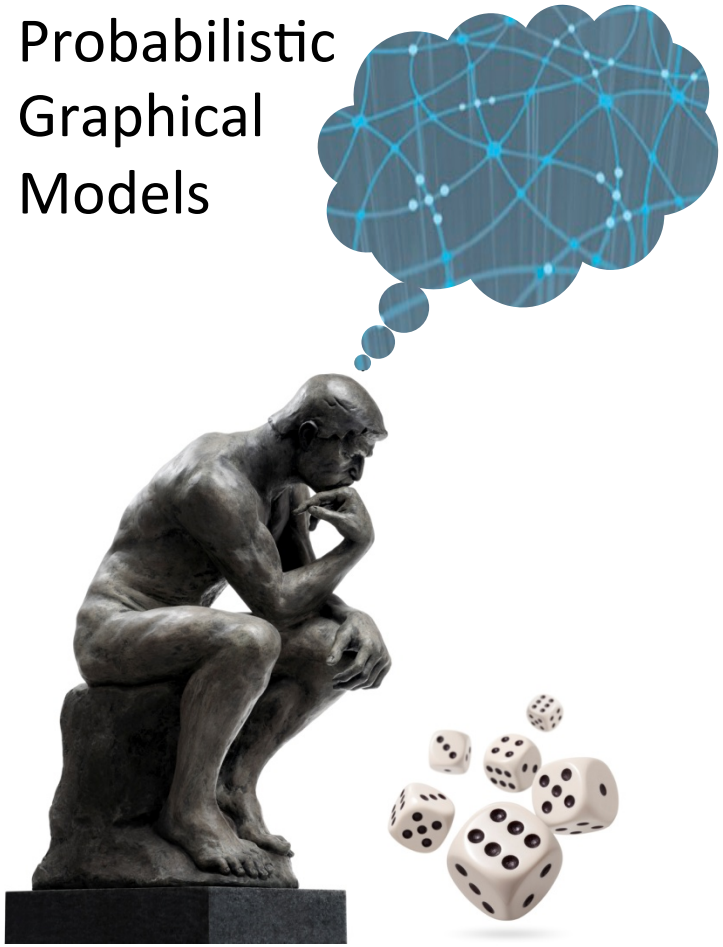
Probabilistic Graphical Models

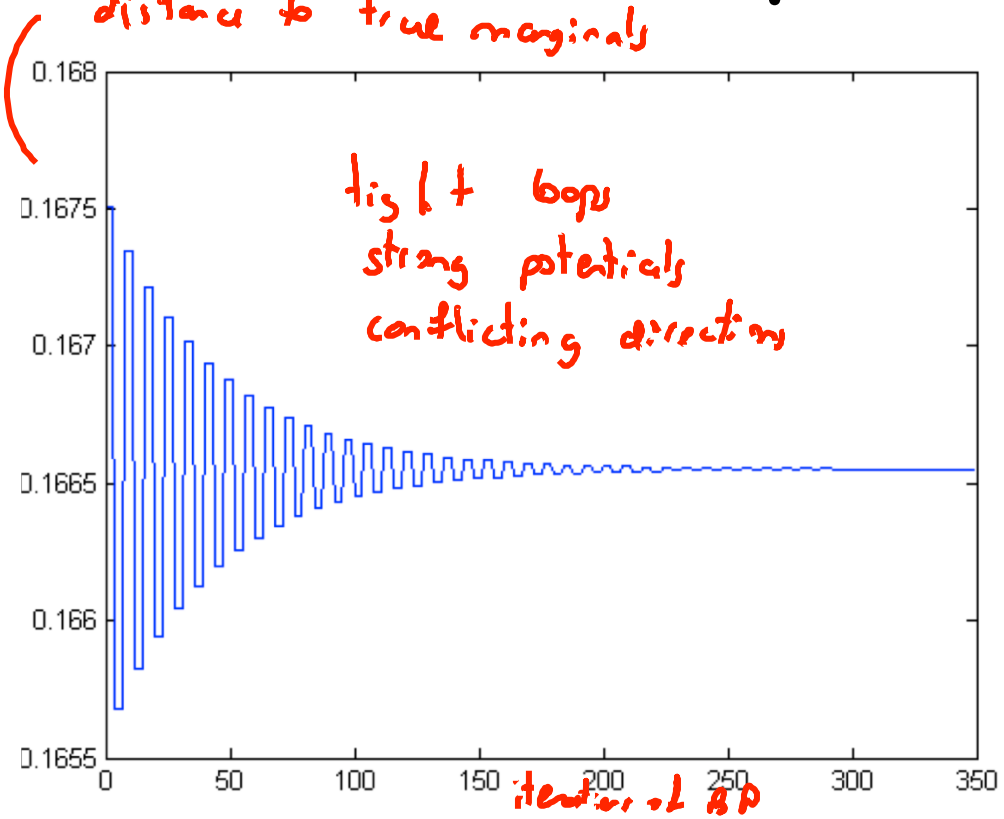


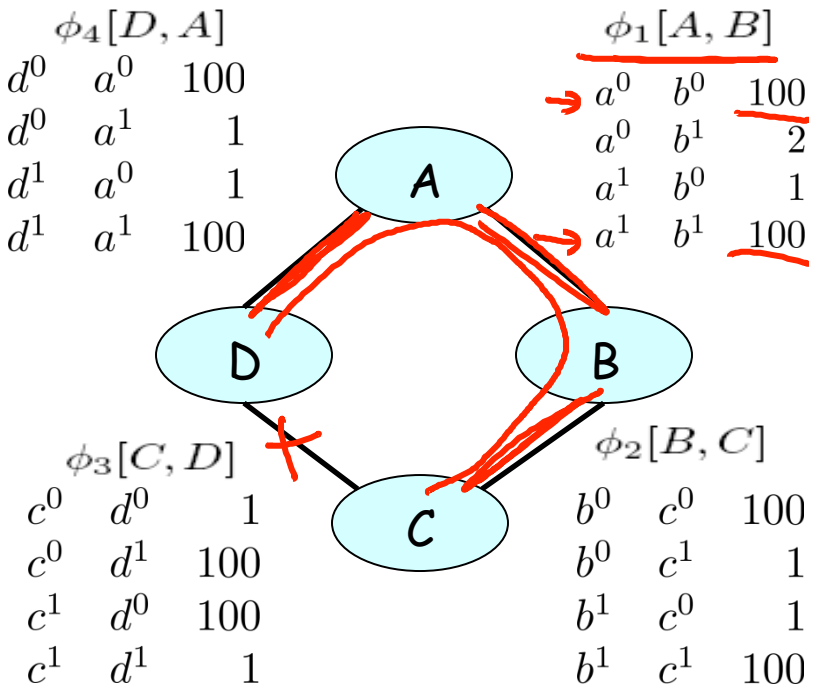
Inference

Message Passing

BP In Practice

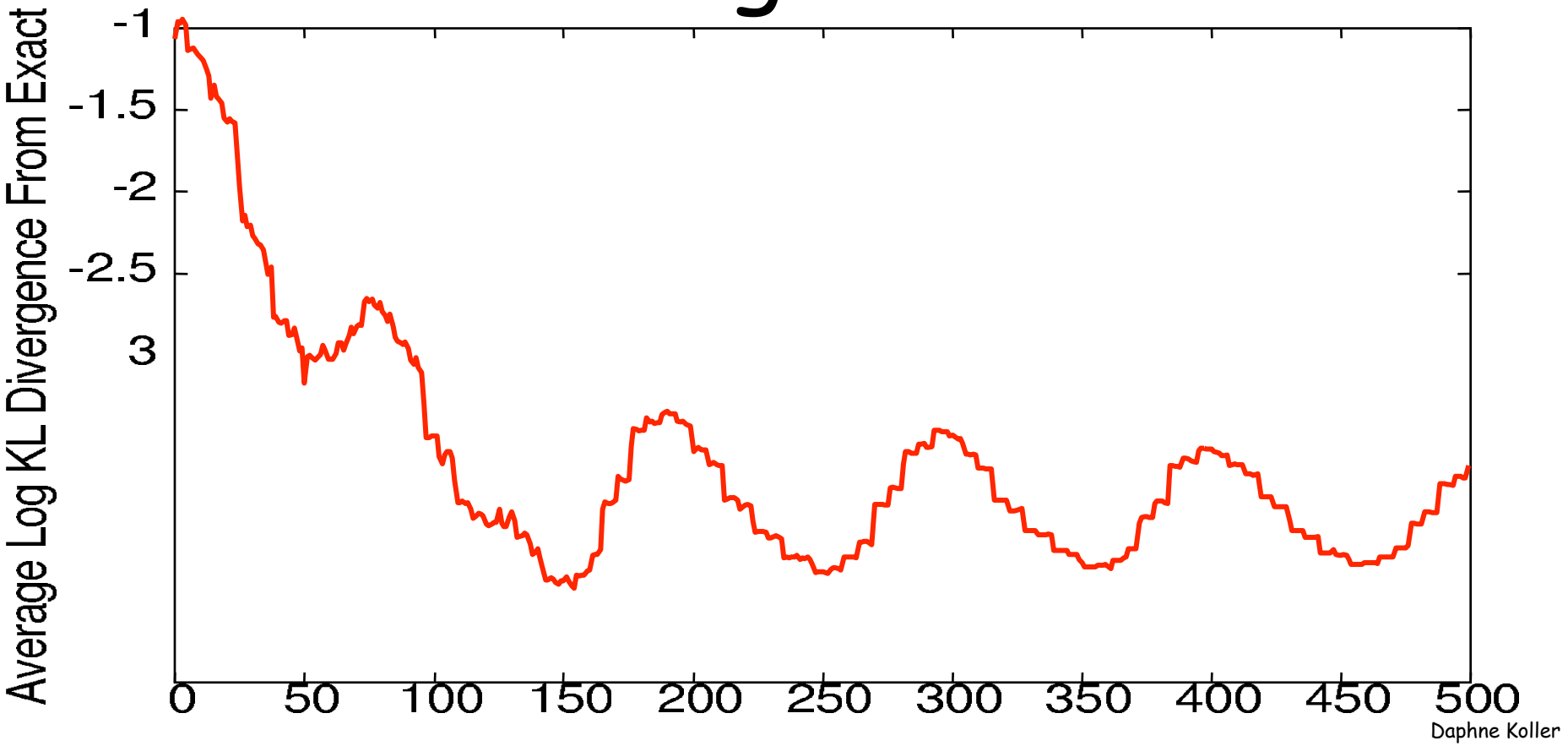
Misconception Revisited





Daphne Koller

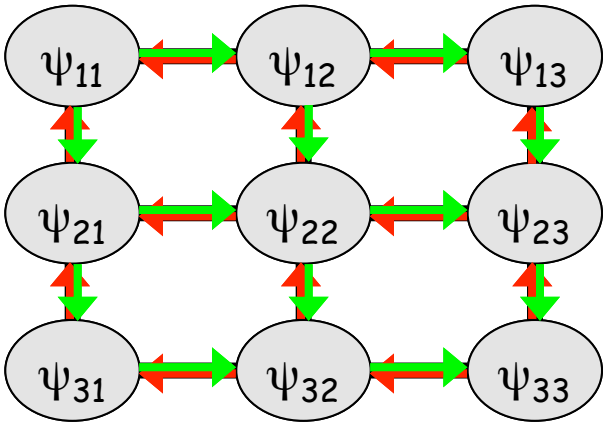
Nonconvergent BP Run

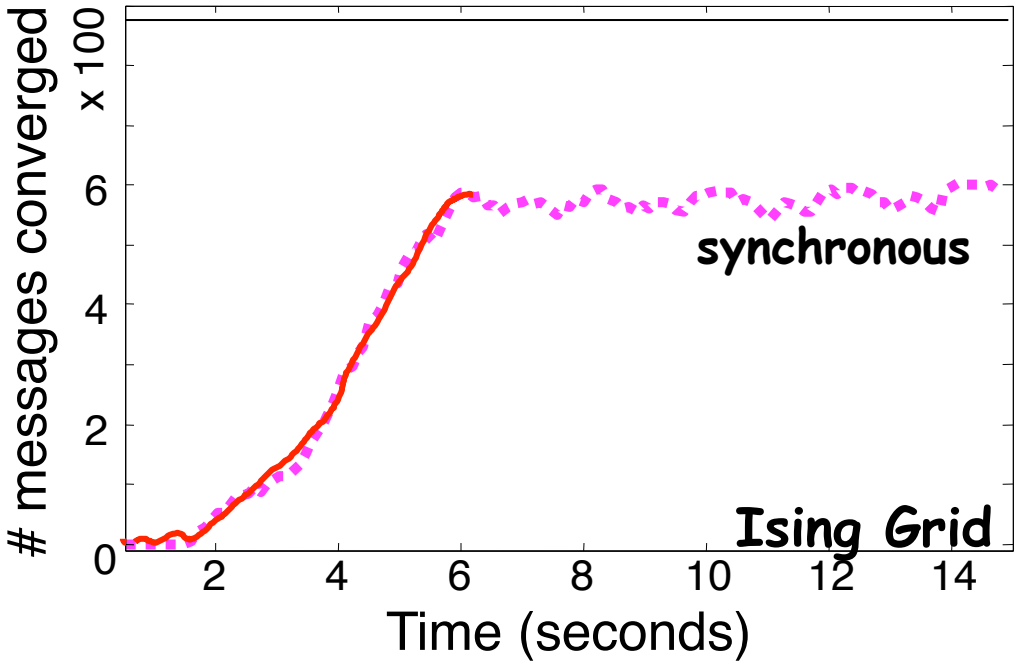


Different Variants of BP

Synchronous BP:

all messages are updated in parallel



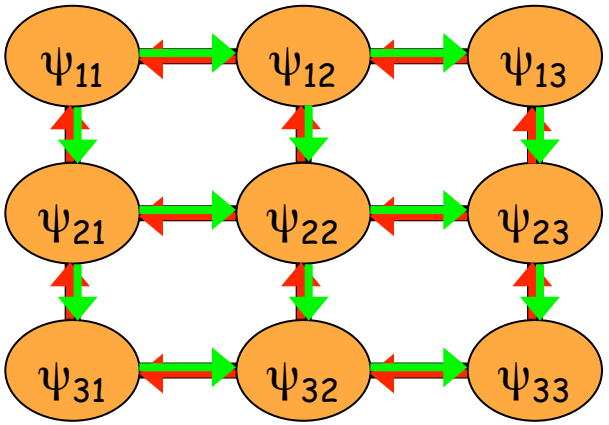


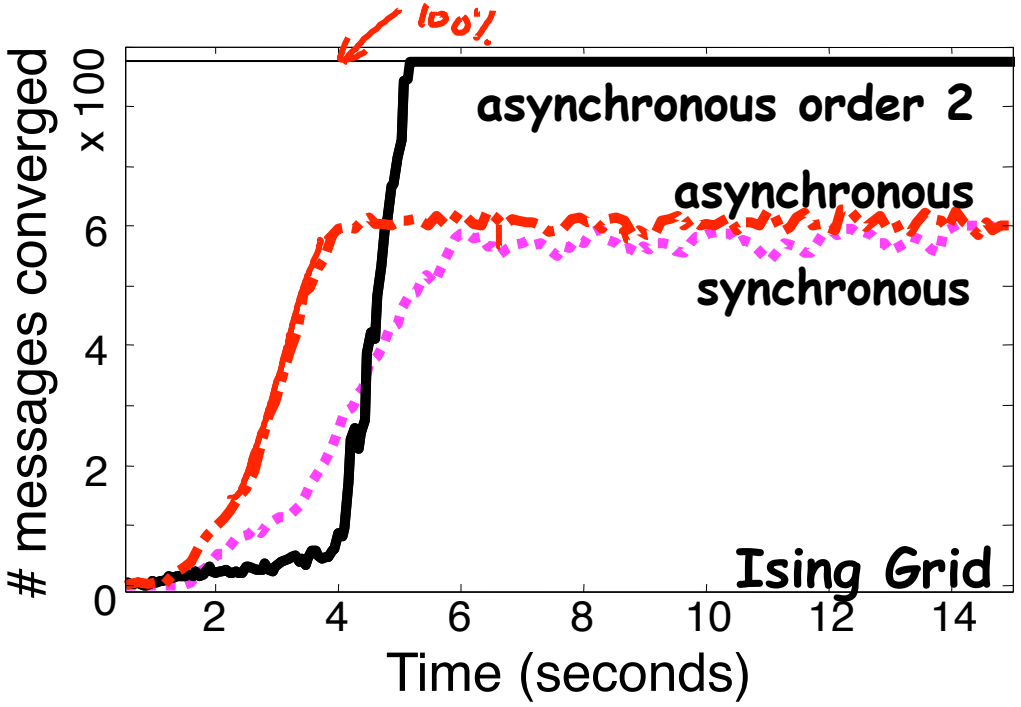
Daphne Koller

Different Variants of BP

Asynchronous BP:

Messages are updated one at a time





Daphne Koller

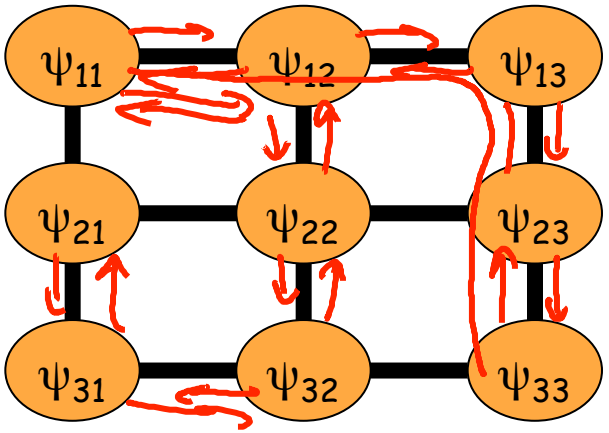
Observations

- Convergence is a local property:
  - some messages converge soon
  - others may never converge
- Synchronous BP converges considerably worse than asynchronous
- Message passing order makes a difference to extent and rate of convergence

Informed Message Scheduling

- Tree reparameterization (TRP)
  - Pick a tree and pass messages to calibrate

trees must cover all edges
improves performance if trees are larger (spanning trees)



Informed Message Scheduling

- Tree reparameterization (TRP)
  - Pick a tree and pass messages to calibrate

- Residual belief propagation (RBP)
  - Pass messages between two clusters whose beliefs over the sepset disagree the most

priority queue of edges

Smoothing (Damping) Messages

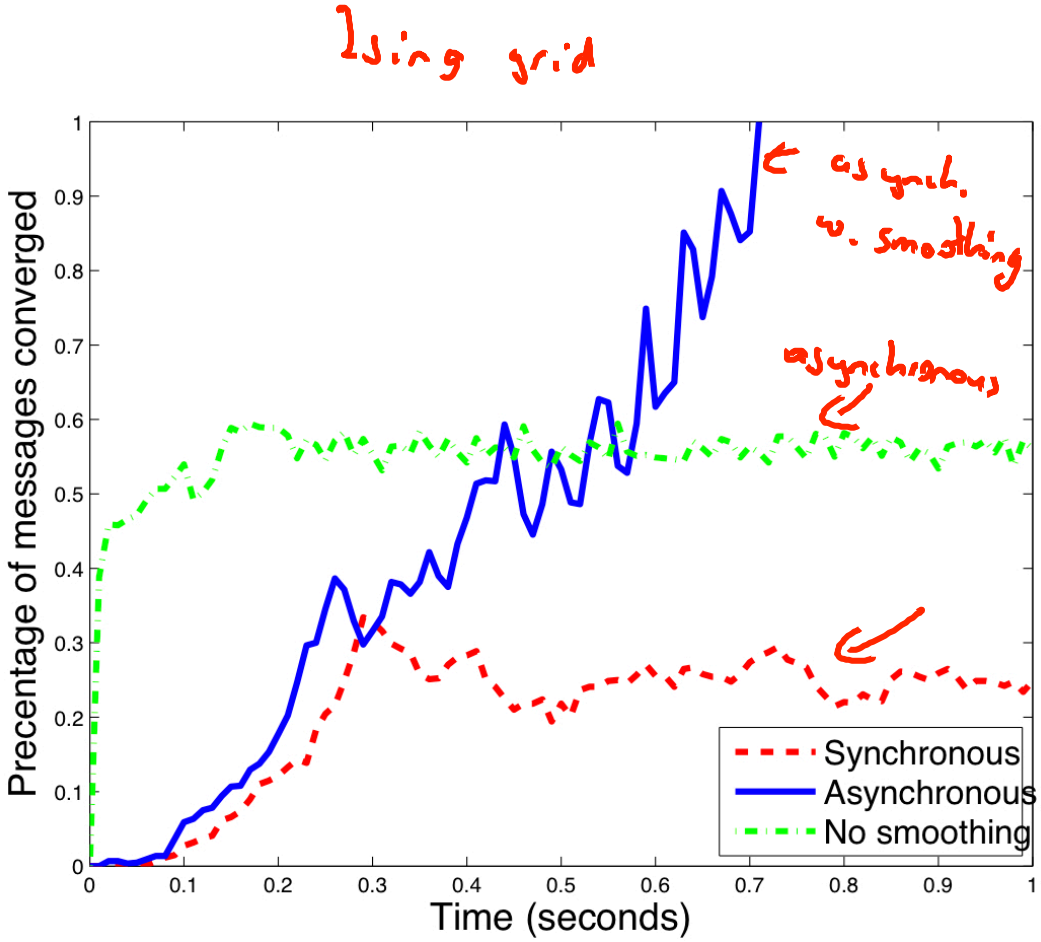
$$\delta_{i\to j} \leftarrow \sum_{\boldsymbol{C}_i - \boldsymbol{S}_{i,j}} \psi_i \prod_{k\neq j} \delta_{k\to i}$$

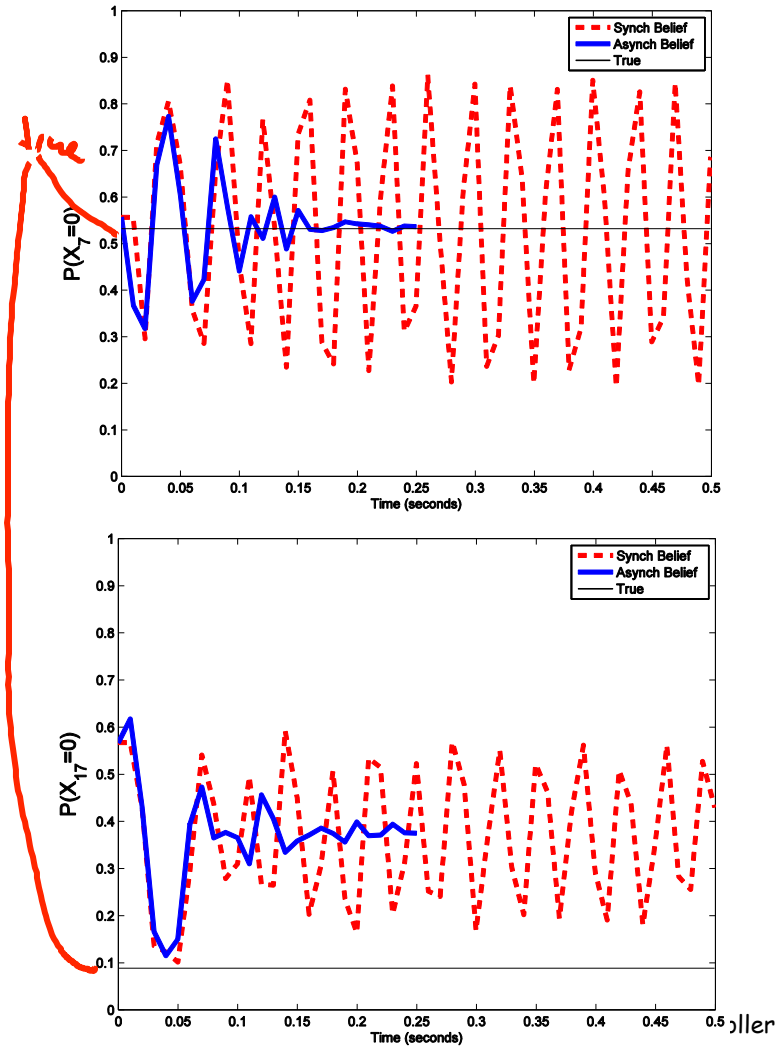
$$\delta_{i\to j} \leftarrow \lambda \left(\sum_{\boldsymbol{C}_i - \boldsymbol{S}_{i,j}} \psi_i \prod_{k\neq j} \delta_{k\to i} \right) + (1-\lambda)\delta^{\text{old}}_{i\to j}$$

new msg

old msg

- Dampens oscillations in messages


Ising grid
as synch. w. smoothing
asynchronous
Precentage of messages converged
Time (seconds)
Synchronous
Asynchronous
No smoothing


true
Synch Belief
Asynch Belief
True
P(X7=0)
P(X17=0)
Time (seconds)

...ller

Summary

- To achieve BP convergence, two main tricks
  - Damping
  - Intelligent message ordering
- Convergence doesn't guarantee correctness
- Bad cases for BP – both convergence & accuracy:
  - Strong potentials pulling in different directions
  - Tight loops
- Some new algorithms have better convergence:
  - Optimization-based view to inference