Finding and Counting Substructures in Graphs and Hypergraphs

Liam Hardiman

December 10, 2020

Quick Definitions

- A **graph** $G = (V, E)$ consists of a (finite) set of **vertices** V and a set E of unordered pairs of vertices called **edges**.

Quick Definitions

- A **graph** $G = (V, E)$ consists of a (finite) set of **vertices** V and a set E of unordered pairs of vertices called **edges**.
- If $X \subseteq V$, then the **subgraph induced by** X, $G[X]$ is the graph with vertex set X and all edges from E that have both ends in X.

Quick Definitions

- A **graph** $G = (V, E)$ consists of a (finite) set of **vertices** V and a set E of unordered pairs of vertices called **edges**.
- If $X \subseteq V$, then the **subgraph induced by** X, $G[X]$ is the graph with vertex set X and all edges from E that have both ends in X.
- The **degree** of a vertex v, denoted $d_G(v)$, is the number of edges of G that v appears in.

Quick Definitions

- A **graph** $G = (V, E)$ consists of a (finite) set of **vertices** V and a set E of unordered pairs of vertices called **edges**.
- If $X \subseteq V$, then the **subgraph induced by** X, $G[X]$ is the graph with vertex set X and all edges from E that have both ends in X.
- The **degree** of a vertex v, denoted $d_G(v)$, is the number of edges of G that v appears in.
- picture here

History and Motivation

Theorem (L. Lovász, T. Gallai - 1979)

Let $G = (V, E)$ be any graph. Then G admits a partitioning of its vertex set into two parts, $V = V_1 \cup V_2$, so that each vertex in $G[V_1]$ and each vertex in $G[V_2]$ has even degree. In particular, any graph on n vertices has an even subgraph of order at least $n/2$.

History and Motivation

Theorem (L. Lovász, T. Gallai - 1979)

Let $G = (V, E)$ be any graph. Then G admits a partitioning of its vertex set into two parts, $V = V_1 \cup V_2$, so that each vertex in $G[V_1]$ and each vertex in $G[V_2]$ has even degree. In particular, any graph on n vertices has an even subgraph of order at least $n/2$.

Proof sketch:

asdf $\square$

History and Motivation

- What about odd subgraphs?

History and Motivation

- What about odd subgraphs?

Theorem (A. Scott - 1992)

Every graph $G(V,E)$ on n vertices, none of which are isolated, contains a set $W \subseteq V(G)$ such that $|W| \geq \frac{n}{900 \log n}$ and $G[W]$ has all degrees odd.

History and Motivation

- What about odd subgraphs?

Theorem (A. Scott - 1992)

Every graph $G(V,E)$ on n vertices, none of which are isolated, contains a set $W \subseteq V(G)$ such that $|W| \geq \frac{n}{900 \log n}$ and $G[W]$ has all degrees odd.

Conjecture (A. Scott - 2001)

There exists some constant $c > 0$ such that every graph $G(V,E)$ on n vertices, none of which are isolated, contains a set $W \subseteq V(G)$ such that $|W| \geq cn$ and $G[W]$ has all degrees odd.

History and Motivation

- What about odd subgraphs?

Theorem (A. Scott - 1992)

Every graph $G(V, E)$ on n vertices, none of which are isolated, contains a set $W \subseteq V(G)$ such that $|W| \geq \frac{n}{900 \log n}$ and $G[W]$ has all degrees odd.

Conjecture (A. Scott - 2001)

There exists some constant $c > 0$ such that every graph $G(V, E)$ on n vertices, none of which are isolated, contains a set $W \subseteq V(G)$ such that $|W| \geq cn$ and $G[W]$ has all degrees odd.

History and Motivation

Definition

The **Erdős-Renyi random graph**, $\mathcal{G}(n, p)$, is the random variable that outputs a graph on n vertices, any two of which are independently connected with probability p.

History and Motivation

Definition

The **Erdős-Renyi random graph**, $\mathcal{G}(n, p)$, is the random variable that outputs a graph on n vertices, any two of which are independently connected with probability p.

Theorem (A. Scott - 1992)

Let $G \sim \mathcal{G}(n, 1/2)$. Then with high probability (that is, with probability $1 - o(1)$), G has an induced subgraph on at least $0.7729n$ vertices with all degrees odd.

Other Moduli and Remainders

Conjecture (A. Scott - 2001)

For any positive integer q at least 2, there exists a constant c_q so that every graph on n vertices without isolated vertices has an induced subgraph on at least $c_q n$ vertices with all degrees $1 \pmod{q}$.

Other Moduli and Remainders

Conjecture (A. Scott - 2001)

For any positive integer q at least 2, there exists a constant c_q so that every graph on n vertices without isolated vertices has an induced subgraph on at least $c_q n$ vertices with all degrees $1 \pmod q$.

Theorem (A. Ferber, H., M. Krivelevich - 2020+)

Let $q \geq 2$ and let r be an integer. Then there exists a constant c_q such that, with high probability, the random graph $G \sim \mathcal{G}(n, 1/2)$ has an induced subgraph on at least $c_q n$ vertices wijth all degrees $r \pmod q$.

Key Idea Behind Proof

- If $G \sim \mathcal{G}(n, 1/2)$, then its adjacency matrix M ($M_{ij} = 1$ if and only vertices i and j are connected and $M_{ij} = 0$ otherwise) is a random symmetric $n \times n$ matrix whose above-diagonal entries are iid Bern$(1/2)$ random variables.

Key Idea Behind Proof

- If $G \sim \mathcal{G}(n, 1/2)$, then its adjacency matrix M ($M_{ij} = 1$ if and only vertices i and j are connected and $M_{ij} = 0$ otherwise) is a random symmetric $n \times n$ matrix whose above-diagonal entries are iid Bern$(1/2)$ random variables.
- In this case, the i-th entry of $M\mathbf{1}$ is the degree of vertex i in G, where $\mathbf{1}$ is the length n all-1 vector.

Key Idea Behind Proof

- If $G \sim \mathcal{G}(n, 1/2)$, then its adjacency matrix M ($M_{ij} = 1$ if and only vertices i and j are connected and $M_{ij} = 0$ otherwise) is a random symmetric $n \times n$ matrix whose above-diagonal entries are iid Bern$(1/2)$ random variables.
- In this case, the i-th entry of $M\mathbf{1}$ is the degree of vertex i in G, where $\mathbf{1}$ is the length n all-1 vector.

Observation

G contains a subgraph with all degrees $r \pmod q$ if and only if its adjacency matrix contains a principal submatrix B satisfying $B\mathbf{1} \equiv r\mathbf{1} \pmod q$.

A Useful Computation

Fix a positive integer $q \geq 2$ and let r be an integer. Let $x_1, \ldots, x_t$ be iid Bern$(1/2)$ random variables.

A Useful Computation

Fix a positive integer $q \geq 2$ and let r be an integer. Let $x_1, \ldots, x_t$ be iid Bern$(1/2)$ random variables.

$$\Pr[x_1 + \cdots + x_t \equiv r] = \mathbb{E}[\delta_0(x_1 + \cdots + x_t - r)]$$

A Useful Computation

Fix a positive integer $q \geq 2$ and let r be an integer. Let $x_1, \ldots, x_t$ be iid Bern$(1/2)$ random variables.

$$\begin{aligned}
\Pr[x_1 + \cdots + x_t \equiv r] &= \mathbb{E}[\delta_0(x_1 + \cdots + x_t - r)] \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} \mathbb{E} \exp\left[\frac{2\pi i \ell}{q}(x_1 + \cdots + x_t - r)\right]
\end{aligned}$$

A Useful Computation

Fix a positive integer $q \geq 2$ and let r be an integer. Let $x_1, \dots, x_t$ be iid Bern$(1/2)$ random variables.

$$
\begin{aligned}
\Pr[x_1 + \cdots + x_t \equiv r] &= \mathbb{E}[\delta_0(x_1 + \cdots + x_t - r)] \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} \mathbb{E} \exp\left[\frac{2\pi i \ell}{q}(x_1 + \cdots + x_t - r)\right] \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} e^{-2\pi i r \ell / q} \prod_{j=1}^{t} \mathbb{E} e^{2\pi i \ell x_j / q}
\end{aligned}
$$

A Useful Computation

Fix a positive integer $q \geq 2$ and let r be an integer. Let $x_1, \ldots, x_t$ be iid Bern$(1/2)$ random variables.

$$\begin{aligned}
\Pr[x_1 + \cdots + x_t \equiv r] &= \mathbb{E}[\delta_0(x_1 + \cdots + x_t - r)] \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} \mathbb{E} \exp \left[\frac{2\pi i \ell}{q} (x_1 + \cdots + x_t - r) \right] \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} e^{-2\pi i r \ell / q} \prod_{j=1}^{t} \mathbb{E} e^{2\pi i \ell x_j / q} \\
&= \frac{1}{q} \sum_{\ell \in \mathbb{Z}_q} e^{-2\pi i r \ell / q} \left(\frac{1 + e^{2\pi i \ell / q}}{2} \right)^t .
\end{aligned}$$

A Useful Computation

- Isolate the $\ell \equiv 0$ term, apply the triangle inequality and Euler's identity.

A Useful Computation

- Isolate the $\ell \equiv 0$ term, apply the triangle inequality and Euler's identity.

$$\left|\Pr[x_1 + \cdots + x_t \equiv r] - \frac{1}{q}\right| \leq \frac{1}{q} \sum_{\ell=1}^{q-1} |\cos(\pi\ell/q)|^t.$$

A Useful Computation

- Isolate the $\ell \equiv 0$ term, apply the triangle inequality and Euler's identity.

$$\left|\Pr[x_1 + \cdots + x_t \equiv r] - \frac{1}{q}\right| \leq \frac{1}{q}\sum_{\ell=1}^{q-1} |\cos(\pi\ell/q)|^t.$$

- Calculus trick: when $|x| \leq \pi/2$, $|\sin x| \geq (2/\pi)|x|$.

A Useful Computation

- Isolate the $\ell \equiv 0$ term, apply the triangle inequality and Euler's identity.

$$\left|\Pr[x_1 + \cdots + x_t \equiv r] - \frac{1}{q}\right| \leq \frac{1}{q}\sum_{\ell=1}^{q-1} |\cos(\pi\ell/q)|^t.$$

- Calculus trick: when $|x| \leq \pi/2$, $|\sin x| \geq (2/\pi)|x|$.
- From this, it follows that $|\cos(\pi\ell/q)| \leq e^{-2/q^2}$ for all $\ell = 1, \ldots, q-1$.

A Useful Computation

- Isolate the $\ell \equiv 0$ term, apply the triangle inequality and Euler's identity.

$$\left|\Pr[x_1 + \cdots + x_t \equiv r] - \frac{1}{q}\right| \leq \frac{1}{q}\sum_{\ell=1}^{q-1} |\cos(\pi\ell/q)|^t.$$

- Calculus trick: when $|x| \leq \pi/2$, $|\sin x| \geq (2/\pi)|x|$.
- From this, it follows that $|\cos(\pi\ell/q)| \leq e^{-2/q^2}$ for all $\ell = 1, \ldots, q-1$.

$$\left|\Pr[x_1 + \cdots + x_t \equiv r] - \frac{1}{q}\right| \leq \frac{q-1}{q} e^{-2t/q^2}.$$

A Useful Computation

- The distribution of a random Bernoulli sum modulo q is asymptotically uniform!

A Useful Computation

- The distribution of a random Bernoulli sum modulo q is asymptotically uniform!
- By independence we have.

Lemma

Let M be an $s \times t$ matrix whose entries are iid Bern$(1/2)$ random variables. Then for any $v \in \mathbb{Z}_q^s$,

$$\Pr[M\mathbf{1} \equiv v] = \frac{1}{q^s}\left(1 + O\left(e^{-2t/q^2}\right)\right)^s.$$

Symmetric Case

- But the adjacency matrix is symmetric!

Symmetric Case

- But the adjacency matrix is symmetric!
- A similar, but more technical argument handles the symmetric case.

Symmetric Case

- But the adjacency matrix is symmetric!
- A similar, but more technical argument handles the symmetric case.

Lemma

Let M be an $m \times m$ symmetric matrix whose diagonal is zero and whose entries above the diagonal are iid $Bern(1/2)$ random variables. Then for any $v \in \mathbb{Z}_q^m$,

$$\Pr[M\mathbf{1} \equiv v] = \begin{cases} \frac{1}{q^m}(1 + O(e^{-m/q^2})) & \text{if } q \text{ is odd} \\ \frac{2}{q^m}(1 + O(e^{-m/q^2})) & \text{if } q \text{ is even and } \sum v_i \text{ is even} \\ 0 & \text{if } q \text{ is even and } \sum v_i \text{ is odd.} \end{cases}$$

Parity?

- We have
$$\sum_i (M\mathbf{1})_i = \sum_{i,j} M_{ij} = 2\sum_{i<j} M_{ij},$$
which is always even modulo q if q is even and can be any residue modulo odd q.

Parity?

- We have
$$\sum_i (M\mathbf{1})_i = \sum_{i,j} M_{ij} = 2\sum_{i<j} M_{ij},$$
which is always even modulo q if q is even and can be any residue modulo odd q.
- The symmetric lemma then says that the distribution of $M\mathbf{1} \pmod{q}$ is asymptotically uniform over all "feasible" values.

Proof (Sketch) of Theorem

Theorem (Chebyshev's Inequality)

Let X be a nonnegative integer-valued random variable with finite variance. Then

$$\Pr[X > 0] \geq 1 - \frac{Var[X]}{(\mathbb{E}[X])^2}.$$

Proof (Sketch) of Theorem

Theorem (Chebyshev's Inequality)

Let X be a nonnegative integer-valued random variable with finite variance. Then

$$\Pr[X > 0] \geq 1 - \frac{Var[X]}{(\mathbb{E}[X])^2}.$$

- Let $k = c_q n$ for some $c_q > 0$ and let X_k be the number of $k \times k$ principal submatrices B of the adjacency matrix of G satisfying $B\mathbf{1} \equiv r\mathbf{1} \pmod{q}$.

Proof (Sketch) of Theorem

Theorem (Chebyshev's Inequality)

Let X be a nonnegative integer-valued random variable with finite variance. Then

$$\Pr[X > 0] \geq 1 - \frac{Var[X]}{(\mathbb{E}[X])^2}.$$

- Let $k = c_q n$ for some $c_q > 0$ and let X_k be the number of $k \times k$ principal submatrices B of the adjacency matrix of G satisfying $B\mathbf{1} \equiv r\mathbf{1} \pmod{q}$.
- Show that c_q can be chosen so that $\mathrm{Var}[X_k] = o(E[X_k]^2)$. Then G has an induced subgraph of size $c_q n$ with high probability.

Proof (Sketch) of Theorem

Maybe have some of the calculations and that overlap diagram.

State theorem for arbitrary distribution of residues mod q.

Packing

- Recall Gallai's result on even subgraphs.

Theorem (L. Lovász, T. Gallai - 1979)

Let $G = (V, E)$ be any graph. Then G admits a partitioning of its vertex set into two parts, $V = V_1 \cup V_2$, so that each vertex in $G[V_1]$ and each vertex in $G[V_2]$ has even degree. In particular, any graph on n vertices has an even subgraph of order at least $n/2$.

Packing

- Recall Gallai's result on even subgraphs.

Theorem (L. Lovász, T. Gallai - 1979)

Let $G = (V, E)$ be any graph. Then G admits a partitioning of its vertex set into two parts, $V = V_1 \cup V_2$, so that each vertex in $G[V_1]$ and each vertex in $G[V_2]$ has even degree. In particular, any graph on n vertices has an even subgraph of order at least $n/2$.

- What else can we say about partitions and remainders?

Packing Theorems

Theorem (A. Scott - 2001)

Let G be a graph. Then G admits a partition into odd subgraphs if and only if every component of G has even order.

Packing Theorems

Theorem (A. Scott - 2001)

Let G be a graph. Then G admits a partition into odd subgraphs if and only if every component of G has even order.

Theorem (A. Scott - 2001)

With high probability, $G \sim \mathcal{G}(n, p)$ for n even admits a partition into three odd subgraphs.

Packing Theorems

Theorem (A. Scott - 2001)

Let G be a graph. Then G admits a partition into odd subgraphs if and only if every component of G has even order.

Theorem (A. Scott - 2001)

With high probability, $G \sim \mathcal{G}(n,p)$ for n even admits a partition into three odd subgraphs.

Conjecture (A. Scott - 2001)

For $q > 1$ and $r \leq q$, there exists a constant c_q such that $G \sim \mathcal{G}(n,p)$ admits a partition into c_q classes such that the degrees in each class are congruent to $r \pmod q$ with high probability.

Packing Theorems

Theorem (A. Ferber, H., M. Krivelevich)

For $q > 1$ and $r \leq q$, $G \sim \mathcal{G}(n, 1/2)$ admits a partition into $q+1$ classes such that the degrees in each class are congruent to $r \pmod{q}$ with high probability.

Sketch of Packing Proof

Explain that it's just another second moment argument. Maybe some calculations, diagrams. Explain need for "both ways" lemma.

1. A Finding Problem
2. A Counting Problem
3. Another Counting Problem

Definitions

Definition

Let $G = (V, E)$ be a graph. A subset of edges $\mathcal{M} \subseteq E$ is called a **matching** if its edges are vertex-disjoint. The matching $\mathcal{M}$ is **perfect** if the vertices that comprise it cover all of V.

Definitions

Definition

Let $G = (V, E)$ be a graph. A subset of edges $\mathcal{M} \subseteq E$ is called a **matching** if its edges are vertex-disjoint. The matching $\mathcal{M}$ is **perfect** if the vertices that comprise it cover all of V.

Definition

Let $G = (V, E)$ be a graph and let $\mathcal{C}$ be a sequence of edges. Then $\mathcal{C}$ is a...

- ...**path** if it joins a sequence of distinct vertices.

Definitions

Definition

Let $G = (V, E)$ be a graph. A subset of edges $\mathcal{M} \subseteq E$ is called a **matching** if its edges are vertex-disjoint. The matching $\mathcal{M}$ is **perfect** if the vertices that comprise it cover all of V.

Definition

Let $G = (V, E)$ be a graph and let $\mathcal{C}$ be a sequence of edges. Then $\mathcal{C}$ is a...

- ...**path** if it joins a sequence of distinct vertices.
- ...**cycle** if it joints a sequence of vertices $v_1, v_2, \ldots, v_k, v_1$.

Definitions

Definition

Let $G = (V, E)$ be a graph. A subset of edges $\mathcal{M} \subseteq E$ is called a **matching** if its edges are vertex-disjoint. The matching $\mathcal{M}$ is **perfect** if the vertices that comprise it cover all of V.

Definition

Let $G = (V, E)$ be a graph and let $\mathcal{C}$ be a sequence of edges. Then $\mathcal{C}$ is a...

- ...**path** if it joins a sequence of distinct vertices.
- ...**cycle** if it joints a sequence of vertices $v_1, v_2, \ldots, v_k, v_1$.
- ...**Hamiltonian cycle** if it is a cycle that visits every vertex in G exactly once.

History and Motivation

- When does a graph have Hamiltonian a cycle (when is it **Hamiltonian**)?

History and Motivation

- When does a graph have Hamiltonian a cycle (when is it **Hamiltonian**)?
- For any graph G, let $\delta(G)$ denote the minimum degree of any vertex.

History and Motivation

- When does a graph have Hamiltonian a cycle (when is it **Hamiltonian**)?
- For any graph G, let $\delta(G)$ denote the minimum degree of any vertex.

Theorem (G. Dirac - 1952)

A graph on $n \geq 3$ vertices is Hamiltonian if $\delta(G) \geq n/2$.

History and Motivation

- When does a graph have Hamiltonian a cycle (when is it **Hamiltonian**)?
- For any graph G, let $\delta(G)$ denote the minimum degree of any vertex.

Theorem (G. Dirac - 1952)

A graph on $n \geq 3$ vertices is Hamiltonian if $\delta(G) \geq n/2$.

- A graph is Hamiltonian if it is sufficiently dense.

Counting Hamiltonian Cycles in Dense Graphs - Intuition

- The expected number of Hamiltonian cycles in $G \sim \mathcal{G}(n, p)$ is

$$\frac{1}{2}(n-1)!p^n.$$

Counting Hamiltonian Cycles in Dense Graphs - Intuition

- The expected number of Hamiltonian cycles in $G \sim \mathcal{G}(n, p)$ is

$$\frac{1}{2}(n-1)!p^n.$$

- Fix $\gamma > 0$. If $p > (1/2 + \gamma)$, then

Cuckler-Kahn

Introduce hypergraphs and notion of cycle

Hypergraph theorems

Our theorem

Proof sketch