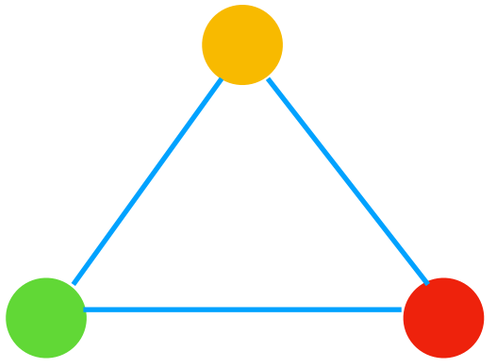
A Deterministic Algorithm for Counting Colorings with 2Δ Colors

Jingcheng Liu (UC Berkeley $\longrightarrow$ Caltech)

Based on joint work with Alistair Sinclair (UC Berkeley) and Piyush Srivastava (TIFR, India)

Approximate counting

- Consider proper q-colorings of a graph G
- Let Δ:=Δ(G) be the maximum degree
- If $q \geq \Delta + 1$:
  - G has a q-coloring, and it can be found efficiently
  - How many q-colorings does G have?

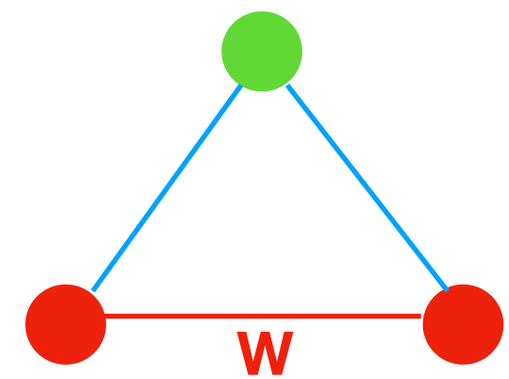


#P-complete, approximation?

Approximate counting

The Potts model

- A graph G=(V,E)
- Configuration $\sigma : V \to [q]$
- The partition function



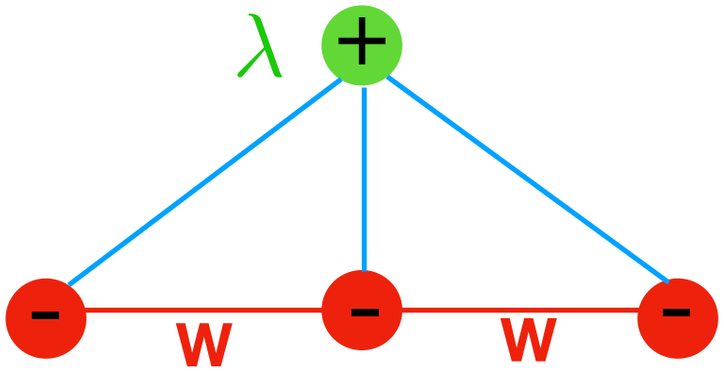
$$Z_G(w) := \sum_{\sigma:V\to[q]} w^{(\#\text{monochromatic edges})}$$
$$= \sum_{\sigma:V\to[q]} w^{|\{\{u,v\}\in E\,:\,\sigma(u)=\sigma(v)\}|}.$$

Example:

- $Z_G(0) = (\#$ valid q-colorings of $G)$
- **q=2: Ising model**

Goal: compute $(1 \pm \varepsilon) \cdot Z_G(w)$ **.**

The Ising model (a.k.a q=2)



- Configuration $\sigma : V \to \{+, -\}$
- External field λ
- The partition function

$$Z_G(w, \lambda) := \sum_{\sigma : V \to \{+,-\}} w^{(\#\text{monochromatic edges})} \lambda^{(\# + \text{-spin vertices})}$$
$$= \sum_{\sigma : V \to \{+,-\}} w^{|E_=(\sigma)|} \lambda^{|V_+(\sigma)|}.$$

- Gibbs distribution: $\Pr[\sigma] = \dfrac{1}{Z_G(w, \lambda)} \cdot w^{|E_=(\sigma)|} \lambda^{|V_+(\sigma)|}$

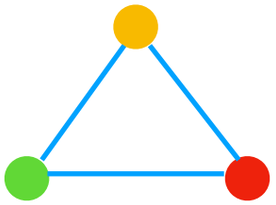
- $w < 1$: Anti-ferromagnetic, favors larger cuts
- $w > 1$: Ferromagnetic, favors smaller cuts

Approximate counting

- Compute $(1 \pm \varepsilon) \cdot Z_G(w)$
- Equivalent to
  - Approximate inference
  - Approximate sampling
  - Approximate root-finding

Our main result

$$Z_G(0) = (\#\text{ valid } q\text{-colorings of } G)$$



Our algorithmic result:

For any $w \in [0,1]$, any integer $q \geq 2\Delta$, and bounded-degree graph G, there exists a deterministic *FPTAS*, which outputs $\hat{Z}$ s.t.

$$|\hat{Z} - Z_G(w)| \leq \epsilon \cdot Z_G(w)$$

in time $\mathrm{poly}(|G|, 1/\epsilon)$ for G

Dobrushin condition

Prior to our work:

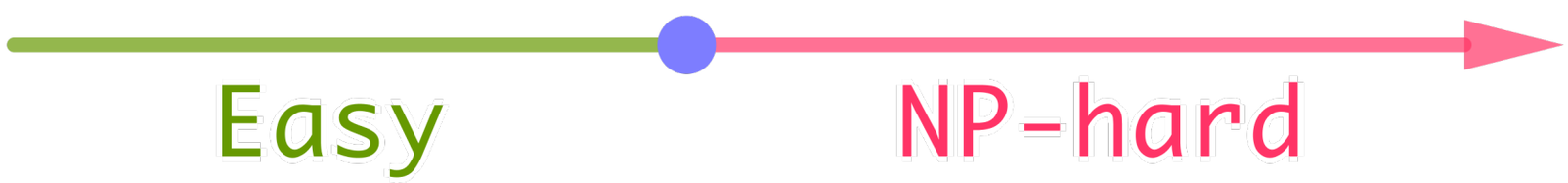
- ***Randomized*** **MCMC for $q \geq 2\Delta + 1$ [Jerrum]**
- ***Randomized*** **MCMC for $q \geq (\frac{11}{6} - \epsilon)\Delta$ [Chen et al.] [Vigoda]**
- **Deterministic algorithm for $q \geq 2.58\Delta + 1$ [Lu and Yin]**
- **Deterministic algorithm for $q = 4, \Delta = 3$ [Lu, Yang, Zhang and Zhu]**

Computational phase transition

This work: exploit formal connection between algorithms and phase transition.

Crucially we exploit both directions

- Many problems have a “computational phase transition” (roughly):

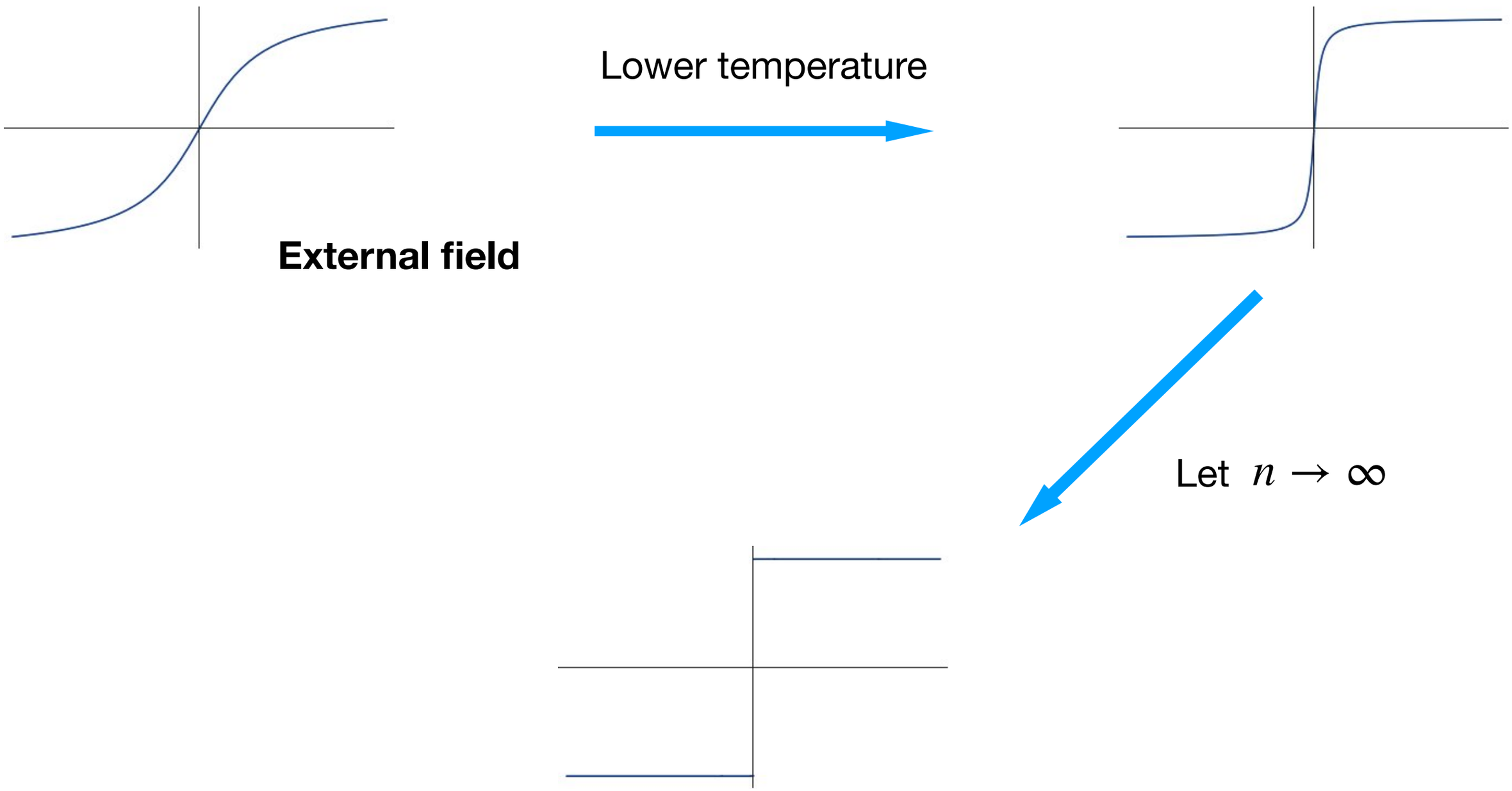


- In some cases, known to coincide with statistical physics phase transition:
  - Hardcore model
  - Anti-ferromagnetic Ising/2-spin model
  - ....

Phase transitions ≈ Discontinuity

In statistical physics

Observable, e.g., mean magnetization :≈ average number of +-spin vertices



Phase transitions formally:

Geometry of polynomials

- Lee-Yang theory:

Phase transition ≈ Non-analyticity of the "free energy" $\log Z$

- Continuity of observables ≈ Analyticity of log Z

- Consider the Ising model: $$Z_G(w, \lambda) := \sum_{\sigma: V \to \{+,-\}} w^{|E_=(\sigma)|} \lambda^{|V_+(\sigma)|}$$

- Observable: the mean magnetization

$$\mathbb{E}_\sigma |V_+(\sigma)| = \sum_{\sigma: V \to \{+,-\}} |V_+(\sigma)| \cdot \Pr[\sigma]$$

continued on next slide…

- Consider the Ising model: $Z_G(w,\lambda) := \sum_{\sigma:V\to\{+,-\}} \lambda^{|V_+(\sigma)|} w^{|E_=(\sigma)|}$

- Observable: the mean magnetization

$$
\begin{aligned}
\mathbb{E}_\sigma |V_+(\sigma)| &= \sum_{\sigma:V\to\{+,-\}} |V_+(\sigma)| \cdot \Pr[\sigma] \\
&= \frac{1}{Z_G(w,\lambda)} \sum_{\sigma:V\to\{+,-\}} |V_+(\sigma)| \cdot \lambda^{|V_+(\sigma)|} w^{|E_=(\sigma)|} \\
&= \frac{\lambda}{Z_G(w,\lambda)} \cdot \frac{\partial Z_G(w,\lambda)}{\partial \lambda} \\
&= \lambda \cdot \frac{\partial \log Z_G(w,\lambda)}{\partial \lambda}
\end{aligned}
$$

- Analyticity of log Z ≈ Continuity of observables

≈ Absence of complex zeros: Z≠0

Lack of phase transition ≈ Lack of complex zeros

I. Algorithms based on Phase transitions

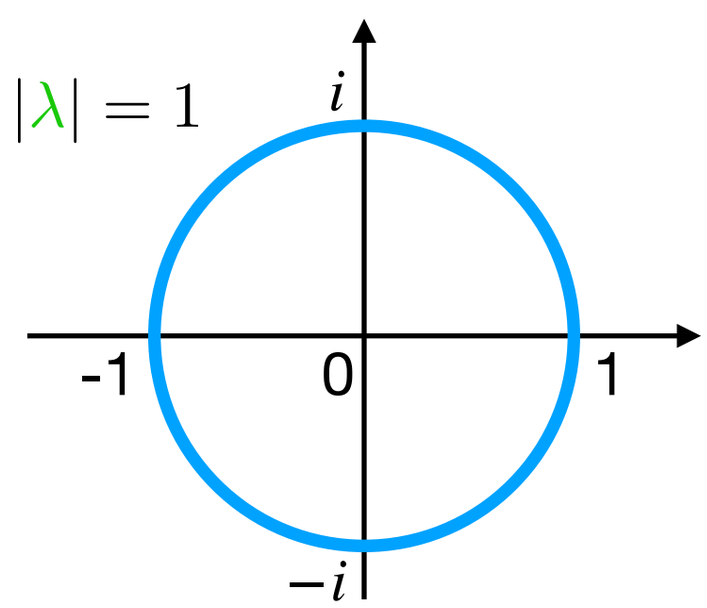
Ising model: Lee-Yang theorem

Recall the Ising partition function: $Z_G(w, \lambda) := \sum_{\sigma: V \to \{+,-\}} w^{|E_=(\sigma)|} \lambda^{|V_+(\sigma)|}$

Fix w, view λ as variable.

Lee-Yang theorem (ferromagnetic):
Let $w \geq 1$**, then**

$$Z_G(w, \lambda) = 0 \implies |\lambda| = 1$$

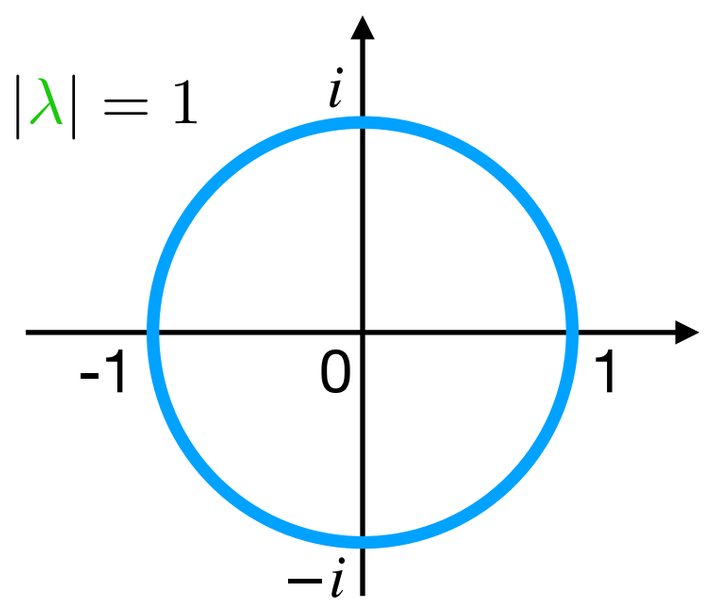


Algorithmic Lee-Yang theory

$$Z_G(w,\lambda) := \sum_{\sigma:V\to\{+,-\}} w^{|E_=(\sigma)|}\lambda^{|V_+(\sigma)|}$$

Lee-Yang theorem:
For any $w \geq 1$, $|\lambda| \neq 1$,

$$Z_G(w,\lambda) \neq 0$$



Our result (Joint w/ A. Sinclair, P. Srivastava):
For any $w \geq 1$, $|\lambda| \neq 1$ **and bounded degree graph G, there exists an *FPTAS*, which outputs** $\hat{Z}$ **s.t.**

$$|\hat{Z} - Z_G(w,\lambda)| \leq \epsilon \cdot Z_G(w,\lambda)$$

in time $\mathrm{poly}(|G|, 1/\epsilon)$

Barvinok's interpolation

Recall the Ising partition function: $Z_G(w,\lambda) := \sum_{\sigma:V\to\{+,-\}} w^{|E_=(\sigma)|}\lambda^{|V_+(\sigma)|}$

Fix w, view λ as variable.

Lee-Yang theorem:
For any $w \geq 1$, $|\lambda| \neq 1$,

$$Z_G(w,\lambda) \neq 0$$

[Barvinok, Barvinok and Soberon]

- **Consider the Taylor expansion of log Z**
- **In a zero free region, log Z can be approximated to** $\pm\epsilon$ **by its k-th order Taylor series for** $k = O(\log(n/\epsilon))$
- $\log Z_G(w,\lambda) \pm \epsilon \iff (1 \pm \epsilon)\cdot Z_G(w,\lambda)$

To apply Barvinok's interpolation:

- **Taylor expansion of** $\log Z_G(w,\lambda)$ around $\lambda = 0$
- **By Lee-Yang theorem,** $|\lambda| < 1$ **is a zero free region**
- **k-th order Taylor series is determined by the first k+1 coefficients of Z**
- **Naively computing the first k+1 coefficients of Z takes time** $O(n^k)$
- $\Longrightarrow$ **Quasi-polynomial time algorithm for** $k = O(\log(n/\epsilon))$

continued next slide…

Faster computation of coefficients

$$Z_G(w,\lambda) := \sum_{\sigma:V\to\{+,-\}} w^{|E_=(\sigma)|}\lambda^{|V_+(\sigma)|}$$

[Patel-Regts]

- **If the coefficients = $\sum$ (induced subgraphs)**
- **Then it suffices to compute $\sum$ (connected induced subgraphs)**

For a graph of maximum degree Δ

- **#(induced subgraphs of size k) =** $O(n^k)$
- **#(connected induced subgraphs of size k) =** $O(n(e\Delta)^k)$

Recall that $k = O(\log(n/\epsilon))$

- **For the Ising model, not exactly "induced subgraphs"**
- **But essentially the same line of argument goes through**

II. Phase transitions from the analysis of algorithms

Locating zeros beyond the Lee-Yang theorem

Fisher zeros in Ising model

Fisher zeros:

fix $\lambda = 1$, view w as variable:

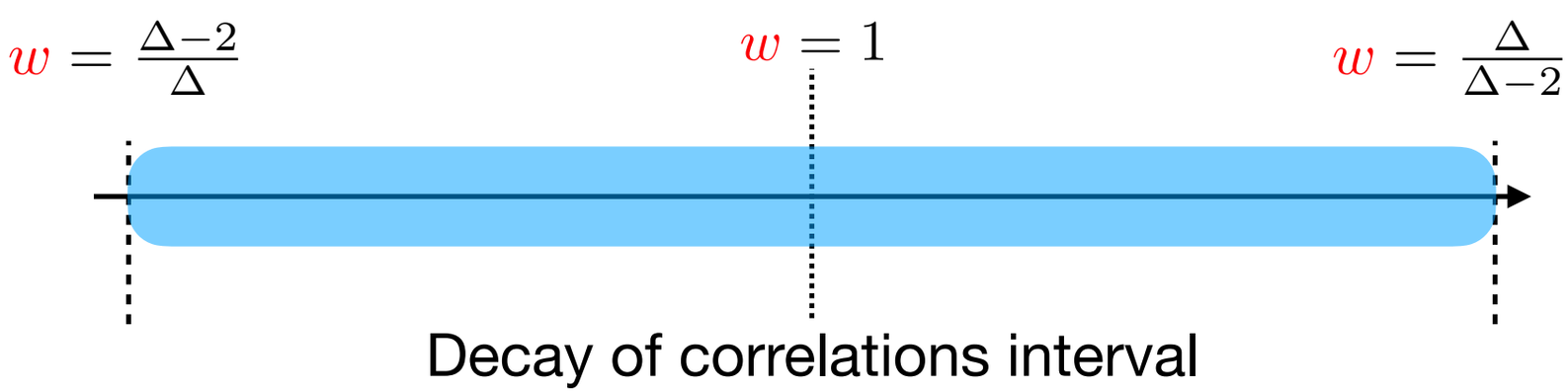
$$Z_G(w) := \sum_{\sigma: V \to \{+,-\}} w^{|E_=(\sigma)|}$$

Lee-Yang zeros:

fix w, view λ as variable:

$$Z_G(w, \lambda) := \sum_{\sigma: V \to \{+,-\}} w^{|E_=(\sigma)|} \lambda^{|V_+(\sigma)|}$$

Our zero-freeness result (Joint w/ A. Sinclair, P. Srivastava):
$Z_G(w)$ does not vanish in a *complex open region* containing the *entire* correlation decay interval.



Potts model

Fix the max. degree = Δ. Recall the partition function:

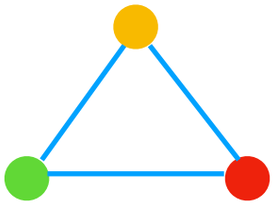
$$Z_G(w) := \sum_{\sigma: V \to [q]} w^{|E_=(\sigma)|}$$

<u>Our zero-freeness result (Joint w/ A. Sinclair, P. Srivastava):</u>
Fix any integer q such that $q \geq 2\Delta$. Then there exists a constant τ_Δ such that $Z_G(w) \neq 0$ when $w \in B([0,1], \tau_\Delta)$

<u>Prior to our work:</u>
- $q \geq e\Delta + 1$ **then** $Z_G(w) \neq 0$ when $w \in B([0,1], \tau_\Delta)$**. [Bencs. et. al.]**

FPTAS for counting graph colorings

$$Z_G(0) = (\#\text{ valid } q\text{-colorings of } G)$$



<u>New algorithmic result as a by-product:</u>

For any $w \in [0,1]$, any integer $q \geq 2\Delta$, and bounded-degree graph G, there exists a deterministic *FPTAS*, which outputs $\hat{Z}$ s.t.

$$|\hat{Z} - Z_G(w)| \leq \epsilon \cdot Z_G(w)$$

in time $\mathrm{poly}(|G|, 1/\epsilon)$ for G

Dobrushin condition

<u>Prior to our work:</u>

- ***Randomized*** **MCMC for $q \geq 2\Delta + 1$ [Jerrum]**
- ***Randomized*** **MCMC for $q \geq (\frac{11}{6} - \epsilon)\Delta$ [Chen et al.] [Vigoda]**
- **Deterministic algorithm for $q \geq 2.58\Delta + 1$ [Lu and Yin]**
- **Deterministic algorithm for $q = 4, \Delta = 3$ [Lu et al.]**

Potts model (triangle-free)

$$Z_G(w) := \sum_{\sigma : V \to [q]} w^{|E_=(\sigma)|}$$

<u>Our zero-freeness result (Joint w/ A. Sinclair, P. Srivastava):</u>
Fix any integer q such that: $q \gtrapprox 1.7633\Delta + \beta$, there exists a constant τ_Δ such that

$$Z_G(w) \neq 0 \text{ if } G \text{ is triangle-free and } w \in B([0,1], \tau_\Delta)$$

$\implies$ **FPTAS for bounded degree graphs**

$\alpha^\star \approx 1.7633$ **is the threshold for spatial mixing, which is the unique positive solution of the equation**

$$xe^{-1/x} = 1$$

Zero-freeness using algorithmic ideas

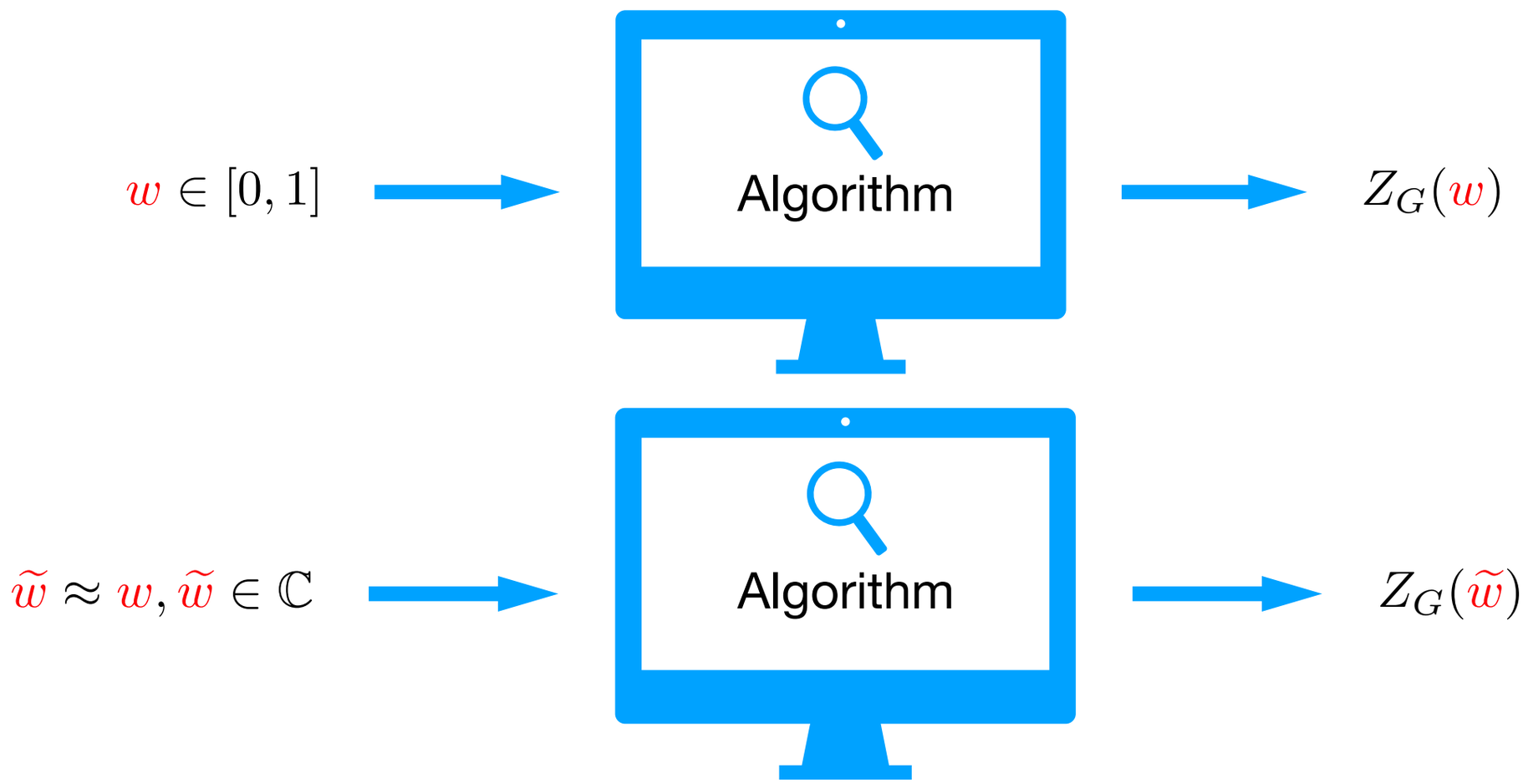
$$Z_G(w) := \sum_{\sigma: V \to [q]} w^{|E_=(\sigma)|}$$

If G has a valid q-coloring, then

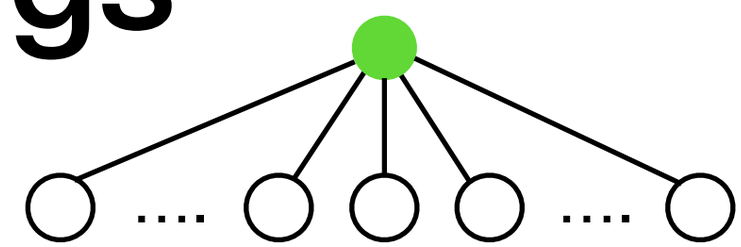
$$Z_G(w) \geq 1$$

It suffices to prove that

$$\left| \frac{Z_G(\widetilde{w})}{Z_G(w)} \right| > 0$$



Ratios and Pinnings



Fix a vertex v

$Z^{(i)}_{G,v}(\omega)$: colorings in which vertex v receives color i

$$Z_G(\omega) = \sum_{i \in L(v)} Z^{(i)}_{G,v}(\omega) = \boxed{Z^{(j)}_{G,v}(\omega)} \cdot \sum_{i \in L(v)} \frac{Z^{(i)}_{G,v}(\omega)}{Z^{(j)}_{G,v}(\omega)}$$

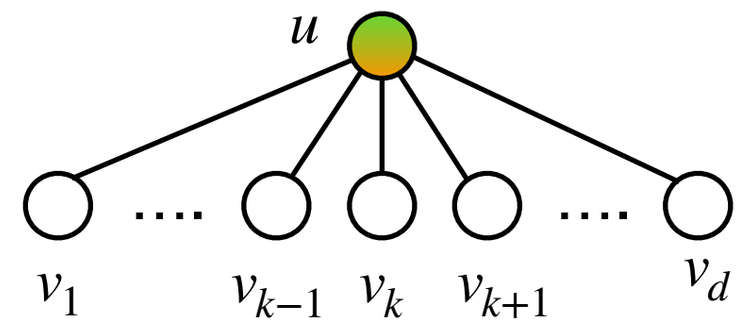
$$Z_G(\widetilde{\omega}) = \sum_{i \in L(v)} Z^{(i)}_{G,v}(\widetilde{\omega}) = \boxed{Z^{(j)}_{G,v}(\widetilde{\omega})} \cdot \sum_{i \in L(v)} \frac{Z^{(i)}_{G,v}(\widetilde{\omega})}{Z^{(j)}_{G,v}(\widetilde{\omega})}$$

Induction

Consider $R^{(i,j)}_{G,v}(\omega) := \dfrac{Z^{(i)}_{G,v}(\omega)}{Z^{(j)}_{G,v}(\omega)}$ inductively

We will show that $R^{(i,j)}_{G,v}(\widetilde{w}) \approx R^{(i,j)}_{G,v}(w)$

Recurrence for colorings ($w = 0$)



$$\frac{\Pr_G[c(u) = i]}{\Pr_G[c(u) = j]} = R_{G,u}^{(i,j)} = \frac{Z_{G,u}^{(i)}}{Z_{G,u}^{(j)}}$$

$$= \frac{Z_{G\setminus u}\left[\forall k \in [d], c(v_k) \neq i\right]}{Z_{G\setminus u}\left[\forall k \in [d], c(v_k) \neq j\right]}$$

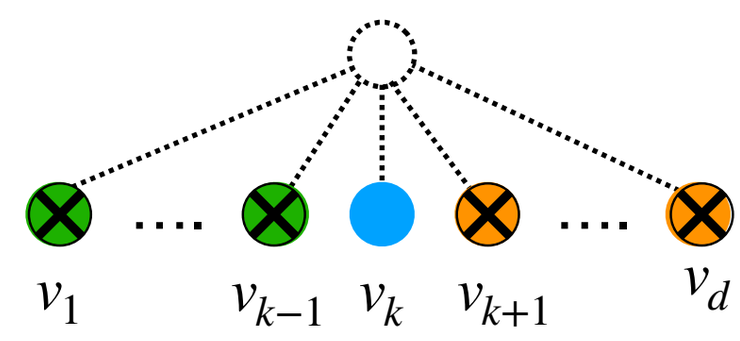
$$= \frac{\Pr_{G\setminus u}\left[\forall k \in [d], c(v_k) \neq i\right]}{\Pr_{G\setminus u}\left[\forall k \in [d], c(v_k) \neq j\right]}$$

$$= \prod_{k=1}^{d} \frac{\Pr_{G\setminus u}\left[c(v_k) \neq i, c(v_t) \neq i \text{ for } t \leq k-1, c(v_t) \neq j \text{ for } t \geq k+1\right]}{\Pr_{G\setminus u}\left[c(v_k) \neq j, c(v_t) \neq i \text{ for } t \leq k-1, c(v_t) \neq j \text{ for } t \geq k+1\right]}$$

$$= \prod_{k=1}^{d} \frac{\Pr_{G\setminus u}\left[c(v_k) \neq i \mid c(v_t) \neq i \text{ for } t \leq k-1, c(v_t) \neq j \text{ for } t \geq k+1\right]}{\Pr_{G\setminus u}\left[c(v_k) \neq j \mid c(v_t) \neq i \text{ for } t \leq k-1, c(v_t) \neq j \text{ for } t \geq k+1\right]}$$

$$= \prod_{k=1}^{d} \frac{\Pr_{G_k}\left[c(v_k) \neq i\right]}{\Pr_{G_k}\left[c(v_k) \neq j\right]}$$

$$= \prod_{k=1}^{d} \frac{1 - \Pr_{G_k}\left[c(v_k) = i\right]}{1 - \Pr_{G_k}\left[c(v_k) = j\right]}$$



Recurrence for Potts model

$$\frac{\Pr_G[c(u)=i]}{\Pr_G[c(u)=j]} = R_{G,u}^{(i,j)} = \prod_{k=1}^{d} \frac{1-\Pr_{G_k}[c(v_k)=i]}{1-\Pr_{G_k}[c(v_k)=j]}$$

Can be generalized to any $w \in [0,1]$, and in fact to any $w \in \mathbb{C}$:

$$\frac{\Pr_{G,\omega}[c(u)=i]}{\Pr_{G,\omega}[c(u)=j]} = R_{G,u}^{(i,j)}(\omega) = \prod_{k=1}^{d} \frac{1-(1-\omega)\cdot\Pr_{G_k,\omega}[c(v_k)=i]}{1-(1-\omega)\cdot\Pr_{G_k,\omega}[c(v_k)=j]}$$

where

$$R_{G,v}^{(i,j)}(\omega) := \frac{Z_{G,v}^{(i)}(\omega)}{Z_{G,v}^{(j)}(\omega)}$$

$$\Pr_{G,\omega}[c(v)=i] := \frac{Z_{G,v}^{(i)}(\omega)}{Z_G(\omega)}.$$

$$\Pr_{G,\omega}[c(v)=j] := \frac{Z_{G,v}^{(j)}(\omega)}{Z_G(\omega)}$$

Tree recurrence

$w \in [0,1]$

$R_{G,v}^{(i,j)}(w)$

At every vertex of the tree, the two computations remain close

$\widetilde{w} \approx w, \widetilde{w} \in \mathbb{C}$

$R_{G,v}^{(i,j)}(\widetilde{w})$

Goal: $R_{G,v}^{(i,j)}(\widetilde{w}) \approx R_{G,v}^{(i,j)}(w)$

Analyzing the recurrence

$$\frac{\Pr_{G,w}[c(u)=i]}{\Pr_{G,w}[c(u)=j]} = \prod_{k=1}^{d} \frac{1-(1-w)\cdot \Pr_{G_k,w}[c(v_k)=i]}{1-(1-w)\cdot \Pr_{G_k,w}[c(v_k)=j]}$$

Apply log

For $w \in [0,1]$

$$A - B = \sum_{k=1}^{d} (f(A_k) - f(B_k))$$

Similarly for $\widetilde{w} \approx w, \widetilde{w} \in \mathbb{C}$

$$\widetilde{A} - \widetilde{B} = \sum_{k=1}^{d} \left(\widetilde{f}(\widetilde{A_k}) - \widetilde{f}(\widetilde{B_k})\right)$$

Goal: $\widetilde{A} - \widetilde{B} \approx A - B$

Induction step: $\widetilde{A} - \widetilde{B} \approx_\epsilon A - B$ provided that $\widetilde{A_k} - \widetilde{B_k} \approx_\epsilon A_k - B_k$

for the same ϵ

where

$$f(x) := \log(1-(1-w)e^x)$$

$$A := \log \Pr_{G,w}[c(v)=i] \quad \text{and} \quad B := \log \Pr_{G,w}[c(v)=j]$$

$$A_k := \log \Pr_{G_k,w}[c(v_k)=i] \quad \text{and} \quad B_k := \log \Pr_{G_k,w}[c(v_k)=j]$$

For linear recurrence

Suppose that f is linear, say $f(x) = C \cdot x$

Moreover $|C| \leq \dfrac{1}{d+1}$

Only place where q has to be large

$$A - B = \sum_{k=1}^{d} (C \cdot A_k - C \cdot B_k) \qquad \widetilde{A} - \widetilde{B} = \sum_{k=1}^{d} \left(\widetilde{C} \cdot \widetilde{A_k} - \widetilde{C} \cdot \widetilde{B_k}\right)$$

Induction step: $\widetilde{A} - \widetilde{B} \approx_\epsilon A - B$ provided that $\widetilde{A_k} - \widetilde{B_k} \approx_\epsilon A_k - B_k$

$$\begin{aligned}
\left|(A - B) - (\widetilde{A} - \widetilde{B})\right| &= \left|C \cdot \sum_{k=1}^{d} \left((A_k - B_k) - \left(\widetilde{A_k} - \widetilde{B_k}\right)\right) + O(|C - \widetilde{C}|)\right| \\
&\leq d|C| \cdot \max_k \left|(A_k - B_k) - \left(\widetilde{A_k} - \widetilde{B_k}\right)\right| + O(d|C - \widetilde{C}|) \\
&\leq \max_k \left|(A_k - B_k) - \left(\widetilde{A_k} - \widetilde{B_k}\right)\right|
\end{aligned}$$

Mean value theorem

On the real line $f(a) - f(b) = f'(\xi)(a - b)$ for some ξ

In the complex plane

<u>New mean value theorem:</u>
Let f be an analytic function that *maps the real line to the real line*.

$$\Re f(a) - \Re f(b) = \Re f'(\xi) \cdot (\Re a - \Re b) + \Im f'(\eta) \cdot (\Im a - \Im b)$$
$$|\Im f(a) - \Im f(b)| \leq |\sup \Re f'| \cdot \max\{|\Im a - \Im b|, |\Im a|, |\Im b|\}$$

negligible

for some ξ, η in the rectangular region given by a, b

negligible terms stay negligible

Linearization

Caveat: Does not linearize to the same multiplier

Niceness condition

$$\mathfrak{R}f' \approx \frac{\Pr_{G,w}[c(v)=i]}{1-\Pr_{G,w}[c(v)=i]} \approx \Pr_{G,w}[c(v)=i]$$

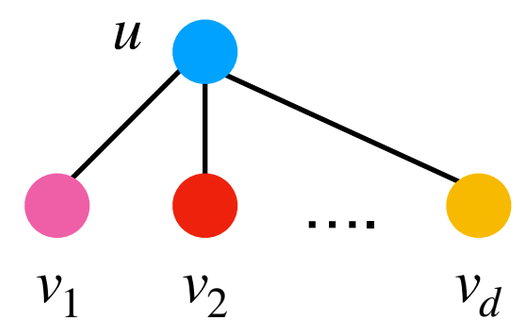
A vertex u is *nice* in G if for any $w \in [0,1]$ and any color i ,

$$\Pr_{G,w}[c(u)=i] \leq \frac{1}{d+2}$$

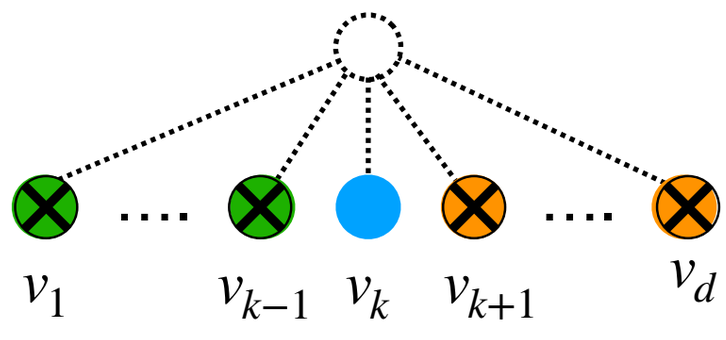
Trivial?

For q-colorings ($w=0$), if $q \geq \Delta+2$, then

$$\Pr_{G}[c(u)=i] = \frac{1}{q} \leq \frac{1}{d+2}$$



Unfortunately, in general we need to handle list-coloring for the induction: recall G_k

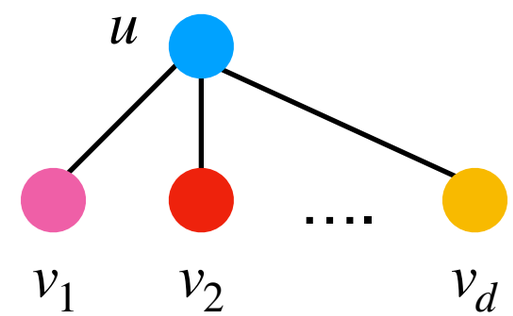


Niceness condition

A vertex u is *nice* in G if for any $w \in [0,1]$ and any color i ,

$$\Pr_{G,w}[c(u) = i] \leq \frac{1}{d+2}$$

Lemma:
If $q \geq 2\Delta + 2$ then every vertex is nice



A more careful argument shows that $q \geq 2\Delta$ suffices for general graphs, and $q > 1.76\Delta + \beta$ for triangle-free graphs

Let's check for proper colorings, i.e. $w = 0$

Color the neighbors of u arbitrarily: denoted by σ'

$$\Pr_G[c(u) = i | \sigma'] \leq \frac{1}{|L(u)| - d} \leq \frac{1}{d+2}$$

$$\implies \Pr_G[c(u) = i] \leq \frac{1}{|L(u)| - d} \leq \frac{1}{d+2}$$

Summary of the induction

$$Z_G(w) = \sum_{i \in L(v)} Z_{G,v}^{(i)}(w)$$

$$= Z_{G,v}^{(j)}(w) \cdot \sum_{i \in L(v)} \frac{Z_{G,v}^{(i)}(w)}{Z_{G,v}^{(j)}(w)}$$

$$Z_G(\widetilde{w}) = \sum_{i \in L(v)} Z_{G,v}^{(i)}(\widetilde{w})$$

$$= Z_{G,v}^{(j)}(\widetilde{w}) \cdot \sum_{i \in L(v)} \frac{Z_{G,v}^{(i)}(\widetilde{w})}{Z_{G,v}^{(j)}(\widetilde{w})}$$

Inductively

$$\left| \frac{Z_{G,v}^{(j)}(\widetilde{w})}{Z_{G,v}^{(j)}(w)} \right| \geq e^{-\epsilon(n-1)}$$

$$R_{G,v}^{(i,j)}(\widetilde{w}) \approx e^{\epsilon} \cdot R_{G,v}^{(i,j)}(w)$$

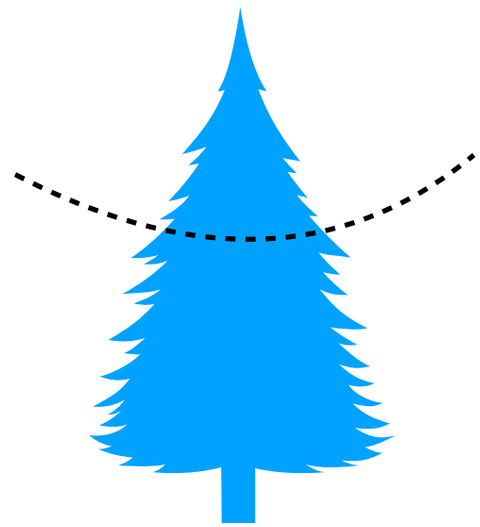
Recap

- One can exploit the analysis of algorithms (tree recurrence) to study phase transitions (prove zero-freeness)
- Conversely, zero-freeness results can also be exploited algorithmically

Discussion: Comparison to decay of correlations

<u>Correlation decay approach</u>

Truncating



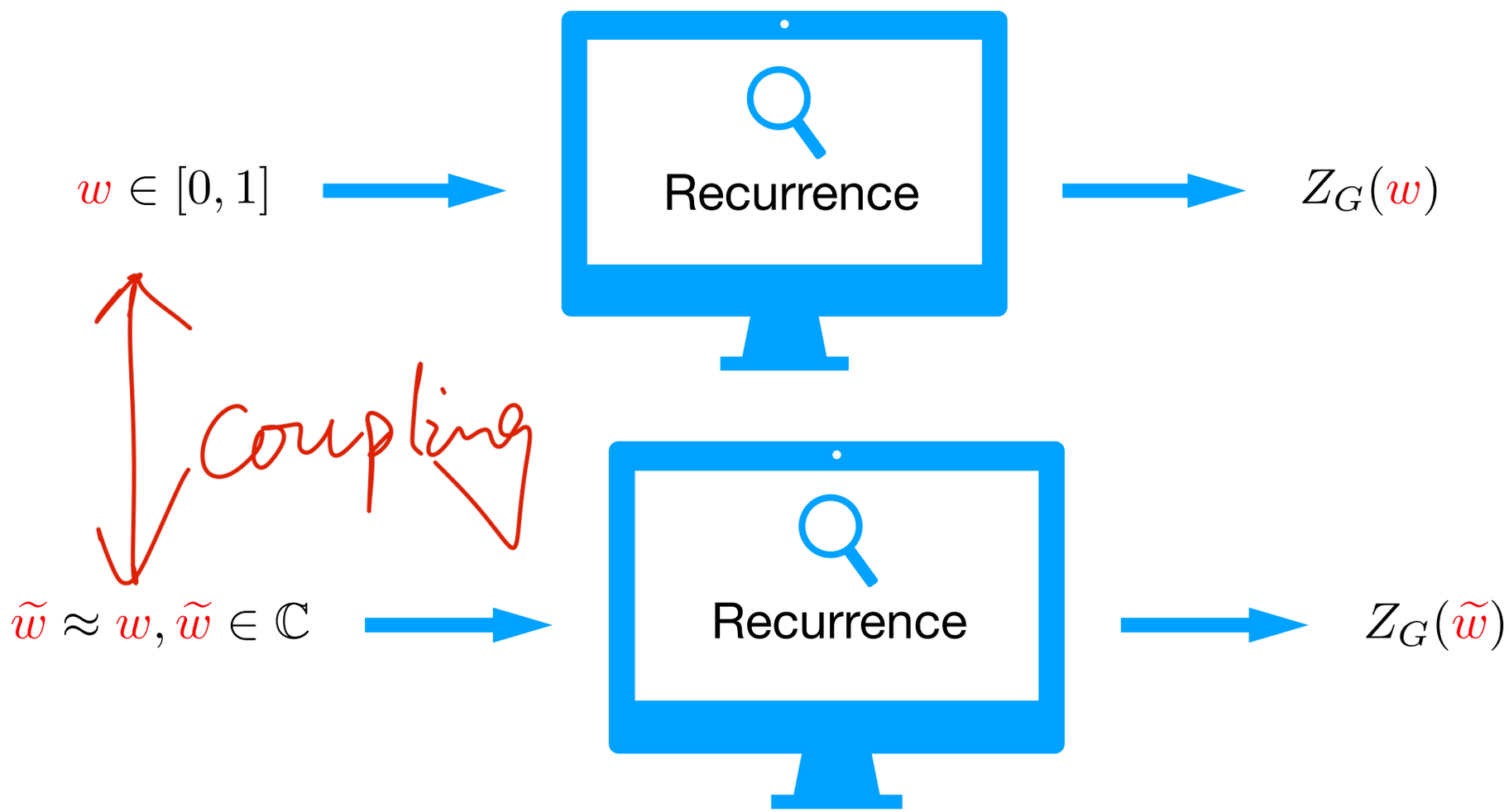
Guess the values

Tree recurrence

Issue: guessing the values explicitly can be hard

In proving zero-freeness, we only need that "*good*" values exist

Comparison to MCMC



Future direction: generalize to more sophisticated coupling argument

Thanks