Matrix product states for modeling dynamical processes on networks

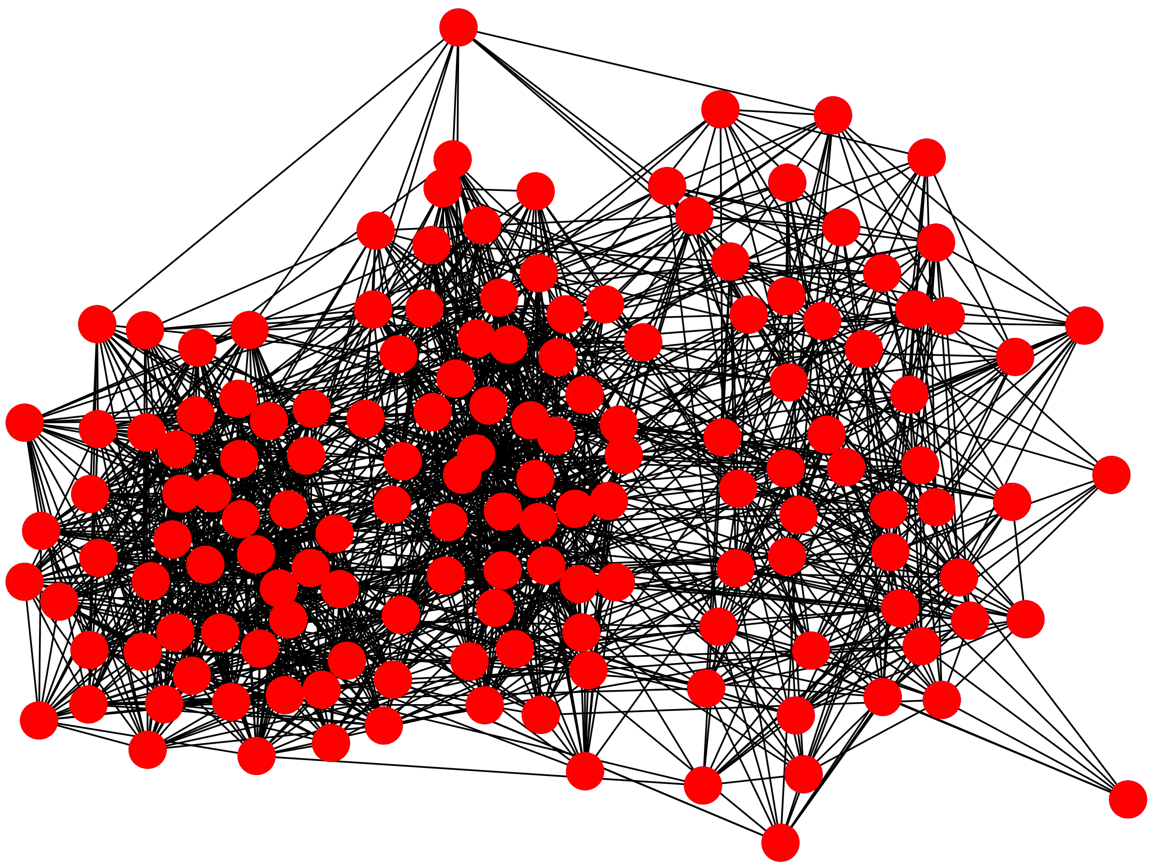
Caterina De Bacco
Cyber Valley, Max Planck Institute for Intelligent Systems

caterina.debacco@tuebingen.mpg.de

Processes on networks
The problem

σ_i **variables on nodes**

$$\sigma_i \in \{blue, red, green\}\ ,\ i = 1, \ldots, N$$



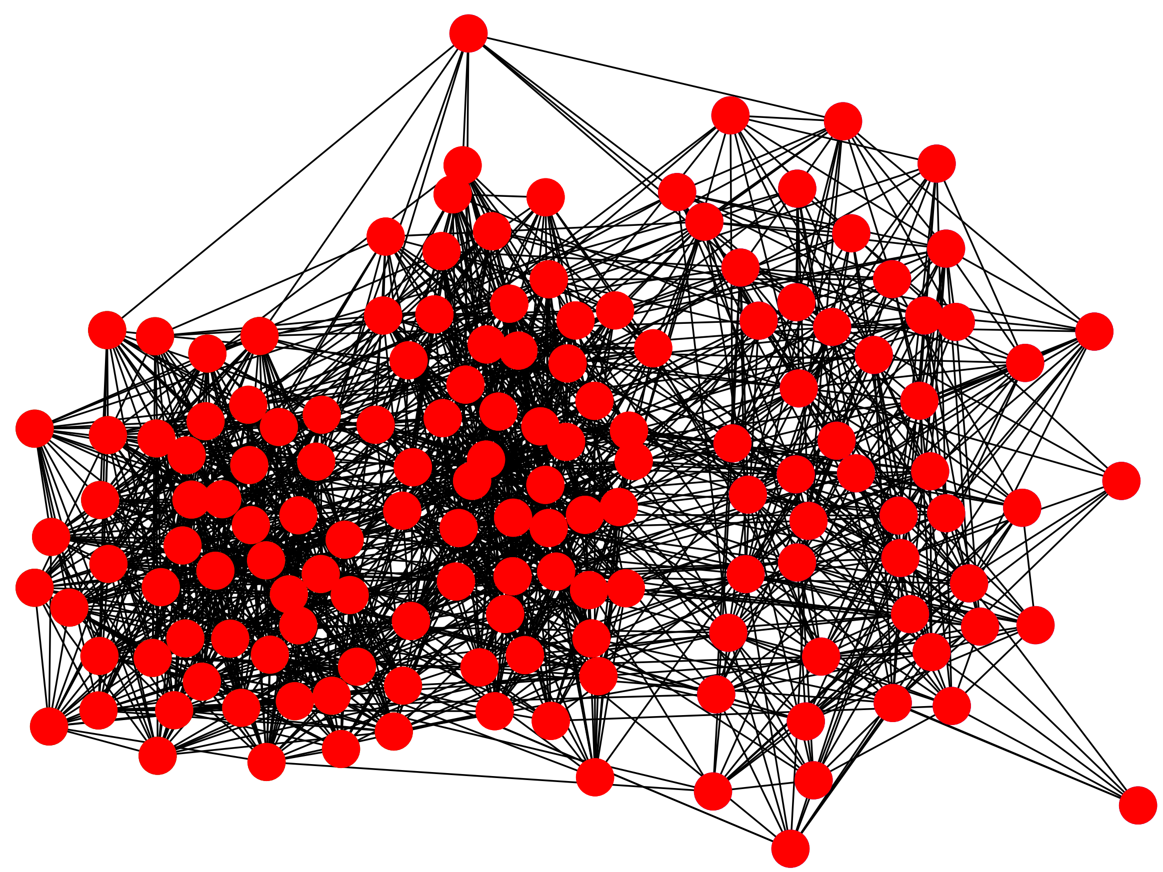
Goal: find the values of σ_i

caterina.debacco@tuebingen.mpg.de

Processes on networks
The problem

σ_i **variables on nodes**

$$\sigma_i \in \{blue, red, green\}\,,\, i = 1, \dots, N$$

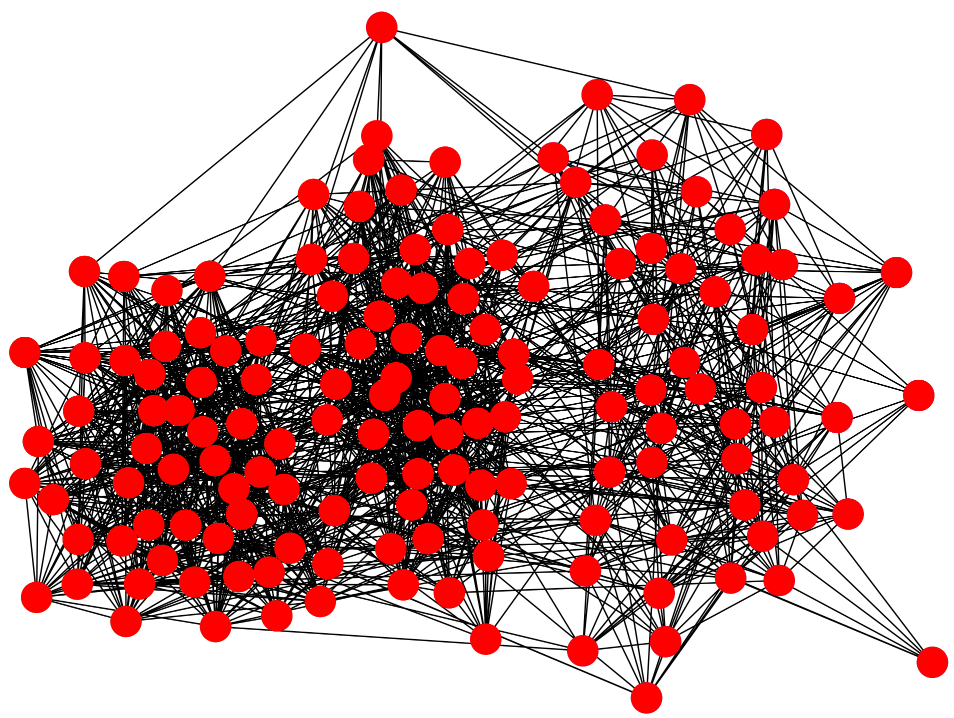


$$P(\sigma|J,C,\gamma) \propto \prod_{i<j} C_{\sigma_i\sigma_j}^{J_{ij}} (1 - C_{\sigma_i\sigma_j})^{(1-J_{ij})} \prod_i \gamma_{\sigma_i}$$

Processes on networks
The problem

σ_i **variables on nodes**

$$\sigma_i \in \{blue, red, green\}\,,\; i = 1, \ldots, N$$



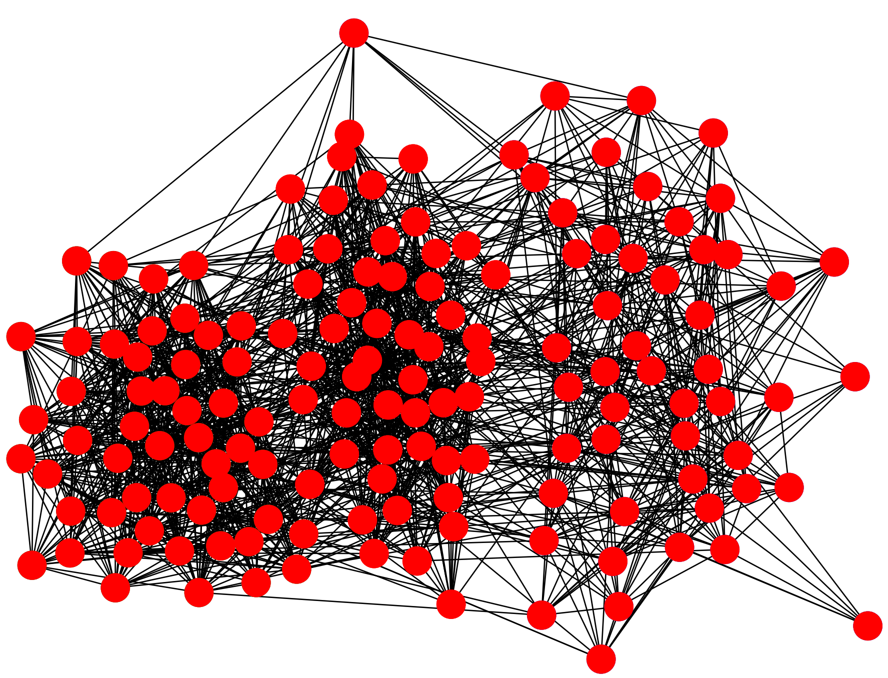
$$P(\sigma|J,C,\gamma) = e^{-H(\sigma|J,C,\gamma)}/Z(J,C,\gamma)$$

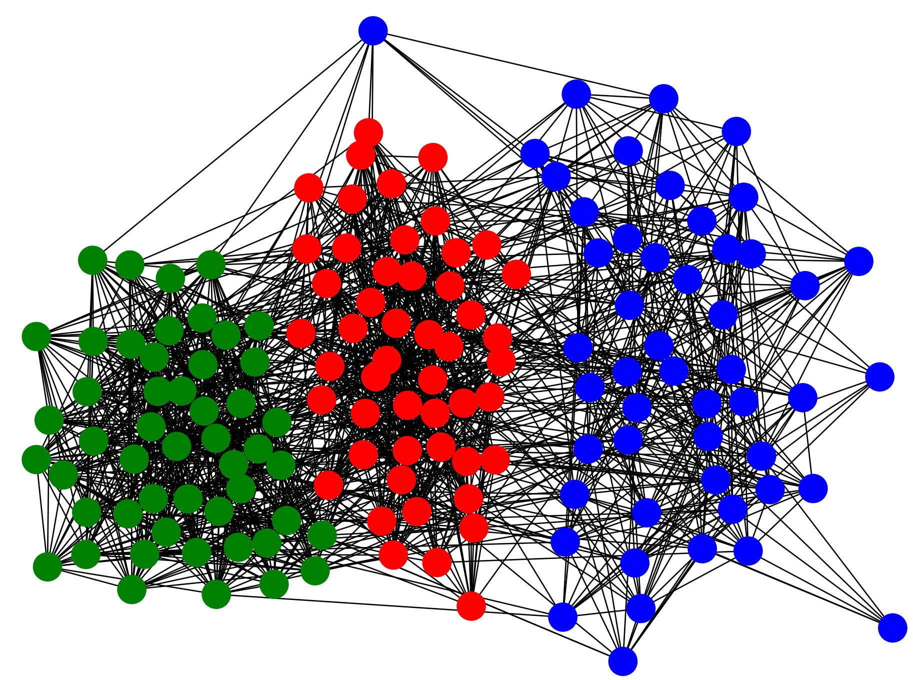
$$H(\sigma|J,C,\gamma) = -\sum_i \log\gamma_i - \sum_{i<j} J_{ij}\log C_{\sigma_i\sigma_j} + (1-J_{ij})\log(1-C_{\sigma_i\sigma_j})$$

Processes on networks
The problem

Goal: find marginal distributions like

$$P(\sigma_i|J,C,\gamma) \qquad\qquad P(\sigma_i,\sigma_j|J,C,\gamma)$$





Processes on networks
The problem

Challenge: performing the marginalization exactly is not possible, **exponential number of configurations**

$$P(\sigma|J,C,\gamma) = e^{-H(\sigma|J,C,\gamma)}/Z(J,C,\gamma) \longrightarrow P(\sigma_i|J,C,\gamma)$$

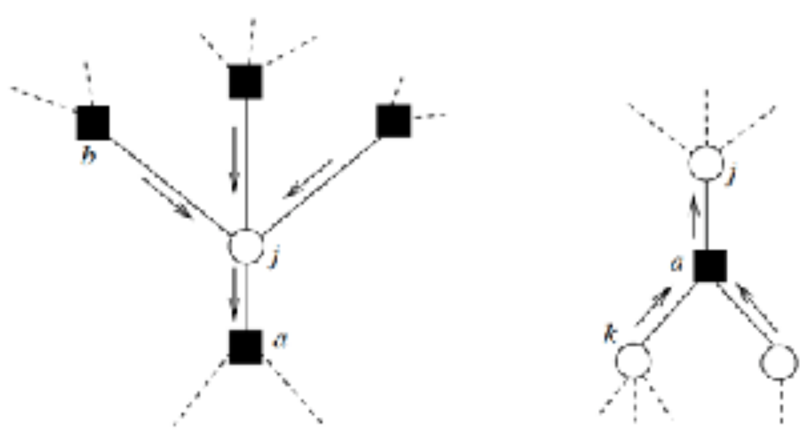
Processes on networks
The message-passing algorithm

MAX-PLANCK-GESELLSCHAFT

- Message-passing (or cavity method, or belief propagation)
- Exact in tree, good approximation on locally tree-like networks

Fast, can be done in linear time in N

$$\mu_{ij}(\sigma_i) \propto \gamma_{\sigma_i} e^{-h_i} \prod_{k\in\partial i\setminus j}\left[\sum_{\sigma_k} C_{\sigma_i\sigma_k}\mu_{ki}(\sigma_k)\right]$$



- Mezard and Parisi 2003
- Yedidia et al 2003
- Kschischang et al 2001
- Kabashima and Saad 1998
- Pearl 1988
- Bethe 1935

caterina.debacco@tuebingen.mpg.de

Processes on networks
The problem

Challenge: performing th[...]ation exactly is not possible,
exponentia[...] of [...]onfigurations

$$P(\sigma|J,C,\gamma) = e^{-H(\sigma|J,C,\gamma ...}Z(J \ldots \longrightarrow P(\sigma_i|J,C,\gamma)$$

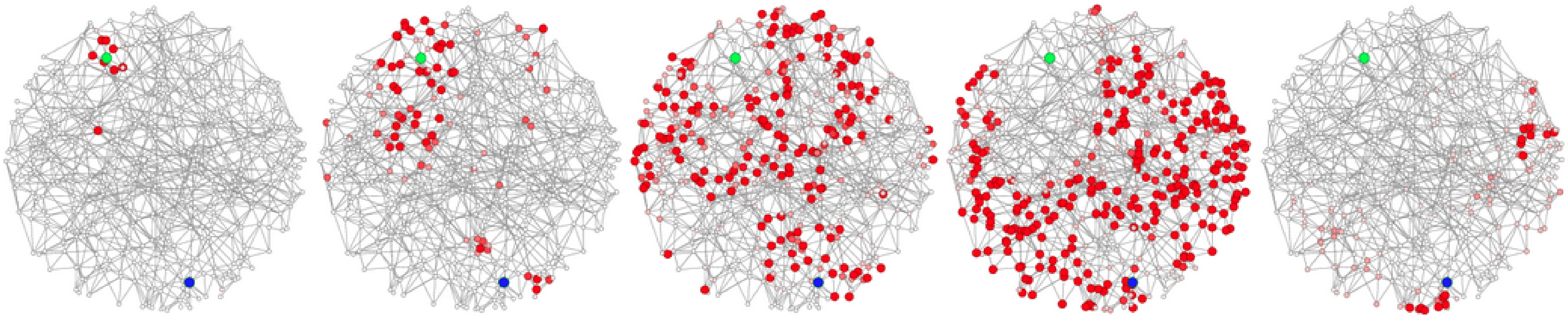
Message-passing

- Mezard and Parisi 2003
- Yedidia et al 2003
- Kschischang et al 2001
- Kabashima and Saad 1998
- Pearl 1988
- Bethe 1935

caterina.debacco@tuebingen.mpg.de

Dynamical Processes on networks
The problem

σ_i^t "on" or "off" depending on the state of its neighbors in time

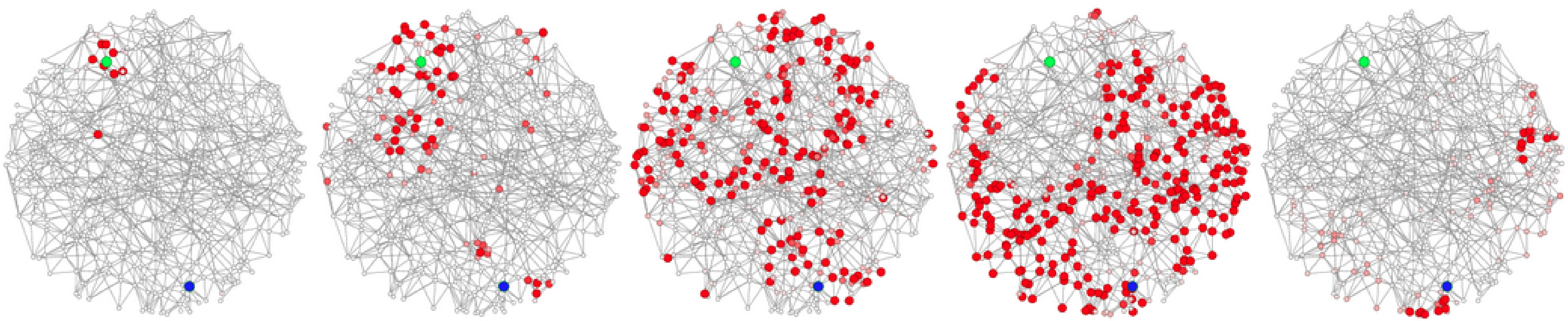


Dynamical Processes on networks

The problem

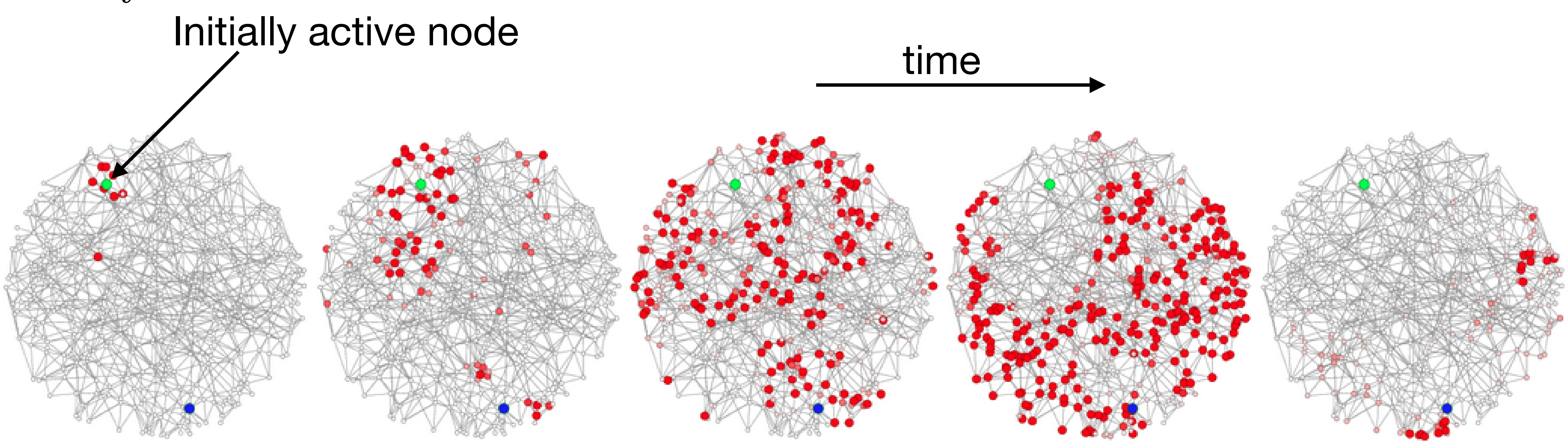
dynamics

σ_i^t "on" or "off" depending on the state of its neighbors in time



Dynamical Processes on networks
The problem

σ_i^t “on” or “off” depending on the state of its neighbors in time



$$W(\sigma_i^{t+1} | \{\sigma_j^t\}_{j \in \partial i}) = \frac{e^{-\beta \sigma_i^{t+1} h_i^t}}{2 \cosh(\beta h_i^t)} \qquad h_i^t = \sum_j J_{ij} \sigma_j^t$$

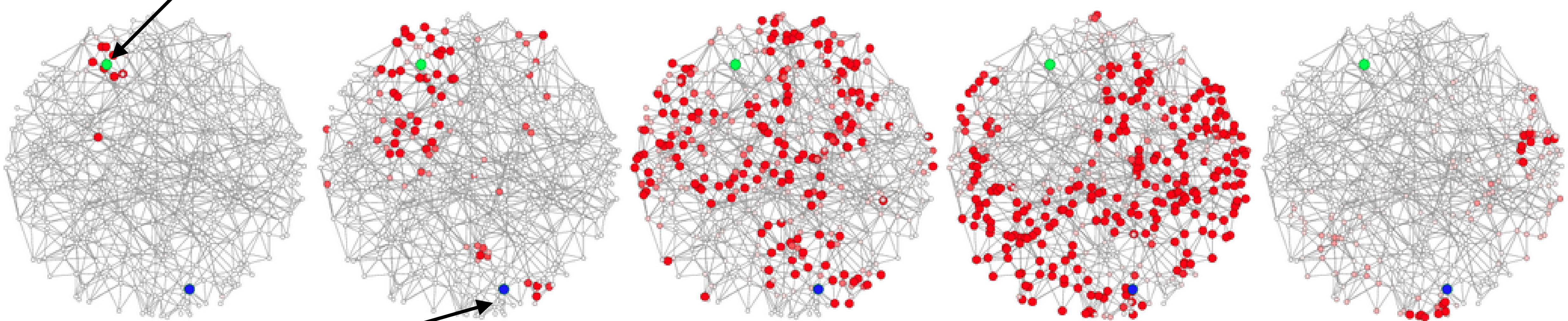
Glauber dynamics, RJ Glauber J Math Phys 1963

Dynamical Processes on networks
The problem

σ_i^t "on" or "off" depending on the state of its neighbors in time

Initially active node

time



is this on or off?

$$W(\sigma_i^{t+1}|\{\sigma_j^t\}_{j\in\partial i}) = \frac{e^{-\beta\sigma_i^{t+1}h_i^t}}{2\cosh(\beta h_i^t)} \qquad h_i^t = \sum_j J_{ij}\sigma_j^t$$

Glauber dynamics, RJ Glauber J Math Phys 1963

Dynamical Processes on networks

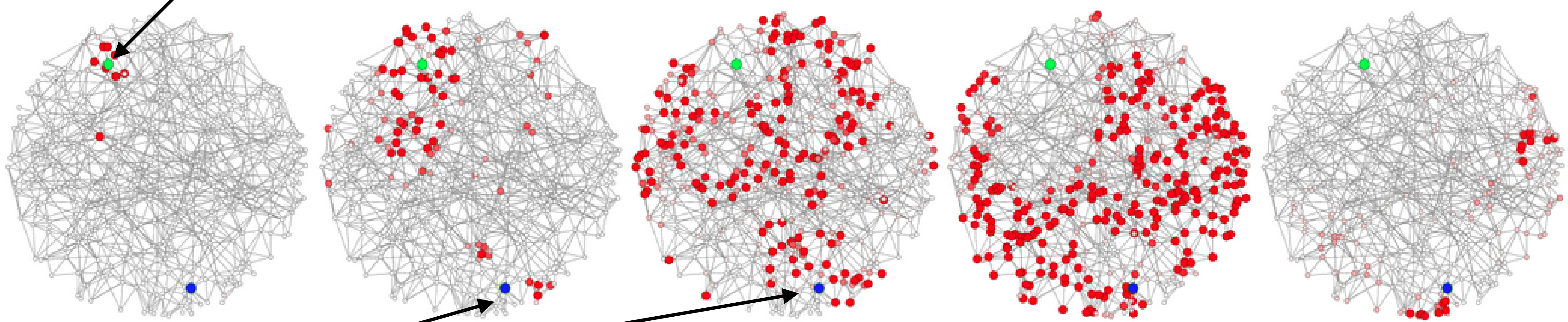
The problem

MAX-PLANCK-GESELLSCHAFT

σ_i^t "on" or "off" depending on the state of its neighbors in time

Initially active node

time



is this on or off?

$$W(\sigma_i^{t+1}|\{\sigma_j^t\}_{j\in\partial i}) = \frac{e^{-\beta\sigma_i^{t+1}h_i^t}}{2\cosh(\beta h_i^t)} \qquad h_i^t = \sum_j J_{ij}\sigma_j^t$$

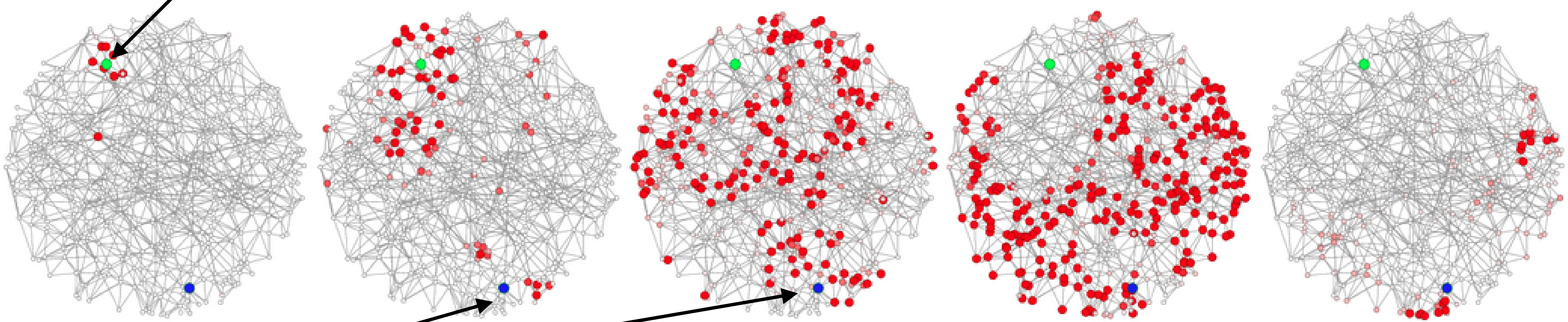
Glauber dynamics, RJ Glauber J Math Phys 1963

Dynamical Processes on networks
The problem

σ_i^t "on" or "off" depending on the state of its neighbors in time

Initially active node

time



is this on or off?

$$P_i(\sigma_i^t)$$ marginal probability ?

caterina.debacco@tuebingen.mpg.de

Dynamical Processes on networks
The problem

$$W(\sigma_i^{s+1} | \sigma_i^s, \{\sigma_j^s\}_{j \in \partial i})$$ Transition probability

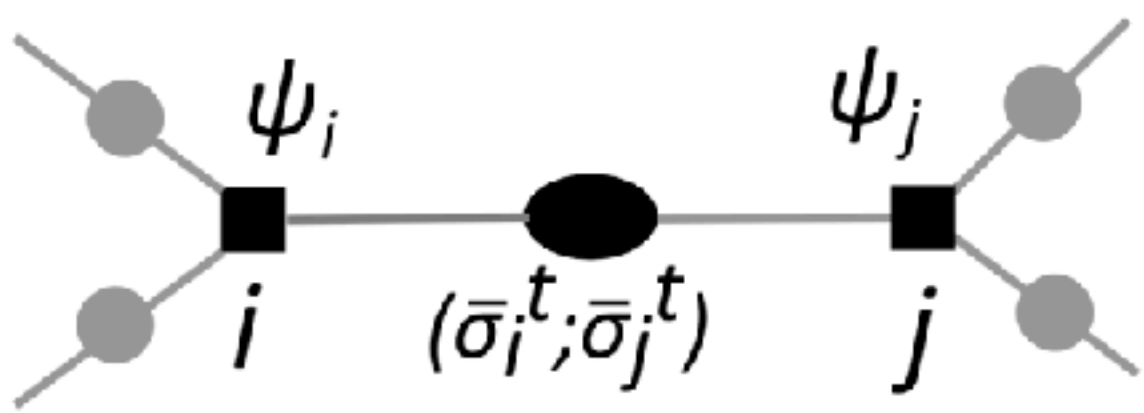
$$P(\bar{\sigma}^t) = P_0(\{\sigma_i^0\}) \prod_{i=1}^{N} \prod_{s=0}^{t-1} W(\sigma_i^{s+1} | \sigma_i^s, \{\sigma_j^s\}_{j \in \partial i})$$

$\bar{\sigma}^t$ is a Nt-dimensional vector denoting the N-variable configuration of time-trajectories

Dynamical Processes on networks
The problem

In analogy with the static case, one can iteratively calculate the messages as:

$$\mu_{ij}(\bar{\sigma}_i^t|\bar{\sigma}_j^{t-1}) = \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t-1} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \prod_{k\in\partial i\setminus j} \mu_{ki}(\bar{\sigma}_k^{t-1}|\bar{\sigma}_i^{t-2})$$



Then the marginal in the original graph is:

$$P_i(\bar{\sigma}_i^t) = \frac{1}{Z} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t-1} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \prod_{k\in\partial i} \mu_{ki}(\bar{\sigma}_k^{t-1};\bar{\sigma}_i^{t-2})$$

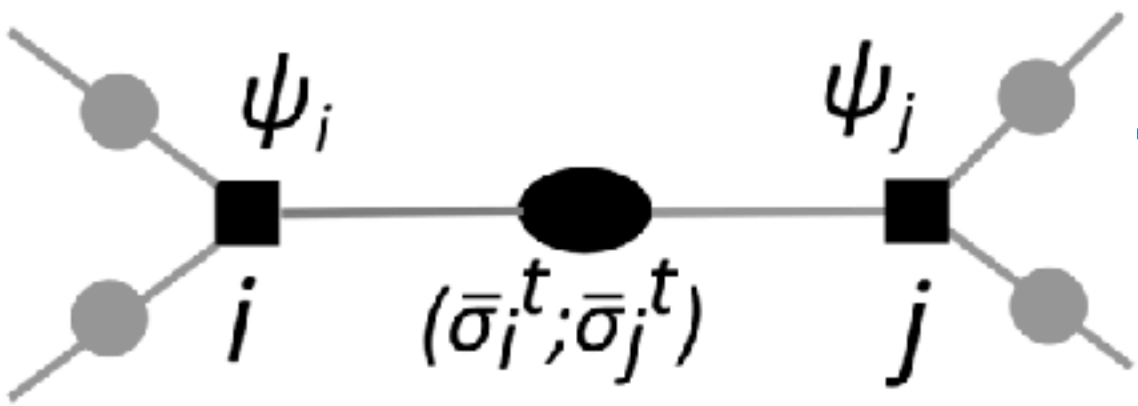
caterina.debacco@tuebingen.mpg.de

Dynamical Processes on networks

The problem

In analogy with the static case, one can iteratively calculate the messages as:

$$\mu_{ij}(\bar{\sigma}_i^t|\bar{\sigma}_j^{t-1}) = \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t-1} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \prod_{k\in\partial i\setminus j} \mu_{ki}(\bar{\sigma}_k^{t-1}|\bar{\sigma}_i^{t-2})$$



This can be done linearly in N!

Then the marginal in the original graph is:

$$P_i(\bar{\sigma}_i^t) = \frac{1}{Z} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t-1} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \prod_{k\in\partial i} \mu_{ki}(\bar{\sigma}_k^{t-1};\bar{\sigma}_i^{t-2})$$

Exponential in t!

Problem.

The single trajectory $\bar{\sigma}_i^t$ can take d^t possible values, where d is the number of values that $\sigma_i(t)$ can take, for example $d = 2$ in the case of Ising spin.

This implies that solving the cavity equation takes an exponential number of operations
$\rightarrow$ need to find an approximation to lower the complexity

Dynamical Processes on networks
The problem

- Approximations: Mean-field, One-time approximation → factorization in time (Neri and Bollé 2009), E. Aurell and H. Mahmoudi 2011,2012); n-step Markov chain (G. Del Ferraro and E. Aurell 2015)
- Unidirectional dynamics: only need to take track of instant time steps when a transition happens → polynomial complexity (SIR, threshold models, etc...) (B. Karrer and M. Newman 2010, F. Altarelli et al. 2013, 2014..., A. Lokhov et al. 2015)

Matrix Product State and message-passing

The ansatz: matrix product edge messages (MPEM)

MAX-PLANCK-GESELLSCHAFT

We propose the Matrix Product ansatz for Dynamical Cavity:

$$\mu_{ij}(\bar{\sigma}_i^t | \bar{\sigma}_j^{t-1}) =$$

$$A_{ij}^{(t+1)}(\sigma_i^t) A_{ij}^{(t)}(\sigma_i^{t-1}) \left[\prod_{s=1}^{t} A_{ij}^{(s)}(\sigma_i^{s-1} | \sigma_j^s) \right] A_{ij}^{(0)}(\sigma_j^0)$$

where the quantities A's have the following shape: $A_{ij}^{(t+1)}(\sigma_i^t)$ is a $1 \times M$ matrix; $A_{ij}^{(t)}(\sigma_i^{t-1})$ and all the $A_{ij}^{(s)}(\sigma_i^{s-1}|\sigma_j^s)$ are $M \times M$ matrices; $A_{ij}^{(0)}(\sigma_j^0)$ is a $M \times 1$ matrix.

Matrix Product State and message-passing

The ansatz: matrix product edge messages (MPEM)

Iterative consistent cavity equation:

$$\begin{aligned}
\mu_{ij}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^{t}) &= B_{ij}^{(t+2)}(\sigma_i^{t+1}) B_{ij}^{(t+1)}(\sigma_i^{t}) \left[\prod_{s=1}^{t} B_{ij}^{(s)}(\sigma_i^{s-1}|\sigma_j^{s})\right] B_{ij}^{(0)}(\sigma_j^{0}) \\
&= \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t} W(\sigma_i^{s+1}|\sigma_i^{s}, \{\sigma_j^{s}\}_{j\in\partial i}) \times \\
&\times \prod_{k\in\partial i\setminus j} \left\{ A_{ki}^{(t+1)}(\sigma_k^{t}) A_{ki}^{(t)}(\sigma_k^{t-1}) \left[\prod_{s=1}^{t} A_{ki}^{(s)}(\sigma_k^{s-1}|\sigma_i^{s})\right] A_{ki}^{(0)}(\sigma_i^{0}) \right\}
\end{aligned}$$

Valid for synchronous parallel update

Matrix Product State and message-passing

The ansatz: matrix product edge messages (MPEM)

MAX-PLANCK-GESELLSCHAFT

Iterative consistent cavity equation:

$$\mu_{ij}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^{t}) = B_{ij}^{(t+2)}(\sigma_i^{t+1})B_{ij}^{(t+1)}(\sigma_i^{t})\left[\prod_{s=1}^{t} B_{ij}^{(s)}(\sigma_i^{s-1}|\sigma_j^{s})\right] B_{ij}^{(0)}(\sigma_j^{0})$$

$$= \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \times$$

$$\times \prod_{k\in\partial i\setminus j} \left\{ A_{ki}^{(t+1)}(\sigma_k^t) A_{ki}^{(t)}(\sigma_k^{t-1}) \left[\prod_{s=1}^{t} A_{ki}^{(s)}(\sigma_k^{s-1}|\sigma_i^{s})\right] A_{ki}^{(0)}(\sigma_i^0) \right\}$$

Valid for synchronous parallel update

Up to here, this is only a re-parametrization, complexity is still exponential

Matrix Product State and message-passing

The ansatz: matrix product edge messages (MPEM)

Iterative consistent cavity equation:

$$\mu_{ij}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^{t}) = B_{ij}^{(t+2)}(\sigma_i^{t+1}) B_{ij}^{(t+1)}(\sigma_i^{t}) \left[\prod_{s=1}^{t} B_{ij}^{(s)}(\sigma_i^{s-1}|\sigma_j^{s}) \right] B_{ij}^{(0)}(\sigma_j^{0})$$

$$= \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t} W(\sigma_i^{s+1}|\sigma_i^s, \{\sigma_j^s\}_{j\in\partial i}) \times$$

$$\times \prod_{k\in\partial i\setminus j} \left\{ A_{ki}^{(t+1)}(\sigma_k^t) A_{ki}^{(t)}(\sigma_k^{t-1}) \left[\prod_{s=1}^{t} A_{ki}^{(s)}(\sigma_k^{s-1}|\sigma_i^{s}) \right] A_{ki}^{(0)}(\sigma_i^0) \right\}$$

Valid for synchronous parallel update

Up to here, this is only a re-parametrization, complexity is still exponential

How do we make it scalable?

Matrix Product State and message-passing
Lower the complexity

MAX-PLANCK-GESELLSCHAFT

1. Singular Value Decomposition of the matrices A's

$$A = U \Lambda V^\dagger$$

where $U^\dagger U = V^\dagger V = \mathbf{1}$ and Λ a diagonal one made of the singular values s_i, with $i = 1, \dots, M$.

2. Truncation of the smallest singular values $\rightarrow$ new dim is $\tilde{M} \leq M$
3. One possible way to estimate the error is :

$$||A - \tilde{A}||_2 = \sqrt{\sum_{i > \tilde{M}} s_i^2}$$

where $||A||_2 = \sqrt{\sum_{ij} A_{ij}^2}$ is the Frobenious norm of the matrix A

Matrix Product State and message-passing

Lower the complexity

MAX-PLANCK-GESELLSCHAFT

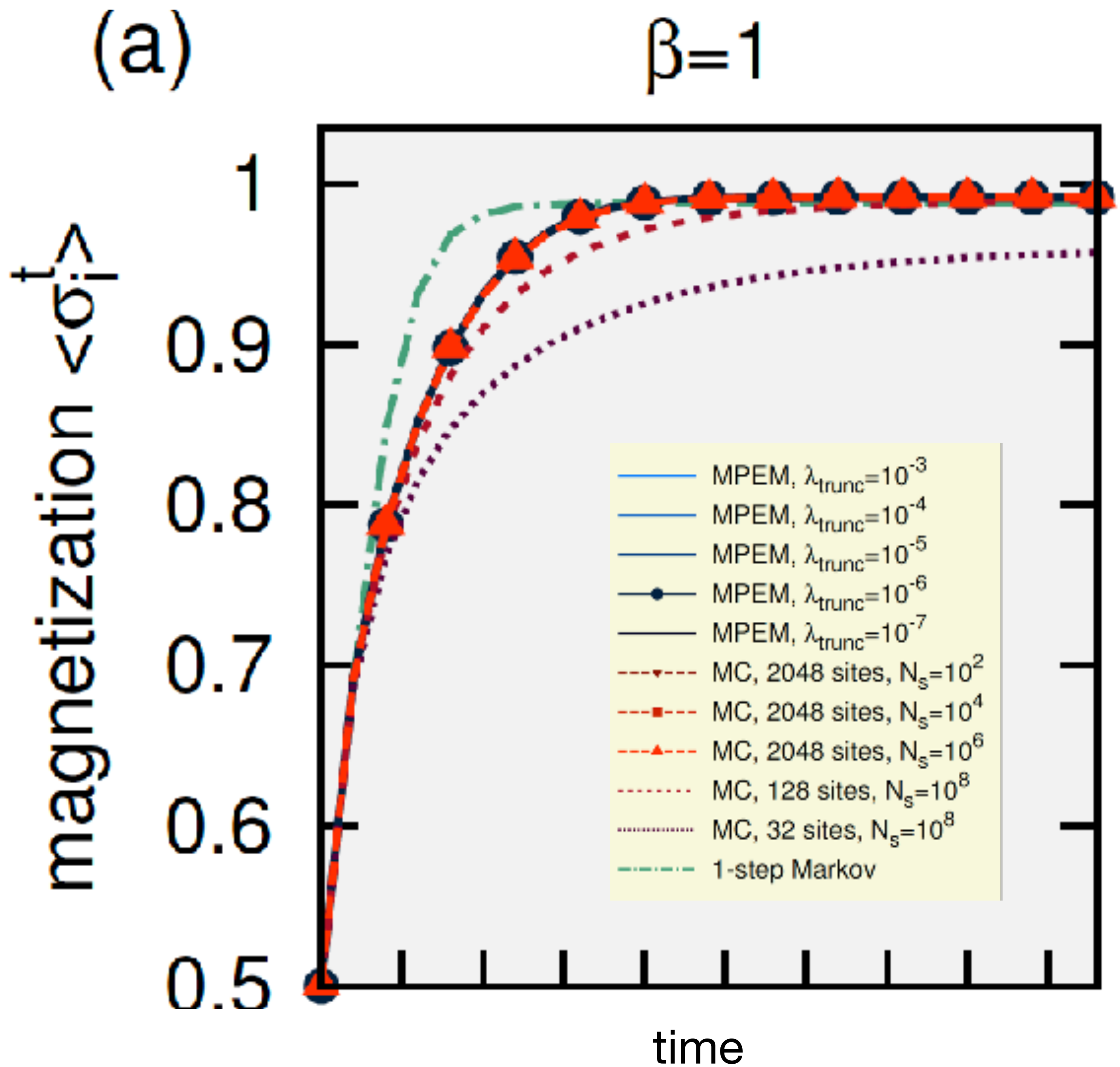
Two truncation criteria:

1. Fixed dimension $\tilde{M} = L$
2. Control the norm. Choose $\tilde{M}$ such that:

$$||A - \tilde{A}||_2 = \sqrt{\sum_{i > \tilde{M}} s_i^2} < L_{max}$$

Matrix Product State and message-passing

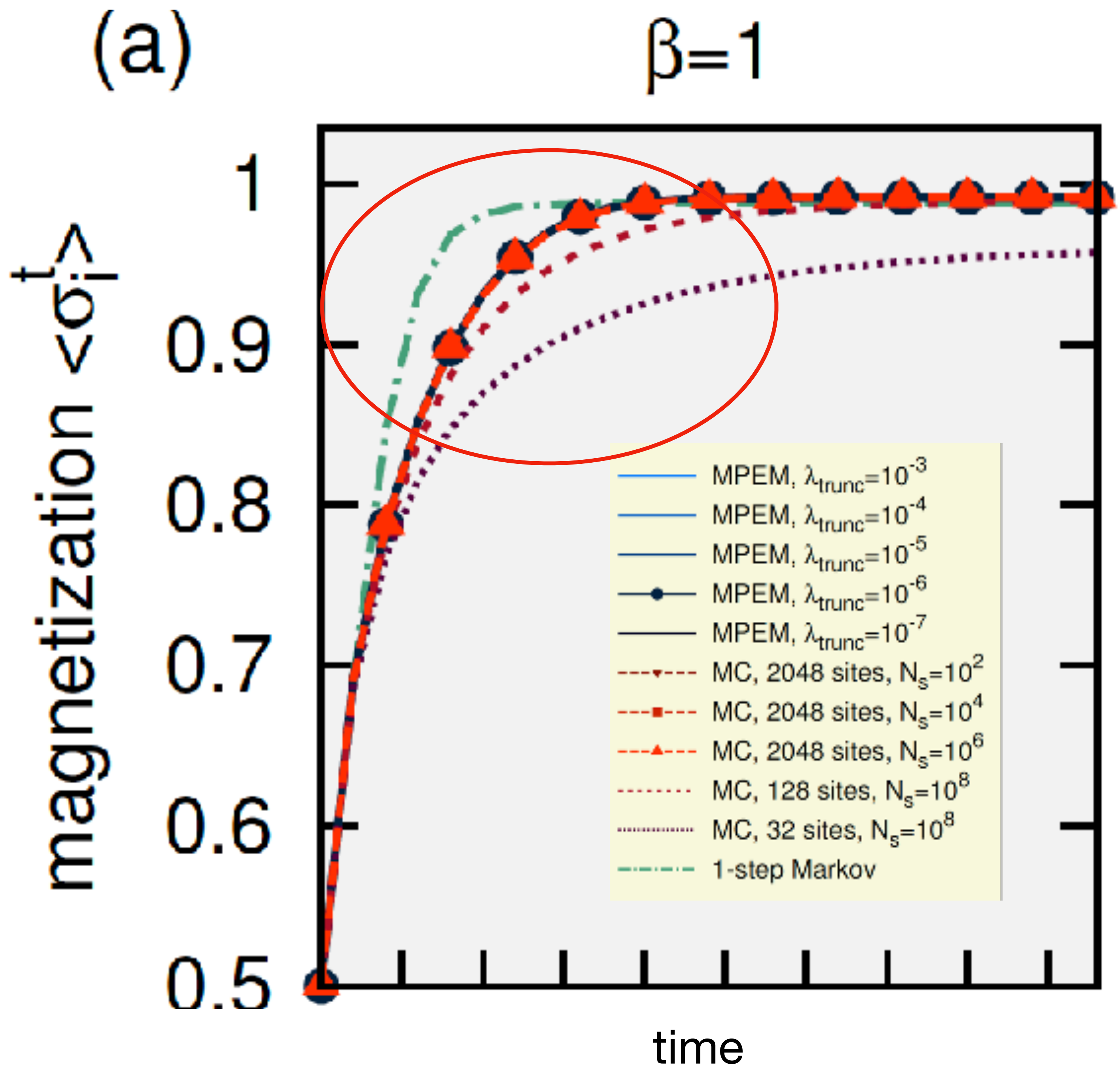
Results on Glauber dynamics



- **Out of equilibrium** we see most discrepancies
- MPEM does well at **every t**
- MCMC has **finite-size** problems

Regular regular graphs z=3
J=1 (ferromagnet),
thermodynamic limit

Matrix Product State and message-passing
Results on Glauber dynamics

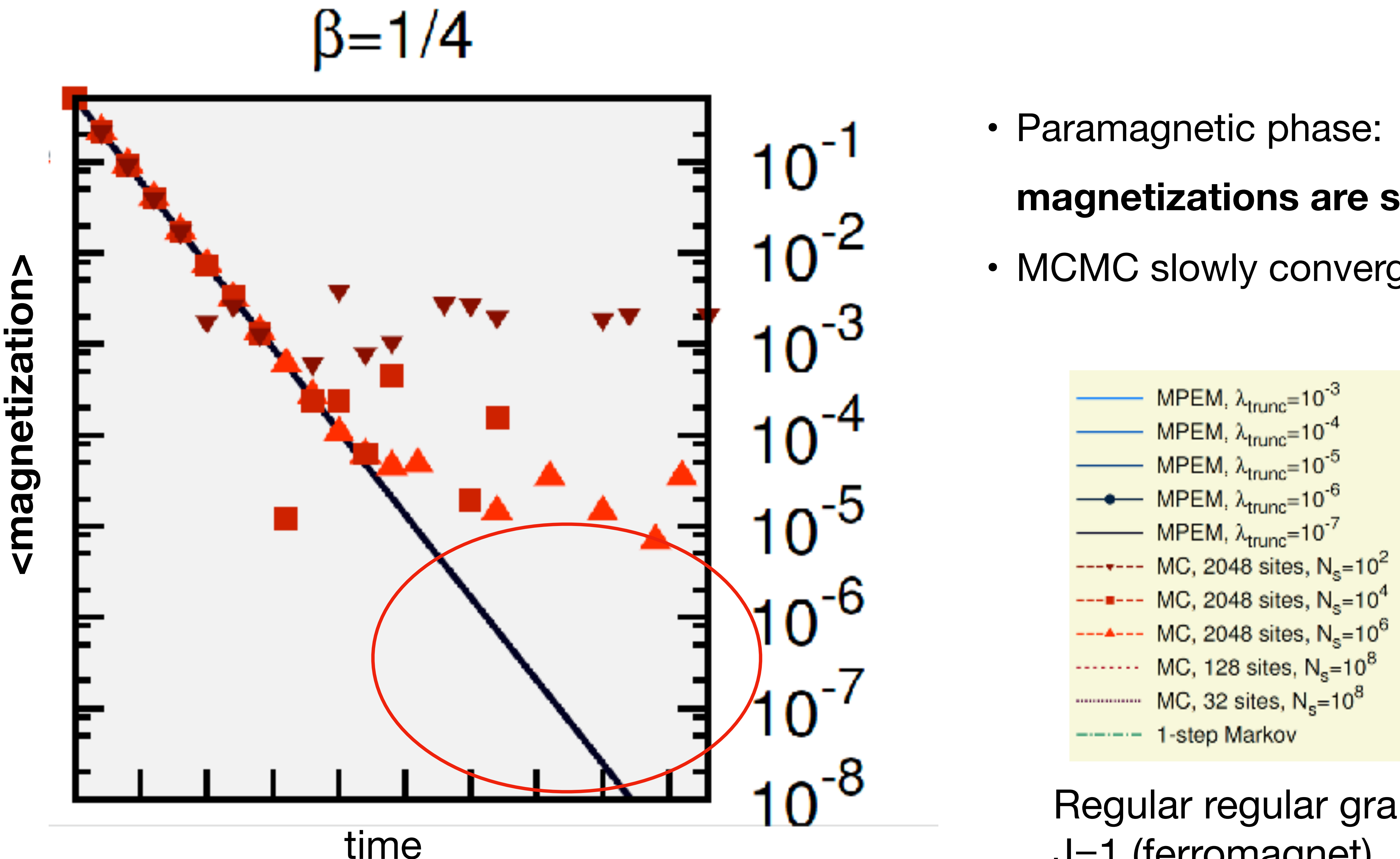


- **Out of equilibrium** we see most discrepancies
- MPEM does well at **every t**
- MCMC has **finite-size** problems

Regular regular graphs z=3
J=1 (ferromagnet),
thermodynamic limit

Matrix Product State and message-passing

Results on Glauber dynamics



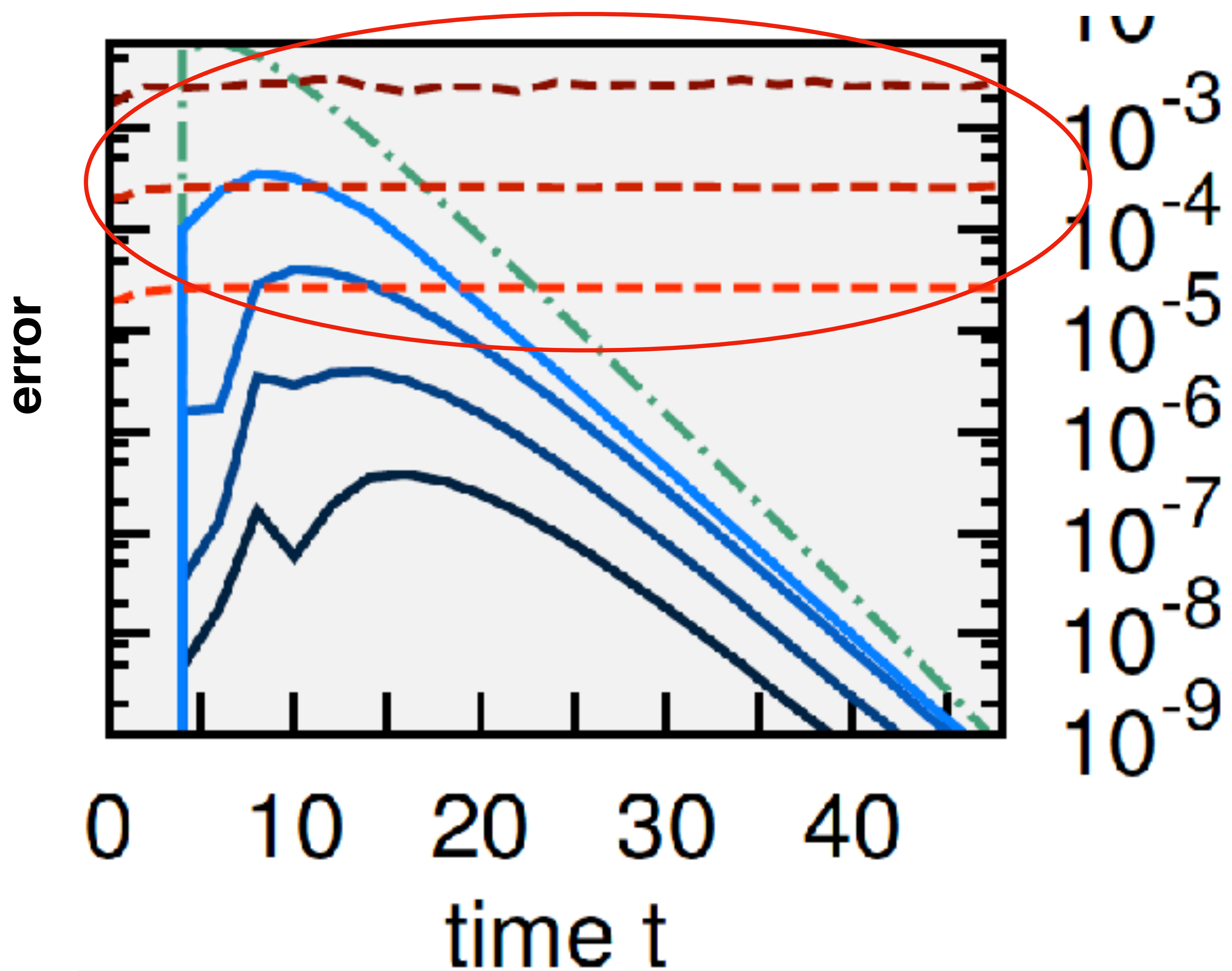
- Paramagnetic phase:

  magnetizations are small
- MCMC slowly converges

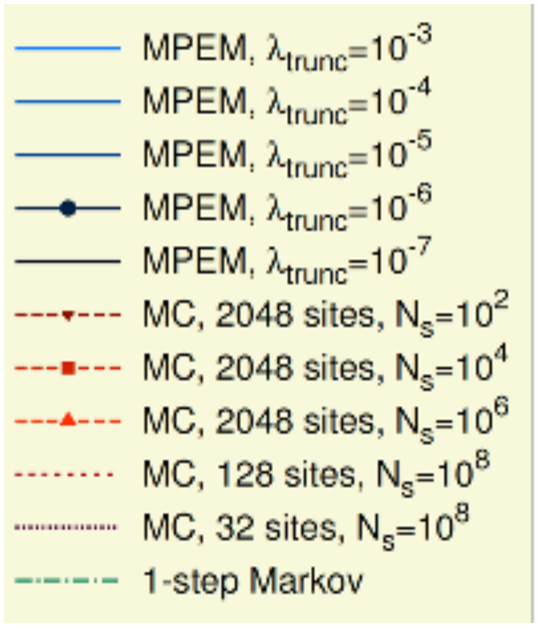
Regular regular graphs z=3
J=1 (ferromagnet),
thermodynamic limit

Matrix Product State and message-passing

Results on Glauber dynamics



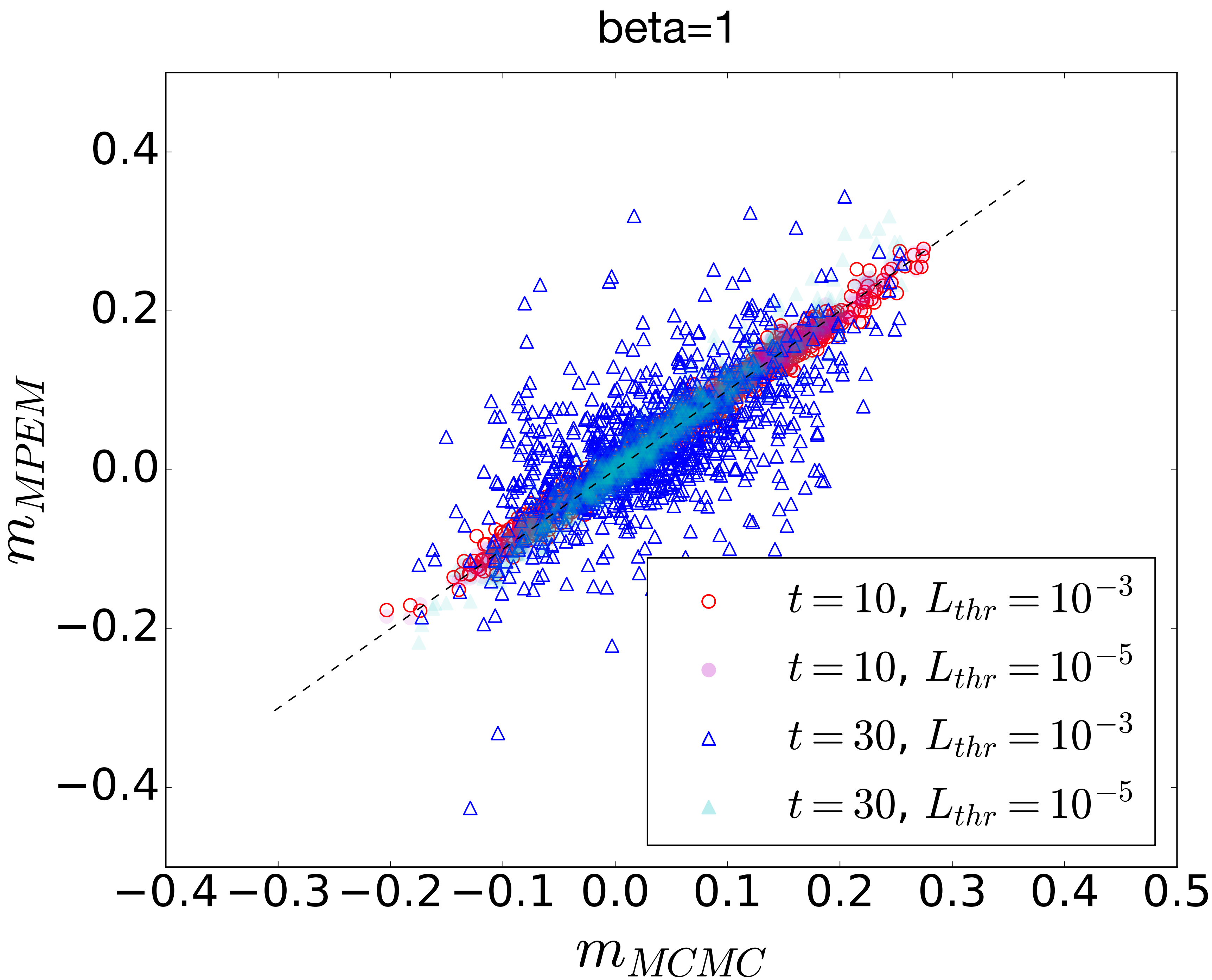
- Paramagnetic phase: **magnetizations are small**
- MCMC has slowly decaying errors



Regular regular graphs z=3
J=1 (ferromagnet),
thermodynamic limit

Matrix Product State and message-passing

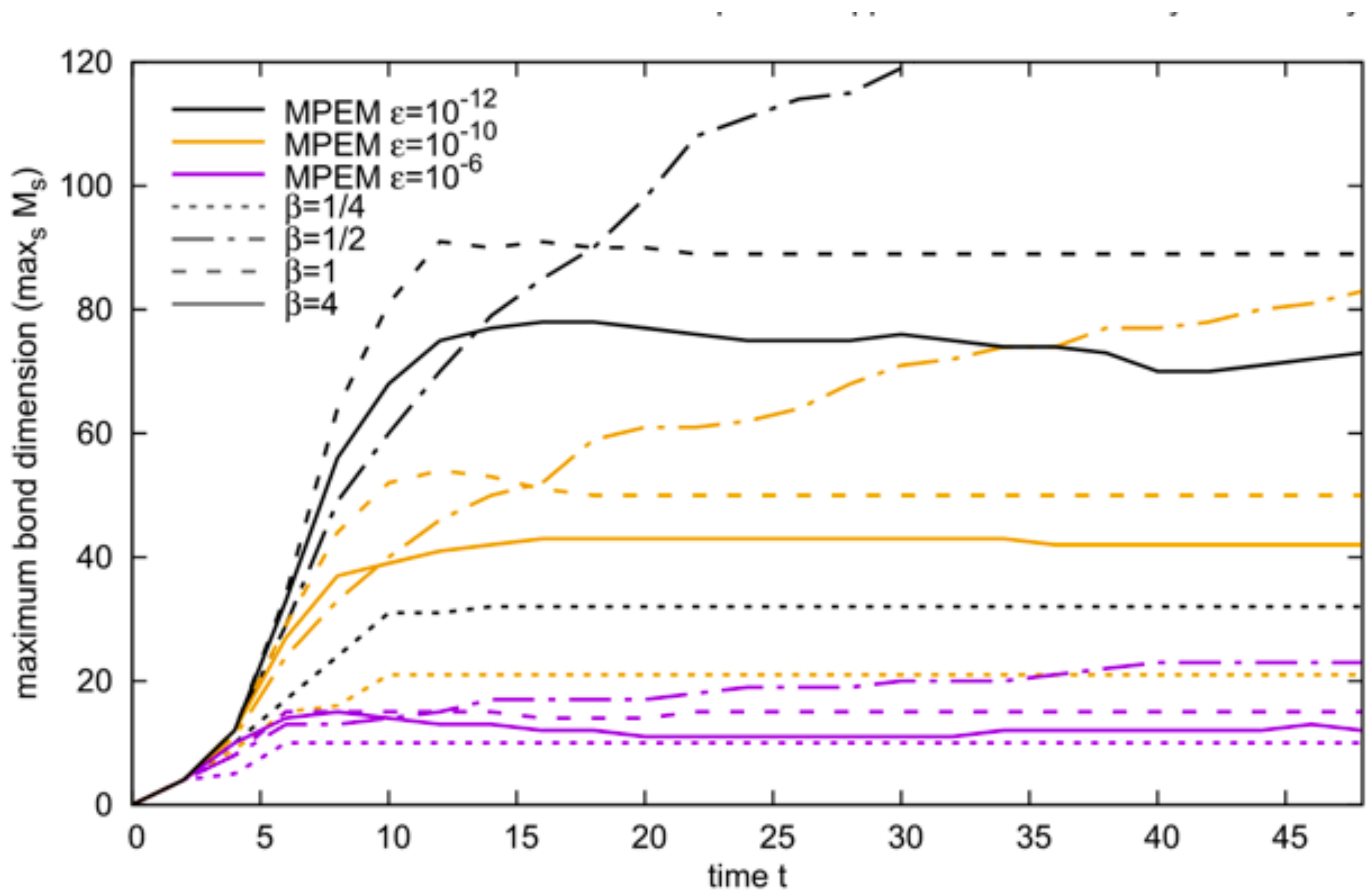
Trade-off accuracy vs complexity



- At **late time** steps we need **higher dimensions,** higher complexity

Matrix Product State and message-passing

Trade-off accuracy vs complexity



- $O(M^{2z-1})$ most expensive step
- M constant after initial steps
- Need to compute t matrices A
- Total cost: **linear in time**

T Barthel, J Stat Mech 013217, 2020

Matrix Product State and message-passing
Open questions

- How to further **optimize** the computational complexity?
    - Is there a way to avoid recomputing at each time all the matrices A from scratch?

This would introduce an approximation, but if we learn (somehow) what are the most relevant correlations, perhaps we can guide this approximation

Iterative consistent cavity equation:

$$
\begin{aligned}
\mu_{ij}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^{t}) &= B_{ij}^{(t+2)}(\sigma_i^{t+1})B_{ij}^{(t+1)}(\sigma_i^{t})\left[\prod_{s=1}^{t} B_{ij}^{(s)}(\sigma_i^{s-1}|\sigma_j^{s})\right] B_{ij}^{(0)}(\sigma_j^{0}) \\
&= \frac{1}{Z_{ij}} P_i(\sigma_i^0) \sum_{\{\bar{\sigma}_k^{t-1}\}} \prod_{s=0}^{t} W(\sigma_i^{s+1}|\sigma_i^{s}, \{\sigma_j^s\}_{j\in\partial i}) \times \\
&\times \prod_{k\in\partial i\setminus j} \left\{ A_{ki}^{(t+1)}(\sigma_k^{t}) A_{ki}^{(t)}(\sigma_k^{t-1}) \left[\prod_{s=1}^{t} A_{ki}^{(s)}(\sigma_k^{s-1}|\sigma_i^{s})\right] A_{ki}^{(0)}(\sigma_i^{0}) \right\}
\end{aligned}
$$

Valid for synchronous parallel update

Matrix Product State and message-passing

Open questions

- How to further **optimize** the computational complexity?
  - Is there a way to avoid recomputing at each time all the matrices from scratch?
  - Can we lower the complexity in terms of z? (node average degree)

Most real networks have z > 3. In addition, there could be “hubs”, nodes with high z. How do we deal with these?

Matrix Product State and message-passing
Open questions

- How to further **optimize** the computational complexity?
  - Is there a way to avoid recomputing at each time all the matrices from scratch?
  - Can we lower the complexity in terms of z? (node average degree)
- How to make it more **flexible**?
  - Only parallel dynamics, how about sequential?

Matrix Product State and message-passing
Open questions

- How to further **optimize** the computational complexity?
  - Is there a way to avoid recomputing at each time all the matrices from scratch?
  - Can we lower the complexity in terms of z? (node average degree)
- How to make it more **flexible**?
  - Only parallel dynamics, how about sequential?
  - How about continuous variables?

E.g. Community detection with mixed-membership

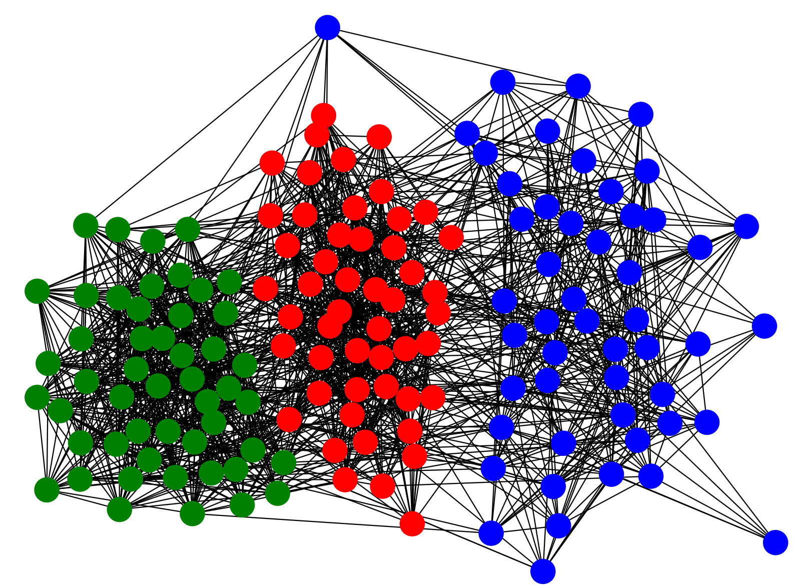
Matrix Product State and message-passing

Dynamic stochastic block model

$$\mu_{ij}(\sigma_i) \propto \gamma_{\sigma_i} e^{-h_i} \prod_{k \in \partial i \setminus j} \left[\sum_{\sigma_k} C_{\sigma_i \sigma_k} \mu_{ki}(\sigma_k) \right]$$

$$W(\sigma_i^t | \sigma_i^{t-1}) = \eta \delta_{\sigma_i^t, \sigma_i^{t-1}} + (1-\eta) \gamma_{\sigma_i^t}$$

Nodes change labels: with probability eta they keep it, with probability 1-eta they switch it randomly



$$\mu_{ij}(\bar{\sigma}_i) \propto \gamma_{\bar{\sigma}_i} e^{-h_i(\bar{\sigma}_i)} \prod_{k \in \partial i \setminus j} \sum_{\bar{\sigma}_k} P(\bar{A}_{ki} | \bar{\sigma}_i, \bar{\sigma}_k) \mu_{ki}(\bar{\sigma}_k)$$

$$= \gamma_{\bar{\sigma}_i} e^{-h_i(\bar{\sigma}_i)} \prod_{k \in \partial i \setminus j} \sum_{\bar{\sigma}_k} \left[\prod_t C_{\sigma_i^t \sigma_k^t} \right] \mu_{ki}(\bar{\sigma}_k)$$

A Ghasemian, 2019; A Ghasemian et al. Physical Review X 6 (3), 031005, 2016

Thanks

Summary: dynamical processes on networks and MPS

- Modeling dynamical processes in network requires estimating marginals: this is expensive
- Message-passing algorithms help to reduce the complexity in space
- MPS help to reduce the complexity in time
- Combining these two (MPEM), we obtain accurate approximation for the marginals, at a tunable computational cost

caterina.debacco@tuebingen.mpg.de