Rapid Mixing of Subset Glauber Dynamics on Graphs of Bounded Tree-Width

Ross J. Kang Magnus Bordewich

School of Engineering and Computing Sciences, Durham University

7 July 2011
ICALP 2011, Zürich

Counting complexity and #BIS

- P v NP-complete
  - Maximum independent set (MIS) is NP-hard.
  - Maximum independent set in a bipartite graph (MBIS) is in P.
- FP v #P-complete
  - Counting spanning trees is in FP (Kirchhoff, 19C).
  - Counting independent sets in a bipartite graph (#BIS) is #P-complete.

Approximate counting complexity

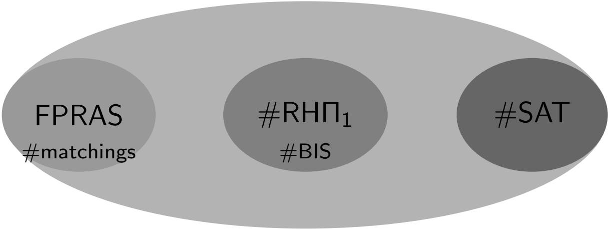
Another viewpoint on #P w.r.t. randomised approximation.

A *randomised approximation scheme* for a function $f : \Sigma^* \to \mathbb{N}$ is a probabilistic Turing machine that takes as input a pair $(x, \varepsilon) \in \Sigma^* \times (0, 1)$ and produces as output an integer random variable Y satisfying the condition $\Pr(e^{-\varepsilon} \le Y/f(x) \le e^{\varepsilon}) \ge 3/4$. It is *fully polynomial* (FPRAS) if it runs in time polynomial in both $|x|$ and ε^{-1}.

In the AP-reducibility framework developed by Dyer, Goldberg, Greenhill and Jerrum (2004), the "FPRASable" class takes on the role of FP (though it clearly contains FP).

Approximate counting complexity and #BIS

- FPRASable v AP-interreducible with #SAT
  - A counting problem that corresponds to an NP-complete decision problem must be #P-complete with respect to AP-reducibility.
  - DGGJ discovered an *intermediate* complexity class[1], all AP-interreducible, denoted $\#RH\Pi_1$, represented by #BIS.



[1] Includes counting problems for downsets of a partial order, configurations in the Widom-Rowlinson model, stable matchings.

FPRASs and rapidly mixing Markov chains

Many FPRASs are obtained using *Markov chain Monte Carlo* (MCMC).

That is, by designing a suitable Markov chain, one that converges to a desired probability distribution, then proving fast convergence, we can derive a FPRAS.

(NB: This derivation is not trivial and there is a deeper link.)

The Tutte polynomial and the random cluster model

The *partition function of the random cluster model* is defined for any $G = (V, E)$ and q, μ as

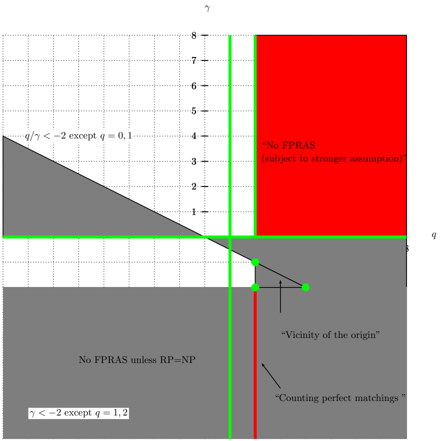
$$Z_{RC}(G; q, \gamma) := \sum_{S \subseteq E} q^{\kappa(S)} \gamma^{|S|},$$

where $\kappa(S)$ is the number of components in (V, S). $Z_{RC}(G; q, \gamma)$ is equivalent to the *Tutte polynomial*, defined for any $G = (V, E)$ and x, y as

$$T(G; x, y) := \sum_{S \subseteq E} (x-1)^{r(E)-r(S)} (y-1)^{|S|-r(S)},$$

where $r(S)$ is the $\mathbb{F}_2$-rank of incidence matrix for (V, S).

The Tutte polynomial and the random cluster model



Stolen from Leslie Goldberg's Dagstuhl slides.

A general form of graph polynomial

An *edge subset expansion formula* for $\mathcal{P}$ is written as follows: for any simple graph $G = (V, E)$,

$$\mathcal{P}(G) = \sum_{S \subseteq E} w((V, S))$$

for some graph function w, where (V, S) denotes the graph with vertex set V and edge set S.

The weight function w shall be assumed to be positive — from a statistical physics viewpoint, this results in a so-called 'soft-core model'.

A general form of graph polynomial

The *partition function of the random cluster model* is defined for any $G = (V, E)$ and q, μ as

$$Z_{RC}(G; q, \gamma) := \sum_{S \subseteq E} q^{\kappa(S)} \gamma^{|S|},$$

where $\kappa(S)$ is the number of components in (V, S). $Z_{RC}(G; q, \gamma)$ is equivalent to the *Tutte polynomial*, defined for any $G = (V, E)$ and x, y as

$$T(G; x, y) := \sum_{S \subseteq E} (x-1)^{r(E)-r(S)} (y-1)^{|S|-r(S)},$$

where $r(S)$ is the $\mathbb{F}_2$-rank of incidence matrix for (V, S).

A general form of graph polynomial

The *partition function of the random cluster model* is defined for any $G = (V, E)$ and q, μ as

$$Z_{RC}(G; q, \gamma) := \sum_{S \subseteq E} q^{\kappa(S)} \gamma^{|S|},$$

where $\kappa(S)$ is the number of components in (V, S). $Z_{RC}(G; q, \gamma)$ is equivalent to the *Tutte polynomial*, defined for any $G = (V, E)$ and x, y as

$$T(G; x, y) := \sum_{S \subseteq E} (x-1)^{r(E)-r(S)} (y-1)^{|S|-r(S)},$$

where $r(S)$ is the $\mathbb{F}_2$-rank of incidence matrix for (V, S).

A general form of graph polynomial

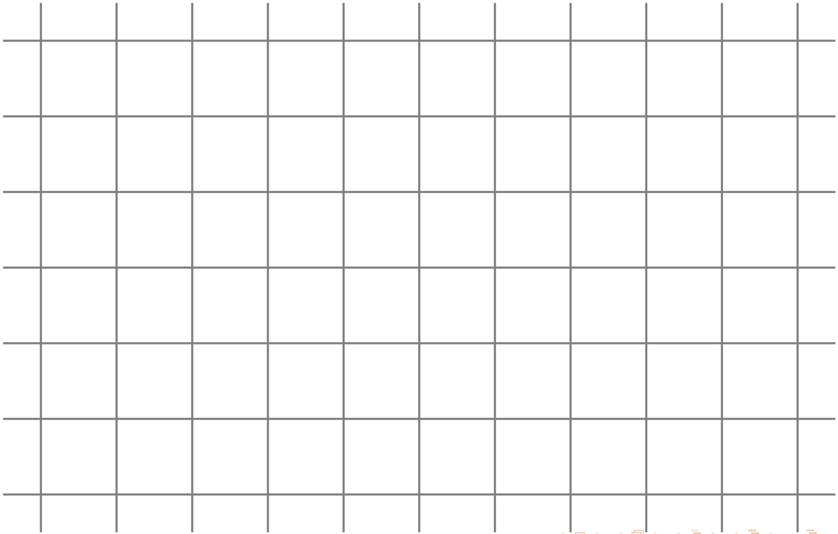
For the partition function of the random cluster model,

$$w((V,S)) = q^{\kappa(S)}\gamma^{|S|}.$$

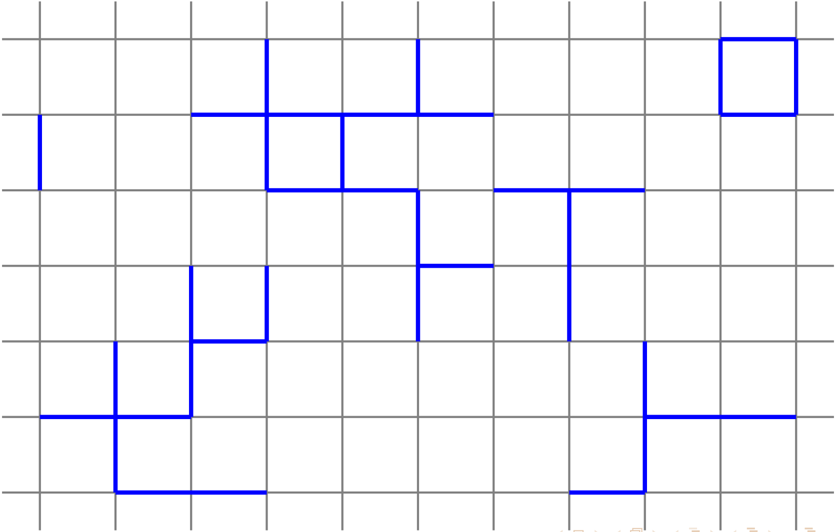
For the Tutte polynomial,

$$w((V,S)) = (x-1)^{r(E)-r(S)}(y-1)^{|S|-r(S)}.$$

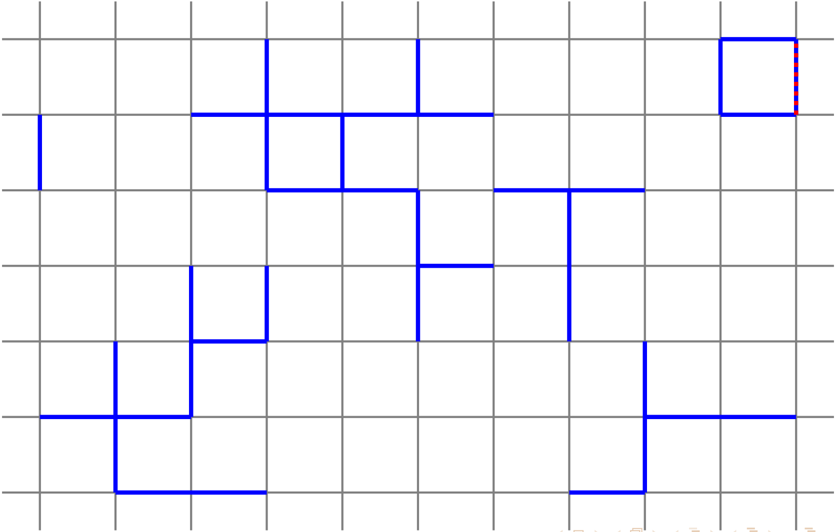
Subset Glauber dynamics



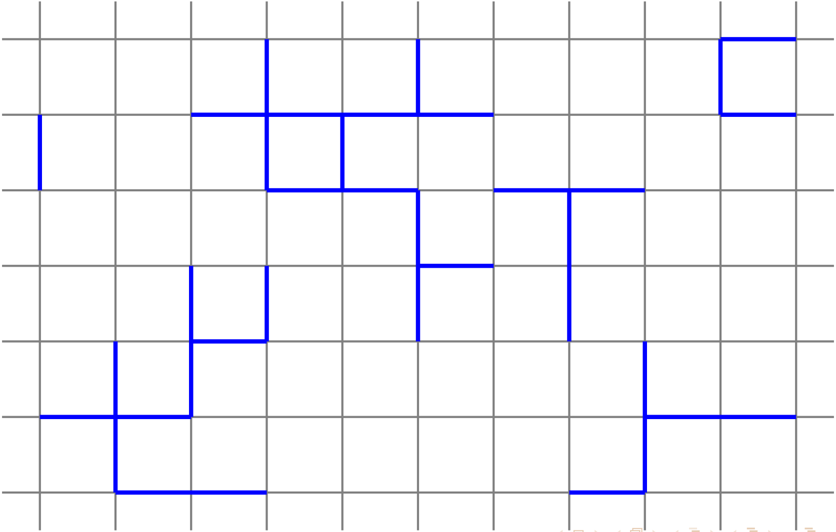
Subset Glauber dynamics



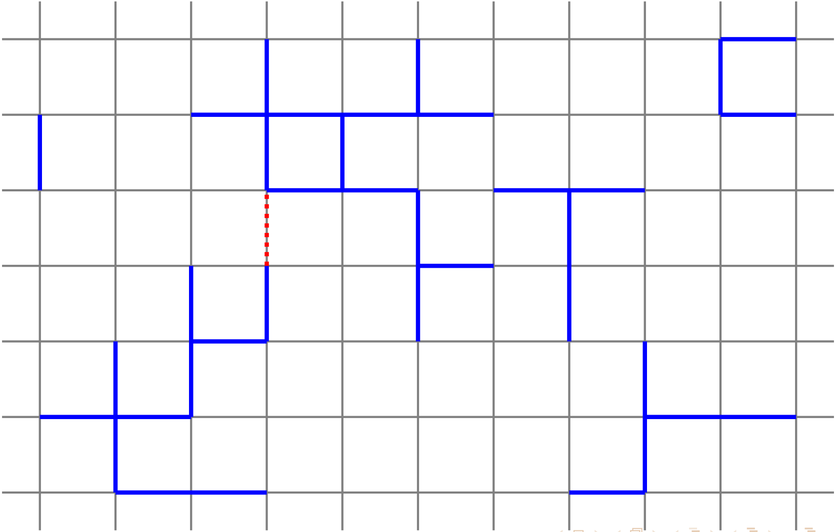
Subset Glauber dynamics



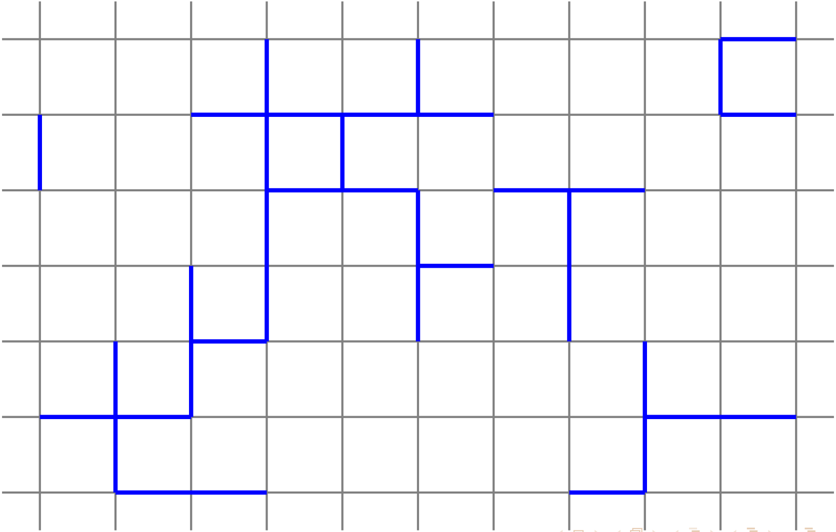
Subset Glauber dynamics



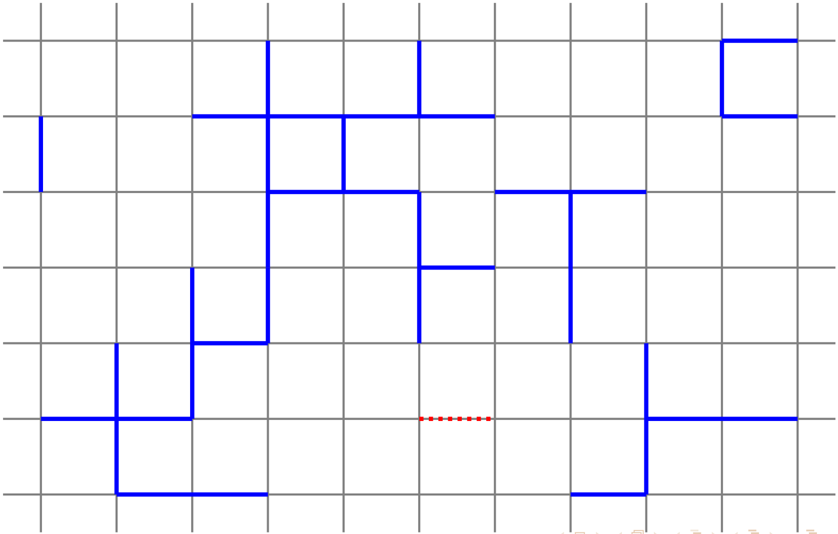
Subset Glauber dynamics



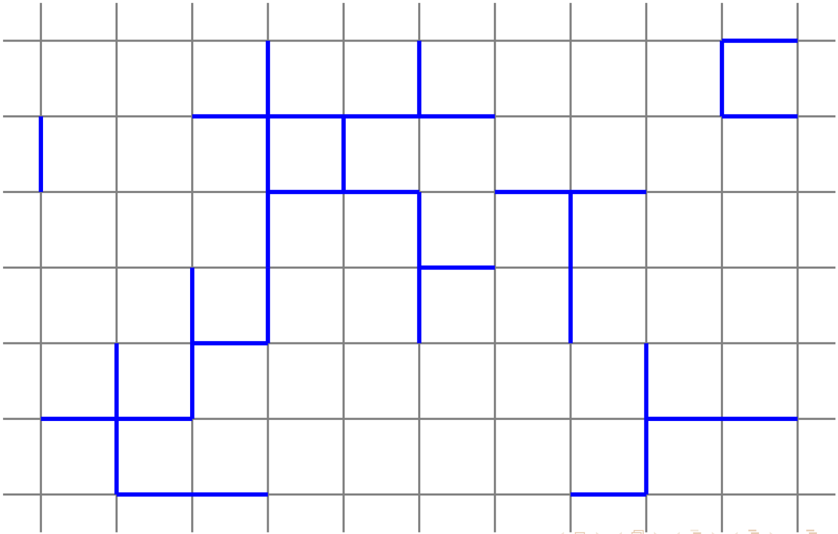
Subset Glauber dynamics



Subset Glauber dynamics



Subset Glauber dynamics



Subset Glauber dynamics

$$\mathcal{P}(G) = \sum_{S \subseteq E} w((V,S))$$

$$\Downarrow$$

Using the weighting w, we define the *single bond flip chain* $\mathcal{M} = (X_t)_{t=0}^{\infty}$. Start with arbitrary $X_0 \subseteq E$ and then repeatedly run the following.

1. Pick an edge $e \in E$ u.a.r.
2. Set $X_{t+1} = X_t \oplus \{e\}$ w.p. $\frac{1}{2} \min\left\{1, \frac{w((V,S))}{w((V,X_t))}\right\}$ (and $X_{t+1} = X_t$ o/w).

$\mathcal{M}$ is a reversible Markov chain with unique stationary distribution $\pi(S) \propto w((V,S))$.

Subset Glauber dynamics

Using the weighting w, we define the *single bond flip chain* $\mathcal{M} = (X_t)_{t=0}^{\infty}$. Start with arbitrary $X_0 \subseteq E$ and then repeatedly run the following.

1. Pick an edge $e \in E$ u.a.r.
2. Set $X_{t+1} = X_t \oplus \{e\}$ w.p. $\frac{1}{2} \min \left\{1, \frac{w((V,S))}{w((V,X_t))}\right\}$ (and $X_{t+1} = X_t$ o/w).

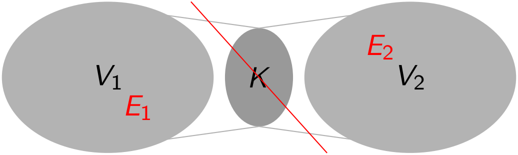
$\mathcal{M}$ is a reversible Markov chain with unique stationary distribution $\pi(S) \propto w((V,S))$.

Rapid mixing of $\mathcal{M}$

$\Downarrow$

FPRAS to compute $\mathcal{P}(G)$

A mild condition on the weight functions



For $G = (V, E)$, a partition (E_1, E_2) of E is *appropriate* for a vertex cut (V_1, K, V_2) if E_1 has no edge adjacent to a vertex in V_2 and E_2 has no edge adjacent to a vertex in V_1.

Fix $\lambda > 0$ and let $\hat{\lambda} := \max\{\lambda, 1/\lambda\}$. Then w is *λ-multiplicative*, if for any $G = (V, E)$, any vertex cut (V_1, K, V_2), any appropriate partition (E_1, E_2),

$$\hat{\lambda}^{-|K|} \leq \frac{w((V_1 \cup K, E_1))w((V_2 \cup K, E_2))}{w(G)} \leq \hat{\lambda}^{|K|}.$$

A mild condition on the weight functions

For the partition function of the random cluster model,

$$w((V,S)) = q^{\kappa(S)}\gamma^{|S|}.$$

For the Tutte polynomial,

$$w((V,S)) = (x-1)^{r(E)-r(S)}(y-1)^{|S|-r(S)}.$$

The main theorem

Theorem

Let $G = (V, E)$ *where* $|V| = n$*. If* w *is* λ*-multiplicative for some* $\lambda > 0$*, then the mixing time of* $\mathcal{M}$ *on* G *satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)|\log\lambda|}\log(1/\varepsilon)\right)$$

where $\mathrm{tw}(G)$ *denotes the tree-width of* G*.*

Canonical paths and linear-width

Proof is via a "canonical paths" argument, a method pioneered by Diaconis and Stroock (1991) and Sinclair (1992).

To define the collection of canonical paths, we use a "linear-width" edge-ordering of the base graph.

For $G = (V, E)$, an ordering $(e_1, \ldots, e_m)$ of E has *linear-width* at most ℓ if for each i there are at most ℓ vertices incident to both an edge in $\{e_1, \ldots, e_{i-1}\}$ and an edge in $\{e_i, \ldots, e_m\}$. The *linear-width* $\mathrm{lw}(G)$ of G is the least ℓ such that such an ordering of E exists.

Importantly, $\mathrm{lw} \lesssim \mathrm{tw} \log n$.

Applications

With appropriate conditions on their respective parameters, we obtain rapid mixing on the associated Glauber dynamics as well as FPRASs for the following, for graphs of bounded tree-width:

- $Z_{RC}(G; q, \gamma)$ and $T(G; x, y)$, in particular, the partition function for the ferromagnetic Potts model;
- $R_2(G; q, \mu)$, the adjacency-rank polynomial of Ge and Štefankovič (2010);
- $Z_{Tutte}(G; q, \vec{v})$, the multivariate Tutte polynomial of Sokal (2005);
- $U(G; \vec{x}, y)$, the U polynomial of Noble and Welsh (1999); and
- $q(G; x, y)$, the interlace polynomial of Arratia, Bollobás and Sorkin (2004).

Concluding remarks

- Ge and Štefankovič used this same methodology to show, one, that the MC for the random cluster model mixes rapidly for bounded tree-width and, two, that the MC for their adjacency-rank polynomial mixes rapidly upon trees.
- Although exponential mixing can occur for the adjacency-rank polynomial upon bipartite graphs, Goldberg and Jerrum (2010), it is still of interest to show/refute rapid mixing for large classes of graphs and polynomials in the framework.
- Efficient exact computation can usually be achieved upon graphs of bounded tree-width with dynamic programming.