STATISTICAL MECHANICS OF THE INVERSE ISING MODEL

Mauro Cirio

Supervisors:

Prof. Michele Caselle | Prof. Riccardo Zecchina

July 2009

INTRODUCTION

SUMMARY OF THE PRESENTATION

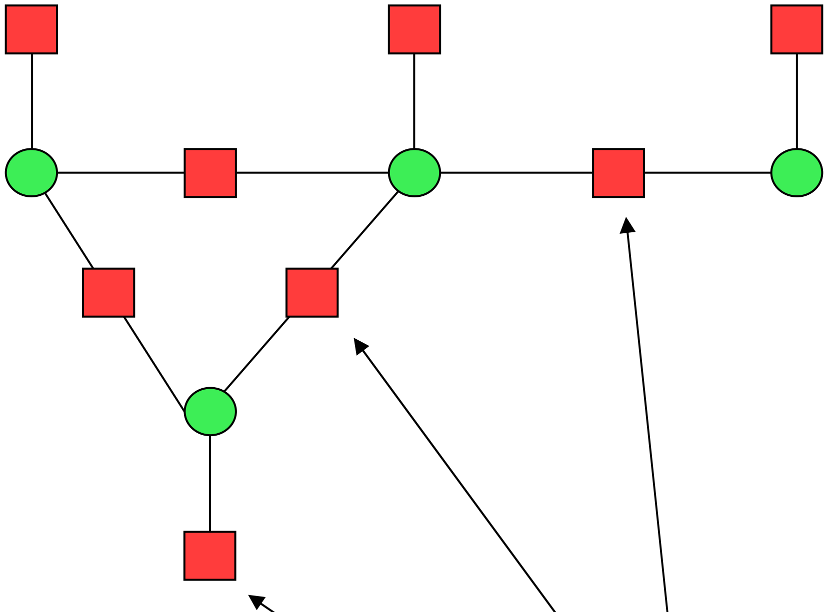
- Definition of the direct and inverse problem
- Approximation methods of the direct problem: variational approaches (Bethe)
- Overview of direct and inverse algorithms
- Simulations

GOALS OF THE THESIS

- Use of algorithms which generalize Bethe approximation (Gbp) in order to solve the inverse problem
- Upgrade of the Gbp algorithm to get a more accurate calculation of the correlations and application to the inverse problem

ISING MODEL

$$H(\{x\}) = -\sum_{<i,j>} J_{ij} x_i x_j - \sum_i h_i x_i$$



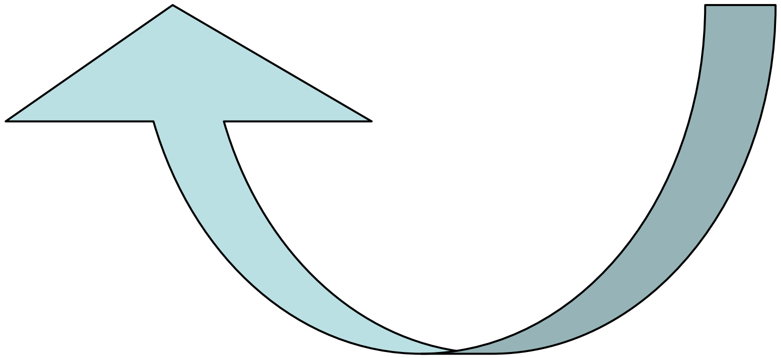
We can represent the usual Boltzmann probability distribution with a factor graph:

$$P(\{x\}) \propto e^{-\beta\, H(\{x\})} = e^{\beta \sum_i h_i x_i}\, e^{\beta \sum_{<i,j>} J_{ij} x_i x_j}$$

ISING MODEL

$$H(\{x\}) = -\sum_{<i,j>} J_{ij} x_i x_j - \sum_i h_i x_i$$

$$(h_i, J_{ij}) \longrightarrow (<x_i>, <x_i x_j>)$$



Inverse Problem!

EXAMPLES OF APPLICATIONS

- Neuron networks reconstruction
  [E.Schneidman, M.J.Berry, R.Segev, W.Bialek (2006)]
- Genes networks reconstruction
  [A.Braustein, A.Pagani, M.Weigt, R.Zecchina (2008)]
- Protein networks recontructions
  [G.Tkacik (2007)]

Why Ising model? ⇒ Maximum entropy principle

VARIATIONAL APPROACHES

Mean field approximations

$$G[P] = < E >_P - T\, S_P \geq G[\tilde{P}] = F$$

Gibbs Free Energy — Boltzmann Distr. — Helmoltz Free Energy

Mean field approximation ⟷ Minimization of G
In a subdomain

⇨ We chose a “form” for P
and we minimize the functional G

VARIATIONAL APPROACHES

Bethe approximation

BELIEFS:

$$\sum_{x_i} b_i(x_i) = 1$$

$$\sum_{x_i, x_j} b_{ij}(x_i, x_j) = 1$$

Local consistency:

$$\sum_{x_j} b_{ij}(x_i, x_j) = b_i(x_i)$$

$$p(\{x\}) = \prod_{<i,j>} b_{ij}(x_i, x_j) \prod_i [b_i(x_i)]^{1-q_i}$$

$$G = \sum_{<i,j>} \sum_{x_i, x_j} b_{ij}(x_i, x_j)\left[E_{ij}(x_i, x_j) + \log b_{ij}(x_i, x_j)\right] - \sum_i (q_i - 1) \sum_{x_i} b_i(x_i)\left[E_i(x_i) + \log b_i(x_i)\right]$$

Minimum equations

Constraints equations $\sum_{x_j} b_{ij}(x_i, x_j) = b_i(x_i)$

$$b_i(x_i) = f_1(m, Couplings)$$

$$b_{ij}(x_i, x_j) = f_2(m, Couplings)$$

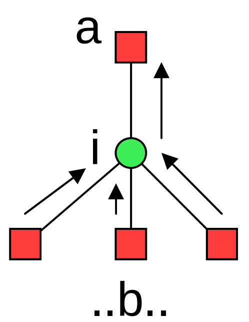
Messages equations at the fix point

BP

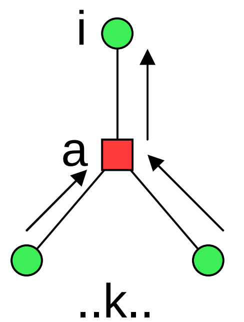
Direct Ising

BP equations

$$\nu_{i\to a}^{(t+1)}(x_i) \cong \prod_{b\in\partial j\setminus a} \eta_{b\to i}^{(t)}(x_i)$$



$$\eta_{a\to i}^{(t)}(x_i) \cong \sum_{x_{\partial a\setminus i}} \psi_a(x_{\partial a}) \prod_{k\in\partial a\setminus i} \nu_{k\to a}^{(t)}(x_k)$$

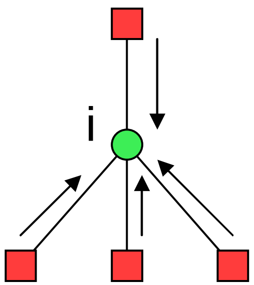


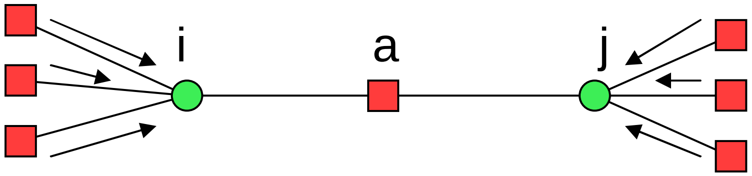
BP

Direct Ising

$$b_i(x_i) \cong \prod_{a \in \partial i} \eta^{(*)}_{a \to i}(x_i)$$

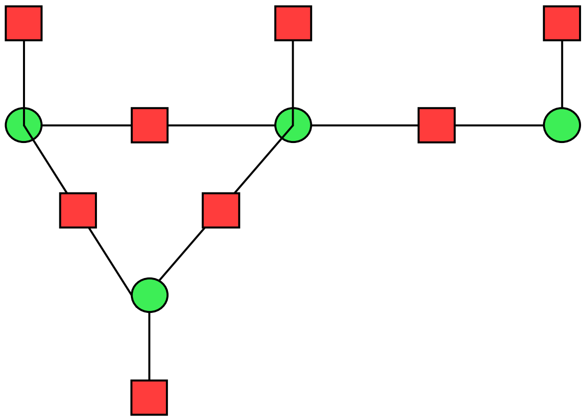
$$b_{ij}(x_i, x_j) \cong \psi_a(x_i, x_j) \prod_{s \in \partial i \setminus a} \eta^{(*)}_{s \to i}(x_i) \prod_{t \in \partial j \setminus a} \eta^{(*)}_{t \to j}(x_j)$$





GBP

[J.S.Yedida, W.T. Freeman, Y.Weiss (2001)]
Direct Ising



Let's define:

$$\begin{cases} U = \sum_R c_R \sum_{x_R} b_R(x_R) E_R(x_R) \\ S = -\sum_R c_R \sum_{x_R} b_R(x_R) Log(b_R(x_R)) \end{cases}$$

R := variables close to a function node → Bethe approximation

$$c_R = 1 - \sum_{U \in A(R)} c_U$$

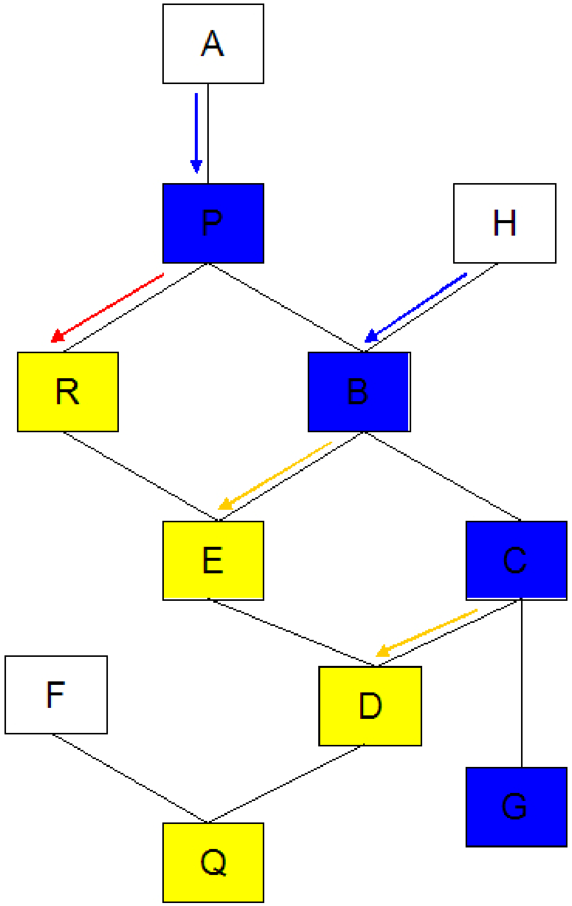
Generalization of regions → New form of G

b({x}) : min(G[b({x}]) → GBP equations

GBP

Direct Ising

$$m_{P\to R}(x_R) \cong \frac{\sum_{x_{P\setminus R}} \prod_{a\in F_{P\setminus R}} \psi_a(x_{\partial a}) \prod_{(I,J)\in N(P,R)} m_{I\to J}(x_J)}{\prod_{(I,J)\in D(P,R)} m_{I\to J}(x_J)}$$

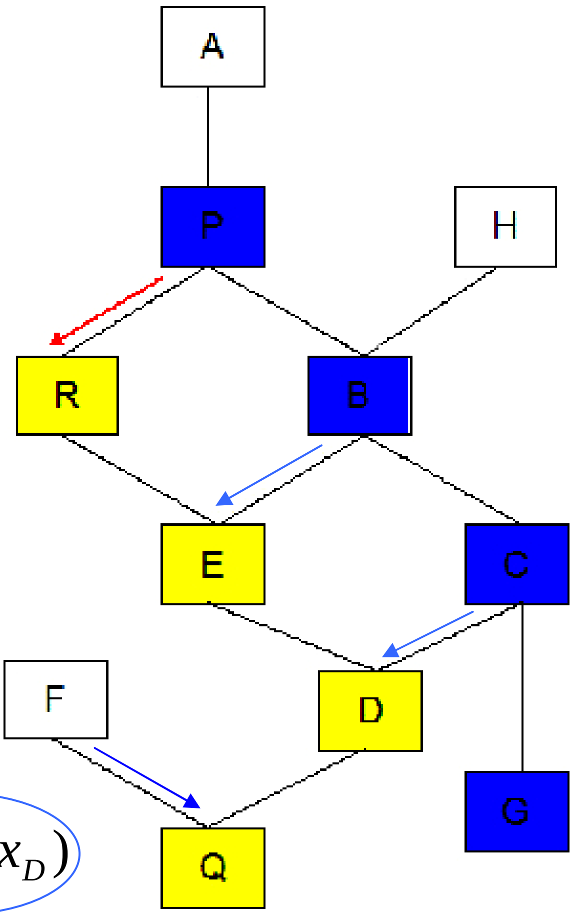


GBP

Direct Ising

$$m_{P\to R}(x_R) \cong \frac{\sum_{x_{P\setminus R}} \prod_{a\in F_{P\setminus R}} \psi_a(x_{\partial a}) \prod_{(I,J)\in N(P,R)} m_{I\to J}(x_J)}{\prod_{(I,J)\in D(P,R)} m_{I\to J}(x_J)}$$

$$b_R(x_R) \cong \prod_{a\in A_R} \Psi_a(x_{\partial a}) \prod_{P\in P(R)} m_{P\to R}(x_R) \prod_{D\in D(R)} \prod_{P'\in P(D)\setminus E(R)} m_{P'\to D}(x_D)$$



GBP

[J.S.Yedida, W.T. Freeman, Y.Weiss (2001)]

Direct Ising

Is G valid? — depends on → regions

c_R

↓

Condition:

$$\sum_R c_R I_R(a) = \sum_R c_R I_R(i) = 1$$

$(I_R(x) = 1$ if $x \in R)$

(We want to count every node only once)

INVERSE ISING

Solving the inverse problem through an iterative method

Complete Graph
+
(m_i, χ_{ij})

Iterative process

Update rules for h_i, J_{ij}

Graph
with:
$(<\sigma_i>, <\sigma_i \sigma_j>) = (m_i, \chi_{ij})$

BP/GBP

Inverse Ising

Self-consistent equations in the messages $\longrightarrow$ Self consistent equations in the inputs

$$m_i = b_i(+) - b_i(-)$$

$$\chi_{ij} = b(+,+) + b(-,-) - b(+,-) - b(-,+)$$

$$m_i = f(M, h_i, J_{ij})$$

$$\chi_{ij} = g(M, h_i, J_{ij})$$

Messages fixed $\longrightarrow$

$$h_i = f^{-1}(M, m_i, \chi_{ij})$$

$$J_{ij} = g^{-1}(M, m_i, \chi_{ij})$$

BP/GBP

Inverse Ising

Obtaining experimental values of m_i, χ_{ij}

Initialize a complete graph with random h_i, J_{ij}

For t = 1,T

Iterations Bp/Gbp $\longrightarrow$ save messages

Function nodes update in the graph

$$h_i = f^{-1}(M, m_i, \chi_{ij})$$

$$J_{ij} = g^{-1}(M, m_i, \chi_{ij})$$

End

BP

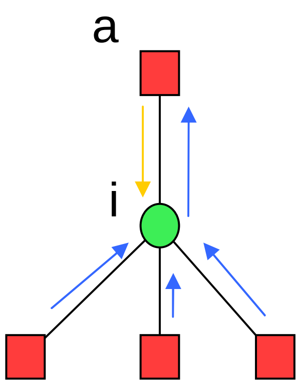
Inverse Ising

EXAMPLE

$$m_i = b_i(+) - b_i(-) \cong \nu_{a\to i}(+)\nu_{i\to a}(+) - \nu_{a\to i}(-)\nu_{i\to a}(-) = \frac{e^{h_i + \tilde{h}_i} - e^{-h_i - \tilde{h}_i}}{e^{h_i + \tilde{h}_i} + e^{-h_i - \tilde{h}_i}}$$

$$= Tanh(h_i + \tilde{h}_i)$$

$$\Rightarrow \quad h_i = ATanh(m_i) - \tilde{h}_i$$



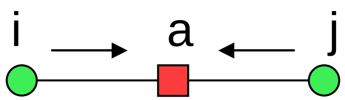
BP

Inverse Ising

EXAMPLE

$$\chi_{ij} = b_{ij}(+,+) + b_{ij}(-,-) - b_{ij}(+,-) - b_{ij}(-,+)$$

$$b_{ij}(x_i, x_j) \cong e^{J_{ij} x_i x_j} \nu_{i\to a}(x_i) \nu_{j \to a}(x_j)$$



$$\begin{cases} \nu_{i\to a}(x_i) = \dfrac{1 + x_i m_{\rightarrow}}{2} \\ \nu_{j\to a}(x_i) = \dfrac{1 + x_i m_{\leftarrow}}{2} \end{cases}$$

$$\chi_{ij} = \frac{Tanh(J_{ij}) + m_{\rightarrow} m_{\leftarrow}}{1 + m_{\rightarrow} m_{\leftarrow} Tanh(J_{ij})} \quad \Longrightarrow \quad J_{ij} = ATanh \frac{\chi_{ij} - m_{\rightarrow} m_{\leftarrow}}{1 - \chi_{ij} m_{\rightarrow} m_{\leftarrow}}$$

Sus.Prop.

[M.Mezard, T.Mora (2008)]
Direct and Inverse Ising

BP limitations → χ_{ij} only if $\exists < ij >$

Fluctuation-Response theorem

$$<\sigma_i \sigma_j> - <\sigma_i><\sigma_j> = \frac{\partial m_i}{\partial h_j} = \frac{\partial b_i(+)}{\partial h_j} - \frac{\partial b_i(-)}{\partial h_j}$$

with $b_i(x_i) \cong \prod_{a \in \partial i} \eta_{a \to i}(x_i)$

⇒ We have to know derivatives of messages!

- BP equations (ν, η)
- BP derivatives $(\nu, \eta, \partial\nu, \partial\eta)$

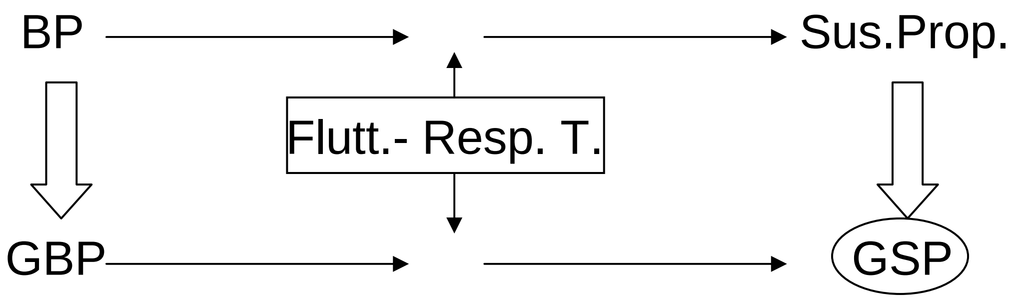
→ $\chi_{ij} = f(m, \partial m, h, J)$

Inverting f → New update rule for J

GSP

Direct Ising

Aim: improve accuracy in χ_{ij} in the GBP scheme



- Find equations for derivatives of messages in GBP → Gsp equations
- Do the derivatives of the 1-beliefs equations → Correlations

GSP

Ising Inverso

Extracting couplings from GSP with arbitrary regions?

Contraint: at least all the 1 and 2 regions

1-regions “see” Bethe approx.

Partial reuse of Sus.Prop. formalism

Small correlations expansion

[R.Monasson, V.Sessak (2008)]

Likelihood maximization

↓

Let's suppose to have n copies of the system

Let's maximize the probability to have these copies given the theory:

$$\max[\frac{1}{n} Log \prod_n P(\{\sigma\}_n \mid h, J)] = \min(S)$$

$$where \; : \; S = Log\,(Z\,(h, J)) - \sum_i h_i m_i - \sum_{i,j} J_{ij}\,(c_{ij} + m_i m_j)$$

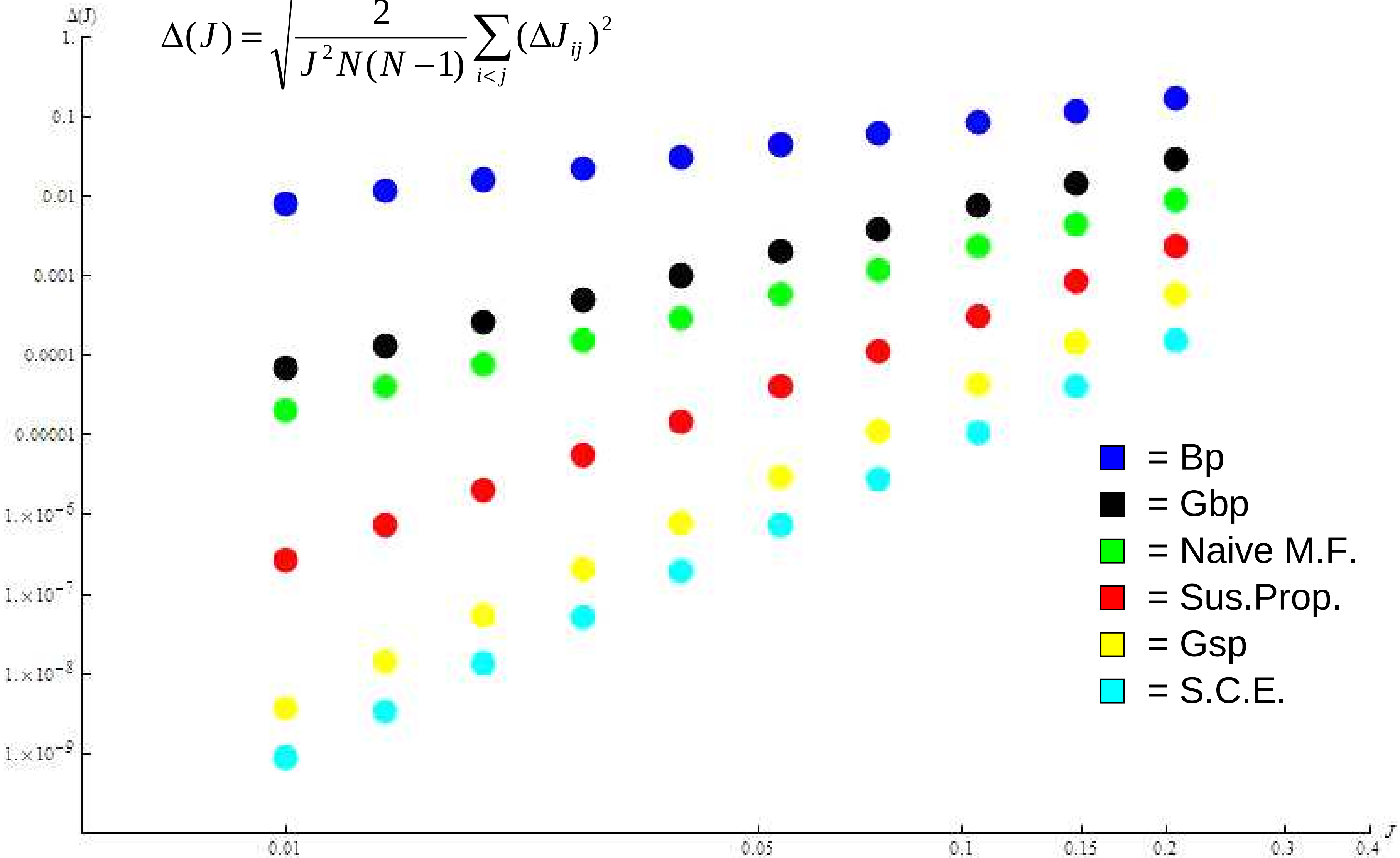
Let's solve the problem for $c_{ij} \to 0$ ⇨ $c_{ij} \to \beta\, c_{ij}$

⇨ Analitical expressions for J e h

Simulations examples

Complete Graph, $N = 20$

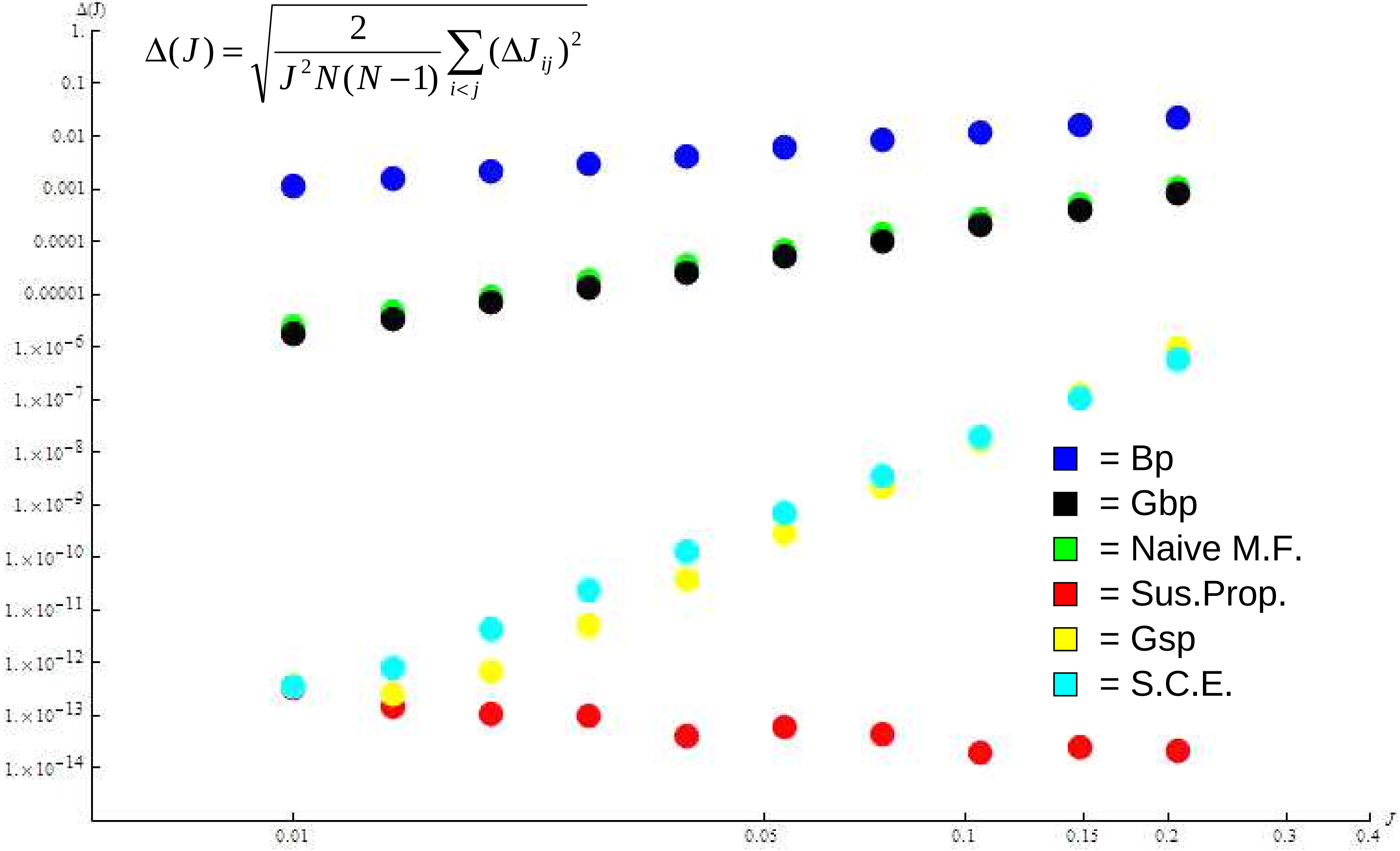
$$\Delta(J) = \sqrt{\frac{2}{J^2 N(N-1)} \sum_{i<j} (\Delta J_{ij})^2}$$



Simulations examples

$$Tree,\quad N = 20$$

$$\Delta(J) = \sqrt{\frac{2}{J^2 N(N-1)} \sum_{i<j} (\Delta J_{ij})^2}$$



Conclusions and future

BP → Sus.Prop.

↓ ↓

GBP → GSP

PROBLEMS

Big correlations → Loopy BP / Small Correlation Expansion

Glassy phase → Necessity to redefine the problem

⇒

- Iteration methods in big correlations regime
- Cavity methods for GBP in 1RSB [T.Rizzo et al. (2009)]
- New approximations for the free energy

Acronyms & Notation

BP	= Belief Propagation
GBP	= Generalized Belief Propagation
Sus.Prop.	= Susceptibility Propagation
GSP	= Generalized Susceptibility Propagation
S.C.E.	= Small Correlation Expansion
Naive M.F.	= Mean Field Theory + Fluctuation Response Theorem
GBP (3)	= GBP with all regions with 1 or 2 or 3 spins
GSP (3)	= GSP with all regions with 1 or 2 or 3 spins
$\cong$	= Identity true except for a normalization factor
∂i	= Set of nodes which are connected to *i* with a link
$\partial i \setminus a$	= Set ∂i without the node *a*
	= node representing a variable (labeled by *i*,*j*,*k*,…)
	= node representing a function (labeled by *a*,*b*,*c*,…)