Block updating of spatial fields

Ingelin Steinsland & Håvard Rue

ingelins@math.ntnu.no & hrue@math.ntnu.no

Norwegian University of Science and Technology

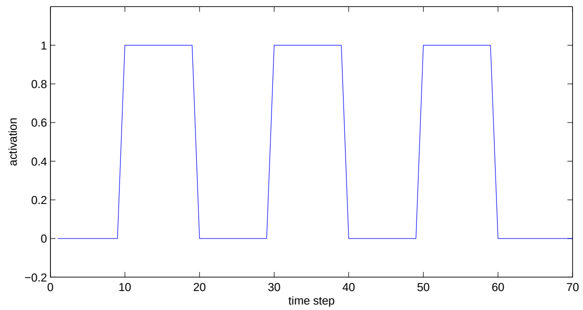
Outline

- Introduce the fMRI example
- Problems with traditional block samplers
- Overlapping blocks
- Toy examples overlapping blocks
- Solution fMRI

fMRI

functional Magnetic Resonance Imaging -
Data from a visual stimulation experiment.

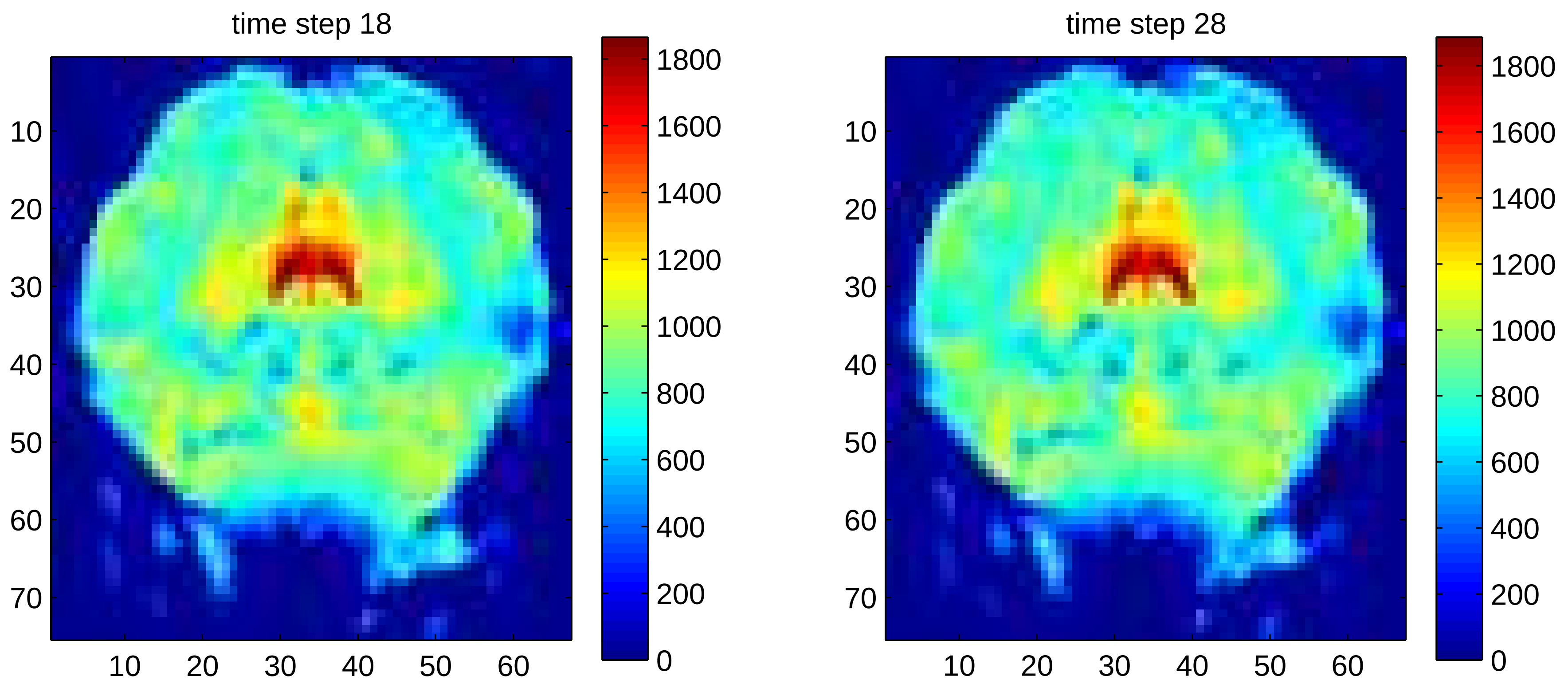
- Stimulus: 8 Hz flickering checkerboard
- 4 period (a 30 sec.) rest, 3 periods stimulus



- Cross section of the brain observed every 3rd sec.
- observe BOLD effects

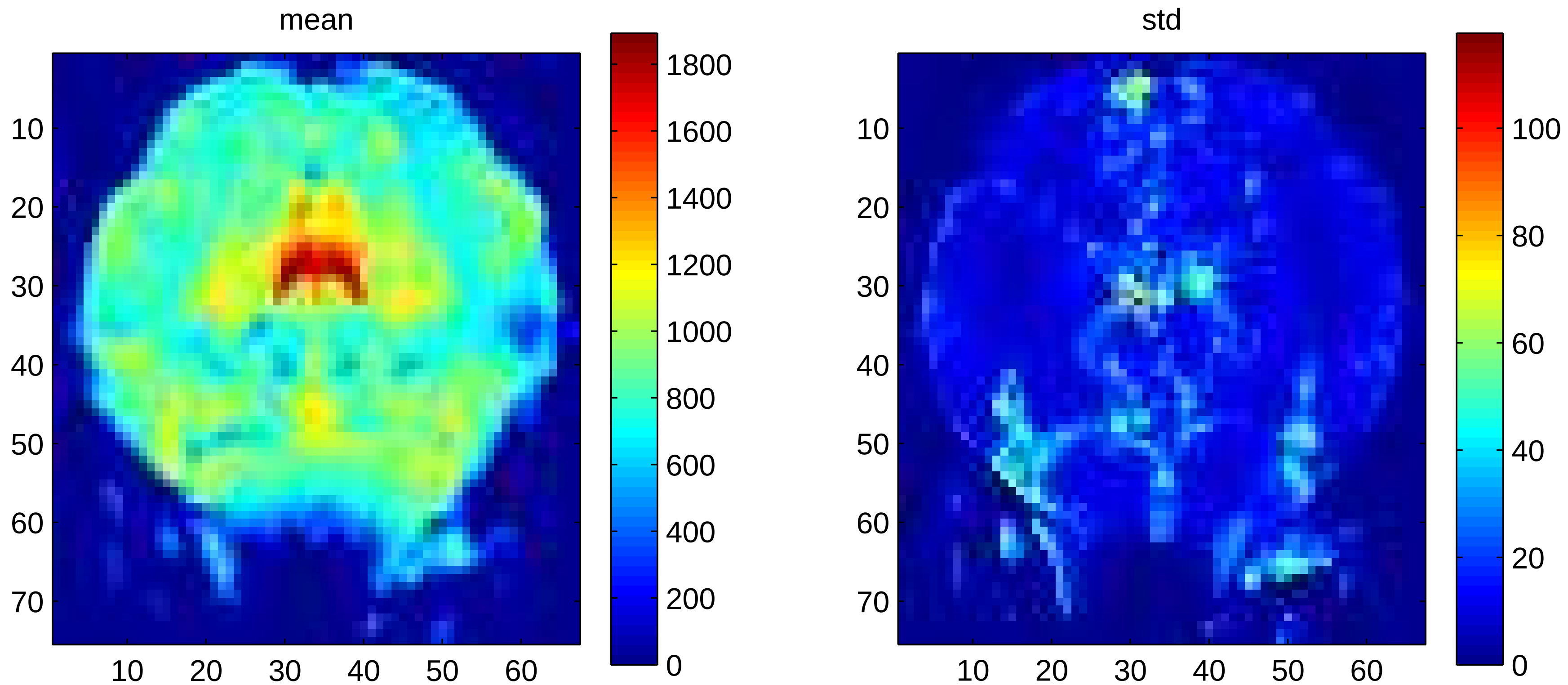
fMRI

functional Magnetic Resonance Imaging -
Data from a visual stimulation experiment.



fMRI

functional Magnetic Resonance Imaging -
Data from a visual stimulation experiment.



Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- y_{it}: Data in pixel i at time step t

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- y_{it}: Data in pixel i at time step t
- a_i: Baseline image, pixel i

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- y_{it}: Data in pixel i at time step t
- a_i: Baseline image, pixel i
- z_t: Transformed stimulus time step t

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- y_{it}: Data in pixel i at time step t
- a_i: Baseline image, pixel i
- z_t: Transformed stimulus time step t
- b_{it}: Activation effect of pixel i time step t

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- y_{it}: Data in pixel i at time step t
- a_i: Baseline image, pixel i
- z_t: Transformed stimulus time step t
- b_{it}: Activation effect of pixel i time step t
- ϵ_{it}: Measurement error of pixel i time step t

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- $\epsilon \sim N(0, \tau_{Data} I) \rightarrow$
  $y_{it} | a, b \sim N(a_i + z_t b_{it}, \tau_{Data})$

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- $y_{it} | a, b \sim N(a_i + z_t b_{it}, \tau_{Data})$
- z use estimate from similar studies

Model

$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- $y_{it}|a, b \sim N(a_i + z_t b_{it}, \tau_{Data})$
- GMRF for a:
$$\pi(a) \propto \exp(-\frac{1}{2}\tau_A \sum_{t=1}^{T} \sum_{i \overset{s}{\sim} j} (a_i - a_j)^2)$$

Model

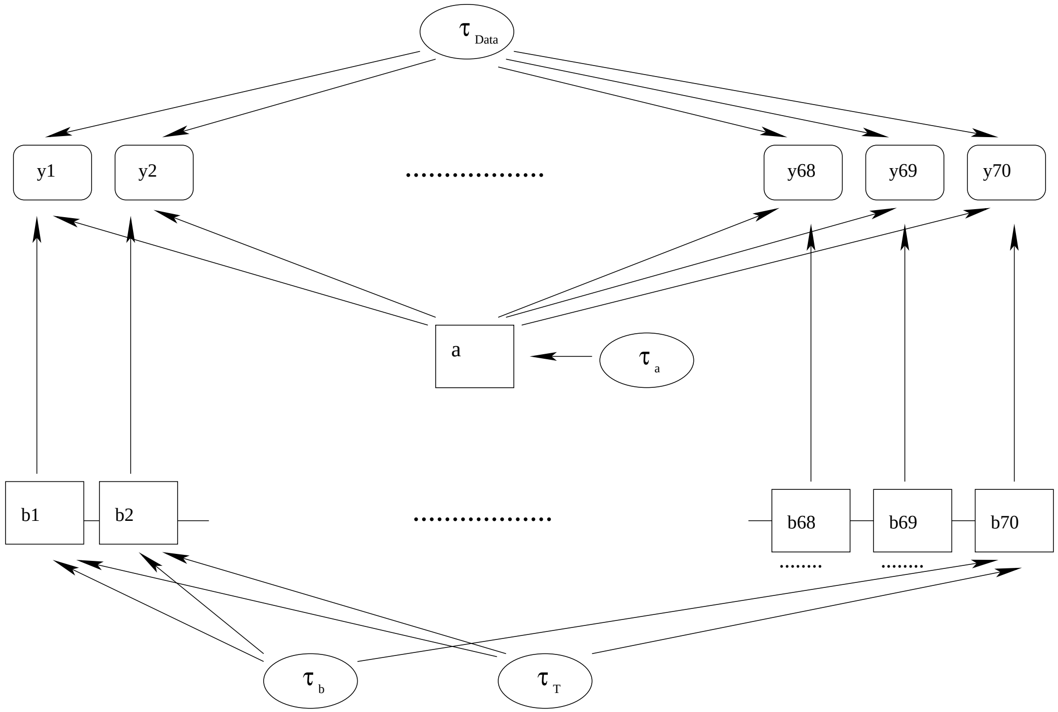
$$y_{it} = a_i + z_t b_{it} + \epsilon_{it}$$

- $y_{it}|a,b \sim N(a_i + z_t b_{it}, \tau_{Data})$
- GMRF for a:
  $$\pi(a) \propto \exp(-\tfrac{1}{2}\tau_A \textstyle\sum_{t=1}^{T} \sum_{i \overset{s}{\sim} j} (a_i - a_j)^2)$$
- Time-space GMRF for b:

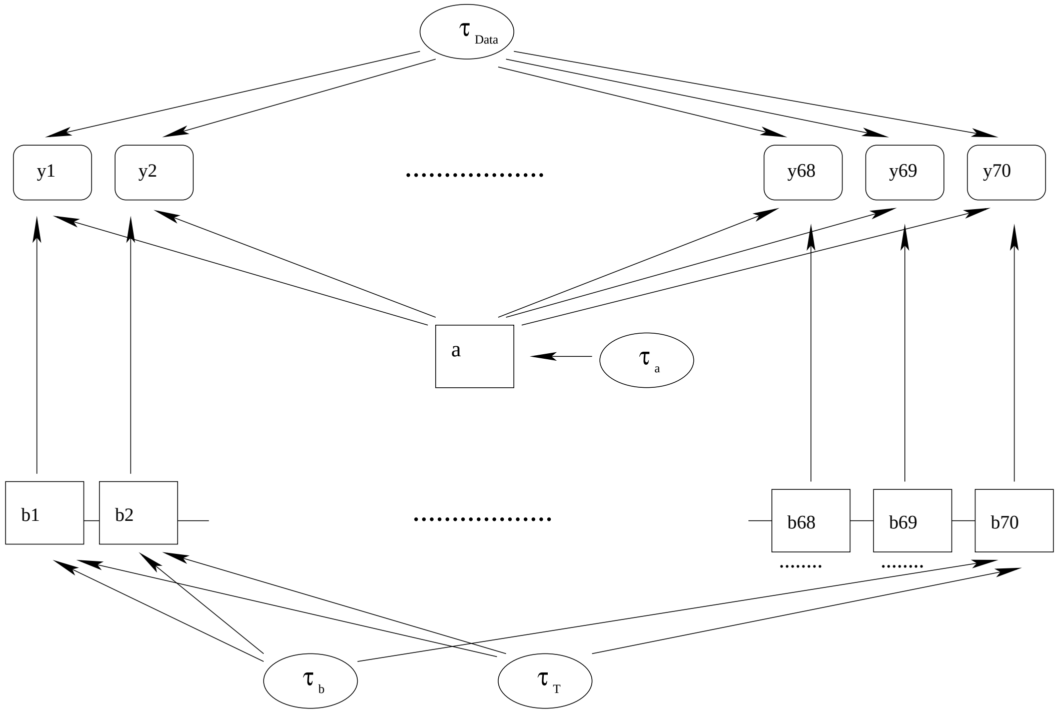
$$\pi(b) \propto \exp(-\frac{1}{2}\tau_B \sum_{t=1}^{T} \sum_{i \overset{s}{\sim} j} (b_{it} - b_{jt})^2)$$

$$\exp(-\frac{1}{2}\tau_T \sum_{i=1}^{N} \sum_{t \overset{t}{\sim} r} (b_{it} - b_{ir})^2)$$

DAG

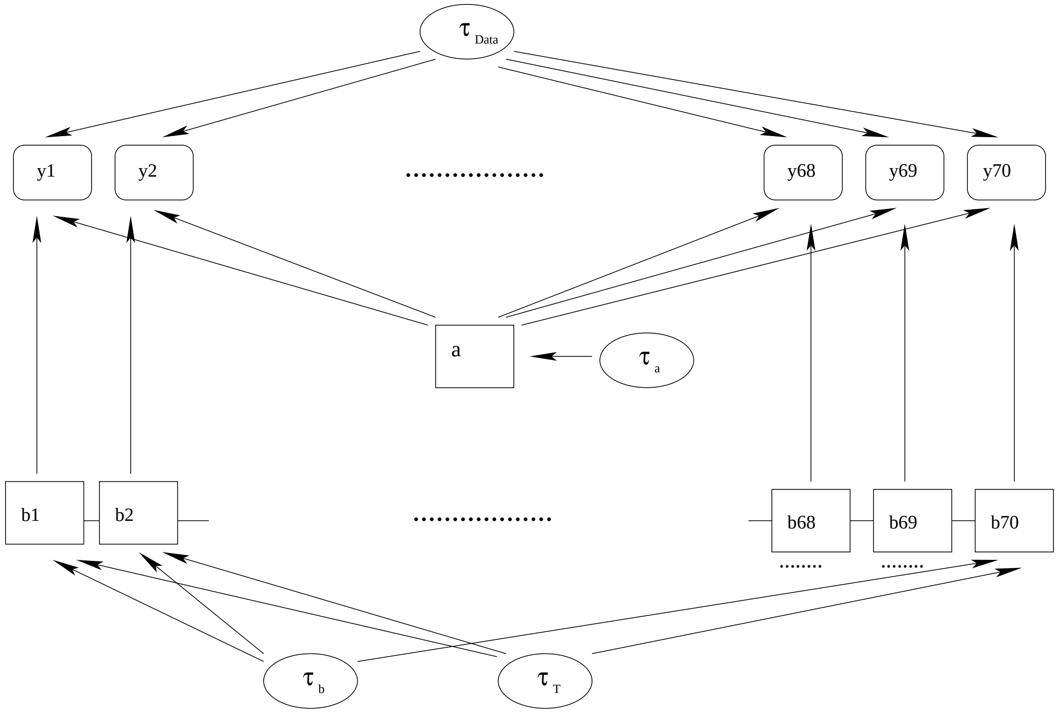


DAG



- Dimension (a, b) full problem 356 775.

DAG



- Dimension (a, b) full problem 356 775.
- Dimension (a, b) reduced problem 111 825.

MCMC

We have a latent field $x = (a, b)$ and hyper-parameters $\theta = (\tau_{data}, \tau_A, \tau_B, \tau_T)$

MCMC

- Given x^0 and θ^0

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$
  - Sample $x^{new} \sim \; q(x|x^j, \theta^{new})$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$
  - Sample $x^{new} \sim \ q(x|x^j, \theta^{new})$
  - accept / reject

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$
  - Sample $x^{new} \sim \ q(x|x^j, \theta^{new})$
  - accept / reject
  - if(accept)
    - $\theta^{j+1} = \theta^{new}$ and $x^{j+1} = x^{new}$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$
  - Sample $x^{new} \sim \; q(x|x^j, \theta^{new})$
  - accept / reject
  - if(accept)
    - $\theta^{j+1} = \theta^{new}$ and $x^{j+1} = x^{new}$
  - else
    - $\theta^{j+1} = \theta^j$ and $x^{j+1} = x^j$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
    - Sample $\theta^{new} \sim q(\theta|\theta^j)$
    - Sample $x^{new} \sim q(x|x^j, \theta^{new})$
    - accept / reject
    - if(accept)
        - $\theta^{j+1} = \theta^{new}$ and $x^{j+1} = x^{new}$
    - else
        - $\theta^{j+1} = \theta^j$ and $x^{j+1} = x^j$
- Return $(x^1, x^2, \ldots, x^{niter})$ and $(\theta^1, \theta^2, \ldots, \theta^{niter})$

MCMC

- Given x^0 and θ^0
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$
  - Sample $x^{new} \sim$ $\boxed{q(x|x^j, \theta^{new})}$
  - accept / reject
  - if(accept)
    - $\theta^{j+1} = \theta^{new}$ and $x^{j+1} = x^{new}$
  - else
    - $\theta^{j+1} = \theta^j$ and $x^{j+1} = x^j$
- Return $(x^1, x^2, \ldots, x^{niter})$ and $(\theta^1, \theta^2, \ldots, \theta^{niter})$

Challenge: To make a good and cheap proposal for x.

Traditional blocking

- x: 100×100 GMRF, $E(x) = 0$ but $x^0 = 3$
- 5×5 neighbourhood
- GMRF a approximation to correlation function:

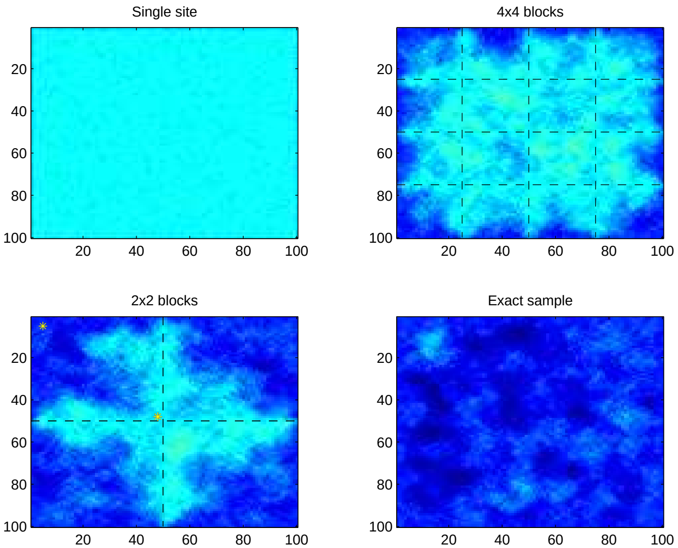
$$\rho(y_i, y_j) = \exp(\frac{-3d(x_i, x_j)}{r})$$

Traditional blocking

- x: 100×100 GMRF, $E(x) = 0$ but $x^0 = 3$
- 5×5 neighbourhood
- GMRF a approximation to correlation function:

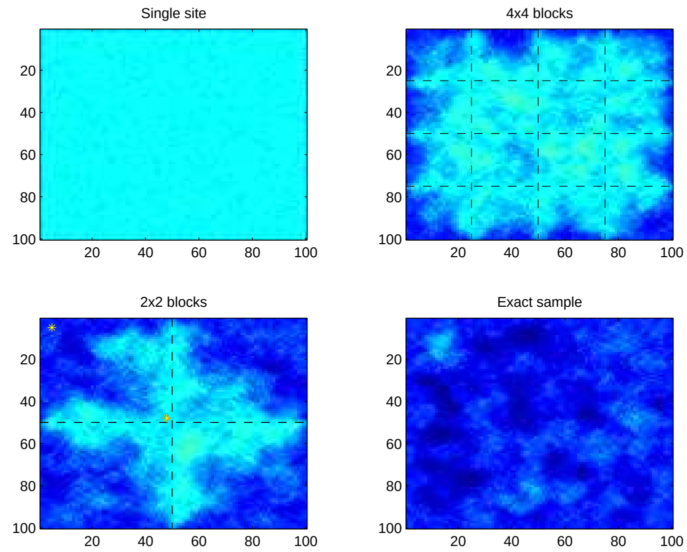
$$\rho(y_i, y_j) = \exp(\frac{-3d(x_i, x_j)}{r})$$

First iteration ($r = 40$)

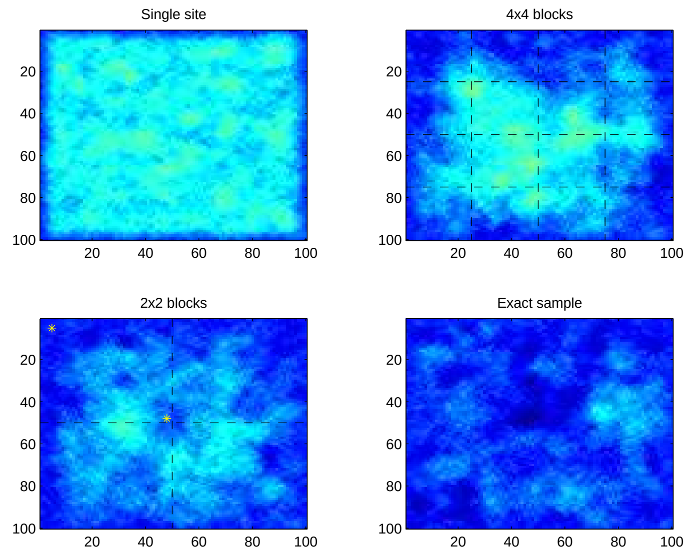


Traditional blocking

First iteration ($r = 40$)

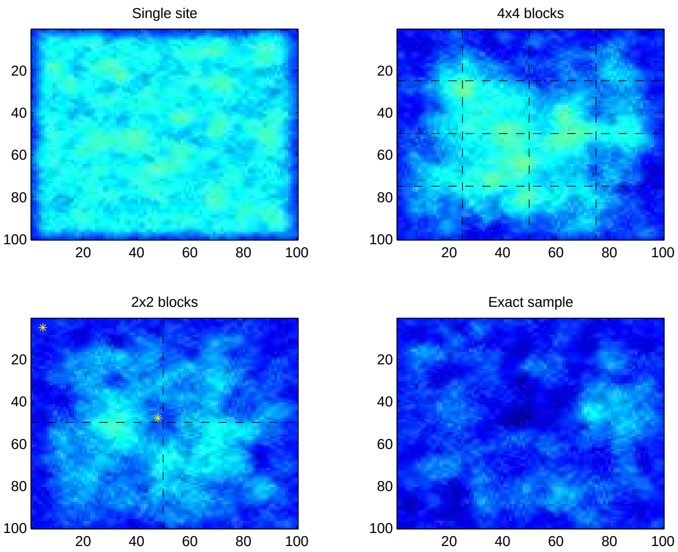


200th iteration ($r = 40$)

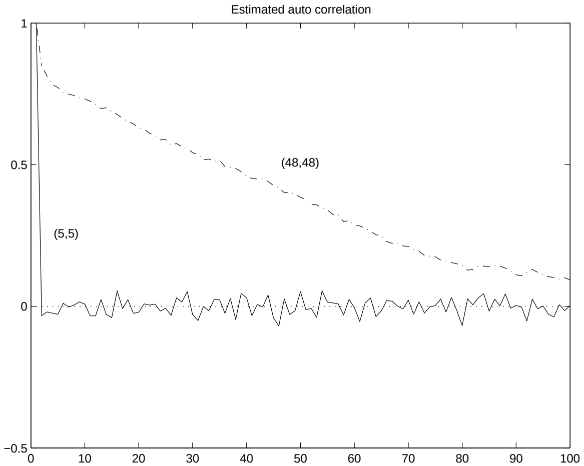


Traditional blocking

200th iteration ($r = 40$)



Estimated autocorrelation:

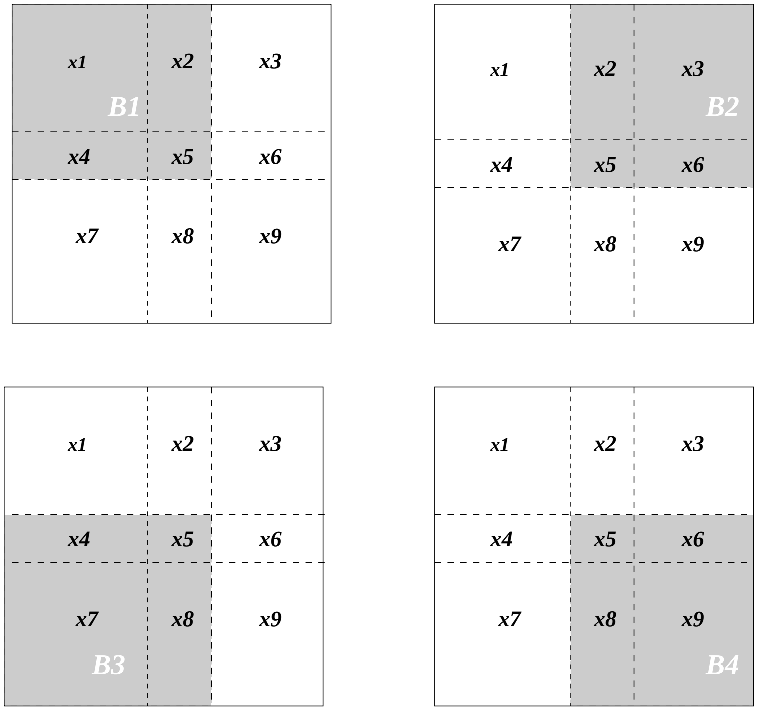


Overlapping blocks

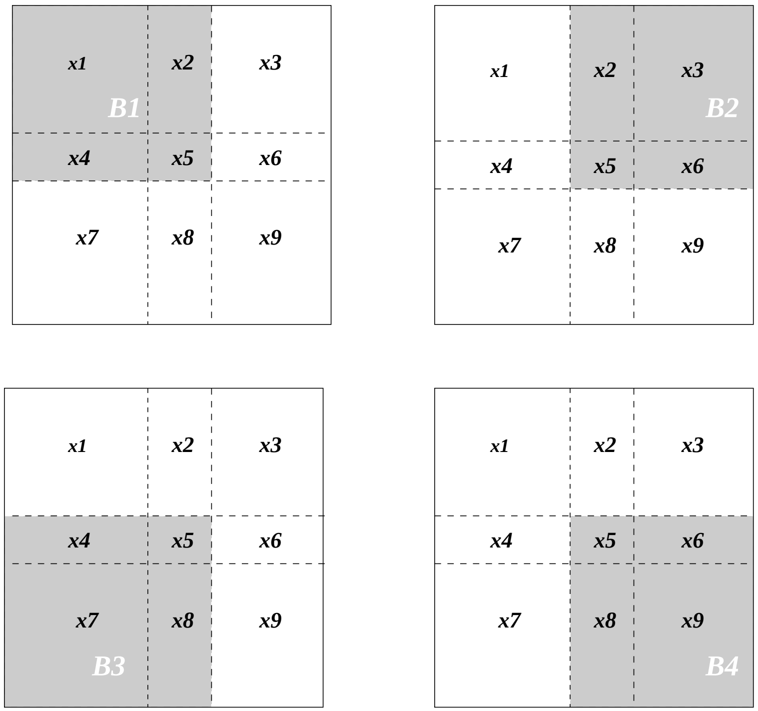
Idea: Let the blocks overlap.

Overlapping blocks

Idea: Let the blocks overlap.

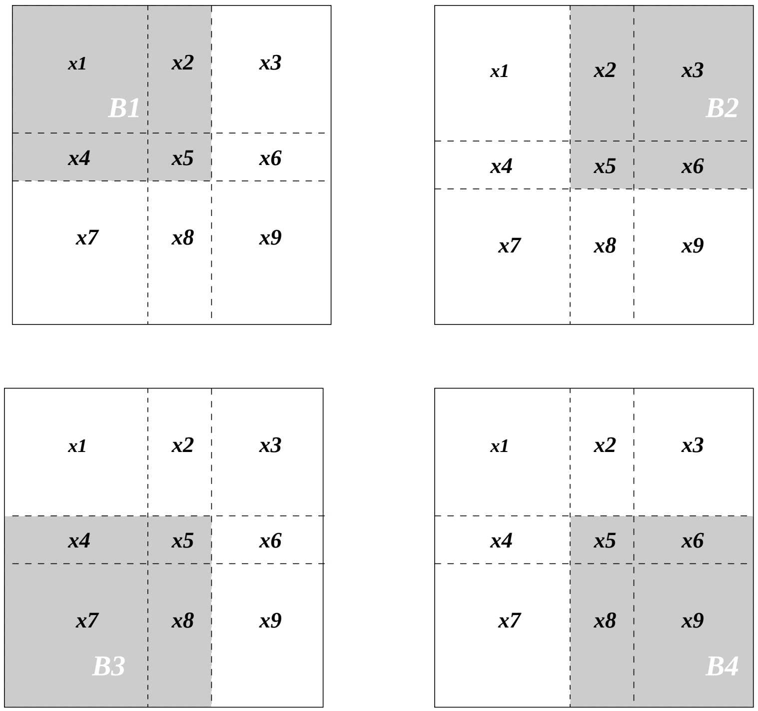


Overlapping blocks



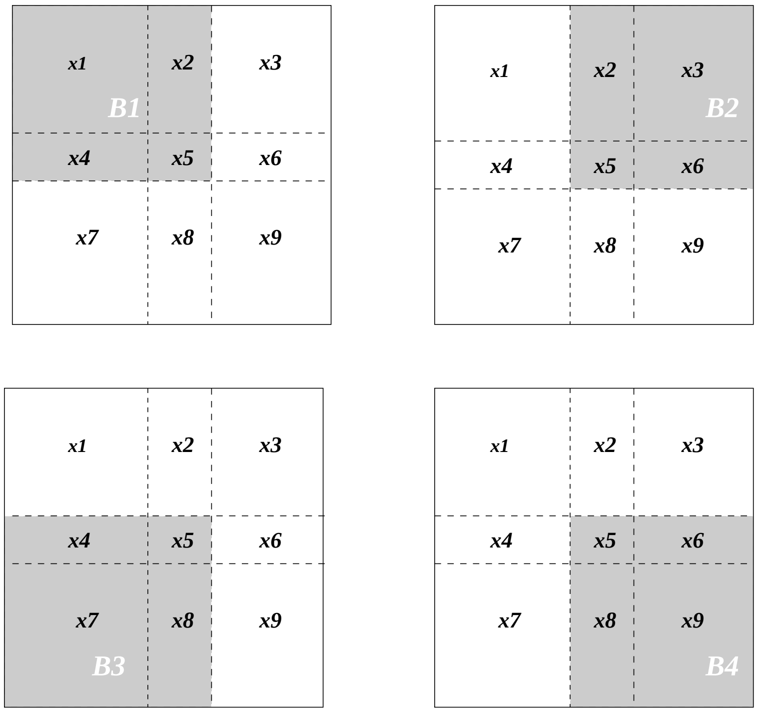
- Given x^i

Overlapping blocks



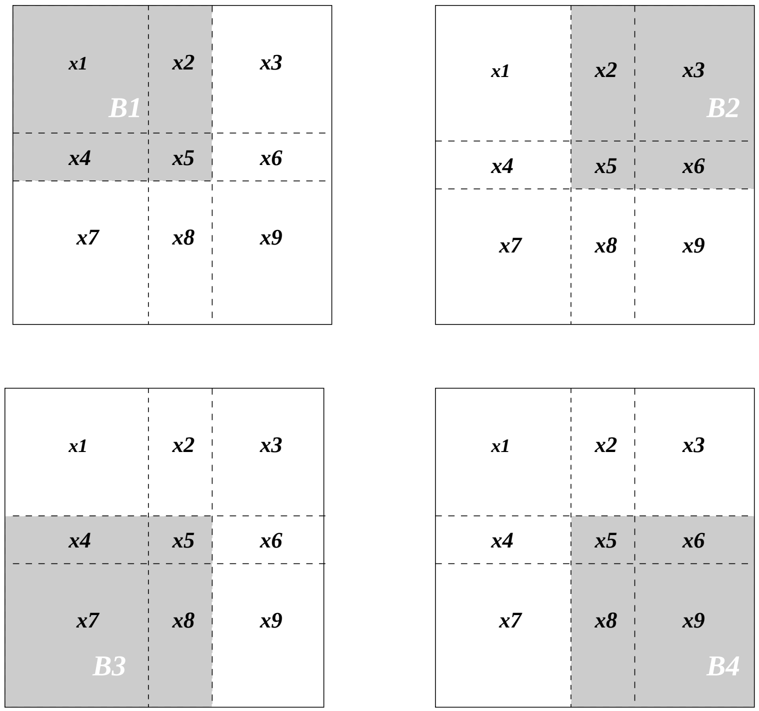
- Given x^i
  - Sample $(x_1^{i+1}, x_2^{B1}, x_4^{B1}, x_5^{B1}) \sim \pi_1(B_1 | x_3^i, x_6^i, x_7^i, x_8^i, x_9^i)$

Overlapping blocks



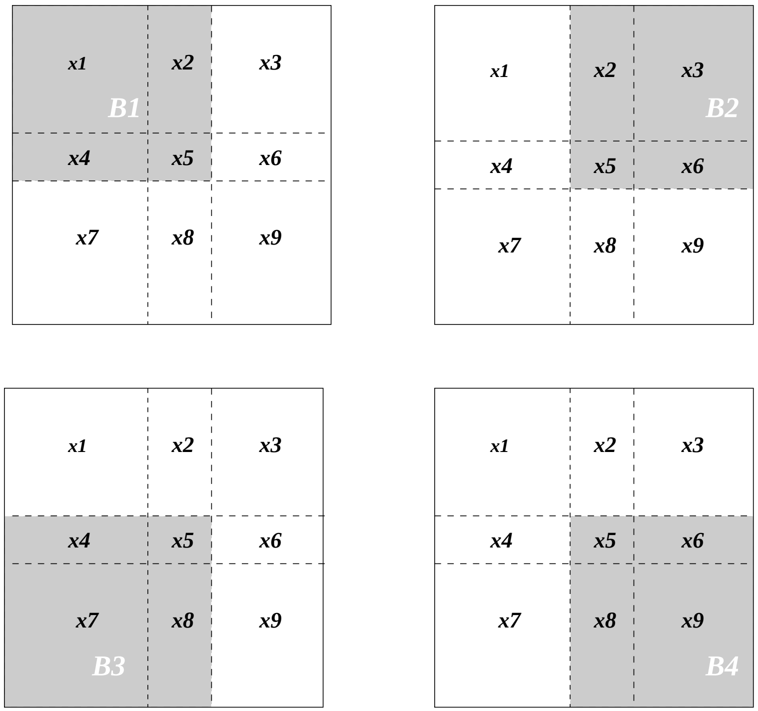
- Given x^i
  - Sample $(x_1^{i+1}, x_2^{B1}, x_4^{B1}, x_5^{B1}) \sim \pi_1(B_1 | x_3^i, x_6^i, x_7^i, x_8^i, x_9^i)$
  - Sample $(x_2^{i+1}, x_3^{i+1}, x_5^{B2}, x_6^{B2}) \sim \pi_2(B_2 | x_1^{i+1}, x_4^{B1}, x_7^i, x_8^i, x_9^i)$

Overlapping blocks



- Given x^i
  - Sample $(x_1^{i+1}, x_2^{B1}, x_4^{B1}, x_5^{B1}) \sim \pi_1(B_1|x_3^i, x_6^i, x_7^i, x_8^i, x_9^i)$
  - Sample $(x_2^{i+1}, x_3^{i+1}, x_5^{B2}, x_6^{B2}) \sim \pi_2(B_2|x_1^{i+1}, x_4^{B1}, x_7^i, x_8^i, x_9^i)$
  - Sample $(x_4^{i+1}, x_5^{B3}, x_7^{i+1}, x_8^{B3}) \sim \pi_3(B_3|x_1^{i+1}, x_2^{i+1}, x_3^{i+1}, x_6^{B2}, x_9^i)$
  - Sample $(x_5^{i+1}, x_6^{i+1}, x_8^{i+1}, x_9^{i+1}) \sim \pi_4(B_4|x_1^{i+1}, x_2^{i+1}, x_3^{i+1}, x_4^{i+1}, x_7^{i+1})$

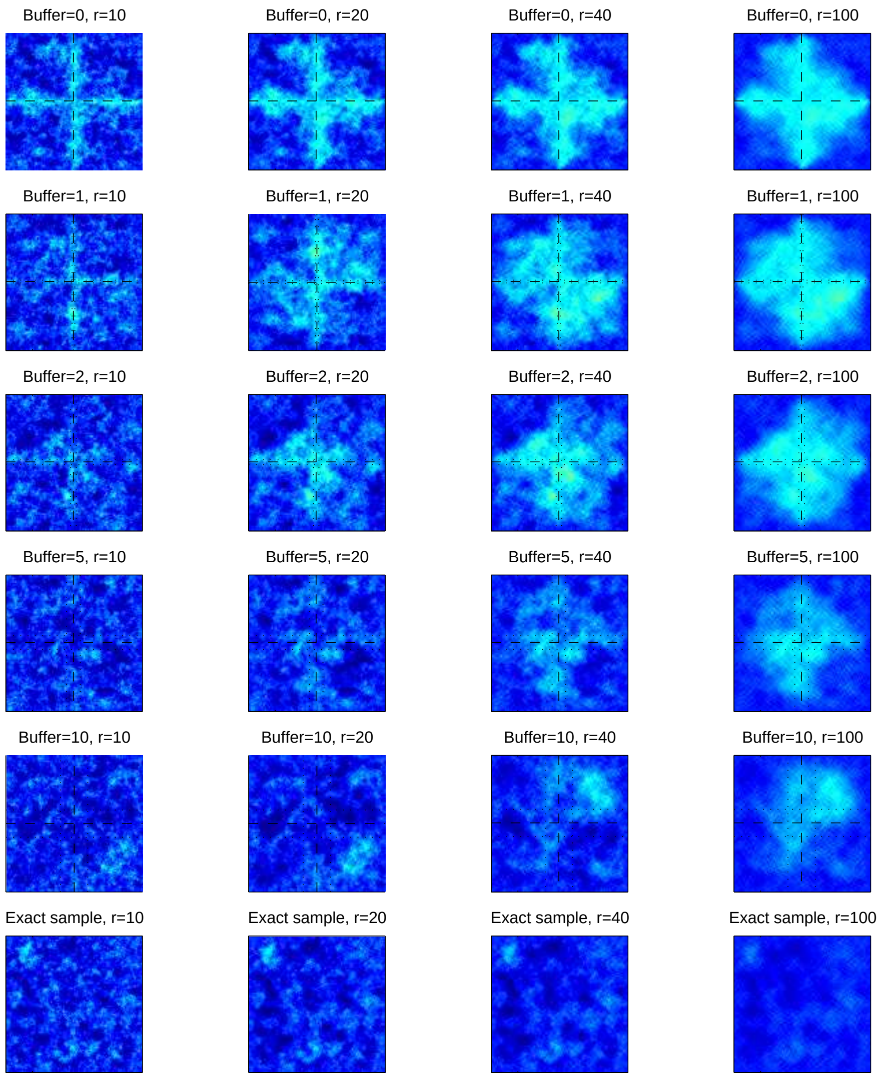
Overlapping blocks



- Given x^i
  - Sample $(x_1^{i+1}, x_2^{B1}, x_4^{B1}, x_5^{B1}) \sim \pi_1(B_1 | x_3^i, x_6^i, x_7^i, x_8^i, x_9^i)$
  - Sample $(x_2^{i+1}, x_3^{i+1}, x_5^{B2}, x_6^{B2}) \sim \pi_2(B_2 | x_1^{i+1}, x_4^{B1}, x_7^i, x_8^i, x_9^i)$
  - Sample $(x_4^{i+1}, x_5^{B3}, x_7^{i+1}, x_8^{B3}) \sim \pi_3(B_3 | x_1^{i+1}, x_2^{i+1}, x_3^{i+1}, x_6^{B2}, x_9^i)$
  - Sample $(x_5^{i+1}, x_6^{i+1}, x_8^{i+1}, x_9^{i+1}) \sim \pi_4(B_4 | x_1^{i+1}, x_2^{i+1}, x_3^{i+1}, x_4^{i+1}, x_7^{i+1})$
- Return x^{i+1}

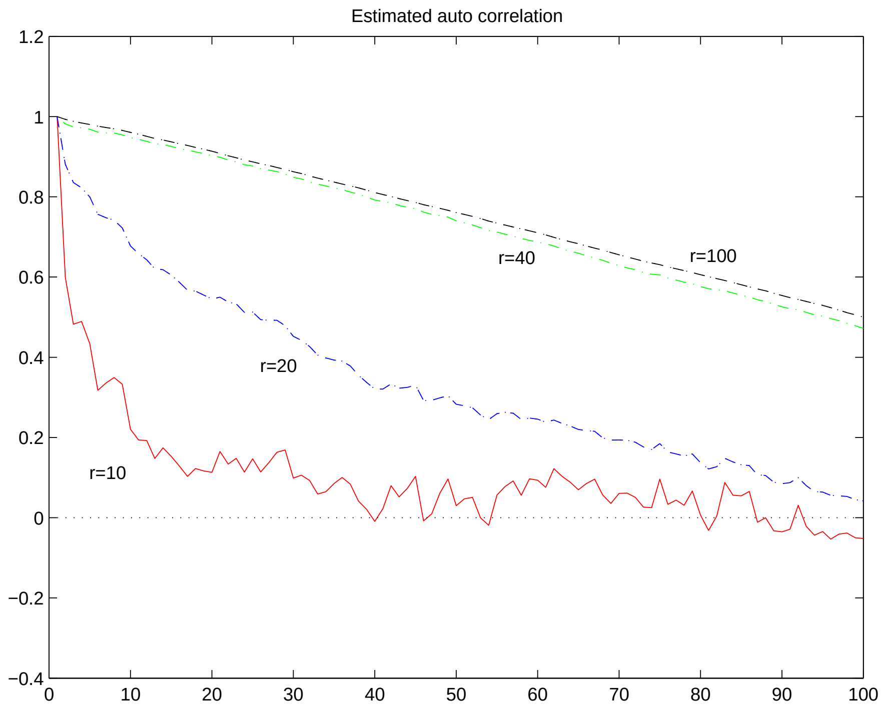
Does it work?

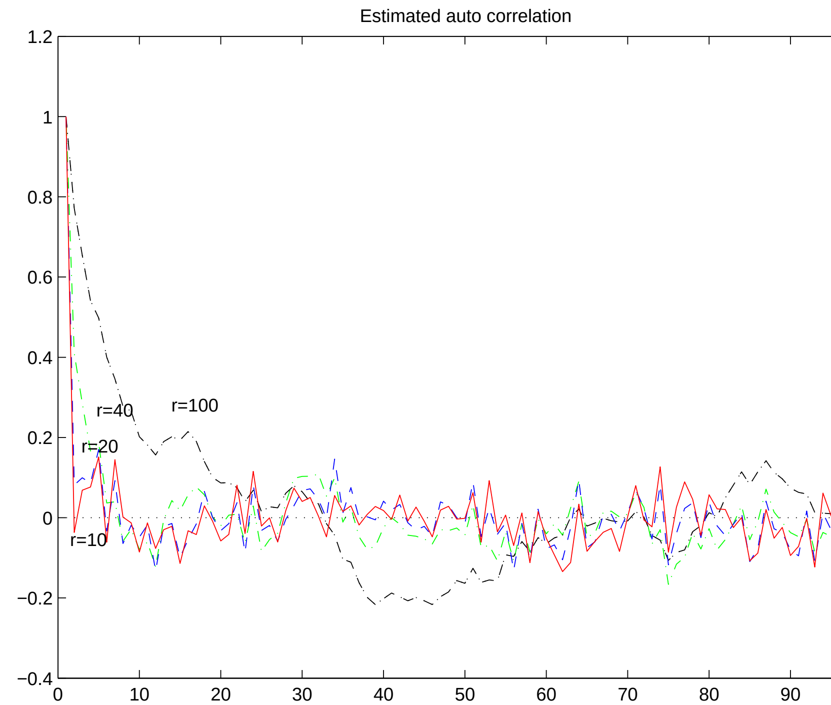
As previous example with $r = \{10, 20, 40, 100\}$ and buffer $= \{0, 1, 2, 3, 10\}$



Does it work?

Estimated auto-correlation function at pixel $(48, 48)$.
Left:block Gibbs without buffers
Right: with buffer five



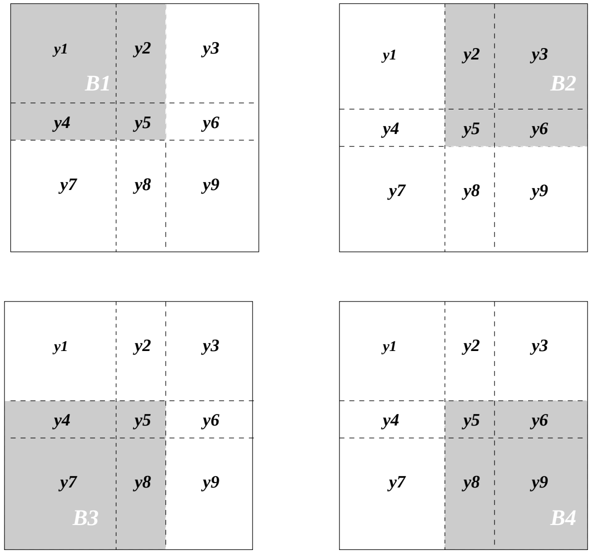


Transition probability

Hard to calculate the transition probability:

Transition probability

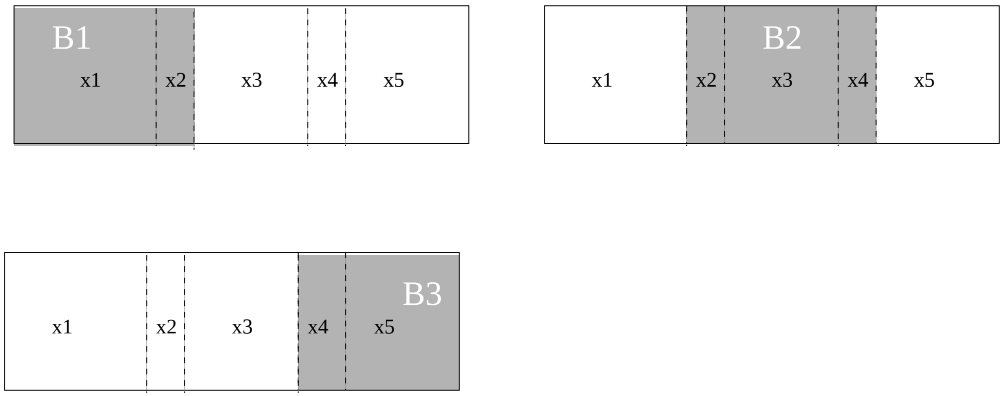
Hard to calculate the transition probability:



$$
\begin{aligned}
q(x|x^{'}) \quad = \quad & \int [\pi_1(x_1^{'}, x_2^{B1}, x_4^{B1}, x_5^{B1} | x_3, x_6, x_7, x_8, x_9) \\
& \pi_2(x_2^{'}, x_3^{'}, x_5^{B2}, x_6^{B2} | x_1^{'}, x_4^{B1}, x_7, x_8, x_9) \\
& \pi_3(x_4^{'}, x_5^{B3}, x_7^{'}, x_8^{B3} | x_1^{'}, x_2^{'}, x_3^{'}, x_6^{B2}, x_9) \\
& \pi_4(x_5^{'}, x_6^{'}, x_8^{'}, x_9^{'} | x_1^{'}, x_2^{'}, x_3^{'}, x_4^{'}, x_7^{'})] dx_2^{B1} dx_4^{B1} dx_5^{B1} dx_5^{B2} dx_5^{B3} dx_6^{B2} dx_8^{B3}
\end{aligned}
$$

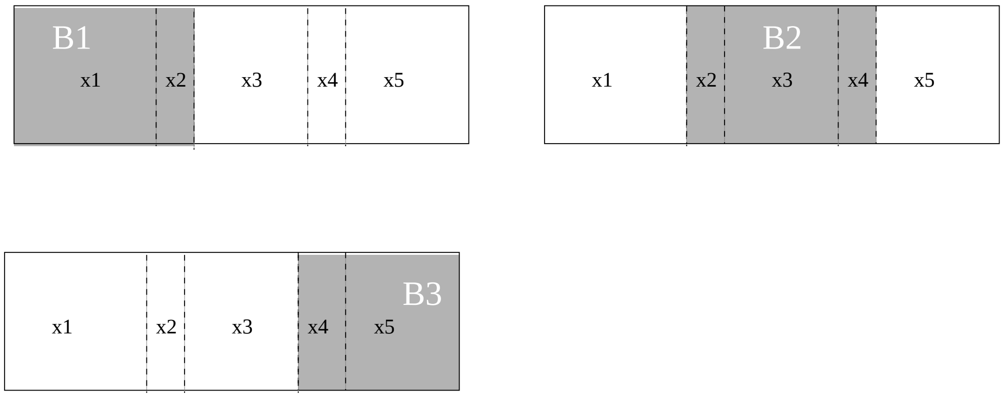
Transition probability

Time series blocking:



Transition probability

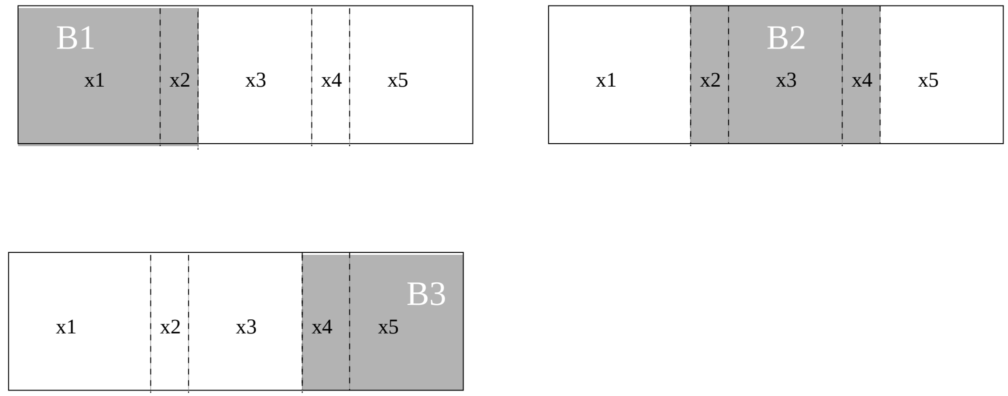
Time series blocking:



$$
\begin{aligned}
q(x|x') \quad = \quad & \int [\pi_1(x_1', x_2^{B1}|x_3, x_4, x_5) \\
& \pi_2(x_2', x_3', x_4^{B2}|x_1', x_5) \\
& \pi_3(x_4', x_5'|x_1', x_2', x_3')] dx_2^{B1} dx_4^{B2}
\end{aligned}
$$

Transition probability

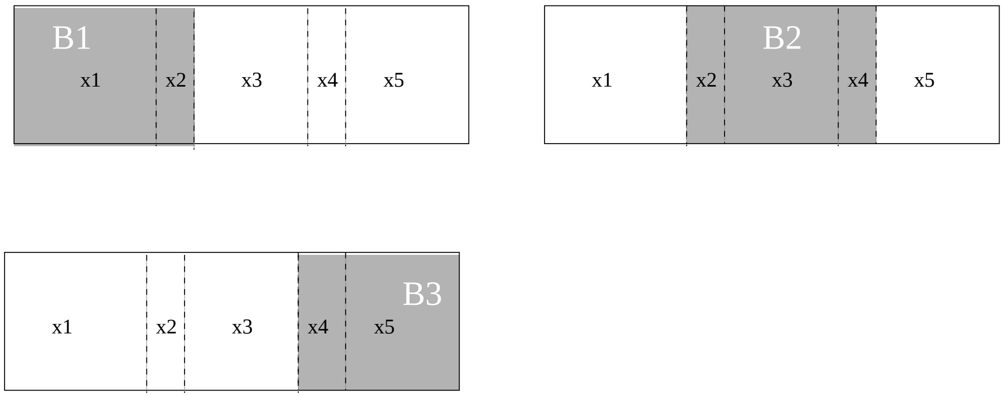
Time series blocking:



$$
\begin{aligned}
q(x|x') &= \int [\pi_1(x_1', x_2^{B1}|x_3, x_4, x_5) \\
&\quad \pi_2(x_2', x_3', x_4^{B2}|x_1', x_5) \\
&\quad \pi_3(x_4', x_5'|x_1', x_2', x_3')] dx_2^{B1} dx_4^{B2} \\
&= \pi(x_1'|x_3, x_4, x_5) \cdot \pi(x_2', x_3'|x_1', x_5) \cdot \pi(x_4', x_5'|x_1', x_2', x_3')
\end{aligned}
$$

Transition probability

Time series blocking:



$$q(x'|x) = \pi(x_1'|x_3, x_4, x_5) \cdot \pi(x_2', x_3'|x_1', x_5) \cdot \pi(x_4', x_5'|x_1', x_2', x_3')$$

Can use that:

$$\pi(x_1'|x_3, x_4, x_5) = \frac{\pi(x_1', x_2^{B1}|x_3, x_4, x_5)}{\pi(x_2^{B1}|x_1', x_3, x_4, x_5)}$$

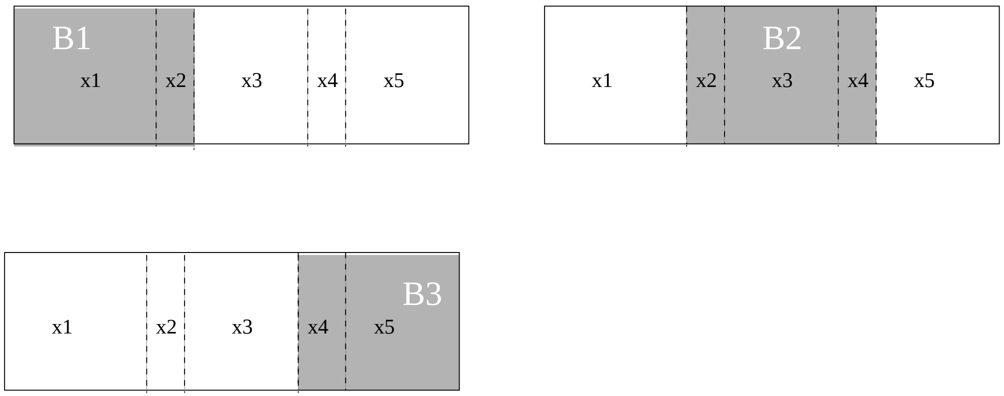
for any x_2^{B1}

Opposite reversal

- A M-H proposal constructed by Gibbs steps doesn't give acceptance rate 1.

Opposite reversal

- A M-H proposal constructed by Gibbs steps doesn't give acceptance rate 1.
- Sample first a direction $i = \{0, 1\}$
  - if $i == 0$ use $q_0 : B_1 \rightarrow B_2 \rightarrow B_3$
  - if $i == 1$ use $q_1 : B_3 \rightarrow B_2 \rightarrow B_1$



Opposite reversal

- A M-H proposal constructed by Gibbs steps doesn't give acceptance rate 1.
- Sample first a direction $i = \{0, 1\}$
  - if $i == 0$ use $q_0 : B_1 \longrightarrow B_2 \longrightarrow B_3$
  - if $i == 1$ use $q_1 : B_3 \longrightarrow B_2 \longrightarrow B_1$
- use acceptance rate

$$\alpha_{i,1-i}(y|x) = \min\left\{1, \frac{\pi(x')q_{1-i}(x|x')}{\pi(x)q_i(x'|x)}\right\}$$

Opposite reversal

- A M-H proposal constructed by Gibbs steps doesn't give acceptance rate 1.
- Sample first a direction $i = \{0, 1\}$
  - if $i == 0$ use $q_0 : B_1 \to B_2 \to B_3$
  - if $i == 1$ use $q_1 : B_3 \to B_2 \to B_1$
- use acceptance rate

$$\alpha_{i,1-i}(y|x) = \min\left\{1, \frac{\pi(x')q_{1-i}(x|x')}{\pi(x)q_i(x'|x)}\right\}$$

- This give $\alpha = 1$ for Gibbs steps

Toy example

Model:

- x: GMRF 100×10 with $E(x) = \beta$ and τ.
- 5×5 neighbourhood
- GMRF a approximation to correlation function:

$$\rho(y_i, y_j) = \exp(\frac{-3d(x_i, x_j)}{r})$$

Toy example

Model:

- x: GMRF 100×10 with $E(x) = \beta$ and τ.
- 5×5 neighbourhood
- GMRF a approximation to correlation function:

$$\rho(y_i, y_j) = \exp\left(\frac{-3d(x_i, x_j)}{r}\right)$$

- Hyper-parameters:
  - $\beta \sim N(0, 1)$
  - $r \sim \text{Unif}(1, 50)$
  - $\tau \sim \text{Gamma}(0.25, 0.05)$

Toy example

Algorithm

Toy example

Algorithm

- Given x^0 and $\theta^0 = (\tau^0, \beta^0, r^0)$

Toy example

Algorithm

- Given x^0 and $\theta^0 = (\tau^0, \beta^0, r^0)$
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$, r.w.

Toy example

Algorithm

- Given x^0 and $\theta^0 = (\tau^0, \beta^0, r^0)$
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$, r.w.
  - Sample $x^{new} \sim q(x|x^j, \theta^{new})$
    2 blocks with/ without overlap

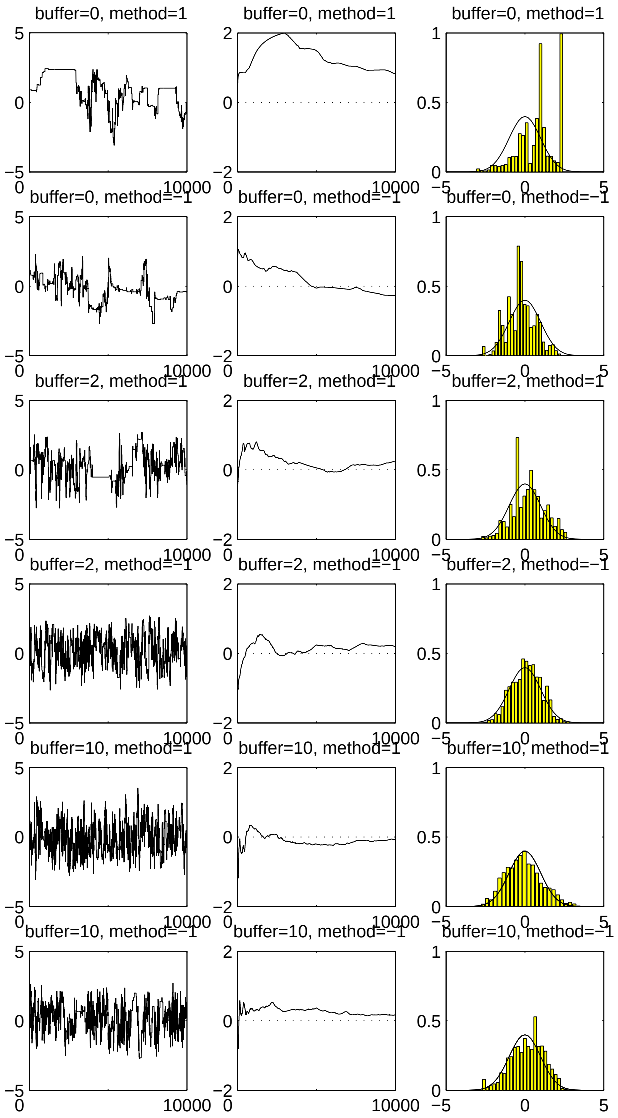
Toy example

Algorithm

- Given x^0 and $\theta^0 = (\tau^0, \beta^0, r^0)$
- for $j = 0 : (niter - 1)$
  - Sample $\theta^{new} \sim q(\theta|\theta^j)$, r.w.
  - Sample $x^{new} \sim q(x|x^j, \theta^{new})$
    2 blocks with/ without overlap
  - accept / reject
  - if(accept)
    - $\theta^{j+1} = \theta^{new}$ and $x^{j+1} = x^{new}$
  - else
    - $\theta^{j+1} = \theta^j$ and $x^{j+1} = x^j$
- Return $(x^1, x^2, \ldots, x^{niter})$ and $(\theta^1, \theta^2, \ldots, \theta^{niter})$

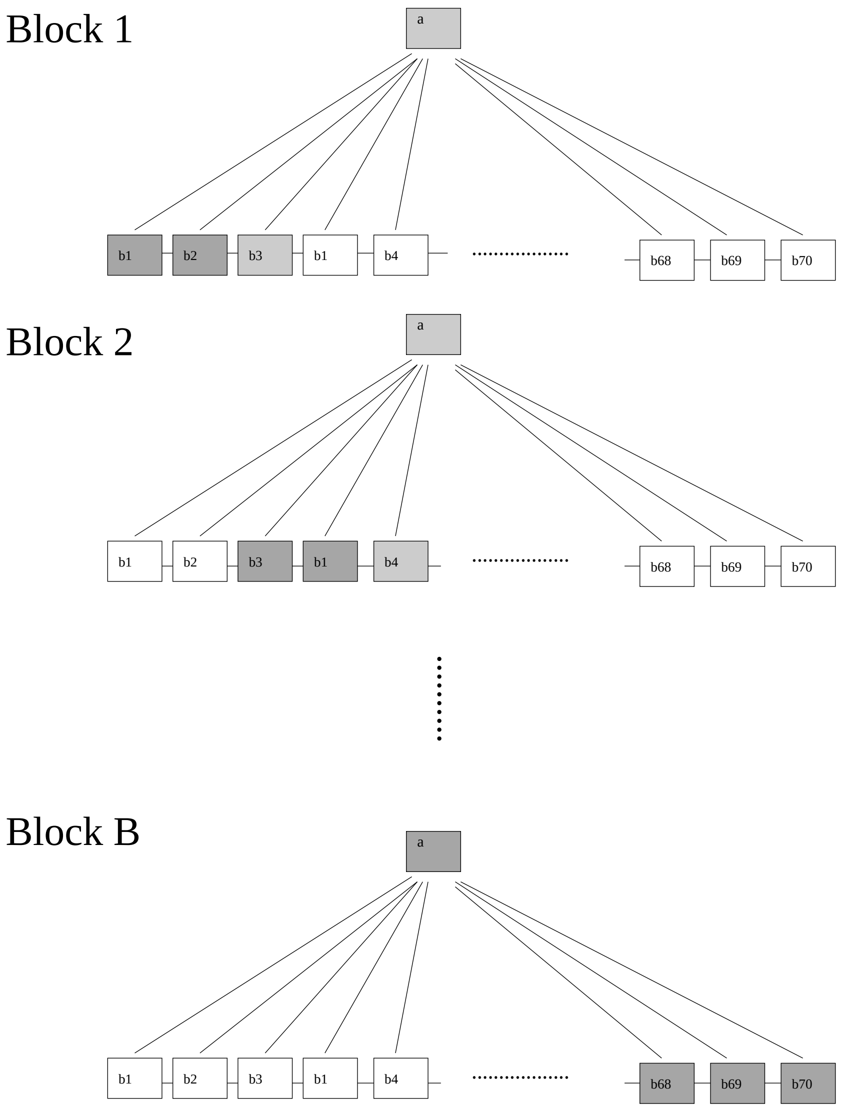
Toy example

Results mixing for β with buffer=$\{0, 2, 10\}$:



Solution fMRI

Sampling scheme:



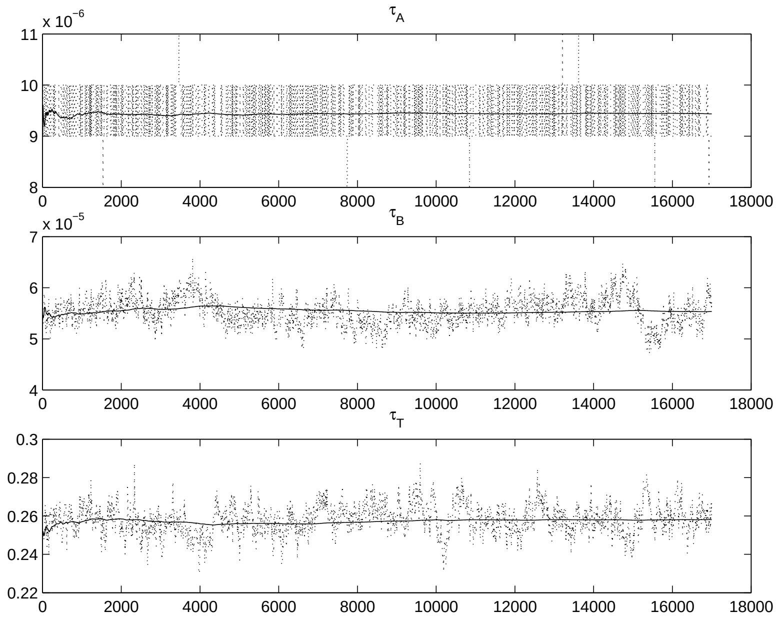
Results fMRI

- r.w. for hyper-parameters, τ_{Data} estimated beforehand
- Each block had a and five b_t
- Overlap: a and two b_t
- 20 000 iterations

Results fMRI

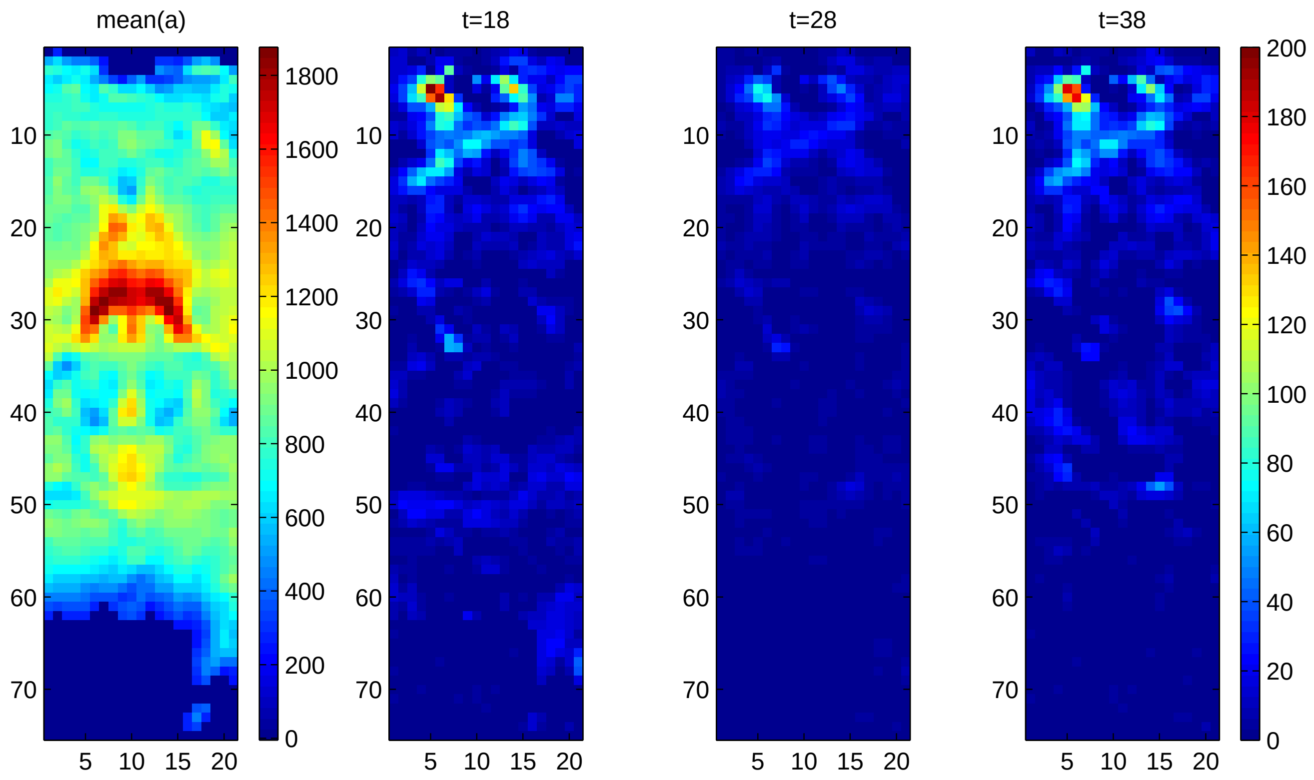
- r.w. for hyper-parameters, τ_{Data} estimated beforehand
- Each block had a and five b_t
- Overlap: a and two b_t
- 20 000 iterations

Trace plots hyper-parameters:



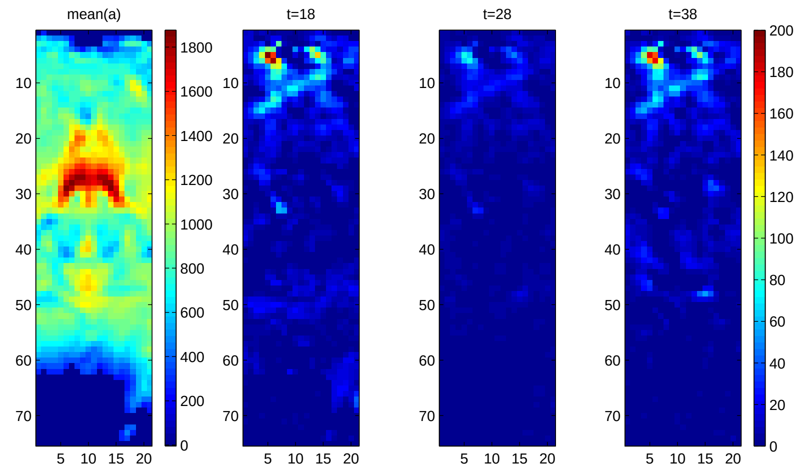
Results fMRI

Estimated mean a and b_{18}, b_{28} and b_{38}:

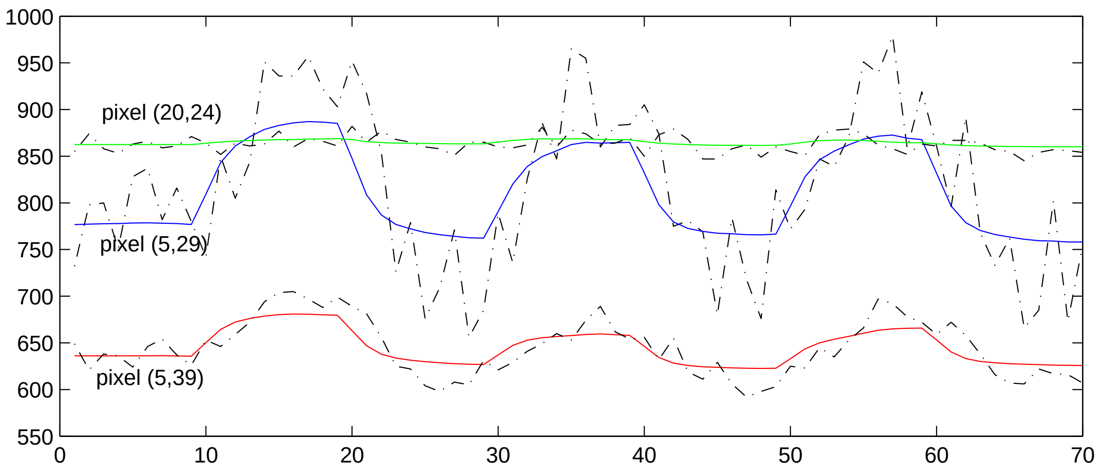


Results fMRI

Estimated mean a and b_{18}, b_{28} and b_{38}:



Estimate for some pixels in time:



Summary

- We have constructed a proposal of dimension 100000 from smaller overlapping blocks.

Summary

- We have constructed a proposal of dimension 100000 from smaller overlapping blocks.
- Enable us to update hyper-parameters and the latent field jointly.