AN OVERVIEW OF QUASIRANDOM GRAPH SEQUENCES

LEAH VASHEVKO

Abstract. In their 1989 paper *Quasi-Random Graphs*, Chung, Graham, and Wilson introduced the notion of quasirandom graph sequences: deterministic sequences that nevertheless exhibit many of the same properties as random graphs. Remarkably, these properties are equivalent, so verifying one guarantees all properties of a quasirandom sequence. This paper examines several of these properties and provides proofs of their equivalence.

Contents

1. Introduction

Consider the random graph $G(n,p)$, which denotes the graph with n vertices and each edge is included independently with probability p. As n grows, certain quantities associated with $G(n,p)$ converge to their expected values.

These quantities may include the number of edges, the average degree, the average codegree of two vertices, the number of cycles of a given length, the number of induced subgraph occurrences of a fixed graph of size $s \leq n$, the eigenvalues of the adjacency matrix, and many others.

Chung, Graham, and Wilson [1] introduced the notion of *quasirandom* graph sequences. A graph sequence is called quasirandom if it satisfies any one of a collection of equivalent properties. Remarkably, verifying one of these properties guarantees all the others.

By possessing these properties, quasirandom graph sequences "look like" random graph sequences as the number of vertices grows, even though the graphs may be completely deterministic.

This paper discusses a portion of their primary results, presents revised versions of their proofs with commentary, and discusses an example of a quasirandom graph sequence.

2. Asymptotic Notation

Since the theorem discussed in this paper applies to sequences of graphs, the statements about properties of quasirandom graph sequences are true as n grows large. These properties are described with the $o(\alpha(n))$ notation, where $\alpha(n)$ is some function of n.

The notation $f(n) = o(1)$ means that $f(n) \to 0$ as $n \to \infty$. More generally, when $f(n) = o(\alpha(n))$, $\frac{f(n)}{\alpha(n)} \to 0$ as $n \to \infty$.

It is common to write $f(n) = (1 + o(1))g(n)$. This means that $f(n)$ is asymptotically equal to $g(n)$, so $\frac{f(n)}{g(n)} \to 1$ as $n \to \infty$.

Therefore, if $f(n) = o(n^j)$ for some j, $f(n) = o(n^k)$ for all $k \leq j$.

3. Properties of Quasirandom Graph Sequences

We will consider the following properties that a sequence of graphs may or may not satisfy.

In particular, this paper focuses on four key properties of quasirandom graph sequences which are asymptotically equivalent for sequences with edge density $\frac{1}{2}$.

While quasirandom graphs may have any edge density, we restrict attention to sequences with edge density $\frac{1}{2}$, i.e. resembling $G(n, \frac{1}{2})$.

Let $\mathcal{G} = (G(n))_{n\geq 1}$ denote a sequence of graphs, where $G(n)$ has n vertices.

Property 3.1 (**Induced Subgraph Count**)**.** Fix $M(s)$ to be some s-vertex graph. We define $N^*(M(s))$ to be the number of times a labelled copy of $M(s)$ occurs as an induced subgraph of $G(n)$. Then, this count satisfies

$$N^*(M(s)) = (1 + o(1))n^s 2^{-\binom{s}{2}}.$$

Property 3.2 (**Edge Count and 4-Cycle Count**)**.** Let M be some graph. We define $N(M)$ to be the number of times a labelled copy of M occurs as a subgraph (non-induced) of G. Therefore, $N(C_4)$ is the number of 4-cycles in G.

We also define $E(G(n))$ to be the number of edges in $G(n)$. Then, we have the following counts.

$$E(G(n)) \geq (1 + o(1))\frac{n^2}{4} \text{ and } N(C_4) \leq (1 + o(1))\left(\frac{n}{2}\right)^4.$$

Property 3.3 (**Codegree Discrepancy**)**.** For vertices v, v', let $\operatorname{codeg}(v, v')$ denote $|N(v) \cap N(v')|$, the number of common neighbors of v and v'. Then the sum of deviations from the expected codegree satisfies

$$\sum_{v,v'} \left|\operatorname{codeg}(v, v') - \frac{n}{4}\right| = o(n^3).$$

Property 3.4 (**Agreement Discrepancy**)**.** For vertices v, v', define

$$s(v, v') = |\{w \in V(G(n)) \mid a(w, v) = a(w, v')\}|,$$

where $a(w,v) = 1$ if w and v are adjacent, and 0 otherwise. Thus, $s(v,v')$ counts the number of vertices that are either adjacent to both v and v' or to neither.

In $G(n, \frac{1}{2})$, the expected value of $s(v,v')$ is $\frac{n}{2}$.

This property asserts that the sum of deviations of $s(v,v')$ from its expected value in a random graph is asymptotically small:

$$\sum_{v,v'} \left| s(v,v') - \frac{n}{2} \right| = o(n^3).$$

Now, we reach the central theorem of this paper.

Theorem 3.5 (**Equivalence of Quasirandom Properties**). *Let $\mathcal{G} = (G(n))_{n\geq 1}$ be a sequence of n-vertex graphs with edge density $1/2$, and let the properties 3.1–3.4 be as defined above. Then, the following are equivalent:*

(1) $\mathcal{G}$ satisfies the Induced Subgraph Count property (Property 3.1),
(2) $\mathcal{G}$ satisfies the Edge Count and 4-Cycle Count property (Property 3.2),
(3) $\mathcal{G}$ satisfies the Codegree Discrepancy property (Property 3.3),
(4) $\mathcal{G}$ satisfies the Agreement Discrepancy property (Property 3.4).

In other words, if $\mathcal{G}$ satisfies any one of these properties, it satisfies all of them and is quasirandom.

4. Expected Occurrence in Random Graph Sequences

In each of the above properties, a quantity of the quasirandom graph sequence is asymptotically equivalent to the expected value of that quantity in a random graph sequence. This is what it means for a quasirandom graph sequence to "look like" a random graph sequence.

To verify this, we will calculate the expected values of each quantity on the random graph sequence $\mathcal{G} = (G(n, \frac{1}{2}))_{n\geq 1}$.

4.1. Induced Subgraph Count.

Let M be a graph on s vertices. We want to find the expected value of $N^*(M)$ for the random graph $G = G(n, \frac{1}{2})$.

Since we are counting *labelled* induced subgraphs, the number of choices of s labelled vertices is $\binom{n}{s} \cdot s!$.

Since $N^*(M)$ counts *induced* subgraphs, the adjacency for each pair of vertices must match M. The number of vertex pairs is $\binom{s}{2}$. Each of the possible edges matches the adjacency of M with probability $\frac{1}{2}$. Therefore, the probability that M is an induced subgraph on a set of s points is $2^{-\binom{s}{2}}$.

Therefore, the expected value

$$\begin{aligned}\mathbb{E}(N^*(M(s))) &= \binom{n}{s} s! \cdot 2^{-\binom{s}{2}} \\ &= n(n-1)...(n-s+1) \cdot 2^{-\binom{s}{2}} \\ &= (1+o(1))n^s \cdot 2^{-\binom{s}{2}}.\end{aligned}$$

Notice that this is exactly the value that quasirandom graph sequences asymptotically approach.

4.2. **Edge and 4-Cycle Count.** First, we will find the expected edge count of $G = G(n, \frac{1}{2})$. The total number of pairs of vertices of G is $\binom{n}{2}$ and the probability that a pair is connected is $\frac{1}{2}$. Thus

$$\begin{aligned}\mathbb{E}(E(G)) = \binom{n}{2} \cdot \frac{1}{2} &= \frac{n(n-1)}{2} \cdot \frac{1}{2} \\ &= (1+o(1))\frac{n^2}{2} \cdot \frac{1}{2} \\ &= (1+o(1))\frac{n^2}{4}.\end{aligned}$$

Next, we will find the expected number of labelled 4-cycles $N(C_4)$.

There are $\binom{n}{4} \cdot 4!$ ways to choose 4 labelled vertices. Then, since we are not counting *induced* 4-cycles, only the 4 outer edges need to be properly connected. Therefore, a choice of 4 vertices contains a 4-cycle with probability 2^{-4}. Thus,

$$\begin{aligned}\mathbb{E}(N(C_4)) &= \binom{n}{4} \cdot 4! \cdot 2^{-4} \\ &= (1+o(1))n^4 \cdot 2^{-4} \\ &= (1+o(1))\left(\frac{n}{2}\right)^4.\end{aligned}$$

Notice that this calculation gives asymptotic estimates for the edge and 4-cycle count, but the statement in Property 3.2 is weaker, requiring only asymptotic upper and lower bounds.

4.3. **Codegree.** Let $v, v' \in V(G)$. Any vertex $w \neq v, v'$ is adjacent to v and v' with probability $\frac{1}{4}$. There are $n-2$ other vertices on G, so

$$\mathbb{E}(\text{codeg}(v, v')) = \frac{n-2}{4} = (1+o(1))\frac{n}{4}.$$

This is the value to which the actual codegree of vertices is compared in Property 3.3.

4.4. **Agreement.** Let $v, v' \in V(G)$. Any vertex $w \neq v, v'$ is adjacent to v and v' with probability $\frac{1}{4}$. The probability that w is adjacent to neither v nor v' is also $\frac{1}{4}$, since adjacency to each vertex is independent with probability $\frac{1}{2}$. Therefore, w counts toward $s(v, v')$ with probability $\frac{1}{2}$.

There are $n-2$ other vertices on G, so

$$\mathbb{E}(s(v, v')) = \frac{n-2}{2} = (1+o(1))\frac{n}{2}.$$

This is the value to which the actual s-value of vertices is compared in Property 3.4.

5. Proofs of Equivalence

In this section, we will prove Theorem 3.5, which states that the quasirandom properties are asymptotically equivalent. Our strategy is to show a cycle of implications between the four properties—that each property implies the next—which will prove that all the properties are equivalent.

5.1. Induced Subgraph Count Implies Edge and 4-Cycle Count.

Proposition 5.1. *If a graph sequence satisfies the Induced Subgraph Count property (Property 3.1), then it satisfies the Edge Count and 4-Cycle Count property (Property 3.2).*

Proof. Assume that Property 3.1 is true for all $s \leq n$. We will show that the edge and 4-cycle counts follow from the induced subgraph counts. This follows by applying Property 3.1 to carefully chosen subgraphs.

Edge count. An edge can be viewed as a 2-vertex graph, so we let M be the 2-vertex graph with an edge as below.

$$M = \quad 1 \text{ —— } 2$$

Then, by Property 3.1, we have that

$$N^*(M) = (1+o(1))n^2 2^{-\binom{2}{2}} = (1+o(1))n^2 \cdot 2^{-1}.$$

However, since $N^*(M)$ counts an induced subgraph for each ordering of the vertices, it double counts all edges. Therefore, we have

$$E(G(n)) = \frac{N^*(M)}{2} = (1+o(1))\frac{n^2}{4}.$$

4-cycle count. Since the 4-cycle count $N(C_4)$ does not require that the 4-cycles be induced, there are four ways that a 4-cycle can appear in an induced subgraph on 4 vertices, which are shown in Figure 1 below.

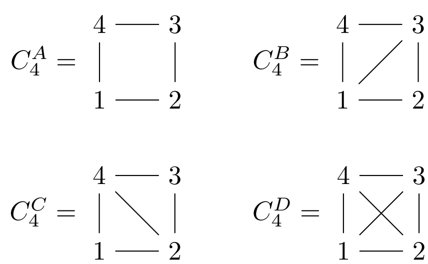

FIGURE 1. All four variants of C_4.

Then, we can count the occurrences of each graph. For each graph, we have the following count by Property 3.1.

$$\begin{aligned} N^*(C_4^A) = N^*(C_4^B) = N^*(C_4^C) = N^*(C_4^D) &= (1+o(1))n^4 2^{-\binom{4}{2}} \\ &= (1+o(1))n^4 2^{-6}. \end{aligned}$$

Then, $N(C_4)$ is the sum of the counts of each type of induced subgraph.

$$\begin{aligned} N(C_4) = N^*(C_4^A) + N^*(C_4^B) + N^*(C_4^C) + N^*(C_4^D) &= 4 \cdot (1+o(1))n^4 2^{-6} \\ &= (1+o(1))n^4 2^{-4} \\ &= (1+o(1))\left(\frac{n}{2}\right)^4. \end{aligned}$$

□

As noted above, this implication gives equalities for the edge and 4-cycle count, which is stronger than Property 3.2 requires.

5.2. Edge and 4-Cycle Count Implies Codegree Discrepancy.

Proposition 5.2. *If a graph sequence satisfies the Edge Count and 4-Cycle Count property (Property 3.2), then it satisfies the Codegree Discrepancy property (Property 3.3).*

Proving this proposition requires using the Cauchy-Schwarz inequality to bound the sum of the codegree discrepancies using the technique shown below.

Fact 5.1 (**Cauchy-Schwarz Inequality**). *For any real numbers $a_1, \dots, a_n$ and $b_1, \dots, b_n$,*

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right).$$

This proof (and the following ones) rely heavily on the Cauchy–Schwarz inequality. Two immediate consequences will be useful. Setting $b_i = 1$ yields

$$\left(\sum_{i=1}^{n} a_i\right)^2 \le n \sum_{i=1}^{n} a_i^2. \tag{5.3}$$

Equivalently, we can rewrite this as

$$\sum_{i=1}^{n} a_i^2 \ge \frac{1}{n}\left(\sum_{i=1}^{n} a_i\right)^2. \tag{5.4}$$

Remark 5.5 (**Technique for bounding discrepancy sums.**). In the following two proofs, we repeatedly use the Cauchy–Schwarz inequality to bound discrepancy sums of the form

$$\sum_{x} |\alpha(x) - C|$$

where $\alpha(x)$ is some quantity associated with a vertex, pair of vertices, or other combinatorial object, and C is its expected value in a random graph.

Lemma 5.6. *For any collection of values $\alpha(x)$ with expected value C,*

$$\sum_{x} |\alpha(x) - C| \le \sqrt{N}\left[\sum_{x} \alpha(x)^2 - 2\sum_{x} C\alpha(x) + \sum_{x} C^2\right]^{\frac{1}{2}} \tag{5.7}$$

where N is the number of terms in the sum.

Proof. Rather than estimating the absolute differences directly, we bound the discrepancy by considering the first and second moments of the quantity $\alpha(x)$, namely

$$\sum_{x} \alpha(x) \text{ and } \sum_{x} \alpha(x)^2.$$

The sum of values $\sum_x \alpha(x)$ shows the average value of the quantity, while the sum of squares $\sum_x \alpha(x)^2$ captures the variance of the quantity.

Applying the Cauchy-Schwarz inequality (5.3) to the square of the discrepancy sum,

$$\left[\sum_x |\alpha(x) - C|\right]^2 \leq N \sum_x |\alpha(x) - C|^2 .$$

where N is the number of terms in the sum. So taking the square root gives

$$\begin{aligned}\sum_x |\alpha(x) - C| &\leq \sqrt{N}\left[\sum_x |\alpha(x) - C|^2\right]^{\frac{1}{2}} \\ &= \sqrt{N}\left[\sum_x \alpha(x)^2 - 2\sum_x C\alpha(x) + \sum_x C^2\right]^{\frac{1}{2}} .\end{aligned}$$

Substituting the values of the first and second moments gives a bound for the discrepancy sum. □

Now, we establish the facts necessary to prove the proposition. To find an estimate for the first and second moments of codegree, we find the number of homomorphisms from C_4 into G, which is closely related to the number of 4-cycles. The following lemmas establish this connection.

Lemma 5.8. *The number of* ***non-injective*** *labelled homomorphisms from C_4 to G is $o(n^4)$.*

Proof. Homomorphisms allow for non-injective maps where two opposite vertices coincide, as shown in Figure 2.

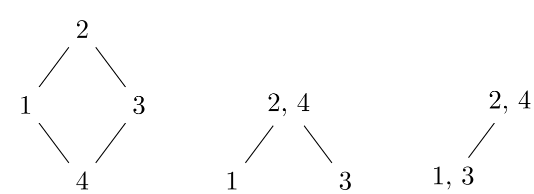

FIGURE 2. Three possible homomorphic images of C_4.

Since the non-injective maps have at most three vertices, the number of them is at most of degree 3 (since the number of ways to choose 3 vertices is of degree 3) and is therefore captured by a $o(n^4)$ term. □

Lemma 5.9. *The number of labelled homomorphisms from C_4 to G equals*

$$\hom(C_4, G) = \sum_{u,v} \operatorname{codeg}(u, v)^2 .$$

Proof. We count a homomorphism by choosing two vertices to be opposite vertices in the cycle. Then, we must choose two other vertices adjacent to both u and v. The number of such common neighbors is precisely $\operatorname{codeg}(u, v)$, and because we count labelled homomorphisms, each pair u, v contributes $\operatorname{codeg}(u, v)^2$ homomorphisms. $\square$

Lemma 5.10. *The first and second moments of codegree satisfy the asymptotic estimates*

$$\sum_{u,v} \operatorname{codeg}(u, v) = (1 + o(1))\frac{n^3}{4} \quad \textit{and} \quad \sum_{u,v} \operatorname{codeg}(u, v)^2 = (1 + o(1))\left(\frac{n}{2}\right)^4.$$

Proof. First, we assume the bound on edge count and 4-cycle count from Property 3.2.

Then, we notice that the bound on $N(C_4)$ also bounds $\hom(C_4, G)$, where $\hom(C_4, G)$ is the number of subgraphs homomorphic to the 4-cycle. Note that $\hom(C_4, G)$ includes injective maps (C_4 itself), and noninjective maps. By Lemma 5.8, the count of noninjective maps is $o(n^4)$. Therefore,

$$\hom(C_4, G) = N(C_4) + o(n^4) \leq (1 + o(1))\left(\frac{n}{2}\right)^4.$$

Then, we use Lemma 5.9 to count homomorphisms through codegree and by Cauchy-Schwarz (5.4) we have that

$$\sum_{u,v} \operatorname{codeg}(u, v)^2 \geq \frac{1}{n^2}\left(\sum_{u,v} \operatorname{codeg}(u, v)\right)^2,$$

as there are n^2 terms in the sum.

The sum of codegrees can be reindexed as the sum of the squares of degrees, as shown below

$$\sum_{u,v} \operatorname{codeg}(u, v) = \sum_{w} \deg(w)^2.$$

Instead of fixing vertices u and v and counting their common neighbors, we sum over all vertices w and count the number of pairs of vertices adjacent to both, which is $\deg(w)^2$ ordered pairs.

The sum of degrees on G is twice the edge count of G, since each edge is double counted:

$$\sum_{w} \deg(w) = 2E(G(n)).$$

Bringing everything together, we have the following inequalities:

$$
\begin{aligned}
(1+o(1))\left(\frac{n}{2}\right)^4 \geq \hom(C_4, G) &\overset{5.9}{=} \sum_{u,v} \operatorname{codeg}(u,v)^2 \\
&\overset{5.1}{\geq} \frac{1}{n^2}\left(\sum_{u,v} \operatorname{codeg}(u,v)\right)^2 = \frac{1}{n^2}\left(\sum_{v} \deg(v)^2\right)^2 \\
&\overset{5.1}{\geq} \frac{1}{n^2}\left(\frac{1}{n}\left[\sum_{v} \deg(v)\right]^2\right)^2 = \frac{1}{n^2}\left(\frac{1}{n}\left[2 \cdot E(G(n))\right]^2\right)^2 \\
&\overset{3.2}{\geq} \left(\frac{1}{n}\left[2 \cdot \left((1+o(1))\frac{n^2}{4}\right)\right]^2\right)^2 = \frac{1}{n^2}\left(\frac{1}{n}\left[(1+o(1))\frac{n^2}{2}\right]^2\right)^2 \\
&= \frac{1}{n^2}\left(\frac{1}{n}\cdot(1+o(1))\frac{n^4}{4}\right)^2 = \frac{1}{n^2}\left((1+o(1))\frac{n^3}{4}\right)^2 = \frac{1}{n^2}(1+o(1))\frac{n^6}{2^4} \\
&= (1+o(1))\left(\frac{n}{2}\right)^4.
\end{aligned}
$$

Rearranging, we have the estimates for first and second moments of codegree. □

Now, we can combine these facts to prove Proposition 5.2.

Proof. We apply Lemma 5.6 to bound the codegree discrepancy sum and substitute the bounds from Lemma 5.10:

$$
\begin{aligned}
\sum_{u,v} \left|\operatorname{codeg}(u,v) - \frac{n}{4}\right| &\overset{5.6}{\leq} n\left[\sum_{u,v} \operatorname{codeg}(u,v)^2 - \frac{n}{2}\sum_{u,v} \operatorname{codeg}(u,v) + \sum_{u,v} \frac{n^2}{2^4}\right]^{\frac{1}{2}} \\
&\overset{5.10}{=} n\left[(1+o(1))\frac{n^4}{2^4} \;-\; \frac{n}{2}(1+o(1))\frac{n^3}{4} \;+\; n^2\frac{n^2}{2^4}\right]^{\frac{1}{2}} \\
&= n\left[(1+o(1))\frac{n^4}{2^4} \;-\; (1+o(1))\frac{n^4}{2^3} \;+\; \frac{n^4}{2^4}\right]^{\frac{1}{2}} \\
&= n\left[\frac{n^4}{2^4} + \frac{n^4}{2^4}o(1) \;-\; \frac{n^4}{2^3} - \frac{n^4}{2^3}o(1) \;+\; \frac{n^4}{2^4}\right]^{\frac{1}{2}} \\
&= n\left[\frac{n^4}{2^4}o(1) - \frac{n^4}{2^3}o(1)\right]^{\frac{1}{2}} \\
&= n\left[o(n^4)\right]^{\frac{1}{2}} = n \cdot o(n^2) = o(n^3).
\end{aligned}
$$

The n^4 terms cancel, leaving $o(n^4)$ in the brackets, which gives the desired bound on codegree discrepancy.

□

5.3. Codegree Discrepancy Implies Agreement Discrepancy.

Proposition 5.11. *If a graph sequence satisfies the Codegree Discrepancy property (Property 3.3), then it also satisfies the Agreement Discrepancy property (Property 3.4).*

We prove this proposition in two steps. First, we use the adjacency matrix of the graph to determine that a small codegree discrepancy implies that the edge count of the graph matches the expected count from Property 3.2. This allows us to use Lemma 5.6 to bound the degree discrepancy from the expected value. Then, we are able to use the codegree discrepancy and degree discrepancy to bound the agreement discrepancy.

Remark 5.12. For a graph G on n vertices, the adjacency matrix

$$A = \Big[a(u,v)\Big]_{u,v}$$

has $a(u,v) = 1$ if u and v are adjacent, and 0 otherwise. This matrix is symmetric, which means it has an orthonormal basis of eigenvectors with real eigenvalues.

The powers of A have combinatorial meaning:

- $(A^2)_{uv}$ counts the number of common neighbors of u and v, i.e. $\mathrm{codeg}(u,v)$.
- More generally, $\mathrm{Tr}(A^k)$ counts the number of closed walks of length k.

Thus, the eigenvalues of A encode codegree and edge distribution information.

We use $\lambda_1, ..., \lambda_n$ to denote the eigenvalues of the adjacency matrix A. The eigenvalues are ordered such that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$.

Lemma 5.13. *If* $\sum_{u,v} \left|\mathrm{codeg}(u,v) - \frac{n}{4}\right| = o(n^3)$, *then the eigenvalues of A satisfy*

$$\lambda_1 = (1 + o(1))\frac{n}{2}, \qquad \lambda_2, \ldots, \lambda_n = o(n).$$

Proof. Let $\mathbf{v} = \big(1, 1, ..., 1\big)^T$, the ones vector.

Since A is symmetric, the spectral theorem guarantees an orthonormal basis of eigenvectors, and the Rayleigh quotient shows that the maximum factor by which A can stretch a vector is its largest eigenvalue λ_1. Therefore, we have that $\lambda_1|v| \geq |Av|$, and therefore

$$\lambda_1^2 |\mathbf{v}|^2 \geq |A\mathbf{v}|^2 = \langle A\mathbf{v}, A\mathbf{v}\rangle = \langle A^2\mathbf{v}, \mathbf{v}\rangle$$

by the linearity of the inner product.

Then, we substitute $\mathbf{v}$ and codegree for A^2 (by Remark 5.12), giving

$$\begin{aligned}\langle A\mathbf{v}, A\mathbf{v}\rangle &= \begin{pmatrix}1 & \ldots & 1\end{pmatrix} \Big[\mathrm{codeg}(u,v)\Big]_{u,v} \begin{pmatrix}1 \\ \vdots \\ 1\end{pmatrix} \\ &= \sum_{u,v} \mathrm{codeg}(u,v) = (1 + o(1))\frac{n^3}{4}.\end{aligned}$$

Since $||\mathbf{v}||^2 = n$, we have $\lambda_1^2 \geq (1 + o(1))\dfrac{n^2}{4}$, so

$$\lambda_1 \geq (1 + o(1))\frac{n}{2}.$$

Now, we can estimate the sum of all of the eigenvalues to find a bound on the remaining ones. We take the sum of the fourth powers of the eigenvalues, since this can be expressed as a sum of codegree.

Since the trace of the adjacency matrix is equal to the sum of the eigenvalues, we deconstruct the trace of the fourth power into a matrix multiplication and substitute the codegree per Remark 5.12:

$$\begin{aligned}\sum_i \lambda_i^4 &= \mathrm{Tr}(A^4) = \mathrm{Tr}(A^2 \cdot A^2) \\ &= \sum_u (A^2 \cdot A^2)_{uu} = \sum_u \sum_v (A^2)_{uv}(A^2)_{vu} \\ &= \sum_{u,v} (A^2)_{uv}^2 \overset{5.12}{=} \sum_{u,v} \mathrm{codeg}(u,v)^2.\end{aligned}$$

Combinatorically, the discrepancy of the codegree $(\mathrm{codeg}(u,v) - \frac{n}{4})$ can be at most $\frac{3}{4}n$, so we can loosely bound

$$\sum_{u,v} (\mathrm{codeg}(u,v) - \frac{n}{4})^2 \leq n \sum_{u,v} (\mathrm{codeg}(u,v) - \frac{n}{4}) = no(n^3) = o(n^4).$$

We can use this bound and the codegree discrepancy to estimate the eigenvalues:

$$\begin{aligned}\sum_i \lambda_i^4 &= \sum_{u,v} \mathrm{codeg}(u,v)^2 \\ &= \sum_{u,v} \left(\frac{n}{4} + (\mathrm{codeg}(u,v) - \frac{n}{4})\right)^2 \\ &= \sum_{u,v} \left(\frac{n^2}{16} + 2\frac{n}{4}(\mathrm{codeg}(u,v) - \frac{n}{4}) + (\mathrm{codeg}(u,v) - \frac{n}{4})^2\right) \\ &= n^2 \left(\frac{n^2}{16}\right) + \frac{n}{2} o(n^3) + o(n^4) \\ &= (1 + o(1)) \frac{n^4}{2^4}.\end{aligned}$$

The first eigenvalue is large enough that it accounts for the entirety of the sum of fourth powers of eigenvalues, which gives a bound on the remaining eigenvalues.

Since the fourth power of the first eigenvalue λ_1 already accounts for the entire value of this sum, the remaining eigenvalues thus must be small:

$$\lambda_2, \lambda_3, ..., \lambda_n = o(n).$$

□

Claim 5.14. *The edge count can be expressed as half the inner product of the ones vector* $\mathbf{v}$ *and the adjacency matrix* A *applied to* $\mathbf{v}$*:*

$$E(G(n)) = \frac{1}{2} \langle \mathbf{v}, A\mathbf{v} \rangle.$$

This is because the inner product counts the number of 1 entries in the adjacency matrix, since $\mathbf{v}$ is the ones vector, which is the sum of degrees of G. The edge count is half of the sum of the degrees.

While the Codegree Discrepancy assumption does not directly allow us to apply the adjacency matrix to a vector, we can diagonalize the matrix and estimate the eigenvectors. Then, we can perform the inner product and obtain an estimate for the edge count.

Lemma 5.15. *Under the same assumption of Codegree Discrepancy (Property 3.3),*

$$E(G(n)) = (1 + o(1))\frac{n^2}{4}.$$

Proof. Since the adjacency matrix is symmetric, it is diagonalizable:

$$A = U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^{-1}.$$

We want to estimate the eigenvectors to find U, which allows us to perform the inner product in Claim 5.14.

Let $\{\mathbf{e_i}\}$ be the orthonormal basis of eigenvectors of A such that $A\mathbf{e_i} = \lambda_i \mathbf{e_i}$. We define

$$\mathbf{u} = \frac{1}{\sqrt{n}}\mathbf{v} = \sum_i a_i \mathbf{e_i}.$$

We want to estimate the coefficients a_i. Since A^2 is the matrix of codegrees, we can rewrite it as

$$A^2 = \begin{bmatrix} \frac{n}{4} & \cdots & \frac{n}{4} \\ \vdots & \ddots & \vdots \\ \frac{n}{4} & \cdots & \frac{n}{4} \end{bmatrix} + E$$

where E is an error matrix and we know from the codegree assumption that

$$\sum_{u,v} |E_{uv}| = o(n^3).$$

Then, we apply A^2 to the vector $\mathbf{u}$, giving

$$A^2\mathbf{u} = \frac{n^2}{4}\mathbf{u} + E\mathbf{u} = \frac{n^2}{4}\mathbf{u} + \mathbf{w} \text{ where } \mathbf{w} := E\mathbf{u}.$$

Next, we can bound the error vector $\mathbf{w}$.

$$\begin{aligned} |\mathbf{w}|^2 = |E\mathbf{u}|^2 &= \frac{1}{n}|E\mathbf{v}|^2 \\ &= \frac{1}{n}\sum_u \left(\sum_v \left|\operatorname{codeg}(u,v) - \frac{n}{4}\right|\right)^2 \\ &\overset{\text{CS}}{\leq} \frac{1}{n}\sum_u n \sum_v \left|\operatorname{codeg}(u,v) - \frac{n}{4}\right|^2 \\ &= \sum_{u,v} \left|\operatorname{codeg}(u,v) - \frac{n}{4}\right|^2 \\ &= o(n^4). \end{aligned}$$

Applying A^2 to $\mathbf{u}$ gives,

$$A^2\mathbf{u} = \sum_i \lambda_i^2 a_i \mathbf{e_i}.$$

Then, by the definition of $\mathbf{w}$, we expand $\mathbf{w}$ in the same eigenbasis:

$$\mathbf{w} = \sum_i (\lambda_i^2 - \frac{n^2}{4}) a_i \mathbf{e_i}.$$

Squaring both sides, we can estimate the sum with the bound on $\mathbf{w}^2$ found above:

$$\sum_i \left(\lambda_i^2 - \frac{n^2}{4}\right)^2 a_i^2 = |\mathbf{w}|^2 = o(n^4).$$

Then, we separate the first eigenvalue from the sum and substitute the value from Lemma 5.13 to have

$$\begin{aligned}
\sum_i \left(\lambda_i^2 - \frac{n^2}{4}\right)^2 a_i^2 &= \left(\lambda_1^2 - \frac{n^2}{4}\right)^2 a_1^2 + \sum_{i>1} \left(\lambda_i^2 - \frac{n^2}{4}\right)^2 a_i^2 \\
&\overset{5.13}{=} \left(\left[(1+o(1))\frac{n}{2}\right]^2 - \frac{n^2}{4}\right)^2 a_1^2 + \sum_{i>1} \left(o(n^2) - \frac{n^2}{4}\right)^2 a_i^2 \\
&= o(n^4) a_1^2 \quad + \quad (1+o(1))\frac{n^4}{16} \sum_{i>1} a_i^2 \\
&= (1+o(1))\frac{n^4}{16} \sum_{i>1} a_i^2 = o(n^4).
\end{aligned}$$

From this sum, we see that $a_2, ..., a_n = o(1)$, and thus $a_1 = 1 + o(1)$.

Finally, we are able to use the bounds on the coefficients to calculate the edge count.

By Claim 5.14,

$$E(G(n)) = \frac{1}{2} \sum_v \deg(v) = \frac{1}{2} \langle \mathbf{v}, A\mathbf{v} \rangle = \frac{n}{2} \langle \mathbf{u}, A\mathbf{u} \rangle.$$

We substitute the decomposition of $\mathbf{u}$, which gives the desired edge count:

$$\begin{aligned}
E(G(n)) &= \frac{n}{2} \langle \sum_i a_i \mathbf{e_i}, \sum_i \lambda_i a_i \mathbf{e_i} \rangle \\
&= \frac{n}{2} \sum_i \lambda_i a_i^2 \\
&= \frac{n}{2} \left((1+o(1)) \left[(1+o(1))\frac{n}{2}\right] o(1) + \sum_{i>1} o(n) \cdot o(1) \right) \\
&= (1+o(1))\frac{n^2}{4}.
\end{aligned}$$

□

Now, we can use the edge count to calculate the degree discrepancy.

Claim 5.16. *We have the following estimate for the sum of codegrees*

$$\sum_{u,v} \operatorname{codeg}(u,v) = \frac{n^3}{4} + o(n^3) = (1+o(1))\frac{n^3}{4}.$$

Proof. Assuming Property 3.3, we can remove the absolute value by the triangle inequality to have

$$\sum_{u,v}\left[\operatorname{codeg}(u,v)-\frac{n}{4}\right]=o(n^3).$$

Rearranging gives the desired count:

$$\sum_{u,v}\operatorname{codeg}(u,v)=\frac{n^3}{4}+o(n^3)=(1+o(1))\frac{n^3}{4}.$$

□

Lemma 5.17. *The degree discrepancy is bounded by $o(n^2)$:*

$$\sum_{x}\left|\deg(x)-\frac{n}{2}\right|=o(n^2).$$

Proof. By the edge count found above in Lemma 5.15 and that the sum of degrees is twice the edge count, we have the first moment of degree as:

$$\sum_{w}\deg(w)=(1+o(1))\frac{n^2}{2}.$$

Then, by the codegree count from Claim 5.16 and reindexing it in terms of degree, we have the second moment of degree as:

$$\sum_{w}\deg(w)^2=(1+o(1))\frac{n^3}{4}.$$

Then, using the discrepancy sum technique from Lemma 5.6, we substitute the first and second moment to get:

$$\begin{aligned}\sum_{x}\left|\deg(x)-\frac{n}{2}\right|&=\sqrt{n}\sqrt{\sum_{x}\deg(x)^2-n\sum_{x}\deg(x)+\sum_{x}\frac{n^2}{4}}\\&=\sqrt{n}\sqrt{\left((1+o(1))\frac{n^3}{4}\right)-n\left((1+o(1))\frac{n^2}{2}\right)+\frac{n^3}{4}}\\&=\sqrt{n}\sqrt{o(n^3)}=o(n^2).\end{aligned}$$

□

We now have all the information to count agreement discrepancy.

Claim 5.18. *The $s(u,v)$ can be expressed in terms of degree and codegree as:*

$$s(u,v)=n-\deg(u)-\deg(v)+2\operatorname{codeg}(u,v).$$

Proof. As shown in the table below, $s(u,v)$ is the sum of the bolded terms.

	adjacent to v	not adjacent to v
adjacent to u	$\boldsymbol{\operatorname{codeg}(u,v)}$	$\deg(u)-\operatorname{codeg}(u,v)$
not adjacent to u	$\deg(v)-\operatorname{codeg}(u,v)$	$\boldsymbol{n-\deg(u)-\deg(v)+\operatorname{codeg}(u,v)}$

□

Finally, we can prove Proposition 5.11.

Proof. Rearranging the terms from Claim 5.18, we have that

$$s(u,v) - \frac{n}{2} = -\Big(\deg(u) - \frac{n}{2}\Big) - \Big(\deg(v) - \frac{n}{2}\Big) + 2\Big(\operatorname{codeg}(u,v) - \frac{n}{4}\Big).$$

Hence by the triangle inequality,

$$\big|s(u,v) - \tfrac{n}{2}\big| \;\le\; \big|\deg(u) - \tfrac{n}{2}\big| + \big|\deg(v) - \tfrac{n}{2}\big| + 2\big|\operatorname{codeg}(u,v) - \tfrac{n}{4}\big|.$$

Now, we sum this inequality over all pairs of vertices:

$$\begin{aligned}\sum_{u,v}\big|s(u,v) - \tfrac{n}{2}\big| &\le \sum_{u,v}\big|\deg(u) - \tfrac{n}{2}\big| + \sum_{u,v}\big|\deg(v) - \tfrac{n}{2}\big| + 2\sum_{u,v}\big|\operatorname{codeg}(u,v) - \tfrac{n}{4}\big| \\ &= n\sum_{x}\big|\deg(x) - \tfrac{n}{2}\big| + n\sum_{x}\big|\deg(x) - \tfrac{n}{2}\big| + 2\sum_{u,v}\big|\operatorname{codeg}(u,v) - \tfrac{n}{4}\big| \\ &= 2n\sum_{x}\big|\deg(x) - \tfrac{n}{2}\big| + 2\sum_{u,v}\big|\operatorname{codeg}(u,v) - \tfrac{n}{4}\big|.\end{aligned}$$

Then, we substitute the bounds on degree discrepancy (Lemma 5.17) and codegree discrepancy (Property 3.3), which shows that:

$$\begin{aligned}\sum_{u,v}\big|s(u,v) - \tfrac{n}{2}\big| \;&\le\; 2n\cdot o(n^2) + 2\cdot o(n^3) = o(n^3) + o(n^3) \\ &= o(n^3).\end{aligned}$$

□

5.4. **Agreement Discrepancy Implies Induced Subgraph Count.**

Proposition 5.19. *If a graph sequence satisfies the Agreement Discrepancy property (Property 3.4), then it satisfies the Induced Subgraph Count property (Property 3.1).*

We will prove this by induction on the number of vertices in the induced subgraph.

Definition 5.20. For some value p, we define the notation $p_{(r)} = p(p-1)(p-2)(p-3)...(p-r+1)$.

Proposition 5.21. *Let $t \in \mathbb{N}$ and fix $M(t)$ to be a graph with vertices $\{v_1, v_2, ..., v_t\}$. Then for $r \le t$, $M_r(t)$ is the induced subgraph of $M(t)$ formed by vertices $v_1, ..., v_r$. Then, the count of $M_r(t)$ as an induced subgraph of G is*

$$N^*(M_r(t)) = (1 + o(1))n_{(r)}2^{-\binom{r}{2}}.$$

We will prove Proposition 5.21 by induction on r, which will prove Proposition 5.19.

For $r = 1$, the base case, the induced subgraph count is

$$N^*(M_1(t)) = (1 + o(1))n_{(1)}2^{-\binom{1}{2}} = n.$$

This is because 1-vertex graph appears n times in $G(n)$ since there are n vertices in $G(n)$.

Now, we prove the inductive step.

Let $\alpha = (\alpha_1, \alpha_2, ..., \alpha_r)$ be a list of r vertices of the larger graph $G(n)$. Let $\epsilon = (\epsilon_1, \epsilon_2, ..., \epsilon_r)$ where $\epsilon_1, ..., \epsilon_r \in \{0, 1\}$.

Definition 5.22. We define, for a fixed copy α of $M_r(t)$ and a given list ϵ, the extension function:

$$f(\alpha, \epsilon) = |\{v \in V \mid v \notin \alpha \text{ and } a(v, \alpha_j) = \epsilon_j \text{ for all } 1 \leq j \leq r\}|$$

In other words, $f(\alpha, \epsilon)$ is the number of vertices in G which are adjacent or not to each vertex in α as prescribed by ϵ.

Assume that Proposition 5.21 is true for some r. Then, we have

$$N^*(M_{r+1}(t)) = \sum_{\alpha \text{ a copy of } M_r(t)} f(\alpha, \epsilon) \text{ where } \epsilon_j = a(v_{r+1}, \alpha_j).$$

Now, we can consider $f(\alpha, \epsilon)$ as a random variable over the space of all sets α and ϵ, of which there are $n_{(r)}2^r$ possibilities. Therefore, we can calculate the expected value of f.

Claim 5.23. *The average extension count is*

$$\mathbb{E}(f) = \frac{n - r}{2^r}.$$

Proof. By the definition of expected value, we sum over all possible values and divide by the number of possibilities:

$$\mathbb{E}(f) = \frac{1}{n_{(r)}2^r} \sum_{\alpha, \epsilon} f(\alpha, \epsilon) = \frac{1}{n_{(r)}2^r} \sum_{\alpha} \sum_{\epsilon} f(\alpha, \epsilon).$$

First, we see that the sum $\sum_{\epsilon} f(\alpha, \epsilon)$ is simply the number of all vertices not in α, which there are $n - r$ of. Then, the number of ways to choose r vertices for α is $n_{(r)}$. Therefore, we have the following expected value.

$$\mathbb{E}(f) = \frac{1}{n_{(r)}2^r} n_{(r)}(n - r) = \frac{n - r}{2^r}.$$

□

The expected value of $f(\alpha, \epsilon)$ gives the first moment, but we also want to calculate the second moment to understand the variance.

Definition 5.24. We define the quantity

$$S_r = \sum_{\alpha, \epsilon} f(\alpha, \epsilon)(f(\alpha, \epsilon) - 1).$$

Claim 5.25. *S_r can equivalently be expressed in terms of the agreement function as*

$$S_r = \sum_{u \neq v} s(u, v)_{(r)}.$$

Proof. S_r can be interpreted as the sum, over all α and ϵ, of the number of ways to choose two distinct vertices which are adjacent to each vertex of α according to ϵ. Since we sum over all possible ϵ, this is equivalent to counting, for each α, the number of pairs of vertices that are adjacent or not adjacent to each vertex in α in the same way.

Equivalently, we can view S_r as the sum over all pairs of vertices u, v of the number of choices of α for which each vertex in α has the same adjacency relation to both u and v. In this form, the value of S_r depends on $s(u, v)$. □

Lemma 5.26. *We have the following estimate for the S_r quantity:*
$$S_r = (1+o(1))n^{r+2}2^{-r}.$$

Proof. Let $\delta_{uv} = s(u,v) - \frac{n}{2}$. Expanding $s(u,v)_{(r)}$ and substituting δ_{uv}, we have

$$\begin{aligned} s(u,v)_{(r)} &= \left(\frac{n}{2} + \delta_u v\right)\left(\frac{n}{2} + \delta_u v - 1\right) ... \left(\frac{n}{2} + \delta_u v - r + 1\right) \\ &= \sum_{k=0}^{r} c_k \left(\frac{n}{2}\right)^k \delta_{uv}^{r-k} \end{aligned}$$

where c_k is an appropriate constant. Then, by substituting into Claim 5.25,

$$\begin{aligned} S_r = \sum_{k=0}^{r}\sum_{u\neq v} c_k \left(\frac{n}{2}\right)^k \delta_{uv}^{r-k} &= \sum_{k=0}^{r} c_k \left(\frac{n}{2}\right)^k \sum_{u\neq v} \delta_{uv}^{r-k} \\ &= \left(\frac{n}{2}\right)^r \sum_{u\neq v} \delta_{uv}^{0} + \sum_{k=0}^{r-1} c_k \left(\frac{n}{2}\right)^k \sum_{u\neq v} \delta_{uv}^{r-k} \\ &= \left(\frac{n}{2}\right)^r n_{(2)} + \sum_{k=0}^{r-1} c_k \left(\frac{n}{2}\right)^k \sum_{u\neq v} \delta_{uv}^{r-k}. \end{aligned}$$

From the assumption of Agreement Discrepancy (Property 3.4), we have that
$$\sum_{u\neq v} \delta_{uv}^{r-k} = o(n^3).$$

Therefore, the second term of S_r is contained in $o(n^{r+2})$. Therefore, we have that:

$$S_r = (1+o(1))n^{r+2}2^{-r}.$$

□

Finally, we can use S_r to estimate the difference between the expected value of $f(\alpha,\epsilon)$ and the actual value.

Lemma 5.27. *We can estimate the difference between the expected value of $f(\alpha,\epsilon)$ as the following:*
$$\sum_{\alpha,\epsilon} |f(\alpha,\epsilon) - \mathbb{E}(f)| = o(n^{r+1}).$$

Proof. By Cauchy-Schwarz, we have the following:

$$\begin{aligned} \left[\sum_{\alpha,\epsilon} |f(\alpha,\epsilon) - \mathbb{E}(f)|\right]^2 &\overset{5.1}{\leq} n_{(r)}2^r \sum_{\alpha,\epsilon} [f(\alpha,\epsilon) - \mathbb{E}(f)]^2 \\ &= n_{(r)}2^r \sum_{\alpha,\epsilon} \left[f(\alpha,\epsilon)^2 - 2f(\alpha,\epsilon)\mathbb{E}(f) + \mathbb{E}(f)^2\right] \\ &= n_{(r)}2^r \left(\sum_{\alpha,\epsilon} \left[f(\alpha,\epsilon)^2\right] - 2\sum_{\alpha,\epsilon} [f(\alpha,\epsilon)\mathbb{E}(f)] + \sum_{\alpha,\epsilon} \left[\mathbb{E}(f)^2\right]\right). \end{aligned}$$

By definition,
$$\mathbb{E}(f) = \frac{1}{n_{(r)}2^r} \sum_{\alpha,\epsilon} f(\alpha,\epsilon).$$

Therefore,

$$\sum_{\alpha,\epsilon}[f(\alpha,\epsilon)\mathbb{E}(f)] = \sum_{\alpha,\epsilon}\left[\mathbb{E}(f)^2\right].$$

Thus, the middle and last terms combine:

$$n_{(r)}2^r\left(\sum_{\alpha,\epsilon}\left[f(\alpha,\epsilon)^2\right] - 2\sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2 + \sum_{\alpha,\epsilon}\left[\mathbb{E}(f)^2\right]\right) = n_{(r)}2^r\left(\sum_{\alpha,\epsilon}\left[f(\alpha,\epsilon)^2\right] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\right).$$

Then, we add and subtract a term, then rearrange and substitute S_r:

$$\begin{aligned}
n_{(r)}2^r\left(\sum_{\alpha,\epsilon}\left[f(\alpha,\epsilon)^2\right] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\right) &= n_{(r)}2^r\Bigg(\sum_{\alpha,\epsilon}\left[f(\alpha,\epsilon)^2\right] - \sum_{\alpha,\epsilon}[f(\alpha,\epsilon)] \\
&\qquad + \sum_{\alpha,\epsilon}[f(\alpha,\epsilon)] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\Bigg) \\
&= n_{(r)}2^r\Bigg(\sum_{\alpha,\epsilon}\left[f(\alpha,\epsilon)^2 - f(\alpha,\epsilon)\right] \\
&\qquad + \sum_{\alpha,\epsilon}[f(\alpha,\epsilon)] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\Bigg) \\
&= n_{(r)}2^r\left(S_r + \sum_{\alpha,\epsilon}[f(\alpha,\epsilon)] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\right).
\end{aligned}$$

Then, we substitute the value of S_r and the expected value of f:

$$\begin{aligned}
n_{(r)}2^r\left(S_r + \sum_{\alpha,\epsilon}[f(\alpha,\epsilon)] - \sum_{\alpha,\epsilon}[\mathbb{E}(f)]^2\right) &= n_{(r)}2^r\Big(\left[(1+o(1))n^{r+2}2^{-r}\right] + \left[n_{(r)}(n-r)\right] \\
&\qquad - \left[n_{(r)}2^r(n-r)^2 2^{-2r}\right]\Big) \\
&= n_{(r)}2^r\Big(\left[n^{r+2}2^{-r}\right] + \left[n^{r+2}2^{-r}\right]o(1) + \left[n_{(r)}(n-r)\right] \\
&\qquad - \left[n_{(r)}(n-r)^2 2^{-r}\right]\Big) \\
&= n_{(r)}2^r\Big(o(n^{r+2})\Big) \\
&= o(n^{2r+2}).
\end{aligned}$$

Taking the square root, we have

$$\sum_{\alpha,\epsilon}|f(\alpha,\epsilon) - \mathbb{E}(f)| = o(n^{r+1}).$$

$\square$

We denote the difference

$$\Delta(\alpha,\epsilon) = f(\alpha,\epsilon) - \mathbb{E}(f).$$

Thus,

$$f(\alpha,\epsilon) = \mathbb{E}(f) + \Delta(\alpha,\epsilon) = \frac{n-r}{2^r} + \Delta(\alpha,\epsilon).$$

Claim 5.28. *Proposition 5.21 is true for $r+1$. Namely,*

$$N^*(M_{r+1}(t)) = (1+o(1))n_{(r+1)}2^{-\binom{r+1}{2}}.$$

Proof. Substituting the Δ notation introduced above, we have

$$N^*(M_{r+1}(t)) = \sum_{\alpha \text{ a copy of } M_r(t)} \left[\frac{n-r}{2^r} + \Delta(\alpha, \epsilon)\right].$$

By the inductive hypothesis, this sum has $(1+o(1))n_{(r)}2^{-\binom{r}{2}}$ terms. Thus,

$$N^*(M_{r+1}(t)) = (1+o(1))n_{(r)}2^{-\binom{r}{2}}\left[\frac{n-r}{2^r}\right] + \sum_{\alpha \text{ a copy of } M_r(t)} \left[\Delta(\alpha, \epsilon)\right].$$

Then, by the fact that

$$\binom{r+1}{2} = r + \binom{r}{2},$$

we can simplify and substitute the bound on the Δ error term from Lemma 5.27:

$$\begin{aligned} N^*(M_{r+1}(t)) &= (1+o(1))n_{(r+1)}2^{-\binom{r+1}{2}} + o(n^{r+1}) \\ &= (1+o(1))n_{(r+1)}2^{-\binom{r+1}{2}}. \end{aligned}$$

□

Therefore, the induction is completed, proving Proposition 5.19.

We have established the following cycle of implications:

$$\begin{tikzcd} \boxed{\text{Property 3.1}} \arrow[r] & \boxed{\text{Property 3.2}} \arrow[d] \\ \boxed{\text{Property 3.4}} \arrow[u] & \boxed{\text{Property 3.3}} \arrow[l] \end{tikzcd}$$

Therefore, we can conclude that all four properties are equivalent, completing the proof of Theorem 3.5. □

6. Example of Quasirandom Graph Sequence

We now give a concrete example of a quasirandom graph sequence, namely the *Paley graph sequence* defined using finite fields and quadratic residues.

Definition 6.1. Let p be a prime with $p \equiv 1 \pmod 4$, and let $\mathbb{F}_p$ denote the finite field of order p.

Definition 6.2. An element $a \in \mathbb{F}_p^\times$ is called a *quadratic residue* if there exists $b \in \mathbb{F}_p^\times$ such that $a = b^2$. Otherwise, a is a *quadratic nonresidue*. Let R denote the set of quadratic residues.

Definition 6.3. The *Paley graph* Q_p is the graph with vertex set numbered by $\mathbb{F}_p$ in which $x, y \in \mathbb{F}_p$ are adjacent if and only if $x - y$ is a quadratic residue in $\mathbb{F}_p$.

The assumption $p \equiv 1 \pmod 4$ guarantees that -1 is a quadratic residue, and therefore adjacency is symmetric: $x-y$ is a quadratic residue if and only if $y-x$ is a quadratic residue. This is because the quadratic residues are a group: since $\frac{x-y}{y-x} = -1$, if the numerator is a quadratic residue, the denominator is, and vice versa.

Theorem 6.4. *The sequence of Paley graphs (Q_p), indexed by primes $p \equiv 1 \pmod 4$, is quasirandom.*

Proof. We verify that this sequence is quasirandom by checking Property 3.4 (Agreement Discrepancy).

Fix vertices x, y and choose a third vertex z. We claim that $a(x,z) = a(y,z)$ if and only if

$$\frac{x-z}{y-z} \in R.$$

If both $x-z$ and $y-z$ are quadratic residues, their quotient is a quadratic residue since the quadratic residues form a multiplicative subgroup. Also, note that if both $x-z$ and $y-z$ are quadratic nonresidues, their quotient is also a quadratic residue.

Therefore, for each quadratic residue $\beta \in R$, we can solve

$$\frac{x-z}{y-z} = \beta \quad \Longrightarrow \quad z = y - \frac{x-y}{\beta - 1}.$$

Thus, there is a bijection between vertices z where $a(x,z) = a(y,z)$ and the set of quadratic residues $R \setminus \{1\}$ (since 1 causes the denominator to be 0). Since there are $\frac{p-1}{2}$ elements in the set of quadratic residues, there are exactly $\frac{p-3}{2}$ such vertices z.

Therefore, the agreement function is

$$s(x,y) = \begin{cases} p & x = y \\ \frac{p-3}{2} & x \neq y \end{cases}.$$

Then, we calculate the agreement discrepancy

$$\begin{aligned} \sum_{x,y} \left| s(x,y) - \frac{p}{2} \right| &= \sum_{x} \left| s(x,x) - \frac{p}{2} \right| + \sum_{x,y} \left| s(x,y) - \frac{p}{2} \right| \\ &= \sum_{x} \left| p - \frac{p}{2} \right| + \sum_{x,y} \left| \frac{p-3}{2} - \frac{p}{2} \right| \\ &= \frac{p^2}{2} + \frac{3}{2} p(p-1) = o(p^3). \end{aligned}$$

This is the correct bound, so Property 3.4 is satisfied, and thus the sequence of Paley graphs is quasirandom. $\square$

Despite being entirely deterministic, the Paley graphs, among many other examples, are quasirandom and therefore exhibit many of the same properties as random graph sequences.

7. Acknowledgments

It is my great pleasure to thank my mentor, Miles Kretschmer, for his help throughout the process of investigating this topic and writing this paper. His support and feedback was invaluable and gave me a thorough and clear understanding of the background for the material and the techniques used in the results. Furthermore, I'd like to thank Peter May for all his work organizing the University of Chicago Mathematics REU and Daniil Rudenko for his excellent apprentice program lectures.

References

[1] F. R. K. Chung, R. L. Graham, and R. M. Wilson Quasi-Random Graphs Combinatorica. 1989.

[2] Joel Spencer and Noga Alon. The Probabilistic Method. Wiley-Interscience. 1990.

[3] Yufei Zhao. Pseudorandom Graphs I: Quasirandomness. Graph Theory and Additive Combinatorics. MIT OpenCourseWare. 2023. https://ocw.mit.edu/courses/18-225-graph-theory-and-additive-combinatorics-fall-2023/resources/lecture-11-pseudorandom-graphs-i-quasirandomness/