Convergence of MCMC and Loopy BP in the Tree Uniqueness Region for the Hard-Core Model

Charis Efthymiou
efthymiou@gmail.com

University of Frankfurt

joint work with:

T. Hayes, D. Štefankovič, E. Vigoda, Y. Yin

Workshop on Random Instances and Phase Transitions
Simons Institute 2-6 May 2016

Hard-Core Model

Hard-Core Model

Gibbs Distribution

Hard-Core Model

Gibbs Distribution

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, for each *independent set* σ we have

$$\mu(\sigma) = \lambda^{|\sigma|}/Z,$$

Hard-Core Model

Gibbs Distribution

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, for each *independent set* σ we have

$$\mu(\sigma) = \lambda^{|\sigma|}/Z,$$

where

$$Z = \sum_{\sigma} \lambda^{|\sigma|}$$

Hard-Core Model

Gibbs Distribution

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, for each *independent set* σ we have

$$\mu(\sigma) = \lambda^{|\sigma|}/Z,$$

where

$$Z = \sum_{\sigma} \lambda^{|\sigma|}$$

$Z = Z(G, \lambda)$ is the *partition function*.

Partition function

Partition function

The problem

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, *compute* the partition function

$$Z(G, \lambda) = \sum_{\sigma} \lambda^{|\sigma|}$$

Partition function

The problem

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, *compute* the partition function

$$Z(G, \lambda) = \sum_{\sigma} \lambda^{|\sigma|}$$

- computationally hard problem

Partition function

The problem

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, *compute* the partition function

$$Z(G, \lambda) = \sum_{\sigma} \lambda^{|\sigma|}$$

- computationally hard problem
  - $\#P$-complete [Valiant 1979]

Partition function

The problem

Given a graph $G = (V, E)$ and *fugacity* $\lambda > 0$, *compute* the partition function

$$Z(G, \lambda) = \sum_{\sigma} \lambda^{|\sigma|}$$

- computationally hard problem
  - $\#P$-complete [Valiant 1979]
- focus on the approximation algorithms

Approximation Algorithms' Approach

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic
- Randomized

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic
  - Compute numerically (estimations of) the probability of a configuration
- Randomized

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic
  - Compute numerically (estimations of) the probability of a configuration
- Randomized
  - Generate Samples (approximately) Gibbs distributed

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic
  - Compute numerically (estimations of) the probability of a configuration
  - Fully Polynomial Time Approximation Scheme (FPTAS)
- Randomized
  - Generate Samples (approximately) Gibbs distributed
  - Fully Polynomial Time Randomized Approximation Scheme (FPRAS)

Approximation Algorithms' Approach

Approach

Compute estimates of the Gibbs distribution

- Deterministic
  - Compute numerically (estimations of) the probability of a configuration
  - Fully Polynomial Time Approximation Scheme (FPTAS)
    - in time $\mathrm{poly}(n)$ and $\mathrm{poly}(\epsilon^{-1})$

$$\hat{Z} \in (1 \pm \epsilon) Z(G, \lambda)$$

- Randomized
  - Generate Samples (approximately) Gibbs distributed
  - Fully Polynomial Time Randomized Approximation Scheme (FPRAS)
    - in time $\mathrm{poly}(n)$, $\mathrm{poly}(\epsilon^{-1})$ and $\mathrm{poly}(\log(\delta^{-1}))$

$$\Pr[\hat{Z} \in (1 \pm \epsilon) Z(G, \lambda)] > 1 - \delta$$

How well can we approximate

How well can we approximate

Hardness of approximation [Sly 2010]

For triangle-free Δ-regular graphs, where $\Delta \geq 3$, and for all $\lambda > \lambda_c(\Delta)$, it is NP-hard to approximate the partition function within factor $2^{\gamma n}$.

How well can we approximate

Hardness of approximation [Sly 2010]

For triangle-free Δ-regular graphs, where $\Delta \geq 3$, and for all $\lambda > \lambda_c(\Delta)$, it is NP-hard to approximate the partition function within factor $2^{\gamma n}$.

- Galanis, Ge, Stefankovic, Vigoda, Yang (2011)
- Sly, Sun (2012)
- Galanis, Stefankovic, Vigoda (2012)

How well can we approximate

Hardness of approximation [Sly 2010]

For triangle-free $\varDelta$-regular graphs, where $\varDelta \geq 3$, and for all $\lambda > \lambda_c(\varDelta)$, it is NP-hard to approximate the partition function within factor $2^{\gamma n}$.

What is $\lambda_c(\varDelta)$? [Kelly 1985]

Critical point for "uniqueness/non-uniqueness" phase transition of the hard-core model on $\varDelta$ regular trees

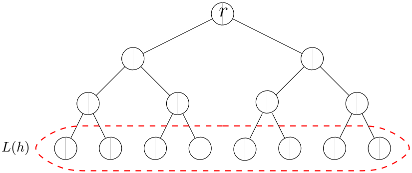
$$\lambda_c(\varDelta) := \frac{(\varDelta - 1)^{\varDelta - 1}}{(\varDelta - 2)^{\varDelta}} \sim \frac{e}{\varDelta}$$

Gibbs Uniqueness

Gibbs Uniqueness

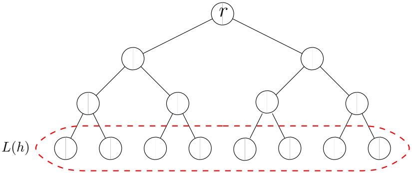
Δ-regular tree T of height h

Gibbs Uniqueness



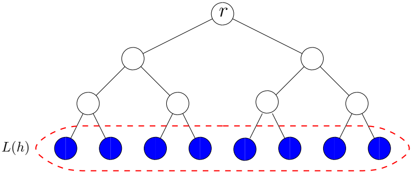
Δ-regular tree T of height h

Gibbs Uniqueness



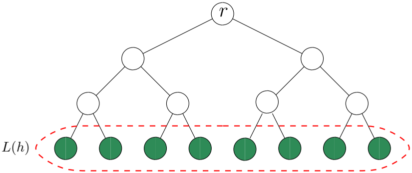
Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

Gibbs Uniqueness



Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

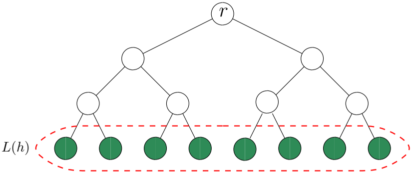
Gibbs Uniqueness



Δ-regular tree T of height h

Take two *extreme* configurations on $L(h)$

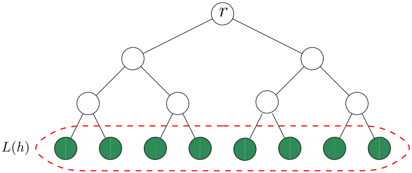
Gibbs Uniqueness



Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

For every λ consider

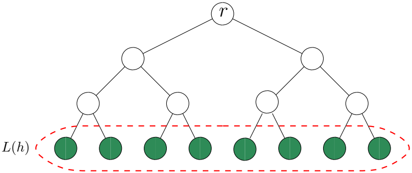
Gibbs Uniqueness



Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

For every λ consider

$$||\mu(\cdot|L(h) \text{ occupied}) - \mu(\cdot|L(h) \text{ unoccupied})||_{\{r\}}$$

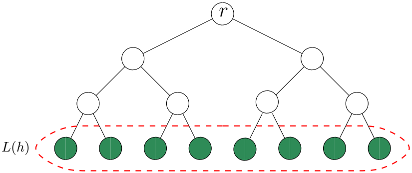


Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h) \text{ occupied}) - \mu(\cdot|L(h) \text{ unoccupied})||_{\{r\}}$$

Gibbs Uniqueness



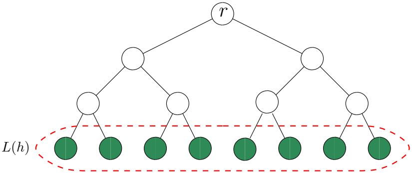
Δ-regular tree T of height h

Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h)\text{ occupied})-\mu(\cdot|L(h)\text{ unoccupied})||_{\{r\}} = \Bigg\{$$

Gibbs Uniqueness

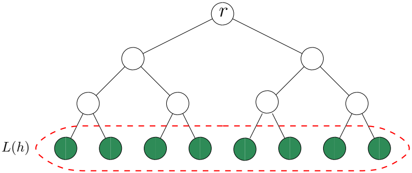


Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h) \text{ occupied}) - \mu(\cdot|L(h) \text{ unoccupied})||_{\{r\}} = \begin{cases} 0 & \\ & \end{cases}$$

Gibbs Uniqueness



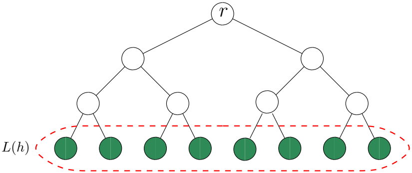
Δ-regular tree T of height h

Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h)\text{ occupied})-\mu(\cdot|L(h)\text{ unoccupied})||_{\{r\}} = \begin{cases} 0 \\ \delta \end{cases}$$

Gibbs Uniqueness



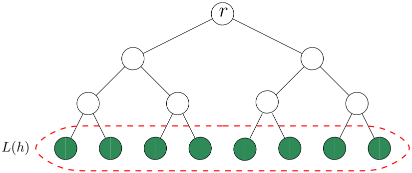
Δ-regular tree T of height h

Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h)\ \text{occupied}) - \mu(\cdot|L(h)\ \text{unoccupied})||_{\{r\}} = \begin{cases} 0 & \text{Unique} \\ \delta & \text{non-Unique} \end{cases}$$

Gibbs Uniqueness



Δ-regular tree T of height h

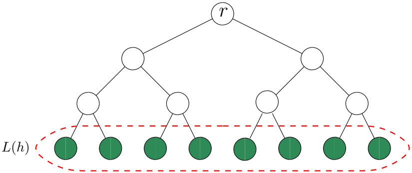
Take two *extreme* configurations on $L(h)$

For every λ consider

$$\lim_{h\to\infty} ||\mu(\cdot|L(h) \text{ occupied})-\mu(\cdot|L(h) \text{ unoccupied})||_{\{r\}} = \begin{cases} 0 & \text{Unique} \\ \delta & \text{non-Unique} \end{cases}$$

$$\lambda < \lambda_c(\Delta) \Leftrightarrow \text{Gibbs measure is Unique}$$

Gibbs Uniqueness



Δ-regular tree T of height h
Take two *extreme* configurations on $L(h)$

For every λ we compare

$$\lim_{h\to\infty} ||\mu(\cdot|L(h) \text{ occupied})-\mu(\cdot|L(h) \text{ unoccupied})||_{\{r\}} = \begin{cases} 0 & \text{Unique} \\ \delta & \text{non-Unique} \end{cases}$$

$$\lambda < \lambda_c(\Delta) \Leftrightarrow \text{we have spatial mixing}$$

Deterministic Algorithms

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.
- the size of computations depends on the size of the tree

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.
- the size of computations depends on the size of the tree
  - in the worst case the tree is exponentially large

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.
- the size of computations depends on the size of the tree
  - in the worst case the tree is exponentially large
- (strong) spatial mixing allows to "prune" the tree and still be accurate.

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.
- the size of computations depends on the size of the tree
  - in the worst case the tree is exponentially large
- (strong) spatial mixing allows to "prune" the tree and still be accurate.
  - this step requires $\lambda < \lambda_c(\Delta)$

Deterministic Algorithms

Weitz's approach [Weitz 2006]

Given G and $\lambda < \lambda_c$,

- uses tree of self avoiding walks, to organize the computations
  - reduces to dynamic programming.
- the size of computations depends on the size of the tree
  - in the worst case the tree is exponentially large
- (strong) spatial mixing allows to "prune" the tree and still be accurate.
  - this step requires $\lambda < \lambda_c(\Delta)$

- L. Li, P. Lu, and Y. Yin (2012), (2013)
- Restrepo, Shin, Tetali, Vigoda, and Yang (2013)
- A. Sinclair, P. Srivastava, and Y. Yin (2013)

Performance Weitz's algorithm

Approximation guarantees

For all $\delta > 0$, there exists constant $C = C(\delta) > 0$, for all Δ all G of maximum degree Δ, all $\lambda < (1-\delta)\lambda_c(\Delta)$ all $\epsilon > 0$ Weitz's algorithm returns an estimation $\hat{Z}$ of the partition function $Z(G,\lambda)$ such that

$$(1-\epsilon)Z(G,\lambda) \leq \hat{Z} \leq (1+\epsilon)Z(G,\lambda)$$

in time $O((n/\epsilon)^{C \log \Delta})$.

Randomized Algorithm

Randomized Algorithm

Markov Chain Monte Carlo

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets
- the equilibrium distribution is the hard-core model with fugacity λ

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets
- the equilibrium distribution is the hard-core model with fugacity λ
- the algorithm simulates the Markov chain

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets
- the equilibrium distribution is the hard-core model with fugacity λ
- the algorithm simulates the Markov chain
- outputs the configuration of the chain after "sufficiently many" steps

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets
- the equilibrium distribution is the hard-core model with fugacity λ
- the algorithm simulates the Markov chain
- outputs the configuration of the chain after "sufficiently many" steps

the output should be close to the equilibrium distribution

Randomized Algorithm

Markov Chain Monte Carlo

Given G and $\lambda > 0$,

- set up an ergodic Markov Chain over the independent sets
- the equilibrium distribution is the hard-core model with fugacity λ
- the algorithm simulates the Markov chain
- outputs the configuration of the chain after "sufficiently many" steps

the output should be close to the equilibrium distribution

it is desirable that the chain mixes "fast"

The dynamics

The dynamics

Glauber dynamics (X_t)

The dynamics

Glauber dynamics (X_t)

$X_t \to X_{t+1}$ is defined as follows:

The dynamics

Glauber dynamics (X_t)

$X_t \to X_{t+1}$ is defined as follows:

1. Choose v uniformly at random from V.

The dynamics

Glauber dynamics (X_t)

$X_t \to X_{t+1}$ is defined as follows:

1. Choose v uniformly at random from V.

$$X' = \begin{cases} X_t \cup \{v\} & \text{with probability } \lambda/(1+\lambda) \\ X_t \setminus \{v\} & \text{with probability } 1/(1+\lambda) \end{cases}$$

The dynamics

Glauber dynamics (X_t)

$X_t \rightarrow X_{t+1}$ is defined as follows:

1. Choose v uniformly at random from V.

$$X' = \begin{cases} X_t \cup \{v\} & \text{with probability } \lambda/(1+\lambda) \\ X_t \setminus \{v\} & \text{with probability } 1/(1+\lambda) \end{cases}$$

2. If X' is independent set, then $X_{t+1} = X'$, otherwise $X_{t+1} = X_t$

The dynamics

Glauber dynamics (X_t)

$X_t \to X_{t+1}$ is defined as follows:

1. Choose v uniformly at random from V.

$$X' = \begin{cases} X_t \cup \{v\} & \text{with probability } \lambda/(1+\lambda) \\ X_t \setminus \{v\} & \text{with probability } 1/(1+\lambda) \end{cases}$$

2. If X' is independent set, then $X_{t+1} = X'$, otherwise $X_{t+1} = X_t$

The chain converges to the hard-core model with fugacity λ.

Our Results

Our Results

Theorem

For all $\delta > 0$*, there exists* $\Delta_0 = \Delta_0(\delta)$ *for all graphs* $G = (V, E)$ *of maximum degree* $\Delta \geq \Delta_0$ *and* girth ≥ 7*, all* $\lambda < (1-\delta)\lambda_c(\Delta)$*, the mixing time of the Glauber dynamics satisfies*

$$T_{mix} = O\left(n \log(n)\right).$$

Our Results

Theorem

For all $\delta > 0$, *there exists* $\Delta_0 = \Delta_0(\delta)$ *for all graphs* $G = (V, E)$ *of maximum degree* $\Delta \geq \Delta_0$ *and* girth ≥ 7, *all* $\lambda < (1-\delta)\lambda_c(\Delta)$, *the mixing time of the Glauber dynamics satisfies*

$$T_{mix} = O\left(n\log(n)\right).$$

Mixing Time . . .

$$T_{mix} = \min\{t : \text{ for all } X_0, d_{tv}(X_t, \mu) \leq 1/4\},$$

Our Results

Theorem

For all $\delta > 0$*, there exists* $\Delta_0 = \Delta_0(\delta)$ *for all graphs* $G = (V, E)$ *of maximum degree* $\Delta \geq \Delta_0$ *and* girth ≥ 7*, all* $\lambda < (1-\delta)\lambda_c(\Delta)$*, the mixing time of the Glauber dynamics satisfies*

$$T_{mix} = O\left(n \log(n)\right).$$

Corollary

The above sampling result yields an FPRAS for estimating the partition function Z. The running time is $O^*(n^2)$.

Our Results

Theorem

For all $\delta > 0$*, there exists* $\Delta_0 = \Delta_0(\delta)$ *for all graphs* $G = (V, E)$ *of maximum degree* $\Delta \geq \Delta_0$ *and* girth ≥ 7*, all* $\lambda < (1-\delta)\lambda_c(\Delta)$*, the mixing time of the Glauber dynamics satisfies*

$$T_{mix} = O\left(n \log(n)\right).$$

Previous work

$T_{mix} = O\left(n \log(n)\right)$ for Glauber dynamics on G of maximum degree Δ and $\lambda < 2/(\Delta - 2)$

- Dyer Greenhill, Luby, Vigoda

$O(n \log n)$ mixing for Random Graphs

$O(n \log n)$ mixing for Random Graphs

Relaxation for girth

" # short cycles in the neighborhood of its vertex in G are not too many"

$O(n \log n)$ mixing for Random Graphs

Relaxation for girth

" $\#$ short cycles in the neighborhood of its vertex in G are not too many"

Corollary

$T_{mix} = O(n \log n)$ for Glauber dynamics with $\lambda \leq (1-\delta)\lambda_c(\Delta)$ for

$O(n \log n)$ mixing for Random Graphs

Relaxation for girth

" # short cycles in the neighborhood of its vertex in G are not too many"

Corollary

$T_{mix} = O(n \log n)$ for Glauber dynamics with $\lambda \leq (1-\delta)\lambda_c(\Delta)$ for

- random Δ-regular graph

$O(n \log n)$ mixing for Random Graphs

Relaxation for girth

" # short cycles in the neighborhood of its vertex in G are not too many"

Corollary

$T_{mix} = O(n \log n)$ for Glauber dynamics with $\lambda \leq (1-\delta)\lambda_c(\Delta)$ for

- random Δ-regular graph
- random Δ-regular bipartite graph

$O(n \log n)$ mixing for Random Graphs

Relaxation for girth

" # short cycles in the neighborhood of its vertex in G are not too many"

Corollary

$T_{mix} = O(n \log n)$ for Glauber dynamics with $\lambda \leq (1-\delta)\lambda_c(\Delta)$ for

- random Δ-regular graph
- random Δ-regular bipartite graph

Mossel, Weitz, Wormald (2009)

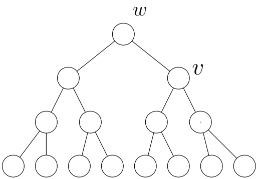
Belief Propagation on trees

Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

Belief Propagation on trees

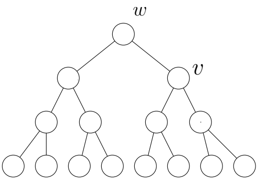
For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$



Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

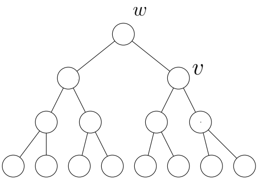
$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$



Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$

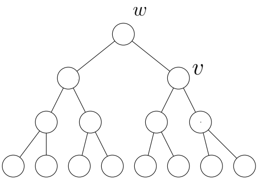


Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$

$$R_{v \to w} = \frac{q_w(v)}{1 - q_w(v)}$$

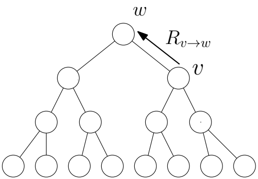


Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied}|w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied}|w \text{ unoccupied})$$

$$R_{v \to w} = \frac{q_w(v)}{1 - q_w(v)}$$

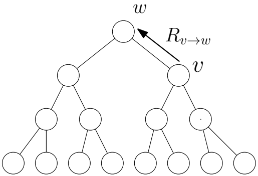


For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$

$$R_{v\to w} = \frac{q_w(v)}{1 - q_w(v)}$$

$$R_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1 + R_{z\to v}}$$



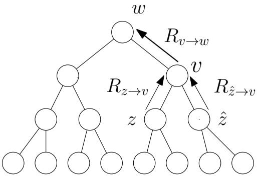
Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$

$$R_{v \to w} = \frac{q_w(v)}{1 - q_w(v)}$$

$$R_{v \to w} = \lambda \prod_{z \in N(v) \setminus \{w\}} \frac{1}{1 + R_{z \to v}}$$



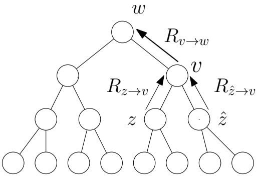
Belief Propagation on trees

For T and λ compute
$\mu(v \text{ occupied} | w \text{ unoccupied})$

$$q_w(v) = \mu(v \text{ occupied} | w \text{ unoccupied})$$

$$R_{v\to w} = \frac{q_w(v)}{1 - q_w(v)}$$

$$R_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1 + R_{z\to v}}$$



BP starts from arbitrary $R^0_{v\to w}$s,
iterates like

$$R^i_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1 + R^{i-1}_{z\to v}}$$

Convergence

Convergence

Convergence on trees

There exists i_0 such that for every $i \geq i_0$ and every $(R^0_{v\to w})_{\{v,w\}\in E}$ we have

$$R^i_{v\to w} = R^*_{v\to w}$$

In turn

$$\mu(v \text{ occupied}|w \text{ unoccupied}) = q^* = \frac{R^*_{v\to w}}{1 + R^*_{v\to w}}$$

Convergence

Convergence on trees

There exists i_0 such that for every $i \geq i_0$ and every $(R^0_{v\to w})_{\{v,w\}\in E}$ we have

$$R^i_{v\to w} = R^*_{v\to w}$$

In turn

$$\mu(v \text{ occupied}|w \text{ unoccupied}) = q^* = \frac{R^*_{v\to w}}{1+R^*_{v\to w}}$$

BP is an elaborate use of *Dynamic Programing* to compute marginal.

(Loopy) Belief Propagation

(Loopy) Belief Propagation

- We do not know whether it converges

(Loopy) Belief Propagation

- We do not know whether it converges
- . . . if does, we do not know where exactly it converges

BP Convergence for girth ≥ 6

BP Convergence for girth ≥ 6

$$R^{i}_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1+R^{i-1}_{z\to v}}$$

BP Convergence for girth ≥ 6

$$R^i_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1+R^{i-1}_{z\to v}} \quad \text{and} \quad q^i_w(v) = \frac{R^i_{v\to w}}{1+R^i_{v\to w}}$$

BP Convergence for girth ≥ 6

$$R^i_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1+R^{i-1}_{z\to v}} \quad \text{and} \quad q^i_w(v) = \frac{R^i_{v\to w}}{1+R^i_{v\to w}}$$

Theorem

For $G=(V,E)$ of maximum degree $\Delta \geq \Delta_0$ and girth ≥ 6, all $\lambda < (1-\delta)\lambda_c(\Delta)$, the following holds: for $i \geq C$, for all $v \in V$, $w \in N(v)$,

$$\left| \frac{q^i_w(v)}{\mu(v \ is\ occupied \mid w \ is\ unoccupied)} - 1 \right| \leq \epsilon$$

BP Convergence for girth ≥ 6

$$R^i_{v\to w} = \lambda \prod_{z\in N(v)\setminus\{w\}} \frac{1}{1+R^{i-1}_{z\to v}} \quad \text{and} \quad q^i_w(v) = \frac{R^i_{v\to w}}{1+R^i_{v\to w}}$$

Theorem

For $G=(V,E)$ of maximum degree $\Delta \geq \Delta_0$ and girth ≥ 6, all $\lambda < (1-\delta)\lambda_c(\Delta)$, the following holds: for $i \geq C$, for all $v \in V$, $w \in N(v)$,

$$\left| \frac{q^i_w(v)}{\mu(v \text{ is occupied} \mid w \text{ is unoccupied})} - 1 \right| \leq \epsilon$$

we also have convergence for the BP estimate of $\mu(v \text{ is occupied})$

Path Coupling for Rapid Mixing

Path Coupling for Rapid Mixing

Path Coupling [Bubley and Dyer 1997]

Path Coupling for Rapid Mixing

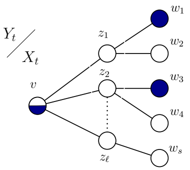
Path Coupling [Bubley and Dyer 1997]

Consider copies $(X_s), (Y_s)$ such that $X_t \oplus Y_t = \{v\}$

Path Coupling for Rapid Mixing

Path Coupling [Bubley and Dyer 1997]

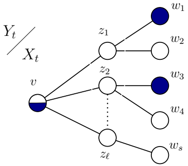
Consider copies $(X_s), (Y_s)$ such that $X_t \oplus Y_t = \{v\}$



Path Coupling [Bubley and Dyer 1997]

Consider copies $(X_s), (Y_s)$ such that $X_t \oplus Y_t = \{v\}$

$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] \leq (1-\gamma)\Phi(X_t, Y_t).$$



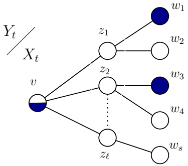
Path Coupling for Rapid Mixing

Path Coupling [Bubley and Dyer 1997]

Consider copies $(X_s), (Y_s)$ such that $X_t \oplus Y_t = \{v\}$

$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] \leq (1-\gamma)\Phi(X_t, Y_t).$$

$\Phi : \Omega \times \Omega \to \mathbb{R}_{\geq 1}$ is a "distance metric"



Path Coupling for Rapid Mixing

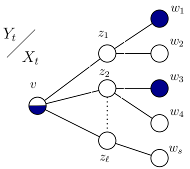
Path Coupling [Bubley and Dyer 1997]

Consider copies $(X_s), (Y_s)$ such that $X_t \oplus Y_t = \{v\}$

$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] \leq (1-\gamma)\Phi(X_t, Y_t).$$

$\Phi : \Omega \times \Omega \to \mathbb{R}_{\geq 1}$ is a "distance metric"

$$\Phi(X, Y) = \sum_{u \in X \oplus Y} \Phi(u)$$



Key Results

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like R^*

Key Results

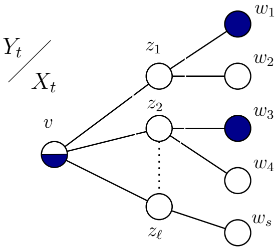
- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like R^*
- Glauber dynamics converges locally to R^*

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like R^*
- Glauber dynamics converges locally to R^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Path Coupling Example

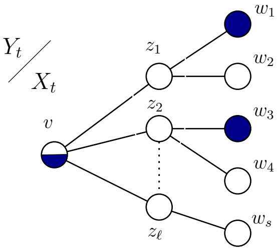
Path Coupling Example



Path Coupling Example

Expected distance

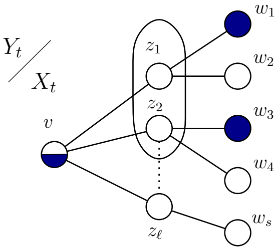
$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] = \left(1 - \frac{1}{n}\right)\Phi(v) + \frac{1}{n}\sum_{z_i} \Pr[z_i \in Y_{t+1}] \cdot \Phi(z_i)$$



Path Coupling Example

Expected distance

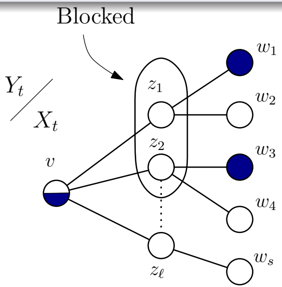
$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] = \left(1 - \frac{1}{n}\right)\Phi(v) + \frac{1}{n}\sum_{z_i} \Pr[z_i \in Y_{t+1}] \cdot \Phi(z_i)$$



Path Coupling Example

Expected distance

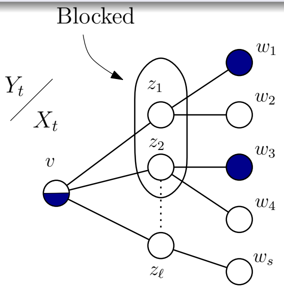
$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] = \left(1 - \frac{1}{n}\right)\Phi(v) + \frac{1}{n}\sum_{z_i}\Pr[z_i \in Y_{t+1}] \cdot \Phi(z_i)$$



Path Coupling Example

Expected distance

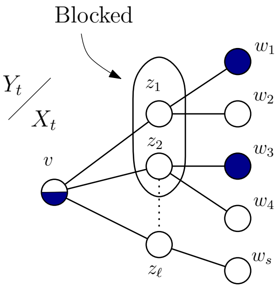
$$\mathbb{E}\left[\Phi(X_{t+1}, Y_{t+1}) \middle| X_t, Y_t\right] = \left(1 - \frac{1}{n}\right)\Phi(v) + \frac{1}{n}\sum_{z_i} \mathbf{1}\{z_i \text{ unblocked}\}\frac{\lambda\Phi(z_i)}{1+\lambda}$$



Path Coupling Example

Path coupling condition

$$\Phi(v) > \frac{\lambda}{1+\lambda} \sum_{z_i} \mathbf{1}\{z_i \text{ unblocked in } Y_t\} \cdot \Phi(z_i)$$



Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like R^*
- Glauber dynamics converges locally to R^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Unblocked Neighbors and loopy BP

Unblocked Neighbors and loopy BP

$$\omega_z^i = \prod_{y \sim z} \frac{1}{1 + \lambda \cdot \omega_y^{i-1}}$$

Unblocked Neighbors and loopy BP

$$\omega_z^i = \prod_{y \sim z} \frac{1}{1 + \lambda \cdot \omega_y^{i-1}}$$

- $\omega^i(z)$ is the loopy BP estimate of z to be unblocked

Unblocked Neighbors and loopy BP

$$\omega_z^i = \prod_{y \sim z} \frac{1}{1 + \lambda \cdot \omega_y^{i-1}}$$

- $\omega^i(z)$ is the loopy BP estimate of z to be unblocked
- converges to a unique fixed point ω^*

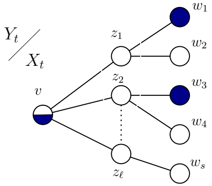
Unblocked Neighbors and loopy BP

$$\omega_z^i = \prod_{y \sim z} \frac{1}{1 + \lambda \cdot \omega_y^{i-1}}$$

- $\omega^i(z)$ is the loopy BP estimate of z to be unblocked
- converges to a unique fixed point ω^*
- $\omega^*(z) \approx \mu(z \text{ is unblocked})$

Back to Path Coupling

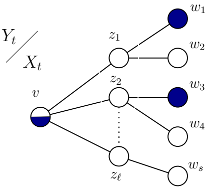
Back to Path Coupling



Back to Path Coupling

worst case condition

$$\Phi(v) > \frac{\lambda}{1+\lambda} \sum_{z_i} \mathbf{1}\{z_i \text{ unblocked}\} \cdot \Phi(z_i)$$



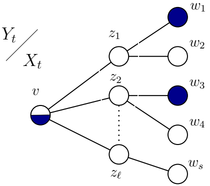
Back to Path Coupling

worst case condition

$$\Phi(v) > \frac{\lambda}{1+\lambda} \sum_{z_i} \mathbf{1}\{z_i \text{ unblocked}\} \cdot \Phi(z_i)$$

when X_t, Y_t "behave" like ω^*

$$\Phi(v) > \frac{\lambda}{1+\lambda} \sum_{z_i} \omega^*(z_i) \cdot \Phi(z_i)$$



Finding Φ

Finding Φ

Reformulation

Finding Φ

Reformulation

For $\rho = 1 - \delta$, there is Φ such that

$$\rho \cdot \Phi(v) \geq \sum_{z_i} \frac{\lambda \omega^*(z_i)}{1 + \lambda \omega^*(z_i)} \cdot \Phi(z_i)$$

Finding Φ

Reformulation

For $\rho = 1 - \delta$, there is Φ such that

$$\rho \cdot \Phi(v) \geq \sum_{z_i} \frac{\lambda \omega^*(z_i)}{1 + \lambda \omega^*(z_i)} \cdot \Phi(z_i)$$

$n \times n$ matrix $\mathcal{C}$

$$\mathcal{C}(v, z) = \begin{cases} \frac{\lambda \omega^*(z)}{1+\lambda \omega^*(z)} & \text{if } z \in N(v) \\ 0 & \text{otherwise} \end{cases}$$

Finding Φ

Reformulation

For $\rho = 1 - \delta$, there is Φ such that

$$\rho \cdot \Phi(v) \geq \sum_{z_i} \frac{\lambda\omega^*(z_i)}{1 + \lambda\omega^*(z_i)} \cdot \Phi(z_i)$$

$n \times n$ matrix $\mathcal{C}$

$$\mathcal{C}(v, z) = \begin{cases} \frac{\lambda\omega^*(z)}{1+\lambda\omega^*(z)} & \text{if } z \in N(v) \\ 0 & \text{otherwise} \end{cases}$$

There is a vector $\Phi \in \mathbb{R}^V_{\geq 1}$ such that

$$\mathcal{C}\Phi \leq \rho \cdot \Phi.$$

Connections with Loopy BP

Connections with Loopy BP

Jacobian of Loopy BP

Connections with Loopy BP

Jacobian of Loopy BP

BP Operator

$$F(\omega_z) = \prod_{y \in N(z)} \frac{1}{1 + \lambda \omega_y}.$$

Connections with Loopy BP

Jacobian of Loopy BP

BP Operator

$$F(\omega_z) = \prod_{y \in N(z)} \frac{1}{1+\lambda\omega_y}.$$

$J^* = J|_{\omega=\omega^*}$ denote the Jacobian of F at the fixed point $\omega = \omega^*$.

Connections with Loopy BP

Jacobian of Loopy BP

BP Operator

$$F(\omega_z) = \prod_{y \in N(z)} \frac{1}{1 + \lambda \omega_y}.$$

$J^* = J|_{\omega=\omega^*}$ denote the Jacobian of F at the fixed point $\omega = \omega^*$.

$$\hat{J} = D^{-1} J^* D,$$

where D is diagonal matrix, with $D(v, v) = \omega^*(v)$

Connections with Loopy BP

Jacobian of Loopy BP

BP Operator

$$F(\omega_z) = \prod_{y \in N(z)} \frac{1}{1+\lambda\omega_y}.$$

$J^* = J|_{\omega=\omega^*}$ denote the Jacobian of F at the fixed point $\omega = \omega^*$.

$$\hat{J} = D^{-1} J^* D,$$

where D is diagonal matrix, with $D(v,v) = \omega^*(v)$

Relation to Path Coupling

$$\hat{J} = \mathcal{C}$$

Covergence from loopy BP

Covergence from loopy BP

Reduction to BP Spectral radius

For $\rho = 1 - \delta$, there is a vector $\Phi \in \mathbb{R}^V$ such that

$$\hat{J}\Phi \leq \rho \cdot \Phi$$

Covergence from loopy BP

Reduction to BP Spectral radius

For $\rho = 1 - \delta$, there is a vector $\Phi \in \mathbb{R}^V$ such that

$$\hat{J}\Phi \leq \rho \cdot \Phi$$

$\hat{J}$ has the same eigenvalues as the Jacobian of BP at the fixed point

Covergence from loopy BP

Reduction to BP Spectral radius

For $\rho = 1 - \delta$, there is a vector $\Phi \in \mathbb{R}^V$ such that

$$\hat{J}\Phi \leq \rho \cdot \Phi$$

$\hat{J}$ has the same eigenvalues as the Jacobian of BP at the fixed point

Spectral radius of BP in uniqueness region

We should expect $\rho(\lambda, \Delta) < 1$, because the fixed point ω^* is attractive

Covergence from loopy BP

Reduction to BP Spectral radius

For $\rho = 1 - \delta$, there is a vector $\Phi \in \mathbb{R}^V$ such that

$$\hat{J}\Phi \leq \rho \cdot \Phi$$

$\hat{J}$ has the same eigenvalues as the Jacobian of BP at the fixed point

Spectral radius of BP in uniqueness region

We should expect $\rho(\lambda, \Delta) < 1$, because the fixed point ω^* is attractive

- $\Phi > 0$ from Perron-Frobenius

Covergence from loopy BP

Reduction to BP Spectral radius

For $\rho = 1 - \delta$, there is a vector $\Phi \in \mathbb{R}^V$ such that

$$\hat{J}\Phi \leq \rho \cdot \Phi$$

$\hat{J}$ has the same eigenvalues as the Jacobian of BP at the fixed point

Spectral radius of BP in uniqueness region

We should expect $\rho(\lambda, \Delta) < 1$, because the fixed point ω^* is attractive

- $\Phi > 0$ from Perron-Frobenius

What is Φ

$$\Phi(v) = \sqrt{\frac{1 + \lambda\omega^*(v)}{\omega^*(v)}}$$

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like R^*
- Glauber dynamics converges locally to R^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like ω^*
- Glauber dynamics converges locally to ω^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like ω^*
- Glauber dynamics converges locally to ω^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Local Uniformity with Loopy BP fixed points

Theorem

Let G be of girth ≥ 7 and maximum degree Δ, for $\Delta > \Delta_0$. Let (X_t) be the Glauber dynamics with $\lambda < (1-\delta)\lambda_c(\Delta)$.
For any vertex v, with probability $1 - \exp\left[-\Delta/C\right]$, it holds that

$$\#\ \text{Unblocked Neighbors of } v \text{ in } X_t < \sum_{z \in N(v)} \omega^*(z) + \epsilon\Delta$$

where $t \geq Cn \log \Delta$.

Key Results

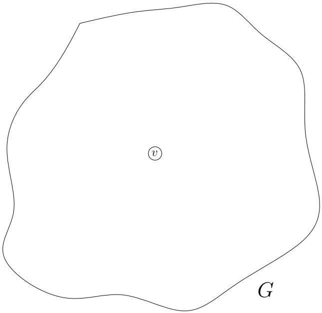
- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- can find Φ when locally X_t, Y_t "behave" like ω^*
  - Φ is from the Jacobian of BP operator
- Glauber dynamics (approximately) converges locally to ω^*
  - locally Glauber dynamics behaves approximately like BP fixed points
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Rapid Mixing with uniformity

Dyer, Frieze, Hayes, Vigoda 2013

Rapid Mixing with uniformity

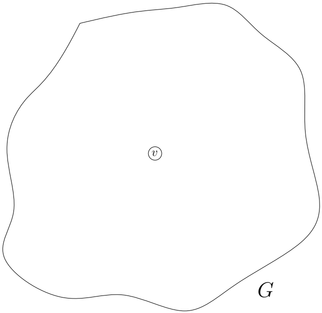
Dyer, Frieze, Hayes, Vigoda 2013



There is a single disagreement at v

Rapid Mixing with uniformity

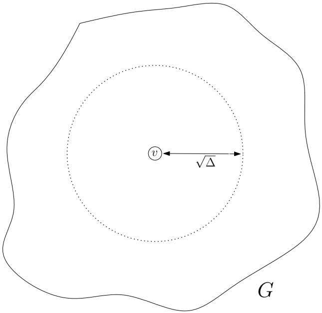
Dyer, Frieze, Hayes, Vigoda 2013



Run the chains for $Cn \log \Delta$ steps, "burn-in"

Rapid Mixing with uniformity

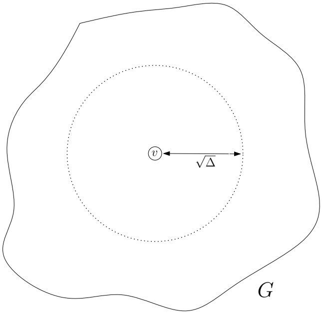
Dyer, Frieze, Hayes, Vigoda 2013



Run the chains for $Cn \log \Delta$ steps, "burn-in"

Rapid Mixing with uniformity

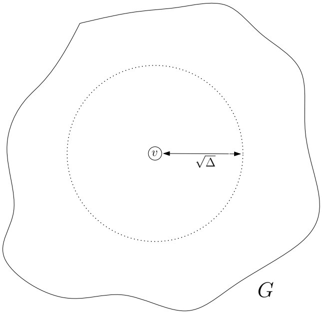
Dyer, Frieze, Hayes, Vigoda 2013



The disagreements spread in the graph during burn-in

Rapid Mixing with uniformity

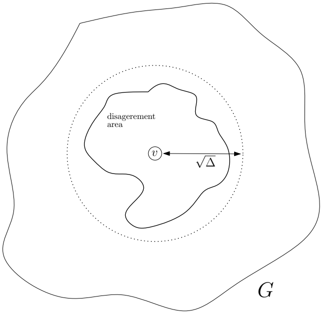
Dyer, Frieze, Hayes, Vigoda 2013



Typically the disagreements do not escape the ball

Rapid Mixing with uniformity

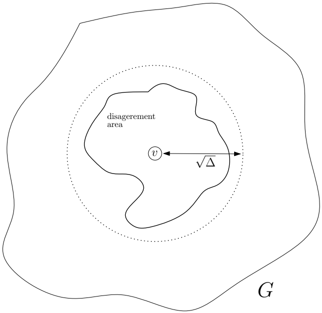
Dyer, Frieze, Hayes, Vigoda 2013



Typically the disagreements do not escape the ball

Rapid Mixing with uniformity

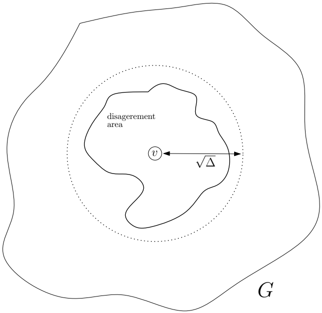
Dyer, Frieze, Hayes, Vigoda 2013



Typically the ball has uniformity.

Rapid Mixing with uniformity

Dyer, Frieze, Hayes, Vigoda 2013

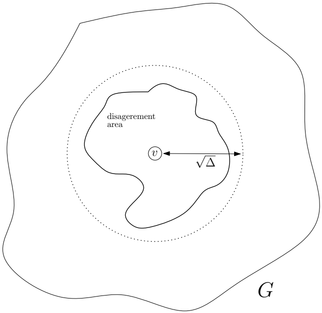


Interpolate and do path coupling for the pairs,

... the pairs now "behave" like ω^*

Rapid Mixing with uniformity

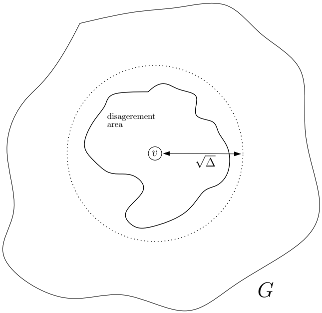
Dyer, Frieze, Hayes, Vigoda 2013



Interpolate and do path coupling for the pairs,
. . . the pairs now "behave" like ω^* and Φ gives convergence

Rapid Mixing with uniformity

Dyer, Frieze, Hayes, Vigoda 2013



$$\mathbb{E}\left[\Phi(X_{C'n\log\Delta}, Y_{C'n\log\Delta})\,\middle|\, X_0, Y_0\right] \leq (1-\gamma)\Phi(X_0, Y_0)$$

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- We can find Φ when $X, Y \sim \omega^*$
- Glauber dynamics converges locally to ω^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Key Results

- We don't know a Φ that gives contraction for worst-case X_t, Y_t.
- We can find Φ when $X, Y \sim \omega^*$
- Glauber dynamics converges locally to ω^*
- Given Φ and convergence of Glauber dynamics we show rapid mixing

Local uniformity I

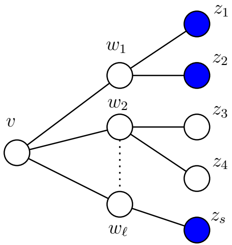
$$\mathbf{R}(\sigma, v) = \prod_{w \sim v} \left(1 - \frac{\lambda}{1+\lambda}\mathbf{1}\{\text{w unblocked by its children}\}\right),$$

Local uniformity I

$$\mathbf{R}(\sigma, v) = \prod_{w \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbf{1}\{\text{w unblocked by its children}\}\right),$$

Local uniformity I

$$\mathbf{R}(\sigma, v) = \prod_{w \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbf{1}\{\text{w unblocked by its children}\}\right),$$



$$\mathbf{R}(\sigma, v) = \Pr_{Y \sim \mu}\left[v \text{ is unblocked in } Y | v \notin Y,\ Y(S_2(v)) = \sigma(S_2(v))\right]$$

Local uniformity I

$$\mathbf{R}(\sigma, v) = \prod_{w \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbf{1}\{\text{w unblocked by its children}\}\right),$$

BP for Gibbs measure

Let G be of girth ≥ 6 and maximum degree $\Delta > \Delta_0$. Let X be distributed as in μ with $\lambda < (1-\delta)\lambda_c(\Delta)$.
Then for any vertex v with probability $\geq 1 - \exp\left(-\Delta/C\right)$ it holds that

$$\left| \mathbf{R}(X, v) - \prod_{z \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbf{R}(X, z)\right) \right| < \gamma.$$

Local uniformity I

$$\mathbf{R}(\sigma, v) = \prod_{w \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbf{1}\{\text{w unblocked by its children}\}\right),$$

BP for Glauber dynamics

Let G be of girth ≥ 7 and maximum degree $\Delta > \Delta_0$. Let (X_t) be the Glauber dynamics with $\lambda < (1-\delta)\lambda_c(\Delta)$.
Then for any vertex v and any $t > Cn \log \Delta$ with probability $\geq 1 - \exp(-\Delta/C)$ it holds that

$$\left| \mathbf{R}(X_t, v) - \prod_{z \sim v} \left(1 - \frac{\lambda}{1+\lambda} \mathbb{E}_{t_z}\left[\mathbf{R}(X_{t_z}, z)\right]\right) \right| < \gamma.$$

Local uniformity II

Local uniformity II

Lemma

Let G be of girth ≥ 7 and maximum degree $\Delta > \Delta_0$. Let (X_t) be the Glauber dynamics with $\lambda < (1-\delta)\lambda_c(\Delta)$.
For all $\mathcal{I} = [t_0, t_1]$, where $t_0 = Cn \log \Delta$, for every $v \in V$ with probability $1 - (1 + |\mathcal{I}|/n) \exp(-\Delta/C)$, we have that

$$(\forall t \in \mathcal{I}) \quad |\mathbf{R}(X_t, v) - \omega^*(v)| \leq \epsilon.$$

Local uniformity II

Lemma

Let G be of girth ≥ 7 and maximum degree $\Delta > \Delta_0$. Let (X_t) be the Glauber dynamics with $\lambda < (1-\delta)\lambda_c(\Delta)$.
For all $\mathcal{I} = [t_0, t_1]$, where $t_0 = Cn \log \Delta$, for every $v \in V$ with probability $1 - (1 + |\mathcal{I}|/n) \exp\left(-\Delta/C\right)$, we have that

$$(\forall t \in \mathcal{I}) \quad |\mathbf{R}(X_t, v) - \omega^*(v)| \leq \epsilon.$$

- Hayes 2012

Iterations in space and time

Convergence with Ψ

Potential function

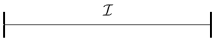
$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Iterations in space and time

Convergence with Ψ

Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

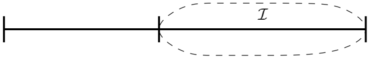


Iterations in space and time

Convergence with Ψ

Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$



Iterations in space and time

Convergence with Ψ

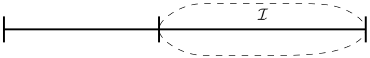
Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq \alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

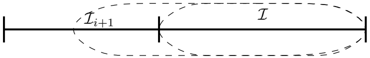
Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq \alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

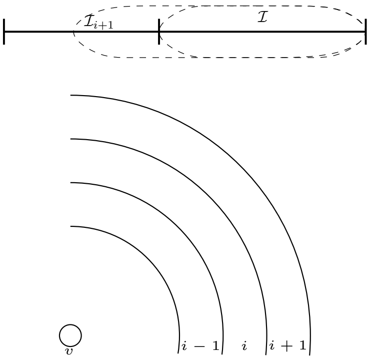
Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq \alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

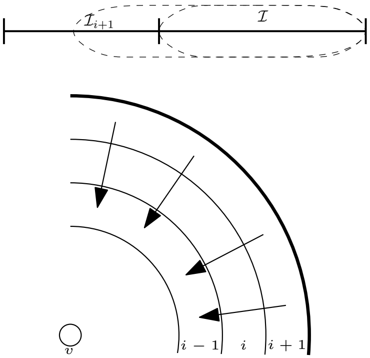
Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \le \alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

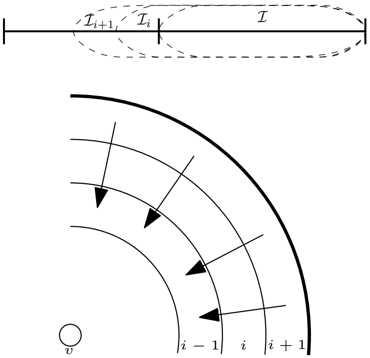
Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq \alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

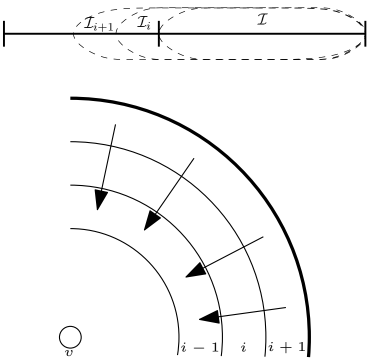
Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \le \alpha_{i+1}$$

$\forall t \in \mathcal{I}_i$, $u \in B(v, i)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \le (1-\delta)\alpha_{i+1}$$



Iterations in space and time

Convergence with Ψ

Potential function

$$\Psi(x) = (\lambda)^{-1}\mathrm{arcsinh}(\sqrt{\lambda x})$$

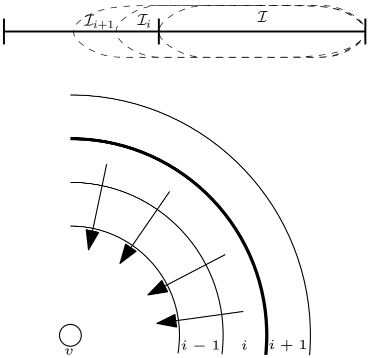
Provided

- $t \in \mathcal{I}'$ approximate BP equation hold in $B(v, R)$
- $\forall t \in \mathcal{I}_{i+1}$, $u \in B(v, i+1)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq \alpha_{i+1}$$

$\forall t \in \mathcal{I}_i$, $u \in B(v, i)$

$$|\Psi(\mathbf{R}(X_t, u)) - \Psi(\omega^*(u))| \leq (1-\delta)\alpha_{i+1}$$



Concluding Remarks

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*
  - proposing "Hamming weights"

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*
  - proposing "Hamming weights"
- Establish a novel connection between Path Coupling and Loopy BP
  - this is important for both uniformity and Hamming weights

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*
  - proposing "Hamming weights"
- Establish a novel connection between Path Coupling and Loopy BP
  - this is important for both uniformity and Hamming weights
- Use experience from Glauber dynamics to analyze Loopy BP
  - for graphs of girth ≥ 6 in the uniqueness region

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*
  - proposing "Hamming weights"
- Establish a novel connection between Path Coupling and Loopy BP
  - this is important for both uniformity and Hamming weights
- Use experience from Glauber dynamics to analyze Loopy BP
  - for graphs of girth ≥ 6 in the uniqueness region
- The connection between Glauber dynamics and Loopy BP is deep

Concluding Remarks

- Rapid mixing for Glauber Dynamics
  - G max degree $\Delta > \Delta_0$ and girth ≥ 7
  - λ in uniqueness
- Approach
  - by establishing *uniformity*
  - proposing "Hamming weights"
- Establish a novel connection between Path Coupling and Loopy BP
  - this is important for both uniformity and Hamming weights
- Use experience from Glauber dynamics to analyze Loopy BP
  - for graphs of girth ≥ 6 in the uniqueness region
- The connection between Glauber dynamics and Loopy BP is deep
  - Allows to establish uniformity and weights in a *systematic way*

The End

THANK YOU!