Rapid Mixing Swendsen-Wang Sampler for Stochastic Partitioned Attractive Models

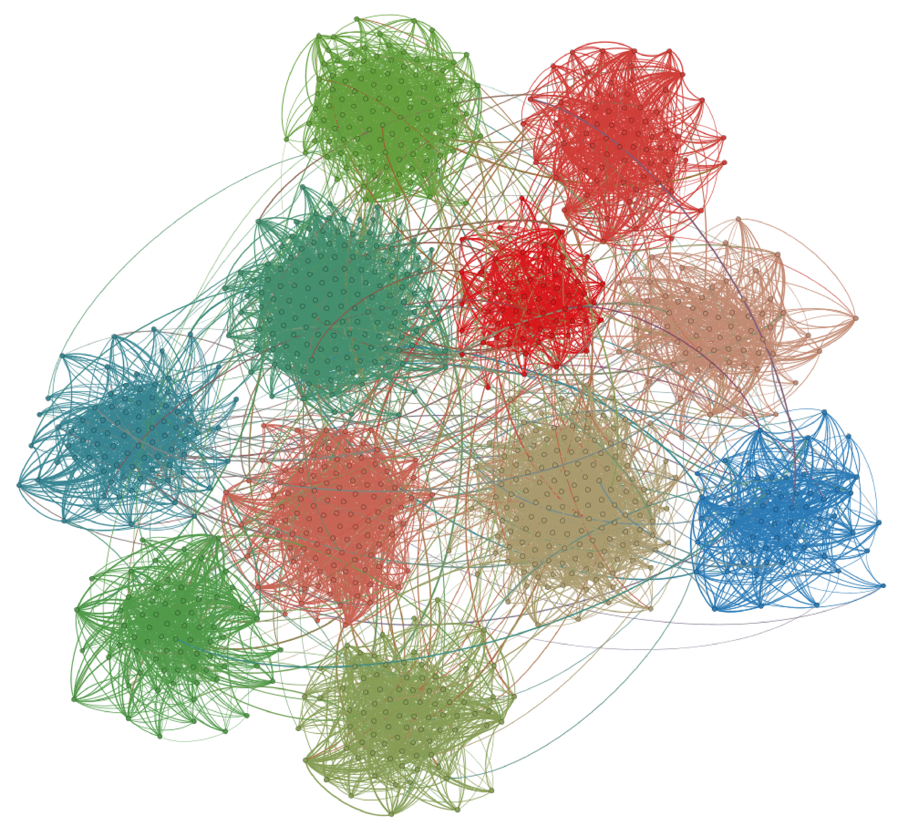
Sejun Park, Yunhun Jang, Andreas Galanis, Jinwoo Shin, Daniel Stefankovic, Eric Vigoda

Motivation and Summary

- Gibbs sampler is one of the most popular Markov chains used for learning and inference problems in graphical models. These tasks are computationally intractable in general, and Gibbs sampler often suffers from slow mixing.
- In this research, we study Swendsen-Wang dynamics (SW) which is a more sophisticated Markov chain for attractive Ising models designed to overcome bottlenecks of Gibbs sampler.
- Recently, Guo et al. proved $O(|V|^{10})$ mixing of SW for Ising models on arbitrary graphs. Rapid mixing of SW has been studied for complete graph (Peres et al., Vigoda et al.), grid graph (Ullrich) which have fully symmetric graphical structures.
- We prove $O(\log|V|)$ mixing of SW for stochastic partitioned models which is not symmetric at all. Also, we prove $O(\log|V|)$ mixing of SW for complete bipartite graphs in high temperature region.
- Experimental results empirically verifies that SW performs better than Gibbs sampler in practical tasks under various settings.

Stochastic Partitioned Model

- Stochastic partitioned model is a random graph model such that given vertices and disjoint communities of vertices $S_1,\ldots,S_r$, any edge between communities S_i and S_j exists with probability p_{ij}.
- Social graphs, Erdős-Rényi random graphs and r-partite graphs are examples of stochastic partitioned models.

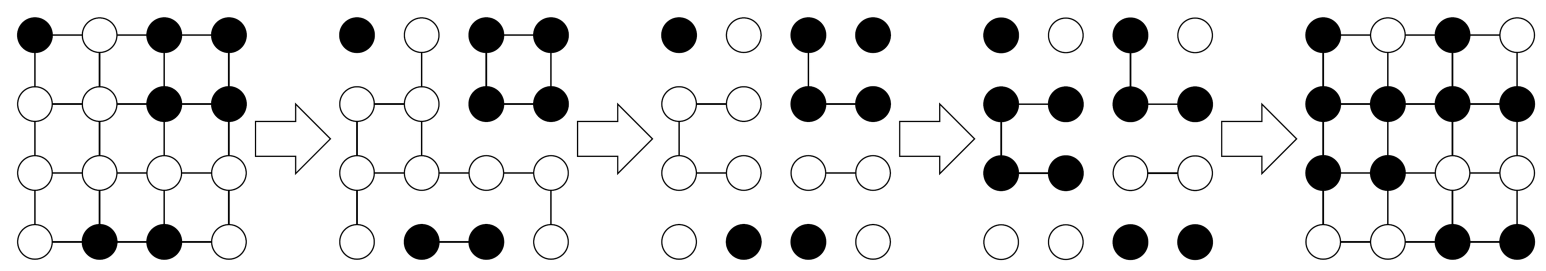


Ising Model and Swendsen-Wang Dynamics

- Given a graph (V,E), Ising model, equivalently pairwise binary model, is a joint distribution on $x\in\{-1,1\}^V$ defined as

$$P(x)\propto\exp\left(\sum_{v\in V}\gamma_v x_v+\sum_{(u,v)\in E}\beta_{uv}x_u x_v\right).$$

- Swendsen-Wang dynamics is a Markov chain running on attractive Ising model (i.e., $\beta_{uv}\geq 0$) which converges to the distribution of Ising model. The transition from X^t to X^{t+1} is as follows:



1. Delete edges between vertices of different spins in X^t.
2. For each remaining edge (u,v), delete it with probability $q(\beta_{uv})$.
3. For each connected component C, assign a spin 1 to vertices in C with probability $q(\gamma_C)$ and -1 with probability $1-q(\gamma_C)$.
4. Denote X^{t+1} as a resulting state.

Main Result

- Our main result is proving the rapid mixing of Swendsen-Wang dynamics for Ising models on stochastic partitioned graphs while Gibbs sampler suffers from exponential running time.

Theorem 1 *The mixing time of Swendsen-Wang chain on a stochastic partitioned model is $O(\log|V|)$ a.a.s. if $\gamma\geq 0$ and either (a) or (b) holds:*

(a) *For all i and for all $u,v\in S_i$, $p_{ii},\beta_{uv}=\Omega(1)$.*

(b) *For all $i\neq j$ and for all $u\in S_i, v\in S_j$, $p_{ij},\beta_{uv}=\Omega(1)$.*

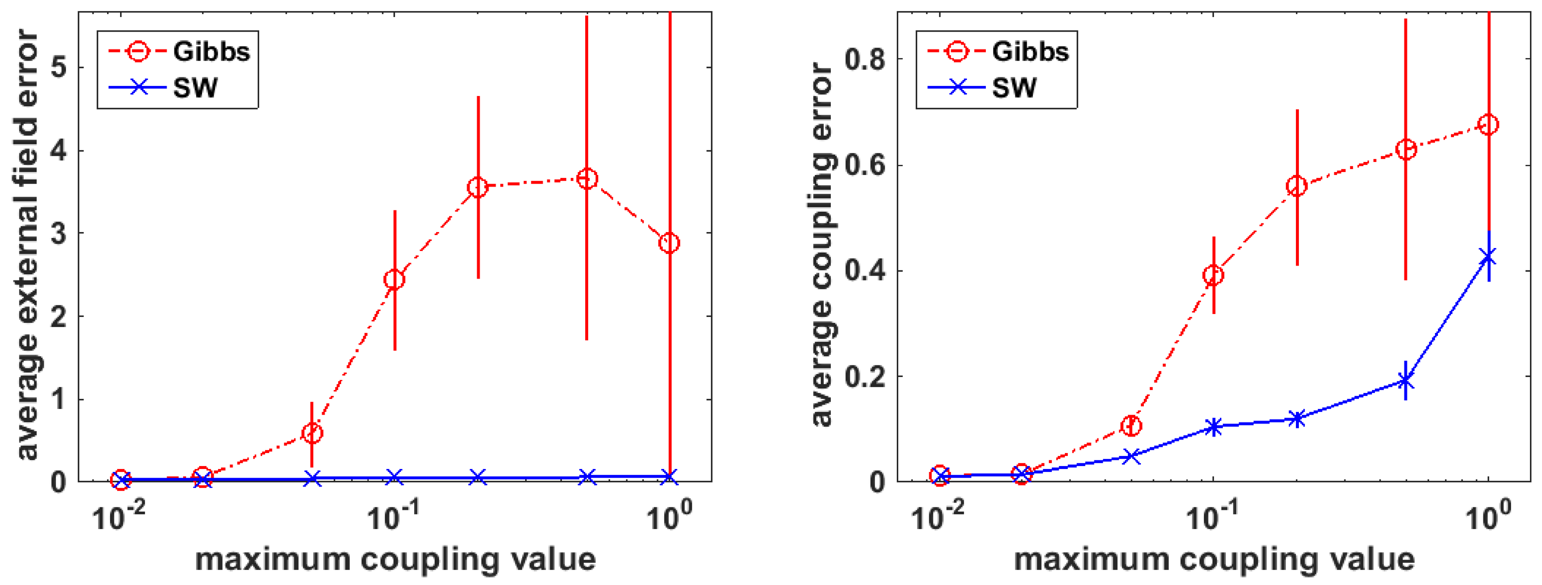
- Theorem 1 states that Swendsen-Wang dynamics rapidly mixes under mild conditions ($\beta,p=\Omega(1)$).

Theorem 2 *The mixing time of Swendsen-Wang chain on a complete bipartite graph is $O(\log|V|)$ if $\beta=k/|V|$ where $k\neq 2$ is any constant.*

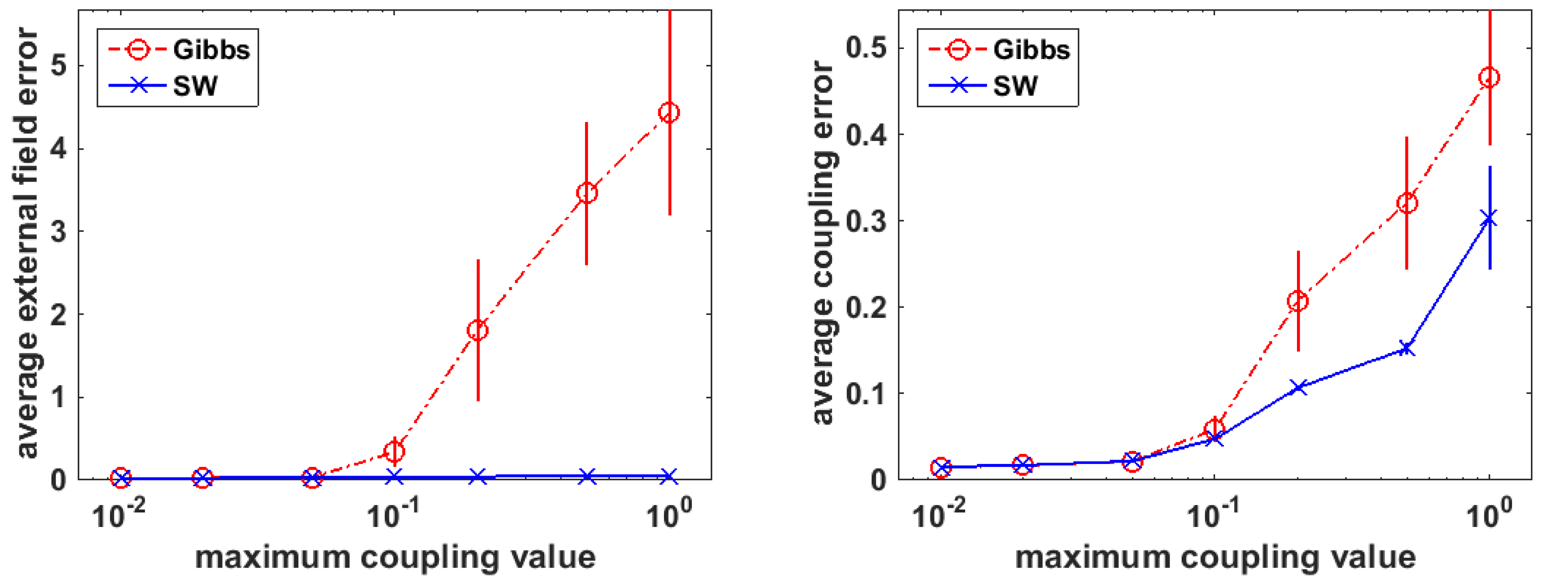
- Theorem 2 states that Swendsen-Wang dynamics on a bipartite graph rapidly mixes in the high temperature region.

Experiments

- We compare the empirical performance of Swendsen-Wang dynamics and Gibbs sampler by learning parameters of Ising models.
- Learning tasks are performed for attractive Ising models on two real world social graphs and synthetic graphs while parameters are randomly chosen from $\gamma_v\sim$ Unif$(-1,1)$ and $\beta_{uv}\sim$ Unif$(0,x)$.
- Experimental results verify that Swendsen-Wang dynamics performs better than Gibbs sampler in various practical settings.
- **Facebook Graph:** 4039 vertices and 88234 edges



- **UCI Graph:** 1899 vertices and 13838 edges



- **Synthetic Graph:** 100-1000 vertices and $\beta_{uv}\sim$ Unif$(0,1)$