An extension of path coupling and its application to the Glauber dynamics for graph colourings

Leslie Ann Goldberg, Warwick

Joint work with

Martin Dyer, Leeds
Catherine Greenhill, Leeds
Mark Jerrum, Edinburgh
Michael Mitzenmacher, Harvard

Sampling via Markov-chain simulation

Markov chain $\mathcal{M}$
finite state space Ω
stationary distribution π

Total variation distance of $\mathcal{M}$ from π at t
(with initial state x)

$$\begin{aligned} d(P_x^t, \pi) &= \frac{1}{2} \sum_{y \in \Omega} |P_x^t(y) - \pi(y)| \\ &= \max_{A \subseteq \Omega} |P_x^t(A) - \pi(A)|. \end{aligned}$$

Mixing time

$$\tau_x(\epsilon) = \min_t\{t \mid d(P_x^{t'}, \pi) \le \epsilon \text{ for all } t' \ge t\}.$$

$$\tau(\epsilon) = \max_x\{\tau_x(\epsilon)\}.$$

Proving Rapid Mixing
Coupling Method

Coupling for $\mathcal{M}$
Markov process (X_t, Y_t) on $\Omega \times \Omega$

- X_t is a faithful copy of $\mathcal{M}$

- Y_t is a faithful copy of $\mathcal{M}$

Coupling Lemma Suppose that Y_0 has the stationary distribution π. Then, if X_t has distribution P^t,

$$d(P^t, \pi) \leq \Pr(X_t \neq Y_t).$$

Proximity function

$\rho : \Omega \times \Omega \to \mathbb{N}$

$\rho(x, y) = 0$ implies $x = y$.

$$\Pr(X_t \neq Y_t) \leq E[\rho(X_t, Y_t)].$$

For some $\gamma < 1$,

$$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (x, y)] \leq \gamma\, \rho(x, y).$$

$$d(P^t, \pi) \leq \gamma^t \rho(X_0, Y_0).$$

$$\tau(\epsilon) \leq \max_{(x,y)} \frac{\log(\rho(x, y)/\epsilon)}{1 - \gamma}.$$

Example: Sampling colourings

state space $\Omega_k(G)$ is the set of proper k-colourings of an n-vertex graph G

"Glauber Dynamics" Markov Chain:

- Choose vertex v and colour i u.a.r.
- Recolour v with i iff this makes v properly coloured

Special Case: G is the edge-free graph

The coupling: Use the same random choice for both copies.

$\rho =$ Hamming distance.

$\gamma = 1 - 1/n$.

$\rho(x, y) \leq n$.

$$\tau(\epsilon) \leq n \log(n/\epsilon).$$

Path Coupling (Bubley and Dyer 1997)

- $S \subseteq \Omega \times \Omega$

- transitive closure of S
  (viewed as a relation) is $\Omega \times \Omega$

- Proximity function ρ for pairs in S

- Define coupling on pairs in S and "lift" to $\Omega \times \Omega$

A **path** from state x to state y

$$\overset{x}{\overset{\|}{w_1}} — w_2 — w_3 — w_4 — \qquad \cdots \qquad — \overset{y}{\overset{\|}{w_r}}$$

Adjacency: $(w_i, w_{i+1}) \in S$

- Define ρ on **all** pairs in $\Omega \times \Omega$, ensuring that the *triangle inequality* holds:

$$\rho(x,y) = \min \sum_{i=1}^{r-1} \rho(w_i, w_{i+1}),$$

  where the "min" is over all paths from x to y.

- Extending the coupling: For each pair (x,y) **fix** a shortest path from x to y and compose the couplings along the path.

By construction of the coupling on S, there is a $\gamma \leq 1$ such that for all $(\chi, \psi) \in S$ we have

$$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (\chi, \psi)] \leq \gamma\, \rho(\chi, \psi).$$

We want this for all $(x, y) \in \Omega \times \Omega$:

$$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (x, y)] \leq \gamma\, \rho(x, y).$$

This follows by linearity of expectation.

Delayed Path Coupling

Czumaj, Kanarek, Kutylowksi, Loryś

Instead of proving, for all $(\chi, \psi) \in S$,

$$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (\chi, \psi)] \leq \gamma\, \rho(\chi, \psi).$$

Choose a positive integer R and prove

$$E[\rho(X_R, Y_R) \mid (X_0, Y_0) = (\chi, \psi)] \leq \gamma\, \rho(\chi, \psi).$$

$$E[\rho(X_{R\,t}, Y_{R\,t}) \mid (X_0, Y_0) = (x, y)] \leq \gamma^t \rho(x, y).$$

$$\tau(\epsilon) \leq \max_{(x,y)} \frac{\log(\rho(x, y)/\epsilon)}{1 - \gamma} \cdot R.$$

Our Approach

If

A. for all $(\chi, \psi) \in S$,
$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (\chi, \psi)] \leq \beta\rho(\chi, \psi)$,
and

B. R is a positive integer, and

C. T is a stopping time such that for $0 \leq s \leq T$ we have $\rho(X_s, Y_s) = \rho(X_0, Y_0)$,

Then

$$
\begin{aligned}
E\ [\ \rho(X_R, Y_R) \mid (X_0, Y_0) = (\chi, \psi)\ ]
&\leq \Pr(T > R)\, \rho(\chi, \psi) \\
&+ \Pr(T \leq R)\, \beta^R \\
&\quad E[\rho(X_T, Y_T) \mid T \leq R \text{ and } (X_0, Y_0) = (\chi, \psi)].
\end{aligned}
$$

Suppose $\rho(\chi, \psi) = 1$. Let γ be

$$\begin{aligned} & \Pr(T > R) \\ + \; & \Pr(T \le R)\, \beta^R \\ & E[\rho(X_T, Y_T) \mid T \le R \text{ and } (X_0, Y_0) = (\chi, \psi)]. \end{aligned}$$

If $0 < \gamma < 1$, we get

$$\tau(\epsilon) \le \max_{(x,y)} \frac{\log(\rho(x,y)/\epsilon)}{1-\gamma} \cdot R.$$

The Glauber Dynamics

state space $\Omega_k(G)$ is the set of proper k-colourings of an n-vertex graph G

- Choose vertex v and colour i u.a.r.

- Recolour v with i iff this makes v properly coloured

- Jerrum; Salas, Sokal (1995)
  $O(n \log n)$ mixing time for $k > 2\Delta$

- Bubley, Dyer (1997)
  $O(n^3)$ mixing time for $k = 2\Delta$

- Vigoda (1999)

  - $O(n \log n)$ mixing time for more complicated chain when $k \geq 11\Delta/6$

  - applied comparison methods of Diaconis and Saloff-Coste (1993) to get mixing time $O(k \log(k) n^2 \log(n))$ for Glauber dynamics

Suppose

- The graph is Δ-regular (for $\Delta \geq 14$) and triangle-free

- $(2 - \eta)\Delta \leq k \leq 2\Delta$

Then the mixing time is $O(n \log n)$

state space $\Omega_k(G)$ is the set of proper k-colourings of a Δ-regular triangle-free n-vertex graph G.

Markov chain: Glauber dynamics.

$S \subseteq \Omega_k(G) \times \Omega_k(G)$.

ρ : Hamming distance.

Define coupling on pairs in S and “lift” to $\Omega_k(G) \times \Omega_k(G)$

Our Approach

If

A. for all $(\chi, \psi) \in S$,
$E[\rho(X_1, Y_1) \mid (X_0, Y_0) = (\chi, \psi)] \leq \beta\rho(\chi, \psi)$,
and

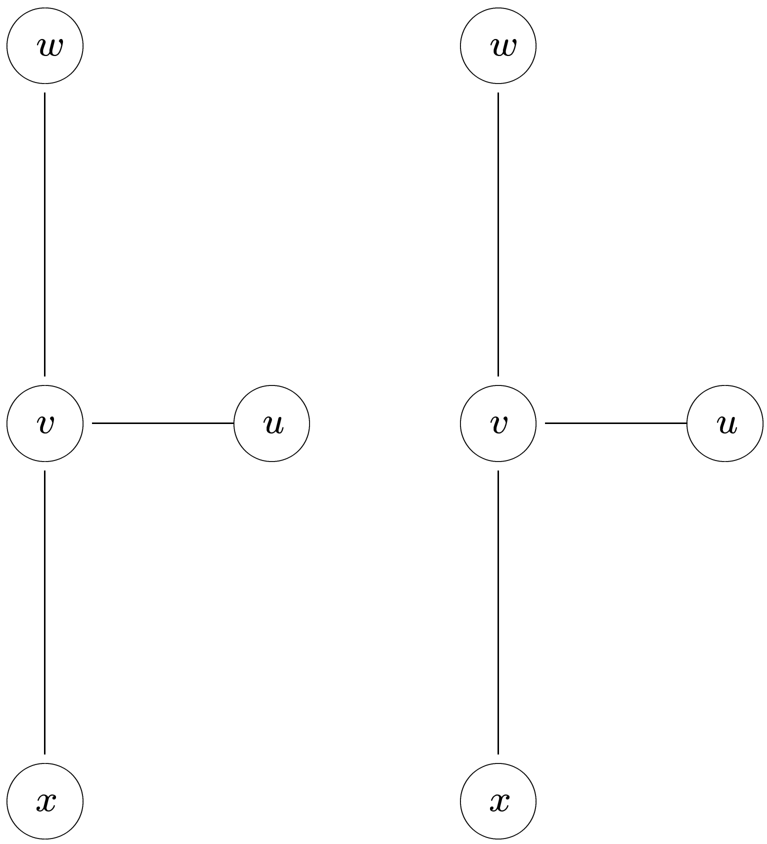
B. R is a positive integer, and

C. T is a stopping time such that for $0 \leq s \leq T$ we have $\rho(X_s, Y_s) = \rho(X_0, Y_0)$,

Then

$$
\begin{aligned}
E\ [\ \rho(X_R, Y_R) \mid (X_0, Y_0) = (\chi, \psi)\ ] \\
\leq\ & \Pr(T > R)\, \rho(\chi, \psi) \\
+\ & \Pr(T \leq R)\, \beta^R \\
& E[\rho(X_T, Y_T) \mid T \leq R \text{ and } (X_0, Y_0) = (\chi, \psi)].
\end{aligned}
$$

χ ψ



The Coupling for pairs in S

If vertex α and colour i are chosen then use this choice in both copies except

- $\alpha = w$, $i = \chi(v)$. Use $(w, \chi(v))$ in χ. Use $(w, \psi(v))$ in ψ. (Both fail, so ρ does not change.)

- $\alpha = w$, $i = \psi(v)$. Use $(w, \psi(v))$ in χ. Use $(w, \chi(v))$ in ψ. (ρ may increase by 1.)

The only other time that ρ could change is when $\alpha = v$. If the colour i is good for v then ρ decreases by 1.

Expected new value of ρ

$$\leq 1 - \frac{A}{kn} + \frac{\Delta}{kn} \leq 1 + \frac{\eta\Delta}{kn},$$

where A is the number of colours **not** used by v's neighbors.

Stopping time T: First time that either

- v is selected, or
- a neighbor w is selected with colour $\Psi(v)$.

Suppose $\rho(\chi, \psi) = 1$. Let γ be

$$\begin{aligned} & \Pr(T > R) \\ + \; & \Pr(T \leq R)\, \beta^R \\ & E[\rho(X_T, Y_T) \mid T \leq R \text{ and } (X_0, Y_0) = (\chi, \psi)]. \end{aligned}$$

If $0 < \gamma < 1$, we get

$$\tau(\epsilon) \leq \max_{(x,y)} \frac{\log(\rho(x,y)/\epsilon)}{1 - \gamma} \cdot R.$$

$$\begin{aligned} \gamma = 1 \;\; - \;\; & \Pr(T \leq R) \\ \Big(\;\; & 1 - \beta^R E[\rho(X_T, Y_T) \mid \\ & \quad T \leq R \text{ and } (X_0, Y_0) = (\chi, \psi)\;] \Big). \end{aligned}$$

Expected new value of ρ

$$\leq 1 - \frac{A}{kn} + \frac{\Delta}{kn}$$

$$E[C \mid T \leq n/8 \text{ and } (X_0, Y_0) = (\chi, \psi)] \geq 2\eta\Delta,$$

where C is the number of colours which occur *more than once* around v just before the stopping time.

$B_i =$ set of neighbors of v at which colour i is acceptable in both χ and ψ

There are at least $k/5$ colours i such that $|B_i| \geq \Delta/3$

Consider the event in which exactly *two* elements of B_i, u and w, are re-coloured with colour i before the stopping time.

- u and w are never re-chosen.

- no neighbor of u or w is chosen with colour i.

$\Pr[i \text{ repeated} \mid T \leq n/8, |B_i| \geq \Delta/3] \geq \frac{10\eta\Delta}{k}$.