# Random numerical semigroups and a simplicial complex of irreducible semigroups

Jesus De Loera

Christopher O'Neill

Department of Mathematics University of California Davis Davis, CA, U.S.A.

deloera@math.ucdavis.edu

Department of Mathematics and Statistics San Diego State University San Diego, CA, U.S.A.

cdoneill@sdsu.edu

Dane Wilburne

Institute for Computational and Experimental Research in Mathematics Brown University Providence, RI, U.S.A.

### dane\_wilburne@brown.edu

Submitted: Apr 17, 2018; Accepted: Nov 1, 2018; Published: Nov 16, 2018 © The authors. Released under the CC BY-ND license (International 4.0).

#### Abstract

We examine properties of random numerical semigroups under a probabilistic model inspired by the Erdős-Rényi model for random graphs. We provide a threshold function for cofiniteness, and bound the expected embedding dimension, genus, and Frobenius number of random semigroups. Our results follow, surprisingly, from the construction of a very natural shellable simplicial complex whose facets are in bijection with irreducible numerical semigroups of a fixed Frobenius number and whose h-vector determines the probability that a particular element lies in the semigroup.

Mathematics Subject Classifications: 20M14, 05E45

# 1 Introduction

A numerical semigroup is a subset  $S \subset \mathbb{Z}_{\geq 0}$  that is closed under addition (we do **not** require S to have finite complement in  $\mathbb{Z}_{\geq 0}$ ). A nonnegative integer n is a gap of S if  $n \notin S$  and we denote the set of gaps of S by G(S). Numerical semigroups appear in several areas of mathematics [5], and there are several interesting combinatorial invariants of a semigroup [18]. Notable numerical semigroup invariants include

- 1. the embedding dimension e(S), which is the size of the (unique) minimal generating set of S,
- 2. the genus g(S) = #G(S), which is the number of gaps of S, and
- 3. the Frobenius number F(S), which is the largest gap of S.

The latter two invariants are usually only defined when S is *cofinite*, that is, when S has finite complement in  $\mathbb{Z}_{\geq 0}$ . The theory of numerical semigroups is a vibrant subject, with connections to algebraic geometry and commutative algebra [1, 6, 9, 10, 15, 20] as well as integer optimization and number theory (see [5] and references therein). In this paper, we investigate invariants of numerical semigroups from a probabilistic point of view.

V. Arnold [4], J. Bourgain and Y. Sinai [8] initiated the study of the "average behavior" of numerical semigroups by analyzing the Frobenius function F(S) for "typical" numerical semigroups (see more recent work in [2]). Each of these papers produced random numerical semigroups using the uniform probability distribution on the collection

$$G(N,T) = \{ \mathbf{a} \in \mathbb{Z}_{>0}^N \colon \gcd(\mathbf{a}) = 1 \text{ and } \|\mathbf{a}\|_{\infty} \leqslant T \}$$

of generating sets, and each proved several interesting statements about the expected value (in the usual probabilistic sense) of the Frobenius number. See the references in [2] for a thorough overview.

In this paper, we study a different model, which generates at random a numerical semigroup  $\mathcal{S}$  according to the following procedure:

- 1. fix a nonnegative integer M and a probability  $p \in [0, 1]$ ;
- 2. for each positive integer  $n \leq M$ , independently choose with probability p whether to include n in a set of generators  $\mathcal{A}$  for  $\mathcal{S}$ .

The notation  $S \sim S(M, p)$  indicates S is a random numerical semigroup produced with this model. A similar model was recently used to produce random monomial ideals [11] (that is, each multivariate monomial with bounded total degree is included in a generating set with probability p). The authors dubbed this model the "ER-type model" because its definition was inspired by the Erdős-Rényi model of random graphs [13]; we will use the same convention.

Unlike previously studied models, which sampled uniformly among numerical semigroups with a fixed number of generators, the ER-type model allows one to specify a probability as input, yielding more refined control over the numerical semigroups produced. Our model is also more closely aligned with the "standard" sampling methods from probabilistic combinatorics, and more compatible with the use of numerical semigroups in integer programming, where non-cofinite semigroups occur alongside cofinite ones.

Our main result is as follows.

**Theorem 1.** Let  $S \sim S(M, p)$ , where p = p(M) is a monotone decreasing function of M. (a) If  $p \ll \frac{1}{M}$ , then  $S = \langle 0 \rangle$  a.a.s. (b) If  $\frac{1}{M} \ll p \ll 1$ , then S is cofinite, i.e., the set of gaps is finite, a.a.s and

$$\lim_{M \to \infty} \mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] = \lim_{M \to \infty} \mathbb{E}\left[\mathbf{g}(\mathcal{S})\right] = \lim_{M \to \infty} \mathbb{E}\left[\mathbf{F}(\mathcal{S})\right] = \infty$$

(c) If  $\lim_{M\to\infty} p(M) > 0$ , then

 $\lim_{M \to \infty} \mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] < \infty, \qquad \lim_{M \to \infty} \mathbb{E}\left[\mathbf{g}(\mathcal{S})\right] < \infty, \qquad and \qquad \lim_{M \to \infty} \mathbb{E}\left[\mathbf{F}(\mathcal{S})\right] < \infty,$ 

and each limit is bounded by explicit rational functions in p (see Theorem 30).

Although part (a) of Theorem 1 follows from standard arguments in probabilistic combinatorics (Theorem 5), parts (b) and (c) follow, surprisingly, from the construction of a very natural shellable simplicial complex (Definition 8) whose facets are in bijection with irreducible numerical semigroups of a fixed Frobenius number (Definition 7). As it turns out, some of the probabilities involved in determining the expected values above require precisely the *h*-vector (in the sense of algebraic combinatorics [21]) for this simplicial complex. Through the *h*-vector, we distinguish parts (b) and (c) of Theorem 1 (Corollary 26) and estimate the finite expectations (Theorem 30).

## 2 Distribution and cofiniteness

In this section, we prove that the threshold function for cofiniteness coincides with the threshold function for nonemptyness (Theorem 5). We begin by briefly establishing some notation and terminology for threshold functions; for a more comprehensive resource on methods in probabilistic combinatorics, we refer the reader the excellent book of Alon and Spencer [3].

If f, g are functions depending on some parameter n, we use the notation  $f \gg g$  to indicate that the ratio  $g/f \to 0$  as  $n \to \infty$  and similarly  $f \ll g$  means that  $f/g \to 0$  as  $n \to \infty$ . With this, clearly  $f \ll 1$  implies  $f \to 0$  as  $n \to \infty$ . In the theory of Erdős-Rényi random graphs, many graph properties tend to appear or not appear with high probability based on the asymptotics of the probability parameter p. This phenomenon is quantified by the notion of a *threshold function* [3, Chapter 10]. Here we define a notion of threshold function tailored to the context of random numerical semigroups. We say that a property  $\mathcal{P}$  of a numerical semigroup S is *monotone* if whenever S has  $\mathcal{P}$ , so does every oversemigroup of S. For example, the property of being cofinite is monotone, whereas the property of having Frobenius number n is not. A *threshold function* for a monotone numerical semigroup property  $\mathcal{P}$  is a function t(M) such that

$$\lim_{M \to \infty} \mathbb{P}\left[\mathcal{S} \text{ has } \mathcal{P}\right] = \begin{cases} 1 & \text{if } p(M) \gg t(M), \\ 0 & \text{if } p(M) \ll t(M) \end{cases}$$

THE ELECTRONIC JOURNAL OF COMBINATORICS 25(4) (2018), #P4.37

for  $S \sim S(M, p)$  as  $M \to \infty$ . If  $\lim_{M\to\infty} \mathbb{P}[S \text{ has } \mathcal{P}] = 1$ , then  $\mathcal{P}$  is said to hold asymptotically almost surely (hereafter abbreviated as *a.a.s.*). Loosely speaking, if *t* is a threshold for a property  $\mathcal{P}$ , when *p* is much smaller than *t*, then S will not have  $\mathcal{P}$  a.a.s., and if *p* is much larger than *t*, S will have  $\mathcal{P}$  a.a.s.

Proposition 3 gives the probability of observing a fixed numerical semigroup in terms of its embedding dimension and gaps.

Remark 2. In what follows, we denote by  $g_M(S)$  the number of gaps n of S such that  $n \leq M$ .

**Proposition 3.** Let  $S \sim S(M, p)$  and let S be a numerical semigroup with minimal generating set  $A \subset [M]$ . Then,

$$\mathbb{P}\left[\mathcal{S}=S\right] = p^{\mathbf{e}(S)}(1-p)^{\mathbf{g}_M(S)}.$$

If, in addition,  $F(S) \leq M$ , then

$$\mathbb{P}\left[\mathcal{S}=S\right] = p^{\mathbf{e}(S)}(1-p)^{\mathbf{g}(S)}.$$

Proof. Let  $\mathcal{A}$  denote the generating set for  $\mathcal{S} \sim S(M, p)$ . The key observation is that  $\mathcal{S} = S$  if and only if  $A \subset \mathcal{A}$  and no gap of S is in  $\mathcal{A}$ . The first claim thus follows by independence, and the second claim then follows from the fact that  $F(S) \leq M$  implies  $g_M(S) = g(S)$ .

**Lemma 4.** Let  $S \sim S(M, p)$ . If  $p(M) \ll 1/M$ , then  $S = \langle 0 \rangle$  a.a.s.

*Proof.* The condition  $p(M) \ll 1/M$  implies  $pM \to 0$  as  $M \to \infty$ . By Markov's inequality,  $\mathbb{P}[|\mathcal{A}| > 0] \leq \mathbb{E}[|\mathcal{A}|]$  and by linearity of expectation,  $\mathbb{E}[|\mathcal{A}|] = pM \to 0$  as  $M \to \infty$ .  $\Box$ 

The proof of the threshold function for cofiniteness relies on counting the number of coprime pairs in the random set  $\mathcal{A}$ .

**Theorem 5.** The function t(M) = 1/M is a threshold function for the property that  $S \sim S(M, p)$  is cofinite.

*Proof.* By Lemma 4, when  $p \ll 1/M$ , S is the non-cofinite numerical semigroup  $\langle 0 \rangle$  a.a.s. Consider the case when  $p \gg 1/M$ . For each pair of coprime integers  $\{i, j\}$  such that  $1 \leq i < j \leq M$ , define the indicator random variable  $1_{\{i,j\}}$  to be

$$1_{\{i,j\}} := \begin{cases} 1, & \text{if } i, j \in \mathcal{A} \\ 0, & \text{otherwise} \end{cases}$$

and let

$$X = \sum_{\substack{i < j \\ (i,j) = 1}} 1_{\{i,j\}},$$

where (i, j) denotes the gcd of i and j. Then,  $\mathbb{E}[1_{\{i,j\}}] = p^2$  and  $\operatorname{Var}[1_{\{i,j\}}] = p^2(1-p^2)$ . Further, if  $\{i, j\} \cap \{i', j'\} = \emptyset$ ,  $\operatorname{Cov}[1_{\{i,j\}}, 1_{\{i',j'\}}] = 0$ , since the events  $1_{\{i,j\}} = 1$  and  $1_{\{i',j'\}} = 1$  are independent. If i = i' and  $j \neq j'$ ,

$$\operatorname{Cov}[1_{\{i,j\}}, 1_{\{i,j'\}}] = \mathbb{E}[1_{\{i,j\}} \cdot 1_{\{i,j'\}}] - \mathbb{E}[1_{\{i,j\}}]\mathbb{E}[1_{\{i,j'\}}] = p^3(1-p).$$

Thus,

$$\mathbb{E}[X] = \sum_{\substack{i < j \\ (i,j) = 1}} p^2 \leqslant M^2 p^2,$$

and

$$\operatorname{Var}[X] = \sum_{\substack{i < j \\ (i,j) = 1}} p^2 (1 - p^2) + \sum_{\substack{i < j < j' \\ (i,j) = 1 \\ (i,j') = 1}} p^3 (1 - p) \leqslant M^2 p^2 (1 - p^2) + M^3 p^3 (1 - p).$$

Hence,

$$\frac{\operatorname{Var}[X]}{\mathbb{E}[X]^2} \leqslant \frac{M^2 p^2 (1-p^2) + M^3 p^3 (1-p)}{\mathbb{E}[X]^2}.$$

Now, to get a handle on the denominator  $\mathbb{E}[X]^2$ , note that

$$\mathbb{E}[X] = \sum_{\substack{i < j \\ (i,j) = 1}} p^2 = \sum_{n=1}^{M-1} \varphi(n) p^2,$$

where  $\varphi(\cdot)$  denotes the Euler totient function. By [16, p. 131],  $\sum_{n=1}^{M-1} \varphi(n) \sim \frac{3}{\pi^2} M^2 + O(M \ln M)$ , from which it follows that  $\mathbb{E}[X]^2 \ge \frac{9}{\pi^4} M^4 p^4$  for M sufficiently large. Hence,  $\operatorname{Var}[X]/\mathbb{E}[X]^2 \to 0$  as  $M \to \infty$ . By the second moment method,  $\mathbb{P}[X > 0] \to 1$  and thus when  $p \gg 1/M$ ,  $\mathcal{A}$  will contain a pair of coprime integers a.a.s, which guarantees that  $\mathcal{S}$  is cofinite a.a.s. in this case.

# 3 The simplicial complex of irreducible semigroups

Before proving the remaining parts of Theorem 1, we introduce in Definition 8 a simplicial complex whose combinatorial properties govern several questions arising from the ER-type model for sampling random numerical semigroups. We prove that this complex is shellable (Proposition 12), in the process uncovering a combinatorial interpretation of its *h*-vector entries (Corollary 15). We begin by recalling the definition of a simplicial complex and some related concepts, as presented in [21, Chapter 2].

**Definition 6** ([21, Chapter 2]). Fix  $n \in \mathbb{Z}_{\geq 1}$ . A simplicial complex with vertices  $[n] = \{1, \ldots, n\}$  is a collection  $\Delta$  of subsets of [n] (called *faces*) such that  $A \in \Delta$  implies  $B \in \Delta$  whenever  $B \subset A$ . The dimension of a face A is dim(A) = |A| - 1 and the dimension

 $d = \dim(\Delta)$  of  $\Delta$  is the largest dimension among the faces of  $\Delta$ . The *facets* of  $\Delta$  are the maximal faces with respect to containment, and  $\Delta$  is *pure* if its facets all have the same dimension. The *f*-vector  $(f_{-1}, f_0, \ldots, f_d)$  of  $\Delta$  has entries giving the number

$$f_i = \#\{A \in \Delta : |A| = i+1\}$$

of *i*-dimensional faces of  $\Delta$ , and the *h*-vector  $(h_0, h_1, \ldots)$  of  $\Delta$  has entries expressed as

$$h_i = \sum_{j=0}^{i} (-1)^{i-j} \binom{d-j}{i-j} f_{j-1}$$

in terms of the f-vector.

**Definition 7** ([18, Chapter 2]). A numerical semigroup S is *irreducible* if it is maximal (with respect to containment) among numerical semigroups with the same Frobenius number F(S).

We now define the simplicial complex  $\Delta_n$  whose facets are in natural bijection with irreducible numerical semigroups with Frobenius number n.

**Definition 8.** Fix  $n \ge 2$ , and let  $S_1, \ldots, S_r$  denote the irreducible numerical semigroups with Frobenius number n. Define  $\Delta_n$  as the simplicial complex on [n-1] with facets  $F_i = S_i \cap [n-1]$  (or, equivalently, a set  $A \subset [n-1]$  is a face  $A \in \Delta_n$  whenever  $n \notin \langle A \rangle$ ). Let  $d_n = \deg h_n(x)$ , where

$$h_n(x) = h_{n,0} + h_{n,1}x + \cdots$$

denotes the polynomial in x whose coefficients are the h-vector  $(h_{n,0}, h_{n,1}, \ldots)$  of  $\Delta_n$ .

**Lemma 9.** A numerical semigoup S with Frobenius number F(S) = n is irreducible if and only if  $n - s \notin S$  implies  $s \in S$  for s < n/2. In particular,  $\Delta_n$  is pure of dimension  $\lfloor (n-1)/2 \rfloor$ .

*Proof.* The first claim follows from [18, Proposition 3.4], and yields a bijection between the gaps of S (excluding n/2) and the elements of S less than n. The second claim now follows.

Remark 10. Lemma 9 implies that for an irreducible numerical semigroup S, the set of minimal generators less than n/2 (so long as it is nonempty) uniquely determines S. In particular, the minimal generators determine which integers less than n/2 lie in S, and the fact that  $i \in S$  if and only if  $n - i \notin S$  determines the remainder of the gaps of S.

**Definition 11** ([21, Definition 2.1]). A shelling order of a pure simplicial complex  $\Delta$  is a total ordering  $F_1, \ldots, F_r$  of the facets of  $\Delta$  so that for every i > 1, the complex  $\{F_i \cap F_1, \ldots, F_i \cap F_{i-1}\}$  is pure of dimension dim $(\Delta) - 1$ . We say  $\Delta$  is shellable if it has a shelling order.

The electronic journal of combinatorics 25(4) (2018), #P4.37

**Proposition 12.** Fix  $n \ge 1$ , let  $S_1, \ldots, S_r$  denote the irreducible numerical semigroups with Frobenius number n, and let  $F_i = S_i \cap [n-1]$  be the facet of  $\Delta_n$  corresponding to  $S_i$ . If  $\sum F_i \ge \sum F_j$  for all i < j, then  $F_1, \ldots, F_r$  is a shelling order for  $\Delta_n$ .

Proof. Fix  $i \ge 2$ . It suffices to prove that whenever  $|F_j \cap F_i| < |F_i| - 1$  for j < i, there exists k < i such that  $F_k \cap F_i \supset F_j \cap F_i$ . Let  $a = \min(F_i \setminus F_j)$  and  $b = \max(F_j \setminus F_i)$ . Then  $S_i \setminus \{a\}$  is closed under addition since  $S_j$  is closed under addition and  $F_j$  and  $F_i$  have identical elements less than a. Additionally,  $S_i \cup \{b\} \setminus \{a\}$  is closed under addition, since its elements greater than b are identical to those of  $S_j$ . In particular,  $S_k = S_i \cup \{b\} \setminus \{a\}$  is an irreducible numerical semigroup with Frobenius number n, and k < i since  $\sum F_k > \sum F_i$ . This completes the proof.

**Theorem 13.** Each h-vector entry  $h_{n,i}$  is nonnegative and equals the number of irreducible numerical semigroups with Frobenius number n with exactly i minimal generators less than n/2.

*Proof.* Fix notation as in Proposition 12. Any shellable simplicial complex has nonnegative *h*-vector entries [21, Corollary 3.2], so the non-negativity of each  $h_{n,i}$  follows from Proposition 12. Moreover, under any shelling order of  $\Delta_n$  (in particular, under any ordering  $F_1, \ldots, F_r$  satisfying Proposition 12),  $F_j \setminus (F_1 \cup \cdots \cup F_{j-1})$  has a unique minimal face  $R_j$  for each j, and

$$h_{n,i} = \#\{R_j : \#R_j = i\}$$

counts the number of such minimal faces with *i* vertices. As such, it suffices to show each  $R_i$  equals the set  $G_i$  of minimal generators of  $S_i$  less than n/2.

Now, clearly  $h_{n,0} = 1$ , so assume  $G_i$  is nonempty. Any facet F containing  $G_i$  agrees with  $F_i$  for all elements less than n/2, and Remark 10 implies  $F_i = F$ . As such,  $F_i$  is the only facet containing the face  $G_i$ , and  $R_i \subset G_i$ . Conversely, suppose  $a \in G_i \setminus R_i$ , and let  $S' = S_i \setminus \{a\} \cup \{n - a\}$ . Since a is a minimal generator of S, the set S' is closed under addition and has the same number of elements less than n as  $S_i$ , meaning S' is an irreducible numerical semigroup. Since  $S_i$  contains  $R_i$  and its corresponding facet  $F_i$ appears before the facet F' corresponding to S' in the shelling order, we have arrived at a contradiction.

**Example 14.** The faces in the simplicial complex  $\Delta_7$  are depicted in Figure 1. The shelling order for  $\Delta_7$  produced by Proposition 12 is (456, 356, 246), and the unique minimal face corresponding to each facet (see the proof of Theorem 13) is double-circled in the figure. Since  $h_{n,i}$  counts the number of such minimal faces with exactly *i* elements, we have  $h_7(x) = 1 + 2x$ .

**Corollary 15.** There is a bijection between the irreducible numerical semigroups with Frobenius number n and the numerical semigroups not containing n whose generators are all less than n/2. In particular,  $h_{n,i}$  equals the number of embedding dimension i semigroups not containing n whose generators are all less than n/2.

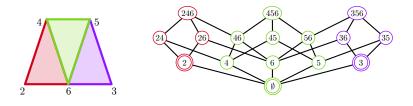


Figure 1: The simplicial complex  $\Delta_7$  (left) and its face poset (right); see Example 14.

*Proof.* For any irreducible numerical semigroup  $S = \langle n_1 < \cdots < n_k \rangle$  with F(S) = n, the semigroup  $T = \langle n_1, \ldots, n_t \rangle$  with  $n_t < n/2 < n_{t+1}$  is contained in S and thus cannot contain n. Conversely, fix a numerical semigroup  $T = \langle n_1 < \cdots < n_t \rangle$  with  $n_t < n/2$ , let

 $B = \{ s \in (n/2, n) : s \notin T \text{ and } n - s \notin T \},\$ 

and let  $S = T \cup B \cup \{n + 1, n + 2, ...\}$ . If  $s + s' \notin T$  for  $s \in B$  and  $s' \in T$ , then  $(n - s) - s' \notin T$ , meaning  $s + s' \in B$ . We conclude S is closed under addition, at which point Lemma 9 implies S is irreducible. The second claim now follows from the first and Theorem 13.

Remark 16. By Theorem 13, the coefficients  $h_{n,i}$  of the polynomial  $h_n(x)$  can be computed using an implementation of [7] in the GAP package numericalsgps [12], which gives an algorithm to compute the set of irreducible numerical semigroups with Frobenius number n. A precomputed list up to n = 90 can be found at the following webpages (the computations for  $n \ge 88$  each take over a day to complete with the authors' personal computers):

https://gist.github.com/coneill-math/c2f12c94c7ee12ac7652096329417b7d http://www.combinatorics.org/ojs/index.php/eljc/article/view/v25i4p37/data

The *h*-vectors of  $\Delta_{89}$  and  $\Delta_{90}$  (some entries of which are given in Table 1) demonstrate an interesting phenomenon: not only are the coefficients  $h_{n,i}$  not necessarily monotone for fixed *i*, but the fewer divisors *n* has, the larger  $h_{n,i}$  tends to be with respect to the surrounding *n*-values. This is likely due in part to the complex  $\Delta_n$  having more vertices in this case. The computations also take considerably longer in such cases; indeed, n = 89took longer than for n = 88 and n = 90 combined.

We remind the reader that the polynomial  $h_n(x)$  does not depend on a given numerical semigroup; rather, there is precisely one polynomial for each  $n \in \mathbb{Z}_{\geq 1}$ , and it encodes information about *all* numerical semigroups with Frobenius number n. Though a wide assortment of posets whose elements are numerical semigroups have been studied elsewhere in the literature [19, 17], we were unable to locate any that consider  $\Delta_n$ .

We are now ready to give the following bounds on the *h*-vector entries of  $\Delta_n$ , which play a crucial role in establishing the final threshold function in Section 4 and estimating several expected values in Section 5.

n = 89				n = 90				n = 500		
i	Lower	Actual	Upper	i	Lower	Actual	Upper	i	Lower	Upper
1	43	43	43	1	34	34	43	10	$2.4 \cdot 10^{12}$	$9.3 \cdot 10^{16}$
3	501	3873	9139	3	403	2442	9139	20	$8.1\cdot10^{18}$	$4.4\cdot10^{27}$
5	3025	27570	324632	5	2023	16065	324632	30	$3.4 \cdot 10^{22}$	$7.7\cdot 10^{34}$
7	6436	39358	2629575	7	3433	21213	2629575	40	$8.0 \cdot 10^{23}$	$1.3\cdot 10^{39}$
9	5005	18186	4686825	9	2002	8343	4686825	50	$1.4 \cdot 10^{23}$	$2.0\cdot 10^{40}$
11	1365	3044	1352078	11	364	1055	1352078	60	$1.8\cdot10^{20}$	$6.4\cdot10^{37}$
13	105	153	27132	13	14	31	27132	70	$5.2\cdot10^{14}$	$1.6\cdot 10^{30}$

Table 1: Comparison of the bounds in Theorem 17 with values computed in Remark 16.

**Theorem 17.** For any  $n \ge 1$ ,  $i \ge 1$ , and  $N \ge 2$ , we have

$$\sum_{j=2}^{N} \binom{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}{i} \leqslant h_{n,i} \leqslant \binom{\lceil n/2 \rceil - 2i}{i}.$$

*Proof.* By Corollary 15,  $h_{n,i}$  counts sets A minimally generating a semigroup not containing n. We claim

$$m = \min(A) \geqslant \begin{cases} 2i+1 & \text{if } n \text{ is even;} \\ 2i & \text{if } n \text{ is odd.} \end{cases}$$

Indeed, since A forms a minimal generating set, each element must be distinct modulo m. Additionally, since  $n \notin \langle A \rangle$ , any element  $a \in A$  cannot satisfy  $a \equiv n \mod m$ , and since  $\max(A) \leq \lfloor (n-1)/2 \rfloor$ , any two elements  $a, b \in A$  cannot satisfy  $a + b \equiv n \mod m$ . If n is odd, this means  $m \geq 2i$ , and if n is even, then additionally no  $a \in A$  satisfies  $a \equiv n/2 \mod m$ , so  $m \geq 2i + 1$ . This proves the claimed lower bound on m. Lastly, we can write

$$\lfloor (n-1)/2 \rfloor - (2i-1) = \lfloor (n+1)/2 \rfloor - 2i = \lceil n/2 \rceil - 2i,$$

from which the upper bound immediately follows.

For the lower bound, we claim that for any  $j \ge 2$ , each set A of i distinct integers chosen from the open interval (n/(j+1), n/j) minimally generates a numerical semigroup S with  $n \notin S$ . Indeed, the sum of any j elements of A is strictly less than n, while the sum of any j + 1 is strictly larger than n. Additionally, since  $j \ge 2$ , the sum of any two elements of A exceeds n/j, ensuring A minimally generates S. This completes the proof.

Remark 18. Proposition 19 implies the lower bound in Theorem 17 is tight for i = 1. The given bounds are tighter than those for more general Cohen-Macaulay simplicial complexes [21] and are sufficient to prove the results in the coming sections, but still leave room for improvement. Table 1 compares values from Theorem 17 with those computed in Remark 16.

We conclude this section by using Theorem 17, which gives some basic properties of  $h_n(x)$ , and Corollary 20, which gives a lower bound on the number of irreducible numerical semigroups with Frobenius number n, improving a previous bound in [14, Proposition 5].

**Proposition 19.** Fix  $n \ge 1$ .

(a) h<sub>n,0</sub> = 1.
(b) h<sub>n,1</sub> = ⌊(n + 1)/2⌋ − τ(n), where τ(n) denotes the number of divisors of n.
(c) d<sub>n</sub> = deg h<sub>n</sub>(x) = ⌊(n − 1)/2⌋ − ⌊n/3⌋.

*Proof.* We proceed using Corollary 15 and characterizing the possible sets  $A \subset (0, n/2)$  of integers minimally generating a numerical semigroup  $S = \langle A \rangle$  with  $n \notin S$ . Since  $A = \emptyset$  generates the semigroup  $S = \{0\}$ , we have  $h_{n,0} = 1$ . Additionally, any non-divisor of n less than n/2 generates a semigroup not containing n, which proves part (b).

Now, taking cases for n modulo 6, the bound on min(A) obtained in the proof of Theorem 17 implies  $d_n \leq \lfloor (n-1)/2 \rfloor - \lfloor n/3 \rfloor$ . Moreover, the set  $A = (n/3, n/2) \cap \mathbb{Z}$ minimally generates a semigroup not containing n since the sum of any two elements is strictly less than n while the sum of any three is strictly larger than n. Since  $|A| = d_n$ , this proves (c).

**Corollary 20.** For any  $N \ge 2$ , the number of irreducible numerical semigroups with Frobenius number n is at least

$$\sum_{j=2}^{N} 2^{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}.$$

## 4 Expected number of minimal generators

The main result of this section is Corollary 26, which states that if  $p \to 0$  as  $M \to \infty$ , then the expected number of generators, expected number of gaps, and expected Frobenius number are all unbounded. Our proof uses a surprising connection between the probability  $a_n(p)$  that a non-negative integer *n* lies outside the chosen semigroup (Definition 22) and the *h*-vector of the simplicial complex  $\Delta_n$  introduced in Section 3; see Remark 24.

Note 21. In the remainder of the paper, we adhere to the convention that if S is a numerical semigroup that is not cofinite, then g(S) = F(S) = 0. We do this so as not to affect  $\mathbb{E}[g(S)]$  and  $\mathbb{E}[F(S)]$ ; indeed, the assumption  $1/M \ll p(M)$  is made everywhere either of those two quantities appears, so Theorem 5 ensures that S is cofinite a.a.s.

**Definition 22.** Let  $S \sim S(M, p)$ . For each integer  $n \in [M]$ , denote by  $a_n(p)$  the probability that  $n \text{ cannot be written as a sum of elements in } S \cap [n-1]$ , that is,

$$a_n(p) := \mathbb{P}\left[n \notin \langle \mathcal{S} \cap [n-1] \rangle\right].$$

**Proposition 23.** For each  $n \ge 1$  with  $n \le M$ ,  $a_n(p) = (1-p)^{\lfloor n/2 \rfloor} h_n(p)$ .

THE ELECTRONIC JOURNAL OF COMBINATORICS 25(4) (2018), #P4.37

*Proof.* The faces of  $\Delta_n$  are precisely the sets  $A \subset [n-1]$  satisfying  $n \notin \langle A \rangle$ . In particular,  $a_n(p)$  is the probability of choosing a generating set for S that lies in  $\Delta_n$ . As such,

$$a_n(p) = \sum_{A \in \Delta_n} p^{|A|} (1-p)^{n-1-|A|} = (1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_n} f_{i-1} p^i (1-p)^{d_n-i} = (1-p)^{\lfloor n/2 \rfloor} h_n(p)$$

follows upon unraveling Definitions 6 and 8.

*Remark* 24. By Proposition 23, we can write

$$a_n(p) = (1-p)^n \mathcal{H}(\exists [x]/I_{\Delta_n}; p)$$

in terms of the Hilbert series of the *Stanley-Reisner ring*  $\exists [x]/I_{\Delta_n}$  (see [21, Chapter 2, Definition 1.1]). We encourage the interested reader to consult [21] for background on Hilbert functions.

**Theorem 25.** Let  $S \sim S(M, p)$ , where p = p(M) is a monotone decreasing function of M. If  $1/M \ll p \ll 1$ , then  $\lim_{M\to\infty} \mathbb{E}[e(S)] = \infty$ .

*Proof.* The probability of an integer n being a minimal generator of S equals  $pa_n(p)$  since the event that n is chosen (probability p) and the event n cannot be written as a sum of other chosen values (probability  $a_n(p)$ ) are independent. As such, by Proposition 23 and the linearity of expectation,

$$\mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] = \sum_{n=1}^{M} p a_n(p) = \sum_{n=1}^{M} p (1-p)^{\lfloor n/2 \rfloor} h_n(p).$$

For any fixed integer  $N \ge 2$ , the lower bound in Theorem 17 gives

$$\mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] \geqslant \sum_{n=1}^{M} p(1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_n} \sum_{j=2}^{N} \binom{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}{i} p^i$$
$$= \sum_{j=2}^{N} \sum_{n=1}^{M} p(1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_n} \binom{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}{i} p^i.$$

We now consider each summand  $\sigma_j$  of the outer sum for a fixed value of j. Using the division algorithm to write each n = k(2j(j+1)) + r for  $k \ge 0$  and  $1 \le r \le 2j(j+1)$ , and supposing M = m(2j(j+1)) for some  $m \in \mathbb{Z}_{\ge 1}$ , we obtain

$$\sigma_{j} = \sum_{n=1}^{M} p(1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_{n}} \left( \frac{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}{i} \right) p^{i}$$
$$= \sum_{n=1}^{M} p(1-p)^{\lfloor n/2 \rfloor} (1+p)^{\lfloor (n-1)/j \rfloor - \lfloor n/(j+1) \rfloor}$$

THE ELECTRONIC JOURNAL OF COMBINATORICS 25(4) (2018), #P4.37

$$= \sum_{k=1}^{m} p(1-p)^{kj(j+1)+O(1)} \sum_{r=1}^{2j(j+1)} (1+p)^{2k(j+1)-2kj+O(1)}$$
  
=  $\left(\sum_{r=1}^{2j(j+1)} (1-p)^{O(1)}(1+p)^{O(1)}\right) \sum_{k=1}^{m} p\left((1-p)^{j(j+1)}(1+p)^2\right)^{k-1}$   
=  $\left(\sum_{r=1}^{2j(j+1)} (1-p)^{O(1)}(1+p)^{O(1)}\right) p\frac{1-\left((1-p)^{j(j+1)}(1+p)^2\right)^m}{1-(1-p)^{j(j+1)}(1+p)^2}.$ 

Since  $p \gg 1/M$ , a simple calculus exercise shows  $((1-p)^{j(j+1)}(1+p)^2)^m \to 0$ . If  $p \to 0$ , we obtain

$$\mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] \geqslant \sum_{j=2}^{N} \sigma_j \to \sum_{j=2}^{N} \frac{2j(j+1)}{j(j+1)-2} \geqslant 2(N-1),$$

which must hold for every  $N \ge 2$ .

**Corollary 26.** Resuming notation from Theorem 25, if  $1/M \ll p \ll 1$ , then

$$\lim_{M \to \infty} \mathbb{E}\left[\mathrm{e}(\mathcal{S})\right] = \lim_{M \to \infty} \mathbb{E}\left[\mathrm{g}(\mathcal{S})\right] = \lim_{M \to \infty} \mathbb{E}\left[\mathrm{F}(\mathcal{S})\right] = \infty.$$

*Proof.* Apply Theorem 25 and the inequalities

$$e(S) - 1 \leq \min(S \setminus \{0\}) - 1 \leq g(S) \leq F(S),$$

which hold for any cofinite numerical semigroup S.

# 5 Approximations

In the final section of this paper, we prove the only remaining case in Theorem 1, namely where p is bounded away from zero (Theorem 30). In this case, it suffices to assume  $p \in (0, 1)$  is constant. In doing so, we provide explicit bounds on  $\mathbb{E}[e(S)]$ ,  $\mathbb{E}[g(S)]$ , and  $\mathbb{E}[F(S)]$  as  $M \to \infty$  using the *h*-vector bounds in Theorem 17; Remark 31 discusses the accuracy of these estimates.

**Lemma 27.** For any cofinite numerical semigroup S and  $M \leq F(S)$ ,  $|S \cap [0, M]| \leq M/2$ .

*Proof.* The key observation is that if  $a \in S$ , then  $F(S) - a \notin S$ , so at least half of the integers less than F(S) lie outside of S. As such, if  $|S \cap [0, M]| > M/2$  and F is the smallest gap of S not less than M, the semigroup  $T = S \cup [F+1, \infty)$  violates the observation.  $\Box$ 

Recall that  $g_M(S)$  denotes the number of gaps n of S such that  $n \leq M$ .

**Lemma 28.** Let  $S \sim S(M, p)$  where  $p \in (0, 1)$  is consant. Then,

$$\lim_{M \to \infty} \mathbb{E}\left[ \mathrm{g}(\mathcal{S}) \right] = \lim_{M \to \infty} \mathbb{E}\left[ \mathrm{g}_{M}(\mathcal{S}) \right].$$

The electronic journal of combinatorics 25(4) (2018), #P4.37

*Proof.* We split the expectation  $\mathbb{E}[g(\mathcal{S})]$  into two parts:

$$\mathbb{E}\left[g(\mathcal{S})\right] = \sum_{n=1}^{\infty} \mathbb{P}\left[n \in G(\mathcal{S})\right] = \sum_{n=1}^{M} \mathbb{P}\left[n \in G(\mathcal{S})\right] + \sum_{n > M} \mathbb{P}\left[n \in G(\mathcal{S})\right].$$

The first term is just  $\mathbb{E}[g_M(\mathcal{S})]$ . For the second term, notice that if n > M is a gap, then no consecutive pair of integers m, m + 1 in the range  $[1, \sqrt{M}]$  can be in  $\mathcal{S}$ , since

$$F(\langle m, m+1 \rangle) = m(m+1) - 2m - 1 \leqslant M.$$

For each even integer  $m \leq \sqrt{M}$ , let  $E_m$  denote the event that m and m-1 are not both in  $\mathcal{A}$ . Note that  $\mathbb{P}[E_m] = 1 - p^2$  and that for  $m \neq m'$ , the events  $E_m$  and  $E_{m'}$  are independent. Hence,

$$\mathbb{P}\left[n \in G(S)\right] \leqslant \mathbb{P}\left[E_2 \cap E_4 \cap \cdots\right] = (1 - p^2)^{\left\lfloor \frac{1}{2}\sqrt{M} \right\rfloor}$$

which implies the desired result.

**Proposition 29.** For constant  $p \in (0, 1)$ ,

$$\frac{6 - 8p + 3p^2}{2p - 2p^3 + p^4} \leqslant \lim_{M \to \infty} \sum_{n=1}^M a_n(p) \leqslant \frac{2 - p^2}{p^2}.$$

*Proof.* We begin with the upper bound. Suppose M = 2m for some  $m \in \mathbb{Z}_{\geq 1}$ . By Theorem 17,

$$\sum_{n=1}^{M} a_n(p) \leqslant \sum_{n=1}^{M} (1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_n} \binom{\lfloor n/2 \rfloor}{i} p^i \leqslant \sum_{n=1}^{M} (1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{\lfloor n/2 \rfloor}{i} p^i$$
$$= \sum_{n=1}^{M} (1-p^2)^{\lfloor n/2 \rfloor} = 1 + 2 \sum_{n=1}^{m} (1-p^2)^n = 1 + 2(1-p^2) \frac{1-(1-p^2)^m}{1-(1-p^2)},$$

which yields the claimed upper bound since  $(1 - p^2)^m \to 0$  as  $m \to \infty$ . For the lower bound, write n = 6k + r for  $r \in [6]$ , let M = 6m for some  $m \in \mathbb{Z}_{\geq 1}$ , and use Theorem 17 with N = 2 to obtain

$$\sum_{n=1}^{M} a_n(p) \ge \sum_{n=1}^{M} (1-p)^{\lfloor n/2 \rfloor} \sum_{i=0}^{d_n} {\binom{d_n}{i}} p^i = \sum_{n=1}^{M} (1-p)^{\lfloor n/2 \rfloor} (1+p)^{d_n}$$
$$= (6-8p+3p^2) \sum_{k=0}^{M} ((1-p)^3(1+p))^k,$$

and since  $((1-p)^3(1+p))^{m+1} \to 0$  as  $m \to \infty$ , the desired lower bound is obtained.  $\Box$ 

THE ELECTRONIC JOURNAL OF COMBINATORICS 25(4) (2018), #P4.37

**Theorem 30.** Let  $S \sim S(M, p)$  where  $p \in (0, 1)$  is constant. Then

$$\frac{6-8p+3p^2}{2-2p^2+p^3} \leqslant \lim_{M \to \infty} \mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] \leqslant \frac{2-p^2}{p},$$
$$\frac{6-14p+11p^2-3p^3}{2p-2p^3+p^4} \leqslant \lim_{M \to \infty} \mathbb{E}\left[\mathbf{g}(\mathcal{S})\right] \leqslant \frac{(1-p)(2-p^2)}{p^2}, \text{ and }$$
$$\frac{6-14p+11p^2-3p^3}{2p-2p^3+p^4} \leqslant \lim_{M \to \infty} \mathbb{E}\left[\mathbf{F}(\mathcal{S})\right] \leqslant \frac{2(1-p)(2-p^2)}{p^2}.$$

*Proof.* Lemma 27 implies  $g(S) \leq F(S) \leq 2g(S)$  for any cofinite numerical semigroup S. Since

$$\mathbb{E}\left[\mathbf{e}(\mathcal{S})\right] = p \sum_{n=1}^{M} a_n(p) \quad \text{and} \quad \mathbb{E}\left[\mathbf{g}_M(\mathcal{S})\right] = (1-p) \sum_{n=1}^{M} a_n(p),$$

each claimed inequality follows from Lemma 28 and Proposition 29.

Remark 31. Neither of the bounds in Proposition 29 use the full strength of Theorem 17. The summands for  $j \ge 3$  in the lower bound yield increasingly complicated rational functions in p, though the resulting sequence of values necessarily converges to 0 by Theorem 30. Additionally, substituting the upper bound for  $h_{n,i}$  given in Theorem 17 yields a sum that is nontrivial to unravel in a way sufficient to compute the limit as  $M \to \infty$ , and doing so would likely only marginally improve the resulting upper bounds in Theorem 30. Indeed, it is the larger values of i whose terms benefit from the improved upper bound, and these terms are rendered negligible by the large exponent in the accompanying value  $p^i$ .

We see here the need for improvements in the upper bound in Theorem 17. Several computed values can be found in Table 2. For instance, the upper bound for  $\mathbb{E}[\mathbf{e}(\mathcal{S})]$  in Theorem 30 could be made better by simply noting that  $\mathbf{e}(S)$  is at most the smallest generator (which has expected value 1/p). Such improvements to Theorem 17 should be possible, given the precise characterization of the *h*-vector of  $\Delta_n$  in Corollary 15.

It is also worth noticing that the polynomials computed in Remark 16 are not sufficient for accuracy for the p values in Table 2. Indeed, with p = 0.01 and p = 0.001, each partial sum for M = 90 fails to reach the lower bound, and even when p = 0.1, the last 10 summands (i.e. for  $n = 81, \ldots, 90$ ) each lie between 0.01 and 0.02, so the next several summands will likely still contribute significantly to the limit.

*Proof of Theorem 1.* Apply Theorems 5, Corollary 26 and Theorem 30.

#### Acknowledgements

The authors would like to thank Iskander Aliev, Martin Henk, Calvin Leng, and Pedro García-Sánchez for several helpful conversations and suggestions, and are grateful to Zachary Spaulding for assisting with the experiments in Table 2. The first and second author were partially supported by NSF grant DMS-1522158 to the University of California Davis. The first author was also partially supported by NSF grant DMS-1440140,

ĺ		Lower	Experi	iments	Upper	Rmk 16
	p	Thm $30$	M = 25000	M = 50000	Thm $30$	M = 90
	0.25	2.21	3.3663	3.3761	3.75	2.767
	0.1	2.64	4.6236	4.6402	19.9	3.782
	0.01	2.96	9.7906	9.776	199.9	0.858
	0.001	2.996	15.3096	16.9539	1999.9	0.089

Table 2: Comparing estimates of  $\mathbb{E}[e(\mathcal{S})]$  using the bounds in Theorem 30, exact computation using the polynomials in Remark 16, and experimental evidence from 100,000 samples.

while he visited the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2017 semester. The third author was supported by NSF collaborative grant DMS-1522662 to Illinois Institute of Technology.

# References

- S. S. Abhyankar, Local rings of high embedding dimension, Amer. J. Math. 89 (1967), no. 4, 1073–1077.
- [2] I. Aliev, M. Henk, and A. Hinrichs, *Expected Frobenius numbers*, J. Combin. Theory Ser. A **118** (2011), no. 2, 525–531.
- [3] N. Alon and J. H. Spencer, The probabilistic method, 4th ed., Wiley, 2016.
- [4] V. I. Arnold, Weak asymptotics of the numbers of solutions of Diophantine equations, Funktsional. Anal. i Prilozhen. 33 (1999), no. 4, 65–66.
- [5] A. Assi and P. A. García-Sánchez, Numerical semigroups and applications, RSME Springer Series, vol. 1, Springer, [Cham], 2016.
- [6] V. Barucci, D. E Dobbs, and M. Fontana, Maximality properties in numerical semigroups and applications to one-dimensional analytically irreducible local domains, vol. 598, American Mathematical Soc., 1997.
- [7] Victor Blanco and José Carlos Rosales, The tree of irreducible numerical semigroups with fixed Frobenius number, Forum Math. 25 (2013), no. 6, 1249–1261.
- [8] J. Bourgain and Ya. G. Sinaĭ, Limit behavior of large Frobenius numbers, Uspekhi Mat. Nauk 62 (2007), no. 4(376), 77–90.
- [9] W. C Brown and J. Herzog, One dimensional local rings of maximal and almost maximal length, J. Alg, 151 (1992), no. 2, 332–347.
- [10] A. Campillo, On saturations of curve singularities, Singularities 40 (1983), no. Part 1, 211.
- [11] J. De Loera, S. Petrović, L. Silverstein, D. Stasi, and D. Wilburne, *Random monomial ideals*, J. Algebra (to appear).

- [12] M. Delgado and P. A. García-Sánchez, numericalsgps, a gap package for numerical semigroups, ACM Commun. Comput. Algebra 50 (2016), no. 1, 12–24.
- [13] P. Erdős and A. Rényi, On random graphs, I, Publ. Math. (Debrecen) 6 (1959), 290–297.
- [14] R. Fröberg, C. Gottlieb, and R. Häggkvist, On numerical semigroups, Semigroup Forum 35 (1987), no. 1, 63–83.
- [15] E. Kunz, The value-semigroup of a one-dimensional gorenstein ring, Proc. Amer. Math. Soc. 25 (1970), no. 4, 748–751.
- [16] T. Nagell, Introduction to number theory, Almqvist & Wiksell, Stockholm, 1951.
- [17] J. C. Rosales, Families of numerical semigroups closed under finite intersections and for the Frobenius number, Houston J. Math. 34 (2008), no. 2, 339–348.
- [18] J. C. Rosales and P. A. Garcí a Sánchez, Numerical semigroups, Developments in Mathematics, vol. 20, Springer, New York, 2009.
- [19] J. C. Rosales, P. A. García-Sánchez, J. I. García-García, and J. A. Jiménez Madrid, *The oversemigroups of a numerical semigroup*, Semigroup Forum **67** (2003), no. 1, 145–158.
- [20] J. D Sally, On the associated graded ring of a local cohen-macaulay ring, J. Math. Kyoto Univ. 17 (1977), no. 1, 19–21.
- [21] R. P. Stanley, Combinatorics and commutative algebra, second ed., Progress in Mathematics, vol. 41, Birkhäuser Boston, Inc., Boston, MA, 1996.