Graphical models, from graphs to trees (and back)

Andrea Montanari

Stanford University

April 14, 2008

Outline

Outline

1. What is this talk about, and why should one care
2. Uniform decorrelation
3. Non-Uniform decorrelation
4. Trees vs graphs: from reconstruction to pure states

Outline

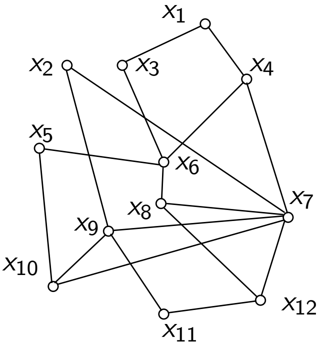
1. What is this talk about, and why should one care
2. Uniform decorrelation
3. Non-Uniform decorrelation
4. Trees vs graphs: from reconstruction to pure states

Outline

1. What is this talk about, and why should one care
2. Uniform decorrelation
3. Non-Uniform decorrelation
4. Trees vs graphs: from reconstruction to pure states

What is this talk about, and why should one care

Graphical models



$$G = (V, E),\ V = [n], \qquad \underline{x} = (x_1, \dots, x_n),\ x_i \in \mathcal{X}$$

$$\mu(\underline{x}) = \frac{1}{Z} \prod_{(ij) \in G} \psi_{ij}(x_i, x_j)\,.$$

This talk

1. G has bounded degree (on average).

2. G has girth $\ell(n) \to \infty$ (apart from $o(n)$ vertices).

3. G is random.

This talk

1. G has bounded degree (on average).

2. G has girth $\ell(n) \to \infty$ (apart from $o(n)$ vertices).

3. G is random.

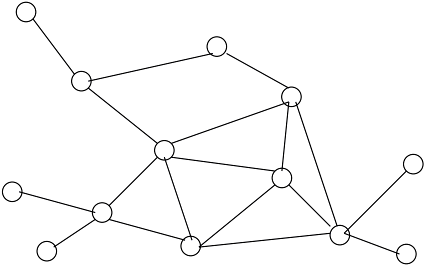
This talk

1. G has bounded degree (on average).

2. G has girth $\ell(n) \to \infty$ (apart from $o(n)$ vertices).

3. G is random.

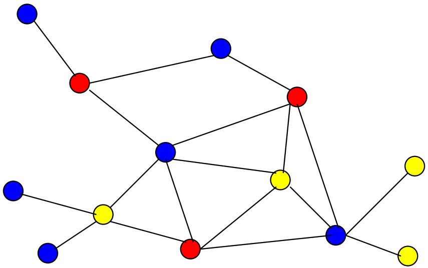
Example 1: q-coloring



$G = (V, E)$ graph.

$\underline{x} = (x_1, x_2, \dots, x_n)$, $x_i \in \{1, \dots, q\}$ variables

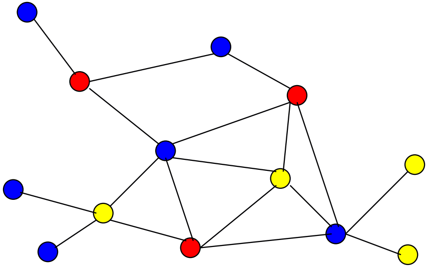
Example 1: q-coloring



$G = (V, E)$ graph.

$\underline{x} = (x_1, x_2, \ldots, x_n)$, $x_i \in \{1, \ldots, q\}$ variables

Uniform measure over proper colorings



$$\mu(\underline{x}) = \frac{1}{Z} \prod_{(i,j)\in E} \psi(x_i, x_j)\,, \qquad\qquad \psi(x,y) = \mathbb{I}(x \neq y)\,.$$

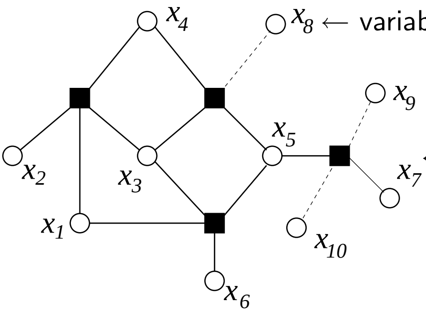
Example 2: k-satisfiability

n variables: $\underline{x} = (x_1, x_2, \dots, x_n)$, $x_i \in \{0, 1\}$

m k-clauses

$$(x_1 \vee \overline{x_5} \vee x_7) \wedge (x_5 \vee x_8 \vee \overline{x_9}) \wedge \cdots \wedge (\overline{x_{66}} \vee \overline{x_{21}} \vee \overline{x_{32}})$$

Uniform measure over solutions



$\leftarrow$ variables $x_i \in \{0, 1\}$

$\leftarrow$ clauses, e.g. $(x_5 \vee x_7 \vee \overline{x_9} \vee \overline{x_{10}})$

$$F = \cdots \wedge \underbrace{(x_{i_1(a)} \vee \overline{x}_{i_2(a)} \vee \cdots \vee x_{i_k(a)})}_{a\text{-th clause}} \wedge \cdots$$

$$\mu(\underline{x}) = \frac{1}{Z} \prod_{a=1}^{M} \psi_a(x_{i_1(a)}, \ldots, x_{i_k(a)})$$

Many other examples

- Communications/signal processing (technologically relevant)
- . . .

Probability, physics, computer science, information theory,. . .

Many other examples

- Communications/signal processing (technologically relevant)
- ...

Probability, physics, computer science, information theory,...

Is this motivating enough? (a personal view)

Emerging conceptual unity:

- Approximation of sparse graph models by trees.
- ...

[cf. Aldous' local weak convergence]

Is this motivating enough? (a personal view)

Emerging conceptual unity:

- Approximation of sparse graph models by trees.
- . . .

[cf. Aldous' local weak convergence]

Is this motivating enough? (a personal view)

Emerging conceptual unity:

- Approximation of sparse graph models by trees.
- ...

[cf. Aldous' local weak convergence]

Is this motivating enough? (a personal view)

Emerging conceptual unity:

- Approximation of sparse graph models by trees.
- ...

[cf. Aldous' local weak convergence]

Uniform decorrelation

Random k-satisfiability

Each clause is uniformly random among the $2^k\binom{n}{k}$ possible ones.

$$n \longrightarrow \infty,\ m = \alpha n$$

Random k-satisfiability

Each clause is uniformly random among the $2^k \binom{n}{k}$ possible ones.

$$n \longrightarrow \infty, \; m = \alpha n$$

Number of 'good' truth assignments

$$Z_n(\beta) = \sum_{\underline{x}} \exp\Big\{ -2\beta \, \#[\text{clauses violated by } \underline{x}] \Big\}$$

Theorem (Montanari, Shah, 2007)

If $\alpha < \alpha_{\rm u}(k) = (2 \log k) k^{-1} \, [1 + o_k(1)]$ *then*

$$\frac{1}{n} \log Z_n(\beta) \xrightarrow{a.s.} \phi(\alpha, \beta) \, ,$$

where ...

. . . where. . .

$$\phi(\alpha,\beta) = -k\alpha\mathbb{E}\log[1+\tanh h\tanh u] + \alpha\mathbb{E}\log\left\{1-\frac{1}{2^k}(1-e^{-\beta})\prod_{i=1}^{k}(1-\tanh h_i)\right\} + \\ +\mathbb{E}\log\left\{\prod_{i=1}^{\ell_+}(1+\tanh u_i^+)\prod_{i=1}^{\ell_-}(1-\tanh u_i^-)+\prod_{i=1}^{\ell_+}(1-\tanh u_i^+)\prod_{i=1}^{\ell_-}(1+\tanh u_i^-)\right\},$$

and h, u are the *unique* solution of

$$h \stackrel{\mathrm{d}}{=} \sum_{a=1}^{l_+} u_a - \sum_{b=1}^{l_-} u'_b\,, \qquad u \stackrel{\mathrm{d}}{=} f_\beta(h_1,\dots,h_{k-1})\,.$$

[Conjectured by Monasson, Zecchina 1999]

Proof: 1st preliminary remark

Sufficient to prove

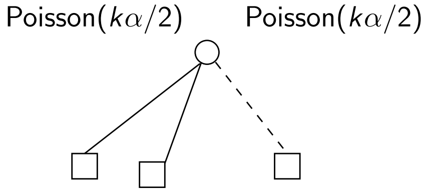
$$\frac{1}{n}\mathbb{E}\log Z_n(\beta) \longrightarrow \phi(\alpha, \beta)\,,$$

Proof: 2nd preliminary remark

The tree ensemble $\mathsf{T}(\ell)$, $\mathsf{T}(\infty)$

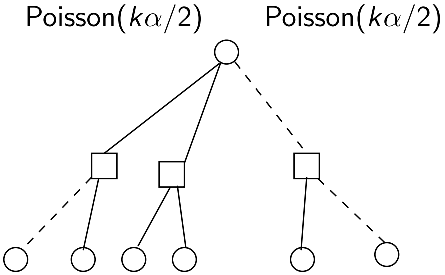
$\text{Poisson}(k\alpha/2)$ $\quad$ $\text{Poisson}(k\alpha/2)$

The tree ensemble $\mathsf{T}(\ell)$, $\mathsf{T}(\infty)$

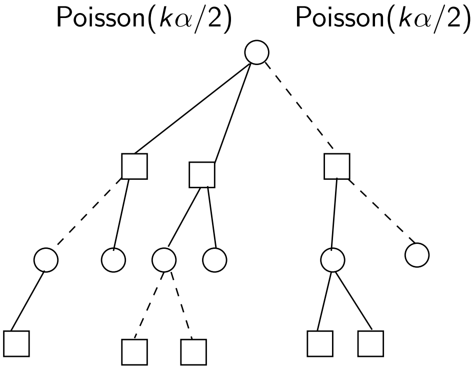


Proof: 2nd preliminary remark

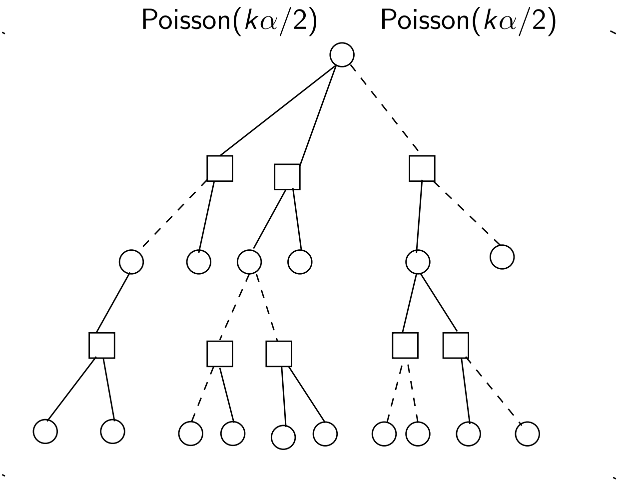
The tree ensemble $\mathsf{T}(\ell)$, $\mathsf{T}(\infty)$



The tree ensemble $\mathsf{T}(\ell)$, $\mathsf{T}(\infty)$



The tree ensemble $\mathsf{T}(\ell)$, $\mathsf{T}(\infty)$



Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E} \log Z_n(0)$.
2. Write $\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta)$ in terms of local expectations. ????

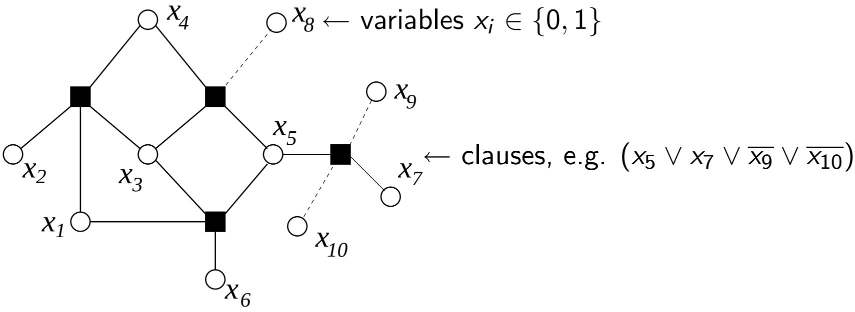
Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E}\log Z_n(0)$.
2. Write $\frac{\mathrm{d}}{\mathrm{d}\beta}\log Z_n(\beta)$ in terms of local expectations. ????

Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E} \log Z_n(0)$.
2. Write $\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta)$ in terms of local expectations. ????

Local expectations ????



$$\mu(\underline{x}) = \frac{1}{Z_n(\beta)} \prod_{a=1}^{m} \psi_{\beta,a}(x_{i_1(a)}, \ldots, x_{i_k(a)})$$

$$\psi_{\beta,a}(\cdots) = \begin{cases} 1 & \text{if clause } a \text{ is satisfied} \\ e^{-2\beta} & \text{otherwise} \end{cases}$$

$$\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta) = -2 \sum_{a=1}^{M} \mu(\text{clause } a \text{ is not satisfied})$$

Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E} \log Z_n(0)$.
2. Express $\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta)$ in terms of local expectations.
3. Prove that local expectations on G converge to expectations on $\mathsf{T}(\infty)$.
4. Show $\frac{\mathrm{d}\phi(\alpha,\beta)}{\mathrm{d}\beta}$ is equal to the same expectations on $\mathsf{T}(\infty)$.

Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E} \log Z_n(0)$.
2. Express $\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta)$ in terms of local expectations.
3. Prove that local expectations on G converge to expectations on $\mathsf{T}(\infty)$.
4. Show $\frac{\mathrm{d}\phi(\alpha,\beta)}{\mathrm{d}\beta}$ is equal to the same expectations on $\mathsf{T}(\infty)$.

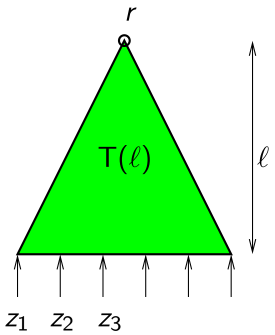
Proof strategy

1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E}\log Z_n(0)$.
2. Express $\frac{\mathrm{d}}{\mathrm{d}\beta}\log Z_n(\beta)$ in terms of local expectations.
3. Prove that local expectations on G converge to expectations on $\mathsf{T}(\infty)$.
4. Show $\frac{\mathrm{d}\phi(\alpha,\beta)}{\mathrm{d}\beta}$ is equal to the same expectations on $\mathsf{T}(\infty)$.

Proof strategy

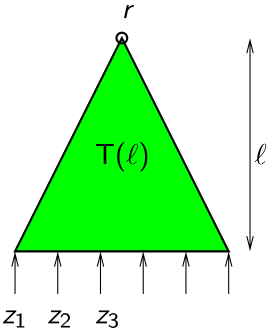
1. Check it for $\beta = 0$: $\phi(\alpha, 0) = \log 2 = n^{-1}\mathbb{E} \log Z_n(0)$.
2. Express $\frac{\mathrm{d}}{\mathrm{d}\beta} \log Z_n(\beta)$ in terms of local expectations.
3. Prove that local expectations on G converge to expectations on $\mathsf{T}(\infty)$.
4. Show $\frac{\mathrm{d}\phi(\alpha,\beta)}{\mathrm{d}\beta}$ is equal to the same expectations on $\mathsf{T}(\infty)$.

Convergence to tree values



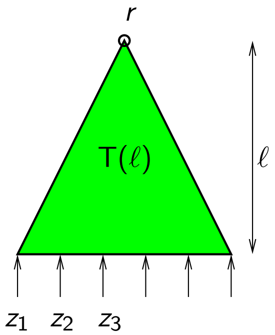
$T(\infty)$	infinite k-SAT tree
$T(\ell)$	first ℓ generations
$\mu^{\ell,z}(\,\cdot\,)$	Boltzmann measure on $T(\ell)$ boundary condition z
$\mu_r^{\ell,z}(\,\cdot\,)$	root variable marginal

Convergence to tree values



$\mathrm{T}(\infty)$	infinite k-SAT tree
$\mathrm{T}(\ell)$	first ℓ generations
$\mu^{\ell,z}(\,\cdot\,)$	Boltzmann measure on $\mathrm{T}(\ell)$ boundary condition z
$\mu_r^{\ell,z}(\,\cdot\,)$	root variable marginal

Convergence to tree values



T(∞)	infinite k-SAT tree
T(ℓ)	first ℓ generations
$\mu^{\ell,z}(\,\cdot\,)$	Boltzmann measure on T(ℓ) boundary condition z
$\mu_r^{\ell,z}(\,\cdot\,)$	root variable marginal

$$\mathbb{E}\left\{\max_{z(1),z(2)} ||\mu_r^{t,z(1)}(\,\cdot\,) - \mu_r^{t,z(2)}(\,\cdot\,)||_{\mathrm{TV}}\right\} \to 0\,.$$

'Easy' sufficient condition

True only at very small α
$(\alpha_{\mathrm{u}}(k) \simeq (2\log k)/k$, conjecture up to $\simeq 2^k \log 2)$

$$\mathbb{E}\left\{\max_{z(1),z(2)} ||\mu_r^{t,z(1)}(\,\cdot\,) - \mu_r^{t,z(2)}(\,\cdot\,)||_{\rm TV}\right\} \to 0\,.$$

'Easy' sufficient condition

True only at very small α
($\alpha_{\rm u}(k) \simeq (2\log k)/k$, conjecture up to $\simeq 2^k \log 2$)

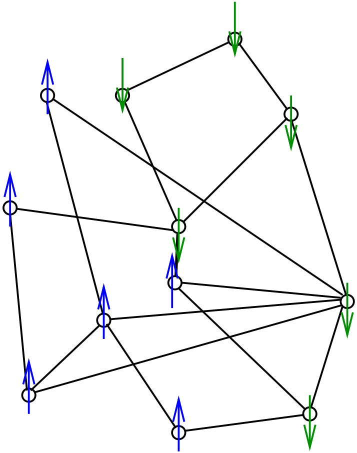
$$\mathbb{E}\left\{\max_{z(1),z(2)} ||\mu_r^{t,z(1)}(\,\cdot\,) - \mu_r^{t,z(2)}(\,\cdot\,)||_{\rm TV}\right\} \to 0\,.$$

'Easy' sufficient condition

True only at very small α
($\alpha_{\rm u}(k) \simeq (2\log k)/k$, conjecture up to $\simeq 2^k \log 2$)

Non-Uniform decorrelation

Ferromagnetic Ising model



Ferromagnetic Ising model

$G_n = (V_n \equiv [n], E_n)$

$x_i \in \{+1, -1\}$

$$\mu(\underline{x}) = \frac{1}{Z_n(\beta, B)} \exp\left\{\beta \sum_{(ij)\in E_n} x_i x_j + B \sum_i x_i\right\}$$

[Johnston, Plecháč 1998, Leone et al 2004]

Ferromagnetic Ising model

$G_n = (V_n \equiv [n], E_n)$

$x_i \in \{+1, -1\}$

$$\mu(\underline{x}) = \frac{1}{Z_n(\beta, B)} \exp\left\{\beta \sum_{(ij)\in E_n} x_i x_j + B \sum_i x_i\right\}$$

[Johnston, Plechác 1998, Leone et al 2004]

Free energy density

Theorem (Dembo, Montanari 2008)

If G_n converges locally to $\mathsf{T}(P)$, then

$$\frac{1}{n}\log Z_n(\beta, B) \xrightarrow{a.s} \phi(P, \beta, B)\,.$$

where...

Free energy density

Theorem (Dembo, Montanari 2008)

If G_n converges locally to $\mathsf{T}(P)$, *then*

$$\frac{1}{n}\log Z_n(\beta, B) \xrightarrow{a.s} \phi(P, \beta, B)\,.$$

where...

$\phi(P, \beta, B)$

For $B \geq 0$, let $\theta \equiv \tanh\beta$, $h^{(0)} > 0$, and

$$h^{(\ell+1)} \stackrel{\rm d}{=} \tanh\left\{B + \sum_{i=1}^{K-1} \mathrm{atanh}(\theta\, h_i^{(\ell)})\right\},$$

Then $h^{(\ell)} \xrightarrow{\rm d} h^*$ and

$$\phi(P,\beta,B) \equiv \log\cosh B + \frac{\overline{P}}{2}\log\cosh\beta - \frac{\overline{P}}{2}\,\mathbb{E}\log(1+\theta h_1^* h_2^*) +$$
$$+\mathbb{E}\,\log\left\{(1+\tanh B)\prod_{i=1}^{L}(1+\theta h_i^*) + (1-\tanh B)\prod_{i=1}^{L}(1-\theta h_i^*)\right\}.$$

For $B \geq 0$, let $\theta \equiv \tanh \beta$, $h^{(0)} > 0$, and

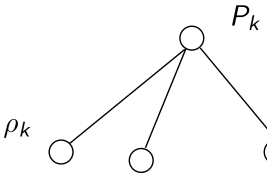
$$h^{(\ell+1)} \stackrel{\mathrm{d}}{=} \tanh \left\{ B + \sum_{i=1}^{K-1} \operatorname{atanh}(\theta\, h_i^{(\ell)}) \right\},$$

Then $h^{(\ell)} \xrightarrow{\mathrm{d}} h^*$ and

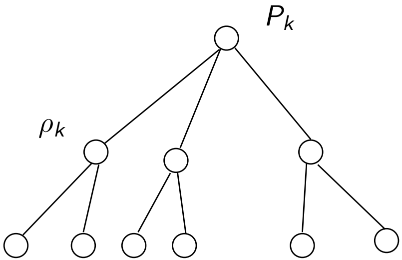
$$\phi(P, \beta, B) \equiv \log\cosh B + \frac{\overline{P}}{2} \log\cosh\beta - \frac{\overline{P}}{2}\, \mathbb{E} \log(1 + \theta h_1^* h_2^*) +$$
$$+ \mathbb{E} \log \left\{ (1 + \tanh B) \prod_{i=1}^{L} (1 + \theta h_i^*) + (1 - \tanh B) \prod_{i=1}^{L} (1 - \theta h_i^*) \right\}.$$

P_k

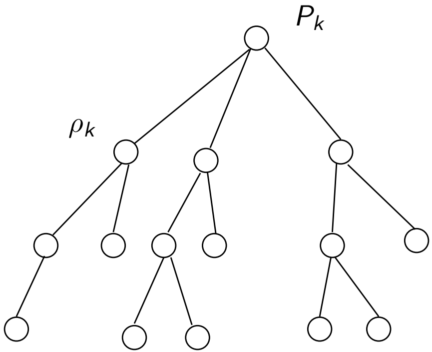
ρ_k


P_k
ρ_k

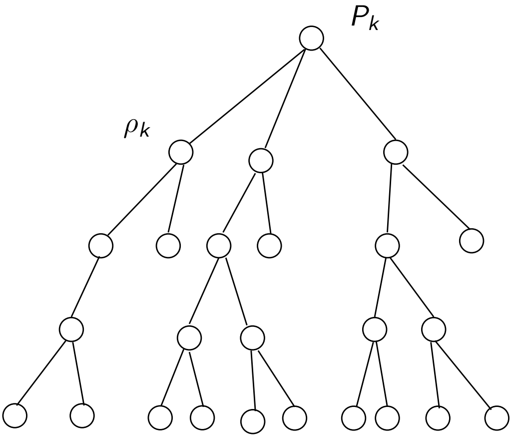
$\mathsf{T}(P, \ell)$



$\mathsf{T}(P, \ell)$



$\mathsf{T}(P, \ell)$



‘Converges locally’

$P \equiv \{P_k\}_{k \geq 0}$	Degree distribution
$\mathsf{T}(P, \ell)$	ℓ-generations Galton-Watson tree
$\mathsf{B}_i(\ell)$	Ball of radius ℓ around uniformly random node

Definition

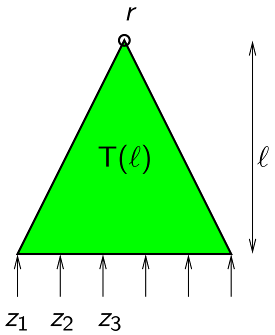
G_n *converges locally* to $\mathsf{T}(P)$ if uniform bound on the edge number distribution and, for any ℓ,

$$\mathsf{B}_i(\ell) \text{ converges in distribution to } \mathsf{T}(P, \ell).$$

Why non-uniform control? Phase transition...

For $\beta > \beta_{\rm c} \equiv {\rm atanh}(1/\overline{\rho})$

$$\lim_{B\to 0+}\lim_{n\to\infty}\mathbb{E}_i\langle x_i\rangle = -\lim_{B\to 0-}\lim_{n\to\infty}\mathbb{E}_i\langle x_i\rangle > 0$$



$$z = (+1, +1, \dots, +1) \qquad \Rightarrow \qquad \lim_{\ell \to \infty} \langle x_r \rangle_\ell > 0$$

$$z = (-1, -1, \dots, -1) \qquad \Rightarrow \qquad \lim_{\ell \to \infty} \langle x_r \rangle_\ell < 0$$

A different case: Ising spin glass

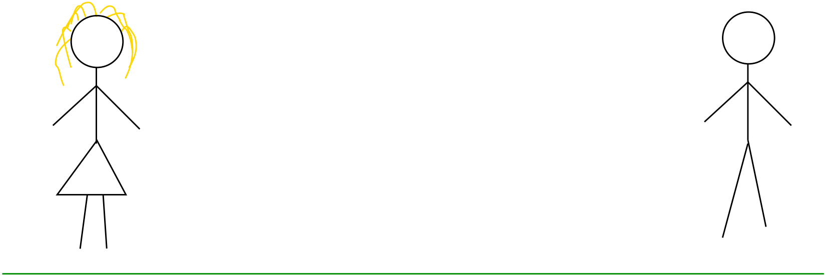
$$\mu(\underline{x}) = \frac{1}{Z} \exp\left\{ \beta \sum_{(ij)\in E_n} J_{ij} x_i x_j + B \sum_i x_i \right\}$$

$J_{ij} \in \{+1, -1\}$ uniformly random

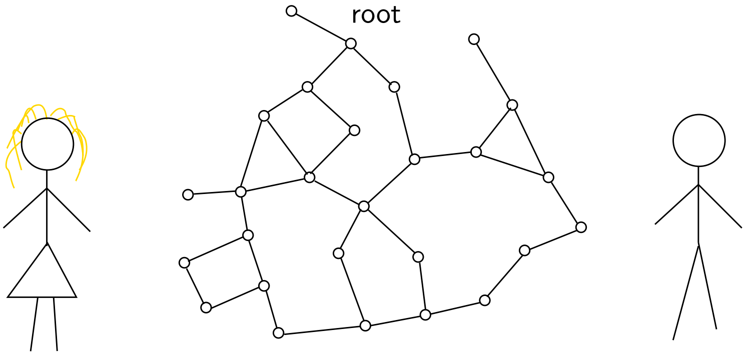
[Viana, Bray 1985]
[Other approaches: Talagrand 2001, Guerra, Toninelli 2003]
[Approximation by trees: Montanari, Gerschenfeld, in preparation]

Trees vs graphs: from reconstruction to pure states

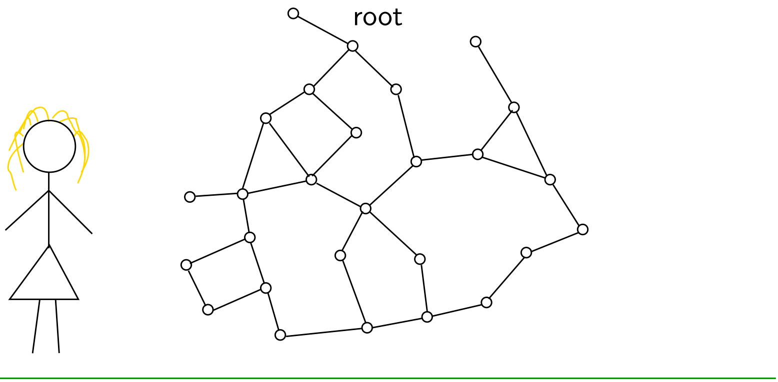
Alice and Bob



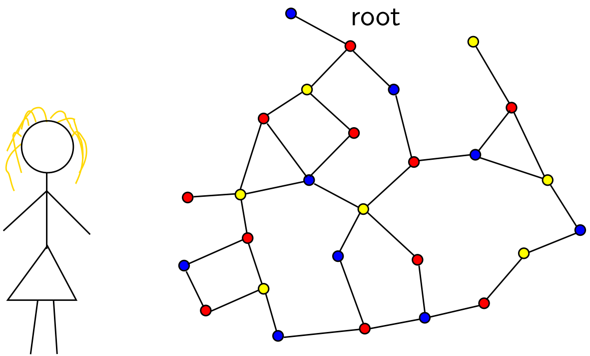
Alice, Bob and G



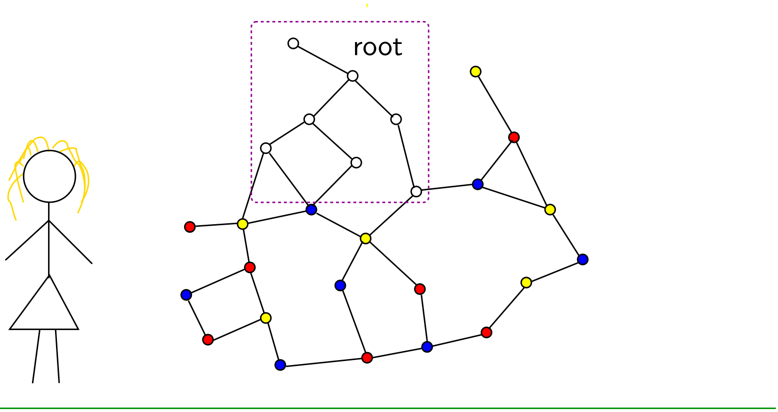
Exit Bob



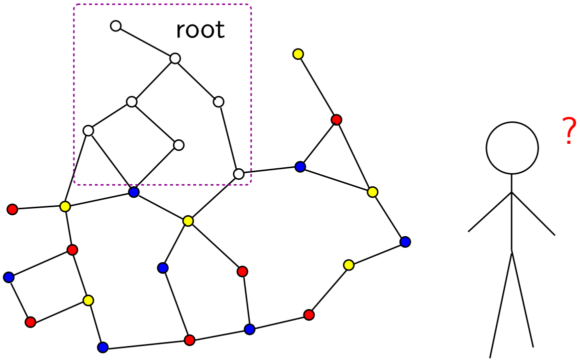
Alice samples a proper coloring (uniformly). . .



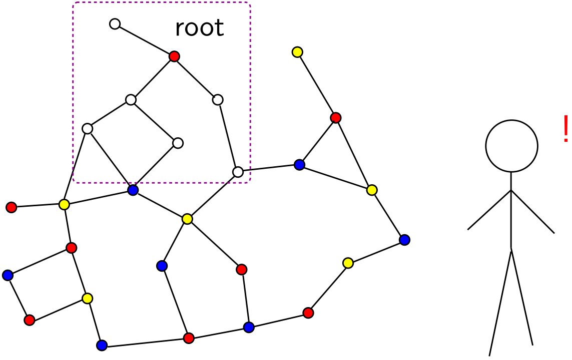
. . . and hides a ball B(root, t)



Bob. . .



. . . guesses right!



The problem

Does Bob have a chance?

Formally

$X = \{X_i : i \in V\}$ uniformly random proper coloring.

$\mu_U(\,\cdot\,|G)$ distribution of $X_U \equiv \{X_i : i \in U \subseteq V\}$

$\overline{\mathrm{B}}(r,t) = \{i \in V : d(i,r) \geq t\}$

Definition

The reconstruction problem is solvable for the sequence of random rooted graphs $G_n = (V_n = [n], E_n)$ if for some $\varepsilon > 0$,

$$||\mu_{r,\overline{\mathrm{B}}(r,t)}(\,\cdot\,,\,\cdot\,|G_n) - \mu_r(\,\cdot\,|G_n)\mu_{\overline{\mathrm{B}}(r,t)}(\,\cdot\,|G_n)||_{\mathrm{TV}} \geq \varepsilon\,,$$

with positive probability (bounded away from 0 as $n \to \infty$).

When G =Tree

$\longrightarrow$ Bleher, Ruiz, Zagrebenov (1995): Ising model on b-ary trees

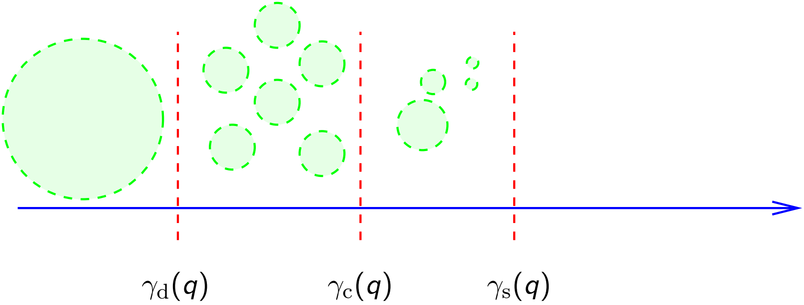
$\longrightarrow$ Evans, Kenyon, Peres, Schulman (2000): Ising on general trees

$\longrightarrow$ Mossel, Peres (2003): Non binary variables

$\longrightarrow$ Brightwell, Winkler (2004), Martin (2004): Independent sets.

$\longrightarrow$ Chayes et al. (2006): Asymmetric Ising.

Pure states decomposition in q-COL



[Biroli, Monasson, Weigt 2001]
[Mézard, Parisi, Zecchina 2003]
[Achlioptas, Ricci 2007]
[Krzakala, Montanari, Ricci, Semerjian, Zdeborova 2007]

Pure states decomposition in q-COL

Conjecture (Mézard, Montanari, 05)

$$\begin{aligned}
\gamma_{\mathrm{d}}(q) &= \\
&= \textit{Multiple pure states} \\
&= \textit{Graph reconstruction threshold} \\
&= \textit{Tree reconstruction threshold}
\end{aligned}$$

A general sufficient condition

Theorem (Gerschenfeld, Montanari, 2007)

If $\mu(\,\cdot\,|G)$ is roughly spherical then

$$\text{Graph solvable} \Leftrightarrow \text{Tree solvable.}$$

If $\mu(\,\cdot\,|G)$ is not roughly spherical then

Graph reconstruction is solvable

A general sufficient condition

Theorem (Gerschenfeld, Montanari, 2007)

If $\mu(\,\cdot\,|G)$ is roughly spherical then

$$\text{Graph solvable} \Leftrightarrow \text{Tree solvable.}$$

If $\mu(\,\cdot\,|G)$ is not roughly spherical then

Graph reconstruction is solvable

A general sufficient condition

Theorem (Gerschenfeld, Montanari, 2007)

If $\mu(\,\cdot\,|G)$ *is roughly spherical then*

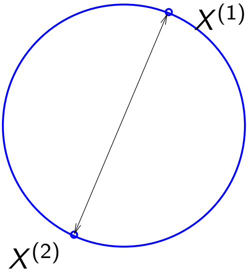
Graph solvable $\Leftrightarrow$ *Tree solvable.*

If $\mu(\,\cdot\,|G)$ *is not roughly spherical then*

Graph reconstruction is solvable

Roughly spherical???

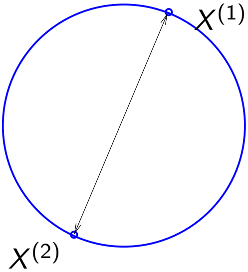
$X_i \in \{0,1\}$.
$X^{(1)} = \{X_i^{(1)}\}$, $X^{(2)} = \{X_i^{(2)}\}$ independent with distribution $\mu(\,\cdot\,|G_n)$



$\mu(\,\cdot\,|G_n)$ is *roughly spherical* if $d(X^{(1)}, X^{(2)}) \approx n/2$ with high probability.

Roughly spherical???

$X_i \in \{0, 1\}$.
$X^{(1)} = \{X_i^{(1)}\}$, $X^{(2)} = \{X_i^{(2)}\}$ independent with distribution $\mu(\,\cdot\,|G_n)$



$\mu(\,\cdot\,|G_n)$ is *roughly spherical* if $d(X^{(1)}, X^{(2)}) \approx n/2$ with high probability.

Can you check this condition?

Theorem

1. *q-coloring,* $\gamma < (q-1)\log(q-1)$*: roughly spherical.*
2. *Ising spin glass* $2\gamma(\tanh\beta)^2 < 1$*: roughly spherical.*
3. *Ising ferromagnet: not roughly spherical*

[Tree reconstruction threshold

1. Bhatayangar, Vera, Vigoda 2008, Sly 2008
2. Evans, Kenyon, Peres, Schulman 2000
3. Tree$\neq$Graph]

Can you check this condition?

Theorem

1. *q-coloring,* $\gamma < (q-1)\log(q-1)$*: roughly spherical.*
2. *Ising spin glass* $2\gamma(\tanh\beta)^2 < 1$*: roughly spherical.*
3. *Ising ferromagnet: not roughly spherical*

[Tree reconstruction threshold

1. Bhatayangar, Vera, Vigoda 2008, Sly 2008
2. Evans, Kenyon, Peres, Schulman 2000
3. Tree$\neq$Graph]

Conclusion

Combinatorics/Probability problems on random sparse graphs.

Unifying approach: approximation by trees.

Naturally leads to analytc questons.

Conclusion

Combinatorics/Probability problems on random sparse graphs.

Unifying approach: approximation by trees.

Naturally leads to analytc questons.

Conclusion

Combinatorics/Probability problems on random sparse graphs.

Unifying approach: approximation by trees.

Naturally leads to analytc questons.