CLUSTER VARIATION METHOD:

FROM STATISTICAL MECHANICS TO MESSAGE PASSING ALGORITHMS

Alessandro Pelizzola

Dipartimento di Fisica, Politecnico di Torino

- Variational free energy
- Exactness, realizability, convexity, bounds
- Message–passing algorithms
- Double–loop algorithms
- Applications to disordered systems

Lav

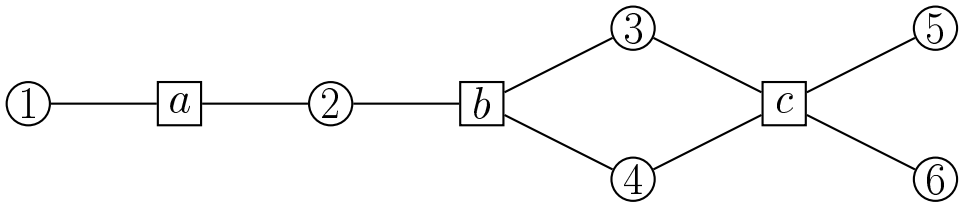
Probabilistic Graphical Models and Statistical Mechanics Models

Factor graph representation (esp. useful for non–pairwise interact

$$H(s_1, s_2, s_3, s_4, s_5, s_6) = H_a(s_1, s_2) + H_b(s_2, s_3, s_4) + H_c(s_3, s_4, s_5,$$

$$\phi_a = \exp(-H_a), \qquad \ldots$$

$$p(s_1, s_2, s_3, s_4, s_5, s_6) = \frac{1}{Z}\phi_a(s_1, s_2)\phi_b(s_2, s_3, s_4)\phi_c(s_3, s_4, s_5, s_6)$$



○ Variable node

□ Factor node

Statistical mechanics models:

$$H(s_1, s_2, \ldots s_N) = \sum_a H_a(s_a),$$

$$p(s_1, s_2, \ldots s_N) = \frac{1}{Z} \exp[-H(s_1, s_2, \ldots s_N)],$$

$$Z = \sum_{\{s_i\}} \exp[-H(s_1, s_2, \ldots s_N)]$$

Probabilistic graphical models (on an undirected graph, for simpl

$$p(s_1, s_2, \ldots s_N) = \frac{1}{Z} \prod_a \phi_a(s_a), \qquad s_a = \{s_i, i \in a\}$$

$$Z = \sum_{\{s_i\}} \prod_a \phi_a(s_a),$$

Directed graphs: replace potentials ϕ_a with conditional probabilities.

Different formulations of the same problem

One is often interested in computing **marginal distributions** and/or **configurations**.

Cluster variation method (CVM)

- Variational principle of equilibrium statistical mechanics:

$$F = -\ln Z = \min_p \mathcal{F}(p) = \min_p \sum_s [p(s)H(s) + p(s)\ln p($$

$$\sum_s p(s) = 1$$

$$\hat{p}(s) = \frac{1}{Z}\exp[-H(s)] = \operatorname{argmin}\mathcal{F}$$

$$\mathcal{F}(p) = F + \sum_s p(s)\ln\frac{p(s)}{\hat{p}(s)}$$

- *Cluster* or *Region*: subset α of the factor graph, such that if a fac
  belongs to α, all its variable nodes belong to α

- Cluster probability, cluster energy, cluster entropy:

$$p_\alpha(s_\alpha) = \sum_{s \setminus s_\alpha} p(s)$$
$$H_\alpha(s_\alpha) = \sum_{a \in \alpha} H_a(s_a)$$
$$S_\alpha = -\sum_{s_\alpha} p_\alpha \ln p_\alpha$$

- Cumulant expansion of the entropy:

$$S_\alpha = \sum_{\beta \subseteq \alpha} \tilde{S}_\beta \qquad \Leftrightarrow \qquad \tilde{S}_\beta = \sum_{\alpha \subseteq \beta} (-1)^{n_\alpha - n_\beta} S_\alpha$$
$$\mathcal{F}(p) = \sum_s p(s) H(s) - \sum_\beta \tilde{S}_\beta$$

- Truncation of the cumulant expansion to a set R of clusters made o
  maximal clusters and all their subclusters

$$\sum_{\beta} \tilde{S}_\beta \simeq \sum_{\beta \in R} \tilde{S}_\beta = \sum_{\alpha \in R} a_\alpha S_\alpha$$

$$\mathcal{F}(\{p_\alpha, \alpha \in R\}) = \sum_{\alpha \in R} a_\alpha \mathcal{F}_\alpha(p_\alpha)$$

$$\mathcal{F}_\alpha(p_\alpha) = \sum_{s_\alpha} \left[p_\alpha(s_\alpha) H_\alpha(s_\alpha) + T p_\alpha(s_\alpha) \ln p_\alpha(s_\alpha) \right]$$

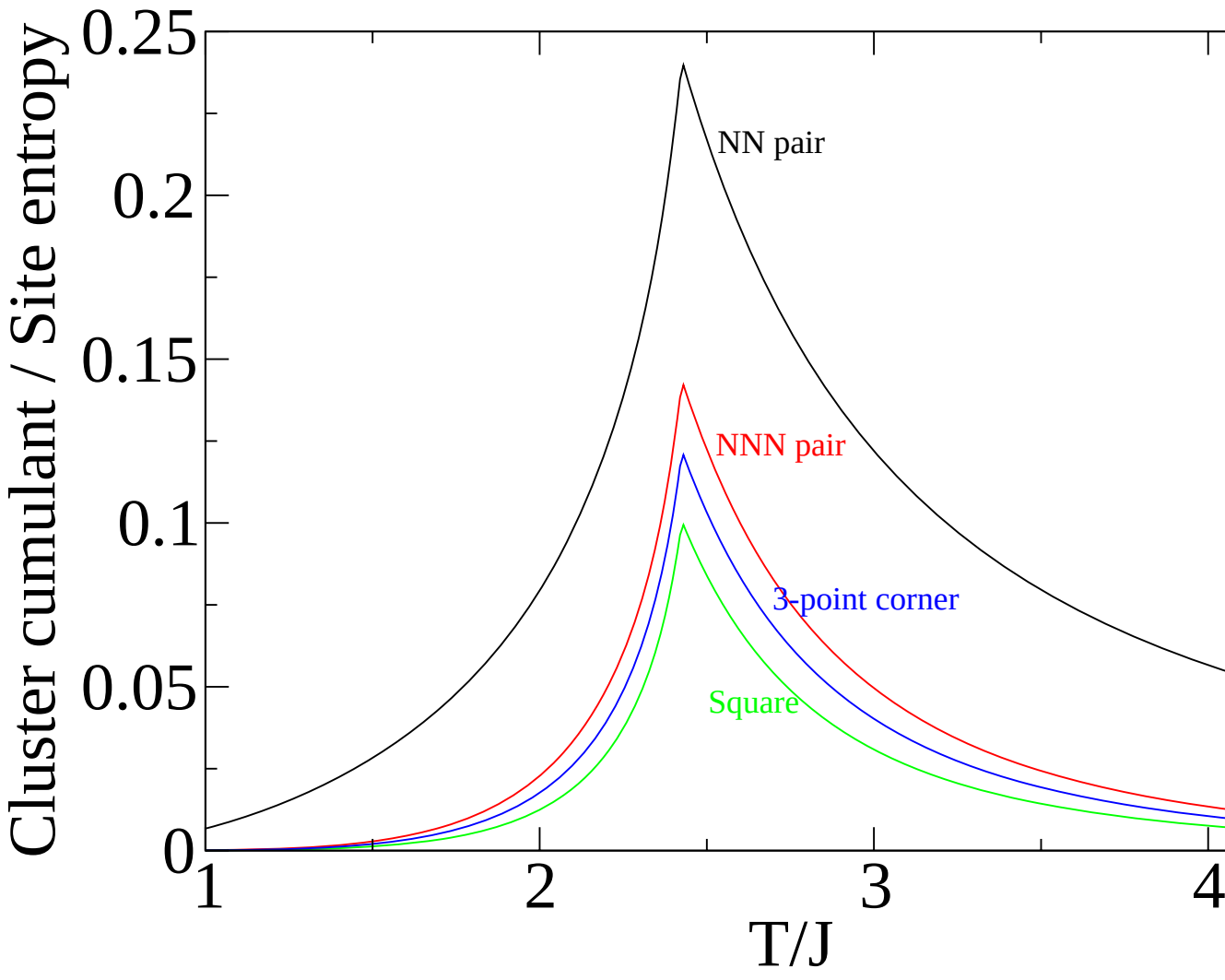
Meaning of cumulants

Two weakly correlated variables

$$p_{12}(s_1, s_2) = p_1^{(0)}(s_1) p_2^{(0)}(s_2) \left[1 + \varepsilon \, \delta p(s_1, s_2) \right], \qquad \varepsilon \ll 1$$

$$\tilde{S}_{12} = S_{12} - S_1 - S_2 = -\langle \ln \left[1 + \varepsilon \, \delta p(s_1, s_2) \right] \rangle = O(\varepsilon)$$

Behaviour: Ising model, square approximation



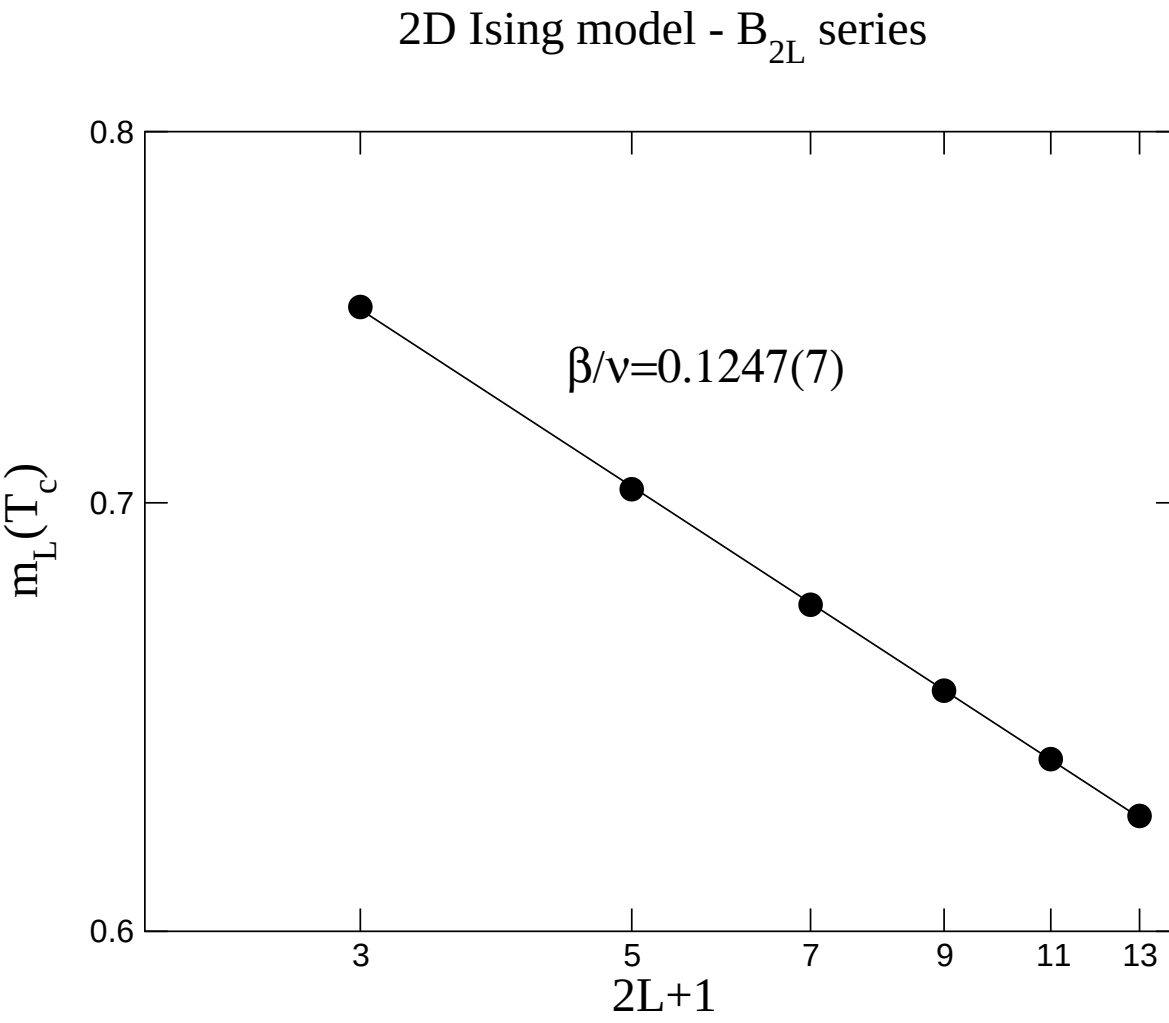
APPLICATIONS TO LATTICE MODELS, CRITICAL

Simple cubic Ising model in zero field, magnetization vs temperat

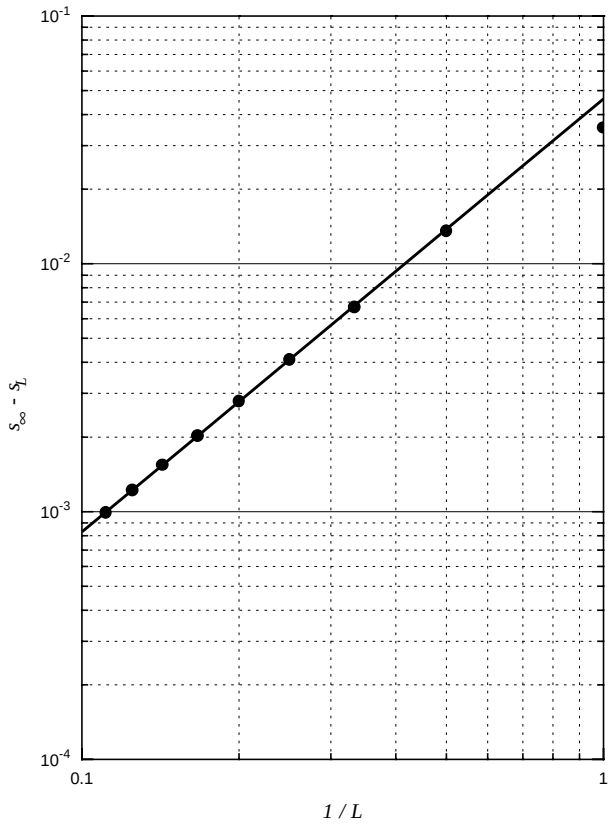
$$T_c \simeq 4.512 \div 4.515$$

T	m (18 site CVM)	m (MCRG by Talapov and Blöte)
3.4	0.8972562	0.8972440
3.5	0.8806417	0.8806366
3.6	0.8616750	0.8616735
3.7	0.8399256	0.8399255
3.8	0.8148173	0.8148161
3.9	0.7855490	0.7855416
4.0	0.7509519	0.7509251
4.1	0.7092094	0.7091249
4.2	0.6572414	0.6569722
4.3	0.5891051	0.5881361
4.4	0.4905811	0.4859045
4.5	0.3067063	0.2378014

Critical magnetization of the square Ising ferromagnet, B_{2L} approxim



Entropy of the triangular Ising antiferromagnet, B_{2L} approximati



Region–based free energy approximations

Given a set R of regions, assign *counting numbers* a_α to regions, such tha
counting of nodes) every factor node and every variable node is counted

$$\sum_{\alpha \in R, a \in \alpha} a_\alpha = 1 \qquad \forall a$$
$$\sum_{\alpha \in R, i \in \alpha} a_\alpha = 1 \qquad \forall i$$

Junction graph: R contains two types of regions, *large regions* a
regions, organized in a directed graph, with edges from large to smal
such that: (i) every edge connects a large region with a small region w
subset of the former; (ii) the subgraph of the regions containing a given
connected tree

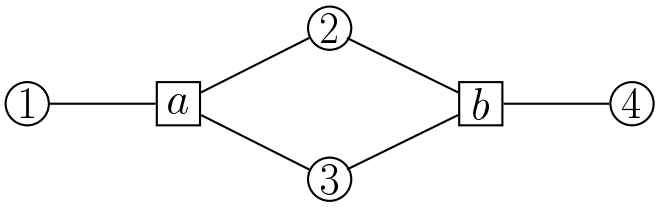
Bethe: large regions are made of a factor node and its variable nod
regions are variable nodes.

CVM: R is closed under intersection, which satisfies the additional c
(single counting of every subcluster)

$$\sum_{\alpha \in R, \beta \subseteq \alpha} a_\alpha = 1 \qquad \forall \beta \in R$$

Bethe is a special case of CVM only if no factor node shares more
variable node with another factor node.

Be careful when applying Bethe to non–pairwise interactions, e.g.



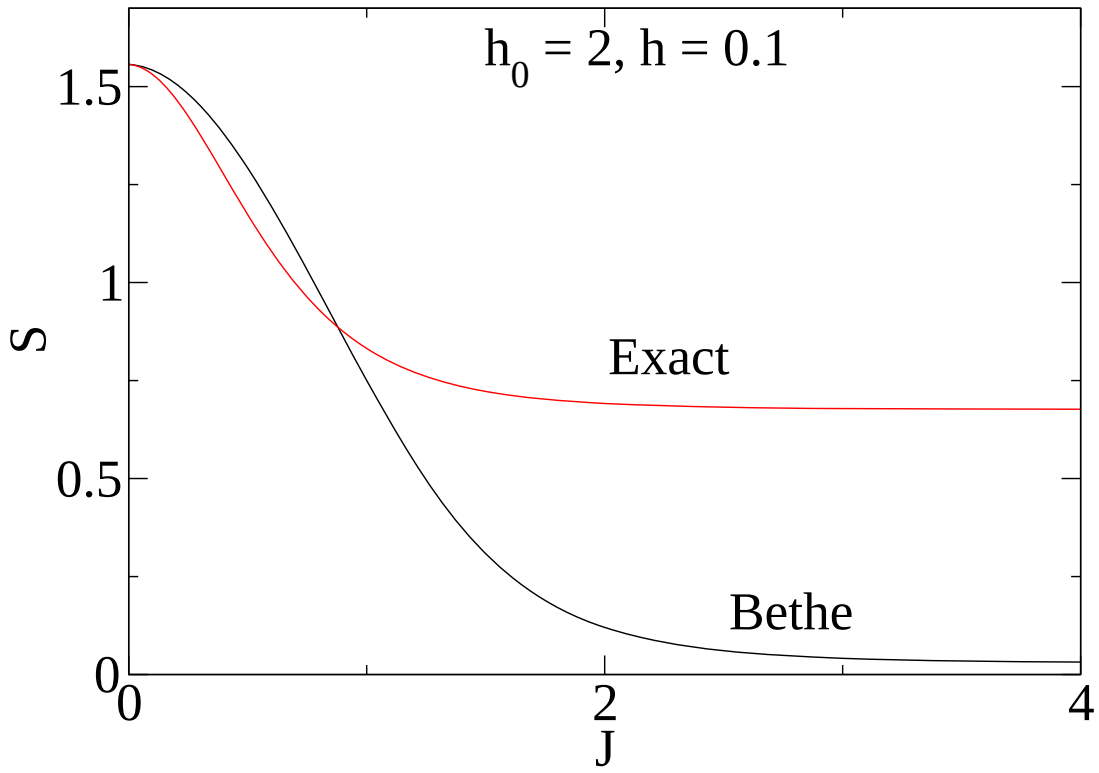
$$s_i = \pm 1, \qquad i = 1, 2, 3, 4$$

$$H_a(s_1, s_2, s_3) = -h_0 s_1 - \frac{h}{2}(s_2 + s_3) - J s_1 s_2 s_3$$

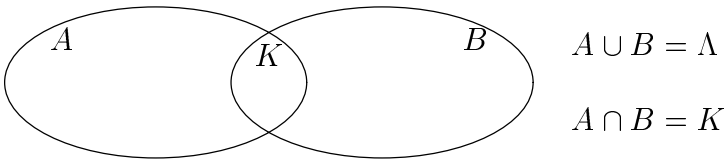
$$H_b(s_2, s_3, s_4) = -h_0 s_4 - \frac{h}{2}(s_2 + s_3) - J s_2 s_3 s_4$$

$$S = S(a123) + S(b234) - S(23)$$

$$S_{\text{Bethe}} = S(a123) + S(b234) - S(2) - S(3)$$



THE SIMPLEST EXAMPLE



- No interactions between $A' = A \setminus K$ and $B' = B \setminus K$: $H = H_A(s_A) +$
- The probability factors: $p_\Lambda = \dfrac{p_A p_B}{p_K}$

$$p_\Lambda = \frac{e^{-H}}{\mathrm{Tr} e^{-H}} = \frac{1}{Z} e^{-H}$$

$$p_A = \sum_{\Lambda \setminus A} p_\Lambda = \frac{1}{Z} e^{-H_A} \sum_{B'} e^{-H_B}$$

$$p_B = \sum_{\Lambda \setminus B} p_\Lambda = \frac{1}{Z} e^{-H_B} \sum_{A'} e^{-H_A}$$

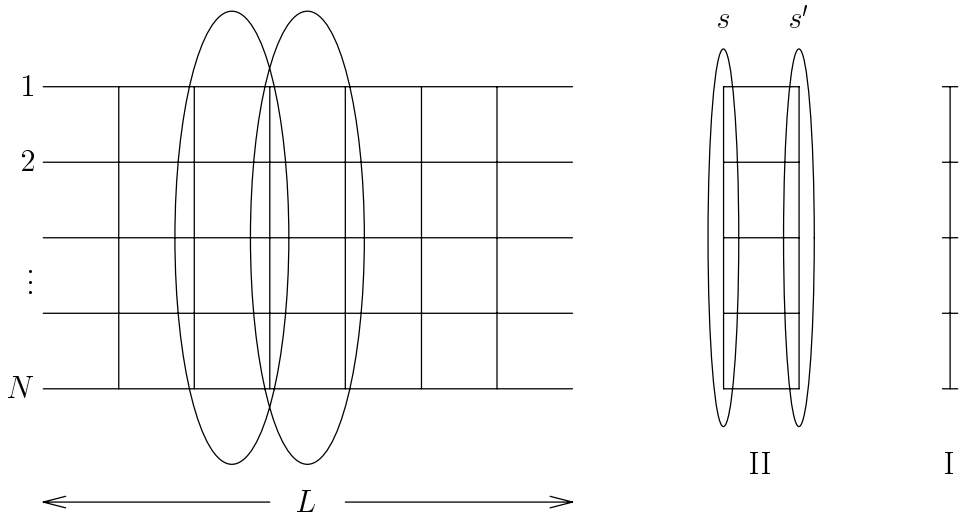
$$p_K = \sum_{A'} p_A = \sum_{B'} p_B = \frac{1}{Z} \sum_{A'} e^{-H_A} \sum_{B'} e^{-H_B}$$

- The cluster variation method with $R = A, B, K$ is exact, since:

$$S_\Lambda = S_A + S_B - S_K$$

ONE DIMENSIONAL SYSTEMS

- Strip of width N:



$$p_\Lambda = \frac{\prod_{\alpha\in \mathrm{II}} p_\alpha}{\prod_{\alpha\in \mathrm{I}} p_\alpha} \qquad \Rightarrow \qquad S_\Lambda = \sum_{\alpha\in \mathrm{II}} S_\alpha - \sum_{\alpha\in \mathrm{I}} S_\alpha$$

The cluster variation method with $R = \mathrm{II}\cup\mathrm{I}$ is exact ($L = 1$: the Beth approximation is exact for the chain).

- In the thermodynamic limit $L \to \infty$ (open boundary conditions) tra
  invariance is recovered:

$$\frac{F}{L} = \min \sum_{\Lambda} \left(p_{\mathrm{II}} H_{\mathrm{II}} + p_{\mathrm{II}} \ln p_{\mathrm{II}} - p_{\mathrm{I}} \ln p_{\mathrm{I}} \right)$$

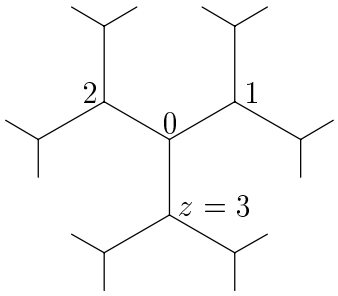
  Solving for p_{II} we recover the transfer matrix formalism:

$$\frac{F}{L} = - \ln \max \left\{ \sum_{s,s'} p_{\mathrm{I}}^{1/2}(s) \exp \left[-H_{\mathrm{II}}(s,s') \right] p_{\mathrm{I}}^{1/2}(s') \right\}$$

$$\sum_{s} p_{\mathrm{I}}(s) = 1$$

TREE-LIKE LATTICES (when Bethe ≡ CVM)

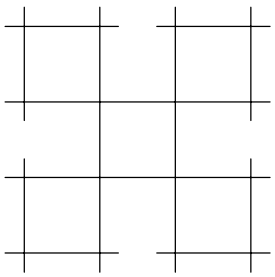
- Bethe lattice (interior of a Cayley tree), NO LOOPS:



$$p_\Lambda = \frac{\prod_{\alpha \in \text{LINKS}} p_\alpha}{\prod_{\alpha \in \text{SITES}} p_\alpha^{z_\alpha - 1}} \quad \Rightarrow \quad S_\Lambda = \sum_{\alpha \in \text{LINKS}} S_\alpha - \sum_{\alpha \in \text{SITES}} (z_\alpha - 1)$$

The cluster variation method with $R = \text{LINKS} \cup \text{SITES}$ is exact (z
Bethe-Peierls approximation is exact for the simple chain).

- Cactus lattice (interior of a Husimi tree):



$$p_\Lambda = \frac{\prod\limits_{\alpha\in\mathrm{PLAQS}} p_\alpha}{\prod\limits_{\alpha\in\mathrm{SITES}} p_\alpha} \quad \Rightarrow \quad S_\Lambda = \sum_{\alpha\in\mathrm{PLAQS}} S_\alpha - \sum_{\alpha\in\mathrm{SITES}} S_\alpha$$

The cluster variation method with $R = \mathrm{PLAQS} \cup \mathrm{SITES}$ is exact.

REALIZABILITY

$$s_i = \pm 1, \qquad i = 1, 2, 3$$

$$p_i(s_i) = 1/2 \qquad i = 1, 2, 3$$
$$p_{ij}(s_i, s_j) = \frac{1 + c s_i s_j}{4} \qquad i < j$$

Q. Is there a global $p(s_1, s_2, s_3)$ marginalizing to the above probabilities

A. Only for $-1/3 \leq c \leq 1$

Q. Is this global p given by

$$\frac{p_{12}(s_1, s_2) p_{13}(s_1, s_3) p_{23}(s_2, s_3)}{p_1(s_1) p_2(s_2) p_3(s_3)} \quad ?$$

A. Only for $c = 0$

Extending local thermodynamic states on a *square lattice*:

$$S \simeq \sum_{[ijkl]} S_{[ijkl]} - \sum_{\langle ij \rangle} S_{ij} + \sum_{i} S_i$$

does not mean that

$$\prod_{[ijkl]} p_{[ijkl]}(s_i, s_j, s_k, s_l) \prod_{\langle ij \rangle} p_{ij}^{-1}(s_i, s_j) \prod_{i} p_i(s_i)$$

is a good approximation to the true equilibrium distribution.

In general such a global p does not marginalize to $p_{[ijkl]}$, p_{ij}, p_i and is normalized $\Rightarrow$ CVM free energy is not, in general, an upper bound.

The above is well defined only if

- odd correlations vanish
- $\langle s_i s_k \rangle_{\langle\langle ik \rangle\rangle} = \langle s_i s_j \rangle^2_{\langle ij \rangle}$

DISORDER POINTS

- Ising model with competitive interactions: for instance, on the square

$$H = -K_1 \sum_{\langle ij \rangle} s_i s_j - K_2 \sum_{\langle\langle ij \rangle\rangle} s_i s_j - K_4 \sum_{[ijkl]} s_i s_j s_k s_l,$$

with K_1 (nearest neighbour coupling) > 0, K_2 (next nearest neighb
pling) < 0, K_4 (plaquette coupling).

- Integrable case:

$$\cosh(2K_1) = \frac{e^{4K_2+2K_4} + e^{-4K_2+2K_4} + 2e^{-2K_2}}{2\left(e^{2K_2} + e^{2K_4}\right)}$$

Free energy:

$$F = -\ln\left[\exp(-K_4) + \exp(K_4 - 2K_2)\right]$$

- The cluster variation method with $R = \{$plaquettes and their subclus
  with maximal clusters larger than the plaquette) is exact in the i
  case.

  The probability factors:

$$p_\Lambda = \prod_{\alpha\in\mathrm{PLAQS}} p_\alpha \left(\prod_{\alpha\in\mathrm{LINKS}} p_\alpha \right)^{-1} \prod_{\alpha\in\mathrm{SITES}} p_\alpha \qquad \Rightarrow$$

$$S_\Lambda = \sum_{\alpha\in\mathrm{PLAQS}} S_\alpha - \sum_{\alpha\in\mathrm{LINKS}} S_\alpha + \sum_{\alpha\in\mathrm{SITES}} S_\alpha$$

- Correlation function $\Gamma(x,y) = \langle s(x_0,y_0)s(x_0+x,y_0+y)\rangle$:

$$\Gamma(x,y) = g^{|x|+|y|}, \qquad g = \frac{\exp(-4K_2) - \cosh(2K_1)}{\sinh(2K_1)}$$

- Many-site correlation functions can be calculated, e.g.

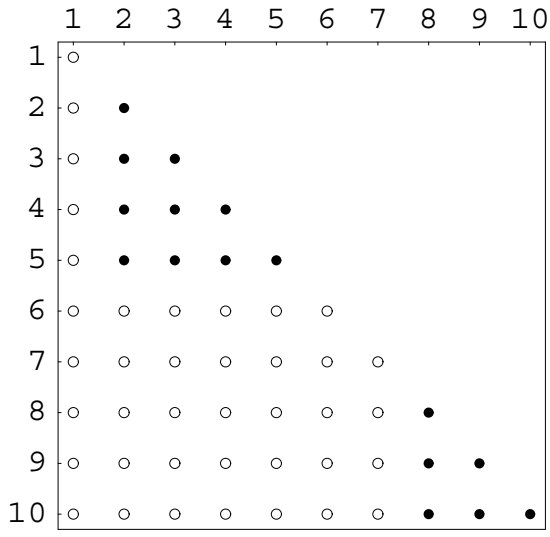
  $q = \langle s_i s_j s_k s_l \rangle_{[ijkl]}$

$$q = \frac{e^{4K_4}\left(1 - e^{8K_2}\right) + 4e^{2K_2}\left(e^{2K_4} - e^{2K_2}\right)}{e^{4K_4}\left(1 - e^{8K_2}\right) + 4e^{2K_2}\left(e^{2K_4} + e^{2K_2}\right)}$$

- Similar results can be derived for other two-dimensional models.

EATON MODEL of PROTEIN FOLDING

$$H = \sum_{i=1}^{L} \sum_{j=i}^{L} h_{i,j} x_{i,j}, \qquad x_{i,j} = \prod_{k=i}^{j} x_k, \qquad x_k = 0, 1$$



$\bullet = 1, \circ = 0$

Constraints: $x_{i,j} = x_{i+1,j} x_{i,j-1}$

Probability factors in a CVM–like way due to constraints and locality o
tions

Local interactions $\Rightarrow$

$$p(\{x_{i,j}\}) = \frac{p^{(1,2)}p^{(2,3)}\dots p^{(L-1,L)}}{p^{(2)}\dots p^{(L-1)}}$$

$$\begin{cases} p^{(j)} = & \text{prob of row } j \\ p^{(j,j+1)} = & \text{joint prob of rows } j \text{ and } j+1 \end{cases}$$

Constraints $\Rightarrow$ (for instance)

$$\begin{aligned} p^{(j)}(0,\dots 0_i, 1_{i+1}, \dots 1) &= p^{(j)}_{i,i+1}(0,1) = \\ &= \frac{p^{(j)}_{1,2}(0,0)\cdots p^{(j)}_{i,i+1}(0,1)\cdots p^{(j)}_{L-1,L}(1,1}{p^{(j)}_2(0)\cdots p^{(j)}_i(0)p^{(j)}_{i+1}(1)\cdots p^{(j)}_{L-1}(1)} \end{aligned}$$

and similarly for $p^{(j,j+1)}$.

It follows

$$p(\{x_{i,j}\}) = \prod_{\alpha\in R} p_\alpha(x_\alpha)^{a_\alpha},$$

where $R = \{$square plaquettes, corners (on the diagonal), and their subc

CVM as an APPROXIMATION

- Leaving apart exactly solvable cases,

$$S \simeq \sum_{\alpha \in R} a_\alpha S_\alpha$$

  can be a reasonable approximation, but

$$\prod_{\alpha \in R} p_\alpha^{a_\alpha}$$

  is not a probability.

- **Upper bound**: the CVM free energy is not an upper bound of the e
  energy (contrast with mean field)

- **Convexity**: due to negative a_α coefficients, convexity is no more gu
  (and hence multiple minima can appear), unless

$$\forall S \subseteq R \qquad \sum_{\alpha \in R_S} a_\alpha \geq 0 \qquad R_S = \{\alpha \in R | \exists \beta \subseteq \alpha, \beta \in S$$

- Anyway, we are left with the problem of finding minima of $\mathcal{F}$ with norm and compatibility constraints on probabilities:

$$\sum_{s_\alpha} p_\alpha(s_\alpha) = 1 \qquad \forall \alpha \in R,$$

$$p_\alpha(s_\alpha) = \sum_{s_\beta \setminus s_\alpha} p_\beta(s_\beta) \qquad \forall \alpha \subset \beta \in R.$$

Algorithms to find min in F_{CVM}

Constraints are always implemented through

- Lagrange multipliers, *or*
- effective/cavity fields, *or*
- messages,

equivalently.

(G)BP: (generalized) belief propagation

1. Solve analytically w.r.t. the p_α's and
2. solve constraints w.r.t. messages by iteration.

Very *fast*, *single loop* algorithm, often *not convergent*.

NIM: natural iteration method

CCCP: concave–convex procedure

HAK: class of algorithms by Heskes, Albers and Kappen

1. Outer loop: update probabilities
2. Inner loop: update multipliers, or messages, to solve constraints

Slower, *double loop* algorithms, *proof of convergence* exists (except in a
cases).

Details: Bethe approximation (Ising example)

$$
\begin{aligned}
\mathcal{F} \;=\; & -\sum_i h_i \sum_{s_i} s_i p_i(s_i) - \sum_{\langle ij\rangle} J_{ij} \sum_{s_i,s_j} s_i s_j p_{ij}(s_i,s_j) \\
& + \sum_{\langle ij\rangle} \sum_{s_i,s_j} p_{ij} \ln p_{ij} - \sum_i (z_i - 1) \sum_{s_i} p_i \ln p_i \\
& + \sum_i \lambda_i (\sum_{s_i} p_i - 1) + \sum_{\langle ij\rangle} \lambda_{ij} (\sum_{s_i,s_j} p_{ij} - 1) \\
& + \sum_{\langle ij\rangle} \Bigg[\sum_{s_i} \mu_{i,j}(s_i)(p_i - \sum_{s_j} p_{ij}) + \sum_{s_j} \mu_{j,i}(s_j)(p_j - \sum_{s_i} p_{ij}
\end{aligned}
$$

Thouless–Anderson–Palmer (TAP)

To obtain the corresponding free energy from the Bethe free energy use

$$p_i(s_i) = \frac{1 + s_i m_i}{2} \qquad p_{ij}(s_i, s_j) = \frac{1 + s_i m_i + s_j m_j + s_i s_j c_{ij}}{4},$$

then solve analytically for the c_{ij}'s and expand to second order in β.

Natural Iteration Method (NIM)

Define site probs as *marginals* of pair probs

$$p_i = \frac{1}{z_i} \sum_{j \,\mathrm{NN}\, i} \sum_{s_j} p_{ij},$$

then solve stationarity w.r.t. pair probs to get the basic iterative eqs.

$$\begin{aligned} p_{ij}(s_i, s_j) \;=\; \exp \Bigg[& -1 - \lambda_{ij} + \frac{1}{z_i} h_i s_i + \frac{1}{z_j} h_j s_j + J_{ij} s_i s_j \\ & + \mu_{i,j}(s_i) + \mu_{j,i}(s_j) - \frac{1}{z_i} \sum_{k \,\mathrm{NN}\, i} \mu_{i,k}(s_i) - \frac{1}{z_j} \sum_{k \,\mathrm{NN}\, j} \mu_{j,k}(s_j \\ & \times \left[p_i(s_i)\right]^{1-1/z_i} \left[p_j(s_j)\right]^{1-1/z_j} \end{aligned}$$

At each iteration, find Lagrange mults, e.g. by iteration:

$$\mu_{i,j}(s_i) = \hat{\mu}_{i,j}(s_i) + b \ln \frac{p_i(s_i)}{\sum_{s_j} p_{i,j}(s_i, s_j)}$$

Effective fields

Stationarity is solved by

$$p_i(s_i) \;=\; \exp\left[F_i + \left(h_i + \sum_{k\,\mathrm{NN}\,i} h_{i,k}\right) s_i\right]$$

$$p_{ij}(s_i, s_j) \;=\; \exp\left[F_{ij} + \left(h_i + \sum_{k\,\mathrm{NN}\,i}^{k\neq j} h_{i,k}\right) s_i + \left(h_j + \sum_{k\,\mathrm{NN}\,j}^{k\neq i} h_{j,k}\right) s_j +\right.$$

and compatibility leads to

$$h_{i,j} = \tanh^{-1}\left[\tanh\left(h_j + \sum_{k\,\mathrm{NN}\,j}^{k\neq i} h_{j,k}\right)\tanh J_{ij}\right].$$

The corresponding iterative algorithm is nothing but **Belief propagatio**

Cavity fields

Cavity bias (factor $\to$ variable): $= h_{i,j}$

Cavity field (variable $\to$ factor): $= \sum_{k\,\mathrm{NN}\,i}^{k\neq j} h_{i,k}$

Belief propagation (BP)

Pair and site probs are rewritten in terms of messages:

$$p_i(s_i) = \exp(h_i s_i) \prod_{j \mathrm{NN} i} m_{j\to i}(s_i)$$

$$p_{ij}(s_i, s_j) = \exp\left(-\lambda + h_i s_i + h_j s_j + J_{ij} s_i s_j\right) \times$$

$$\times \prod_{k \mathrm{NN} i, k\neq j} m_{k\to i}(s_i) \prod_{k \mathrm{NN} j, k\neq i} m_{k\to j}(s_j).$$

The messages are related to Lagrange mults by

$$\exp\left[\mu_{i,j}(s_i)\right] = \exp(h_i s_i) \prod_{k \mathrm{NN} i, k\neq j} m_{k\to i}(s_i).$$

As the effective fields, messages can be determined iteratively by impos
patibility

$$m_{j\to i}(s_i) \propto \sum_{s_j} \exp\left(h_j s_j + J_{ij} s_i s_j\right) \prod_{k \mathrm{NN} j, k\neq i} m_{k\to j}(s_j).$$

When convergent, fixed point is a minimum of the Bethe free energy.

Generalized belief propagation (GBP)

$$p_\alpha(s_\alpha) \propto \left[\exp\left(-\sum_{a\in\alpha} H_a(s_a)\right)\right] \prod_{\beta\subseteq\alpha} \prod_{\gamma\in \mathrm{Par}(\beta)}^{\gamma\not\subseteq\alpha} m_{\gamma\to\beta}(s_\beta)$$

where

$$m_{\gamma\to\beta}(s_\beta) = \exp(\mu_{\gamma\to\beta}(s_\beta))$$

and $\mu_{\gamma\to\beta}(s_\beta)$ is the Lagrange multiplier for the constraint

$$\sum_{\beta\subseteq\alpha\in R} a_\alpha p_\alpha(s_\beta) = \sum_{\gamma\subseteq\alpha\in R} a_\alpha p_\alpha(s_\beta).$$

Concave convex procedure (CCCP)

The free energy is split into concave + convex:

$$\mathcal{F}[\{p_\alpha\}] = \mathcal{F}_{\rm vex}[\{p_\alpha\}] + \mathcal{F}_{\rm cave}[\{p_\alpha\}].$$

Iterative equations are derived from:

$$\nabla \mathcal{F}_{\rm vex}[\{p_\alpha^{(t+1)}\}] = -\nabla \mathcal{F}_{\rm cave}[\{p_\alpha^{(t)}\}].$$

In the Bethe case one chooses

$$\mathcal{F}_{\rm cave}[\{p_i\}] = -\sum_i z_i \sum_{s_i} \left[p_i(H_i + \ln p_i)\right],$$
$$\mathcal{F}_{\rm vex}[\{p_i, p_{ij}\}] = \mathcal{F} - \mathcal{F}_{\rm cave}.$$

Algorithm by Heskes, Albers and Kappen (HAK)

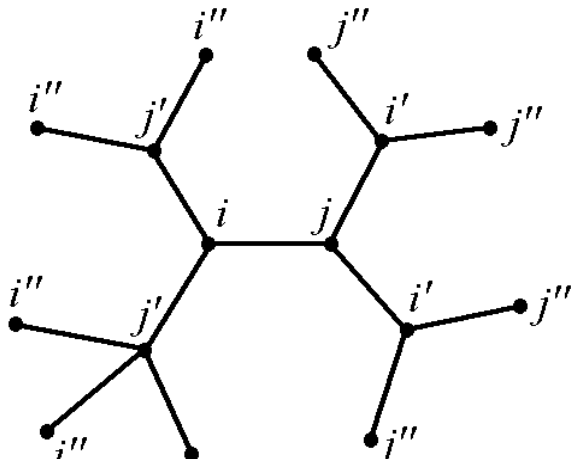
$$p^{(t+1)} = \text{argmin}_p \, \mathcal{F}_{\text{vex}}(p, p^{(t)})$$

where $p \equiv \{p_\alpha, \alpha \in R\}$ and $\mathcal{F}_{\text{vex}}(p, p')$ is a *convex* function of p properties

- $\mathcal{F}_{\text{vex}}(p, p') \geq \mathcal{F}(p)$
- $\mathcal{F}_{\text{vex}}(p, p) = \mathcal{F}(p)$

The free energy *decreases* at each iteration

CP (conditional probability) algorithm



$$p_{ij\{i'j'\}}(x_i, x_j, \{x_{i'}, x_{j'}\}) = p_{ij}(x_i, x_j) \prod_{i'} \frac{p_{ji'}(x_j, x_{i'})}{\sum_{\tilde{x}_{i'}} p_{ji'}(x_j, \tilde{x}_{i'})} \prod_{j'} \frac{p_{ij'}(x_i$$

$$p_{\{i'j'\}}(\{x_{i'}, x_{j'}\}) = \sum_{x_i, x_j} p_{ij\{i'j'\}}(x_i, x_j, \{x_{i'}, x_{j'}\})$$

$$p_{ij}(x_i, x_j) = \sum_{\{x_{i'}, x_{j'}\}} \frac{w_{ij\{i'j'\}}(x_i, x_j, \{x_{i'}, x_{j'}\})}{\sum_{\tilde{x}_i, \tilde{x}_j} w_{ij\{i'j'\}}(\tilde{x}_i, \tilde{x}_j, \{x_{i'}, x_{j'}\})} \; p_{\{i'j'\}}(\{x_{i'}, x_j$$

CP fixed points are minima of the **Bethe** free energy

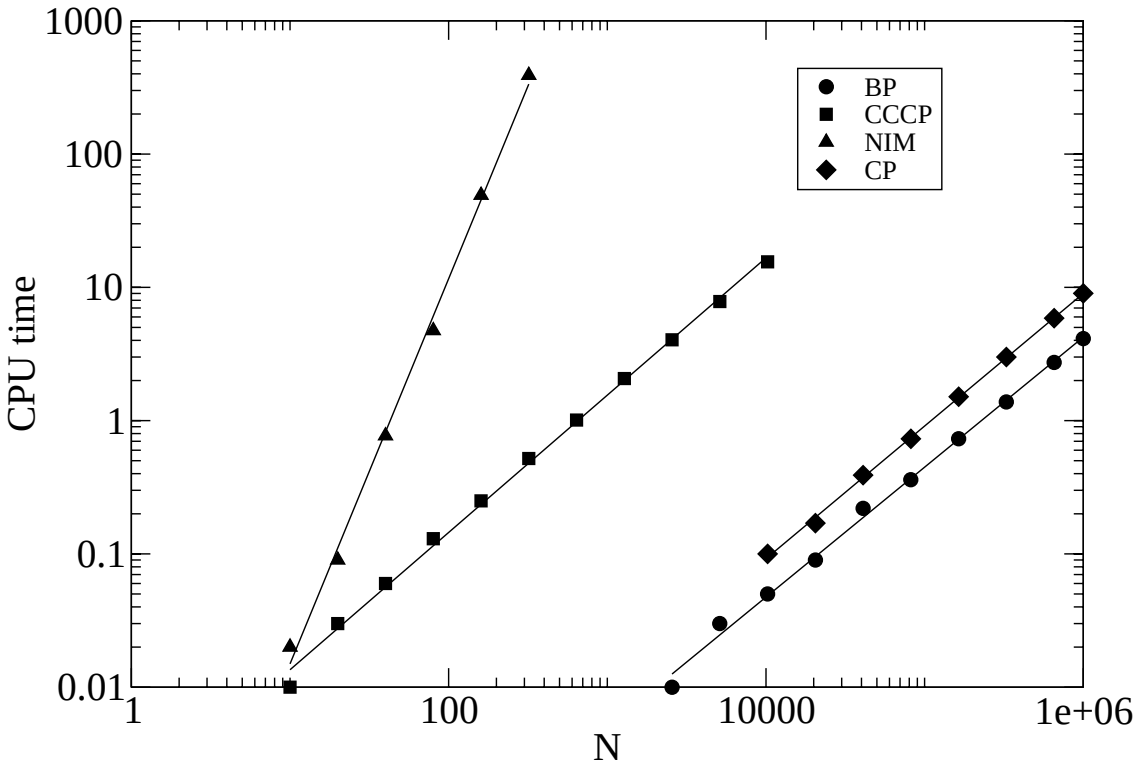
Analog to *hard–spin mean–field*

Converges more often than BP

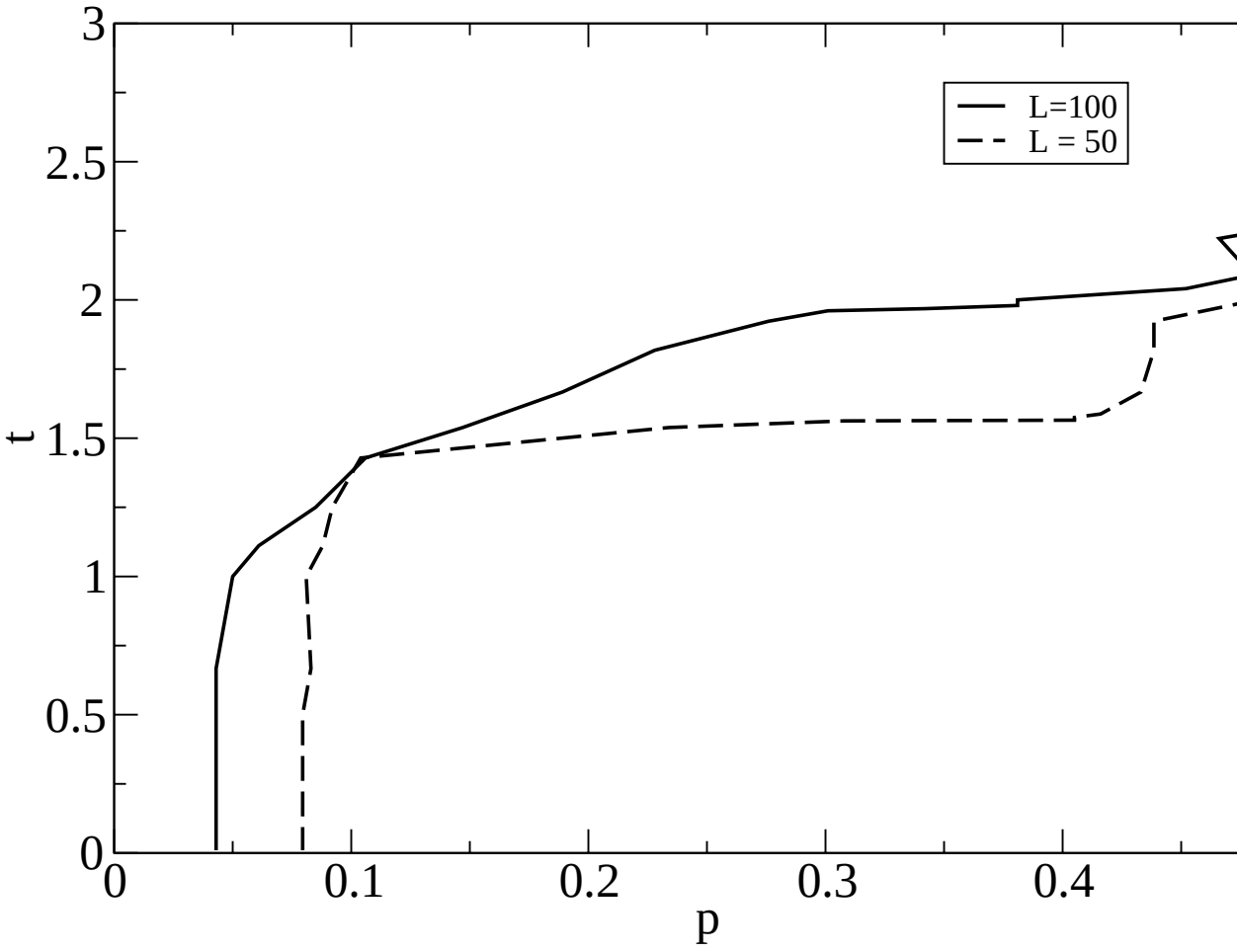
Comparable with best double loop algorithms, **slower** ($\sim$ 1 order of ma
than BP

SPEED OF ALGORITHMS

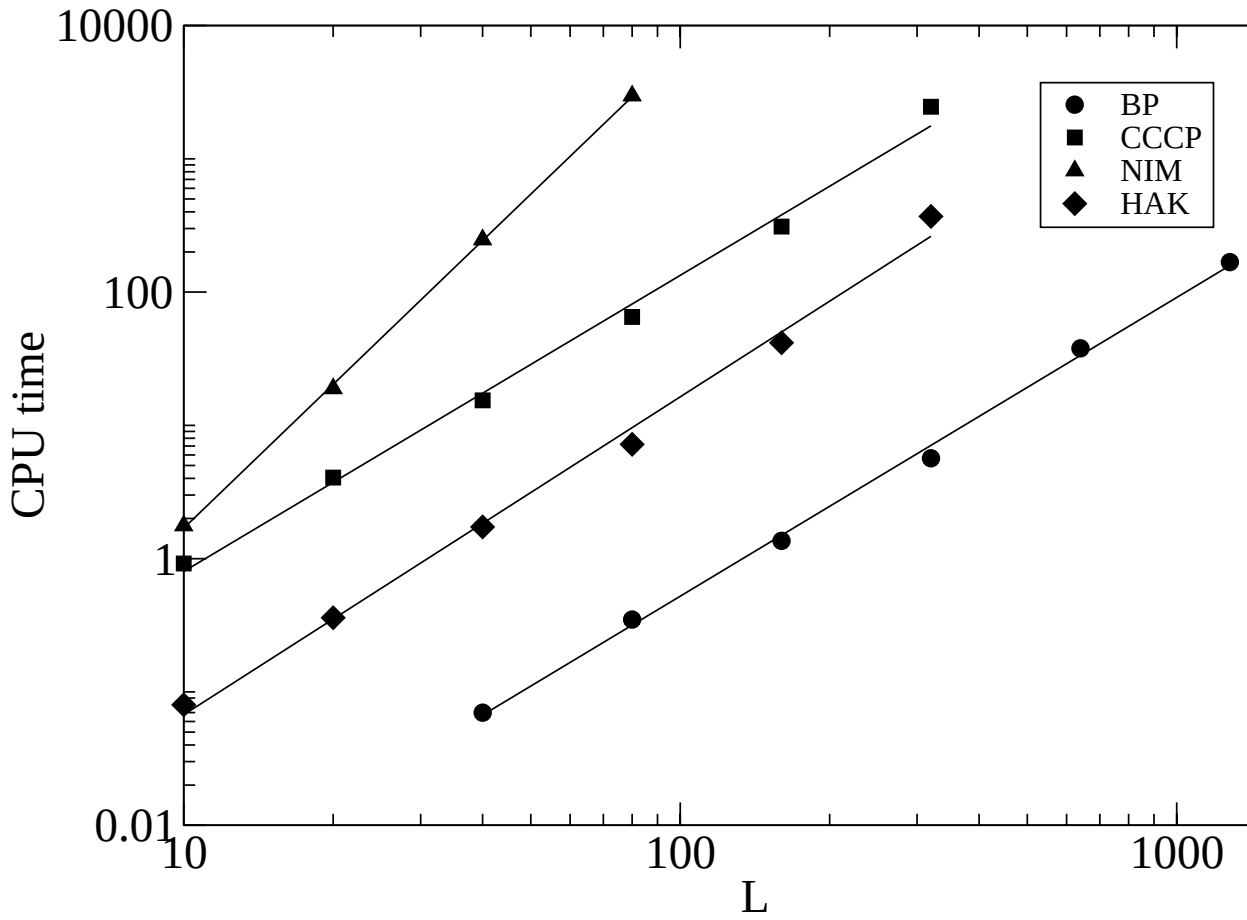
1D Ising model in random field, exact solution through Bethe approxima
energy is *convex over the constraint set* (HAK cannot be distinguished
on this graph)



2D Edwards–Anderson model with a fraction p of AF bonds: convergenc
for the BP algorithm with ferromagnetic initial condition



CPU times for the 2d model: paramagnetic case (CP $\sim$ HAK)



CPU times for the 2d model: ferromagnetic case (CP $\sim$ HAK)