CS265 Final Cheatsheet

1 L1–8: Foundations (Brief)

1.1 Schwartz-Zippel (L1)

$P(x_1,\dots,x_n) \neq 0$, degree d, pick $r_i \in S$ u.a.r.: $\Pr[P(r_1,\dots,r_n)=0] \le d/|S|$.

Las Vegas: always correct, random runtime ($\mathbb{E} < \infty$). **Monte Carlo:** bounded runtime, may err. 1-sided (RP) or 2-sided (BPP).

Repeat 2-sided MC t times + majority $\Rightarrow$ error $\le e^{-2\epsilon^2 t}$.

1.2 Linearity of Expectation, Karger (L2)

$\mathbb{E}[\sum X_i] = \sum \mathbb{E}[X_i]$ (no independence needed).

Karger's Min-Cut: Pick random edge, contract, until 2 vertices. Prob of finding specific min-cut $\ge \binom{n}{2}^{-1}$. Repeat $O(n^2 \log n)$ times $\Rightarrow$ high prob.

Quicksort: Random pivot $\Rightarrow \mathbb{E}[\text{comparisons}] = 2n\ln n + O(n)$.

1.3 Primality Testing (L3)

Fermat: if p prime, $a^{p-1} \equiv 1 \pmod p$ for all a. Miller-Rabin: test witnesses $\Rightarrow$ poly-time randomized primality test.

1.4 Markov, Chebyshev (L4)

Markov: $X \ge 0$: $\Pr[X \ge t] \le \mathbb{E}[X]/t$.

Chebyshev: $\Pr[|X - \mathbb{E}X| \ge t] \le \mathrm{Var}(X)/t^2$.

Sampling-based median: Sample $O(\sqrt{n}\log n)$ elements, sort sample, use as approx median.

1.5 Chernoff Bounds (L5)

$X = \sum X_i$, independent $X_i \in \{0,1\}$:

- $\Pr[X \ge (1+\delta)\mu] \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$; for $\delta \in (0,1]$: $\le e^{-\delta^2\mu/3}$
- $\Pr[X \le (1-\delta)\mu] \le e^{-\delta^2\mu/2}$ for $\delta \in (0,1]$
- $c \ge 6$: $\Pr[X \ge c\mu] \le 2^{-c\mu}$

Hoeffding: $X_i \in [a_i,b_i]$ indep, $S = \sum X_i$: $\Pr[|S - \mathbb{E}S| \ge t] \le 2\exp\left(\frac{-2t^2}{\sum(b_i-a_i)^2}\right)$.

Poisson tail: $X \sim \mathrm{Poi}(\lambda)$: $\Pr[|X-\lambda| \ge c] \le 2\exp\left(\frac{-c^2}{2(c+\lambda)}\right)$.

MGF technique: $\Pr[X \ge t] = \Pr[e^{sX} \ge e^{st}] \le \mathbb{E}[e^{sX}]/e^{st}$. Factor MGF by independence, optimize s.

Hypercube routing: n-dim hypercube, fix-from-left routing. Phase 1: random intermediate, Phase 2: route to dest. $O(n)$ steps w.h.p. via Chernoff + union bound.

1.6 Balls-in-Bins, Hashing (L6)

n balls, n bins: max load $\Theta(\log n/\log\log n)$ w.h.p.

Coupon Collector: $\mathbb{E}[\text{collect all } n] = nH_n \approx n\ln n$. Each new coupon: geometric wait, $\mathbb{E} = n/(n-i)$ for i-th coupon.

Birthday Paradox: $\Theta(\sqrt{n})$ samples for collision.

Power of 2 Choices: max load $O(\log\log n)$ w.h.p.

Bloom Filters: k hash functions, m bits. False positive $\approx (1-e^{-kn/m})^k$. Optimal $k = m\ln 2/n$.

1.7 Metric Embeddings (L7)

Embed metric (X,d) into (Y,d') with distortion D: $d(x,y) \le d'(f(x),f(y)) \le D\cdot d(x,y)$.

Bourgain's theorem: any n-point metric embeds into ℓ_1 with $O(\log n)$ distortion.

1.8 JL Lemma, LSH (L8)

Johnson-Lindenstrauss: n points in $\mathbb{R}^d$ can be projected to $\mathbb{R}^k$, $k = O(\log n/\epsilon^2)$, preserving all pairwise distances within $(1\pm\epsilon)$. Random projection works.

LSH: Hash family where $\Pr[h(x)=h(y)]$ depends on $d(x,y)$. Amplify via AND/OR composition.

2 L9: Probabilistic Method

Core: Show object C exists by: (1) define r.v. X, show $\Pr[X=C] > 0$; or (2) show $\mathbb{E}[f(X)] \ge \alpha \Rightarrow \exists$ object with $f \ge \alpha$.

2.1 Ramsey Numbers

R_k: smallest n s.t. any 2-coloring of K_n has monochromatic k-clique.

Thm: $R_k \in (2^{k/2}, 2^{2k})$.

Lower bound ($R_k > 2^{k/2}$): $n = 2^{k/2}$, color edges randomly. For k verts: $\Pr[\text{mono clique}] = 2^{-\binom{k}{2}+1}$. Union bound: $\binom{n}{k}\cdot 2^{-\binom{k}{2}+1} \le \frac{n^k}{k!}\cdot 2^{-k^2/2+k/2+1} = \frac{2^{k^2/2}}{k!}\cdot 2^{-k^2/2+k/2+1} = \frac{2^{k/2+1}}{k!} < 1$.

Upper bound ($R_k < 2^{2k}$): $R_{a,b} \le 1 + R_{a-1,b} + R_{a,b-1}$. Fix vertex v, partition neighbors into red (S_r) and blue (S_b) sets. $|S_r| \ge R_{a-1,b}$ or $|S_b| \ge R_{a,b-1}$. Induction: $R_{a,b} < 2^{a+b}$.

2.2 Independent Sets via Prob. Method

Thm: Graph with n vertices, $m \ge n/2$ edges $\Rightarrow \exists$ independent set of size $\ge n^2/(4m)$.

Proof: (1) Remove each vertex w.p. $1 - n/(2m)$. (2) For each remaining edge, delete one endpoint. Let $X = \#$surviving vertices, $Y = \#$surviving edges. $\mathbb{E}[X] = n\cdot\frac{n}{2m} = \frac{n^2}{2m}$, $\mathbb{E}[Y] = m\cdot(\frac{n}{2m})^2 = \frac{n^2}{4m}$. $\mathbb{E}[\text{returned}] \ge \mathbb{E}[X-Y] = \frac{n^2}{4m}$.

2.3 Max-Cut & Derandomization

Random partition into A,B: each edge cut w.p. $1/2$. $\mathbb{E}[\text{cut}] = |E|/2$.

Derandomization via Conditional Expectation: For $v_1,\dots,v_n$: assign v_t to whichever of A,B maximizes $\mathbb{E}[\text{cut}|v_1,\dots,v_t]$. Since $\mathbb{E}[\text{cut}|v_{\le t-1}] = \frac{1}{2}\mathbb{E}[\text{cut}|v_t \in A, v_{\le t-1}] + \frac{1}{2}\mathbb{E}[\text{cut}|v_t \in B, v_{\le t-1}]$, at least one term $\ge$ LHS. This is equivalent to: put v_t in set with fewer of its already-assigned neighbors = greedy algorithm!

2.4 k-SAT

Random assignment satisfies each k-clause w.p. $1-1/2^k$. $\mathbb{E}[\text{satisfied}] \ge (1-1/2^k)m$. Derand: assign x_t to maximize conditional expected # satisfied clauses.

3 L10: Second Moment Method & LLL

3.1 Second Moment Method

$\Pr[X=0] \le \mathrm{Var}[X]/(\mathbb{E}[X])^2$. (From Chebyshev with $t = |\mathbb{E}[X]|$.)

Use when X counts something, $\mathbb{E}[X]$ is large, want $\Pr[X \ge 1]$ is high.

4-Clique Threshold: $G_{n,p}$, $X = \#$4-cliques. $\mathbb{E}[X] = \binom{n}{4}p^6$. Upper: $p = c_1 n^{-2/3}$: $\mathbb{E}[X] \le c_1^6 < 0.1$ so $\Pr[X \ge 1] \le \mathbb{E}[X] < 0.1$. Lower: $p = c_2 n^{-2/3}$: Show $\mathrm{Var}[X] \ll (\mathbb{E}[X])^2$ by bounding correlated pairs ($O(n^7)$ dependent vs $O(n^8)$ independent). Second moment method gives $\Pr[X=0] < 0.1$ for large c_2.

3.2 Lovász Local Lemma (LLL)

Events $A_1,\dots,A_n$, $\Pr[A_i] \le p$, each A_i **mutually independent** of all but $\le d$ others.

Mutual independence: B mutually indep of $\{B_1,\dots,B_k\}$ if $\Pr[B] = \Pr[B|\bigcap_{j\in J} B_j]$ for all $J \subseteq \{1,\dots,k\}$.

Version I: $pd \le 1/4 \Rightarrow \Pr[\bigcap_i \overline{A_i}] \ge (1-2p)^n > 0$.

Version II: $p(d+1) \le 1/e \Rightarrow \Pr[\bigcap_i \overline{A_i}] \ge (1-\frac{1}{d+1})^n > 0$.

Proof (Version I): Key Lemma: $\Pr[A_i|\bigcap_{j\in S}\overline{A_j}] \le 2p$ for all S, $i \notin S$.

Induction on $|S|$. Base: $|S|=0$: $\Pr[A_i] \le p \le 2p$. Step: partition $S = S_{\text{ind}} \cup S_{\text{dep}}$ where A_i mutually indep of S_{ind}.

$$\Pr[A_i|\bigcap_{j\in S}\overline{A_j}] = \frac{\Pr[A_i \cap (\bigcap_{j\in S_{\text{dep}}}\overline{A_j})|\bigcap_{j\in S_{\text{ind}}}\overline{A_j}]}{\Pr[\bigcap_{j\in S_{\text{dep}}}\overline{A_j}|\bigcap_{j\in S_{\text{ind}}}\overline{A_j}]}$$

Numerator $\le \Pr[A_i|\bigcap_{j\in S_{\text{ind}}}\overline{A_j}] = \Pr[A_i] \le p$ (mutual independence). Denominator $\ge 1 - \sum_{j\in S_{\text{dep}}}\Pr[A_j|\bigcap_{j\in S_{\text{ind}}}\overline{A_j}] \ge 1 - d\cdot 2p \ge 1 - 1/2 = 1/2$ (inductive hypothesis, $|S_{\text{ind}}| \le |S|-1$). So ratio $\le p/(1/2) = 2p$.

Then $\Pr[\bigcap_i \overline{A_i}] = \prod_i(1 - \Pr[A_i|\overline{A_1}\cap\cdots\cap\overline{A_{i-1}}]) \ge (1-2p)^n$.

3.3 Asymmetric LLL

$\exists r_i \in [0,1)$: $\Pr[A_i] \le r_i\prod_{j\notin S_i}(1-r_j) \Rightarrow \Pr[\bigcap_i \overline{A_i}] \ge \prod_i(1-r_i)$.

3.4 LLL Application: k-SAT

Each clause has exactly k vars, each var in $\leq 2^{k-2}/k$ clauses $\Rightarrow$ satisfiable!

$p = 1/2^k$, $d = k \cdot 2^{k-2}/k = 2^{k-2}$. Check: $pd = 2^{-k} \cdot 2^{k-2} = 1/4$. ✓

Example: 10-SAT, each var in $\leq 2^6 = 64$ clauses $\Rightarrow$ satisfiable.

4 L12: Constructive LLL (Moser-Tardos)

Setup: V = finite set of indep random vars. $\mathcal{A}$ = events determined by V. $\mathrm{vbl}(A)$ = variables determining A. $\Gamma(A) = \{B : \mathrm{vbl}(A) \cap \mathrm{vbl}(B) \neq \emptyset\}$.

Algorithm: (1) Random assignment to all vars. (2) While $\exists A \in \mathcal{A}$ that is true: pick any such A, resample $\mathrm{vbl}(A)$.

Thm (Moser-Tardos): If $\exists x : \mathcal{A} \to (0,1)$ s.t. $\Pr[A] \leq x(A) \prod_{B \in \Gamma(A) \setminus \{A\}} (1 - x(B))$ for all A, then algorithm terminates. Expected #resamplings $\leq \sum_A \frac{x(A)}{1-x(A)}$.

Symmetric Corollary: $|\Gamma(A)| \leq d+1$, $\Pr[A] \leq \frac{1}{e(d+1)} \Rightarrow$ expected resamplings $O(|\mathcal{A}|/(d+1))$.

Proof: Set $x(A) = 1/(d+1)$. $x(A) \prod_{B \in \Gamma(A)\setminus\{A\}} (1 - x(B)) \geq \frac{1}{d+1}(1 - \frac{1}{d+1})^d \geq \frac{1}{e(d+1)} \geq \Pr[A]$. ✓

4.1 Moser's Entropic Proof (k-SAT)

Idea: If algorithm runs too long $\Rightarrow$ print statements compress random bits $\Rightarrow$ contradiction with incompressibility.

Incompressibility: For any injective f from t-bit strings: $\Pr[|f(s)| \leq t - c] \leq 1/2^{c-1}$.

FIX runs for T steps using $Tk + n$ random bits. Print statements encode these in $\log m + n + T(\log(d+1) + O(1))$ bits. Contradiction if $d + 1 \leq 2^{k-c_3}$ for constant c_3.

$\Rightarrow O(m \log m)$ resamplings w.h.p., poly time in n, m.

5 L13: Markov Chains & 2-SAT

5.1 Markov Chain Definitions

$X_0, X_1, \ldots$ is **Markov chain** if $\Pr[X_t | X_0, \ldots, X_{t-1}] = \Pr[X_t | X_{t-1}]$.

Time homogeneous: $\Pr[X_t = a | X_{t-1} = b] = P_{b,a}$ (indep of t).

Transition matrix P**:** row stochastic. Distribution after t steps from v: vP^t.

5.2 Randomized 2-SAT Algorithm

Input: 2-SAT formula, n vars. Start with any assignment A_0.

For $t = 1$ to cn^2: pick unsatisfied clause, flip one of its two vars (uniformly). If formula satisfied, return.

Analysis: Fix satisfying assignment S. $X_t = \#$ vars in A_t agreeing with S.

At each step: clause has ≥ 1 "wrong" var. Flip it $\Rightarrow X_t$ increases (w.p. $\geq 1/2$) or decreases (w.p. $\leq 1/2$).

Coupling: Define $Y_t \leq X_t$ with $\Pr[Y_t = Y_{t-1} \pm 1] = 1/2$, $Y_0 = X_0$. ($X_t - Y_t$ always even.)

Hitting time lemma: $r_i = \mathbb{E}[\text{time for } Y \text{ to hit } n \,|\, Y_0 = i]$. $r_n = 0$, $r_0 = 1 + r_1$, $r_i = 1 + \frac{1}{2} r_{i-1} + \frac{1}{2} r_{i+1}$.

Solution: $r_i = r_{i+1} + 2i + 1$, so $r_i = \sum_{j=i}^{n-1} (2j+1)$, $r_0 = n^2$.

By Markov: w.p. $\geq 1/2$, Y hits n within $2n^2$ extra steps. Over $c/2$ blocks: $\Pr[\text{fail}] \leq (1/2)^{c/2}$.

Thm: If satisfiable, algorithm returns satisfying assignment w.p. $\geq 1 - 1/2^{c/2}$.

5.3 Schöning's 3-SAT

Repeat: random assignment, then $3n$ greedy steps (pick unsatisfied clause, flip random var). Progress prob $\geq 1/3$ per step, but net drift is negative. Random restarts save: expected runtime $O((4/3)^n)$.

6 L14: Fundamental Thm & Metropolis

6.1 Classification of States

Irreducible: $\forall i, j$: $\sum_{t \geq 0} \Pr[X_t = j | X_0 = i] > 0$ (can reach any state from any state).

Recurrent: $r_i = \Pr[\text{return to } i \,|\, \text{start at } i] = 1$. **Transient:** $r_i < 1$. Expected visits $= 1/(1 - r_i)$.

Aperiodic: $\gcd\{t : \Pr[X_t = i | X_0 = i] > 0\} = 1$ for all i.

Fact: Self-loop at any state + irreducible $\Rightarrow$ aperiodic.

Gambler's Ruin: ± 1 fair walk on $\mathbb{Z}$. State 0 is recurrent (will return w.p. 1), but $\mathbb{E}[\text{return time}] = \infty$.

6.2 Fundamental Theorem of Markov Chains

Finite state space, irreducible, aperiodic $\Rightarrow$:

1. $\exists!$ stationary distribution π: $\lim_{t\to\infty} \Pr[X_t = i | X_0 = j] = \pi_i$ for all i, j
2. $\pi_i = 1/\mathbb{E}[\min\{t \geq 1 : X_t = i\} | X_0 = i]$ (inverse return time)
3. $\pi P = \pi$ (left eigenvector, eigenvalue 1)

Uniqueness $\Rightarrow$ to find π: guess and verify $\pi P = \pi$!

Symmetric transitions ($P_{ij} = P_{ji}$): π = uniform.

Proof: $(\pi P)(i) = \sum_j \pi_j P_{ji} = \sum_j \pi_j P_{ij} = \frac{1}{|S|} \sum_j P_{ij} = \frac{1}{|S|} = \pi_i$.

Random walk on undirected graph: $\pi(v) = \deg(v)/(2|E|)$.

Proof: Prob entering v: $\sum_{u \sim v} \frac{\deg(u)}{2|E|} \cdot \frac{1}{\deg(u)} = \frac{\deg(v)}{2|E|} = \pi(v)$.

Not irreducible: Decompose into strongly connected components, each with own stationary dist.

Periodic: Still has stationary dist ($\pi P = \pi$), but lim doesn't exist (depends on $t \mod \text{period}$). Time-average still converges.

6.3 Card Shuffling

Choose 2 cards at random, swap positions. Symmetric transitions $\Rightarrow$ stationary = uniform over $n!$ orderings.

6.4 Metropolis Algorithm (MCMC)

Target distribution π, given connected graph G on states, $d > \max\deg(G)$:

$$P_{ij} = \begin{cases} 0 & i, j \text{ not neighbors} \\ \frac{1}{d} \min\left(1, \frac{\pi(j)}{\pi(i)}\right) & i \neq j, \text{ neighbors} \\ 1 - \sum_{\ell \neq i} P_{i\ell} & i = j \end{cases}$$

Properties: Irreducible (connected graph), aperiodic ($d >$ max degree $\Rightarrow$ self-loops), stationary distribution $= \pi$.

Proof π is stationary: Show flow balance at each node. Prob mass leaving i: $\pi(i) \sum_{j \sim i} \frac{1}{d} \min(1, \frac{\pi(j)}{\pi(i)})$. Prob mass entering i: $\sum_{j \sim i} \pi(j) \frac{1}{d} \min(1, \frac{\pi(i)}{\pi(j)})$. These are equal (case split on $\pi(i) \gtrless \pi(j)$).

Key advantage: Only need $\pi(i)/\pi(j)$, not π itself (e.g., for distributions with unknown normalizing constants).

Graph coloring example: Choose random vertex + color. Accept if valid coloring, else stay. Symmetric transitions $\Rightarrow$ uniform stationary dist over proper colorings.

7 L15: Mixing Times & Coupling

7.1 Total Variation Distance

$$\begin{aligned} \|D_1 - D_2\|_{\mathrm{TV}} &= \tfrac{1}{2} \textstyle\sum_s |D_1(s) - D_2(s)| \\ &= \max_{A \subseteq S} |\Pr_{D_1}[A] - \Pr_{D_2}[A]| \end{aligned}$$

Coupling characterization: $\|D_1 - D_2\|_{\mathrm{TV}} = \min_{(X,Y)} \Pr[X \neq Y]$ over joint distributions with marginals D_1, D_2. For *any* coupling: $\|D_1 - D_2\| \leq \Pr[X \neq Y]$.

Proof sketch: Let $p = \sum_x \min(D_1(x), D_2(x))$. $\|D_1 - D_2\| = 1 - p$. Optimal coupling: w.p. p, $X = Y$ from shared mass; w.p. $1 - p$, draw independently from excess.

7.2 Mixing Time

$\Delta(t) = \max_s \|\pi - P_s^t\|_{\mathrm{TV}}$ (worst-case distance to stationarity).

$\tau_{\mathrm{mix}} = \min\{t : \Delta(t) \leq 1/(2e)\}$.

Key facts: $\Delta(t)$ is non-increasing. $\Delta(c \cdot \tau_{\mathrm{mix}}) \leq 1/e^c$ (exponential decay).

$\Delta(t) \leq \max_{s,s'} \|P_s^t - P_{s'}^t\| \leq 2\Delta(t)$.

7.3 Coupling for Mixing Time

Coupling: Joint process (X_t, Y_t) s.t. (1) each marginal follows the Markov chain, (2) once $X_t = Y_t$, they stay together.

Coupling Lemma: $\Delta(t) \le \max_{s,s'} \Pr[T_{s,s'} > t]$ where $T_{s,s'} = \min\{t : X_t = Y_t\}$.

Strategy: Design coupling that makes chains meet quickly for *worst-case* starting states.

Example: Proper Graph Colorings. Graph max degree d, k colors, $k \ge 4d + 1$.

Coupling: X, Y choose same vertex v and same color c. Let $Z_t = \#$ differing vertices.

Prob of progress ($Z_{t+1} = Z_t - 1$): $\ge Z_t(k-2d)/(nk)$. Prob of regression ($Z_{t+1} = Z_t + 1$): $\le 2dZ_t/(nk)$.

$\mathbb{E}[Z_{t+1}|Z_t] \le Z_t(1 - \frac{k-4d}{kn})$. So $\mathbb{E}[Z_t] \le ne^{-t/(kn)}$.

$\Pr[Z_t > 0] \le \mathbb{E}[Z_t] \le ne^{-t/(kn)}$.

For $t \ge kn(2 + \log n)$: $\Pr[Z_t > 0] < 1/(2e)$. $\Rightarrow \tau_{\text{mix}} \le kn(2 + \log n)$.

7.4 Strong Stationary Stopping Times

T: random variable (depending on $X_0, \dots, X_T$) s.t. $\Pr[X_t = s \,|\, t \ge T] = \pi(s)$.

$\Delta(t) \le \Pr[T > t]$.

Top-in-at-Random Shuffle: Remove top card, insert at random position. $T =$ step after bottom card reaches top. Cards below bottom card are uniformly random. $\mathbb{E}[T] = n/1 + n/2 + \cdots \approx n \ln n$. Markov: $\Pr[T > 2en \ln n] \le 1/(2e)$. $\tau_{\text{mix}} \le 2en \ln n$.

Random-to-Top Shuffle: Pick random card, move to top. Coupling: both chains pick same card. Time to couple = coupon collector $\approx n \ln n$.

8 L16: Martingales & Azuma-Hoeffding

8.1 Martingale Definition

$\{Z_t\}$ is a **martingale** w.r.t. $\{X_t\}$ if:

1. Z_t is a function of $X_0, \dots, X_t$
2. $\mathbb{E}[|Z_t|] < \infty$
3. $\mathbb{E}[Z_t | X_0, \dots, X_{t-1}] = Z_{t-1}$

Consequence: $\mathbb{E}[Z_t] = \mathbb{E}[Z_0]$ for any fixed t.

$Y_t = Z_t - Z_{t-1}$: martingale differences. $\mathbb{E}[Y_t | X_0, \dots, X_{t-1}] = 0$.

8.2 Doob Martingale

Given r.v. A and sequence $\{X_t\}$: $Z_t = \mathbb{E}[A | X_0, \dots, X_t]$ is a martingale w.r.t. $\{X_t\}$.

$\mathbb{E}[\mathbb{E}[A|X_0, \dots, X_t] \,|\, X_0, \dots, X_{t-1}] = \mathbb{E}[A | X_0, \dots, X_{t-1}] = Z_{t-1}$. ✓

Edge exposure: Reveal edges one at a time. #terms = $\binom{n}{2}$.

Vertex exposure: Reveal neighborhoods one vertex at a time. #terms = n. Usually better for Azuma (fewer terms $\Rightarrow$ tighter bound).

8.3 Azuma-Hoeffding Inequality

$\{Z_t\}$ martingale w.r.t. $\{X_t\}$, $|Z_i - Z_{i-1}| \le c_i$ for all $i \le n$:

$$\Pr[|Z_n - Z_0| \ge \lambda] \le 2 \exp\left(\frac{-\lambda^2}{2\sum_{i=1}^n c_i^2}\right)$$

Key Lemma: If $Y \in [-c, c]$, $\mathbb{E}[Y] = 0$: $\mathbb{E}[e^{tY}] \le e^{t^2c^2/2}$.

Proof of lemma: Convexity: $e^{tx} \le \frac{(1-x)e^{-t} + (1+x)e^t}{2}$. Take expectation: $\mathbb{E}[e^{tY}] \le \frac{e^{-t}+e^t}{2} = \cosh(t) \le e^{t^2/2}$ (compare Taylor series term by term).

Proof of theorem: $\Pr[Z_n - Z_0 \ge \lambda] \le \mathbb{E}[e^{s(Z_n - Z_0)}]/e^{s\lambda}$. Iteratively peel off: $\mathbb{E}[e^{s(Z_n - Z_0)}] \le e^{s^2 \sum c_i^2/2}$. Set $s = \lambda / \sum c_i^2$.

8.4 Method of Bounded Differences

$X_1, \dots, X_n$ **independent**, $A = A(X_1, \dots, X_n)$. If $|A(\dots, x_i, \dots) - A(\dots, x_i', \dots)| \le c_i$ for all inputs:

$$\Pr[|A - \mathbb{E}A| > \lambda] \le 2 \exp\left(\frac{-\lambda^2}{2\sum c_i^2}\right)$$

Without independence: Replace condition with $|\mathbb{E}[A|X_{\le t-1}, X_t = a] - \mathbb{E}[A|X_{\le t-1}, X_t = a']| \le c_t$ for all a, a'.

8.5 Applications

Chromatic # of $G_{n,p}$: Vertex exposure martingale. $|Z_t - Z_{t-1}| \le 1$ (changing one vertex's neighborhood changes χ by ≤ 1). $\Pr[|\chi - \mathbb{E}[\chi]| \ge c\sqrt{n}] \le 2e^{-c^2/2}$.

Compare: edge exposure would give $c_i = 1$ but $n^2/2$ terms $\Rightarrow$ much weaker: $\Pr[\cdots] \le 2e^{-c^2/n}$.

Gambling: Fair game, bet $\le B$ per round. $Z_t =$ net winnings. $|Z_t - Z_{t-1}| \le B$. $\Pr[|Z_t| \ge \lambda] \le 2e^{-\lambda^2/(2tB^2)}$. Holds for *any adaptive* betting strategy.

Empty bins (HW8): n balls, m bins, $Z = \#$empty bins. Doob martingale w.r.t. ball placements. $|Z_{k+1} - Z_k| \le 1$. $\Pr[|Z - \mathbb{E}Z| \ge \epsilon n] \le 2e^{-\epsilon^2 n/2}$.

Coin flips: $X_i \in \{0, 1\}$ iid, $S = \sum X_i$. $|Z_i - Z_{i-1}| = |X_i - 1/2| = 1/2$. $\Pr[|S - n/2| \ge c\sqrt{n}] \le 2e^{-2c^2}$.

9 L17: Martingale Stopping Thm

9.1 Stopping Time

T is a **stopping time** for $\{X_t\}$ if $\{T = i\}$ depends only on $X_0, \dots, X_i$ (can't look into future).

Examples: $T_1 = \min\{t : Z_t = 10\}$ (✓). $T_2 = \min\{t : Z_t \notin [-5, 10]\}$ (✓). $T_3 = \max\{t : Z_t > 0\}$ (×, requires future).

9.2 Martingale Stopping Theorem (MST)

$\{Z_t\}$ martingale w.r.t. $\{X_t\}$, T stopping time $\Rightarrow \mathbb{E}[Z_T] = \mathbb{E}[Z_0]$ if **any one** of:

1. $\exists c$: $|Z_i| \le c$ for all i (bounded values)
2. $\exists c$: $T < c$ w.p. 1 (bounded time)
3. $\mathbb{E}[T] < \infty$ AND $\exists c$: $\mathbb{E}[|Z_{i+1} - Z_i| \,|\, X_0, \dots, X_i] < c$ (bounded steps + finite expected time)

Counterexample: $T' = \min\{t : Z_t = 10\}$ for fair ± 1 walk starting at 0. $\mathbb{E}[Z_{T'}] = 10 \ne 0$. None of the conditions hold: T' unbounded, values unbounded, $\mathbb{E}[T'] = \infty$.

9.3 Hitting Times: Fair Walk

$Z_0 = 0$, fair ± 1 walk. $T = \min\{t : Z_t \in \{-a, b\}\}$.

$$\Pr[Z_T = b] = \frac{a}{a+b}, \qquad \mathbb{E}[T] = ab.$$

Proof (probability): Z_t bounded in $[-a, b] \Rightarrow$ Condition 1. MST: $0 = \mathbb{E}[Z_T] = b \cdot p - a \cdot (1-p)$. Solve: $p = a/(a+b)$.

Proof ($\mathbb{E}[T]$): $Y_t = Z_t^2 - t$ is martingale: $\mathbb{E}[Y_t | Z_{t-1}] = \frac{1}{2}(Z_{t-1}+1)^2 + \frac{1}{2}(Z_{t-1}-1)^2 - t = Z_{t-1}^2 + 1 - t = Z_{t-1}^2 - (t-1) = Y_{t-1}$.

Check Condition 3: $|Y_{t+1} - Y_t| = |Z_t^2 - Z_{t-1}^2 - 1| \le 1 + 2\max(a, b)$. $\mathbb{E}[T] < \infty$ because w.p. $\ge 1/2^{a+b}$, game ends in $a + b$ steps.

MST: $0 = \mathbb{E}[Y_T] = \mathbb{E}[Z_T^2] - \mathbb{E}[T]$. $\mathbb{E}[Z_T^2] = a^2 \cdot \frac{b}{a+b} + b^2 \cdot \frac{a}{a+b} = \frac{ab(a+b)}{a+b} = ab$. So $\mathbb{E}[T] = ab$.

9.4 Biased Random Walk

Win \$1 w.p. p, lose \$1 w.p. $1 - p$, $p \ne 1/2$. Let $c = \frac{1-p}{p}$.

$Y_t = c^{Z_t}$ is martingale: $\mathbb{E}[Y_t | Z_{t-1}] = pc^{Z_{t-1}+1} + (1-p)c^{Z_{t-1}-1} = c^{Z_{t-1}}(pc + (1-p)/c) = c^{Z_{t-1}} = Y_{t-1}$.

$T = \min\{t : Z_t \in \{-a, b\}\}$. MST (bounded values): $\mathbb{E}[Y_T] = Y_0 = 1$. $c^{-a}\Pr[Z_T = -a] + c^b \Pr[Z_T = b] = 1$. Solve:

$$\Pr[Z_T = -a] = \frac{1 - c^b}{c^{-a} - c^b}.$$

For $\mathbb{E}[T]$: $Q_t = Z_t - (2p-1)t$ is martingale. MST: $\mathbb{E}[Q_T] = 0$, so $\mathbb{E}[T] = \frac{(-a)\Pr[Z_T = -a] + b\Pr[Z_T = b]}{2p-1}$.

9.5 Ballot Theorem

n votes, $N_A > N_B$, tallied randomly. $\Pr[A \text{ ahead at all steps}] = \frac{N_A - N_B}{n}$.

$Z_t = \frac{X_{n-t}}{n-t}$ (backward count) is martingale. Stopping time T: first t with $Z_t = 0$ (or $n - 1$). MST: $\mathbb{E}[Z_T] = (N_A - N_B)/n$. If A always ahead: $Z_T = Z_{n-1} = X_1/1 = 1$. Else $Z_T = 0$.

$\Pr[A \text{ always ahead}] \cdot 1 + \Pr[\text{not}] \cdot 0 = (N_A - N_B)/n$.

10 L18: Pseudorandomness

10.1 Extractors

Min-entropy: $H_\infty(X) = \min_x \log(1/\Pr[X = x])$. k-source: $\Pr[X = x] \le 2^{-k}$ for all x.

(k, ϵ)-extractor: $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ s.t. $\|\text{Ext}(X, U_d) - U_m\| \le \epsilon$ for all k-sources X.

Seedless impossible: for $k \le n - 1$: $\exists$ k-source with first output bit always 1.

Existence (probabilistic method): $m = k + d - 2\log(1/\epsilon) - O(1)$, $d = \log(n-k) + 2\log(1/\epsilon) + O(1)$.

For $\epsilon = O(1)$, $k \ll n$: $d \approx \log n$, $m \approx k + d$. Can exhaust over $2^d \approx n$ seeds in $O(n)$ time.

10.2 Expander Graphs

G: Δ-regular, N vertices. M = normalized adjacency matrix. $\lambda_1 = 1 \ge \lambda_2 \ge \cdots \ge \lambda_N$.

$\lambda(G) := \max\{|\lambda_2|, |\lambda_N|\}$ (spectral gap parameter).

Thm: $\|\pi - P_i^t\|_{\text{TV}} \le \lambda(G)^t \sqrt{N}$.

$\lambda(G) < 1 - \delta$ for constant $\delta \Rightarrow \tau_{\text{mix}} = O(\log N)$.

Proof: $\sigma = \pi + v$ where $v \perp \pi$. $\|M^t v\|_2 \le \lambda(G)^t \|v\|_2$. Convert $\ell_2 \to \ell_1$ via Cauchy-Schwarz.

Alon-Boppana: $\lambda(G) \ge 2\sqrt{\Delta - 1}/\Delta$ (best possible). Ramanujan graph achieves this.

10.3 Extractors from Expanders

ℓ-step walk on $N = 2^n$-vertex expander, starting from k-source. Output = final vertex. Seed = walk directions ($d = \ell \log \Delta$). $\ell = n/2 - k/2 + \log(1/\epsilon) + O(1)$.

$\|M^\ell \sigma - \pi\|_{\text{TV}} \le \frac{\sqrt{N}}{2} \lambda(G)^\ell \cdot 2^{-k/2+1} \le \epsilon$.

11 Key Problem Patterns

11.1 When to Use What

Prob. Method: Exists a good object $\Rightarrow$ random construction, show good w.p. > 0 or $\mathbb{E}[\text{quality}] \ge \alpha$.

2nd Moment: $\Pr[X \ge 1]$ when X counts something. Need $\mathbb{E}[X]$ large, $\text{Var}[X]$ small.

LLL: Many bad events, each low probability and only locally dependent. Check $pd \le 1/4$.

Derand: Prob method gives existence $\Rightarrow$ sequentially fix variables to maintain conditional expectation.

Coupling: Bound mixing time. Design joint process, bound Pr[not coupled by t].

SST: Bound mixing time via a natural "milestone" event.

Azuma: Concentration of function of independent (or Markov-dependent) variables. Identify martingale, bound differences c_i.

MST: Find $\mathbb{E}$[hitting time] or Pr[outcome]. Construct martingale (often $Z_t^2 - t$ or c^{Z_t}), apply stopping theorem.

12 All Useful Inequalities & When to Use

12.1 Basic Probability Inequalities

Union Bound: $\Pr[\bigcup_i A_i] \le \sum_i \Pr[A_i]$.

When: Upper bound on prob that *any* bad event occurs. Often combined with Chernoff/Markov to bound each $\Pr[A_i]$.

Markov's Inequality: $X \ge 0 \Rightarrow \Pr[X \ge t] \le \mathbb{E}[X]/t$.

When: Only know $\mathbb{E}[X]$, X non-negative. Weak but universal. Also: $\Pr[X \ge c\mathbb{E}[X]] \le 1/c$.

Chebyshev's Inequality: $\Pr[|X - \mathbb{E}X| \ge t] \le \text{Var}(X)/t^2$.

When: Know $\mathbb{E}[X]$ and $\text{Var}(X)$. Works for *any* distribution. Only needs pairwise independence for $\text{Var}[\sum X_i] = \sum \text{Var}[X_i]$. Gives $\Pr[|X - \mu| \ge c\sigma] \le 1/c^2$.

Second Moment Method: $\Pr[X = 0] \le \text{Var}[X]/(\mathbb{E}[X])^2$.

When: X counts objects (integer-valued, ≥ 0), want to show $X > 0$ w.h.p. Need $\mathbb{E}[X]$ large and $\text{Var}[X]$ not too large relative to $(\mathbb{E}[X])^2$.

12.2 Chernoff-type Bounds

Chernoff (Multiplicative): $X = \sum X_i$, *independent* $X_i \in \{0,1\}$, $\mu = \mathbb{E}[X]$:

- $\Pr[X \ge (1+\delta)\mu] \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$
- For $\delta \in (0,1]$: $\Pr[X \ge (1+\delta)\mu] \le e^{-\delta^2\mu/3}$
- $\Pr[X \le (1-\delta)\mu] \le e^{-\delta^2\mu/2}$ for $\delta \in (0,1]$
- $c \ge 6$: $\Pr[X \ge c\mu] \le 2^{-c\mu}$

When: Sum of **independent** bounded r.v.'s. Gives exponential tails. Use multiplicative form when deviation is proportional to μ.

Hoeffding's Inequality: $X_i \in [a_i, b_i]$ *independent*, $S = \sum X_i$:

$$\Pr[|S - \mathbb{E}S| \ge t] \le 2\exp\left(\frac{-2t^2}{\sum_i (b_i - a_i)^2}\right)$$

When: Sum of independent bounded r.v.'s but not necessarily 0/1. Use additive form when deviation is absolute. Typical: $X_i \in [0,1]$ gives $\Pr[|S - \mu| \ge t] \le 2e^{-2t^2/n}$, so $t = O(\sqrt{n})$ suffices.

Poisson Tail: $X \sim \text{Poi}(\lambda)$: $\Pr[|X - \lambda| \ge c] \le 2\exp\left(\frac{-c^2}{2(c+\lambda)}\right)$.

When: Poisson r.v. (e.g., limit of balls-in-bins).

12.3 Martingale Concentration

Azuma-Hoeffding: $\{Z_t\}$ martingale w.r.t. $\{X_t\}$, $|Z_i - Z_{i-1}| \le c_i$:

$$\Pr[|Z_n - Z_0| \ge \lambda] \le 2\exp\left(\frac{-\lambda^2}{2\sum_{i=1}^n c_i^2}\right)$$

When: Martingale with bounded differences. Key use: Doob martingale $Z_t = \mathbb{E}[A|X_1, \ldots, X_t]$ where A depends on independent (or sequentially revealed) X_i's. *Does not require independence of X_i's!*

Method of Bounded Differences: $X_1, \ldots, X_n$ **independent**, $A = A(X_1, \ldots, X_n)$, changing X_i changes A by $\le c_i$:

$$\Pr[|A - \mathbb{E}A| > \lambda] \le 2\exp\left(\frac{-\lambda^2}{2\sum c_i^2}\right)$$

When: Function of independent r.v.'s with bounded "Lipschitz" condition. This is Azuma applied to the Doob martingale. *Requires independence* for the simple deterministic condition. Without independence, need: $|\mathbb{E}[A|X_{\le t-1}, X_t{=}a] - \mathbb{E}[A|X_{\le t-1}, X_t{=}a']| \le c_t$.

12.4 Probabilistic Method Inequalities

LLL (Version I): $\Pr[A_i] \le p$, each mutually indep of all but $\le d$: $pd \le 1/4 \Rightarrow \Pr[\bigcap_i \overline{A_i}] > 0$.

LLL (Version II): $p(d+1) \le 1/e \Rightarrow \Pr[\bigcap_i \overline{A_i}] \ge (1 - \frac{1}{d+1})^n > 0$.

When: Many bad events with local dependencies. Union bound fails ($np \gg 1$), but dependencies are limited.

Asymmetric LLL: $\exists r_i \in [0,1)$: $\Pr[A_i] \le r_i \prod_{j \notin S_i}(1 - r_j)$ $\Rightarrow \Pr[\bigcap_i \overline{A_i}] \ge \prod_i (1 - r_i)$.

When: Events have different probabilities or different dependency degrees.

12.5 Mixing & Spectral Bounds

Coupling bound: $\Delta(t) \le \max_{s,s'} \Pr[T_{s,s'} > t]$.

SST bound: $\Delta(t) \le \Pr[T > t]$.

Spectral: $\|\pi - P_i^t\|_{\text{TV}} \le \lambda(G)^t \sqrt{N}$.

$\Delta(c \cdot \tau_{\text{mix}}) \le 1/e^c$ (exponential decay after mixing).

12.6 Algebraic & Combinatorial

- $1 - x \le e^{-x}$ for all x (very common trick)
- $(1 - 1/n)^n \approx 1/e$; more generally $(1-p)^{1/p} \to 1/e$
- $(n/k)^k \le \binom{n}{k} \le (en/k)^k$; also $\binom{n}{k} \le n^k/k!$
- $\sum_{i=1}^n 1/i = \ln n + \gamma + O(1/n) = \Theta(\log n)$
- $\sum_{i=1}^n 1/i^c = O(1)$ for $c > 1$
- $n! = e^{\Theta(n \log n)}$ (Stirling: $n! \approx \sqrt{2\pi n}(n/e)^n$)
- $\ln \frac{1}{1-x} = x + x^2/2 + x^3/3 + \cdots$ for $|x| < 1$
- $\sum_{j=0}^k 2^{-j} = 2 - 2^{-k}$
- $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
- $X = \sum X_i$: $\text{Var}[X] = \sum \text{Var}[X_i] + 2\sum_{i<j} \text{Cov}(X_i, X_j)$
- Pairwise indep $\Rightarrow \text{Var}[\sum X_i] = \sum \text{Var}[X_i]$
- Indicator: $\text{Var}[X_i] = p_i(1 - p_i) \le 1/4$

12.7 Distributions

- $X \sim \text{Ber}(p)$: $X \in \{0,1\}$, $\Pr[X{=}1] = p$, $\mathbb{E}[X] = p$, $\text{Var}[X] = p(1-p)$
- $X \sim \text{Bin}(n,p)$: $\mathbb{E}[X] = np$, $\text{Var}[X] = np(1-p)$
- $X \sim \text{Poi}(\lambda)$: $\Pr[X{=}k] = e^{-\lambda}\lambda^k/k!$, $\mathbb{E}[X] = \text{Var}[X] = \lambda$
- $X \sim \text{Geom}(p)$: $\mathbb{E}[X] = 1/p$, $\text{Var}[X] = (1-p)/p^2$
- $X \sim N(\mu,\sigma^2)$: pdf $\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$

12.8 Inequality Selection Guide

- Know only $\mathbb{E}[X]$, $X \ge 0$? → **Markov**
- Know $\mathbb{E}[X]$ and $\text{Var}(X)$? → **Chebyshev**
- Sum of indep bounded? → **Chernoff/Hoeffding** (exponential tail)
- Only pairwise indep? → **Chebyshev** (can't use Chernoff!)
- Function of indep vars? → **Azuma** (Doob + bounded differences)
- Dependent vars but bounded martingale differences? → **Azuma** directly
- Want $\Pr[X > 0]$ for a count X? → **Second moment** ($\Pr[X{=}0] \le \text{Var}/\mu^2$)
- Many local-dep bad events? → **LLL** (check $pd \le 1/4$)

13 Additional Topics & Techniques

13.1 Variance Calculations

$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

$X = \sum X_i$: $\text{Var}[X] = \sum \text{Var}[X_i] + 2\sum_{i<j}\text{Cov}(X_i,X_j)$.

Pairwise independent ⇒ $\text{Var}[\sum X_i] = \sum \text{Var}[X_i]$.

X_i indicator: $\text{Var}[X_i] = p_i(1-p_i) \le 1/4$.

13.2 Tail Bounds Summary

For $S = \sum_{i=1}^n X_i$, $X_i \in \{0,1\}$ iid w/ $\mathbb{E}[X_i] = p$:

- $\Pr[S \ge pn + t]$: need $t = O(\sqrt{n})$ for const prob (Hoeffding)
- $\Pr[|Y - \mathbb{E}Y| \ge t]$ where $Y = \sqrt{\sum Y_i^2}$, $Y_i \in [0,1]$ iid: use Azuma with $c_i = 1 \Rightarrow \le 2e^{-t^2/(2n)}$
- Pairwise indep $Y_i \in \{0,1\}$, $\mathbb{E}[Y_i] = 1/2$: $\Pr[\sum Y_i = 0] \le O(1/k)$ (2nd moment)
- Graph n verts, $m = 20n$ edges: ∃ indep set $\Omega(n)$ (prob method, $n^2/(4m) = n/80$)

13.3 Common Martingale Constructions

- Fair walk Z_t: Z_t itself is martingale
- Fair walk Z_t: $Z_t^2 - t$ is martingale (for $\mathbb{E}[T]$)
- Biased walk ($p \ne 1/2$): c^{Z_t} where $c = (1-p)/p$
- Biased walk: $Z_t - (2p-1)t$ is martingale
- Doob: $\mathbb{E}[A|X_1,\dots,X_t]$ for any function A
- Voter model: $\sum_v \deg(v)\mathbf{1}[\sigma(v) = s]$
- Hat problem: $Y_j = \sum_{i=1}^j (X_i - \mathbb{E}[X_i|X_{<i}])$

13.4 MST Condition Checklist

1. Check: is $\{Z_t\}$ actually a martingale?
2. Find/construct a stopping time T
3. Verify one of the 3 conditions:
   - Bounded values? ($|Z_i| \le c$)
   - Bounded time? ($T < c$ w.p. 1)
   - Finite $\mathbb{E}[T]$ + bounded steps? ($\mathbb{E}[|Z_{i+1} - Z_i| | \cdots] < c$)
4. Compute $\mathbb{E}[Z_T]$ from boundary values and apply $\mathbb{E}[Z_T] = \mathbb{E}[Z_0]$

13.5 Practice Final Q5: Hat Problem

n people, random permutations, X_i = people finding hats in round i. R = total rounds.

$\mathbb{E}[X_i|X_1,\dots,X_{i-1}] = 1$ (linearity: each of m remaining people has $1/m$ prob of finding hat, sum = 1).

$Y_j = \sum_{i=1}^j (X_i - 1)$ is martingale. MST (condition 3, $|Y_{i+1} - Y_i| \le 2$, $\mathbb{E}[R] < \infty$): $0 = \mathbb{E}[Y_R] = \mathbb{E}[\sum(X_i - 1)] = n - \mathbb{E}[R]$. So $\mathbb{E}[R] = n$.

13.6 Stationary Distribution Computation

For finite MC: write $\pi P = \pi$ componentwise + $\sum \pi_i = 1$. Solve linear system.

For doubly stochastic ($\sum_i P_{ij} = 1$ for all j): π = uniform.

For detailed balance: if $\pi_i P_{ij} = \pi_j P_{ji}$ for all i,j ("reversible"), then π is stationary.

14 Coupling Tricks & Examples

14.1 What Is a Coupling?

Joint process (X_t, Y_t) where: (1) each marginal follows the Markov chain independently, (2) once $X_t = Y_t$ they stay together forever. Goal: $\Pr[X_t \ne Y_t]$ small ⇒ $\Delta(t)$ small.

Coupling Lemma: $\Delta(t) \le \max_{s,s'} \Pr[T_{s,s'} > t]$ where $T_{s,s'} = \min\{t : X_t = Y_t\}$.

14.2 Trick 1: Use Same Randomness

Both chains make the *same random choice* at each step. Differences can only shrink.

Graph Coloring (L15): Both chains pick same vertex v and same color c. If v was a vertex where colorings differed, coloring it the same color fixes it. Track Z_t = #disagreements. $\mathbb{E}[Z_{t+1}|Z_t] \le Z_t(1 - \frac{k-4d}{kn})$ when $k \ge 4d+1$. $\mathbb{E}[Z_t] \le ne^{-t/(kn)}$. Use Markov: $\Pr[Z_t \ge 1] \le \mathbb{E}[Z_t] \le ne^{-t/(kn)} < 1/(2e)$ when $t = kn(2 + \log n)$.

Random-to-Top Shuffle (L15): Both chains pick same *card* (not position). After selecting a card, it goes to top in both decks. Chains agree on all selected cards (top of deck). Time to couple ≤ time until all n cards selected = coupon collector $\approx n \ln n$.

Hypercube Walk (HW7): Lazy walk on $\{0,1\}^n$. Both chains pick same coordinate i and same replacement bit b. After step, coord i agrees in both chains. Disagree on coord i only if it was never picked. $\mathbb{E}[\text{\#disagreements at } t] \le n(1-1/n)^t \le ne^{-t/n}$. Mixing time $O(n\log n)$.

14.3 Trick 2: Shared Event Coupling

Force both chains to visit a common state with probability $\ge \epsilon$ at each step.

Practice Final Q3: MC with $\Pr[X_{t+1} = 1|X_t = j] \ge \epsilon$ for all j.

Coupling: If $X_t \ne Y_t$, w.p. ϵ set $X_{t+1} = Y_{t+1} = 1$. W.p. $1-\epsilon$, each evolves independently with modified transitions: $\Pr[i \to j] = \frac{P_{ij}-\epsilon}{1-\epsilon}$ for $j = 1$; $\frac{P_{ij}}{1-\epsilon}$ for $j \ne 1$.

Verify marginals: $\epsilon \cdot \mathbf{1}[j=1] + (1-\epsilon)\cdot\frac{P_{ij} - \epsilon\mathbf{1}[j=1]}{1-\epsilon} = P_{ij}$. ✓

$\Pr[\text{not coupled by } t] \le (1-\epsilon)^t \le e^{-\epsilon t}$. Set $t = \log(2e)/\epsilon$: $\tau_{\text{mix}} \le O(1/\epsilon)$.

Key: Any MC where all states have $\ge \epsilon$ prob of reaching a common "bottleneck" state mixes in $O(1/\epsilon)$.

14.4 Trick 3: Monotone/Path Coupling

Track a distance metric $d(X_t, Y_t)$ and show it decreases in expectation.

General pattern: Define $Z_t = f(X_t, Y_t) \ge 0$. Show $\mathbb{E}[Z_{t+1}|Z_t] \le (1-\alpha)Z_t$. Then $\mathbb{E}[Z_t] \le Z_0 e^{-\alpha t}$. Since $\Pr[X_t \ne Y_t] = \Pr[Z_t \ge 1] \le \mathbb{E}[Z_t]$, get mixing in $O(\frac{1}{\alpha}\log Z_0)$.

Example: For graph coloring, Z_t = #differing vertices, $\alpha = \frac{k-4d}{kn}$, $Z_0 \le n$. Mixing $\le O(kn\log n/(k-4d))$.

14.5 Trick 4: Independent Until Meeting

Run X_t, Y_t independently until they happen to be in the same state. Simplest coupling but often gives bad bounds (e.g., $> n!$ for card shuffling).

When to use: Only when state space is small or chains naturally concentrate.

14.6 Coupling Validity Checklist

1. Marginals correct? Each chain individually follows the original transition probabilities
2. Stays together? If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$ (use same randomness for coupled state)
3. Worst case? Bound holds for *all* starting pairs (s, s')

14.7 LLL Application Template

1. Define random construction (usually uniform/independent)
2. Identify bad events A_i and compute $\Pr[A_i] \le p$
3. Determine dependency: A_i depends on variables shared with $\le d$ other events
4. Verify $pd \le 1/4$ (or $p(d+1) \le 1/e$)

5. Conclude existence (or use Moser-Tardos for algorithm)

14.8 Entropy/Compression Argument (L12)

If algorithm uses R random bits and runs for T steps, and we can recover R bits from C bits of output with $C < R$, then by incompressibility, this can happen with prob $\leq 2^{-(R-C-1)}$. Contradiction if algorithm runs for T steps too often.

14.9 Mixing Time Lower Bounds

Bottleneck: If $\exists S$ with $\pi(S) \leq 1/2$ and few edges between S and $\bar{S}$, then mixing is slow.

Conductance: $\Phi = \min_{S:\pi(S)\leq 1/2} \frac{\sum_{i\in S, j\notin S} \pi_i P_{ij}}{\pi(S)}$. $\tau_{\text{mix}} \geq \frac{1}{2\Phi}$.

14.10 Spectral Gap & Mixing

λ_2 = second largest eigenvalue of transition matrix. $\tau_{\text{mix}} = \Theta(1/(1-\lambda_2))$ for reversible chains.

For expanders: $\lambda(G) < 1 \Rightarrow$ rapid mixing. $\tau_{\text{mix}} \leq \frac{\log N}{1-\lambda(G)}$.

14.11 Key Proof Techniques Recap

- **Union bound:** $\Pr[\bigcup A_i] \leq \sum \Pr[A_i]$
- **Linearity of expectation:** always applies
- **Indicator r.v.'s:** convert counting to probabilities
- **Markov/Chebyshev:** first/second moment bounds
- **Chernoff:** exponential tail for sums of independent r.v.'s
- **Azuma-Hoeffding:** Chernoff-like for martingales
- **Prob method:** show $\mathbb{E}[X] \geq \alpha$ or $\Pr[X > 0] > 0$
- **2nd moment method:** $\Pr[X = 0] \leq \text{Var}[X]/(\mathbb{E}[X])^2$
- **LLL:** local dependencies, $pd \leq 1/4$
- **Conditional expectation:** derandomize
- **Coupling:** bound TV distance/mixing time
- **MST:** find expected hitting times, outcome probs