Adding loops to mean field approximation for disordered models

Federico Ricci-Tersenghi
Sapienza University of Rome

- Replica cluster variational method,
  T. Rizzo, A. Lage-Castellanos, R. Mulet and F. Ricci-Tersenghi,
  J. Stat. Phys. 139, 375 (2010).

- Inference algorithm for finite-dimensional spin glasses: Belief Propagation on the dual lattice,
  A. Lage-Castellanos, R. Mulet, F. Ricci-Tersenghi and T. Rizzo,
  Phys. Rev. E 84, 046706 (2011).

- Characterizing and Improving Generalized Belief Propagation Algorithms on the 2D Edwards-Anderson Model,
  E. Dominguez, A. Lage-Castellanos, R. Mulet, F. Ricci-Tersenghi and T. Rizzo,
  J. Stat. Mech. P12007 (2011).

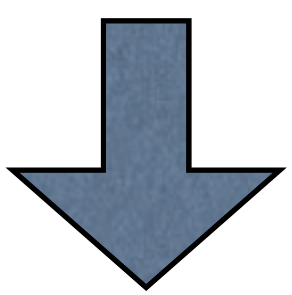
- Replica cluster variational method: the replica symmetric solution for the 2d random bond Ising model,
  A. Lage-Castellanos, R. Mulet, F. Ricci-Tersenghi and T. Rizzo,
  arxiv:1204.0439

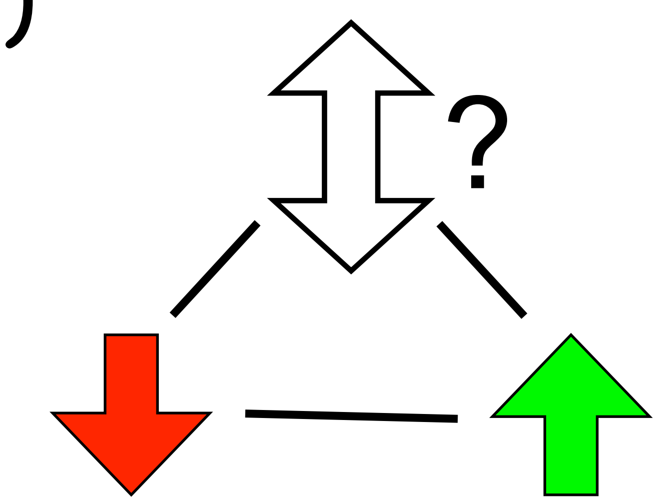
Models of interest

- SPIN GLASSES on D-dimensional lattices

$$H = -\sum_{(i,j)} J_{ij}\sigma_i\sigma_j$$

disorder (random couplings)
frustration





COMPLEXITY

What we want to know

- Physical properties:
  (free-)energy and correlations
  -> nature of low T phase

- Single sample:
  site specific
  marginals
  $$\begin{cases} p_i(\sigma_i) = \sum_{\vec{\sigma}\setminus\sigma_i} e^{-\beta H}/Z \\ p_{ij}(\sigma_i,\sigma_j) = \sum_{\vec{\sigma}\setminus\sigma_i,\sigma_j} e^{-\beta H}/Z \end{cases}$$

- Average case: global quantities

$$f = \langle \ln Z \rangle_J \quad m = \frac{1}{N}\sum_i m_i \quad q = \frac{1}{N}\sum_i m_i^{(1)} m_i^{(2)}$$

Mean field approximations

Variational approach

$$F = U - TS = \sum_{\vec{\sigma}} H(\vec{\sigma}) P(\vec{\sigma}) + T \sum_{\vec{\sigma}} P(\vec{\sigma}) \ln P(\vec{\sigma})$$

Short description, few parameters:

magnetizations

(MF)

$$P_{MF}(\vec{\sigma}) = \prod_i b_i(\sigma_i) = \prod_i \frac{1 + m_i \sigma_i}{2}$$

and correlations

(Bethe)

$$P_{Bethe}(\vec{\sigma}) = \prod_{ij} \frac{b_{ij}(\sigma_i, \sigma_j)}{b_i(\sigma_i) b_j(\sigma_j)} \prod_i b_i(\sigma_i)$$

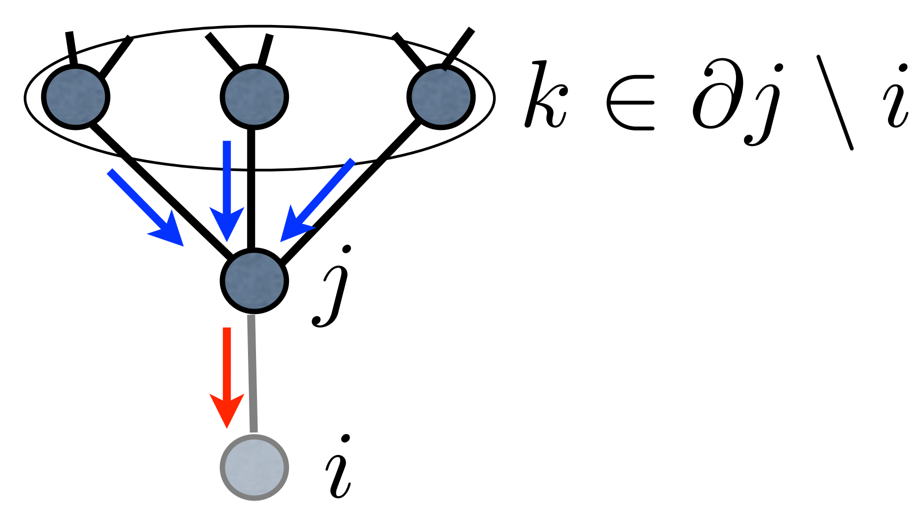
Bethe approximation

- Exact on trees
- Marginalization conditions $\sum_{\sigma_j} b_{ij}(\sigma_i, \sigma_j) = b_i(\sigma_i)$

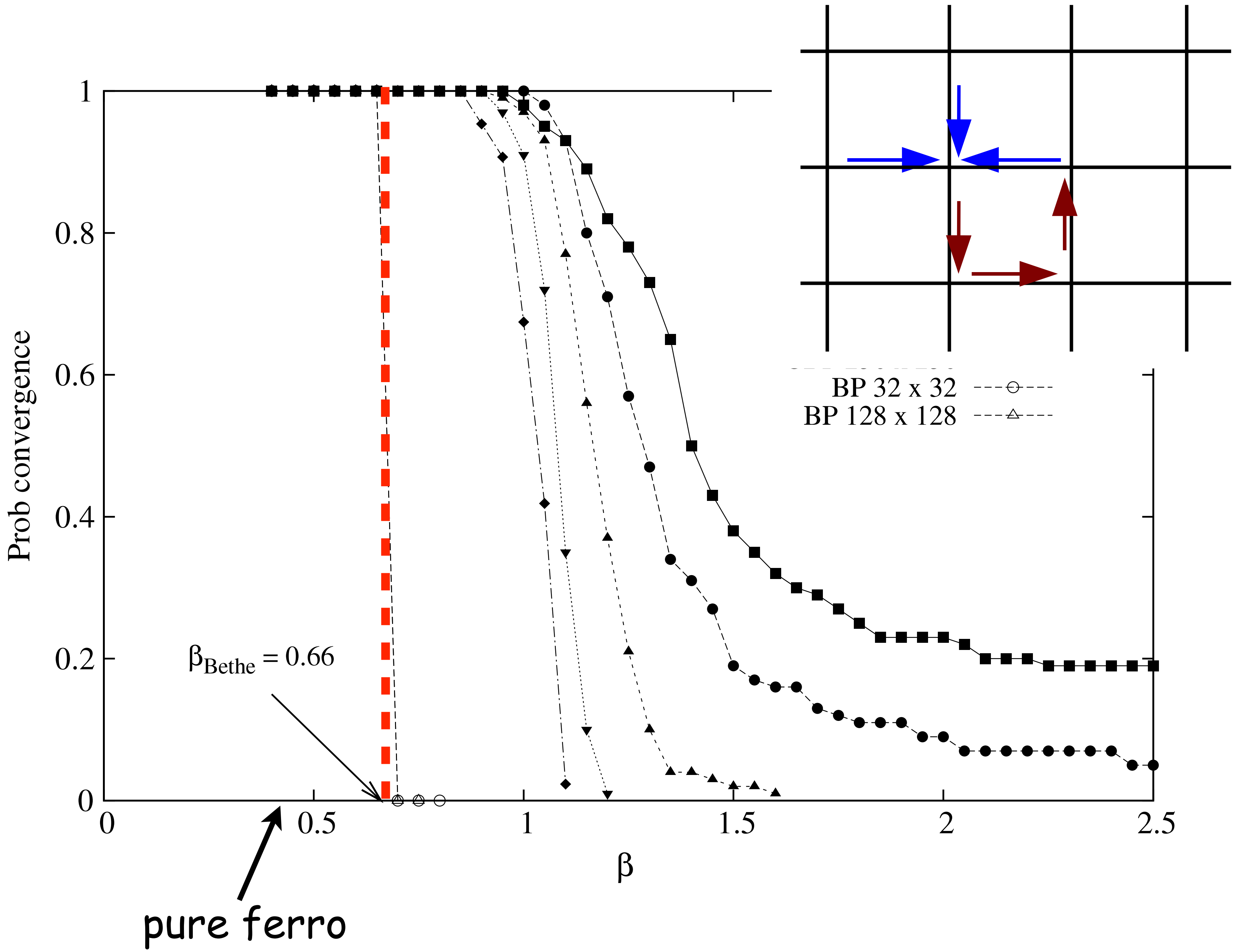
$$\tanh(\beta u_{j\to i}) = \tanh(\beta J_{ij}) \tanh(\beta \sum_{k\in\partial j\setminus i} u_{k\to j})$$

$$m_i = \tanh(\beta \sum_{j\in\partial i} u_{j\to i})$$

- Belief Propagation (BP)



BP for SG on a 2D lattice



Improving BP (loop corrections)

- Montanari, Rizzo, JSTAT 2005
- Parisi, Slanina, JSTAT 2006
- Chertkov, Chernyak, JSTAT 2006 (loop calculus)

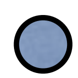
all based on existence of BP fixed point !

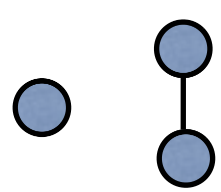
Cluster Variation Method

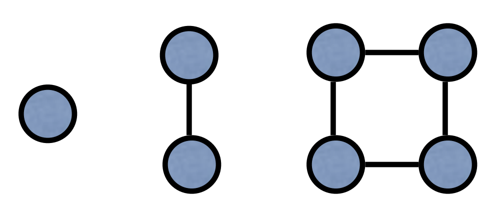
Kikuchi 1951

$$F = \sum_{\vec{\sigma}} H(\vec{\sigma}) P(\vec{\sigma}) + T \sum_{\vec{\sigma}} P(\vec{\sigma}) \ln P(\vec{\sigma})$$

energy (easy) entropy (hard)

- Mean field  $P_{MF}(\vec{\sigma}) = \prod_i b_i(\sigma_i) = \prod_i \frac{1 + m_i \sigma_i}{2}$

- Bethe  $P_{Bethe}(\vec{\sigma}) = \prod_{ij} \frac{b_{ij}(\sigma_i, \sigma_j)}{b_i(\sigma_i) b_j(\sigma_j)} \prod_i b_i(\sigma_i)$

- Plaquette CVM  $\{b_i, b_{ij}, b_{ijkl}\}$

Region graph approximation

Yedidia Freeman Weiss, IEEE-IT 2005

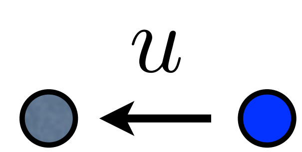
- Disorder -> heterogeneities
- Choose an arbitrary set of regions (containing all interactions)

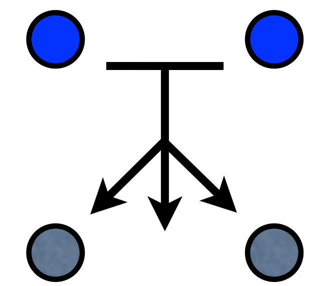
$$F_{RGA} = \sum_{r \in R} c_r \left(\sum_{x_r} b_r E_r + \sum_{x_r} b_r \ln b_r \right)$$

- Find an extremum by Generalized Belief Propagation (GBP)

Square plaquette GBP

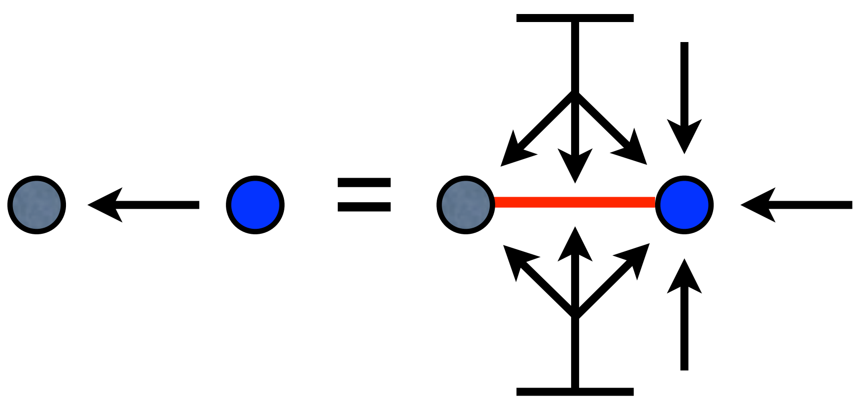
- 2 kind of messages (marginalizations)

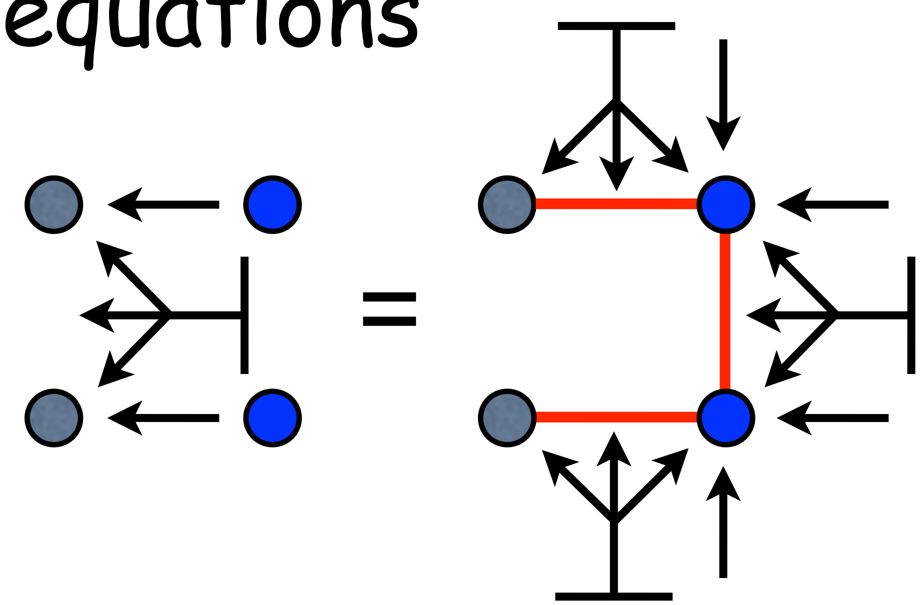




(U, u_1, u_2)

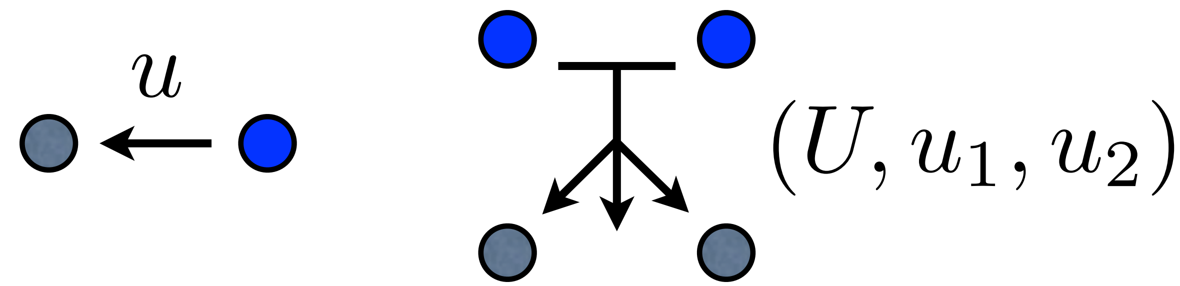
- 2 kind of self-consistency equations



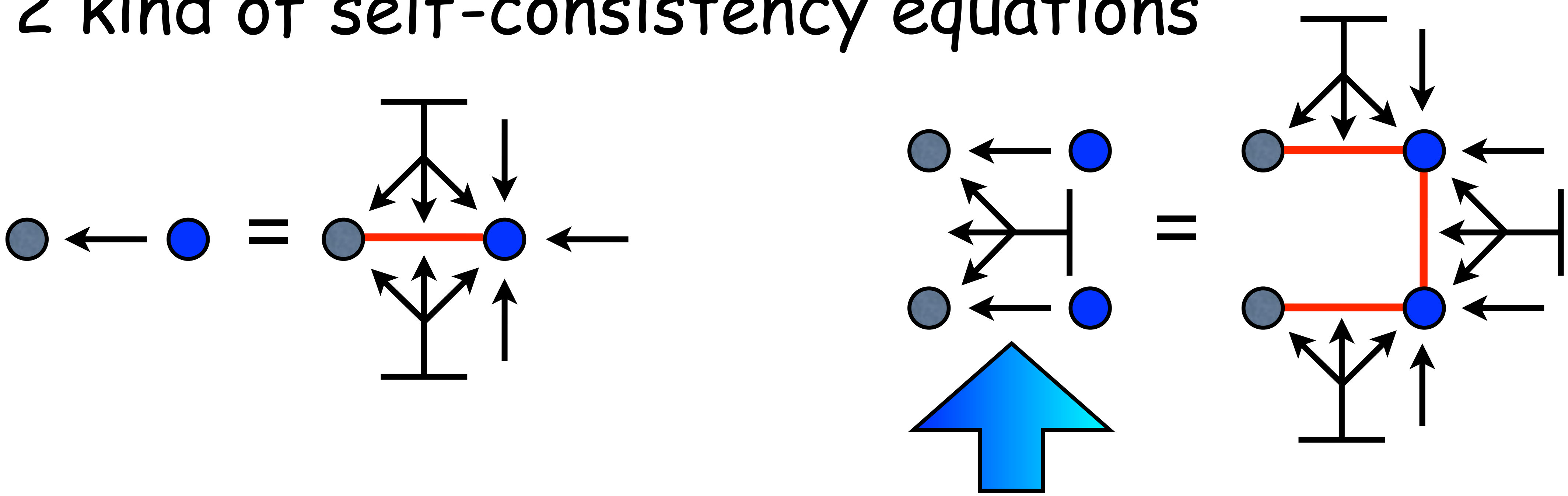


Square plaquette GBP

- 2 kind of messages (marginalizations)

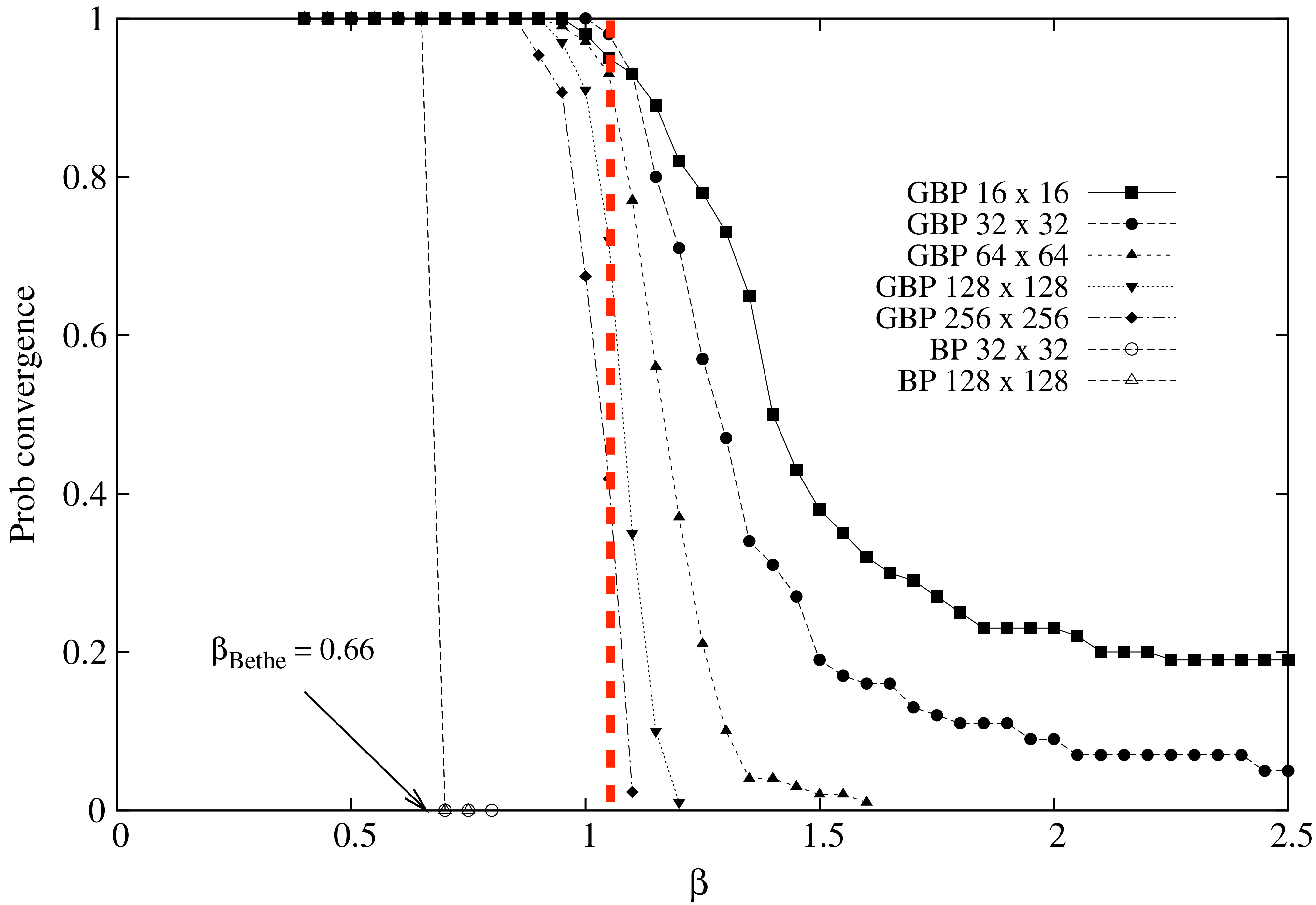


- 2 kind of self-consistency equations



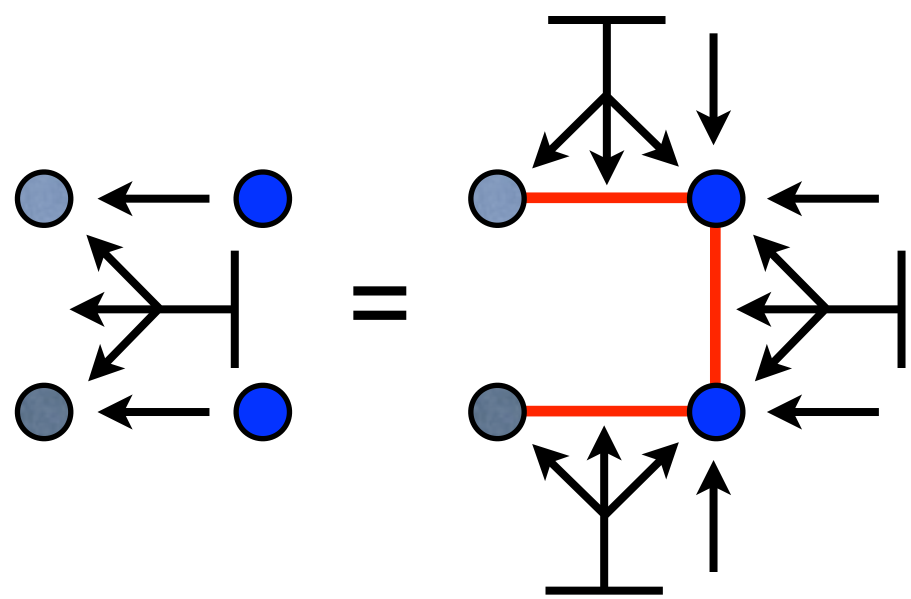
Single and triple messages appear together!

GBP for a SG on a 2D lattice



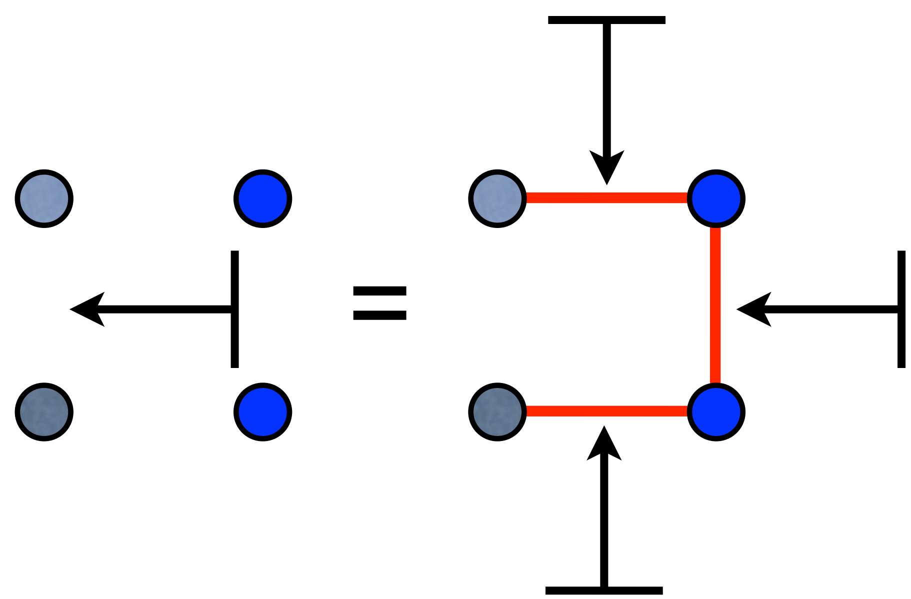
Improving GBP convergence

Exploit symmetries: $m_i = 0 \implies u = 0$



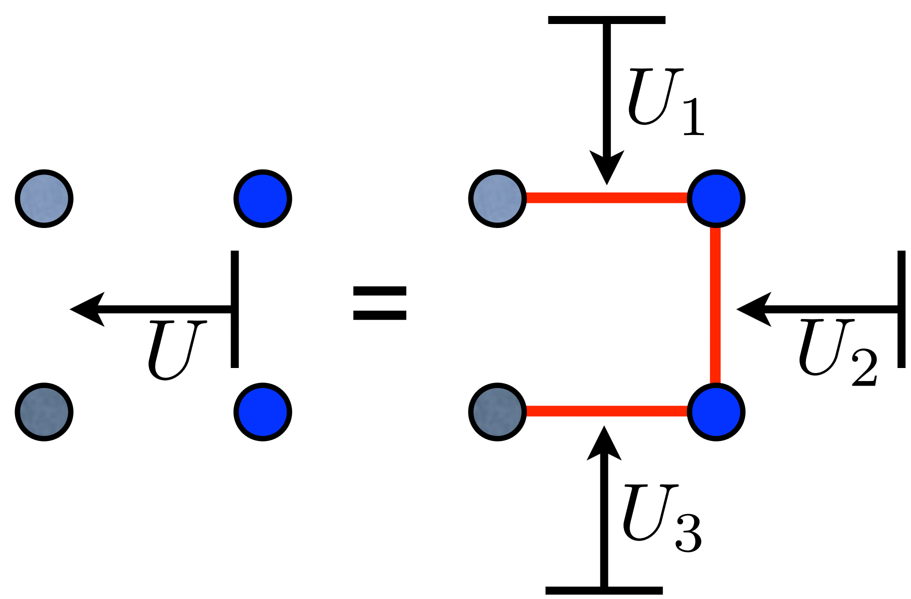
Improving GBP convergence

Exploit symmetries: $m_i = 0 \implies u = 0$



Improving GBP convergence

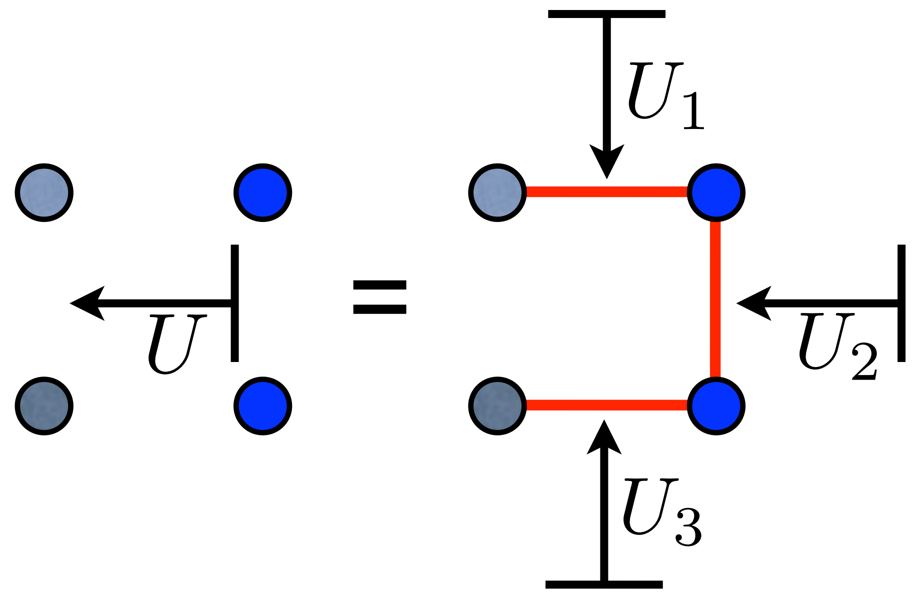
Exploit symmetries: $m_i = 0 \implies u = 0$



$$\tanh(\beta U) = \tanh(\beta(J_1 + U_1))\tanh(\beta(J_2 + U_2))\tanh(\beta(J_3 + U_3))$$

Improving GBP convergence

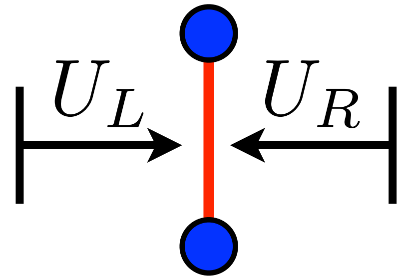
Exploit symmetries: $m_i = 0 \implies u = 0$



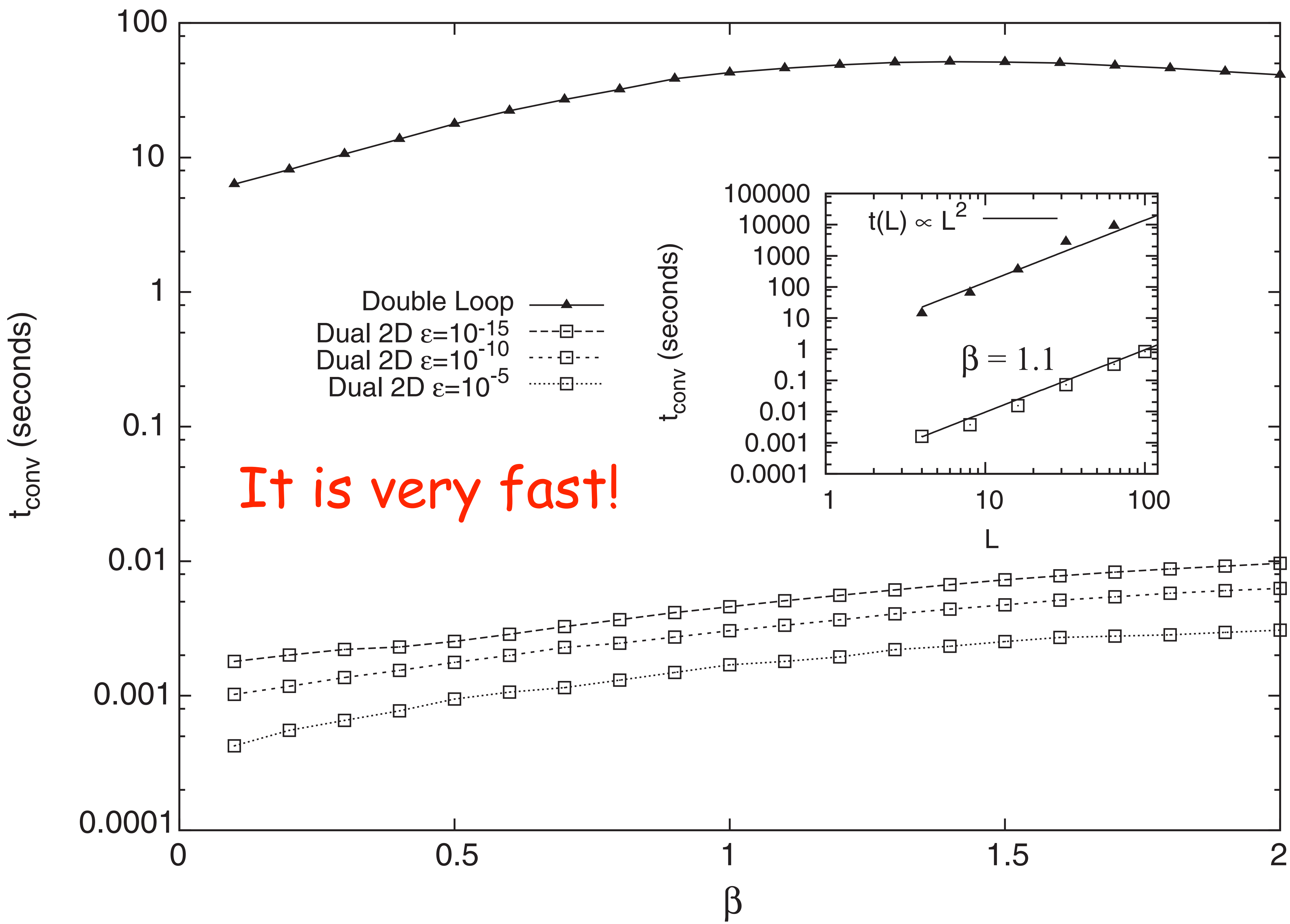
Dual algorithm

$$\tanh(\beta U) = \tanh(\beta(J_1 + U_1))\tanh(\beta(J_2 + U_2))\tanh(\beta(J_3 + U_3))$$

$$c_{ij} = \tanh(\beta(J_{ij} + U_L + U_R))$$



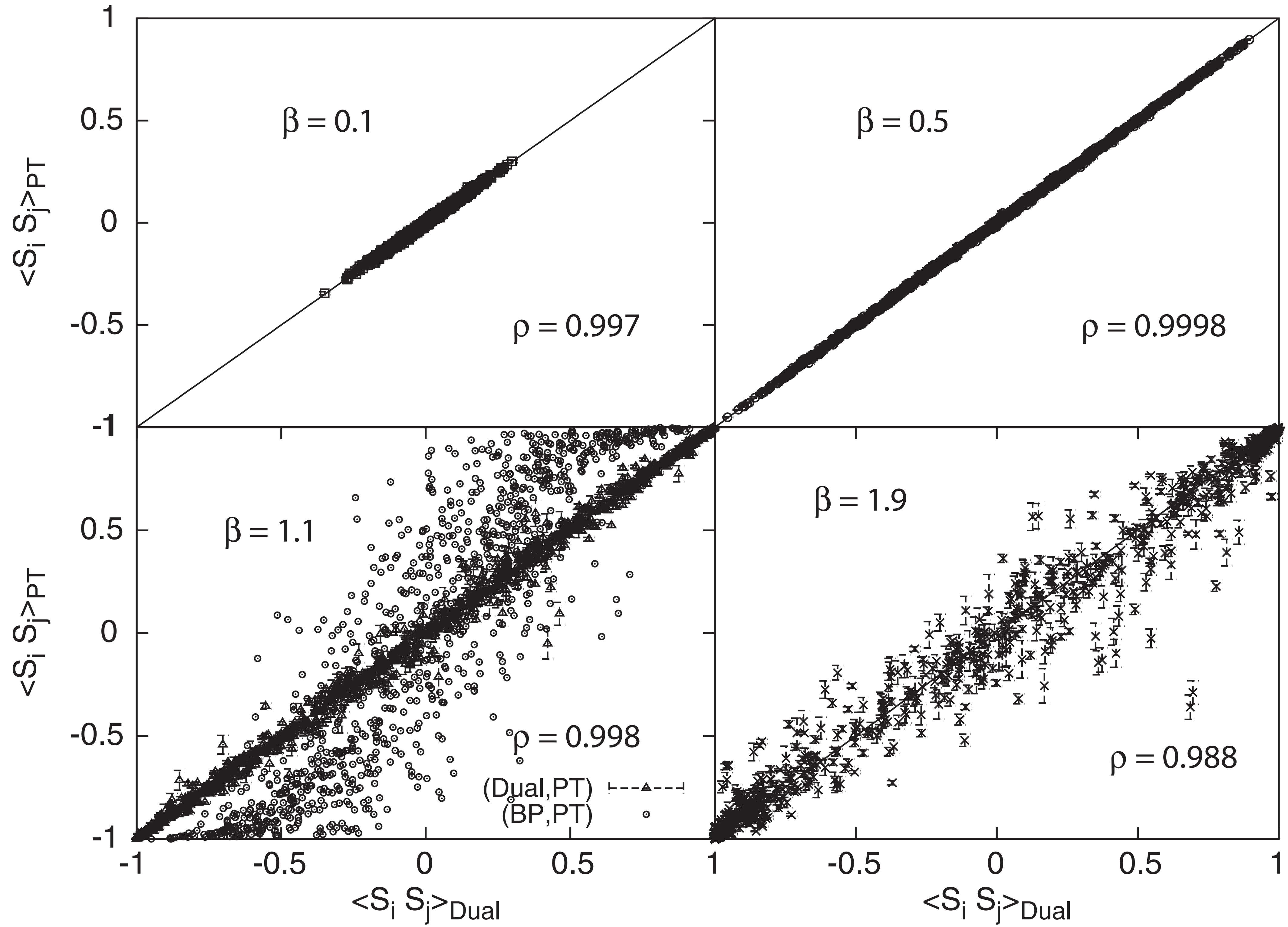
Dual algorithm



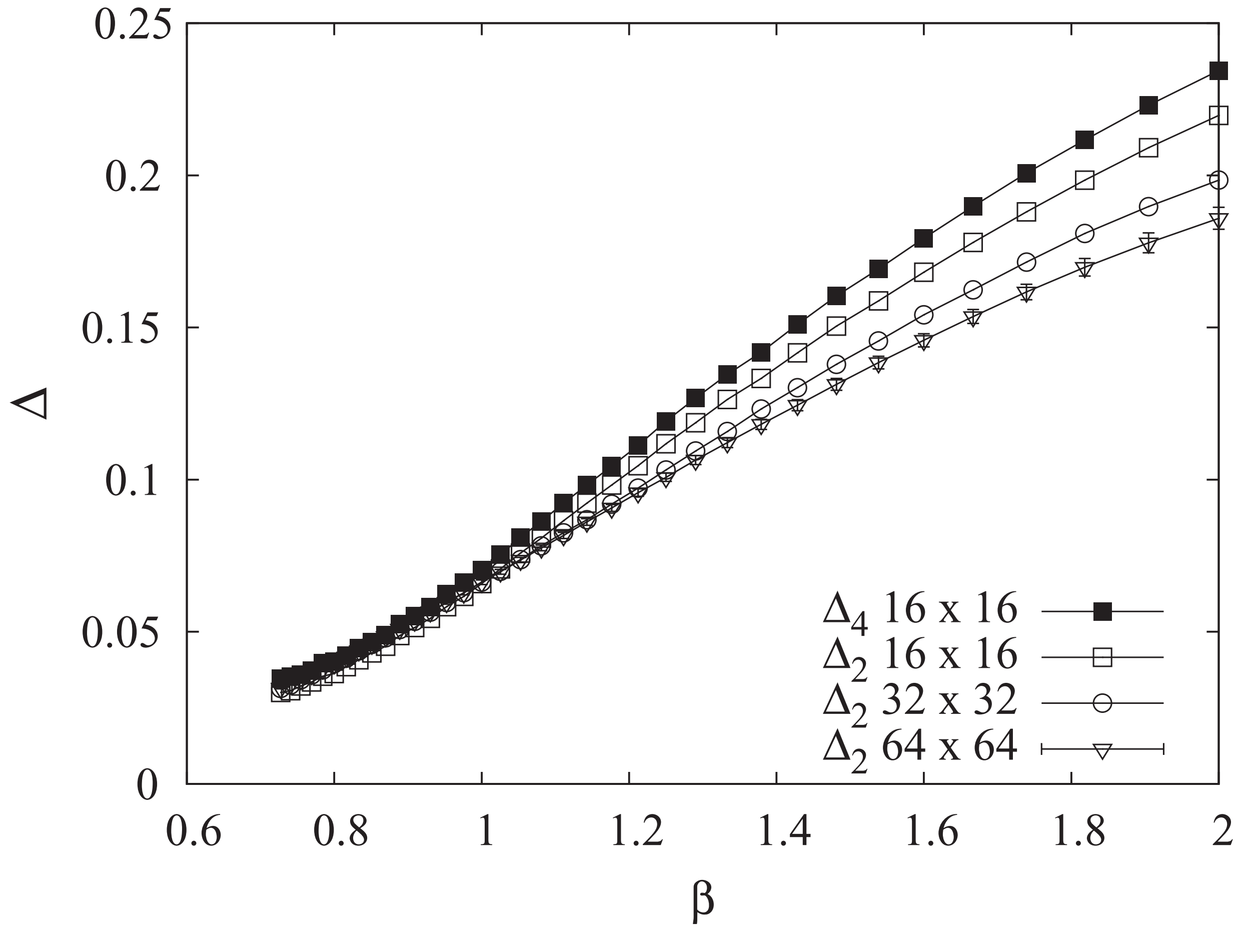
It is very fast!

Dual algorithm

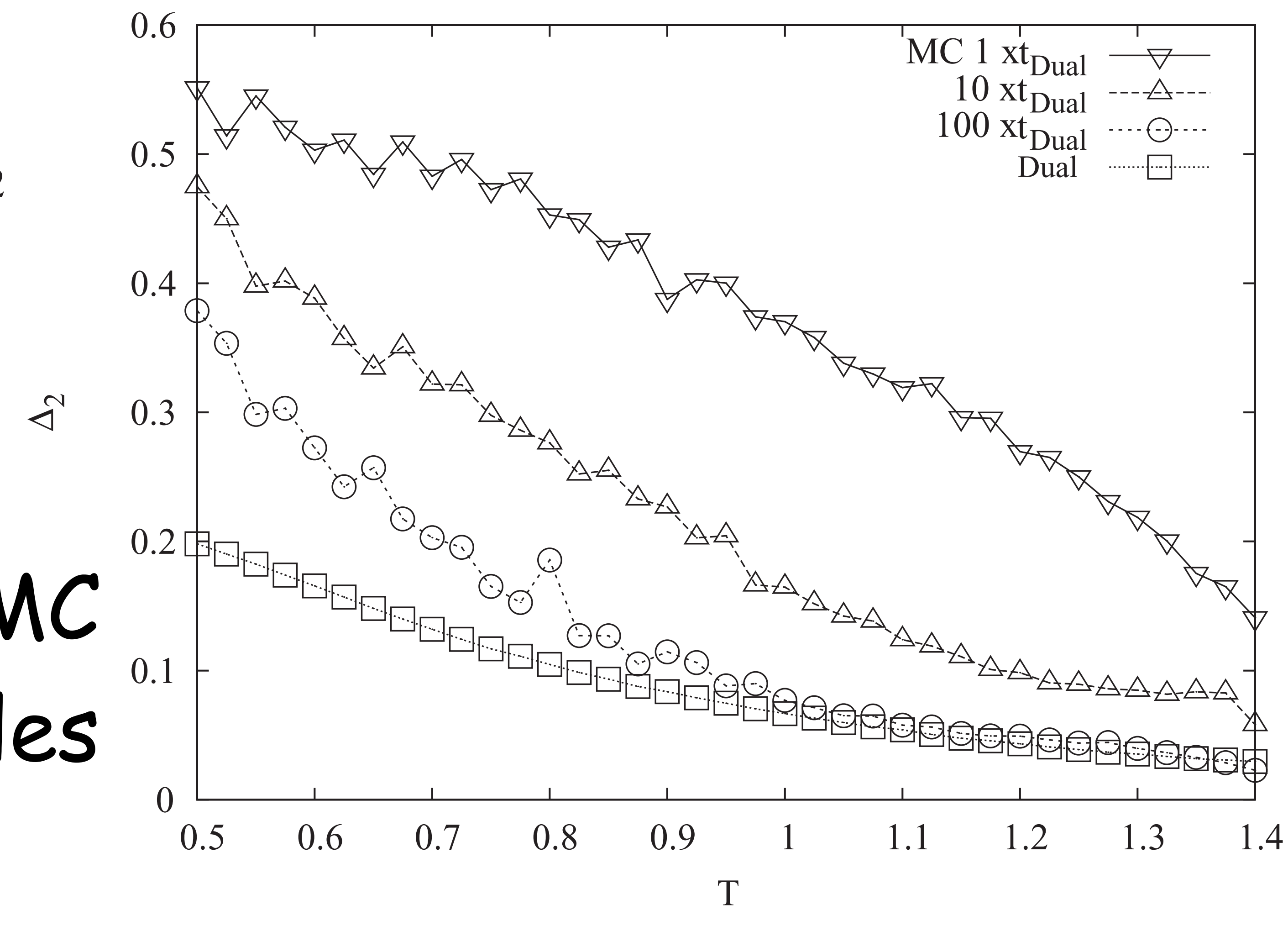
Rather accurate results



Dual algorithm



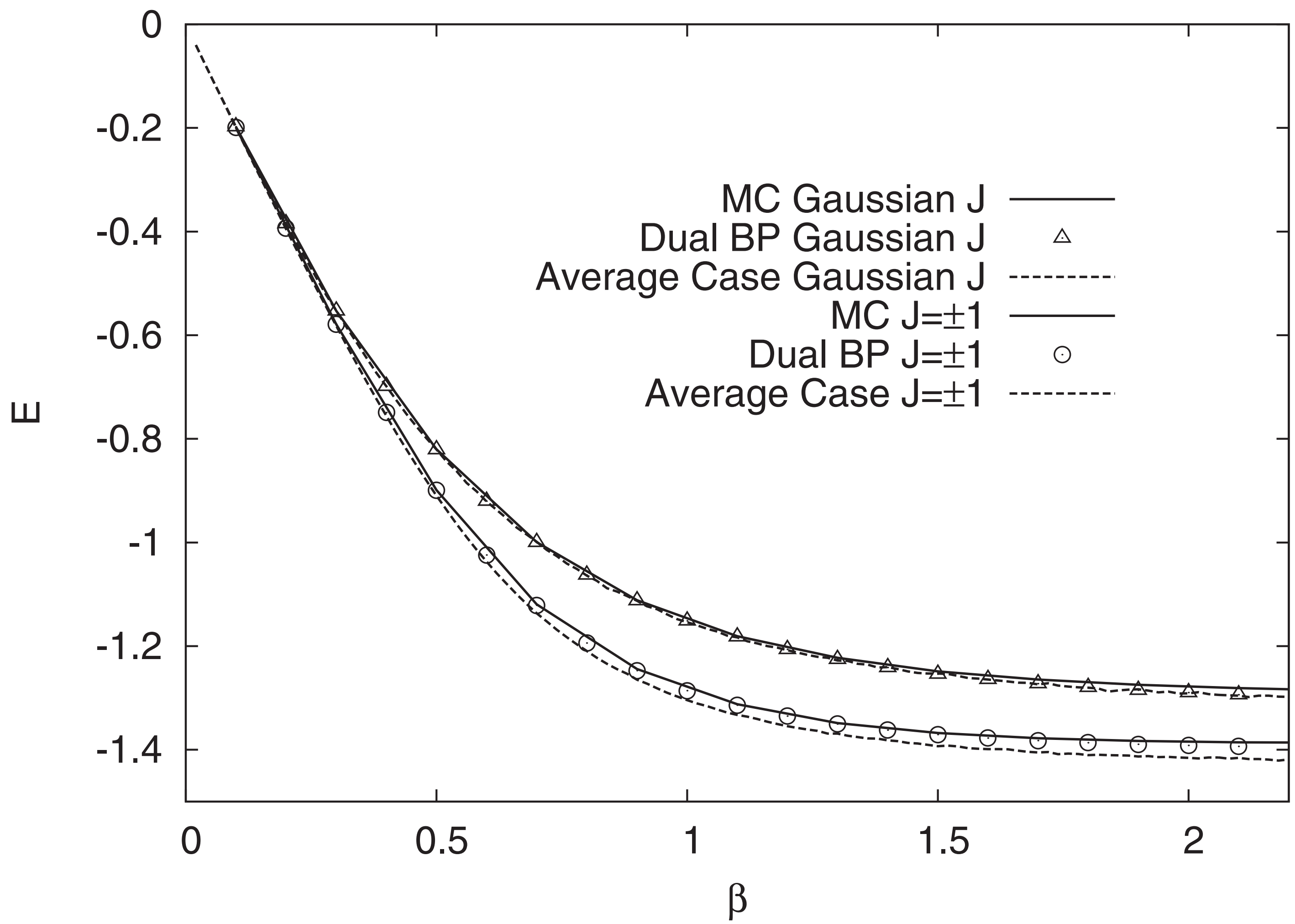
Errors improve
for larger sizes



Better than MC
on short time scales

Dual algorithm

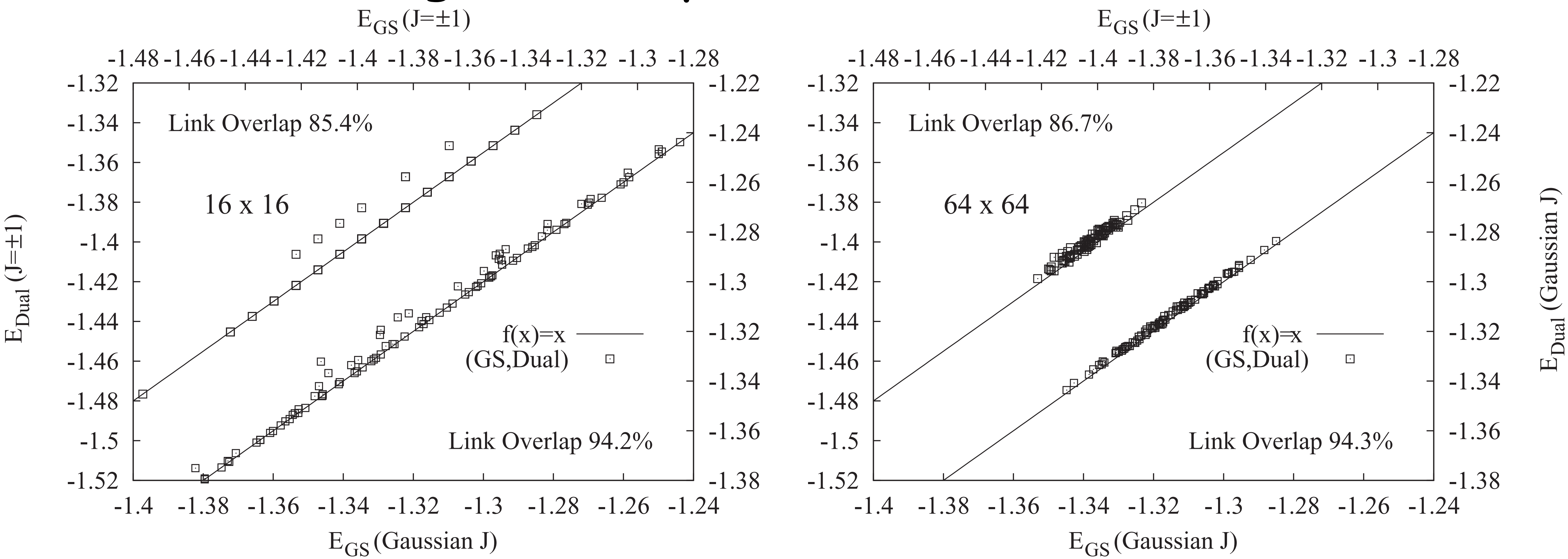
Not exact (even on the long run)
but very fast and good enough for many applications



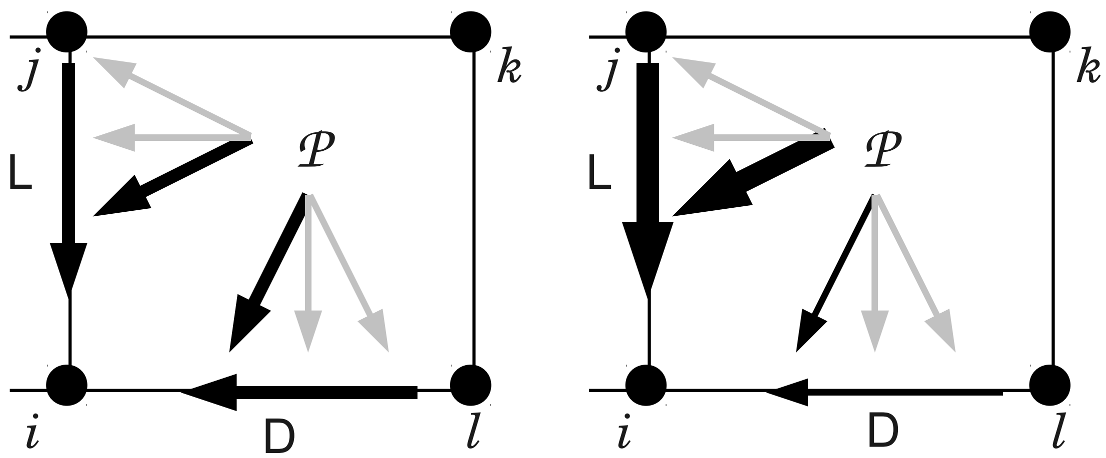
Dual algorithm

Not exact (even on the long run)
but very fast and good enough for many applications

Computing ground states by fixing variables
according to dual-predicted correlations

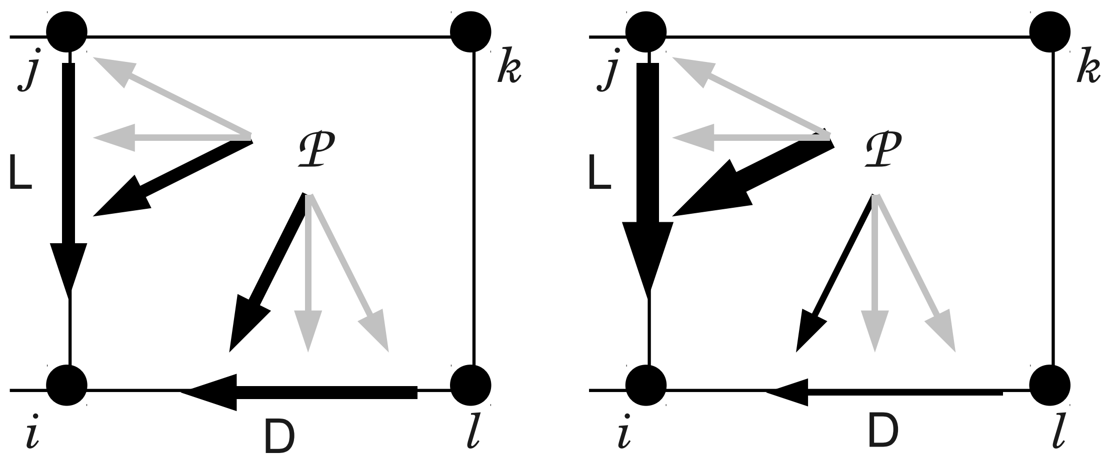


Improving GBP convergence



Fixing a gauge symmetry

Improving GBP convergence

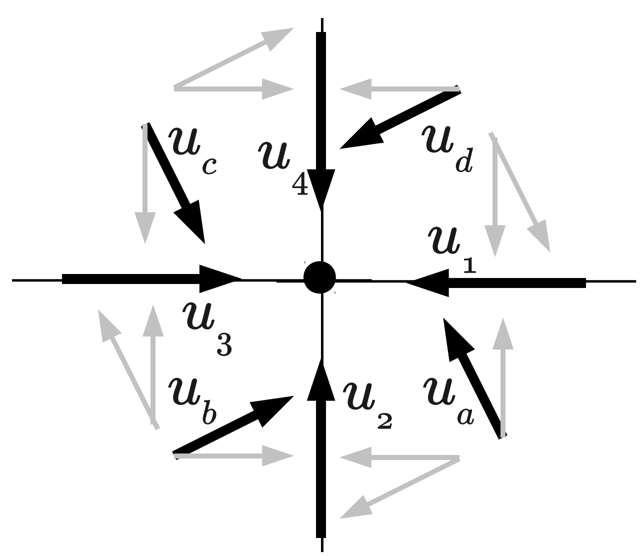


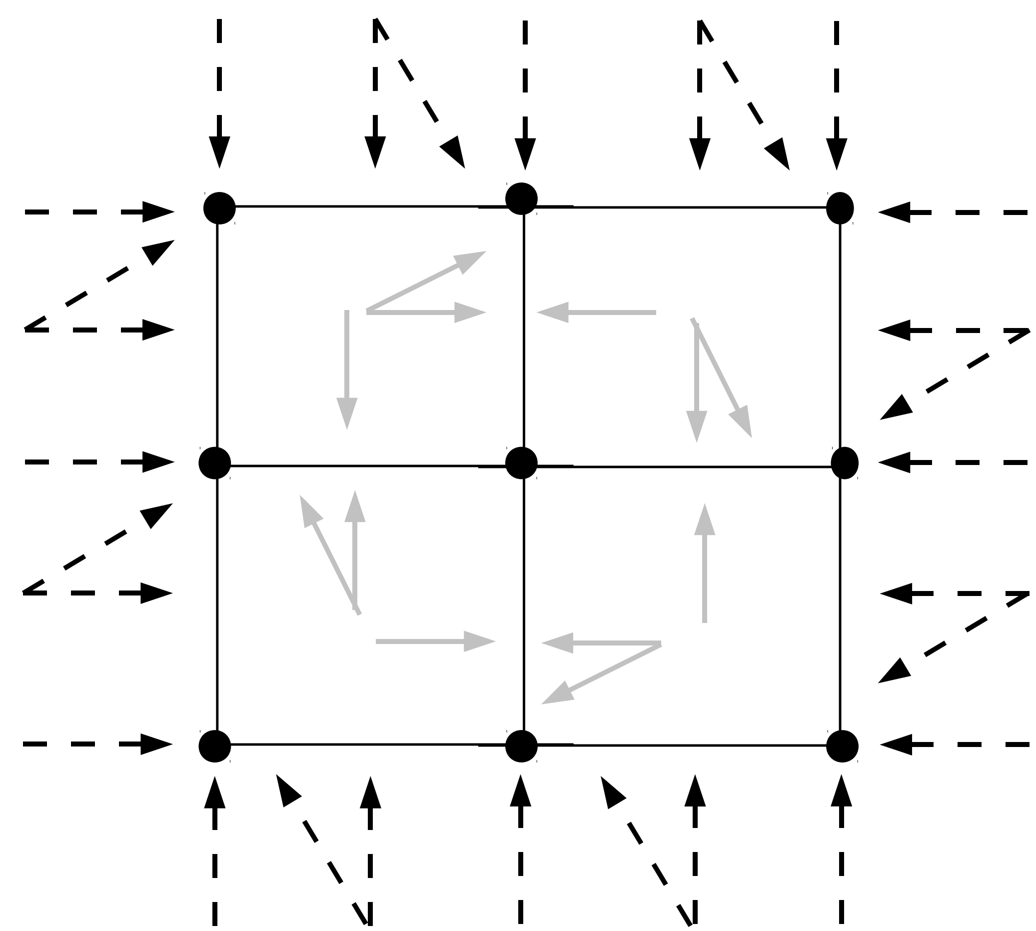
Fixing a gauge symmetry

A better updating scheme

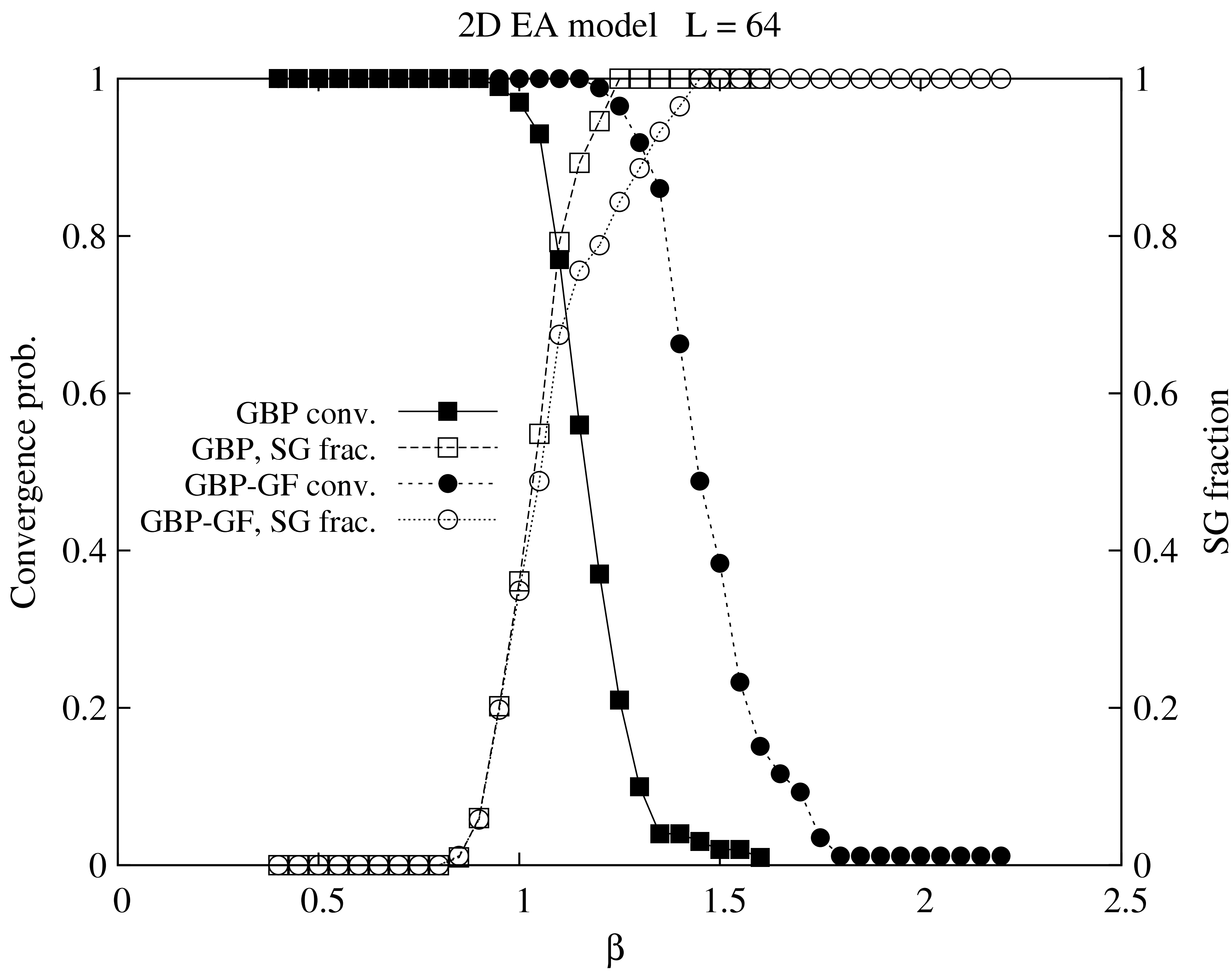
Gauge-fixing

GF-GBP

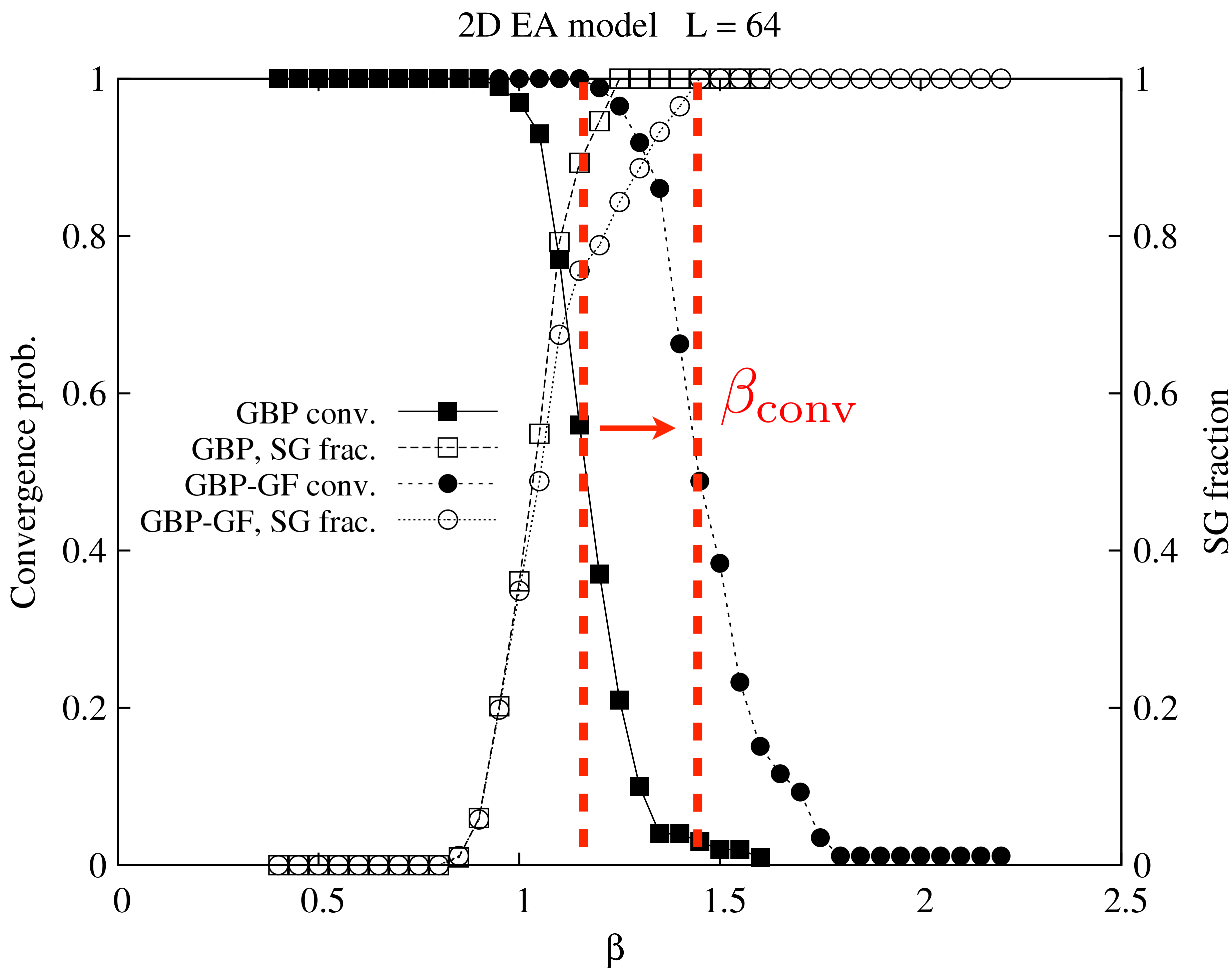




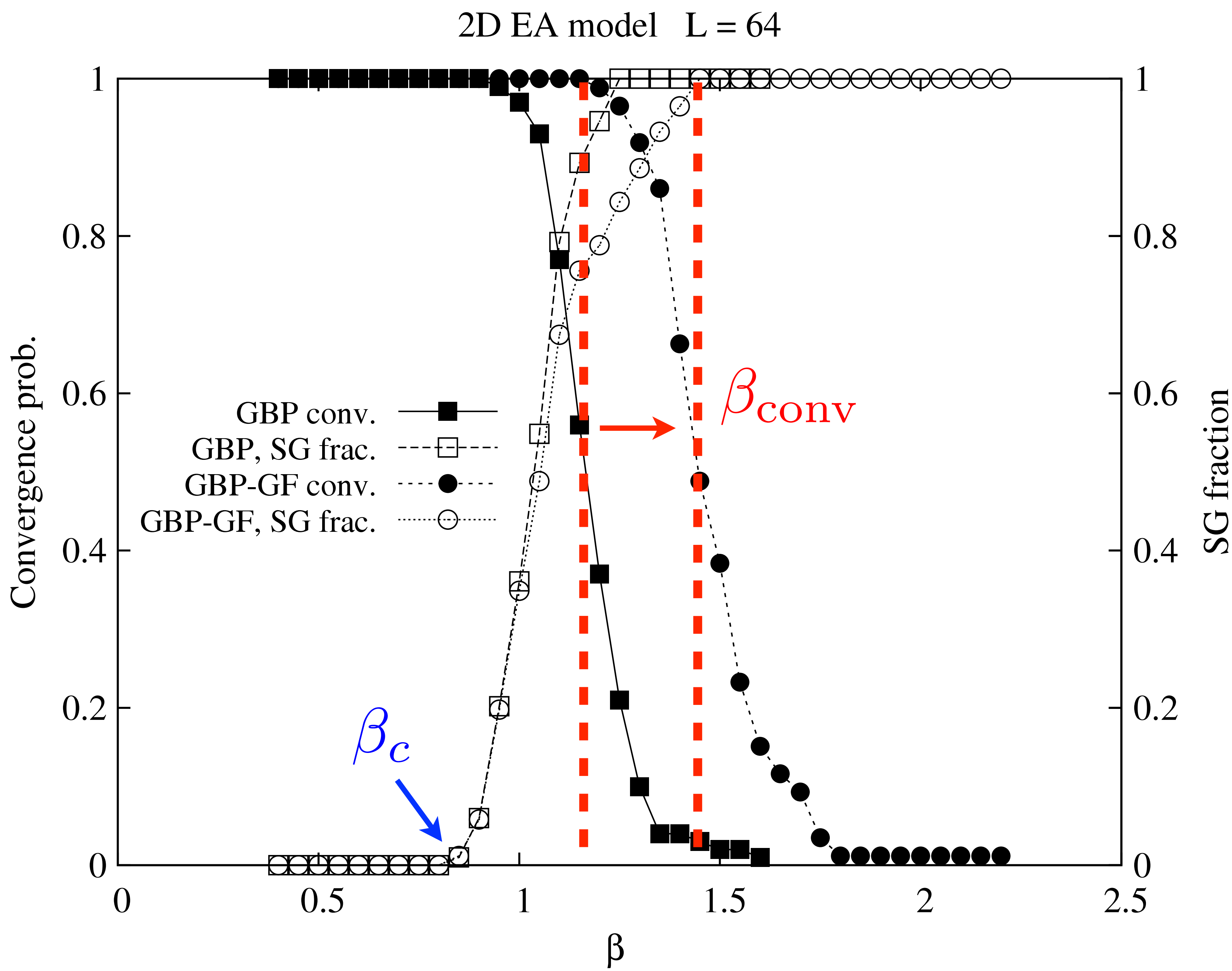
GF-GBP does it better



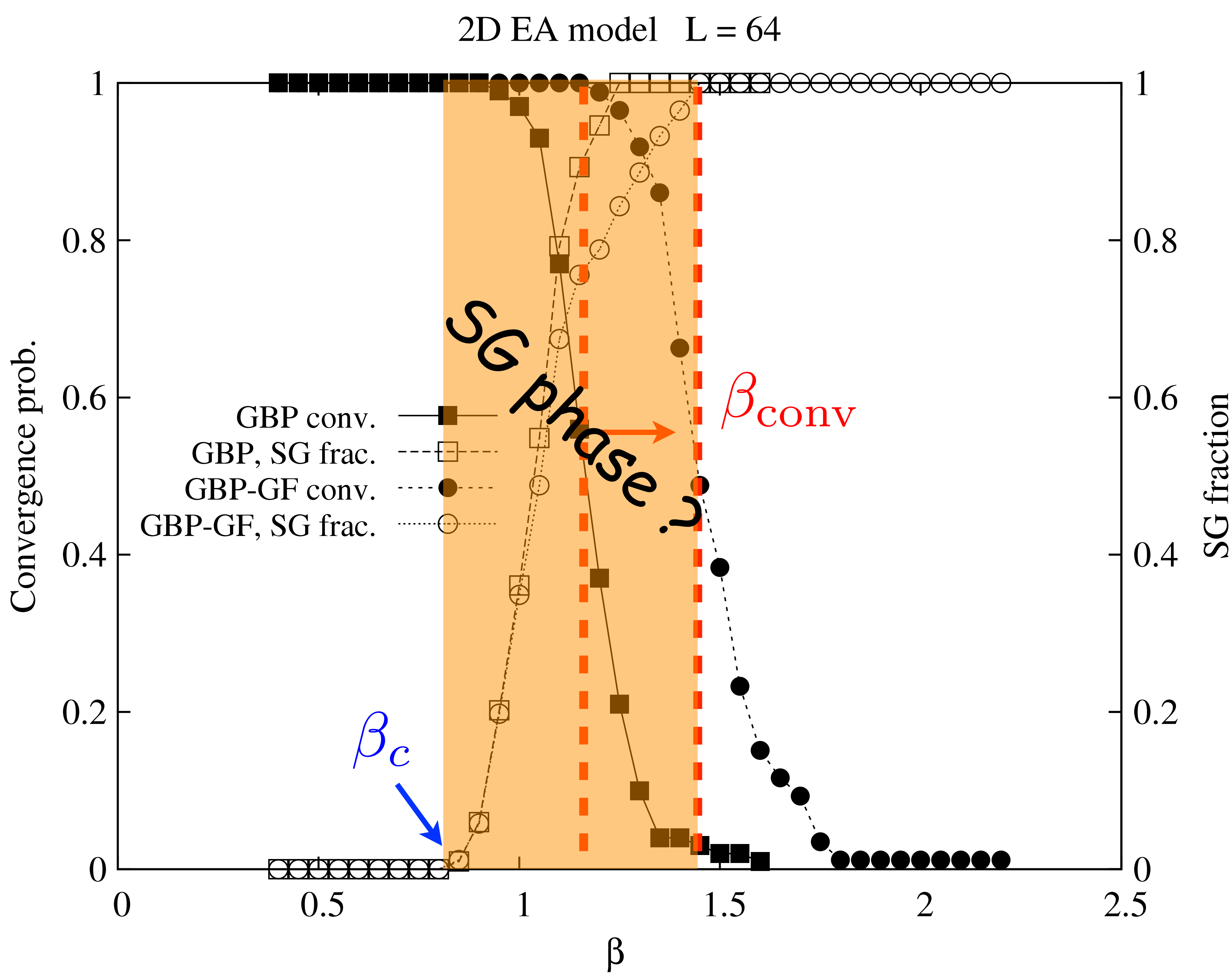
GF-GBP does it better



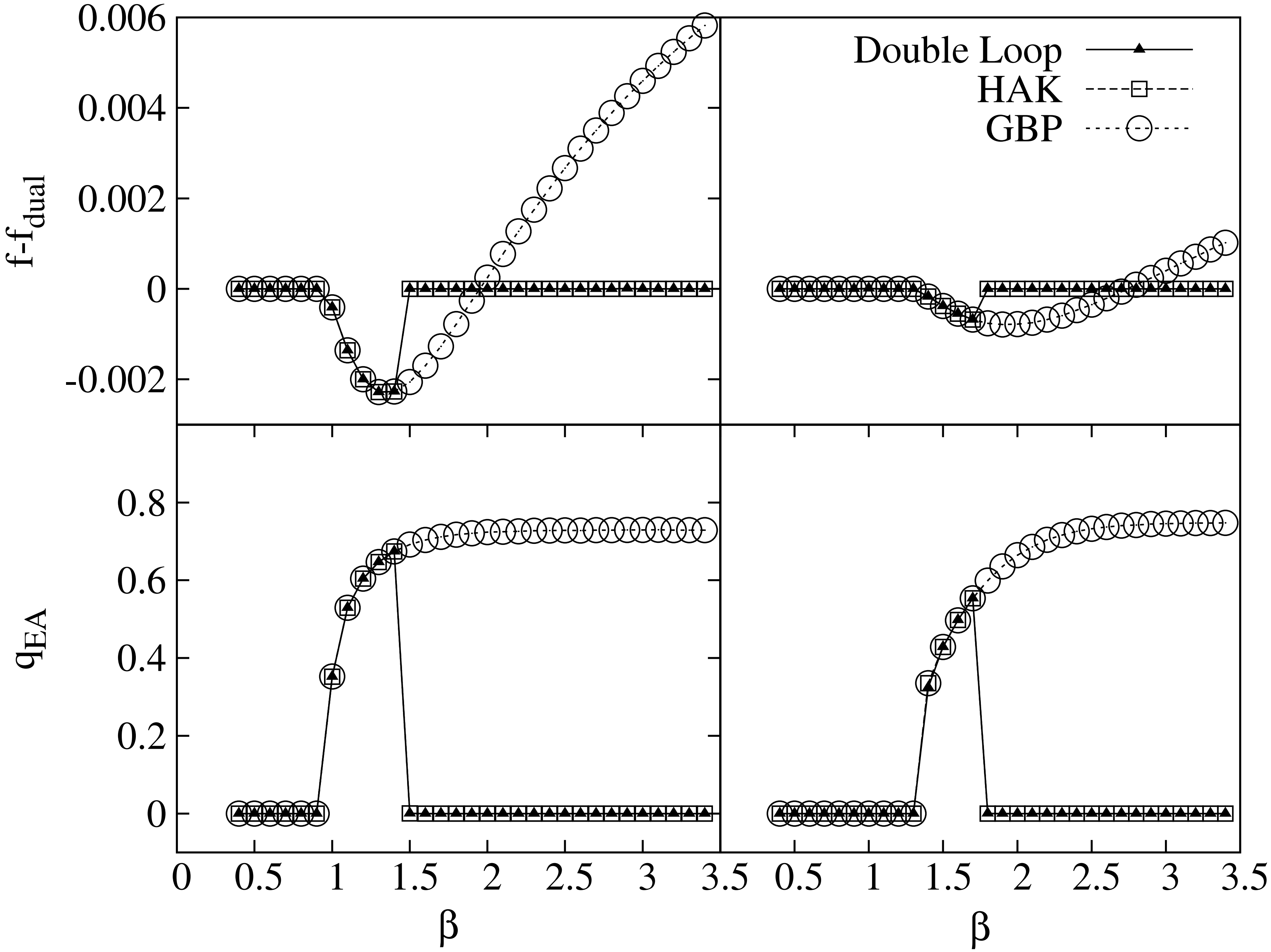
GF-GBP does it better



GF-GBP does it better

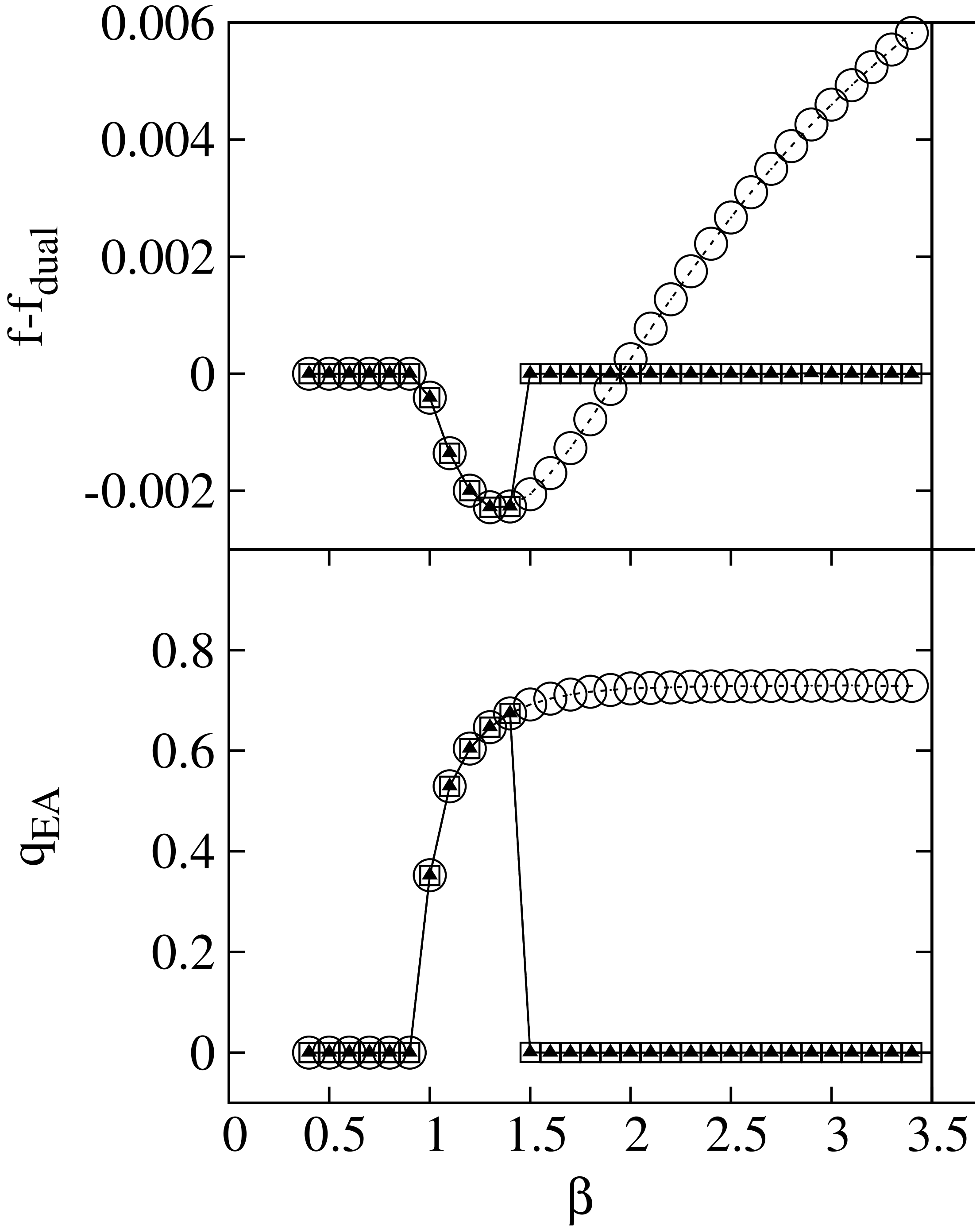


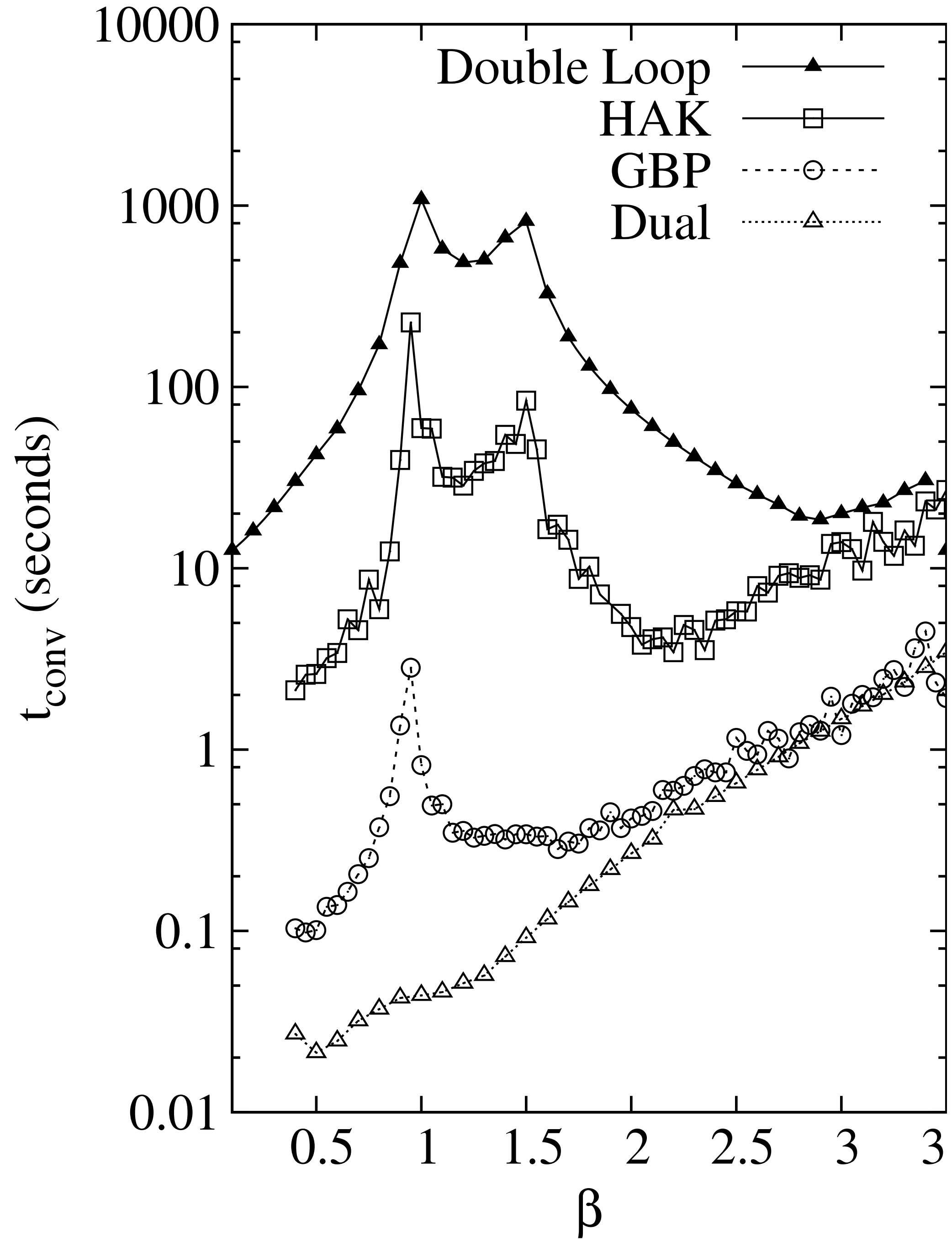
GF-GBP does it better



GF-GBP does it better

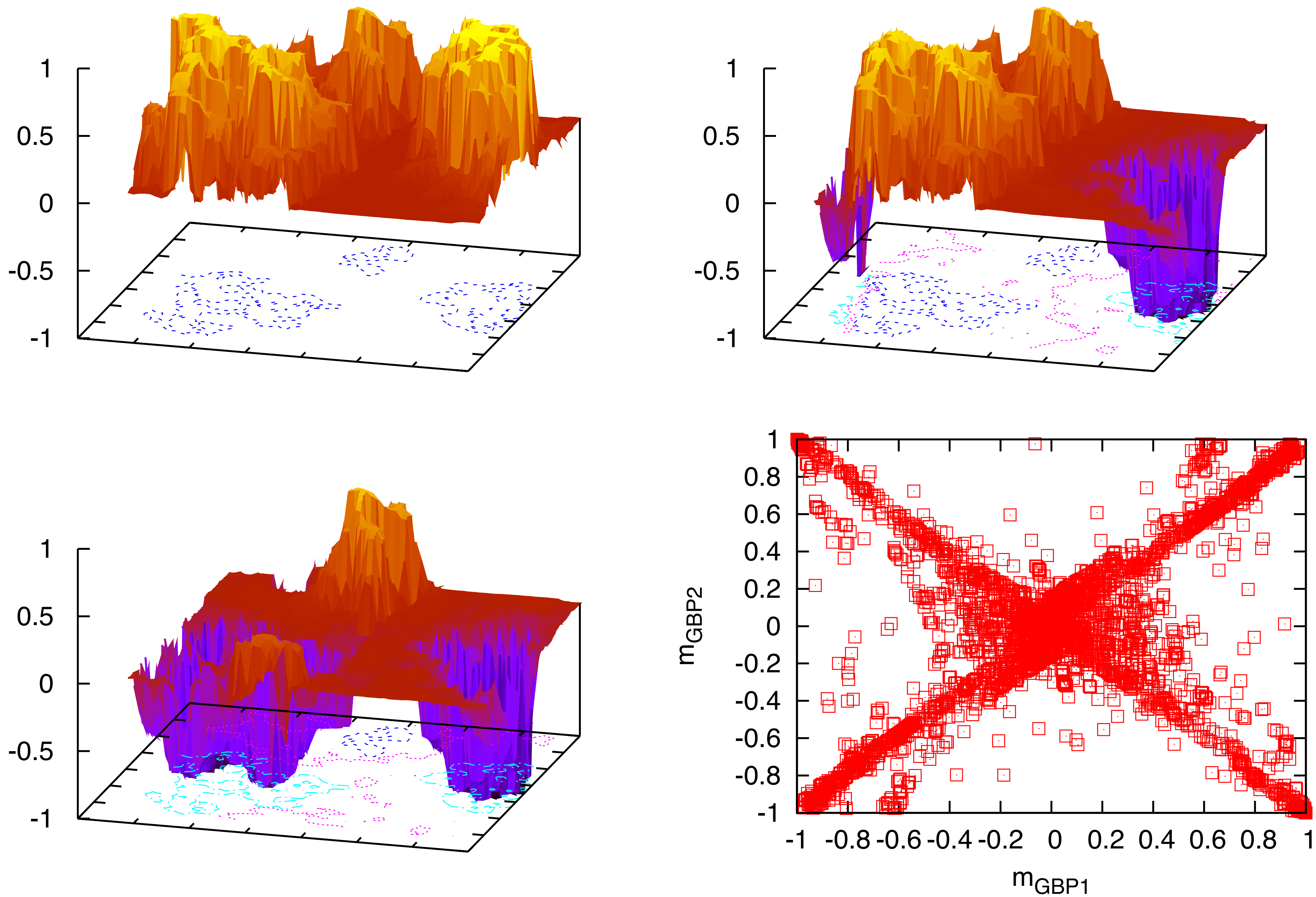
...and faster!



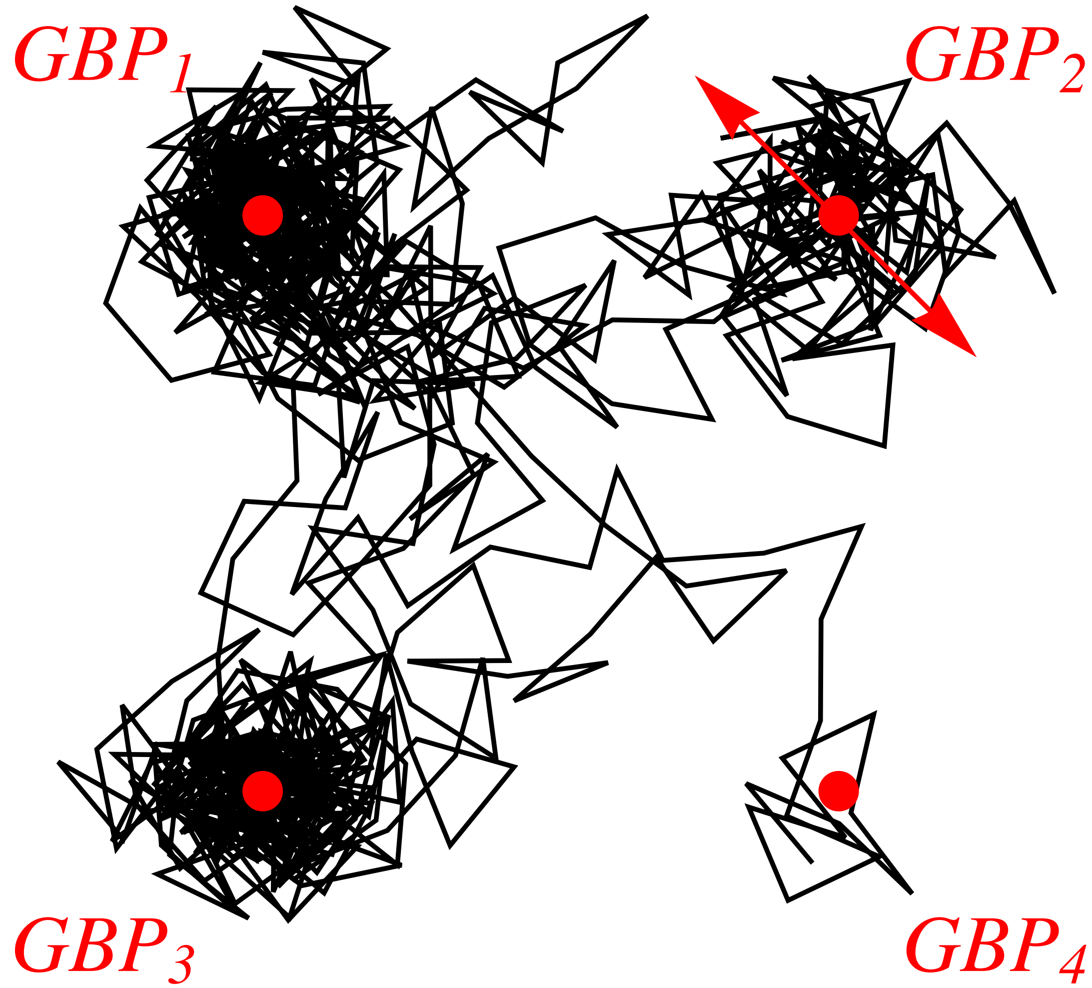


Relevance of SG solutions

$$T_c < T < T_{\text{conv}}$$

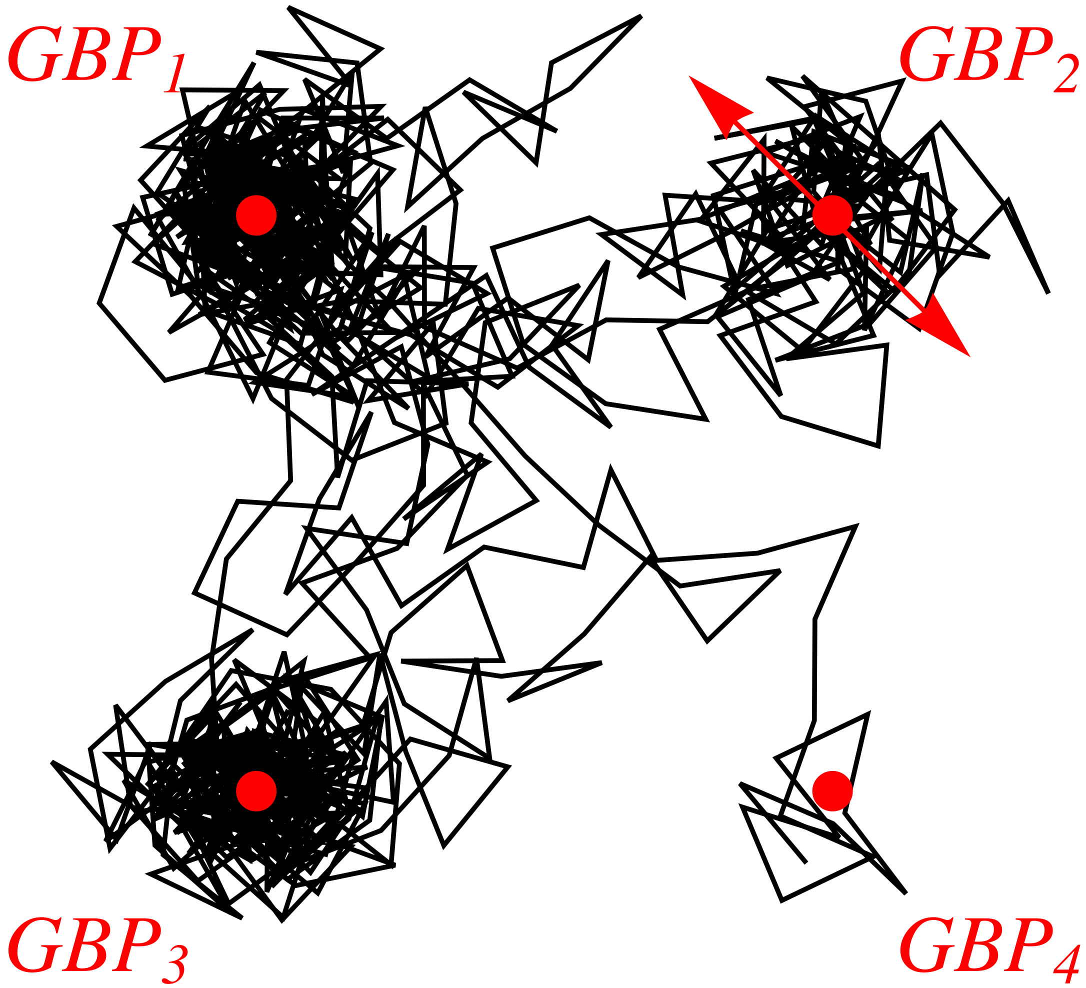


Relevance of SG solutions

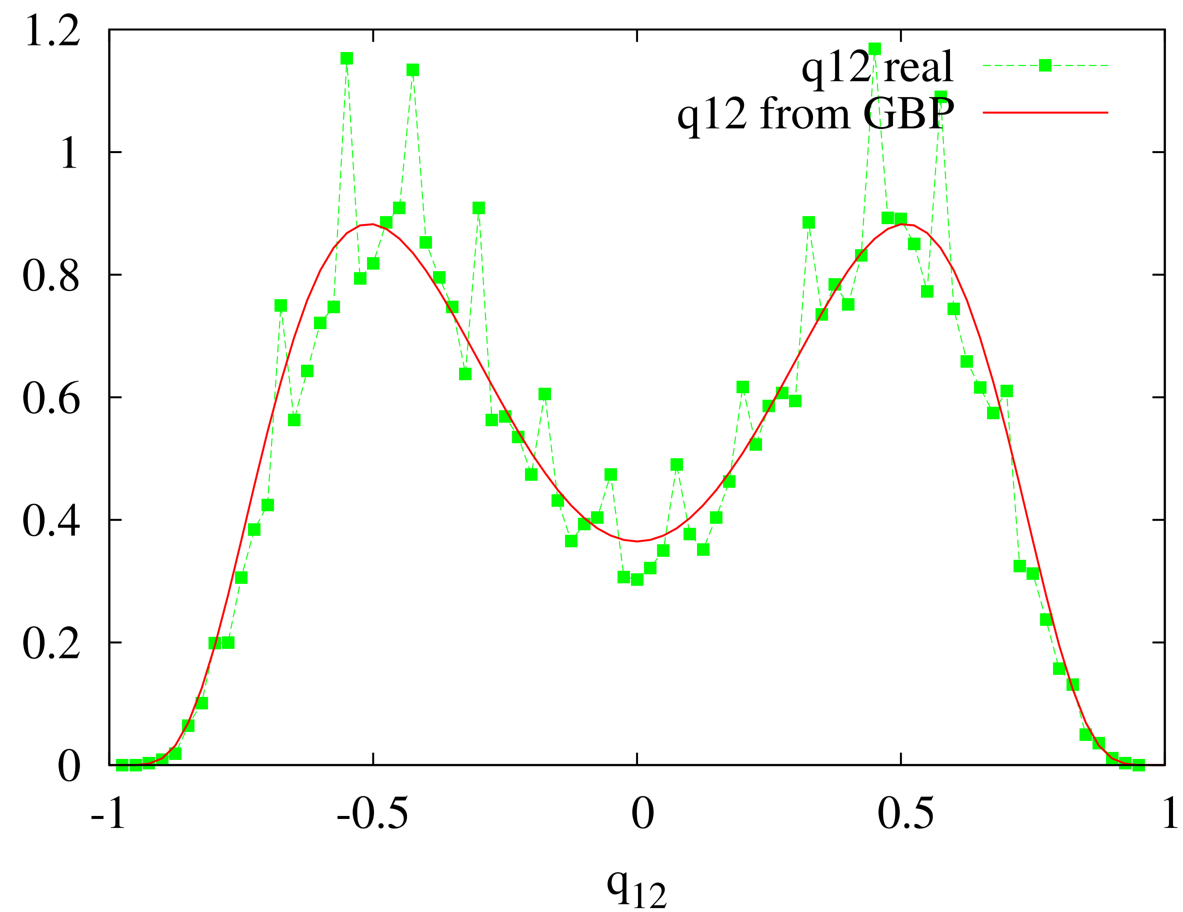


$$P(\vec{\sigma}) \simeq \sum_{\alpha \in \text{GBP f.p.}} e^{-\beta F_\alpha} P_\alpha(\vec{\sigma})$$

Relevance of SG solutions



$$P(\vec{\sigma}) \simeq \sum_{\alpha \in \text{GBP f.p.}} e^{-\beta F_\alpha} P_\alpha(\vec{\sigma})$$



Replica CVM

a technical slide :-)

n-replicated free energy

$$\Phi(n) = -\frac{1}{n\beta N} \ln \mathrm{Tr} \left\langle \exp\Big(\sum_{(ij)} \beta J_{ij} \sum_{a=1}^{n} s_i^a s_j^a \Big) \right\rangle_J$$

n-replicated spins $\sigma_i \equiv \{s_i^1, \ldots, s_i^n\}$

energy of region r

$$\psi_r(\sigma_r) \equiv \prod_{i,j \in r} \langle \exp(\beta J_{ij} \sum_a s_i^a s_j^a) \rangle_J$$

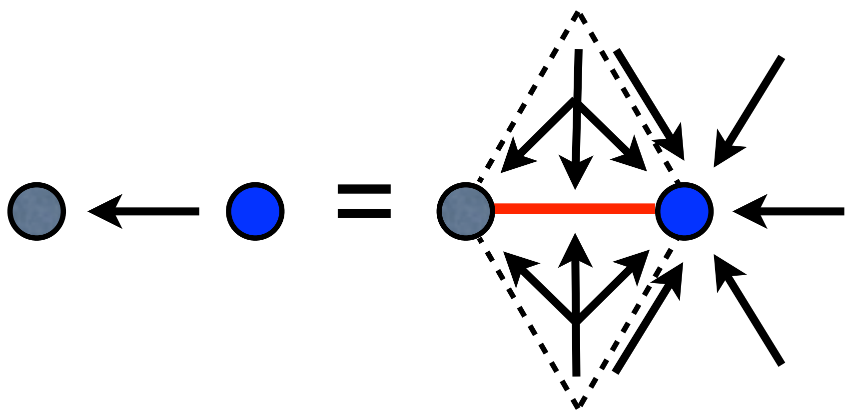
$$E_r = -\ln \prod_{ij} \psi_{ij}(\sigma_i, \sigma_j) - \ln \prod_i \psi_i(\sigma_i)$$

n-replicated CVM free energy

$$F_K = \sum_{r \in R} c_r \left(\sum_{x_r} b_r E_r + \sum_{x_r} b_r \ln b_r \right)$$

Replica CVM

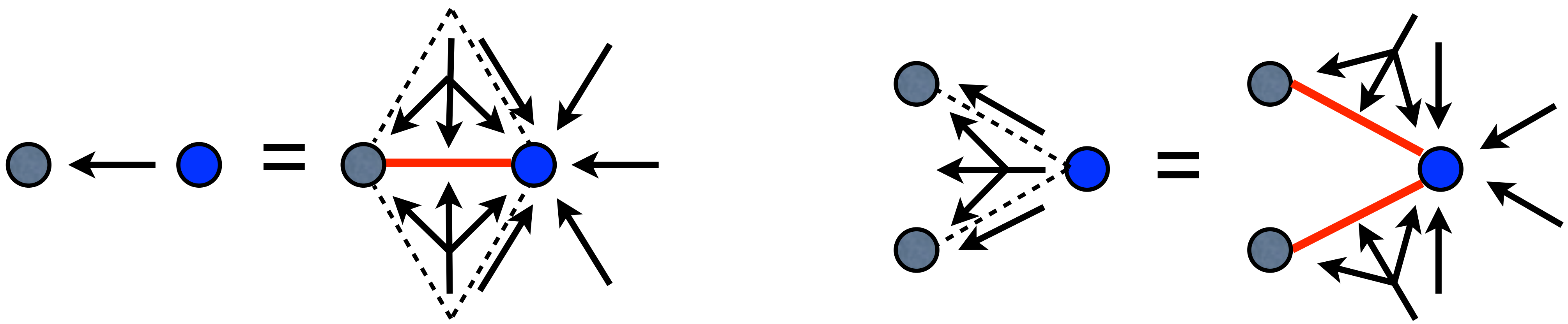
a second technical slide :-|



$$m_{(ij)\to j}(\sigma_j) \propto \sum_{\sigma_i} \psi_{(ij)}(\sigma_i,\sigma_j) M_{\alpha\to(ij)}(\sigma_i,\sigma_j) M_{\beta\to(ij)}(\sigma_i,\sigma_j) \prod_{k\in\partial i\setminus j} m_{(ki)\to i}(\sigma_i)$$

Replica CVM

a second technical slide :-|



$$m_{(ij)\to j}(\sigma_j) \propto \sum_{\sigma_i} \psi_{(ij)}(\sigma_i,\sigma_j) M_{\alpha\to(ij)}(\sigma_i,\sigma_j) M_{\beta\to(ij)}(\sigma_i,\sigma_j) \prod_{k\in\partial i\setminus j} m_{(ki)\to i}(\sigma_i)$$

$$M_{(ijk)\to(ij)}(\sigma_i,\sigma_j) m_{(ik)\to i}(\sigma_i) m_{(jk)\to j}(\sigma_j) \propto \sum_{\sigma_k} \psi_{(ik)}(\sigma_i,\sigma_k)\psi_{(jk)}(\sigma_j,\sigma_k)$$

$$\prod_{\alpha\in\partial(ik)\setminus(ijk)} M_{\alpha\to(ik)}(\sigma_i,\sigma_k) \prod_{\beta\in\partial(jk)\setminus(ijk)} M_{\beta\to(jk)}(\sigma_j,\sigma_k) \prod_{l\in\partial k\setminus\{i,j\}} m_{l\to k}(\sigma_k)$$

Replica CVM

a third technical slide :-(

RS ansatz on the messages

$$m(\sigma_i) = \int du\, q(u) \exp\left[\beta u \sum_{a=1}^{n} \sigma_i^a\right] (2\cosh \beta u)^{-n}$$

$$M(\sigma_i, \sigma_j) \propto \int dU\, du_i\, du_j\, Q(U, u_i, u_j) \exp\left[\beta U \sum_{a=1}^{n} \sigma_i^a \sigma_j^a + \beta u_i \sum_{a=1}^{n} \sigma_i^a + \beta u_j \sum_{a=1}^{n} \sigma_j^a\right]$$

Replica CVM

a third technical slide :-(

RS ansatz on the messages

$$m(\sigma_i) = \int du\, q(u) \exp\left[\beta u \sum_{a=1}^{n} \sigma_i^a\right] (2\cosh\beta u)^{-n}$$

$$M(\sigma_i, \sigma_j) \propto \int dU\, du_i\, du_j\, Q(U, u_i, u_j) \exp\left[\beta U \sum_{a=1}^{n} \sigma_i^a \sigma_j^a + \beta u_i \sum_{a=1}^{n} \sigma_i^a + \beta u_j \sum_{a=1}^{n} \sigma_j^a\right]$$

analytic continuation for n->0

$$q(u) = \int \prod_i^k dq_i \prod_\alpha^p dQ_\alpha \left\langle \delta(u - \hat{u}(\#)) \right\rangle_J \,,$$

$$R(U, u_a, u_b) \equiv \int du_i\, du_j\, Q(U, u_i, u_j) q(u_a - u_i) q(u_b - u_j) =$$

$$= \int \prod_i^K dq_i \prod_\alpha^P dQ_\alpha \left\langle \delta(U - \hat{U}(\#)) \delta(u_a - \hat{u}_a(\#)) \delta(u_b - \hat{u}_b(\#)) \right\rangle_J$$

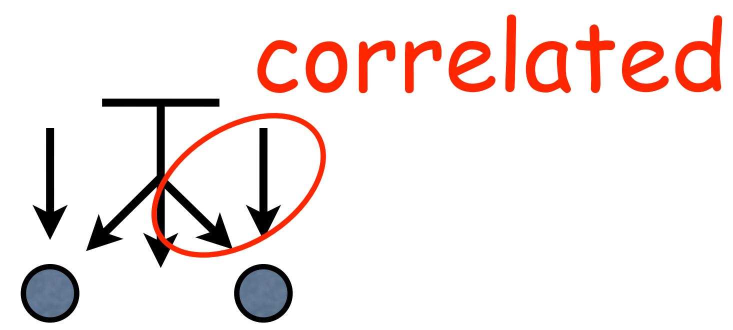
$$q(u) = \int \prod_i^k dq_i \prod_\alpha^p dQ_\alpha \, \langle \delta(u - \hat{u}(\#)) \rangle_J \, ,$$

$$R(U, u_a, u_b) \equiv \int du_i \, du_j \, Q(U, u_i, u_j) q(u_a - u_i) q(u_b - u_j) =$$

$$= \int \prod_i^K dq_i \prod_\alpha^P dQ_\alpha \, \langle \delta(U - \hat{U}(\#)) \delta(u_a - \hat{u}_a(\#)) \delta(u_b - \hat{u}_b(\#)) \rangle_J$$

Very hard to solve:

- convolution in $R(U, u_a, u_b)$
- non-positive defined $Q(U, u_i, u_j)$
- no population dynamics



Hard to solve because are the right equations...?

Paramagnetic solution stability analysis

$$q(u) = \delta(u) \qquad Q(U, u_i, u_j) = a_0(U)\delta(u_1)\delta(u_2)$$

first two moments

$$m = \int q(u)\, u\, \mathrm{d}u \ , \qquad a = \int q(u)\, u^2 \mathrm{d}u \ ,$$

$$M_i(U) = \iint Q(U, u_1, u_2)\, u_i\, \mathrm{d}u_i \ , \qquad a_{ij}(U) = \iint Q(U, u_1, u_2)\, u_i\, u_j\, \mathrm{d}u_1 \mathrm{d}u_2 \ ,$$

three phases:

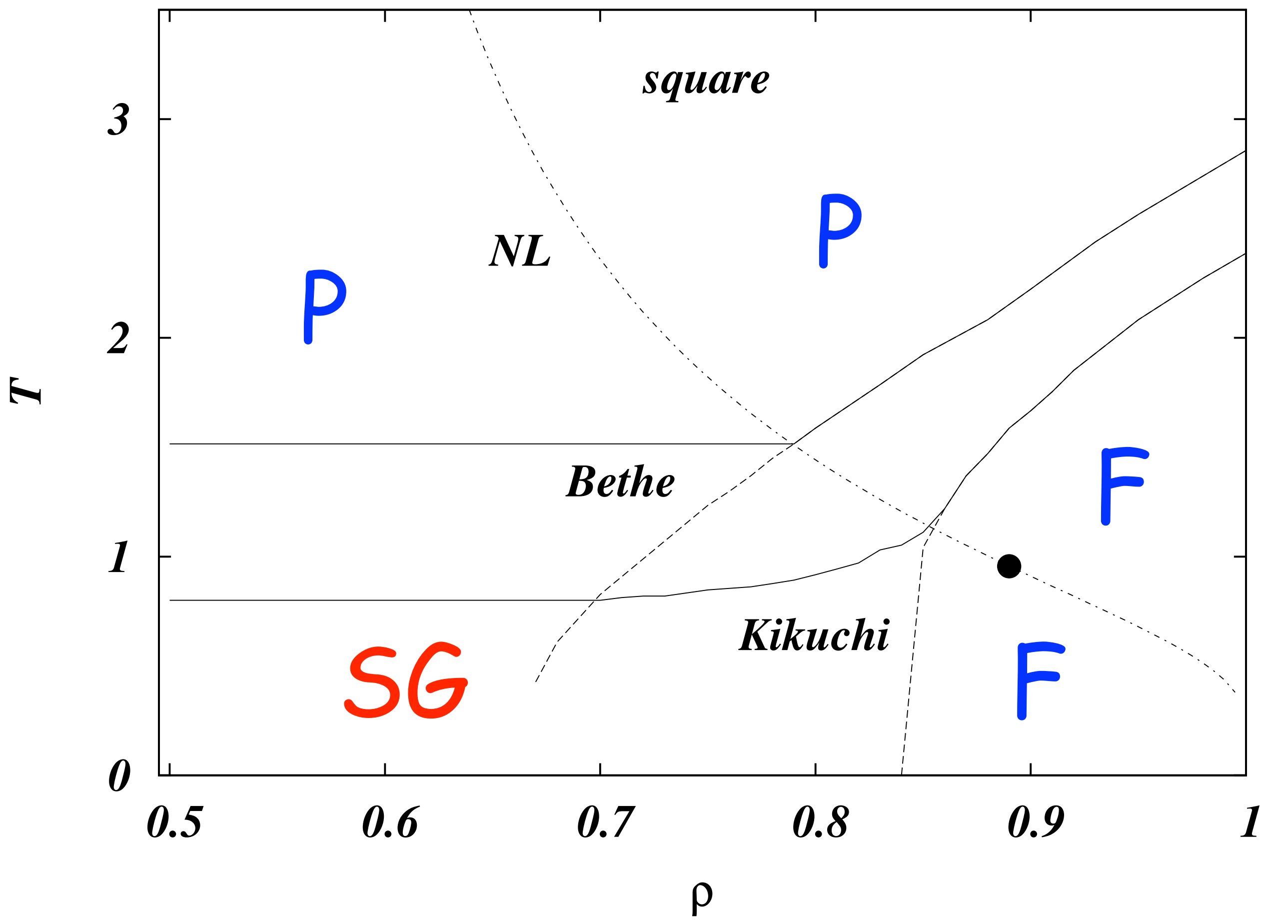
paramagnetic (P) $m = 0 \quad a = 0$

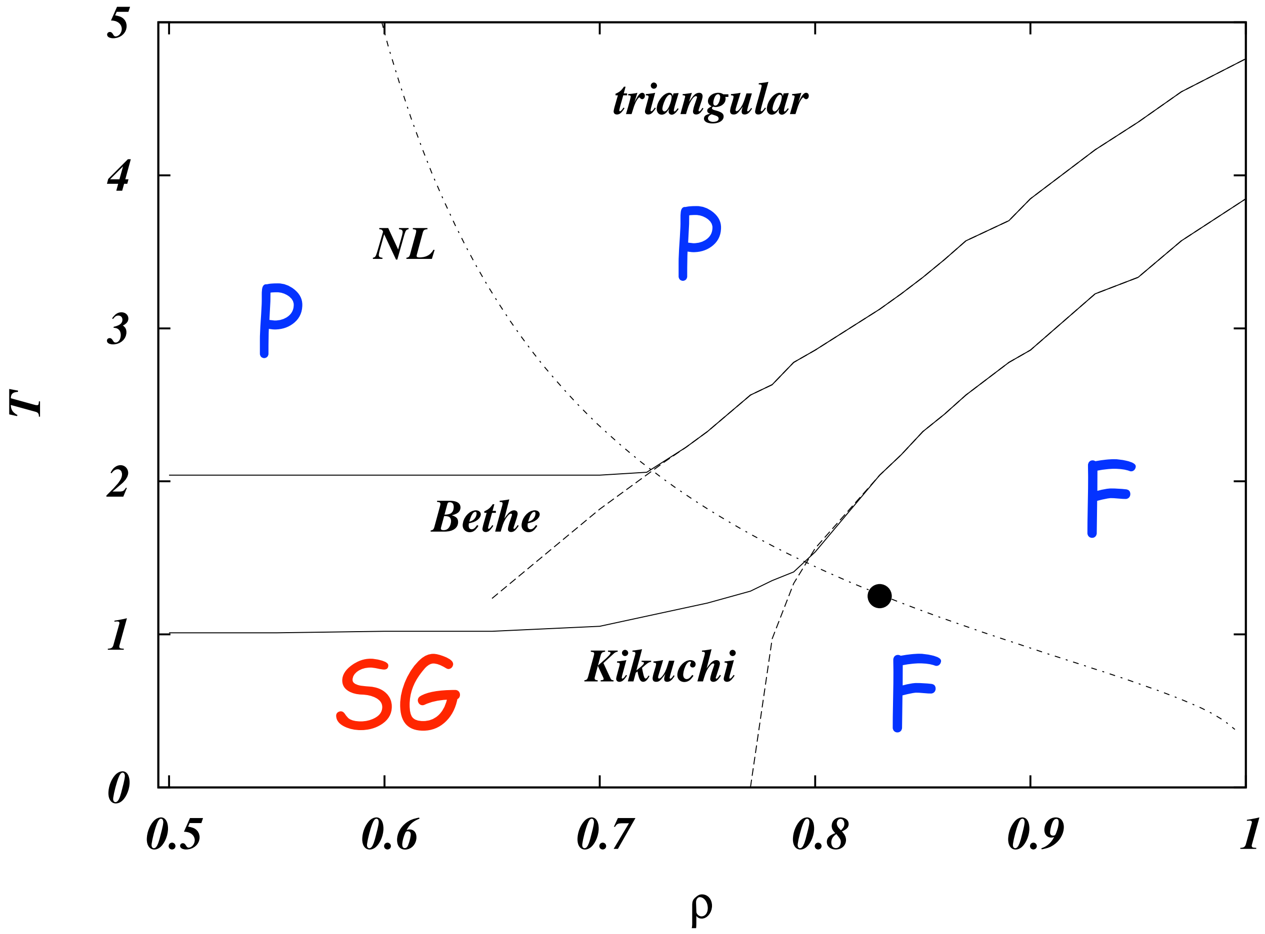
spin glass (SG) $m = 0 \quad a \neq 0$

ferromagnetic (F) $m \neq 0 \quad a \neq 0$

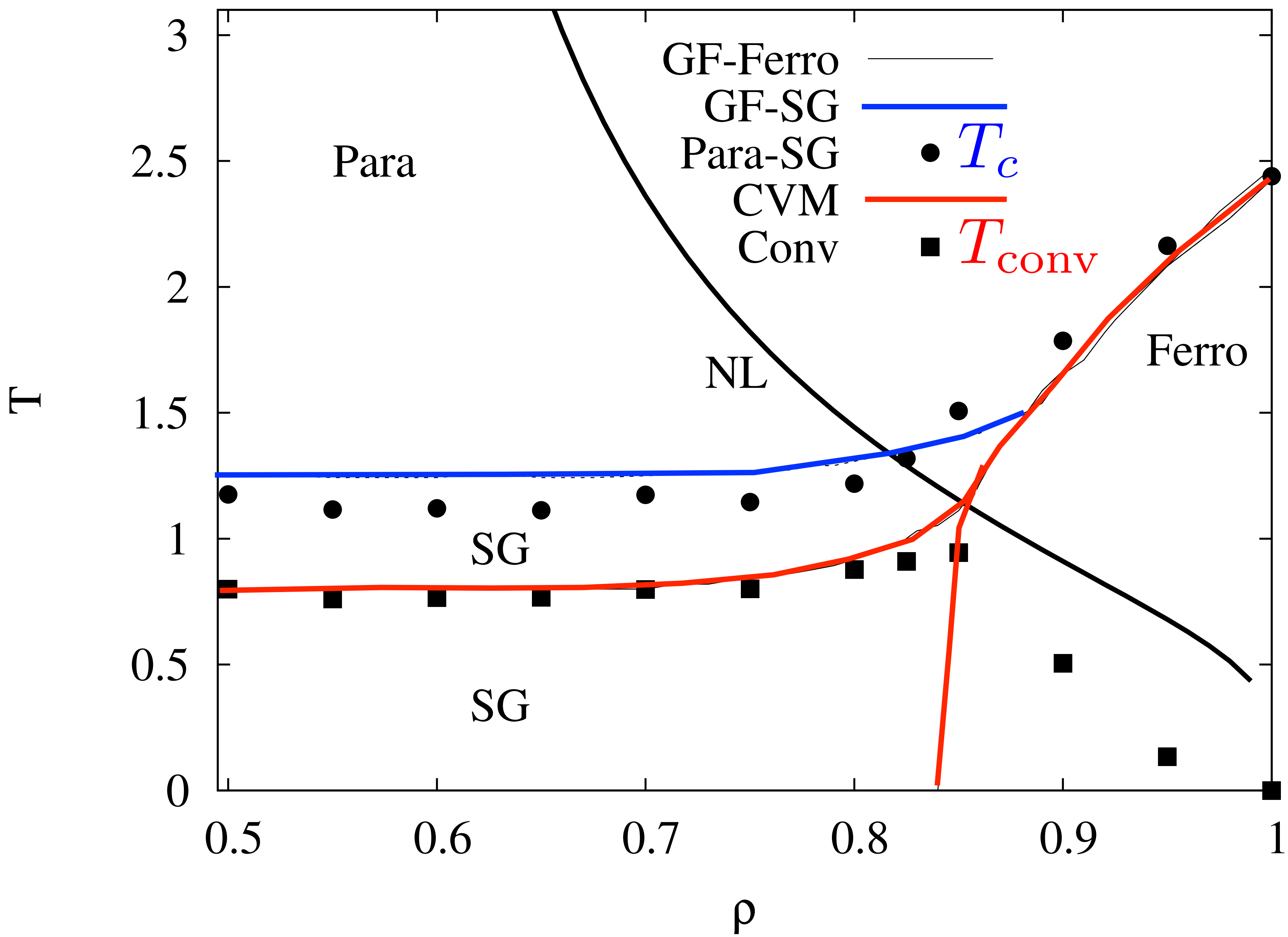
2D Plaquette Replica CVM

$$P(J) = \rho\,\delta(J-1) + (1-\rho)\delta(J+1)$$

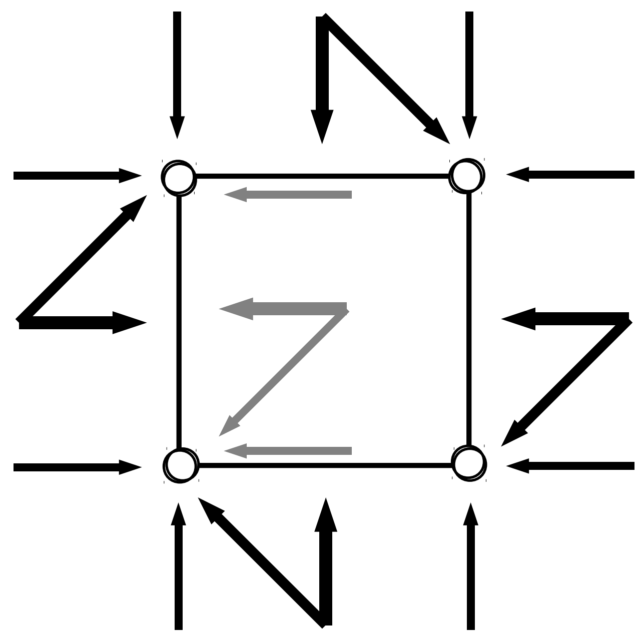




Replica CVM "explains" the behavior of GBP on given samples

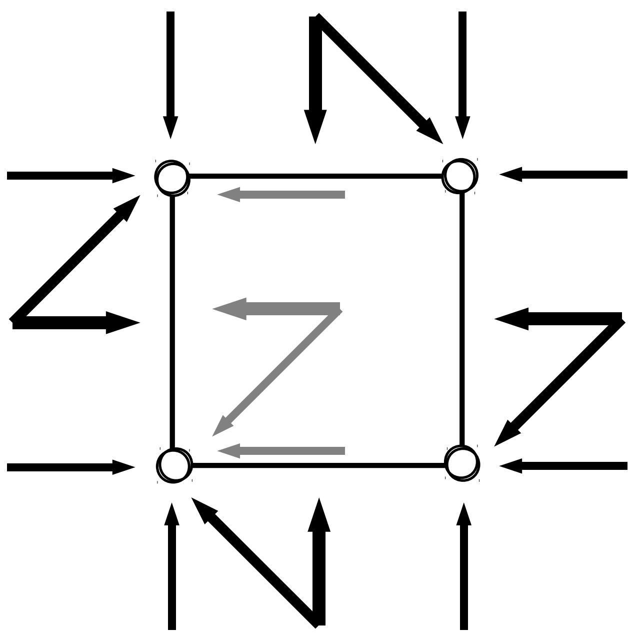


Population dynamics for gauge-fixed messages

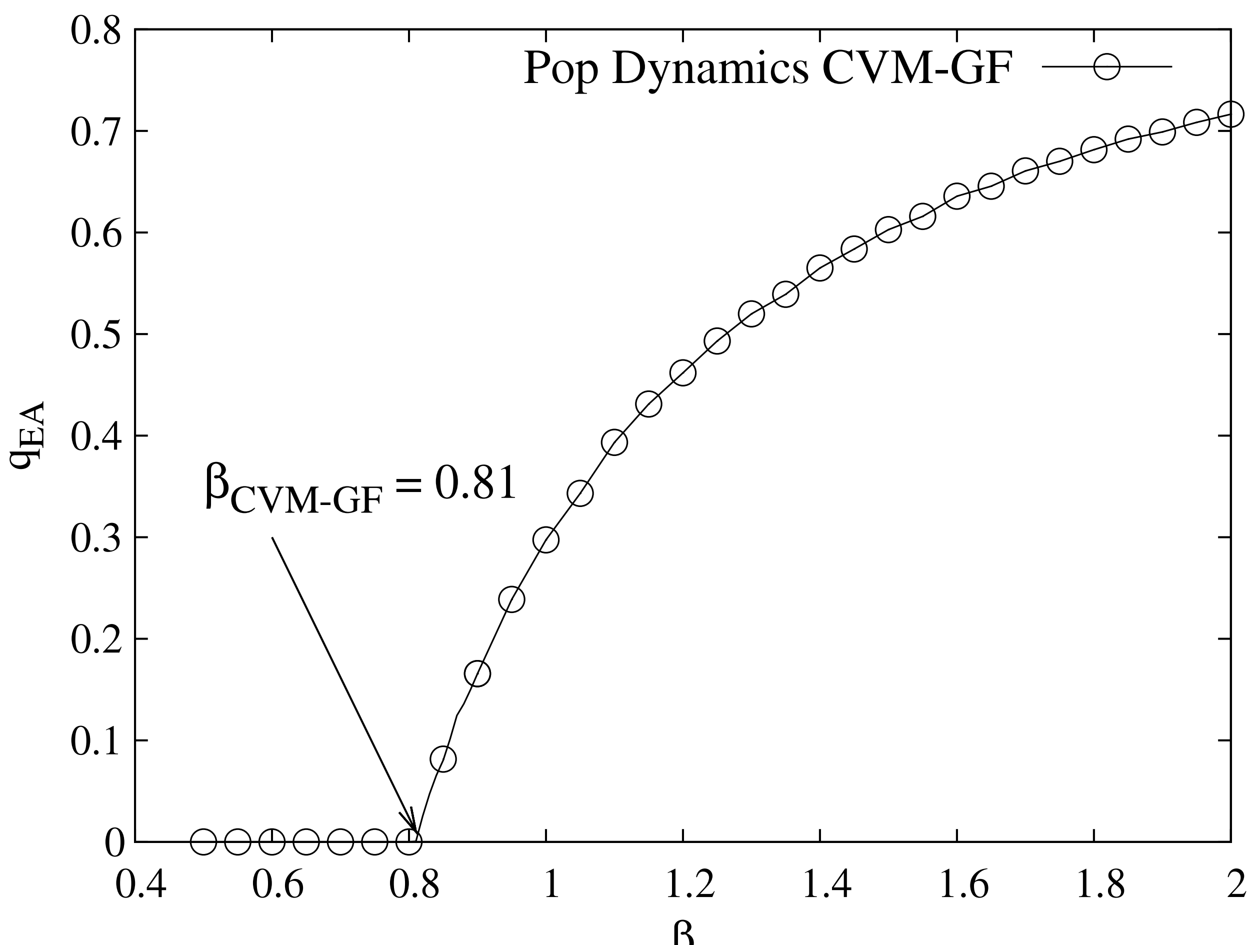


Choose randomly the external 4-fields and compute the internal 4-field. Repeat until convergence.

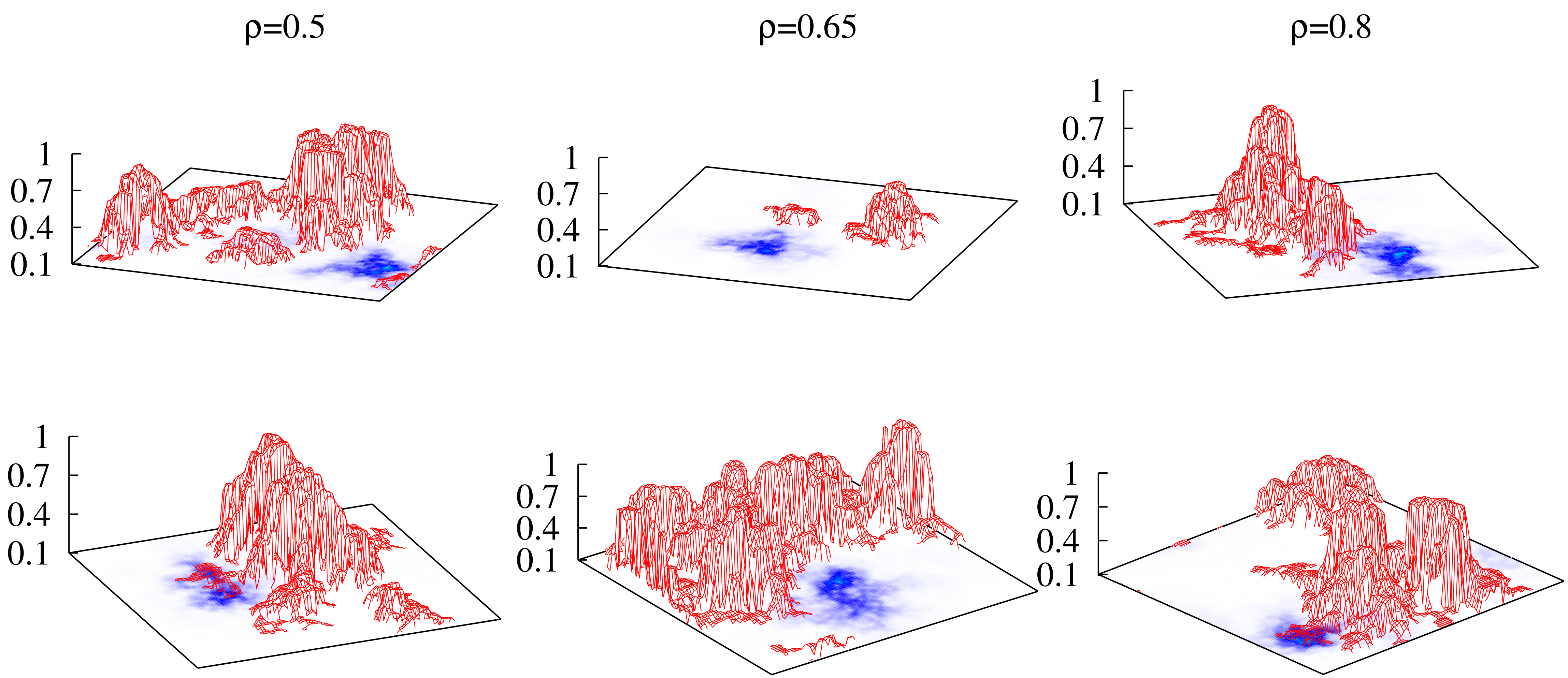
Population dynamics for gauge-fixed messages



Choose randomly the external 4-fields and compute the internal 4-field. Repeat until convergence.



Why T_{conv} and T_{CVM} are close?



Summary

- BP fails on regular lattices...
  and so improved methods based on BP
- GBP works better...
  much better if improved
  (dual algorithm, gauge fixing)
- SG solutions found by GBP are faithful
- Replica CVM improves largely Bethe
  (e.g. phase diagrams)
  and can "explain" GBP behavior