- PhD student since 2008
- Main focus is on construction of discrete Markov random field models from training images (Toftaker and Tjelmeland, 2012).
- In the fall of 2010 I was in Aalborg, working with professor Jesper Møller on anisotropic spatial point patterns (Møller and Toftaker, 2012).
- Now we work on another approach to construct discrete models from training images.
- The thesis will be finished in November 2012.

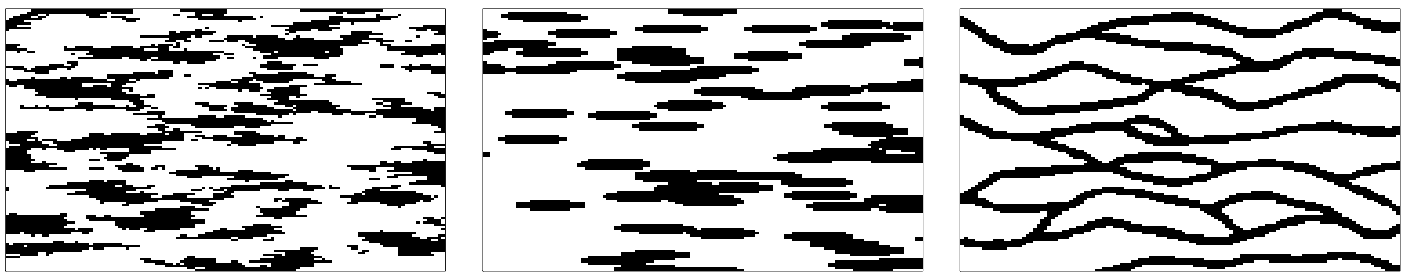
Construction of Markov random field prior models.

Håkon Toftaker

NTNU

February 8, 2012

- Construct realistic Markov random field prior models for rock types in petroleum reservoirs.
- Models should be constructed from training data such as these.



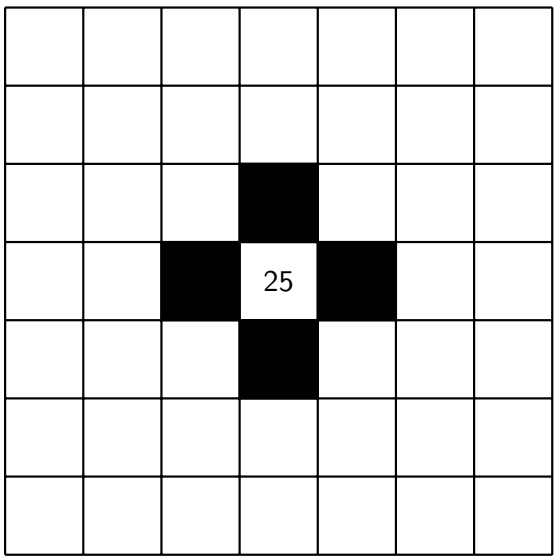
Markov random fields

- The Markov random field $x = (x_1, ..., x_n)$ is a set of binary variables defined on a lattice of n nodes

x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
x_{15}	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}
x_{22}	x_{23}	x_{24}	x_{25}	x_{26}	x_{27}	x_{28}
x_{29}	x_{30}	x_{31}	x_{32}	x_{33}	x_{34}	x_{35}
x_{36}	x_{37}	x_{38}	x_{39}	x_{40}	x_{41}	x_{42}
x_{43}	x_{44}	x_{45}	x_{46}	x_{47}	x_{48}	x_{49}

Markov random fields

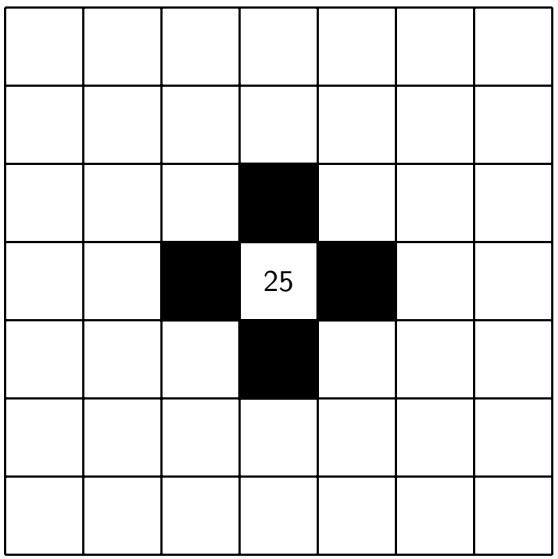
- The Markov random field $x = (x_1, ..., x_n)$ is a set of binary variables defined on a lattice of n nodes
- To each node k we associate a neighborhood ∂k



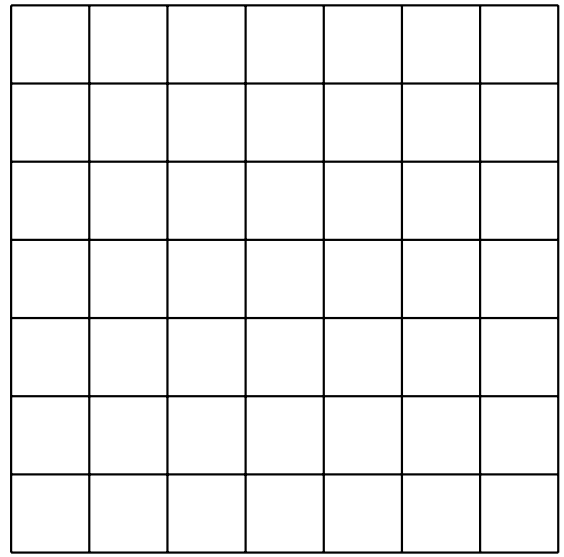
Markov random fields

- The Markov random field $x = (x_1, ..., x_n)$ is a set of binary variables defined on a lattice of n nodes
- To each node k we associate a neighborhood ∂k
- The probability distribution of x has a Markov property, which means

$$p_\theta(x_k|x_{-k}) = p_\theta(x_k|x_{\partial k})$$

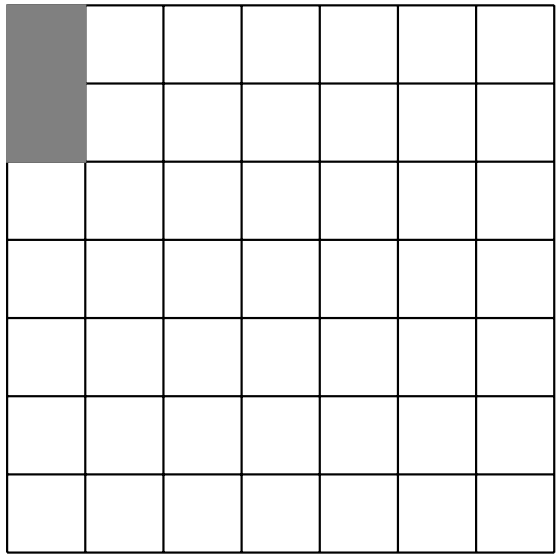


- A clique is defined as a set of nodes, which are all neighbors of each other.



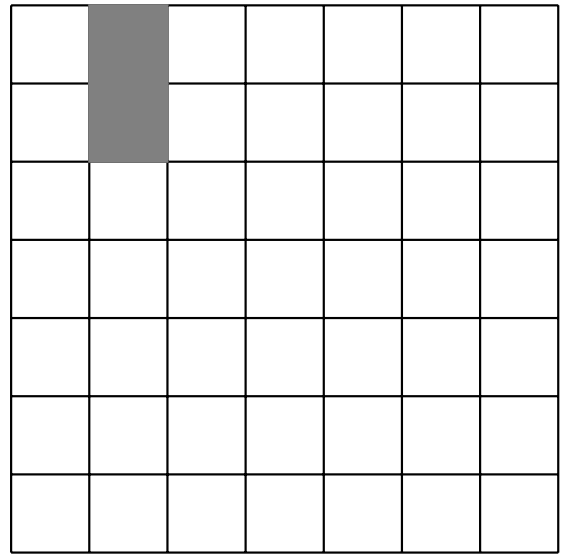
- A clique is defined as a set of nodes, which are all neighbors of each other.
- The corresponding joint distribution is

$$p_\theta(x) = c(\theta) \exp\left\{ - \sum_{C \in \mathcal{C}} V(x_C, \theta) \right\}$$
$$= c(\theta) \exp\left\{ -U_\theta(x) \right\}.$$



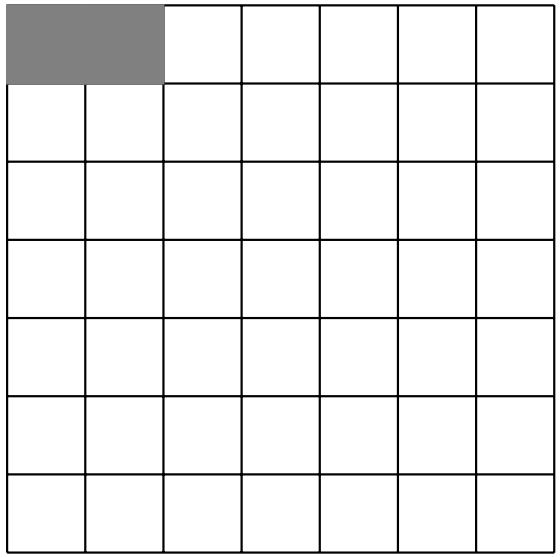
- A clique is defined as a set of nodes, which are all neighbors of each other.
- The corresponding joint distribution is

$$p_\theta(x) = c(\theta) \exp\left\{ - \sum_{C \in \mathcal{C}} V(x_C, \theta) \right\}$$
$$= c(\theta) \exp\left\{ -U_\theta(x) \right\}.$$

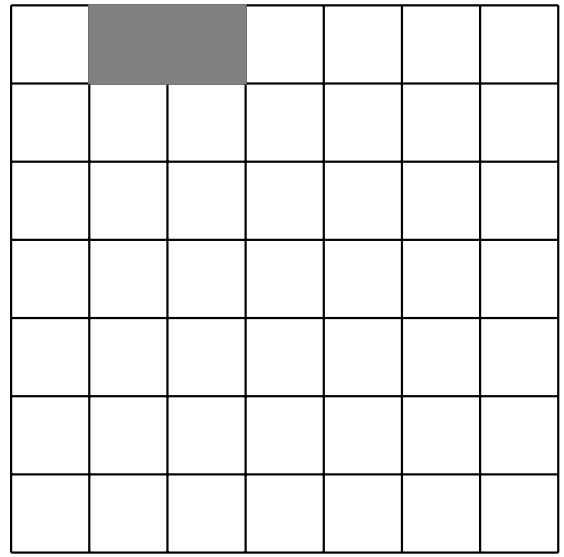


- A clique is defined as a set of nodes, which are all neighbors of each other.
- The corresponding joint distribution is

$$p_\theta(x) = c(\theta) \exp\left\{ - \sum_{C \in \mathcal{C}} V(x_C, \theta) \right\}$$
$$= c(\theta) \exp\left\{ -U_\theta(x) \right\}.$$



- A clique is defined as a set of nodes, which are all neighbors of each other.
- The corresponding joint distribution is



$$p_\theta(x) = c(\theta) \exp\left\{ - \sum_{C \in \mathcal{C}} V(x_C, \theta) \right\}$$

$$= c(\theta) \exp\left\{ -U_\theta(x) \right\}.$$

- Here $c(\theta)$ is a normalizing constant

$$c(\theta) = \sum_x \exp\left\{ -U_\theta(x) \right\},$$

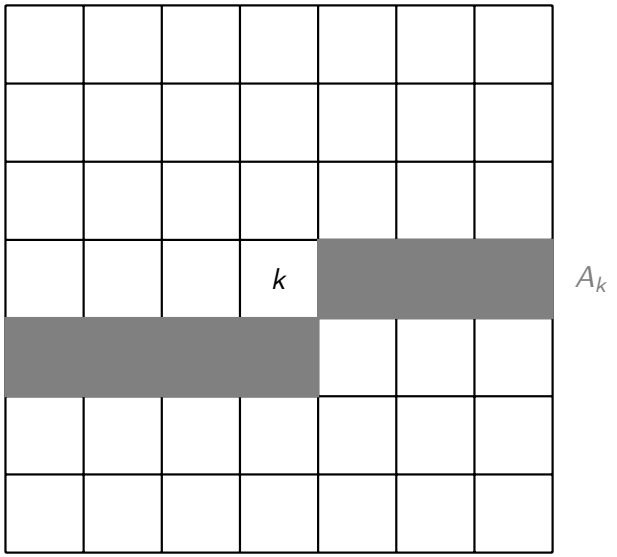
and $U_\theta(x)$ is called an energy function.

Forward-backward algorithm

- By a forward-backward algorithm (Friel and Rue, 2005) we can sum out the components of x, and sequentially obtain

$$p_\theta(x_k|x_{k+1:n}) = p_\theta(x_k|x_{A_k})$$

and $c(\theta)$.

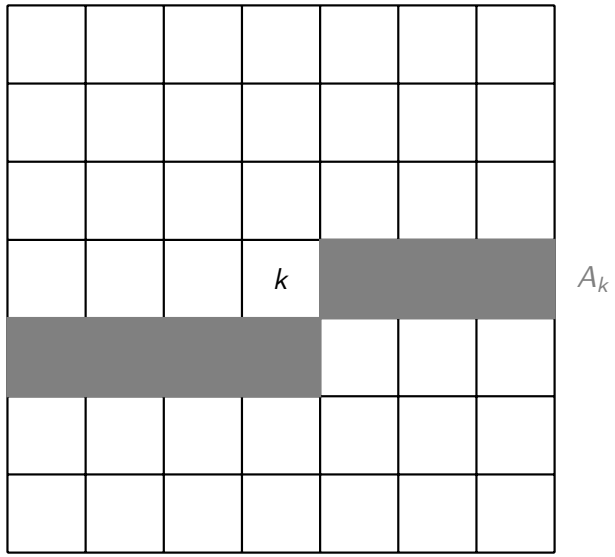


Forward-backward algorithm

- By a forward-backward algorithm (Friel and Rue, 2005) we can sum out the components of x, and sequentially obtain

$$p_\theta(x_k|x_{k+1:n}) = p_\theta(x_k|x_{A_k})$$

and $c(\theta)$.



- The probability distribution is available as

$$p_\theta(x) = p_\theta(x_n) \prod_{k=1}^{n-1} p_\theta(x_k|x_{A_k}),$$

and it is simple to sample from $p_\theta(x)$.

- Computational complexity is $2^{|A_k|+1} = 2^{7+1}$

The approximation

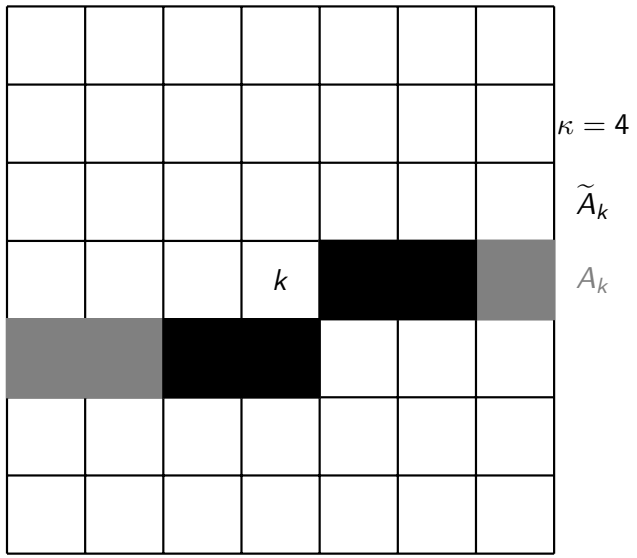
- Austad and Tjelmeland, 2011, develop a computationally feasable approximation of the above algorithm.

The approximation

- Austad and Tjelmeland, 2011, develop a computationally feasable approximation of the above algorithm.
- An approximate factorization is obtained

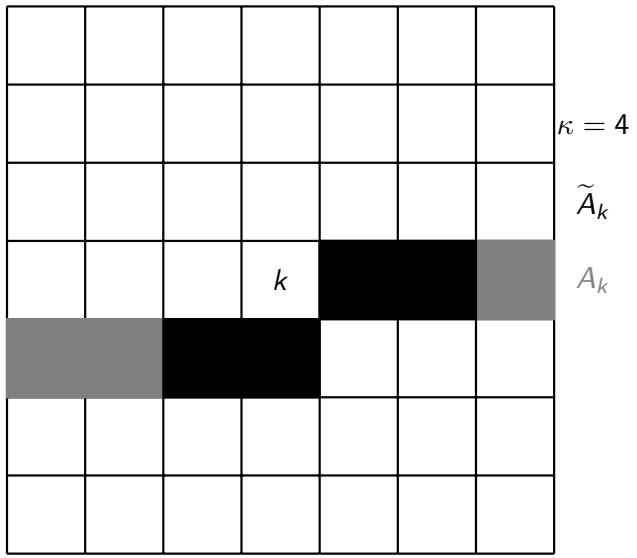
$$\widetilde{p}_\theta(x) = \widetilde{p}_\theta(x_n) \prod_{k=1}^{n-1} \widetilde{p}_\theta(x_k | x_{\tilde{A}_k}),$$

where $\tilde{A}_k \subseteq A_k$, and $|\tilde{A}_k| < \kappa$.



The approximation

- Austad and Tjelmeland, 2011, develop a computationally feasable approximation of the above algorithm.
- An approximate factorization is obtained



$$\widetilde{p}_\theta(x) = \widetilde{p}_\theta(x_n) \prod_{k=1}^{n-1} \widetilde{p}_\theta(x_k | x_{\tilde{A}_k}),$$

where $\tilde{A}_k \subseteq A_k$, and $|\tilde{A}_k| < \kappa$.

- The computational complexity is reduced to $2^{\kappa+1}$

Multi-grid MRF

- A large neighborhood is required to capture the structure in the training images.
- A large neighbourhood increases the size of A_k and increases the error of the approximation for a given value of κ.
- We split the lattice into T sublattices $S_1, ..., S_T$.

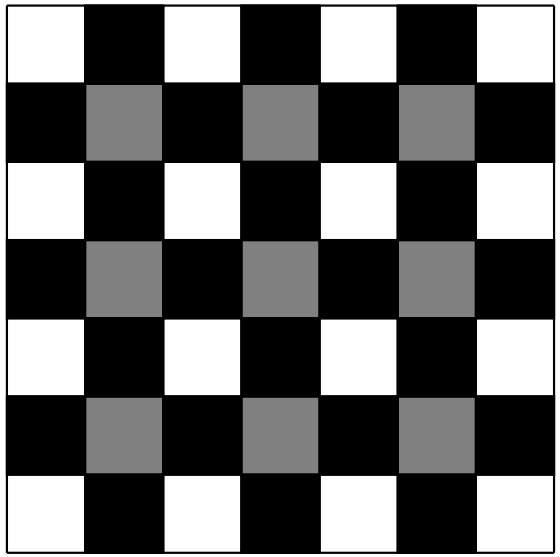


Figure: T = 3

- The joint distribution for x is specified by the marginal of x_{S_1} and the conditionals of x_{S_t} given $x_{S_{1:t-1}}$

$$\begin{aligned} p_\theta(x) = & \quad p_{\theta_1}(x_{S_1}) \prod_{t=2}^{T} p_{\theta_t}(x_{S_t} | x_{S_{1:t-1}}) \\ = & \quad c(\theta_1) \exp\left\{-U_{\theta_1}(x_{S_1})\right\} \\ & \prod_{t=2}^{T} c(\theta_t, x_{S_{1:t-1}}) \exp\left\{-U_{\theta_t}(x_{S_1}, x_{S_{1:t-1}})\right\} \end{aligned}$$

- The energy function has the form

$$U_{\theta_t}(x_{S_t}, x_{S_{1:t-1}}) = U^1_{\theta_t,1:9}(x_{S_t}) + \sum_{i \in S_t} U^2_{\theta_t,10:19}(x_i, x_{B_i})$$

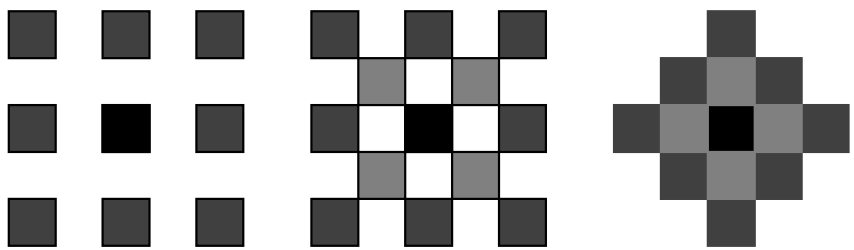
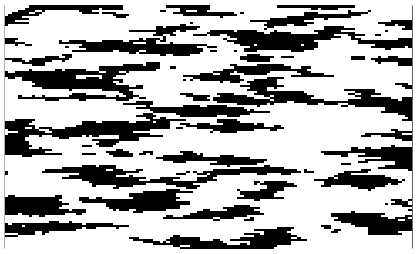
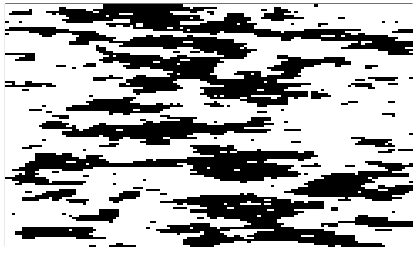


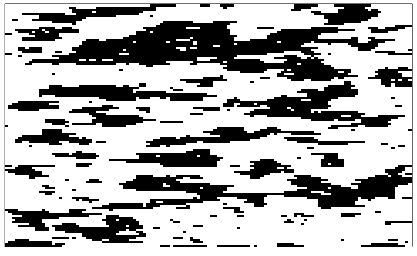
Figure: The dark gray nodes in S_t which are neighbors of the black node. The light gray are nodes from $S_{1:t-1}$ in B_i

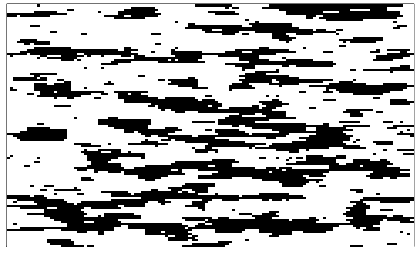
- Approximate simulation from $p_\theta(x)$ is done sequentially by first drawing x_{S_1} from $\tilde{p}_{\theta_1}(x_{S_1})$ and then x_{S_2} from $\tilde{p}_{\theta_2}(x_{S_2}|x_{S_1})$ and so on.
- In the examples we have used $T = 7$ sublattices.

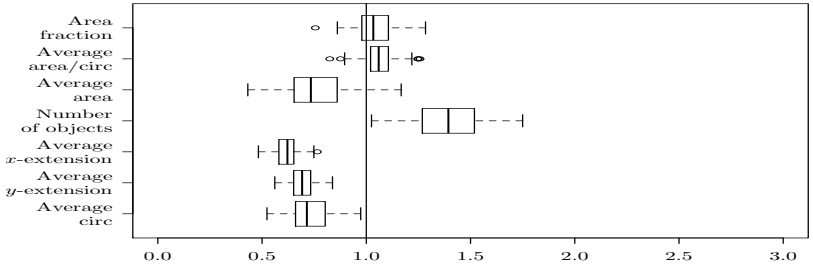
Unconditional simulation

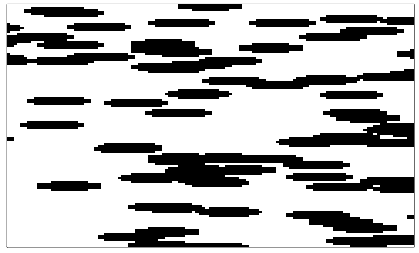


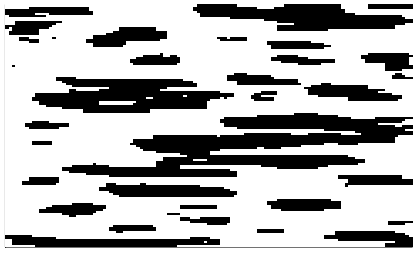


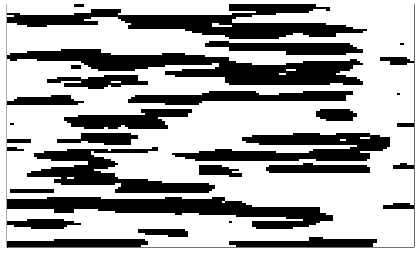


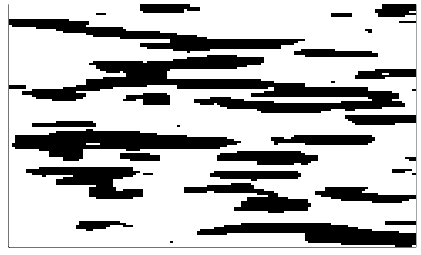


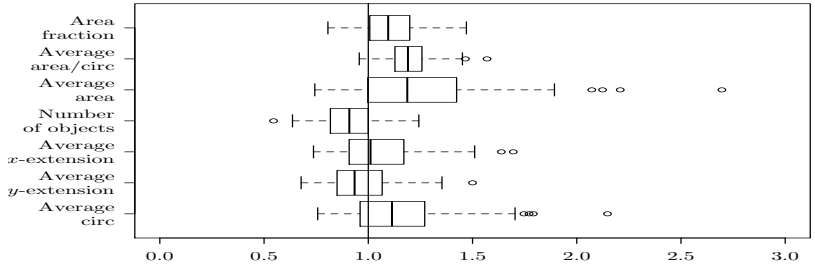


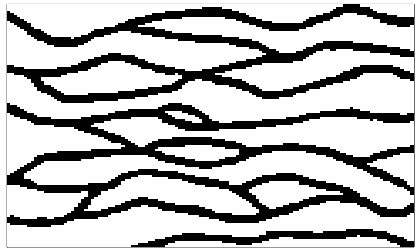


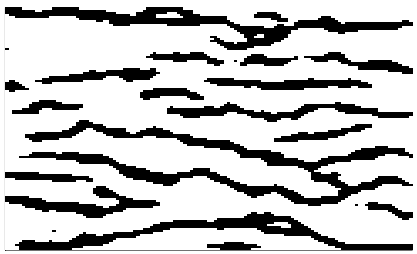


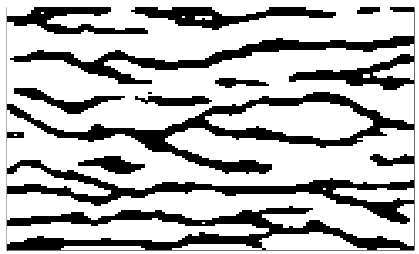


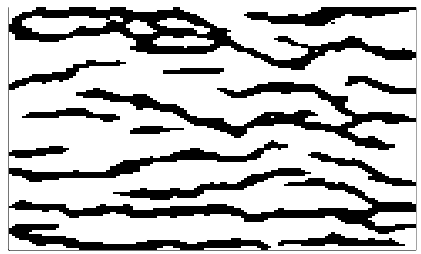


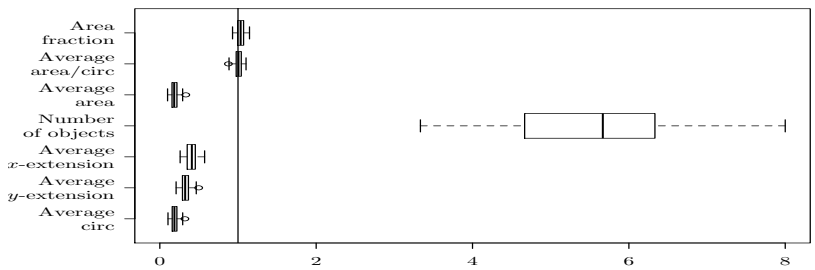

Area fraction
Average area/circ
Average area
Number of objects
Average x-extension
Average y-extension
Average circ
0.0
0.5
1.0
1.5
2.0
2.5
3.0










Area fraction
Average area/circ
Average area
Number of objects
Average x-extension
Average y-extension
Average circ
0
2
4
6
8

Conditional simulation

- Conditional simulation is complicated by the unknown normalizing constants $c(\theta_t, x_{S_{1:t-1}})$.
- To solve this problem we apply the approximate forward-backward algorithm to $p_{\theta_t}(x_{S_t}|x_{S_{1_t-1}})$, and obtain $\tilde{c}(\theta_t, x_{S_{1:t-1}})$
- The resulting approximation is

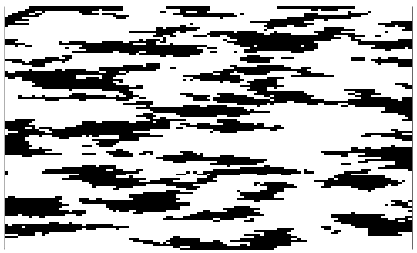
$$\widetilde{p}_\theta(x) = \tilde{c}(\theta_1)\exp\left\{-U_{\theta_1}(x_{S_1})\right\}\prod_{t=2}^{T}\widetilde{c}(\theta_t, x_{S_{1:t-1}})\exp\left\{-U_{\theta_t}(x_{S_{1:t}})\right\},$$

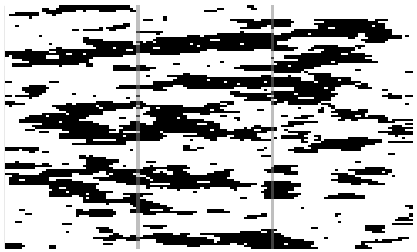
- let z denote a vector of observed quantities which is related to x by a likelihood function $h(z|x)$.
- For the approximation $\widetilde{p}_\theta(x)$ we get a corresponding approximations to $p_\theta(x|z)$

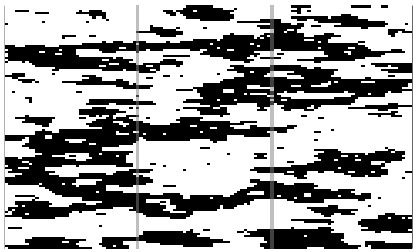
$$\widetilde{p}_\theta(x|z) \propto \widetilde{p}_\theta(x)h(z|x)$$

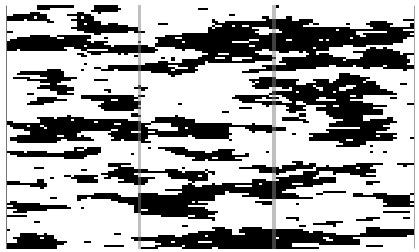
- The approximate distribution $\widetilde{p}_\theta(x|z)$ is an MRF, and we can apply the approximate forward-backward algorithm again.

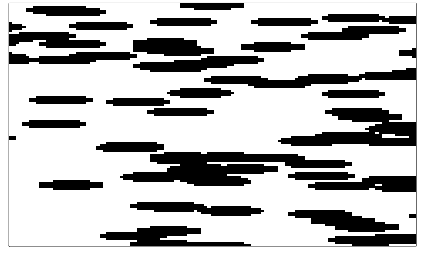
Conditioning on 2 wells

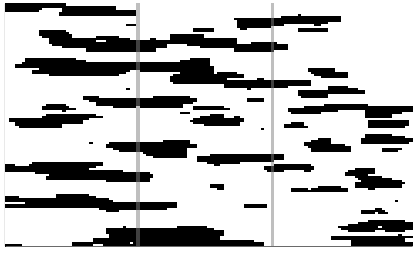


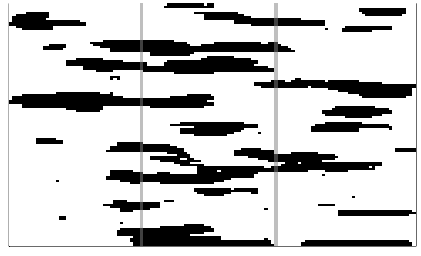


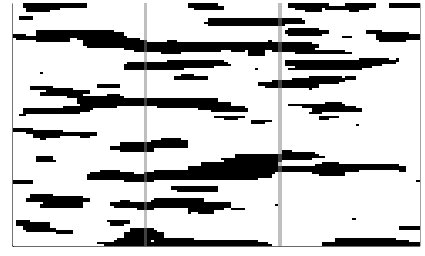


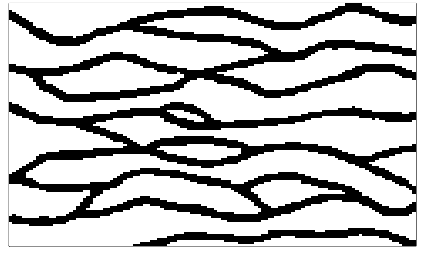


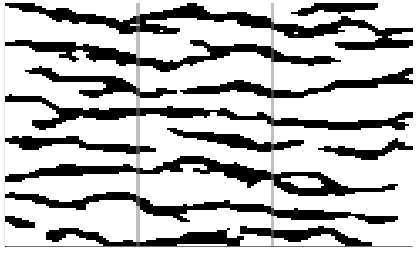


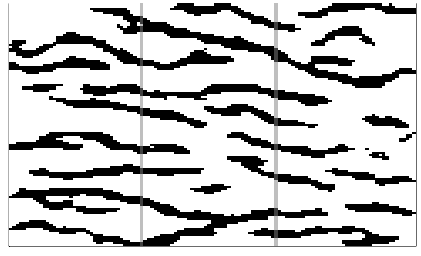


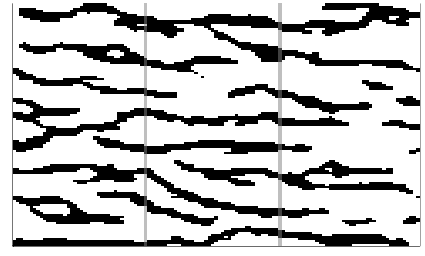












- The method seems to be quite flexible.
- Conditional simulation is straight forward if the likelihood is easy to compute.
- A Generalization to 3D is simple, but would probably be computationally unfeasible.
- More than 2 colors is also straight forward but computational complexity increases fast.