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.
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 [10] (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” for its resemblance to the Erdős-Rényi model of random graphs [11]; we will use the same convention.

Unlike previously used 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 1.1.** 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

$$\lim_{M \to \infty} E[e(S)] = \lim_{M \to \infty} E[g(S)] = \lim_{M \to \infty} E[F(S)] = \infty.$$

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

$$\lim_{M \to \infty} E[e(S)], \lim_{M \to \infty} E[g(S)], \lim_{M \to \infty} E[F(S)] < \infty,$$

and each limit is bounded by explicit rational functions in $p$ (see page 11).

Although part (a) of Theorem 1.1 follows from standard arguments in probabilistic combinatorics (Theorem 3.4), parts (b) and (c) follow, surprisingly, from the construction of a very natural shellable simplicial complex (Definition 4.3) whose facets are in bijection with irreducible numerical semigroups of a fixed Frobenius number (Definition 4.2). 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 [18]) for this simplicial complex. Through the $h$-vector, we distinguish parts (b) and (c) of Theorem 1.1 (Corollary 5.6) and estimate the finite expectations (Theorem 6.4).

**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, 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.

2. Background

We begin by recalling some basic notions from probability theory and establishing notation used throughout the paper. For a more comprehensive resource on methods in probabilistic combinatorics, we refer the reader the excellent book of Alon and Spencer [3]. The **expected value** of a discrete random variable $X$ taking values in $\Omega \subseteq \mathbb{N}$ is $E[X] = \sum_{n \in \Omega} n \cdot P[X = n]$, and the **variance** of $X$ is given by $Var[X] = E[X^2] - E[X]^2$. The most useful property of the expectation operator
is its linearity: $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$ for random variables $X, Y$ and scalars $a, b$. The covariance of two random variables $X, Y$ is given by $\text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. The variance of a sum of random variables $X = X_1 + \ldots + X_n$ can be computed via the formula

$$\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] + \sum_{i<j} \text{Cov}[X_i, X_j].$$

The indicator random variable $1_E$ of an event $E$ takes the value 1 if $E$ occurs and is 0 otherwise. If $E$ occurs with probability $p$, then $\mathbb{E}[1_E] = p$ and $\text{Var}[1_E] = p(1-p)$. If $E$ depends on a parameter $n$ and $\lim_{n \to \infty} \mathbb{P}[E] = 1$, we say that $E$ holds asymptotically almost surely (abbreviated a.a.s.).

In Theorem 3.4 we will bound the probability that a nonnegative integer random variable $X$ is non-zero using the so-called moment method techniques. The first moment method, a consequence of Markov’s inequality \cite{3}, says that the probability $X$ is non-zero is bounded above by its expectation, i.e. $\mathbb{P}[X \neq 0] \leq \mathbb{E}[X]$. On the other hand, the second moment method provides an upper bound for the probability that $X$ is zero in terms of its variance, $\mathbb{P}[X = 0] \leq \text{Var}[X]/\mathbb{E}[X]^2$, and follows from Chebyshev’s inequality \cite[Theorem 4.1.1]{3}.

For functions $f, g$ 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 (see \cite[Chapter 10]{3}). 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 $S$ has $\mathcal{P}$ and if $S \cup \{s\}$ for $s \in \mathbb{N}$ is still a numerical semigroup, then $S \cup \{s\}$ also has $\mathcal{P}$. 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

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

for $S \sim S(M, p)$. Loosely speaking, if $t$ is a threshold for a property $\mathcal{P}$, when $p$ is much smaller than $t$, then $S \sim S(M, p)$ will not have $\mathcal{P}$ a.a.s., and if $p$ is much larger than $t$, $S$ will have $\mathcal{P}$ a.a.s.

3. DISTRIBUTION AND COFINITENESS

In this section, we prove that the threshold function for cofiniteness coincides with the threshold function for nonemptyness (Theorem 3.4). First, we give Theorem 3.2 which states the probability of observing a fixed numerical semigroup in terms of its embedding dimension and gaps.

**Notation 3.1.** In what follows, we denote by $g_M(S)$ the number of gaps $n$ of $S$ such that $n \leq M$.

**Theorem 3.2.** Let $S \sim S(M, p)$ and let $S$ be a numerical semigroup with minimal generating set $A \subset [M]$. Then,

$$\mathbb{P}[S = S] = p^{e(S)}(1 - p)^{g_M(S)}.$$

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

$$\mathbb{P}[S = S] = p^{e(S)}(1 - p)^{g(S)}.$$
Thus, independent. If \( i \neq j \),

\[
\mathbb{P}[\mathcal{A} = B] = p^{|B|}(1-p)^{M-|B|} = p^{c(S)}p^{|C|}(1-p)^{g_M(S)}(1-p)^{M-\epsilon(S)-g_M(S)-|C|}.
\]

Hence,

\[
\mathbb{P}[\mathcal{S} = \mathcal{S}] = \sum_{B \supseteq \mathcal{A}} \mathbb{P}[\mathcal{A} = B] = p^{c(S)}(1-p)^{g_M(S)} \sum_{C \subseteq \mathbb{N}} p^{|C|}(1-p)^{M-\epsilon(S)-g_M(S)-|C|}.
\]

By the binomial theorem, \( \sum_{C \subseteq \mathbb{N}} p^{|C|}(1-p)^{M-\epsilon(S)-g_M(S)-|C|} = 1 \), which completes the proof of the first claim. The second claim follows since if \( F(S) \leq M \), then \( g_M(S) = g(S) \).

**Lemma 3.3.** Let \( \mathcal{S} \sim S(M,p) \). If \( p \ll 1/M \), then \( \mathcal{S} = \{0\} \) a.a.s.

**Proof.** The condition \( p \ll 1/M \) implies \( pM \to 0 \) as \( M \to \infty \). Thus, \( \mathbb{E}[|\mathcal{A}|] = pM \to 0 \) as \( M \to \infty \) and the result follows by applying Markov’s inequality. \( \square \)

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

**Theorem 3.4.** The function \( t(M) = 1/M \) is a threshold function for the property that \( \mathcal{S} \sim S(M,p) \) is cofinite.

**Proof.** By Lemma 3.3, when \( p \ll 1/M \), \( \mathcal{S} \) is the non-cofinite numerical semigroup \( \{0\} \) a.a.s. Consider the case when \( p \gg 1/M \). For each pair \( \{i,j\} \) such that \( 1 \leq i < j \leq M \), define the indicator random variable \( 1_{\{i,j\}} \) as follows:

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

Let \( X \) be the number of pairs of coprime integers in \( \mathcal{A} \), so that

\[
X = \sum_{\substack{i < j \\{(i,j)\neq 1}}} 1_{\{i,j\}}.
\]

By the basic properties of indicator random variables, \( \mathbb{E}[1_{\{i,j\}}] = p^2 \) and \( \text{Var}[1_{\{i,j\}}] = p^2(1-p^2) \).

If \( \{i,j\} \cap \{i',j'\} = \emptyset \), \( \text{Cov}[1_{\{i,j\}}, 1_{\{i',j'\}}] = 0 \), since the events \( 1_{\{i,j\}} = 1 \) and \( 1_{\{i',j'\}} = 1 \) are independent. If \( i = i' \) and \( j \neq j' \),

\[
\text{Cov}[1_{\{i,j\}}, 1_{\{i,j'\}}] = \mathbb{E}[1_{\{i,j\}} \cdot 1_{\{i,j'\}}] - \mathbb{E}[1_{\{i,j\}}]\mathbb{E}[1_{\{i,j'\}}] = p^2(1-p).
\]

Thus,

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

and

\[
\text{Var}[X] = \sum_{\substack{i < j \\{(i,j)\neq 1}}} p^2(1-p^2) + \sum_{\substack{i < j < j' \\{(i,j) = 1 \text{ and } (i,j') = 1}}} p^3(1-p) \leq M^2 p^2(1-p^2) + M^3 p^3(1-p).
\]
Hence,
\[
\frac{\text{Var}[X]}{\mathbb{E}[X]^2} \leq \frac{M^2 p^2 (1 - p^2) + M^3 p^3 (1 - p)}{M^4 p^4} \to 0 \quad \text{as } M \to \infty.
\]
By the second moment method, \( P[X > 0] \to 1 \) and thus when \( p \gg 1/M, A \) will contain a pair of coprime integers a.a.s, which guarantees that \( S \) is cofinite a.a.s. in this case. \( \square \)

### 4. The simplicial complex of irreducible semigroups

Before proving the remaining parts of Theorem 1.1 we introduce in Definition 4.1 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 4.7), in the process uncovering a combinatorial interpretation of its \( h \)-vector entries (Corollary 4.10).

We begin by recalling the definition of a simplicial complex and some related concepts, as presented in [18, Chapter 2].

**Definition 4.1** ([18, 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** \((h_{-1}, h_0, \ldots)\) 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 4.2** ([14, 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 4.3.** Fix \( n \geq 2 \), 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 \( F_i = S_i \cap [n-1] \), and 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 4.4.** A numerical semigroup \( S \) 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 [14, 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. \( \square \)

**Remark 4.5.** Lemma 4.4 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 4.6** ([18] 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 $\{F_1 \cap F_i, \ldots, F_{i-1} \cap F_i\}$ is pure of dimension $\dim(\Delta) - 1$. We say $\Delta$ is shelly if it has a shelling order.

**Proposition 4.7.** Fix $n \geq 1$, 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 $\sum F_i \geq \sum F_j$ for all $i < j$, then $F_1, \ldots, F_r$ is a shelling order for $\Delta_n$.

**Proof.** Fix $i \geq 2$. It suffices to prove that whenever $|F_j \cap F_i| < |F_i| - 1$ for $j < i$, there exists $k < i$ such that $F_k \cap F_i \supseteq F_j \cap F_i$. 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 $S_i$ is closed under addition and $F_j$ and $F_i$ have identical elements less than $a$. Additionally, $S_i \setminus \{b\}$ is closed under addition, since its elements greater than $b$ are identical to those of $S_j$. In particular, $S_k = S_i \setminus \{b\}$, so the non-negativity of each $h$-vector entry follows from Proposition 4.7. This completes the proof. □

**Example 4.8.** The faces in the simplicial complex $\Delta_7$ are depicted in Figure 1. The shelling order for $\Delta_7$ produced by Proposition 4.7 is $(456, 356, 246)$, and the unique minimal face corresponding to each facet is double-circled in the figure. Since $h_{n,i}$ counts the number of such minimal faces with exactly $i$ elements, we have $h_7(x) = 1 + 2x$.

**Theorem 4.9.** 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 4.7. Any shellable simplicial complex has non-negative $h$-vector entries ([18] Corollary 3.2), so the non-negativity of each $h_{n,i}$ follows from Proposition 4.7. Moreover, under any shelling order of $\Delta_n$ (in particular, under any ordering $F_1, \ldots, F_r$ satisfying Proposition 4.7), $F_j \setminus (F_1 \cup \cdots \cup F_j)$ 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 4.5 implies $F_i = F$. As such, $F_i$ is the only facet containing the face $G_i$, and $R_i \subseteq G_i$. Conversely, suppose $a \in G_i \setminus R_i$, and let $S' = S_i \setminus \{a\} \setminus \{n - a\}$. Since $a$ is a minimal generator of $S_i$, 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 appears before $S_i$ in the shelling order, we have arrived at a contradiction. □

**Corollary 4.10.** 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 $n/2$. 
Proof. For any irreducible numerical semigroup \( S = \langle n_1, \ldots, n_k \rangle \) 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 \( T = \langle n_1, \ldots, n_t \rangle \) with \( n_t < n/2 \), let
\[
B = \{ s \in (n/2, n) : s \notin T \text{ and } n - s \notin T \},
\]
and let \( S = T \cup B \). 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 4.4 implies \( S \) is irreducible.

The second claim now follows from the first and Theorem 4.9. □

Remark 4.11. By Theorem 4.9, the coefficients \( h_{n,i} \) of the polynomial \( h_n(x) \) can be computed using [10], 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 webpage (the computations for \( n \geq 88 \) each take over a day to complete with the authors’ personal computers):

https://gist.github.com/coneill-math/c2f12c94c7ee12ac7652096329417b7d

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, \( n = 89 \) took 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 [15] [13], we were unable to locate any that consider \( \Delta_n \).

We now give some basic properties of the \( h \)-vector of \( \Delta_n \).

Proposition 4.12. Fix \( n \geq 1 \).

(a) \( h_{n,0} = 1 \).
(b) \( h_{n,1} = [n + 1]/2 - \tau(n) \), where \( \tau(n) \) denotes the number of divisors of \( n \).
(c) \( d_n = \deg h_n(x) = [n - 1]/2 - [n/3] \).

Proof. We proceed using Corollary 4.11 and characterizing the possible sets \( A \subseteq (0, n/2) \) 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, Theorem 4.13 implies \( d_n \leq [(n - 1)/2] - [n/3] \). 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). □

We conclude this section with the following bounds on the \( h \)-vector entries of \( \Delta_n \), which play a crucial role in establishing the final threshold function in Section 5 and estimating several expected values in Section 6.

Theorem 4.13. For any \( n \geq 1 \), \( i \geq 1 \), and \( N \geq 2 \), we have
\[
\sum_{j=2}^{N} \binom{\lceil (n - 1)/j \rceil}{i} \leq h_{n,i} \leq \binom{\lceil n/2 \rceil - 2i}{i}.
\]
Proof. By Corollary 4.10, \( h_{n,i} \) counts sets \( A \) minimally generating a semigroup not containing \( n \). We claim \( m = \min(A) \geq 2i \). 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 \( \max(A) \leq \lceil n/2 \rceil \), any two elements \( a, b \in A \) cannot satisfy \( a + b \equiv n \mod m \). The upper bound immediately follows.

For the lower bound, we claim that for any \( j \geq 2 \), 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 \( j \geq 2 \), the sum of any two elements of \( A \) exceeds \( n/j \), ensuring \( A \) minimally generates \( S \). This completes the proof.

Remark 4.14. Proposition 4.12 implies the lower bound in Theorem 4.13 is tight for \( i = 1 \). The given bounds are tighter than those for more general Cohen-Macaulay simplicial complexes [18] and are sufficient to prove the results in the coming sections, but still leave room for improvement.

Table 1 compares values from Theorem 4.13 with those computed in Remark 4.11.

| \( n = 89 \) | \( n = 90 \) | \( n = 500 \) |
|---|---|---|
| \( i \) | 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 \( \cdot \) 10^{18} | 4.4 \( \cdot \) 10^{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 \( \cdot \) 10^{20} | 6.4 \( \cdot \) 10^{37} |
| 13 | 105 | 153 | 27132 | 13 | 14 | 31 | 27132 | 70 | 5.2 \( \cdot \) 10^{14} | 1.6 \( \cdot \) 10^{30} |

Table 1. Comparison of the bounds in Theorem 4.13 with values computed in Remark 4.11.

5. Expected number of minimal generators

The main result of this section is Corollary 5.6 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 in the chosen semigroup (Definition 5.2) and the \( h \)-vector of the simplicial complex \( \Delta_n \) introduced in Section 4; see Remark 5.4.

Convention 5.1. 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 3.4 ensures that \( S \) is cofinite a.a.s.

Definition 5.2. Let \( S \sim S(M,p) \). For each integer \( n \in [M] \), denote by \( a_n(p) \) the probability that \( n \) cannot be written as a sum of elements in \( S \cap [n-1] \), that is,

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

Proposition 5.3. For each \( n \geq 1 \), \( a_n(p) = (1-p)^{\lfloor n/2 \rfloor} h_n(p) \).
Proof. The faces of $\Delta_n$ are precisely the sets $A \subset [n-1]$ satisfying $n \notin (A)$. As such,

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

follows upon unraveling Definitions 4.1 and 4.3.

Remark 5.4. By Proposition 5.3 we can write

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

in terms of the Hilbert series of the Stanley-Reisner ring $k[x]/I_{\Delta_n}$ (see [18, Chapter 2, Definition 1.1]). We encourage the interested reader to consult [18] for background on Hilbert functions.

Theorem 5.5. 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. By Proposition 5.3 and the linearity of expectation,

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

For any fixed integer $N \geq 2$, the lower bound in Theorem 4.13 gives

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

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 $k \geq 0$ and $1 \leq r \leq 2j(j+1)$, and supposing $M = m(2j(j+1))$ for some $m \in \mathbb{Z}_{>1}$, we obtain

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

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

$$= \sum_{k=1}^{m} p(1-p)^{kj(j+1)+O(1)} (1+p)^{2kj(j+1)-2kj+O(1)}$$

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

$$= \left( \sum_{r=1}^{2j(j+1)} (1-p)^{O(1)} (1+p)^{O(1)} \right) \sum_{k=1}^{m} \left( \frac{1 - ((1-p)^{j(j+1)} (1+p)^2)^m}{1 - (1-p)^{j(j+1)} (1+p)^2} \right).$$
Since \( p \gg 1/M \), a simple calculus exercise shows \((1 - p)^{\frac{j(j+1)}{2}}(1 + p)^2)^m \to 0\). If \( p \to 0 \), we obtain
\[
\mathbb{E}[e(S)] \geq \sum_{j=2}^{N} \sigma_j \to \sum_{j=2}^{N} \frac{2j(j+1)}{j(j+1)-2} \geq 2(N-1),
\]
which must hold for every \( N \geq 2 \).

\[\]

**Corollary 5.6.** Resuming notation from Theorem 5.5, if \( 1/M \ll p \ll 1 \), then
\[
\lim_{M \to \infty} \mathbb{E}[e(S)] = \lim_{M \to \infty} \mathbb{E}[g(S)] = \lim_{M \to \infty} \mathbb{E}[F(S)] = \infty.
\]

**Proof.** Apply Theorem 5.5 and the inequalities
\[\]
\[e(S) - 1 \leq \min(S \setminus \{0\}) \leq g(S) \leq F(S),\]
which hold for any cofinite numerical semigroup \( S \).

\[\]

6. **Approximations**

In the final section of this paper, we prove the only remaining case in Theorem 1.1, namely where \( p \) is bounded away from zero (Theorem 6.4). 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 4.13. Remark 6.5 discusses the accuracy of these estimates.

**Lemma 6.1.** 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 \( F(S) - a \not\in S \), 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.

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

**Lemma 6.2.** Let \( S \sim S(M,p) \) where \( p \in (0,1) \) is constant. Then,
\[
\lim_{M \to \infty} \mathbb{E}[g(S)] = \lim_{M \to \infty} \mathbb{E}[g_M(S)].
\]

*Proof.** We split the expectation \( \mathbb{E}[g(S)] \) into two parts:
\[
\mathbb{E}[g(S)] = \sum_{n=1}^{\infty} \mathbb{P}[n \in G(S)] = \sum_{n=1}^{M} \mathbb{P}[n \in G(S)] + \sum_{n>M} \mathbb{P}[n \in G(S)].
\]
The first term is just \( \mathbb{E}[g_M(S)] \). 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 \( S \), since
\[
F((m, m+1)) = m(m+1) - 2m - 1 \leq M.
\]
For each even integer \( m \leq \sqrt{M} \), let \( E_m \) denote the event that \( m \) and \( m-1 \) are not both in \( A \). 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,
\[
\mathbb{P}[n \in G(S)] \leq \mathbb{P}[E_2 \cap E_4 \cap \cdots] = (1 - p^2)^{\left\lfloor \sqrt{M} \right\rfloor},
\]
which implies the desired result.

\[\]
Proposition 6.3. For constant \( p \in (0, 1) \),
\[
\frac{6 - 8p + 3p^2}{2p - 2p^3 + p^4} \leq \lim_{M \to \infty} \sum_{n=1}^{M} a_n(p) \leq \frac{2 - p^2}{p^2}.
\]

Proof. We begin with the upper bound. Suppose \( M = 2m \) for some \( m \in \mathbb{Z}_{\geq 1} \). By Theorem 4.13,
\[
\sum_{n=1}^{M} a_n(p) \leq \sum_{n=1}^{M} (1 - p)^{\lceil n/2 \rceil} \sum_{i=0}^{d_n} \binom{\lceil n/2 \rceil}{i} p^i \leq \sum_{n=1}^{M} (1 - p)^{\lceil n/2 \rceil} \sum_{i=0}^{\lceil n/2 \rceil} \binom{n/2}{i} p^i
\]
\[=
\sum_{n=1}^{M} (1 - p^2)^{\lceil n/2 \rceil} = 1 + 2 \sum_{n=1}^{m} (1 - p^2)^n = 1 + 2(1 - p^2) \frac{1 - (1 - p^2)^m}{1 - (1 - p^2)},
\]
which yields the claimed upper bound since \( (1 - p^2)^m \to 0 \) as \( m \to \infty \). For the lower bound, writing \( n = 6k + r \) for \( r \in [6] \) and letting \( M = 6m \) for some \( m \in \mathbb{Z}_{\geq 1} \), we use Theorem 4.13 to obtain
\[
\sum_{n=1}^{M} a_n(p) \geq \sum_{n=1}^{M} (1 - p)^{\lceil n/2 \rceil} \sum_{i=0}^{d_n} \binom{\lceil n/2 \rceil}{i} p^i = \sum_{n=1}^{M} (1 - p)^{\lceil n/2 \rceil} (1 + p)^{d_n}
\]
\[=
(6 - 8p + 3p^2) \sum_{k=0}^{m} ((1 - p)^3(1 + p))^k,
\]
and since \( ((1 - p)^3(1 + p))^{m+1} \to 0 \) as \( m \to \infty \), the desired lower bound is obtained. \( \square \)

Theorem 6.4. Let \( S \sim S(M, p) \) where \( p \in (0, 1) \) is constant. Then
\[
\frac{6 - 8p + 3p^2}{2 - 2p^2 + p^4} \leq \lim_{M \to \infty} \mathbb{E}[e(S)] \leq \frac{2 - p^2}{p},
\]
\[
\frac{6 - 14p + 11p^2 - 3p^3}{2p - 2p^3 + p^4} \leq \lim_{M \to \infty} \mathbb{E}[g(S)] \leq \frac{(1 - p)(2 - p^2)}{p^2}, \text{ and}
\]
\[
\frac{6 - 14p + 11p^2 - 3p^3}{2p - 2p^3 + p^4} \leq \lim_{M \to \infty} \mathbb{E}[F(S)] \leq \frac{2(1-p)(2-p^2)}{p^2}.
\]

Proof. Lemma 6.1 implies \( g(S) \leq F(S) \leq 2g(S) \) for any cofinite numerical semigroup \( S \). Since
\[
\mathbb{E}[e(S)] = p \sum_{n=1}^{M} a_n(p) \quad \text{and} \quad \mathbb{E}[g_M(S)] = (1 - p) \sum_{n=1}^{M} a_n(p),
\]
each claimed inequality follows from Lemma 6.2 and Proposition 6.3. \( \square \)

Remark 6.5. Neither of the bounds in Proposition 6.3 use the full strength of Theorem 4.13. The summands for \( j \geq 3 \) in the lower bound yield increasingly complicated rational functions in \( p \), though the resulting sequence of values necessarily converges to 0 by Theorem 6.4. Additionally, substituting the upper bound for \( h_n \) given in Theorem 4.13 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 6.4. 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 4.13. Several computed values can be found in Table 2. For instance, the upper bound for \( \mathbb{E}[e(S)] \) in Theorem 6.4 could be
Table 2. Comparing estimates of $E[e(S)]$ using the bounds in Theorem 6.4, exact computation using the polynomials in Remark 4.11 and experimental evidence from 100,000 samples.

|     | Lower Experiments | Upper Experiments |
|-----|-------------------|--------------------|
| $p$ | Thm 6.4 $M = 25000$ | Thm 6.4 $M = 50000$ | Thm 6.4 $M = 90$ |
| 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       |

made better by simply noting that $e(S)$ is at most the smallest generator (which has expected value $1/p$). Such improvements to Theorem 4.13 should be possible, given the precise characterization of the $h$-vector of $\Delta_n$ in Corollary 4.10.

It is also worth noticing that the polynomials computed in Remark 4.11 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 1.1. Apply Theorems 3.4, Corollary 5.6 and Theorem 6.4. □

References

1. 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. MR 2739501
3. N. Alon and J. H. Spencer, The probabilistic method, 4th ed., Wiley, 2016.
4. V. I. Arnol’d, Weak asymptotics of the numbers of solutions of Diophantine equations, Funktsional. Anal. i Prilozhen. 33 (1999), no. 4, 65–66. MR 1746430
5. A. Assi and P. A. García-Sánchez, Numerical semigroups and applications, RSME Springer Series, vol. 1, Springer, [Cham], 2016. MR 3558713
6. V. Barucci, D. E Dobbs, and M. Fontana, Maximal properties in numerical semigroups and applications to one-dimensional analytically irreducible local domains, vol. 598, American Mathematical Soc., 1997.
7. J. Bourgain and Ya. G. Sinai, Limit behavior of large Frobenius numbers, Uspekhi Mat. Nauk 62 (2007), no. 4(376), 77–90. MR 2358737
8. W. C Brown and J. Herzog, One dimensional local rings of maximal and almost maximal length, J. Alg, 151 (1992), no. 2, 332–347.
9. A. Campillo, On saturations of curve singularities, Singularities 40 (1983), no. Part 1, 211.
10. J. De Loera, S. Petrović, L. Silverstein, D. Stasi, and D. Wilburne, Random monomial ideals, arXiv: 1701.07130 (2017).
11. P. Erdős and A. Rényi, On random graphs, I, Publ. Math. (Debrecen) 6 (1959), 290–297.
12. E. Kunz, The value-semigroup of a one-dimensional gorenstein ring, Proc. Amer. Math. Soc. 25 (1970), no. 4, 748–751.
13. 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. MR 2417395
14. J. C. Rosales and P. A. García Sánchez, Numerical semigroups, Developments in Mathematics, vol. 20, Springer, New York, 2009. MR 2549780
15. 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. MR 1978528
16. , *Fundamental gaps in numerical semigroups*, J. Pure Appl. Algebra, vol. 189, no. 1-3, 301–313. MR 2038577

17. J. D Sally, *On the associated graded ring of a local cohen-macaulay ring*, J. Math. Kyoto Univ., vol. 17, no. 1, 19–21.

18. R. P. Stanley, *Combinatorics and commutative algebra*, second ed., Progress in Mathematics, vol. 41, Birkhäuser Boston, Inc., Boston, MA, 1996. MR 1453579