, & \text{if event } A \text{ occurs,} \\ 0, & \text{otherwise.} \end{cases}$$

The expected value and variance of an indicator random variable satisfy

$$(1.1.8) \quad \mathbb{E}[\mathbf{1}_A] = \mathbb{P}[A],$$

$$(1.1.9) \quad \text{Var}[\mathbf{1}_A] = \mathbb{P}[A](1 - \mathbb{P}[A]).$$

The *union bound* says that if A_1, \dots, A_r are a collection of events, then the probability that at least one of them occurs (i.e., the probability of their union), satisfies

$$(1.1.10) \quad \mathbb{P}\left[\bigcup_{i=1}^r A_i\right] \leq \sum_{i=1}^r \mathbb{P}[A_i].$$

If the A_i are independent and identically distributed (abbreviated *i.i.d.*), with $\mathbb{P}[A_i] = p$ for all i , then the union bound implies the following useful inequality:

$$1 - (1 - p)^r \leq rp.$$

We often consider a sequence of probabilistic events $A(n)$, for n in the natural numbers \mathbb{N} , and analyze the *asymptotic behavior* of $\lim_{n \rightarrow \infty} A(n)$. The following asymptotic notation and abbreviations will be used throughout the text.

DEFINITION 1.1.11 (Asymptotically almost surely). If $A(n)$ is a sequence of probabilistic events, then A happens *asymptotically almost surely*, abbreviated *a.a.s.*, if

$$\lim_{n \rightarrow \infty} \mathbb{P}[A(n)] = 1.$$

DEFINITION 1.1.12 (Asymptotic notation). For two non-negative, real-valued functions $f(x), g(x)$ of the same variable, we

- (1) write $f = \mathcal{O}(g)$, and say “ f is ‘big o’ of g ,” if there exists a constant $c < \infty$ such that $\lim_{x \rightarrow \infty} f(x)/g(x) \leq c$,
- (2) write $f = o(g)$, and say “ f is ‘little o’ of g ,” if $\lim_{x \rightarrow \infty} f(x)/g(x) = 0$,
- (3) write $f = \Omega(g)$, and say “ f is ‘big omega’ of g ,” if there exists a constant $C < \infty$ such that $\lim_{x \rightarrow \infty} f(x)/g(x) \geq C$,
- (4) write $f = \omega(g)$, and say “ f is ‘little omega’ of g ,” if $\lim_{x \rightarrow \infty} f(x)/g(x) = \infty$, and

(5) write $f = \Theta(g)$, and say “ f is ‘big theta’ of g ,” if both $f = \mathcal{O}(g)$ and $f = \Omega(g)$.

Note that the equality signs are a traditional part of this notation, and do not represent literal mathematical equality.

1.1.2. Random graphs. Though usually called the Erdős-Rényi random graph, the following model was introduced by Gilbert [Gil59] as well as Erdős and Rényi [ER59].

DEFINITION 1.1.13 (Erdős-Rényi random graph). Given parameters $n \in \mathbb{N}$ and $p \in (0, 1)$, a random graph \mathfrak{G} has vertex set $[n] = \{1, \dots, n\}$, and edge set given by

$$(1.1.14) \quad \mathbb{P}[e \in E(\mathfrak{G})] = p \quad \text{for all } e \in E(K_n).$$

The resulting probability distribution, on the space of all graphs with vertex set $[n]$, is denoted by $\mathcal{G}(n, p)$. We write $\mathfrak{G} \sim \mathcal{G}(n, p)$ to indicate that \mathfrak{G} is a graph sampled according to $\mathcal{G}(n, p)$.

Figure 1.2 displays some Erdős-Rényi graphs produced by randomly selecting edges to draw, according to Equation (1.1.14), when compiling this document. A useful feature of the Erdős-Rényi random graph model is that the events $e \in E(\mathfrak{G})$ are i.i.d. This often simplifies computing probabilities of events described in terms of sets of edges, such as in the following two propositions. The proofs are explained in detail, as an illustration of some of the probabilistic arguments used to prove the main theorems of this dissertation.

PROPOSITION 1.1.15. Fix a labeled graph H on n vertices, and let $\mathfrak{G} \sim \mathcal{G}(n, p)$. Then

$$\mathbb{P}[\mathfrak{G} = H] = p^{\#E(H)}(1 - p)^{\binom{n}{2} - \#E(H)}.$$

PROOF. Let $\mathbf{1}_e$ be the indicator random variable for the event “ $e \in E(\mathfrak{G})$.” Since \mathfrak{G}, H are both graphs on vertex set $[n]$, the event “ $\mathfrak{G} = H$ ” occurs if and only if $E(\mathfrak{G}) = E(H)$, which occurs if and only if

$$\mathbf{1}_e = \begin{cases} 1 & e \in E(H), \\ 0 & e \notin E(H), \end{cases}$$

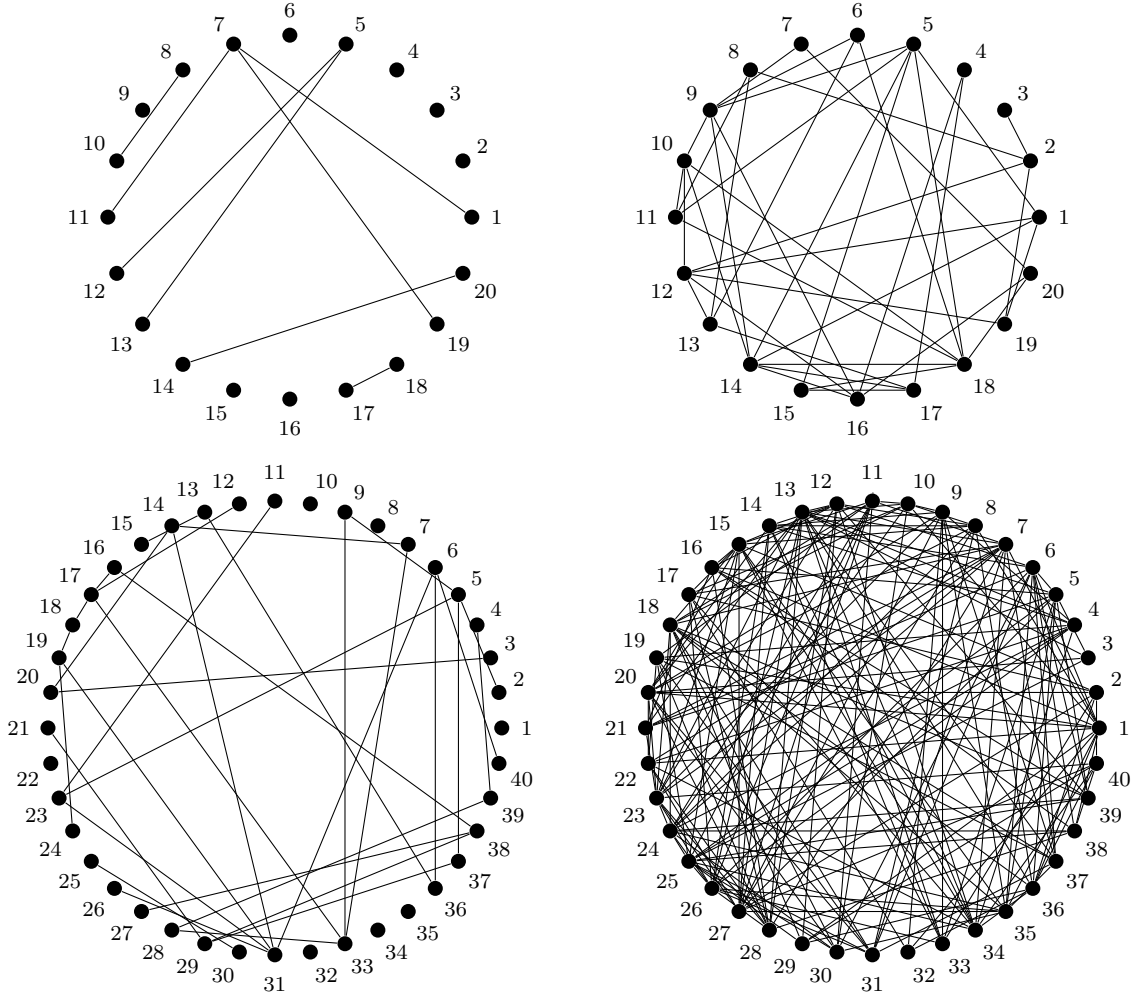


FIGURE 1.2. Erdős-Rényi random graphs sampled from the distributions $\mathcal{G}(20, 0.05)$ (top left), $\mathcal{G}(20, 0.25)$ (top right), $\mathcal{G}(40, 0.05)$ (bottom left), and $\mathcal{G}(40, 0.25)$ (bottom right).

for every edge $e = \{i, j\}$ of the complete graph K_n . Since the $\mathbf{1}_e$ are independent,

$$\begin{aligned}
 \mathbb{P}[\mathfrak{G} = H] &= \mathbb{P} \left[\bigwedge_{e \in H} \{\mathbf{1}_e = 1\} \wedge \bigwedge_{e \notin E(H)} \{\mathbf{1}_e = 0\} \right] \\
 &= \prod_{e \in E(H)} \mathbb{P}[\mathbf{1}_e = 1] \prod_{e \notin E(H)} \mathbb{P}[\mathbf{1}_e = 0] \\
 &= \prod_{e \in E(H)} p \prod_{e \notin E(H)} (1 - p).
 \end{aligned}$$

The first product has $\#E(H)$ factors, by definition, and since there are $\binom{n}{2}$ edges in K_n , the second product has $\binom{n}{2} - \#E(H)$ factors, and the proposition follows. \square

PROPOSITION 1.1.16. Let $\mathfrak{G} \sim \mathcal{G}(n, p)$. Then for every $i \in [n]$,

$$\mathbb{E}[\deg(i)] = np - p.$$

PROOF. The degree of vertex i is equal to the number of edges of \mathfrak{G} that contain i , so

$$\begin{aligned} \mathbb{E}[\deg(i)] &= \mathbb{E}\left[\sum_{j \neq i} \mathbf{1}_{\{i,j\}}\right] \\ &= \sum_{j \neq i} \mathbb{E}[\mathbf{1}_{\{i,j\}}] \\ &= \sum_{j \neq i} \mathbb{P}[e \in \mathfrak{G}] = \sum_{j \neq i} p = (n-1)p. \end{aligned}$$

The second line follows from linearity of expectation (Equation (1.1.1)), and the third from the fact about indicator variables, $\mathbb{E}[\mathbf{1}_A] = \mathbb{P}[A]$ (Equation (1.1.8)). \square

DEFINITION 1.1.17 (Monotone property). For two graphs G, H on vertex set $[n]$, we say that $G \leq H$ if the edges of G are a subset of the edges of H . We think of a graph theoretic property \mathcal{A} as the collection of all graphs with that property. Property \mathcal{A} is called *monotone* if $G \in \mathcal{A}$ and $G \leq H$ implies that $H \in \mathcal{A}$.

For example, the property of being *connected* is monotone, because adding edges to a connected graph always results in a graph that is still connected.

DEFINITION 1.1.18 (Threshold function). For a graph property \mathcal{A} , we say that $t(n)$ is a *threshold function* for \mathcal{A} if

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{G}(n, p(n)) \in \mathcal{A}] = \begin{cases} 0, & p(n) = o(t(n)), \\ 1, & p(n) = \omega(t(n)). \end{cases}$$

The following proposition is to demonstrate the usefulness of the first- and second moment methods in proofs of threshold behavior. This example is taken from [AS92, Chapter 10], although a different proof is given here.

PROPOSITION 1.1.19. For $\mathfrak{G} \sim \mathcal{G}(n, p)$, the function $t(n) = 1/n$ is a threshold for the property of being triangle-free.

PROOF. For each triple of vertices $\{i, j, k\} \subset [n]$, let $\mathbf{1}_{ijk}$ be the indicator random variable for the event “ \mathfrak{G} contains the triangle with vertices $\{i, j, k\}$.” Then

$$\mathbb{P}[\mathbf{1}_{ijk}] = \mathbb{P}[\{i, j\} \in E(\mathfrak{G}) \wedge \{j, k\} \in E(\mathfrak{G}) \wedge \{i, k\} \in E(\mathfrak{G})] = p^3$$

by independence. Now let X be the random variable that counts the number of triangles in \mathfrak{G} , so

$$(1.1.20) \quad X = \sum_{\{i,j,k\} \subseteq [n]} \mathbf{1}_{ijk}$$

and therefore

$$\mathbb{E}[X] = \mathbb{E} \left[\sum_{\{i,j,k\} \subseteq [n]} \mathbf{1}_{ijk} \right] = \sum_{\{i,j,k\} \subseteq [n]} \mathbb{E}[\mathbf{1}_{ijk}] = \binom{n}{3} p^3.$$

By the first moment method,

$$\lim_{n \rightarrow \infty} \mathbb{P}[X > 0] \leq \lim_{n \rightarrow \infty} \mathbb{E}[X] = \lim_{n \rightarrow \infty} \binom{n}{3} p^3.$$

The event $X > 0$ means \mathfrak{G} has at least one triangle. To establish the first side of the threshold, suppose $p = o(1/n)$, so by definition $\lim_{x \rightarrow \infty} np = 0$. Then

$$\lim_{n \rightarrow \infty} \binom{n}{3} p^3 \leq \lim_{n \rightarrow \infty} n^3 p^3 = \lim_{n \rightarrow \infty} (np)^3 = 0,$$

so \mathfrak{G} is triangle-free a.a.s. To compute $\text{Var}[X]$, we apply its definition (Equation (1.1.2)) to Equation (1.1.20) to get:

$$(1.1.21) \quad \text{Var}[X] = \sum_{\{i,j,k\} \subseteq [n]} \text{Var}[\mathbf{1}_{ijk}] + \sum_{\{i,j,k\} \neq \{i',j',k'\}} \text{Cov}[\mathbf{1}_{ijk}, \mathbf{1}_{i'j'k'}].$$

By Equation (1.1.8), each term in the first sum is $\text{Var}[\mathbf{1}_{ijk}] = p^3(1 - p^3)$. For the second sum, we note that if the triangles on vertices $\{i, j, k\}$ and vertices $\{i', j', k'\}$ have no edges in common, then $\mathbf{1}_{ijk}$ and $\mathbf{1}_{i'j'k'}$ are independent and have zero covariance. Since the triangles

cannot have two or three edges in common without being identical, the only interesting possibility when there is exactly one edge in common, i.e., exactly two vertices in common. In this case $\mathbb{P}[\mathbf{1}_{ijk} \wedge \mathbf{1}_{ij'k'}] = p^5$, since there are five edges which must be included in \mathfrak{G} with independent probability p each, and so $\text{Cov}[\mathbf{1}_{ijk}, \mathbf{1}_{ij'k'}] = p^5 - p^6$. For each set $\{i, j, k\} \subseteq [n]$, there are three choices of $\{i', j', k'\}$ with this situation, so Equation (1.1.21) becomes

$$(1.1.22) \quad \text{Var}[X] = \binom{n}{3}(p^3 - p^6) - 3\binom{n}{3}(p^5 - p^6).$$

By the second moment method,

$$\mathbb{P}[X = 0] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2} = \frac{\binom{n}{3}(p^3 - p^6) - 3\binom{n}{3}(p^5 - p^6)}{\binom{n}{3}^2 p^6} = \frac{1 - 3p^2 + 2p^3}{\binom{n}{3} p^3}.$$

Let $p = \omega(1/n)$, then $\lim_{n \rightarrow \infty} \mathbb{P}[X = 0] = 0$, and therefore \mathfrak{G} has a nonzero number of triangles a.a.s. □

A much more famous threshold for random graphs is the threshold for connectedness proved in [ER59]:

THEOREM 1.1.23. For $\mathfrak{G} \sim \mathcal{G}(n, p)$, the function $t(n) = \ln(n)/n$ is a threshold for the property of being connected.

The proof of Theorem 1.1.23 is beyond the scope of this introduction. However, it is useful for demonstrating one way to visualize a family of distributions governed by a threshold law. See Figure 1.3. Each square in the image corresponds to an (n, p) pair, for every $n \in \{20, \dots, 50\}$, and values of p between 0 and 0.3, taken at intervals of 0.01. For each (n, p) pair of parameters, the color of the square shows the proportion of random samples $\mathfrak{G} \sim \mathcal{G}(n, p)$ that were connected, with darker shades representing higher sample probabilities of connectedness. For each fixed n , the probability of being connected increases as p increases, from 0 to 1, and as n grows the shift in the phase transition follows the threshold function $p = \ln(n)/n$ (dotted white curve). Furthermore, the transition becomes sharper as n increases.

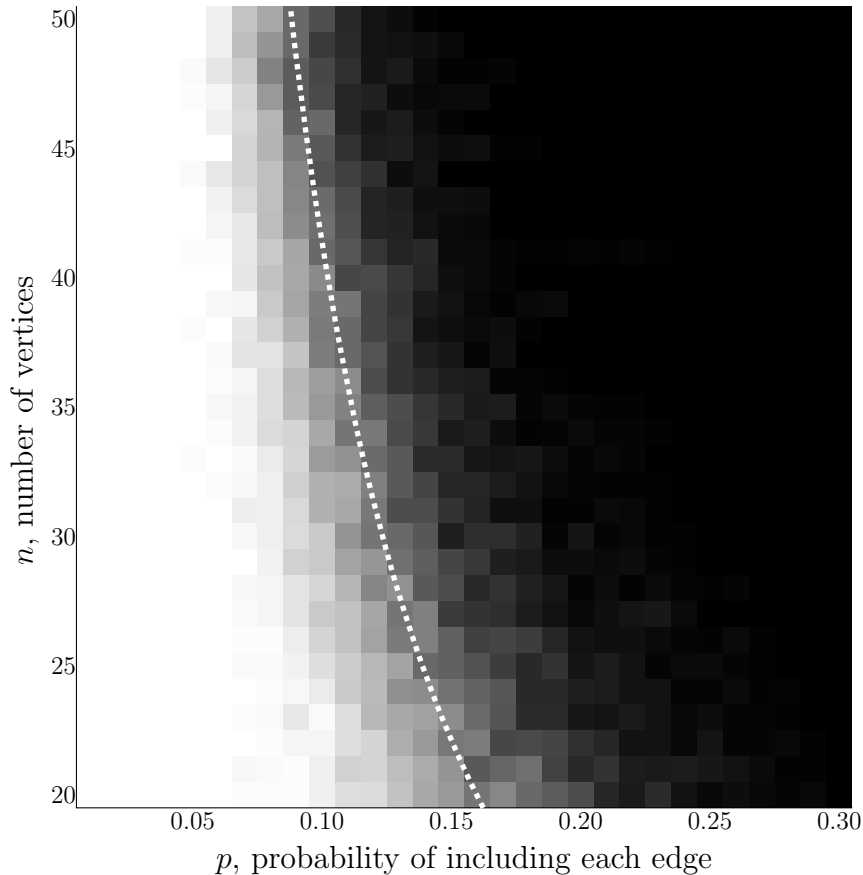


FIGURE 1.3. Visualizing the connectedness threshold of Erdős-Rényi random graphs. Each square in the image corresponds to a value n between 20 and 50, and a value of p between 0 and 0.3. For each (n, p) pair of parameters, 100 random graphs $\mathfrak{G} \sim \mathcal{G}(n, p)$ were sampled and tested for connectedness using the `Graphs` package in `Macaulay2`. The color of the square indicates the sample probability of bring connected, from white (probability zero) to black (probability one). The known threshold function $p = \ln(n)/n$ is the dotted line, and follows the approximate center of the empirical phase transition.

Figure 1.3 demonstrates that although a threshold function is defined asymptotically, it describes a pattern that often emerges at small finite values. That is, knowing whether p is below or above $\ln(n)/n$ gives a good prediction of whether \mathfrak{G} is connected, even for small n .

1.1.3. Random simplicial complexes.

DEFINITION 1.1.24 (Simplicial complex). An (abstract) *simplicial complex* on $[n]$ is a collection Δ of subsets of $[n]$ that is closed under taking subsets. That is, if $\sigma \in \Delta$ and

$\tau \subseteq \sigma$, then $\tau \in \Delta$. Each $\sigma \in \Delta$ is called a *face* of Δ . The elements of $[n]$ are called the *vertices* of Δ . The *facets* of Δ are the faces that are maximal with respect to containment.

DEFINITION 1.1.25 (Dimension of a simplicial complex). If σ is a face of the simplicial complex Δ , we define $\dim \sigma = \#\sigma - 1$. The dimension of Δ is the maximum dimension of one of its faces, i.e.,

$$(1.1.26) \quad \dim \Delta = \max_{\sigma \in \Delta} (\dim \sigma) = \max_{\sigma \in \Delta} (\#\sigma - 1).$$

DEFINITION 1.1.27 (*f*-vector). The *f*-vector of a simplicial complex is defined by $f_i = \#$ faces of dimension i .

EXAMPLE 1.1.28. Let Δ be the simplicial complex on $[6]$ with facets $\{1, 2, 3, 4\}$, $\{4, 5, 6\}$, and $\{1, 5\}$. Figure 1.4a shows the poset of faces of Δ , ordered by containment, while Figure 1.4b depicts a geometric realization of Δ . The dimension of Δ is three. Its *f*-vector is

$$f(\Delta) = (1, 6, 11, 5, 1).$$

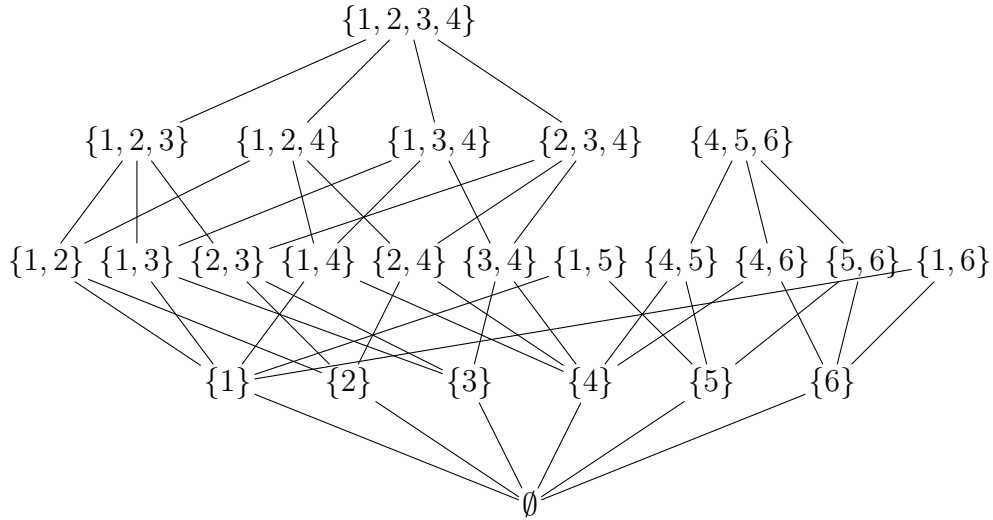
DEFINITION 1.1.29 (*d*-skeleton). The *d*-skeleton of a simplicial complex Δ is written Δ^d , and consists of the collection of faces of Δ of dimension no more than d .

EXAMPLE 1.1.30. The 1-skeleton of a simplicial complex contains its vertices along with subsets of the form $\{i, j\}$. The 1-skeleton is therefore a graph, sometimes called the *underlying graph* of the simplicial complex. Figure 1.4c depicts the 1-skeleton of Δ from Example 1.1.28.

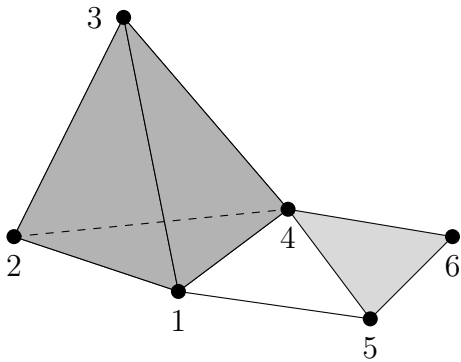
Several random models for simplicial complexes have been studied in recent years, including the following.

DEFINITION 1.1.31 (Linial-Meshulam model [LM06, BHK11]). Start with the complete 1-skeleton on $[n]$. A random two-dimensional simplicial complex \mathfrak{X} is sampled by including triangles $T = \{i, j, k\} \subseteq [n]$ according to

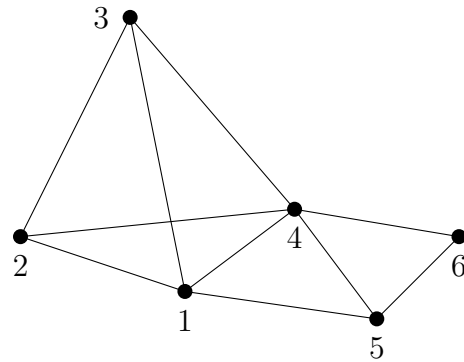
$$\mathbb{P}[T \in E(\mathfrak{X})] = p \text{ for all } T.$$



(A) The poset of faces of Δ , ordered by containment.



(B) A geometric realization of Δ .



(C) The 1-skeleton of Δ .

FIGURE 1.4. Several representations of the simplicial complex Δ from Example 1.1.28, which has facets $\{1, 2, 3, 4\}$, $\{4, 5, 6\}$, and $\{1, 5\}$.

This model is analogous to the Erdős-Rényi graph model after “moving up a dimension.” A natural next step is to generalize to an arbitrary dimension.

DEFINITION 1.1.32 (Meshulam-Wallach model [MW09]). Start with the complete $(d - 1)$ -skeleton on $[n]$. A random d -dimensional simplicial complex \mathfrak{X} is sampled by including d -faces f according to

$$\mathbb{P}[f \in E(\mathfrak{X})] = p \text{ for all } F \subseteq [n].$$

The next model, on the other hand, generates a random simplicial complex directly from an Erdős-Rényi random graph.

DEFINITION 1.1.33 (Random clique complex model [Kah07, Kah09]). Given n and p , sample $\mathfrak{G} \sim \mathcal{G}(n, p)$. The clique complex \mathfrak{X} is defined by attaching a d -face to every clique of size $d + 1$ in \mathfrak{G} . In other words, \mathfrak{X} is the largest simplicial complex that has \mathfrak{G} as its 1-skeleton.

Properties of interest for random simplicial complexes include higher-dimensional analogues of connectivity, the appearance and disappearance of nontrivial homology groups, and their fundamental groups [BHK11].

In [CF16], Costa and Farber introduced a *multi-parameter* model for random simplicial complexes that includes each of these models as a special case.

DEFINITION 1.1.34 (Costa-Farber model). Select a vector $\mathbf{p} = (p_0, p_1, p_2, \dots)$, with $0 \leq p_i \leq 1$ for all i . The faces of a simplicial complex \mathfrak{X} on $[n]$ are sampled one dimension at a time. Each vertex is included according to

$$\mathbb{P}[\{i\} \in E(\mathfrak{X})] = p_0.$$

The edges are then chosen according to

$$\mathbb{P}[\{i, j\} \in E(\mathfrak{X})] = \begin{cases} p_1 & \{i\}, \{j\} \in E(\mathfrak{X}), \\ 0 & \text{otherwise.} \end{cases}$$

and so forth. For each dimension d , possible d -faces are included with independent probability p_d , given that they are supported on the lower-dimensional faces already sampled.

Viewed as special cases of the Costa-Farber model, the random models already mentioned can be described by the following particular choices of parameters:

- Erdős-Rényi: $\mathbf{p} = (1, p, 0, 0, 0, \dots)$.
- Linial-Meshulam: $\mathbf{p} = (1, 1, p, 0, 0, 0, \dots)$.
- Meshulam-Wallach: $\mathbf{p} = (1, \underbrace{1, \dots, 1}_d, p, 0, 0, \dots)$.
- Clique complex: $\mathbf{p} = (1, p, 1, 1, 1, \dots)$.

1.2. Ideals and varieties

Throughout this dissertation \mathbb{k} is a field, and S refers to $S = \mathbb{k}[x_1, \dots, x_n]$, the ring of polynomials in the variables x_1, \dots, x_n , with coefficients in \mathbb{k} . Given a set of polynomials $\{f_1, f_2, \dots, f_m\}$, the *ideal* generated by f_1, \dots, f_m is written $I = \langle f_1, f_2, \dots, f_m \rangle$, and equals

$$I = \{p_1 f_1 + p_2 f_2 + \dots + p_m f_m : p_i \in S\},$$

the set of S -linear combinations of the generating set. For F an infinite set of polynomials, the ideal $I = \langle F \rangle$ is defined as the smallest subset of S that contains F , is closed under addition, and “absorbs products,” i.e., is closed under multiplication by elements of S . However, by the *Hilbert basis theorem* (see, e.g., Chapter 1 of [BWK93]), every ideal of S is *finitely generated*. In other words, there exists a finite set of polynomials $\{f_1, \dots, f_m\}$ such that $I = \langle F \rangle = \langle f_1, \dots, f_m \rangle$.

DEFINITION 1.2.1 (Affine variety). The *variety* of I , written $\mathcal{V}(I)$, is the set

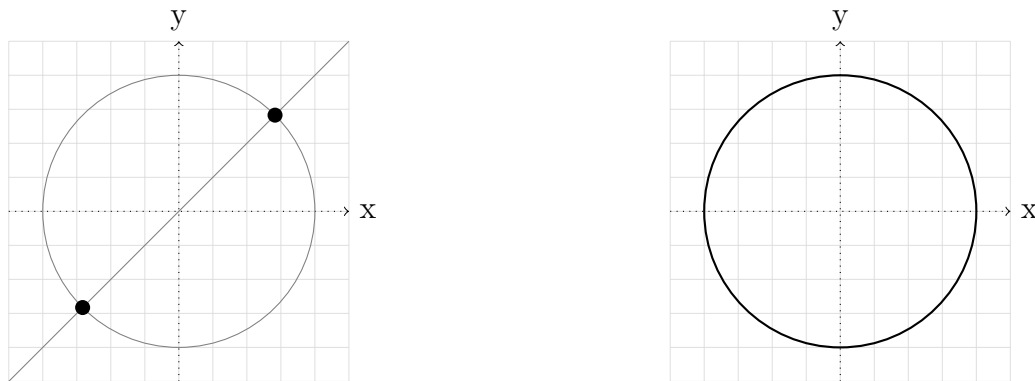
$$\{\vec{v} \in \mathbb{k}^n : f(\vec{v}) = 0 \text{ for all } f \in I\}.$$

Equivalently, if $\{f_1, \dots, f_m\}$ generate I , then $\mathcal{V}(I)$ is the set of simultaneous solutions over \mathbb{k} to the system $\{f_1 = 0, f_2 = 0, \dots, f_m = 0\}$.

EXAMPLE 1.2.2. Let $I = \langle x^2 + y^2 - 1, x - y \rangle \subset \mathbb{R}[x, y]$. Then $\mathcal{V}(I) = \{(\sqrt{2}/2, \sqrt{2}/2), (-\sqrt{2}/2, -\sqrt{2}/2)\}$, the two points of \mathbb{R}^2 that lie both on the circle $x^2 + y^2 = 1$ and on the line $x = y$. See Figure 1.5a. On the other hand, let $J = \langle x^2 + y^2 - 1 \rangle \subset \mathbb{R}[x, y]$. Then $\mathcal{V}(J)$ contains infinitely many points; namely, all points of \mathbb{R}^2 that lie on the circle $x^2 + y^2 = 1$. See Figure 1.5b.

In addition to defining the variety of an ideal, we can also define the *ideal of a variety*. Given V a subset of \mathbb{k}^n , the ideal of V is the set of all polynomials that simultaneously vanish at every $\vec{v} \in V$:

$$\mathcal{I}(\mathcal{V}) = \{f \in S : f(\vec{v}) = 0 \text{ for all } \vec{v} \in V\}.$$



(A) The variety of $I = \langle x^2 + y^2 - 1, x - y \rangle \subset \mathbb{R}[x, y]$ consists of the two points (in black) that simultaneously satisfy $x^2 + y^2 - 1 = 0$ and $x - y = 0$.

(B) The variety of $J = \langle x^2 + y^2 - 1 \rangle \subset \mathbb{R}[x, y]$ is a 1-dimensional curve (circle) in \mathbb{R}^2 .

FIGURE 1.5. First examples of varieties of polynomial ideals.

The *dimension of a variety* is its dimension as a subspace of \mathbb{k}^n . Intuitively, if $\mathcal{V}(I)$ is a finite set of points, then $\dim \mathcal{V}(I) = 0$, if $\mathcal{V}(I)$ a union of lines, curves, and points, then $\dim \mathcal{V}(I) = 1$, and so forth.

EXAMPLE 1.2.3. For $I = \langle x^2 + y^2 - 1, x - y \rangle \subset \mathbb{R}[x, y]$, $\dim \mathcal{V}(I) = 0$ since $\mathcal{V}(I)$ is finite. For $J = \langle x^2 + y^2 - 1 \rangle \subset \mathbb{R}[x, y]$, $\dim \mathcal{V}(J) = 1$ since $\mathcal{V}(J)$ is a one-dimensional curve.

The formal definition of dimension uses the algebraic notion of the *Krull dimension* of a ring, and is explained in Section 1.2.3.

The following standard definitions and theorems appear in any algebra text; e.g., [Rot02]. We review them briefly without much comment.

DEFINITION 1.2.4 (Radical of an ideal and radical ideals). The *radical of an ideal* I is defined by

$$\sqrt{I} = \{f \in S : f^m \in I \text{ for some } m \geq 1\}.$$

An ideal I is called a *radical ideal* if $\sqrt{I} = I$.

DEFINITION 1.2.5 (Prime and primary ideals). An ideal $I \subset S$ is *prime* if it is *proper* (i.e., not equal to the ring itself), and $ab \in I$ implies that $a \in I$ or $b \in I$.

An ideal $I \subset S$ is *primary* if it is proper and $ab \in I$ and $b \notin I$ imply that $a^n \in I$ for some $n \geq 1$.

PROPOSITION 1.2.6. If I is primary, then \sqrt{I} is prime.

DEFINITION 1.2.7 (Primary decomposition, associated primes). A *primary decomposition* of I is a finite set of primary ideals J_1, \dots, J_r satisfying

$$(1.2.8) \quad I = J_1 \cap J_2 \cap \cdots \cap J_r.$$

A primary decomposition is called *irredundant* if no J_i can be omitted from Equation (1.2.8). The prime ideals $\sqrt{J_1}, \dots, \sqrt{J_r}$ are called the *associated primes* of I .

THEOREM 1.2.9 (Lasker-Noether theorem). Every proper ideal $I \subset S$ has a primary decomposition, and the set of associated primes of I is unique.

DEFINITION 1.2.10 (Minimal primes). A prime ideal \mathfrak{p} is a *minimal prime* of I if $I \subseteq \mathfrak{p}$ and there is no prime \mathfrak{p}' such that $I \subseteq \mathfrak{p}' \subsetneq \mathfrak{p}$. Equivalently, the minimal primes of I are the minimal elements in the set of associated primes of I .

DEFINITION 1.2.11 (Ideal quotient). Given $I \subset S$ and $J \subset S$, the *ideal quotient* $\langle I : J \rangle$, also called *colon ideal* is an ideal defined by

$$\langle I : J \rangle = \{f \in S : fj \in I \text{ for all } j \in J\}.$$

We will see in Section 1.4.2 that the ideal quotient $\langle I : J \rangle$ arises quite naturally as the kernel of the quotient map from S/I to $S/(I + J)$.

When \mathbb{k} is *algebraically closed*, then (by definition) every root of a polynomial with coefficients in \mathbb{k} is an element of \mathbb{k} . This leads to the following important theorem.

THEOREM 1.2.12 (Hilbert's Nullstellensatz). Let I be an ideal of $S = \mathbb{k}[x_1, \dots, x_n]$ for \mathbb{k} algebraically closed. Then

- (1) $\mathcal{V}(I) = \emptyset$ if and only if $I = S$.
- (2) $\mathcal{I}(\mathcal{V}(I)) = \sqrt{I}$.

Statements (1) and (2) are often known as the *weak Nullstellensatz* and *strong Nullstellensatz*, respectively. Rotman [Rot02] has a nice and simple proof for the case $\mathbb{k} = \mathbb{C}$ (Chapter 5), in addition to the more general version (Chapter 10).

EXAMPLE 1.2.13. For an example of why algebraically closure is necessary in Theorem 1.2.12, let $I = \langle x^2 + 1 \rangle \subset \mathbb{R}[x]$. Then $\mathcal{V}(I) = \emptyset$ even though I is not equal to $\mathbb{R}[x]$.

REMARK 1.2.14. Version (1) of Hilbert’s Nullstellensatz can be viewed as a method for certifying the infeasibility of various combinatorial problems, by expressing combinatorial constraints as polynomials and then proving that the ideal they generate is the entire polynomial ring. This view, and its fascinating connections to computational complexity and NP-completeness, was the subject of the 2008 Ph.D. dissertation of my academic sister, Susan Margulies [Mar08].

REMARK 1.2.15. The results proved in this dissertation hold over arbitrary fields, unless explicitly stated otherwise. Although algebraic closure and characteristic are generally crucial to proofs in commutative algebra, the combinatorial nature of monomial ideals (Section 1.2.1) often allows us to skirt these considerations. For instance, both versions of Theorem 1.2.12 still hold for monomial ideals even if \mathbb{k} is not algebraically closed, so results about dimension such as Theorem 3.3.2 are true even over fields like \mathbb{R} . Furthermore, the *characteristic* of \mathbb{k} , on which invariants like the minimal free resolution of an ideal (Section 1.3) usually depend, even for monomial ideals, will turn out to be unimportant in the main results on resolutions, Theorems 4.2.2 and 5.2.1. In particular, the last non-zero total Betti number (Definition 1.3.14) of the resolution of a monomial ideal, and hence its projective dimension, is invariant under characteristic. (One way to see this is via *Alexander duality* for monomial ideals, which is not covered in this introduction but is explained in great detail in [MS04] and [HH11].)

1.2.1. Monomial ideals. Recall that S always denotes $\mathbb{k}[x_1, \dots, x_n]$. A *monomial* of S is a polynomial with a single term. For example, x^3y^7z and x^9z^2 are two monomials in $\mathbb{k}[x, y, z]$. In general, we will write x^α as shorthand for the monomial $x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2} \cdots x_n^{\alpha_n}$.

The vector α is called the *exponent vector* of this monomial. For example, the exponent vector of $x^9z^2 \in \mathbb{k}[x, y, z]$ is $(9, 0, 2)$. We call $I \subseteq S$ a *monomial ideal* if I can be generated by $G = \{g_1, \dots, g_m\}$ where every g_i is a monomial.

EXAMPLE 1.2.16. The ideal $I = \langle x^2 + 3y, x^2 - 2y^2, y^2 - y \rangle \subset \mathbb{k}[x, y]$ is a monomial ideal with monomial generating set $\{x^2, y\}$.

PROPOSITION 1.2.17 (Unique minimal generators). Every monomial ideal has a unique minimal monomial set of generators.

PROPOSITION 1.2.18 (Prime and irreducible monomial ideals). A monomial ideal of $S = \mathbb{k}[x_1, \dots, x_n]$ is prime (Definition 1.2.5) if and only if it is generated by a subset of the variables $\{x_1, \dots, x_n\}$.

A monomial ideal is called *irreducible* if it cannot be written as the proper intersection of two monomial ideals. A monomial ideal I is irreducible if and only if it is generated by *pure powers* of the variables, i.e.

$$I = \langle x_{i_1}^{a_1}, x_{i_2}^{a_2}, \dots, x_{i_r}^{a_r} \rangle.$$

1.2.2. Squarefree monomial ideals and the Stanley-Reisner correspondence.

DEFINITION 1.2.19 (Squarefree). A monomial x^α is called *squarefree* if every $\alpha_i \in \alpha$ is either one or zero. A *squarefree monomial ideal* is an ideal with a generating set of squarefree monomials. The *support* of x^α is the set $\{x_i : \alpha_i > 0\}$, written $\text{supp}(x^\alpha)$. The *squarefree part* of x^α is $\prod_{x_i \in \text{supp}(x^\alpha)} x_i$, and is written $\sqrt{x^\alpha}$.

EXAMPLE 1.2.20. The ideal $\langle x_1x_3x_4, x_1x_2, x_3x_4x_5 \rangle \subset \mathbb{k}[x_1, \dots, x_5]$ is a squarefree monomial ideal. The ideal $\langle x_1^3x_3x_4, x_1x_2^3, x_3x_4x_5 \rangle$ is not. The support of $x_1^3x_3x_4$ is $\text{supp}(x_1^3x_3x_4) = \{x_1, x_3, x_4\}$, and its squarefree part is $\sqrt{x_1^3x_3x_4} = x_1x_3x_4$.

A few useful facts about squarefree monomial ideals are collected here, all of which can be found, with proofs, in Chapter 1 of [HH11].

PROPOSITION 1.2.21. If I is a monomial ideal with minimal generators $G = \{g_1, \dots, g_r\}$, then \sqrt{I} is a squarefree monomial ideal generated by $\{\sqrt{g_1}, \dots, \sqrt{g_r}\}$.

COROLLARY 1.2.22. A monomial ideal is radical if and only if it squarefree.

EXAMPLE 1.2.23. Even when G is a minimal generating set for I , the set $\{\sqrt{g} : g \in G\}$ may not be a *minimal* generating set for \sqrt{I} . For example, let $G = \{x^3yz, xy^3z, xyz^3, x^2y^2z^2w^2, w^3\}$, $I = \langle G \rangle \subset \mathbb{k}[x, y, z, w]$. Although these five monomials minimally generate I , their squarefree parts are, respectively, xyz , xyz , xyz , $xyzw$, and w . Therefore $\{\sqrt{g} : g \in G\} = \{xyz, xyzw, w\}$ even though \sqrt{I} is minimally generated by $\{xyz, w\}$. This example also shows that we can have two minimal generators of I such that $g \neq g'$ but $\sqrt{g} = \sqrt{g'}$.

PROPOSITION 1.2.24. A squarefree monomial ideal is the intersection of its minimal primes (Definition 1.2.10), each of which is a monomial prime ideal.

Monomial ideals have a complementary structure to simplicial complexes in the following sense: a simplicial complex is closed “downwards,” with a unique maximal set of facets along with everything smaller in the partial order of containment, while a monomial ideal is closed “upwards,” with a unique minimal set of monomials along with everything larger in the partial order of divisibility. When a monomial ideal is squarefree, its generators are essentially subsets of the variables, and this complementarity becomes literal.

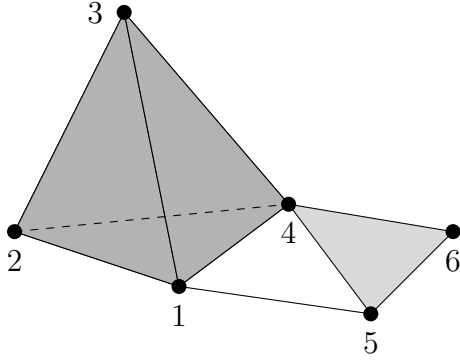
DEFINITION 1.2.25 (Stanley-Reisner ideal). If Δ is a simplicial complex on $[n]$, then the *Stanley-Reisner ideal* of Δ , I_Δ , is the squarefree monomial ideal of S defined by

$$(1.2.26) \quad I_\Delta = \langle x_{i_1} \cdots x_{i_r} : \{i_1, \dots, i_r\} \notin \Delta \rangle.$$

The *Stanley-Reisner ring* of Δ , which encodes all \mathbb{k} -linear combinations of the faces of Δ , is S/I_Δ .

THEOREM 1.2.27. The definition of I_Δ in Equation (1.2.26) defines a bijection between simplicial complexes on $[n]$ and squarefree monomial ideals in $\mathbb{k}[x_1, \dots, x_n]$.

See, for example, [MS04, Theorem 1.7].



$$I_{\Delta} = \langle x_1x_6, x_2x_5, x_2x_6, x_3x_5, x_3x_6, x_1x_4x_5 \rangle$$

FIGURE 1.6. The simplicial complex Δ from Example 1.1.28 (left). The subset $\{1, 4, 6\}$ is a non-face, but not a minimal one, since it properly contains the non-face $\{1, 6\}$. The monomials $x_1x_4x_6$ and x_1x_6 are both elements of I_{Δ} by definition, but $x_1x_4x_6$ is a redundant generator because it is a multiple of x_1x_6 . The minimal generators of the Stanley-Reisner ideal I_{Δ} (right) are exactly those that come from minimal non-faces of the complex.

EXAMPLE 1.2.28. Equation (1.2.26) says that I_{Δ} is generated by the monomials corresponding to the *non-faces* of Δ . Recall the simplicial complex from Example 1.1.28 which had facets $\{1, 2, 3, 4\}$, $\{4, 5, 6\}$, and $\{1, 5\}$. The subset $\{1, 4, 6\}$ is not a face of this complex, so the monomial $x_1x_4x_6$ belongs to I_{Δ} . Similarly $\{1, 6\}$ is not a face, so $x_1x_6 \in I_{\Delta}$. Since x_1x_6 divides $x_1x_4x_6$, the latter monomial is not a minimal generator. This motivates the definition of a *minimal non-face* of Δ : a non-face all of whose proper subsets are faces. Because containment of faces corresponds to divisibility of monomials in I_{Δ} , the minimal generators of I_{Δ} are given by the minimal non-faces of Δ . In an example like Figure 1.6, they can be listed by simply inspecting the figure.

Through the Stanley-Reisner correspondence, any random model for graphs or simplicial complexes can be viewed as a random model for square-free monomial ideals, and vice versa (see Section 2.4). The random models in Chapter 2 are much more general, as they are distributions over all monomial ideals and not just squarefree ones. Even so, squarefree monomial ideals will appear in many computations in later chapters. For instance, computing the probabilistic dimension of a random monomial ideal relies on a hypergraph defined by its (squarefree) radical in Sections 3.2 and 3.3, and the irreducible decomposition of a monomial

ideal is described in ?? by first converting a general monomial ideal into a squarefree one in a new set of variables. This process is called *polarization*.

DEFINITION 1.2.29 (Polarization of a monomial ideal). Let I be a monomial ideal of $S = \mathbb{k}[x_1, \dots, x_n]$ with minimal generators $G = \{g_1, \dots, g_r\}$, with $g_i = x_1^{\alpha_{i,1}} \cdots x_n^{\alpha_{i,n}}$. The polarization of I is a squarefree monomial ideal defined by generators z_1, \dots, z_r , where z_i is obtained from g_i by replacing each power $x_j^{\alpha_{i,j}}$ with the product $x_{j,1}x_{j,2} \cdots x_{j,\alpha_{i,j}}$. Formally,

$$\text{polarization}(I) = \left\langle \prod_{j=1}^n \prod_{k=1}^{\alpha_{i,j}} x_{j,k} \right\rangle.$$

The polarization is an ideal of $\mathbb{k}[x_{1,1}, \dots, x_{1,m_1}, x_{2,1}, \dots, x_{2,m_2}, \dots, x_{n,1}, \dots, x_{n,m_n}]$, where $m_j = \max \alpha_{i,j}$.

EXAMPLE 1.2.30. The actual process of polarization is simpler than the process of writing down the correct notation in the definition. For instance, if $I = \langle x^2y^2, y^3, xz^2 \rangle$, then

$$\text{polarization}(I) = \langle x_1x_2y_1y_2, y_1y_2y_3, x_1z_1z_2 \rangle,$$

and lives in the ring $\mathbb{k}[x_1, x_2, y_1, y_2, y_3, z_1, z_2]$.

1.2.3. The dimension of a monomial ideal. The *Krull dimension* of a ring R is defined as the supremum of the lengths of chains of distinct prime ideals in R . For instance, the Krull dimension of the polynomial ring $S = \mathbb{k}[x_1, \dots, x_n]$ is n . A maximal chain of prime ideals in S is

$$\langle 0 \rangle \subset \langle x_1 \rangle \subset \langle x_1, x_2 \rangle \subset \cdots \subset \langle x_1, \dots, x_n \rangle,$$

which has n proper containments. Note that the ring S itself is not considered a prime ideal. That there are no chains of greater length requires additional justification, see Chapter 8 of [Eis95].

If I is a prime ideal of R , its *codimension* is the supremum of lengths of chains of distinct primes descending from I . For instance, the prime ideal $I = \langle x_1, \dots, x_r \rangle$ of S , for $r \leq n$, has

codimension r , evidenced by the chain

$$\langle 0 \rangle \subset \langle x_1 \rangle \subset \langle x_1, x_2 \rangle \subset \cdots \subset \langle x_1, \dots, x_r \rangle.$$

Hence $\dim I = n - r$.

For I not assumed prime, $\text{codim } I$ is the minimum codimension of a prime containing I .

EXAMPLE 1.2.31. The ideal $J = \langle xy^3, x^3y \rangle \subset \mathbb{k}[x, y]$ is not prime. J is contained by the prime ideals $\langle x \rangle$ and $\langle y \rangle$, which each have codimension 1, and by $\langle x, y \rangle$, which has codimension 2. Therefore $\text{codim } J = 1$. Since $\dim \mathbb{k}[x, y] = 2$, $\dim J = 1$.

Example 1.2.31 illustrates how $\text{codim } I$ can be computed combinatorially in the case that I is a monomial ideal of $S = \mathbb{k}[x_1, \dots, x_n]$. Let \mathfrak{p}_T denote the prime ideal generated by $\{x_i \in T\}$. Since the prime monomial ideals of $\mathbb{k}[x_1, \dots, x_n]$ are exactly those given by \mathfrak{p}_T for $T \subseteq S$, the codimension of I is equal to

$$(1.2.32) \quad \text{codim } I = \min\{\#T : I \subseteq \mathfrak{p}_T\}.$$

Furthermore, for every $T \subseteq \{x_1, \dots, x_n\}$, $I \subseteq \mathfrak{p}_T$ if and only if for every monomial $g \in I$, at least one $x_i \in T$ divides g . So computing the codimension of a monomial ideal is equivalent to finding the smallest subset T with this property. It is convenient to recast this as finding a minimum vertex cover of a hypergraph defined by the generators of I . This is explained in detail in Section 3.1.

1.3. Resolutions of monomial ideals

1.3.1. Minimal free resolutions. Minimal free resolutions are an important and central topic in commutative algebra. For instance, in the setting of modules over finitely generated graded \mathbb{k} -algebras, the numerical data of these resolutions encode the Hilbert series, Castelnuovo-Mumford regularity and other fundamental invariants. Minimal free resolutions also provide a starting place for a myriad of homology and cohomology computations. An essential resource on minimal free resolutions in the general commutative algebra setting is [Eis95].

DEFINITION 1.3.1 (Minimal free resolution). Let M be a finitely generated S -module. A *free resolution* of M is an exact chain complex

$$0 \longleftarrow M \xleftarrow{\delta_0} F_0 \xleftarrow{\delta_1} F_1 \xleftarrow{\delta_2} F_2 \xleftarrow{\delta_3} \dots$$

where the F_i are free S -modules. The S -module homomorphisms δ_i are sometimes called *differentials* for reasons that have nothing to do with combinatorial commutative algebra. When S has a grading, we define a *graded free resolution* by adding the requirement that all differentials be homogeneous of degree zero. This is accomplished by “twisting” or grading the summands of F_i appropriately, so each F_i is written

$$(1.3.2) \quad F_i = \bigoplus_{j=1}^r S(-a_{i,j}).$$

The integer r is called the *rank* of F_i . A *minimal (graded) free resolution* is one where the rank of each F_i is minimal over all free resolutions of I .

This definition requires a minimal free resolution to *simultaneously* minimize the rank of every F_i ; *a priori*, it is not clear that such an object exists, let alone is unique. The next theorem is quite wonderful.

THEOREM 1.3.3. The minimal (graded) free resolution of a finitely generated S -module M exists, is unique up to isomorphism, and is a direct summand of any free (graded) resolution of M .

For a proof see, e.g., [Eis95, Theorem 20.2]. In light of Theorem 1.3.3, it is common to say “the” minimal free resolution of M when referring to any isomorphic copy.

EXAMPLE 1.3.4. Let $I = \langle x^2y, yz^3, x^2z^2 \rangle \subseteq \mathbb{k}[x, y, z]$. To build a free resolution of S/I , we begin by setting $F_0 = S$, and letting δ_0 be the canonical quotient map $\pi : S \rightarrow S/I$.

$$0 \longleftarrow S/I \xleftarrow{\pi} S \longleftarrow \dots$$

The kernel of π is I . Since I consists of all S -linear combinations of x^2y , yz^3 , and x^2z^2 , we can write $I = \left\{ [x^2y \ yz^3 \ x^2z^2] \cdot [f_1 \ f_2 \ f_3]^T : [f_1 \ f_2 \ f_3]^T \in S^3 \right\}$, and recognize I as the

image of $\delta_1 : S^3 \rightarrow S$ where δ_1 is left-multiplication by $[x^2y \ yz^3 \ x^2z^2]$. We write this

$$(1.3.5) \quad 0 \longleftarrow S/I \xleftarrow{\pi} S \xleftarrow{[x^2y \ yz^3 \ x^2z^2]} S^3 \longleftarrow \dots$$

Call the standard basis elements of $F_1 = S^3$ $e_1 = [1 \ 0 \ 0]^T$, $e_2 = [0 \ 1 \ 0]^T$, and $e_3 = [0 \ 0 \ 1]^T$. As defined in Equation (1.3.5), δ_1 sends e_1 , a degree-zero element of S^3 , to $x^2y \in S$, which has degree three. To compensate, we redefine $\deg(e_1) = -3$. Similarly, we redefine $\deg(e_2) = \deg(e_3) = -4$. After these shifts, δ_1 is homogeneous of degree 0, and we record the shifts this way:

$$0 \longleftarrow S/I \xleftarrow{\pi} S \xleftarrow{[x^2y \ yz^3 \ x^2z^2]} \begin{array}{c} S(-3) \\ \oplus \\ S(-4) \\ \oplus \\ S(-4) \end{array} \longleftarrow \dots$$

Next, we need to find $\ker \delta_1$, and define F_2 and δ_2 so that $\delta_2 : F_2 \rightarrow S^3$ satisfies $\text{im } \delta_2 = \ker \delta_1$. The module F_2 is called the *(first) syzygy module* of I and consists of all relations among the generators of I . For instance, the generators x^2y and yz^3 have the relation $z^3 \cdot x^2y - x^2 \cdot yz^3 = 0$. This means that the column vector $z^3e_1 - x^2e_2 = [z^3 \ -x^2 \ 0]^T$ of S^3 belongs to $\ker \delta_1$, since $\delta_1(z^3e_1 - x^2e_2) = [x^2y \ yz^3 \ x^2z^2] \cdot [z^3 \ -x^2 \ 0]^T = 0$. Similarly, from the relations $z^2 \cdot x^2y - y \cdot x^2z^2 = 0$, and $x^2 \cdot yz^3 - yz \cdot x^2z^2 = 0$, we find two more elements of $\ker \delta_1$: $z^2e_1 - ye_3$ and $x^2e_2 - yze_3$. The reader familiar with the Buchberger algorithm for computing Gröbner bases will notice that we are computing *S-pairs*. [CLO05] is highly recommended for its computational explanation of free resolutions based on Gröbner basis theory.

We define F_2 to have three basis elements and δ_2 as the map that sends them to $z^3e_1 - x^2e_2$, $z^2e_1 - ye_3$ and $x^2e_2 - yze_3$, respectively. For reasons that will become clear in the next section, we will name the basis elements of F_2 as $e_{\{1,2\}}$, $e_{\{1,3\}}$, and $e_{\{2,3\}}$, respectively. Finally we

shift the basis elements to make δ_2 degree zero.

$$\begin{array}{ccccccc}
0 \longleftarrow S/I \xleftarrow{\pi} S \xleftarrow{[x^2y \ yz^3 \ x^2z^2]} & S(-3) & \xleftarrow{\begin{bmatrix} z^3 & z^2 & 0 \\ -x^2 & 0 & x^2 \\ 0 & -y & -yz \end{bmatrix}} & S(-6) & \longleftarrow \dots \\
& \oplus & & \oplus & \\
& S(-4) & & S(-5) & \\
& \oplus & & \oplus & \\
& S(-4) & & S(-6) &
\end{array}$$

The next step in the resolution is finding $\ker \delta_2$, which is the module of relations on the syzygy module F_2 , or the *second syzygy module*. Pairwise, the columns of $\begin{bmatrix} z^3 & z^2 & 0 \\ -x^2 & 0 & x^2 \\ 0 & -y & -yz \end{bmatrix}$ are S -linearly independent, but there is a relation on all three of them, since the first column is equal to z times the second column minus the third. In other words, $\ker \delta_2$ is generated by $e_{\{1,2\}} - ze_{\{1,3\}} + e_{\{2,3\}}$, so we define $\delta_3 = \begin{bmatrix} 1 \\ -z \end{bmatrix}$ to take a single basis element of F_3 to this syzygy on F_2 . Since δ_3 is injective, it has trivial kernel, so the construction terminates and we have found a free resolution of S/I .

$$\begin{array}{ccccccc}
(1.3.6) \quad 0 \longleftarrow S/I \xleftarrow{\pi} S \xleftarrow{[x^2y \ yz^3 \ x^2z^2]} & S(-3) & \xleftarrow{\begin{bmatrix} z^3 & z^2 & 0 \\ -x^2 & 0 & x^2 \\ 0 & -y & -yz \end{bmatrix}} & S(-6) & \xleftarrow{\begin{bmatrix} 1 \\ -z \\ 1 \end{bmatrix}} & S(-6) & \longleftarrow 0. \\
& \oplus & & \oplus & & & \\
& S(-4) & & S(-5) & & & \\
& \oplus & & \oplus & & & \\
& S(-4) & & S(-6) & & &
\end{array}$$

But wait a second! Since we can write one of the columns of $\begin{bmatrix} z^3 & z^2 & 0 \\ -x^2 & 0 & x^2 \\ 0 & -y & -yz \end{bmatrix}$ as a S -linear combination of the other two, that means that two columns were sufficient to generate all syzygies. If we lose the redundant syzygy and define δ_2 by the smaller matrix $\begin{bmatrix} z^2 & 0 \\ 0 & x^2 \\ -y & -yz \end{bmatrix}$, then we can define F_2 as a free module of rank 2 instead of 3, while still satisfying $\text{im } \delta_2 = \ker \delta_1$. Now there are no nontrivial relations on the generators of F_2 , so the resolution ends after the first syzygy module, with a smaller free resolution:

$$\begin{array}{ccccccc}
(1.3.7) \quad 0 \longleftarrow S/I \xleftarrow{\pi} S \xleftarrow{[x^2y \ yz^3 \ x^2z^2]} & S(-3) & \xleftarrow{\begin{bmatrix} z^2 & 0 \\ 0 & x^2 \\ -y & -yz \end{bmatrix}} & S(-5) & \longleftarrow 0. \\
& \oplus & & \oplus & \\
& S(-4) & & S(-6) & \\
& \oplus & & & \\
& S(-4) & & &
\end{array}$$

Since I cannot have fewer than three generators, and at least two syzygies are needed to generate all of their relations, it is impossible to find a smaller free resolution. Therefore Equation (1.3.7) is the (up to isomorphism) minimal free resolution of S/I .

In general the minimal free resolution of a monomial ideal may be *characteristic-dependent*. As explained in Remark 1.2.15, we can safely ignore this issue in the proofs of this thesis.

1.3.2. Projective dimension and Cohen-Macaulayness.

DEFINITION 1.3.8 (Projective dimension). The *projective dimension* of an S -module M , written $\text{pdim } M$ is defined as the minimum number of nonzero syzygy modules in a (graded) free resolution, equivalently the number of nonzero syzygy modules in the minimal (graded) free resolution of M .

In general algebra settings, projective dimension refers to the minimum length of a *projective resolution*: one where all modules of the chain complex are required to be projective, but not necessarily free. This distinction is unnecessary in our setting of finitely generated modules over polynomial rings, since in this case a module is projective if and only if it is free. This powerful result was conjectured by Serre and eventually proved by Quillen [Qui76] and, independently, Suslin [Sus76].

EXAMPLE 1.3.9. Continuing Example 1.3.4, let $S = \mathbb{k}[x, y, z]$ and let $I = \langle x^2y, yz^3, x^2z^2 \rangle$. Then $\text{pdim}(S/I) = 2$.

For finitely generated modules over polynomial rings, Hilbert proved an elegant upper bound on the number of nonzero syzygy modules in a minimal resolution.

THEOREM 1.3.10 (Hilbert syzygy theorem). Let $S = \mathbb{k}[x_1, \dots, x_n]$. For every finitely generated S -module M , $\text{pdim } M \leq n$. In other words, M has a free resolution of the form

$$0 \longleftarrow M \longleftarrow F_0 \longleftarrow F_1 \longleftarrow \cdots \longleftarrow F_m \longleftarrow 0,$$

where $m \leq n$.

Appendix A3 of [Eis95] proves a more general statement using modern homological algebra. There is also a lovely computational proof of Theorem 1.3.10 in [CLO05] based on Gröbner bases of modules.

DEFINITION 1.3.11 (Cohen-Macaulay). An S -module is called *Cohen-Macaulay* if its Krull dimension equals its *depth*. For S a polynomial ring, $\text{depth } S/I = n - \text{pdim } S/I$ by the Auslander-Buchsbaum theorem [Eis95, Corollary 19.10]. Therefore for an ideal I of a polynomial ring S , S/I is Cohen-Macaulay if and only if

$$(1.3.12) \quad \dim(S/I) = \text{pdim}(S/I).$$

Since ideals of polynomial rings are the focus of this thesis, we will avoid going into full *depth* about Cohen-Macaulayness, and take Equation (1.3.12) as its definition.

EXAMPLE 1.3.13. Recall $\text{pdim}(S/I) = 2$ for $I = \langle x^2y, yz^3, x^2z^2 \rangle \subseteq S = \mathbb{k}[x, y, z]$. To find $\dim(S/I)$, use Proposition 3.1.2: the support hypergraph of I is the complete graph on vertex set $\{x, y, z\}$, which has transversal number 2. Since S is generated by three variables, $\dim(S/I) = 3 - 2 = 1$. Therefore S/I is not Cohen-Macaulay.

1.3.3. Graded Betti numbers and regularity.

DEFINITION 1.3.14 (Total and graded Betti numbers). Let \mathcal{F} be a minimal graded free resolution of a graded S -module M , with the free modules of the chain complex written, as before, in the form

$$F_i = \bigoplus_j S(-a_{i,j}).$$

The index i is called the *homological degree* of F_i . Because \mathcal{F} is unique up to isomorphism (Theorem 1.3.3), both the rank of F_i and the multiset of grades $\{\{a_{i,j} : S(-a_{i,j}) \in F_i\}\}$ are invariants of M . The ranks of the free modules are called the *total Betti numbers* of M , and written with a single index:

$$\beta_i(M) = \text{rk}(F_i).$$

The *graded Betti numbers* of M count the number of free S -modules of each grade in each homological degree. They are written with two indices:

$$\beta_{i,j}(M) = \# \text{ of copies of } S(-j) \text{ in } F_i.$$

EXAMPLE 1.3.15. Let $M = \mathbb{k}[x, y, z]/\langle x^2y, yz^3, x^2z^2 \rangle$ whose minimal graded resolution appears in Example 1.3.4. The total Betti numbers of M are $\beta_0 = 1$, $\beta_1 = 3$, and $\beta_2 = 2$, with $\beta_i = 0$ for all other i . The graded Betti numbers of M are $\beta_{0,0} = 1$, $\beta_{1,3} = 1$, $\beta_{1,4} = 2$, $\beta_{2,5} = 1$, and $\beta_{2,6} = 1$, with $\beta_{i,j} = 0$ for all other pairs i, j .

The graded Betti numbers of M are often collected in a table called the *Betti table* of M . The columns correspond to homological degrees, and the rows correspond to grades. To save space, we take into account the fact that $\beta_{i,j} = 0$ whenever $j < i$, and place $\beta_{i,i+k}$ in the i th column and k th row. Indexing starts at $(0, 0)$ in the top left.

EXAMPLE 1.3.16. The Betti table of M from the Example 1.3.15 is

	0	1	2
total:	1	3	2
0:	1	.	.
1:
2:	.	1	.
3:	.	2	1
4:	.	.	1

REMARK 1.3.17. In `Macaulay2`, the command `betti res M` will create a Betti table in exactly this format, and the command `tex betti res M` conveniently returns the \LaTeX code for typesetting it.

REMARK 1.3.18. The projective dimension of M is the largest column index in the Betti table of M .

DEFINITION 1.3.19 (Regularity). The *Castelnuovo-Mumford regularity* (or simply *regularity*) of M is written $\text{reg } M$, and defined as the greatest integer k such that the $\beta_{i,i+j}(M) = 0$ for all $j > k$.

Bayer and Stillman showed that (under a few assumptions) the degree complexity of a Gröbner basis with respect to reverse lexicographic order is bounded above by the regularity of the ideal [BS87a], so understanding this invariant has great computational significance.

REMARK 1.3.20. The regularity of M is the largest row index in the Betti table of M .

EXAMPLE 1.3.21. From the Betti table in Example 1.3.16, we see by the number of rows that $M = \mathbb{k}[x, y, z]/\langle x^2y, yz^3, x^2z^2 \rangle$ has regularity 4.

1.3.4. The Taylor complex. When I is a monomial ideal, the syzygies on its generators are determined by *least common multiples* (lcm's) of generators. For instance, in Example 1.3.4, the monomials yz^3 and x^2z^2 are related by $x^2 \cdot yz^3 - yz \cdot x^2z^2 = 0$ and thus the module homomorphism defined by $e_1 \mapsto x^2y$, $e_2 \mapsto yz^3$, $e_3 \mapsto x^2z^2$ contains $x^2e_2 - yze_3$ in its kernel. The fact that $\text{lcm}(yz^3, x^2z^2) = x^2yz^3$ has degree 6 means that the syzygy is in degree 6, and the basis element of F_2 mapped to $x^2e_2 - yze_3$ needs a degree shift of 6. This syzygy corresponds to one of the direct summands of F_2 equal to $S(-6)$ in Equation (1.3.6).

In general, if I is a monomial ideal minimally generated by $\{g_1, \dots, g_s\}$, then every two minimal generators g_i, g_j satisfy the relation

$$(1.3.22) \quad \frac{\text{lcm}(g_i, g_j)}{g_j} \cdot g_i - \frac{\text{lcm}(g_i, g_j)}{g_i} \cdot g_j = 0.$$

Thus $\frac{\text{lcm}(g_i, g_j)}{g_j}e_i - \frac{\text{lcm}(g_i, g_j)}{g_i}e_j$ will belong to the first syzygy module. Suppose $e_{\{i,j\}}$ is a basis element of F_2 and δ_2 is defined so $\delta_2 e_{\{i,j\}} = \frac{\text{lcm}(g_i, g_j)}{g_j}e_i - \frac{\text{lcm}(g_i, g_j)}{g_i}e_j$; then this basis element must be shifted by the degree of $\text{lcm}(g_i, g_j)$, so the free module $S(-\deg \text{lcm}(g_i, g_j))$ shows up as a direct summand of F_2 .

As we saw in Example 1.3.4, including a basis element $e_{\{i,j\}}$ for every syzygy on a g_i, g_j pair is sure to generate the entire syzygy module F_2 . The caveat is that redundant relations may be included.

These constructions readily extend to higher syzygy modules and motivate the definition of the *Taylor complex* of a monomial ideal, which encodes all possible syzygies on sets of the generators and gives a universal recipe for a free resolution of any monomial ideal.

In general, though, the Taylor complex has many redundant generators and is far from a minimal resolution.

DEFINITION 1.3.23 (Taylor complex). Let $I = \langle G \rangle$ be a monomial ideal with minimal generating set $G = \{g_1, \dots, g_r\}$. For each subset L of $\{1, \dots, r\}$ let $m_L = \text{lcm}(g_i \mid i \in L)$. Let $\alpha_L \in \mathbb{N}^n$ be the exponent vector of m_L and let $S(-\alpha_L)$ be the free S -module with one generator in multidegree α_L . The *Taylor complex* of S/I is the \mathbb{Z}^n -graded module

$$(1.3.24) \quad \mathcal{F} = \bigoplus_{L \subseteq \{1, \dots, r\}} S(-\alpha_L)$$

with basis denoted by $\{e_L\}_{L \subseteq \{1, \dots, r\}}$, and equipped with the differential:

$$(1.3.25) \quad \delta(e_L) = \sum_{i \in L} \text{sign}(i, L) \cdot \frac{m_L}{m_{L \setminus i}} \cdot e_{L \setminus i},$$

where $\text{sign}(i, L)$ is $(-1)^{j+1}$ if i is the j th element in the ordering of L . This is a free resolution of S/I over S with 2^r terms; the terms are in bijection with the 2^r subsets of G , and the term corresponding to $L \subseteq G$ appears in homological degree $\#L$.

EXAMPLE 1.3.26. The resolution in Equation (1.3.6) is actually the Taylor complex of $M = \mathbb{C}[x, y, z] / \langle x^2y, yz^3, x^2z^2 \rangle$.

EXAMPLE 1.3.27. Let $I = \langle xw, z^3, y^3w, xyz^2 \rangle \subset \mathbb{C}[x, y, z, w]$. The Taylor complex is depicted in Figure 1.7. The differentials δ are omitted from the figure for readability, but are easily described using Equation (1.3.25). For instance, δ_1 , using the order of basis elements in the figure, is given by the matrix

$$(1.3.28) \quad \begin{bmatrix} z^3 & y^3 & 0 & yz^2 & 0 & 0 \\ -xw & 0 & xy & 0 & y^3w & 0 \\ 0 & -x & 0 & 0 & -z^3 & xz^2 \\ 0 & 0 & -z & -w & 0 & -y^2w \end{bmatrix}.$$

Since the Taylor complex is a free resolution, it contains the minimal free resolution as a direct summand. The minimal free resolution is shown in black in Figure 1.7, with the rest

$$\begin{array}{ccccccc}
& & & & S(-\begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \end{bmatrix}) & & \\
& & & & \oplus & & \\
& & S(-\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix}) \\
& & \oplus & & \oplus & & \oplus \\
& & S(-\begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 1 \\ 3 \\ 0 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}) \\
0 \longleftarrow S \longleftarrow & \oplus & \longleftarrow & \oplus & \longleftarrow & \oplus & \longleftarrow S(-\begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}) \longleftarrow 0, \\
& \oplus & & \oplus & & \oplus & \\
& S(-\begin{bmatrix} 0 \\ 3 \\ 0 \\ 1 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}) & \\
& \oplus & & \oplus & & \oplus & \\
& S(-\begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}) & & S(-\begin{bmatrix} 0 \\ 3 \\ 3 \\ 1 \end{bmatrix}) & & S(-\begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}) & \\
& & & \oplus & & & \\
& & & S(-\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}) & & &
\end{array}$$

FIGURE 1.7. The Taylor complex of $I = \langle xw, z^3, y^3w, xyz^2 \rangle \subset \mathbb{C}[x, y, z, w]$ (see Definition 1.3.23). The minimal free resolution of I is strictly contained in the Taylor complex, and is supported on the free modules drawn in black. The Taylor complex additionally includes the free modules drawn in gray.

of the Taylor complex in gray. Note that the redundant submodule of the Taylor complex in homological degree 1 is $S(-\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix})$. This corresponds to the redundancy of the last column of δ_1 in Equation (1.3.28), which is an S -linear combination of the second and fourth columns.

There is a common property of the modules in Figure 1.7 that appear in the Taylor complex but not in the minimal free resolution: they have multidegrees that appear more than once. This is not a coincidence, by the following theorem.

THEOREM 1.3.29. Any free resolution of an S -module M can be decomposed as the direct sum of the minimal free resolution of M and sequences of the form

$$0 \longleftarrow S(-\alpha) \longleftarrow S(-\alpha) \longleftarrow 0,$$

i.e. pairs of free modules in consecutive homological degrees, shifted by identical multidegrees.

See [Eis06].

REMARK 1.3.30. As a corollary, any multidegree which appears exactly once in the Taylor resolution must be part of the minimal free resolution. The converse is not true, as seen in Example 1.3.27. A copy of $S(-\begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix})$ appears in homological degree 3 of the minimal free resolution, even though this multidegree appears three times in the Taylor resolution. The copies of S graded by unique multidegrees are a subset, possibly strict, of the copies of S that appear in the minimal free resolution. This fact motivates the definition of the *Scarf complex* of a monomial ideal.

1.3.5. The Scarf complex.

DEFINITION 1.3.31. Let $I = \langle G \rangle$ be a monomial ideal with minimal generating set $G = \{g_1, \dots, g_r\}$, and all other notation as in Definition 1.3.23. The *Scarf complex* of I is defined as the subcomplex of the Taylor complex with basis

$$\{e_L : L \subseteq \{1, \dots, r\} \text{ and } m_L \neq m_K \text{ for all } K \subseteq \{1, \dots, r\}, K \neq L\}.$$

In other words, the Scarf complex is the Taylor complex restricted to modules whose multidegree shifts are unique.

REMARK 1.3.32. The Scarf complex is named for mathematical economist Herbert Scarf, and should not be confused with the non-capitalized *scarf complex*, which is a disturbing emotional obsession with neckwear.

By Remark 1.3.30, the Scarf complex is always contained in the minimal free resolution. In general, it may not be a resolution. Its maps are inherited from the Taylor complex, but after restriction, the chain complex need not be exact. If the Scarf complex *is* exact, it is a resolution of S/I contained in the minimal free resolution and therefore equal to the minimal free resolution.

DEFINITION 1.3.33. If S/I is resolved by the Scarf complex of I , we say that the ideal I *is Scarf*.

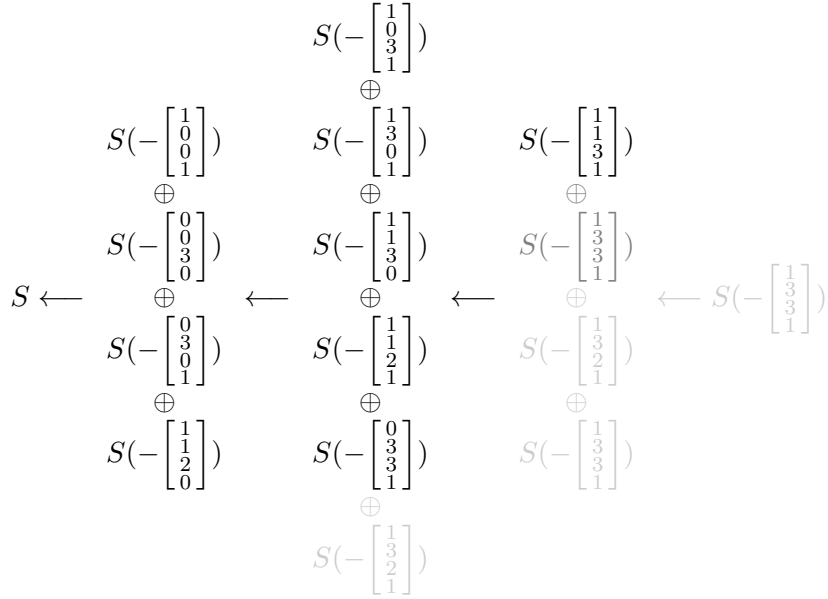


FIGURE 1.8. The Scarf complex of $I = \langle xw, z^3, y^3w, xyz^2 \rangle \subset \mathbb{C}[x, y, z, w]$, indicated in black, is the subcomplex of the Taylor complex supported on all submodules $S(-a)$ such that a appears exactly once in the Taylor complex. The maps between modules are inherited from the Taylor complex. The minimal free resolution of S/I strictly contains the Scarf complex of I in this example; it is additionally supported on the submodule $S(-\begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix})$ in homological degree 3 (shown in dark gray).

EXAMPLE 1.3.34. The Scarf complex of $I = \langle xw, z^3, y^3w, xyz^2 \rangle \subset \mathbb{k}[x, y, z, w]$, from Example 1.3.27, is the subcomplex of the Taylor resolution indicated in black in Figure 1.8. In this example, I is not Scarf, since it is strictly contained in the minimal free resolution.

Computationally, the Taylor complex has the desirable feature that its modules and differentials have an explicit general formula. Naively, one could simply construct the Taylor complex first, then reduce each syzygy module to a minimal set of generators, much in the same way that we constructed the full 4×6 matrix in Equation (1.3.28), then noticed that one of the columns was already in the module generated by the other five. This idea falls apart with consideration of just how extreme the non-minimality of the Taylor complex can be. For instance, consider an ideal with 20 monomial generators in 4 variables. The Taylor complex has 2^{20} faces, one for each subset of the generators. On the other hand, by the Hilbert syzygy theorem (Theorem 1.3.10), the length of the minimal resolution is

at most 4. Therefore the superfluous basis elements of the Taylor complex must include *at least* all $\binom{20}{5}$ basis elements in homological degree 5, all $\binom{20}{6}$ in homological degree 6, and so forth—totaling over a million!

The Scarf complex also has an explicit general formula that makes it easy to write down, *and* it is at most the size of the minimal resolution, so we never do any “extra” work. Unfortunately, it isn’t guaranteed to be a resolution. The best situation would be to somehow know whether or not the Scarf complex will be a resolution before computing it. In fact, there is a simple condition on the minimal generators of a monomial ideal I that guarantees that I is Scarf. This is the topic of the next section.

1.3.6. Genericity and strong genericity.

DEFINITION 1.3.35 (Strongly generic monomial ideal). A monomial ideal I is *strongly generic* if no variable appears with the same nonzero exponent in two distinct minimal generators of I .

EXAMPLE 1.3.36. In $\mathbb{k}[x, y, z, w]$, $J_1 = \langle x^5yz^2, x^2w^6, y^3, zw^7 \rangle$ is strongly generic, but $J_2 = \langle x^5yz, x^2w^6, y^3, zw^7 \rangle$ is not, because z^1 is the highest power of z dividing two different minimal generators of J_2 .

In [BPS98], Bayer, Peeva and Sturmfels proved that strongly generic monomial ideals are always Scarf.

EXAMPLE 1.3.37. Because $J_1 = \langle x^5yz^2, x^2w^6, y^3, zw^7 \rangle \subset \mathbb{k}[x, y, z, w]$ is strongly generic, we can determine the Betti numbers of J_1 combinatorially. Every subset of size zero and one

has a unique lcm, so $\beta_0 = 1$ and $\beta_1 = 4$. The lcms of the other subsets are:

$$\begin{aligned}
\text{lcm}(x^5yz^2, x^2w^6) &= x^5yz^2w^6, & \text{lcm}(x^5yz^2, x^2w^6, y^3) &= x^5y^3z^2w^6, \\
\text{lcm}(x^5yz^2, y^3) &= x^5y^3z^2, & \text{lcm}(x^5yz^2, x^2w^6, zw^7) &= x^5yz^2w^7, \\
\text{lcm}(x^5yz^2, zw^7) &= x^5yz^2w^7, & \text{lcm}(x^5yz^2, y^3, zw^7) &= x^5y^3z^2w^7, \\
\text{lcm}(x^2w^6, y^3) &= x^2y^3w^6, & \text{lcm}(x^2w^6, y^3, zw^7) &= x^2y^3zw^7, \\
\text{lcm}(x^2w^6, zw^7) &= x^2zw^7, & & \\
\text{lcm}(y^3, zw^7) &= y^3zw^7, & \text{lcm}(x^5yz^2, x^2w^6, y^3, zw^7) &= x^5y^3z^2w^7.
\end{aligned}$$

The repeated lcm's are shown in gray. Discarding these, we find that $\beta_2 = 5$, $\beta_3 = 2$, and $\beta_4 = 0$.

The more general definition of a *generic* monomial ideal is found in Miller and Sturmfels [MS04]. First we need a related definition.

DEFINITION 1.3.38. For monomials $m = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $m' = x_1^{\beta_1} \cdots x_n^{\beta_n}$, we say that m *strongly divides* m' if $\alpha_i < \beta_i$ whenever $\alpha_i \neq 0$.

EXAMPLE 1.3.39. The monomial x_1x_3 strongly divides $x_1^2x_3^3$, but not $x_1x_3^3$.

DEFINITION 1.3.40 (Generic monomial ideal). A monomial ideal I is *generic* if whenever two distinct minimal generators g_i and g_j have the same positive degree in some variable, a third generator g_k strongly divides $\text{lcm}(m_i, m_j)$.

Miller and Sturmfels proved that monomial ideals that are generic in this broader sense are always Scarf, too.

EXAMPLE 1.3.41. The ideal $J_2 = \langle x^5yz, x^2w^6, y^3, zw^7 \rangle$ from Example 1.3.36 is generic but not strongly generic. Even though x^5yz and zw^7 share the same positive power of z , their lcm is x^5yzw^7 , and there is another minimal generator, namely x^2w^6 , which strongly

divides this. An enumeration of the lcd's of subsets reveals that J_2 has the same total Betti numbers as J_1 , namely $\beta = (1, 4, 5, 2)$.

Although being generic/strongly generic is *sufficient* for a monomial ideal to be Scarf, it is not necessary. For example, the ideal

$$I = \langle x_1^2 x_6 x_8^2, x_1 x_4 x_7^2 x_8, x_1 x_3 x_4 x_5 x_8, x_1 x_2 x_5 x_6 x_7, x_1^2 x_3^2 x_5 \rangle$$

of $\mathbb{k}[x_1, \dots, x_8]$ is resolved by its Scarf complex even though it is not generic. To learn how to generate these kinds of examples, see Section 5.2.

1.4. Hilbert functions and series

Let \mathbb{k} be a field, and R a \mathbb{k} -algebra.

DEFINITION 1.4.1 (Graded ring). A \mathbb{Z} -grading on R is a decomposition

$$(1.4.2) \quad R = \bigoplus_{d \in \mathbb{Z}} R_d,$$

that is compatible with the algebra multiplication: for $f \in R_a$ and $g \in R_b$, $fg \in R_{a+b}$. We say that R is a *graded ring* or *graded algebra*.

An example of a graded ring is the polynomial ring $S = \mathbb{k}[x_1, \dots, x_n]$, where each S_d is the set of homogeneous polynomials of degree d . This grading is called the *standard grading* of $\mathbb{k}[x_1, \dots, x_n]$. If I is a homogeneous ideal of S , then the quotient ring $R = S/I$ is also graded, with the direct sum decomposition:

$$(1.4.3) \quad R = \bigoplus_{d \geq 0} (S/I)_d = \bigoplus_{d \geq 0} S_d / (I \cap S_d).$$

DEFINITION 1.4.4 (Hilbert function). The *Hilbert function* of a graded ring R , written $h(R, d)$, is the function $h(R, d) : \mathbb{Z} \rightarrow \mathbb{Z}_{\geq 0}$ defined by

$$(1.4.5) \quad h(R, d) = \dim_{\mathbb{k}} R_d.$$

In other words, the Hilbert function maps d to the dimension, as a vector space over \mathbb{k} , of the d -graded part of the ring R .

DEFINITION 1.4.6. The *Hilbert series* of R is the generating function of the Hilbert function $h(R, d)$. That is,

$$(1.4.7) \quad H(R, t) = \sum_{d \in \mathbb{Z}} h(R, d)t^d.$$

Of particular interest to us is the case when $R = S/I$ for a homogenous ideal I of a polynomial ring S . In this case the Hilbert function always has a rational generating function which takes the form

$$(1.4.8) \quad H(R, t) = \frac{P(t)}{(1-t)^k},$$

where $P(t)$ is a polynomial, and $k = \dim R$ [Sta07].

When I is a monomial ideal, the Hilbert function of S/I counts the number of monomials in each degree that are not in I . These monomials are called the *standard monomials* of S/I . The set of standard monomials can be visualized geometrically as a staircase in the lattice of monomials, as illustrated in Figure 1.9.

1.4.1. Bounds on Hilbert functions. Not all non-negative integer sequences can be the Hilbert function of a graded \mathbb{k} -algebra. In 1927, Macaulay gave a complete characterization of allowable Hilbert functions [Mac27], which is briefly summarized here. First, we need a lemma about the so-called *Macaulay expansion* of an integer:

LEMMA 1.4.9. Let j be a positive integer. Then each positive integer a has a unique expansion

$$a = \binom{a_j}{j} + \binom{a_{j-1}}{j-1} + \cdots + \binom{a_k}{k},$$

where $a_j \geq a_{j-1} \geq a_k \geq k \geq 1$.

For a proof of this and Theorem 1.4.11, see the original paper [Mac27], or a more modern exposition such as [HH11].

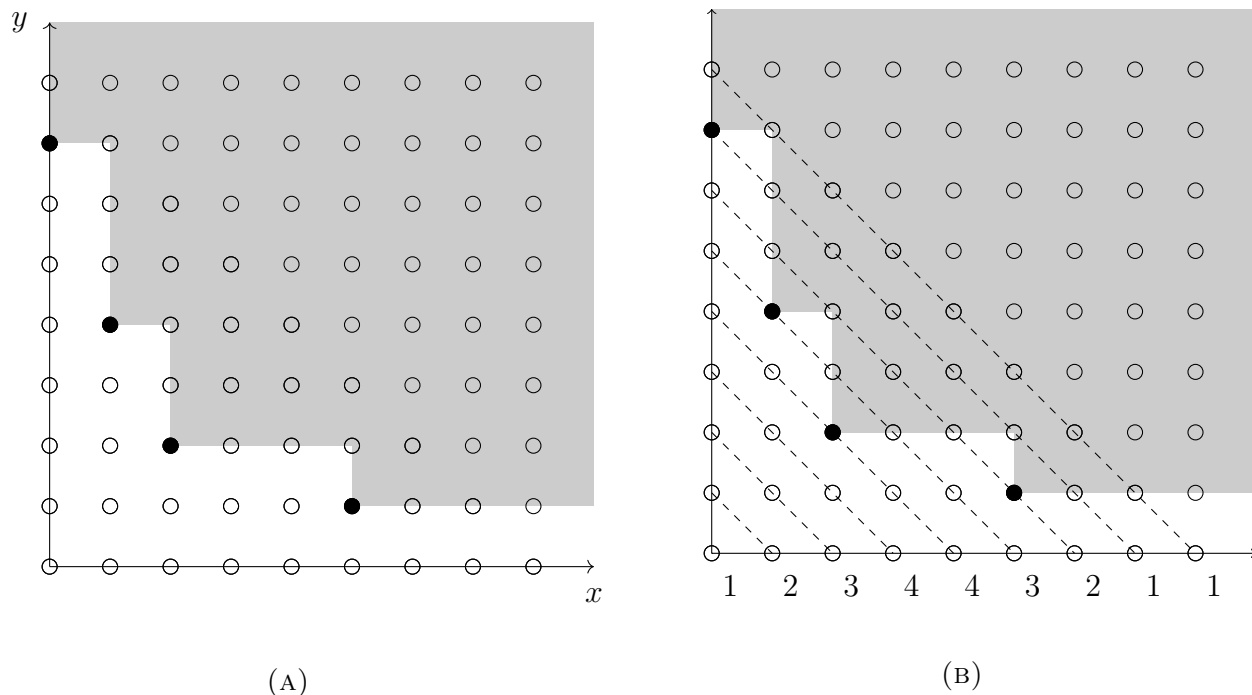


FIGURE 1.9. The staircase diagram of a monomial ideal I in two variables x, y (left). Every lattice point (a, b) that is “under the staircase” corresponds to the standard monomial $x^a y^b$ of S/I . The Hilbert function of S/I counts the number of standard monomials of each degree. This can be visualized as counting lattice points “under the staircase” in each degree (right).

DEFINITION 1.4.10. Let $a = \binom{a_j}{j} + \binom{a_{j-1}}{j-1} + \cdots + \binom{a_k}{k}$ be the Macaulay expansion of a with respect to j . Define

$$a^{(j)} = \binom{a_j + 1}{j + 1} + \binom{a_{j-1} + 1}{j} + \cdots + \binom{a_k + 1}{k + 1}.$$

THEOREM 1.4.11 (Macaulay’s Theorem). Let $h : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$. The following are equivalent:

- (1) h is the Hilbert function of a standard graded \mathbb{k} -algebra
- (2) $h(0) = 1$, and $h(j + 1) \leq h(j)^{(j)}$ for all $j > 0$.

Macaulay’s theorem is actually constructive—his proof describes how to create a particular kind of monomial ideal, called a *lex-segment ideal*, with Hilbert function h , where h is any function satisfying Theorem 1.4.11. A monomial ideal I is *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 consists

of the $\dim_{\mathbb{k}} I_j$ lexicographically largest monomials in S_j . Macaulay's theorem shows that for any Hilbert function h of a homogenous ideal, there exists a unique lex-segment ideal I_{lex} such that $h(S/I_{lex}) = h$.

Further consequences of Macaulay's theorem come from the intimate relationship between the Hilbert function and the Betti numbers of a graded \mathbb{k} -algebra. Macaulay's paper also proved that for a given Hilbert function, the lex-segment ideal with that Hilbert function will always have maximum total Betti numbers among the set of ideals with the Hilbert function. In 1993 Bigatti [Big93] and, independently, Hulett [Hul93] established that lex-segment ideals also maximize *graded* Betti numbers, as long as \mathbb{k} has characteristic zero. This established explicit bounds on the growth of graded Betti numbers of a monomial ideal. In 1996 Pardue [Par96] extended this theorem to an arbitrary field.

THEOREM 1.4.12 (Bigatti-Hulett-Pardue theorem). Let I be an ideal of the polynomial ring $S = \mathbb{k}[x_1, \dots, x_n]$ and let I_{lex} be the unique lex-segment ideal such that $h(S/I) = h(S/I_{lex})$. Then, the corresponding graded Betti numbers satisfy the inequality $\beta_{i,j}(I) \leq \beta_{i,j}(I_{lex})$ for all i, j .

1.4.2. Recursive computation of Hilbert series. By computing an initial ideal, finding the Hilbert series $H(S/I, t)$ of any ideal reduces to finding the Hilbert series of a monomial ideal. State-of-the-art implementations for computing Hilbert series are based on a recursive algorithm introduced in [BS92], and improved in [Big97] with the use of *pivots*.

In each step of the algorithm, a pivot monomial \mathcal{P} is chosen which strictly divides at least one minimal generator of I . Then by the exact sequence

$$(1.4.13) \quad 0 \longrightarrow S/\langle I : \mathcal{P} \rangle \xrightarrow{\cdot \mathcal{P}} S/I \longrightarrow S/(I + \langle \mathcal{P} \rangle) \longrightarrow 0,$$

the computation is decomposed, with respect to \mathcal{P} :

$$(1.4.14) \quad H(S/I, t) = H(S/(I + \langle \mathcal{P} \rangle), t) + t^{\deg \mathcal{P}} H(S/\langle I : \mathcal{P} \rangle, t).$$

The recursion continues until hitting *base cases*, monomial ideals with one of the following two structures:

- (1) Ideals generated only by *pure powers* of the variables; i.e. $I = \langle x_{i_1}^{a_1}, \dots, x_{i_k}^{a_k} \rangle$. In this case the numerator of the Hilbert series of I is $\prod_{j=1}^k 1 - t^{a_j}$.
- (2) Ideals generated by pure powers along with one non-trivial power product $x_1^{b_1} \cdots x_n^{b_n}$. In this case the Hilbert series of I is $\prod_{j=1}^k (1 - t^{a_j}) - t^{|b|} \left(\prod_{j=1}^k (1 - t^{a_j - b_j}) \right)$.

There are many choices for \mathcal{P} at every step, and many “pivot rules” for making this choice. With any pivot rule the correct answer is eventually computed, but the number of base cases and timing of the algorithm are highly dependent on the rule, and no single heuristic performs best over all kinds of input. In [Big97] Bigatti demonstrates this variability with several explicit examples. In Section 7.2 we will explain the various pivot rule choices in detail, and show how machine learning can be applied to the algorithm selection problem of how to choose a pivot rule that performs this recursive computation using the least number of steps/base cases.

1.5. Supervised machine learning

Machine learning is rapidly becoming ubiquitous in artificial intelligence, data analysis, and applied mathematics. The fundamental principle underlying machine learning is that meaningful patterns and predictions can be extrapolated from large quantities of empirical data, even in the absence of any known logical or scientific framework.

For example, a dermatologist learning to diagnose skin cancer is typically given logical instructions for differentiating suspicious skin lesions and moles, like, “if the shape is asymmetrical, perform a biopsy.” These classification rules are imperfect, but simple to state and remember. Perhaps we suspect that cancerous lesions have other characteristic features too subtle for humans to recognize, and want to train a computer, equipped with a camera far superior to the human eye, to attempt this classification task. Rather than programming it with logical rules, the computer’s software would instead be fed a *training set* of possibly millions of images, each with a label of 1=benign or -1 =malignant. Each image x is a

matrix of pixel values in some space X of real matrices. The software’s task is to find, i.e. “learn”, a function $h : X \rightarrow \{1, -1\}$ that “fits” the training images in the sense that $h(x)$ returns the correct label for x for as many of the training images x as possible. Now, when the camera sees an unknown image z , the value of $h(z)$ predicts whether this new image is benign or malignant. This is an example of *supervised machine learning*, since the training data was correctly labeled using outside supervision (for instance, actual biopsy results). An immense, rapidly-growing body of literature is concerned with the theory and best practices of how to fit a *hypothesis function*, h , to the given training data to maximize predictive accuracy. Section 1.5.1 only introduces some basics.

In any supervised learning problem there is an indispensable, but often unspoken, assumption being made: that the training examples have patterns that are representative of the larger class of unknown instances. For identifying skin cancer, it seems reasonable to expect that the features of malignant skin growths in a sample of human patients are representative of their features in the general human population. If the training images were of rats, this might not be a reasonable assumption. A different kind of unreasonableness would be to attempt to train an algorithm to predict winning lottery numbers. No matter how many training examples we provide, of sequences of numbers labeled with whether or not they won that day, and no matter how expertly we designed the training model and tuned its parameters, there is no meaningful pattern or correlation that could predict numbers from a uniformly random distribution.

In 1993, Bayer and Mumford famously asked: *What can be computed in algebraic geometry?* [BM92]. One question posed in this thesis is, instead: *What can be predicted in algebraic geometry?* In other words, are the algebraic properties of rings, ideals, and varieties governed by subtle patterns that, though they have thus far evaded rigorous mathematical proof, can nonetheless be learned from data and used to make predictions for new examples? Or do the invariants of these objects behave so randomly and unpredictably that machine learning is impossible?

Chapter 7 explores this question in two ways. In Section 7.2, we consider *algorithm selection* in computer algebra, for the specific problem of learning how to choose good pivots for the recursive Hilbert series computation described in Section 1.4.2. We show that training a neural network to make these decisions on the fly results in better performance than using any fixed choice of pivot rule. In Section 7.3, we try to directly predict the values of several NP-hard invariants: dimension, projective dimension, and regularity. The predictions are not always correct, but are very often no more than one away from the correct answer.

Both tasks are modeled as supervised learning problems, where the training inputs are statistics on minimal generating sets of monomial ideals (see Section 7.1), and the training labels are either the best choice of pivot, or the correct value of the algebraic invariant, as appropriate, for each ideal. The success of both experiments establishes beyond a doubt the feasibility of machine learning for problems in commutative algebra and algebraic geometry. The particular methodology chosen—supervised learning using hand-crafted features to train a neural network with the particular architecture and loss function described in the next section—found patterns in the training data that successfully extrapolated to new unseen examples. There are other machine learning methods and variations that could be applied to these problems, and the pursuit of better performance and robustness is just beginning. However, as a proof of concept, Chapter 7 of this thesis establishes for the first time the feasibility of machine learning in commutative algebra.

We now explain the particular methodology used.

1.5.1. Supervised learning on neural networks. A *neural network*, whose name derives from its origins as a simple model of neuronal activity in the human brain, is a directed graph consisting of nodes (neurons) and weighted directed edges. See Figure 1.10. The first layer of the network is the entries of $x^{(0)} = (x_1^{(0)}, \dots, x_d^{(0)})$, a vector in \mathbb{R}^d representing the input to the problem. For learning images, for example, the value of $x_i^{(0)}$ could be the intensity of the i th pixel.

The *hidden layers* of the network contain nodes whose values are determined by the weights of the directed edges and a chosen *activation function* θ . Let $x_i^{(l)}$ represent the i th

node in the l th layer, and let $w_{ij}^{(l)}$ represent the weight on the directed edge from $x_i^{(l)}$ to $x_j^{(l+1)}$. Then the value of $x_j^{(l+1)}$ is calculated via

$$x_j^{(l+1)} = \theta(w_j^{(l)} \cdot x^{(l)}) = \theta \left(\sum_{i=1}^{m_l} w_{ij}^{(l)} x_i^{(l)} \right),$$

where m_l is the number of nodes in layer l . See Figure 1.10 for an illustration. It is typical to add a first entry $x_0^{(l)} = 1$ to each layer so that $w^{(l)} \cdot x^{(l)}$ can be an affine function of $x^{(l)}$, and not just linear.

In the first neural networks, which used the so-called *perceptron* model, the activation function θ was the sign function:

$$\theta(x) = \text{sign}(x) = \begin{cases} 1 & x \geq 0, \\ -1 & x < 0. \end{cases}$$

A friendly introduction to perceptron learning is the textbook [AMMIL12]. The sign function is reminiscent of the way that a biological neuron will either fire or not, depending on if it receives enough electrical signals to surpass some threshold, with no middle ground. Since neural networks have been repurposed for machine learning, several other activation functions have become popular, including the *sigmoid function*,

$$\theta(x) = \sigma(x) = \frac{e^x}{1 + e^x},$$

the *hyperbolic tangent* function,

$$\theta(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

and the *rectified linear unit function* (*ReLU*),

$$\theta(x) = \text{ReLU}(x) = \begin{cases} x & x \geq 0, \\ 0 & x < 0. \end{cases}$$

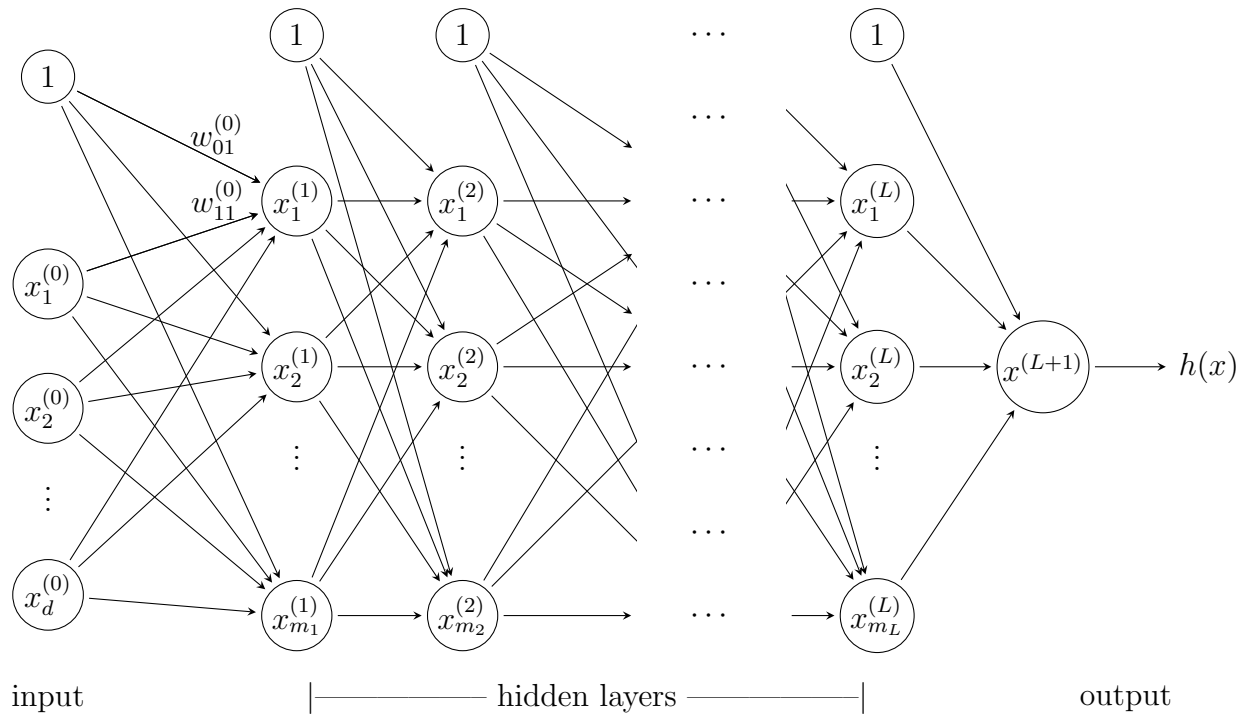
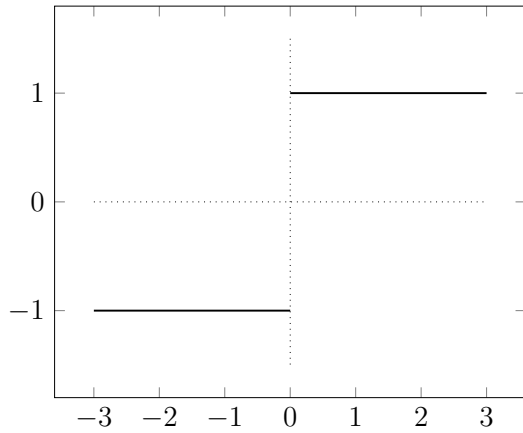
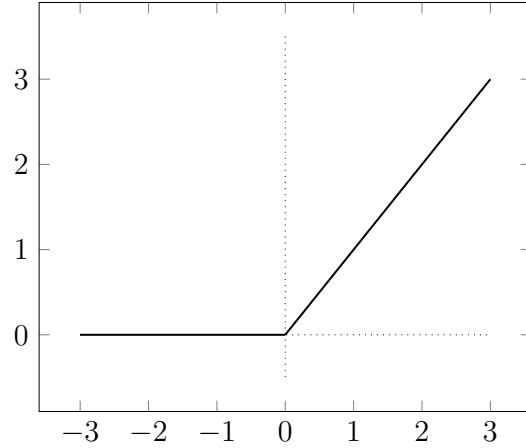


FIGURE 1.10. The anatomy of a neural network. The directed edge from node $x_i^{(l)}$ to node $x_j^{(l+1)}$ has weight $w_{ij}^{(l)}$, and the value of node $x_j^{(l+1)}$ is calculated by $\theta(w_j^{(l)} \cdot x^{(l)})$, for a chosen activation function θ . Only a few weights have been labeled in the figure, for better readability. Input nodes (on the left) are the data or features of each input x , and the output, $h(x)$, is calculated by $\theta(x^{(L+1)})$. The goal is to find the weights that maximize the number of correct outputs over the set of training examples.

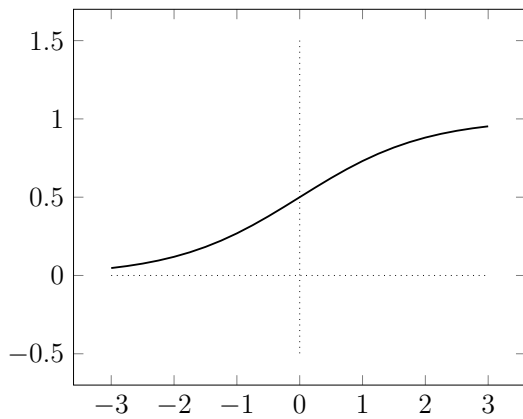
All four of these common choices for θ are depicted in Figure 1.11. One advantage of the sigmoid and hyperbolic tangent functions is that they are everywhere differentiable. An advantage of ReLU is that it maintains the piecewise linear simplicity of the perceptron model, but allows each neuron to be “more or less” activated depending on the signals it receives, instead of only on or off. The ReLU function, though not differentiable at $x = 0$, nonetheless has an easily computed piecewise linear derivative.



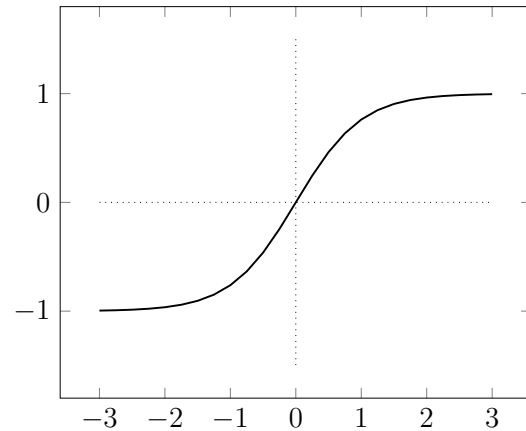
(A) Sign function, $\text{sign}(x)$



(B) Rectified linear unit function, $\text{ReLU}(x)$



(C) Sigmoid function, $\sigma(x)$



(D) Hyperbolic tangent function, $\text{tanh}(x)$

FIGURE 1.11. Some common choices for activation functions inside neural networks.

The importance of differentiability is its key role in *training* the neural network, which is the process of finding optimal values for the weights $w_{ij}^{(l)}$ in the network. By optimal, we mean the weights that minimize an objective function called the *loss function* of the network. There are multiple ways to define the loss, or error, of the network on the training data set. For classification problems, the loss function could be as simple as the number of misclassified training examples:

$$(1.5.1) \quad L(X) = \sum_{x \in X} \mathbf{1}_{h(x) \neq t(x)},$$

where X is the set of training inputs, $h(x)$ is the label for $x \in X$ that was output by the neural network, and $t(x)$ is the true label for x (which is known by us, the omniscient supervisors). The loss function in Equation (1.5.1) makes sense if $h(x)$ is, for instance, the output from a perceptron node, which always equals -1 or 1 .

In Chapter 7, we consider problems in the realm of *multi-class classification*, as opposed to simpler *binary classification* such as benign vs. malignant tumors. For instance, to use an actual example from Chapter 7, suppose the input is monomial ideals generated in no more than 20 variables, and the goal is predicting their projective dimension. By the Hilbert Syzygy Theorem (Theorem 1.3.10), the possible values/labels are integers between 0 and 20. (Recall that $\text{pdim}(S/I) = 0$ if and only if $I = \langle 0 \rangle$.) To predict the right label, the last layer of the network contains 21 nodes, one for each label, equipped with sigmoid activation functions, so that we output a vector s containing 21 real numbers between 0 and 1. For instance:

$$(1.5.2) \quad s(I) = [0, 0, 0, 0.05, 0.08, 0.23, 0.48, 0.98, 0.07, 0.12, 0, 0.02, 0, 0.01, 0, 0, 0, 0, 0, 0].$$

The value of the i th entry is interpreted as a prediction of how “confident” the network is that I has projective dimension i . Of course, to interpret $h(I)$ as a probability distribution, the vector has to be normalized in some way. We do this by applying the so-called *softmax* function:

$$(1.5.3) \quad f(s)_i = \frac{e^{s_i}}{\sum_{j=1}^{21} e^{s_j}}.$$

For instance, applying a softmax to Equation (1.5.2) results in the probability distribution:

$$(1.5.4) \quad h(I) = f(s(I)) = [0.04, 0.04, 0.04, 0.04, 0.05, 0.05, 0.07, 0.11, 0.04, 0.05, \\ 0.04, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04],$$

after rounding to two decimal places. From either Equation (1.5.2) or Equation (1.5.4) we take the maximum entry as the predicted output; in this case, we predict that $\text{pdim}(S/I) = 7$.

Now, rather than simply tallying the number of correct versus incorrect predictions, we use a *cross-entropy loss function*, defined on a single example x by:

$$(1.5.5) \quad L_{CE}(x) = - \sum_{i=1}^{21} t(x)_i \ln h(x)_i,$$

where $t(x)$ is the true labeling of x (with a one in the entry corresponding to the correct label, and zeroes elsewhere) and $h(x)$ is the distribution predicted by the softmax function applied to the output, as in Equation (1.5.4). Note that by the definition of the softmax function, all entries of $h(x)$ are strictly positive and thus Equation (1.5.5) is well-defined.

In our example, if the true projective dimension were indeed 7, then $t(I)$ would be the standard basis vector e_8 , and $L_{CE}(I)$ would be $-\ln(0.11) \approx 2.2$. If, on the other hand, the true projective dimension were 8, we would calculate $L_{CE}(I) = -\ln(0.05) \approx 3.0$.

The loss for a training set X is the sum over all $x \in X$ of $L_{CE}(x)$:

$$(1.5.6) \quad L_{CE}(X) = \sum_{x \in X} L_{CE}(x),$$

and our final optimization problem is: find the weights $w_{ij}^{(l)}$ of the neural network which minimize $L_{CE}(X)$.

To find (or at least approximate) the optimal weights, we use the standard machine learning technique of *gradient descent*. Informally, the partial derivatives of the loss function, with respect to the individual neurons in the network, show the contribution each weight makes to the total error; the weights are then nudged in the direction of decreasing the error. This process is iterated until the gradient is zero, demonstrating that we have found a local (and hopefully global) minimizer of the loss function. Since gradient descent is a central topic of machine learning theory, we refer the interested reader to introductory texts such as [AMMIL12, Str19]. One of the concerns of this theory is how to choose good *step sizes* for the gradient descent algorithm, determining the amounts that the weights are adjusted in each iteration, so that the algorithm converges quickly without getting stuck in a shallow local minimum. See, e.g., [TMDQ16] and the references therein.

CHAPTER 2

Random models for monomial ideals

This chapter introduces some natural combinatorial models for studying the probabilistic behavior of monomial ideals and their coordinate rings.

This is not the first time that problems in commutative algebra have been studied probabilistically. Since at least the 1930's, when Littlewood and Offord studied the expected number of real roots of a polynomial with random coefficients [LO38], algebraists and algebraic geometers have studied the typical properties of polynomials with *generic* coefficients. Work now considered classical includes the study of random varieties, defined by random coefficients on a fixed Newton polytope support, as in [Kac43, Kou76, Stu98] and the references therein. The field of *smooth analysis* studies how algorithmic performance varies under random perturbations of the problem input. Contributions to algebraic geometry based on this method of analysis include [BP09] and [BC11]. More recently, Ein, Erman, and Lazarsfeld [EEL15] (see also [EL12, EEL16]), studied the Betti numbers of modules defined by uniformly random Boij-Söderberg coefficients [ES09].

Common to all of these examples, however, are probability models based on random *coefficients*. In contrast, this thesis studies discrete random models, inspired by the combinatorial theory of commutative algebra developed in texts including [Sta07, BS92, BS87b, BPS98, MS04, GS93, Hà14, HH11, CP17]. The heart of combinatorial commutative algebra is that algebraic properties of polynomial ideals and rings may be derived from combinatorial relationships among the monomials appearing in polynomial generating sets.

Via the *Stanley-Reisner correspondence* (Section 1.2.2), monomial ideals are a generalization of simplicial complexes, and thus of graphs. Ever since the seminal paper of Erdős and Rényi [ER59], probabilistic methods have been indispensable to graph theory (see, e.g., [AS92, Bol01] and the references therein). The astonishing discovery of Erdős and Rényi

was that global properties of random combinatorial objects, even ones which are hard to characterize in general, display emergent behaviors like *phase transitions* and *thresholds* when studied asymptotically. This work continues in their tradition by proving that important global properties of random monomial ideals, and their coordinate rings, are also governed by phrase transitions and thresholds. Identifying these emergent phenomena sheds light on the theory of combinatorial commutative algebra, and informs computation by quantifying “typical behavior.”

2.1. The Erdős-Rényi-type model

DEFINITION 2.1.1 (Erdős-Rényi-type random monomial ideals). Let \mathbb{k} be a field and $S = \mathbb{k}[x_1, \dots, x_n]$.

Given an integer D and a parameter $p = p(n, D)$, $0 \leq p \leq 1$, sample a random generating set of monomials \mathfrak{B} according to

$$(2.1.2) \quad \mathbb{P}[x^\alpha \in \mathfrak{B}] = p \text{ for all } x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in \mathbb{k}[x_1, \dots, x_n] \text{ with } 1 \leq |\alpha| \leq D$$

Denote by $\mathcal{B}(n, D, p)$ the resulting Erdős-Rényi-type distribution on sets of monomials of total degree at most D in $\mathbb{k}[x_1, \dots, x_n]$. 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

$$(2.1.3) \quad \mathbb{P}[\mathfrak{B} = B] = p^{\#B} (1 - p)^{\binom{D+n}{D} - \#B - 1}.$$

Since an ideal may be generated by many different sets B , $\mathcal{B}(n, D, p)$ induces a different distribution on ideals of $\mathbb{k}[x_1, \dots, x_n]$, which will be denoted by $\mathcal{I}(n, D, p)$. The notation $\mathfrak{J} \sim \mathcal{I}(n, D, p)$ indicates that $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $\mathfrak{J} = \langle \mathfrak{B} \rangle$.

THEOREM 2.1.4. Let $I \subseteq S$ be a fixed monomial ideal generated in degree at most D , and let $\mathfrak{J} \sim \mathcal{I}(n, D, p)$. Then

$$(2.1.5) \quad \mathbb{P}[\mathfrak{J} = I] = p^{\beta_1(S/I)} (1 - p)^{\sum_{d=1}^D h(S/I, d)}.$$

PROOF. Fix $I \subseteq S$ generated in degree at most D and let G be the unique minimal set of generators of I . Then, $\mathfrak{J} = I$ if and only if $\mathfrak{B} \supseteq G$ and no monomial x^α such that $x^\alpha \notin I$ is in \mathfrak{B} . Let A_1 denote the event that each of the $\beta_1(S/I)$ elements of G is in \mathfrak{B} and let A_2 denote the event that no monomial x^α such that $x^\alpha \notin I$ is in \mathfrak{B} . Then, the event $\mathfrak{J} = I$ is equivalent to the event $A_1 \cap A_2$. Since the events A_1 and A_2 are independent, $\mathbb{P}[\mathfrak{J} = I] = \mathbb{P}[A_1 \cap A_2] = \mathbb{P}[A_1] \mathbb{P}[A_2]$. Observe that $\mathbb{P}[A_1] = p^{\beta_1(S/I)}$, since each of the $\beta_1(S/I)$ elements of G is chosen to be in \mathfrak{B} independently with probability p and $\mathbb{P}[A_2] = (1 - p)^{\sum_{d=1}^D h_I(d)}$, since there are exactly $\sum_{d=1}^D h(S/I, d)$ monomials of degree at most D not contained in I (recall we do not consider the constant monomial), and each of them is excluded from \mathfrak{B} independently with probability $1 - p$. \square

2.2. The graded model

DEFINITION 2.2.1. Fix a degree bound $D > 0$. The *graded model* for random monomials selects generators according to probabilities that depend on total degree. That is, a set of D parameters $\mathbf{p}(n, D) = (p_1(n, D), \dots, p_D(n, D))$ with $0 \leq p_i(n, D) < 1$ for each i , is selected, and then the random generating set \mathfrak{B} is sampled according to:

$$\mathbb{P}[p_\alpha \in \mathfrak{B}] = p_\alpha.$$

With respect to any particular total degree $|\alpha|$, the probability of each monomial with that degree being chosen is the same as any other.

The analogue of Theorem 2.1.4 for $\mathcal{G}(n, D, \mathbf{p})$, which has an almost identical proof, is as follows:

THEOREM 2.2.2. Fix n, D , and the graded model parameters $\mathbf{p}(n, D) = (p_1, \dots, p_D)$. For any fixed monomial ideal $I \subseteq S$, random monomial ideals from the graded model distribution $\mathfrak{J} \sim \mathcal{G}(n, D, \mathbf{p})$ satisfy the following:

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

where $\beta_{1,d}(S/I)$ is the number of degree- d minimal generators of I (see Definition 1.3.14), and $h(S/I, d)$ is its Hilbert function.

2.2.1. A special case of the graded model. In Chapters 4 and 5, we consider random monomial ideals in n variables, where each monomial of degree D has the same probability p of appearing as a generator. We define this as follows:

DEFINITION 2.2.3 (Graded model for random monomial ideals). A generating set \mathfrak{G} is sampled according to

$$\mathbb{P}[x^\alpha \in \mathfrak{G}] = \begin{cases} p & |\alpha| = D, \\ 0 & \text{otherwise,} \end{cases}$$

for all $x^\alpha \in S = \mathbb{k}[x_1, \dots, x_n]$. We then set $\mathfrak{M} = \langle \mathfrak{G} \rangle$.

It is important to note that in Definition 2.2.3, \mathfrak{G} is always a minimal generating set. Given the three parameters n , D , and p , we denote this model by $\mathcal{M}(n, D, p)$, and write $\mathfrak{M} \sim \mathcal{M}(n, D, p)$.

This notation is chosen to make it easy to distinguish ideals coming from the ER-type model, $\mathfrak{J} \sim \mathcal{I}(n, D, p)$, from ideals in this special graded model, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$. Though the symbol \mathfrak{G} was used in Section 1.1.2 to indicate a random graph, this should cause no confusion with its use in later chapters to indicate a random minimal generating set in the special graded model.

2.3. The general model

Fix a degree bound $D > 0$. To *each* monomial $x^\alpha \in \mathbb{k}[x_1, \dots, x_n]$ 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 :

$$(2.3.1) \quad \mathbb{P}[x^\alpha] = p_\alpha.$$

Hence the general model has many parameters, namely $\{p_a : a \in \mathbb{N}^n \setminus \{0\}, |a| \leq D\}$. 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 $p_a = p(n, D)$ does not depend on a .

2.4. Specialization to random simplicial complexes and graphs

A corollary of Theorem 2.1.4 follows directly from the well-known formulas for Stanley-Reisner rings:

COROLLARY 2.4.1. 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.

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_a in (2.3.1) such that the resulting distribution on square-free monomial ideals in $S = \mathbb{k}[x_1, \dots, x_n]$ is precisely the distribution on the abstract simplicial complexes on $[n]$ under the Costa-Farber model.

THEOREM 2.4.2 (See [Wil18]). Let $\tilde{\mathbf{p}} = (\tilde{p}_0, \tilde{p}_1, \dots, \tilde{p}_r, 0, \dots, 0)$ 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_a , where $a \in \mathbb{N}^n$ and $0 < \|a\|_1 \leq r + 1$, for the general monomial model (2.3.1):

$$(2.4.3) \quad p_a = \begin{cases} 1 - \tilde{p}_{\deg(x^a)-1}, & \text{if } 0 \neq a \in \{0, 1\}^n, \\ 0, & \text{otherwise.} \end{cases}$$

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 (2.3.1).

In other words, the specification of probabilities in Theorem 2.4.2 recovers the Costa-Farber model on random simplicial complexes as a sub-model of the model (2.3.1) 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 [CF16, 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 (2.4.3) 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 (2.4.3), so $p_j = 1 - \tilde{p}_{j-1}$. We now apply Theorem 2.2.2 to conclude that

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

□

REMARK 2.4.4. There are other models that do not sample individual monomials, but instead sample ensembles of monomials all at the same time. Examples of these include methods to generate random integer partitions and Ferrers diagrams [Pit97], and random lattice polytopes [BM05].

CHAPTER 3

Krull dimension

3.1. Dimension as a vertex cover problem

Recall that a *hypergraph* $H = (V, E)$ is defined as a set V , called the *vertex set* of H , together with a collection E of subsets of V . We call E the *edges* of H . This definition is a strict generalization of the definition of a graph, since a graph is just a hypergraph where every edge has cardinality two. A *clutter* is a hypergraph where no edge is contained in any other. For the basics of hypergraphs see [Cor01].

A *vertex cover* of H , also called a *hitting set* or a *transversal*, is a set $T \subseteq V$ such that every edge in E has a nonempty intersection with T . A vertex cover T is *minimal* if no proper subset of T is a vertex cover. The *vertex cover number* of H , also called *transversal number*, is denoted $c(H)$ and defined as the minimum cardinality of a vertex cover of H .

DEFINITION 3.1.1 (Support hypergraph). Recall that the *support* of $x_1^{a_1} \cdots x_n^{a_n}$ is the set $\{x_i : a_i > 0\}$. Given a set of monomials $B = \{m_1, \dots, m_r\}$ in $\mathbb{k}[x_1, \dots, x_n]$, define the *support hypergraph* of B , written \mathcal{H}_B , by:

- (1) $V(\mathcal{H}_B) = \{x_1, \dots, x_n\}$, and
- (2) $E(\mathcal{H}_B) = \{\text{supp}(m_i) : 1 \leq i \leq r\}$.

PROPOSITION 3.1.2. Let $B = \{m_1, \dots, m_r\}$ be a set of monomials in $S = \mathbb{k}[x_1, \dots, x_n]$, and let $I = \langle B \rangle$. Then:

- (1) T is a vertex cover of \mathcal{H}_B if and only if $I \subseteq \mathfrak{p}_T$.
- (2) T is a minimal vertex cover of \mathcal{H}_B if and only if \mathfrak{p}_T is a minimal prime of I .
- (3) $\text{codim } I = c(\mathcal{H}_B)$.
- (4) $\dim I = n - c(\mathcal{H}_B)$.

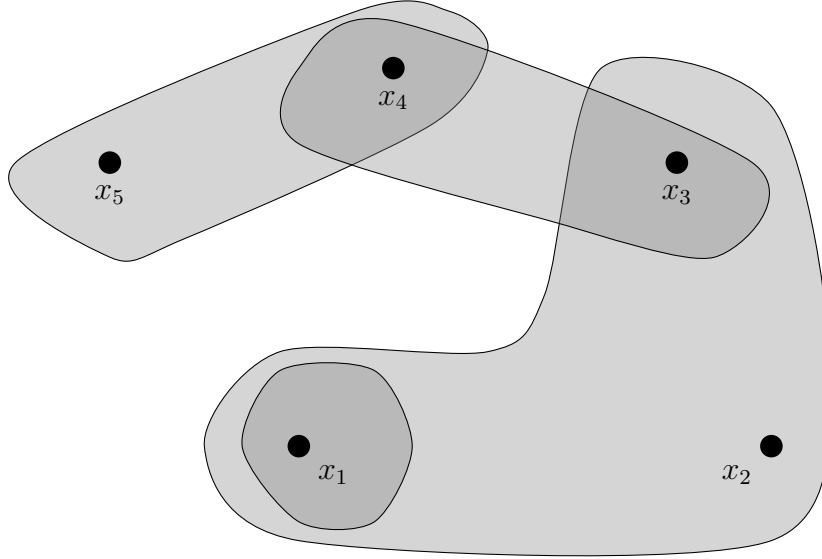


FIGURE 3.1. The support hypergraph of the monomials $B = \{x_1^3x_2^2x_3, x_1^4, x_3^2x_4^2, x_4x_5^3\} \subset \mathbb{k}[x_1, \dots, x_5]$. Since the vertex cover number of \mathcal{H}_B is 2, the codimension of $I = \langle B \rangle$ is also 2. Therefore $\dim I = 3$.

PROOF. To prove statement (1), let T be a vertex cover of \mathcal{H}_B . For each $m \in B$, there exists $x_i \in T \cap \text{supp}(m)$. This means that x_i divides m and therefore $m \in \mathfrak{p}_T$. Since this holds for every generator m of the ideal I , $I \subseteq \mathfrak{p}_T$. On the other hand, if $I \subseteq \mathfrak{p}_T$ then every monomial in I is divisible by at least one $x_i \in T$. In particular every $m \in B$ is divisible by at least one $x_i \in T$ and thus $x_i \in T \cap \text{supp}(m)$. As this holds for all $m \in B$, T is a vertex cover of \mathcal{H}_B . Since $T_1 \subseteq T_2$ if and only if $\mathfrak{p}_{T_1} \supseteq \mathfrak{p}_{T_2}$, T is minimal if and only if \mathfrak{p}_T is, proving statement (2). From Equation (1.2.32), $\text{codim } I = \min\{\#T : I \subseteq \mathfrak{p}_T\}$, and this quantity equals $c(\mathcal{H}_B) = \min\{\#T : T \text{ is a vertex cover of } \mathcal{H}_B\}$ by statement (1). This proves statement (3), and since we are working in the ring S which has dimension n , statement (4) immediately follows. \square

EXAMPLE 3.1.3. Suppose $B = \{x_1^3x_2^2x_3, x_1^4, x_3^2x_4^2, x_4x_5^3\} \subset \mathbb{k}[x_1, \dots, x_5]$ and $I = \langle B \rangle$. The support hypergraph of B , \mathcal{H}_B is illustrated in Figure 3.1. The minimal vertex covers of the hypergraph are $\{x_1, x_4\}$ and $\{x_1, x_3, x_5\}$, so the vertex cover number is 2. Therefore $\dim I = 3$ by Proposition 3.1.2.

Proposition 3.1.2 was already well-established in the literature for hypergraphs with edges defined by minimal generating sets of squarefree monomial ideals; see for example Chapter 2 of [GV11]. Instead of the previous proof, we could have justified extending to arbitrary generating sets of arbitrary monomial ideals with a few observations. To argue that B can be non-minimal, note that the unique minimal generators G of a monomial ideal are a subset of any monomial generating set B . Thus any vertex cover of \mathcal{H}_B is also a vertex cover of \mathcal{H}_G . On the other hand, suppose T is a vertex cover of \mathcal{H}_G , and let $b \in B$. Since $I = \langle G \rangle$, b is divisible by at least one $g \in G$ and hence $\text{supp}(g) \subseteq \text{supp}(b)$. Since $T \cap \text{supp}(g) \neq \emptyset$, it follows that $T \cap \text{supp}(b) \neq \emptyset$. Therefore T is a vertex cover of \mathcal{H}_B as well. This shows that \mathcal{H}_G and \mathcal{H}_B have the same vertex covers and therefore Proposition 3.1.2 holds for any monomial generating set of I .

To extend to non-squarefree monomial ideals, note that $\mathcal{V}(\sqrt{I}) = \mathcal{V}(I)$ and therefore $\dim \sqrt{I} = \dim I$. Since $B' = \{\text{SF}(b) : b \in B\}$ is a set of monomials generating \sqrt{I} , by Proposition 1.2.21, $\dim I = n - c(\mathcal{H}_{B'})$. Now we simply observe that $\mathcal{H}_B = \mathcal{H}_{B'}$ by the construction of the support hypergraph.

It is very convenient that Proposition 3.1.2 gives a combinatorial description of the dimension and codimension of $I = \langle B \rangle$ for any choice of B . We will see in Chapter 3 that when B comes from a random distribution, calculating the probabilistic dimension is simplified by being able to sidestep questions about minimality.

On the other hand, it is sometimes preferable to work with an irredundant set of hypergraph edges; for instance in the proof of Theorem 3.2.1. For this reason we will make an additional definition.

DEFINITION 3.1.4 (Reduced support hypergraph). For a monomial set $B = \{m_1, \dots, m_r\}$ in $\mathbb{k}[x_1, \dots, x_n]$, define the *reduced support hypergraph of B* , written $\mathcal{H}_B^{\text{red}}$, as the hypergraph obtained from \mathcal{H}_B by deleting all edges in $E(\mathcal{H}_B)$ that strictly contain another edge in $E(\mathcal{H}_B)$.

EXAMPLE 3.1.5. Suppose $B = \{x_1^3 x_2^2 x_3, x_1^4, x_3^2 x_4^2, x_4 x_5^3\} \subset \mathbb{k}[x_1, \dots, x_5]$, whose support hypergraph \mathcal{H}_B is illustrated in Figure 3.1. Since the edge $\{x_1, x_2, x_3\}$ of \mathcal{H}_B strictly contains

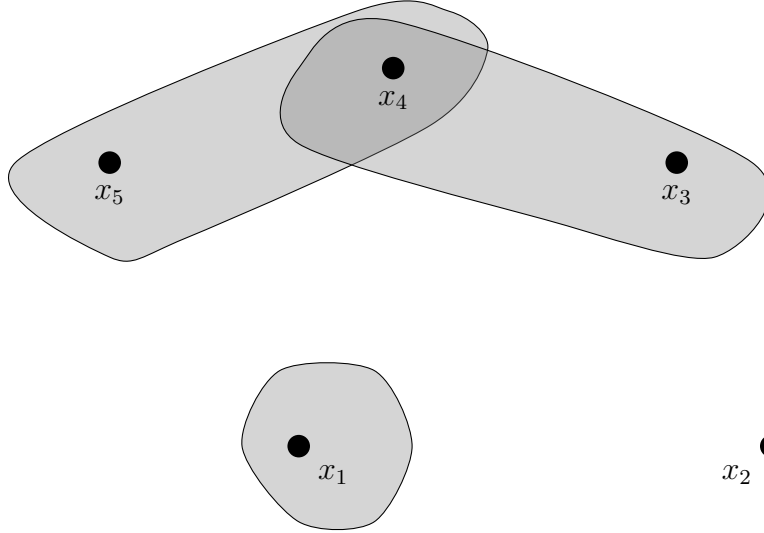


FIGURE 3.2. The reduced support hypergraph $\mathcal{H}_B^{\text{red}}$ for $B = \{x_1^3x_2^2x_3, x_1^4, x_3^2x_4^2, x_4x_5^3\} \subset \mathbb{k}[x_1, \dots, x_5]$. Let $I = \langle B \rangle$, then \sqrt{I} has minimal generating set $G = \{x_1, x_3x_4, x_4x_5\}$, and $\mathcal{H}_B^{\text{red}} = \mathcal{H}_G$.

another, we delete it to create the *reduced* support hypergraph of B , $\mathcal{H}_B^{\text{red}}$. The reduced support hypergraph is illustrated in Figure 3.2.

REMARK 3.1.6. By construction, the reduced support hypergraph of any set of monomials is a *clutter*.

REMARK 3.1.7. Let $I \subset S$ be a monomial ideal, and let G be the unique minimal generating set of \sqrt{I} . Then for *any* set of monomials B such that $I = \langle B \rangle$, the reduced support hypergraph of B is equal to the support hypergraph of G .

The minimum vertex cover problem is a classic example of an NP-complete problem. It was one of the original 21 problems that Karp proved is NP-complete [Kar72], using a reduction from Boolean satisfiability, which Cook established the NP-completeness of only one year prior [Coo71]. From this section it follows that bounding the size of a minimum vertex cover reduces to bounding the codimension of a monomial ideal. An explicit reduction is given in [BS92], along with a justification for why the problem belongs to NP. Bayer and Stillman state the result this way:

PROPOSITION 3.1.8. The following problem is NP-complete: Given a monomial ideal $J \subset \mathbb{k}[x_1, \dots, x_n]$, and an integer K , is the codimension of $J < K$?

REMARK 3.1.9. In Proposition 3.1.8, we are considering the computational complexity with respect to increasing n .

REMARK 3.1.10. The minimum vertex cover problem is already NP-complete for graphs, not even hypergraphs. So codimension/dimension of monomial ideals is already NP-complete for squarefree monomial ideals generated in degree two, which have only $\mathcal{O}(n^2)$ minimal generators.

3.2. Krull dimension probabilities in the ER-type model

The relationship between dimension and hypergraph transversals leads to a complete characterization of the probability of producing a monomial ideal with a particular fixed dimension in the ER-type model (Theorem 3.2.1). Explicitly computing the values given by Theorem 3.2.1 requires an exhaustive combinatorial enumeration, which is demonstrated for several special cases in Theorem 3.2.5. Nevertheless, where the size of the problem makes enumeration prohibitive, Theorem 3.2.1 gives that $\mathbb{P}[\dim \mathfrak{J} = t]$ is always a polynomial in p of a specific degree depending on n and D , and thus can be approximated by numerically evaluating (or statistically estimating) $\mathbb{P}[\dim \mathfrak{J} = t]$ for a sufficient number of values of p , then interpolating. We will see in the next section that asymptotically, dimension can be classified via threshold functions, bypassing these computational considerations.

THEOREM 3.2.1. Let $\mathfrak{J} \sim \mathcal{I}(n, D, p)$. For any integer t , $0 \leq t \leq n$, the probability that \mathfrak{J} has dimension t is given by a polynomial in p of degree $\sum_{i=1}^{t+1} \binom{D}{i} \binom{n}{i}$. More precisely,

$$\mathbb{P}[\dim \mathfrak{J} = t] = \sum_{C \in \mathcal{C}_{n-t}} \prod_{\sigma \in E(C)} 1 - (1-p)^{\binom{D}{\#\sigma}} \prod_{\substack{\sigma' \subset \{x_1, \dots, x_n\} \\ \sigma \not\subseteq \sigma' \forall \sigma \in E(C)}} (1-p)^{\binom{D}{\#\sigma}},$$

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

PROOF. Let \mathfrak{B} be the random set of monomials generating \mathfrak{J} . By Proposition 3.1.2, $\mathbb{P}[\dim \mathfrak{J} = t] = \mathbb{P}[c(\mathcal{H}_{\mathfrak{B}}^{\text{red}}) = n - t]$. Let \mathcal{C}_{n-t} denote the set of clutters on $\{x_1, \dots, x_n\}$ with

transversal number $n - t$. Then $c(\mathcal{H}_{\mathfrak{B}}^{\text{red}}) = n - t$ if and only if $\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C$ for some $C \in \mathcal{C}_{n-t}$.

As these are disjoint events, we have:

$$(3.2.2) \quad \mathbb{P}[\dim \mathfrak{J} = t] = \sum_{C \in \mathcal{C}_{n-t}} \mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C].$$

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

$$(3.2.3) \quad \mathbb{P}[\sigma \in E(\mathcal{H}_{\mathfrak{B}})] = 1 - (1 - p)^{\binom{D}{\#\sigma}}, \text{ and } \mathbb{P}[\sigma \notin E(\mathcal{H}_{\mathfrak{B}})] = (1 - p)^{\binom{D}{\#\sigma}}.$$

Both expressions are polynomials in p of degree exactly $\binom{D}{\#\sigma}$. Now $\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C$ if only if:

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

Note that condition (2) is equivalent to: Any edge of $\mathcal{H}_{\mathfrak{B}}$ which is not an edge of C is not minimal. Since edges are included in $\mathcal{H}_{\mathfrak{B}}$ independently, this shows that

$$(3.2.4) \quad \mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C] = \prod_{\sigma \in E(C)} \mathbb{P}[\sigma \in E(\mathcal{H}_{\mathfrak{B}})] \prod_{\substack{\sigma' \subset \{x_1, \dots, x_n\} \\ \sigma \not\subseteq \sigma' \forall \sigma \in E(C)}} \mathbb{P}[\sigma' \notin E(\mathcal{H}_{\mathfrak{B}})].$$

We now prove two useful properties of the index sets in (3.2.4). First, *every* $\sigma \in E(C)$ satisfies $\#\sigma \leq t + 1$. To show this, 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 $T = \{x_1, \dots, x_n\} \setminus \sigma$ 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 $C \in \mathcal{C}_{n-t}$.

The second property is: *if* σ' *is a subset of* $\{x_1, \dots, x_n\}$ *satisfying* $\sigma \not\subseteq \sigma'$ *for all* $\sigma \in E(C)$, *then* $\#\sigma' \leq t$. To prove this, 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 $T = \{x_1, \dots, x_n\} \setminus \sigma'$. Hence T is a transversal of C with $\#T = n - \#\sigma' < n - t$, a contradiction.

Thus the first product in (3.2.4) is taken over subsets of cardinality at most $t + 1$, while the second is taken over subsets of cardinality at most t . Furthermore, no $\sigma \subseteq \{x_1, \dots, x_n\}$

can appear in both index sets. Since each subset that does appear contributes a polynomial in p of degree $\binom{D}{\#\sigma}$, the result is that $\mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C]$ is a polynomial in p of degree no greater than $\sum_{i=1}^{t+1} \binom{D}{i} \binom{n}{i}$. Since (3.2.2) is a sum of such polynomials, $\mathbb{P}[\dim \mathfrak{J} = t]$ has the same degree bound.

To prove this bound is achieved, we show there is a particular clutter for which Equation (3.2.4) has degree exactly $\sum_{i=1}^{t+1} \binom{D}{i} \binom{n}{i}$, and for every other clutter the expression is of strictly lower degree. 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

$$\begin{aligned} \mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = K_n^{t+1}] &= \prod_{\substack{\sigma \in \{x_1, \dots, x_n\} \\ \#\sigma = t+1}} \mathbb{P}[\sigma \in E(\mathcal{H}_{\mathfrak{B}})] \prod_{\substack{\sigma \in \{x_1, \dots, x_n\} \\ |\sigma| \leq t}} \mathbb{P}[\sigma \notin E(\mathcal{H}_{\mathfrak{B}})] \\ &= \left(1 - (1-p)^{\binom{D}{t+1}}\right)^{\binom{n}{t+1}} (1-p)^{Dn + \binom{D}{2} \binom{n}{2} + \dots + \binom{D}{t} \binom{n}{t}} \end{aligned}$$

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 $\#\sigma \leq t$. 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 Equation (3.2.4). In particular there are positively many subsets of cardinality $t+1$ which do not contribute factors to $\mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C]$. \square

Using the formula of Theorem 3.2.1 requires an enumeration of 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 3.2.5. Let $\mathfrak{J} \sim \mathcal{I}(n, D, p)$. Then,

- (a) $\mathbb{P}[\dim \mathfrak{J} = 0] = (1 - (1-p)^D)^n$.
- (b) $\mathbb{P}[\dim \mathfrak{J} = 1] = \sum_{j=0}^{n-1} \binom{n}{j} (1 - (1-p)^D)^j (1-p)^{D(n-j)} \left(1 - (1-p)^{\binom{D}{2}}\right)^{\binom{n-j}{2}}$.
- (c) $\mathbb{P}[\dim \mathfrak{J} = n-1] = -(1-p)^{\binom{n+D}{n}-1} + \sum_{j=1}^n (-1)^{j-1} \binom{n}{j} (1-p)^{\binom{n+D}{n}-1 - \binom{n+D-j}{n}}$.
- (d) $\mathbb{P}[\dim \mathfrak{J} = n] = (1-p)^{\binom{n+D}{n}-1}$.

PROOF. *Part (a):* For $\mathfrak{B} \sim \mathcal{B}(n, D, p)$, $\mathfrak{J} = \langle \mathfrak{B} \rangle$, \mathfrak{J} is zero dimensional if and only if $\mathcal{H}_{\mathfrak{B}}^{\text{red}} \in \mathcal{C}_n$. 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 3.2.1,

$$(3.2.6) \quad \mathbb{P}[\dim \mathfrak{J} = 0] = \prod_{i=1}^n 1 - (1-p)^{\binom{D}{\#(x_i)}} = (1 - (1-p)^D)^n.$$

Part (b): \mathfrak{J} is one-dimensional if and only if $\mathcal{H}_{\mathfrak{B}}^{\text{red}} \in \mathcal{C}_{n-1}$. 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 $0 \leq j \leq n-1$, and denote by V the set of these j vertices. Then V is a subset of any transversal of C . Let $W = \{x_1, \dots, x_n\} \setminus V$, then it can be shown that $c(C) = n-1$ if and only if $E(C) = \{\{x_i\} : x_i \in V\} \cup \{\{x_i, x_k\} \mid x_i, x_k \in W, x_i \neq x_k\}$. Hence $\mathbb{P}[\mathcal{H}_{\mathfrak{B}}^{\text{red}} = C]$ equals

$$(3.2.7) \quad \prod_{x_i \in V} \mathbb{P}[\{x_i\} \in E(\mathcal{H}_{\mathfrak{B}})] \prod_{x_i, x_k \in W} \mathbb{P}[\{x_i, x_k\} \in E(\mathcal{H}_{\mathfrak{B}})] \prod_{x_i \in W} \mathbb{P}[\{x_i\} \notin E(\mathcal{H}_{\mathfrak{B}})] \\ = (1 - (1-p)^D)^j \left(1 - (1-p)^{\binom{D}{2}}\right)^{\binom{n-j}{2}} (1-p)^{D(n-j)}.$$

The expression for $\mathbb{P}[\dim \mathfrak{J} = 1]$ is obtained by summing over all $\binom{n}{j}$ ways of selecting the j 1-edges, for each $0 \leq j \leq n-1$.

Part (c): For the case of $(n-1)$ -dimensionality, Theorem 3.2.1 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 $\mathbb{P}[F_i] = (1-p)^{\binom{n+D}{D}-1-\binom{n+D-1}{D-1}}$. 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.

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

REMARK 3.2.8. For $n \leq 4$, Theorem 3.2.5 specifies the complete probability distribution of Krull dimension induced by $\mathcal{I}(n, D, p)$, for any integer D .

3.3. Krull dimension thresholds in the ER-type model

For arbitrary t , explicitly computing $\mathbb{P}[\dim \mathfrak{J} = t]$ is difficult. In fact, it follows from Proposition 3.1.8 that any formula for \mathcal{C}_{n-t} , the set of clutters with cover number $n - t$, is exponential (unless P=NP). For this reason, the results of this section are quite powerful. Even though bounding the dimension of a monomial ideal is an NP-complete problem as n increases, in the case of $\mathfrak{J} \sim \mathcal{I}(n, D, p)$, the dimension of \mathfrak{J} behaves according to threshold laws that hold for arbitrarily large n .

LEMMA 3.3.1. Fix integers n and t , with $1 \leq t \leq n$. If $\mathfrak{B} \sim \mathcal{B}(n, D, p)$ and $p(D) = o(D^{-t})$, 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 $p(D) = o(D^{-t})$ and n and t are constants, thus establishing the lemma. \square

THEOREM 3.3.2. Fix integers n and t with $1 \leq t \leq n$, and let $p = p(D)$. Then for $\mathfrak{J} \sim \mathcal{I}(n, D, p)$, D^{-t} is a threshold function for the property that $\dim \mathfrak{J} \leq t - 1$. In other words,

$$(3.3.3) \quad \lim_{D \rightarrow \infty} \mathbb{P}[\dim \mathfrak{J} \leq t - 1] = \begin{cases} 0, & \text{if } p = o(D^{-t}), \\ 1, & \text{if } p = \omega(D^{-t}). \end{cases}$$

PROOF. Let $p = o(D^{-t})$. Consider $\mathcal{H}_{\mathfrak{B}}$, the support hypergraph of \mathfrak{B} . By Lemma 3.3.1, a.a.s. every monomial in \mathfrak{B} has support of size $t + 1$ or more, so every edge of $\mathcal{H}_{\mathfrak{B}}$ has size

$t + 1$ or more. This implies that every set of $(n - t)$ variables is a vertex cover of $\mathcal{H}_{\mathfrak{B}}$; hence $c(\mathcal{H}_{\mathfrak{B}}) \leq n - t$. By Proposition 3.1.2, $\dim \mathfrak{J} \geq t$ a.a.s.

Now, let $p = \omega(D^{-t})$. For each $\sigma \subset \{x_1, \dots, x_n\}$ with $\#\sigma = t$, define $X_\sigma = \#\{b \in \mathfrak{B} : \text{supp}(b) = \sigma\}$. For each D , the number of possible such b is $\binom{D}{t}$. This is because each b is of the form $\prod\{x_i \in \sigma\}$ multiplied by a monomial in t variables (the variables of σ) of degree no greater than $D - t$.

Thus, $\mathbb{E}[X_\sigma] = \binom{D}{t}p$ and this tends to infinity as D does, whenever $p = \omega(D^{-t})$. Further, $\text{Var}[X_\sigma] \leq \mathbb{E}[X_\sigma]$, as X_σ is a sum of independent indicator random variables. Hence we can apply the second moment method to conclude that as D tends to infinity,

$$\mathbb{P}[X_\sigma = 0] \leq \frac{\text{Var}[X_\sigma]}{\mathbb{E}[X_\sigma]^2} \leq \frac{1}{\binom{D}{t}p} \rightarrow 0.$$

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 $\{x_1, \dots, x_n\}$, it follows that the same holds for every choice of σ .

By the union bound, $\mathbb{P}[X_\sigma = 0 \text{ for at least one } \sigma \text{ with } \#\sigma = t]$ is bounded above by the sum of their probabilities and tends to zero, as there are finitely many such σ . Then, a.a.s. $\mathcal{H}_{\mathfrak{B}}$ contains *every* edge of size t . This means that for any set T of $n - t$ or fewer vertices, $\mathcal{H}_{\mathfrak{B}}$ contains an edge disjoint from T . Thus $c(\mathcal{H}_{\mathfrak{B}}) \geq n - t + 1$, and we use Proposition 3.1.2 again to conclude that $\dim \mathfrak{J} \leq t - 1$ a.a.s. □

COROLLARY 3.3.4. Fix integers n and t with $0 \leq t \leq n$. Then for $\mathfrak{J} \sim \mathcal{I}(n, D, p(D))$, if $p(D)$ satisfies both $p = \omega(D^{-(t+1)})$ and $p = o(D^{-t})$, as $D \rightarrow \infty$, then $\dim \mathfrak{J} = t$ asymptotically almost surely.

PROOF. Theorem 3.3.2 establishes a threshold result for each choice of constant t . But, if both $p = \omega(D^{-(t+1)})$ and $p = o(D^{-t})$ hold, then the theorem gives that events $\dim(\mathfrak{J}) > t$ and $\dim(\mathfrak{J}) < t$ each hold with probability tending to 0. Therefore the probability of their union also tends to 0, establishing the result. □

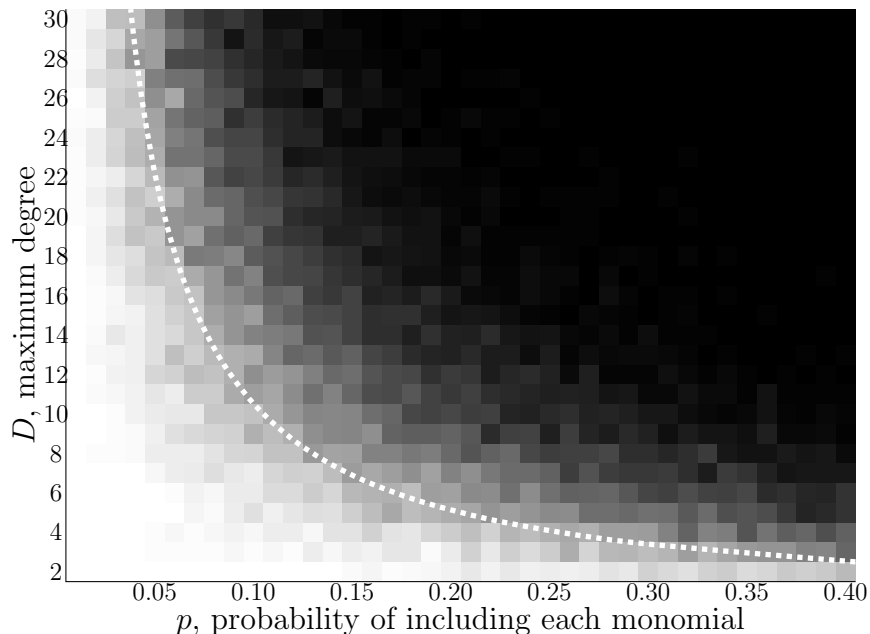


FIGURE 3.3. Visualizing the zero-dimensional threshold of ER-type model random monomial ideals in three variables. Each square in the image corresponds to a maximum degree D between 2 and 30, and a value of p between 0 and 0.4. For each (D, p) pair of parameters, 100 random monomial ideals $\mathcal{J} \sim \mathcal{I}(3, D, p)$ were sampled using the `RandomMonomialIdeals` package in `Macaulay2`, and their dimension computed with the methods of Section 3.4. The color of the square indicates the sample probability of dimension zero, from white (probability zero) to black (probability one). The phase transition follows the threshold function $p = 1/D$ (dotted white curve) proved in Theorem 3.3.2.

3.4. Fast dimension computations using integer linear programming

In addition to its theoretical use in the proofs of Sections 3.2 and 3.3, the vertex cover characterization of dimension is practical for faster computation of this invariant. Using an integer linear programming formulation of the minimum vertex cover problem to compute the dimension of a monomial ideal outperforms, often dramatically, the algorithms for computing dimension that are currently available in `Macaulay2`.

PROPOSITION 3.4.1. Let $I \subseteq \mathbb{k}[x_1, \dots, x_n]$ be a monomial ideal generated by $G = \{g_1, \dots, g_r\}$. Then $\text{codim}(I)$ is equal to the optimal objective value of the following integer

program:

$$\begin{aligned} & \text{minimize: } \sum_{i=1}^n x_i \\ & \text{subject to: } x_i \in \{0, 1\}, & 1 \leq i \leq n, \\ & \sum_{x_i \in \text{supp}(g_j)} x_i \geq 1, & 1 \leq j \leq r, \end{aligned}$$

which has n binary variables and r linear constraints.

PROOF. For any feasible solution (x_1, \dots, x_n) , set $T = \{x_i : x_i = 1\}$. Then T is a vertex cover of \mathcal{H}_G , because for each $g_j \in G$, the inequality $\sum_{x_i \in \text{supp}(g_j)} x_i \geq 1$ enforces that at least one $x_i \in \text{supp}(g_j)$ is in T . The vertex cover number of \mathcal{H}_G is thus equal to the minimum number of x_i 's in any feasible solution. \square

Figures 3.4 and 3.5 compares the native dimension computation of `Macaulay2` with the integer linear programming reformulation. A few clarifying comments are in order. In `Macaulay2`, the command `dim I` calls different functions in `Macaulay2` depending on whether `I` is an object of type `Ideal` or `MonomialIdeal`; in the latter case, faster combinatorial algorithms are used rather than the most general dimension algorithm. All ideals in these experiments were defined as type `MonomialIdeal`. Once each instance `I` was defined in `Macaulay2`, two computations were timed:

- (1) Wall time (using the `Macaulay2` function `elapsedTiming`) to return `dim I`,
- (2) Wall time (using the `Macaulay2` function `elapsedTiming`) to write the integer program corresponding to the dimension of `I` to an external ZIMPL file, call SCIP to solve the IP, write the solution to a file, and read the answer into `Macaulay2` as an integer of type `ZZ`.

The reason for measuring wall time rather than CPU time is that the latter is handled in `Macaulay2` by the command `timing`, which does not register the CPU usage of the external SCIP command. Though SCIP has its own timing functions, calling these would not account for the time used to reformulate the problems and handle file input/output, and would

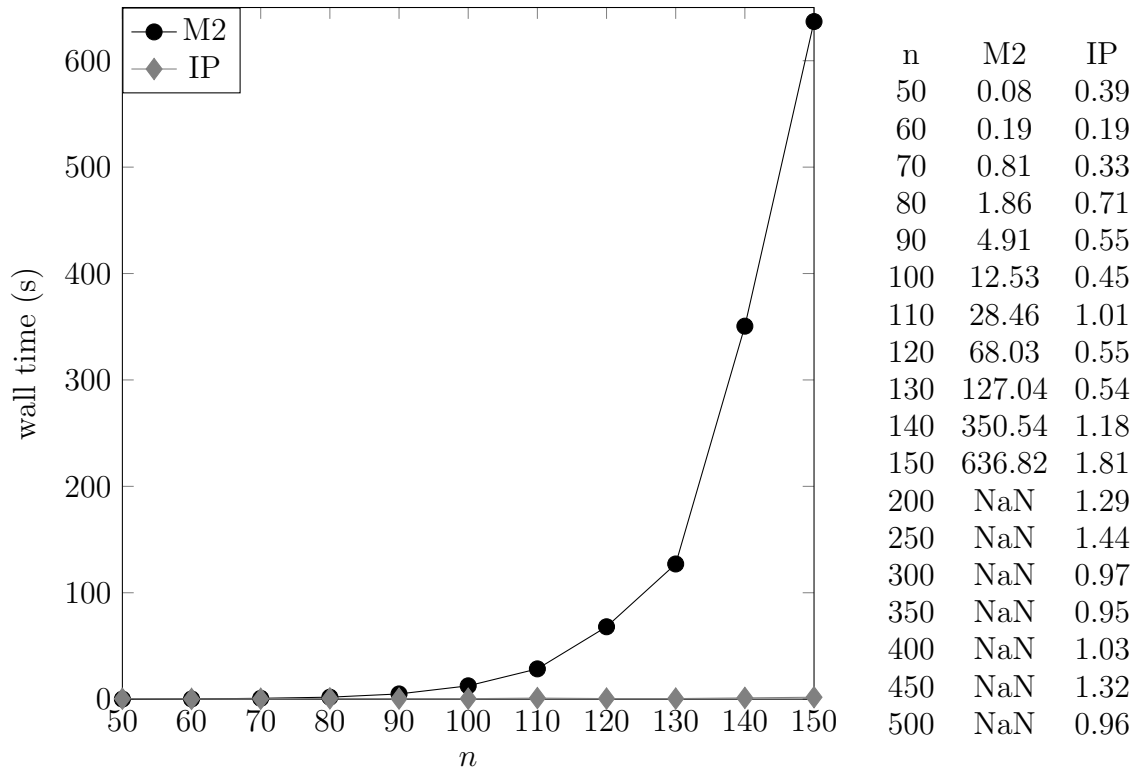


FIGURE 3.4. Time to compute the dimension of squarefree monomial ideals with 50 generators of degree 20 in n variables, as n increases, using `Macaulay2` alone (M2) versus writing the integer program of Proposition 3.4.1 to a ZIMPL file, calling `SCIP` to solve it, and reading the solution back into `Macaulay2` (IP). For each value of n , timing for each method was averaged over (the same) 10 random instances. These averages, displayed in the table on the right, were used to generate the plot on the left. To demonstrate the performance of the IP method, the table includes its average timing for even larger instances. See Appendix A.2 for details about source code and reproducing this test.

make the speed of the `SCIP` method seem even more dramatic. Since the purpose of these experiments is to demonstrate a practical improvement for the computer algebra user, it seems most fair to measure the user’s experience running each method from start to finish from within a `Macaulay2` session.

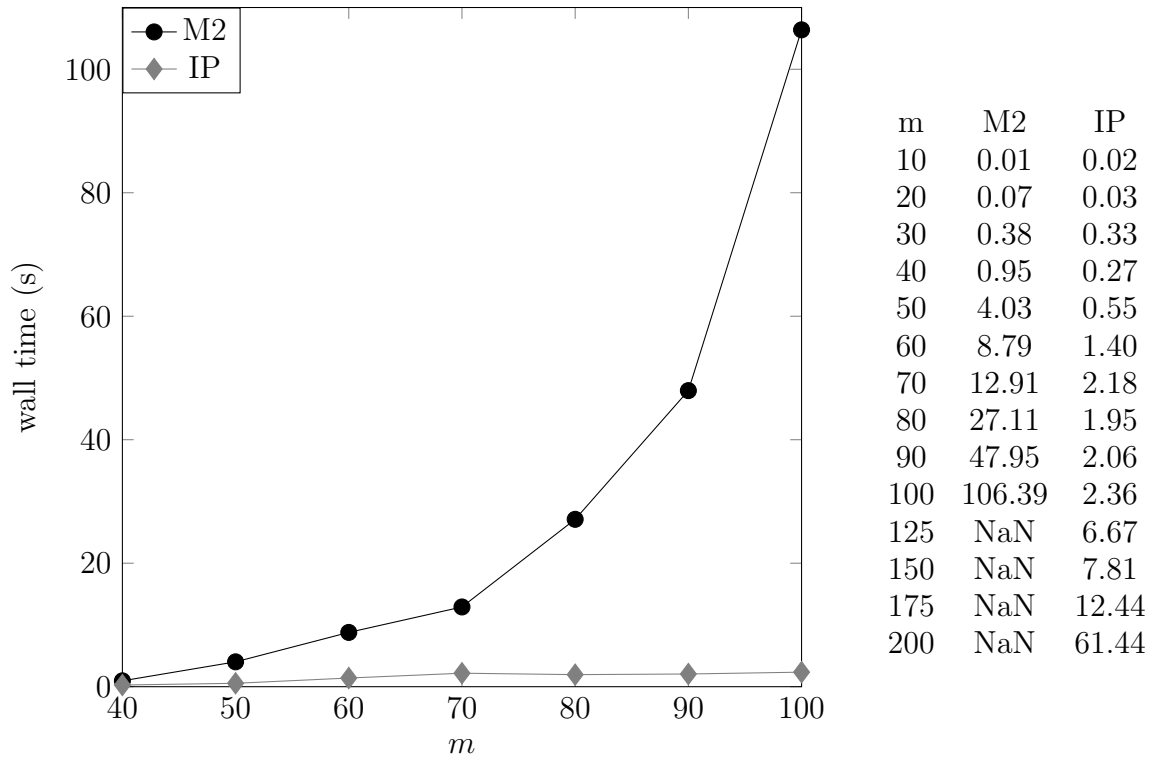


FIGURE 3.5. Time to compute the dimension of squarefree monomial ideals with m generators of degree 30 in 100 variables, as m increases, using `Macaulay2` alone (M2) versus writing the integer program of Proposition 3.4.1 to a ZIMPL file, calling SCIP to solve it, and reading the solution back into `Macaulay2` (IP). For each value of m , timing for each method was averaged over (the same) 10 random instances. These averages, displayed in the table on the right, were used to generate the plot on the left. To demonstrate the performance of the IP method, the table includes its average timing for even larger instances. See Appendix A.2 for details about source code and reproducing this test.

Projective dimension

4.1. Monomial ideals with large projective dimension

Let $I = \langle G \rangle \subseteq S$ be a monomial ideal with minimal generating set G . In 2017 Alesandroni proved a combinatorial criterion for G that is equivalent to the statement $\text{pdim}(S/I) = n$. See [Ale17a, Ale17b] for details and proofs. Only the statement is given here.

Let L be a set of monomials. An element $m = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in L$ is a *dominant monomial* (in L) if there is a variable x_i such that the x_i exponent of m , α_i , is strictly larger than the x_i exponent of any other monomial in L . If every $m \in L$ is a dominant monomial, then L is a *dominant set*. For example, $L_1 = \{x_1^3 x_2 x_3^2, x_2^2 x_3, x_1 x_3^3\}$ is a dominant set in $\mathbb{k}[x_1, x_2, x_3]$, but $L_2 = \{x_1^3 x_2 x_3^2, x_2^2 x_3, x_1^3 x_3^3\}$ is not.

THEOREM 4.1.1. [Ale17b, Theorem 5.2, Corollary 5.3] Let $I \subseteq S$ be a monomial ideal minimally generated by G . Then $\text{pdim}(S/I) = n$ if and only if there is a subset L of G with the following properties:

- (1) L is dominant.
- (2) $\#L = n$.
- (3) No element of G strongly divides $\text{lcm}(L)$. (Definition 1.3.38.)

More precisely, if $L \subseteq G$ satisfies conditions (1), (2) and (3), then the minimal free resolution of S/I has a basis element with multidegree $\text{lcm}(L)$ in homological degree n . On the other hand, if there is a basis element with multidegree x^α and homological degree n , then G must contain some L' satisfying 1, 2, 3 and the condition $\text{lcm}(L') = x^\alpha$.

4.2. Most monomial ideals have large projective dimension

DEFINITION 4.2.1. When L is any set of minimal generators of M that satisfies the three conditions of Theorem 4.1.1, then L witnesses $\text{pdim}(S/M) = n$, and we will call L a *witness set*. The monomial $x^\alpha \in S$ is a *witness lcm* if L is a witness set and $x^\alpha = \text{lcm}(L)$.

The distinction between witness sets and witness lcm's is important, as several witness sets can have a common lcm. We found it useful to think of the event “ x^α is a witness lcm” in geometric terms, as illustrated in Figure 4.1 for the case of $n = 3$.

The monomials of total degree D are represented as lattice points in a regular $(n - 1)$ -simplex with side lengths D . Given $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, the n inequalities $x_1 \leq \alpha_1, \dots, x_n \leq \alpha_n$ determine a new regular simplex Δ_α (shaded). If L is a dominant set that satisfies $\#L = n$ and $\text{lcm}(L) = x^\alpha$, then L must contain exactly one lattice point from the interior of each facet of Δ_α . (Monomials on the boundary of a facet are dominant in more than one variable.) Meanwhile, the strong divisors of x^α are the lattice points in the interior of Δ_α . The event “ x^α is a witness lcm” occurs when at least one generator is chosen in the interior of each facet of Δ_α , and no generators are chosen in the interior of Δ_α .

Understanding the probability of witness sets and witness lcm's is crucial to the proof of the following theorem, which establishes a threshold for the projective dimension of random monomial ideals.

THEOREM 4.2.2. Let $S = \mathbb{k}[x_1, \dots, x_n]$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, and $p = p(D)$. As $D \rightarrow \infty$, $p = D^{-n+1}$ is a threshold for the projective dimension of S/\mathfrak{M} . If $p = o(D^{-n+1})$ then $\text{pdim}(S/\mathfrak{M}) = 0$ a.a.s., and if $p = \omega(D^{-n+1})$ then $\text{pdim}(S/\mathfrak{M}) = n$ a.a.s.

In other words, the case of equality in Hilbert's syzygy theorem is the most typical situation for non-trivial ideals.

The rest of this section comprises the proof of Theorem 4.2.2 and two of its consequences. First we show that for p below the threshold, usually $\text{pdim}(S/\mathfrak{M}) = 0$. Let

$$m_n(D) = \binom{D + n - 1}{n - 1}$$

denote the number of monomials in n variables of degree D . This is a polynomial in D of degree $n - 1$ and can be bounded, for D sufficiently large, by

$$(4.2.3) \quad \frac{1}{(n-1)!} D^{n-1} \leq m_n(D) \leq \frac{2}{(n-1)!} D^{n-1}.$$

PROPOSITION 4.2.4. If $p = o(D^{-n+1})$ then $\text{pdim}(S/\mathfrak{M}) = 0$ a.a.s. as $D \rightarrow \infty$.

PROOF. For each $x^\alpha \in S$, let X_α be the random variable indicating that $x^\alpha \in \mathfrak{G}$ ($X_\alpha = 1$) or $x^\alpha \notin \mathfrak{G}$ ($X_\alpha = 0$). We define $X = \sum_{\alpha \in S} X_\alpha$, so that X records the cardinality of the random minimal generating set \mathfrak{G} . By Markov's inequality,

$$\mathbb{P}[X > 0] = \mathbb{P}[X \geq 1] \leq \mathbb{E}[X] = \sum_{\substack{\alpha \in S \\ |\alpha|=D}} \mathbb{E}[X_\alpha] = m_n(D)p.$$

Letting $D \rightarrow \infty$, we have

$$\lim_{D \rightarrow \infty} \mathbb{P}[X > 0] = \lim_{D \rightarrow \infty} m_n(D)p = 0,$$

since $p = o(D^{-n+1})$. So $\#\mathfrak{G} = 0$, equivalently $\mathfrak{M} = \langle 0 \rangle$, with probability converging to 1 as $D \rightarrow \infty$. Therefore below the threshold D^{-n+1} , almost all random monomial ideals have $\text{pdim}(S/\mathfrak{M}) = 0$. \square

For the case $p = \omega(D^{-n+1})$, we use the second moment method. Recall that $x^\alpha \in S$ is a witness lcm to $\text{pdim}(S/\mathfrak{M}) = n$ if and only if there is a dominant set $L \subseteq \mathfrak{G}$ with $\#L = n$, $\text{lcm}(L) = x^\alpha$, and no generator in \mathfrak{G} strongly divides x^α . For each α , we define an indicator random variable w_α that equals 1 if x^α is a witness lcm and 0 otherwise. Next we define W_a , for integers $a > 1$, and W by

$$W_a = \sum_{\substack{|\alpha|=D+a \\ \alpha_i \geq a \forall i}} w_\alpha, \quad W = \sum_{a=n-1}^A W_a,$$

where $A = \lfloor (p/2)^{-\frac{1}{n-1}} \rfloor - n$. The random variable W_a counts *most* witness lcm's of degree $D+a$. The reason for the restriction $\alpha_i \geq a$ is easily explained geometrically. In general, the probability that x^α is a witness lcm depends only on the side length of the simplex Δ_α (see

Figure 4.1). If, however, the facet defining inequalities of Δ_α intersect outside of the simplex of monomials with degree D , the situation is more complicated and has many different cases. The definition of W_a bypasses these cases, and this does not change the asymptotic analysis.

In Lemma 4.2.7, we compute the order of $\mathbb{P}[w_\alpha]$ and use this to prove that $\mathbb{E}[W] \rightarrow \infty$ as $D \rightarrow \infty$ in Lemma 4.2.9. Then in Lemma 4.2.10, we prove $\text{Var}[W] = o(\mathbb{E}[W]^2)$ and thus that the right-hand side of Equation (1.1.6) goes to 0 as $D \rightarrow \infty$. In other words, $\mathbb{P}[W > 0] \rightarrow 1$, meaning that $\mathfrak{M} \sim \mathcal{M}(n, D, p)$ will have at least one witness to $\text{pdim}(S/M) = n$ with probability converging to 1 as $D \rightarrow \infty$. This proves the second side of the threshold and establishes the theorem.

We first give the value of $\mathbb{P}[w_\alpha]$ for an exponent vector α with $|\alpha| = D + a$ and $\alpha_i \geq a$ for all i . The monomials of degree D that divide x^α form the simplex Δ_α , and those that strongly divide x^α form the interior of Δ_α . Thus there are $m_n(a)$ divisors and $m_n(a - n)$ strong divisors of x^α in degree D . Recall that for x^α to be a witness lcm, for each variable x_i there must be at least one monomial x^β in \mathfrak{G} with x^β in the relative interior of the facet of Δ_α parallel to the subspace $\{x_i = 0\}$. In other words, there must be an $x^\beta \in \mathfrak{G}$ satisfying $\beta_i = \alpha_i$ and $\beta_j < \alpha_j$ for all $j \neq i$. Therefore $x^{\alpha - \beta}$ is a monomial of degree a without x_i and with positive exponents for each of the other variables. See Figure 4.1. The number of such monomials is $m_{n-1}(a - n + 1)$. The relative interiors of the facets of Δ_α are disjoint, so the events that a monomial appears in each one are independent. Additionally, \mathfrak{G} must not contain any monomials that strongly divide x^α , and the probability of this is $q^{m_n(a-n)}$ where $q = 1 - p$. Therefore, for α with $|\alpha| = D + a$ and $\alpha_i \geq a$ for all i ,

$$(4.2.5) \quad \mathbb{P}[w_\alpha] = (1 - q^{m_{n-1}(a-n+1)})^n q^{m_n(a-n)}.$$

By linearity of expectation, a consequence of this formula is

$$(4.2.6) \quad \mathbb{E}[W_a] = m_n(D + a - na) (1 - q^{m_{n-1}(a-n+1)})^n q^{m_n(a-n)},$$

because the number of exponent vectors α with $|\alpha| = D + a$ and $\alpha_i \geq a$ for all i is $m_n(D + a - na)$.

LEMMA 4.2.7. Let α be an exponent vector with $a = |\alpha| - D \leq p^{-\frac{1}{n-1}}$ and $\alpha_i \geq a$ for all i . Then,

$$(4.2.8) \quad \frac{1}{2}p^n (m_{n-1}(a-n+1))^n \leq \mathbb{P}[w_\alpha] \leq p^n (m_{n-1}(a-n+1))^n.$$

PROOF. The union-bound implies that

$$1 - q^{m_{n-1}(a-n+1)} \leq pm_{n-1}(a-n+1).$$

The upper bound on $\mathbb{P}[w_\alpha]$ follows from applying this inequality to the expression in (4.2.5). For the lower-bound, note that $\mathbb{P}[w_\alpha]$ is bounded below by the probability that exactly one monomial is chosen to be in \mathfrak{G} from the relative interior of each facet of Δ_α , and no other monomials are chosen in Δ_α . The probability of this latter event is given by

$$p^n (m_{n-1}(a-n+1))^n q^{m_n(a)-n}$$

since there are $m_{n-1}(a-n+1)$ choices for the monomial picked in each facet. Now we use the fact that $m_n(a) \leq m_n(A) \leq p/2$ (and this is the reason for the choice of $A = \lfloor (p/2)^{-\frac{1}{n-1}} \rfloor - n$) to conclude

$$q^{m_n(a)-n} \geq 1 - (m_n(a) - n)p \geq 1 - \frac{(a+n)^{n-1}}{(n-1)!} p \geq \frac{1}{2}.$$

□

LEMMA 4.2.9. If $p = \omega(D^{-n+1})$ then $\lim_{D \rightarrow \infty} \mathbb{E}[W] = \infty$.

PROOF. If $\lim_{D \rightarrow \infty} p > 0$, then $\mathbb{E}[W_{n-1}] \geq m_n(D-1)p^n$ which goes to infinity in D . Instead assume that $p = \omega(D^{-n+1}), p = o(1)$. From Lemma 4.2.7, we have

$$\mathbb{P}[w_\alpha] \geq \frac{1}{2}p^n (m_{n-1}(a-n+1))^n \geq \frac{1}{2}p^n \left(\frac{(a-n)^{n-2}}{(n-2)!} \right)^n.$$

For $n-1 \leq a \leq A$ with $A = \lfloor (p/2)^{-\frac{1}{n-1}} \rfloor - n$, one gets $a = o(D)$, and hence for D sufficiently large, $na < D/2$, which means $D + a - na > D/2$. Therefore

$$m_n(D + a - na) \geq \frac{D^{n-1}}{2^{n-1}(n-1)!}.$$

Since $m_n(D + a - na)$ is the number of exponent vectors α with $|\alpha| = D + a$ and $\alpha_i \geq a$ for all i ,

$$\mathbb{E}[W_a] = \sum_{\substack{|\alpha|=D+a \\ \alpha_i \geq a \forall i}} \mathbb{P}[w_\alpha] \geq c_n D^{n-1} p^n (a - n)^{n(n-2)}$$

where $c_n > 0$ is a constant that depends only on n . Summing up over a gives the bound

$$\mathbb{E}[W] = \sum_{a=n-1}^A \mathbb{E}[W_a] \geq c_n D^{n-1} p^n \sum_{a=n-1}^A (a - 2n)^{n^2-2n}.$$

The function $f(A) = \sum_{a=n-1}^A (a - 2n)^{n^2-2n}$ is a polynomial in A with leading term $t = A^{n^2-2n+1}/(n^2 - 2n + 1)$. Since A is proportional to $p^{-\frac{1}{n-1}}$, for p sufficiently small $f(A) \geq t/2$ and so

$$\mathbb{E}[W] \geq c_n D^{n-1} p^n \frac{p^{-\frac{n^2-2n+1}{n-1}}}{2(n^2 - 2n + 1)} = c'_n D^{n-1} p$$

and $D^{n-1} p$ goes to infinity as $D \rightarrow \infty$. □

LEMMA 4.2.10. If $p = \omega(D^{-n+1})$ then

$$\lim_{D \rightarrow \infty} \frac{\text{Var}[W]}{\mathbb{E}[W]^2} = 0.$$

PROOF. Since W is a sum of indicator variables w_α , we can bound $\text{Var}[W]$ by

$$\text{Var}[W] \leq \mathbb{E}[W] + \sum_{(\alpha, \beta)} \text{Cov}[w_\alpha, w_\beta].$$

The covariance is easy to analyze in the following two cases. If the degree of $\gcd(x^\alpha, x^\beta)$ is at most D , then w_α and w_β depend on two sets of monomials being in \mathfrak{G} which share at most one monomial. In this case w_α and w_β are independent so $\text{Cov}[w_\alpha, w_\beta] = 0$. The second case is that $x^\alpha | x^\beta$ and $\alpha \neq \beta$. If $w_\alpha = 1$, then \mathfrak{G} contains a monomial that strictly divides x^β . In this case w_α and w_β are mutually exclusive, so $\text{Cov}[w_\alpha, w_\beta] < 0$. The cases with $\text{Cov}[w_\alpha, w_\beta] \leq 0$ are illustrated geometrically, for $n = 3$, in Figure 4.2.

Thus we focus on the remaining case, when $\deg \gcd(x^\alpha, x^\beta) > D$ and neither of x^α and x^β divides the other. In other words Δ_α and Δ_β have intersection of size > 1 and neither is contained in the other.

Let $a = \deg(x^\alpha) - D$, $b = \deg(x^\beta) - D$, which are the edge lengths of the simplices Δ_α and Δ_β respectively. Let $c = \deg(\gcd(x^\alpha, x^\beta)) - D$, which is the edge length of the simplex $\Delta_\alpha \cap \Delta_\beta$. Note that $0 < c < a$ due to assumptions made on α and β . The number of common divisors of x^α and x^β of degree D is given by $m_n(c)$. Let $\delta_{\alpha,i}$ and $\delta_{\beta,i}$ denote the relative interiors of the i th facets of Δ_α and Δ_β , respectively. The type of intersection of Δ_α and Δ_β is characterized by signs of the entries of $\alpha - \beta$, which is described by a 3-coloring C of $[n]$ with color classes $C_\alpha, C_\beta, C_\gamma$ for positive, negative, and zero, respectively.

Since w_α is a binary random variable, $\text{Cov}[w_\alpha, w_\beta] = \mathbb{P}[w_\alpha w_\beta] - \mathbb{P}[w_\alpha] \mathbb{P}[w_\beta]$, and hence it is bounded by $\mathbb{P}[w_\alpha w_\beta]$. Therefore we will focus on bounding this quantity. Let $w_{\alpha,i}$ be the indicator variable for the event that \mathfrak{G} contains a monomial $x_1^{u_1} \cdots x_n^{u_n}$ with $u_i = \alpha_i$ and $u_j < \alpha_j$ for each $j \neq i$. Then

$$\mathbb{P}[w_\alpha w_\beta] \leq \mathbb{P}\left[\prod_{i=1}^n w_{\alpha,i} w_{\beta,i}\right].$$

For $i \in C_\alpha$, the facet $\delta_{\alpha,i}$ does not intersect Δ_β . See Figure 4.3a. For each $i \in C_\alpha$, we have

$$\mathbb{P}[w_{\alpha,i}] = 1 - q^{m_{n-1}(a-n+1)} \leq m_{n-1}(a-n+1)p \leq a^{n-2}p \leq A^{n-2}p \leq p^{1/(n-1)}.$$

Similarly for $i \in C_\beta$, $\mathbb{P}[w_{\beta,i}] \leq p^{1/(n-1)}$.

For each pair $i \in C_\beta$ and $j \in C_\alpha$, facets $\delta_{\alpha,i}$ and $\delta_{\beta,j}$ intersect transversely. Let H be the bipartite graph on $C_\beta \cup C_\alpha$ formed by having $\{i, j\}$ as an edge if and only if there is a monomial in \mathfrak{G} in $\delta_{\alpha,i} \cap \delta_{\beta,j}$. Let $e_{i,j}$ be the event that $\{i, j\}$ is an edge of H . Let V denote the subset of $C_\beta \cup C_\alpha$ not covered by H . If $w_\alpha w_\beta$ is true, then for each $i \in V \cap C_\beta$, there must be a monomial in \mathfrak{G} in $\delta_{\alpha,i} \setminus \bigcup_{j \in C_\alpha} \delta_{\beta,j}$, and let v_i be this event. Similarly for each $j \in V \cap C_\alpha$, there must be a monomial in \mathfrak{G} in $\delta_{\beta,j} \setminus \bigcup_{i \in C_\beta} \delta_{\alpha,i}$, and let v_j be this event. See Figure 4.3 for the geometric intuition behind these definitions.

Note that all events $e_{i,j}$ and v_i are independent since they involve disjoint sets of variables.

Therefore

$$\mathbb{P}\left[\prod_{i \in C_\alpha} w_{\alpha,i} \prod_{i \in C_\beta} w_{\beta,i}\right] \leq \sum_H \prod_{\{i,j\} \in E(H)} \mathbb{P}[e_{i,j}] \prod_{i \in V} \mathbb{P}[v_i].$$

For any $(i, j) \in C_\beta \times C_\alpha$,

$$\#(\delta_{\alpha,i} \cap \delta_{\beta,j}) \leq m_{n-2}(c) \leq c^{n-3} \leq p^{-\frac{n-3}{n-1}}.$$

Therefore

$$\mathbb{P}[e_{i,j}] = 1 - q^{\#(\delta_{\alpha,i} \cap \delta_{\beta,j})} \leq p^{\#(\delta_{\alpha,i} \cap \delta_{\beta,j})} \leq p^{\frac{2}{n-1}}.$$

We also know that for $i \in C_\beta$, $\mathbb{P}[v_i] \leq \mathbb{P}[w_{\alpha,i}] \leq p^{1/(n-1)}$, and similarly for $i \in C_\alpha$. So then

$$\sum_H \prod_{\{i,j\} \in E(H)} \mathbb{P}[e_{i,j}] \prod_{i \in V} \mathbb{P}[v_i] \leq \sum_H p^{\frac{2\#E(H) + \#V}{n-1}}.$$

The number of graphs H is $2^{\#C_\beta \cdot \#C_\alpha} \leq 2^{n^2}$ and for any graph H , $2\#E(H) + \#V \geq \#C_\beta + \#C_\alpha$ since every element of $C_\beta \cup C_\alpha$ must be covered by H or in V . Then

$$\mathbb{P} \left[\prod_{i \in C_\alpha} w_{\alpha,i} \prod_{i \in C_\beta} w_{\beta,i} \right] \leq 2^{n^2} p^{\frac{\#C_\beta + \#C_\alpha}{n-1}}.$$

Finally for each $i \in C_\gamma$, facets $\delta_{\alpha,i}$ and $\delta_{\beta,i}$ have full dimensional intersection. Again \mathfrak{G} may contain distinct monomials in $\delta_{\alpha,i}$ and $\delta_{\beta,i}$, or just one in their intersection. However, $\delta_{\alpha,i}$ does not intersect any other facets of Δ_β so there are only two cases.

$$\mathbb{P}[w_{\alpha,i} w_{\beta,i}] \leq (1 - q^{m_{n-1}(a-n+1)})^2 + 1 - q^{m_{n-1}(c-n+1)} \leq p^{2/(n-1)} + p^{1/(n-1)} \leq 2p^{1/(n-1)}.$$

Combining these results, we have

$$\mathbb{P}[w_\alpha w_\beta] \leq 2^{n^2} p^{\frac{\#C_\beta + \#C_\alpha}{n-1}} \prod_{i \in C_\alpha} p^{\frac{1}{n-1}} \prod_{j \in C_\beta} p^{\frac{1}{n-1}} \prod_{i \in C_\gamma} 2p^{\frac{1}{n-1}} \leq 2^{n^2 + \#C_\gamma} p^{\frac{2n - \#C_\gamma}{n-1}}.$$

To sum up over all pairs α, β with potentially positive variance, we must count the number of pairs of each coloring C . To do so, first fix C and α and count the number of β such that the intersection of Δ_α and Δ_β have type C . Note that the signs of the entries of $\alpha - \beta$ are prescribed, and that the entries of $\alpha - \beta$ are bounded by $p^{-\frac{1}{n-1}}$ because the degrees of x^α and x^β are each within $p^{-\frac{1}{n-1}}$ of the degree of their gcd. A rough bound then on the number of values of β is $(p^{-\frac{1}{n-1}})^{n - \#C_\gamma}$. The number of values of α for each choice

of a is $m_n(D + a - na) \leq D^{n-1}$, so summing over all possible values of a , the number of α values is bounded by $p^{-\frac{1}{n-1}} D^{n-1}$. Therefore

$$\begin{aligned} \sum_{(\alpha, \beta) \text{ of type } C} \text{Cov}[w_\alpha, w_\beta] &\leq \#\{(\alpha, \beta) \text{ of type } C\} 2^{n^2 + \#C_\gamma} p^{\frac{2n - \#C_\gamma}{n-1}} \\ &\leq p^{-\frac{1}{n-1}} D^{n-1} (p^{-\frac{1}{n-1}})^{n - \#C_\gamma} 2^{n^2 + \#C_\gamma} p^{\frac{2n - \#C_\gamma}{n-1}} \leq 2^{n^2 + n} D^{n-1} p \leq \frac{2^{n^2 + n}}{c'_n} \mathbb{E}[W]. \end{aligned}$$

Then summing over all colorings C , of which there are less than 3^n , shows that $\text{Var}[W] \leq c''_n \mathbb{E}[W]$ for $c''_n > 0$ depending only on n . Therefore

$$\lim_{D \rightarrow \infty} \frac{\text{Var}[W]}{\mathbb{E}[W]^2} \leq \lim_{D \rightarrow \infty} \frac{c''_n}{c'_n} = 0.$$

□

PROOF OF THEOREM 4.2.2. If $p = o(D^{-n+1})$, Proposition 4.2.4 implies $\text{pdim}(S/\mathfrak{M}) = 0$. If $p = \omega(D^{-n+1})$, Lemma 4.2.9 proves that $\mathbb{E}[W] \rightarrow \infty$ as $D \rightarrow \infty$. Since Lemma 4.2.10 shows that $\mathbb{P}[W > 0] \rightarrow 1$, we conclude that there is a witness set asymptotically almost surely. This is equivalent to $\text{pdim}(S/\mathfrak{M}) = n$. □

4.3. Cohen-Macaulayness

COROLLARY 4.3.1. Let $S = k[x_1, \dots, x_n]$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, and $p = p(D)$. If $p = \omega(D^{-n+1})$ and $p = o(1)$, then asymptotically almost surely S/\mathfrak{M} is not Cohen-Macaulay.

PROOF. For a monomial ideal $M \subseteq S$, the Krull dimension of S/M is zero if and only if for each $i = 1, \dots, n$, M contains a minimal generator of the form x_i^j for $j = 1, \dots, n$. For $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, this can only occur if every monomial in the set $\{x_1^D, x_2^D, \dots, x_n^D\}$ is chosen as a minimal generator, an event that has probability p^n . Thus for fixed n and $p = o(1)$, $\mathbb{P}[\dim \mathfrak{M} = 0] = p^n \rightarrow 0$ as $D \rightarrow \infty$. If also $p = \omega(D^{-n+1})$, then by Theorem 4.2.2, $\mathbb{P}[\text{pdim}(S/\mathfrak{M}) = n] \rightarrow 1$. Together, these imply that $\mathbb{P}[S/\mathfrak{M} \text{ is Cohen-Macaulay}] \rightarrow 0$ as $D \rightarrow \infty$. □

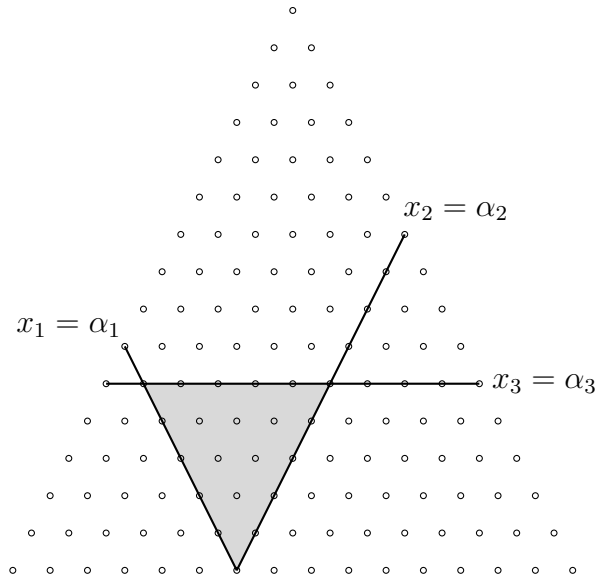
REMARK 4.3.2. A different investigation of “average” homological behavior was that of Ein, Erman and Lazarsfeld [EEL15], who studied the ranks of syzygy modules of smooth projective varieties. Their conjecture—that these ranks are asymptotically normal as the positivity of the embedding line bundle grows—is supported by their proof of asymptotic normality for the case of random Betti tables. Their random model is based on the elegant Boij-Söderberg theory established by Eisenbud and Schreyer [ES09]; for a fixed number of rows, they sample by choosing Boij-Söderberg coefficients independently and uniformly from $[0, 1]$, then show that with high probability the Betti table entries become normally distributed as the number of rows goes to infinity. Further support to this conjecture is the paper of Erman and Yang [EY17], which uses the probabilistic method to exhibit concrete examples of families of embeddings that demonstrate this asymptotic normality.

In [EY17], Erman and Yang consider random squarefree monomial ideals in n variables, defined as the Stanley-Reisner ideals of random flag complexes on n vertices, and study their asymptotic behavior as $n \rightarrow \infty$. Though the model is different, they find a similar result: for many choices of their model parameter, Cohen-Macaulayness essentially never occurs.

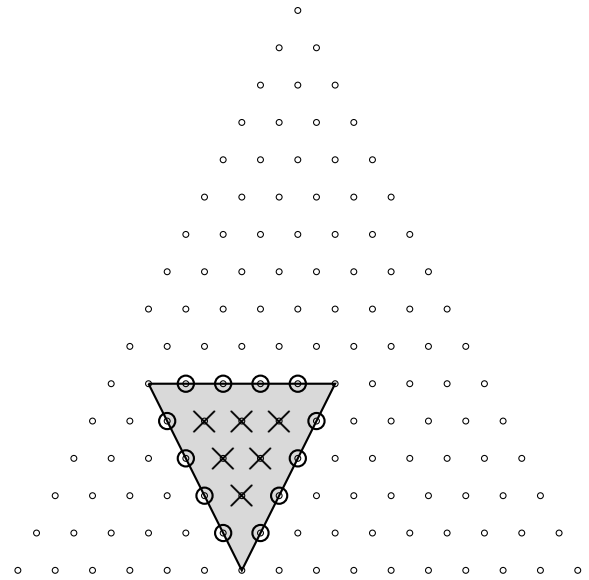
Another corollary to Theorem 4.2.2 is about Betti numbers. By the results of Brun and Römer [BR04], which extended those of Charalambous [Cha91] (see also [BS18]), a monomial ideal with projective dimension d will satisfy $\beta_i(S/M) \geq \binom{d}{i}$ for all $1 \leq i \leq d$. In the special case $d = n$, Alesandroni gives a combinatorial proof of the implied inequality $\sum_{i=0}^n \beta_i(S/M) \geq 2^n$ [Ale17b]. These inequalities are of interest because they relate to the long-standing *Buchsbaum-Eisenbud-Horrocks conjecture* [BE77, Har79], that $\beta_i(N) \geq \binom{c}{i}$ for N an S -module of codimension c . In 2017, Walker [Wal17] settled the BEH conjecture outside of the characteristic 2 case. Here we show that a probabilistic result, which holds regardless of characteristic, follows easily from Theorem 4.2.2.

COROLLARY 4.3.3. Let $M \sim \mathcal{M}(n, D, p)$ and $p = p(D)$. If $p = \omega(D^{-n+1})$, then asymptotically almost surely $\beta_i(S/M) \geq \binom{n}{i}$ for all $1 \leq i \leq n$.

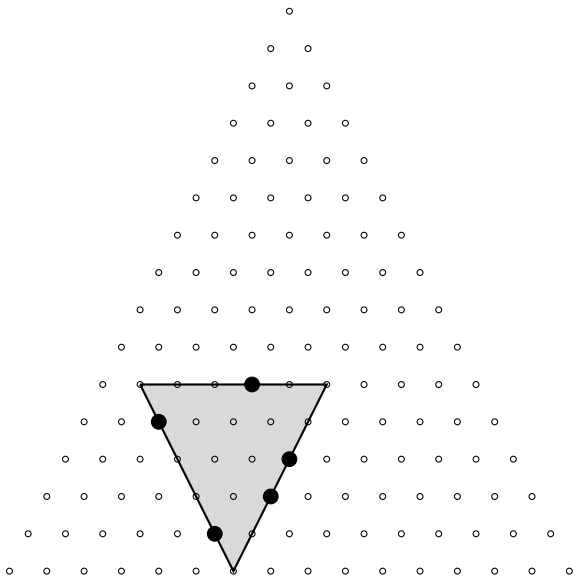
PROOF. Follows immediately from [BR04, Theorem 1.1] and Theorem 4.2.2. □



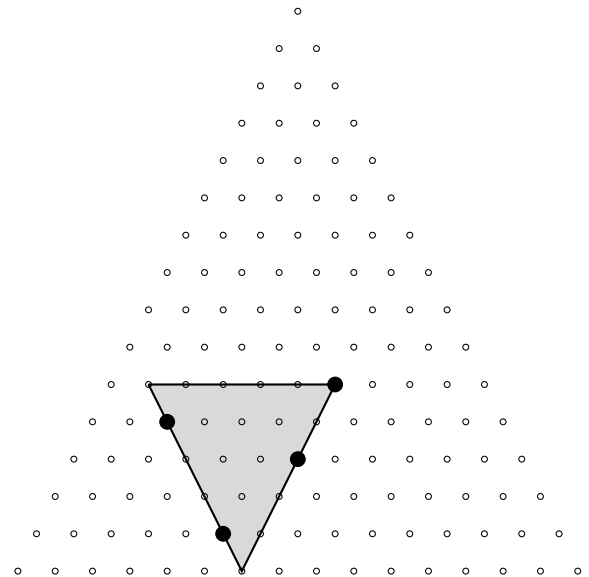
(A) The simplex Δ_α , associated with a possible witness lcm $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, has facet-defining inequalities $x_i \leq \alpha_i$ for $i = 1, \dots, n$.



(B) For x^α to be a witness lcm, at least one monomial on the interior of each facet (circled) must be chosen, and none of the interior monomials (crossed out) can be chosen.

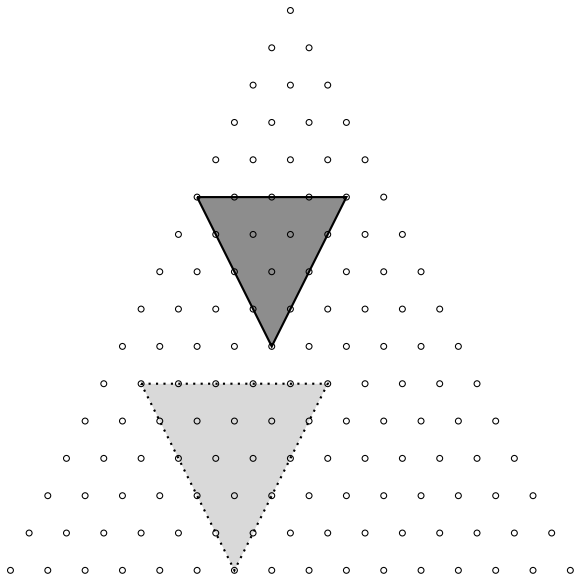


(C) A situation where x^α is a witness lcm with four different witness sets.

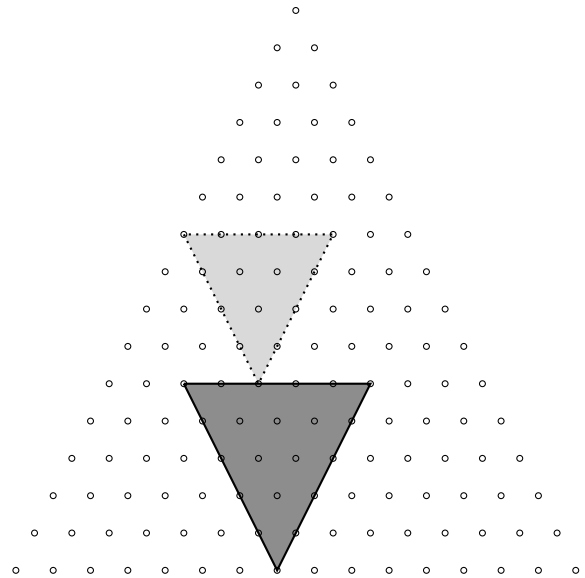


(D) Not a witness set, since there is not a monomial on the *interior* of each facet of Δ_α .

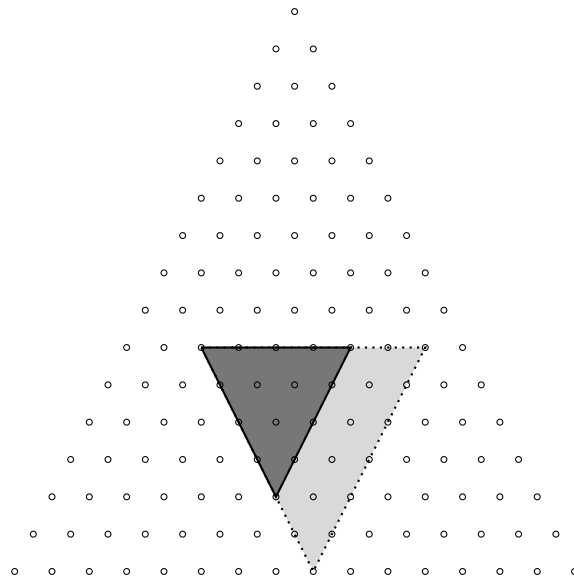
FIGURE 4.1. Geometric interpretation of witnesses to large projective dimension (see Definition 4.2.1).



(A) If $\gcd(x^\alpha, x^\beta)$ has degree $< D$, then the intersection of Δ_α and Δ_β is the empty set, so $\text{Cov}[w_\alpha, w_\beta] = 0$.

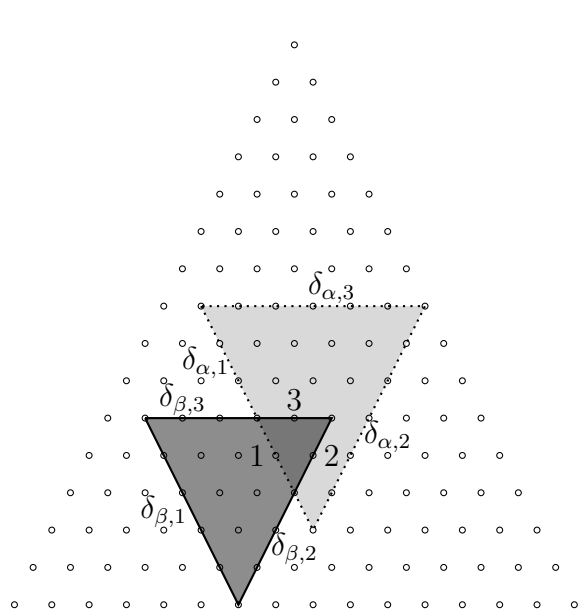


(B) If $\gcd(x^\alpha, x^\beta)$ has degree $= D$, then $\Delta_\alpha \cap \Delta_\beta$ contains a single point, which lies on a lower-dimensional face of at least one of the simplices, so $\text{Cov}[w_\alpha, w_\beta] = 0$.

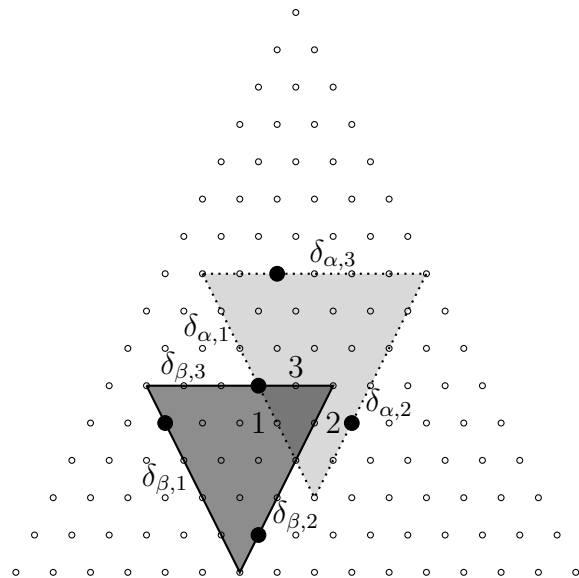


(C) If $x^\alpha | x^\beta$, then $\Delta_\alpha \subseteq \Delta_\beta$. In this case, $\text{Cov}[w_\alpha, w_\beta] < 0$.

FIGURE 4.2. Situations where pairs of witness lcm's have negative or zero covariance.



(A) An intersection of Δ_α (light gray/dotted) and Δ_β (dark gray/solid). The facets of the intersection are labeled 1, 2, 3, and the color class of [3] associated with this intersection is $(-, +, +)$. This encodes that $C_\alpha = \{2, 3\}$, $C_\beta = \{1\}$ and $C_\gamma = \emptyset$. Since $1 \in C_\beta$, the facet $\delta_{\beta,1}$ does not intersect Δ_α . Similarly, since $C_\alpha = \{2, 3\}$, the facets $\delta_{\alpha,2}$ and $\delta_{\alpha,3}$ do not intersect Δ_β .



(B) A set of five generators (above, in black), for which $w_\alpha w_\beta = 1$. Since one generator belongs to the intersection of facets 1 and 3, the associated bipartite graph H (below) has edge $\{1, 3\}$. Here $V = \{2\}$, indicating that G must contain a generator in $\delta_{\beta,2} \setminus (\delta_{\alpha,1} \cup \delta_{\alpha,3})$.

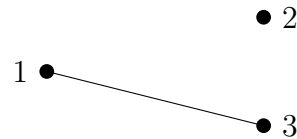


FIGURE 4.3. An illustration of intersection types, color classes, the graph H , and the set V .

Genericity and Scarf complexes in the graded model

5.1. A threshold for genericity

Recall (Definition 1.3.35) that an ideal with minimal generating set G is strongly generic if for every distinct pair of monomials x^α and x^β in G , either $\alpha_i = 0$ or $\alpha_i \neq \beta_i$ for all $i = 1, \dots, n$. If a monomial ideal is generated in degree D only, it is generic (Definition 1.3.40) if and only if it is strongly generic.

THEOREM 5.1.1. Let $S = \mathbb{k}[x_1, \dots, x_n]$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, and $p = p(D)$. As $D \rightarrow \infty$, $p = D^{-n+3/2}$ is a threshold for \mathfrak{M} being (strongly) generic. In other words, if $p(D) = o(D^{-n+3/2})$ then \mathfrak{M} is (strongly) generic a.a.s., and if $p(D) = \omega(D^{-n+3/2})$ then \mathfrak{M} is (strongly) generic a.a.s.

PROOF. Let V be the indicator variable that \mathfrak{M} is strongly generic. For each variable x_i and each exponent c , let $v_{i,c}$ denote the indicator variable for the event that there is at most one monomial in \mathfrak{G} with x_i exponent equal to c , and let $V_i = \prod_{c=1}^D v_{i,c}$. Then

$$V = \prod_{i=1}^n V_i.$$

Given a set Γ of monomials of degree D in S with $\#\Gamma = m$, the probability that \mathfrak{G} contains at most one monomial in Γ is

$$\mathbb{P}[\#(\Gamma \cap \mathfrak{G}) \leq 1] = q^m + mpq^{m-1} \geq 1 - mp + mp(1 - (m-1)p) \geq 1 - m^2p^2.$$

On the other hand

$$\mathbb{P}[\#(\Gamma \cap \mathfrak{G}) \leq 1] \leq \mathbb{P}[\#(\Gamma \cap \mathfrak{G}) \neq 2] = 1 - \binom{m}{2} p^2 q^{m-2}.$$

Assuming that $p = o(m^{-1})$ then for p sufficiently small, $q^{m-2} \geq 1/2$ so

$$(5.1.2) \quad \mathbb{P}[\#(\Gamma \cap \mathfrak{G}) \leq 1] \leq 1 - \frac{(m-1)^2}{4} p^2.$$

The above gives bounds on $\mathbb{P}[v_{i,c}]$ by taking Γ to be the set of monomials of degree D with x_i degree equal to c . Then $\#\Gamma = m_{n-1}(D-c) \leq D^{n-2}$, hence

$$\mathbb{P}[v_{i,c}] \geq 1 - D^{2n-4} p^2.$$

By the union-bound,

$$\mathbb{P}[V] \geq 1 - \sum_{i=1}^n \sum_{c=1}^D (1 - \mathbb{P}[v_{i,c}]) \geq 1 - np^2 D^{2n-3}.$$

Therefore, for $p = o(D^{-n+3/2})$, $\mathbb{P}[V]$ goes to 1.

For a lower bound on $\mathbb{P}[V_i]$, let U_i be the random variable that counts the number of values of c for which $v_{i,c}$ is false. Assuming that $p = o(D^{-n+2})$ and p sufficiently small, and using the upper bound on $\mathbb{P}[v_{i,c}]$ established in Equation (5.1.2), we get

$$\mathbb{E}[U_i] = \sum_{c=1}^D (1 - \mathbb{P}[v_{i,c}]) \geq \frac{p^2}{4} \sum_{c=1}^D (m_{n-1}(D-c) - 1)^2.$$

The function $f(D) = \sum_{c=1}^D (m_{n-1}(D-c) - 1)^2$ is a polynomial in D with lead term $t = D^{2n-3}/(n-2)!^2(2n-3)$. Thus for D sufficiently large, $f(D) \geq t/2$ so

$$\mathbb{E}[U_i] \geq \frac{p^2 D^{2n-3}}{8(n-2)!^2(2n-3)}.$$

Therefore, for $p = \omega(D^{-n+3/2})$, $p = o(D^{-n+2})$,

$$\lim_{D \rightarrow \infty} \mathbb{E}[U_i] = \infty.$$

Since the indicator variables $v_{i,1}, \dots, v_{i,D}$ are independent, $\text{Var}[U_i] \leq \mathbb{E}[U_i]$. By the second moment method,

$$0 = \lim_{D \rightarrow \infty} \mathbb{P}[U_i = 0] = \lim_{D \rightarrow \infty} \mathbb{P}[V_i] \geq \lim_{D \rightarrow \infty} \mathbb{P}[V].$$

Finally, note that for D fixed, $\mathbb{P}[V]$ is monotonically decreasing in p . Therefore $\mathbb{P}[V]$ goes to 0 as D goes to infinity for all $p = \omega(D^{-n+3/2})$. \square

5.2. Almost a threshold for being Scarf

The threshold for genericity (Theorem 5.1.1) implies that for $p = o(D^{-n+3/2})$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$ is asymptotically almost always Scarf. The main result of this section is that $D \rightarrow \infty$, \mathfrak{M} is almost never Scarf when $p = \omega(D^{-n+2-1/n})$. This leaves a gap where we do not know the asymptotic behavior.

THEOREM 5.2.1. Let $S = \mathbb{k}[x_1, \dots, x_n]$, $\mathfrak{M} \sim \mathcal{M}(n, D, p)$, and $p = p(D)$. If $p = \omega(D^{-n+2-1/n})$ then \mathfrak{M} is not Scarf asymptotically almost surely.

The proof of Theorem 5.2.1 is long, so first we will give a logical outline. Recall that every face of the Scarf complex of S/\mathfrak{M} corresponds to a subset of \mathfrak{G} whose lcm is unique over all subsets. Suppose that $L \subseteq \mathfrak{G}$ is a witness set to $\text{pdim}(S/\mathfrak{M}) = n$. By Theorem 4.1.1, the free module $S(-\alpha_L)$ appears in the minimal free resolution of S/\mathfrak{M} in homological degree n . Suppose further that there exists $g \in \mathfrak{G} \setminus L$, such that g divides $\text{lcm}(L)$. Then $\text{lcm}(L) = \text{lcm}(L \cup \{g\})$, so by definition $S(-\alpha_L)$ does *not* appear in the Scarf complex. Thus, the minimal free resolution strictly contains the Scarf complex, and \mathfrak{M} is not Scarf. When this occurs, we call $L \cup \{g\}$ a *non-Scarf witness set*. To prove Theorem 5.2.1, then, we prove that for $p = \omega(D^{-n+2-1/n})$, there is at least one non-Scarf witness set a.a.s.

For each $x^\alpha \in S$, define y_α as the indicator random variable:

$$y_\alpha = \begin{cases} 1 & x^\alpha \text{ is the lcm of a non-Scarf witness set} \\ 0 & \text{otherwise.} \end{cases}$$

For each integer $a \geq 1$, define the random variable Y_a that counts the monomials of degree $D + a$ that are lcm's of non-Scarf witness sets. Let Y be the sum of these variables over a certain range of a :

$$Y_a = \sum_{\substack{|\alpha|=D+a \\ \alpha_i \geq a \forall i}} y_\alpha, \quad Y = \sum_{a=2}^A Y_a$$

where $A = \lfloor (p/2)^{-\frac{1}{n-1}} \rfloor - n$.

For y_α to be true, there must be a monomial in \mathfrak{G} in the relative interior of each facet of the simplex Δ_α and one of the facets must have at least two monomials in \mathfrak{G} . Additionally \mathfrak{G} must have no monomials in the interior of Δ_α . For $x^\alpha \in S$ with $|\alpha| = D + a$, and $\alpha_i \geq a$ for $i = 1, \dots, n$,

(5.2.2)

$$\mathbb{P}[Y_a] = m_n(D+a-na) \left((1 - q^{m_{n-1}(a-n+1)})^n - (m_{n-1}(a-n+1)pq^{m_{n-1}(a-n+1)-1})^n \right) q^{m_n(a-n)}.$$

This follows from the same argument as the formula 4.2.5, subtracting the case that exactly one monomial lies on each facet. The relevant bound is

LEMMA 5.2.3. Let α be an exponent vector with $a = |\alpha| - D \leq p^{-\frac{1}{n-1}}$ and $\alpha_i \geq a$ for all i . Then,

$$(5.2.4) \quad \frac{1}{4}p^{n+1}m_{n-1}(a-n+1)^{n+1} \leq \mathbb{P}[y_\alpha] \leq \frac{1}{2}p^{n+1}m_{n-1}(a-n+1)^{n+1}.$$

PROOF. The union-bound implies that

$$1 - q^{m_{n-1}(a-n+1)} \leq pm_{n-1}(a-n+1).$$

The upper bound on $\mathbb{P}[y_\alpha]$ follows from applying this inequality to the expression in Equation (4.2.5).

For the lower-bound, note that $\mathbb{P}[y_\alpha]$ is bounded below by the probability that exactly two monomials are chosen to be in \mathfrak{G} from the relative interior of one of the facets of Δ_α and exactly one is chosen from each other facet, and no other monomials are chosen in Δ_α .

The probability of this event is given by

$$\binom{m_n(a-n)}{2} m_n(a-n)^{n-1} p^{n+1} q^{m_n(a)-n-1}$$

since there are $m_n(a - n)$ choices for the monomial chosen in each facet. Also by the union-bound we have

$$q^{m_n(a)-n-1} \geq 1 - (m_n(a) - n - 1)p \geq 1 - \frac{(a + n)^{n-1}}{(n - 1)!}p \geq \frac{1}{2}.$$

□

We can then find a threshold for p where non-Scarff witness sets are expected to appear frequently.

LEMMA 5.2.5. If $p = \omega(D^{-n+2-1/n})$ then $\lim_{D \rightarrow \infty} \mathbb{E}[Y] = \infty$.

PROOF. We follow the same argument as in the proof of Lemma 4.2.9. If $\lim_{D \rightarrow \infty} p > 0$, then $\mathbb{E}[Y_n] \geq m_n(D - 2)p^{n+1}q$ which goes to infinity in D . Instead assume that $p = \omega(D^{-n+2-1/n})$, $p = o(1)$ and take $n - 1 \leq a \leq p^{-\frac{1}{n-1}}$. As in the proof of Lemma 4.2.9, for D sufficiently large

$$m_n(D + a - na) \geq \frac{D^{n-1}}{2^{n-1}(n - 1)!}.$$

Therefore

$$\mathbb{E}[Y_a] \geq c_n D^{n-1} p^{n+1} a^{(n+1)(n-2)}$$

where $c_n > 0$ is constants that depends only on n . Summing up over a gives the bound

$$\mathbb{E}[Y] \geq c'_n D^{n-1} p^{\frac{n}{n-1}}$$

and $D^{n-1} p^{\frac{n}{n-1}}$ goes to infinity as $D \rightarrow \infty$. □

LEMMA 5.2.6. If $p = \omega(D^{-n+2-1/n})$ then

$$\lim_{D \rightarrow \infty} \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} = 0.$$

PROOF. The proof follows the same structure as that of Lemma 4.2.10. We bound $\text{Var}[Y]$ by

$$\text{Var}[Y] \leq \mathbb{E}[V] + \sum_{(\alpha, \beta)} \text{Cov}[y_\alpha, y_\beta].$$

For the pair of exponent vectors (α, β) , y_α and y_β are independent or mutually exclusive in the same set of cases as for w_α and w_β , in which case $\text{Cov}[y_\alpha, y_\beta]$ is non-positive. The remaining case is when the simplices Δ_α and Δ_β intersect and neither is contained in the other. Let $C = (C_\alpha, C_\beta, C_\gamma)$ be the coloring corresponding to this pair.

Define indicators e_i , $v_{i,j}$ and graph H as in the proof of Lemma 4.2.10. It was shown that $\mathbb{P}[w_\alpha w_\beta]$ is bounded above by

$$B = 2^{n^2 + \#C_\gamma} p^{\frac{2n - \#C_\gamma}{n-1}}.$$

For $y_\alpha y_\beta$ to be true, it must be that $w_\alpha w_\beta$ is true, plus an extra monomial appears in some facet of Δ_α and the same for Δ_β . We will enumerate the cases of how this can occur, and modify the bound B in each case to give a bound on $\mathbb{P}[y_\alpha y_\beta]$. Recall that for a set Γ of size m , we have that the probability of at least 2 monomials in \mathfrak{G} being chosen from Γ is bounded

$$\mathbb{P}[\#(\Gamma \cap \mathfrak{G}) \geq 2] \leq m^2 p^2.$$

There are two cases where a single monomial in \mathfrak{G} is the extra one for both y_α and y_β :

- For some $i \in C_\gamma$, there are at least two monomials in $\delta_{\alpha,i} \cap \delta_{\beta,i}$. The probability that this occurs is bounded by $m_{n-1}(A)^2 p^2 \leq p^{\frac{2}{n-1}}$ and this replaces a factor in the original bound B of $p^{\frac{1}{n-1}}$, so the probability of $y_\alpha y_\beta$ being true and this occurring for some fixed choice of i is bounded by $B p^{\frac{1}{n-1}}$.
- For some edge (i, j) of H , there are at least two monomials in $\delta_{\alpha,i} \cap \delta_{\beta,j}$. The probability that this occurs is bounded by $m_{n-2}(A)^2 p^2 \leq p^{\frac{4}{n-1}}$ and this replaces a factor in B of $p^{\frac{2}{n-1}}$.

In the rest of the cases the extra monomial for v_α is distinct from the extra one for v_β . For $v_\alpha v_\beta$ to be true, two of these cases must be paired. We describe the situation for v_α , but the v_β case is symmetric.

- For some $i \in C_\beta$, the vertex in the graph H has degree at least 2. In this case $2\#E(H) + \#V \geq \#C_\alpha + \#C_\beta + 1$, one greater than the bound in the original computation of B . Thus we pick up an extra factor of $p^{\frac{1}{n-1}}$ over B .
- For $i \in C_\alpha$ or $i \in C_\beta \cap V$ or $i \in C_w$ with no monomial in $\delta_{\alpha,i} \cap \delta_{\beta,i}$, there are at least two monomials in $\delta_{\alpha,i} \setminus \bigcup_j \delta_{\beta,j}$. We replace a factor of $p^{\frac{1}{n-1}}$ in B by $p^{\frac{2}{n-1}}$.
- For $i \in C_\beta \setminus V$ or $i \in C_w$ with a monomial in $\delta_{\alpha,i} \cap \delta_{\beta,i}$, there is a monomial in $\delta_{\alpha,i} \setminus \bigcup_j \delta_{\beta,j}$. Thus in the bound we pick up an extra factor of $p^{\frac{1}{n-1}}$ over B .

The probability of the first case being true is bounded by $Bp^{\frac{1}{n-1}}$, while in all others it is bounded by $Bp^{\frac{2}{n-1}}$, and the former bound dominates. The total number of cases among all the situations above is some finite N (depending only on n) so we can conclude that

$$\mathbb{P}[y_\alpha y_\beta] \leq NBp^{\frac{1}{n-1}}.$$

The remainder of the proof is identical to that of Lemma 4.2.10, and so we arrive at

$$\text{Var}[Y] \leq N2^{n^2+n}D^{n-1}p^{\frac{n}{n-1}} \leq \frac{N2^{n^2+n}}{c'_n} \mathbb{E}[Y],$$

and therefore

$$\lim_{D \rightarrow \infty} \frac{\text{Var}[Y]}{\mathbb{E}[Y]^2} \leq \lim_{D \rightarrow \infty} \frac{c''_n}{\mathbb{E}[Y]} = 0.$$

□

PROOF OF THEOREM 5.2.1. If $p = \omega(D^{-n+2-1/n})$, Lemma 5.2.5 proves that $\mathbb{E}[Y] \rightarrow \infty$ as $D \rightarrow \infty$. By the second moment method, Lemma 5.2.6 implies that $\mathbb{P}[Y > 0] \rightarrow 1$. We conclude that there is a non-Scarf witness set a.a.s., in which case \mathfrak{M} is not Scarf. □

Notice that Theorem 5.2.1 does not provide a threshold result for being Scarf. Nevertheless, taken together with Theorem 5.1.1 it indicates that being Scarf is almost equivalent to being generic in our probabilistic model. Monomial ideals that are not generic but Scarf live in the small range $D^{-n+3/2} \ll p \ll D^{-n+2-1/n}$. This narrow “twilight zone” can be seen in Figure 5.1 as the transition region where black, grey, and white are all present.

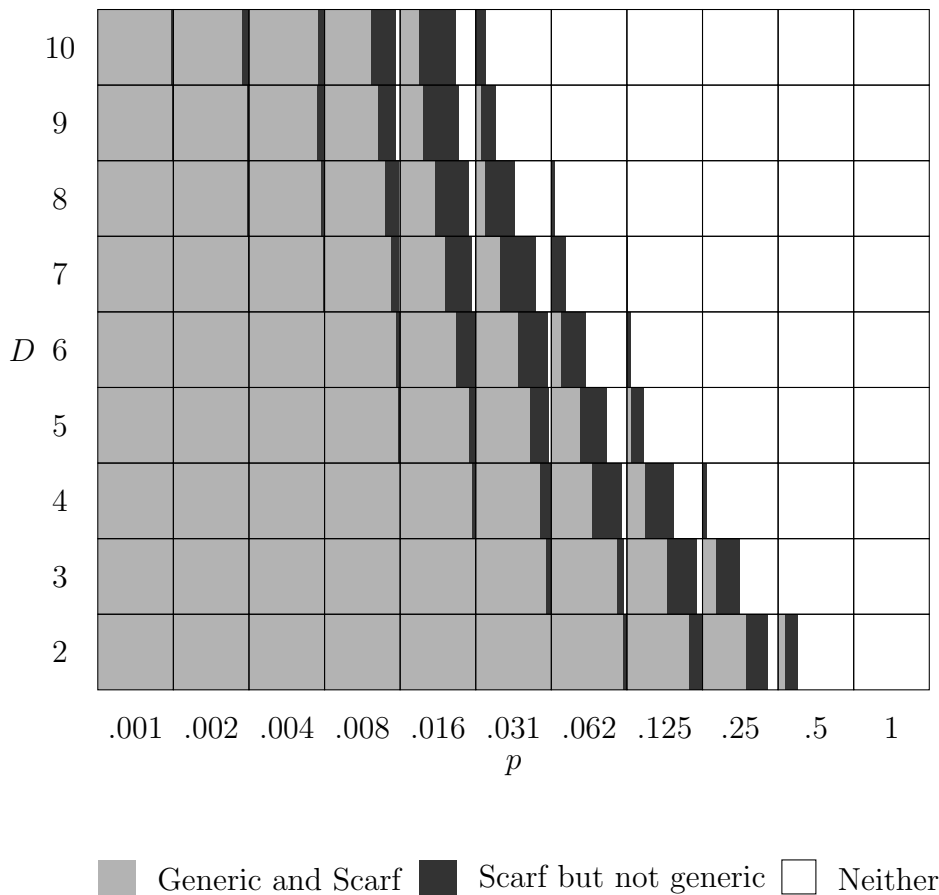


FIGURE 5.1. Generic versus Scarf monomial ideals in computer simulations of the graded model.

As an application of the probabilistic method, by choosing parameters in the twilight zone, we can generate countless examples of ideals with the unusual property of being Scarf but not generic. An example found while creating Figure 5.1 is $I = \langle x_1^4 x_3 x_5^5, x_1 x_2^2 x_3^2 x_6^4 x_8, x_2^3 x_5^2 x_6^3 x_7 x_8, x_1^3 x_5^2 x_7^2 x_8^3, x_2 x_3 x_4^3 x_6 x_8 x_9^3, x_1 x_3^4 x_4 x_6^2 x_8 x_{10}, x_1 x_3 x_4^2 x_5 x_6 x_8^3 x_{10}, x_2 x_3 x_6^3 x_8^4 x_{10}, x_4 x_5^5 x_7 x_{10}^3, x_1 x_5^4 x_{10}^5 \rangle \subseteq \mathbb{k}[x_1, \dots, x_{10}]$, which has the following total Betti numbers:

$$\begin{array}{c|cccccccc}
 i & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
 \beta_i & 1 & 10 & 45 & 114 & 168 & 147 & 75 & 20 & 2,
 \end{array}$$

and is indeed Scarf. Creating—or even verifying—such examples by hand would be a rather difficult task!

CHAPTER 6

Distributions of Hilbert functions

The *Hilbert function*, and the *Hilbert polynomial* that it determines, are important tools in the classification of ideals via the Hilbert scheme [Eis95]. Of course, monomial ideals are key to this enterprise: given an arbitrary homogeneous ideal I and a monomial order on $\mathbb{k}[x_1, \dots, x_n]$, 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 $h(S/I, d)$ is the number of standard monomials of degree d ; i.e. the number of monomials of degree d that are not contained in I .

Since the Erdős-Rényi-type model specifies a distribution on monomial ideals, it provides a formal probabilistic procedure to generate random Hilbert functions. Many natural questions arise, such as:

- *What is the probability of observing a particular Hilbert function?*
- *What is the most/least likely Hilbert function for a given set of parameters?*
- *If we fix a Hilbert function, are all monomial ideals with that Hilbert function equally likely to appear?*
- *How does the expected Hilbert function change as the parameter p changes?*

Recall from Theorem 2.1.4 that the Erdős-Rényi distribution of monomial ideals is intimately related to Hilbert functions, via Equation (2.1.5):

$$\mathbb{P}[\mathfrak{J} = I] = p^{\beta_1(S/I)}(1 - p)^{\sum_{d=1}^D h(S/I, d)}.$$

It follows that two monomial ideals with the same Hilbert function and the same number of minimal generators have the same probability of occurring in the model. If we could

enumerate all the monomial ideals (in n variables, with generators of degree less than or equal to D) that share a given Hilbert function, we could use Equation (2.1.5) to calculate the probability of obtaining said Hilbert function in the model. This strategy leads to a closed formula for the induced distribution of Hilbert functions under the ER-type model (Theorem 6.1.5).

6.1. The De Loera polytope

LEMMA 6.1.1. Fix a (partial) Hilbert function $\mathbf{h} = (h_0, h_1, \dots, h_D)$. Let $\text{NMon}(n, D, \mathbf{h})$ equal the number of monomial ideals I satisfying:

- (1) $I \subseteq \mathbb{k}[x_1, \dots, x_n]$,
- (2) $h(S/I, d) = h_d$ for all $0 \leq d \leq D$, and
- (3) $\beta_{1,d}(S/I) = 0$ for all $d > D$.

Let $DL(n, D, \mathbf{h})$ denote the convex polytope defined by:

$$(6.1.2) \quad \sum_{|\alpha|=d} x_\alpha = \binom{n+d-1}{d} - h_d, \quad 1 \leq d \leq D,$$

$$(6.1.3) \quad x_\alpha \leq x_\gamma, \quad \forall \alpha \leq \gamma \text{ with } |\alpha| + 1 = |\gamma|,$$

$$(6.1.4) \quad 0 \leq x_\alpha \leq 1 \quad \forall a$$

where a, γ range over all exponent vectors of monomials in S with positive total degree no more than D . (Thus, there are $\binom{n+D}{D} - 1$ variables.) Then $\text{NMon}(n, D, \mathbf{h})$ equals the number of 0 – 1 vertices of the convex polytope $DL(n, D, \mathbf{h})$.

PROOF. Given a vertex $v \in V(DL(n, D, \mathbf{h}))$, define I_v according to the 0 – 1 coordinates of v : if $x_\alpha = 1$, then $x_\alpha \in I_v$, and if $x_\alpha = 0$, then $x_\alpha \notin I_v$. The inequalities in Equation (6.1.3) ensure that if x^α is in I_v , then any multiple of x^α is also in I_v , and thus that I_v is an ideal. The equalities in Equation (6.1.2) ensure that the values of the Hilbert function at each degree d are correct. On the other hand, any monomial ideal I with Hilbert function \mathbf{h} can be encoded as a 0 – 1 solution of $DL(n, D, \mathbf{h})$, by assigning ones to the x^α variables

whenever $x^\alpha \in I$, and zeros otherwise, noting that the linear constraints are automatically satisfied. \square

The De Loera polytope is named for my advisor, in the tradition of Buchberger vis-à-vis Gröbner bases.

THEOREM 6.1.5. Let $\mathbf{h} = (h_0, h_1, \dots, h_D)$ be a partial Hilbert function, and let $Q(n, D, \mathbf{h})$ be the polytope from Lemma 6.1.1. Then the probability of \mathbf{h} in the ER-type model (in other words the probability that $\mathfrak{J} \sim \mathcal{I}(n, D, p)$ satisfies $h(S/\mathfrak{J}, d) = h_d$ for all $0 \leq d \leq D$) is given by:

$$(6.1.6) \quad \mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}] = (1-p)^{\sum_{d=1}^D h_d} \sum_{v \in V(Q(n, D, \mathbf{h}))} p^{\beta_1(v)},$$

where $\beta_1(v)$ is shorthand for $\beta_1(S/I_v)$, and I_v is the monomial ideal associated to v via the bijection in the proof of Lemma 6.1.1.

PROOF. The random monomial ideal \mathfrak{J} has Hilbert function \mathbf{h} if and only if \mathfrak{J} is one of the ideals in bijection with the vertices of $Q(n, D, \mathbf{h})$. Since these are disjoint events,

$$(6.1.7) \quad \mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}] = \sum_{v \in V(Q(n, D, \mathbf{h}))} \mathbb{P}[\mathfrak{J} = I_v].$$

For each choice of v , $\mathbb{P}[\mathfrak{J} = I_v]$ is given by Theorem 2.1.4 and is completely determined by $\sum_{d=1}^D h_d$ and by $\beta_1(v)$. The factor $(1-p)^{\sum_{d=1}^D h_d}$ is common to each term since all these ideals have the same Hilbert function. The first Betti number may be different for different choices of v . \square

An extension of Lemma 6.1.1 defines a polytope that counts monomial ideals with both a prescribed Hilbert function *and* prescribed first graded Betti numbers.

LEMMA 6.1.8. Denote by $\text{NMon}(n, D, \mathbf{h}, \beta_1)$ the number of monomial ideals I satisfying:

- (1) $I \subseteq \mathbb{k}[x_1, \dots, x_n]$,
- (2) $h(S/I, d) = h_d$ for all $0 \leq d \leq D$, and

(3) $\beta_{1,d}(S/I) = \beta_{1,d}$ for all $1 \leq d \leq D$, and

(4) $\beta_{1,d}(S/I) = 0$ for all $d > D$.

Let $Q(n, D, \mathbf{h}, \beta_1)$ denote the convex polytope defined by:

$$(6.1.9) \quad \sum_{|\alpha|=d} x_\alpha = \binom{n+d-1}{d} - h_d, \quad 1 \leq d \leq D,$$

$$(6.1.10) \quad \sum_{|\alpha|=d} y_\alpha = \beta_d, \quad 1 \leq d \leq D,$$

$$(6.1.11) \quad x_\alpha \leq x_\gamma, \quad \forall \alpha \leq \gamma \text{ with } |\alpha| + 1 = |\gamma|,$$

$$(6.1.12) \quad x_\alpha - \sum_{|\gamma|+1=|\alpha|} x_\alpha \leq y_\alpha \leq x_\alpha, \quad \forall \alpha,$$

$$(6.1.13) \quad y_\alpha + y_\gamma \leq 1, \quad \forall \alpha \leq \gamma$$

$$(6.1.14) \quad 0 \leq x_\alpha, y_\alpha \leq 1 \quad \forall \alpha$$

where α, γ range over all exponent vectors of monomials in S with positive total degree no more than D . (Thus, there are $2^{\binom{n+D}{D}} - 2$ variables.) Then the number of 0 – 1 vertices of the convex polytope $Q(n, D, \mathbf{h}, \beta_1)$ is equal to $\text{NMon}(n, D, \mathbf{h}, \beta_1)$.

PROOF. This proof is very similar to the proof of Lemma 6.1.1, except that in addition to the x_α variables that indicate whether each monomial is in the ideal, we add a set of y_α variables that indicate the monomials that are minimal generators of the ideal.

Given a vertex $v \in V(Q(n, D, \mathbf{h}, \beta_1))$, the ideal I_v is, as in the proof of Lemma 6.1.1, defined by the 0 – 1 coordinates of the x_α variables: if $x_\alpha = 1$, then $x_\alpha \in I_v$, and if $x_\alpha = 0$, then $x_\alpha \notin I_v$. Equations (6.1.9) and (6.1.11) again ensure that I_v is in fact an ideal, and has the correct Hilbert function.

The y_α variables of vertex v “mark” the minimal generators of I_v in that $y_\alpha = 1$ in v if and only if x^α is a minimal generator of I_v . This is due to the additional constraints in Equations (6.1.12) and (6.1.13). If x^α is a minimal generator of I_v , then $x^\alpha \in I_v$ but no monomial dividing x^α is in I_v . Because of Equation (6.1.11), it is enough to check the

monomials that have total degree exactly one less than x^α . If none of these is in I_v , but x^α is, then

$$x_\alpha - \sum_{|\gamma|+1=|\alpha|} x_\gamma = x_\alpha = 1,$$

so Equation (6.1.12) forces $y_\alpha = 1$. If $x^\alpha \notin I_v$, then $x_\alpha - \sum_{|\gamma|+1=|\alpha|} x_\gamma = x_\alpha = 0$ so the same inequalities force $y_\alpha = 0$. The only other situation is that $x^\alpha \in I_v$ but is not minimal, in which case at least one of the immediate divisors is also in I_v . In this case $x_\alpha - \sum_{|\gamma|+1=|\alpha|} x_\gamma$ is zero or negative, while $x_\alpha = 1$, so Equation (6.1.12) is redundant with Equation (6.1.14). However, in this case at least one monomial x^γ dividing x^α is a minimal generator, so Equation (6.1.13) enforces that only one of y_γ, y_α can equal one. Since the actually minimal one *must* equal one, as we've already shown, the non-minimal one cannot. This establishes that $y_\alpha = 1$ in v if and only if x^α is a minimal generator of I_v .

Finally, Equation (6.1.10) enforces the correct number of minimal generators of each degree.

□

REMARK 6.1.15. 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* [OS99].

6.2. Enumerating monomial ideals with specified Hilbert functions

Let $S = \mathbb{k}[x, y, z]$, and consider monomial ideals of S generated in degrees up to 4. For each non-negative integer vector $\mathbf{h} = (1, h_1, h_2, h_3, h_4)$ that could be the start of a Hilbert function of a graded ideal, by Macaulay's criterion, let $\text{NMon}(3, 4, \mathbf{h})$ denote the number of monomial ideals of S that satisfy $h(S/I, d) = h_d$ for $1 \leq d \leq 4$, and that are generated in degree no more than 4. Table 6.1 gives the value of $\text{NMon}(3, 4, \mathbf{h})$ for every allowable partial Hilbert function \mathbf{h} .

TABLE 6.1. All Hilbert functions of ideals of $\mathbb{k}[x, y, z]$, generated in degree up to four, along with the number of distinct monomial ideals with that Hilbert function.

h	NMon(3, 4, h)	h	NMon	h	NMon
1, 0, 0, 0, 0	1	1, 1, 0, 0, 0	3	1, 1, 1, 0, 0	3
1, 1, 1, 1, 0	3	1, 1, 1, 1, 1	3	1, 2, 0, 0, 0	3
1, 2, 1, 0, 0	9	1, 2, 1, 1, 0	6	1, 2, 1, 1, 1	6
1, 2, 2, 0, 0	9	1, 2, 2, 1, 0	18	1, 2, 2, 1, 1	12
1, 2, 2, 2, 0	9	1, 2, 2, 2, 1	18	1, 2, 2, 2, 2	9
1, 2, 3, 0, 0	3	1, 2, 3, 1, 0	12	1, 2, 3, 1, 1	6
1, 2, 3, 2, 0	18	1, 2, 3, 2, 1	27	1, 2, 3, 2, 2	9
1, 2, 3, 3, 0	12	1, 2, 3, 3, 1	36	1, 2, 3, 3, 2	36
1, 2, 3, 3, 3	12	1, 2, 3, 4, 0	3	1, 2, 3, 4, 1	15
1, 2, 3, 4, 2	30	1, 2, 3, 4, 3	30	1, 2, 3, 4, 4	15
1, 2, 3, 4, 5	3	1, 3, 0, 0, 0	1	1, 3, 1, 0, 0	6
1, 3, 1, 1, 0	3	1, 3, 1, 1, 1	3	1, 3, 2, 0, 0	15
1, 3, 2, 1, 0	21	1, 3, 2, 1, 1	15	1, 3, 2, 2, 0	9
1, 3, 2, 2, 1	18	1, 3, 2, 2, 2	9	1, 3, 3, 0, 0	20
1, 3, 3, 1, 0	55	1, 3, 3, 1, 1	30	1, 3, 3, 2, 0	54
1, 3, 3, 2, 1	87	1, 3, 3, 2, 2	36	1, 3, 3, 3, 0	22
1, 3, 3, 3, 1	66	1, 3, 3, 3, 2	66	1, 3, 3, 3, 3	22
1, 3, 3, 4, 0	3	1, 3, 3, 4, 1	15	1, 3, 3, 4, 2	30
1, 3, 3, 4, 3	30	1, 3, 3, 4, 4	15	1, 3, 3, 4, 5	3
1, 3, 4, 0, 0	15	1, 3, 4, 1, 0	69	1, 3, 4, 1, 1	30
1, 3, 4, 2, 0	126	1, 3, 4, 2, 1	156	1, 3, 4, 2, 2	54
1, 3, 4, 3, 0	114	1, 3, 4, 3, 1	279	1, 3, 4, 3, 2	231
1, 3, 4, 3, 3	66	1, 3, 4, 4, 0	51	1, 3, 4, 4, 1	207

Continued on next page

Table 6.1 – *Continued from previous page*

h	NMon(3, 4, h)	h	NMon	h	NMon
1, 3, 4, 4, 2	324	1, 3, 4, 4, 3	240	1, 3, 4, 4, 4	81
1, 3, 4, 4, 5	9	1, 3, 4, 5, 0	9	1, 3, 4, 5, 1	54
1, 3, 4, 5, 2	135	1, 3, 4, 5, 3	180	1, 3, 4, 5, 4	135
1, 3, 4, 5, 5	54	1, 3, 4, 5, 6	9	1, 3, 5, 0, 0	6
1, 3, 5, 1, 0	42	1, 3, 5, 1, 1	15	1, 3, 5, 2, 0	126
1, 3, 5, 2, 1	123	1, 3, 5, 2, 2	36	1, 3, 5, 3, 0	210
1, 3, 5, 3, 1	396	1, 3, 5, 3, 2	267	1, 3, 5, 3, 3	66
1, 3, 5, 4, 0	210	1, 3, 5, 4, 1	654	1, 3, 5, 4, 2	786
1, 3, 5, 4, 3	450	1, 3, 5, 4, 4	117	1, 3, 5, 4, 5	9
1, 3, 5, 5, 0	126	1, 3, 5, 5, 1	591	1, 3, 5, 5, 2	1131
1, 3, 5, 5, 3	1125	1, 3, 5, 5, 4	609	1, 3, 5, 5, 5	168
1, 3, 5, 5, 6	18	1, 3, 5, 6, 0	42	1, 3, 5, 6, 1	279
1, 3, 5, 6, 2	792	1, 3, 5, 6, 3	1245	1, 3, 5, 6, 4	1170
1, 3, 5, 6, 5	657	1, 3, 5, 6, 6	204	1, 3, 5, 6, 7	27
1, 3, 5, 7, 0	6	1, 3, 5, 7, 1	54	1, 3, 5, 7, 2	216
1, 3, 5, 7, 3	504	1, 3, 5, 7, 4	756	1, 3, 5, 7, 5	756
1, 3, 5, 7, 6	504	1, 3, 5, 7, 7	216	1, 3, 5, 7, 8	54
1, 3, 5, 7, 9	6	1, 3, 6, 0, 0	1	1, 3, 6, 1, 0	10
1, 3, 6, 1, 1	3	1, 3, 6, 2, 0	45	1, 3, 6, 2, 1	36
1, 3, 6, 2, 2	9	1, 3, 6, 3, 0	120	1, 3, 6, 3, 1	183
1, 3, 6, 3, 2	102	1, 3, 6, 3, 3	22	1, 3, 6, 4, 0	210
1, 3, 6, 4, 1	525	1, 3, 6, 4, 2	510	1, 3, 6, 4, 3	241
1, 3, 6, 4, 4	51	1, 3, 6, 4, 5	3	1, 3, 6, 5, 0	252
1, 3, 6, 5, 1	945	1, 3, 6, 5, 2	1440	1, 3, 6, 5, 3	1149
1, 3, 6, 5, 4	507	1, 3, 6, 5, 5	114	1, 3, 6, 5, 6	9

Continued on next page

Table 6.1 – *Continued from previous page*

\mathbf{h}	NMon(3, 4, \mathbf{h})	\mathbf{h}	NMon	\mathbf{h}	NMon
1, 3, 6, 6, 0	210	1, 3, 6, 6, 1	1113	1, 3, 6, 6, 2	2490
1, 3, 6, 6, 3	3056	1, 3, 6, 6, 4	2235	1, 3, 6, 6, 5	984
1, 3, 6, 6, 6	245	1, 3, 6, 6, 7	27	1, 3, 6, 7, 0	120
1, 3, 6, 7, 1	861	1, 3, 6, 7, 2	2694	1, 3, 6, 7, 3	4809
1, 3, 6, 7, 4	5376	1, 3, 6, 7, 5	3885	1, 3, 6, 7, 6	1806
1, 3, 6, 7, 7	519	1, 3, 6, 7, 8	84	1, 3, 6, 7, 9	6
1, 3, 6, 8, 0	45	1, 3, 6, 8, 1	423	1, 3, 6, 8, 2	1782
1, 3, 6, 8, 3	4428	1, 3, 6, 8, 4	7182	1, 3, 6, 8, 5	7938
1, 3, 6, 8, 6	6048	1, 3, 6, 8, 7	3132	1, 3, 6, 8, 8	1053
1, 3, 6, 8, 9	207	1, 3, 6, 8, 10	18	1, 3, 6, 9, 0	10
1, 3, 6, 9, 1	120	1, 3, 6, 9, 2	660	1, 3, 6, 9, 3	2200
1, 3, 6, 9, 4	4950	1, 3, 6, 9, 5	7920	1, 3, 6, 9, 6	9240
1, 3, 6, 9, 7	7920	1, 3, 6, 9, 8	4950	1, 3, 6, 9, 9	2200
1, 3, 6, 9, 10	660	1, 3, 6, 9, 11	120	1, 3, 6, 9, 12	10
1, 3, 6, 10, 0	1	1, 3, 6, 10, 1	15	1, 3, 6, 10, 2	105
1, 3, 6, 10, 3	455	1, 3, 6, 10, 4	1365	1, 3, 6, 10, 5	3003
1, 3, 6, 10, 6	5005	1, 3, 6, 10, 7	6435	1, 3, 6, 10, 8	6435
1, 3, 6, 10, 9	5005	1, 3, 6, 10, 10	3003	1, 3, 6, 10, 11	1365
1, 3, 6, 10, 12	455	1, 3, 6, 10, 13	105	1, 3, 6, 10, 14	15
1, 3, 6, 10, 15	1				

Generating Table 6.1 took approximately 37 seconds. See Appendix A.2 for details.

6.3. Explicit Hilbert functions probabilities

In addition to merely counting the monomial ideals with each Hilbert function, we can generate and save the lists of ideals. It took approximately 166 seconds to generate the list of

ideal	graded Betti numbers				
		0	1	2	3
	total:	1	9	12	4
$\langle z^3, yz^2, xz^2, y^2z, x^2z, y^3, xy^2, x^2y, x^3 \rangle$	0:	1	.	.	.
	1:
	2:	.	9	12	3
	3:	.	.	.	1
$\langle z^3, yz^2, xz^2, y^2z, xyz, y^3, xy^2, x^2y, x^3 \rangle$					
$\langle z^3, yz^2, xz^2, y^2z, xyz, x^2z, y^3, xy^2, x^3 \rangle$					
$\langle z^3, yz^2, xz^2, xyz, x^2z, y^3, xy^2, x^2y, x^3 \rangle$	total:	1	9	13	5
$\langle z^3, yz^2, xz^2, y^2z, xyz, x^2z, y^3, x^2y, x^3 \rangle$	0:	1	.	.	.
	1:
	2:	.	9	12	4
	3:	.	.	1	1
$\langle z^3, yz^2, y^2z, xyz, x^2z, y^3, xy^2, x^2y, x^3 \rangle$					
$\langle z^3, xz^2, y^2z, xyz, x^2z, y^3, xy^2, x^2y, x^3 \rangle$					
$\langle z^3, yz^2, xz^2, y^2z, xyz, x^2z, xy^2, x^2y, x^3, y^4 \rangle$		0	1	2	3
$\langle yz^2, xz^2, y^2z, xyz, x^2z, y^3, xy^2, x^2y, x^3, z^4 \rangle$	total:	1	10	15	6
$\langle z^3, yz^2, xz^2, y^2z, xyz, x^2z, y^3, xy^2, x^2y, x^4 \rangle$	0:	1	.	.	.
	1:
	2:	.	9	13	5
	3:	.	1	2	1

TABLE 6.2. All monomial ideals of $\mathbb{k}[x, y, z]$ with Hilbert function $\mathbf{h} = (1, 3, 6, 1, 0)$, along with their graded Betti tables.

ideals for every Hilbert function from Table 6.1. For example, there are 10 monomial ideals of $\mathbb{k}[x, y, z]$, generated in degree at most 4, with the Hilbert function $\mathbf{h} = (1, 3, 6, 1, 0, 0, \dots)$. Table 6.2 displays each of these ideals along with its Betti table. This example illustrates how monomial ideals with the same Hilbert function may have different graded (and total) Betti numbers.

The first seven ideals in Table 6.2 have nine minimal generators, so by Theorem 2.1.4 each has probability $p^9(1-p)^{10}$ of appearing in the ER-type model $\mathcal{I}(3, 4, p)$. The last three ideals in Table 6.2 have ten minimal generators, so each of these appears with probability $p^{10}(1-p)^{10}$. We conclude that for $\mathfrak{J} \sim \mathcal{I}(3, 4, p)$,

$$(6.3.1) \quad \mathbb{P}[h(S/\mathfrak{J}, d) = (1, 3, 6, 1, 0, \dots)] = (1-p)^{10}(3p^{10} + 7p^9).$$

By repeating this computation for all integer sequences allowable by Macaulay's theorem, we can explicitly render the probability of every Hilbert function in for any choice of n and D . For instance Table 6.3 displays the probability, as a polynomial in p , of every Hilbert function that can be generated in $\mathcal{I}(2, 5, p)$. In addition, the polynomial is evaluated for three choices of the parameter p : 0.5, 0.25, and 0.125. Table 6.3 took less than 2 seconds to produce.

TABLE 6.3. Probability of each partial Hilbert function of $\mathbb{k}[x, y]$, for monomial ideals generated in degree 5 or less.

\mathbf{h}	$\mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}]$	$p = 0.5$	$p = 0.25$	$p = 0.125$
1,0,0,0,0	p^2	.25	.0625	.01562
1,1,0,0,0	$2p^2(1-p)$.25	.09375	.02734
1,1,1,0,0	$2p^2(1-p)^2$.125	.07031	.02393
1,1,1,1,0	$2p^2(1-p)^3$.0625	.05273	.02094
1,1,1,1,1	$2p^2(1-p)^4$.03125	.03955	.01832
1,1,1,1,1	$2p(1-p)^5$.03125	.1187	.1282
1,2,0,0,0	$p^3(1-p)^2$.03125	.008789	.001495
1,2,1,0,0	$p^2(1-p)^3(2p+1)$.0625	.03955	.01308
1,2,1,1,0	$2p^3(1-p)^4$.01562	.009888	.00229
1,2,1,1,1	$2p^3(1-p)^5$.007812	.007416	.002004
1,2,1,1,1	$2p^2(1-p)^6$.007812	.02225	.01402
1,2,2,0,0	$3p^3(1-p)^4$.02344	.01483	.003435

Continued on next page

Table 6.3 – *Continued from previous page*

\mathbf{h}	$\mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}]$	$p = 0.5$	$p = 0.25$	$p = 0.125$
1,2,2,1,0,0	$2p^2(1-p)^5(2p+1)$.03125	.04449	.02004
1,2,2,1,1,0	$4p^3(1-p)^6$.007812	.01112	.003506
1,2,2,1,1,1	$4p^2(1-p)^7$.007812	.03337	.02454
1,2,2,2,0,0	$3p^3(1-p)^6$.005859	.008343	.00263
1,2,2,2,1,0	$2p^2(1-p)^7(2p+1)$.007812	.02503	.01534
1,2,2,2,1,1	$4p^2(1-p)^8$.003906	.02503	.02148
1,2,2,2,2,0	$3p^3(1-p)^8$.001465	.004693	.002013
1,2,2,2,2,1	$6p^2(1-p)^9$.00293	.02816	.02819
1,2,2,2,2,2	$3p(1-p)^{10}$.001465	.04224	.09865
1,2,3,0,0,0	$p^4(1-p)^5$.001953	.000927	.0001252
1,2,3,1,0,0	$2p^3(1-p)^6(p+1)$.005859	.006952	.001972
1,2,3,1,1,0	$2p^4(1-p)^7$.0009766	.001043	.0001917
1,2,3,1,1,1	$2p^3(1-p)^8$.0009766	.003129	.001342
1,2,3,2,0,0	$3p^3(1-p)^7(p+1)$.004395	.007821	.002589
1,2,3,2,1,0	$p^2(1-p)^8(2p+1)^2$.003906	.01408	.008389
1,2,3,2,1,1	$2p^2(1-p)^9(2p+1)$.001953	.01408	.01174
1,2,3,2,2,0	$3p^4(1-p)^9$.0003662	.0008799	.0002202
1,2,3,2,2,1	$6p^3(1-p)^{10}$.0007324	.005279	.003083
1,2,3,2,2,2	$3p^2(1-p)^{11}$.0003662	.007919	.01079
1,2,3,3,0,0	$4p^4(1-p)^8$.0009766	.001564	.0003356
1,2,3,3,1,0	$6p^3(1-p)^9(p+1)$.002197	.008799	.003964
1,2,3,3,1,1	$6p^3(1-p)^{10}$.0007324	.005279	.003083
1,2,3,3,2,0	$6p^3(1-p)^{10}(p+1)$.001099	.006599	.003468
1,2,3,3,2,1	$6p^2(1-p)^{11}(2p+1)$.001465	.02376	.02698
1,2,3,3,2,2	$6p^2(1-p)^{12}$.0003662	.01188	.01888

Continued on next page

Table 6.3 – *Continued from previous page*

h	$\mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}]$	$p = 0.5$	$p = 0.25$	$p = 0.125$
1,2,3,3,3,0	$4p^4(1-p)^{11}$.0001221	.0006599	.0002248
1,2,3,3,3,1	$12p^3(1-p)^{12}$.0003662	.005939	.004721
1,2,3,3,3,2	$12p^2(1-p)^{13}$.0003662	.01782	.03305
1,2,3,3,3,3	$4p(1-p)^{14}$.0001221	.01782	.07711
1,2,3,4,0,0	$p^5(1-p)^9$.00006104	.00007332	.000009175
1,2,3,4,1,0	$p^4(1-p)^{10}(2p+3)$.0002441	.0007699	.0002087
1,2,3,4,1,1	$2p^4(1-p)^{11}$.00006104	.00033	.0001124
1,2,3,4,2,0	$p^3(1-p)^{11}(3p^2+6p+1)$.0002899	.001774	.0008079
1,2,3,4,2,1	$6p^3(1-p)^{12}(p+1)$.0002747	.003712	.002655
1,2,3,4,2,2	$3p^3(1-p)^{13}$.00004578	.001114	.001033
1,2,3,4,3,0	$2p^4(1-p)^{12}(2p+3)$.0001221	.0008662	.0003196
1,2,3,4,3,1	$12p^3(1-p)^{13}(p+1)$.0002747	.005568	.004647
1,2,3,4,3,2	$6p^2(1-p)^{14}(2p+1)$.0001831	.01002	.01807
1,2,3,4,3,3	$4p^2(1-p)^{15}$.00003052	.003341	.008433
1,2,3,4,4,0	$5p^5(1-p)^{13}$.00001907	.000116	.00002689
1,2,3,4,4,1	$20p^4(1-p)^{14}$.00007629	.001392	.000753
1,2,3,4,4,2	$30p^3(1-p)^{15}$.0001144	.006264	.007906
1,2,3,4,4,3	$20p^2(1-p)^{16}$.00007629	.01253	.0369
1,2,3,4,4,4	$5p(1-p)^{17}$.00001907	.009396	.06457
1,2,3,4,5,0	$p^6(1-p)^{14}$	9.537e-7	.00000435	5.883e-7
1,2,3,4,5,1	$6p^5(1-p)^{15}$.000005722	.0000783	.00002471
1,2,3,4,5,2	$15p^4(1-p)^{16}$.00001431	.0005873	.0004324
1,2,3,4,5,3	$20p^3(1-p)^{17}$.00001907	.002349	.004035
1,2,3,4,5,4	$15p^2(1-p)^{18}$.00001431	.005285	.02119
1,2,3,4,5,5	$6p(1-p)^{19}$.000005722	.006342	.05932

Continued on next page

Table 6.3 – *Continued from previous page*

\mathbf{h}	$\mathbb{P}[h(S/\mathfrak{J}) = \mathbf{h}]$	$p = 0.5$	$p = 0.25$	$p = 0.125$
1,2,3,4,5,6	$(1 - p)^{20}$	9.537e-7	.003171	.06921

6.4. Betti tables with the same Hilbert function

In Table 6.2, we saw that monomial ideals with a common Hilbert function may have different numbers of minimal generators, and different higher Betti numbers. In that example, there was only one set of graded Betti numbers (i.e., only one Betti table) observed for each set of total Betti numbers. In general, there may be several Betti tables that correspond to the same Hilbert function and the same total Betti numbers. For example,

$$(6.4.1) \quad \mathbb{P}[h(S/\mathfrak{J}) = (1, 3, 5, 5, 4, \dots)] = (1 - p)^{17} (270p^5 + 300p^4 + 39p^3),$$

because there are 609 monomial ideals of $\mathbb{k}[x, y, z]$ with a Hilbert function that begins $(1, 3, 5, 4, 4, \dots)$: 270 with first Betti number 5, 300 with first Betti number 4, and 39 with first Betti number 3. However, there are *sixteen* different sets of graded Betti numbers. These are displayed, in the form of Betti tables, in Table 6.4.

TABLE 6.4. All Betti tables of ideals of $\mathbb{k}[x, y, z]$, generated in degree up to four, that have Hilbert function beginning $\mathbf{h} = (1, 3, 5, 5, 4)$. The second column displays the number of distinct monomial ideals with that Betti table.

Betti table	# mon. ideals	Betti table	# mon. ideals																																																																	
<table style="border-collapse: collapse; margin-left: 20px;"> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>3</td><td>3</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>1</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>.</td><td>2</td><td>1</td></tr> </table>		0	1	2	3	total:	1	3	3	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	1	.	3:	.	.	2	1	27	<table style="border-collapse: collapse; margin-left: 20px;"> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>3</td><td>3</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>1</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>.</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	3	3	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	1	.	3:	.	.	1	.	4:	.	.	1	1	12
	0	1	2	3																																																																
total:	1	3	3	1																																																																
0:	1	.	.	.																																																																
1:	.	1	.	.																																																																
2:	.	2	1	.																																																																
3:	.	.	2	1																																																																
	0	1	2	3																																																																
total:	1	3	3	1																																																																
0:	1	.	.	.																																																																
1:	.	1	.	.																																																																
2:	.	2	1	.																																																																
3:	.	.	1	.																																																																
4:	.	.	1	1																																																																

Continued on next page

Table 6.4 – *Continued from previous page*

Betti table	# mon. ideals	Betti table	# mon. ideals																																																																																
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>4</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>2</td><td>1</td></tr> </table>		0	1	2	3	total:	1	4	4	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	2	1	174	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>4</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	4	4	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	1	.	4:	.	.	1	1	78															
	0	1	2	3																																																																															
total:	1	4	4	1																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	2	1																																																																															
	0	1	2	3																																																																															
total:	1	4	4	1																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	1	.																																																																															
4:	.	.	1	1																																																																															
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>4</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>2</td><td>.</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>.</td><td>1</td></tr> </table>		0	1	2	3	total:	1	4	4	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	.	.	4:	.	.	2	.	5:	.	.	.	1	6	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>4</td><td>1</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	4	4	1	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	1	.	4:	5:	.	.	1	1	12
	0	1	2	3																																																																															
total:	1	4	4	1																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	.	.																																																																															
4:	.	.	2	.																																																																															
5:	.	.	.	1																																																																															
	0	1	2	3																																																																															
total:	1	4	4	1																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	1	.																																																																															
4:																																																																															
5:	.	.	1	1																																																																															
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>5</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>3</td><td>2</td></tr> </table>		0	1	2	3	total:	1	4	5	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	.	.	4:	.	.	3	2	15	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>5</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>1</td><td>.</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>2</td><td>2</td></tr> </table>		0	1	2	3	total:	1	4	5	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	.	.	4:	.	.	1	.	5:	.	.	2	2	9					
	0	1	2	3																																																																															
total:	1	4	5	2																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	.	.																																																																															
4:	.	.	3	2																																																																															
	0	1	2	3																																																																															
total:	1	4	5	2																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	.	.																																																																															
4:	.	.	1	.																																																																															
5:	.	.	2	2																																																																															
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>4</td><td>5</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>3:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>2</td><td>1</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	4	5	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	2	.	3:	.	1	.	.	4:	.	.	2	1	5:	.	.	1	1	6	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>6</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	6	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	3	1	150										
	0	1	2	3																																																																															
total:	1	4	5	2																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	2	.																																																																															
3:	.	1	.	.																																																																															
4:	.	.	2	1																																																																															
5:	.	.	1	1																																																																															
	0	1	2	3																																																																															
total:	1	5	6	2																																																																															
0:	1	.	.	.																																																																															
1:	.	1	.	.																																																																															
2:	.	2	3	1																																																																															
3:	.	2	3	1																																																																															

Continued on next page

Table 6.4 – *Continued from previous page*

Betti table	# mon. ideals	Betti table	# mon. ideals																																																																																					
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>6</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	6	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	2	.	4:	.	.	1	1	54	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>6</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	6	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	2	.	4:	5:	.	.	1	1	24										
	0	1	2	3																																																																																				
total:	1	5	6	2																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	2	.																																																																																				
4:	.	.	1	1																																																																																				
	0	1	2	3																																																																																				
total:	1	5	6	2																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	2	.																																																																																				
4:																																																																																				
5:	.	.	1	1																																																																																				
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>6</td><td>2</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>2</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>6:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	6	2	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	2	.	4:	5:	6:	.	.	1	1	6	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>7</td><td>3</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>3</td><td>2</td></tr> </table>		0	1	2	3	total:	1	5	7	3	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	1	.	4:	.	.	3	2	12					
	0	1	2	3																																																																																				
total:	1	5	6	2																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	2	.																																																																																				
4:																																																																																				
5:																																																																																				
6:	.	.	1	1																																																																																				
	0	1	2	3																																																																																				
total:	1	5	7	3																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	1	.																																																																																				
4:	.	.	3	2																																																																																				
<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>7</td><td>3</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>2</td><td>1</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	7	3	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	1	.	4:	.	.	2	1	5:	.	.	1	1	12	<table> <tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td></tr> <tr><td>total:</td><td>1</td><td>5</td><td>7</td><td>3</td></tr> <tr><td>0:</td><td>1</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>1:</td><td>.</td><td>1</td><td>.</td><td>.</td></tr> <tr><td>2:</td><td>.</td><td>2</td><td>3</td><td>1</td></tr> <tr><td>3:</td><td>.</td><td>2</td><td>1</td><td>.</td></tr> <tr><td>4:</td><td>.</td><td>.</td><td>2</td><td>1</td></tr> <tr><td>5:</td><td>.</td><td>.</td><td>.</td><td>.</td></tr> <tr><td>6:</td><td>.</td><td>.</td><td>1</td><td>1</td></tr> </table>		0	1	2	3	total:	1	5	7	3	0:	1	.	.	.	1:	.	1	.	.	2:	.	2	3	1	3:	.	2	1	.	4:	.	.	2	1	5:	6:	.	.	1	1	12
	0	1	2	3																																																																																				
total:	1	5	7	3																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	1	.																																																																																				
4:	.	.	2	1																																																																																				
5:	.	.	1	1																																																																																				
	0	1	2	3																																																																																				
total:	1	5	7	3																																																																																				
0:	1	.	.	.																																																																																				
1:	.	1	.	.																																																																																				
2:	.	2	3	1																																																																																				
3:	.	2	1	.																																																																																				
4:	.	.	2	1																																																																																				
5:																																																																																				
6:	.	.	1	1																																																																																				

Examination of Table 6.4 reveals an interesting trend: for each possible set of total Betti numbers, the Betti tables with lower regularity are much more frequent than the ones with higher regularity. Among the ideals with Betti numbers $(1, 4, 4, 1)$, there are 174 with regularity 3, 93 with regularity 4, and only 15 with regularity 5. Among the ideals with Betti numbers $(1, 5, 6, 2)$, there are 150 with regularity 3, 54 with regularity 4, 24 with regularity 5, and only 6 with regularity 6. Since each monomial ideal with the same Hilbert function and

first Betti number is equally likely to appear in the random case, this means that ideals with low regularity are much more likely to appear for these cases. This trend was seen in other computations, and helps to explain, at least experimentally, why the typical complexity of polynomial computations is often much better than worst-case complexity.

CHAPTER 7

Supervised machine learning in commutative algebra

Please refer to Section 1.5.1 for the basics of the machine learning methodology used in this chapter. Let us also take a moment to introduce how the accuracy of machine learning experiments will be reported in Section 7.3, with a figure called a *confusion matrix*. For multi-class classification problems, both index sets of the matrix list the labels for all classes. The value of the (i, j) entry of the matrix contains the number of test examples whose *true* label was i and whose *predicted* label was j . Thus, a perfectly accurate classifier would have nonzero entries only on the diagonal. For example, returning to the image classification of benign/malignant tumors used as the main example of Section 1.5.1, suppose we test a neural network trained on this task with 100 images whose true classifications are known to the supervisor, but withheld from the network. The confusion matrix in Figure 7.1 indicates that 70 benign tumors were correctly identified as benign, 7 benign tumors were incorrectly classified as malignant, 3 malignant tumors were incorrectly classified as benign, and 20 malignant tumors were correctly identified.

predicted	malig.	3	20
	benign	70	7
		benign	malig.
		actual	

FIGURE 7.1. An example of a confusion matrix that would describe the accuracy of image classification on 100 test images. In this (made-up) example, 70 benign tumors were correctly identified as benign, 7 benign tumors were incorrectly classified as malignant, 3 malignant tumors were incorrectly classified as benign, and 20 malignant tumors were correctly identified.

A confusion matrix supplies much more information than reporting the accuracy as a single value (in this example, 90% of examples correctly classified). For binary classification, this breakdown lets us calculate the false positive rate versus the false negative rate, which sometimes are of vastly different importance depending on the application. (E.g., giving unnecessary biopsies is preferable to missing deadly cancers.) For multi-class classification, the values of the confusion matrix show us exactly where the neural network was “confused” between different classes, hence the name. In addition, the confusion matrix encodes the frequency of the true classes encountered in the test set. For instance, in the previous example, the network saw disproportionately more benign examples than malignant ones in the test set. Throughout Chapter 7, we test neural networks with a randomly selected 10% of the training data (which is reserved and not seen by the network prior to testing). Thus, the frequency of classes in the test data is a reasonable approximation of their frequency in the data used to train. To supplement this, Figures 7.2c, 7.3c and 7.4c also display the actual histogram of class frequencies throughout the training data.

The squares in Figure 7.1 are also shaded according to the entries of the matrix, with darker shades corresponding to higher values. For the much larger confusion matrices of Section 7.3, the numerical values are omitted in some cases for readability, and a key to the shading scale provided instead. This makes it easier to visualize the overall performance of the classifiers.

7.1. Monomial ideal features and training data

Any supervised learning problem needs training data representative of the inputs on which the learned classifier will later operate. The monomial ideals used as training data in this chapter are those that the author found interesting: the Erdős-Rényi-type random monomial ideals defined in Section 2.1 and studied in Chapters 3 and 6 and [DPS⁺19], the graded model of random monomial ideals defined in Section 2.2 and studied in Chapters 4 and 5 and [DHKS19], squarefree monomial ideals (see Section 1.2.2) defined as edge ideals of graphs and hypergraphs, etc. Many of the concrete examples the author works with are

in no more than 5 or 10 variables, generated in degrees around the same order of magnitude, with a few to a few dozen generators. These examples are large enough to be interesting, yet small enough for a home laptop running `Macaulay2` to comfortably compute with in large quantities (for instance, when generating data for graphics like Figures 3.3 and 5.1).

There is no reason to expect the trained classification networks to perform well on monomial ideals with extremely many generators, generated in extremely high degrees, for instance. However, we can expect the same *methodology* to produce a good classifier for different classes of ideals, as long as representative training data is used.

Table 7.1 describes in detail the families of ideals considered in this chapter, along with the number of random samples from each family, and some descriptive statistics (explained in the table’s caption). For further explanation of the families, and code to generate more random examples, see Appendix A.2.

TABLE 7.1. Training data used for Sections 7.2 and 7.3. Each row describes a random family of ideals incorporated into the overall data set of 30,000 examples of monomial ideals. The first column (**vars**) is the number of variables of the polynomial ring to which the family of ideals belongs. The second column (**degs**) gives the range of total degrees of minimal generators. A single value means that the ideals are generated in a single degree. The column labeled **gens** gives the number of minimal generators, often a range for each family, with the average number across the family in parentheses. The **SF** column has entry “Y” if the monomial ideals are squarefree, and an “N” otherwise. Finally, the row has a brief description of the family, and the last column (**#**) indicates the number of ideals from that family. For more details about these families, and for `Macaulay2` code to generate more examples, see Appendix A.2.

vars	degs	gens (avg)	SF	description	#
5	1–20	1–12 (5)	N	ER-type random monomial ideals $\mathfrak{J} \sim \mathcal{I}(5, 20, 0.0001)$.	100
5	2–20	15–55 (34)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(5, 20, 0.001)$.	500
10	1–5	2–27 (14)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(10, 5, 0.005)$.	500
10	2–10	1–22 (9)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(10, 10, 0.00005)$.	500
10	4–10	7–28 (18)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(10, 10, 0.0001)$.	64
15	2–5	2–17 (8)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(15, 5, 0.0005)$.	500
20	1–5	1–12 (5)	N	ER-type, $\mathfrak{J} \sim \mathcal{I}(20, 5, 0.0001)$.	500

Continued on next page

Table 7.1 – *Continued from previous page*

vars	degs	gens (avg)	SF	description	#
5	5	1–14 (6)	N	Graded model random monomial ideals $\mathfrak{M} \sim \mathcal{M}(5, 5, 0.05)$.	1000
5	5	3–25 (13)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(5, 5, 0.1)$.	1000
5	10	2–20 (10)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(5, 10, 0.01)$.	1000
5	15	1–12 (4)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(5, 15, 0.001)$.	1000
5	15	8–39 (19)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(5, 15, 0.005)$.	462
10	5	1–20 (10)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(10, 5, 0.005)$.	461
10	5	7–31 (20)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(10, 5, 0.01)$.	283
10	10	2–20 (10)	N	Graded model, $\mathfrak{M} \sim \mathcal{M}(10, 10, 0.0001)$.	165
10	2	2–44 (23)	Y	Edge ideals of random graphs on 10 vertices.	880
10	2	9	Y	Edge ideals of spanning trees on 10 vertices.	1000
10	4	45	N	I^2 for I from the previous dataset.	1000
10	3	2–91 (43)	Y	Edge ideals of 3-uniform hypergraphs on 10 vertices.	1000
10	3	8	Y	Edge ideals of spanning 3-hypergraphs on 10 vertices.	1000
10	6	36	N	I^2 for I from the previous dataset	1000
10	4	2–91 (31)	Y	Edge ideals of 4-uniform hypergraphs on 10 vertices.	1000
10	5	2–50 (26)	Y	Edge ideals of 5-uniform hypergraphs on 10 vertices.	500
15	2	14	Y	Edge ideals of spanning trees on 15 vertices.	1000
15	3	10–30 (20)	Y	Edge ideals of 3-uniform hypergraphs on 15 vertices.	202
15	3	13	Y	Edge ideals of spanning 3-hypergraphs on 15 vertices.	1000
15	5–10	10–20 (14)	Y	Edge ideals of hypergraphs on 15 vertices with edges of various sizes.	198
20	2	19	Y	Edge ideals of spanning trees on 20 vertices.	482
16	2	5–15 (10)	Y	Edge ideals generated by knight moves on a 4x4 chessboard.	1100
20	2	5–15 (10)	Y	Edge ideals generated by knight moves on a 4x5 chessboard.	1100

Continued on next page

Table 7.1 – *Continued from previous page*

vars	degs	gens (avg)	SF	description	#
5	2–40	4	N	Edge ideals of spanning trees on 5 vertices, with random exponents (uniform on $\{1, \dots, 20\}$).	1000
5	2–42	3–7 (5)	N	Three steps of random Markov process on each ideal in the previous dataset.	3000
5	10–50	5–10	N	Edge ideals of random graphs on 5 vertices, with random exponents (uniform on $\{5, \dots, 25\}$).	120
10	2–10	9	N	Edge ideals of spanning trees on 10 vertices, with random exponents (uniform on $\{1, \dots, 5\}$).	531
10	2–13	8–12 (11)	N	Three steps of random Markov process on each ideal in the previous dataset.	1591
5	1–10	1–29 (11)	N	Markov process starting with (x_1, \dots, x_5) , with maximum degree 10.	1000
10	1–5	3–24 (10)	N	Markov process starting with (x_1, \dots, x_{10}) , with maximum degree 5.	1000
10	1–10	7–52 (26)	N	Markov process starting with (x_1, \dots, x_{10}) , with maximum degree 10.	182
5	3–10	5–16 (8)	N	Zero-dimensional ideals in 5 variables.	579
10	6–10	10–22 (13)	N	Zero-dimensional ideals in 10 variables.	375
15	7–10	15–20 (17)	N	Zero-dimensional ideals in 15 variables.	125

A neural network, as described in Section 1.5.1, takes a vector of fixed size as its input. The “natural” way to describe a monomial ideal is with a list of (exponent vectors of) its minimal generators, but such lists have varying lengths. Even for a fixed number of variables, n , and a fixed maximum degree, D , the length of this list changes with the number of generators. Additionally, we would like to learn patterns among monomial ideals as n and D vary! For this reason, every monomial ideal is first converted from its “natural” representation as a list of generators to a fixed number of descriptive *features*.

Many of the features chosen for monomial ideals were inspired by features of SAT problems used in [NDSL04, XHHL08]. For example, those papers considered a *constraint graph* on the clauses, with two clauses connected if the same variable appears in both, and additionally a graph on the variables, with two variables connected if they appear in the a clause together. A slight modification of these definitions works for monomial ideals:

DEFINITION 7.1.1 (Constraint graphs of an ideal). Given a monomial ideal $I \subset S$ with minimal generating set G , define the following.

- (1) The *variable graph* of I has a node for each variable x_i generating S , and an edge between x_i and x_j if $x_i x_j$ divides at least one $g \in G$.
- (2) The *generator graph* of I has a node for each minimal generator $g \in G$, and an edge between g_i and g_j if $\gcd(g_i, g_j) > 1$.

Of course, the constraint graphs vary in size, too, so following [NDSL04] we report exactly four values for each constraint graph: the mean, variance, minimum and maximum values of its vertex degrees.

A total of 23 features, listed in Table 7.2, were chosen, so that every monomial ideal is represented as a vector $x \in \mathbb{R}^{23}$ in the neural network input.

7.2. Algorithm selection for Hilbert series computations

We set out to train a neural network to choose the best pivot rule for each instance of the Hilbert series problem. We considered the nine pivot rules described in Table 7.3, which are a superset of those discussed in [BS92, Big97]. Throughout the table, the term “non-trivial generators” refers to those generators of I that are not pure powers of variables.

For each ideal in the training data sets, we computed the number of base cases for each pivot rule. To account for the effect of randomization, for each rule except **BS** the number of base cases is averaged over five runs.

index	description of feature
0	# Variables
1	# Minimal generators
2—5	Support statistics: mean, variance, minimum, and maximum sizes of support sets of minimal generators.
6—9	Degree statistics: mean, variance, minimum, and maximum total degrees of minimal generators.
10—13	Divisibility statistics: mean, variance, minimum, and maximum number of minimal generators divisible by each variable.
14	# Pure power generators.
15—18	Variable graph statistics: mean, variance, minimum, and maximum degrees of vertices in variable graph (Definition 7.1.1).
19—22	Generator graph statistics: mean, variance, minimum, and maximum degrees of vertices in generator graph (Definition 7.1.1).

TABLE 7.2. Features of monomial ideals used for training data in Sections 7.2 and 7.3.

The number of base cases is not the only measure of algorithm performance. For example empirical timing is another good measure. We chose base cases to ensure “comparing apples to apples,” since pivot rule **M2** is compiled into **Macaulay2**’s C++ engine, with timing advantages over any pivot rules implemented in higher-level **Macaulay2** language. We thank Mike Stillman for raising this issue. A portion of Bigatti’s paper on pivot rules considers the timing advantages of specialized monomial ideal data structures and rules for sorting and interreducing lists of generators [Big97]. This means that minimizing the number of base cases is not necessarily equivalent to minimizing timing.

We discovered that for pivot rule selection, classifying feature vectors requires more nuance than simply labeling each feature vector with the “best” rule for than ideal. The reason is that the “best” rule is not always unique, and even when it is, other rules may

- BS** This is the strategy used in the original Bayer-Stillman paper [BS92], and depends on a different interpretation of Equation 1.4.13. The non-trivial generator least in reverse lexicographic order is taken as \mathcal{P} , and the ideal is written $I + \langle \mathcal{P} \rangle$, i.e. I is the ideal obtained by excluding \mathcal{P} as a generator. Then Eq. 1.4.14 is interpreted as $H(S/(I + \langle \mathcal{P} \rangle), t) = H(S/I, t) - t^{\deg \mathcal{P}} H(S/\langle I : \mathcal{P} \rangle, t)$.
- PV** \mathcal{P} is the “popular variable,” i.e., the x_i that appears in the most non-trivial generators of I .
- M2** \mathcal{P} is the popular variable, raised to the smallest nonzero power with which it appears in a non-trivial generator. This is the pivot rule currently implemented in `Macaulay2`.
- PP1** One of the non-trivial generators divisible by the popular variable is chosen at random, and \mathcal{P} is the popular variable raised to the power with which it appears in this generator.
- PP2** Two of the non-trivial generators divisible by the popular variable are chosen at random, and \mathcal{P} is the popular variable raised to the power with which it appears in their GCD.
- PP3** Same as PP2, but choosing three of these generators.
- GC2** Two of the non-trivial generators divisible by the popular variable are chosen at random, and their GCD, which may have multi-variable support, is used at \mathcal{P} .
- GC3** Same as GC2, but with three of these generators chosen at random.
- GC4** Same as GC2, but with four of these generators chosen at random.

TABLE 7.3. The nine pivot rules considered in the algorithm selection problem for recursive Hilbert series computation. Note that “non-trivial generators” refers to those generators of I that are not pure powers of variables.

BS	PV	M2	PP1	PP2	PP3	GC2	GC3	GC4
1167	321	147	117	113	125	135	131	136
398	55	37	33	40	41	33	37	32
881	162	105	93	97	102	111	100	96
1808	234	136	149	125	134	154	130	134
1893	586	247	203	200	205	251	239	238
14	6	6	6	6	6	6	6	6
16	12	12	12	12	12	12	12	12

TABLE 7.4. Number of base cases for each pivot rule, for five ER-type model random monomial ideals $\mathcal{J} \sim \mathcal{I}(5, 20, 0.0001)$ (top), and two edge ideals of random graphs on ten vertices (bottom). Choices considered **optimal** are bold and in black, while choices considered good are black only.

be close to optimal. To illustrate this, Table 7.4 displays the scores for all pivot rules for five monomial ideals in data set A0, and two from data set B0. For the first ideal, **PP2** is the winning rule with 113 base cases. But if this ideal were in our test set, and the neural network chose rule **PP1**, with 117 base cases, that is significantly different than choosing **BS**, which takes 1167 base cases. Thus we must examine not only how often a wrong choice is made, but *how* wrong it is.

For square-free monomial ideals, which constitute 13 of the 41 random families considered, pivot rules **PV**, **M2**, **PP1**, **PP2**, and **PP3** are always equally good. Sometimes, as with edge ideals of graphs, the GCD of at least two random monomials containing the popular variable will also always equal the popular variable itself, and these rules will also be equivalent. This means that for many of the 30,000 ideals considered for this project, several or even most of the pivot rules are optimal.

This problem was modeled using the classification setup described in Section 1.5.1, with some modifications to account for the many ties and close seconds just described. To assign labels to the training data, ties were broken by overall performance of the pivot rule (see Table 7.5). In other words, if **M2** and **PP2** were both optimal for a particular input, we

assigned the label **PP2** since that strategy had better performance than **M2** on average. In this way our labels were biased toward strategies with good average performance (which seems like a reasonable bias to have.)

After training the neural network with these labels, we evaluate the case-by-case choices of pivot rules for the test ideals using three different measures of performance:

- (1) How often the pivot rule used is optimal; i.e., uses the minimum number of base cases for all pivot rules considered.
- (2) How often the pivot rule used is good; i.e., uses no more than 10% extra base cases, as a percentage of the minimum over all rules.
- (3) On average, what percentage of base cases above the minimum were needed by the rule chosen.

In Table 7.5, the performance of each fixed pivot rule is evaluated according to these three measures. For instance, the **M2** row tells us that `Macaulay2`'s hard-coded pivot rule is optimal for 38% of ideals in the data set, is good for 59% of ideals in the data set, and on average requires 17% more base cases than an optimal choice.

The neural network's case-by-case choice (denoted **DKSZ** for De Loera, Krone, Silverstein, and Zhao) is superior to any fixed choice of pivot rule, by all three measures. It chooses an optimal pivot rule 67% of the time, a good pivot rule 85% of the time, and on averages uses only 4.6% more bases cases than the minimum possible. These values are for the "full" DKSZ classifier, which uses all 23 features in Table 7.2. More impressively, we achieved similar results with a "cheap" DKSZ classifier based on only the five features listed in Table 7.6.

The ability to predict pivot rules based on the features in Table 7.6 alone echoes a general observation from the study of random ideals in Chapters 3 to 5: typical properties of monomial ideals appear to be very predictable from the simple data of ratios among number of variables, number of minimal generators, and degrees of generators. This trend appears to generalize from the specific models used in those chapters, to the more diverse families of ideals considered in the training data for this chapter.

	% optimal	% good	% above optimal
BS	13	15	212
PV	25	39	283
M2	38	59	17
PP1	49	54	20
PP2	43	63	15
PP3	40	63	16
GC2	45	65	10
GC3	53	77	7.8
GC4	59	78	7.4
DKSZ (full)	67	85	4.6
DKSZ (cheap)	64	83	5.6

TABLE 7.5. Suitability of each fixed pivot rule for the entire data set, compared to the performance of our case-by-case choices (DKSZ), using both the full set of features and a cheap set of five features. The DKSZ rules were trained on a random 90% of the data, and tested on the reserved 10%. For each pivot rule, the first column displays how often the strategy is optimal. The second column displays how often the strategy is “good” (no more than 10% from optimal). The third column displays how many extra base cases the strategy needs, on average, as a percentage of the optimal number.

index	description of feature
0	# variables
1	# minimal generators
6–7	mean and variance of degrees of generators
14	# pure power generators.

TABLE 7.6. The five features of monomial ideals used for the “cheap” pivot rule classifier.

Both neural networks (based on the full and cheap feature transforms) most frequently predicted rule **GC4**, the pivot rule with the best average performance across all training data. This was unsurprising. Very surprising, however, was that the second most common prediction of both was **BS**, even though this rule has very poor average performance. It seems that the superiority of DKSZ is largely attributable to its ability to identify instances where **BS** has optimal performance, and choose this pivot rule for *just* those instances. This is a great illustration of the importance of selecting a computer algebra algorithm for each *instance* of a problem rather than for an entire class of problems.

7.3. Fast predictions of algebraic invariants

Now we move to discuss the machine learning prediction of three invariant properties of monomial ideals.

Certain properties of monomial ideals are known to be difficult to compute. Our primary example is the codimension, which is NP-hard [**BS92**]. An alternative to directly computing such information is to use machine learning to build an efficient prediction algorithm to guess the dimension of monomial ideals and with a large proportion correct predictions. Our classifier is a neural network on feature vectors that we train on examples generated from a random model for monomial ideals of bounded degree in a fixed polynomial ring. Our approach is similar to the successful machine learning study of SAT problems [**XHHLB08**]; since SAT and monomial ideal dimension are closely related problems, adapting this strategy is a natural fit. We begin with a description of the training data.

7.3.1. Dimension. We randomly selected 10% of the data to set aside before training, then tested the dimension classifier on these unseen examples. The correct dimension was identified exactly for 82% of the examples (2463 out of 3000). The predicted dimension was within 1 of the correct dimension for 99% of the examples (2969 out of 3000). Figure 7.2a shows the *confusion matrix* for this test set.

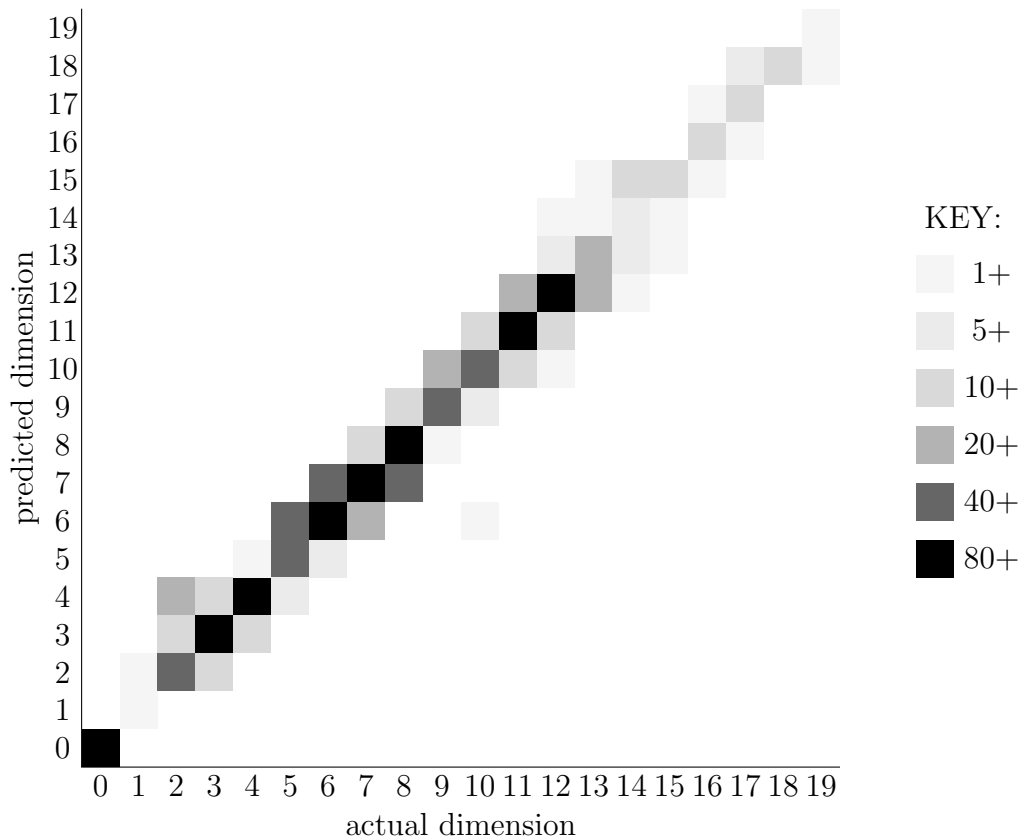
Certain values for dimension were well-represented in our data, while others occurred in fewer of the training examples. A histogram depicting the values observed in the training data can be seen in Figure 7.2c.

Based on Figure 7.2c, we considered four larger classes of ideals: $\{I : \dim I = 0\}$, $\{I : 1 \leq \dim I \leq 4\}$, $\{I : 5 \leq \dim I \leq 9\}$, and $\{I : \dim I \geq 10\}$. Our same learning framework now identifies the right class for 96.6% of a reserved testing set of 3000 ideals. This confusion matrix is pictured in Figure 7.2b. Because the larger size of the figure allows, every nonzero entry of the matrix is labeled with the exact number of observations.

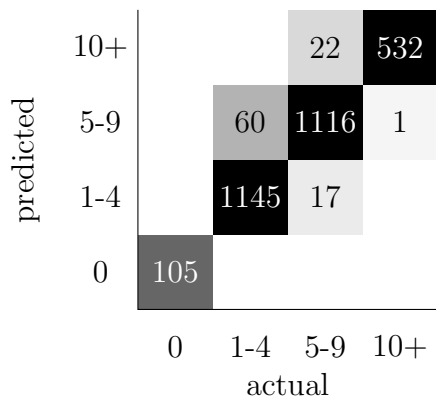
Notice that zero-dimensional ideals are identified perfectly, with no false positives or false negatives in the predictions. Recall that an ideal of $k[x_1, \dots, x_n]$ is zero-dimensional if and only if it contains a generator x_i^j , with $j > 0$, for every $1 \leq i \leq n$; in other words if it contains n pure power generators. Since both n and the number of pure power generators are values in our feature vector, we know that the feature vector theoretically contains the information necessary to exactly predict zero-dimensionality. That the neural network was able to recover this predictive power, with no theoretical understanding, is an interesting proof of concept.

7.3.2. Projective dimension. We performed the same experiments for learning projective dimension, which took values between 0 and 15 in our training data. The lack of examples of higher values is due to our focus on ideals in 5 or 10 variables. A histogram for this invariant is given in Figure 7.3c. As with dimension, we trained neural networks on two classification tasks, based only on the ideal’s feature vector: identifying the exact projective dimension, and identifying the approximate value. For exact classification, the accuracy was 67%, with 98% of the test ideals being no more than 1 away. Figure 7.3a shows this confusion matrix.

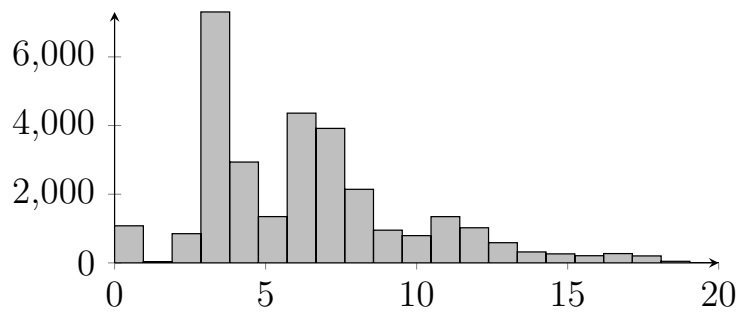
Figure 7.3b is the confusion matrix for identifying ideals as belonging to the sets $\{I : \text{pdim}(S/I) \leq 3\}$, $\{I : 4 \leq \text{pdim}(S/I) \leq 5\}$, $\{I : 6 \leq \text{pdim}(S/I) \leq 8\}$ and $\{I : \text{pdim}(S/I) \geq 9\}$. This accuracy was 88.4%. The neural network actually did *worse* at predicting this “approximate projective dimension” than it did at getting within 1 of the correct value



(A) Confusion matrix for predicting the dimension of monomial ideals, treating each value as a separate class. 99% of predictions were within 1 of correct.



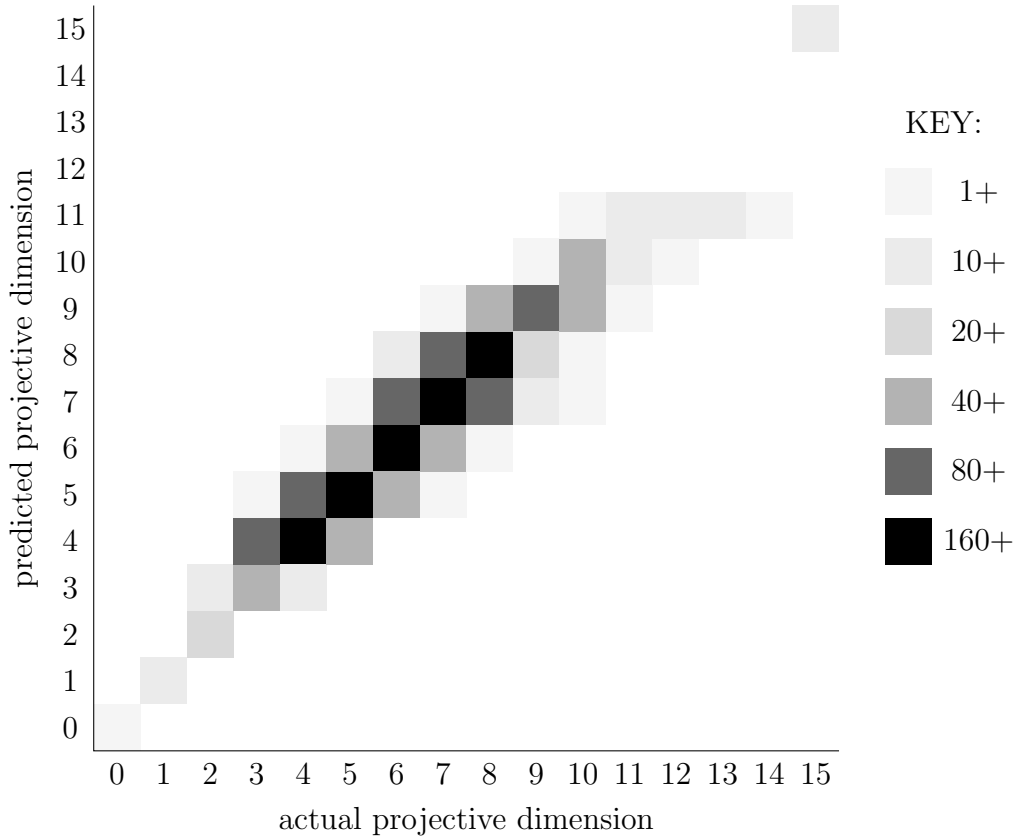
(B) Confusion matrix for predicting dimension, using four classes. Accuracy was 96.6%.



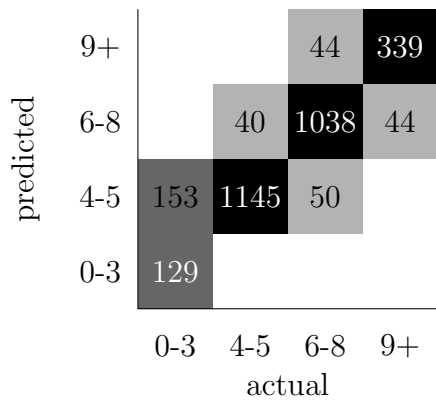
(C) Distribution of actual values of the dimension in training data.

FIGURE 7.2. Learning the dimension of a monomial ideal.

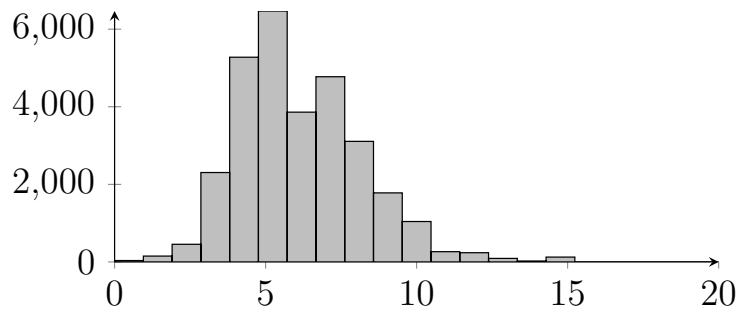
when making exact predictions. Even though fewer classes seems like an “easier” problem



(A) Confusion matrix for projective dimension. 98% of predictions were within 1 of correct.



(B) Confusion matrix for the four-class projective dimension problem. Accuracy was 88.4%.



(C) Distribution of projective dimension in training data.

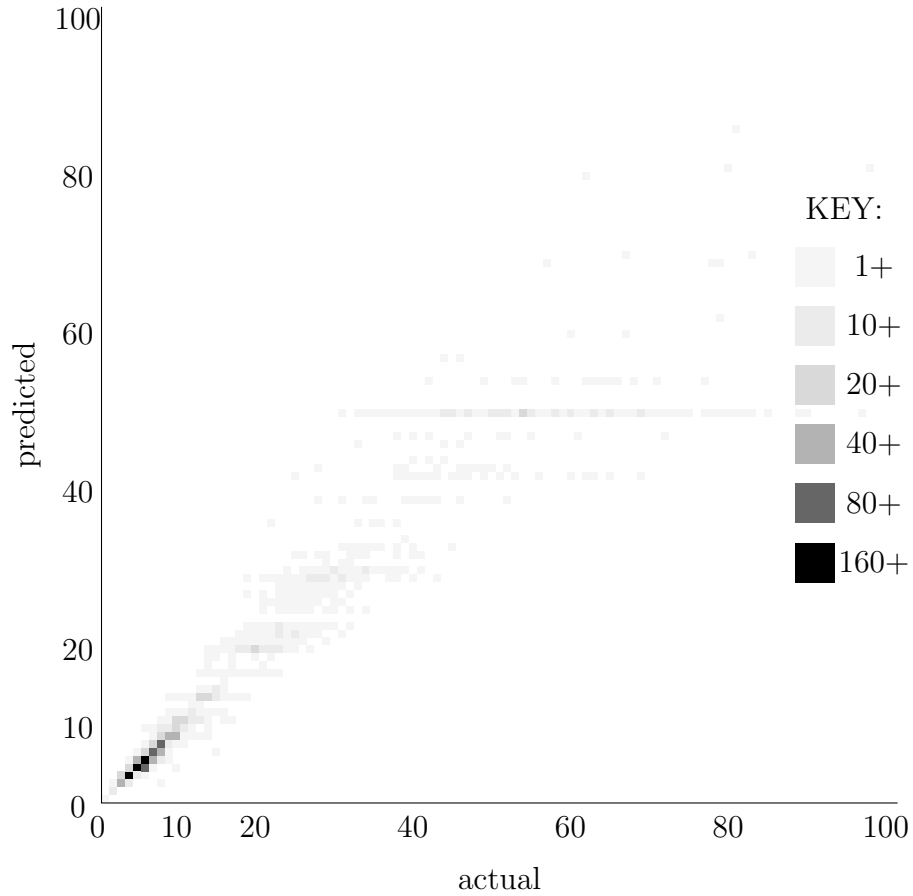
FIGURE 7.3. Learning the projective dimension of a monomial ideal.

for a neural network, the finer classification is in fact the better approach to approximating projective dimension.

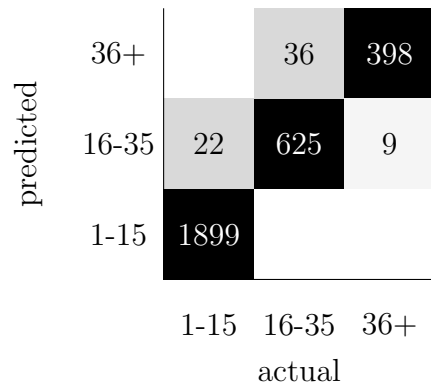
7.3.3. Regularity. For regularity, we observed more than 100 unique values over the 30,000 ideals in the data set. A few small values—integers from 3 to 10, inclusive—make up the bulk of examples, with over 1,000 instances each. The histogram for the distribution of regularity, Figure 7.4c, suggests a clustering of *low*-, *medium*-, and *high-regularity* ideals defined by $\{I : \text{reg}(S/I) \leq 15\}$, $\{I : 16 \leq \text{reg}(S/I) \leq 35\}$, and $\{I : \text{reg}(S/I) \geq 36\}$, respectively. Note that technically $\text{reg}(S/\langle 0 \rangle) = -\infty$, following the convention that the degree of the zero polynomial is $-\infty$. For computational purposes we took $\text{reg}(S/\langle 0 \rangle) = 0$.

Figure 7.4a illustrates the neural network’s accuracy at predicting the exact regularity of an ideal from its features. With over 100 classes to choose from, the accuracy is unsurprisingly much lower than for the other invariants, with 42% of predictions correct. In fact, 42% is quite impressive given the difficulty of the problem and the large number of classes. The confusion matrix has a linear trend despite many off-diagonal entries, and 90% of the predictions were within 5 of correct. The linear trend worsens dramatically around regularity 35, where the number of training examples per class drops in Figure 7.4c. For classes with many examples, the accuracy was much better.

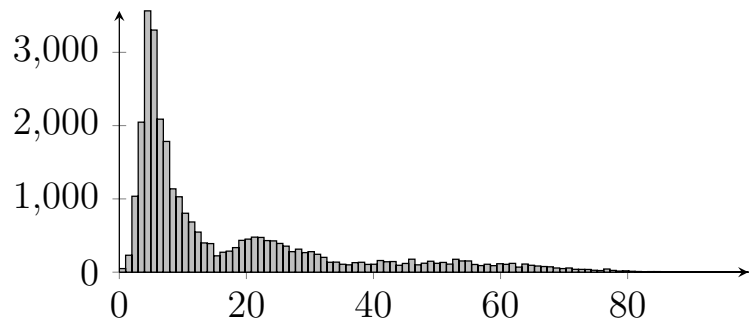
Classifying ideals as low-, medium-, and high-regularity was very successful. Ideals were correctly sorted into these three classes with 97.4% accuracy, and only 9 out of 3,000 test ideals were incorrectly sorted into a lower-regularity class.



(A) Confusion matrix for regularity, considering each value as a separate class. 42% of predictions were exact, while 90% were within 5.



(B) Confusion matrix for the three-class regularity problem. Accuracy was 97.4%.



(C) Distribution of regularity in training data.

FIGURE 7.4. Learning the regularity of a monomial ideal.

APPENDIX A

Computational details

Most of the algebra computations for this dissertation were done in `Macaulay2` [GS], including in particular the `EdgeIdeals` package [FHV09], `Graphs` package [BCJ⁺], `LexIdeals` package [Fra], and `RandomMonomialIdeals` package [PSW17]. The previous sentence was created with the help of the `PackageCitations` package [Dal]. For a brief moment, the idea of creating a package for automatically citing `PackageCitations` whenever `PackageCitations` is used to cite packages was entertained, but the author decided that a self-referential joke about “the idea of creating a package for automatically citing `PackageCitations` whenever `PackageCitations` is used to cite packages” would suffice.

A.1. Code for discrete optimization techniques in commutative algebra

As described in Section 3.4, integer programming computations used the SCIP Optimization Suite [G⁺18], including the ZIMPL modeling language [Koc04]. Most of the optimization problems were systematically generated from within `Macaulay2`. For example, the list of Betti tables in Section 6.4 was created by defining a polynomial ring and Hilbert function in `Macaulay2`, generating a description of the corresponding De Loera polytope in ZIMPL format, and running a SCIP command to write all integer lattice points to a file. From `Macaulay2` each line of the file was parsed as a 0-1 vector and the corresponding monomial ideal determined. Finally the Betti table of each ideal was computed and the tally of their frequencies returned.

To implement this and similar computations featured in the dissertation, which exist at the interface of commutative algebra and discrete optimization, the author is releasing a `Macaulay2` package called `MonomialIntegerPrograms`. This software is also available at

<https://github.com/lilysilverstein/MonIP>.

In addition to enumerative problems related to Hilbert functions and Betti numbers, the package provides the fast method of computing dimension from Section 3.4, and related ILP-based functions for computing degree and minimal primes. Other discrete optimization algorithms for commutative algebra are in development.

A.2. Code and data for machine learning on monomial ideals

The artificial neural network project described in Chapter 7 was implemented in TensorFlow [Aea15] via the Keras Python API [C⁺15]. Warm thanks and praise to Zekai Zhao for his work on this code.

All code and data related to this portion of the dissertation is available at the following repository:

<https://github.com/lilysilverstein/MonLearning/>.

This includes:

- *MonLearning.py*: Keras/TensorFlow code defining the neural network architecture and implementing training and testing routines.
- *TrainingData/*: A directory containing the complete training data, organized into subdirectories corresponding to the families of monomial ideals listed in Table 7.1.
- *HilbertML.m2*: Macaulay2 code implementing flexible pivot rule choices for Hilbert series computations, counting base cases, etc.
- *IdealML.m2*: Miscellaneous Macaulay2 code for generating random monomial ideals of various kinds, computing invariants, file I/O and organization for training data, etc.

Bibliography

- [Aea15] M. Abadi and et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from [tensorflow.org](https://www.tensorflow.org).
- [Ale17a] G. Alesandroni. Minimal resolutions of dominant and semidominant ideals. *J. Pure Appl. Algebra*, pages 780–798, 2017.
- [Ale17b] G. Alesandroni. Monomial ideals with large projective dimension. *arXiv preprint arXiv:1710.05124*, 2017.
- [AMMIL12] Y. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin. *Learning From Data*. AMLBook, 2012.
- [AS92] N. Alon and J. Spencer. *The probabilistic method*. Wiley-Interscience series in discrete mathematics and optimization. J. Wiley & sons, New York, Chichester, Brisbane, 1992.
- [BC11] P. Bürgisser and F. Cucker. On a problem posed by Steve Smale. *Annals of Mathematics*, 174(3):1785–1836, 2011.
- [BCJ⁺] J. Burkart, D. Cook, C. Jansen, A. Taylor, and A. O’Keefe. Graphs: Package for processing graphs and directed graphs (digraphs). Version 0.3.1. A *Macaulay2* package available at <https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages>.
- [BE77] D. Buchsbaum and D. Eisenbud. Algebra structures for finite free resolutions, and some structure theorems for ideals of codimension 3. *Amer. J. Math.*, 99(3):447–485, 1977.
- [BHK11] E. Babson, C. Hoffman, and M. Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, January 2011.
- [Big93] A. Bigatti. Upper bounds for the Betti numbers of a given Hilbert function. *Comm. Algebra*, 21(7):2317–2334, 1993.
- [Big97] A. Bigatti. Computation of Hilbert-Poincaré series. *Journal of Pure and Applied Algebra*, 119(3):237–253, 1997.
- [Bil95] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley, 1995.
- [BK18] O. Bobrowski and M. Kahle. Topology of random geometric complexes: a survey. *Journal of Applied and Computational Topology*, 1:331–364, june 2018.
- [BM92] D. Bayer and D. Mumford. What can be computed in algebraic geometry? In *Computational algebraic geometry and commutative algebra*, pages 1–48. University Press, April 1992.

- [BM05] I. Bárány and J. Matoušek. The randomized integer convex hull. *Discrete & Computational Geometry*, 33(1):3–25, 2005.
- [BNVW17] M.-F. Balcan, V. Nagarajan, E. Vitercik, and C. White. Learning-theoretic foundations of algorithm configuration for combinatorial partitioning problems. *Proceedings of Machine Learning Research*, 65:1–62, 2017.
- [Bol01] B. Bollobás. *Random graphs*. Cambridge University Press, 2nd edition, 2001.
- [BP09] C. Beltrán and L. Pardo. Smale’s 17th problem: average polynomial time to compute affine and projective solutions. *Journal of the AMS*, 22(2):363–385, 2009.
- [BPS98] D. Bayer, I. Peeva, and B. Sturmfels. Monomial resolutions. *Math. Res. Lett.*, 5(1–2):31–46, 1998.
- [BR04] M. Brun and T. Römer. Betti numbers of \mathbb{Z}^n -graded modules. *Communications in Algebra*, 32(12):4589–4599, 2004.
- [BS87a] D. Bayer and M. Stillman. A criterion for detecting m-regularity. *Inventiones mathematicae*, 87:1–12, 1987.
- [BS87b] D. Bayer and M. Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Math. J.*, 55(2):321–328, 06 1987.
- [BS88] D. Bayer and M. Stillman. On the complexity of computing syzygies. *Journal of Symbolic Computation*, 6(2):135 – 147, 1988.
- [BS92] D. Bayer and M. Stillman. Computation of Hilbert functions. *J. Symbolic Computation*, 14:31–50, 1992.
- [BS18] A. Boocher and J. Ser. Lower bounds for Betti numbers of monomial ideals. *Journal of Algebra*, 508:445–460, 2018.
- [BWK93] T. Becker, V. Weispfenning, and H. Kredel. *Gröbner bases: a computational approach to commutative algebra*. Graduate texts in mathematics. Springer-Verlag, 1993.
- [C+15] François Chollet et al. Keras, 2015.
- [CF16] A. Costa and M. Farber. Random simplicial complexes. In *Configuration spaces - Geometry, Topology, and Representation Theory*, INdAM Series., pages 129–153. Springer, 2016.
- [Cha91] H. Charalambous. Betti numbers of multigraded modules. *Journal of Algebra*, 137(2):491–500, 1991.
- [CLO05] D.A. Cox, J. Little, and D. O’Shea. *Using Algebraic Geometry*. Graduate Texts in Mathematics. Springer New York, 2005.
- [CLO07] D. Cox, J.. Little, and D. O’Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2007.

- [Coo71] S. Cook. The complexity of theorem-proving procedures. In *Proc 3rd FOCS*, pages 151–158, 01 1971.
- [Cor01] G. Cornuéjols. *Combinatorial Optimization: Packing and Covering*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, 2001.
- [CP17] D. Cifuentes and P. Parrilo. Chordal networks of polynomial ideals. *SIAM J. Appl. Algebra Geometry*, 1:73–110, 2017.
- [Dal] A. Dall. PackageCitations: citation of *Macaulay2* packages. Version 0.1. A *Macaulay2* package available at <https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages>.
- [DHKS19] J. De Loera, S. Hoşten, R. Krone, and L. Silverstein. Average behavior of minimal free resolutions of monomial ideals. *Proceedings of the Amer. Math. Soc.*, April 2019.
- [DKZ⁺17] H. Dai, E. Khalil, Y. Zhang, B. Dilkina, and L. Song. Learning combinatorial optimization algorithms over graphs. In *Neural Information Processing Systems*, 2017.
- [DPS⁺19] J. De Loera, S. Petrović, L. Silverstein, D. Stasi, and D. Wilburne. Random monomial ideals. *Journal of Algebra*, 519:440 – 473, 2019.
- [EEL15] L. Ein, D. Erman, and R. Lazarsfeld. Asymptotics of random Betti tables. *J. Reine Angew. Math.*, 702:55–75, 2015.
- [EEL16] L. Ein, D. Erman, and R. Lazarsfeld. A quick proof of nonvanishing for asymptotic syzygies. *Algebraic Geometry*, 3:211–222, 2016.
- [Eis95] D. Eisenbud. *Commutative algebra: with a view toward algebraic geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
- [Eis06] D. Eisenbud. *The Geometry of Syzygies: A Second Course in Algebraic Geometry and Commutative Algebra*. Graduate Texts in Mathematics. Springer New York, 2006.
- [EL12] L. Ein and R. Lazarsfeld. Asymptotic syzygies of algebraic varieties. *Inventiones mathematicae*, 190(3):603–646, 2012.
- [ER59] P. Erdős and A. Rényi. On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.
- [ES09] D. Eisenbud and F. Schreyer. Betti numbers of graded modules and cohomology of vector bundles. *J. Amer. Math. Soc.*, 22(3):859–888, 2009.
- [EY17] D. Erman and J. Yang. Random flag complexes and asymptotic syzygies. *arXiv preprint arXiv:1706.01488*, 2017.
- [FHV09] C. Francisco, A. Hoefel, and A. Van Tuyl. EdgeIdeals: A package for (hyper)graphs. *Journal of Software for Algebra and Geometry*, 1:1–4, 2009.

- [Fra] C. Francisco. LexIdeals: A *Macaulay2* package for manipulating lexicographic-type monomial ideals. . available at <https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages>.
- [G⁺18] A. Gleixner et al. The SCIP Optimization Suite 6.0. Technical report, Optimization Online, July 2018.
- [Gil59] E. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [GJ79] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1979.
- [GR15] R. Gupta and T. Roughgarden. A PAC approach to application-specific algorithm selection. *ArXiv e-prints*, November 2015.
- [GS] D. Grayson and M. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at faculty.math.illinois.edu/Macaulay2/.
- [GS93] P. Gritzmann and B. Sturmfels. Minkowski addition of polytopes: Computational complexity and applications to Gröbner bases. *SIAM J. Disc. Math*, 6(2):246–269, May 1993.
- [GV11] I. Gitler and R. Villarreal. *Graphs, rings and polyhedra*, volume 35 of *Aportaciones Matemáticas: Textos [Mathematical Contributions: Texts]*. Sociedad Matemática Mexicana, México; Instituto de Matemáticas, UNAM, México, 2011.
- [Hà14] H. Hà. Regularity of squarefree monomial ideals. In *Connections between algebra, combinatorics, and geometry*, pages 251–276. Springer, 2014.
- [Har79] R. Hartshorne. Algebraic vector bundles on projective spaces: a problem list. *Topology*, 18(2):117–128, 1979.
- [HH11] J. Herzog and T. Hibi. *Monomial Ideals*, volume 260 of *Graduate Texts in Mathematics*. Springer, 2011.
- [Hul93] H. Hulett. Maximum Betti numbers of homogeneous ideals with a given Hilbert function. *Comm. Algebra*, 21(7):2335–2350, 1993.
- [Kac43] M. Kac. On the average number of real roots of a random algebraic equation. *Bulletin of the American Mathematical Society*, 49:314–320, 1943.
- [Kah07] M. Kahle. The neighborhood complex of a random graph. *J. Combin. Theory Ser. A*, 114(2):380–387, 2007.
- [Kah09] M. Kahle. Topology of random clique complexes. *Discrete mathematics*, 309(6):1658–1671, April 2009.
- [Kar72] R. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer US, Boston, MA, 1972.

- [KDN⁺17] E. Khalil, B. Dilkina, G. Nemhauser, S. Ahmed, and Y. Shao. Learning to run heuristics in tree search. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI'17*, pages 659–666. AAAI Press, 2017.
- [Koc04] T. Koch. *Rapid Mathematical Prototyping*. PhD thesis, Technische Universität Berlin, 2004.
- [Kou76] A. Kouchnirenko. Polyèdres de Newton et nombres de Milnor. *Inventiones mathematicae*, 32:1–32, 1976.
- [Lan11] J. Landsberg. *Tensors: Geometry and Applications*. Graduate studies in mathematics. American Mathematical Society, 2011.
- [LM06] N. Linial and R. Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, August 2006.
- [LO38] J. Littlewood and A. Offord. On the number of real roots of a random algebraic equation. *Journal of the London Mathematical Society*, S1-13(4):288, 1938.
- [Mac27] F. Macaulay. Some properties of enumeration in the theory of modular systems. *Proceedings of the London Mathematical Society*, 2(1):531–555, 1927.
- [MALW14] A. Marcos Alvarez, Q. Louveaux, and L. Wehenkel. A supervised machine learning approach to variable branching in branch-and-bound. In *IN ECML*, 2014.
- [Mar08] S. Margulies. *Computer Algebra, Combinatorics, and Complexity: Hilbert’s Nullstellensatz and NP-complete Problems*. PhD thesis, University of California, Davis, 2008.
- [MM82] E. Mayr and A. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. Math.*, 46:305–329, 1982.
- [MS04] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227 of *Graduate Texts in Mathematics*. Springer New York, 2004.
- [MW09] R. Meshulam and N. Wallach. Homological connectivity of random k-dimensional complexes. *Random Structures & Algorithms*, 34(3):408–417, 2009.
- [NDSL04] E. Nudelman, A. Devkar, Y. Shoham, and K. Leyton-Brown. Understanding random SAT: Beyond the clauses-to-variables ratio. In *Tenth International Conference on Principles and Practice of Constraint Programming (CP'04)*, pages 438–452, 2004.
- [OS99] S. Onn and B. Sturmfels. Cutting corners. *Advances in applied mathematics*, 23(1):29–48, 1999.
- [Par96] K. Pardue. Deformation classes of graded modules and maximal Betti numbers. *Illinois J. Math.*, 40(4):564–585, 1996.
- [Pit97] B. Pittel. On a likely shape of the random Ferrer’s diagram. *Advances in Applied Mathematics*, 18(4):432 – 488, 1997.

- [PS98] C. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Dover Books on Computer Science. Dover Publications, 1998.
- [PSW17] S. Petrović, D. Stasi, and D. Wilburne. Random Monomial Ideals Macaulay2 Package. *ArXiv e-prints*, November 2017.
- [Qui76] D. Quillen. Projective modules over polynomial rings. *Invent. Math.*, 36:167–171, 1976.
- [Rot02] J. Rotman. *Advanced Modern Algebra*. Prentice Hall, 2002.
- [Sta07] R. Stanley. *Combinatorics and Commutative Algebra*. Progress in Mathematics. Birkhäuser Boston, 2007.
- [Str19] G. Strang. *Linear Algebra and Learning from Data*. Wellesley-Cambridge Press, January 2019.
- [Stu98] B. Sturmfels. Polynomial equations and convex polytopes. *The American Mathematical Monthly*, 105(10):907–922, 1998.
- [Sus76] A. Suslin. Projective modules over polynomial rings are free. *Dokl. Akad. Nauk SSSR*, 229(5):1063–1066, 1976.
- [TMDQ16] C. Tan, S. Ma, Y.-H. Dai, and Y. Qian. Barzilai-borwein step size for stochastic gradient descent. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 685–693. Curran Associates, Inc., 2016.
- [Wal17] M. Walker. Total Betti numbers of modules of finite projective dimension. *Ann. of Math. (2)*, 186(2):641–646, 2017.
- [Wil18] D. Wilburne. *Quantifying Uncertainty in Random Algebraic Objects Using Discrete Methods*. PhD thesis, Illinois Institute of Technology, 2018.
- [XHHLB08] L. Xu, F. Hutter, H. Hoos, and K. Leyton-Brown. SATzilla: portfolio-based algorithm selection for SAT. *Journal of artificial intelligence research*, 32:565–606, 2008.