UPPSALA UNIVERSITET

LINKÖPING UNIVERSITY

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

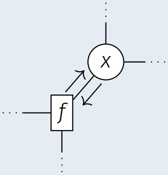
Graphical model inference: Sequential Monte Carlo meets deterministic approximations

Fredrik Lindsten (Linköping University and Uppsala University)
Jouni Helske (Linköping University)
Matti Vihola (University of Jyväskylä)

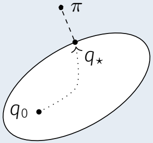
Approximate Bayesian inference

Deterministic methods

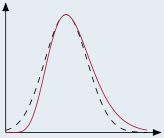
Message passing



Variational inference

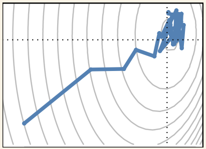


Laplace's method

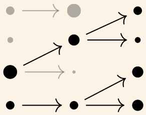


Monte Carlo methods

Markov chain Monte Carlo



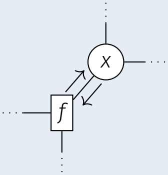
Sequential Monte Carlo



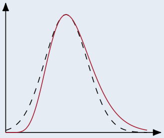
Approximate Bayesian inference

Deterministic methods

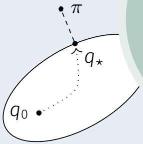
Message passing

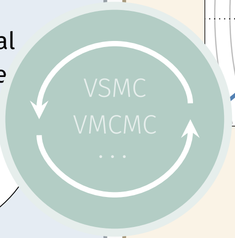


Laplace's method



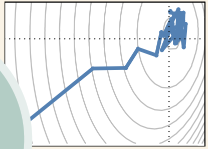
Variational inference



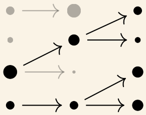


Monte Carlo methods

Markov chain Monte Carlo



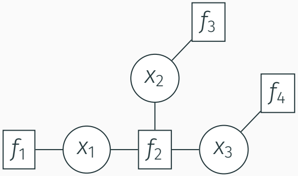
Sequential Monte Carlo



Probabilistic graphical models

We consider inference in *factor graphs* with joint distribution

$$\pi(x_{1:T}) = \frac{1}{Z} \prod_{j \in \mathcal{F}} f_j(x_{\mathcal{I}_j}).$$



Task:

- Compute expectations w.r.t. $\pi(x_{1:T})$.
- Compute the normalizing constant Z.

Sequential Monte Carlo

Sequential Monte Carlo (SMC) can be used for probabilistic graphical model inference via *sequential graph decompositions:*

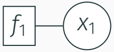
Christian A. Naesseth, Fredrik Lindsten and Thomas B. Schön. **Sequential Monte Carlo methods for graphical models.** *Advances in Neural Information Processing Systems 27*, December, 2014.

Sequential Monte Carlo (SMC) can be used for probabilistic graphical model inference via *sequential graph decompositions:*

Christian A. Naesseth, Fredrik Lindsten and Thomas B. Schön. **Sequential Monte Carlo methods for graphical models.** *Advances in Neural Information Processing Systems 27*, December, 2014.

Define intermediate SMC targets: $\gamma_t(x_{1:t}) = \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j})$.

Iteration $t = 1$



$\gamma_1(x_1)$

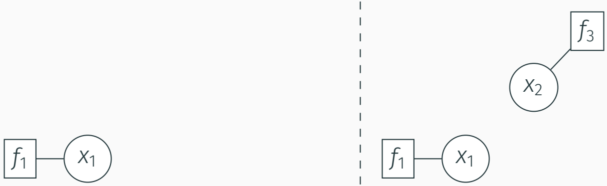
Sequential Monte Carlo

Sequential Monte Carlo (SMC) can be used for probabilistic graphical model inference via *sequential graph decompositions:*

Christian A. Naesseth, Fredrik Lindsten and Thomas B. Schön. **Sequential Monte Carlo methods for graphical models.** *Advances in Neural Information Processing Systems 27*, December, 2014.

Define intermediate SMC targets: $\gamma_t(x_{1:t}) = \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j})$.

Iteration $t = 1$ | **Iteration** $t = 2$



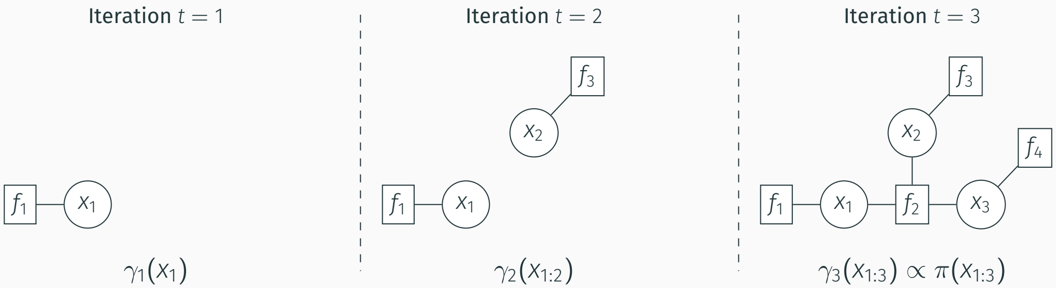
$\gamma_1(x_1)$ | $\gamma_2(x_{1:2})$

Sequential Monte Carlo

Sequential Monte Carlo (SMC) can be used for probabilistic graphical model inference via *sequential graph decompositions:*

Christian A. Naesseth, Fredrik Lindsten and Thomas B. Schön. **Sequential Monte Carlo methods for graphical models.** *Advances in Neural Information Processing Systems 27*, December, 2014.

Define intermediate SMC targets: $\gamma_t(x_{1:t}) = \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j})$.



Twisted SMC

Dependencies on “future variables” are not taken into account!

Twisted intermediate targets:

$$\gamma_t^{\psi}(x_{1:t}) := \psi_t(x_{1:t})\gamma_t(x_{1:t}) = \psi_t(x_{1:t}) \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j}).$$

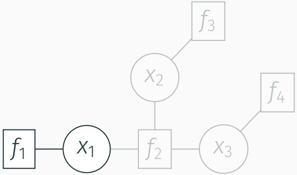
Twisted SMC

Dependencies on “future variables” are not taken into account!

Twisted intermediate targets:

$$\gamma_t^{\psi}(x_{1:t}) := \psi_t(x_{1:t})\gamma_t(x_{1:t}) = \psi_t(x_{1:t}) \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j}).$$

Iteration $t = 1$



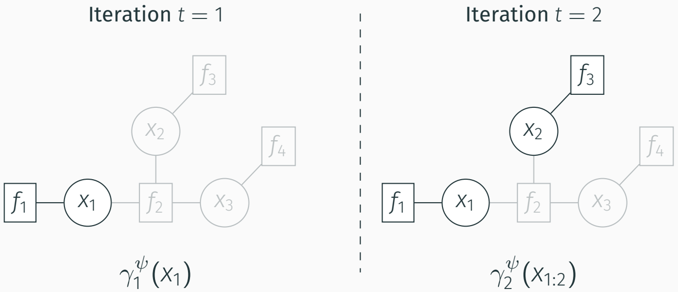
$\gamma_1^{\psi}(x_1)$

Twisted SMC

> Dependencies on "future variables" are not taken into account!

Twisted intermediate targets:

$$\gamma_t^{\psi}(x_{1:t}) := \psi_t(x_{1:t})\gamma_t(x_{1:t}) = \psi_t(x_{1:t}) \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j}).$$



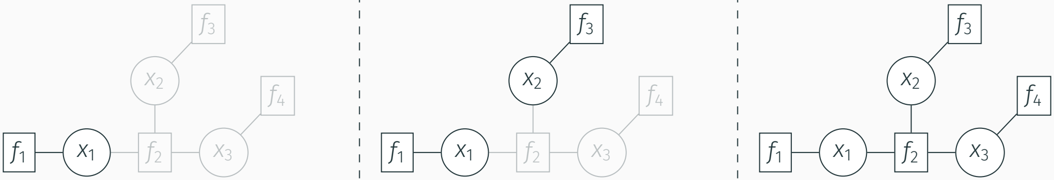
Twisted SMC

Dependencies on "future variables" are not taken into account!

Twisted intermediate targets:

$$\gamma_t^{\psi}(x_{1:t}) := \psi_t(x_{1:t})\gamma_t(x_{1:t}) = \psi_t(x_{1:t}) \prod_{j \in \mathcal{F}_t} f_j(x_{\mathcal{I}_j}).$$

Iteration $t = 1$ | **Iteration** $t = 2$ | **Iteration** $t = 3$



$\gamma_1^{\psi}(x_1)$ | $\gamma_2^{\psi}(x_{1:2})$ | $\gamma_3^{\psi}(x_{1:3}) \propto \pi(x_{1:3})$

How do we choose the twisting functions?

Proposition (Optimal twisting). With

$$\psi_t^*(x_{1:t}) = \int \prod_{j \in \mathcal{F} \setminus \mathcal{F}_t} f_j(x_{\mathcal{I}_j}) dx_{t+1:T},$$

the SMC algorithm outputs i.i.d. draws from π and the normalizing constant estimate is exact; $\widehat{Z} = Z$ w.p.1.

How do we choose the twisting functions?

Proposition (Optimal twisting). With

$$\psi_t^*(x_{1:t}) = \int \prod_{j \in \mathcal{F} \setminus \mathcal{F}_t} f_j(x_{\mathcal{I}_j}) dx_{t+1:T},$$

the SMC algorithm outputs i.i.d. draws from π and the normalizing constant estimate is exact; $\widehat{Z} = Z$ w.p.1.

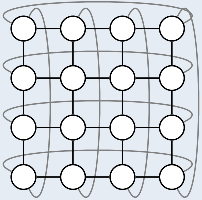
Optimal twisting functions are intractable, but:

- $\psi_t \approx \psi_t^*$ can be computed by various *deterministic inference methods*
- Sub-optimality only affects efficiency, not consistency or unbiasedness
- Can be seen as a bias post-correction

Twisting functions via deterministic approximations

Loopy Belief Propagation

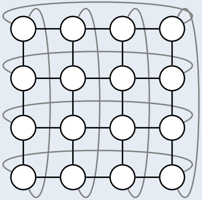
ex) Square lattice Ising model



Twisting functions via deterministic approximations

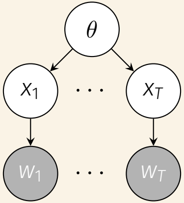
Loopy Belief Propagation

ex) Square lattice Ising model



Expectation Propagation

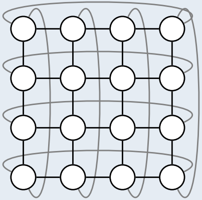
ex) Topic model likelihood evaluation



Twisting functions via deterministic approximations

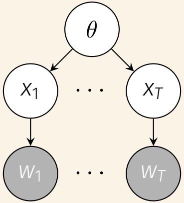
Loopy Belief Propagation

ex) Square lattice Ising model



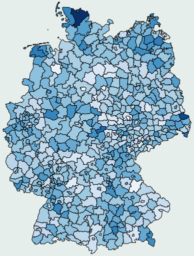
Expectation Propagation

ex) Topic model likelihood evaluation



Laplace Approximation

ex) Gaussian Markov random field



Thank you for listening!

Come see the poster: #51

Code available at:

- `github.com/freli005/smc-pgm-twist`
- `github.com/helske/particlefield`