Making maximum entropy and linear response estimates consistent

Federico Ricci-Tersenghi
Physics Department
Sapienza University, Roma

joint work with Jack Raymond

Two very old problems...

$$P(s_1, \ldots, s_N) = \frac{1}{Z(\boldsymbol{J}, \boldsymbol{h})} \exp\left[\sum_{i \neq j} J_{ij} s_i s_j + \sum_i h_i s_i\right]$$

- DIRECT problem:
  given model parameters $\{J_{ij}, h_i\}$, compute average values $\langle s_i \rangle$ and correlations $\langle s_i s_j \rangle$ (or equivalently marginal probabilities)

- INVERSE problem:
  given measured mean values and correlations, estimate model parameters (previously known as Boltzmann machine learning)

...with a common difficulty

- Compute the free-energy

$$F(\boldsymbol{J}, \boldsymbol{h}) = \log Z(\boldsymbol{J}, \boldsymbol{h}) = \log \sum_{\{s_i\}} \exp\left(\sum J_{ij} s_i s_j + \sum_i h_i s_i\right)$$

  The sum is over exponentially many terms

- Resort to mean field approximations (MFA)
  - few parameters to be fixed self-consistently
  - fast to compute, but inexact!

MFA to the free-energy

- naive mean-field (nMF)
  only N parameters $m_i = \langle s_i \rangle$

$$H(x) \equiv -x \ln(x)$$

$$F_{\mathrm{nMF}} = \sum_i \left[H\left(\frac{1+m_i}{2}\right) + H\left(\frac{1-m_i}{2}\right) \right] + \sum_i h_i m_i + \sum_{i \neq j} J_{ij} m_i m_j$$

$$\frac{\partial F_{\mathrm{nMF}}}{\partial m_i} = \sum_j J_{ij} m_j + h_i - \mathrm{atanh}(m_i) = 0$$

$$\boxed{m_i = \tanh\left[h_i + \sum_j J_{ij} m_j \right]}$$

MFA to the free-energy

- nMF + Onsager reaction term (TAP)

$$F_{\mathrm{TAP}} = \sum_i \left[H\left(\frac{1+m_i}{2}\right) + H\left(\frac{1-m_i}{2}\right)\right] + \\ + \sum_i h_i m_i + \sum_{i\neq j} \left(J_{ij} m_i m_j + \frac{1}{2} J_{ij}^2 (1-m_i^2)(1-m_j^2) \right)$$

$$\boxed{m_i = \tanh\left[h_i + \sum_j J_{ij} \Big(m_j \underbrace{- J_{ij}(1-m_j^2) m_i}_{\text{reaction term}} \Big) \right]}$$

MFA to the free-energy

- Plefka expansion in small J

$$F_{\mathrm{nMF}} = \sum_i \left[H\left(\frac{1+m_i}{2}\right) + H\left(\frac{1-m_i}{2}\right)\right] + \sum_i h_i m_i + \sum_{i\neq j} J_{ij} m_i m_j$$

$$F_{\mathrm{TAP}} = \sum_i \left[H\left(\frac{1+m_i}{2}\right) + H\left(\frac{1-m_i}{2}\right)\right] + \\ + \sum_i h_i m_i + \sum_{i\neq j} \left(J_{ij} m_i m_j + \frac{1}{2} J_{ij}^2 (1-m_i^2)(1-m_j^2) \right)$$

nice discussion in T. Tanaka, PRE, 1998

MFA to the free-energy

- Bethe approximation (BA)
  nMF + nearest neighbors correlations parameters

$$c_{ij} = \langle s_i s_j \rangle - m_i m_j$$

$$\begin{aligned} F_{\mathrm{BA}} = \sum_{i \neq j} \Bigg[& H\left(\frac{(1+m_i)(1+m_j)+c_{ij}}{4}\right) + H\left(\frac{(1-m_i)(1-m_j)+c_{ij}}{4}\right) + \\ & + H\left(\frac{(1+m_i)(1-m_j)-c_{ij}}{4}\right) + H\left(\frac{(1-m_i)(1+m_j)-c_{ij}}{4}\right)\Bigg] + \\ + \sum_i (1-d_i) \left[H\left(\frac{1+m_i}{2}\right) + H\left(\frac{1-m_i}{2}\right)\right] & + \sum_i h_i m_i + \sum_{i \neq j} J_{ij}(c_{ij} + m_i m_j) \ , \end{aligned}$$

MFA to the free-energy

- Bethe approximation (BA)

$$f(m_1, m_2, t) = \frac{1 - t^2 - \sqrt{(1-t^2)^2 - 4t(m_1 - m_2 t)(m_2 - m_1 t)}}{2t(m_2 - m_1 t)}$$

$$\boxed{m_i = \tanh\left[h_i + \sum_j \operatorname{atanh}\Big(t_{ij} f(m_j, m_i, t_{ij})\Big)\right]}$$

MFA to the free-energy

- Bethe approximation (BA)

$$f(m_1, m_2, t) = \frac{1 - t^2 - \sqrt{(1-t^2)^2 - 4t(m_1 - m_2 t)(m_2 - m_1 t)}}{2t(m_2 - m_1 t)}$$

$$\boxed{m_i = \tanh\left[h_i + \sum_j \operatorname{atanh}\Big(t_{ij} f(m_j, m_i, t_{ij})\Big)\right]}$$

Small J expansion gives nMF, TAP, ...

$$h_i + \sum_j \operatorname{atanh}\Big(t_{ij} f(m_j, m_i, t_{ij})\Big) \simeq h_i + \sum_j \Big(J_{ij} m_j - J_{ij}^2 (1 - m_j^2) m_i + \dots\Big)$$

Computing correlations by linear response (LR)

- Correlations are trivial in MFA

$C_{ij} = 0$ in nMF, TAP and BA (between distant spins)

- Non trivial correlations can be obtained by using the linear response (Kappen Rodriguez, 1998)

$$C_{ij} = \frac{\partial m_i}{\partial h_j}\,, \qquad (C^{-1})_{ij} = \frac{\partial h_i}{\partial m_j}$$

Computing correlations by linear response (LR)

- Correlations are trivial in MFA

$C_{ij} = 0$ in nMF, TAP and BA (between distant spins)

- Non trivial correlations can be obtained by using the linear response (Kappen Rodriguez, 1998)

$$C_{ij} = \frac{\partial m_i}{\partial h_j} \,, \qquad (C^{-1})_{ij} = \frac{\partial h_i}{\partial m_j}$$

$$(C^{-1}_{\mathrm{nMF}})_{ij} = \frac{\delta_{ij}}{1-m_i^2} - J_{ij} \,,$$

$$(C^{-1}_{\mathrm{TAP}})_{ij} = \left[\frac{1}{1-m_i^2} + \sum_k J_{ik}^2 (1-m_k^2)\right] \delta_{ij} - \left(J_{ij} + 2J_{ij}^2 m_i m_j\right)$$

Computing correlations by linear response in BA

- Analytic expression for the correlations
  (FRT, JSTAT, 2012)

$$(C_{\mathrm{BA}}^{-1})_{ij} \; = \; \left[\frac{1}{1-m_i^2} - \sum_k \frac{t_{ik} f_2(m_k, m_i, t_{ik})}{1 - t_{ik}^2 f(m_k, m_i, t_{ik})^2} \right] \delta_{ij} - \frac{t_{ij} f_1(m_j, m_i, t_{ij})}{1 - t_{ij}^2 f(m_j, m_i, t_{ij})^2}$$

- Coincide with the fixed point of Susceptibility Propagation

- No need to run any algorithm!

MFA to the free-energy

From correlations to marginals...or better "beliefs"

$$b_i(\sigma_i) = \frac{1 + m_i \sigma_i}{2}$$

$$b_{ij}(\sigma_i, \sigma_j) = \frac{(1 + m_i \sigma_i)(1 + m_j \sigma_j) + C_{ij} \sigma_i \sigma_j}{4}$$

$$b_{ijk}(\sigma_i, \sigma_j, \sigma_k) = \ldots$$

$$b_{ijkl}(\sigma_i, \sigma_j, \sigma_k, \sigma_l) = \ldots$$

Useful if variables are correlated (e.g. form a clique or a loop)

MFA to the free-energy

Cluster Variational Method (CVM) or Region Graph Approx.

$$F = E - S$$

$$\begin{aligned} E(b, J, H) &= -\sum_{i<j} \mathrm{Tr}\, b_{ij} J_{ij} \sigma_i \sigma_j - \sum_i \mathrm{Tr}\, b_i H_i \sigma_i \\ &= -\sum_{i<j} J_{ij}(C_{ij} + m_i m_j) - \sum_i H_i m_i \end{aligned}$$

$$S(b) = -\sum_R c_R \mathrm{Tr}\, b_R \log b_R$$

Limits of MFA

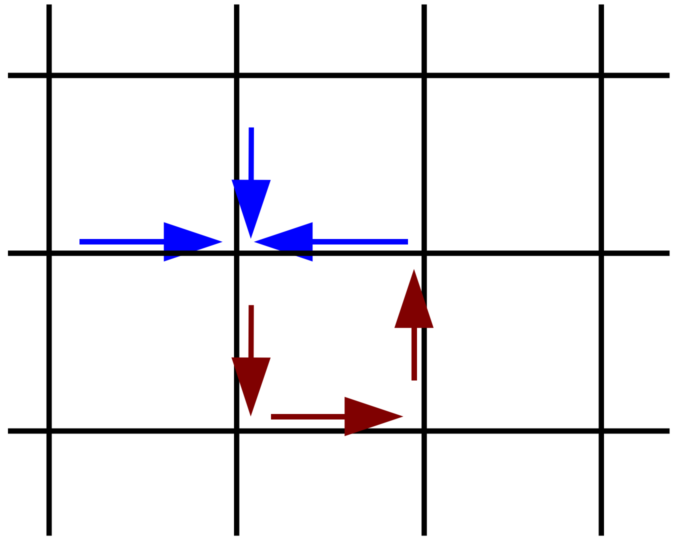
- Impossible to use large regions in CVM
  (too many parameters to optimize over)
  in general the largest region used is the smallest loop/clique in the graph
- Strong need for corrections to MFA that are able to take into account the effect of short loops and small local structures

How loops make MFA fail

e.g. Bethe approximation, high temperature phase

- maximum entropy

$$\langle \sigma_i \sigma_j \rangle_c^{\mathrm{BA}} = c^*_{ij} = \tanh(\beta J_{ij}) \;\; < \langle \sigma_i \sigma_j \rangle_c^{\mathrm{exact}}$$



- linear response

$$\chi_{ii} = 1 + \sum_{j \in \partial i} u_{j \to i, i} \neq 1$$

ferromagnet $\chi_{ij}^{\mathrm{BA}} > \langle \sigma_i \sigma_j \rangle_c^{\mathrm{exact}}$

Make MFA more consistent

"Consistency is more important than truth" (S. Ting)

- Add to the free-energy Lagrange multipliers to enforce (at the fixed point)

$$\boxed{\chi_{ii} = 1 - m_i^2 \qquad \chi_{ij} = c_{ij}}$$

- The entropy term changes to (for nMF and Bethe)

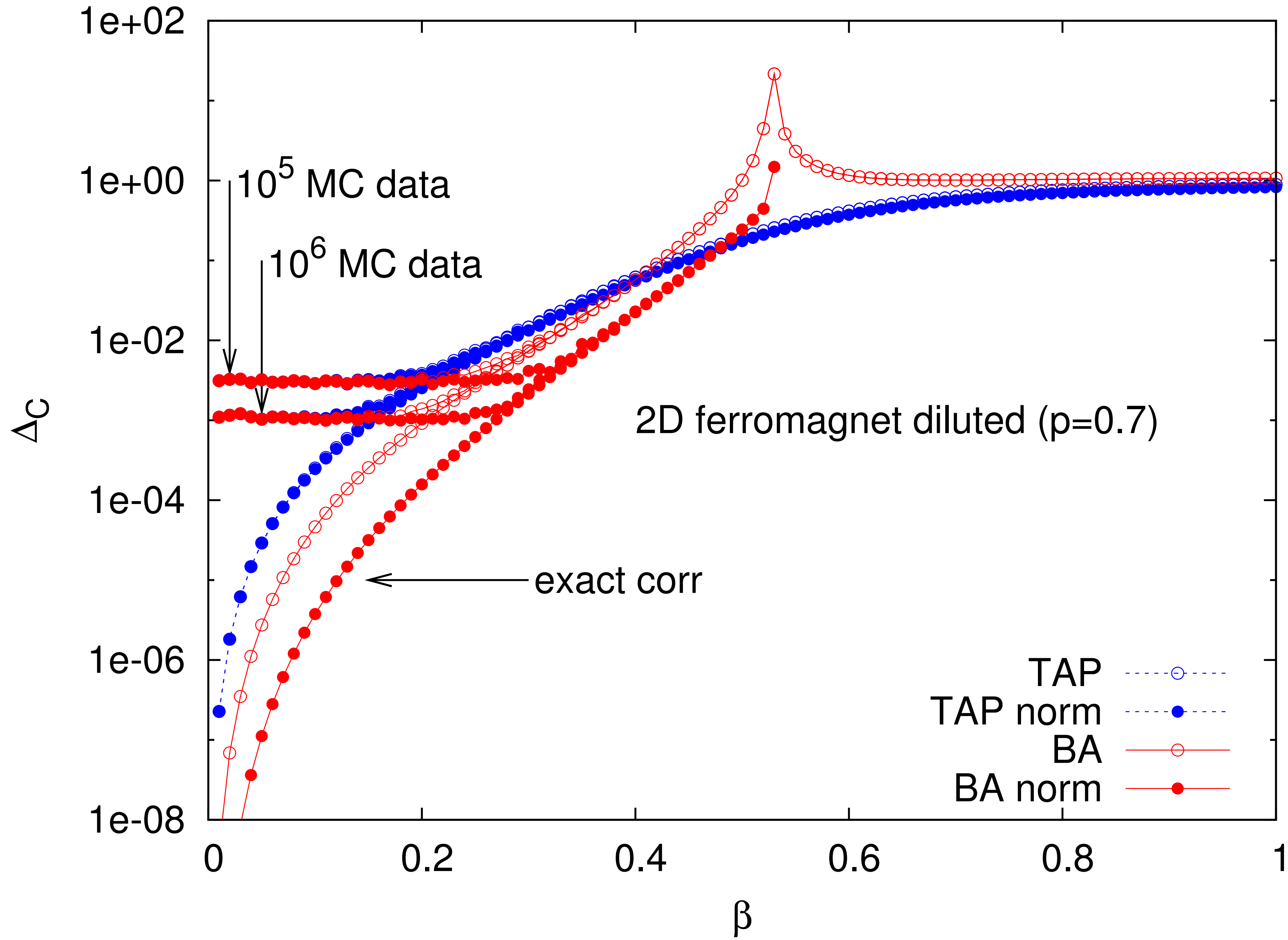
$$S_\lambda^N = -\sum_i \mathrm{Tr}\left[b_i \log b_i\right] - \sum_i \lambda_i \left(\left(1 - m_i^2\right) - \chi_{ii}\right)/2$$

$$S_\lambda^B = S_\lambda^N - \sum_{ij \in I} \left\{ \mathrm{Tr}\left[b_{ij} \log\left(\frac{b_{ij}}{b_i b_j}\right)\right] - \lambda_{ij}\left(C_{ij} - \chi_{ij}\right) \right\}$$

Previous proposals for fixing the diagonal terms χ_{ii}

- Kappen, Rodriguez (Neur. Comp., 1998)
  MF + self-couplings J_{ii}
- Opper, Winter (PRL, PRE, 2001)
  TAP + λ_i
- FRT (JSTAT, 2012)
  normalized correlations useful for inverse pb.
  $$\widehat{\chi}_{ij} \equiv \frac{\chi_{ij}}{\sqrt{\chi_{ii}\chi_{jj}}}$$
- Yasuda, Tanaka (PRE, 2013)
  Bethe + λ_i

Improvement by normalizing correlations (diluted 2D Ising)



General framework

$$\boxed{F_\lambda = F_{\mathrm{MFA}}(\{m_i\},\{C_{ij}\},\ldots) + \sum_i \lambda_i m_i^2 + \sum_{i<j} \lambda_{ij} C_{ij}}$$

Your preferred MFA

can be set to zero to recover known approx. or to satisfy

$$\boxed{\chi_{ii} = 1 - m_i^2 \qquad \chi_{ij} = c_{ij}}$$

General framework

$$F_\lambda = F_{\mathrm{MFA}}(\{m_i\},\{C_{ij}\},\ldots) + \sum_i \lambda_i m_i^2 + \sum_{i<j} \lambda_{ij} C_{ij}$$

Maximum Entropy equations
from free-energy minimization

$$\partial_{m_i} F_\lambda = \mathrm{atanh}(m_i) + \partial_{m_i} S^{2+} - H_i - \sum_j J_{ij} m_j + \lambda_i m_i = 0$$

$$\partial_{C_{ij}} F_\lambda = \partial_{C_{ij}} S^{2+} - J_{ij} + \lambda_{ij} = 0$$

$$\partial_{C_{ijk}} F_\lambda = \partial_{C_{ijk}} S^{2+} = 0$$

Higher order correlation parametrs
are fixed by standard maximum entropy

General framework

$$F_\lambda = F_{\mathrm{MFA}}(\{m_i\},\{C_{ij}\},\ldots) + \sum_i \lambda_i m_i^2 + \sum_{i<j} \lambda_{ij} C_{ij}$$

Linear Response equations from quadratic exp. around the free-energy mininum

$$\partial^2 F_\lambda \quad \Rightarrow \quad [\chi^{-1}]_{ij} = \Phi_{ij} - J_{ij}$$

Some comments before showing results...

- Bethe/CVM free-energies are not convex adding parameters it not obvious to improve!
- λ parameters measure how wrong is the MFA e.g. on a random graph with Bethe approx. $\lambda \to 0$
- $J_{ij} \to J_{ij} + \lambda_{ij}$ would naively imply stronger correlations and also more unstable Susc. Prop. but this is not the case!

Models studied

We are starting with exactly solvable models to check the new method

- Ising variables
- 2D topology (square and triangular lattices)
- homogeneous -> Fourier transform solution
- both ferromagnetic and **frustrated**

...willing to arrive to a general purpose algorithm!

Ferromagnet on a 2D square lattice

$$u = \widehat{u}_{J+\lambda}(h) \equiv \mathrm{atanh}[\tanh(J+\lambda)\tanh(h)]$$

$$h = 3u + (\lambda_0 - 4\lambda)m$$

Modified BP

$$m = \tanh(h+u)$$

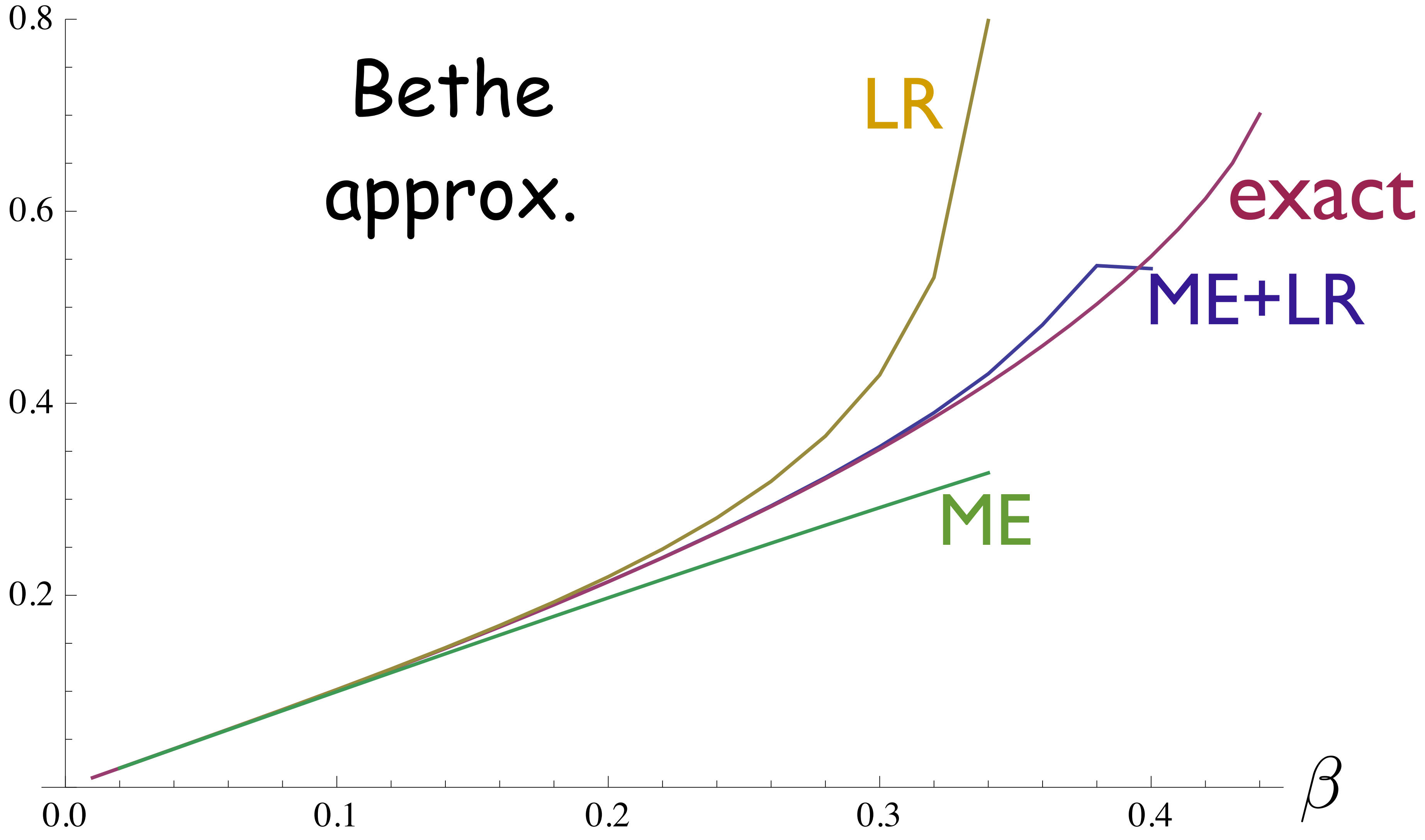
Modified SuscP leading to

$$\chi_x = \frac{1-m^2}{(2\pi)^2}\int_{-\pi}^{\pi} dq_1 \int_{-\pi}^{\pi} dq_2 \frac{\cos(q_1)^x}{\phi_0 - 2\phi_1[\cos(q_1)+\cos(q_2)]}$$

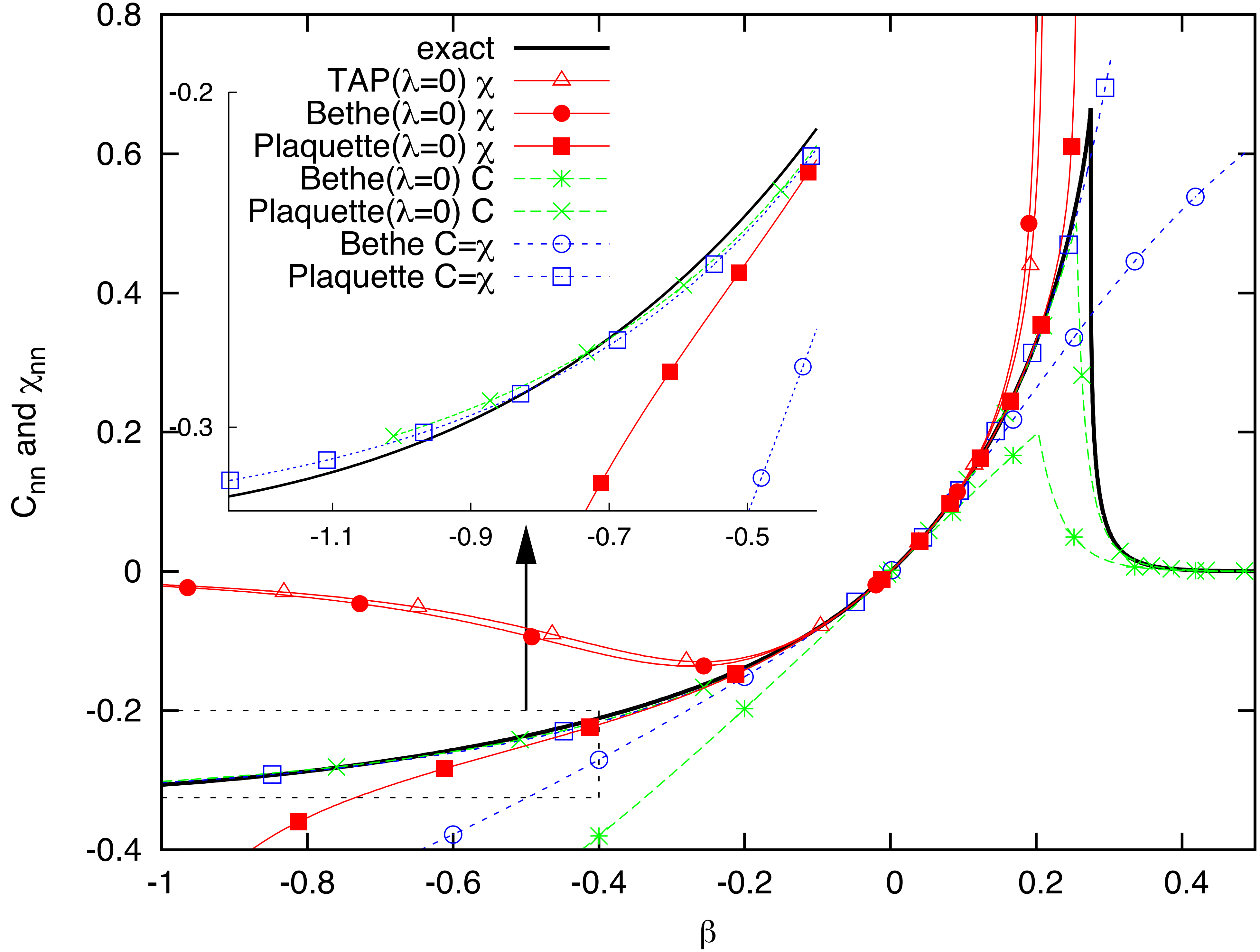
$$\phi_0 = 1 + 4\frac{\widehat{u}'^2}{1-\widehat{u}'^2} - (1-m^2)\lambda_0$$

$$\phi_1 = \frac{\widehat{u}'}{1-\widehat{u}'^2} - (1-m^2)\lambda$$

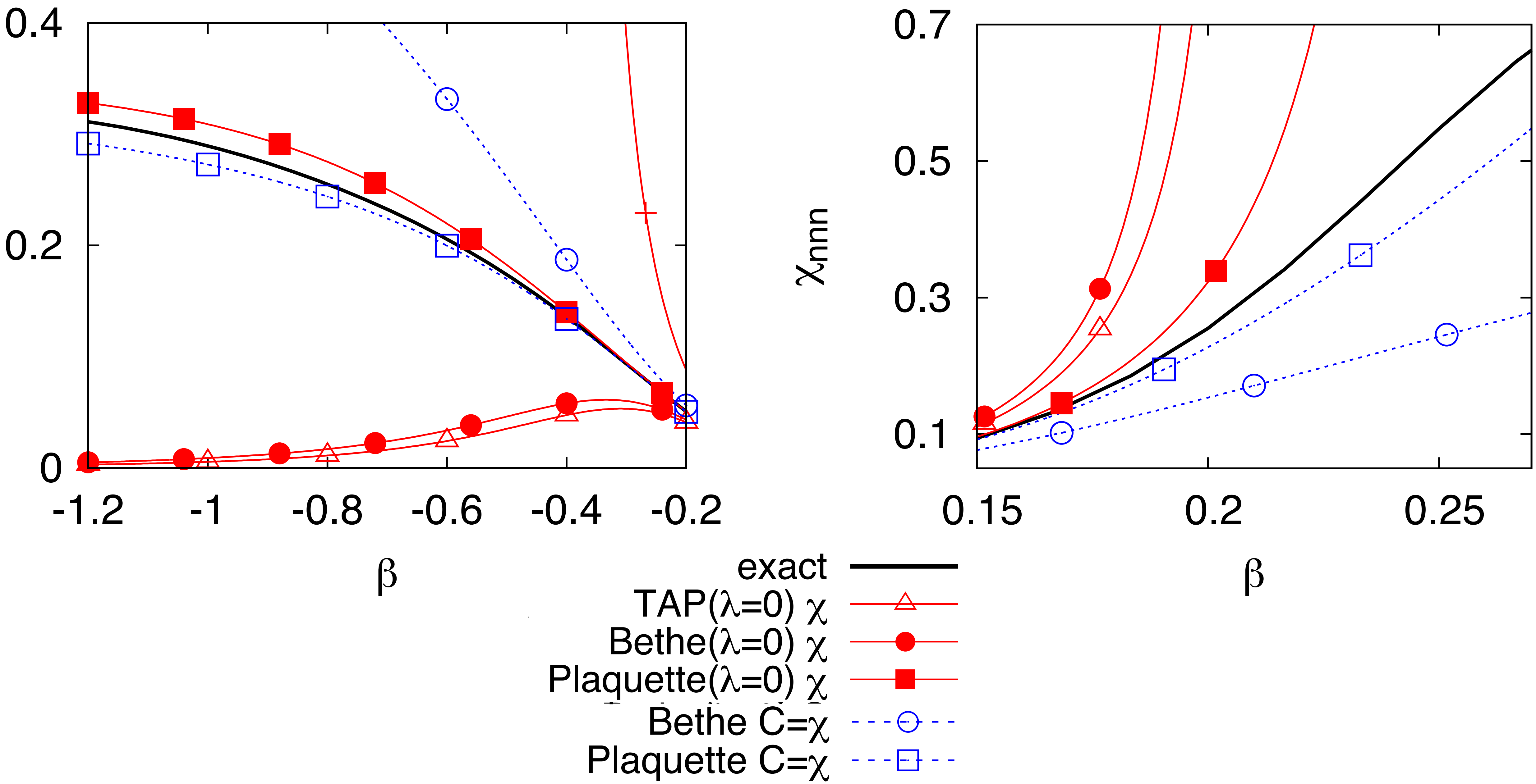
Nearest-neighbor correlation (2D square lattice)



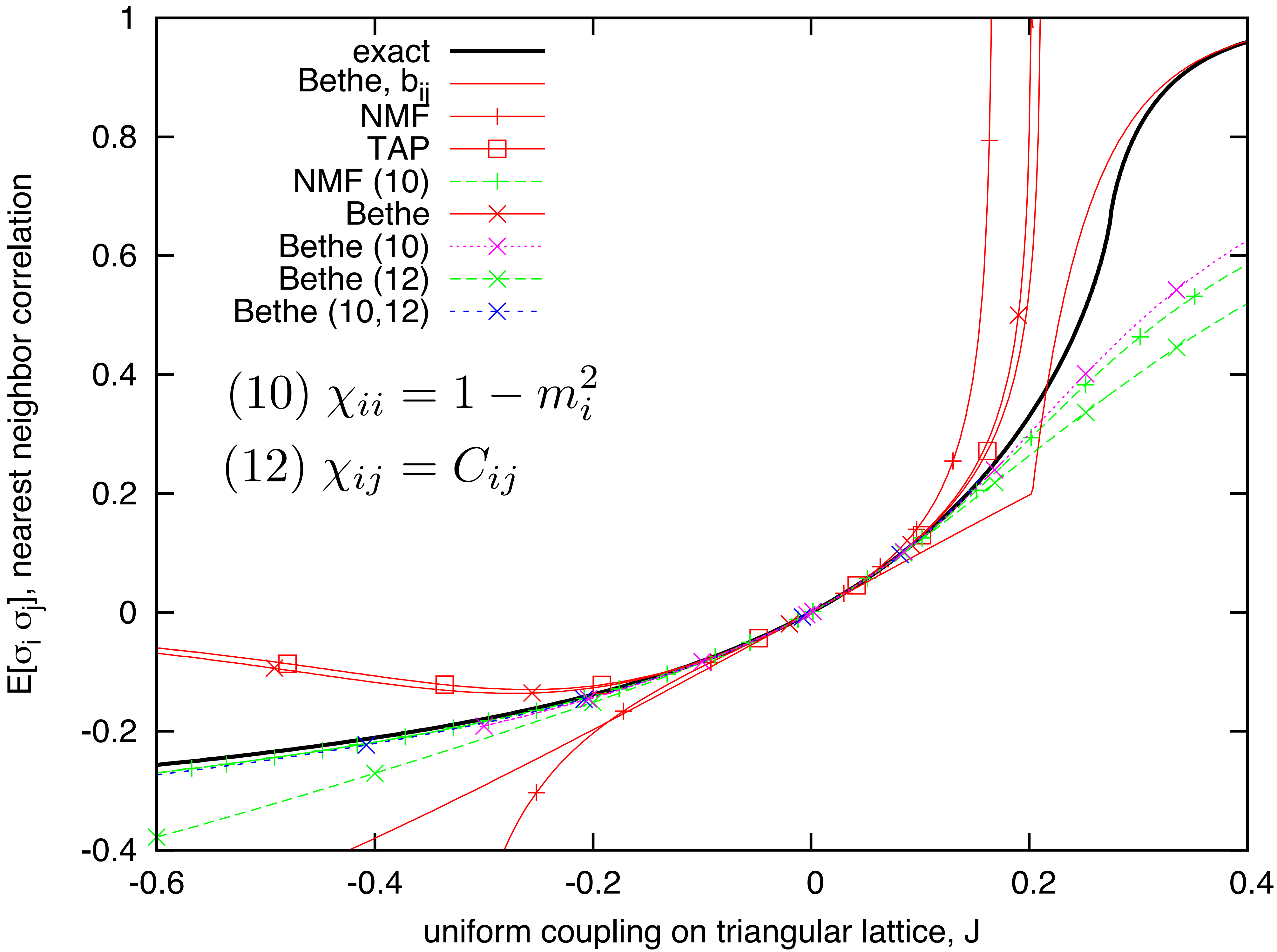
Nearest-neighbor correlation (2D triangular lattice, $\lambda_0 = 0$)



Next nearest-neighbor correlation (2D triangular lattice, $\lambda_0 = 0$)



Nearest-neighbor correlation (2D triangular lattice)

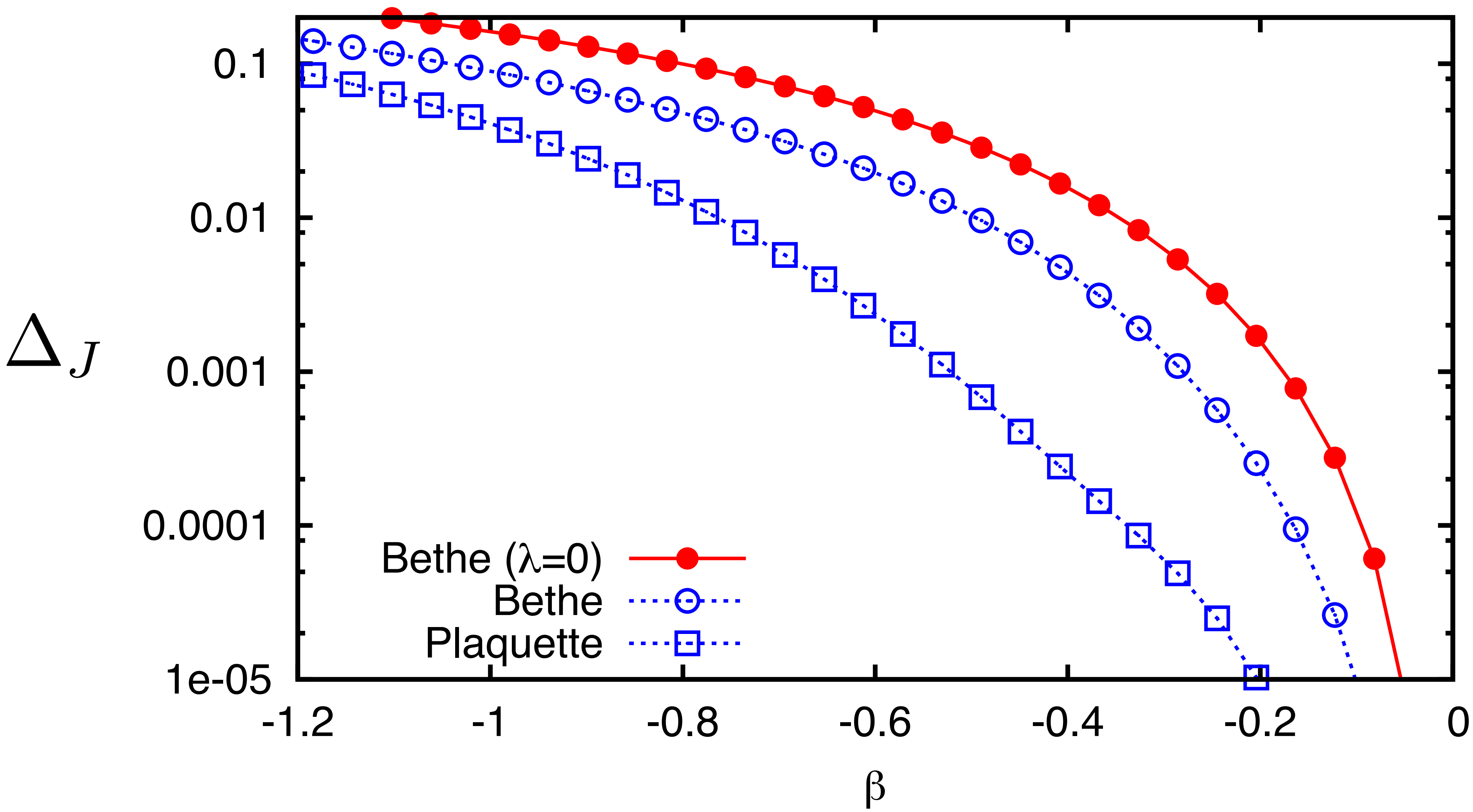


Inverse problem

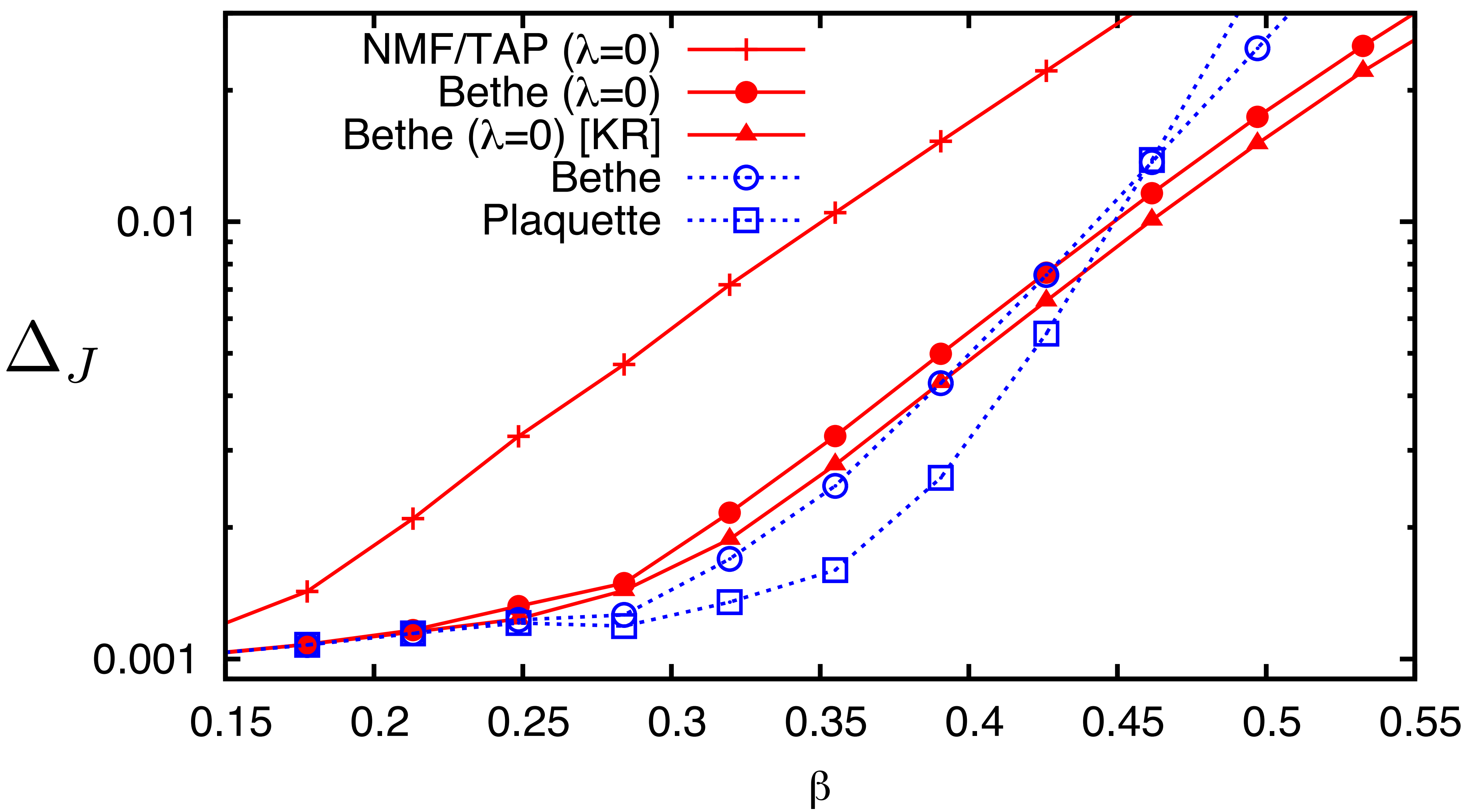
- Use the CVM for the free-energy including all pairs (Bethe) or all triplets (Plaquette)
- Solve equations in J_{ij}, H_i and eventually λ given (exact) correlations and magnetizations
- Measured the error on the inferred couplings by

$$\Delta_J = \sqrt{\frac{\sum_{i<j}(J'_{ij} - J_{ij})^2}{\sum_{i<j} J_{ij}^2}}.$$

Inverse problem from exact statistics (2D triangular lattice, $\lambda_0 = 0$)



Inverse problem from MC data (diluted 2D square lattice, $\lambda_0 = 0$)



Summary

New general framework to force in CVM the consistency between maximum entropy estimates and linear response ones.

Several improvements achieved

- better high temperature expansions
- better correlation estimates
- smaller errors in estimating couplings

...still working for a fast message passing algorithm ;-)