Randomized Algorithms

南京大学

尹一通

Random Walk

- stationary:
  - convergence;
  - stationary distribution;
- hitting time: time to reach a vertex;
- cover time: time to reach all vertices;
- mixing time: time to converge.

Mixing Time

Markov chain: $\mathfrak{M} = (\Omega, P)$

- mixing time: time to be close to the stationary distribution

Total Variation Distance

- two probability measures p, q over Ω:

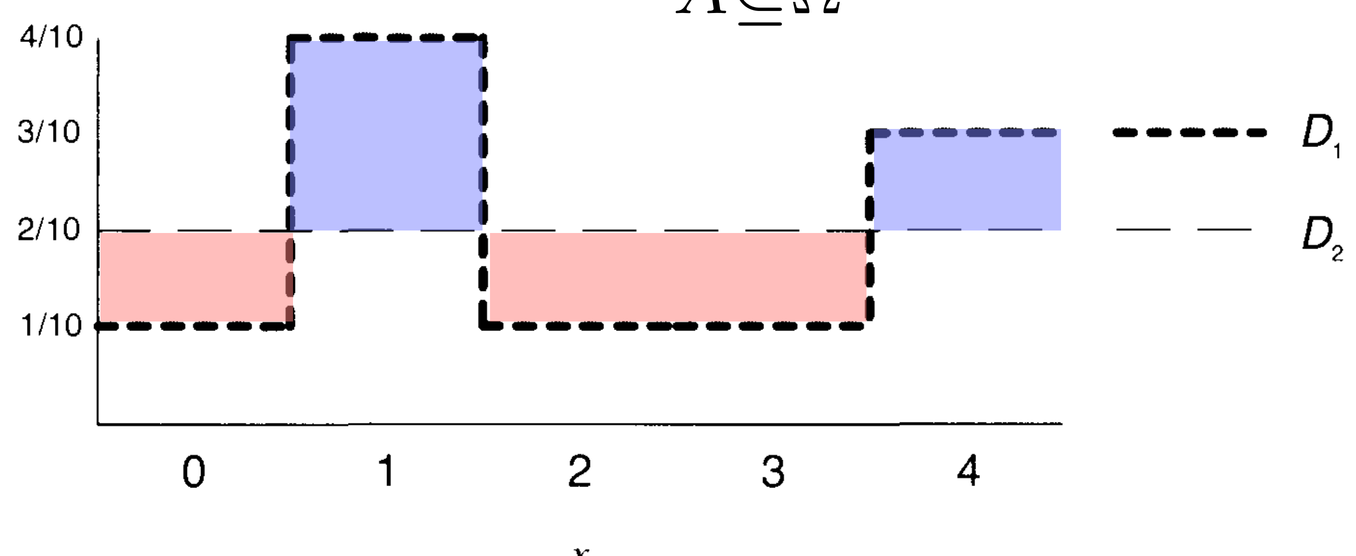
$$p, q \in [0,1]^{\Omega} \qquad \sum_{x\in\Omega} p(x) = 1 \quad \sum_{x\in\Omega} q(x) = 1$$

- total variation distance between p and q:

$$\|p - q\|_{TV} = \frac{1}{2}\|p-q\|_1 = \frac{1}{2}\sum_{x\in\Omega}|p(x)-q(x)|$$

- equivalent definition:

$$\|p - q\|_{TV} = \max_{A\subseteq\Omega} |p(A) - q(A)|$$



Mixing Time

Markov chain: $\mathfrak{M} = (\Omega, P)$

stationary distribution: π

$p_x^{(t)}$: distribution at time t when initial state is x

$$\Delta_x(t) = \|p_x^{(t)} - \pi\|_{TV} \qquad \Delta(t) = \max_{x\in\Omega} \Delta_x(t)$$

$$\tau_x(\epsilon) = \min\{t \mid \Delta_x(t) \le \epsilon\} \qquad \tau(\epsilon) = \max_{x\in\Omega} \tau_x(\epsilon)$$

- mixing time: $\tau_{\text{mix}} = \tau(1/2\mathrm{e})$

  rapid mixing: $\tau_{\text{mix}} = (\log|\Omega|)^{O(1)}$

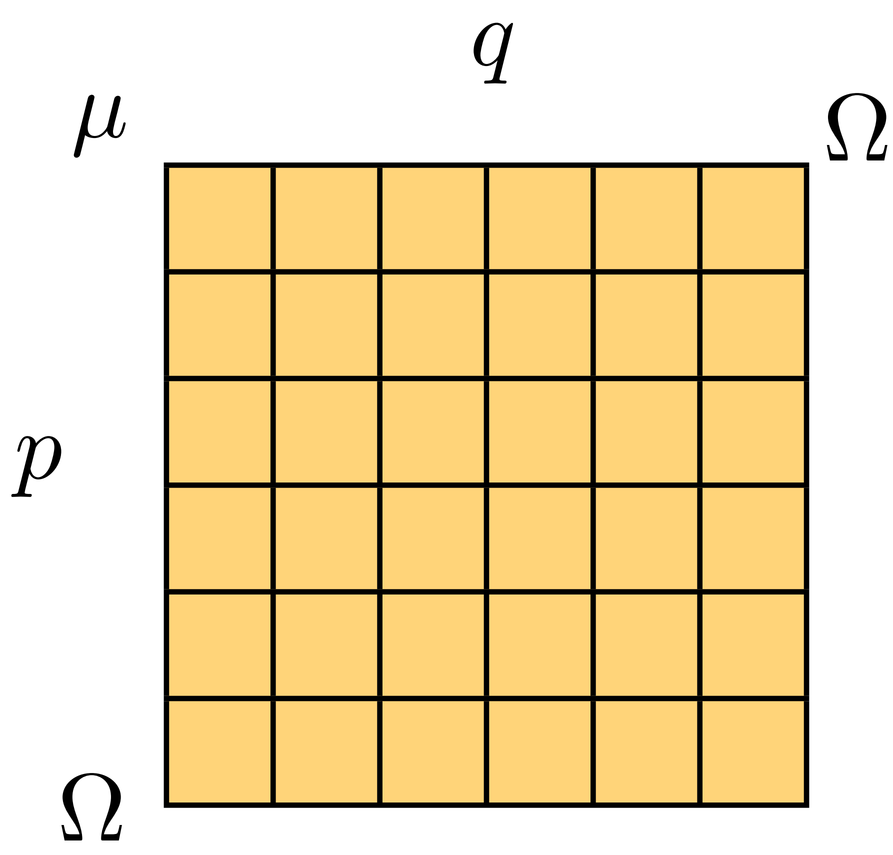
$$\Delta(k\cdot\tau_{\text{mix}}) \le \mathrm{e}^{-k} \quad \text{and} \quad \tau(\epsilon) \le \tau_{\text{mix}} \cdot \left\lceil \ln\frac{1}{\epsilon} \right\rceil$$

Coupling

p,q : distributions over Ω

a distribution μ over $\Omega \times \Omega$ is a coupling of p,q

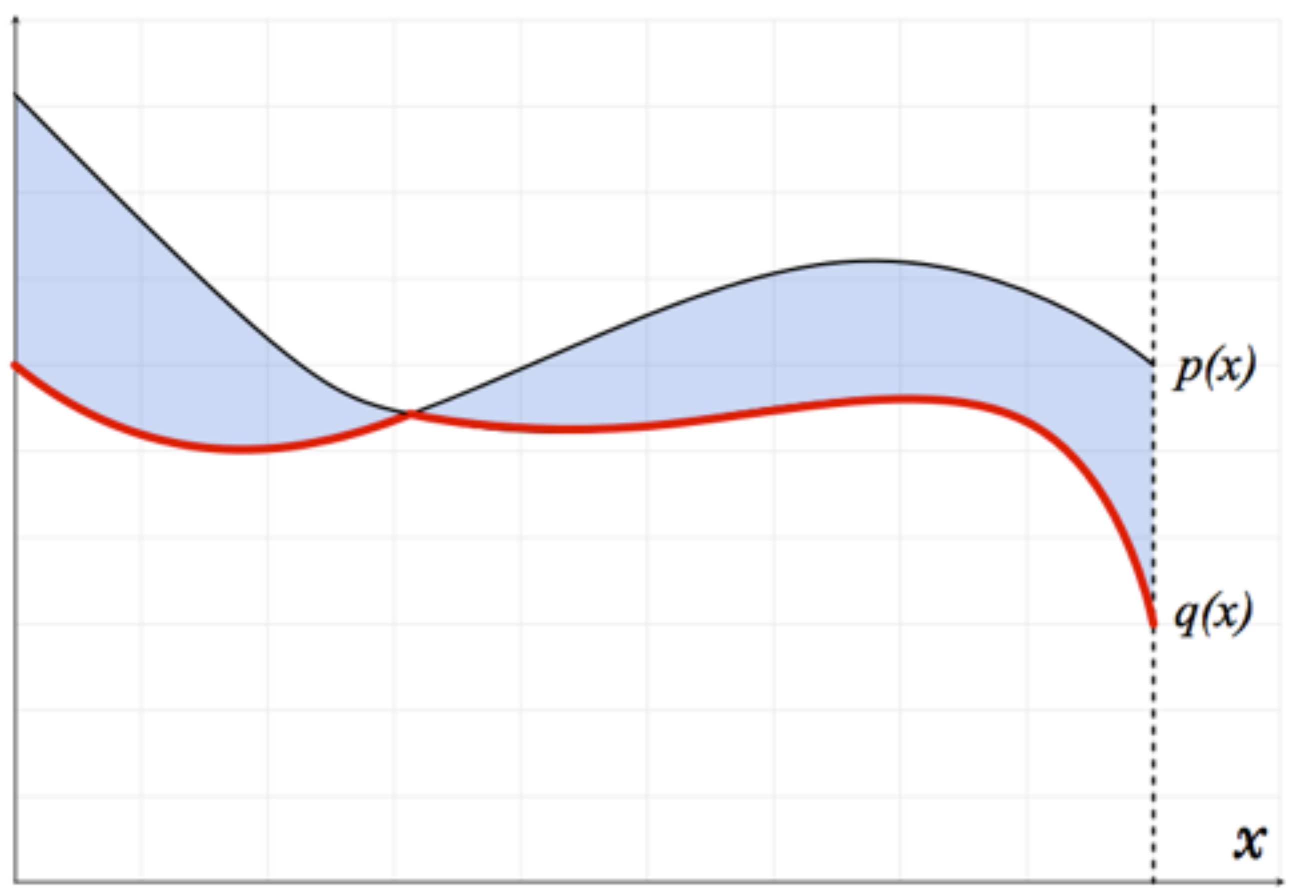
if $\quad p(x) = \sum_{y \in \Omega} \mu(x,y) \qquad q(x) = \sum_{y \in \Omega} \mu(y,x)$



Coupling Lemma

Coupling Lemma

1. (X,Y) is a coupling of p,q ⇨ $\Pr[X \neq Y] \geq \|p-q\|_{TV}$
2. $\exists$ a coupling (X,Y) of p,q s.t. $\Pr[X \neq Y] = \|p-q\|_{TV}$



Coupling of Markov Chains

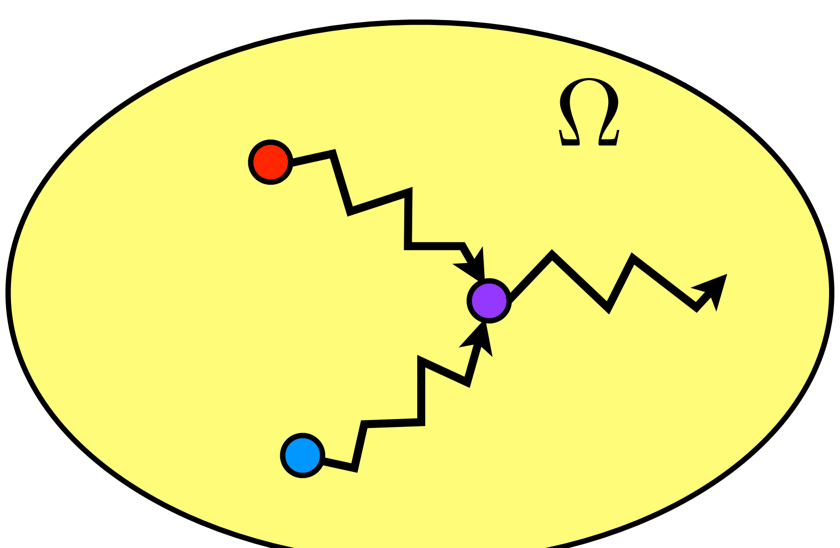
a coupling of $\mathfrak{M} = (\Omega, P)$ is a Markov chain (X_t, Y_t) of state space $\Omega \times \Omega$ such that:

- both are faithful copies of the chain

$$\Pr[X_{t+1} = y \mid X_t = x] = \Pr[Y_{t+1} = y \mid Y_t = x] = P(x, y)$$

- once collides, always makes identical moves

$$X_t = Y_t \Rightarrow X_{t+1} = Y_{t+1}$$



Markov Chain Coupling Lemma

Markov chain: $\mathfrak{M} = (\Omega, P)$

stationary distribution: π

$p_x^{(t)}$: distribution at time t when initial state is x

$$\Delta_x(t) = \|p_x^{(t)} - \pi\|_{TV} \qquad \Delta(t) = \max_{x \in \Omega} \Delta_x(t)$$

Markov Chain Coupling Lemma:

(X_t, Y_t) is a coupling of $\mathfrak{M} = (\Omega, P)$ ⇒

$$\Delta(t) \le \max_{x,y \in \Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y]$$

$p_x^{(t)}$: distribution at time t when initial state is x

Markov Chain Coupling Lemma:

(X_t, Y_t) is a coupling of $\mathfrak{M} = (\Omega, P)$ ⇒

$$\Delta(t) \le \max_{x,y\in\Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y]$$

$$\Delta(t) = \max_{x\in\Omega} \|p_x^{(t)} - \pi\|_{TV}$$

$$\le \max_{x,y\in\Omega} \|p_x^{(t)} - p_y^{(t)}\|_{TV}$$

$$\le \max_{x,y\in\Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y]$$

(coupling lemma)

$\mathfrak{M} = (\Omega, P)$ stationary distribution: π

$p_x^{(t)}$: distribution at time t when initial state is x

$$\Delta_x(t) = \|p_x^{(t)} - \pi\|_{TV} \qquad \Delta(t) = \max_{x \in \Omega} \Delta_x(t)$$

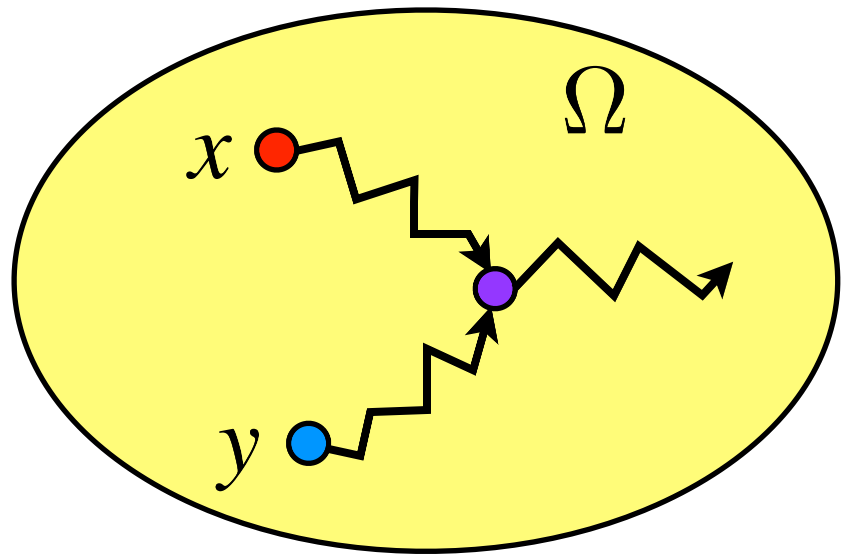
$$\tau_x(\epsilon) = \min\{t \mid \Delta_x(t) \leq \epsilon\} \qquad \tau(\epsilon) = \max_{x \in \Omega} \tau_x(\epsilon)$$

Markov Chain Coupling Lemma:

(X_t, Y_t) is a coupling of $\mathfrak{M} = (\Omega, P)$ $\Rightarrow$

$$\Delta(t) \leq \max_{x,y \in \Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y]$$

$$\max_{x,y \in \Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y] \leq \epsilon \quad \Rightarrow \quad \tau(\epsilon) \leq t$$



Markov Chain Coupling Lemma:

(X_t, Y_t) is a coupling of $\mathfrak{M} = (\Omega, P)$ ⇨

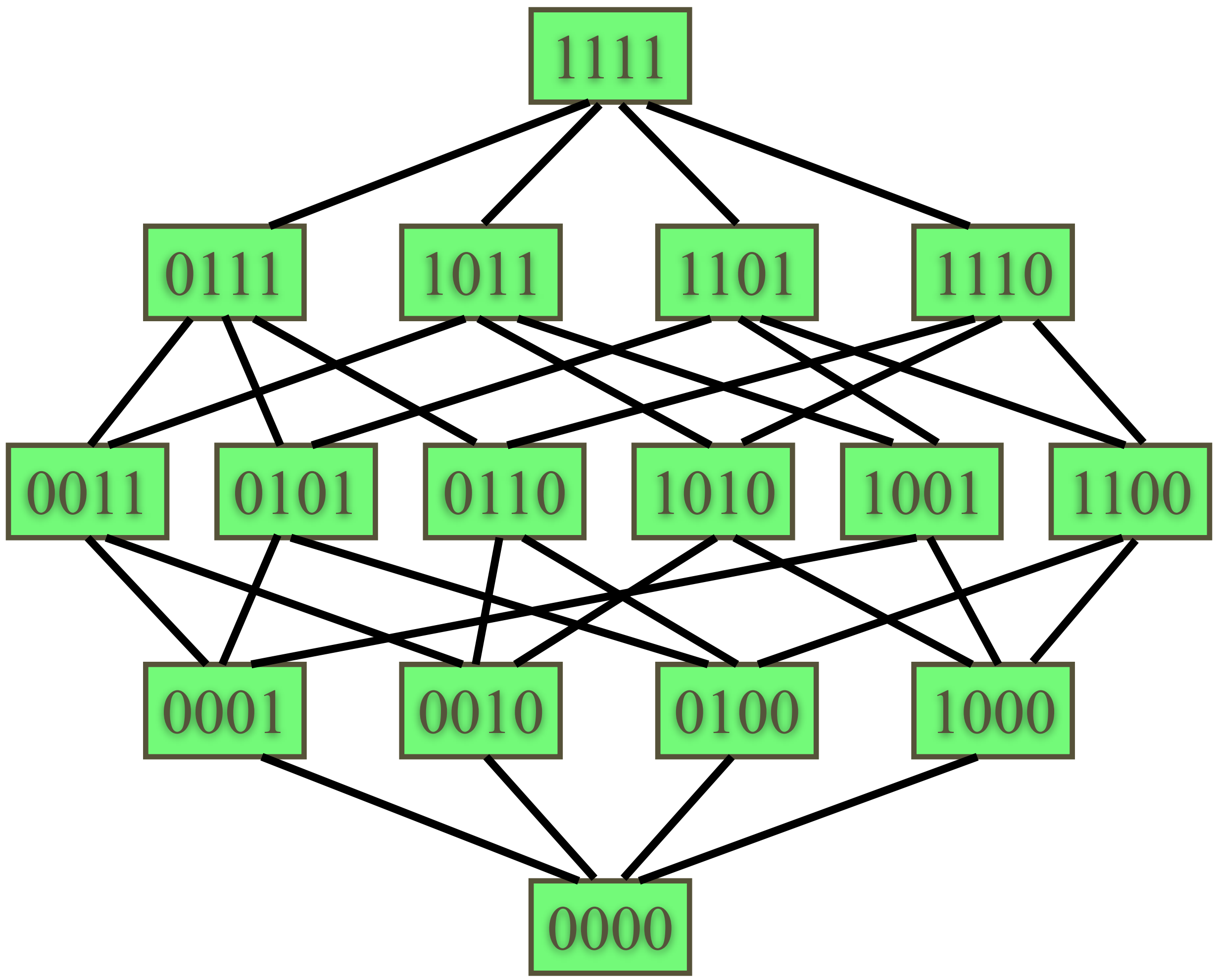
$$\Delta(t) \le \max_{x,y\in\Omega} \Pr[X_t \ne Y_t \mid X_0 = x, Y_0 = y]$$

$$\max_{x,y\in\Omega} \Pr[X_t \ne Y_t \mid X_0 = x, Y_0 = y] \le \epsilon \quad ⇨ \quad \tau(\epsilon) \le t$$

Random Walk on Hypercube

n-dimensional hypercube

$$\Omega = \{0,1\}^n$$



lazy random walk:

current state $x \in \{0,1\}^n$

- with prob. $1/2$ do nothing;
- pick a uniform random $i \in \{1,\ldots,n\}$ and flip x_i;

aperiodic;
irreducible;
uniform stationary distribution;

Random Walk on Hypercube

n-dimensional hypercube $\Omega = \{0,1\}^n$

current state $x \in \{0,1\}^n$

- with prob. $1/2$ do nothing;
- pick a uniform random $i \in \{1,\ldots,n\}$ and flip x_i;

equivalent to:

current state $x \in \{0,1\}^n$

- pick a uniform random $i \in \{1,\ldots,n\}$ and a uniform random bit $b \in \{0,1\}$;
- let $x_i = b$;

n-dimensional hypercube $\Omega = \{0,1\}^n$

current state $x \in \{0,1\}^n$

- pick a uniform random $i \in \{1,\dots,n\}$ and a uniform random bit $b \in \{0,1\}$;
- let $x_i = b$;

coupling rule: $(X_t, Y_t) \in \Omega \times \Omega$

each step, choose the same i and b

coupled if all indices in $\{1,\dots,n\}$ have been fixed

Markov Chain coupling lemma:

$$\Delta(t) \le \max_{x,y\in\Omega} \Pr[X_t \ne Y_t \mid X_0 = x, Y_0 = y]$$

$$\le \Pr[n \text{ coupons are not collected in } t \text{ trials}]$$

n-dimensional hypercube $\Omega = \{0,1\}^n$

current state $x \in \{0,1\}^n$

- pick a uniform random $i \in \{1,\ldots,n\}$ and a uniform random bit $b \in \{0,1\}$;
- let $x_i = b$;

$$\Delta(t) \le \Pr[n \text{ coupons are not collected in } t \text{ trials}]$$

$$\le \mathrm{e}^{-c} \qquad \text{for } t = n \ln n + cn$$

$$\Delta(n \ln n + cn) \le \mathrm{e}^{-c}$$

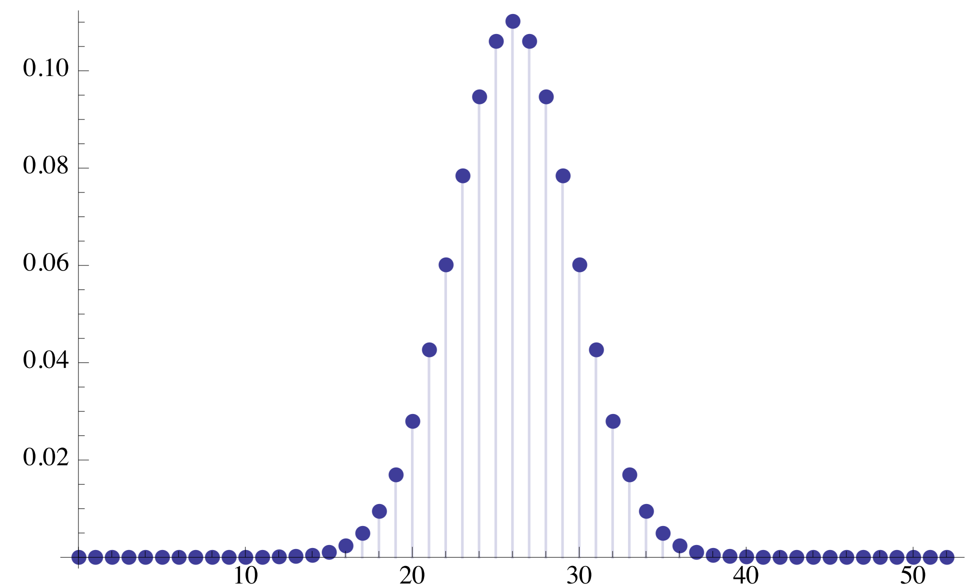
$$\Rightarrow \quad \tau(\epsilon) \le n \ln n + n \ln \frac{1}{\epsilon}$$

Card Shuffling

Riffle Shuffle
(Gilbert-Shannon-Reeds)

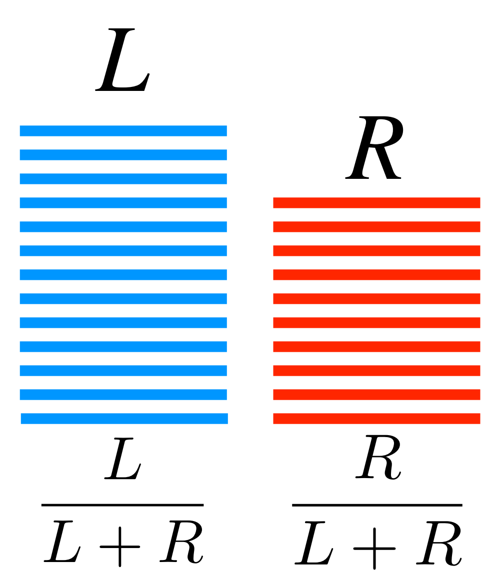
1. split (cut) : n cards

binomial distribution $\mathrm{Bin}(n,1/2)$



2. merge (interleaving):

drops cards in sequence,
proportional to the current weights

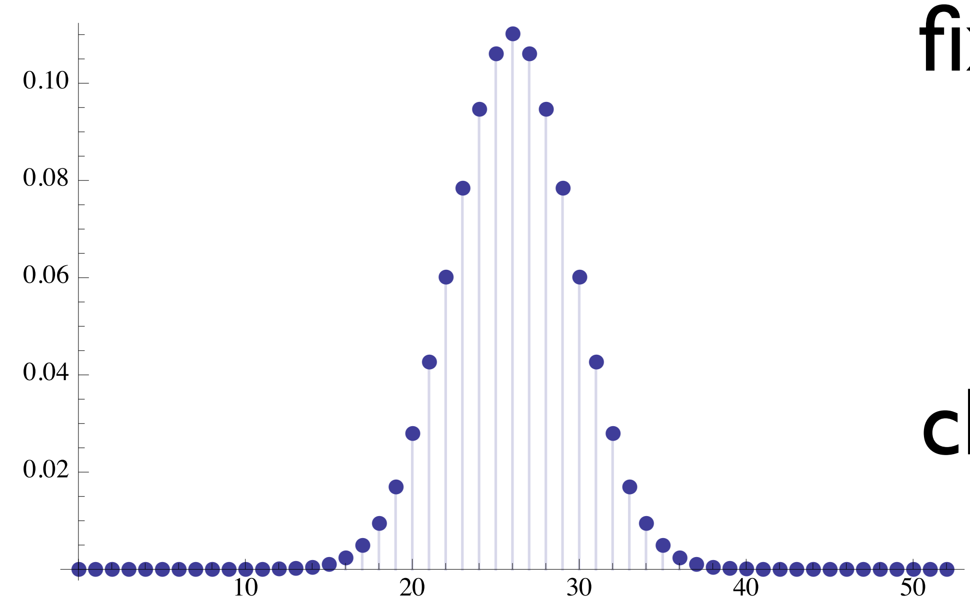


Card Shuffling

Riffle Shuffle
(Gilbert-Shannon-Reeds)

1. split (cut) : n cards

binomial distribution $\mathrm{Bin}(n,1/2)$



fix k:

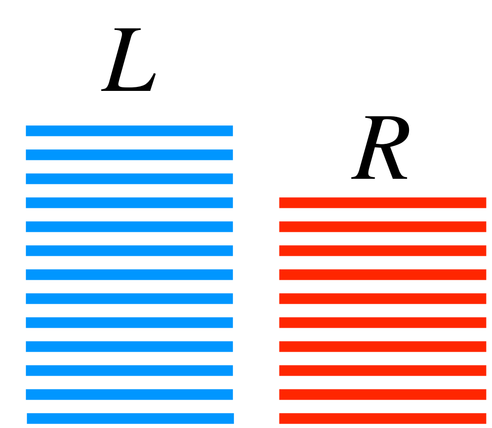
$$\frac{\binom{n}{k}}{2^n}$$

choices

2. merge (interleaving):

uniform random interleaving

fix a cut: $\frac{1}{\binom{n}{k}}$ choices



$$\frac{L}{L+R} \qquad \frac{R}{L+R}$$

any (cut-interleaving) pair: 2^{-n} prob.

Inverse Riffle Shuffle

Inverse **Riffle Shuffle:**

- label each card with a bit from $\{0,1\}$ uniformly and independently at random;
- move all 0 cards above all 1 cards, respecting the relative order within.

inverse of the Riffle shuffle

same uniform stationary distribution and same mixing time

Inverse Riffle Shuffle:

- label each card with a bit from $\{0,1\}$ uniformly and independently at random;
- move all 0 cards above all 1 cards, respecting the relative order within.

coupling rule:

in each round, choose the same random bit for every card

coupling rule:

in each round, choose the same random bit for every card

Lemma

After each round, the cards are sorted according to the binary codes.

⇒ coupled if all cards have distinct labels

coupling rule:

in each round, choose the same random bit for every card

⇒ coupled if all cards have distinct labels

Markov Chain coupling lemma:

$$\Delta(t) \leq \max_{x,y\in\Omega} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y]$$

$$\leq \Pr_{f:[n]\to\{0,1\}^t}[|f([n])| < n] \quad = 1/2\mathrm{e}$$

birthday

⇒ $2^t = O(n^2)$ $\quad\quad \tau_{\mathrm{mix}} \leq 2\log_2 n + O(1)$

coupling rule:

in each round, choose the same random bit for every card

$$\tau_{\text{mix}} \le 2\log_2 n + O(1)$$

state space Ω: all permutations of n cards

$$|\Omega| = n! \qquad \log|\Omega| = \Theta(n\log n)$$

n=52

t	1	2	3	4	5	6	7	8	9
$\Delta(t)$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.003