Pseudorandom Graphs with Sub-random Algorithmic Entropy: Properties and Applications DRAFT

...

Written by Svetlana Pack
Advised by Professor Alex Iosevich
University of Rochester

April 10 2024

This represents a reasonable draft of my senior thesis, written to fulfill the upper-level writing requirement for a Honors degree in Mathematics at the University of Rochester. Research is still ongoing, but this document reviews our findings and objectives thus far.

1 Abstract

Of increasing interest in the era of big data, in which computational resources are increasingly at a premium, is the idea of 'complexity'- the amount of information you need to completely determine a set of data. Several competing ideas have emerged in the last century to formalize this notion. Among them is Kolmogorov complexity or algorithmic entropy, the length of the shortest program describing an object.

The high complexity of random objects have been leveraged in many algorithms. For instance, a class of recurrent neural networks called Echo State Networks (ESNs) use a complex 'reservoir' such as a randomly connected graph to capture the rich nonlinear dynamics of chaotic systems. However, even sparse random graphs have high algorithmic entropy and are accordingly resource intensive to utilize.

It is thus of interest to study 'pseudorandom' graphs, which share important spectral and connectivity properties with random graphs but are deterministically constructed and thus often have lower algorithmic entropy. In this paper we examine Paley graphs, a pseudorandom subset of Cayley graphs [4], and establish graphs satisfy the spectral conditions for pseudorandomness, discuss the resulting connectivity properties. However, we also present a novel proof that all Cayley graphs have at most loglinear algorithmic entropy, meaning Paley graphs exemplify pseudorandom graphs that nevertheless demonstrate subrandom algorithmic entropy.

Our immediate next aim is to prove that our pseudorandom gaphs graphs are capable of serving as effective reservoirs for an ESN, in terms of both memory capacity and empirically validatation of their performance on a chaotic time series forecasting task relative to a traditional erdos renyi- graph Echo State Network.

2 Introduction

One of the most contentious philosophical kerfuffles of the last century is that over the idea of 'randomness'; what does it mean for an object to be truly random? Is such a thing even possible? Likewise, what does it mean for something to be 'complex'?

These questions are not entirely the sort of navel-gazey philosophical dalliance perpetuated by students in the math lounge while they procrastinate their Probability Theory homework. Characterizing the complexity

of objects such as networks is of great interest for a myriad of applications, from simulating gene expression and interaction, to detecting epileptic seizures in electroencephalogram data [1].

In 1948, Claude Shannon catalysed the information theory revolution by introducing "information entropy" as a measure of the number of bits needed to specify an element from a discrete probability distribution. This allowed for people to talk about the information capacity and complexity of probability distributions in a mathematically rigorous way. About a decade later, Kolmogorov (concurrently with but independently of Chaitin and Solomonoff) defined of 'algorithmic entropy', using Turing's theory of computation to adapt the concept of entropy to individual realizations of data instead of larger statistical ensembles.

Alternatively called 'Kolmogorov Complexity' or 'Chaitin-Kolmogorov-Solomonoff complexity', the algorithmic entropy of an object can be intuitively thought of as the minimum length of a computer program needed to completely specify that object. For instance, a sequence of n bits generated from a Bernoulli distribution with probability parameter $p = \frac{1}{2}$ would have $O(n)$ algorithmic entropy in that it would take n bits to uniquely specify the sequence. Meanwhile, a sequence of n bits all equal to one would (on a machine that already knows the value of 'n') would have $O(1)$ algorithmic entropy as the program 'repeat the value '1' n times', which does not scale with n. (Note: if the machine was not already provided with the parameter n then we would need to encode the value of n, taking $O(log(n))$ bits and thus significantly increasing the complexity of our program). This is quite consistent with the $p = \frac{1}{2}$ Bernoulli distribution underlying our first sequence having maximal Shannon entropy compared to all other parameter p Bernoulli distributions, while constant sequences have zero entropy.

Algorithmic entropy has both pronounced strengths and weaknesses as a measure of an object's information content. On one hand, a result termed the Invariance Theorem establishes that - unlike alternate measures of entropy- algorithmic entropy is invariant under object representation up to constant overhead[2]. On the other, algorithmic entropy is cumbersome to work with in that it is not explicitly computable [9]. However, algorithmic complexity is upper semicomputable, and we can use lossless compression algorithms on data to derive upper bounds for their algorithmic entropy.

3 Preliminaries

In this section we summarize the notation and basic graph theory that we will use to define algorithmic entropy and pseudorandomness. We also introduce Cayley graphs, one of the main focuses of this paper.

3.1 Big O and little o notation

Let f, g be function $f : \mathbb{R}_+ \to \mathbb{R}_+$. Then if there exists $x_o \in \mathbb{R}_+, M \in \mathbb{R}_+$ such that $f(x) \leq Mg(x) \quad \forall x_o \leq x$ then we write $f(x) = O(g(x))$.

If $\forall \epsilon > 0$ there exists $x_o > 0$ such that $f(x) \leq \epsilon g(x) \quad \forall x \geq x_o$ then we write $f(X) = o(g(x))$. Note this is equivalent to $\lim_{n\to\infty} \frac{f(x)}{g(x)} = 0$

3.2 Graph Notation

Let $G = (V, E)$ denote a graph on vertex set V with $E \subset V \times V$.

We will restrict our attention to simple graphs, ie. graphs with no more than one edge between vertices, and graphs without self-loops. Then, all graphs satisfying $|V| = n$ G will have an $n \times n$ adjacency matrix representation A_G below, with $[A_G]_{ii} = 0 \; \forall i \in [n]$:

$$[A_G]_{ij} = \begin{cases} 1 \text{ if } (v_i, v_j) \in E \\ 0 \text{ if not} \end{cases}$$

Let $d : V \to [|V|]$ denote the degree function of graph $G = (V, E)$ defined $d(v_i) = |\{(v_i, v_j) \in E : v_j \in V\}$ $\forall v_i \in V$. Then, $G = (V, E)$ with $|V| = n$ has degree list representation $D_G = \{d(v_1), ..., d(v_n)\}$. When $d(v_i) = d \forall i \in [|V|]$ we call a graph G d-regular.

We assume G is an undirected graph and thus A_G symmetric unless explicitly specified otherwise. Then, by spectral theorem A_G has an orthogonal eigenvector basis $\{v_1, ..., v_n\}$ corresponding to eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_n\}$, listed in descending order.

Remark 3.1 (Trivial Eigenvalue) *Let A_G be the adjacency matrix of d regular undirected graph, and $u = (1, ..., 1)$ denote the $n \times 1$ all 1s vector.*

Observe $[A_G u]_i = \sum_{j=1}^{n} A_{ij} * 1 = d * 1 = \forall i \in [n]$, so $A_G u = du$. So, u is an eigenvector of A_G, and because all entries of A_G are in $\{0, 1\}$ it is clear the corresponding eigenvalue d is the largest possible eigenvalue.

Then, because $v_1 = u$ and $\lambda_1 = d$ for all d regular undirected graphs, we term λ_1 the trivial eigenvalue.

3.3 Graph Properties

Here we define properties of graphs and families of graphs, sequences of graphs $\{G_n = (V_n, E_n)\}$ that unless otherwise specified will have $|V_n| = n \quad \forall n \in \mathbb{Z}_+$, that will be referenced throughout this paper:

Definition 3.1 ((n, d, λ) graph) *A graph $G = (V, E)$ is called (n, d, λ) if it has degree d with $|V| = n$ and has and the greatest absolute value of its nontrivial eigenvalues is λ, i.e., $\lambda = max_{i \in \{2,...,n\}}\{|\lambda_i|\} = max(\lambda_2, |\lambda_n|)$. We refer to λ as the spectral gap pf G.*

Definition 3.2 (Dense Family of Graphs) *We call a family of simple undirected graphs $\{G_n = (V_n, E_n)\}$ dense if there exists a constant $\rho \in (0, 1)$ such that $\forall n \in \mathbb{Z}_+$ we have that $|E_n| = (\rho + o(1))\binom{n}{2}$. We say that $\{G_n\}$ has constant order edge density ρ.*

Definition 3.3 (Sparse Family of Graphs) *We call a family of simple undirected graphs $\{G_n = (V_n, E_n)\}$ sparse if we have $\lim_{n \to +\infty} \frac{|E_n|}{\binom{n}{2}} = 0$.*

3.4 Cayley Graphs

Cayley graphs are families of graphs constructed from algebraic groups. This imbues Cayley Graphs with algebraic structure that, as we shall see, makes them a rich and easy to work with source of pseudorandom graphs.

Definition 3.4 (Cayley Graphs) *Let S be a generating set of a group on set H with group operation $* : H \times H \to H$. Then the Cayley graph $C(H, S)$ a graph with vertex set $V = H$ and edge set $E = \{(u, v) | v = s * u, s \in S\}$.*

A useful and well-known property of Cayley graphs on additive groups $(\mathbb{Z}_n, +)$ is the equivalence of the eigenspectrum of the adjacency matrix A_C of a Cayley graph $C(\mathbb{Z}_n, S)$ and the Fourier spectrum of the indicator function, as well as the eigenvectors of A and the characters in the Fourier basis of $\mathbb{Z}_{\text{ı}}$. This property will be of great use in proving the pseudorandomness of Paley graphs.

Lemma 3.1 *Let $C({\mathbb{Z}_p}^n, S)$ be a Cayley graph on the group $(\mathbb{Z}_p, +)$ for prime $p \in \mathbb{Z}$ with generating set $S \subset {\mathbb{Z}_p}^n$ and adjacency matrix A, and let $\chi_S : \mathbb{Z}_n \to \{0, 1\}$ be the indicator function of S . Let $\{\lambda_1 \geq \lambda_2 \geq ..., \geq \lambda_n\}$ be the eigenspectrum of the adjacency matrix of G (with multiplicity), $v_1, v_2, ..., v_n$ be the associated eigenvectors. Let $\{\{\omega(x)^k\}_{k \in \{0,...,p-1\}}$ (where $\omega(x) = e^{i\frac{2\pi x}{p}}$) be the Fourier basis of $\mathbb{Z}_p$ and $\{\hat{\chi_S}(k) : k \in \{0, ..., p-1\}\}$ are the associated Fourier characters. Then, $\lambda_k = \hat{\chi_S}(k)$ and the associated eigenvector $v_k = \omega^k : \mathbb{Z}_p \to \mathbb{R}$ $\forall k \in \mathbb{Z}_p$*

PROOF OF 3.1:

We want to show that for any $k \in \mathbb{Z}_p$ we have $\lambda_{k+1}\omega^k = A\omega^k$ Observe that

$$[A\omega^k]_l = \sum_{j=0}^{p-1} A_{lj}\omega(j)^k = \sum_{j=0}^{p-1} \chi_S(l-j)(\omega(j)^k)$$

χ_S has Fourier expansion $\chi_S(j) = \frac{1}{|\mathbb{Z}_p|}\sum_{m=0}^{p-1} \hat{\chi_S}(m)(\omega(j))^m = \frac{1}{p}\sum_{m=0}^{p-1} \hat{\chi_S}(m)(\omega(j))^m$. Thus,

$$\sum_{j=0}^{p-1} \chi_S(l-j)(\omega(j))^k = \sum_{j=0}^{p-1} \frac{1}{p}(\sum_{m=0}^{p-1} \hat{\chi_S}(m)(\omega(l-j))^m)(\omega(j))^k = \sum_{j=0}^{p-1} \frac{1}{p}(\sum_{m=0}^{p-1} \hat{\chi_S}(m)((\omega(l-j))^m(\omega(j))^k))$$

$$= \frac{1}{p}\sum_{j=0}^{p-1}(\sum_{m=0}^{p-1} \hat{\chi_S}(m)((e^{\frac{i2\pi(l-j)m}{p}})(e^{\frac{i2\pi(j)k}{p}})) = \frac{1}{p}\sum_{j=0}^{p-1}(\sum_{m=0}^{p-1} \hat{\chi_S}(m)(e^{\frac{i2\pi(l+j(k-m))}{p}})$$

$$= \frac{1}{p}\sum_{m=0}^{p-1} e^{\frac{i2\pi lm}{p}}(\hat{\chi_S}(m)\sum_{j=0}^{p-1} e^{\frac{i2\pi(j(k-m))}{p}})$$

By orthogonality of the Fourier basis $\sum_{j=0}^{p-1} e^{\frac{i2\pi(j(k-m))}{p}} = 0$ when $(k-m) \neq 0$, so

$$= \frac{1}{p}\sum_{m=0}^{p-1} e^{\frac{i2\pi lm}{p}}(\hat{\chi_S}(m)\sum_{j=0}^{p-1} e^{\frac{i2\pi(j(k-m))}{p}}) = \frac{1}{p}e^{\frac{i2\pi lk}{p}}(\hat{\chi_S}(k))\sum_{j=0}^{p-1} 1 = \hat{\chi_S}(k)e^{\frac{i2\pi lk}{p}} = \hat{\chi_S}(k)(\omega(l))^k = \hat{\chi_S}(k)[\omega^k]_l$$

Thus, $[A\omega^k]_l = \hat{\chi_S}(k)[\omega^k]_l \ \forall k \in \mathbb{Z}_p$. QED

One widely used example of a Cayley graph on the additive group $(\mathbb{Z}_n, +)$ is the cycle graph.

Definition 3.5 (Cycle graph) *A n vertex cycle graph is the Cayley graph on the additive group $(\mathbb{Z}_n, +)$ with generating set $\{\pm 1\}$. The cycle graph on $(\mathbb{Z}_{13}, +)$ is depicted in figure 1.*

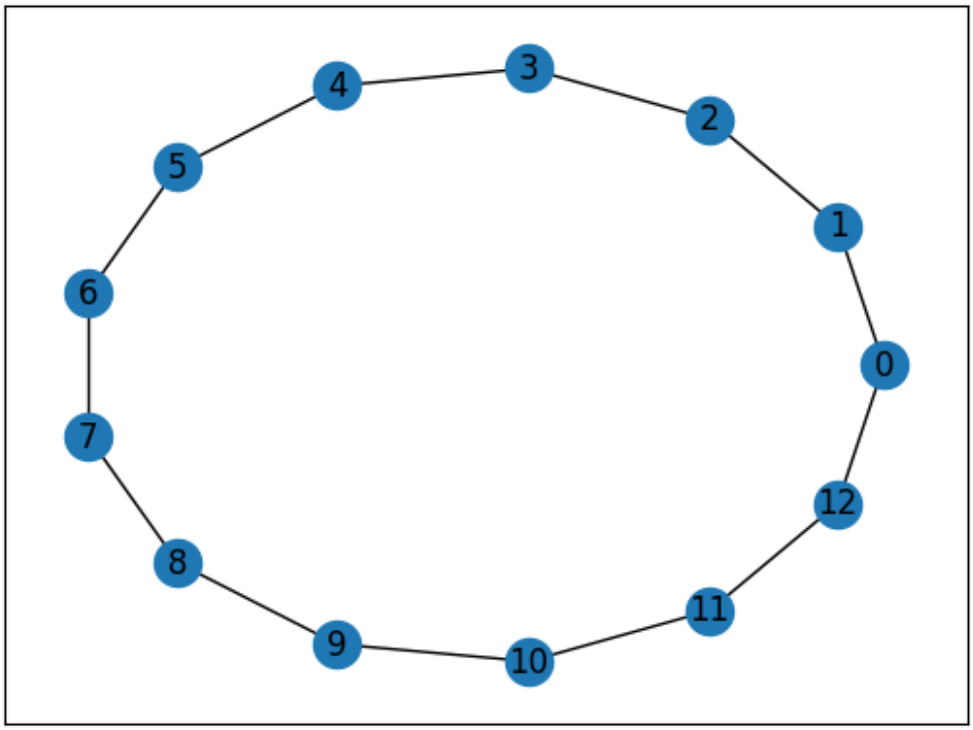

Figure 1: The Cycle graph $C((\mathbb{Z}_{13}, +), U = \{\pm 1\})$

Using Lemma 3.1 we can calculate the eigenspectrum of a Cayley graph on $(\mathbb{Z}_n, +)$ as $\lambda_k = e^{\frac{i2\pi k}{n}} + e^{-\frac{i2\pi k}{n}} = 2cos(\frac{2\pi k}{n}) \quad \forall k \in [n]$. We note the trivial eigenvalue $\lambda_1 = 2cos(\frac{2\pi 0}{n}) = 2(1) = |\{\pm 1\}|$, consistent with cycle graphs having two generators and thus being 2-regular.

Remark 3.2 (Cycle Graphs) *We claim Cycle graphs are the sparsest possible connected Cayley graphs on $(\mathbb{Z}_p, +)$ up to isomorphism when p prime, with $d = 2$.*

It is trivial to observe that cycle graphs are indeed connected $\forall n \in \mathbb{Z}_+$ as $\{\pm 1\}$ generates $(\mathbb{Z}_+, +)$.

We note that the bipartite ladder graph is the Cayley graph on $(\mathbb{Z}_2 n, +)$ for $n \in \mathbb{Z}_+$ with generating set $\{\pm\frac{2n}{2}\} = \{\pm n\} = \{2n - n, n\} = \{n\}$ is the sparsest possible undirected Cayley graph, as $d = 1$. But the bipartite ladder graph is not connected for $n > 1$ as $\{n\}$ has order 2 in $(\mathbb{Z}_{2n}, +)$ as $n + n = 2n \equiv 0$ and so n does not by itself generate $(\mathbb{Z}_{2n}, +)$. Thus the resulting Cayley graph is not connected.

Then, we show that cycle graphs are the sparsest possible undirected Cayley graph on $(\mathbb{Z}_p, +)$ up to isomorphism by showing all undirected $d = 2$ Cayley graphs on $\mathbb{Z}_p$ are isomorphic to the cycle graph when p prime.

For a Cayley graph on $(\mathbb{Z}_p, +)$ to be undirected and degree 2 we must have the generating set take form $\{\pm a\}$ for some $a \in \mathbb{Z}_p$, as undirected Cayley graphs must have generating sets closed under inverses to ensure edges are symmetric.

Cayley graphs on $(\mathbb{Z}_p, +)$ with p prime and generating set $\{\pm a\}$ for $a \in \mathbb{Z}_p - \{0\}$ are identical to a cycle graph up to isomorphism. This is because when p is prime, $a \neq ka$ in $(\mathbb{Z}_p, +)$ for any $k \in \mathbb{Z}_p - \{0, 1\}$, so $\{\pm a\}$ generates $(\mathbb{Z}_p, +) \quad \forall a \in \mathbb{Z}_p - \{0\}$.

Then, if C is the cycle graph of on $(\mathbb{Z}_p, +)$ and G is the Cayley graph of on $(\mathbb{Z}_p, +)$ with p prime and generating set $\{\pm a\}$, we observe that C is isomorphic to G under the vertex relabelling $l : (\mathbb{Z}_p, +) \to (\mathbb{Z}_p, +)$ defined $l(i) = ai$.

Thus we can say that on graphs of prime vertex size, cycle graphs are the sparsest connected Cayley graphs up to isomorphism.

4 Background

4.1 Algorithmic Entropy

The algorithmic entropy of an object, the length of the shortest program describing the object, is formally defined thusly;

Definition 4.1 (Algorithmic Entropy) *Let s be a string of finite length defined on a finite alphabet, T be a universal Turing machine, and P be any program for which T returns s upon halting. The algorithmic entropy of s is defined*

$$K_T(s) = min\{|P| : T(P) = s\}$$

The Invariance Theorem is a well known result establishing that the Kolmogorov complexity of a string does not depend on the representation chosen, so we can fix any arbitrary Universal Turing Machine T for our purposes.

4.1.1 Lempel Ziv Complexity

Kolmogorov complexity is not computable, but is upper-semicomputable [7]. One may obtain upper bounds for the Kolmogorov complexity by applying tools such as a lossless compression algorithm to the string under consideration and calculating the size of the resulting compression [9]. One such algorithm we will make use of is the $LZ76$ algorithm to calculate the Lempel Ziv complexity of a string. Intuitively we can think of Lempel-Ziv complexity as measuring the number of unique non- repeating substrings in a string, formally described in the definition below [10]:

Definition 4.2 *Let $s = \{s_i\}_{i\in[n]}$ with $s_i \in \{0, 1\} \quad \forall i \in [n]$ be a binary string of length $n \in \mathbb{Z}_+$. Let ${s_j}^k$ denote the substring $\{s_i\}_{i=j}^k$ of s starting at index j and ending at index k, for some $1 \leq j \leq k \leq n$*

Use the LZ76 algorithm to recursively partitioning s into some number $p \leq n$ of disjoint substrings, called 'blocks' $\{B_j\}_{j\in[p]}$, each representing the shortest substring that is not contained anywhere in the substring preceding it. The LZ76 Algorithm is formally described in Algorithm 1.

We then define the Lempel-Ziv complexity $LZ(s)$ of s as the number $LZ(s) := p = |\{B_j\}_{j\in[p]}|$ [10].

Algorithm 1 LZ76 Compression

Require: binary string of length $n \in \mathbb{Z}_+$, $s = \{s_i\}_{i\in[n]}$ with $s_i \in \{0,1\} \quad \forall i \in [n]$

Define the starting block of our encoding $B_1 \leftarrow x_1^1$.

Suppose after k steps we have encoded the first n_k bits of s as $B_1, B_2, ..., B_k = x_1^1, x_2^{n_2}, ..., x_{n_{k-1}+1}^{n_k}$

...

Set the following parameters:

uniqueSubstring $\leftarrow FALSE$

$n_{k+1} \leftarrow n_k + 1$

...

Let 'FindSubstring(pattern, string)' be a method implementing convolution-based string pattern matching that returns TRUE when pattern is a substring of string, FALSE otherwise.

while not(uniqueSubstring) **do**

if FindSubstring($x^{n_{k+1}}{}_{n_k+1}, x^{n_k}{}_1$) returns $FALSE$ **then**

$B_{k+1} \leftarrow x^{n_{k+1}}{}_{n_k+1}$

uniqueSubstring $\leftarrow TRUE$

else

$n_{k+1} \leftarrow n_{k+1} + 1$

end if

end while

Repeat until $n_k \geq n$

—

return $\{B_j\}_{j\in k}$

4.1.2 LZ76 Compression Algorithm Example Calculation

To illustrate the LZ76 algorithm, Algorithm 1, we perform an example computation on, say, the indicator function $\chi_U : \mathbb{Z}_{13} \to \{0,1\}$ of the set of nonzero quadratic residues of $(\mathbb{Z}_{13}, +)$, $U = \{u^2 mod 13 : u \in \mathbb{Z}_{13}$. This example is chosen strategically; U is the generating set of the Paley graph G on $(\mathbb{Z}_{13}, +)$, one of the archetypal examples of a dense pseudorandom graph. Accordingly, χ_U is the first row of the adjacency matrix of G, as this row corresponds to the indicator function of elements g of $\mathbb{Z}_{13}$ such that $0 + g \in U$. This kind of computation will turn out to be pivotal for our later proof that Cayley graphs general have at most loglinear Kolmogorov Complexity.

$(\mathbb{Z}_{13}, +)$ contains six nonzero quadratic residues, $U = \{1^2 \text{mod} 13 = 1,\ 2^2 \text{mod} 13 = 4,\ 3^2 \text{mod} 13 = 9,\ 4^2 \text{mod} 13 = 3,\ 5^2 \text{mod} 13 = 12,\ 6^2 \text{mod} 13 = 10\}$. Note that 13 is prime and $13 \equiv 1 \text{mod} 4$, so U is closed under additive inverses as $-1 \equiv 12 \text{mod} 13 \in U$ and the set of nonzero quadratic residues form an equivalence relation on the multiplicative group $(\mathbb{Z}_{13} - \{0\}, \cdot)$.

Writing χ_U as an n dimensional vector indexed by $\mathbb{Z}_{13}$ we see that

$$\begin{bmatrix} Indices: & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \chi_U: & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Then the first block $B_1 = x_1^1$ is the length 1 string containing the fist bit

$\chi_U = \begin{bmatrix} 0| & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$

B_2 is the shortest substring of $s = \chi_U$, starting at the index right after the previous block B_1 ended (so starting at 2), that does not occur as a substring of s_1^1. This turns out to be $B_2 = s_2^2 = \{1\}$. Thus,

$\chi_U = \begin{bmatrix} 0| & 1| & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$

Likewise B_3 is the shortest substring of $s = \chi_U$, starting at the index right after the previous block B_2 ended (so starting at 3), that does not occur as a substring of the substring of s from 1 to the ending index of the previous block s_1^2. This turns out to be $B_3 = s_3^4 = \{01\}$. Thus,

$\chi_U = \begin{bmatrix} 0| & 1| & 0 & 1| & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$

Proceeding like this up until the end of s yields the following partition into blocks:

$\chi_U = \begin{bmatrix} 0| & 1| & 0 & 1| & 1 & 0| & 0 & 0| & 0 & 1 & 1| & 0 & 1 \end{bmatrix}$

$\{B_1 = \{0\}, B_2 = \{1\}, B_3 = \{01\}, B_4 = \{10\}, B_5 = \{00\}, B_6 = \{011\}, B_7 = \{01\}\}$.

We conclude that χ_U by itself has a Lempel -Ziv complexity of seven.

4.2 Pseudorandomness and Randomness

4.2.1 Pseudorandom Graphs

'Pseudorandomness' can refer to a myriad of graph theoretic properties that random graphs provably have with high probability. One of the first groups to describe pseudo- or quasi-random graphs were Chung, Graham, and Wilson, who originally enumerated a list of properties random graphs have with high probability and proved they are equivalent for dense graphs in 1989 [11]. We will focus on two of pseudorandomness properties: discrepancy pseudorandomness, which concerns the connectivity of a graph, and spectral pseudorandomness, which concerns the eigenspectrum of the graph.

We will first develop the notion of pseudorandomness for families of graphs with dense connections from Chung, Graham, and Wilson's work, and then dense discrepancy and spectral pseudorandomness properties are equivalent for d regular graphs (although it is worth noting Chung et' al.'s proof works for any simple undirected graph)[11][8].

We will then adapt the discrepancy and spectral pseudorandomness conditions to sparse graphs using later work by Conlon, Fox, and Zhao in 2014, and then summarize their proof that sparse discrepancy and spectral pseudorandomness properties are equivalent for sparse Cayley graphs [12][8].

Definition 4.3 (Dense Discrepancy Pseudorandomness) *Let $\{G_n = (V_n, E_n)\}$ be a family of dense graphs with constant order edge density ρ. Then $\{G_n\}$ satisfies the discrepancy pseudorandomness property if for vertex subsets $S, T \subset V_n$ we have that $|e(S,T) - \rho|S||T|| = o(n^2)$*

The discrepancy pseudorandomness condition is motivated by the behavior of binomial random graphs. In a a binomial random graph $G = (V, E)$ with probability parameter $\rho \in (0, 1)$ we have that $\rho|S||T|$ is the expected number of edges between disjoint subsets S and T of the vertex set V. So intuitively, a graph with edge density ρ having a relatively small discrepancy $|e(S,T) - \rho|S||T||$ for all subsets S, T of V, including disjoint subsets, behaves similarly to a binomial random graph with probability ρ.

Definition 4.4 (Dense Spectral Pseudorandomness) *Let $\{G_n = (V_n, E_n)\}$ be a family of dense graphs. Then $\{G_n\}$ satisfies the spectral pseudorandomness property if the eigenvalues of the adjacency matrix of G_n, listed $\{\lambda_i\}_{i\in[n]}$ in descending order, satisfy $\lambda_1 = \rho n + o(n)$ and $\lambda = max_{i\in\{2,\dots,n\}}\{|\lambda_i|\} = max(\lambda_2, |\lambda_n|) = o(n)$*

Theorem 4.1 (Equivalence of Discrepancy and Spectral Pseudorandomness for Dense Graphs) *Let $\{G_n = (V_n, E_n)\}$ be a family of dense d-regular graphs with constant order edge density ρ. Then $\{G_n\}$ satisfies the discrepancy pseudorandomness condition if and only if it satisfies the spectral pseudorandomness condition. First proved in [11]*

To prove this theorem, we first need to prove the Expander-Mixing Lemma, an important result relating the discrepancy of a d regular graph and spectral gap (i.e., the maximal absolute value λ of the nontrivial eigenvalues of the adjacency matrix) [4]. The essence of the lemma is that if spectral gap of a graph is small, then the graph will have small discrepancy and so be quite well connected relative to the density of its edges, similar to what you would expect from a binomial random graph of the same edge density.

Lemma 4.2 (Expander Mixing Lemma) *For any $(n, d, \lambda$ graph $G = (V, E)$, if $S \subset V$ and $T \subset V$ such that $S \cap T = \emptyset$ and $e(S,T)$ denotes the set of edges with one vertex in S and one in T, then*

$$|e(S,T) - \frac{d}{n}|S||T|| \leq \lambda\sqrt{|S||T|}$$

PROOF OF 4.2: Let $(n, d, \lambda$ graph $G = (V, E)$ have adjacency matrix A. Let us write $V = [n]$ by enumerating the n vertices of V. G is a simple, undirected graph so A is a symmetric matrix, thus by spectral theorem A has an orthonormal eigenvector basis $\{v_i\}_{i\in[n]}$ associated to it's eigenvalues $\{\lambda_i\}_{i\in[n]}$ (written in decreasing order with multiplicity). As observed in remark 3.1, G degree d implies that $\lambda_1 = d$ and the associated eigenvector $v_i = \frac{1}{\sqrt{n}}(1, ..., 1)$, the normalized all ones $nx1$ vector.

Let $S, T \subset V = [n]$ and let $\chi_S : [n] \to \{0, 1\}$ and $\chi_T : [n] \to \{0, 1\}$ be the indicator functions of S and T respectively. Note that we can write χ_S as a length n vector in $\{0, 1\}^n$ where $chi_{Si} = 1 \iff i \in S \subset [n]$ (likewise for chi_T).

Now, we compute $e(S.T)$ from A and the indicator functions of S and T. We have that $A\chi_{T_i} = \sum_{j\in[n]} A_{ij}\chi_{Tj}$, so ${\chi_S}^\top A\chi_T = \sum_{i\in[n]}\sum_{j\in[n]} A_{ij}\chi_{Si}\chi_{Tj}$. Observer that $\forall i,j \in [n]$ we have that $A_{ij}\chi_{Si}\chi_{Tj} = 1 \iff (i,j) \in E, i \in S$, and $j \in T$, and $A_{ij}\chi_{Si}\chi_{Tj} = 0$ otherwise. Then,

$$ {\chi_S}^\top A\chi_T = \sum_{i\in[n]}\sum_{j\in[n]} A_{ij}\chi_{Si}\chi_{Tj} = e(S,T) $$

.

Our orthonormal eigenbasis $\{v_i\}_{i\in[n]}$ of A spans $\mathbb{R}^n$, so we can write χ_S and χ_T in terms of $\{v_i\}_{i\in[n]}$: $\chi_S = \sum_{i\in[n]} a_i v_i$ and $\chi_T = \sum_{i\in[n]} b_i v_i$ where $a_i := \chi_S^\top v_i$ and $b_i := \chi_T^\top v_i$.

Thus, $A\chi_T = A(\sum_{i\in[n]} b_i v_i) = \sum_{i\in[n]} b_i(Av_i) = \sum_{i\in[n]} b_i(\lambda_i v_i) = d(b_1)(v_1) + \sum_{i\in\{2,\dots,n\}} b_i(\lambda_i v_i)$.

By the orthonormality of $\{v_i\}_{i\in[n]}$ we have that $v_i^\top v_j = 1 \forall i = j$ and $v_i^\top v_j = 0 \forall i \neq j$. Therefore,

$$ e(S,T) = {\chi_S}^\top A\chi_T = \sum_{i\in[n]}\sum_{j\in[n]} \lambda_j a_i b_j (v_i^\top v_j) = \sum_{i\in[n]} \lambda_i a_i b_i = \sum_{i\in[n]} \lambda_i(\chi_S^\top v_i)(\chi_T^\top v_i) $$

Observing that $\lambda_1(\chi_S^\top v_1)(\chi_T^\top v_1) = d(\sum i \in [n] \frac{1}{\sqrt{(n)}} * \chi_S(i))(\sum_{j\in[n]} \frac{1}{\sqrt{(n)}} * \chi_T(j)) = \frac{d}{n}|S||T|$. Therfore we have that $e(S,T) = \frac{d}{n}|S||T| + \sum_{i\in\{2,\dots,n\}\lambda_i(\chi_S^\top v_i)(\chi_T^\top v_i)}$ and by Cauchy Schwarz we obtain,

$$ |e(S,T) - \frac{d}{n}|S||T|| = |\sum_{i\in\{2,\dots,n\}} \lambda_i(\chi_S^\top v_i)(\chi_T^\top v_i)| \leq \lambda \left[\sum_{i\in[n]} (\chi_S^\top v_i)^2\right]^{\frac{1}{2}} \left[\sum_{i\in[n]} (\chi_T^\top v_i)^2\right]^{\frac{1}{2}} $$

Let B be a matrix with column vectors equal to the vectors in the orthonormal eigenbasis $\{v_i\}_{i\in[n]}$. Then we have that $\chi_S^\top B_i = {\chi_S}^\top v_i \quad \forall i \in [n]$. Thus, we have that $\sum_{i\in[n]}(\chi_S^\top v_i)^2 = \sum_{i\in[n]} \chi_S^\top B_i^2 = (\chi_S^\top B)(\chi_S^\top B)^\top = \chi_S^\top B B^\top \chi_S$. But B is an orthonormal matrix so $BB^\top = I$, thus $\sum_{i\in[n]}(\chi_S^\top v_i)^2 = \chi_S^\top\chi_S = \sum_{i\in[n]}(\chi_{Si})^2$. Further, by definition of the characteristic function $\sum_{i\in[n]}(\chi_{Si})^2 = \sum_{i\in S\subset[n]} 1^2 = \sum_{i\in S\subset[n]} 1^2 = |S|$. By identical argument, $\sum_{i\in[n]}(\chi_T^\top v_i)^2 = \sum_{i\in[n]}(\chi_{Ti})^2 = |T|$. Then we obtain

$$ |e(S,T) - \frac{d}{n}|S||T|| \leq \lambda \left[\sum_{i\in[n]} (\chi_S^\top v_i)^2\right]^{\frac{1}{2}} \left[\sum_{i\in[n]} (\chi_T^\top v_i)^2\right]^{\frac{1}{2}} = \lambda\sqrt{|S||T|} $$

QED

With the expander mixing lemma in hand, we can now prove the dense pseudorandom equivalence theorem for d regular graphs.

PROOF OF 4.1:

First observe that in a family $\{G_n\}_{n\in\mathbb{Z}_+}$ of d- regular graphs have edge density $\rho = \frac{d}{n}$. This is because $|E_n| = \frac{dn}{2}$ (d edges multiplied by n vertices, dividing by 2 so the undirected edges aren't double-counted) and so $|E_n| = \frac{nd}{2} = (\binom{n}{2})(\frac{d}{n} + \frac{d}{n(n-1)})$. $\lim_{n\to\infty} \frac{d}{n(n-1)} = 0$ so $\frac{d}{n(n-1)} = o(1)$ and we have $|E_n| = \binom{n}{2}(\frac{d}{n} + o(1))$.

Thus, $\lambda_1 = d = \frac{dn}{n} = \rho n + 0$ so to show a d-regular family of graphs satisfies spectral pseudorandomness it is sufficient to prove that the spectral gap of$\{G_n\}_{n\in\mathbb{Z}_+}$ grows sub-linearly, i.e., $\lambda = o(n)$.

Dense spectral pseudorandomness $\implies$ dense discrepancy pseudorandomness: Suppose $\lambda = o(n)$. Then by the Expander- Mixing Lemma we have that $\forall S,T \subset V_n$ we have $|e(S,T) - \frac{d}{n}|S||T|| = |e(S,T) - \rho|S||T|| \leq \lambda\sqrt{|S||T|} \leq \lambda\sqrt{n*n} = \lambda n = o(n^2)$.

QED

Dense discrepancy pseudorandomness $\implies$ dense spectral pseudorandomness: Proof omitted, not relevant to any work we have done so far.

Next, we define sparse pseudorandomness.

Definition 4.5 (Sparse ϵ- Discrepancy Pseudorandomness) *Let $\{G_n = (V_n, E_n)\}$ be a family of sparse d-regular graphs (i,e., $d = o(n^2)$) and let $\epsilon > 0$. Then $\{G_n\}$ has the ϵ- discrepancy pseudorandomness property if for vertex subsets $S,T \subset V_n$ we have that $|e(S,T) - \rho|S||T|| \leq \epsilon dn$*

Definition 4.6 (Sparse ϵ- Spectral Pseudorandomness) *Let $\{G_n = (V_n, E_n)\}$ be a family of sparse d-regular graphs (i,e., $d = o(n^2)$) and let $\epsilon > 0$. Then $\{G_n\}$ has the ϵ- spectral pseudorandomness property if the eigenvalues of the adjacency matrix of G_n, listed $\{\lambda_i\}_{i\in[n]}$ in descending order, satisfy $\lambda_1 = \rho n + o(n)$ and $\lambda = max_{i\in\{2,\dots,n\}}\{|\lambda_i|\} = max(\lambda_2, |\lambda_n|) \leq \epsilon d$*

Theorem 4.3 *Let $\{C_n\}$ be a family of Cayley graphs on the group $(\Gamma_n, \star)$ of size $|\Gamma_n| = n$ and with group operation $\star : \Gamma_n \times \Gamma_n \to \Gamma_n$, with generating set $U \subset \Gamma_n$ of size $|U| = d$ that is closed under inverses. Then, C_n is undirected and d regular. Let $\{C_n\}$ be sparse, i.e., $\{C_n\}$ has edge density $lim_{n\to\infty}\frac{|E_n|}{\binom{n}{2}} = lim_{n\to\infty}\frac{d}{n} = 0$.*

Then, $\{C_n\}$ has the ϵ - - spectral pseudorandomness property $\implies$ $\{C_n\}$ has the ϵ- discrepancy pseudo randomness property, and $\{C_n\}$ has the ϵ - discrepancy pseudorandomness property $\implies$ $\{C_n\}$ has the 8ϵ - spectral pseudorandomness property. First Proved in [12].

PROOF OF 4.3:

ϵ- spectral pseudorandomness $\implies$ ϵ- discrepancy pseudorandomness: Similar to the previous proof, this result follows almost immediately from the Expander Mixing Lemma.

Suppose $\lambda \leq \epsilon d$. Then by the Expander- Mixing Lemma we have that $\forall S, T \subset V_n$ we have $|e(S,T) - \frac{d}{n}|S||T|| = |e(S,T) - \rho|S||T|| \leq \lambda\sqrt{|S||T|} \leq \lambda\sqrt{n * n} = \lambda n \leq \epsilon dn$.

QED

ϵ- discrepancy pseudorandomness $\implies$ ϵ- spectral pseudorandomness: Proof omitted, not relecant to anything we have done so far.

4.2.2 Random Graphs

For the purposes of this research we will adopt the definition of a truly random graph G, employed in other literature on graph algorithmic entropy, as one that is not compressible, in the sense that there is no lossless encoding of the graph with fewer bits than its adjacency matrix representation (which has $\binom{n}{2}$ bits when the graph G is undirected and $n(n-1)$ bits when G is directed).

Definition 4.7 (Random Graphs) *A graph G is random if $K(G) = O(n^2)$*

This definition is meaningful in the sense that it has been proven that the probability of uniformly randomly choosing a binary string s of length l out of the set of all such binary stings $\{0,1\}^l$ such that s has strictly less than $O(l)$ algorithmic entropy, i.e., $K(s) = o(l)$, converges to 0 in probability as $l \to \infty$ [9]. We formally state and prove this in the following lemma:

Lemma 4.4 *For any $l \in \mathbb{Z}_+$ and any constant $c \in [l]$ we have that there are $2^l - 2^{l-c}$ binary strings of length l $s \in \{0,1\}^l$ such that $K(s) \geq l - c$ [9].*

It follows that the probability of uniformly randomly choosing s from $\{0,1\}^l$ such that $K(s) < l - c$ is at most $\frac{1}{2^c}$ $\quad \forall c \in [l]$.

PROOF OF 4.4:

Consider the set of binary strings of length $l, \{0,1\}^l$. Then for any constant $c \in [l]$, there are exactly $|\{0,1\}^{l-c}| = 2^{l-c}$ binary strings of length $l - c$, so there are only 2^{l-c} ways to compress a string of length l by c bits. Then there must be at least $2^l - 2^{l-c}$ binary strings of length l that do not compress by any more than c bits, i.e, $|\{s \in \{0,1\}^l : K(s) \geq l - c\}| \geq 2^l - 2^{l-c}$.

Then if we chose s form $\{0,1\}^l$ with uniform probability, $\forall c \in [n]$ we have that $K(s) \geq l - c$ with probability at least $\frac{2^{l-c}(2^c-1)}{2^l} = \frac{2^c-1}{2^c}$, and $K(s) < l - c$ with probability at most $\frac{1}{2^c}$. QED

We can represent undirected graphs as binary strings length $\binom{n}{2})$, most of which have algorithmic entropy $\binom{O(n}{2)=O(n^2)}$ by Lemma 4.4. Thus graphs of subrandom Kolmogorov complexity are quite rare for large n.

Our goal will be to demonstrate that the pseudorandom graphs under consideration nevertheless have strictly less than quadratic Algorithmic entropy, making them candidates for lower-complexity replacements for random graphs.

5 Results

5.1 Cayley Graphs

As we established in the preliminary section, Cayley graphs have highly desirable algebraic structure that imparts a variety of useful properties, such as an adjacency matrix with an eigenspectrum consisting of the graph's discrete Fourier components (Theorem with Theorem 3.1). In this section we use this property to prove a variety of both dense and sparse Cayley graphs display spectral pseudorandomness characteristics (and thus discrepancy pseudorandomness) [8].

Intuitively, we would expect Cayley graphs to have sub-random algorithmic complexity, as they are determined entirely by the group $(H, *)$ and the generating set $U \subset H$ they are defined on, which should not take at most a linear in n number of bits to encode (possibly with algorithmic overhead to keep track of labels). We present a (to my knowledge) novel proof that this is indeed the case for Paley graphs on the additive group $(\mathbb{Z}_n, +)$ for any $n \in \mathbb{Z}_+$. This gives us a rich repository of easy to construct and work with graphs with both proven pseudorandom properties and sub-random algorithmic entropy.

5.1.1 Paley Graphs

Consider the Cayley graph $C((\mathbb{Z}_p, +), U)$ for p prime with generating set $U = \{u \in \mathbb{Z}_{p-\{0\}} : u \equiv a^2 mod(p)$ for some $a \in \mathbb{Z}$. Then, $(u, v) \in E \iff v = s + u$ where s is a nonzero quadratic residue in $\mathbb{Z}_p$. We will restrict p such that $p \equiv 1 mod 4$, so that u square in $\mathbb{Z}_p \iff -u$ square in $\mathbb{Z}_p$ and thus, $v = s + u \implies u = -s + v$ and s quadratic residue $\iff -s$ quadratic residue give us that $(u, v) \in E \iff (v, u) \in E$ (ie. G undirected). Fig. 2 depicts the Paley graph $C((\mathbb{Z}_{13}, +), S)$

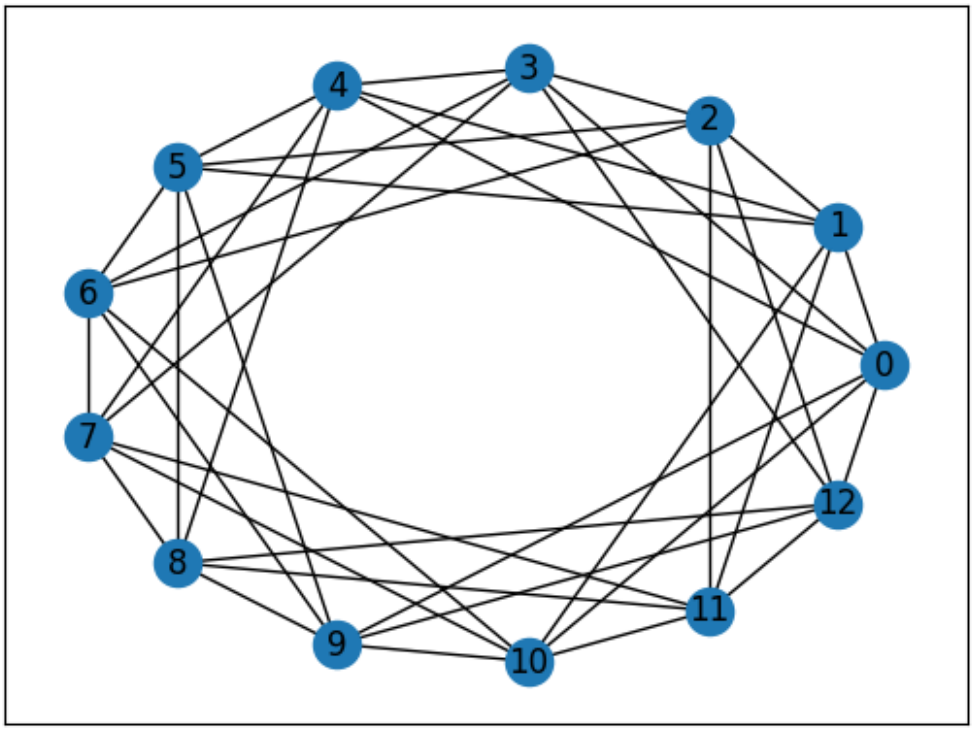

Figure 2: The Paley graph $C((\mathbb{Z}_{13}, +), U = \{a^2 : a \in \mathbb{Z}_{13}\})$

Let $C(\mathbb{Z}_p, U)$ be a Paley graph with U be the set of nonzero quadratic residues as above. Being a nonzero quadratic residue is provably an equivalence relation on $\mathbb{Z}_p - \{0\}$ with equivalence classes U and $\mathbb{Z}_p - \{0\} - U$, so $|U| = \frac{|\mathbb{Z}_p - \{0\}|}{2} = \frac{p-1}{2}$. Thus, there are $\frac{p-1}{2}$ distinct squares in $\mathbb{Z}_p - \{0\}$, so $C(\mathbb{Z}_p, U)$ is $\frac{p-1}{2}$- regular.

Despite their clear symmetry, we claim Paley graphs are both pseudorandom and have an adjacency matrix with high Shannon energy. The high Shannon entropy follows from the $\frac{p-1}{2} \approx \frac{p}{2}$, observing that the Shannon entropy of adjacency matrix is maximized in $\frac{|V|}{2}$-regular graphs. The pseudorandomness of Paley graphs is a well-known result that is outlined in the theorem below.

Theorem 5.1 *Let $C(\mathbb{Z}_p, U)$ be a Paley graph on the group $(\mathbb{Z}_p, +)$ for prime $p \equiv 1 mod(4)$. Let $\{\lambda_1 \geq \lambda_2 \geq ..., \geq \lambda_n\}$ be the eigenspectrum of the adjacency matrix A_G of G. Then, $max(\lambda_2, |\lambda_n|) = O(\sqrt{p})$. Thus, $C(\mathbb{Z}_p, U)$ is pseudorandom 8.*

PROOF OF 5.1:

Per Lemma 3.1 we know that the Fourier characters $\{\omega^k\}$ are the eigenvectors corresponding to eigenvalues $\lambda_{k+1} = \hat{\chi_S}(k)\ \forall k \in \mathbb{Z}_p$, where $\chi_S : \mathbb{Z}_p \to \{0, 1\}$ is the indicator function of $S = \{u \in \mathbb{Z}_p - \{0\} : u \equiv a^2 mod(p)$. We can compute the Fourier components of χ_S using the Fourier transform $\hat{\chi_S}(k) = \sum_{j=0}^{p-1} \chi_S(j)\omega(j)^{-k} = \sum_{j=0}^{p-1} \chi_S(j) e^{\frac{-i2\pi kj}{p}}$. Thus,

$$\lambda_{k+1} = \hat{\chi_S}(k) = \sum_{j=0}^{p-1} \chi_S(j) e^{\frac{i2\pi kj}{p}} = \sum_{j \in S} 1 * e^{\frac{i2\pi kj}{p}}$$

Observe that for any $l \in \{1, ..., \frac{p-1}{2}\}$ we have $(p-l)^2 mod(p) = (p^2 - 2pl + l^2) mod(p) \equiv l^2 mod(p)$, so the sequence $\{l^2 mod(p)\}_{l=0}^{p-1} = \{0\}, \{l^2 mod(p)\}_{l=1}^{\frac{p-1}{2}}, \{(p-l)^2 mod(p)\}_{l=\frac{p-1}{2}}^{1}$ iterates through $\{0\}$ and then S twice. Thus,

$$\sum_{j \in S} e^{\frac{i2\pi kj}{p}} = \frac{\sum_{l=1}^{p-1} e^{\frac{i2\pi kl^2}{p}}}{2} = \frac{(\sum_{l=0}^{p-1} e^{\frac{i2\pi kl^2}{p}}) - 1}{2}$$

.

Using Gauss sums we see that $p \equiv 1 mod(4) \implies \sum_{l=0}^{p-1} e^{\frac{i2\pi kl^2}{p}} = \sqrt{(}p)\ \forall k \in \mathbb{Z}_p - \{0\}$. Thus,

$$\lambda_{k+1} = \hat{\chi_S}(k) = \frac{\sqrt{p} - 1}{2} \quad \forall k \in \mathbb{Z}_p - \{0\}$$

Thus $\lambda = O(\sqrt{p})$

QED

Example: Bounding the Algorithmic Entropy of Paley Graphs

Theorem 5.2 (Algorithmic Entropy of Undirected Paley Graphs) *Let $\{P_n = C(\mathbb{Z}_p, U_n))$ be the family of Paley graphs on additive groups $(\mathbb{Z}_p, +)$ with p prime and $p \equiv 1 mod 4$, with $U_n := \{u^2 mod p : u \in \mathbb{Z}_p\}$. Then Paley graphs have lienar Lempel-Ziv complexity, $LZ(C(\mathbb{Z}_p, U_n)) = O(n)$. This gives us loglinear algorithmic entropy, which is sub-random.*

PROOF OF 5.2

Let A be the adjacency matrix of $P_n = C(\mathbb{Z}_p, U_n)$ with entries $a_{ij} = 1 \iff i - j \in S$.

Observe that $(i, j) \in E \iff i = j + s$ for $s \in S \iff i + 1 = (j + 1) + s \iff (i + 1, j + 1) \in E$, so $a_{ij} = a_{(i+1)(j+1)}\ \forall i, j \in \mathbb{Z}_p$. Then, $\forall i \in \{0, ..., p-1\}$ the ith row vector $A_i = (a_{i0}, ..., a_{i(p-1)})$ of A satisfies

$$(a_{i0}, a_{i1} ..., a_{i,(p-2)}, a_{i(p-1)}) = (a_{(i+1)1} a_{(i+2)1}, ..., a_{(i+1)(p-1)}, a_{(i+1)0})$$

.

Therefore, A_{i+1} is a one bit circular shift of $A_i\ \forall i \in \{0, ...p-2\}$. Thus, any row A_i of A for $i \in \{1, ..., p-1\}$ is reproducible from the first row of the adjacency matrix $A_0 = (a_{00}, ...a_{0p-1})$ via i circular shifts. Specifically, if we flatten the adjacency matrix A into a $1 \times \binom{n}{2}$ vector v_A by concatenating

$$A_{0[1:p-1]}, A_{1[2:p-1]}, ...A_{i[i+1,p-1]}, ..., A_{p-2_{[p-1]}}$$

. The partitioning of the upper triangular part of the adjacency matrix of P_n into blocks/codewords is illustrated in figure 3.

Then, applying the LZ76 algorithm to resultant v_A with a dictionary and look ahead buffer of length $p-1$ each yields at most $O(p)$ code words, as $A_{i[i+1,p-1]}$ is a substring of $A_{i-1[i,p-1]}$ for all $i \in [p-1]$. A_{i-1} has length $p-i$ and so is contained entirely in its dictionary window by construction. Thus, each row vector A_i encoded after A_0 only adds a pointer to a substring of A_0 to the encoding (illustrated using the red and orange lines in fig. 3).

A_0 can be encoded in exactly p bits, plus an additional constant number of bits for each row after. Then, $LZ(C(\mathbb{Z}_p, S)) = O(p)$

Then, adding logarithmic overhead for storing the labels of the codewords themselves, this gives us $O(nlogn)$ algorithmic entropy- less than the $O(n^2)$ algorithmic entropy of the adjacency matrices of random graphs.

QED

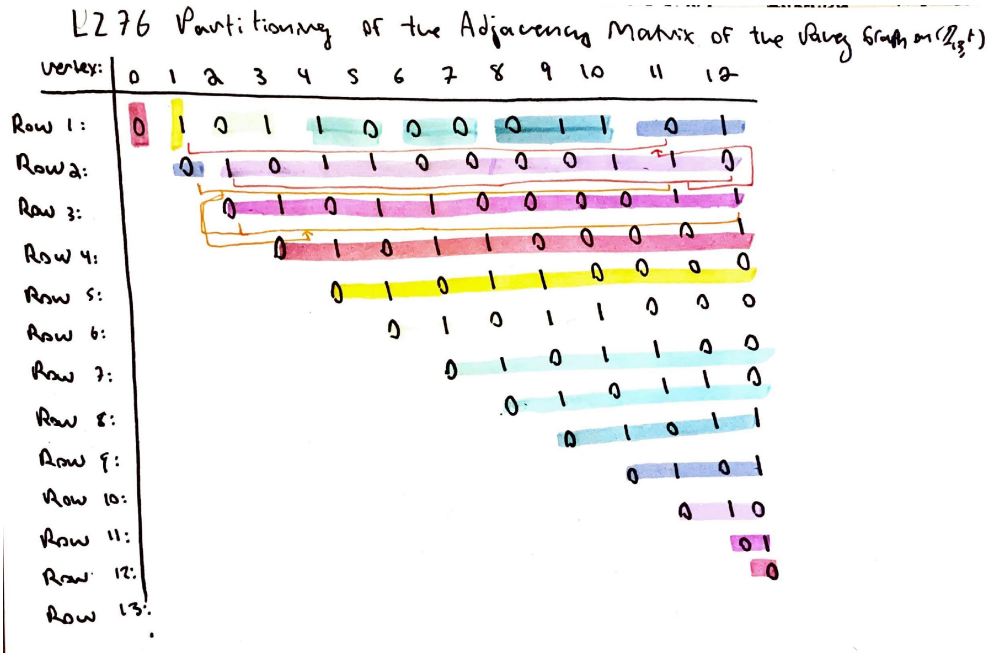

Figure 3: The LZ76 partitioning of the upper triangular adjacency matrix of the Paley graph $C((\mathbb{Z}_{13}, +), U = \{a^2 : a \in \mathbb{Z}_{13}\})$, with blocks highlighted in different colors and two example pointers to previous rows illustrated in red and orange lines. NOTE: THIS WILL BE REPLACED WITH A NICE MACHINE GENERATED FIGURE IN TIME

Now we use the argument in Theorem 5.2 to prove that any undirected Cayley graph on $\mathbb{Z}_n$ for any $n \in \mathbb{Z}_+$ has sub-random Kolmogorov complexity.

Theorem 5.3 (Algorithmic Entropy of Undirected Additive Cayley Graphs) *Let $\{C_n\}$ be a family of undirected Cayley graphs on the additive group $(\mathbb{Z}_n, +)$ with generating sets $U_n \subset \mathbb{Z}_n$. Then the adjacency matrices of $\{C_n\}$ have $O(n)$ Lempel- Ziv complexity. This gives us loglinear algorithmic entropy, which is sub-random.*

PROOF OF 5.3:

Let A denote the adjacency matrix of C_n. Observe that each row of the adjacency matrix, A_i corresponds to the vertex $i-1 \in \mathbb{Z}_n$, and so has $A_{ij} = 1 \iff j - i \in U_n$. Thus, A_i is the indicator function χ_{S_i} of the set $S_i := \{s \in \mathbb{Z}_n : (i-1) + s \in U\}$. In particular, A_1 corresponds to 0, the identity element, and so $A_1 = \chi_U$ as $U = S_i := \{s \in \mathbb{Z}_n : 0 + s = s \in U\}$. Using LZ76 we can encode A_1 in at most n codewords, often less.

Because we're working with undirected Cayley graphs we see that the adjacency matrix of C_n is symmetric and we only need encode the upper triangular part, $\{A_1, A_{2[2:n]}, ...A_{i[i:n]}, ...A_{n-1}[(n-1):n], A_{nn}$

We see that $\forall i \in [n-1]$ consecutive row vectors A_i and A_{i+1} are single bit circular permutations of each other, as $A_{i+1} = \chi_{\{s\in[n]:i+s\in U\}} = \chi_{\{s\in[n]:(i-1)+(s+1)\in U\}}$. All addition is taken modulo n so this gives us $A_{ij} = \chi_{S_i}(j-1) = \chi_{S_{i+1}}(j) = A_{(i+1)(j+1)}$ modulo n as desired.

Then, $\forall i \in [n-1]$, we have that $A_{(i+1)[(i+1):n]} = A_{i[i:n-1]}$, i.e, $A_{(i+1)[(i+1):n]}$ is a direct substring of $A_{i[i:n]}$, so applying the LZ76 algorithm encodes every row after the first row as its own codeword/block, resulting in only a linear number of codewords and thus linear Lempel Ziv complexity.

Then, adding logarithmic overhead for storing the labels of the codewords themselves, this gives us $O(nlogn)$ algorithmic entropy- less than the $O(n^2)$ algorithmic entropy of the adjacency matrices of random graphs.

QED

Empirical Lempel Ziv Complexity Validation Our theoretical calculation for the algorithmic entropy of undirected Cayley graphs on additive groups is validated by empirical computation of the Lempel-Ziv compression on both Paley graphs and cycle graphs using the Lempel-Siv Markov chain algorithm (LZMA). LZMA is an optimized version of the LZ77 algorithm, based on the LZ76 algorithm, that achieves "higher compression rate, faster decompression, and lower memory requirements" 13.

Recall from remark 3.2 that cycle are the sparsest possible undirected, connected Cayley graphs on $(\mathbb{Z}_p, +)$ for p prime up to isomorphism. Then comparing Paley graphs, which are dense, to cycle graphs captures the breadth of undirected Cayley graph connectivity.

Validation was performed using code produced in python for the TRIPODS 2023 Summer Research Program using the numpy, lzma, and networkX packages, cited in the bibliography at [14]. Figures 4 and 5 depict the compression size vs original number of graph nodes curve for cycle graphs and Paley graphs respectively. For the sake of comparison, both graphs were evaluated for prime nodes congruent 1mod4.

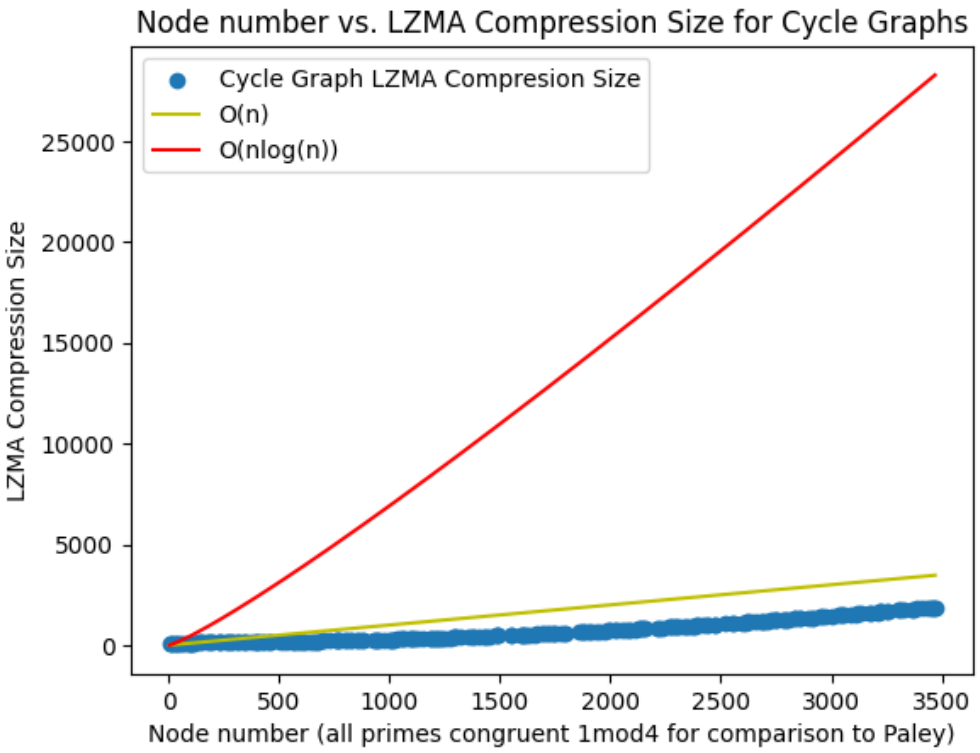

Figure 4: The LZMA compression size versus number of nodes in a cycle graph

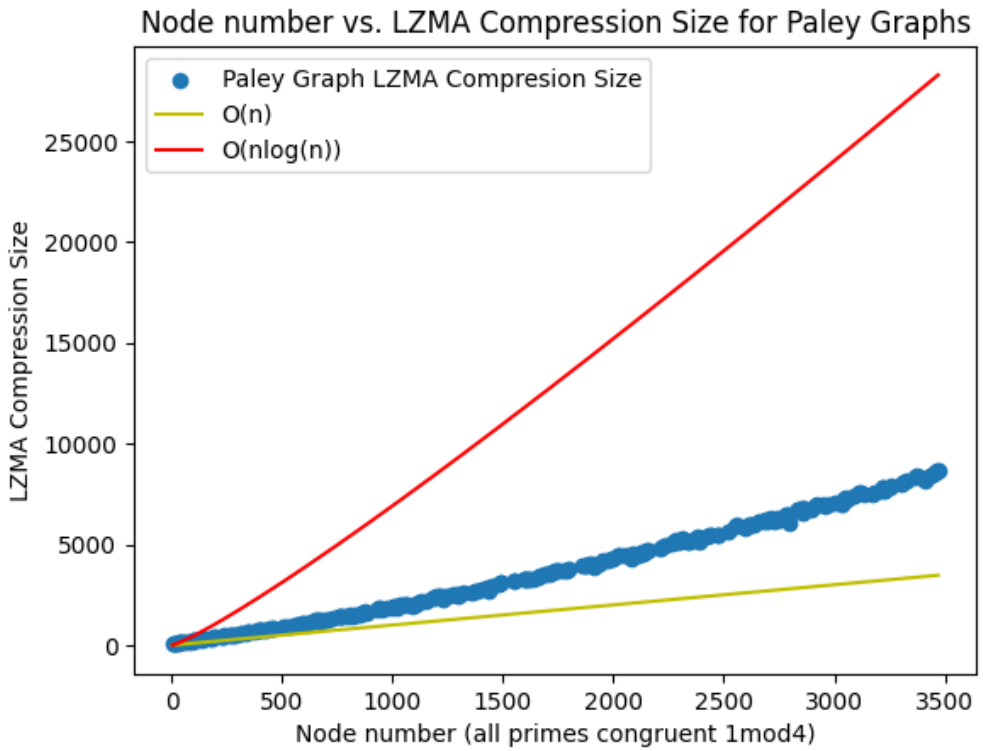

Figure 5: The LZMA compression size versus number of nodes in a Paley graph

Both curves are convex, consistent with our loglinear bound we proved for algorithmic entropy in Theorem 5.3. Note however the relatively flat (and indeed, sub-linear) curve of the cycle graph compression compared to the Paley graph compression.

6 Current Work

Our ongoing work comprises two directions: first, we are working on generalizing Cayley graphs results to expander graphs from constructed from Cayley graphs using the Zig-Zag graph product. Such as such graphs have pseudorandom spectral properties, and are (unlike many of the Cayley graphs we looked at, such as Paley graphs) quite sparse. This makes them good reservoir candidates in Echo State Networks, and also facilitates the networks having low algorithmic entropy.

The second is to apply our pseudorandom reservoir candidates to echo state networks, which traditionally use binomial graph reservoirs to achieve a nonlinear embedding of an input time series in a higher dimensional space in order to forecast the time series [15]. We hope to use pseudorandom graphs to "save on bits of randomness" expended in the construction of reservoirs for echo state networks. As part of this endeavour, we are studying the memory capacity of echo state networks build using our pseudorandom graphs. We then plan to perform empirical tests on the ability of simple echo state networks with pseudorandom reservoirs to forecast the Mackey-Glass time series, single variable chaotic time series. Our end goal is to identify a pseudorandom graph with low algorithmic entropy that outperforms a binomial random graph of the same edge density in the chaotic time series forecasting task.

References

[1] Zenil, Héctor, Narsis A. Kiani, and Jesper Tegnér. 2016. "Methods of Information Theory and Algorithmic Complexity for Network Biology." Seminars in Cell and Developmental Biology 51 (March): 32–43.

https://doi.org/10.1016/j.semcdb.2016.01.011.

[2] Mikołaj Morzy, Tomasz Kajdanowicz, and Przemysław Kazienko, "On Measuring the Complexity of Networks: Kolmogorov Complexity versus Entropy," Complexity 2017 (January 1, 2017): 1–12, https://doi.org/10.1155/2017/3250301.

[3] Zenil, Kiani, and Tegnér, 2017. "Low-Algorithmic-Complexity Entropy-Deceiving Graphs." Physical Review 96 (1). https://doi.org/10.1103/physreve.96.012308.

[4] Krivelevich, Michael, and Benny Sudakov. 2006. "Pseudo-Random Graphs." In Bolyai Society Mathematical Studies, 199–262. https://doi.org/10.1007/978-3-540-32439-3_10.

[5] Trevisan, Luca. 2017. Lecture Notes on Graph Partitioning, Expanders and Spectral Methods.

[6] Mowshowitz, Abbe, and Matthias Dehmer. 2012. "Entropy and the Complexity of Graphs Revisited." Entropy 14 (3): 559–70. https://doi.org/10.3390/e14030559.

[7] Li, Ming, and Paul Vitányi. 2008. An Introduction to Kolmogorov Complexity and Its Applications. Texts in Computer Science. https://doi.org/10.1007/978-3-030-11298-1.

[8] Zhao, Yufei. Graph Theory and Additive Combinatorics. Fall 2019. Massachusetts Institute of Technology: MIT OpenCouseWare, https://ocw.mit.edu/

[9] Trachtenberg, Ari. 2018. "Empirical Kolmogorov Complexity." IEEE Conference Publication — IEEE Xplore, February. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8503120.

[10] Amigó, José M., Janusz Szczepański, Eligiusz Wajnryb, and Maria V. Sanchez-Vives. 2004. "Estimating the Entropy Rate of Spike Trains via Lempel-Ziv Complexity." Neural Computation 16 (4): 717–36. https://doi.org/10.1162/089976604322860677.

[11] Chung, Fan, Ronald L. Graham, and R. Wilson. 1989. "Quasi-random Graphs." Combinatorica 9 (4): 345–62. https://doi.org/10.1007/bf02125347.

[12] Conlon, David, Jacob Fox, and Yufei Zhao. 2014. "Extremal Results in Sparse Pseudorandom Graphs." Advances in Mathematics (New York. 1965) 256 (May): 206–90. https://doi.org/10.1016/j.aim.2013.12.004.

[13] Horita, Augusto Y., Ricardo Bonna, Denis S. Loubach, Ingo Sander, and Ingemar Söderquist. 2019. "Lempel-Ziv-Markov Chain Algorithm Modeling Using Models of Computation and ForSyDe." Linköping Electronic Conference Proceedings (Print), October. https://doi.org/10.3384/ecp19162017.

[14] Pack, Svetlana, Anuraag Kumar, Joshua Iosevich, Alhussein Khalil, Azita Mayeli, and Alex Iosevich. 2023. "Graph LZMA Complexity Empirical Validation." Software. https://colab.research.google.com/drive/1Zn68SkA_T_brFgZnd-zN7dtZhWA7bMiY?usp=sharing.

[15] Aceituno, Pau Vilimelis, Gang Yan, and Yang-Yu Liu. 2020. "Tailoring Echo State Networks for Optimal Learning." iScience (Cambridge) 23 (9): 101440. https://doi.org/10.1016/j.isci.2020.101440.