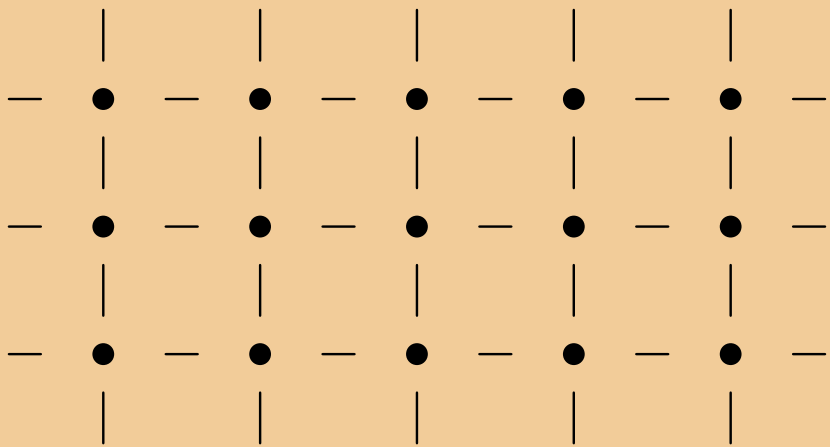
Markov Chain Monte Carlo and mixing rates

Yevgeniy Kovchegov
Department of Mathematics
Oregon State University

Ising Model. Every vertex v of $G = (V, E)$ is assigned a spin $\sigma(v) \in \{-1, +1\}$. The probability of a configuration $\sigma \in \{-1, +1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$



Ising Model. Every vertex v of $G = (V, E)$ is assigned a spin $\sigma(v) \in \{-1, +1\}$. The probability of a configuration $\sigma \in \{-1, +1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta \mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$

$$\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & {\color{red}+1} & - & {\color{blue}-1} & - & {\color{blue}-1} & - & {\color{blue}-1} & - & {\color{red}+1} & - \\
 & | & & | & & | & & | & & | & \\
- & {\color{red}+1} & - & {\color{blue}-1} & - & {\color{red}+1} & - & {\color{red}+1} & - & {\color{blue}-1} & - \\
 & | & & | & & | & & | & & | & \\
- & {\color{red}+1} & - & {\color{blue}-1} & - & {\color{blue}-1} & - & {\color{red}+1} & - & {\color{blue}-1} & - \\
 & | & & | & & | & & | & & | &
\end{array}$$

Ising Model. $\forall \sigma \in \{-1,+1\}^V$, the Hamiltonian

$$\mathcal{H}(\sigma) = -\frac{1}{2}\sum_{u,v:\ u\sim v} \sigma(u)\sigma(v) = \ -\sum_{edges\ \ e=[u,v]} \sigma(u)\sigma(v)$$

and probability of a configuration $\sigma \in \{-1,+1\}^V$ is

$$\pi(\sigma) = \frac{e^{-\beta\mathcal{H}(\sigma)}}{Z(\beta)}, \quad \text{where} \quad \beta = \frac{1}{T}$$

$Z(\beta) = \sum_{\sigma\in\{-1,+1\}^V} e^{-\beta\mathcal{H}(\sigma)}$ - normalizing factor.

Ising Model: local Hamiltonian

$$\mathcal{H}(\sigma) = -\frac{1}{2} \sum_{u,v:\ u \sim v} \sigma(u)\sigma(v) = -\sum_{edges\ e=[u,v]} \sigma(u)\sigma(v)$$

The local Hamiltonian

$$\mathcal{H}_{local}(\sigma, v) = -\sum_{u:\ u \sim v} \sigma(u)\sigma(v)\ .$$

Observe: conditional probability for $\sigma(v)$ is given by $\mathcal{H}_{local}(\sigma, v)$:

$$\mathcal{H}(\sigma) = \mathcal{H}_{local}(\sigma, v) - \sum_{e=[u_1,u_2]:\ u_1,u_2 \neq v} \sigma(u_1)\sigma(u_2)$$

Ising Model via Glauber dynamics.

$$
\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & -1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & \sigma(v) & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}
$$

Observe: conditional probability for $\sigma(v)$ is given by $\mathcal{H}_{local}(\sigma, v)$:

$$\mathcal{H}(\sigma) = \mathcal{H}_{local}(\sigma, v) - \sum_{e=[u_1,u_2]:\ u_1,u_2 \neq v} \sigma(u_1)\sigma(u_2)$$

Ising Model via Glauber dynamics.

$$
\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & -1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & ? & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & -1 & - & -1 & - & +1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}
$$

Randomly pick $v \in G$, erase the spin $\sigma(v)$.
Choose σ_+ or σ_-:

$$\mathbf{Prob}(\sigma \to \sigma_+) = \frac{e^{-\beta\mathcal{H}(\sigma_+)}}{e^{-\beta\mathcal{H}(\sigma_-)} + e^{-\beta\mathcal{H}(\sigma_+)}}$$

$$= \frac{e^{-\beta\mathcal{H}_{local}(\sigma_+,v)}}{e^{-\beta\mathcal{H}_{local}(\sigma_-,v)} + e^{-\beta\mathcal{H}_{local}(\sigma_+,v)}} = \frac{e^{-2\beta}}{e^{-2\beta} + e^{2\beta}} \; .$$

Glauber dynamics: Rapid mixing.

Glauber dynamics - a random walk on state space S (here $\{-1,+1\}^V$) s.t. needed π is stationary w.r.t. Glauber dynamics.

In high temperatures (i.e. $\beta = \frac{1}{T}$ small enough) it takes $O(n \log n)$ iterations to get "ε-close" to π. Here $|V| = n$.

Need: $\boxed{\max_{v \in V} deg(v) \cdot \tanh(\beta) < 1}$

Thus the Glauber dynamics is a fast way to generate π. It is an important example of **Gibbs sampling**.

Close enough distribution and mixing time.

What is "ε-close" to π? Start with σ_0:

$$
\begin{array}{ccccccccccc}
 & | & & | & & | & & | & & | & \\
- & +1 & - & +1 & - & +1 & - & +1 & - & +1 & - \\
 & | & & | & & | & & | & & | & \\
- & +1 & - & +1 & - & +1 & - & -1 & - & -1 & - \\
 & | & & | & & | & & | & & | & \\
- & -1 & - & -1 & - & -1 & - & -1 & - & -1 & - \\
 & | & & | & & | & & | & & | &
\end{array}
$$

If $P_t(\sigma)$ is the probability distribution after t iterations, the total variation distance

$$\|P_t - \pi\|_{TV} = \frac{1}{2} \sum_{\sigma \in \{-1,+1\}^V} |P_t(\sigma) - \pi(\sigma)| \leq \varepsilon \ .$$

Who researched mixing times? D.Aldous, P.Diaconis, J.A.Fill, M.Jerrum, A.Sinclair and many more names.

Close enough distribution and mixing time.

Total variation distance:

$$\|\mu-\nu\|_{TV} := \frac{1}{2}\sum_{x\in S} |\mu(x)-\nu(x)| = \sup_{A\subset S} |\mu(A)-\nu(A)|$$

Mixing time:

$$t_{mix}(\varepsilon) := \inf\{t \;:\; \|P_t - \pi\|_{TV} \leq \varepsilon, \quad \text{all } \sigma_0\} \;.$$

In high temperature, $t_{mix}(\varepsilon) = O(n \log n)$.

Coupling Method.

S - sample space

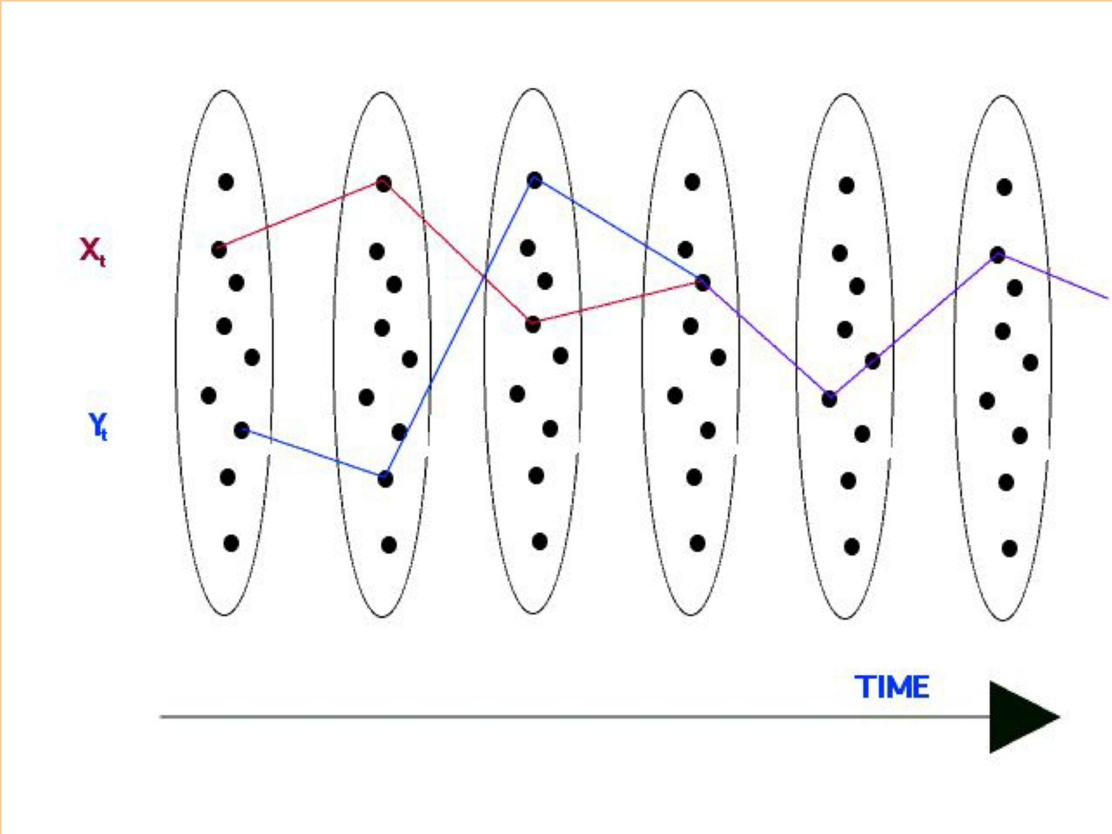
$\{p(i,j)\}_{i,j\in S}$ - transition probabilities

Construct process $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ on $S \times S$ such that

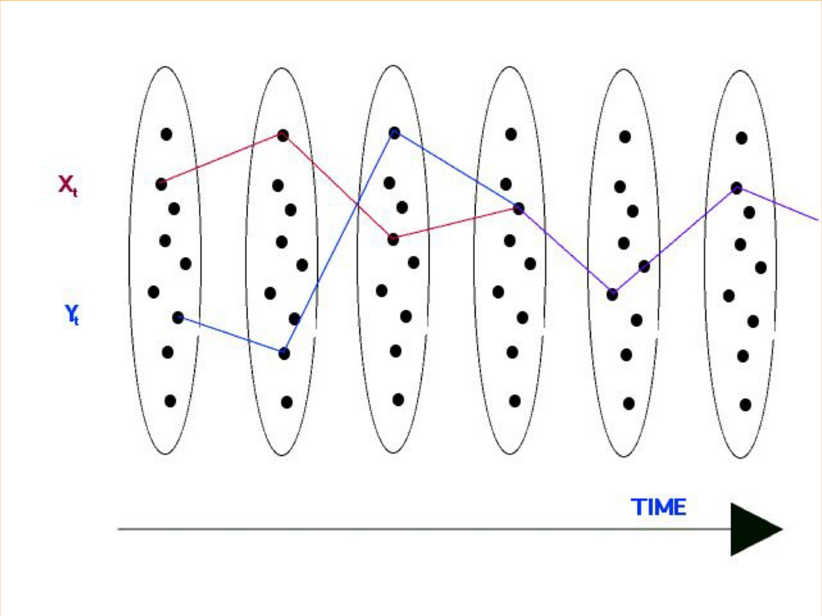
X_t is a $\{p(i,j)\}$-Markov chain
Y_t is a $\{p(i,j)\}$-Markov chain

Once $X_t{=}Y_t$, let $X_{t+1}{=}Y_{t+1}$, $X_{t+2}{=}Y_{t+2}$,...

Coupling Method.



Coupling Method.



Coupling time: $T_{coupling} = \min\{t : \ X_t = Y_t\}$

Successful coupling: $\mathbf{Prob}(T_{coupling} < \infty) = 1$

Mixing times via coupling.

Let $T_{i,j}$ be coupling time for $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ given $X_0 = i$ and $Y_0 = j$. Then

$$\|P_{X_t} - P_{Y_t}\|_{TV} \leq P[T_{i,j} > t] \leq \frac{E[T_{i,j}]}{t}$$

Now, if we let $Y_0 \sim \pi$, then for any $X_0 \in S$,

$$\|P_{X_t} - \pi\|_{TV} = \|P_{X_t} - P_{Y_t}\|_{TV} \leq \frac{\max_{i,j\in S} E[T_{i,j}]}{t} \leq \varepsilon$$

whenever $t \geq \frac{\max_{i,j\in S} E[T_{i,j}]}{\varepsilon}$.

Mixing times via coupling.

$\|P_{X_t} - \pi\|_{TV} \leq \varepsilon$ whenever $t \geq \frac{\max_{i,j \in S} E[T_{i,j}]}{\varepsilon}$.

Thus

$$t_{mix}(\varepsilon) = \inf \left\{ t : \ \|P_{X_t} - \pi\|_{TV} \leq \varepsilon \right\} \leq \frac{\max_{i,j \in S} E[T_{i,j}]}{\varepsilon}.$$

So,

$$O(t_{mix}) \leq O(T_{coupling}) \ .$$

Thus constructing a coupled process that minimizes $E[T_{coupling}]$ gives an effective upper bound on mixing time.

Coupon collector.

n types of coupons: $\boxed{1}$, $\boxed{2}$,. . . ,$\boxed{n}$

Collecting coupons: coupon / unit of time, each coupon type is equally likely.

Goal: To collect a coupon of each type.

Question: How much time will it take?

Coupon collector.

$$\begin{array}{cccccccccc} \boxed{\text{c1}} & \boxed{\text{c2}} & \boxed{\text{c1}} & \boxed{\text{c3}} & \boxed{\text{c3}} & \boxed{\text{c2}} & \boxed{\text{c4}} & \boxed{\text{c5}} & \ldots \\ \uparrow & \uparrow & & \uparrow & & & \uparrow & \uparrow & \\ \tau_1 = 1 & \tau_2 & & \tau_3 & & & \tau_4 & \tau_5 & \end{array}$$

Here $\tau_1 = 1$, $E[\tau_2 - \tau_1] = \frac{n}{n-1}$,
$E[\tau_3 - \tau_2] = \frac{n}{n-2}, \ldots, E[\tau_n - \tau_{n-1}] = n$.

Hence

$$E[\tau_n] = n\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) = n \log n + O(n)$$

Coupon collector and random card-to-random location shuffling.

Shuffling a deck of n different cards:

$$\boxed{1}\quad\boxed{2}\quad\boxed{3}\quad\boxed{4}\quad\boxed{5}\quad\boxed{6}\quad\boxed{7}\quad\boxed{8}$$

Coupon collector and random card-to-random location shuffling.

Shuffling a deck of n different cards:

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$ $\boxed{6}$ $\boxed{7}$ $\boxed{8}$

Pick a card at random:

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{\mathbf{5}}$ $\boxed{6}$ $\boxed{7}$ $\boxed{8}$

Coupon collector and random card-to-random location shuffling.

Shuffling a deck of n different cards:

$$\boxed{1}\quad\boxed{2}\quad\boxed{3}\quad\boxed{4}\quad\boxed{5}\quad\boxed{6}\quad\boxed{7}\quad\boxed{8}$$

Pick a card at random:

$$\boxed{1}\quad\boxed{2}\quad\boxed{3}\quad\boxed{4}\quad\boxed{\mathbf{5}}\quad\boxed{6}\quad\boxed{7}\quad\boxed{8}$$

Pool it out, and place it anywhere in the deck:

$$\boxed{1}\quad\boxed{2}\quad\boxed{\mathbf{5}}\quad\boxed{3}\quad\boxed{4}\quad\boxed{6}\quad\boxed{7}\quad\boxed{8}$$

Iterate. **Question:** $t_{mix}(\varepsilon) =?$

Random card-to-random location:

Cover time: T_{cover} - each card was selected at least once.

$$\boxed{2}\ \boxed{\mathbf{5}}\ \boxed{3}\ \boxed{\mathbf{1}}\ \boxed{6}\ \boxed{8}\ \boxed{\mathbf{7}}\ \boxed{\mathbf{4}}$$

Coupon collector $\Rightarrow E[T_{cover}] = n \log n + O(n)$

Cover time as well as coupling time (both strong stationary time) provides an effective upper and lower bound on mixing time.

Here $\quad E[t_{mix}(\varepsilon)] = n \log n + O(n)$

Shuffling by random transpositions.

Pick two cards at random:

$$\boxed{1}\quad\boxed{2}\quad\boxed{3}\quad\boxed{4}\quad\boxed{\mathbf{5}}\quad\boxed{6}\quad\boxed{\mathbf{7}}\quad\boxed{8}$$

Transpose them:

$$\boxed{1}\quad\boxed{2}\quad\boxed{3}\quad\boxed{4}\quad\boxed{\mathbf{7}}\quad\boxed{6}\quad\boxed{\mathbf{5}}\quad\boxed{8}$$

Iterate:

$$\boxed{\mathbf{3}}\quad\boxed{2}\quad\boxed{\mathbf{1}}\quad\boxed{4}\quad\boxed{\mathbf{7}}\quad\boxed{6}\quad\boxed{\mathbf{5}}\quad\boxed{8}$$

$$\boxed{\mathbf{3}}\quad\boxed{2}\quad\boxed{\mathbf{6}}\quad\boxed{4}\quad\boxed{\mathbf{7}}\quad\boxed{\mathbf{1}}\quad\boxed{\mathbf{5}}\quad\boxed{8}$$

. . . etc.

Shuffling by random transpositions.

$\boxed{\mathbf{3}}$ $\boxed{2}$ $\boxed{\mathbf{6}}$ $\boxed{4}$ $\boxed{\mathbf{7}}$ $\boxed{\mathbf{1}}$ $\boxed{\mathbf{5}}$ $\boxed{8}$

Here coupon collector gives only a lower bound of $\frac{1}{2}n \log n$.

Goal: get $O(n \log(n))$ upper bound with coupling method - hidden coupon collector.

Obtaining $O(n \log(n))$ via super-fast coupling. **(Joint work with R.Burton)**

Diaconis and Shahshahani (early 80's): The mixing time for shuffling a deck of n cards by random transpositions is of order $O(n \log(n))$ with cut-off asymptotics at $\frac{1}{2} n \log(n)$.

Method used: **representation theory.**

We answer an **open problem** (Y. Peres): Provide a coupling proof of $O(n \log(n))$ mixing rate.

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Even in case of **two discrepancies** ($d = 2$) at d_1 and d_2:

$$
\begin{array}{llcccccccccl}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a}} & \boxed{8} & \boxed{7} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{7} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & & & \\
 & & & & d_2 & & d_1 & & & &
\end{array}
$$

Label-to-location coupling:

$$E[T_{coupling}] = \frac{n^2}{4} \quad \text{– too large.}$$

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Case of **two discrepancies** $(d = 2)$:

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{3}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{7} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{3}} & \boxed{8} & \boxed{7} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & & & \\
 & & & & d_2 & & d_1 & & & &
\end{array}
$$

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Case of **two discrepancies** ($d = 2$):

$$
\begin{array}{llcccccccccc}
A_t: & \dots & \boxed{\mathbf{4}} & \boxed{6} & \boxed{\mathbf{3}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{7} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{\mathbf{4}} & \boxed{6} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{3}} & \boxed{8} & \boxed{7} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

$$
\begin{array}{llcccccccccc}
A_t: & \dots & \boxed{7} & \boxed{6} & \boxed{\mathbf{3}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{7} & \boxed{6} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{3}} & \boxed{8} & \boxed{4} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & & & \\
 & & & & d_2 & & d_1 & & & &
\end{array}
$$

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Case of **two discrepancies** ($d = 2$):

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{7} & \boxed{6} & \boxed{\mathbf{3}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{7} & \boxed{6} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{3}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\
 & & & \uparrow & \uparrow & & \uparrow & & & & \\
 & & & i_2 & d_2 & & d_1 & & & &
\end{array}
$$

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{6}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{6}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & & & \\
 & & & & d_2 & & d_1 & & & &
\end{array}
$$

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Case of **two discrepancies** ($d = 2$):

$$\begin{array}{lccccccccccc} A_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{6}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\ B_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{5}} & \boxed{9} & \boxed{\mathbf{6}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\ & & & & & & \uparrow & & & & \\ & & & & & & i_3 & & & & \end{array}$$

$$\begin{array}{lccccccccccc} A_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{6}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \\ B_t: & \ldots & \boxed{7} & \boxed{3} & \boxed{\mathbf{6}} & \boxed{9} & \boxed{\mathbf{5}} & \boxed{8} & \boxed{4} & \boxed{2} & \ldots \end{array}$$

Label-to-location coupling: $E[T_{coupling}] = \frac{n^2}{4}$

A coupling. (Aldous and Fill) $\ll \boxed{\mathbf{a}}, i \gg$: moves card $\boxed{\mathbf{a}}$ to location i in both processes, A_t and B_t.

Mixing: order $O(n^2)$ instead of $O(n \log n)$;

$$E[T_{coupling}] \approx \sum_{d=2}^{n} \frac{n^2}{d^2} \approx \left(\frac{\pi^2}{6} - 1\right) n^2$$

Problem: slows down significantly when the number of discrepancies is small enough.

Obtaining $O(n \log(n))$ via super-fast coupling. (Joint work with R.Burton)

Tunneling into the future approach.

Coupling time:

$$E[T_{coupling}] \leq \left[\frac{1}{2\varepsilon} + \frac{\kappa}{(1-\kappa)(\kappa-\varepsilon)}\right] \cdot n \log n$$

$$\text{for any} \quad 0 < \varepsilon < \kappa < 1.$$

Two discrepancies ($d = 2$) at d_1 and d_2:

$$
\begin{array}{lccccccccccc}
A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

Label-to-location coupling:

$$E[T_{coupling}] = \frac{n^2}{4} \quad \text{– too large.}$$

Jump $\ll \boxed{\mathbf{a}}, i_1 \gg$ of $\boxed{\mathbf{a}}$ to random location i_1 at exponential time t_1:

From

$$\begin{array}{lccccccccccc}
A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}$$

to

$$\begin{array}{lccccccccccc}
A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}$$

Different way of saying the same:

Start with

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2/i_1 & & i_1/d_1 & & d_1/d_2 & &
\end{array}
$$

where at time t_1 the locations relabel according to

$$
\boxed{\begin{array}{l}
d_1/d_2 \longrightarrow i_1 \\
i_1/d_1 \longrightarrow d_1 \\
d_2/i_1 \longrightarrow d_2
\end{array}}.
$$

Jump $\ll \boxed{\textbf{a}}, i_1 \gg$ at time $t_1 \sim \textbf{exponential}\left(\frac{1}{n}\right)$.

From

$$
\begin{array}{lcccccccccc}
A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{b}} & \boxed{9} & \boxed{\textbf{a1}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2/i_1 & & i_1/d_1 & & d_1/d_2 & &
\end{array}
$$

to

$$
\begin{array}{lcccccccccc}
A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{b}} & \boxed{9} & \boxed{\textbf{a1}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\
B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

The following **association map** will determine jumps of $\boxed{\mathbf{a1}}$.

$$
\begin{array}{lccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots
\end{array}
$$

Card $\boxed{\mathbf{a1}}$ will jump to position i_2 on the assoc. map at time t_2, even if $t_2 < t_1$.

$$
\begin{array}{lcccccccccc}
A_t : & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t : & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2^* & & d_1^* & & i_1^* & &
\end{array}
$$

Now $i_2 \neq i_1^*$ and

$$t_2 \sim \textbf{exponential}\left((1 - 1/n) \cdot \frac{1}{n}\right)$$

$\ll \boxed{\mathbf{a1}}, i_1^* \gg = \ll \boxed{\mathbf{a1}}, \boxed{\mathbf{a}} \gg$ is label-to-label, we can skip.

If $i_1 = d_1$ or d_2, discrepancies cancel at t_1;

if $i_2^* = d_1^*$ or d_2^*, discrepancies cancel on the assoc. map at t_2.

If $t_1 < t_2$, $\boxed{\text{assoc. map} \longrightarrow \text{real picture}}$ at t_1, we create one more assoc. map.

Case $t_2 < t_1$, and $i_2^* = d_2^*$. **On association map:**

Start with

$$
\begin{array}{lcccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2/i_1 & & i_1/d_1 & & d_1/d_2 & &
\end{array}
$$

At time t_2:

$$
\begin{array}{lcccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2/i_1 & & i_1/d_1 & & d_1/d_2 & &
\end{array}
$$

At time t_2:

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2/i_1 & & i_1/d_1 & & d_1/d_2 & &
\end{array}
$$

At time t_1:

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

Case $t_2 < t_1$, and $i_2^* = d_2^*$. **Same evolution, original association:**

Start with

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

At time t_2:

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{a}} & \boxed{8} & \boxed{\mathbf{b}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{4} & \boxed{6} & \boxed{\mathbf{a}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a1}} & \boxed{2} & \ldots \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & d_2 & & d_1 & & i_1 & &
\end{array}
$$

At time t_2:

$$\begin{array}{lcccccccccc} A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{a}} & \boxed{8} & \boxed{\textbf{b}} & \boxed{2} & \dots \\ B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a1}} & \boxed{2} & \dots \\ & & & & \uparrow & & \uparrow & & \uparrow & & \\ & & & & d_2 & & d_1 & & i_1 & & \end{array}$$

At time t_1:

$$\begin{array}{lcccccccccc} A_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\ B_t: & \dots & \boxed{4} & \boxed{6} & \boxed{\textbf{a1}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \dots \\ & & & & \uparrow & & \uparrow & & \uparrow & & \\ & & & & d_2 & & d_1 & & i_1 & & \end{array}$$

Here

$$E[T_{coupling}] \approx \frac{n^2}{8}$$

Chain of association maps:

$\ll \boxed{\mathbf{a1}}, i_2 \gg$ occurs at t_2

$$
\begin{array}{llcccccccccc}
A_t: & \ldots & \boxed{\mathbf{a2}} & \boxed{6} & \boxed{\mathbf{b}} & \boxed{9} & \boxed{\mathbf{a1}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{\mathbf{a2}} & \boxed{6} & \boxed{\mathbf{a1}} & \boxed{9} & \boxed{\mathbf{b}} & \boxed{8} & \boxed{\mathbf{a}} & \boxed{2} & \ldots \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
 & & i_2 & & d_2^* & & d_1^* & & i_1^* & &
\end{array}
$$

d_1^* is i_1/d_1 before t_1, and d_1 after t_1;

d_2^* is d_2/i_1 before t_1, and d_2 after t_1;

i_1^* is d_1/d_2 before t_1, and i_1 after t_1.

New association map:

$$
\begin{array}{llcccccccccc}
A_t: & \ldots & \boxed{\textbf{a1}} & \boxed{6} & \boxed{\textbf{b}} & \boxed{9} & \boxed{\textbf{a2}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{\textbf{a1}} & \boxed{6} & \boxed{\textbf{a2}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
 & & d_1^*/d_2^* & & d_2^*/i_2 & & i_2/d_1^* & & i_1^* & &
\end{array}
$$

where at t_2,

$$
\boxed{\begin{array}{l}
d_1^*/d_2^* \longrightarrow i_2 \\
i_2/d_1^* \longrightarrow d_1^* \\
d_2^*/i_2 \longrightarrow d_2^*
\end{array}}\,.
$$

$\boxed{\textbf{a2}}$ will do label-to-location jump w.r.t. the following assoc. map

$$
\begin{array}{lccccccccccc}
A_t: & \ldots & \boxed{\textbf{a1}} & \boxed{6} & \boxed{\textbf{b}} & \boxed{9} & \boxed{\textbf{a2}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
B_t: & \ldots & \boxed{\textbf{a1}} & \boxed{6} & \boxed{\textbf{a2}} & \boxed{9} & \boxed{\textbf{b}} & \boxed{8} & \boxed{\textbf{a}} & \boxed{2} & \ldots \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\
 & & i_2^{**} & & d_2^{**} & & d_1^{**} & & i_1^{*} & &
\end{array}
$$

d_1^{**} is i_2/d_1^* before t_2, and d_1^* after t_2;
d_2^{**} is d_2^*/i_2 before t_2, and d_2^* after t_2;
i_2^{**} is d_1^*/d_2^* before t_2, and i_2 after t_2.

$\ll \boxed{\textbf{a2}}, i_3 \gg$ occurs at $t_3 \sim \textbf{exponential}\left((1-2/n)\cdot\frac{1}{n}\right)$

And so on, creating a **chain** of $k = \lfloor \varepsilon n \rfloor$ association maps.

In case of $d = 2$ discrepancies, the avg. time of discrepancy cancelation on one of assoc. maps is

$$E[T_2] = \frac{n^2}{4(k+1)} = \frac{n}{4\varepsilon}.$$

General d:

$$E[T_d] = \frac{n^2}{2(k+1)d} \approx \frac{n}{2\varepsilon d}.$$

Coupling time (all discrepancies):

$$E[T_{coupling}] \leq \left[\frac{1}{2\varepsilon} + \frac{\kappa}{(1-\kappa)(\kappa-\varepsilon)}\right] \cdot n \log n$$

$$\text{for any} \quad 0 < \varepsilon < \kappa < 1.$$