

# Asymptotics of random Betti tables

By *Lawrence Ein* at Chicago, *Daniel Erman* at Ann Arbor and  
*Robert Lazarsfeld* at Ann Arbor

**Abstract.** This paper is motivated by the question of understanding the asymptotic behavior of the Betti numbers of the resolution of the ideal of a projective variety as the positivity of the embedding line bundle grows. We present a conjecture asserting that these invariants approach a Gaussian distribution, and we verify this in the case of curves. Then we work out the asymptotics of “random” Betti tables with a fixed number of rows, sampled according to a uniform choice of Boij–Söderberg coefficients. This analysis suggests that the normal distribution of Betti numbers is in any event the typical behavior from a probabilistic viewpoint.

## Introduction

The purpose of this paper is twofold. First, we present a conjecture to the effect that the ranks of the syzygy modules of a smooth projective variety become normally distributed as the positivity of the embedding line bundle grows. Then, in an attempt to render the conjecture plausible, we prove a result suggesting that this is in any event the typical behavior from a probabilistic point of view. Specifically, we consider a “random” Betti table with a fixed number of rows, sampled according to a uniform choice of Boij–Söderberg coefficients. We compute the asymptotics of the entries as the length of the table goes to infinity, and show that they become normally distributed with high probability.

Turning to details, we start by discussing at some length the geometric questions underlying the present work. Let  $X$  be a smooth projective variety of dimension  $n$  defined over some field  $k$ , and for  $d > 0$  put

$$L_d = dA + P,$$

where  $A$  is a fixed ample divisor and  $P$  is an arbitrary divisor on  $X$ . We assume that  $d$  is sufficiently large so that  $L_d$  defines a normally generated embedding

$$X \subseteq \mathbf{P}^{r_d}, \quad \text{where } r_d = r_d^{\text{def}} = h^0(X, L_d) - 1 \in \Theta(d^n).$$

(A function  $f(d)$  is in  $\Theta(g(d))$  if there are constants  $c_1, c_2 > 0$  with  $c_1 g(d) \leq f(d) \leq c_2 g(d)$  for all  $d \gg 0$ .)

---

Research of the first author partially supported by NSF grant DMS-1001336. Research of the second author partially supported by the Simons Foundation. Research of the third author partially supported by NSF grant DMS-1159553.

Write  $S = \text{Sym } H^0(X, L_d)$  for the homogeneous coordinate ring of  $\mathbf{P}^{r_d}$ , denote by  $J = J_X \subseteq S$  the homogeneous ideal of  $X$ , and consider the minimal graded free resolution  $E_\bullet = E_\bullet(X; L_d)$  of  $J$ :

$$0 \longleftarrow J_X \longleftarrow \bigoplus S(-a_{1,j}) \longleftarrow \bigoplus S(-a_{2,j}) \longleftarrow \cdots \longleftarrow \bigoplus S(-a_{r,j}) \longleftarrow 0.$$

$$\begin{array}{ccccccc} & & \parallel & & \parallel & & \parallel \\ & & E_1 & & E_2 & & E_r \end{array}$$

As customary, it is convenient to write

$$E_p(X; L_d) = \bigoplus_q K_{p,q}(X; L_d) \otimes_k S(-p-q).$$

Thus  $K_{p,q}(X; L_d)$  is a finite-dimensional vector space, whose non-zero elements are generators in degree  $p+q$  of the  $p^{\text{th}}$  module of syzygies of  $S/J_X$ . We refer to such an element as a  $p^{\text{th}}$  syzygy of weight  $q$ . The dimensions

$$k_{p,q}(X; L_d) \stackrel{\text{def}}{=} \dim K_{p,q}(X; L_d)$$

are the *Betti numbers* of  $L_d$ ; they are the entries of the *Betti table* of  $L_d$ . The basic problem motivating the present paper (one that alas we do not solve) is to understand the asymptotic behavior of these numbers as  $d \rightarrow \infty$ .

Elementary considerations of Castelnuovo–Mumford regularity show that if  $d \gg 0$ , then

$$K_{p,q}(X; L_d) = 0 \quad \text{for } q > n + 1.$$

Furthermore  $K_{p,n+1}(X; L_d) \neq 0$  if and only if

$$r_d - n - (p_g - 1) \leq p \leq r_d - n,$$

where  $p_g = h^0(X, \omega_X)$ . So the essential point is to understand the groups  $K_{p,q}(X; L_d)$  for  $1 \leq q \leq n$  and  $p \in [1, r_d]$ . The main result of [6] is that as  $d \rightarrow \infty$  these groups become non-zero for “essentially all” values of the parameters. Specifically, there exist constants  $C_1, C_2 > 0$  (depending on  $X$  and the choice of the divisors  $A, P$  appearing in the definition of  $L_d$ ) such that if one fixes  $1 \leq q \leq n$ , then

$$K_{p,q}(X; L_d) \neq 0$$

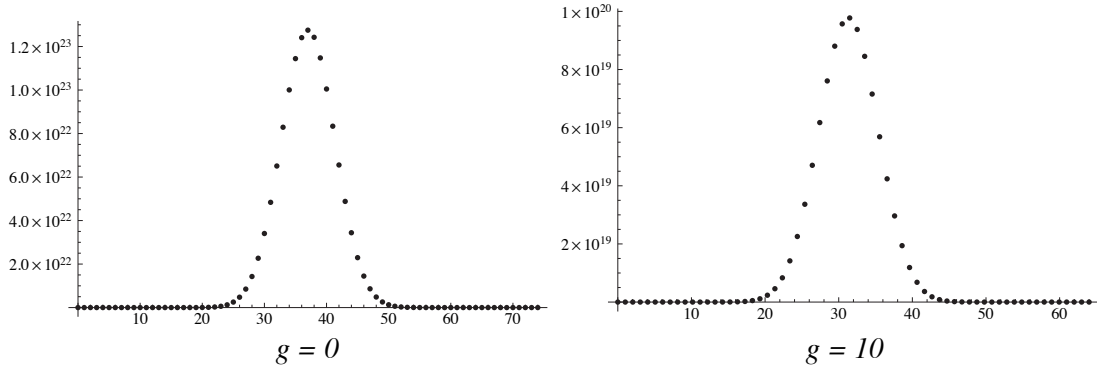
for every value of  $p$  satisfying

$$C_1 \cdot d^{q-1} \leq p \leq r_d - C_2 \cdot d^{n-1}.$$

However the results appearing in [6] do not give quantitative information about the asymptotics in  $p$  of the corresponding Betti numbers  $k_{p,q}(X; L_d)$  for fixed weight  $q \in [1, n]$  and  $d \gg 0$ .

The question is already interesting in the case  $n = 1$  of curves: here it is only the  $k_{p,1}$  that come asymptotically into play.<sup>1)</sup> Figure 1 shows plots of these Betti numbers for a divisor of degree 75 on a curve of genus 0, and on a curve of genus 10. The figure suggests that the  $k_{p,1}$  become normally distributed, and we prove that this is indeed the case.

<sup>1)</sup> If  $X$  is a curve of genus  $g > 2$  and  $L_d$  is a divisor of degree  $d \gg 0$  on  $X$ , then the Betti numbers  $k_{p,1}(X; L_d)$  depend on the geometry of  $X$  and  $L_d$  – in a manner that is not completely understood – when  $r_d - g \leq p \leq r_d - 1$ . However elementary estimates show that when  $d$  is large the invariants in question are exponentially smaller than the  $k_{p,1}$  for  $p \approx \frac{r_d}{2}$ , and so they do not show up in the asymptotic picture. See Remark 2.4.

Figure 1. Betti numbers  $k_{p,1}$  on curves of genus 0 and 10.

**Proposition A.** Let  $L_d$  be a divisor of large degree  $d$  on a smooth projective curve  $X$  of genus  $g$ , so that  $r_d = d - g$ . Choose a sequence  $\{p_d\}$  of integers such that

$$p_d \rightarrow \frac{r_d}{2} + a \cdot \frac{\sqrt{r_d}}{2}$$

for some fixed number  $a$  (i.e.  $\lim_{d \rightarrow \infty} \frac{2p_d - r_d}{\sqrt{r_d}} = a$ ). Then as  $d \rightarrow \infty$ ,

$$\left( \frac{1}{2^{r_d}} \cdot \sqrt{\frac{2\pi}{r_d}} \right) \cdot k_{p_d,1}(X; L_d) \rightarrow e^{-a^2/2}.$$

At the risk of some recklessness, we conjecture that the picture seen in dimension one holds universally:

**Conjecture B.** Let  $X$  be a smooth projective variety of dimension  $n$ , and fix a weight  $1 \leq q \leq n$ . Then there is a normalizing function  $F_q(d)$  (depending on  $X$  and geometric data) such that

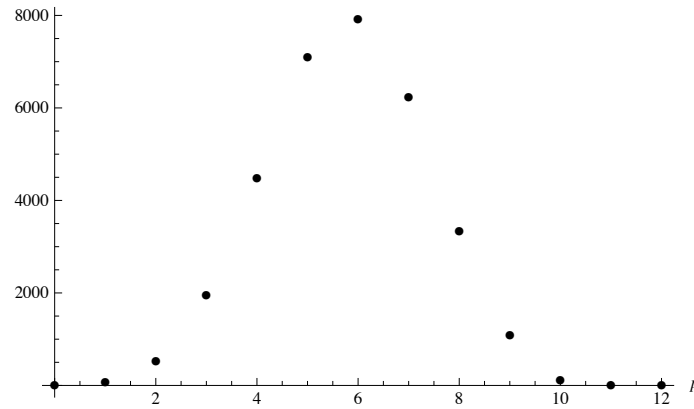
$$F_q(d) \cdot k_{p_d,q}(X; L_d) \rightarrow e^{-a^2/2}$$

as  $d \rightarrow \infty$  and  $p_d \rightarrow \frac{r_d}{2} + a \cdot \frac{\sqrt{r_d}}{2}$ .

In other words, the prediction is that as  $d \rightarrow \infty$  one gets the same sort of normal distribution of the Betti numbers  $k_{p,q}(X; L_d)$ , considered as functions of  $p$ , as that which occurs in the case of curves. Put another way, the conjecture asserts that the rows of the Betti table of any very positive embedding display roughly the pattern that one would see in a large Koszul complex.

As an illustration, we plot in Figure 2 the Betti numbers  $k_{p,1}$  for the embedding  $\mathbf{P}^2 \subset \mathbf{P}^{14}$  defined by  $\mathcal{O}_{\mathbf{P}^2}(4)$ , which is the largest example we were able to run on Macaulay2. We hope that the reader will agree that these data at least seem consistent with the conjectured picture. We remark that in general the methods of [6] yield upper and lower bounds for  $k_{p,q}(X; L_d)$  having the shape predicted by the conjecture, but unfortunately the normalizing functions do not match up.

Concerning the conjecture, the first point to stress is that we do not know how to verify it for any variety of dimension  $n \geq 2$ . For example, it already seems a very interesting and

Figure 2. Betti numbers  $k_{p,1}$  of 4-fold Veronese embedding of  $\mathbf{P}^2$ .

challenging problem to compute the asymptotics of the Betti numbers of the  $d$ -fold Veronese embeddings of  $\mathbf{P}^2$ . In view of this state of affairs, it then becomes natural to ask whether one can give any indirect evidence supporting the conjecture. For example, if one considers resolutions with syzygies having fixed weights  $1 \leq q \leq n$  and lets the length of the resolution grow, are the asymptotics predicted by the conjecture “typical” in some sense? Our main purpose here is to prove a result suggesting that this is indeed the case: with high probability, the entries in the rows of a “random” Betti table become normally distributed as the length of the table goes to infinity.

In order to make this precise we start by introducing a model for the resolutions in question; we will then apply the Boij–Söderberg theory established by Eisenbud and Schreyer [10] to construct a sample space for studying random Betti tables. Specifically, fix a natural number  $n \geq 2$  and let  $R = k[x_1, \dots, x_{r-n}]$  be a polynomial ring in  $r - n$  variables, where  $r$  is a large integer that will later go to infinity. Consider now a finite-length graded  $R$ -module  $M$  with the property that

$$(0.1) \quad K_{p,q}(M) = 0 \quad \text{for } q \notin [1, n]$$

and for every  $0 \leq p \leq n - r$ . Note that if  $X$  carries a line bundle  $B$  such that  $H^i(X, B) = 0$  for all  $i$  – for instance  $B = \mathcal{O}_{\mathbf{P}^n}(-1)$  on  $X = \mathbf{P}^n$  – then the  $S$ -module associated to  $B$  will be a Cohen–Macaulay module, and hence modules as in (0.1) can be constructed by modding out by  $n + 1$  general linear forms. We thus think of  $M$  as having a resolution that models in a simplified manner the resolutions that occur for very positive embeddings of smooth projective varieties of dimension  $n$ .

The theory of Eisenbud–Schreyer asserts that the Betti numbers  $k_{p,q}(M)$  are determined by the Betti numbers of modules having a *pure resolution*. By definition, this is a module  $\Pi$  with the property that the  $p^{\text{th}}$  syzygies of  $\Pi$  all occur in a single weight  $q = q(p)$ . A module with a pure resolution satisfying (0.1) determines an  $(n - 1)$ -element subset

$$I \subseteq [r] \stackrel{\text{def}}{=} \{1, \dots, r\}$$

encoding the values of  $p$  at which the weight  $q(p)$  jumps: we denote by  $\binom{[r]}{n-1}$  the collection of all such  $I$ . The corresponding module  $\Pi_I$  is not unique, but its Betti numbers  $k_{p,q}(\Pi_I)$  are determined up to a scalar that is normalized by fixing the multiplicity of each  $\Pi_I$ . The main

result of [10] implies that given  $M$  as above, there exist non-negative rational numbers  $x_I \in \mathbb{Q}$  such that

$$(0.2) \quad k_{p,q}(M) = \sum_{I \in \binom{[r]}{n-1}} x_I \cdot k_{p,q}(\Pi_I)$$

for all  $p \in [0, r-n]$  and  $q \in [1, n]$ .<sup>2)</sup> In other words, the resolution of  $M$  is numerically a  $\mathbb{Q}$ -linear combination of the resolutions of the  $\Pi_I$ . Conversely, after possibly scaling one can find a module  $M$  that realizes any given non-negative rational numbers  $x_I$ . Thus up to scaling, the possible numerical types of  $n$ -weight resolutions are parameterized by vectors

$$x = \{x_I\}_{I \in \binom{[r]}{n-1}}$$

of Boij–Söderberg coefficients  $x_I \in \mathbb{Q}_{\geq 0}$ .<sup>3)</sup> In order to emphasize the dependence on  $r$ , which will shortly become important, we will henceforth write  $x_r = \{x_{r,I}\}$  to denote these coefficients and  $\Pi_{r,I}$  to denote the corresponding module.

We assume now (by scaling) that each  $x_{r,I} \leq 1$ , and since we are interested in numerical questions, we allow the  $x_{r,I}$  to be real. Denote by

$$\Omega_r = \Omega_{r,n} \stackrel{\text{def}}{=} [0, 1]^{\binom{r}{n-1}}$$

the cube parameterizing the resulting coefficient vectors  $x_r = \{x_{r,I}\}$ . Then given

$$x_r = \{x_{r,I}\} \in \Omega_r,$$

set

$$k_{p,q}(x_r) = \sum_{I \in \binom{[r]}{n-1}} x_{r,I} \cdot k_{p,q}(\Pi_{r,I}).$$

Thus the  $k_{p,q}$  are functions on  $\Omega_r$  computing the Betti numbers of a module described by a Boij–Söderberg coefficient vector  $x_r$ .<sup>4)</sup>

We next imagine choosing  $x_r \in \Omega_r$  uniformly at random. The resulting real numbers  $k_{p,q}(x_r)$  can then be thought of as the entries of a random (and hence “typical”) Betti table with  $n$  rows and  $r+1-n$  columns. This is illustrated when  $n=2$  in Figure 3, which shows plots of  $k_{p,1}(x_r)$  for random vectors  $x_r \in \Omega_r$  with  $r=14$  and  $r=60$ . Fixing  $q \in [1, n]$ , our main result describes the asymptotics in  $p$  of the numbers  $k_{p,q}(x_r)$  for such a randomly chosen coefficient vector  $x_r$  as  $r$  gets very large. It implies in particular that when  $r \rightarrow \infty$ , the Betti numbers  $k_{p,q}(x_r)$  become normally distributed with high probability.

We start with a somewhat informal statement.

<sup>2)</sup> The set of indices  $I$  that arise here form in a natural way the vertices of a simplicial complex, and if one takes into account the simplicial structure, one can arrange that the expression in (0.2) is unique. However for reasons that we will discuss in Remark 1.3, we prefer to allow the indices  $I$  to vary independently. Therefore when  $n \geq 3$ , the coefficients  $x_I$  are not uniquely determined by  $M$ .

<sup>3)</sup> As explained in the previous footnote, this parameterization involves some repetitions.

<sup>4)</sup> In the following, we choose normalizations in such a way that each  $\Pi_{r,I}$  has formal multiplicity equal to 1. There may be no actual module with this property, but since we are working only up to scale, this doesn’t cause any problems.

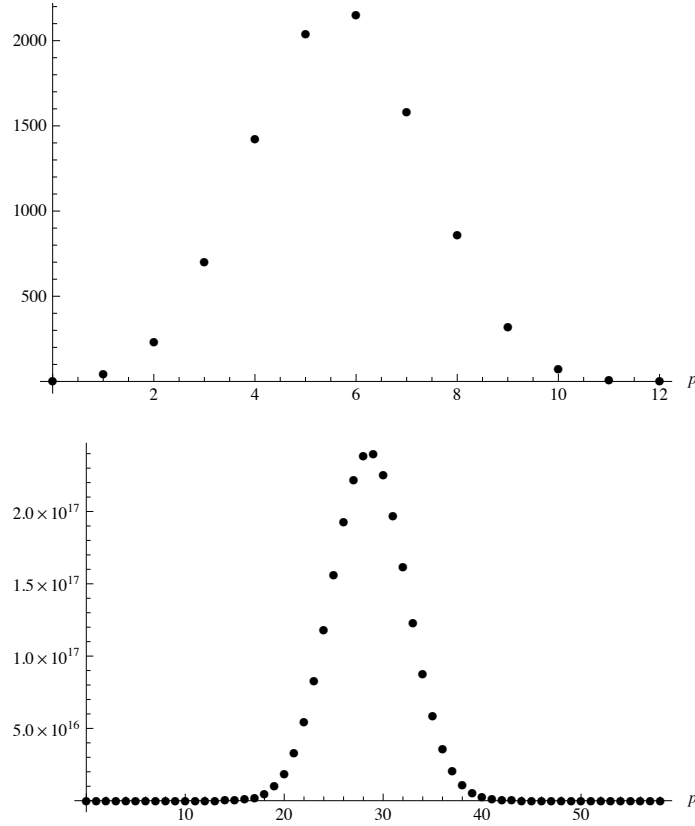


Figure 3. Betti numbers  $k_{p,1}$  of random Betti tables with  $r = 14$  and with  $r = 60$ .

**Theorem C.** Fix a weight  $q \in [1, n]$ . If one chooses  $x_r \in \Omega_r$  randomly, then as  $r \rightarrow \infty$  one expects

$$(0.3) \quad k_{p,q}(x_r) \sim \frac{1}{2^n} \cdot \binom{r-n}{p} \cdot \left( \frac{p^{q-1} \cdot (r-p-n)^{n-q}}{(q-1)! \cdot (n-q)!} \right)$$

provided that  $\frac{p}{r}$  is bounded away from 0 and 1.

More precisely, given  $\varepsilon > 0$  denote by

$$\Sigma_{p,r,n}(\varepsilon) \subseteq \Omega_r$$

the set of all coefficient vectors  $x_r = \{x_{r,I}\} \in \Omega_r$  such that

$$\left| \frac{k_{p,q}(x_r)}{(\text{RHS of (0.3)})} - 1 \right| > \varepsilon.$$

Now suppose that  $\{p_r\}$  is a sequence of positive integers  $0 \leq p_r \leq r - n$  with the property that

$$c < \frac{p_r}{r} < 1 - c$$

for some small  $c > 0$ . Put the standard Lebesgue measure on the cube  $\Omega_r$ , so that  $\text{vol}(\Omega_r) = 1$ .

Then the assertion of the theorem is that

$$\lim_{r \rightarrow \infty} \text{vol}(\Sigma_{p_r, r, n}(\varepsilon)) = 0,$$

and that moreover the convergence is uniform in  $p_r$  (given the bounding constant  $c$ ).

The crucial term in (0.3) is the binomial coefficient. Stirling's formula then implies that as functions of  $p$ , the  $k_{p,q}(x)$  display the sort of normal distribution predicted by the conjecture with high probability.

**Corollary D.** *Fix  $n$  and  $q \in [1, n]$ . There exists a function  $F(r) = F_{q,n}(r)$  with the property that if  $\{p_r\}$  is a sequence with*

$$p_r \rightarrow \frac{r}{2} + a \cdot \frac{\sqrt{r}}{2}$$

for some  $a$ , then

$$F(r) \cdot k_{p_r, q}(x_r) \rightarrow e^{-a^2/2}$$

for “most” choices of  $x_r \in \Omega_r$ .

We refer to Section 2 for the precise meaning of the last phrase.

Returning to the setting of Conjecture B, one most likely would not expect the modules arising from geometry to be described by the sort of uniform choice of Boij–Söderberg coefficients built into Theorem C and Corollary D. However it seems that the qualitative picture coming out of our model is rather robust, and persists for other probability measures on  $\Omega_r$ . We prove a specific result in this direction in Section 3.1.

The sort of probabilistic model discussed in the present paper is very different from the notion of genericity that traditionally comes up in commutative algebra. Specifically, the standard perspective is to consider moduli spaces parameterizing flat families of graded modules (or graded free resolutions) with fixed numerical invariants, and to examine what happens at a general point of this moduli space. The random Betti tables studied here behave quite differently than what one expects in the algebraic setting. For example, these two approaches display essentially opposite semicontinuity properties. A random Betti table in our sense will be drawn from the interior of the cone of Betti tables, and thus “more random Betti table = less sparse Betti table.” By contrast, in the classical moduli perspective each individual Betti number is lower semicontinuous, and thus “more generic resolution = more sparse Betti table.” We analyze this tension more explicitly in a particular case in Section 3.2. From this viewpoint, Conjecture B predicts that the resolutions arising from geometry are extremely non-generic algebraically.<sup>5)</sup>

We close with a philosophical remark. Based on experience with the case of curves, it was expected in some quarters that high degree embeddings of algebraic varieties would display rather sparse Betti tables. However the results of [6] showed that this is not at all the case. On the other hand, the results of that paper do suggest that the syzygies of very positive embeddings exhibit uniform asymptotic behavior. If one believes this, it becomes somewhat hard to guess – given the computations here for random Betti tables – what one might expect other than a statement along the lines of Conjecture B.

<sup>5)</sup> It is not surprising that this should be so. For instance, the ideal of a very positive embedding of any variety is generated by quadrics, but of course the subspace spanned by these quadrics sits in very special position within the linear system of all quadrics on the ambient projective space.

Concerning the organization of the paper, in Section 1 we introduce random Betti tables and carry out the computations leading to Theorem C. These are recast using Stirling's formula in Section 2, where we prove Corollary D and Proposition A. Finally, Section 3 is devoted to some variants and open problems.

**Acknowledgement.** The computer algebra programs Macaulay2 and Mathematica (see [13, 19]) provided useful assistance in studying examples. We also thank David Eisenbud, Milena Hering and Claudiu Raicu for valuable suggestions. We have particularly profited from discussions with Kyungyong Lee, who helped us to understand the direction this project should take. We also thank the anonymous referee for a close reading that improved the paper.

## 1. Random Betti tables

In this section we introduce some notation and definitions concerning Betti tables, and carry out the main computations.

We fix once and for all a positive integer  $n \geq 2$ , and denote by  $r$  a large natural number that will eventually go to infinity. Generally speaking we will render dependence on  $r$  visible in the notation, but leave dependence on  $n$  implicit. We shall be concerned with  $n \times (r + 1 - n)$  matrices of real numbers: the columns will be numbered  $0, \dots, r - n$  and the rows  $1, \dots, n$ . As in the Introduction such *Betti tables* arise upon tabulating the Betti numbers of a finite-length graded module over the polynomial ring  $k[x_1, \dots, x_{r-n}]$ , but in fact we will only be concerned with the tables themselves. Given a Betti table  $B$ , we denote by  $k_{p,q}(B)$  the entry of  $B$  in the  $q^{\text{th}}$  row and  $p^{\text{th}}$  column. (Some authors work instead with  $\beta_{p,j}(B) = k_{p,j-p}(B)$ .)

The theorem of Eisenbud–Schreyer [10] implies that the cone of all Betti tables of modules as in the Introduction is spanned by so-called *pure tables*  $\pi(r, I)$  constructed as follows. Fix an  $(n - 1)$ -element subset

$$I = (i_1 < \dots < i_{n-1})$$

of  $[r] \stackrel{\text{def}}{=} \{1, \dots, r\}$ , and write

$$\{1, \dots, r\} - I = \{d_0, \dots, d_{r-n}\},$$

with  $d_0 < d_1 < \dots < d_{r-n}$ . Then  $\pi(r, I)$  is the Betti table with  $k_{p,j} = 0$  for  $j \neq d_p - p$ , and

$$(1.1) \quad k_{p,d_p-p}(\pi(r, I)) = (r - n)! \cdot \left( \prod_{\ell \neq p} \frac{1}{|d_\ell - d_p|} \right).$$

In other words,  $\pi(r, I)$  is (up to scaling) the Betti table of a module all of whose syzygy modules are generated in a single degree. The integers  $(d_0 < \dots < d_{r-n})$  are the *degree sequence* of  $\pi(r, I)$ , and the subset  $I \subseteq [r]$  determines the values of  $p$  at which the degree sequence skips an integer. The numerical factors in (1.1) are such that  $\pi(r, I)$  has (formal) multiplicity equal to 1. (This follows from the formula for the multiplicity of a module with a pure resolution; see, for instance, [2, p. 88].)

**Example 1.1.** Consider the table  $\pi(7, \{2, 4\})$ . This corresponds to the degree sequence

$$\{1, 2, 3, 4, 5, 6, 7\} \setminus \{2, 4\} = \{1, 3, 5, 6, 7\},$$

and one has

$$\pi(7, \{2, 4\}) = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{pmatrix} \frac{1}{10} & - & - & - & - \\ - & \frac{1}{2} & - & - & - \\ - & - & \frac{3}{2} & \frac{8}{5} & \frac{1}{2} \end{pmatrix} \end{matrix}.$$

It will be useful to have an alternative expression for the entries of  $\pi(r, I)$ . As a matter of notation, for  $m \in \mathbb{Z}$ , write

$$m_+ = \max\{m, 0\}.$$

Then (1.1) yields the following, whose proof we leave to the reader:

**Lemma 1.2.** *The entries of  $\pi(r, I)$  are given by*

$$(1.2) \quad k_{p,q}(\pi(r, I)) = (r-n)! \cdot \frac{\prod_{\alpha=1}^{q-1} (p+q-i_{\alpha})_+ \cdot \prod_{\alpha=q}^{n-1} (i_{\alpha}-p-q)_+}{(p+q-1)! \cdot (r-p-q)!}.$$

We next consider random Betti tables. For each  $I \in \binom{[r]}{n-1}$ , let  $X_I$  be a random variable uniformly distributed on  $[0, 1]$ . We take the  $X_I$  to be independent. Then put

$$(1.3) \quad B_r = B_{r,n} = \sum_{I \in \binom{[r]}{n-1}} X_I \cdot \pi(r, I).$$

Heuristically, one thinks of  $B_r$  as the Betti table obtained by choosing independent random Boij–Söderberg coefficients  $x_I \in [0, 1]$ . More formally,  $B_r$  is a random  $n \times (r+1-n)$  matrix, and for each  $p, q$  the entry  $k_{p,q}(B_r)$  is a random variable. We write  $\mathbf{e}_{p,q}(B_r)$  for the expected value of this variable, i.e.

$$(1.4) \quad \mathbf{e}_{p,q}(B_r) \stackrel{\text{def}}{=} \mathbf{E}(k_{p,q}(B_r)).$$

**Remark 1.3.** When  $n = 2$ , the pure tables are indexed by a single integer

$$i \in [1, r] = \binom{[r]}{1}.$$

In this case the  $\pi(r, i)$  are linearly independent, and there is a unique way to express every Betti table as a linear combination of pure tables. However when  $n \geq 3$ , chains in the set of degree sequences determine the simplices of a simplicial complex whose structure must be taken into account to get a unique decomposition. (See [2] or [10] for further details.) Here we have chosen to ignore this additional structure. This choice is primarily motivated by a desire to simplify the statements and computations, by avoiding some of the complicated geometry of the cone of Betti tables. However there is also emerging evidence from other directions that it can be advantageous to deal with all possible pure Betti tables, instead of just collections from a single simplex. For instance, the recent work [12] provides a pleasingly simple description of a pure table decomposition of any complete intersection, but this description relies on a collection of pure Betti tables that do not come from a single simplex. Lastly, due to the seeming robustness of Corollary D (see Section 3.1), we expect that taking the simplicial structure into consideration would not significantly alter the outcome.

Our main result computes  $\mathbf{e}_{p,q}(\mathbf{B}_r)$ , and shows that if we fix  $q$  and let  $r$  tend to infinity, then  $k_{p,q}(\mathbf{B}_{r,n})$  converges (up to an essentially polynomial factor) to a binomial distribution in  $p$ .

**Theorem 1.4.** *The expected entries of  $\mathbf{B}_r$  are given by the formula*

$$(1.5) \quad \mathbf{e}_{p,q}(\mathbf{B}_r) = \frac{1}{2^n} \cdot \binom{r-n}{p} \cdot \left( \frac{p^{q-1} \cdot (r-p-n)^{n-q}}{(q-1)! \cdot (n-q)!} + o(p, r-p)^{n-1} \right),$$

where the error term on the right indicates a quantity  $g$  where  $\lim_{p,r \rightarrow \infty} \frac{g}{p^i (r-p)^j} = 0$  for some nonnegative integers  $i, j$  such that  $i + j = n - 1$ .

**Remark 1.5.** Note that we could replace the factor  $(r - p - n)$  in (1.5) by  $(r - p)$  without changing the statement. However we prefer to emphasize the symmetry between  $(p, q)$  and  $(r - n - p, n + 1 - q)$  inherent in the situation.

With  $n$  and  $q$  fixed, write

$$(1.6) \quad \mu(r, p) = \mu_{q,n}(r, p) = \frac{1}{2^n} \cdot \frac{p^{q-1} \cdot (r - p - n)^{n-q}}{(q-1)! \cdot (n-q)!}$$

for the expression appearing on the right in (1.5). The law of large numbers then implies that  $k_{p,q}(\mathbf{B}_{r,n})$  converges to  $\mathbf{e}_{p,q}(\mathbf{B}_{r,n})$ . In fact:

**Corollary 1.6.** *Fix a weight  $q \in [1, n]$ , and let  $\{p_r\}$  be any sequence of positive integers  $0 \leq p_r \leq r - n$  such that*

$$c \leq \frac{p_r}{r} \leq 1 - c$$

for some small  $c > 0$ . Then as  $r \rightarrow \infty$ ,

$$\frac{k_{p_r,q}(\mathbf{B}_r)}{\binom{r-n}{p_r} \cdot \mu(r, p_r)} \rightarrow 1$$

in probability, and moreover the convergence is uniform in  $p_r$  (for given  $c$ ).

In other words, given  $\varepsilon, \eta > 0$ , one can find an integer  $R$  such that if  $r \geq R$ , then

$$\mathbf{P}\left(\left|\frac{k_{p_r,q}(\mathbf{B}_r)}{\binom{r-n}{p_r} \cdot \mu(r, p_r)} - 1\right| > \varepsilon\right) < \eta,$$

and furthermore one can take  $R$  to be independent of  $p_r$  provided that  $c < p_r/r < 1 - c$ .<sup>6)</sup> A more concrete interpretation of this assertion is spelled out following the statement of Theorem C in the Introduction.

Turning to the proof of Theorem 1.4, we begin with an elementary lemma. As a matter of notation, given positive integers  $a \geq b$ , and a  $b$ -element subset  $J \in \binom{[a]}{b}$  of  $[a]$ , write  $J = \{j_1 < \cdots < j_b\}$  for the elements of  $J$  in increasing order.

<sup>6)</sup> Here and subsequently, we are writing  $\mathbf{P}$  to denote the probability of an event.

**Lemma 1.7.** For fixed  $b \geq 0$ , consider the function  $\sigma_b : \mathbb{Z}_{\geq b} \rightarrow \mathbb{Z}$  given by

$$\sigma_b(a) = \sum_{J \in \binom{[a]}{b}} j_1 j_2 \cdots j_b.$$

Then  $\sigma_b(a)$  is a polynomial of degree  $2b$  with leading coefficient  $\frac{1}{b!2^b}$ .

*Proof.* If  $b = 0$ , the statement is trivial. We then observe that

$$\sigma_b(a+1) - \sigma_b(a) = (a+1)\sigma_{b-1}(a).$$

and the general case follows by induction on  $b$ .  $\square$

*Proof of Theorem 1.4.* Throughout the proof we fix  $n$  and  $q \in [1, n]$ . Note first from Lemma 1.2 that

$$k_{p,q}(\pi(r, I)) = \left( \frac{(p+q-i_1)_+ \cdots (p+q-i_{q-1})_+ \cdot (i_q-p-q)_+ \cdots (i_{n-1}-p-q)_+}{(p+q-1) \cdots (p+1) \cdot (r-p-q) \cdots (r-p-(n-1))} \right) \cdot \binom{r-n}{p}.$$

Given  $I \in \binom{[r]}{n-1}$ , define

$$(1.7) \quad Y_{I,r,p} = \left( \frac{(p+q-i_1)_+ \cdots (p+q-i_{q-1})_+ \cdot (i_q-p-q)_+ \cdots (i_{n-1}-p-q)_+}{(p+q-1) \cdots (p+1) \cdot (r-p-q) \cdots (r-p-(n-1))} \right) \cdot X_I.$$

Thus for fixed  $p, q$  and  $r$ :

$$(1.8) \quad \frac{k_{p,q}(\mathbf{B}_r)}{\binom{r-n}{p} \cdot \mu(r, p)} = \frac{\sum_{I \in \binom{[r]}{n-1}} Y_{I,r,p}}{\mu(r, p)},$$

where  $\mu(r, p)$  is the quantity from (1.6).

Now since  $\mathbf{E}(X_I) = \frac{1}{2}$ , for fixed  $r, p$  and  $I$ , we have

$$\mathbf{E}(Y_{I,r,p}) = \frac{1}{2} \cdot \left( \frac{(p+q-i_1)_+ \cdots (p+q-i_{q-1})_+ \cdot (i_q-p-q)_+ \cdots (i_{n-1}-p-q)_+}{(p+q-1) \cdots (p+1) \cdot (r-p-q) \cdots (r-p-(n-1))} \right).$$

Hence for fixed  $p$  and  $r$ :

$$\begin{aligned} \mathbf{e}_{p,q}(\mathbf{B}_{r,n}) &= \binom{r-n}{p} \cdot \sum_{I \in \binom{[r]}{n-1}} \mathbf{E}(Y_{I,r,p}) \\ &= \binom{r-n}{p} \cdot \frac{1}{2} \cdot \left( \frac{\sum_{I \in \binom{[r]}{n-1}} (p+q-i_1)_+ \cdots (p+q-i_{q-1})_+}{(p+q-1) \cdots (p+1) \cdot (r-p-q) \cdots (r-p-(n-1))} \right. \\ &\quad \left. \cdot (i_q-p-q)_+ \cdots (i_{n-1}-p-q)_+ \right) \\ &= \binom{r-n}{p} \cdot \frac{1}{2} \cdot \frac{\left( \sum_{J \in \binom{[p+q-1]}{q-1}} j_1 j_2 \cdots j_{q-1} \right) \left( \sum_{J' \in \binom{[r-p-q]}{n-q}} j'_1 \cdots j'_{n-q} \right)}{(p+q-1) \cdots (p+1) \cdot (r-p-q) \cdots (r-p-(n-1))}, \end{aligned}$$

where the last equality is given by a change of coordinates

$$I \mapsto (J, J') = ((j_1, \dots, j_{q-1}), (j'_1, \dots, j'_{n-q}))$$

where  $j_\ell = p + q - i_\ell$  for  $\ell = 1, \dots, q-1$  and  $j'_{\ell'-q-1} = i_{\ell'} - p - q$  for  $\ell' = q, \dots, n-1$ . This last expression further simplifies to

$$= \binom{r-n}{p} \cdot \frac{1}{2} \cdot \left( \frac{\sum_{J \in \binom{[p+q-1]}{q-1}} j_1 j_2 \cdots j_{q-1}}{(p+q-1) \cdots (p+1)} \right) \cdot \left( \frac{\sum_{J' \in \binom{[r-p-q]}{n-q}} j'_1 \cdots j'_{n-q}}{(r-p-q) \cdots (r-p-(n-1))} \right).$$

Lemma 1.7 then implies that

$$(1.9) \quad \mathbf{e}_{p,q}(\mathbf{B}_r) = \binom{r-n}{p} \cdot \mu(r, p) \cdot \left( \frac{1 + \frac{A_{2q-3}(p)}{p^{2q-2}}}{1 + \frac{B_{2q-3}(p)}{p^{2q-2}}} \right) \cdot \left( \frac{1 + \frac{C_{2n-2q-1}(r-p)}{(r-p)^{2n-2q}}}{1 + \frac{D_{2n-2q-1}(r-p)}{(r-p)^{2n-2q}}} \right),$$

where  $A, B, C$ , and  $D$  are polynomials of the indicated degrees whose coefficients depend on  $n$  and  $q$ . The theorem follows.  $\square$

**Remark 1.8.** For later reference, we record a consequence of the proof just completed. Specifically, suppose given a sequence  $\{p_r\}$  with

$$c \leq \frac{p_r}{r} \leq 1 - c$$

for some  $c > 0$ . Then it follows from (1.9) that

$$(1.10) \quad \lim_{r \rightarrow \infty} \left( \frac{\sum \mathbf{E}(Y_{I,r,p_r})}{\mu(r, p_r)} \right) = 1,$$

and that moreover the convergence is uniform in  $p_r$  given  $c$ .

*Proof of Corollary 1.6.* This is essentially just the weak law of large numbers, but for keeping track of the dependence in  $p$  it is quickest to go through the argument leading to that result, as in [4, Section 5.1]. As before, we fix  $n$  and  $q$  at the outset. Note to begin with that the coefficient of  $X_I$  in (1.7) is  $\leq 1$ , and hence

$$\text{Var}(Y_{I,r,p}) \leq \frac{1}{3}$$

for every  $I, r$  and  $p$ . Furthermore, since the  $X_I$  are independent, for given  $r$  and  $p$  the  $Y_{I,r,p}$  are uncorrelated. Therefore, with  $r$  and  $p$  fixed and  $\delta > 0$ , Chebyshev's inequality yields

$$\begin{aligned} \mathbf{P}\left(\left|\frac{\sum Y_{I,r,p} - \mathbf{E}(\sum Y_{I,r,p})}{\mu(r, p)}\right| > \delta\right) &\leq \frac{1}{\delta^2 \cdot \mu(r, p)^2} \cdot \text{Var}\left(\sum Y_{I,r,p}\right) \\ &\leq \frac{1}{\delta^2 \cdot \mu(r, p)^2} \cdot \frac{\binom{r}{n-1}}{3}. \end{aligned}$$

Now fix a sequence  $\{p_r\}$  with  $\frac{p_r}{r}$  bounded away from 0 and 1. Then

$$C_1 \cdot r^{n-1} \leq \mu(r, p_r) \leq C_2 \cdot r^{n-1}$$

for suitable positive constants  $C_1, C_2$  independent of  $p_r$ . Hence there is a constant  $C_3$ , inde-

pendent of  $p_r$ , such that

$$(1.11) \quad \mathbf{P}\left(\left|\frac{\sum Y_{I,r,p_r} - \mathbf{E}(\sum Y_{I,r,p_r})}{\mu(r, p_r)}\right| > \delta\right) \leq \frac{C_3}{\delta^2 \cdot r^{n-1}}.$$

Note that as  $r \rightarrow \infty$ , the term on the right goes to 0. Moreover, as we saw in Remark 1.8,

$$\lim_{r \rightarrow \infty} \frac{\mathbf{E}(\sum Y_{I,r,p_r})}{\mu(r, p_r)} = 1,$$

uniformly in  $p_r$ . It follows that given  $\varepsilon > 0$ :

$$\lim_{r \rightarrow \infty} \mathbf{P}\left(\left|\frac{\mathbf{E}(\sum Y_{I,r,p_r})}{\mu(r, p_r)} - 1\right| > \varepsilon\right) = 0,$$

uniformly in  $p_r$ . Recalling finally that

$$\frac{\mathbf{E}(\sum Y_{I,r,p_r})}{\mu(r, p_r)} = \frac{\mathbf{e}_{p_r,q}(\mathbf{B}_{r,n})}{\binom{r-n}{p} \cdot \mu(r, p_r)},$$

we arrive at the required statement.  $\square$

## 2. Betti asymptotics

This section contains two applications of Stirling's formula. First, we recast the computations of the previous section to give asymptotic expressions for the entries of a random Betti table. Then we return to the case of curves, and prove Proposition A from the Introduction.

In order to get clean statements for the Betti tables it will be helpful to replace convergence in probability with almost everywhere convergence. So we start with some definitions and observations in this direction. As before we fix once and for all an integer  $n \geq 2$ . Denote by  $\binom{[\infty]}{n-1}$  the set of all  $(n-1)$ -element subsets of  $\mathbb{Z}_{>0}$ , so that

$$\binom{[\infty]}{n-1} = \bigcup_{r \geq n-1} \binom{[r]}{n-1}.$$

Then put

$$\Omega = [0, 1]^{\binom{[\infty]}{n-1}}.$$

This is a countable product of copies of the unit interval, and there are natural projections

$$\rho_r : \Omega \rightarrow \Omega_r.$$

By a standard procedure (Kolmogorov's extension theorem, [5, Theorem A.3.1]) there is a unique probability measure on  $\Omega$  compatible with pull-backs of the standard measures on the  $\Omega_r$ . Via composition with  $\rho_r$ , the various functions considered in the previous section – notably  $X_I$ ,  $Y_{I,r,p}$  and  $k_{p,q}(\mathbf{B}_r)$  – determine measurable functions on  $\Omega$ . In other words, all these quantities are random variables on  $\Omega$ , and the computations of expectations and probabilities carried out in the previous section remain valid in this new setting. As a matter of notation, for  $x \in \Omega$ , we write  $x_r = \rho_r(x) \in \Omega_r$ , and set

$$k_{p,q}(x; r) \stackrel{\text{def}}{=} k_{p,q}(\mathbf{B}_r)(x_r).$$

Thus  $k_{p,q}(x; r)$  is the indicated entry of the finite Betti table determined by the Boij–Söderberg coefficient vector  $x_r \in \Omega_r$ .

The computations of the previous section then lead to the following:

**Theorem 2.1.** *Fix a weight  $q \in [1, n]$ , and a sequence  $\{p_r\}$  with*

$$c \leq \frac{p_r}{r} \leq 1 - c$$

*for some small  $c > 0$ . Then for almost all  $x \in \Omega$ :*

$$\lim_{r \rightarrow \infty} \frac{k_{p_r, q}(x; r)}{\binom{r-n}{p_r} \cdot \mu(r, p_r)} = 1.$$

*Sketch of Proof.* Assuming first that  $n \geq 3$ , one can simply modify slightly the first part of the proof of [4, Theorem 5.1.2]. In fact, keeping the notation of the previous section, put

$$Z_r \stackrel{\text{def}}{=} \frac{\sum_{I \in \binom{[r]}{n-1}} (Y_{I, r, p_r} - \mathbf{E}(Y_{I, r, p_r}))}{\mu(r, p_r)}.$$

In view of equations (1.8) and (1.10), it suffices to show that  $Z_r \rightarrow 0$  almost everywhere on  $\Omega$ . But when  $n \geq 3$ , it follows directly from (1.11) that given  $\delta > 0$ ,

$$\sum_r \mathbf{P}(|Z_r| > \delta) < \infty,$$

and as in [4] this implies the desired convergence. When  $n = 2$ , this argument does not work because the right-hand side of (1.11) then has order  $r^{-1}$ . However in this case one can for example adapt in a similar fashion the proof (cf. [5, Theorem 2.3.5]) of the strong law of large numbers for independent random variables with finite second and fourth moments. We leave details to the reader.  $\square$

Stirling's formula now implies the following, which in particular proves Corollary D from the Introduction:

**Corollary 2.2.** *Fix  $n$  and  $q$ , and let  $\{p_r\}$  be a sequence such that*

$$p_r \rightarrow \frac{r}{2} + a \cdot \frac{\sqrt{r}}{2}$$

*for some  $a \in \mathbb{R}$ . Then for almost all  $x \in \Omega$ :*

$$\left( \frac{(q-1)! \cdot (n-q)! \cdot \sqrt{2\pi r}}{2^{r+2-3n} \cdot r^{n-1}} \right) \cdot k_{p_r, q}(x; r) \rightarrow e^{-a^2/2}.$$

**Remark 2.3.** To simplify the formulas a bit, we have chosen to break the symmetry between  $(p, q)$  and  $(r-n-p, n+1-q)$ . One could obtain a similar formula for

$$p_r \rightarrow \frac{r-n}{2} + b \cdot \frac{\sqrt{r-n}}{2}.$$

*Proof of Corollary 2.2.* First, note that

$$\binom{r-n}{p} = \binom{r}{p} \cdot \frac{(r-p-n+1) \cdots (r-p-1) \cdot (r-p)}{(r-n+1) \cdots (r-1) \cdot r},$$

and hence

$$\lim_{r \rightarrow \infty} \frac{\binom{r-n}{p_r}}{\binom{r}{p_r}} = \frac{1}{2^n}.$$

Next, observe that

$$\lim_{r \rightarrow \infty} \frac{\mu(r, p_r)}{r^{n-1}} = \frac{1}{2^{2n-1} \cdot (q-1)! \cdot (n-q)!}.$$

Finally, recall (cf. [5, Theorem 3.1.2]) that

$$\lim_{r \rightarrow \infty} \frac{\sqrt{2\pi r}}{2^{r+1}} \cdot \binom{r}{p_r} = e^{-a^2/2}.$$

The corollary follows upon putting these together with Theorem 2.1.  $\square$

Changing gears slightly, we conclude this section by working out the asymptotics for the Betti numbers of large degree embeddings of curves. In particular, we prove Proposition A from the Introduction.

*Proof of Proposition A.* Let  $X$  be a smooth projective curve of genus  $g$ , and let  $L_d$  be a divisor of degree  $d \gg 0$  on  $X$ , so that

$$r_d \stackrel{\text{def}}{=} h^0(X, L_d) - 1 = d - g.$$

The first point is to compute  $k_{p,1}(X; L_d)$ . This appears in [18, Proposition 3.2], but for the benefit of the reader we recall the derivation. In fact, let  $M_d$  be the rank  $r_d$  vector bundle on  $X$  defined as the kernel of the evaluation mapping

$$\text{ev} : H^0(X, L_d) \otimes \mathcal{O}_X \rightarrow L_d.$$

Then, a well-known argument with Koszul cohomology (cf. [6, Section 3]) shows that

$$\begin{aligned} k_{p,1}(X; L_d) &= h^0(\Lambda^p M_d \otimes L_d) - \dim \Lambda^{p+1} H^0(L_d), \\ k_{p-1,2}(X; L_d) &= h^1(\Lambda^p M_d \otimes L_d). \end{aligned}$$

Furthermore,

$$k_{p-1,2}(X, L_d) = 0 \quad \text{for } p \leq r_d - g$$

thanks to a theorem of Green [14, Theorem 4.a.1]. Thus for  $p \leq r_d - g$ :

$$(*) \quad k_{p,1}(X; L_d) = \chi(X, \Lambda^p M_d \otimes L_d) - \binom{r_d + 1}{p + 1}.$$

On the other hand, the Euler characteristic in  $(*)$  can be computed by Riemann–Roch. In fact,  $\mu(M_d) = \frac{-d}{d-g}$ , and hence

$$\begin{aligned} \chi(\Lambda^p M_d \otimes L_d) &= \text{rank}(\Lambda^p M_d) \cdot (p \cdot \mu(M_d) + \mu(L_d) + 1 - g) \\ &= \binom{d-g}{p} \left( \frac{-pd}{d-g} + (d+1-g) \right). \end{aligned}$$

Writing

$$\binom{r_d + 1}{p + 1} = \binom{d - g}{p} \cdot \frac{d + 1 - g}{p + 1},$$

we find finally that

$$k_{p,1}(X; L_d) = \binom{d - g}{p} \left( \frac{-pd}{d - g} + (d + 1 - g) - \frac{d + 1 - g}{p + 1} \right)$$

provided that  $p \leq r_d - g$ . Proposition A from the Introduction then follows as in the proof of Corollary 2.2 by an application of Stirling's formula.  $\square$

**Remark 2.4.** When  $p > r_d - g = d - 2g$ , the Betti number  $k_{p,1}(X; L_d)$  depends on  $X$  and  $L_d$  in a manner that is not yet completely understood. On the other hand, the computation of syzygies as Koszul cohomology groups shows that  $K_{p,1}(X; L_d)$  is a sub-quotient of  $\Lambda^p H^0(L_d) \otimes H^0(L_d)$ , and therefore

$$k_{p,1}(X; L_d) \leq \binom{d + 1 - g}{p} \cdot (d + 1 - g).$$

For  $p > d - 2g$  this becomes a polynomial upper bound that is much smaller than the value of  $k_{p,1}$  for  $p \approx \frac{d-g}{2}$ . Thus as far as asymptotics are concerned, the  $K_{p,1}$  with  $p > d - 2g$  are essentially negligible. Of course in the statement of Proposition A these ambiguities do not even enter the picture.

### 3. Complements and open questions

In this section we establish some related results, and propose some open problems for further research.

**3.1. Weight functions.** Here we discuss the possibility of generalizing our results by taking non-uniform probability measures on  $\Omega_r$ . We will assume for simplicity that  $n = 2$ : in this case pure tables are labeled by a single integer  $i \in [1, r] = \binom{[r]}{1}$ .

Fix a function

$$h : [0, 1] \rightarrow [0, 1].$$

Then we can use  $h$  to define the *weighted* random Betti table

$$B_r^h \stackrel{\text{def}}{=} \sum_{i=1}^r h\left(\frac{i}{r}\right) \cdot X_i \cdot \pi(r, i),$$

where  $X_i$  ( $1 \leq i \leq r$ ) as above are independent random variables uniformly distributed on the interval  $[0, 1]$ . Equivalently, we are weighting the pure table  $\pi(r, i)$  with a random variable uniformly distributed on the interval  $[0, h(\frac{i}{r})]$ . This procedure is illustrated in Figure 4, which shows a random Betti table with  $r = 500$  weighted according to the somewhat whimsical choice

$$h(t) = \sin^2(2\pi(t - .35)).$$

The upper plot displays random coefficients  $x_i$  ( $1 \leq i \leq 500$ ) chosen uniformly in the interval  $[0, h(i/500)]$ . The resulting Betti numbers  $k_{p,1}$  are plotted at the bottom.

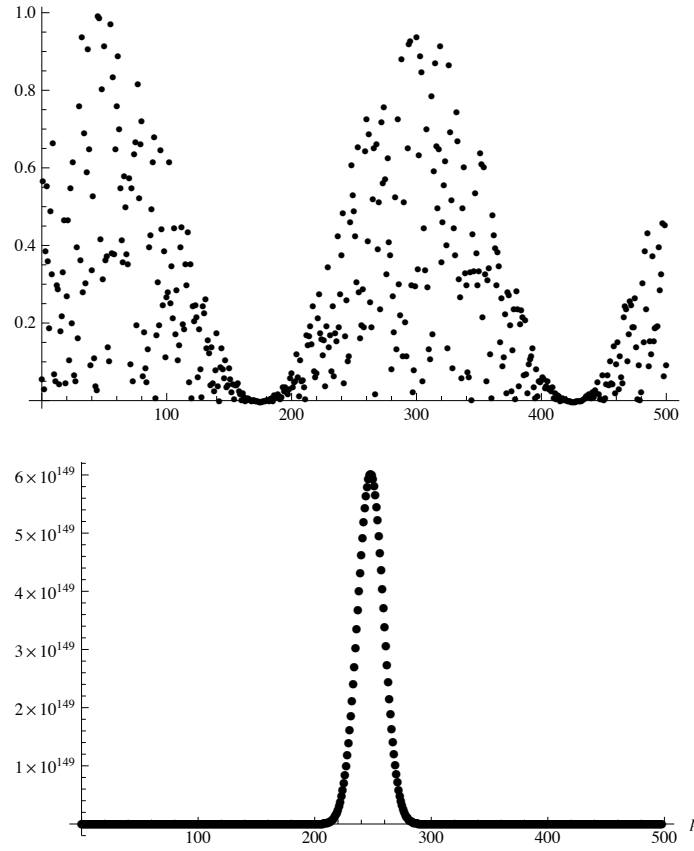


Figure 4. Weighted random Betti table with  $h(t) = \sin^2(2\pi(t - .35))$  and  $r = 500$ .

Now for arbitrary  $h$  one cannot expect the rows of  $B_r^h(x)$  to be normally distributed for most  $x$ . For example if  $\text{supp}(h) \subset [0, \frac{1}{2}]$ , then

$$k_{p,1}(B_r^h(x)) = 0$$

for every  $x$  whenever  $p > \frac{r}{2}$ . However if  $h$  is smooth, and if one rules out the sort of problem just illustrated, then the qualitative results established in the previous section – which correspond to the case  $h(t) \equiv 1$  – do remain valid.

**Theorem 3.1.** *Assume that  $h : [0, 1] \rightarrow [0, 1]$  is smooth, and that it is not identically zero on  $[0, \frac{1}{2}]$  or on  $[\frac{1}{2}, 1]$ . There exist functions  $F_1(r), F_2(r)$  with the property that if  $p_r$  is a sequence with*

$$p_r \rightarrow \frac{r}{2} + a \cdot \frac{\sqrt{r}}{2}$$

*for some  $a \in \mathbb{R}$ , then*

$$F_1(r) \cdot k_{p_r,1}(B_r^h(x)) \rightarrow e^{-a^2/2}, \quad F_2(r) \cdot k_{p_r,2}(B_r^h(x)) \rightarrow e^{-a^2/2}$$

*for almost all  $x \in \Omega$ .*

This suggests that the picture established in the previous section is actually quite robust.

**Remark 3.2.** Alternatively, one could consider the “deterministic” Betti table

$$\sum_{i=1}^r h\left(\frac{i}{r}\right) \cdot \pi(r, i).$$

An evident analogue of the theorem remains true in this setting. See Problem 3.6 for the potential interest in statements of this sort.

**Remark 3.3.** When  $n \geq 3$ , one expects that a similar result holds if one uses a smooth function

$$h : [0, 1]^{n-1} \rightarrow [0, 1]$$

to weight  $X_I$  ( $I \in \binom{[r]}{n-1}$ ) by  $h(\frac{i_1}{r}, \dots, \frac{i_{n-1}}{r})$ .

*Sketch of Proof of Theorem 3.1.* We will assume  $q = 1$ , the case  $q = 2$  being similar. Set

$$(3.1) \quad H(r, p) = \frac{1}{2(1 - \frac{p+1}{r})} \int_{\frac{p+1}{r}}^1 h(t) \cdot \left(t - \frac{p+1}{r}\right) dt.$$

Observe that  $H(r, \frac{r}{2}) \neq 0$  for  $r \gg 0$  thanks to the fact that  $h$  is smooth and does not vanish identically on  $[\frac{1}{2}, 1]$ . Furthermore, if  $p_r \rightarrow \frac{r}{2} + a \cdot \frac{\sqrt{r}}{2}$ , then  $H(r, p_r) \rightarrow H(r, \frac{r}{2})$ . Now, as in the proof of Theorem 1.4,

$$\begin{aligned} \mathbf{e}_{p,1}(\mathbf{B}_r^h) &= \binom{r-2}{p} \cdot \frac{\sum_{i=1}^r h(\frac{i}{r})(i-p-1)_+}{2(r-p-1)} \\ &= \binom{r-2}{p} \cdot \frac{\sum_{i=p+2}^r h(\frac{i}{r})(i-p-1)}{2(r-p-1)} \\ &= \binom{r-2}{p} \cdot \left(\frac{1}{2(1 - \frac{p+1}{r})}\right) \cdot \left(\sum_{i=p+2}^r h\left(\frac{i}{r}\right) \left(\frac{i}{r} - \frac{p+1}{r}\right)\right). \end{aligned}$$

But the product of the two right-hand terms is a Riemann sum for  $r \cdot H(r, p)$ , and one checks that with  $p_r$  as above one in fact has

$$\frac{\mathbf{e}_{p_r,1}(\mathbf{B}_r^h)}{\binom{r-2}{p_r} \cdot r \cdot H(r, \frac{r}{2})} \rightarrow 1 \quad \text{as } r \rightarrow \infty.$$

On the other hand, as in Theorem 2.1,

$$\frac{k_{p_r,1}(\mathbf{B}_r^h(x))}{\mathbf{e}_{p_r,1}(\mathbf{B}_r^h)} \rightarrow 1$$

for almost all  $x \in \Omega$ . One then concludes with an application of Stirling’s formula as in the proof of Corollary 2.2.  $\square$

**3.2. Some generic 2-regular modules.** As noted in the Introduction, it is interesting to compare the “random” Betti tables studied here with those arising from modules that are generic in a more traditional sense. Once again we will assume that  $n = 2$ , where the algebraic situation is particularly clear.

Consider then a finite-length graded module  $M$  over the polynomial ring

$$R = k[x_1, \dots, x_{r-2}],$$

with

$$k_{p,q}(M) = 0 \quad \text{for } q \neq 1, 2.$$

Such a module is 2-regular, and so is given by two vector spaces  $M_1, M_2$ , say of dimensions  $m_1, m_2$ , together with a mapping

$$(*) \quad M_1 \otimes R_1 \rightarrow M_2$$

determining the  $R$ -module structure. After choosing bases,  $(*)$  is in turn equivalent to specifying an  $m_2 \times m_1$  matrix  $\phi$  of linear forms. We will write  $M_\phi$  for the module corresponding to a matrix  $\phi$ . Note that the possible choices of  $\phi$  are parameterized by an irreducible variety.

Returning for a moment to the setting of Section 1, fix now a Boij–Söderberg coefficient vector

$$x = (x_1, \dots, x_r) \in \Omega_r = [0, 1]^r,$$

and consider the corresponding Betti table. This is a sum of  $r$  pure tables, and the expected value of each  $x_i$  is  $\frac{1}{2}$ . Thus the expected formal multiplicity of the table in question is  $\frac{r}{2}$ , and the expected formal Hilbert function has values  $\frac{r}{4}$  in degrees 1 and 2, and 0 for all other degrees. This leads us to consider modules  $M$  with

$$\dim M_1 = \dim M_2 = s$$

for some integer  $s$ , which are described by an  $s \times s$  matrix  $\phi$  of linear forms.

We prove:

**Proposition 3.4.** *There are arbitrarily large integers  $s$  with the property that if  $\phi$  is a general  $s \times s$  matrix of linear forms, then the Betti table of  $M_\phi$  is a sum of pure tables of type  $\pi(r, \{\lfloor \frac{r+1}{2} \rfloor\})$  and  $\pi(r, \{\lceil \frac{r+1}{2} \rceil\})$ . In particular, if  $r$  is odd, then  $M_\phi$  has a pure resolution.*

We note that a conjecture of Eisenbud–Fløystad–Weyman [9, Conjecture 6.1] implies that the statement should hold for all sufficiently large  $s$ . In any event, the proposition shows that genericity in the module-theoretic sense can lead to completely different behavior than that which occurs for the random tables considered above. Observe that this does not contradict Conjecture B: in fact, the results of [6] imply that the resolutions arising in the geometric setting are very far from pure.

*Proof of Proposition 3.4.* For any given  $s$ , it suffices (by the semicontinuity of Betti numbers) to produce one example where the statement holds. Any module whose Betti table is a multiple of  $\pi(r, \{\lfloor \frac{r+1}{2} \rfloor\}) + \pi(r, \{\lceil \frac{r+1}{2} \rceil\})$  will yield such an example, and the statement then follows either from the existence of pure resolutions [9, 10], or even from the special cases of pure resolutions constructed in [8, A2.6] or [3].  $\square$

**3.3. Open questions.** We conclude by proposing a few problems.

First, the reader will note that all our arguments are purely numerical in nature – they do not give an a priori sense why one would expect to see normal distribution of Betti numbers.

**Problem 3.5.** Find a probabilistic (or other) model that explains the behavior that we have established for random Betti tables.

Returning to the geometric questions motivating the present work, consider as in the Introduction a smooth projective variety

$$X \subseteq \mathbf{P}^{r_d}$$

of dimension  $n$  embedded by  $L_d = dA + P$ . Assuming for instance that  $H^i(X, \mathcal{O}_X) = 0$  for  $0 < i < n$ , so that  $X$  is arithmetically Cohen–Macaulay when  $d \gg 0$ , one can consider the Boij–Söderberg decomposition of the resolution of  $X$ . We pose the somewhat vague

**Problem 3.6.** Can one find a “nice” function  $h$  that governs this decomposition as in Section 3.1? If so, what are its properties?

The problem is most immediately meaningful in the case  $n = 2$ , in which case Boij–Söderberg coefficients  $x_i$  are well-defined for each  $i \in [1, r_d - 2]$ . The question then becomes whether they interpolate a fixed smooth or continuous function  $h$ .<sup>7)</sup> At the moment an affirmative solution seems out of reach, since a good answer to the problem would presumably imply Conjecture B. On the other hand, the question is philosophically in keeping with recent work on linear series (e.g. [15, Chapter 2.2.C], [7, 16]) and Hilbert series (e.g. [1, 17]), where it has become apparent that one can often define asymptotic invariants that behave surprisingly well. The asymptotic Boij–Söderberg coefficients of large degree embeddings of curves have been analyzed by the second author [11].

Finally, we expect that Conjecture B, if true, is essentially an algebraic fact.

**Problem 3.7.** Find a purely algebraic statement that implies, or runs parallel to, Conjecture B.

## References

- [1] *M. Beck* and *A. Stapledon*, On the log-concavity of Hilbert series of Veronese subrings and Ehrhart series, *Math. Z.* **264** (2010), no. 1, 195–207.
- [2] *M. Boij* and *J. Söderberg*, Graded Betti numbers of Cohen–Macaulay modules and the multiplicity conjecture, *J. Lond. Math. Soc. (2)* **78** (2008), 85–106.
- [3] *D. A. Buchsbaum* and *D. Eisenbud*, Remarks on ideals and resolutions, in: *Convegno di algebra commutativa* (Rome 1971), *Sympos. Math.* **11**, Academic Press, London (1973), 193–204.
- [4] *K. L. Chung*, *A course in probability theory*, 3rd ed., Academic Press, San Diego 2001.
- [5] *R. Durrett*, *Probability: Theory and examples*, 4th ed., Cambridge University Press, Cambridge 2010.
- [6] *L. Ein* and *R. Lazarsfeld*, Asymptotic syzygies of algebraic varieties, *Invent. Math.* **190** (2012), 603–646.
- [7] *L. Ein*, *R. Lazarsfeld*, *M. Mustață*, *M. Nakamaye* and *M. Popa*, Asymptotic invariants of line bundles, *Pure Appl. Math. Q.* **1** (2005), 397–403.
- [8] *D. Eisenbud*, *Commutative algebra. With a view toward algebraic geometry*, *Grad. Texts in Math.* **150**, Springer-Verlag, New York 1995.
- [9] *D. Eisenbud*, *G. Fløystad* and *J. Weyman*, The existence of equivariant pure free resolutions, *Ann. Inst. Fourier (Grenoble)* **61** (2011), 905–926.
- [10] *D. Eisenbud* and *F. Schreyer*, Betti numbers of graded modules and cohomology of vector bundles, *J. Amer. Math. Soc.* **22** (2009), 859–888.

---

<sup>7)</sup> It is possible that the scaling chosen above is not the appropriate one to use.

- [11] *D. Erman*, Betti tables of curves are asymptotically pure, in preparation.
- [12] *C. Gibbons, J. Jeffries, S. Mayes, C. Raicu, B. Stone and B. White*, Non-simplicial decompositions of Betti diagrams of complete intersections, preprint 2013, <http://arxiv.org/abs/1301.3441>.
- [13] *D. R. Grayson and M. E. Stillman*, Macaulay2. A software system for research in algebraic geometry, available at <http://www.math.uiuc.edu/Macaulay2/>.
- [14] *M. Green*, Koszul cohomology and the geometry of projective varieties, *J. Differential Geom.* **19** (1984), 125–171.
- [15] *R. Lazarsfeld*, Positivity in algebraic geometry. I, Springer-Verlag, Berlin 2004.
- [16] *R. Lazarsfeld and M. Mustață*, Convex bodies associated to linear series, *Ann. Sci. Éc. Norm. Supér.* (4) **42** (2009), 783–835.
- [17] *A. McCabe and G. G. Smith*, Log-concavity of asymptotic multigraded Hilbert series, preprint 2011, <http://arxiv.org/abs/1109.4135>.
- [18] *M. Teixidor i Bigas*, Syzygies using vector bundles, *Trans. Amer. Math. Soc.* **359** (2007), 897–908.
- [19] *Wolfram Research Inc.*, *Mathematica*, Version 8.0, 2010.

---

Lawrence Ein, Department of Mathematics, University of Illinois at Chicago,  
851 South Morgan St., Chicago, IL 60607, USA  
e-mail: ein@uic.edu

Daniel Erman, Department of Mathematics, University of Michigan,  
Ann Arbor, MI 48109, USA  
e-mail: erman@umich.edu

Robert Lazarsfeld, Department of Mathematics, University of Michigan,  
Ann Arbor, MI 48109, USA  
e-mail: rlaz@umich.edu

Eingegangen 12. August 2012, in revidierter Fassung 21. März 2013