Phase transitions in large graphical models: from physics to information theory and computer science

Andrea Montanari

Stanford University

March 15, 2007
Outline

1. An instructive story and many questions

2. The general theme: Phase transitions and Graphical models

3. A couple of applications (for time limits)
   - Modern coding theory
   - Random constraint satisfaction problems
Outline

1. An instructive story and many questions

2. The general theme: Phase transitions and Graphical models

3. A couple of applications (for time limits)
   - Modern coding theory
   - Random constraint satisfaction problems
Outline

1. An instructive story and many questions

2. The general theme: Phase transitions and Graphical models

3. A couple of applications (for time limits)
   - Modern coding theory
   - Random constraint satisfaction problems

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory

Modern coding theory

Random constraint satisfaction problems
Outline

1. An instructive story and many questions

2. The general theme: Phase transitions and Graphical models

3. A couple of applications (for time limits)
   - Modern coding theory
   - Random constraint satisfaction problems
Outline

1. An instructive story and many questions

2. The general theme: Phase transitions and Graphical models

3. A couple of applications (for time limits)
   - Modern coding theory
   - Random constraint satisfaction problems
An instructive story and many questions
Given a graph...
...we want to partition its vertices...
...to maximize the number of edges across.
The physics version

Localized magnetic moments (spins)
Antiferromagnetic interaction (graph)
The physics version

Localized magnetic moments (spins)
Antiferromagnetic interaction (graph)
MAXCUT

NP-hard to approximate
A few questions I like

- What is the structure of low energy configurations/optimal cuts?
- How does Nature find the optimum? How would we find it?
- Is there a ‘physics theory’ to describe low energy configurations?
  Is there an ‘efficient algorithm’ to find optimal cuts?
A few questions I like

- What is the structure of *low energy configurations/optimal cuts*?

- How does Nature *find the optimum*? How would we find it?

- Is there a *physics theory* to describe low energy configurations?
  
  Is there an *efficient algorithm* to find optimal cuts?
A few questions I like

- What is the structure of low energy configurations/optimal cuts?

- How does Nature find the optimum? How would we find it?

- Is there a ‘physics theory’ to describe low energy configurations?
  Is there an ‘efficient algorithm’ to find optimal cuts?
A few questions I like

- What is the structure of low energy configurations/optimal cuts?

- How does Nature find the optimum? How would we find it?

- Is there a ‘physics theory’ to describe low energy configurations? Is there an ‘efficient algorithm’ to find optimal cuts?
Start with ‘simple’ model

Connect each pair of vertices with probability 0.5 (independently).

A random partition yields

$$|\text{CUT}| \approx \frac{1}{2} |\text{EDGES}|.$$

SK (1972): How better is the optimal partition?

$$|\text{CUT}| = \frac{1}{2} |\text{EDGES}| + \frac{1}{4} \Delta |\text{NODES}|^{3/2} + \cdots.$$
Start with ‘simple’ model

Connect each pair of vertices with probability 0.5 (independently)

A random partition yields

\[ |\text{CUT}| \approx \frac{1}{2} |\text{EDGES}|. \]

SK (1972): How better is the optimal partition?

\[ |\text{CUT}| = \frac{1}{2} |\text{EDGES}| + \frac{1}{4} \Delta |\text{NODES}|^{3/2} + \cdots \]
Where

\[ \Delta = \frac{1}{4} \inf_q \left\{ \int_0^\infty (1 - q^2(x)) - \phi_q(0,0) \right\} \]

\[ \frac{\partial \phi(y; x)}{\partial x} = -\frac{1}{2} q'(x) \left[ \frac{\partial^2 \phi(y; x)}{\partial y^2} + x \left( \frac{\partial \phi(y; x)}{\partial y} \right)^2 \right] \]

\[ \phi(y; \infty) = |y| \]

Conjecture : Parisi (1979)

Proof : Guerra, Talagrand (2004)
\[ \Delta = \inf_q F[q] \]
Is there any hidden duality in the problem?

Flipping (spins 1 and 2) \( \approx \) Flipping (1)+ Flipping (2)
Can this fact be exploited algorithmically?

Physical dynamics is ‘local’
How do local optimization algorithms work?
\[ \Delta = \inf_q \mathcal{F}[q] \]

Is there any hidden duality in the problem?

Flipping (spins 1 and 2) \( \approx \) Flipping (1) + Flipping (2)

Can this fact be exploited algorithmically?

Physical dynamics is ‘local’

How do local optimization algorithms work?
\[ \Delta = \inf_q \mathcal{F}[q] \]

Is there any hidden duality in the problem?

Flipping (spins 1 and 2) \( \approx \) Flipping (1) + Flipping (2)

Can this fact be exploited algorithmically?

Physical dynamics is ‘local’

How do local optimization algorithms work?
Phase transitions and Graphical models
What is a phase transition?

An **abrupt change** in the state of a ‘large’ system as some control parameter is varied.

Example: water is liquid at 0.01°C and solid at −0.01°C.
A phase transition is accompanied by the emergence of **long range correlations**.

Example: water is liquid at 0.01°C and solid at −0.01°C.
\[ \mu(x) = \frac{1}{Z} \prod_{(ij) \in E} \psi_{ij}(x_i, x_j), \quad x = (x_1, \ldots, x_n). \]
\[ \mu(x) = \frac{1}{Z} \prod_{(ij) \in E} \exp\{-\beta x_i x_j\}, \quad x_i \in \{+1, -1\}. \]
Generic computational tasks

- Optimization

\[ x_\star = \arg \max_{x} \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- Partition function

\[ Z = \sum_{x} \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- Marginals

\[ \mu(x_i) = \sum_{x \sim i} \mu(x). \]

- Sampling.
Generic computational tasks

- **Optimization**

\[ X_* = \arg \max \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Partition function**

\[ Z = \sum_x \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Marginals**

\[ \mu(x_i) = \sum_{x \sim i} \mu(x). \]

- **Sampling.**
Generic computational tasks

- **Optimization**
  \[ x_\star = \arg \max \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Partition function**
  \[ Z = \sum_x \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Marginals**
  \[ \mu(x_i) = \sum_{x \sim i} \mu(x). \]

- **Sampling.**
Generic computational tasks

- **Optimization**
  \[ x_\star = \arg \max \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Partition function**
  \[ Z = \sum_{\mathbf{x}} \prod_{(ij) \in E} \psi_{ij}(x_i, x_j). \]

- **Marginals**
  \[ \mu(x_i) = \sum_{\mathbf{x} \sim i} \mu(\mathbf{x}). \]

- **Sampling.**
Are far apart variables/particles strongly correlated?
Can we approximate marginals $\mu(x_i)$ using only local information?

Can the system be found in different phases?
What is the ‘qualitative’ structure of $\mu(\cdot)$?
(conductance/concentration)

Does it relax rapidly to equilibrium?
Can we sample/optimize with local algorithms?
Are far apart variables/particles strongly correlated?
Can we approximate marginals $\mu(x_i)$ using only local information?

Can the system be found in different phases?
What is the ‘qualitative’ structure of $\mu(\cdot)$?
(conductance/concentration)

Does it relax rapidly to equilibrium?
Can we sample/optimize with local algorithms?
Are far apart variables/particles strongly correlated? 
Can we approximate marginals $\mu(x_i)$ using only local information?

Can the system be found in different phases? 
What is the ‘qualitative’ structure of $\mu(\cdot)$? 
(conductance/concentration)

Does it relax rapidly to equilibrium? 
Can we sample/optimize with local algorithms?
A couple of applications
Modern coding theory
To be concrete:
coding over binary memoryless symmetric channels.
encoder $\iff$ constraints over message bits
LDPC codes [Gallager 1963, MacKay 1995]

\[ x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0 \quad \ldots \quad x_5 \oplus x_6 \oplus x_8 = 0 \]

constraints over message bits ⇔ graphical representation
\[ x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0 \quad \cdots \quad x_5 \oplus x_6 \oplus x_8 = 0 \]

\[ \mu(x|y) = \frac{1}{Z(y)} \mathbb{I}(x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0) \cdots \mathbb{I}(x_5 \oplus x_6 \oplus x_8 = 0) \cdot \]

\[ \cdot Q(y_1|x_1) \cdots Q(y_8|x_8) \]
From $10^2$ to $10^5$ message bits

Random graph

Iterative message passing decoding
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari
Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science

Andrea Montanari
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science.
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari
Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science

Andrea Montanari
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science.
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science

Andrea Montanari
Message Passing + Density evolution analysis

Andrea Montanari
Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari
Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory and computer science.
Message Passing + Density evolution analysis
Message Passing + Density evolution analysis

Phase transitions in large graphical models: from physics to information theory and computer science
Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

channel noise

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

Decoding error probability

Example:
- LDPC codes
- \( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- \( n = 2500 \)

channel noise

Average curve

Waterfall region

Error floor

Average curve

\( n = 2500 \)
Average (Message Passing) Performance

decoding error probability

Example:
- LDPC codes
- $n = 2500$
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$
Decoding error probability vs. channel noise for LDPC codes with $n = 2500$. The figure shows the average performance curve and the waterfall region. The error floor is also indicated.
Average (Message Passing) Performance

decoding error probability

Example:

- LDPC codes
- $n = 2500$

$n = 2500$
Asymptotic Average (Message Passing) Performance

decoding error probability

channel noise

LDPC($n = 100$)

$n = 2500$

$n = 100$
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC(n = 250)

channel noise

Andrea Montanari
Phase transitions in large graphical models: from physics to information theory and computer science
Asymptotic Average (Message Passing) Performance

Decoding error probability

LDPC\((n = 500)\)

Channel noise

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory
Asymptotic Average (Message Passing) Performance

decoding error probability

channel noise

LDPC($n = 1000$)

$n = 100$

$n = 250$

$n = 500$

$n = 1000$

$n = 2500$
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC\((n = 2500)\)
Asymptotic Average (Message Passing) Performance

decoding error probability

channel noise

LDPC\((n = 5000)\)
Asymptotic Average (Message Passing) Performance

decoding error probability

![Graph showing decoding error probability vs. channel noise for different values of \( n \) including \( n = 100, 250, 500, 1000, 2500, 5000, 10000 \). The graph illustrates the phase transitions in large graphical models, from physics to information theory, and computer science.

LDPC(\( n = 10000 \))
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC\((n = 25000)\)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC\((n = 50000)\)

channel noise
Asymptotic Average (Message Passing) Performance

Decoding error probability

LDPC\( (n = 100000) \)

Channel noise

Andrea Montanari

Phase transitions in large graphical models: from physics to information theory
Asymptotic Average (Message Passing) Performance

decoding error probability

![Graph showing decoding error probability vs. channel noise for different values of n.](image)

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

Decoding error probability

Channel noise

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
decoding error probability

\[ \text{channel noise} \]

\[ \text{LDPC}(n = +\infty) \]
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC\((n = +\infty)\)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)

channel noise
Asymptotic Average (Message Passing) Performance

decoding error probability

\[
\begin{array}{c}
\log_{10}(0.0) \\
\log_{10}(0.02) \\
\log_{10}(0.04) \\
\log_{10}(0.06) \\
\log_{10}(0.08) \\
\log_{10}(0.1) \\
\end{array}
\]

channel noise

LDPC(\(n = +\infty\))
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = \infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
Asymptotic Average (Message Passing) Performance

decoding error probability

LDPC($n = +\infty$)
Constraint satisfaction problems
$N$ variables: $\mathbf{x} = (x_1, x_2, \ldots, x_N), \; x_i \in \{0, 1\}$

$M$ $k$-clauses

$$(x_1 \lor \overline{x}_5 \lor x_7) \land (x_5 \lor x_8 \lor \overline{x}_9) \land \cdots \land (\overline{x}_{66} \lor \overline{x}_{21} \lor \overline{x}_{32})$$

Hereafter $k \geq 4$ (ask me why at the end)
Uniform measure over solutions

\[ F = \cdots \land (x_{i_1}(a) \lor \overline{x}_{i_2}(a) \lor \cdots \lor x_{i_k}(a)) \land \cdots \]

\( a \)-th clause

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

\[ \psi_a(x_{i_1}(a), \ldots, x_{i_k}(a)) = \begin{cases} 1 & \text{clause } a \text{ satisfied} \\ 0 & \text{otherwise} \end{cases} \]
(Factor) graph representation

Here: $N = 10$, $M = 4$

Distance: $i, j \in \{1, \ldots, N\} \mapsto d(i, j)$
Random $k$-satisfiability

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

$N, M \to \infty$ with $\alpha = M/N$ fixed.
Set of solutions of $F$ (cavity method):

\[ \alpha_d(k) \quad \alpha_c(k) \quad \alpha_s(k) \]
Qualitative changes in the correlation strength

\( \alpha < \alpha_d(k) \Rightarrow \text{Weak correlations.} \)

\( \alpha_d(k) < \alpha < \alpha_c(k) \Rightarrow \text{Strong point-to-set correlations.} \)

\( \alpha_c(k) < \alpha < \alpha_s(k) \Rightarrow \text{Strong point-to-point correlations.} \)
\[ \alpha < \alpha_c(k) \Rightarrow \text{Belief propagation is asymptotically correct.} \]

\[ \alpha_c(k) < \alpha < \alpha_s(k) \Rightarrow \text{Survey propagation.} \]
Conclusion 1: Peoples I should have cited/I should thank


Conclusion 2: If you want to know more about this...

- M. Mézard, A. M., *Upcoming book*
- 2007 IT Symposium → Statistical Physics tutorial
- 2007 StatPhys symposium → IT Plenary Talk
- google ee374