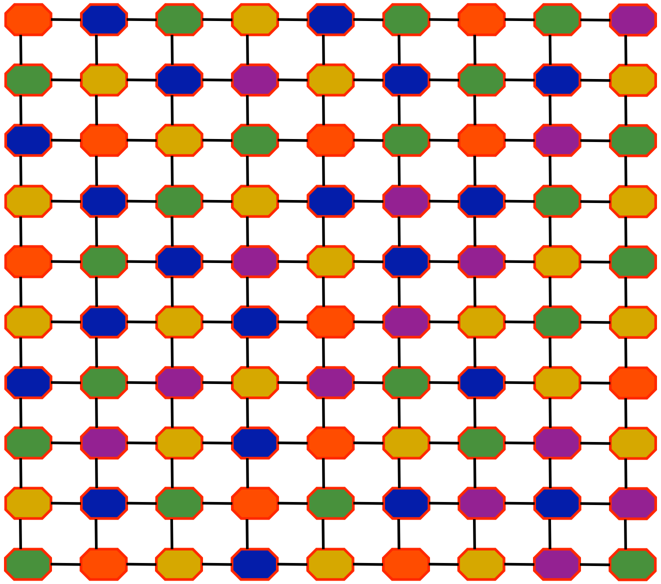
Rapid mixing of Gibbs sampling on graphs that are sparse on average

Allan Sly
Elchanan Mossel
UC Berkeley

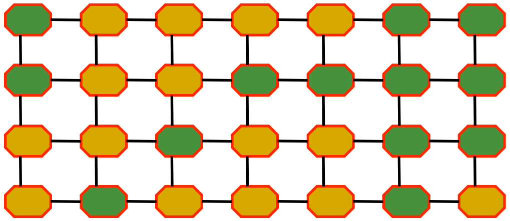
Colouring Model

- Graphs: $G = (V,E)$
- Set of colours A.
- Model = probability distribution over configurations A^V:
  - $P(\sigma) = Z^{-1} \prod_{(u,\, v)\in E} 1(\sigma_u \neq \sigma_v)$

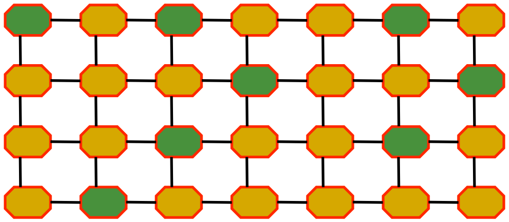


Other Models of Interest

- The theory will apply to other Gibbs measures
- Ising Model ($A = \{-1,1\}$)
  - $P(\sigma) = Z^{-1} \exp(\beta \sum_{(u, v) \in E} \sigma_u \sigma_v)$



- Hardcore Model (Independent Sets) ($A = \{0,1\}$)
  - $P(\sigma) = Z^{-1} \lambda^{\Sigma \sigma_u} \Pi_{(u, v) \in E} 1(1 \neq \sigma_u \sigma_v)$



Temporal Mixing-MCMC

- Glauber Dynamics (also Gibbs sampler, heat bath)
- A Markov process on states σ defined as follows
  - Update the state at v according to the distribution

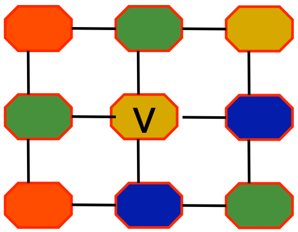
$$P(\sigma_v \mid \sigma_{V-\{v\}}) \text{ at rate 1.}$$

Temporal Mixing–MCMC

- Glauber Dynamics (also Gibbs sampler, heat bath)
- A Markov process on states σ defined as follows
  - Update the state at v according to the distribution

$$P(\sigma_v \mid \sigma_{V-\{v\}}) \text{ at rate 1.}$$

Colouring Model



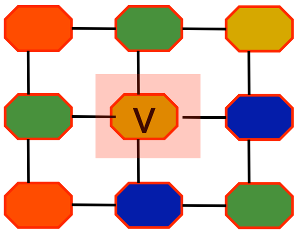
To update v, pick colour randomly from the permissible colours.

Temporal Mixing–MCMC

- Glauber Dynamics (also Gibbs sampler, heat bath)
- A Markov process on states σ defined as follows
  - Update the state at v according to the distribution

$$P(\sigma_v \mid \sigma_{V-\{v\}}) \text{ at rate } 1.$$

Colouring Model



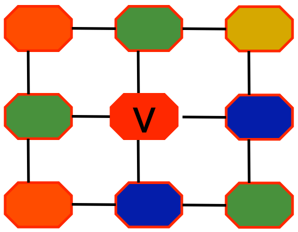
To update v, pick colour randomly from the permissible colours.

Temporal Mixing–MCMC

- Glauber Dynamics (also Gibbs sampler, heat bath)
- A Markov process on states σ defined as follows
  - Update the state at v according to the distribution

$$P(\sigma_v \mid \sigma_{V-\{v\}}) \text{ at rate 1.}$$

Colouring Model



To update v, pick colour randomly from the permissible colours.

Temporal Mixing–Questions

- The performance of the algorithm is measured by the *mixing time*
  - $\tau_{mix} := \max_{\sigma} \inf\{t: d_{TV}(X_t, P) < e^{-1}, X_0 = \sigma\}$

- When is the mixing time/relaxation time polynomial in n, the number of vertices?

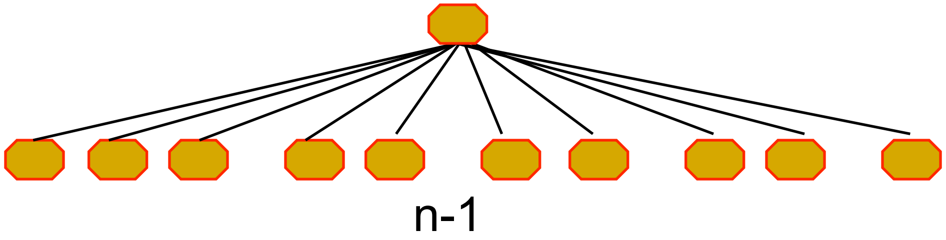
General Results

- For maximum degree Δ, using a range of sophisticated coupling arguments:
  - [Jerrum 95] $q>2\Delta$.
  - [Vigoda 99] $q>11\Delta/6$
  - [Dyer, Frieze 03] $q > 1.76\Delta$, $\Delta = \Omega(\log n)$, large girth.
  - [Molloy 04] $q > 1.49\Delta$, $\Delta = \Omega(\log n)$, large girth.
  - [Hayes and Vigoda 03] $q > (1+\varepsilon)\Delta$ if $\Delta = \Omega(\log n)$, girth > 9.
  - [DFHV 04] $q > \max\{1.49\Delta, q_0\}$ girth at least 6.
  - [HVV 07] $q > \Omega(\Delta/\log \Delta)$ on planar graphs.
  - Conjecture: $q>1+\Delta$ on any graph.
- All these results are based on the ***maximum*** degree of the graph.

Can we use average degree

- Question: Instead of the maximum degree can we prove results in terms of the average degree?

Answer: In general no. Maybe no colourings, may not be ergodic. Take a star with n vertices



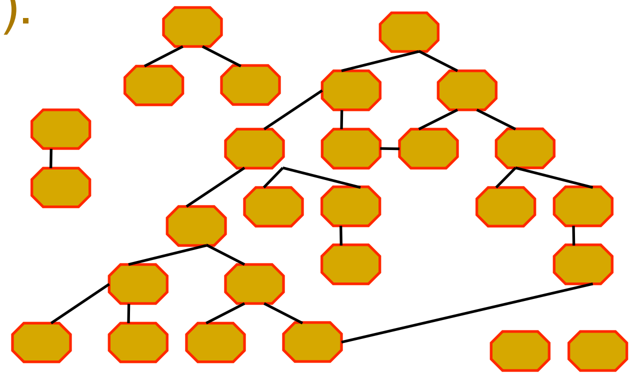
Mixing time $\exp(\Omega(n))$

This graph has slow mixing, we need extra conditions.

Random Graphs

- Erdos-Renyi random graphs are typical graphs of average degree d. The Erdos-Renyi random graph $G(n,d/n)$ is constructed from n vertices with each edge added with probability d/n.
- Standard properties of $G(n,d/n)$ asymptotically almost surely.
  - When $d>1$, G has a giant component.
  - Locally tree-like (locally a GW branching process).
  - The subgraph $B(v, (\frac{1}{2}-\varepsilon)\log_d n) = \{u \in G : d(u,v) \le (\frac{1}{2}-\varepsilon)\log_d n\}$ is a tree plus a bounded number of edges.
  - Maximum degree $O(\log n/\log\log n)$.

[Dyer Flaxman Frieze Vigoda 06] Rapid Mixing of colourings with $O(\log\log n/\log\log\log n)$ colours.



Glauber Dynamics of the Colouring model on ER Random

- [DFFV 06] asked the question, does the Glauber dynamics on G(n,d/n) mix in polynomial time for large enough but constant number of colours?

- Mixing time is greater than npolylog(n).
- Main technical hurdle: Nodes of degree O(log n/ log log n).

- **Is it polynomial in n for any q(d)?**

Our Results

Theorem [Mossel, S. 08]

When $q(d)$ is sufficiently large the Glauber dynamics on q-colourings mixes in poly(n) time on asymptotically almost every graph $G(n,d/n)$.

Similar results apply for the Ising model and the hardcore model.

Related Result

- [Spirakis, Efthymiou 08] showed that a belief propagation algorithm approximately samples random colurings on ER random graphs when q(d) is sufficiently large.

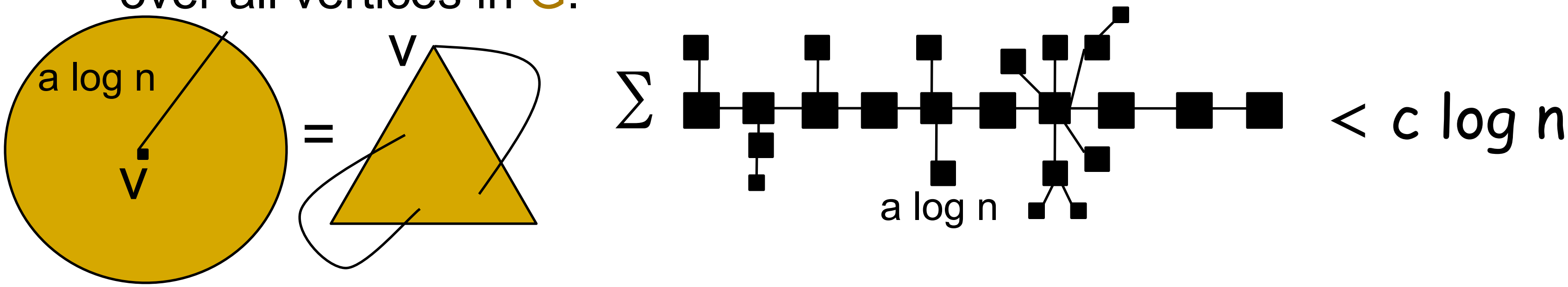
Properties of G(n,d/n) used:

- For $0 < \alpha < 1$ and a graph G define

  $m_\alpha(s)=\max_\Gamma \Sigma_{\{u \in\Gamma\}} \Sigma_{\{v \in G\}} \alpha^{d(v,u)}$

  over all non-self intersecting paths of length at most s.

Let $t(r) = \max_v$ #edges(B(v,r)) - #vertices(B(v,r)) + 1 where the max is over all vertices in G.



Theorem [Mossel, S. 08]

If $t(a \log n) \leq k$ and $m_\alpha (a \log n) \leq c \log n$ then for sufficiently many colours $q(a,k,c,\alpha)$ the Glauber dynamics mixes rapidly.

Proof (Divide and Conquer)

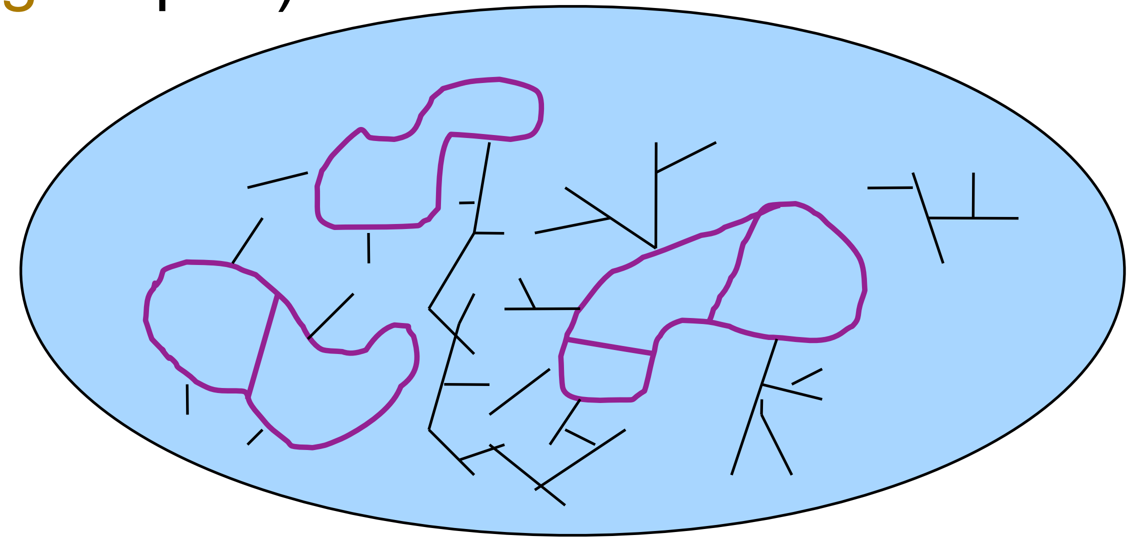
- Main strategy is to partition the graph into blocks with good properties.
- We call a vertex u (α,ε)-good if

  $\Sigma_{v \in G} \alpha^{d(u,v)} < \varepsilon$

- We can partition the graph into blocks $\{V_i\}$ so that all boundary points of the V_i are good.
  - Two bad vertices are in the same block if there is a path joining them that does not contain two consecutive good vertices.
  - For large enough ε, block diameters $< L \log n$.

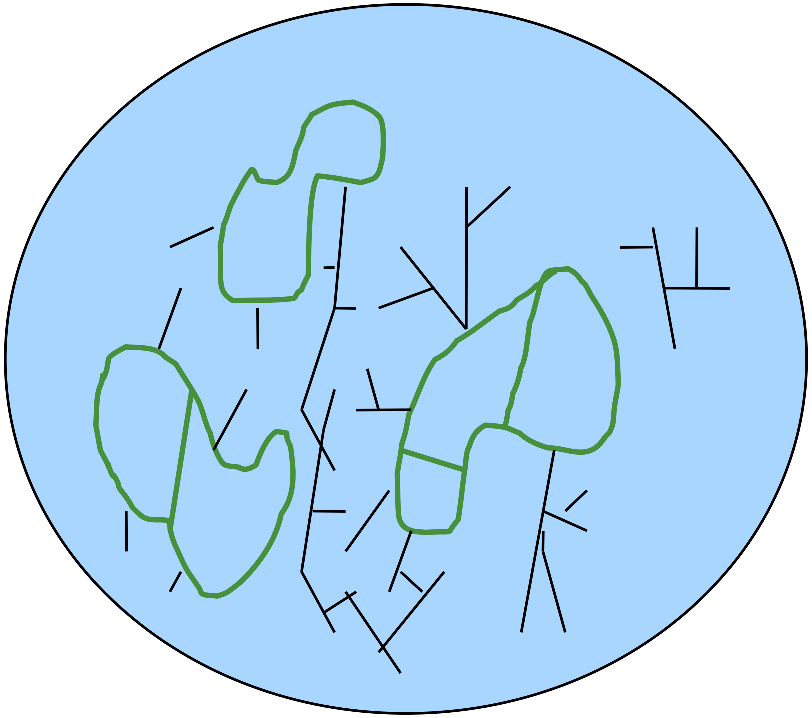
Blocks continued

- We can construct subsets of V which we call "skeletons", subgraphs of degree at most 2k and size at most 5kL log n, containing all small cycles (5L log n) so that skeletons are well divided (at least 5L log n apart).



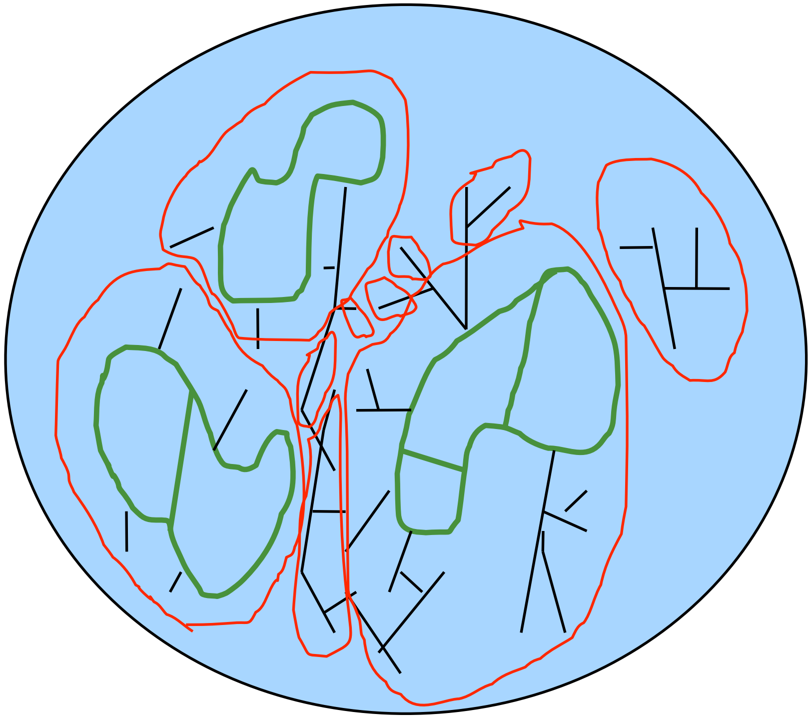
Blocks continued

- We can partition the graph into blocks such that:
  - Each block is either
    - A tree.
    - A tree plus at most κ edges and any internal cycle is at least distance $L \log n$ from the block boundary.
  - No two blocks meet at more than 1 edge.
  - Blocks have diameter at most $5\kappa L \log n$.
  - Boundary blocks are good.



Blocks continued

- We can partition the graph into blocks such that:
  - Each block is either
    - A tree.
    - A tree plus at most κ edges and any internal cycle is at least distance L log n from the block boundary.
  - No two blocks meet at more than 1 edge.
  - Blocks have diameter at most 5κL log n.
  - Boundary blocks are good.



Small boundary effect

- With sufficiently many colours, the boundary of the block does not have a strong effect on the marginal distribution at a vertex inside a block.
- For any block W and vertex u in W we have
  - $\sup \sum_j |P(\sigma_u = j | \sigma_{V-W}) - 1/q| < \beta$ for arbitrarily small $\beta > 0$.

Block Dynamics

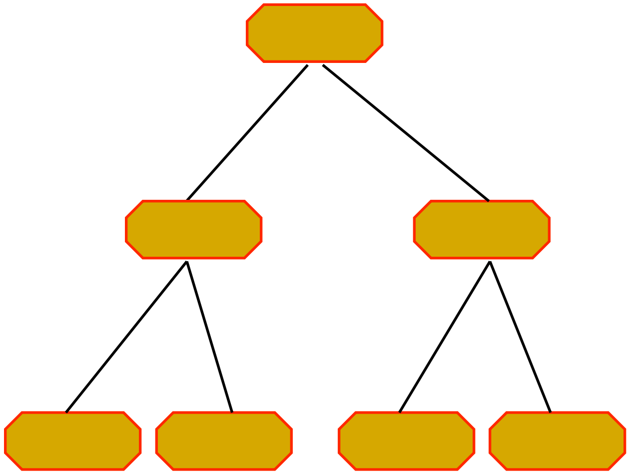
- Block Dynamics
  - Partition V into subsets $\{V_i\}$.
  - Define the *block dynamics* as the new Markov chain where a block is chosen uniformly at random and updated according to the distribution $P(\sigma_{V_i} \mid \sigma_{V-V_i})$.
- Fact:
  - If τ'_{block} is the relaxation time of the block dynamics and τ'_i is the relaxation time of the Glauber dynamics on the block V_i then

$$\tau' \leq \tau'_{block} \max_i \tau'_i$$

Relaxation Time for Trees

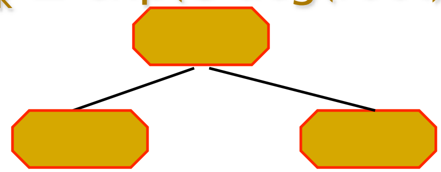
- Let $M(T) = \max \Sigma_{u \in \Gamma} \deg(u)$ over all paths Γ in T from the root.
  - Split into blocks {v}, and each of the sub-trees from V.
  - By induction using block dynamics

$\tau' \leq \exp(C\, M(T))$



Block Dynamics equivalent to "a Glauber dynamics" on the root and its neighbors.

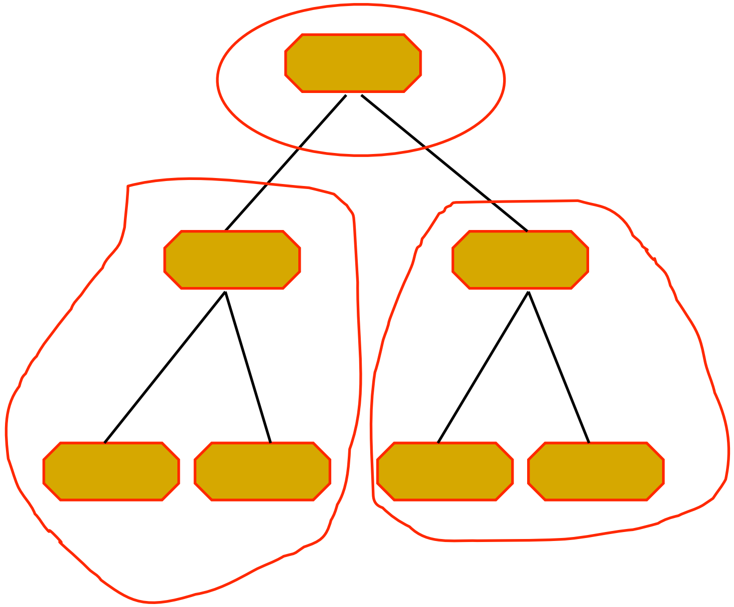
$t'_{block} \leq \exp(C \deg(root))$



Relaxation Time for Trees

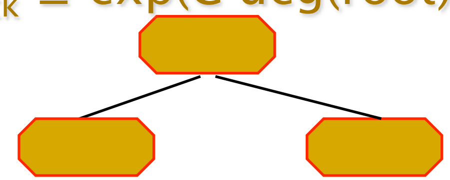
- Let $M(T) = \max \Sigma_{u \in \Gamma} \deg(u)$ over all paths Γ in T from the root.
  - Split into blocks {v}, and each of the sub-trees from V.
  - By induction using block dynamics

$\tau' \leq \exp(C\, M(T))$



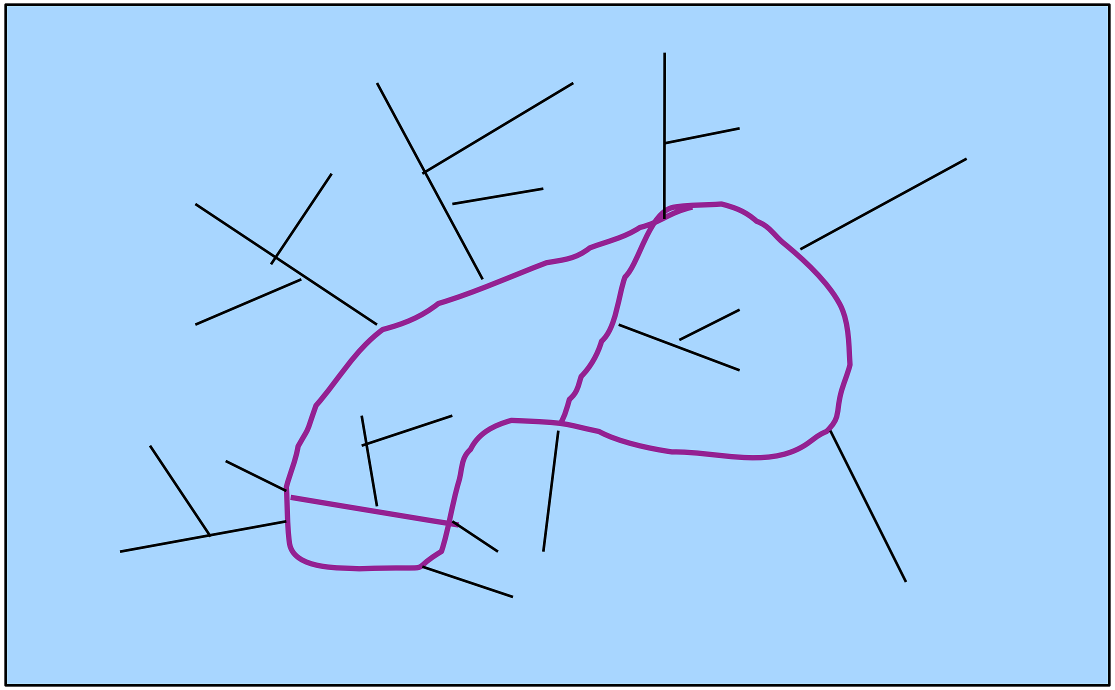
Block Dynamics equivalent to "a Glauber dynamics" on the root and its neighbors.

$t'_{block} \leq \exp(C \deg(root))$



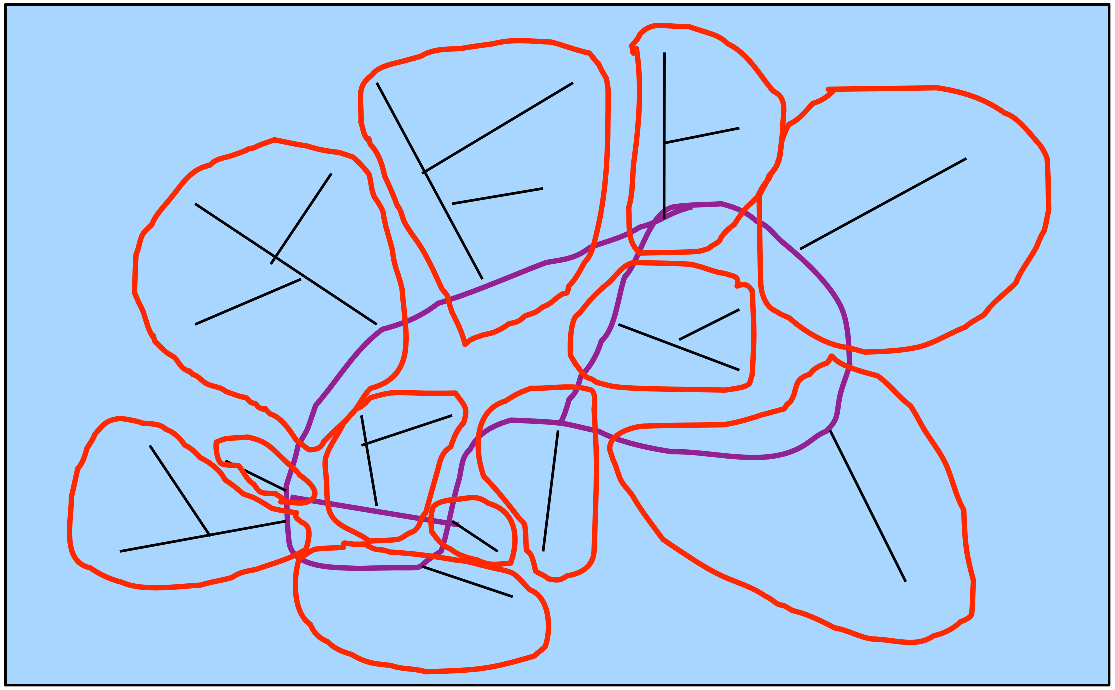
Relaxation Time for Blocks

- Divide the block into sub-blocks, each a tree containing a point of the skeleton.
- Apply Block Dynamics Lemma.



Relaxation Time for Blocks

- Divide the block into sub-blocks, each a tree containing a point of the skeleton.
- Apply Block Dynamics Lemma.

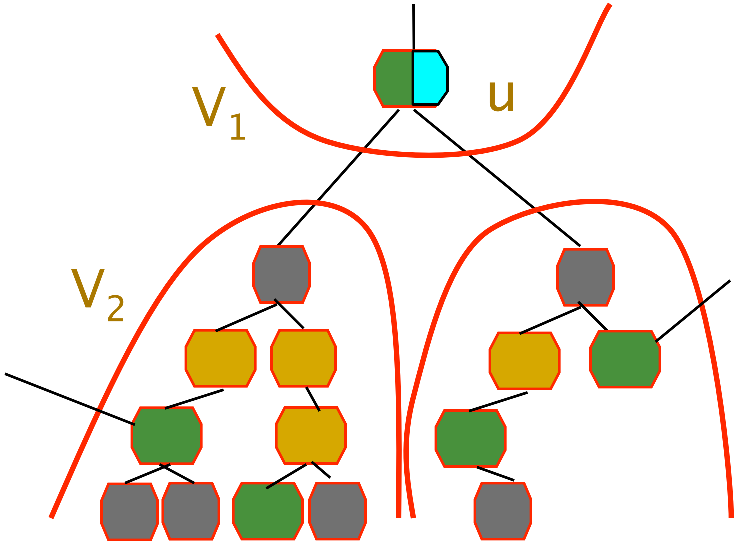


Block Dynamics

- Coupling
  - We couple so that

$E(d_H(X_1, Y_1)) \leq (1-\delta/n) d_H(X_0, Y_0) \Rightarrow \tau_{block} = O(n \log(n))$

  - By path coupling we only need consider the case $d_H(X_0, Y_0) = 1$.



Number of new disagreements when V_2 is updated is bounded by

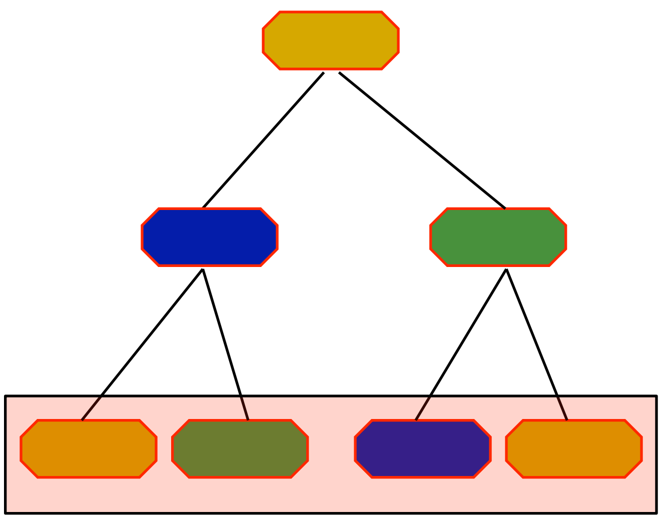
$$\Sigma_{v \in V_2} \beta^{d(u,v)}$$

$$< \beta/\alpha \, \Sigma_{v \in G} \alpha^{d(u,v)}$$

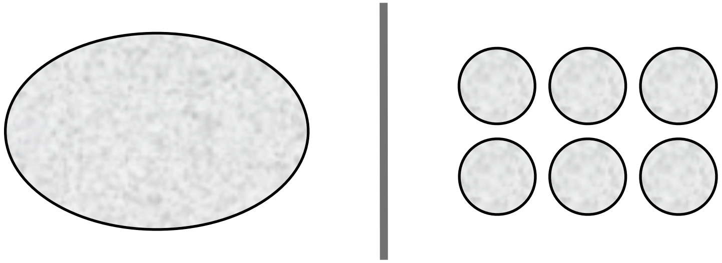
$$< \beta\varepsilon/\alpha$$

Threshold for Polynomial Mixing

- It is conjectured that the threshold for polynomial mixing coincides with the reconstruction threshold.



- <u>Conjecture (Stat Phys):</u> Space of colourings becomes clustered at reconstruction threshold.
- Glauber Dynamics no longer ergodic.



Open Problem

- For which random graph models is mixing on G(n,d/n) determined by the uniqueness or reconstruction threshold on the tree?