

A Scan Markov Chain for Sampling Colourings

Kasper Pedersen

`kasper@dcs.warwick.ac.uk`

Department of Computer Science
University of Warwick

British Colloquium for Theoretical Computer Science, 2006

Outline

- 1 Introduction
 - Graph Colourings and Markov Chains
 - Previous Work
- 2 Systematic Scan
 - General Condition for Rapid Mixing
 - Application: Rapid Mixing for $q \geq 2\Delta$

Proper Colouring of Graphs

Computational problem

Want to sample **efficiently** from the (nearly) uniform distribution of proper q -colourings of a graph with maximum vertex degree Δ using a **systematic approach**.

Definition

A **proper** colouring of a graph is an assignment of a colour to each site where no two adjacent sites have the same colour.

Definition

Ω is the set of all proper colourings.

Proper Colouring of Graphs

Computational problem

Want to sample **efficiently** from the (nearly) uniform distribution of proper q -colourings of a graph with maximum vertex degree Δ using a **systematic approach**.

Definition

A **proper** colouring of a graph is an assignment of a colour to each site where no two adjacent sites have the same colour.

Definition

Ω is the set of all proper colourings.

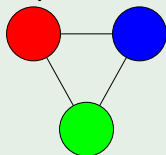
Proper Colourings

Definition

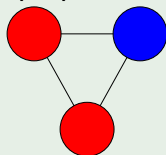
A **proper** colouring of a graph is an assignment of a colour to each site where no two adjacent sites have the same colour.

Example

A proper colouring



An improper colouring



Markov Chains and Sampling

A Markov chain is a random process X_0, X_1, \dots

- each state X_k takes a value in Ω (**state space**), and
- the transition at any time depends **only** on the current state

If q is sufficiently large then a Markov chain converges to its **stationary distribution** (subject to some conditions!)

Definition

\mathcal{M} is a Markov chain with **state space** Ω and **stationary distribution** the uniform distribution on Ω .

Markov Chains and Sampling

A Markov chain is a random process X_0, X_1, \dots

- each state X_k takes a value in Ω (**state space**), and
- the transition at any time depends **only** on the current state

If q is sufficiently large then a Markov chain converges to its **stationary distribution** (subject to some conditions!)

Definition

\mathcal{M} is a Markov chain with **state space** Ω and **stationary distribution** the uniform distribution on Ω .

Markov Chains and Sampling

A Markov chain is a random process X_0, X_1, \dots

- each state X_k takes a value in Ω (**state space**), and
- the transition at any time depends **only** on the current state

If q is sufficiently large then a Markov chain converges to its **stationary distribution** (subject to some conditions!)

Definition

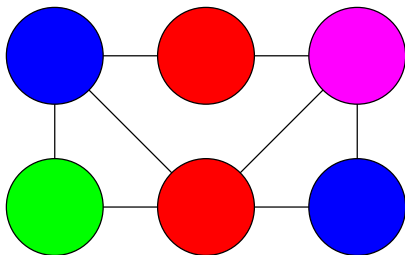
\mathcal{M} is a Markov chain with **state space** Ω and **stationary distribution** the uniform distribution on Ω .

Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

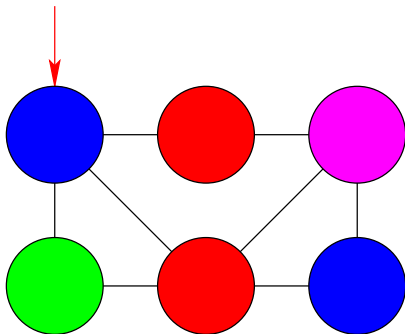


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

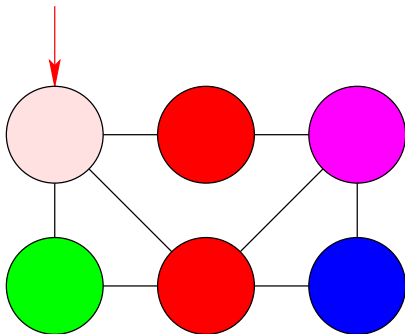


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

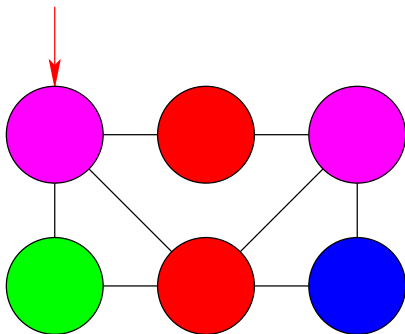


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

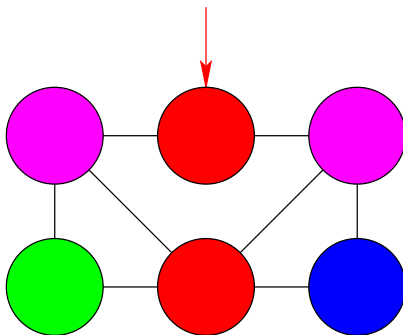


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

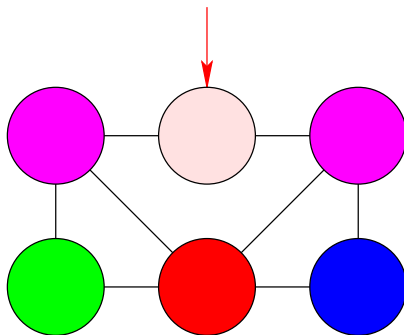


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

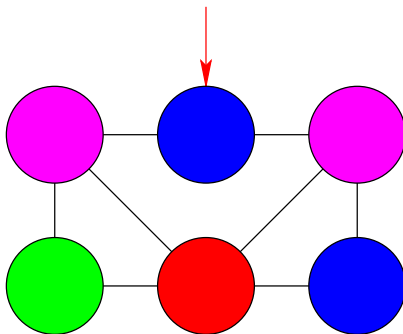


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

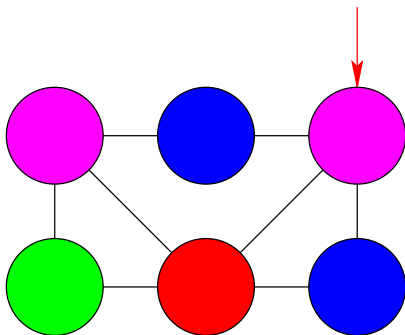


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

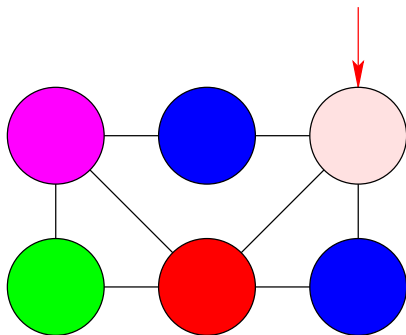


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.

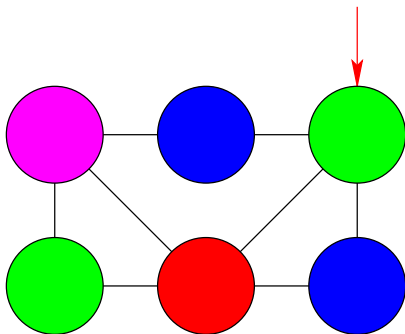


Scan Markov Chains

Definition

A Markov chain on Ω is called a **scan** if the sites are updated in a deterministic order.

- An **update** is a randomised procedure. E.g. heat bath.



Mixing Time of Markov Chains

Definition

- The **mixing time** of a Markov chain is how long it takes to become sufficiently close to its stationary distribution.
- A Markov chain is **rapidly mixing** if the mixing time is at most polynomial in the size of the graph.

Computational question

Given a systematic scan \mathcal{M}

- For what values of q (in terms of Δ) is \mathcal{M} rapidly mixing?

Mixing Time of Markov Chains

Definition

- The **mixing time** of a Markov chain is how long it takes to become sufficiently close to its stationary distribution.
- A Markov chain is **rapidly mixing** if the mixing time is at most polynomial in the size of the graph.

Computational question

Given a systematic scan \mathcal{M}

- For what values of q (in terms of Δ) is \mathcal{M} rapidly mixing?

Previous Work

Systematic scan:

- $O(n \log n)$ mixing, $q > 2\Delta$. General graphs (Dyer, Goldberg and Jerrum, 2005)
- $\text{poly}(n)$ mixing, $q = 2\Delta$. General graphs (Dyer, Goldberg and Jerrum, 2005)
- $O(n \log n)$ mixing, $q > f(\Delta)$ where $f(\Delta) \rightarrow \beta\Delta$ as $\Delta \rightarrow \infty$ and $\beta \approx 1.76$. Bipartite graphs (Bordewich, Dyer and Karpinski, 2005)

Previous Work

Systematic scan:

- $O(n \log n)$ mixing, $q > 2\Delta$. General graphs (Dyer, Goldberg and Jerrum, 2005)
- $\text{poly}(n)$ mixing, $q = 2\Delta$. General graphs (Dyer, Goldberg and Jerrum, 2005)
- $O(n \log n)$ mixing, $q > f(\Delta)$ where $f(\Delta) \rightarrow \beta\Delta$ as $\Delta \rightarrow \infty$ and $\beta \approx 1.76$. Bipartite graphs (Bordewich, Dyer and Karpinski, 2005)

Random update:

- $O(n \log n)$ mixing, $q > \frac{11}{6}\Delta$. General graphs (Vigoda, 2000)

Overview of Results

General Mixing Result

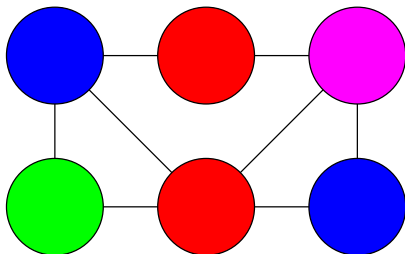
We show a condition for $O(n \log n)$ mixing of an arbitrary systematic scan.

Application

- $O(n \log n)$ mixing when $q \geq 2\Delta$ on general graphs

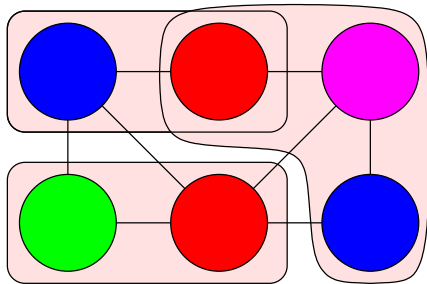
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



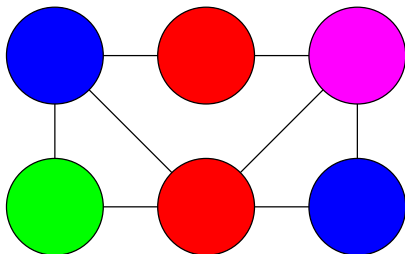
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



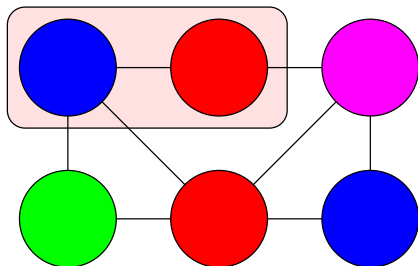
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



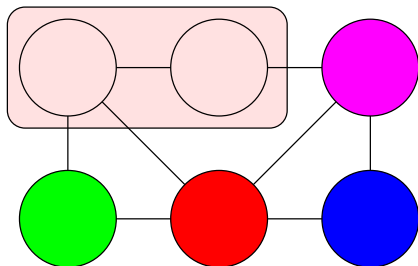
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



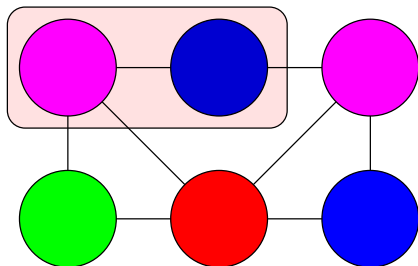
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



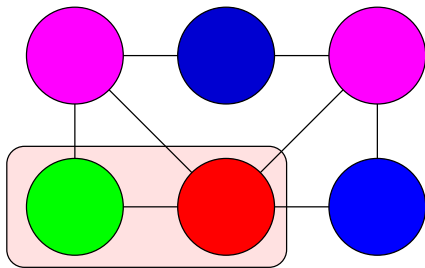
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



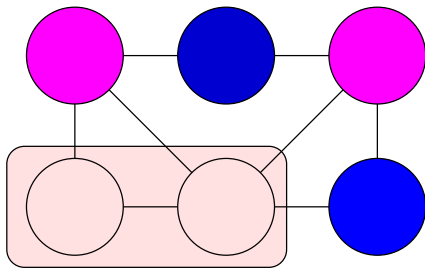
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



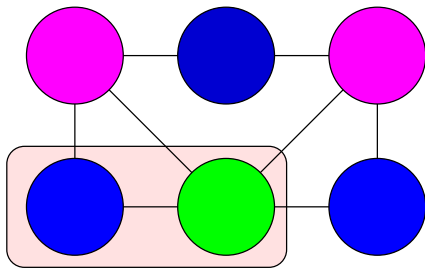
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



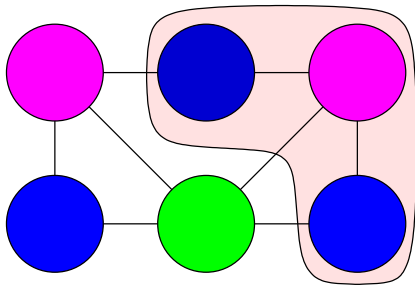
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



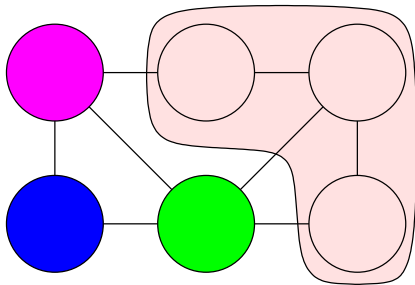
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



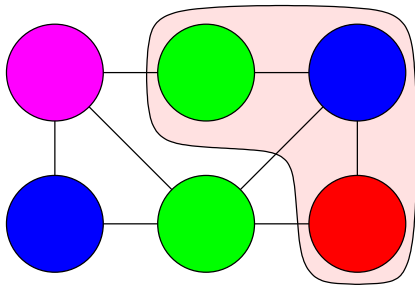
Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



Block Moves

- A **block move** considers a set of sites for simultaneous update.
- Scan: set of blocks Θ must cover V . A systematic scan updates the blocks in a deterministic order.



Coupling

Definition

A **coupling** of two distributions X and Y on Ω is a distribution ψ on $\Omega \times \Omega$ such that when (x, y) is drawn from ψ then the distribution of x is X and the distribution of y is Y .

Example

Let $\Omega = \{S_1, S_2\}$

X, Y distributions on Ω .

- $\Pr_X(S_1) = 1/2$ and
 $\Pr_X(S_2) = 1/2$
- $\Pr_Y(S_1) = 1/3$ and
 $\Pr_Y(S_2) = 2/3$

Coupling

Definition

A **coupling** of two distributions X and Y on Ω is a distribution ψ on $\Omega \times \Omega$ such that when (x, y) is drawn from ψ then the distribution of x is X and the distribution of y is Y .

Example

Let $\Omega = \{S_1, S_2\}$

X, Y distributions on Ω .

- $\Pr_X(S_1) = 1/2$ and
 $\Pr_X(S_2) = 1/2$
- $\Pr_Y(S_1) = 1/3$ and
 $\Pr_Y(S_2) = 2/3$

\Pr_ψ	X	Y
1/3	S_1	S_1
1/2	S_2	S_2
1/6	S_1	S_2

Influence *on* a site

Consider two copies X and Y of the same M.C

- X and Y differ only at the colour of site i
- block Θ_k is going to be updated
- $P^{[k]}$ is the transition matrix for updating block Θ_k

Definition

The **influence** site i has **on** site j is the maximum probability that the two copies of the M.C will differ at site j in **some coupling** of the distributions $P^{[k]}(X, \cdot)$ and $P^{[k]}(Y, \cdot)$. This is denoted by $\rho_{i,j}^k$.

Total Influence *on* a site

- Total influence **on** site j (under block Θ_k) is

$$\sum_{i \in V} \rho_{i,j}^k$$

Definition

The maximum influence **on any** site is

$$\alpha = \max_{k \in \Theta} \max_{j \in V} \sum_{i \in V} \rho_{i,j}^k$$

Mixing Time of Systematic Scan

Definition

$$\alpha = \max_{k \in \Theta} \max_{j \in V} \sum_{i \in V} \rho_{i,j}^k$$

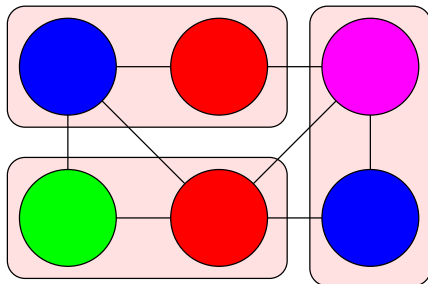
Theorem

\mathcal{M} is any systematic scan with block set Θ . If $\alpha < 1$ then

$$\text{Mix}(\mathcal{M}, \epsilon) \leq O\left(\frac{n \log(n\epsilon^{-1})}{1 - \alpha}\right)$$

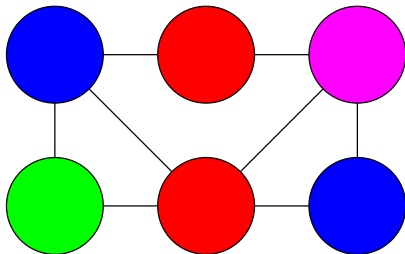
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



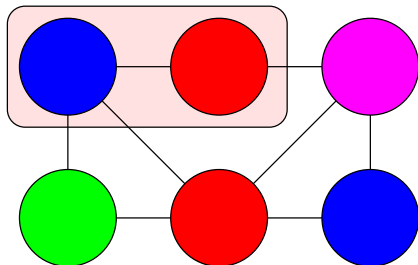
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



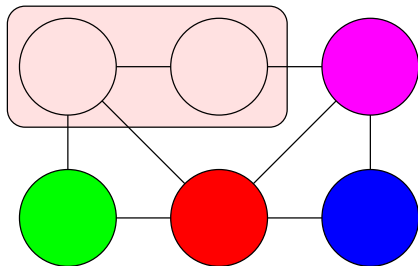
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



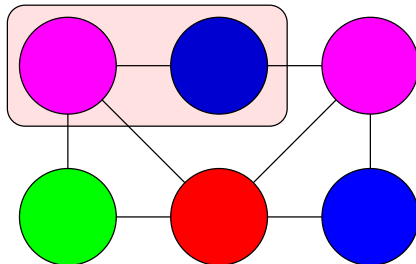
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



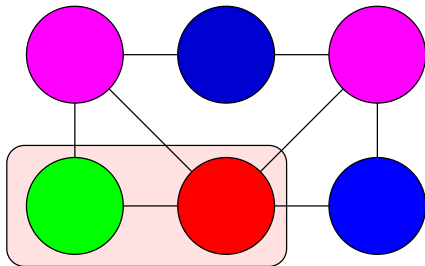
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



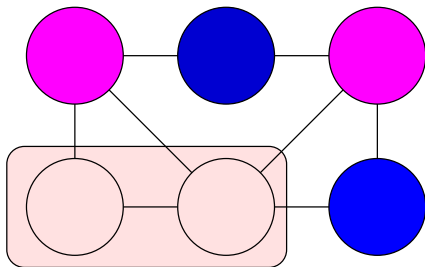
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



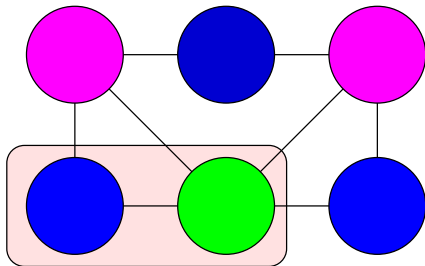
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



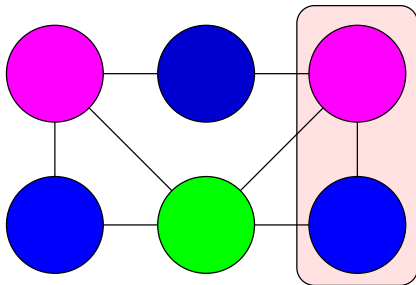
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



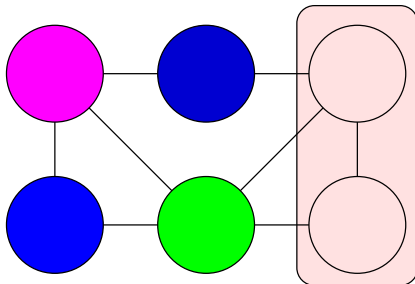
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



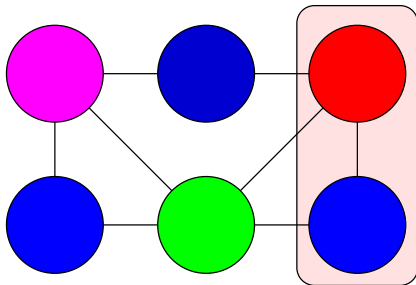
Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



Application: Edge Scan

- $G = (V, E)$ is any graph
- $\Theta \subseteq E$
- Update rule is heat bath



Mixing Time of Edge Scan

Case analysis gives

$$\alpha < 1 - \frac{1}{\Delta^2}$$

when $q \geq 2\Delta$ so

$$\text{Mix}(\mathcal{M}_-, \epsilon) \leq \Delta^2 n \log(n\epsilon^{-1})$$

Remark

This is the first general $O(n \log n)$ mixing result for scan when $q = 2\Delta$.

Summary

- Showed a condition for $O(n \log n)$ mixing of systematic scan
- Used the condition to prove $O(n \log n)$ mixing for general graphs provided $q \geq 2\Delta$.