Sampling up to the *Uniqueness* Threshold

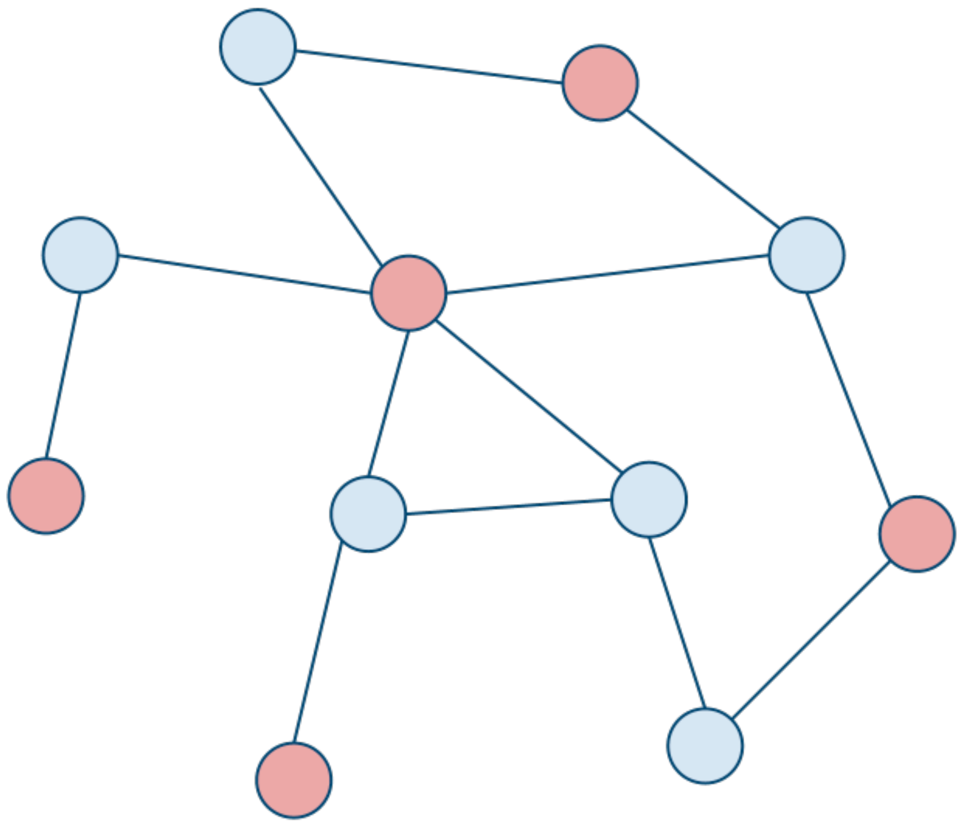
Yitong Yin
Nanjing University

Joint work with: Charis Efthymiou (Goethe-Universität Frankfurt)
Thomas P. Hayes (UNM)
Daniel Štefankovič (Rochester)
Eric Vigoda (Georgia Tech)

Sampling Independent Set

undirected graph $G(V, E)$ of max-degree Δ

- Sample (*nearly*) uniform random independent set.

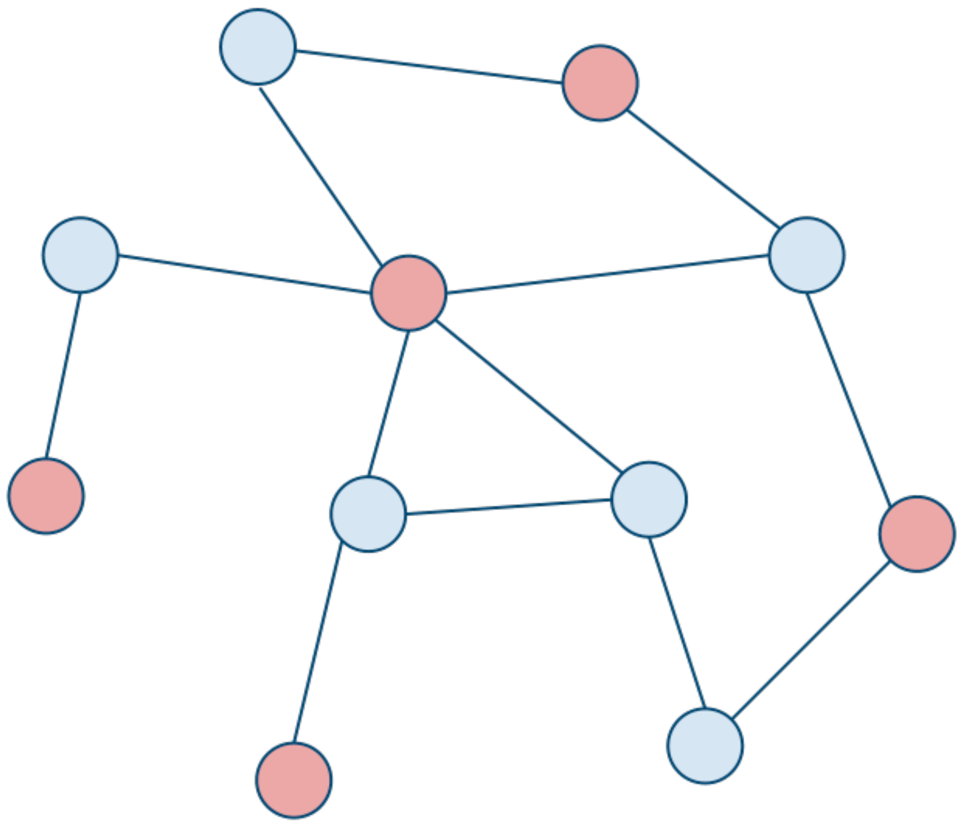


Sampling Independent Set

undirected graph $G(V,E)$ of max-degree Δ

- Sample (*nearly*) uniform random independent set.
- $\Delta \leq 5 \Rightarrow$ poly-time

  Conjecture: $O(n \log n)$ time
- $\Delta \geq 6 \Rightarrow$ NP-hard



Hardcore Model

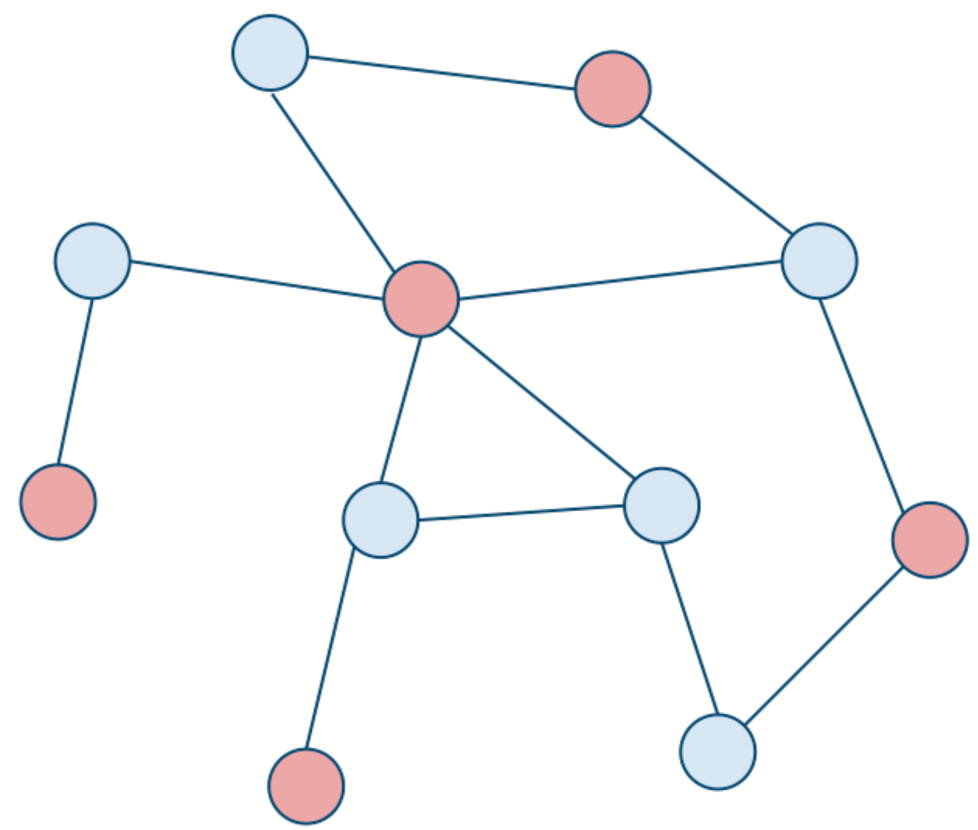
undirected graph $G(V, E)$ of max-degree Δ

- fugacity parameter $\lambda > 0$
- each independent set I in G is assigned a weight:

$$w(I) = \lambda^{|I|}$$

- distribution μ over all independent sets in G:

$$\mu(I) = \frac{w(I)}{\sum_I w(I)}$$

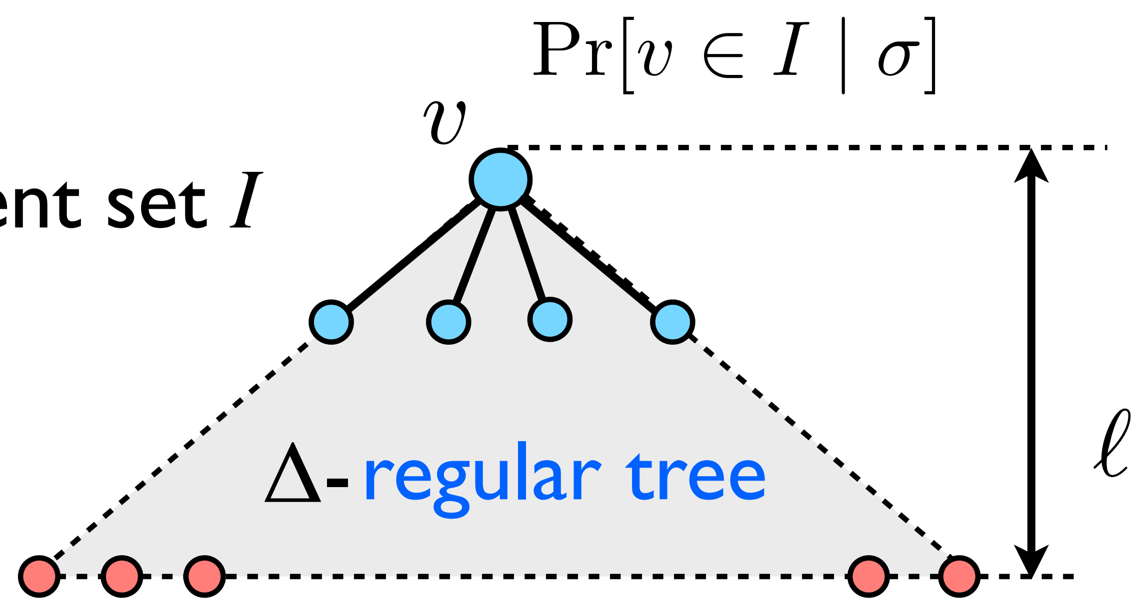


- $\lambda = 1$: uniform distribution over independent sets

Uniqueness Threshold

random independent set I

$\mu(I) \propto \lambda^{|I|}$

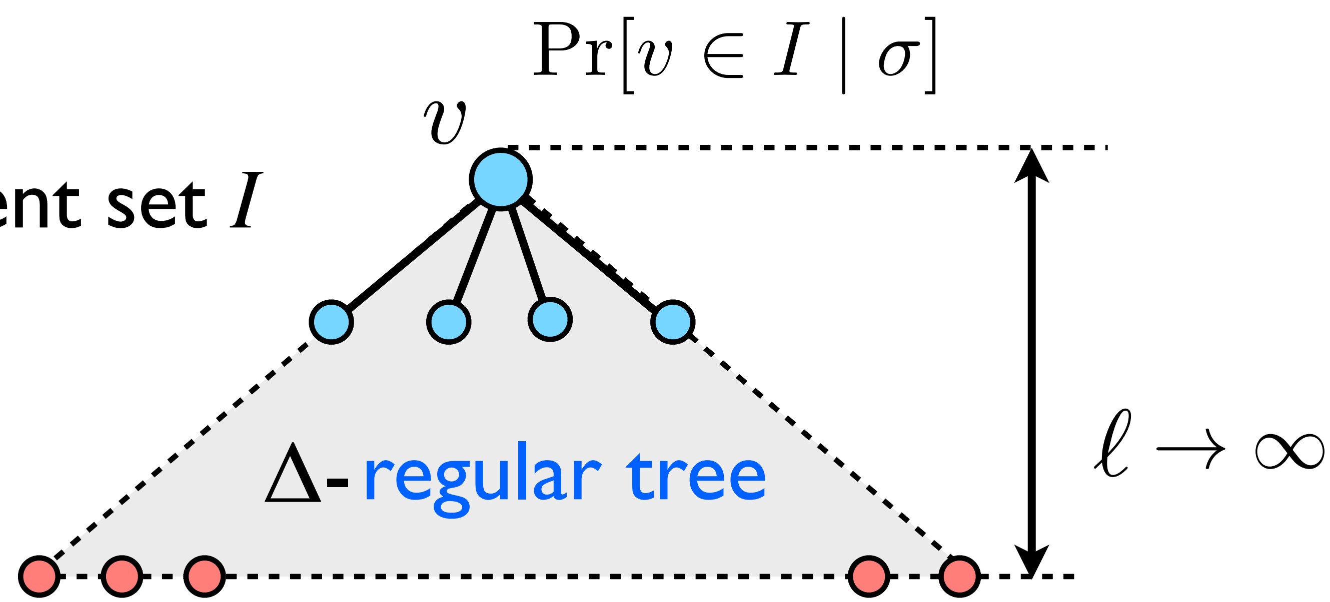


σ: boundary condition fixing each leaf $\in I$ or $\notin I$

Uniqueness Threshold

random independent set I

$\mu(I) \propto \lambda^{|I|}$



σ: boundary condition fixing each leaf $\in I$ or $\notin I$

Critical phenomenon:

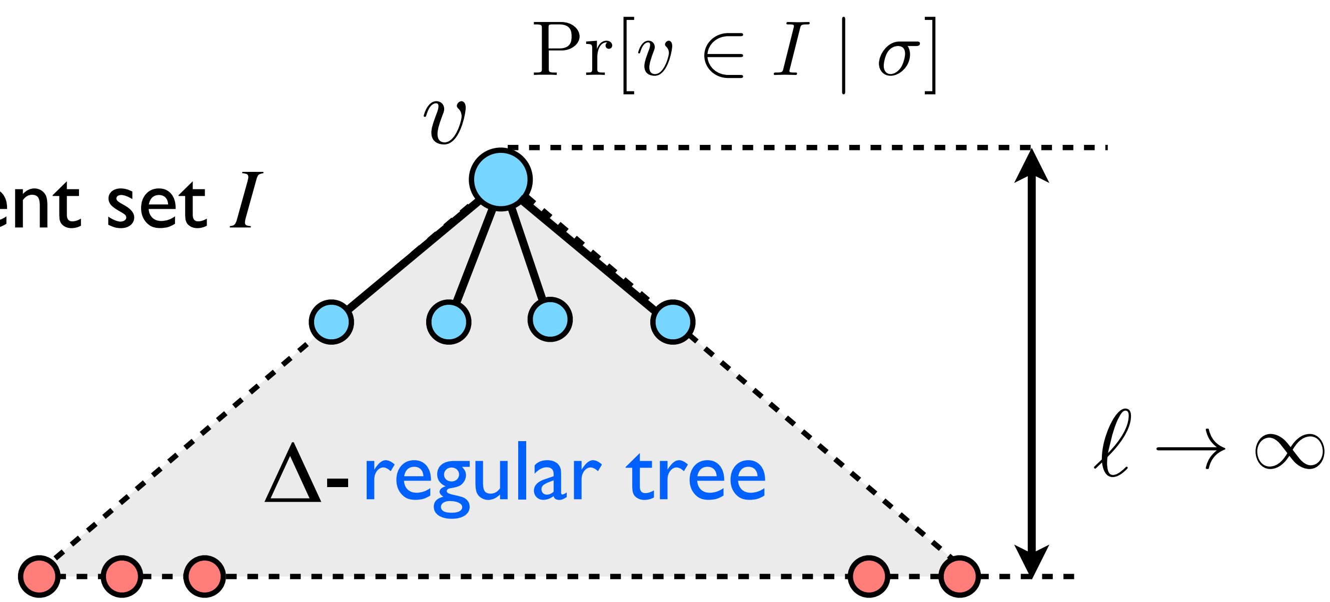
$\Pr[v \in I \mid \sigma]$ is independent of σ when $l \to \infty$

iff $\lambda \leq \lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}}$

Uniqueness Threshold

random independent set I

$\mu(I) \propto \lambda^{|I|}$



σ: boundary condition fixing each leaf $\in I$ or $\notin I$

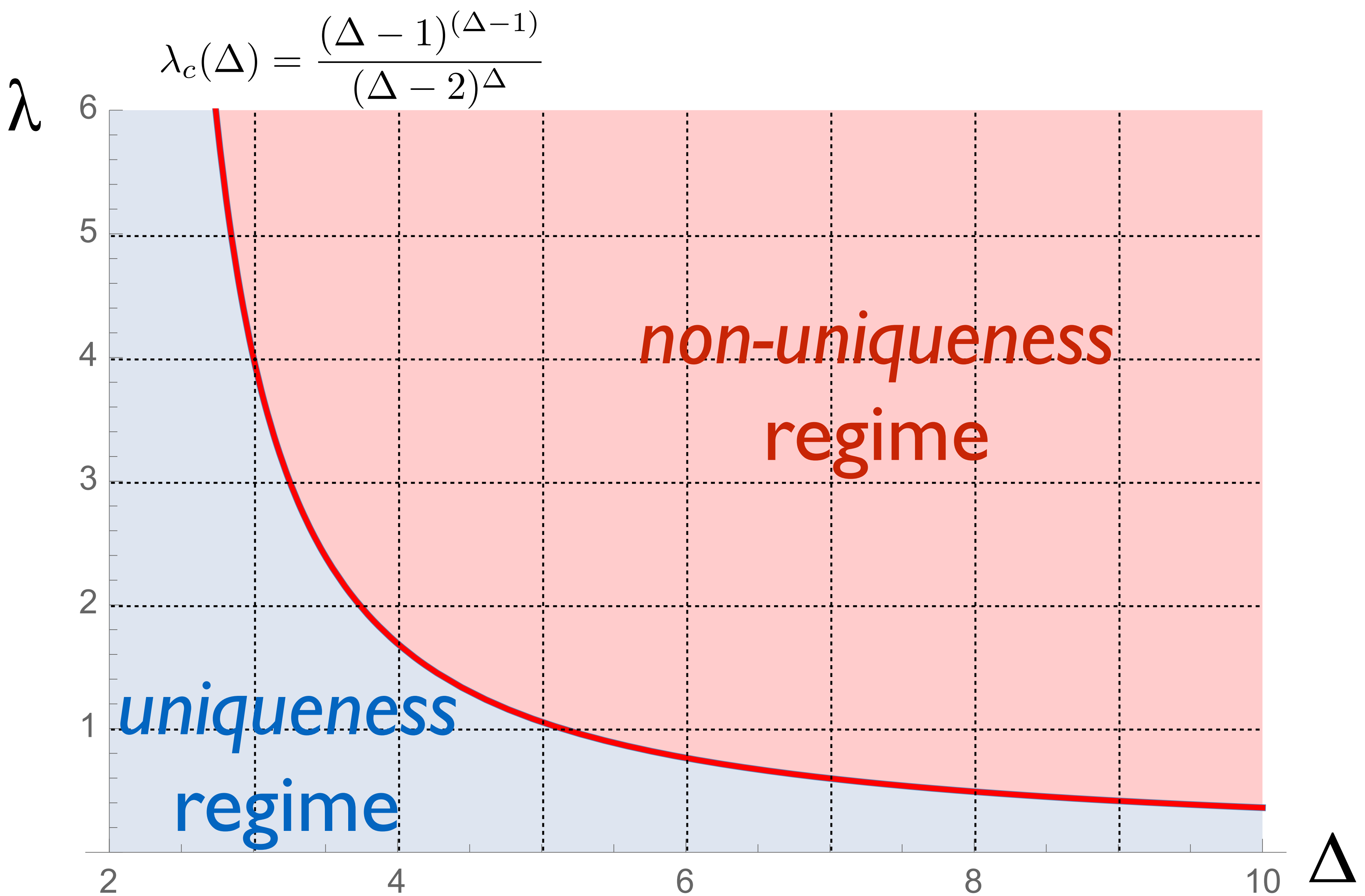
Critical phenomenon:

$\Pr[v \in I \mid \sigma]$ is independent of σ when $l \to \infty$

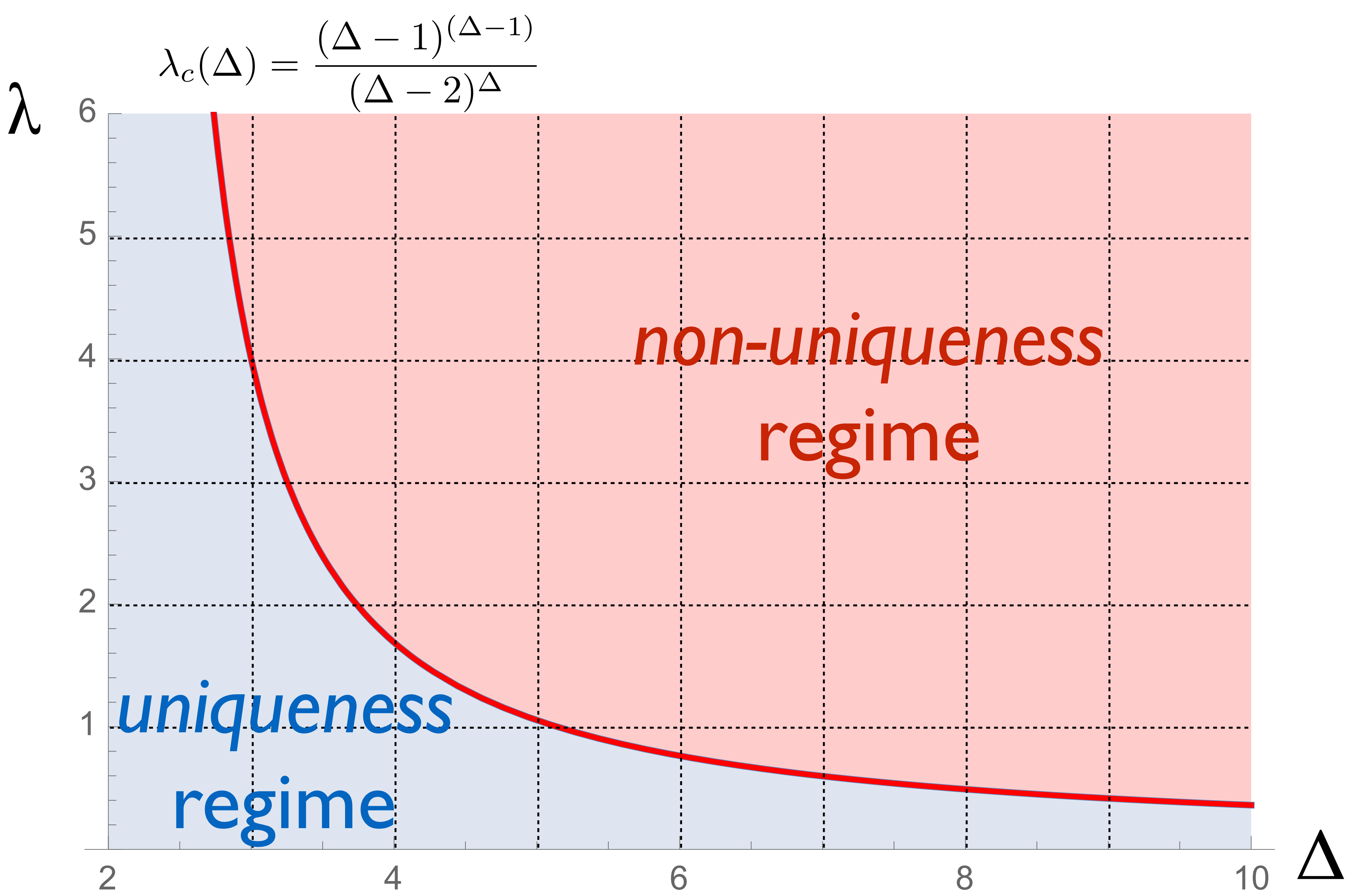
iff $\lambda \leq \lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta - 2}$

Uniqueness threshold

Phase Transition



Phase Transition



sampling (with bounded error) from the hardcore model on graphs of max-degree Δ and fugacity $\lambda>0$

Glauber Dynamics

[Glauber'63]

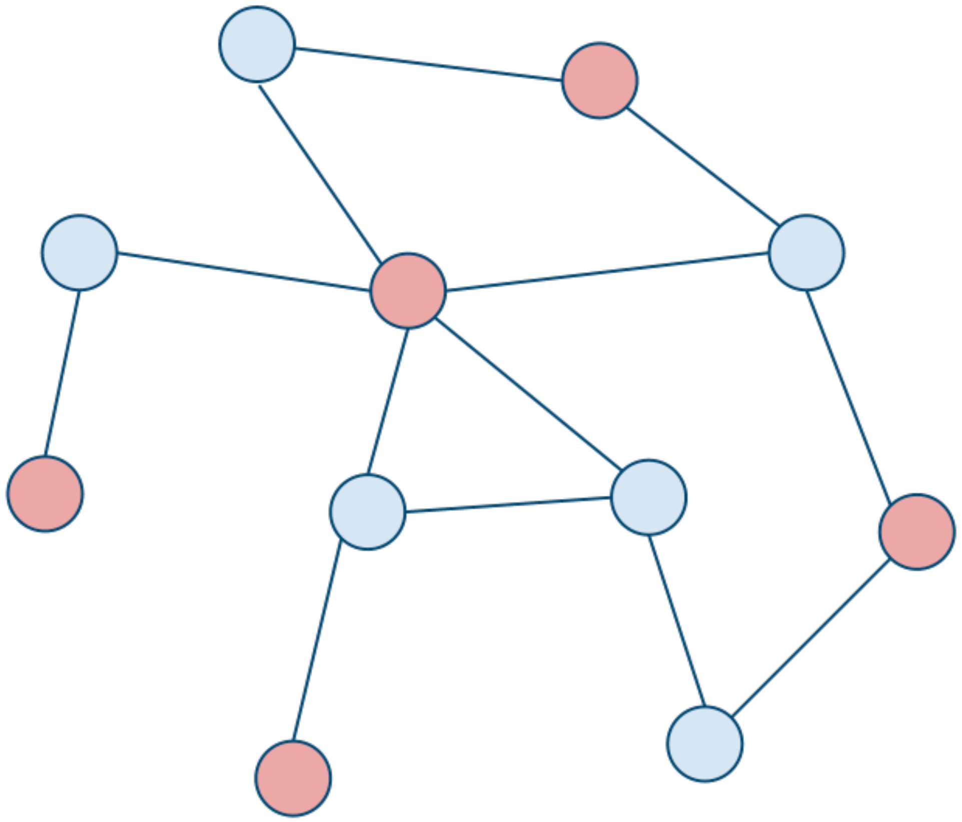
Markov chain $\{X_t\}_{t=0,1,2,\ldots}$ over independent sets in $G(V,E)$

transition $X_t \to X_{t+1}$:

pick a uniform random vertex v;
if $u \notin X_t$ for all v's neighbors u:

$$X_{t+1} = \begin{cases} X_t \cup \{v\} & \text{with prob. } \frac{\lambda}{1+\lambda} \\ X_t \setminus \{v\} & \text{with prob. } \frac{1}{1+\lambda} \end{cases}$$

else $X_{t+1} = X_t$;



Glauber Dynamics

[Glauber'63]

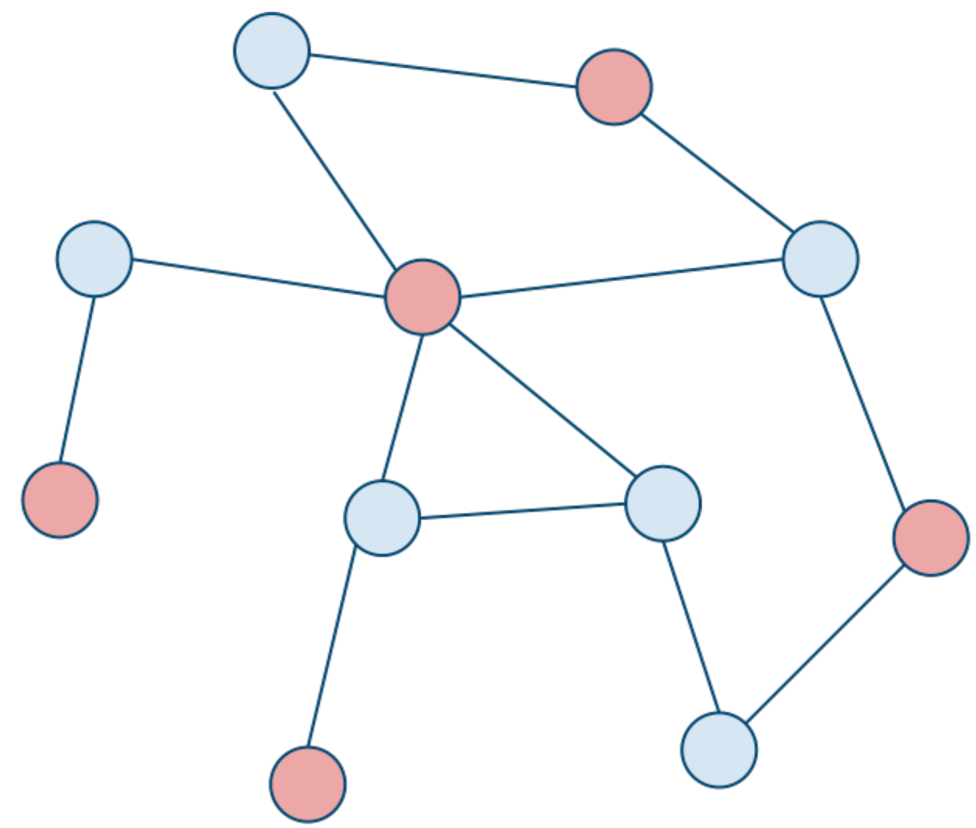
Markov chain $\{X_t\}_{t=0,1,2,\ldots}$ over independent sets in $G(V,E)$

transition $X_t \to X_{t+1}$:

pick a uniform random vertex v;
if $u \notin X_t$ for all v's neighbors u:

$$X_{t+1} = \begin{cases} X_t \cup \{v\} & \text{with prob. } \frac{\lambda}{1+\lambda} \\ X_t \setminus \{v\} & \text{with prob. } \frac{1}{1+\lambda} \end{cases}$$

else $X_{t+1} = X_t$;



ergodic, *irreducible*
time-reversible w.r.t. μ

Markov chain convergence Thm ⟹

$X_t \to \mu$ as $t \to \infty$

Glauber Dynamics

[Glauber'63]

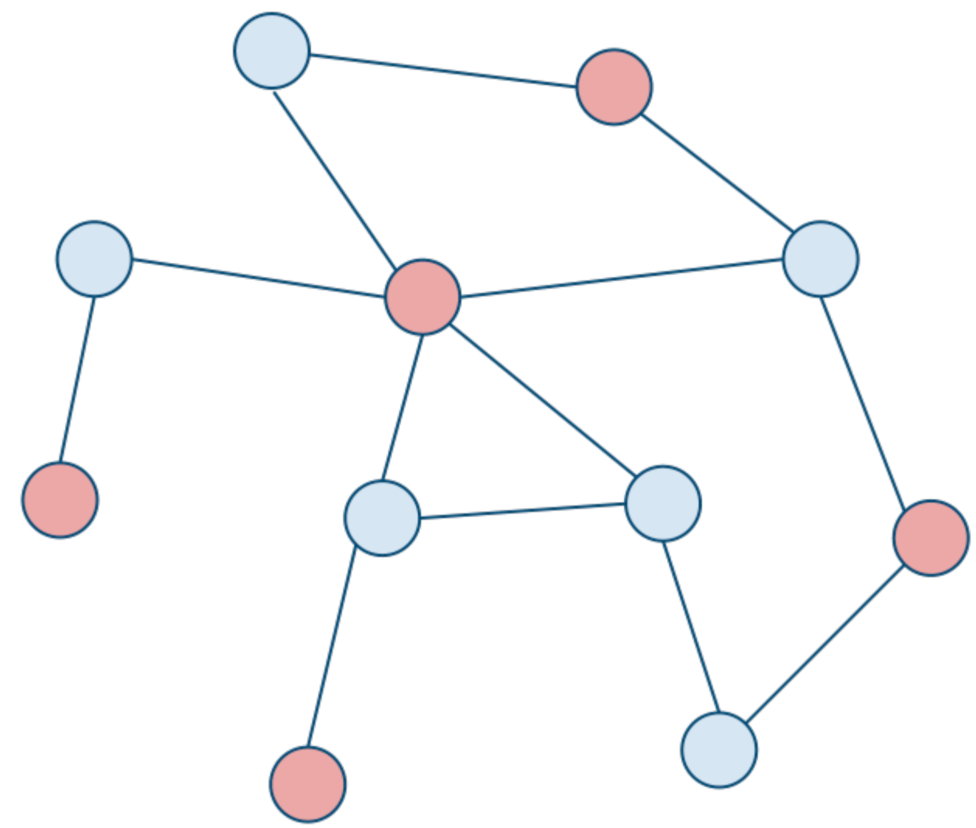
Markov chain $\{X_t\}_{t=0,1,2,\ldots}$ over independent sets in $G(V,E)$

transition $X_t \to X_{t+1}$:

pick a uniform random vertex v;
if $u \notin X_t$ for all v's neighbors u:

$$X_{t+1} = \begin{cases} X_t \cup \{v\} & \text{with prob. } \frac{\lambda}{1+\lambda} \\ X_t \setminus \{v\} & \text{with prob. } \frac{1}{1+\lambda} \end{cases}$$

else $X_{t+1} = X_t$;



ergodic, *irreducible*
time-reversible w.r.t. μ

Markov chain convergence Thm ⟹ $X_t \to \mu$ as $t \to \infty$

mixing time: $\tau_{\text{mix}} = \max_{X_0} \min\{t \mid d_{\text{TV}}(X_t, \mu) < 0.1\}$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] poly-time sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] poly-time ($n^{O(\log\Delta)}$-time) sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] ~~poly-time~~ $n^{O(\log \Delta)}$-time sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$
- [Mossel-Weitz-Wormald'09] slow mixing of Glauber dynamics when $\lambda>\lambda_c(\Delta)$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{\mathrm{e}}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] poly-time $n^{O(\log\Delta)}$-time sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$
- [Mossel-Weitz-Wormald'09] slow mixing of Glauber dynamics when $\lambda>\lambda_c(\Delta)$
- [Sly'10] NP-hardness of sampling when $\lambda>\lambda_c(\Delta)$

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] poly-time $n^{O(\log \Delta)}$-time sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$
- [Mossel-Weitz-Wormald'09] slow mixing of Glauber dynamics when $\lambda>\lambda_c(\Delta)$
- [Sly'10] NP-hardness of sampling when $\lambda>\lambda_c(\Delta)$

Theorem (Efthymiou-Hayes-Štefankovič-Vigoda-Y.'16):

$\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<\lambda_c(\Delta)$,

if Δ is sufficiently large, and there is no small cycles.

sampling (with bounded error) from the hardcore model on graphs with max-degree Δ and fugacity $\lambda>0$

uniqueness threshold: $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}} \approx \frac{e}{\Delta-2}$

- [Glauber'63] discovery of Glauber dynamic
- [Luby-Vigoda'99][Dyer-Greenhill'00][Vigoda'01] $\tau_{\text{mix}}=O(n\log n)$ for Glauber dynamics when $\lambda<2/(\Delta-2)$
- [Weitz'06] poly-time sampler when $\lambda<\lambda_c(\Delta)$ for $\Delta=O(1)$ ($n^{O(\log \Delta)}$-time)
- [Mossel-Weitz-Wormald'09] slow mixing of Glauber dynamics when $\lambda>\lambda_c(\Delta)$
- [Sly'10] NP-hardness of sampling when $\lambda>\lambda_c(\Delta)$

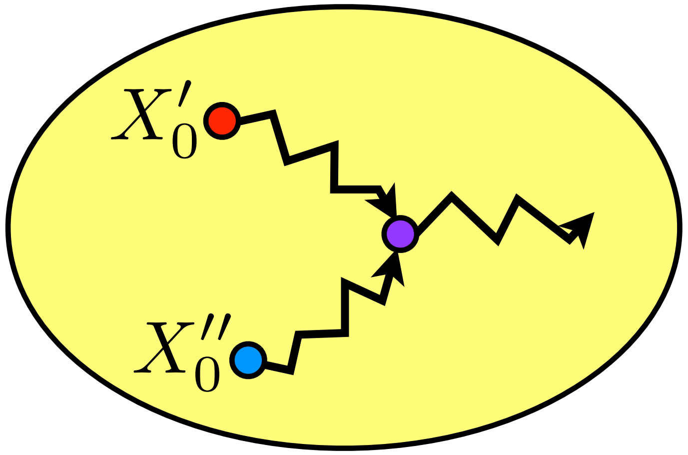
Theorem (Efthymiou-Hayes-Štefankovič-Vigoda-Y.'16):
$\forall\delta>0$, $\exists\Delta_0$ s.t. $\tau_{\text{mix}}=O(n \log n)$ for all graphs of max-degree $\Delta\geq\Delta_0$ and girth ≥ 7 when $\lambda\leq(1-\delta)\lambda_c(\Delta)$.

Coupling

a coupling of a Markov chain X_t is a joint Markov chain (X'_t, X''_t):

- both X_t' and X_t'' follow the same transition rule as X_t
- once collides, makes identical moves thereafter

$$X'_t = X''_t \Rightarrow X'_{t+1} = X''_{t+1}$$

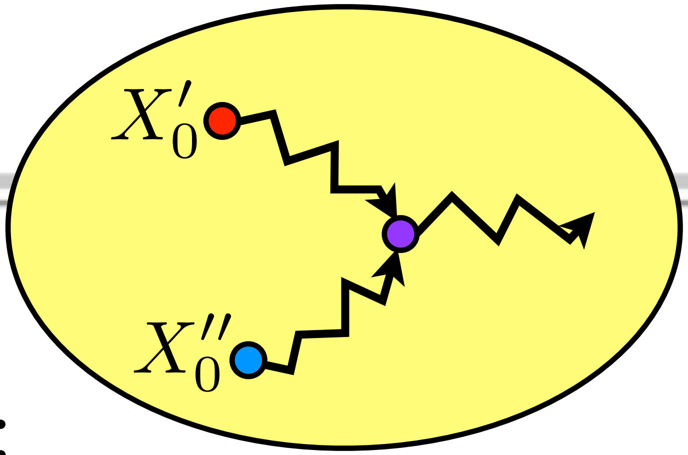


Coupling

a coupling of a Markov chain X_t is a joint Markov chain (X'_t, X''_t):

- both X_t' and X_t'' follow the same transition rule as X_t
- once collides, makes identical moves thereafter

$$X'_t = X''_t \Rightarrow X'_{t+1} = X''_{t+1}$$



Theorem (Griffeath'78, Aldous'83):

$\forall$ coupling (X_t', X_t'') of Markov chain X_t :

$$\tau_{\text{mix}} \le \max_{X'_0, X''_0} \min\{t \mid \Pr[X'_t \neq X''_t] < 0.1\}$$

Path Coupling

Theorem (Bubley-Dyer'97):

If there is a metric $\Phi(\cdot,\cdot)$ over all states for the chain and a coupling $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ for *adjacent* (X_t, Y_t)

s.t.: $$\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$$

$$\Rightarrow \quad \tau_{\mathsf{mix}} = O\left(\frac{n}{\alpha}\log n\right)$$

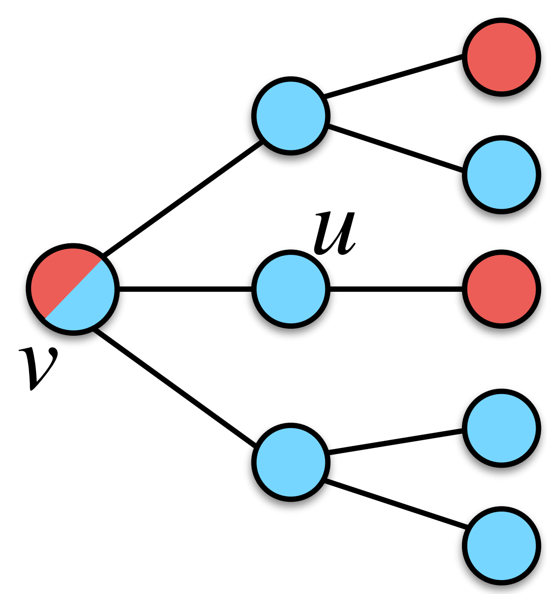
Path Coupling

Theorem (Bubley-Dyer'97):

If there is a metric $\Phi(\cdot,\cdot)$ over all states for the chain and a coupling $(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$ for *adjacent* (X_t, Y_t)

s.t.: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \le (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

$$\Longrightarrow \quad \tau_{\text{mix}} = O\left(\frac{n}{\alpha} \log n\right)$$



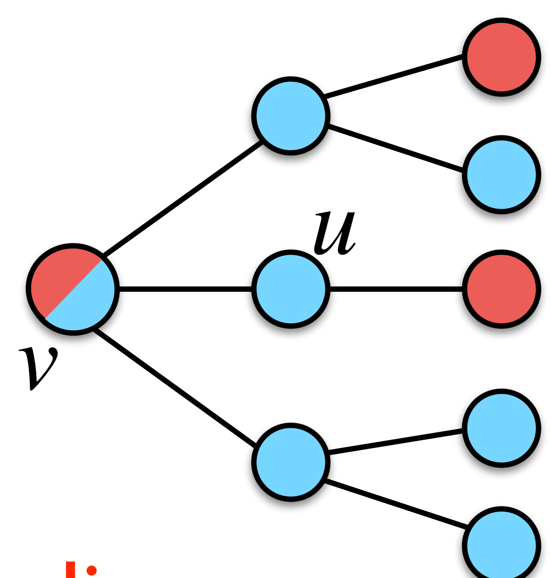
Path Coupling

Theorem (Bubley-Dyer'97):

If there is a metric $\Phi(\cdot,\cdot)$ over all states for the chain and a coupling $(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$ for *adjacent* (X_t, Y_t)

s.t.: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

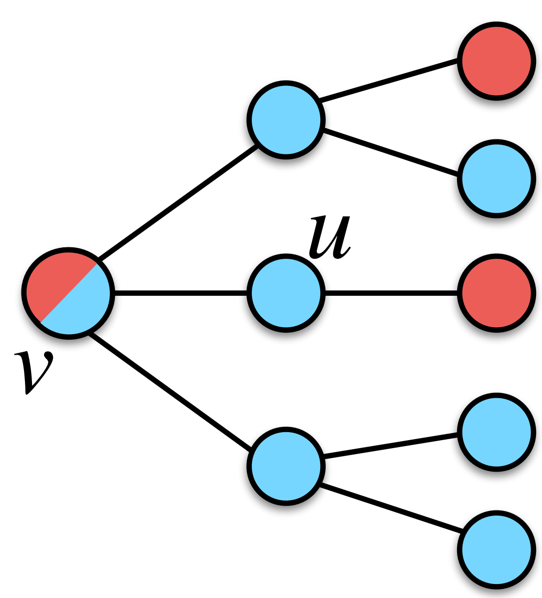
$\Rightarrow$ $\tau_{\text{mix}} = O\left(\frac{n}{\alpha} \log n\right)$



one-step optimal coupling:

$\Pr[X_t(u) \neq X_t(u) \mid u \text{ is picked in step } t]$ is minimized one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

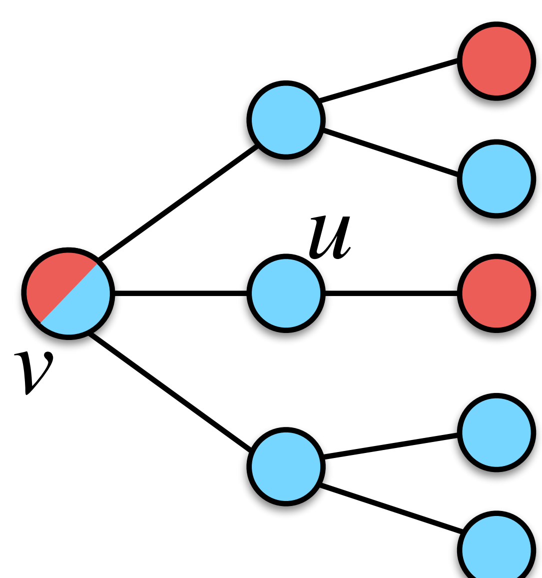


metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v \in X \oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

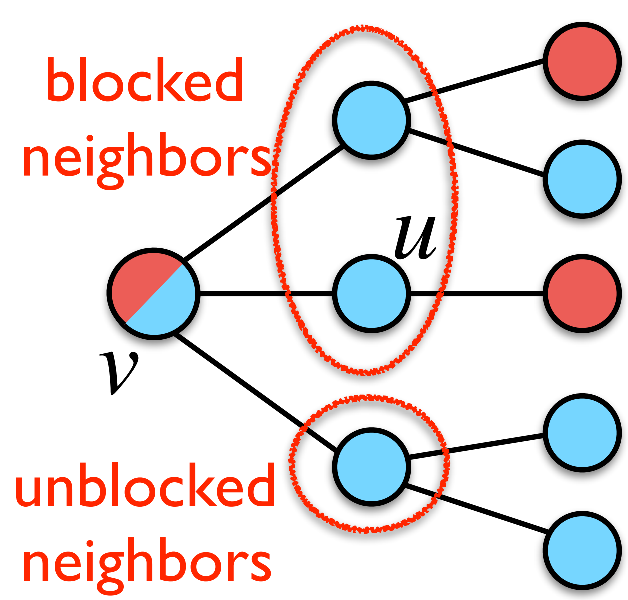


metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v \in X \oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$



$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$

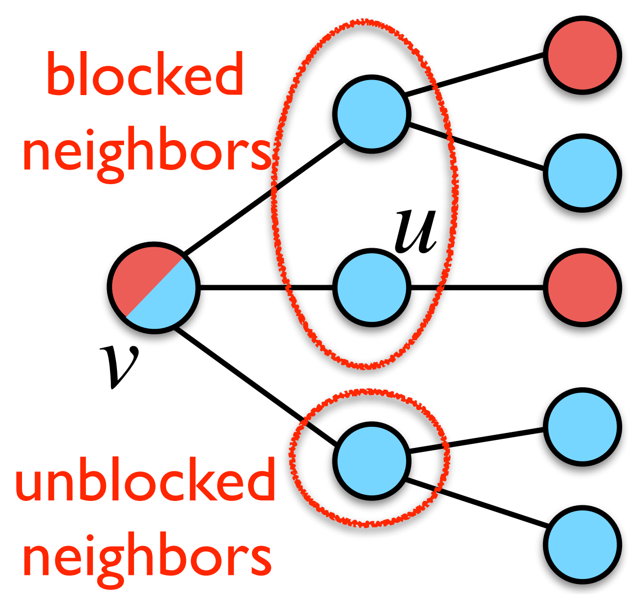
metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v \in X \oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

$$\sum_{u \in N(v)} \frac{\lambda}{1+\lambda} W_u^{(t)} \Phi_u \leq (1-\alpha)\Phi_v$$



$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

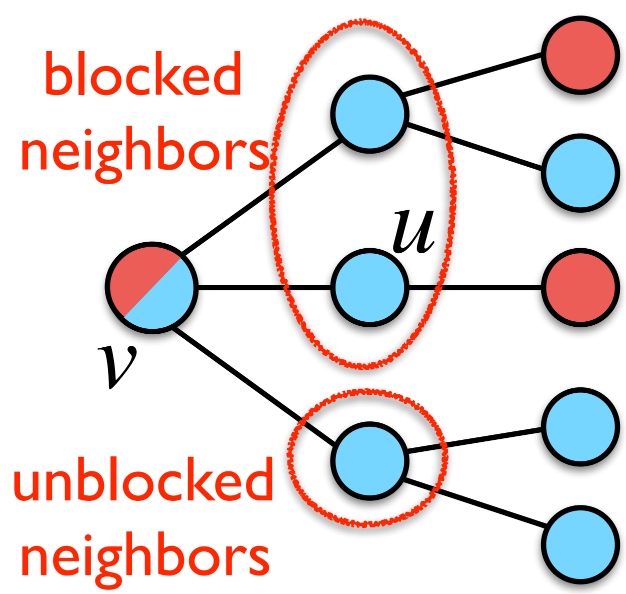
metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v\in X\oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1-\frac{\alpha}{n})\Phi(X_t, Y_t)$

$$\sum_{u\in N(v)} \frac{\lambda}{1+\lambda} W_u^{(t)} \Phi_u \leq \sum_{u\in N(v)} \frac{\lambda W_u^{(t)}}{1+\lambda W_u^{(t)}} \Phi_u \leq (1-\alpha)\Phi_v$$



$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

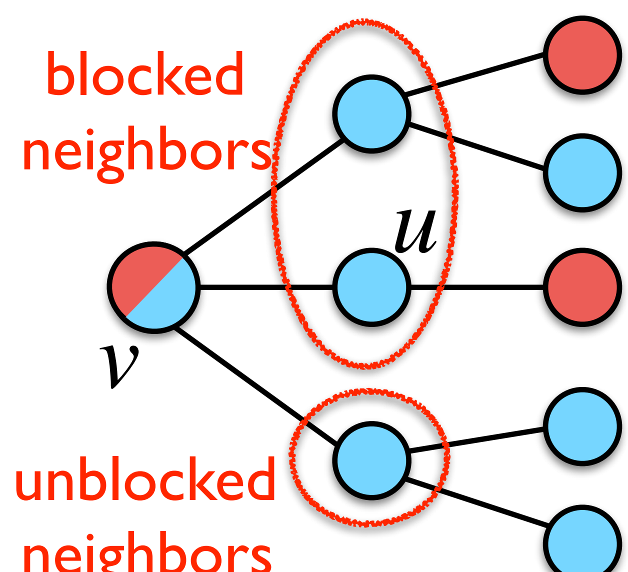
metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v\in X\oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1},Y_{t+1}) \mid X_t, Y_t] \leq (1-\frac{\alpha}{n})\Phi(X_t,Y_t)$

$$\sum_{u\in N(v)} \frac{\lambda W_u^{(t)}}{1+\lambda W_u^{(t)}}\Phi_u \leq (1-\alpha)\Phi_v$$



$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

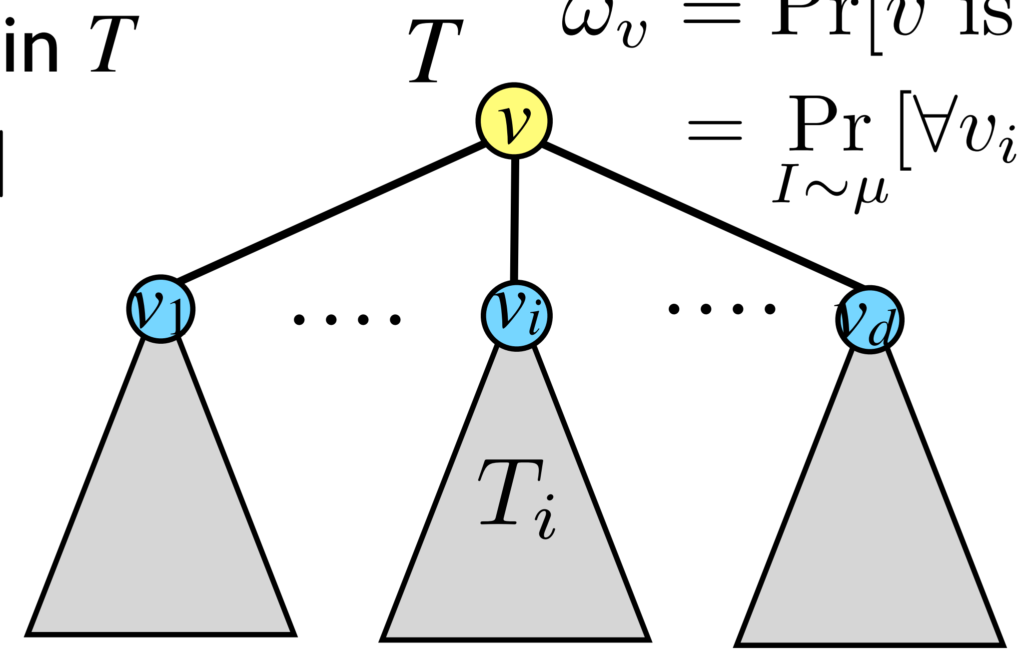
Belief Propagation

[Pearl'82]

independent set I in T

$$\mu(I) \propto \lambda^{|I|}$$

$$\omega_v = \Pr[v \text{ is unblocked } \mid v \notin I]$$
$$= \Pr_{I\sim\mu}[\forall v_i \in N(v), v_i \notin I \mid v \notin I]$$

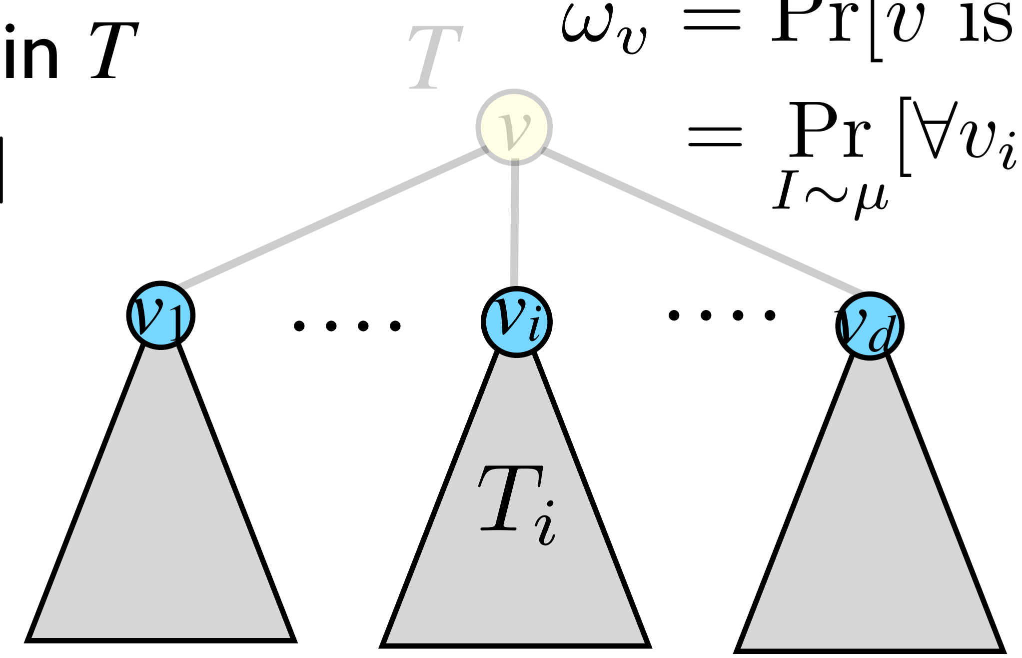


Belief Propagation

[Pearl'82]

independent set I in T

$$\mu(I) \propto \lambda^{|I|}$$



$$\omega_v = \Pr[v \text{ is unblocked } \mid v \notin I]$$
$$= \Pr_{I\sim\mu}[\forall v_i \in N(v), v_i \notin I \mid v \notin I]$$

$\omega_i = \Pr[v_i \text{ is unblocked } \mid v \notin I]$ in T_i

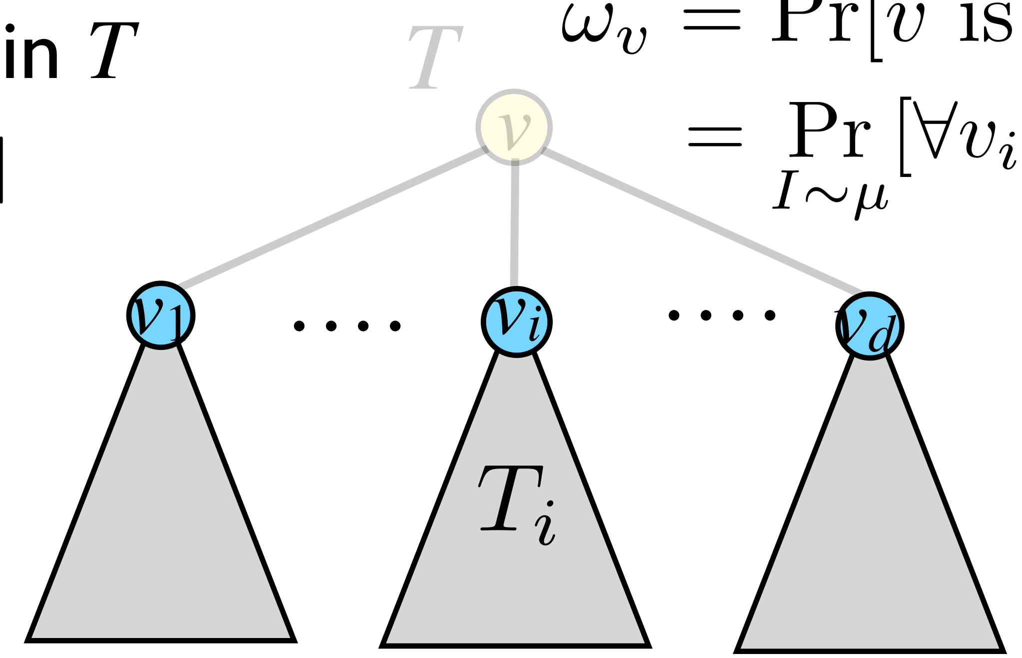
Belief Propagation

[Pearl'82]

independent set I in T

$$\mu(I) \propto \lambda^{|I|}$$

$$\omega_v = \Pr[v \text{ is unblocked } \mid v \notin I] = \Pr_{I \sim \mu}[\forall v_i \in N(v), v_i \notin I \mid v \notin I]$$



$\omega_i = \Pr[v_i \text{ is unblocked } \mid v \notin I]$ in T_i

Belief propagation (BP):

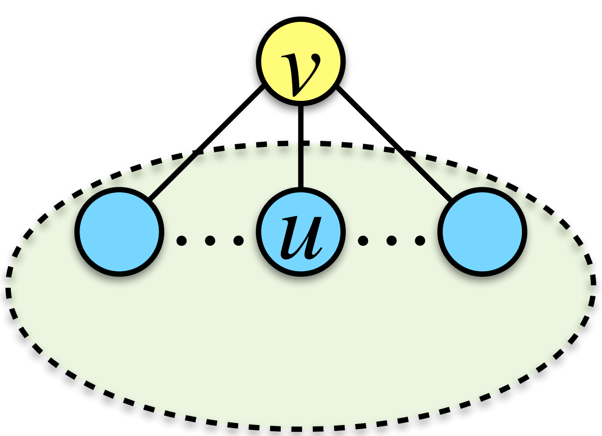
$$\omega_v = \prod_{i=1}^{d} \frac{1}{1 + \lambda \omega_i}$$

Loopy Belief Propagation

graph $G(V, E)$ with max-degree ≤Δ $\qquad \vec{\omega} \in [0,1]^V$

loopy BP: $\vec{\omega}^{(t+1)} = F(\vec{\omega}^{(t)})$

where ∀v: $\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$

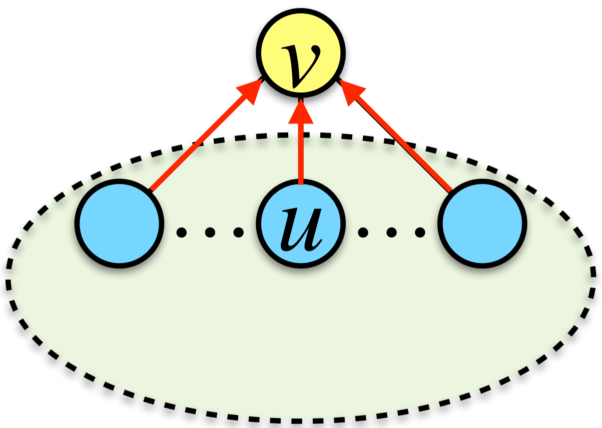


Loopy Belief Propagation

graph $G(V, E)$ with max-degree $\leq\Delta$ $\qquad \vec{\omega} \in [0,1]^V$

loopy BP: $\vec{\omega}^{(t+1)} = F(\vec{\omega}^{(t)})$

where $\forall v$: $\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$

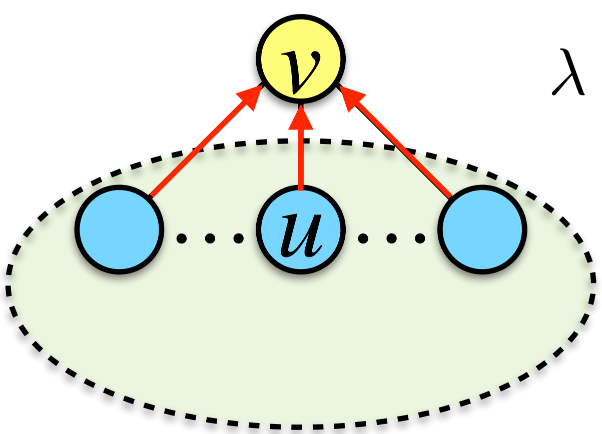


Loopy Belief Propagation

graph $G(V,E)$ with max-degree $\leq\Delta$ $\qquad \vec{\omega}\in[0,1]^V$

loopy BP: $\vec{\omega}^{(t+1)} = F(\vec{\omega}^{(t)})$

where $\forall v$: $\omega_v^{(t+1)} = \prod_{u\in N(v)} \frac{1}{1+\lambda\omega_u^{(t)}}$



$\lambda < \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}}$

$\approx \frac{\mathrm{e}}{\Delta-1}$

$\forall$initial values $\vec{\omega}^{(0)}\in[0,1]^V$

$\vec{\omega}^{(t)}$ converges to a unique fixpoint $\omega^* = F(\omega^*)$

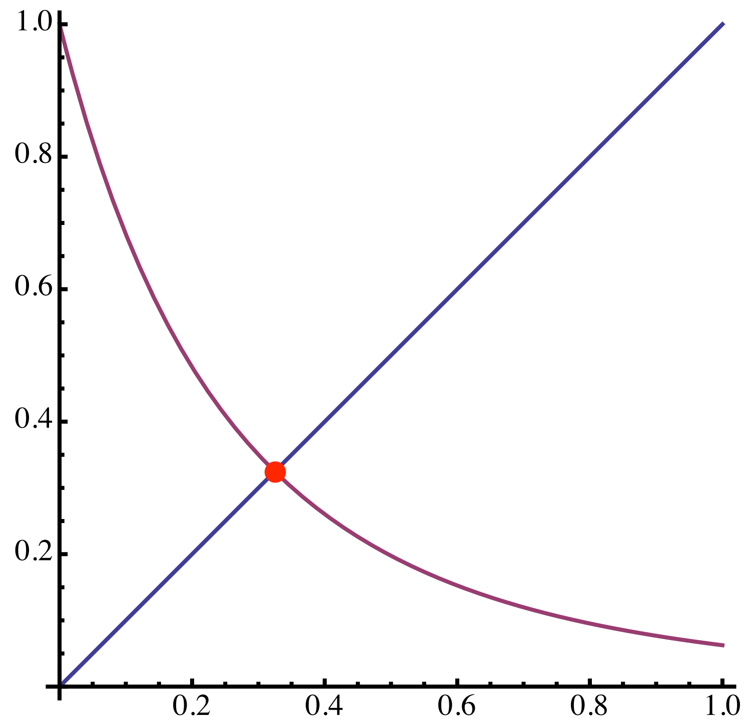
at an exponential rate.

BP Convergence

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

univariate dynamical system:

$$f(x) = (1 + \lambda x)^{-\Delta}$$

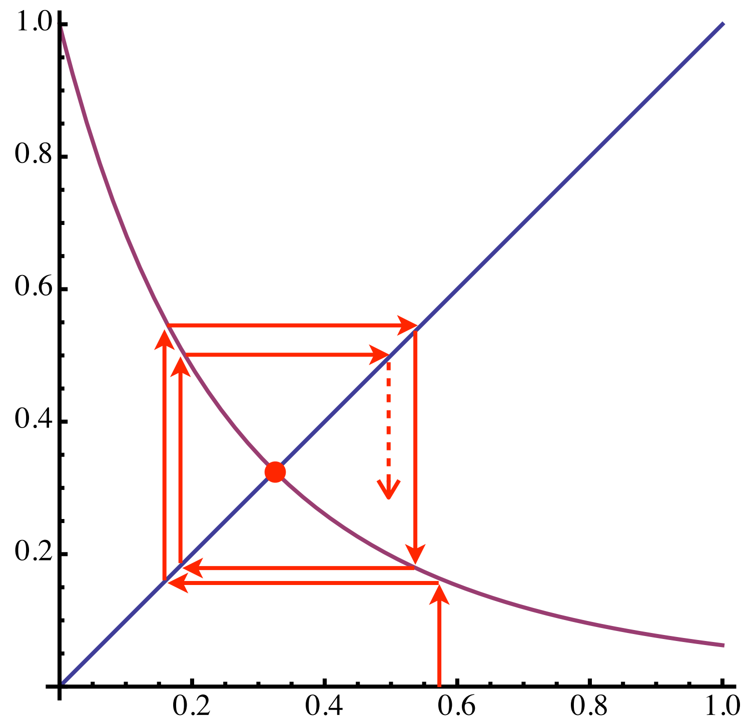


BP Convergence

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

univariate dynamical system:

$$f(x) = (1 + \lambda x)^{-\Delta}$$



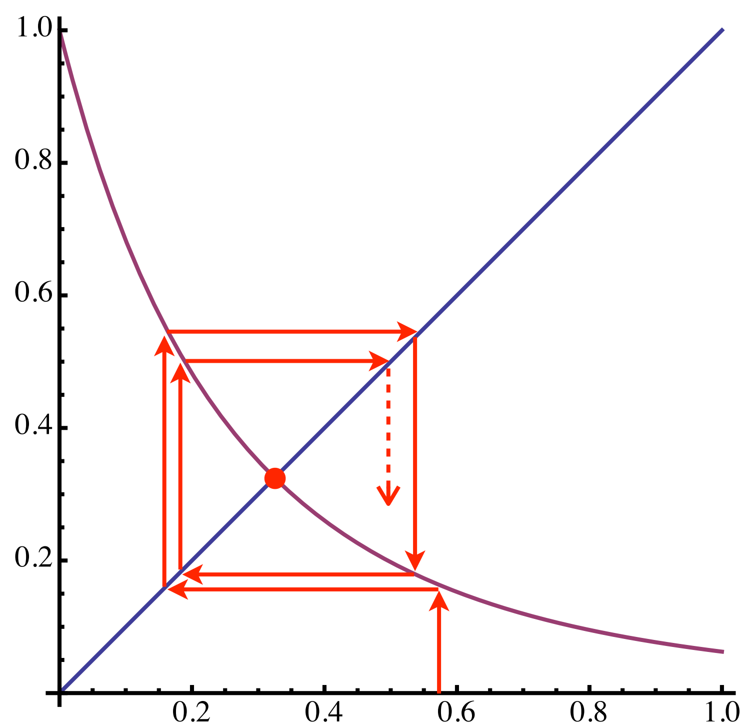
$$\lambda < \frac{\Delta^\Delta}{(\Delta - 1)^{\Delta+1}}$$

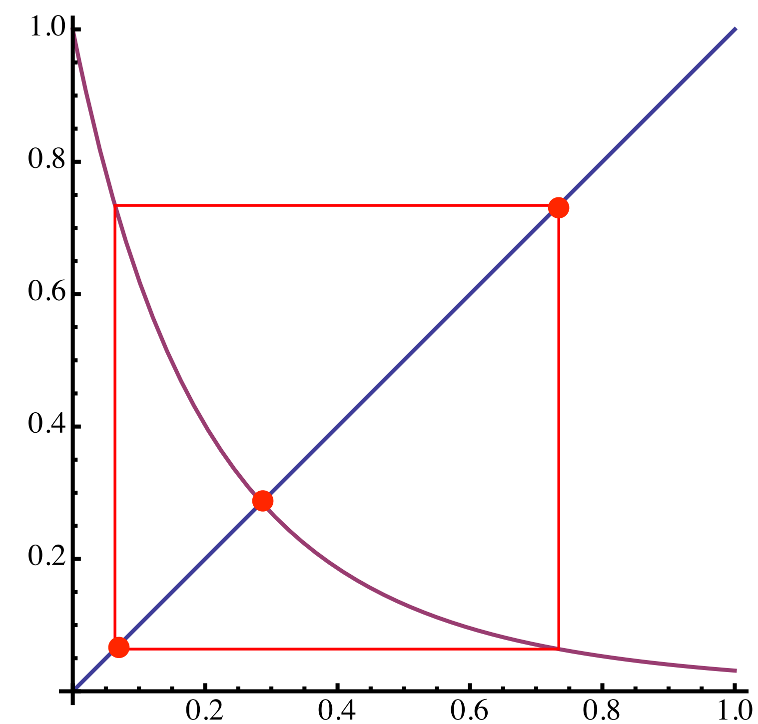
BP Convergence

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1+\lambda \omega_u^{(t)}}$$

univariate dynamical system:

$$f(x) = (1+\lambda x)^{-\Delta}$$





$$\lambda < \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}}$$

$$\lambda > \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}}$$

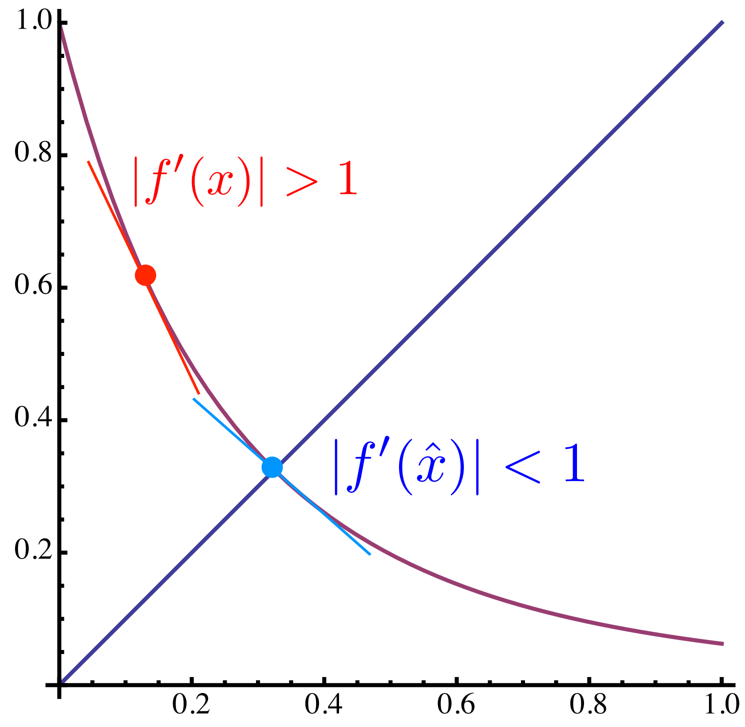
Potential Method

$$f(x) = (1 + \lambda x)^{-\Delta}$$

$\uparrow$

x

$$f(x) = (1 + \lambda x)^{-\Delta}$$



Potential Method

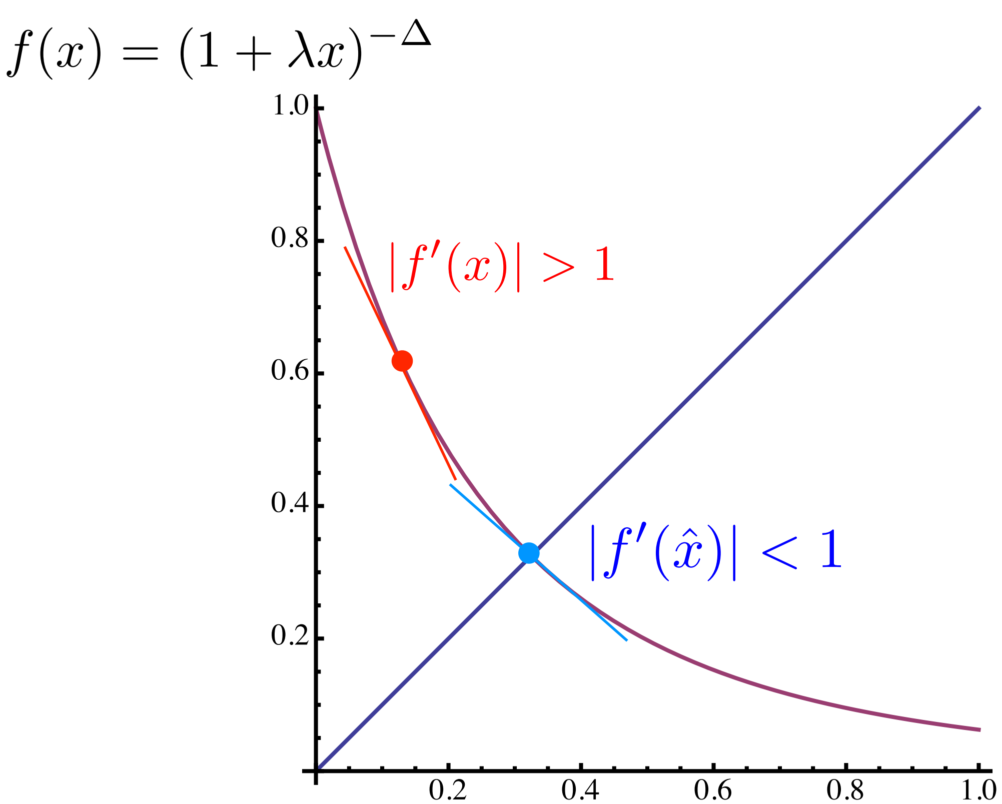
$$f(x) = (1+\lambda x)^{-\Delta} \xrightarrow{\ \psi\ } g(y)$$

$$\uparrow$$

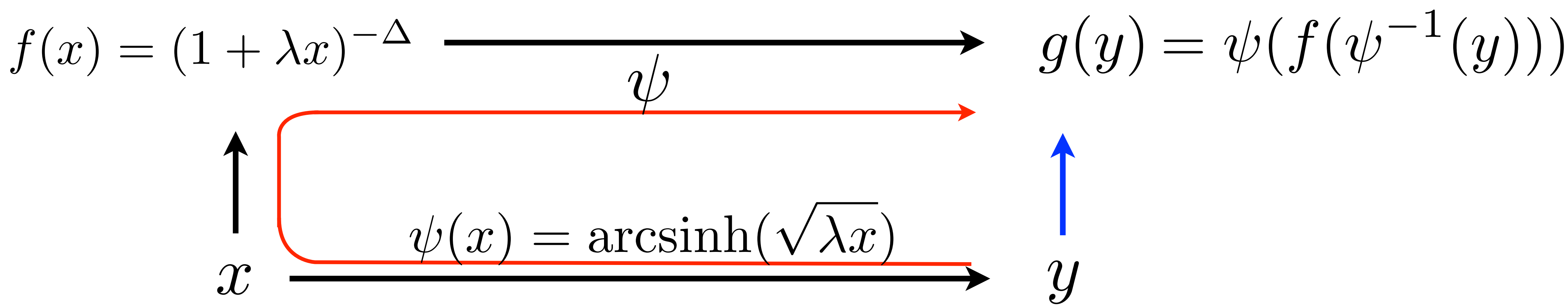
$$x \xrightarrow{\ \psi(x) = \operatorname{arcsinh}(\sqrt{\lambda x})\ } y$$

used in

[Li-Lu-Y.'13] [Sinclair-Srivastava-Y.'13] [Sinclair-Srivastava-Štefankovič-Y.'15]

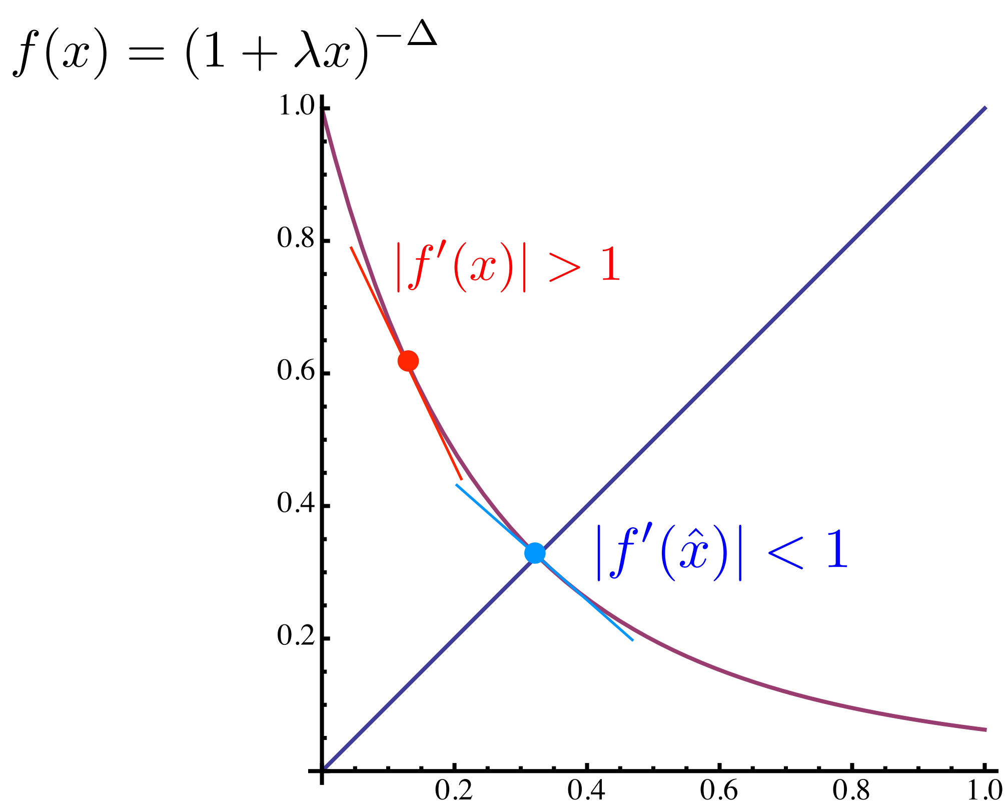


Potential Method



used in

[Li-Lu-Y.'13] [Sinclair-Srivastava-Y.'13] [Sinclair-Srivastava-Štefankovič-Y.'15]



Potential Method

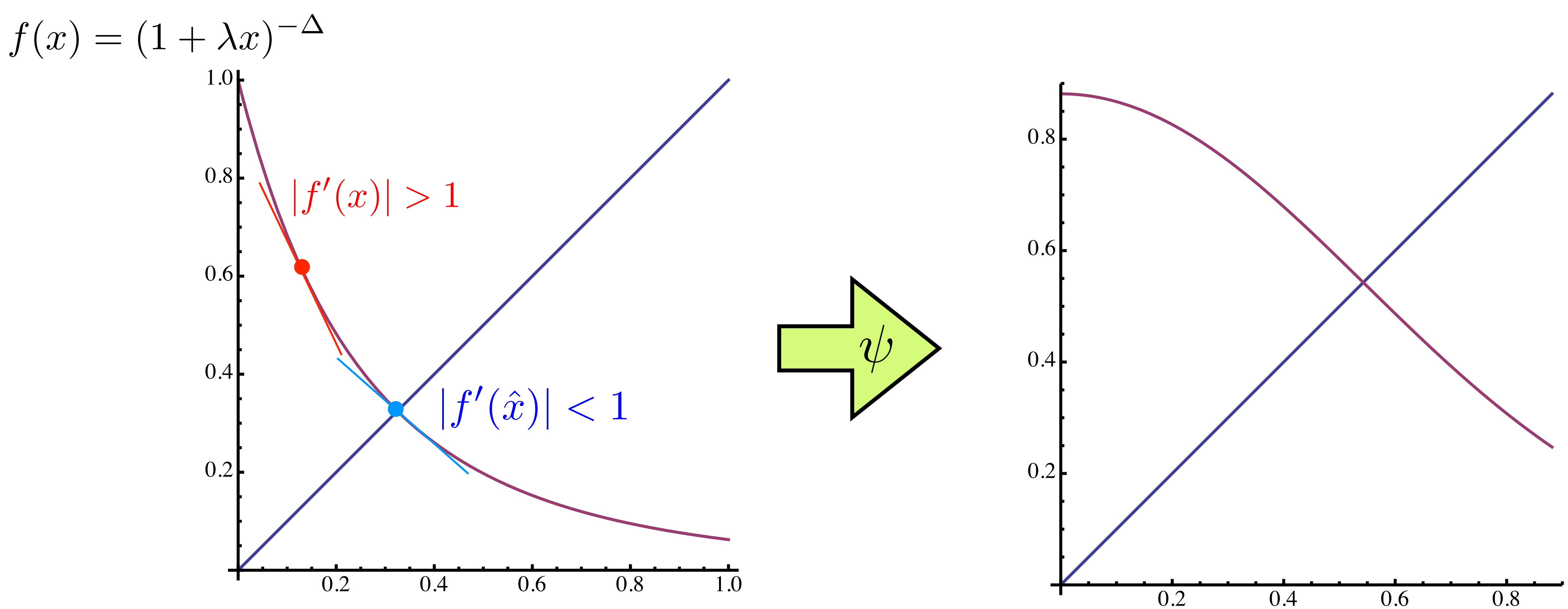
$$f(x) = (1+\lambda x)^{-\Delta} \xrightarrow{\psi} g(y) = \psi(f(\psi^{-1}(y)))$$

$$x \longrightarrow y$$

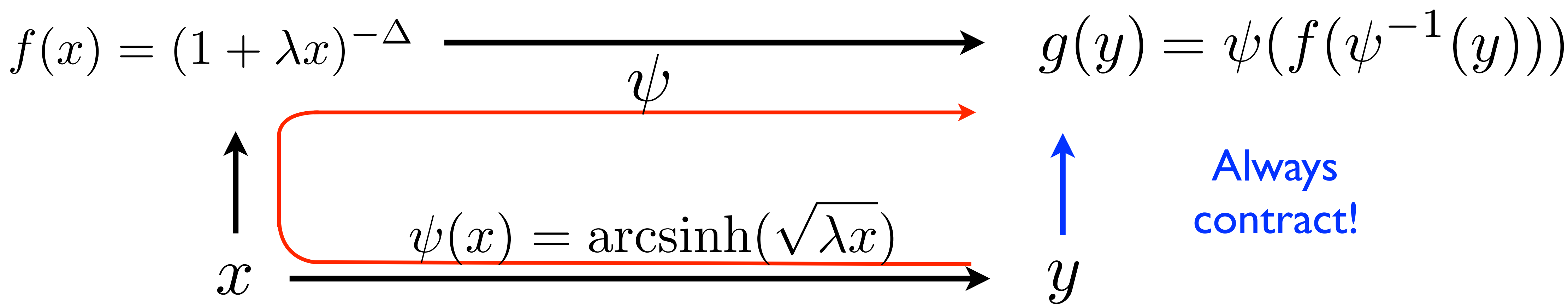
$$\psi(x) = \operatorname{arcsinh}(\sqrt{\lambda x})$$

used in

[Li-Lu-Y.'13] [Sinclair-Srivastava-Y.'13] [Sinclair-Srivastava-Štefankovič-Y.'15]

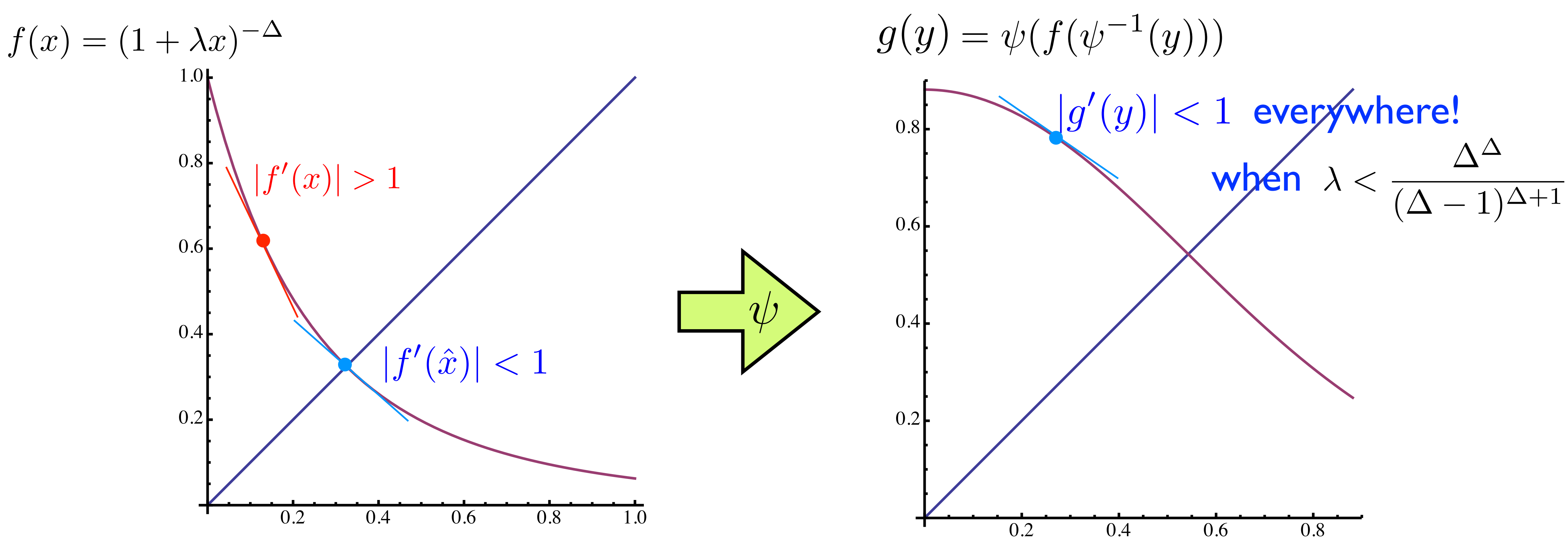


Potential Method



used in

[Li-Lu-Y.'13] [Sinclair-Srivastava-Y.'13] [Sinclair-Srivastava-Štefankovič-Y.'15]



original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

potentials: $\chi_v = \psi(\omega_v) = \mathrm{arcsinh}(\sqrt{\lambda \omega})$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

BP for potentials:

$$\chi_v^{(t+1)} = \psi\left(\prod_{u \in N(v)} \frac{1}{1 + \lambda \cdot \psi^{-1}(\chi_u^{(t)})} \right)$$

potentials: $\chi_v = \psi(\omega_v) = \operatorname{arcsinh}(\sqrt{\lambda \omega})$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

BP for potentials:

$$\chi_v^{(t+1)} = \psi\left(\prod_{u \in N(v)} \frac{1}{1 + \lambda \cdot \psi^{-1}(\chi_u^{(t)})} \right)$$

potentials: $\chi_v = \psi(\omega_v) = \text{arcsinh}(\sqrt{\lambda\omega})$

Jacobian for potentials : $J_{vu} = \frac{\partial \chi_v^{(t+1)}}{\partial \chi_v^{(t)}} = \frac{\partial \omega_v^{(t+1)}}{\partial \omega_v^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})}$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1 + \lambda \omega_u^{(t)}}$$

BP for potentials:

$$\chi_v^{(t+1)} = \psi\left(\prod_{u \in N(v)} \frac{1}{1 + \lambda \cdot \psi^{-1}(\chi_u^{(t)})}\right)$$

potentials: $\chi_v = \psi(\omega_v) = \text{arcsinh}(\sqrt{\lambda \omega})$

Jacobian for potentials : $J_{vu} = \frac{\partial \chi_v^{(t+1)}}{\partial \chi_v^{(t)}} = \frac{\partial \omega_v^{(t+1)}}{\partial \omega_v^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})}$

Theorem (Sinclair-Srivastava-Štefankovič-Y.'15):

$\lambda < \frac{\Delta^\Delta}{(\Delta - 1)^{\Delta+1}}$ ⇨ $\forall \vec{\omega}^{(t)} \in [0,1]^V$,

$$\|J\|_\infty = \max_v \sum_{u \in N(v)} \frac{\lambda \omega_v^{(t+1)}}{1 + \lambda \omega_u^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})} < 1$$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1+\lambda \omega_u^{(t)}}$$

BP for potentials:

$$\chi_v^{(t+1)} = \psi\left(\prod_{u \in N(v)} \frac{1}{1+\lambda \cdot \psi^{-1}(\chi_u^{(t)})} \right)$$

potentials: $\chi_v = \psi(\omega_v) = \operatorname{arcsinh}(\sqrt{\lambda \omega})$

Jacobian for potentials : $J_{vu} = \frac{\partial \chi_v^{(t+1)}}{\partial \chi_v^{(t)}} = \frac{\partial \omega_v^{(t+1)}}{\partial \omega_v^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})}$

Theorem (Sinclair-Srivastava-Štefankovič-Y.'15):

$\lambda < \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}}$ ⇒ $\forall \vec{\omega}^{(t)} \in [0,1]^V$,

$$\|J\|_\infty = \max_v \sum_{u \in N(v)} \frac{\lambda \omega_v^{(t+1)}}{1+\lambda \omega_u^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})} < 1$$

$\vec{\omega}^{(t)}$ converges to a unique fixpoint $\omega^* = F(\omega^*)$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1+\lambda \omega_u^{(t)}}$$

potentials:

$$\chi_v = \psi(\omega_v) = \operatorname{arcsinh}(\sqrt{\lambda \omega})$$

Theorem (Sinclair-Srivastava-Štefankovič-Y.'15):

$$\lambda < \frac{\Delta^{\Delta}}{(\Delta-1)^{\Delta+1}} \Rightarrow \forall \vec{\omega}^{(t)} \in [0,1]^V,$$

$$\|J\|_\infty = \max_v \sum_{u \in N(v)} \frac{\lambda \omega_v^{(t+1)}}{1+\lambda \omega_u^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})} < 1$$

$\vec{\omega}^{(t)}$ converges to a unique fixpoint $\omega^* = F(\omega^*)$

$$\sum_{u \in N(v)} \frac{\lambda \omega_u^*}{1+\lambda \omega_u^*} \frac{1}{\omega_u^* \psi'(\omega_u^*)} < \frac{1}{\omega_v^* \psi'(\omega_v^*)}$$

original BP:

$$\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1+\lambda \omega_u^{(t)}}$$

potentials:

$$\chi_v = \psi(\omega_v) = \text{arcsinh}(\sqrt{\lambda\omega})$$

Theorem (Sinclair-Srivastava-Štefankovič-Y.'15):

$$\lambda < \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}} \Rightarrow \forall \vec{\omega}^{(t)} \in [0,1]^V,$$

$$\|J\|_\infty = \max_v \sum_{u \in N(v)} \frac{\lambda \omega_v^{(t+1)}}{1+\lambda\omega_u^{(t)}} \frac{\psi'(\omega_v^{(t+1)})}{\psi'(\omega_u^{(t)})} < 1$$

$\vec{\omega}^{(t)}$ converges to a unique fixpoint $\omega^* = F(\omega^*)$

$$\sum_{u \in N(v)} \frac{\lambda \omega_u^*}{1+\lambda\omega_u^*} \frac{1}{\omega_u^* \psi'(\omega_u^*)} < \frac{1}{\omega_v^* \psi'(\omega_v^*)}$$

Φ_u Φ_v

define $\Phi_v \triangleq \frac{1}{\omega_v^* \cdot \psi'(\omega_v^*)} = 2\sqrt{\frac{1+\lambda\omega_v^*}{\lambda\omega_v^*}}$

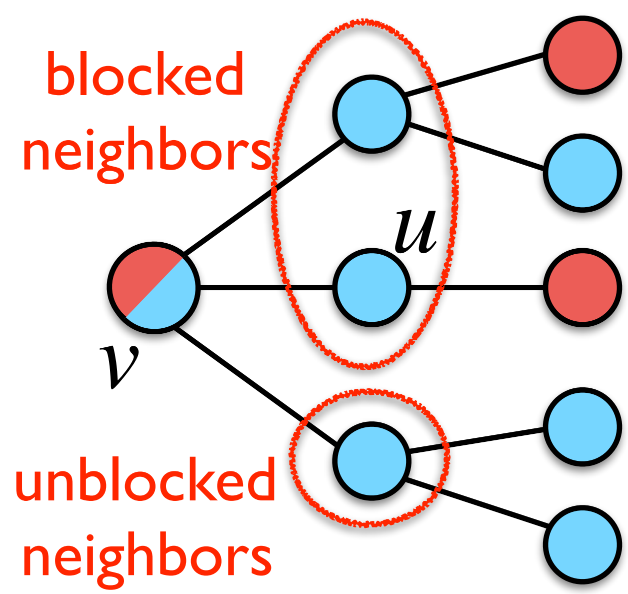
metric $\Phi(\cdot,\cdot)$ is a weighted Hamming distance:

$$\Phi(X,Y) = \sum_{v \in X \oplus Y} \Phi_v$$ where $\Phi_v \geq 1$ is v's weight

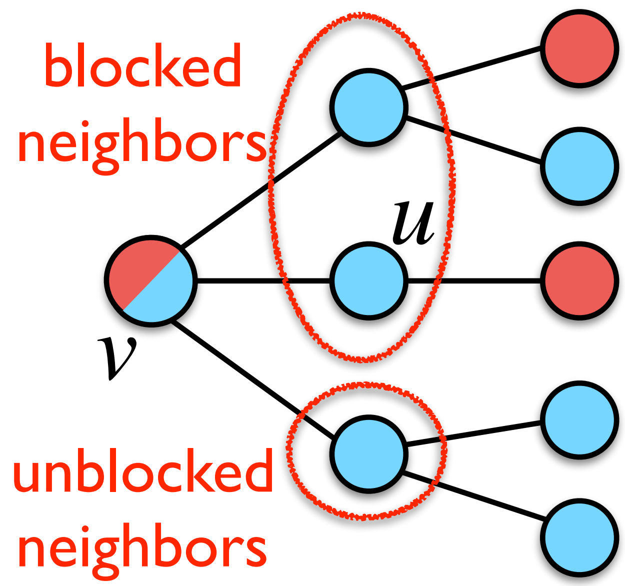
one-step optimal coupling for (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

Goal: $\mathbb{E}[\Phi(X_{t+1}, Y_{t+1}) \mid X_t, Y_t] \leq (1 - \frac{\alpha}{n})\Phi(X_t, Y_t)$

$$\sum_{u \in N(v)} \frac{\lambda W_u^{(t)}}{1 + \lambda W_u^{(t)}} \Phi_u \leq (1-\alpha)\Phi_v$$



$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$

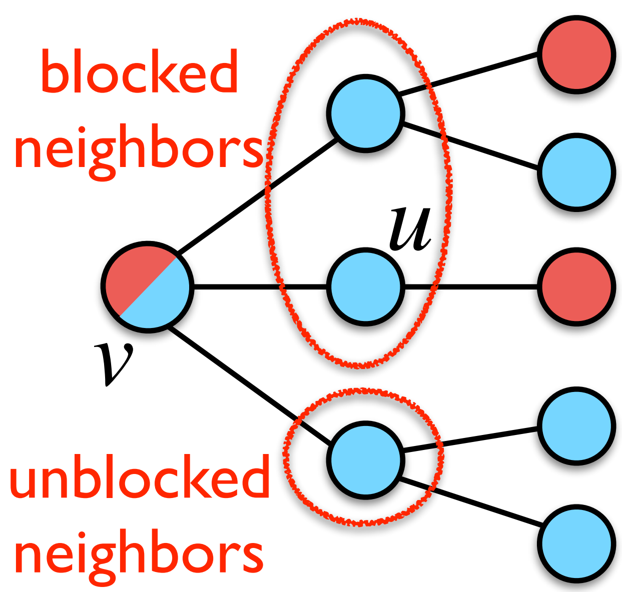


for coupling of Glauber dynamics (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

$\exists$ vertex weights Φ_v :

$$\sum_{u \in N(v)} \frac{\lambda W_u^{(t)}}{1 + \lambda W_u^{(t)}} \Phi_u < \Phi_v$$



for coupling of Glauber dynamics (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

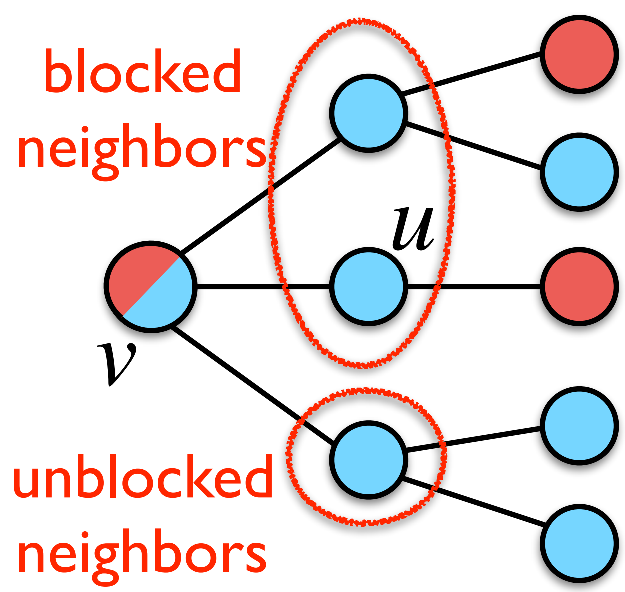
$\exists$ vertex weights Φ_v :

$$\sum_{u \in N(v)} \frac{\lambda W_u^{(t)}}{1 + \lambda W_u^{(t)}} \Phi_u < \Phi_v \quad \Rightarrow \quad \tau_{\text{mix}} = O(n \log n)$$

loopy BP: $\omega_v^{(t+1)} = \prod_{u \in N(v)} \frac{1}{1+\lambda\omega_u^{(t)}}$ fixpoint: $\omega_v^{*} = \prod_{u \in N(v)} \frac{1}{1+\lambda\omega_u^{*}}$

$$\lambda < \frac{\Delta^\Delta}{(\Delta-1)^{\Delta+1}} \Rightarrow \sum_{u \in N(v)} \frac{\lambda\omega_u^*}{1+\lambda\omega_u^*}\Phi_u < \Phi_v$$

where $\Phi_v = 2\sqrt{\frac{1+\lambda\omega_v^*}{\lambda\omega_v^*}}$



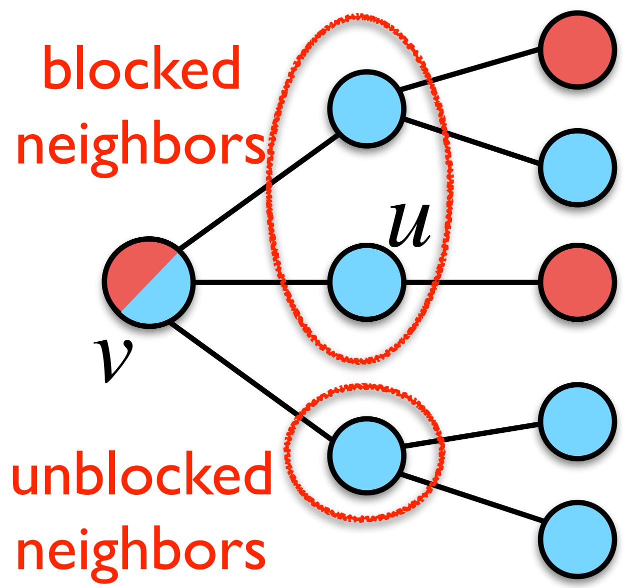
for coupling of Glauber dynamics (X_t, Y_t) that $X_t \oplus Y_t = \{v\}$

$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

$\exists$ vertex weights Φ_v:

$$\sum_{u \in N(v)} \frac{\lambda W_u^{(t)}}{1+\lambda W_u^{(t)}}\Phi_u < \Phi_v \Rightarrow \tau_{\text{mix}} = O(n \log n)$$

Local Uniformity



Glauber dynamics X_t

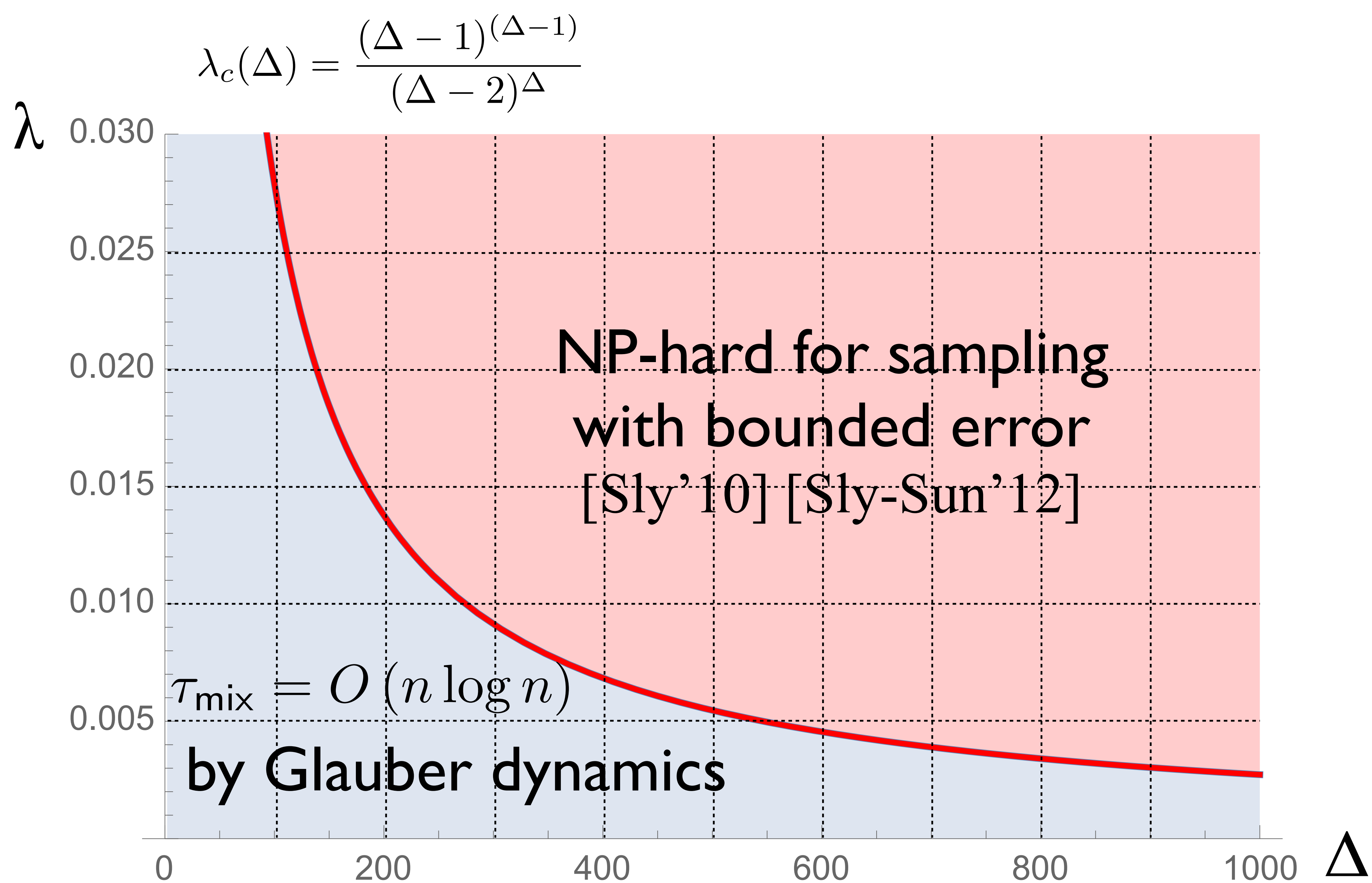
$$W_u^{(t)} = I[u \text{ is unblocked in } X_t]$$

$\forall \epsilon, \delta > 0$, $\exists \Delta_0$ and $\mathcal{I} = [O(n \log \Delta), n \exp(\Omega(\Delta))]$ s.t.
for all graphs of max-degree $\Delta \geq \Delta_0$ and girth ≥ 7:
$\lambda \leq (1-\delta)\lambda_c(\Delta)$ ⇒ $\forall$ vertex v:

$$\Pr\left[\forall t \in \mathcal{I}: \sum_{u \in N(v)} \frac{\lambda W_u^{(t)}}{1+\lambda W_u^{(t)}} \Phi_u < \sum_{u \in N(v)} \frac{\lambda \omega_u^*}{1+\lambda \omega_u^*} \Phi_u + \epsilon\right] \geq 1 - \exp(-\Omega(\Delta))$$

proved by concentration using the arguments in:

[Dyer-Frieze-Hayes-Vigoda'04] [Hayes-Vigoda'05] [Hayes'13]



$\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}}$
λ
0.030
0.025
0.020
0.015
0.010
0.005
0
200
400
600
800
1000
Δ
NP-hard for sampling
with bounded error
[Sly'10] [Sly-Sun'12]
$\tau_{\text{mix}} = O(n \log n)$
by Glauber dynamics

Summary

- Optimal mixing time for Glauber dynamics for the hardcore model in the *uniqueness* regime (when degree is big enough and there is no small cycle).

- Connecting rapid mixing of MCMC sampling with BP convergence:
  - If BP converges, there probably is contraction for Markovian coupling.
  - Path coupling "*from art to science*".

- Open problem: get rid of the degree and girth requirements.

Charis Efthymiou, Thomas P. Hayes, Daniel Štefankovič, Eric Vigoda, Yitong Yin. *Convergence of MCMC and Loopy BP in the Tree Uniqueness Region for the Hard-Core Model*. **FOCS'16**. arxiv: 1604.01422.

Thank you!

Any questions?