MCMC sampling colourings and independent sets of $G(n, d/n)$ near the uniqueness threshold.

Charis Efthymiou
efthymiou@math.uni-frankfurt.de

Goethe University, Frankfurt

(appeared in SODA'14)

Phase Transitions in Discrete Structures and Computational Problems
Warwick, May 2014

Gibbs Distributions and the Sampling Problem

Gibbs Distributions and the Sampling Problem

Gibbs Distribution

Given a graph $G = (V, E)$ and some integer $k > 0$ and $\lambda > 0$ we let

Colouring Model: For each proper k-colouring σ we have

$$\mu(\sigma) = 1/Z_{G,k}$$

Hard-Core Model: For each independent set σ

$$\mu(\sigma) = \lambda^{|\sigma|}/Z_{G,\lambda}.$$

Gibbs Distributions and the Sampling Problem

Gibbs Distribution

Given a graph $G = (V, E)$ and some integer $k > 0$ and $\lambda > 0$ we let

Colouring Model: For each proper k-colouring σ we have

$$\mu(\sigma) = 1/Z_{G,k}$$

Hard-Core Model: For each independent set σ

$$\mu(\sigma) = \lambda^{|\sigma|}/Z_{G,\lambda}.$$

Sampling Problem

Input: A graph $G = (V, E)$ and a target distribution $\mu(\cdot)$, e.g. Colouring or Hard-Core Model.

Output: A configuration distributed as in $\mu(\cdot)$.

Sampling Problem on $G(n, d/n)$

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

We consider the problem with underlying graph $G(n, d/n)$

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

We consider the problem with underlying graph $G(n, d/n)$

- A graph on n vertices and each edge appears independently with probability d/n, where d is fixed

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

We consider the problem with underlying graph $G(n, d/n)$

- A graph on n vertices and each edge appears independently with probability d/n, where d is fixed
- The focus is on "*typical instances*" of $G(n, d/n)$

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

We consider the problem with underlying graph $G(n, d/n)$

- A graph on n vertices and each edge appears independently with probability d/n, where d is fixed
- The focus is on "*typical instances*" of $G(n, d/n)$
- $\mathcal{E}_n$ occurs *with high probability* (w.h.p.) if $\lim_{n\to\infty} Pr[\mathcal{E}_n] = 1$

Sampling Problem on $G(n, d/n)$

Hardness of Sampling

The Sampling Problem is "computationally hard"

- We don't expect to have an efficient algorithm
- The main focus is on approximate algorithms

The Average Case Scenario

We consider the problem with underlying graph $G(n, d/n)$

- A graph on n vertices and each edge appears independently with probability d/n, where d is fixed
- The focus is on "*typical instances*" of $G(n, d/n)$
- $\mathcal{E}_n$ occurs *with high probability* (w.h.p.) if $\lim_{n\to\infty} Pr[\mathcal{E}_n] = 1$

Remark

... for "*typical instances*" of $G(n, d/n)$ we do not expect to have exact algorithms, too.

Markov Chain Monte Carlo Sampling

A rough idea....

Markov Chain Monte Carlo Sampling

A rough idea....

- Consider an appropriately defined Markov Chain $X_0, X_1, \ldots$ over the configurations of G, e.g. k-colouring.

Markov Chain Monte Carlo Sampling

A rough idea....

- Consider an appropriately defined Markov Chain $X_0, X_1, \ldots$ over the configurations of G, e.g. k-colouring.
  - It is ergodic, i.e. it converges to a unique stationary distribution

Markov Chain Monte Carlo Sampling

A rough idea....

- Consider an appropriately defined Markov Chain $X_0, X_1, \ldots$ over the configurations of G, e.g. k-colouring.
  - It is ergodic, i.e. it converges to a unique stationary distribution
  - The stationary distribution should be the Gibbs distribution, $\mu(\cdot)$

Markov Chain Monte Carlo Sampling

A rough idea....

- Consider an appropriately defined Markov Chain $X_0, X_1, \ldots$ over the configurations of G, e.g. k-colouring.
  - It is ergodic, i.e. it converges to a unique stationary distribution
  - The stationary distribution should be the Gibbs distribution, $\mu(\cdot)$
- The algorithm simulates the chain and outputs X_T, for sufficiently large T.

The Markov Chain

The Markov Chain

"Glauber Block Dynamics"

- We are given a partition of the vertex set $\mathcal{B} = \{B_1, \ldots, B_N\}$.
- $X_0 = \sigma$ for arbitrary σ.
- Given X_t, we get X_{t+1} as follows:
  - Choose block B *uniformly at random* among all the blocks in $\mathcal{B}$
  - Set $X_{t+1}(u) = X_t(u)$, for every vertex $u \notin B$
  - Set $X_{t+1}(B)$ according to distribution μ conditional on $X_{t+1}(V \backslash B)$.

Convergence of Glauber Dynamics

Convergence of Glauber Dynamics

Ergodicity

The chain converges to $\mu(\cdot)$ given that the followings hold:

Convergence of Glauber Dynamics

Ergodicity

The chain converges to $\mu(\cdot)$ given that the followings hold:

- Aperiodic

Convergence of Glauber Dynamics

Ergodicity

The chain converges to $\mu(\cdot)$ given that the followings hold:

- Aperiodic
- The state space of the chain is "connected"

Convergence of Glauber Dynamics

Ergodicity

The chain converges to $\mu(\cdot)$ given that the followings hold:

- Aperiodic
- The state space of the chain is "connected"

Remark

For the chains we consider here ergodicity is well known to hold [DFFV'05].

Rate of Convergence

Rate of Convergence

Mixing Time

The number of transitions needed for the chain to reach within *total variation distance* $1/e$ from $\mu(\cdot)$. Regardless of the initial state.

Rate of Convergence

Mixing Time

The number of transitions needed for the chain to reach within *total variation distance* $1/e$ from $\mu(\cdot)$. Regardless of the initial state.

Total Variation Distance

For two distributions ν, μ over Ω, we define their total variation distance as follows:

$$||\nu - \mu||_{TV} = \max_{A \subseteq \Omega} |\nu(A) - \mu(A)|.$$

Rate of Convergence

Mixing Time

The number of transitions needed for the chain to reach within *total variation distance* $1/e$ from $\mu(\cdot)$. Regardless of the initial state.

Rapid Mixing

The mixing time τ_{mix} is polynomial in n, the number of the vertices of G.

- If $T(err)$ is the minimum number of transitions to get within error err from μ, then

$$T(err) \leq \ln\left(\frac{1}{err}\right)\tau_{mix}.$$

Rapid Mixing and Maximum Degree Δ

Rapid Mixing and Maximum Degree Δ

Maximum Degree Bounds for colourings

Vigoda (1999) $k > \frac{11}{6}\Delta$ for general G

Hayes,Vera,Vigoda (2007) $k = \Omega(\Delta/\log\Delta)$ for planar G

Goldberg, Martin, Paterson (2004) $k \geq (1.763 + \epsilon)\Delta$ for G triangle free and amenable

Dyer, Frieze, Hayes, Vigoda (2004) $k \geq (1.48 + \epsilon)\Delta$ for G of girth $g \geq 7$

Frieze, Vera (2006) $k \geq (1.763 + \epsilon)\Delta$ for G locally sparse.

Rapid Mixing and Maximum Degree Δ

Maximum Degree Bounds for colourings

Vigoda (1999) $k > \frac{11}{6}\Delta$ for general G

Hayes,Vera,Vigoda (2007) $k = \Omega(\Delta / \log \Delta)$ for planar G

Goldberg, Martin, Paterson (2004) $k \geq (1.763 + \epsilon)\Delta$ for G triangle free and amenable

Dyer, Frieze, Hayes, Vigoda (2004) $k \geq (1.48 + \epsilon)\Delta$ for G of girth $g \geq 7$

Frieze, Vera (2006) $k \geq (1.763 + \epsilon)\Delta$ for G locally sparse.

Hard-Core

The situation is very similar for the parameter λ in the Hard-Core Model .

The interesting case of $G(n, d/n)$

The interesting case of $G(n, d/n)$

Degrees in $G(n, d/n)$

The interesting case of $G(n, d/n)$

Degrees in $G(n, d/n)$

- The maximum degree in $G(n, d/n)$ is $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$ w.h.p.

The interesting case of $G(n, d/n)$

Degrees in $G(n, d/n)$

- The maximum degree in $G(n, d/n)$ is $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$ w.h.p.
- The "vast majority" of the vertices are of degree in $(1 \pm \epsilon)d$ w.h.p.

The interesting case of $G(n, d/n)$

Degrees in $G(n, d/n)$

- The maximum degree in $G(n, d/n)$ is $\Theta\left(\frac{\ln n}{\ln \ln n}\right)$ w.h.p.
- The "vast majority" of the vertices are of degree in $(1 \pm \epsilon)d$ w.h.p.

Remark

It seems "natural" to have the bounds on k, λ for rapid mixing depending on the *expected degree d* rather than maximum degree Δ.

Statistical Physics Perspective

Statistical Physics Perspective

Conjecture Bounds for rapid mixing

Statistical Physics Perspective

Conjecture Bounds for rapid mixing

- For *colouring* we need $k > d$

Statistical Physics Perspective

Conjecture Bounds for rapid mixing

- For *colouring* we need $k > d$
- For *hard core* we need $\lambda < \frac{(d-1)^{d-1}}{(d-2)^d} \approx \frac{e}{d}$.

Statistical Physics Perspective

Conjecture Bounds for rapid mixing

- For *colouring* we need $k > d$
- For *hard core* we need $\lambda < \frac{(d-1)^{d-1}}{(d-2)^d} \approx \frac{e}{d}$.

Otherwise

... there are exceptional initial states, from which the mixing is slow or there is no mixing at all

Past Work

Previous Bounds for Rapid Mixing

Past Work

Previous Bounds for Rapid Mixing

- Dyer, Flaxman, Frieze, Vigoda (2005): $k \geq \Theta\left(\frac{\ln \ln n}{\ln \ln \ln n}\right)$

Past Work

Previous Bounds for Rapid Mixing

- Dyer, Flaxman, Frieze, Vigoda (2005): $k \geq \Theta\left(\frac{\ln \ln n}{\ln \ln \ln n}\right)$
  - k is exponentially smaller than the max-degree but still depends on n

Past Work

Previous Bounds for Rapid Mixing

- Dyer, Flaxman, Frieze, Vigoda (2005): $k \geq \Theta\left(\frac{\ln \ln n}{\ln \ln \ln n}\right)$
  - k is exponentially smaller than the max-degree but still depends on n
- Mossel, Sly (2008): $k \geq f(d)$ and $\lambda \leq h(d)$.

Past Work

Previous Bounds for Rapid Mixing

- Dyer, Flaxman, Frieze, Vigoda (2005): $k \geq \Theta\left(\frac{\ln\ln n}{\ln\ln\ln n}\right)$
  - k is exponentially smaller than the max-degree but still depends on n
- Mossel, Sly (2008): $k \geq f(d)$ and $\lambda \leq h(d)$.
  - ... $f(d) = d^c$ and $h(d) = d^{-c'}$, for some $c, c' > 4$.

Main Result

Main Result

Result for Rapid Mixing

W.h.p. over the instances of $G(n, d/n)$ the graph admits a partition of the vertex set into a set of "simple structured" blocks $\mathcal{B}$ s.t. the following holds: Let $\mathcal{M}_c$ and $\mathcal{M}_{hc}$ denote the Glauber block dynamics for the *colouring model* and the *hard core model*, respectively, with set of blocks $\mathcal{B}$.

- For $k \geq \frac{11}{2}d$ the mixing time of $\mathcal{M}_c$ is $O(n \ln n)$
- For $\lambda \leq \frac{1-\epsilon}{2d}$ the mixing time of $\mathcal{M}_{hc}$ is $O(n \ln n)$.

Main Result

Result for Rapid Mixing

W.h.p. over the instances of $G(n, d/n)$ the graph admits a partition of the vertex set into a set of "simple structured" blocks $\mathcal{B}$ s.t. the following holds: Let $\mathcal{M}_c$ and $\mathcal{M}_{hc}$ denote the Glauber block dynamics for the *colouring model* and the *hard core model*, respectively, with set of blocks $\mathcal{B}$.

- For $k \geq \frac{11}{2}d$ the mixing time of $\mathcal{M}_c$ is $O(n \ln n)$
- For $\lambda \leq \frac{1-\epsilon}{2d}$ the mixing time of $\mathcal{M}_{hc}$ is $O(n \ln n)$.

For efficient sampling we need to have efficient...

- construction of $\mathcal{B}$
- implementation of the updates
- algorithms that provide initial configurations for both chains.

Rapid Mixing and High Degrees

Rapid Mixing and High Degrees

Technical Challenge

Rapid Mixing and High Degrees

Technical Challenge

- Standard analysis for rapid mixing *cannot* be employed here due to the high degree vertices ($deg > (1+\epsilon)d$)

Rapid Mixing and High Degrees

Technical Challenge

- Standard analysis for rapid mixing *cannot* be employed here due to the high degree vertices ($deg > (1 + \epsilon)d$)
  - Even though they are relatively few, high degree vertices appear everywhere in the graph!

Rapid Mixing and High Degrees

Technical Challenge

- Standard analysis for rapid mixing *cannot* be employed here due to the high degree vertices ($deg > (1+\epsilon)d$)
  - Even though they are relatively few, high degree vertices appear everywhere in the graph!
- Dyer et al. (2005) "Hide the high degree vertices well inside the blocks"

Rapid Mixing and High Degrees

Technical Challenge

- Standard analysis for rapid mixing *cannot* be employed here due to the high degree vertices ($deg > (1+\epsilon)d$)
  - Even though they are relatively few, high degree vertices appear everywhere in the graph!
- Dyer et al. (2005) "Hide the high degree vertices well inside the blocks"

The crux is ...

It is all about creating an appropriate set of blocks.

Rapid Mixing and High Degrees

Technical Challenge

- Standard analysis for rapid mixing *cannot* be employed here due to the high degree vertices ($deg > (1 + \epsilon)d$)
  - Even though they are relatively few, high degree vertices appear everywhere in the graph!
- Dyer et al. (2005) "Hide the high degree vertices well inside the blocks"

The crux is ...

It is all about creating an appropriate set of blocks.

- ... it is highly non-trivial!

Block Construction I

Block Construction I

Weights for Vertices and Paths

- We assign weight to each vertex u of degree deg_u as follows:

$$W(u) = \begin{cases} (1+\gamma)^{-1} & deg_u \leq (1+\epsilon)d \\ d^c \cdot deg_u & \text{otherwise} \end{cases}$$

- Every path L is assigned $\prod_{u \in L} W(u)$

Block Construction I

Weights for Vertices and Paths

- We assign weight to each vertex u of degree deg_u as follows:

$$W(u) = \begin{cases} (1+\gamma)^{-1} & deg_u \leq (1+\epsilon)d \\ d^c \cdot deg_u & \text{otherwise} \end{cases}$$

- Every path L is assigned $\prod_{u \in L} W(u)$

"Break Points"

Let $\mathbb{P}(v)$ denote the set of paths of length at most $\frac{\ln n}{d^{2/5}}$ that emanate from v. We call "break point" every vertex v s.t.

$$\max_{L \in \mathbb{P}(v)} \left\{ \prod_{u \in L} W(u) \right\} \leq 1.$$

Blocks Construction II

Creating Blocks

Blocks Construction II

Creating Blocks

- Find the break-points

Blocks Construction II

Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$

Blocks Construction II

Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do

Blocks Construction II

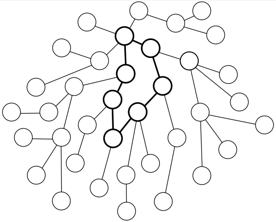
Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C

Blocks Construction II

Creating Blocks

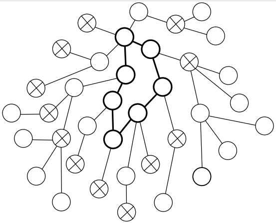
- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C



Blocks Construction II

Creating Blocks

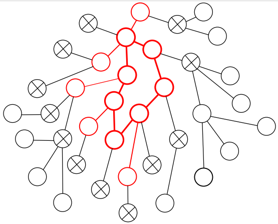
- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C



Blocks Construction II

Creating Blocks

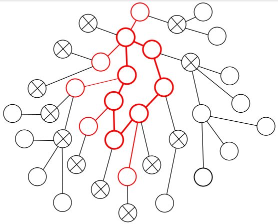
- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C



Blocks Construction II

Creating Blocks

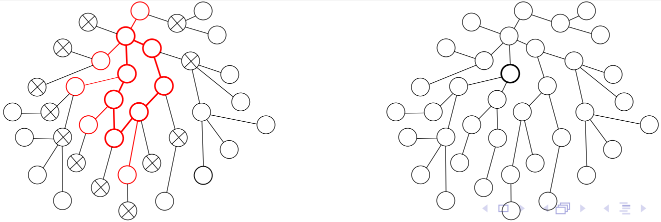
- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C
  - Pick a vertex v (non-break point) that does not belong to a block. The block is the maximal connected subgraph that contains v and no break point



Blocks Construction II

Creating Blocks

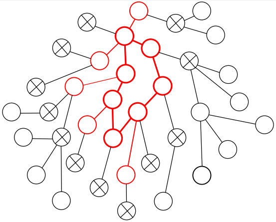
- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C
  - Pick a vertex v (non-break point) that does not belong to a block. The block is the maximal connected subgraph that contains v and no break point

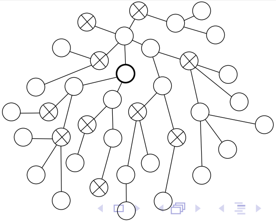


Blocks Construction II

Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C
  - Pick a vertex v (non-break point) that does not belong to a block. The block is the maximal connected subgraph that contains v and no break point

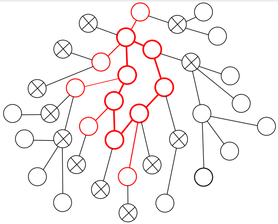


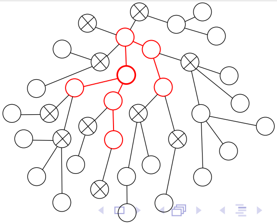


Blocks Construction II

Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C
  - Pick a vertex v (non-break point) that does not belong to a block. The block is the maximal connected subgraph that contains v and no break point

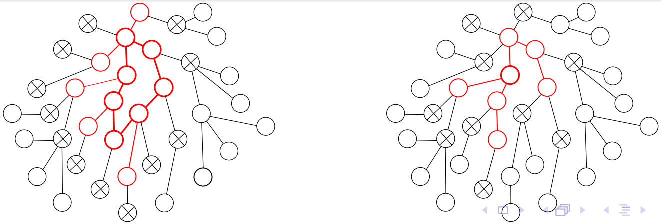




Blocks Construction II

Creating Blocks

- Find the break-points
- $\mathcal{C}$ contains all cycles of length at most $4\frac{\ln n}{\ln^5 d}$
- Given the break points and $\mathcal{C}$ do
  - For each $C \in \mathcal{C}$ construct a block B which is the maximal connected subgraph that contains C but no break points outside C
  - Pick a vertex v (non-break point) that does not belong to a block. The block is the maximal connected subgraph that contains v and no break point
  - If vertex v is a break point then v is a block itself



About the blocks...

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

- The set $\mathcal{B}$ contains blocks which are trees with at most one extra edge

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

- The set $\mathcal{B}$ contains blocks which are trees with at most one extra edge
- The creation of $\mathcal{B}$ can be implemented in polynomial time

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

- The set $\mathcal{B}$ contains blocks which are trees with at most one extra edge
  - The blocks are not extended
- The creation of $\mathcal{B}$ can be implemented in polynomial time

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

- The set $\mathcal{B}$ contains blocks which are trees with at most one extra edge
  - The blocks are not extended
  - No cycles in $\mathcal{C}$ end up in the same block
- The creation of $\mathcal{B}$ can be implemented in polynomial time

About the blocks...

Theorem

W.h.p. over the graph instances $G(n, d/n)$ the following is true:

- The set $\mathcal{B}$ contains blocks which are trees with at most one extra edge
  - The blocks are not extended
  - No cycles in $\mathcal{C}$ end up in the same block
- The creation of $\mathcal{B}$ can be implemented in polynomial time
  - We can check in polynomial time whether some vertex is break-point.

Technique for Rapid Mixing

Technique for Rapid Mixing

Path Coupling, [Bubley, Dyer 1997]

- Consider two copies of the chain at configuration X_0 and Y_0 such that $H(X_0, Y_0) = 1$
- Couple the transitions of the two chains
- For rapid mixing it suffices to have that
$$E[H(X_1, Y_1)|X_0, Y_0] = 1 - \Theta(1/n).$$

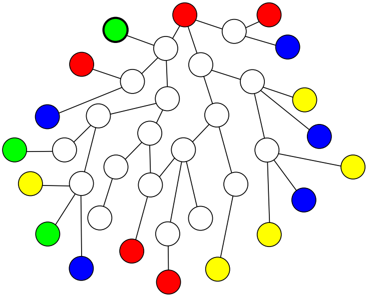
… more concretely

… more concretely

- $\forall B \in \mathcal{B}$ consider arbitrary $\sigma(\partial B)$ and $\tau(\partial B)$ s.t. $H(\sigma(\partial B), \tau(\partial B)) = 1$

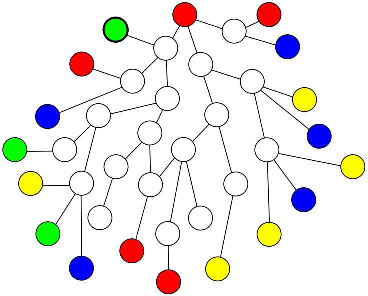
... more concretely

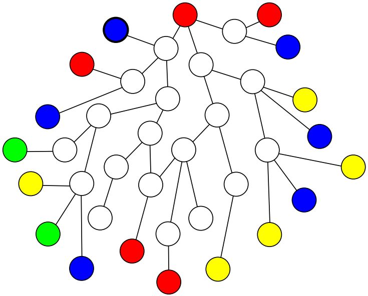
- $\forall B \in \mathcal{B}$ consider arbitrary $\sigma(\partial B)$ and $\tau(\partial B)$ s.t. $H(\sigma(\partial B), \tau(\partial B)) = 1$



… more concretely

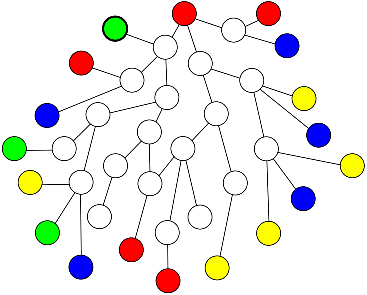
- $\forall B \in \mathcal{B}$ consider arbitrary $\sigma(\partial B)$ and $\tau(\partial B)$ s.t. $H(\sigma(\partial B), \tau(\partial B)) = 1$
- Take $X(B) \sim \mu(\cdot|\sigma(\partial B))$ and $Y(B) \sim \mu(\cdot|\tau(\partial B))$

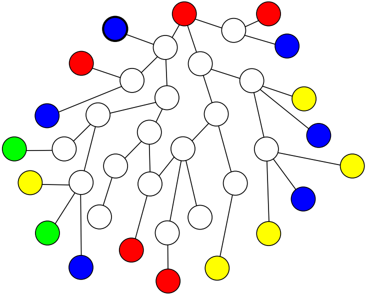




… more concretely

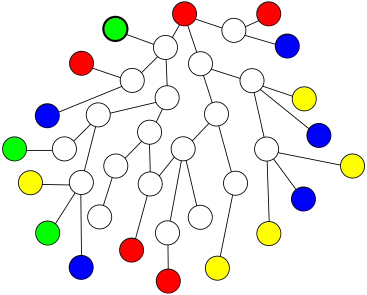
- $\forall B \in \mathcal{B}$ consider arbitrary $\sigma(\partial B)$ and $\tau(\partial B)$ s.t. $H(\sigma(\partial B), \tau(\partial B)) = 1$
- Take $X(B) \sim \mu(\cdot|\sigma(\partial B))$ and $Y(B) \sim \mu(\cdot|\tau(\partial B))$
- Couple $X(B)$ and $Y(B)$ so as minimize $E[H(X(B), Y(B))]$

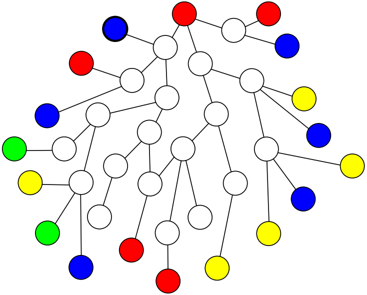




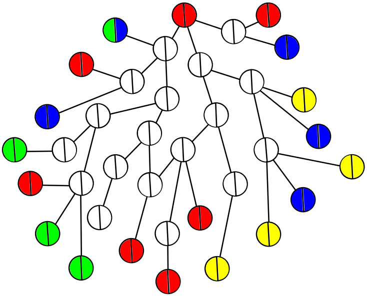
... more concretely

- $\forall B \in \mathcal{B}$ consider arbitrary $\sigma(\partial B)$ and $\tau(\partial B)$ s.t. $H(\sigma(\partial B), \tau(\partial B)) = 1$
- Take $X(B) \sim \mu(\cdot|\sigma(\partial B))$ and $Y(B) \sim \mu(\cdot|\tau(\partial B))$
- Couple $X(B)$ and $Y(B)$ so as minimize $E[H(X(B), Y(B))]$
- We should have sufficiently small $E[H(X(B), Y(B))]$



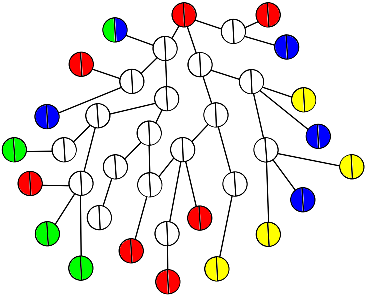


The coupling of $X(B)$ and $Y(B)$



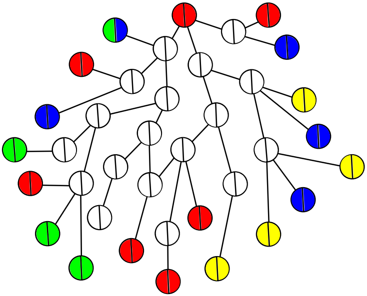
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time



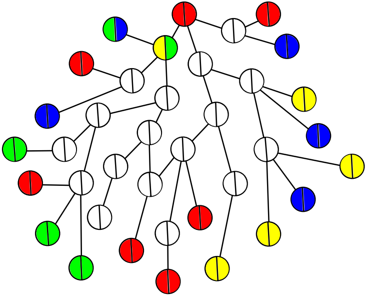
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



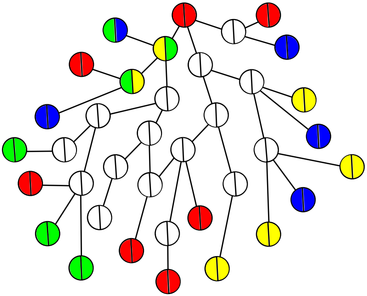
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



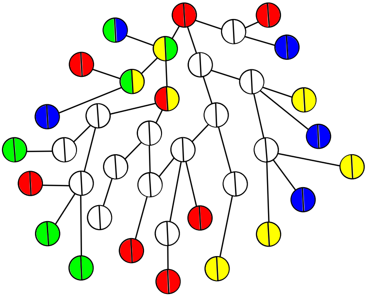
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



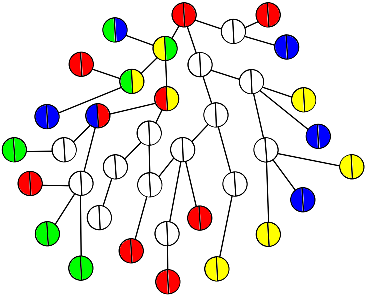
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



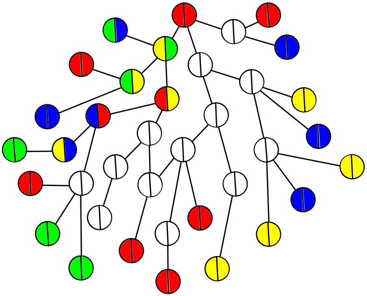
The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



The coupling of $X(B)$ and $Y(B)$

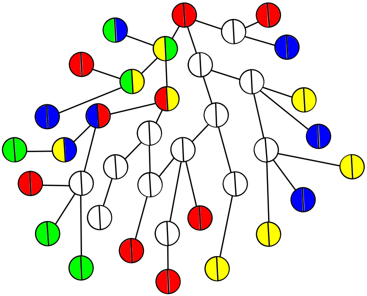
- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

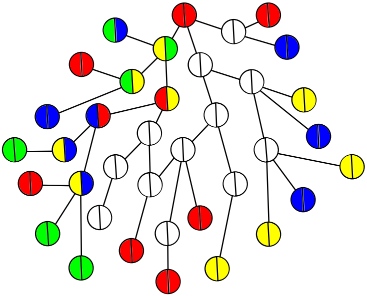
$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

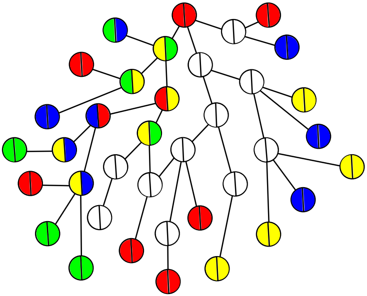
$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

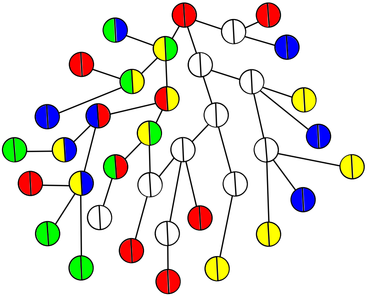
$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

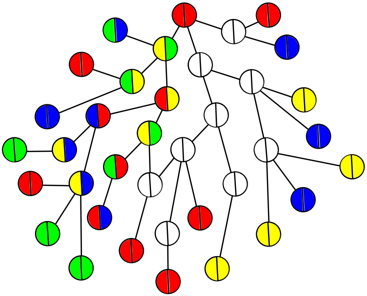
$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

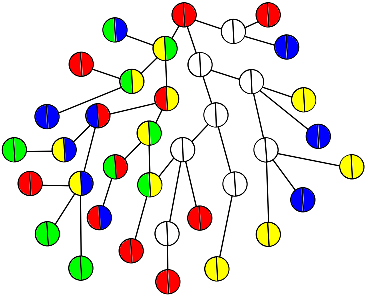
$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$



The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

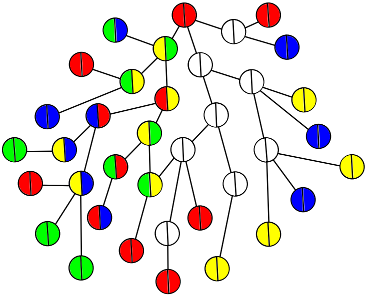


The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- disagreements graph is connected

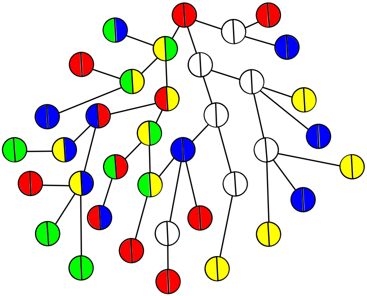


The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- disagreements graph is connected

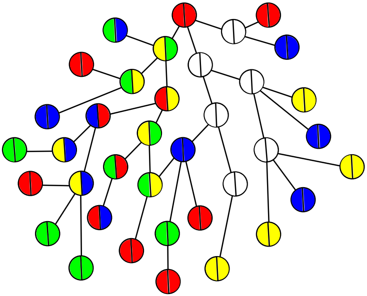


The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k-deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- disagreements graph is connected

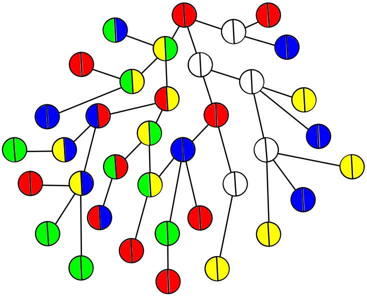


The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k-deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- disagreements graph is connected

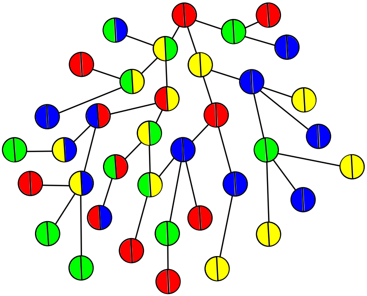


The coupling of $X(B)$ and $Y(B)$

- couple $X(B)$ and $Y(B)$ one vertex at a time
- every time pick a vertex next to disagreement
- each vertex is disagreeing with probability at most

$$\varrho_v \leq \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- disagreements graph is connected



Reduction to Independent Process, [DFFV'05]

Reduction to Independent Process, [DFFV'05]

Disagreement Percolation [van de Berg, Maes, '94]

Reduction to Independent Process, [DFFV'05]

Disagreement Percolation [van de Berg, Maes, '94]

- Product measure $\mathcal{P} : \{\text{"agree"}, \text{" disagree"}\}^B \to [0,1]$ s.t. $\forall v \in B$

$$\mathcal{P}(v := \text{"disagree "}) = \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

Reduction to Independent Process, [DFFV'05]

Disagreement Percolation [van de Berg, Maes, '94]

- Product measure $\mathcal{P} : \{\text{“agree”}, \text{“ disagree”}\}^B \rightarrow [0,1]$ s.t. $\forall v \in B$

$$\mathcal{P}(v := \text{“disagree ”}) = \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- *Path of disagreement* is a path with all of its vertices disagreeing

Reduction to Independent Process, [DFFV'05]

Disagreement Percolation [van de Berg, Maes, '94]

- Product measure $\mathcal{P} : \{\text{“agree”}, \text{“ disagree”}\}^B \to [0,1]$ s.t. $\forall v \in B$

$$\mathcal{P}(v := \text{“disagree ”}) = \begin{cases} \frac{2}{k - deg_v} & deg_v \leq k-2 \\ 1 & \text{otherwise} \end{cases}$$

- *Path of disagreement* is a path with all of its vertices disagreeing
- It holds that

$$\max_{\sigma(\partial B), \tau(\partial B)} E[H(X(B), Y(B))] \leq \sum_{l \in \mathbb{L}} \mathcal{P}(l \text{ is a path of disagreement})$$

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.
  - The root r is the *vertex next to the disagreement* of the boundary

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.
    - The root r is the *vertex next to the disagreement* of the boundary
    - $\forall w \in B$, if w and r are connected with a path of length i inside B, then w belongs to level i of T

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.
  - The root r is the *vertex next to the disagreement* of the boundary
  - $\forall w \in B$, if w and r are connected with a path of length i inside B, then w belongs to level i of T
- L_i^T is the expected number of paths of disagreement in T from the root to level i, (probabilities are w.r.t. measure $\mathcal{P}$)

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.
  - The root r is the *vertex next to the disagreement* of the boundary
  - $\forall w \in B$, if w and r are connected with a path of length i inside B, then w belongs to level i of T
- L_i^T is the expected number of paths of disagreement in T from the root to level i, (probabilities are w.r.t. measure $\mathcal{P}$)
- We will need that

$$L_i^T \leq c(1-\delta)^i \qquad i \geq 0.$$

Bounding the Expected # Paths of Disagreements (I)

Expected number of Paths of disagreements

- We care only for *self-avoiding paths* in each block B.
- Let T be a tree of self-avoiding paths.
  - The root r is the *vertex next to the disagreement* of the boundary
  - $\forall w \in B$, if w and r are connected with a path of length i inside B, then w belongs to level i of T
- L_i^T is the expected number of paths of disagreement in T from the root to level i, (probabilities are w.r.t. measure $\mathcal{P}$)
- We will need that

$$L_i^T \leq c(1-\delta)^i \qquad i \geq 0.$$

- If the root is of degree s, the condition reduces to the subtrees of r

$$L_{i-1}^{T'} \leq c(1-\delta)^i/(s \cdot \varrho_{root})$$

Bounding the Expected # Paths of Disagreements (II)

Expected number of Paths of disagreements

- Unfold down to level i of T. For every w at level i of T we have

$$L_0^w \leq \frac{c(1-\delta)^i}{\prod_x deg_x \cdot \varrho_x}$$

Bounding the Expected # Paths of Disagreements (II)

Expected number of Paths of disagreements

- Unfold down to level i of T. For every w at level i of T we have

$$L_0^w \leq \frac{c(1-\delta)^i}{\prod_x deg_x \cdot \varrho_x}$$

- ... then, we know that the expected number of disagreements is ϱ_w. That is,

$$\varrho_w \leq \frac{c(1-\delta)^i}{\prod_x deg_x \cdot \varrho_x}$$

Bounding the Expected # Paths of Disagreements (II)

Expected number of Paths of disagreements

- Unfold down to level i of T. For every w at level i of T we have

$$L_0^w \leq \frac{c(1-\delta)^i}{\prod_x deg_x \cdot \varrho_x}$$

- ... then, we know that the expected number of disagreements is ϱ_w. That is,

$$\varrho_w \leq \frac{c(1-\delta)^i}{\prod_x deg_x \cdot \varrho_x}$$

- Using appropriate parameters for the weighting schema as well as appropriate k (or λ) the above condition is satisfied.

The sampling algorithm ...

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
- The implementation of the updates

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
- The implementation of the updates
- Algorithms which provide initial configurations for both chains

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
  - This takes polynomial time.
- The implementation of the updates
- Algorithms which provide initial configurations for both chains

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
  - This takes polynomial time.
- The implementation of the updates
  - W.h.p. over $G(n, d/n)$ every $B \in \mathcal{B}$ is a tree with at most one extra edge
- Algorithms which provide initial configurations for both chains

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
  - This takes polynomial time.
- The implementation of the updates
  - W.h.p. over $G(n, d/n)$ every $B \in \mathcal{B}$ is a tree with at most one extra edge
  - ... for such blocks we can use standard algorithms
- Algorithms which provide initial configurations for both chains

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
  - This takes polynomial time.
- The implementation of the updates
  - W.h.p. over $G(n, d/n)$ every $B \in \mathcal{B}$ is a tree with at most one extra edge
  - ... for such blocks we can use standard algorithms
- Algorithms which provide initial configurations for both chains
  - Simple greedy algorithm is sufficient for colouring

The sampling algorithm ...

For the sampling algorithm we also need to have efficient...

- Construction of $\mathcal{B}$
  - This takes polynomial time.
- The implementation of the updates
  - W.h.p. over $G(n, d/n)$ every $B \in \mathcal{B}$ is a tree with at most one extra edge
  - ... for such blocks we can use standard algorithms
- Algorithms which provide initial configurations for both chains
  - Simple greedy algorithm is sufficient for colouring
  - We can start from the empty independent set

Concluding Remarks

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks
  - Weighting Schema

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks
  - Weighting Schema
- Is it possible to prove rapid mixing with site updates?

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks
  - Weighting Schema
- Is it possible to prove rapid mixing with site updates?
  - Comparison techniques

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks
  - Weighting Schema
- Is it possible to prove rapid mixing with site updates?
  - Comparison techniques
- How can we improve the bounds?

Concluding Remarks

- Considered Glauber Block Dynamics for sampling "k-colouring" and "hard-cores" with underlying graph $G(n, d/n)$
  - The structure of the blocks is simple
  - For the colourings we need to have $k \geq \frac{11}{2}d$ for rapid mixing
    - the lower bound on k is off by a factor $\frac{11}{2}$
  - For the "hard-cores" we need to have $\lambda \leq \frac{1-\epsilon}{2d}$
    - λ if off by a factor of $\frac{2}{e}$
- We introduced a new technique of creating blocks
  - Weighting Schema
- Is it possible to prove rapid mixing with site updates?
  - Comparison techniques
- How can we improve the bounds?
  - We will need to speak about "spatial mixing"

The End

THANK YOU!