# Randomized Algorithms

## The Probabilistic Method (I)

Joseph Chuang-Chieh Lin

Department of Computer Science & Engineering,
National Taiwan Ocean University

Spring 2026

# Outline

# Outline

# Motivation

- Prove the existence of objects.

# Motivation

- Prove the existence of objects.
- If the probability of selecting an object with the required properties is positive, then the sample space must contain such an object.

# Outline

### Example

- Coloring the edges of a graph with two **colors**.

  **Constraint:** no large cliques with all edges having the same color.

### Theorem 1

If $\binom{n}{k}2^{-\binom{k}{2}+1} < 1$, then it is possible to color the edges of $K_n$ with two colors so that it has no monochromatic $K_k$ subgraph.

### Theorem 1

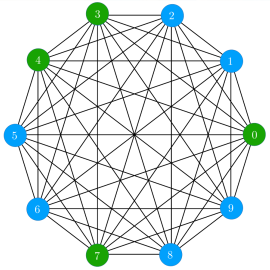
If $\binom{n}{k}2^{-\binom{k}{2}+1} < 1$, then it is possible to color the edges of $K_n$ with two colors so that it has no monochromatic $K_k$ subgraph.

- **Note:**
  - There are $2^{\binom{n}{2}}$ possible colorings of $K_n$.
  - There are $\binom{n}{k}$ different $K_k$ cliques of $K_n$.

### Theorem 1

If $\binom{n}{k}2^{-\binom{k}{2}+1} < 1$, then it is possible to color the edges of $K_n$ with two colors so that it has no monochromatic $K_k$ subgraph.

- **Note:**
  - There are $2^{\binom{n}{2}}$ possible colorings of $K_n$.
  - There are $\binom{n}{k}$ different $K_k$ cliques of $K_n$.
- Flip a fair coin independently to determine the color of each edge.



## General Idea

- Let $A_i$ be the event that clique $i$ is monochromatic.
- **Goal:** Prove that the probability $\Pr\left[\bigcap_{i=1}^{\binom{n}{k}} \overline{A_i}\right] > 0$.

# General Idea

- Let $A_i$ be the event that clique $i$ is monochromatic.
- **Goal:** Prove that the probability $\Pr\left[\bigcap_{i=1}^{\binom{n}{k}} \overline{A_i}\right] > 0$.
    - That is,

$$1 - \Pr\left[\bigcup_{i=1}^{\binom{n}{k}} A_i\right] > 0.$$

# Proof

- Once the first edge in clique $i$ is colored, the remaining $\binom{k}{2} - 1$ edges must all be given the same color. So,

# Proof

- Once the first edge in clique $i$ is colored, the remaining $\binom{k}{2} - 1$ edges must all be given the same color. So,

$$\Pr(A_i) = 2^{-\binom{k}{2}+1}.$$

# Proof

- Once the first edge in clique $i$ is colored, the remaining $\binom{k}{2} - 1$ edges must all be given the same color. So,

$$\Pr(A_i) = 2^{-\binom{k}{2}+1}.$$

- By the union bound,

$$\Pr\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \leq \sum_{I=1}^{\binom{n}{k}} \Pr(A_i) = \binom{n}{k} 2^{-\binom{k}{2}+1} < 1,$$

  - The last inequality: the assumptions of the theorem.

# Proof

- Once the first edge in clique $i$ is colored, the remaining $\binom{k}{2} - 1$ edges must all be given the same color. So,

$$\Pr(A_i) = 2^{-\binom{k}{2}+1}.$$

- By the union bound,

$$\Pr\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \leq \sum_{I=1}^{\binom{n}{k}} \Pr(A_i) = \binom{n}{k} 2^{-\binom{k}{2}+1} < 1,$$

  - The last inequality: the assumptions of the theorem.

- Hence

$$\Pr\left(\bigcap_{i=1}^{\binom{n}{k}} \overline{A_i}\right) = 1 - \Pr\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) > 0.$$

- Our calculations can be simplified if we note that:
  For $n \leq 2^{k/2}$ and $k \geq 3$,

$$\begin{aligned}\binom{n}{k}2^{-\binom{k}{2}+1} &\leq \frac{n^k}{k!}2^{-(k(k-1)/2)+1} \\ &\leq \frac{2^{k/2+1}}{k!} \\ &< 1.\end{aligned}$$

# A Monte Carlo construction

- Sample a coloring by coloring each edge independently at random.
  - How many samples must we generate before obtaining a sample satisfying our requirement?

# A Monte Carlo construction

- Sample a coloring by coloring each edge independently at random.
  - How many samples must we generate before obtaining a sample satisfying our requirement?
- $p$: The probability of obtaining a sample we desire.

# A Monte Carlo construction

- Sample a coloring by coloring each edge independently at random.
  - How many samples must we generate before obtaining a sample satisfying our requirement?
- $p$: The probability of obtaining a sample we desire.
- If $p = 1 - o(1)$, then the sampling algorithm is incorrect with probability $o(1)$.

# A Monte Carlo construction

- Sample a coloring by coloring each edge independently at random.
  - How many samples must we generate before obtaining a sample satisfying our requirement?
- $p$: The probability of obtaining a sample we desire.
- If $p = 1 - o(1)$, then the sampling algorithm is incorrect with probability $o(1)$.
- Example: Finding a coloring on a graph of 1,000 vertices with no monochromatic $K_{20}$.
  - The probability that a random coloring has a monochromatic $K_{20}$ is $\leq$

$$\frac{2^{20/2+1}}{20!} < 8.5 \cdot 10^{-16}.$$

# A Monte Carlo construction

- Sample a coloring by coloring each edge independently at random.
  - How many samples must we generate before obtaining a sample satisfying our requirement?
- $p$: The probability of obtaining a sample we desire.
- If $p = 1 - o(1)$, then the sampling algorithm is incorrect with probability $o(1)$.
- Example: Finding a coloring on a graph of 1,000 vertices with no monochromatic $K_{20}$.
  - The probability that a random coloring has a monochromatic $K_{20}$ is $\leq$

$$\frac{2^{20/2+1}}{20!} < 8.5 \cdot 10^{-16}.$$

$\Rightarrow$ A Monte Carlo algorithm with a very small probability of failure.

# Question

- What if we want a Las Vegas algorithm? Can we still have an expected polynomial running time randomized algorithm for generating such a sample?

# Question

- What if we want a Las Vegas algorithm? Can we still have an expected polynomial running time randomized algorithm for generating such a sample?
- Expected number of samples: $1/p$.

# Question

- What if we want a Las Vegas algorithm? Can we still have an expected polynomial running time randomized algorithm for generating such a sample?
- Expected number of samples: $1/p$.
- For a fixed (constant) $k$, check all $\binom{n}{k}$ cliques and make sure they are not monochromatic.

# Question

- What if we want a Las Vegas algorithm? Can we still have an expected polynomial running time randomized algorithm for generating such a sample?
- Expected number of samples: $1/p$.
- For a fixed (constant) $k$, check all $\binom{n}{k}$ cliques and make sure they are not monochromatic.
  - Not polynomial time when $k$ grows with $n$!

# The average argument

- **Intuitive idea:** Say a discrete random variable $X$ has $\mathbb{E}[X] = \mu$. Then there must be some value $a \leq \mu$ and $b \geq \mu$ such that $\Pr[X = a] > 0$ and $\Pr[X = b] > 0$.

## The average argument

- **Intuitive idea:** Say a discrete random variable $X$ has $\mathbb{E}[X] = \mu$. Then there must be some value $a \le \mu$ and $b \ge \mu$ such that $\Pr[X = a] > 0$ and $\Pr[X = b] > 0$.

**Lemma 1**

Suppose we have a probability space $S$ such and a random variable $X$ defined on $S$ such that $\mathbb{E}[X] = \mu$. Then

$$\Pr[X \ge \mu] > 0 \quad \text{and} \quad \Pr[X \le \mu] > 0.$$

**Proof:** Let $\mathcal{X} = \{x \in \mathbb{R} : \Pr[X = x] > 0\}$ be the support of $X$. Suppose that

$$\mu = \mathbb{E}[X] = \sum_{x \in \mathcal{X}} \Pr[X = x].$$

**Proof:** Let $\mathcal{X} = \{x \in \mathbb{R} : \Pr[X = x] > 0\}$ be the support of $X$. Suppose that

$$\mu = \mathbb{E}[X] = \sum_{x \in \mathcal{X}} \Pr[X = x].$$

If $\Pr[X \geq \mu] = 0$, then

$$\mu = \sum_{x \in \mathcal{X}} x \Pr[X = x] = \sum_{x < \mu} x \Pr[X = x] < \sum_{x < \mu} \mu \Pr[X = x] = \mu.$$

$(\Rightarrow\!\Leftarrow)$

**Proof:** Let $\mathcal{X} = \{x \in \mathbb{R} : \Pr[X = x] > 0\}$ be the support of $X$. Suppose that

$$\mu = \mathbb{E}[X] = \sum_{x \in \mathcal{X}} \Pr[X = x].$$

If $\Pr[X \geq \mu] = 0$, then

$$\mu = \sum_{x \in \mathcal{X}} x \Pr[X = x] = \sum_{x < \mu} x \Pr[X = x] < \sum_{x < \mu} \mu \Pr[X = x] = \mu.$$

$(\Rightarrow\!\Leftarrow)$

Similarly, if $\Pr[X \leq \mu] = 0$ then

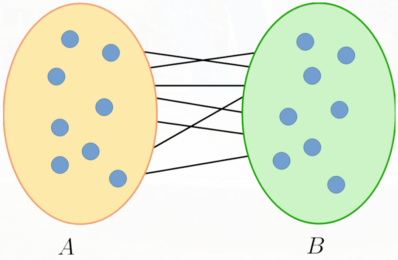
$$\mu = \sum_{x \in \mathcal{X}} x \Pr[X = x] = \sum_{x > \mu} x \Pr[X = x] > \sum_{x > \mu} \mu \Pr[X = x] = \mu.$$

$(\Rightarrow\!\Leftarrow)$

# Application: Finding a Large Cut

**Theorem 2**

Given an undirected graph $G = (V, E)$ with $|V| = n$ and $|E| = m$. There is a partition of $V$ into disjoint sets $A$ and $B$ such that at least $m/2$ edges connect a vertex in $A$ to a vertex in $B$. That is, there is a cut with value at least $m/2$.



# Proof

- Construct sets $A$ and $B$ by randomly and independently assigning each vertex to one of the two sets.
- Let $e_1, \ldots, e_m$ be an arbitrary enumeration of the edges of $G$.
- For $i = 1, \ldots, m$, define $X_i$ such that

$$X_i = \begin{cases} 1 & \text{if edge } i \text{ connects } A \text{ to } B \\ 0 & \text{otherwise} \end{cases}$$

# Proof

- Construct sets $A$ and $B$ by randomly and independently assigning each vertex to one of the two sets.
- Let $e_1, \ldots, e_m$ be an arbitrary enumeration of the edges of $G$.
- For $i = 1, \ldots, m$, define $X_i$ such that

$$X_i = \begin{cases} 1 & \text{if edge } i \text{ connects } A \text{ to } B \\ 0 & \text{otherwise} \end{cases}$$

- The probability that edge $e_i$ connects a vertex in $A$ to a vertex in $B$ is $\frac{1}{2}$, and thus

$$\mathbb{E}[X_i] = \frac{1}{2}$$

# Proof (contd.)

- Let $C = (A, B)$ be a random variable denoting the value of the cut corresponding to the sets $A$ and $B$. Then

$$\mathbb{E}[C(A,B)] = \mathbb{E}\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} \mathbb{E}[X_i] = m \cdot \frac{1}{2} = \frac{m}{2}.$$

# Proof (contd.)

- We need "a lower bound" on the probability that a random partition has a cut of value at least $m/2$.
- To derive such a bound, let $p = \Pr\left[C(A,B) \geq \frac{m}{2}\right]$, and observe that $C(A,B) \leq m$, Then

$$
\begin{aligned}
\frac{m}{2} &= \mathbb{E}[C(A,B)] \\
&= \sum_{i<m/2} i \Pr[C(A,B)=i] + \sum_{i\geq m/2} i \Pr[C(A,B)=i] \\
&\leq (1-p)\left(\frac{m}{2}-1\right) + pm,
\end{aligned}
$$

## Proof (contd.)

- We need "a lower bound" on the probability that a random partition has a cut of value at least $m/2$.
- To derive such a bound, let $p = \Pr\left[C(A,B) \geq \frac{m}{2}\right]$, and observe that $C(A,B) \leq m$, Then

$$
\begin{aligned}
\frac{m}{2} &= \mathbb{E}[C(A,B)] \\
&= \sum_{i<m/2} i \Pr[C(A,B)=i] + \sum_{i \geq m/2} i \Pr[C(A,B)=i] \\
&\leq (1-p)\left(\frac{m}{2}-1\right) + pm,
\end{aligned}
$$

which implies that $p \geq \frac{1}{m/2+1}$

## Proof (contd.)

- We need "a lower bound" on the probability that a random partition has a cut of value at least $m/2$.
- To derive such a bound, let $p = \Pr\left[C(A,B) \geq \frac{m}{2}\right]$, and observe that $C(A,B) \leq m$, Then

$$\begin{aligned} \frac{m}{2} &= \mathbb{E}[C(A,B)] \\ &= \sum_{i<m/2} i \Pr[C(A,B)=i] + \sum_{i \geq m/2} i \Pr[C(A,B)=i] \\ &\leq (1-p)\left(\frac{m}{2}-1\right) + pm, \end{aligned}$$

which implies that $p \geq \frac{1}{m/2+1}$ $\Rightarrow$ expected number of samples before finding a cut: $\leq m/2+1$.

# Application: Maximum Satisfiability

The following expression is an instance of SAT:

$$(x_1 \vee \overline{x_2} \vee \overline{x_3}) \wedge (\overline{x_1} \vee \overline{x_3}) \wedge (x_1 \vee x_2 \vee x_4) \wedge (x_4 \vee \overline{x_3}) \wedge (x_4 \vee \overline{x_1}).$$

# Application: Maximum Satisfiability

The following expression is an instance of SAT:

$$(x_1 \vee \overline{x_2} \vee \overline{x_3}) \wedge (\overline{x_1} \vee \overline{x_3}) \wedge (x_1 \vee x_2 \vee x_4) \wedge (x_4 \vee \overline{x_3}) \wedge (x_4 \vee \overline{x_1}).$$

**Theorem 3**

Given a set of $m$ clauses, let $k_i$ be the number of literals in the $i$th clause for $i = 1, \ldots, m$. Let $k = \min_{i \in \{1,\ldots,m\}} k_i$. Then there is a truth assignment that satisfies at least

$$\sum_{i=1}^{m} (1 - 2^{-k_i}) \geq m(1 - 2^{-k}) \quad \text{clauses.}$$

# Proof

- Assign values independently and uniformly at random to the variables.
- The probability that the $i$th clause with $k_i$ literals is satisfied is at least $(1 - 2^{-k})$.
- The expected number of satisfied clauses is therefore at least

$$\sum_{i=1}^{m}(1 - 2^{-k_i}) \geq m(1 - 2^{-k}),$$

and there must be an assignment that satisfies at least that many clauses.

# Outline

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.
- Again, let us consider the problem of finding a large cut.

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.
- Again, let us consider the problem of finding a large cut.
  - Imagine that we **place the vertices deterministically**, one at a time, in an arbitrary order $v_1, v_2, \ldots, v_n$.

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.
- Again, let us consider the problem of finding a large cut.
  - Imagine that we **place the vertices deterministically**, one at a time, in an arbitrary order $v_1, v_2, \ldots, v_n$.
  - $x_i$: the set where $v_i$ is placed (so $x_i$ is either $A$ or $B$).

# Derandomization

- The probabilistic method can yield insight into how to construct **deterministic** algorithms.
- Derandomization using method of conditional expectations.
- Again, let us consider the problem of finding a large cut.
  - Imagine that we **place the vertices deterministically**, one at a time, in an arbitrary order $v_1, v_2, \ldots, v_n$.
  - $x_i$: the set where $v_i$ is placed (so $x_i$ is either $A$ or $B$).
  - **Assumption:** The first $k$ vertices have been placed. Consider the expected value of the cut if the remaining vertices are then placed independently and uniformly into one of the two sets.

# Derandomization by Conditional Expectation (1/4)

- We show inductively how to place the next vertex so that

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \leq \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}]$$

# Derandomization by Conditional Expectation (1/4)

- We show inductively how to place the next vertex so that

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \le \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}]$$

It follows that

$$\mathbb{E}[C(A,B)] \le \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_n].$$

# Derandomization by Conditional Expectation (1/4)

- We show inductively how to place the next vertex so that

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \le \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}]$$

It follows that

$$\mathbb{E}[C(A,B)] \le \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_n].$$

- **The right-hand side:** the value of the cut determined by our placement algorithm.
- Hence our algorithm returns a cut whose value $\ge \mathbb{E}[C(A,B]) \ge m/2$.

# Derandomization by Conditional Expectation (2/4)

- The induction base: $\mathbb{E}[C(A,B) \mid x_1] = \mathbb{E}[C(A,B)]$

# Derandomization by Conditional Expectation (2/4)

- The induction base: $\mathbb{E}[C(A, B) \mid x_1] = \mathbb{E}[C(A, B)]$
  - Placing the first vertex doesn't matter (symmetric).

# Derandomization by Conditional Expectation (2/4)

- The induction base: $\mathbb{E}[C(A,B) \mid x_1] = \mathbb{E}[C(A,B)]$
  - Placing the first vertex doesn't matter (symmetric).
- We now prove the inductive step:

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \leq \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}]$$

# Derandomization by Conditional Expectation (2/4)

- The induction base: $\mathbb{E}[C(A,B) \mid x_1] = \mathbb{E}[C(A,B)]$
  - Placing the first vertex doesn't matter (symmetric).
- We now prove the inductive step:

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \leq \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}]$$

- Consider placing $v_{k+1}$ randomly, so that it is placed in $A$ or $B$ with probability $1/2$ each.
- Let $Y_{k+1}$ be a random variable representing the set where it is placed.

# Derandomization by Conditional Expectation (3/4)

- Then

$$\begin{aligned}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] &= \frac{1}{2}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k, Y_{k+1} = A] \\ &+ \frac{1}{2}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k, Y_{k+1} = B].\end{aligned}$$

# Derandomization by Conditional Expectation (3/4)

- Then

$$\begin{aligned}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] &= \frac{1}{2}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k, Y_{k+1} = A] \\ &+ \frac{1}{2}\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k, Y_{k+1} = B].\end{aligned}$$

- It follows that

$$\max\Big(\mathbb{E}[C(A,B) \mid x_1, \ldots, x_k, Y_{k+1} = A], \ \mathbb{E}[C(A,B) \mid x_1, \ldots, x_k, Y_{k+1} = B]\Big)$$

$$\geq \ \mathbb{E}[C(A,B) \mid x_1, \ldots, x_k].$$

# Derandomization by Conditional Expectation (4/4)

- Compute the two quantities
  - $\mathbb{E}[C(A,B) \mid x_1,\ldots,x_k, Y_{k+1} = A]$
  - $\mathbb{E}[C(A,B) \mid x_1,\ldots,x_k, Y_{k+1} = B]$

  and then place $v_{k+1}$ in the set that yields the larger expectation.

- Once we do this, we will have a placement satisfying

$$\mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_k] \leq \mathbb{E}[C(A,B) \mid x_1, x_2, \ldots, x_{k+1}].$$

**Computation of $\mathbb{E}[C(A,B) \mid x_1,\ldots,x_k, Y_{k+1} = A \text{ (or } B)]$**

- For the first $k+1$ vertices, compute the number of edges that contribute to the cut.
- For all the other edges, each one contributes to the cut with prob. $1/2$.

# Derandomization by Conditional Expectation (4/4)

- Compute the two quantities
  - $\mathbb{E}[C(A, B) \mid x_1, \ldots, x_k, Y_{k+1} = A]$
  - $\mathbb{E}[C(A, B) \mid x_1, \ldots, x_k, Y_{k+1} = B]$

  and then place $v_{k+1}$ in the set that yields the larger expectation.

- Once we do this, we will have a placement satisfying

$$\mathbb{E}[C(A, B) \mid x_1, x_2, \ldots, x_k] \leq \mathbb{E}[C(A, B) \mid x_1, x_2, \ldots, x_{k+1}].$$

### Computation of $\mathbb{E}[C(A, B) \mid x_1, \ldots, x_k, Y_{k+1} = A \text{ (or } B)]$

- For the first $k + 1$ vertices, compute the number of edges that contribute to the cut.
- For all the other edges, each one contributes to the cut with prob. $1/2$.
  - Depends on whether $v_{k+1}$ has more neighbors in $A$ or in $B$.

# Outline

# Sample and Modify

## Sample and Modify (Two Stages)

- 1st Stage: Construct a random structure that does not have the required properties.
- 2nd Stage: Modify the structure so that the required properties are satisfied.

# Sample and Modify

## Sample and Modify (Two Stages)

- 1st Stage: Construct a random structure that does not have the required properties.
- 2nd Stage: Modify the structure so that the required properties are satisfied.

- In some cases, it is easy to work using this indirect approach.

# Application: Independent Sets

**Theorem 4**

Let $G = (V, E)$ be a graph on $n$ vertices with $m$ edges. Then $G$ has an independent set with at least $n^2/4m$ vertices.

**Proof:** Let $d = 2m/n$ be the average degree of the vertices in $G$. Consider the following randomized algorithm.

# Application: Independent Sets

**Theorem 4**

Let $G = (V, E)$ be a graph on $n$ vertices with $m$ edges. Then $G$ has an independent set with at least $n^2/4m$ vertices.

**Proof:** Let $d = 2m/n$ be the average degree of the vertices in $G$. Consider the following randomized algorithm.

1. Delete each vertex of $G$ (together with its incident edges) independently with probability $1 - 1/d$.
2. For each remaining edge, remove it and one of its adjacent vertices.

# Proof (2/3)

- Let $X$ be the number of vertices that survive the first step of the algorithm.
- Since the graph has $n$ vertices and each vertex survives with probability $1/d$, it follows that

$$\mathbb{E}[X] = \frac{n}{d}.$$

# Proof (2/3)

- Let $X$ be the number of vertices that survive the first step of the algorithm.
- Since the graph has $n$ vertices and each vertex survives with probability $1/d$, it follows that

$$\mathbb{E}[X] = \frac{n}{d}.$$

- Let $Y$ be the number of edges that survive the first step. There are $nd/2$ edges in the graph, and an edge survives iff its two adjacent vertices survive.

# Proof (2/3)

- Let $X$ be the number of vertices that survive the first step of the algorithm.
- Since the graph has $n$ vertices and each vertex survives with probability $1/d$, it follows that

$$\mathbb{E}[X] = \frac{n}{d}.$$

- Let $Y$ be the number of edges that survive the first step. There are $nd/2$ edges in the graph, and an edge survives iff its two adjacent vertices survive.
- Thus,

$$\mathbb{E}[Y] = \frac{nd}{2}\left(\frac{1}{d}\right)^2 = \frac{n}{2d}.$$

# Proof (3/3)

- The second step of the algorithm removes all the remaining edges and at most $Y$ vertices.
- When the algorithm terminates, it outputs an independent set of size at least $X - Y$, and

$$\mathbb{E}[X - Y] = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d}.$$

- The expected size of the independent set generated by the algorithm is $n/2d$, so the graph has an independent set with at least $n/2d = n^2/4m$ vertices.

# Application: Independent Sets

**Theorem 5**

For any integer $k \geq 3$, there is a graph with $n$ nodes, at least $\frac{1}{4}n^{1+1/k}$ edges, and girth at least $k$.

- *girth*: the length of a shortest cycle contained in the graph.

# Application: Independent Sets

**Theorem 5**

For any integer $k \geq 3$, there is a graph with $n$ nodes, at least $\frac{1}{4}n^{1+1/k}$ edges, and girth at least $k$.

- *girth*: the length of a shortest cycle contained in the graph.

**Proof:** We first sample a random graph $G \in G_{n,p}$ with $p = n^{1/k-1}$. Let $X$ be the number of edges in the graph. Then

$$\mathbb{E}[X] = p\binom{n}{2} = \frac{1}{2}\left(1 - \frac{1}{n}\right) n^{1/k+1}.$$

- $G_{n,p}$: a set of graphs of $n$ vertices where every $(u, v)$ exists with prob. $p$.

# Proof (2/3)

- Let $Y$ be the number of cycles in the graph of length at most $k-1$.
- Any specific possible cycle of length $i$, where $3 \le i \le k-1$, occurs with probability $p^i$.

# Proof (2/3)

- Let $Y$ be the number of cycles in the graph of length at most $k-1$.
- Any specific possible cycle of length $i$, where $3 \leq i \leq k-1$, occurs with probability $p^i$.
- Also, there are $\binom{n}{i}\frac{(i-1)!}{2}$ possible cycles of length $i$.
    - First, consider choosing the $i$ vertices, then consider the possible orders, and finally keep in mind that reversing the order yields the same cycle.

# Proof (2/3)

- Let $Y$ be the number of cycles in the graph of length at most $k-1$.
- Any specific possible cycle of length $i$, where $3 \le i \le k-1$, occurs with probability $p^i$.
- Also, there are $\binom{n}{i}\frac{(i-1)!}{2}$ possible cycles of length $i$.
  - First, consider choosing the $i$ vertices, then consider the possible orders, and finally keep in mind that reversing the order yields the same cycle.
- Hence,

$$\begin{aligned}\mathbb{E}[Y] = \sum_{i=3}^{k-1}\binom{n}{i}\frac{(i-1)!}{2}p^i &\le \sum_{i=3}^{k-1} n^i p^i \\ &= \sum_{i=3}^{k-1} n^{i/k} < kn^{(k-1)/k}.\end{aligned}$$

# Proof (3/3)

- We modify the original randomly chosen graph $G$ by eliminating one edge from each cycle of length up to $k-1$.
- The modified graph has girth at least $k$ ($\because$ shorter cycles are destroyed).
- When $n$ is sufficiently large, the expected number of edges in the resulting graph is

$$\mathbb{E}[X-Y] \geq \frac{1}{2}\left(1-\frac{1}{n}\right)n^{1/k+1} - kn^{(k-1)/k} \geq \frac{1}{4}n^{1/k+1}.$$

- Hence there exists a graph with at least $\frac{1}{4}n^{1+1/k}$ edges and girth at least $k$.

# Outline

# The Second Moment Method

- In the $G_{n,p}$ model, it is often the case that there is a threshold function $f$ such that
  - when $p = O(f(n))$ or $p = o(f(n))$, **almost no** graph has the desired property;
  - when $p = \Omega(f(n))$ or $p = \omega(f(n))$, **almost every** graph has the desired property.

# The Second Moment Method

- The following theorem from Chebyshev's inequality is often used.

**Theorem 6 [Derived from Chebyshev's Inequality]**

$X$ is a nonnegative integer-valued random variable, then

$$\Pr[X = 0] \le \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2}.$$

# The Second Moment Method

- The following theorem from Chebyshev's inequality is often used.

**Theorem 6 [Derived from Chebyshev's Inequality]**

$X$ is a nonnegative integer-valued random variable, then

$$\Pr[X = 0] \leq \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2}.$$

**Proof:**

$$\Pr[X = 0] \leq \Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]] \leq \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2}.$$

# Application: Threshold Behavior in Random Graphs

**Theorem 7**

Consider $G_{n,p}$, suppose that $p = f(n)$.

Let $A$ be the event that a random graph chosen from $G_{n,p}$ has a clique of four or more vertices. Then, for any $\varepsilon > 0$ and sufficiently large $n$,

$$\Pr[A] < \varepsilon \quad \text{if } f(n) = o(n^{-2/3}).$$

Similarly, if $f(n) = \omega(n^{-2/3})$ then, for sufficiently large $n$,

$$\Pr[\bar{A}] < \varepsilon.$$

# Proof (1/4)

- We first consider the case in which $p = f(n)$ and $f(n) = o(n^{-2/3})$. Let $C_1, \ldots, C_{\binom{n}{4}}$ be an enumeration of all the subsets of four vertices in $G$.
- Let

$$X_i = \begin{cases} 1 & \text{if } C_i \text{ is a 4-clique} \\ 0 & \text{otherwise.} \end{cases}$$

- Let

$$X = \sum_{i=1}^{\binom{n}{4}} X_i,$$

so that (A 4-clique has 6 edges)

$$\mathbb{E}[X] = \binom{n}{4} p^6.$$

# Proof (2/4)

- Consider the case that $\mathbb{E}[X] = o(1)$

# Proof (2/4)

- Consider the case that $\mathbb{E}[X] = o(1) \Rightarrow \mathbb{E}[X] < \varepsilon$ for sufficiently large $n$.

## Proof (2/4)

- Consider the case that $\mathbb{E}[X] = o(1) \Rightarrow \mathbb{E}[X] < \varepsilon$ for sufficiently large $n$.
- Since $X$ is a nonnegative integer-valued random variable, it follows that $\Pr[X \geq 1] \leq \mathbb{E}[X] < \varepsilon$.

# Proof (2/4)

- Consider the case that $\mathbb{E}[X] = o(1) \Rightarrow \mathbb{E}[X] < \varepsilon$ for sufficiently large $n$.
- Since $X$ is a nonnegative integer-valued random variable, it follows that $\Pr[X \geq 1] \leq \mathbb{E}[X] < \varepsilon$.
- Hence, the probability that a random graph chosen from $G_{n,p}$ has a clique of four or more vertices is less than $\varepsilon$.

# Proof (3/4)

- Next, consider the case when $p = f(n) = \omega(n^{-2/3})$.
- $\mathbb{E}[X] \to \infty$ as $n$ grows large.
  - Not sufficient to conclude that, with high probability, a graph chosen random from $G_{n,p}$ has a clique of at least four vertices.

# Proof (3/4)

- Next, consider the case when $p = f(n) = \omega(n^{-2/3})$.
- $\mathbb{E}[X] \to \infty$ as $n$ grows large.
  - Not sufficient to conclude that, with high probability, a graph chosen random from $G_{n,p}$ has a clique of at least four vertices.
- We can, however, use the second moment method to prove that $\Pr[X = 0] = o(1)$ in this case.

# Proof (3/4)

- Next, consider the case when $p = f(n) = \omega(n^{-2/3})$.
- $\mathbb{E}[X] \to \infty$ as $n$ grows large.
  - Not sufficient to conclude that, with high probability, a graph chosen random from $G_{n,p}$ has a clique of at least four vertices.
- We can, however, use the second moment method to prove that $\Pr[X = 0] = o(1)$ in this case.
- Then, we must show that $\mathrm{Var}[X] = o((\mathbb{E}[X])^2)$.

# Proof (3/4)

- Next, consider the case when $p = f(n) = \omega(n^{-2/3})$.
- $\mathbb{E}[X] \to \infty$ as $n$ grows large.
  - Not sufficient to conclude that, with high probability, a graph chosen random from $G_{n,p}$ has a clique of at least four vertices.
- We can, however, use the second moment method to prove that $\Pr[X = 0] = o(1)$ in this case.
- Then, we must show that $\mathrm{Var}[X] = o((\mathbb{E}[X])^2)$.
- Consider the following lemma:

# Proof (4/4)

## Lemma 2

Let $Y_i, i = 1, \ldots, m$, be 0-1 random variables, and let $Y = \sum_{i=1}^{m} Y_i$. Then

$$\mathrm{Var}[Y] \leq \mathbb{E}[Y] + \sum_{1 \geq i,j \geq m;\ i \neq j} \mathrm{Cov}(Y_i, Y_j).$$

# Proof of Lemma 2

- For any sequence of random variables $Y_1, \ldots, Y_m$,

$$\mathrm{Var}\left[\sum_{i=1}^{m} Y_i\right] = \sum_{i=1}^{m} \mathrm{Var}[Y_i] + \sum_{1 \le i,j \le m; i \ne j} \mathrm{Cov}(Y_i, Y_j).$$

- This is the generalization from two variables to $m$ variables.
- When $Y_i$ is a 0-1 random variable, $\mathbb{E}[Y_i^2] = \mathbb{E}[Y_i]$ and so

$$\mathrm{Var}[Y_i] = \mathbb{E}[Y_i^2] - (\mathbb{E}[Y_i])^2 \le \mathbb{E}[Y_i],$$

giving the lemma.

# Finishing the proof of Theorem 7

- We wish to compute Variance of the number of 4-cliques

$$\mathrm{Var}[X] = \mathrm{Var}\left(\sum_{i=1}^{\binom{n}{4}} X_i\right).$$

## Using covariance

- Since $X_i$ is an indicator,

$$\mathrm{Var}[X_i] \le \mathbb{E}[X_i] = p^6, \quad \Rightarrow \quad \sum_i \mathrm{Var}(X_i) \le \binom{n}{4} p^6.$$

- The remaining term is the total covariance. It depends on how much the two 4-cliques $C_i$ and $C_j$ overlap.
- Write $C_i \cap C_j$ for their intersection. We consider the cases $|C_i \cap C_j| = 0, 1, 2, 3$.

# Pairs of cliques: overlap cases (1/2)

- **Case** $|C_i \cap C_j| = 0$ **or** $1$**.**
  - The sets of edges involved in the two cliques are disjoint.
  - Thus $X_i$ and $X_j$ are independent, so $\mathbb{E}[X_iX_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] = 0$.
- **Case** $|C_i \cap C_j| = 2$**.**
  - The cliques share one edge; altogether 11 distinct edges must be present.
  - Hence $\mathbb{E}[X_iX_j] \le p^{11} \quad \Rightarrow \quad \mathbb{E}[X_iX_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] \le p^{11}$.
  - The number of ordered pairs $(C_i, C_j)$ with $|C_i \cap C_j| = 2$ is

$$\binom{n}{6}\binom{6}{2,2,2}.$$

    Choose 6 vertices and then split them into $C_i \cap C_j$ (2 vertices), 2 vertices for $C_i \setminus C_j$, and 2 for $C_j \setminus C_i$.

# Pairs of cliques: overlap cases (2/2)

- **Case $|C_i \cap C_j| = 3$.**
  - The cliques share three vertices (a triangle), hence 9 distinct edges must appear.
  - Thus
    $$\mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] \le \mathbb{E}[X_i X_j] \le p^9.$$
  - There are
    $$\binom{n}{5}\binom{5}{3,1,1}$$
    ordered pairs $(C_i, C_j)$ with $|C_i \cap C_j| = 3$.

# Completing the bound on Var($X$)

- Collecting all contributions,

$$\mathrm{Var}(X) \le \binom{n}{4} p^6 + \binom{n}{6}\binom{6}{2,2,2} p^{11} + \binom{n}{5}\binom{5}{3,1,1} p^9.$$

- Using $p = f(n) = \omega(n^{-2/3})$ and $\mathbb{E}[X] = \binom{n}{4} p^6$, one checks that

$$\mathrm{Var}[X] = o\big((\mathbb{E}[X])^2\big).$$

  By the second moment method, this implies

$$\Pr[X = 0] = o(1),$$

# Completing the bound on Var($X$)

- Collecting all contributions,

$$\mathrm{Var}(X) \le \binom{n}{4} p^6 + \binom{n}{6}\binom{6}{2,2,2} p^{11} + \binom{n}{5}\binom{5}{3,1,1} p^9 .$$

- Using $p = f(n) = \omega(n^{-2/3})$ and $\mathbb{E}[X] = \binom{n}{4} p^6$, one checks that

$$\mathrm{Var}[X] = o\big((\mathbb{E}[X])^2\big).$$

By the second moment method, this implies

$$\Pr[X = 0] = o(1),$$

$\Rightarrow$ with high probability $G$ contains a copy of $K_4$.

# Outline

- For a sum of Bernoulli random variables, there is an alternative to the second moment method.

**Theorem 8**

Let $X = \sum_{i=1}^{n} X_i$, where each $X_i$ is a 0-1 random variable. Then

$$\Pr[X > 0] \geq \sum_{i=1}^{n} \frac{\Pr[X_i = 1]}{\mathbb{E}[X \mid X_i = 1]}.$$

# Proof

- Let $Y = 1/X$ if $X > 0$, with $Y = 0$ otherwise. Then

$$\Pr[X > 0] = \mathbb{E}[XY].$$

# Proof

- Let $Y = 1/X$ if $X > 0$, with $Y = 0$ otherwise. Then
$$\Pr[X > 0] = \mathbb{E}[XY].$$
- However,

$$\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}\left[\sum_{i=1}^{n} X_i Y\right] = \sum_{i=1}^{n} \mathbb{E}[X_i Y] \\
&= \sum_{i=1}^{n} \left(\mathbb{E}[X_i Y \mid X_i = 1] \Pr[X_i = 1] + \mathbb{E}[X_i Y \mid X_i = 0] \Pr[X_i = 0]\right) \\
&= \sum_{i=1}^{n} \mathbb{E}[Y \mid X_i = 1] \Pr[X_i = 1] = \sum_{i=1}^{n} \mathbb{E}[1/X \mid X_i = 1] \Pr[X_i = 1] \\
&\geq \sum_{i=1}^{n} \frac{\Pr[X_i = 1]}{\mathbb{E}[X \mid X_i = 1]}. \quad \text{(by Jensen's inequality; } \mathbb{E}[f(Z)] \geq f(\mathbb{E}[Z])\text{)}
\end{aligned}$$

# An alternative proof of the 4-clique existence (1/3)

- Let $X = \sum_{i=1}^{\binom{n}{4}} X_i$, where $X_i$ is 1 if the subset of four vertices $C_i$ is a 4-clique and 0 otherwise.
- For a specific $X_j$, we have $\Pr[X_j = 1] = p^6$. Using the linearity of expectations, we compute

$$\mathbb{E}[X \mid X_j = 1] = \mathbb{E}\left[\sum_{i=1}^{\binom{n}{4}} X_i \,\middle|\, X_j = 1\right] = \sum_{i=1}^{\binom{n}{4}} \mathbb{E}[X_i \mid X_j = 1].$$

- Conditioning on $X_j = 1$, we now compute $\mathbb{E}[X_i \mid X_j = 1]$ by using that, for a 0-1 random variable,

$$\mathbb{E}[X_i \mid X_j = 1] = \Pr[X_i = 1 \mid X_j = 1].$$

# An alternative proof of the 4-clique existence (2/3)

- There are $\binom{n-4}{4}$ sets of vertices $C_i$ that do not interest $C_j$. Each corresponding $X_i$ is 1 with probability $p^6$.

# An alternative proof of the 4-clique existence (2/3)

- There are $\binom{n-4}{4}$ sets of vertices $C_i$ that do not interest $C_j$. Each corresponding $X_i$ is 1 with probability $p^6$.
- Similarly, $X_i = 1$ with probability $p^6$ for the $4\binom{n-4}{3}$ sets $C_i$ that have one vertex in common with $C_j$.

# An alternative proof of the 4-clique existence (2/3)

- There are $\binom{n-4}{4}$ sets of vertices $C_i$ that do not interest $C_j$. Each corresponding $X_i$ is 1 with probability $p^6$.
- Similarly, $X_i = 1$ with probability $p^6$ for the $4\binom{n-4}{3}$ sets $C_i$ that have one vertex in common with $C_j$.
- For the remaining cases, we have $\Pr[X_i = 1 \mid X_j = 1] = p^5$ for the $6\binom{n-4}{2}$sets $C_i$ that have two vertices in common with $C_j$ and $\Pr[X_i = 1 \mid X_j = 1] = p^3$ for the $4\binom{n-4}{1}$ sets $C_i$ that have three vertices in common with $C_j$.

# An alternative proof of the 4-clique existence (3/3)

- Overall, we have

$$
\begin{aligned}
\mathbb{E}[X \mid X_j = 1] &= \sum_{i=1}^{\binom{n}{4}} \mathbb{E}[X_i \mid X_j = 1] \\
&= 1 + \binom{n-4}{4} p^6 + 4\binom{n-4}{3} p^6 + 6\binom{n-4}{2} p^5 \\
&+ 4\binom{n-4}{1} p^3
\end{aligned}
$$

- Applying Theorem 8 yields

$$
\Pr[X > 0] \geq \frac{\binom{n}{4} p^6}{1 + \binom{n-4}{4} p^6 + 4\binom{n-4}{3} p^6 + 6\binom{n-4}{2} p^5 + 4\binom{n-4}{1} p^3},
$$

which approaches 1 as $n \uparrow$ when $p = f(n) = \omega(n^{-2/3})$.

# Discussions