A message-passing scheme for non-equilibrium stationary states

NORDITA Seminar
December 2, 2010

E.A., H. Mahmoudi (2010) [in preparation]

I. Neri, D. Bollé, J. Stat Mech. (2009) P08009
Y. Kanoria, A. Montanari (2009) arXiv:0907.0449

Message-passing schemes are ways to approximately compute marginals of probability distributions.

In Artificial Intelligence known as Belief Propagation (BP); in Information Theory known as iterative decoding.

KTH/CSC

Bethe-Peierls approximation = Belief Propagation

JS Yedidia, WT Freeman, Y Weiss (2001)
M Mézard, A Montanari Physics, Oxford University Press (2009)

Configuration space; N discrete variables $\vec{s} \in \{1, ..., q\}^N$

Cost or energy function; sum of local terms $\mathcal{H}(\vec{s}) = \sum_{a=1}^{M} e_a(\vec{s}_a)$

Marginal (one-spin) probabilities $p_i(s_i) = \frac{1}{Z} \sum_{\vec{s} \setminus s_i} e^{-\beta \mathcal{H}(\vec{s})}$

In statistical physics, artificial intelligence and information theory, such marginals (magnetizations, correlation functions etc) are typically what one wants to compute (or use)

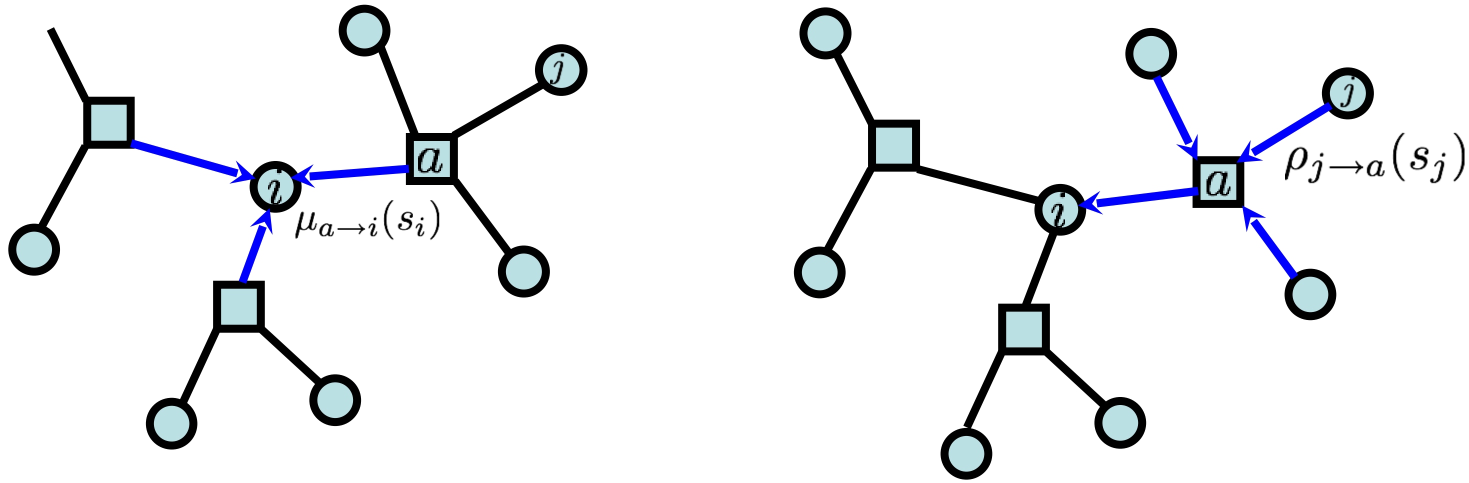
$$\begin{cases} p(\vec{s}) & \propto & \frac{\prod_{a=1}^{M} p_a(\vec{s}_a)}{\prod_{i=1}^{N} (p_i(s_i))^{d_i - 1}} \\ p_i(s_i) & = & \sum_{\vec{s}_a \setminus s_i} p_a(\vec{s}_a) \end{cases}$$

Bethe-Peierls approximation

Belief propagation

KTH/CSC

Factor graph representation illustrates canonical BP



The "messages" $\mu_{a\to i}(s_i)$ and $\rho_{j\to a}(s_j)$ are Lagrange multipliers.

They enforce the conditions that the marginal probabilities p_a over the spins partaking in one interaction a, summed over all those spins except i, is the marginal probability p_i, over that last spin.

$$
\begin{aligned}
p_i(s_i) &= D_i \prod_{a\in\partial i} \mu_{a\to i}(s_i) \\
\mu_{a\to i}(s_i) &= C_{a\to i} \sum_{\vec{s}_a\setminus i} \exp\left(-\beta\, e_a(\vec{s}_a)\right) \prod_{j\in\partial a\setminus i} \rho_{j\to a}(s_j) \\
\rho_{i\to a}(s_i) &= D_{i\to a} \prod_{b\in\partial i\setminus a} \mu_{b\to i}(s_i)
\end{aligned}
$$

- Iterative procedure to find fixed points of BP equations
- Exact on trees and an approximation on loopy graphs
- No guarantee to converge

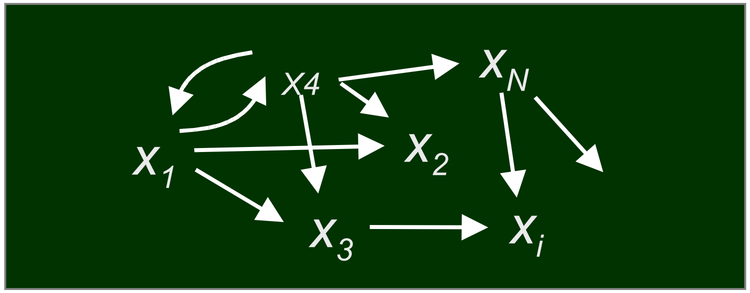
Non-equilibrium is the domain of kinetic theory.

There is no Gibbs measure as there is in equilibrium.

But can one nevertheless compute marginals by a message-passing scheme?

KTH/CSC

Today's case study: diluted asymmetric spin glass



S Kauffman (1969)
J Hopfield (1982)

B Derrida (1987)
A Crisanti, H Sompolinski (1988)
….and many others

$$\frac{d}{dt}p(\vec{\sigma}(t)) = W(t)\,p(\vec{\sigma}(t))$$

Fully asynchronous updates (master equations); not yet done

$$p(\vec{\sigma}(0)\ldots\vec{\sigma}(t)) = \prod_{s=1}^{t} W[\,\vec{\sigma}(s)\,|\,\vec{h}(s)]\;p(\vec{\sigma}(0))$$

$$W[\,\vec{\sigma}(s)\,|\,\vec{h}(s)] = \prod_{i=1}^{N} \frac{\exp(\beta\,h_i(s)\,\sigma_i(s))}{2\cosh(\beta\,h_i(s))}$$

Synchronous updates; which can yet be varied in several ways (here in two ways)

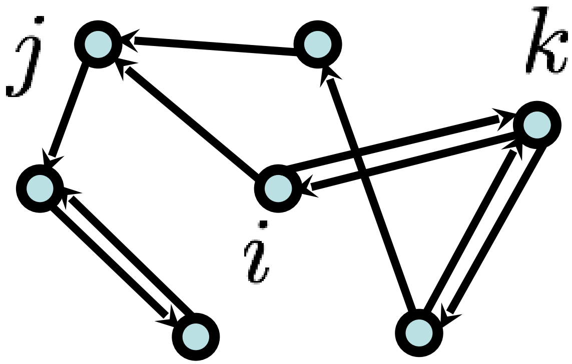
$$h_i(t) = \sum_{j\in\partial i} c_{ji}\,J_{ji}\,\sigma_j(t-1) + \theta_i(t)$$

- parallel : simultaneously update all spins
- sequential : one spin updated at a time

$$\begin{aligned}\sigma_i(t+1) &= +1 \quad \text{with probability} \quad \{1+\exp(2\beta\,h_i(t))\}^{-1}\\ \sigma_i(t+1) &= -1 \quad \text{with probability} \quad \{1+\exp(-2\beta\,h_i(t))\}^{-1}\end{aligned}$$

KTH/CSC

Dilution, asymmetry, and interaction strength



$$\text{connectivity} \quad c_{ij} \quad (i \longrightarrow j)$$

$$p(c_{ij}) = \frac{c}{N}\delta_{c_{ij},1} + (1 - \frac{c}{N})\delta_{c_{ij},0}$$

Starts as directed Erdős-Renyi graphs; where average connectivity would be ***c***

$$\text{connectivivty} \quad \rightarrow \quad c$$

$$\text{asymmetry} \quad \rightarrow \quad \epsilon$$

Connections $i \rightarrow j$ and $j \rightarrow i$ dependent

Derrida (1987) parametrization

$$p(c_{ji}|c_{ij}) = \epsilon\,\delta_{c_{ij},c_{ji}} + (1-\epsilon)\,p(c_{ji})$$

Gaussian or binary J_{ij} $\Longrightarrow$ $h_i(t) = \sum_{j \in \partial i} c_{ji}\,J_{ji}\,\sigma_j(t-1) + \theta_i(t)$

KTH/CSC

Mean-field theory for magnetizations and correlations

$$p(\sigma_i(0),\ldots,\sigma_i(t)) = \sum_{\vec{\sigma}_{\setminus i}(0),\ldots,\vec{\sigma}_{\setminus i}(t)} p(\vec{\sigma}(0),\ldots,\vec{\sigma}(t)) \longrightarrow m_i(t)$$

$$p_{ij}(\sigma_i(0),\ldots\sigma_i(t),\sigma_j(0),\ldots,\sigma_j(t')) = \sum_{\vec{\sigma}_{\setminus i,j}(0),\ldots,\vec{\sigma}_{\setminus i,j}(t)} p(\vec{\sigma}(0),\ldots,\vec{\sigma}(t)) \longrightarrow c_{ij}(t,t')$$

$$\frac{d\,m_i(t)}{dt} = -m_i(t) + \langle \tanh(\beta H_i(t)) \rangle$$

$$\frac{d\,c_{ij}(t,t')}{dt} = -c_{ij}(t,t') + \langle \tanh(\beta \sigma_j(t') H_i(t)) \rangle$$

Crisanti, Sompolinski (1988)
Kappen, Spanjers (2000)

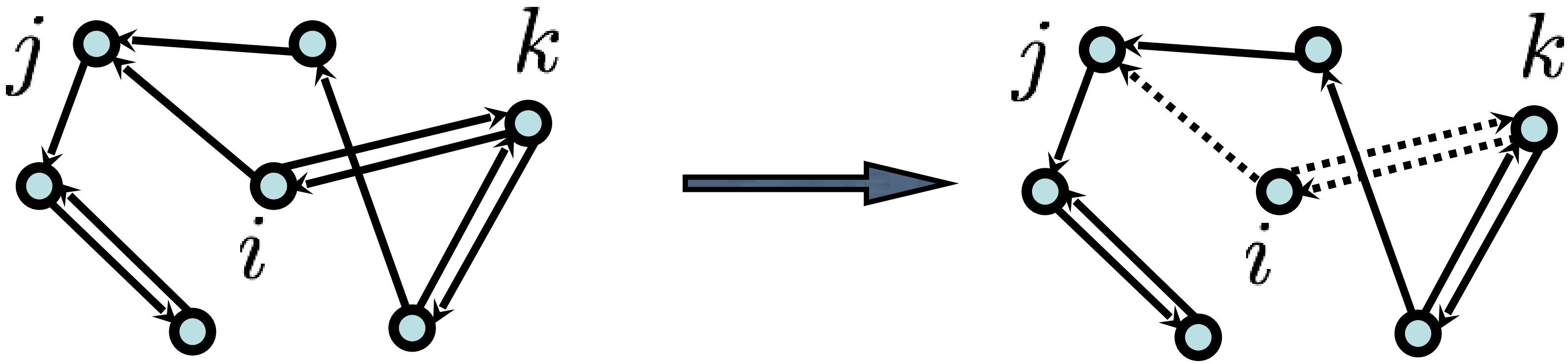
Hertz et al (2010)
Hertz, Roudi arXiv:1009.5946
Zeng et al arXiv:1011.6216

nMF $\quad m_i = \tanh\left[\beta(\sum_j J_{ji} m_j + \theta_i)\right]$

TAP $\quad m_i = \tanh\left[\beta(\sum_j J_{ji} m_j + \theta_i - m_i \sum_j J_{ji}^2 (1-m_j)^2)\right]$

Bethe - Peierls approximation for spin histories

Cavity graph



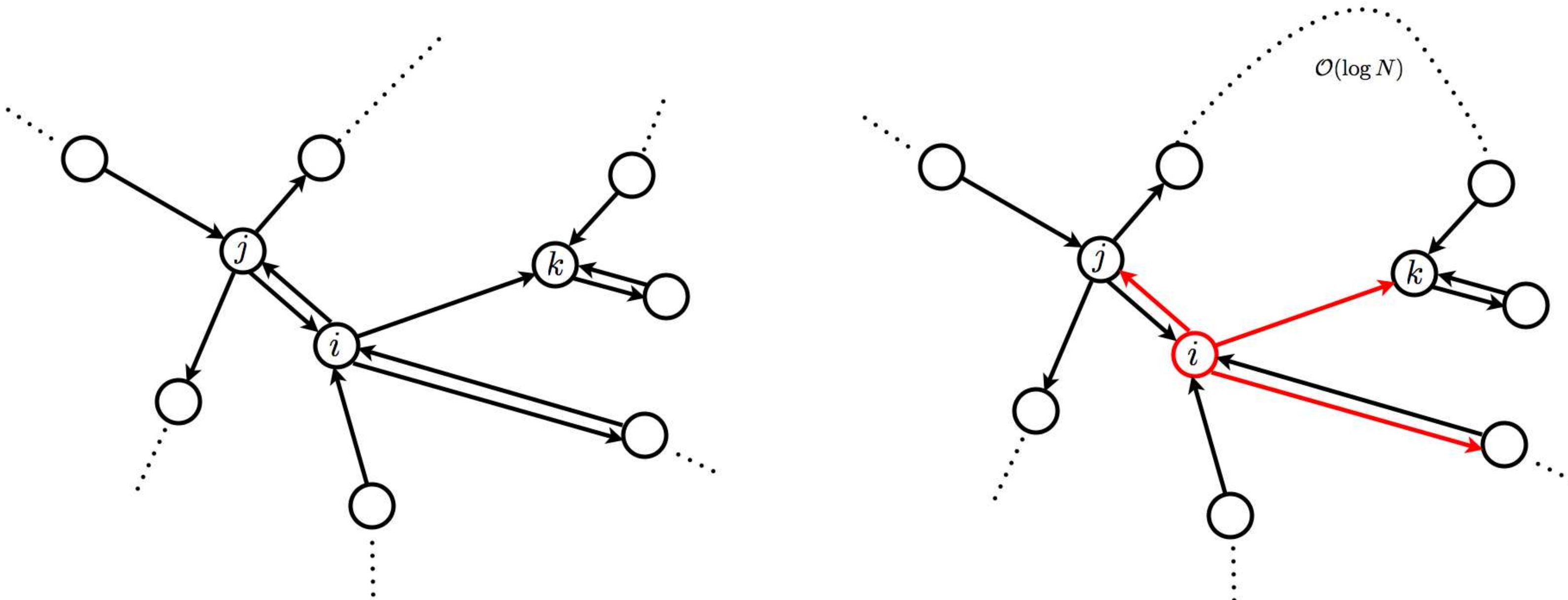
$$p(\vec{\sigma}(0)...\vec{\sigma}(t)) = p^{(i)}(\vec{\sigma}(0),...,\vec{\sigma}(t)\,|\sigma_i(0),...,\sigma_i(t)) \quad p_i(\sigma_i(0),...,\sigma_i(t))$$

The cavity assumption: spins in cavity independent

$$p^i(\vec{\sigma}(0),...,\vec{\sigma}(t)|\sigma_i(0),...,\sigma_i(t)) = \prod_{k\in\partial i} m_{k\to i}(\sigma_k(0),...,\sigma_k(t)|\sigma_i(0),...,\sigma_i(t))$$

The "messages" $m_{k\to i}$ are conditional probabilities

The (dynamic) cavity equation (parallel updates)



$$m_{i\to j}(\sigma_i(0),\ldots,\sigma_i(t)) = \sum_{\vec{\sigma}_{\partial i\setminus j}(0),\ldots,\vec{\sigma}_{\partial i\setminus j}(t)} \prod_{k\in\partial i\setminus j} m_{k\to i}(\sigma_k(0),\ldots,\sigma_k(t)|\sigma_i(0),\ldots,\sigma_i(t))$$

$$\prod_{s=1}^{t} \frac{e^{\beta\, h_{i\to j}(s)\sigma_i(s)}}{2\cosh(\beta\, h_{i\to j}(s))}\; m_{i\to j}(\sigma_i(0))$$

In static (equilibrium) case this would be BP,
Here history dependent, hence hard to use directly.

KTH/CSC

The time-factorized approximation (parallel updates)

$$m_{i\to j}(\sigma_i(0), ..., \sigma_i(t)) = \prod_{s=0}^{t} m_{i\to j}(\sigma_i(s))$$

$$m_{i\to j}(\sigma_i(t)) = \sum_{\sigma_i(t-2), \vec{\sigma}_{\partial i\setminus j}(t-1)} \prod_{k\in\partial i\setminus j} m_{k\to i}(\sigma_k(t-1)|\sigma_i(t-2)) \frac{e^{\beta(\sum_{j\in\partial i} J_{ji}\,\sigma_j(t-1)+\theta_i)\sigma_i(t)}}{2\cosh(\beta(\sum_{j\in\partial i} J_{ji}\,\sigma_j(t-1)+\theta_i))} m_{i\to j}(\sigma_i(t-2))$$

I Neri, D Bolle (2009)
Y Kanoria, A Montanari (2009)

(formally) time-dependent magnetization

Aurell, Mahmoudi (2010)

$$m_i(t) = \sum_{\vec{\sigma}_{\partial i}(t-1)} \prod_{j\in\partial i} m_{j\to i}(\sigma_j(t-1)) \tanh\left[\beta(\sum_{j\in\partial i} \frac{c_{ij}}{c} J_{ji}\,\sigma_j(t-1)+\theta_i)\right]$$

The sequential update model (poor man's Master eq.)

$$\frac{dp(\vec{\sigma}(t))}{dt} = \sum_{i=1}^{N} \{w_i\,(F_i\vec{\sigma}(t))\; p(F_i\vec{\sigma}(t)) - w_i(\vec{\sigma})(t)\, p(\vec{\sigma})(t)$$

Fully asynchronous updates (Master equations) is not yet done. In sequential updates one (randomly picked) spin is updated at a time.

$$p(\vec{\sigma}(t)) = \sum_{\vec{\sigma}(t-1)} W(\vec{\sigma}(t)\,|\,\vec{\sigma}(t-1))\, p(\vec{\sigma}(t-1))$$

$$W(\vec{\sigma}(t)\,|\,\vec{\sigma}(t-1)) = \delta_{\vec{\sigma}(t),\vec{\sigma}(t-1)} + \frac{1}{N}\sum_{i=1}^{N} \{w_i(F_i\vec{\sigma}(t))\,\delta_{\vec{\sigma}(t),F_i\vec{\sigma}(t-1)} \;\; - w_i(\vec{\sigma}(t))\,\delta_{\vec{\sigma}(t),\vec{\sigma}(t-1)}\}$$

When *N* goes to infinity, it is reasonable to expect that sequential update tends to fully asynchronous updates. But at finite *N* there will be a difference.

The dynamic cavity equations for sequential updates

The full dynamic cavity equations :

$$
\begin{aligned}
m_{i\to j}(\sigma_i(0),\ldots,\sigma_i(t)) &= \sum_{\vec{\sigma}_{\partial i\setminus j}(0),\ldots,\vec{\sigma}_{\partial i\setminus j}(t)} \prod_{k\in\partial i\setminus j} m_{k\to i}(\sigma_k(0),\ldots,\sigma_k(t)|\sigma_i(0),\ldots,\sigma_i(t)) \\
& \prod_{s=1}^{t}\left[\frac{1}{N}\frac{e^{\beta h_i^{(j)}(s)\sigma_i(s)}}{2\cosh(\beta h_i^{(j)}(s))} + (1-\frac{1}{N})\delta_{\sigma_i(s),\sigma_i(s-1)}\right] m_{i\to j}(\sigma_i(0))
\end{aligned}
$$

The time factorized approximation :

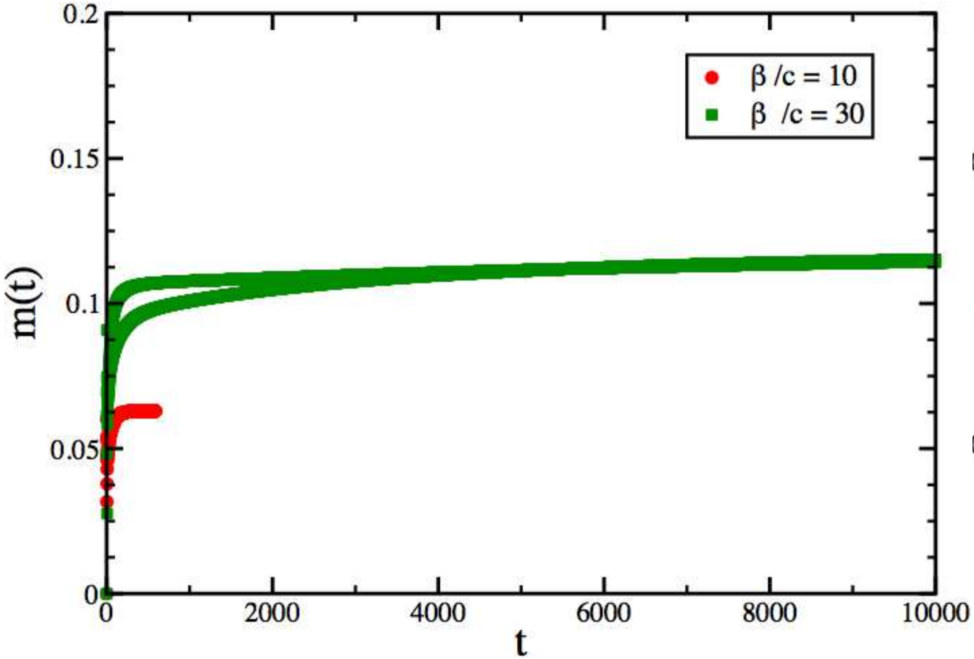
$$
\begin{aligned}
m_{i\to j}(\sigma_i(t)) &= \frac{1}{N}\sum_{\vec{\sigma}_{\partial i\setminus j}(t-1),\sigma_i(t-2)} \prod_{k\in\partial i\setminus j} m_{k\to i}(\sigma_k(t-1)|\sigma_i(t-2)) \\
& \frac{e^{\beta h_i^{(j)}(t)\sigma_i(t)}}{2\cosh(\beta h_i^{(j)}(t))} m_{i\to j}(\sigma_i(t-2)) + (1-\frac{1}{N})\sum_{\sigma_i(t-1)} m_{i\to j}(\sigma_i(t-1))\delta_{\sigma_i(t),\sigma_i(t-1)}
\end{aligned}
$$

Aurell, Mahmoudi (2010)

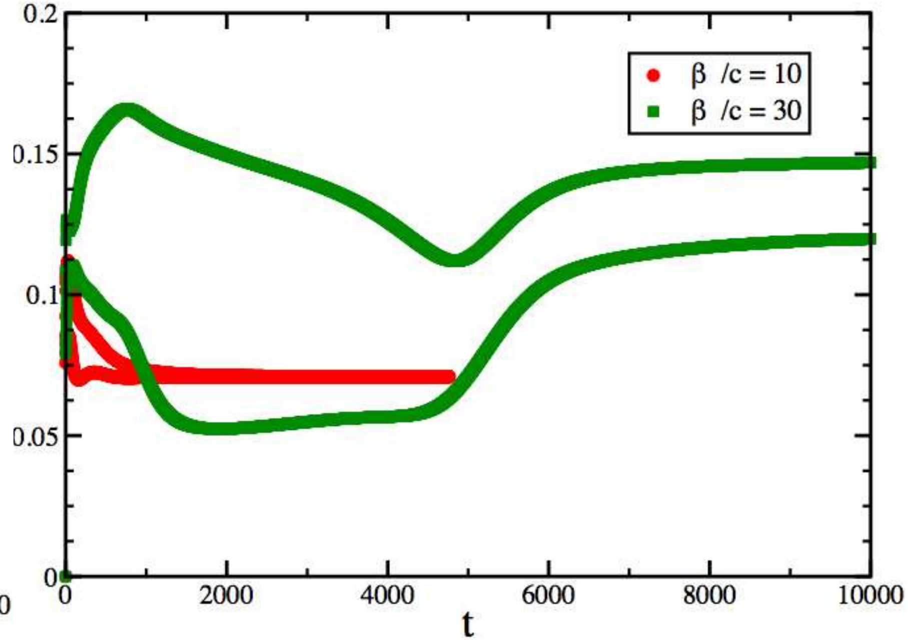
Comparing dynamic cavity equations to Monte Carlo

work in progress…

Parallel updates, different temperatures



fully asymmetric

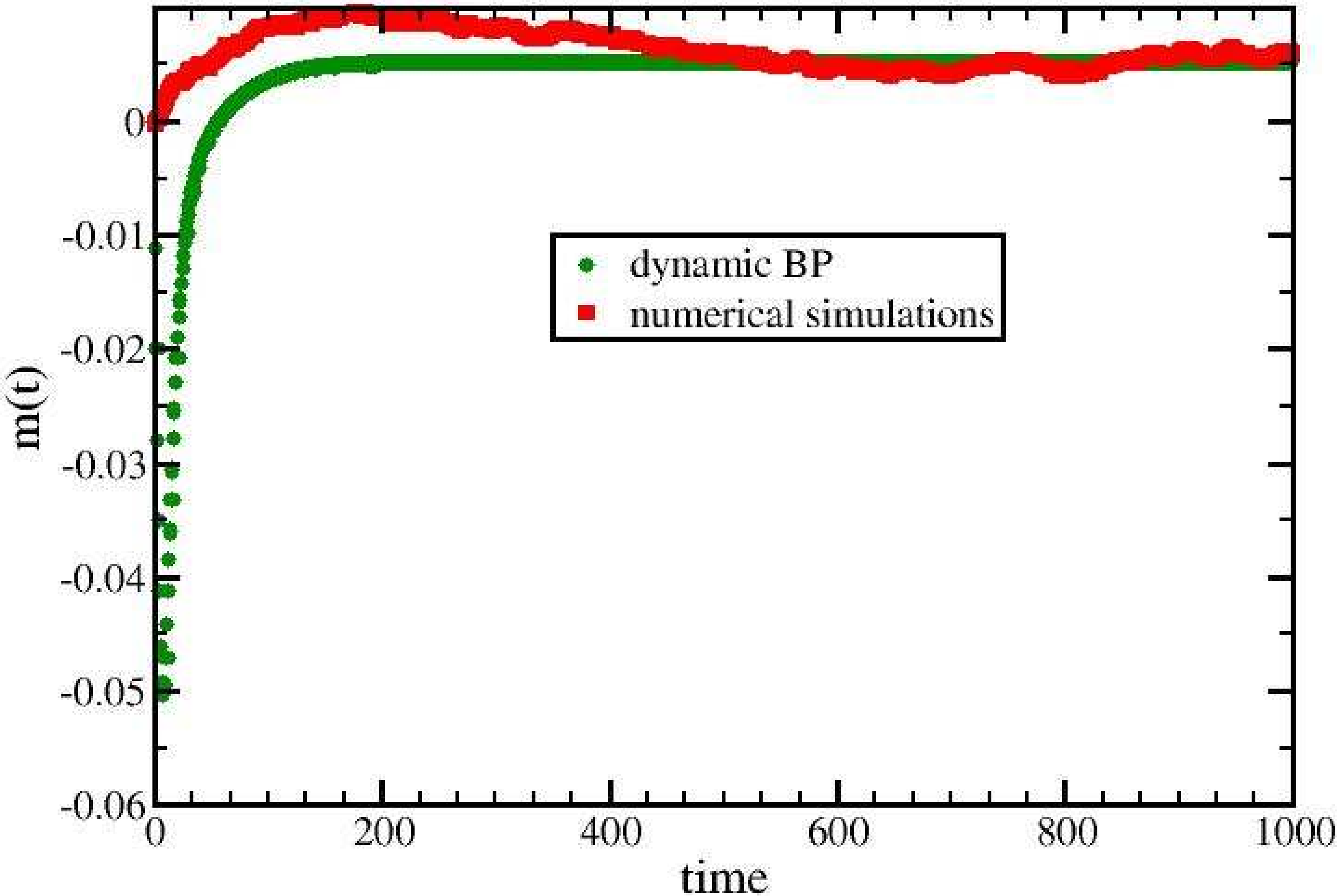


fully symmetric

Average magnetization measured

$$m(t) = \frac{1}{N} \sum_{i=1}^{N} m_i(t)$$

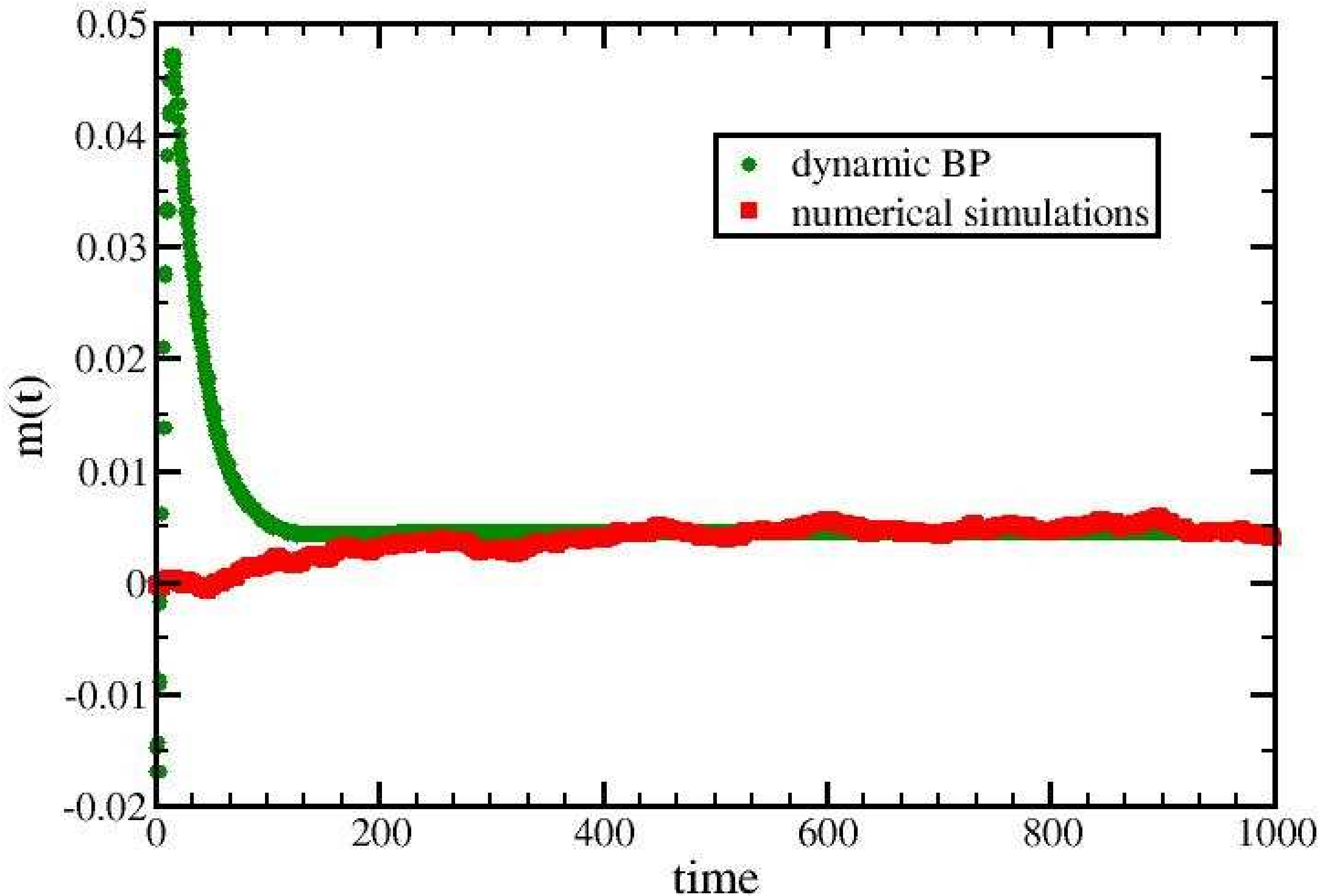
Sequential updates, N=100, β=15, c=3



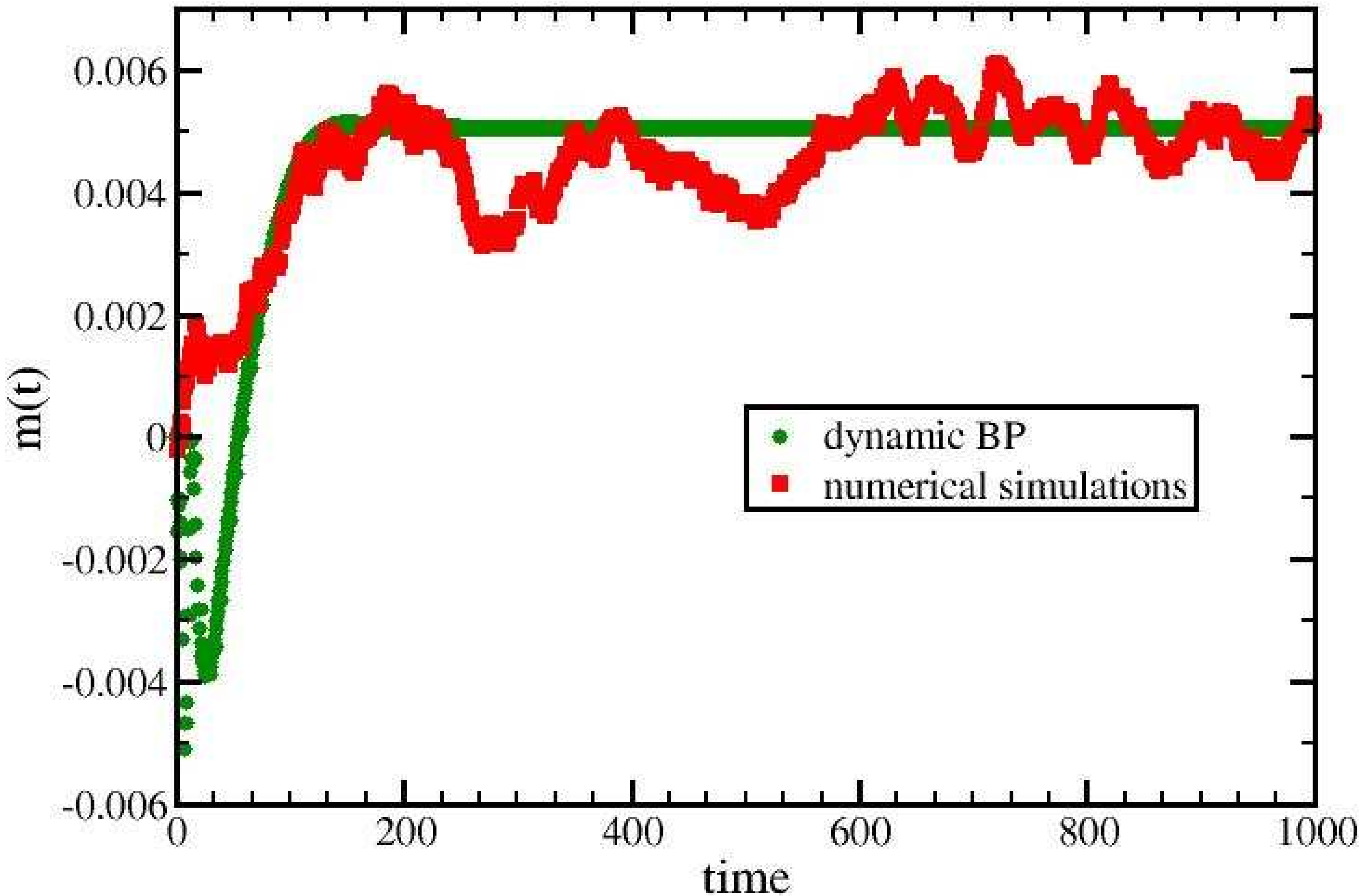
Monte Carlo averaged over very many realizations.
Dynamic cavity method run once.

KTH/CSC

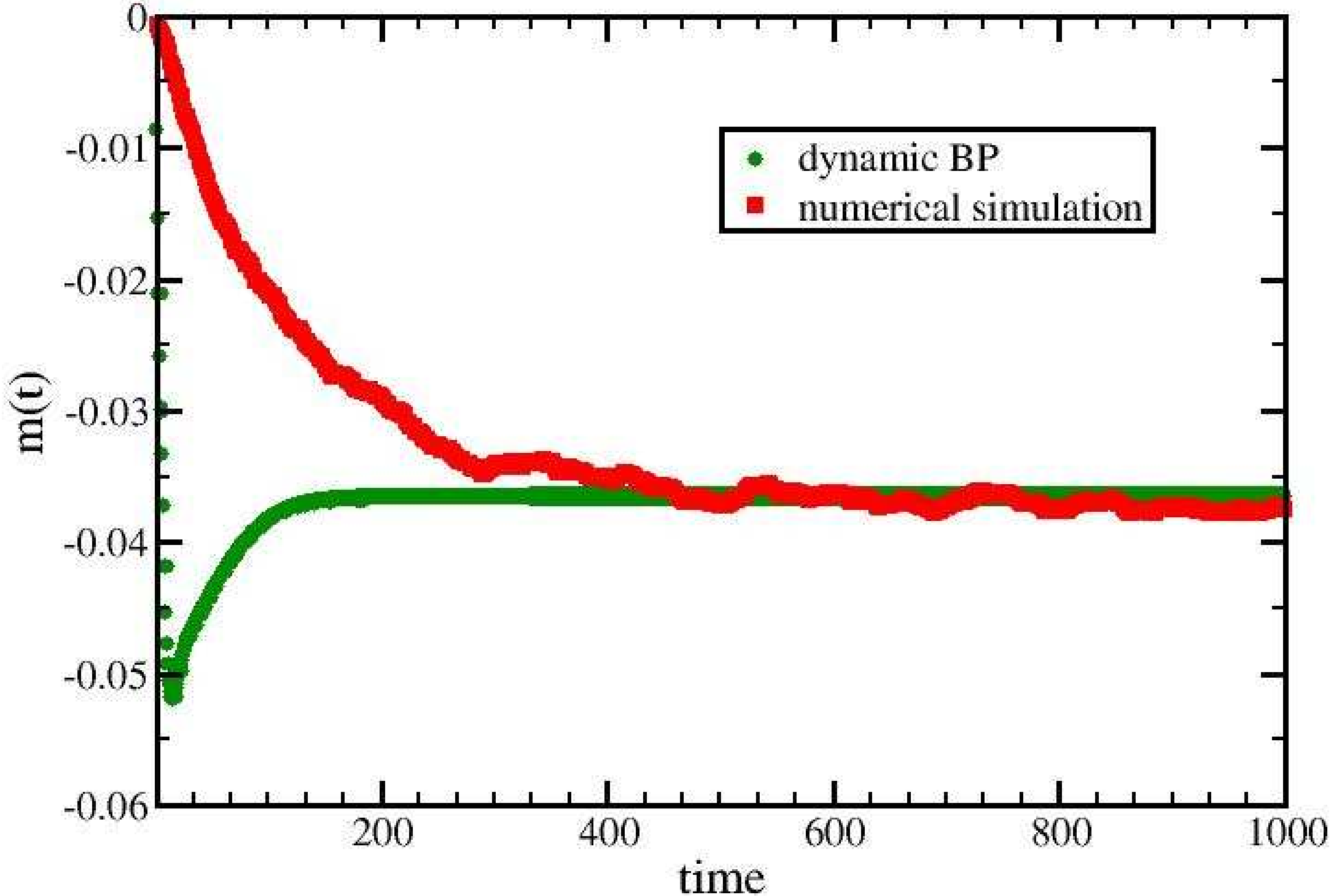
Some more sequential update cases….



Yet some more sequential update cases….



And yet more still.



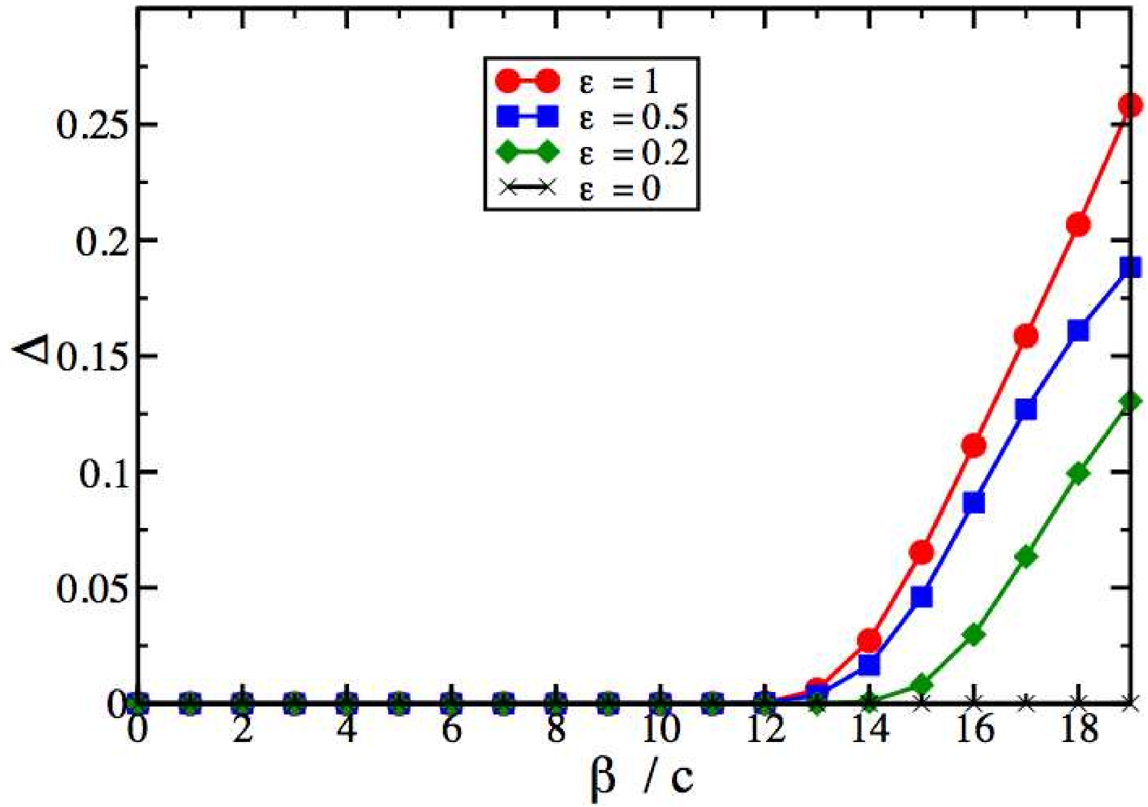
There seems to be something to it.

Therefore…to be continued…

KTH/CSC

Parallel updates can have cyclic behaviour

$$\Delta(t) = \frac{1}{N}\sum_{i=1}^{N}(m_i(t) - m_i(t-1))^2$$



Thanks