# Towards Real-Time Bayesian Inference for Magnetoencephalography

Alberto Sorrentino, **Adam M. Johansen**, John A. D. Aston, Tom Nichols, Wilfrid S. Kendall

University of Warwick
a.m.johansen@warwick.ac.uk

JSM: August 12th, 2015

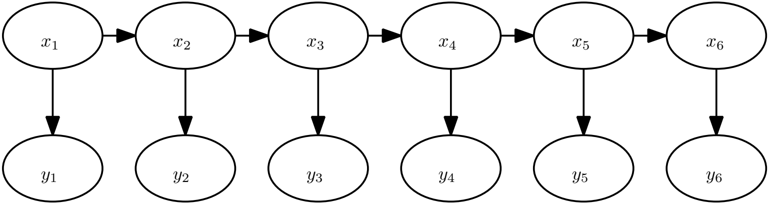
# Outline

- Background
  - Magnetoencephalography
  - Hidden Markov Models / State Space Models
  - Particle Filters / Sequential Monte Carlo
- Models for Filtering of MEG Data
- Examples
- Computational Considerations
- Conclusions

# Magnetoencephalography

- Noninvasive imaging technique which measures magnetic fields induced by neural currents.
- Sampling rate $\sim$ 100Hz.
- Ill-posed inverse problem: turn measurements into estimates of brain activity.
- Problem addressed: estimation of current brain activity given observations received.

# Hidden Markov Models




- Unobserved Markov chain $\{X_n\}$ transition $f$.
- Observed process $\{Y_n\}$ conditional density $g$.
- Density:

$$p(x_{1:n}, y_{1:n}) = f_1(x_1)g(y_1|x_1)\prod_{i=2}^{n} f(x_i|x_{i-1})g(y_i|x_i).$$

# Formal Solutions

- Filtering and Prediction Recursions:

$$p(x_n|y_{1:n}) = \frac{p(x_n|y_{1:n-1})g(y_n|x_n)}{\int p(x'_n|y_{1:n-1})g(y_n|x'_n)dx'_n}$$

$$p(x_{n+1}|y_{1:n}) = \int p(x_n|y_{1:n})f(x_{n+1}|x_n)dx_n$$

- Smoothing:

$$p(x_{1:n}|y_{1:n}) = \frac{p(x_{1:n-1}|y_{1:n-1})f(x_n|x_{n-1})g(y_n|x_n)}{\int g(y_n|x'_n)f(x'_n|x_{n-1})p(x'_{n-1}|y_{1:n-1})dx'_{n-1:n}}$$

# Importance Sampling in This Setting

- Given $p(x_{1:n}|y_{1:n})$ for $n = 1, 2, \ldots$.
- Could sample from a sequence $q_1(x_1), \ldots, q_n(x_{1:n})$
- If $q_n(x_{1:n}) = q_n(x_n|x_{1:n-1})q_{n-1}(x_{1:n-1})$
  we can re-use samples.
- Weight:

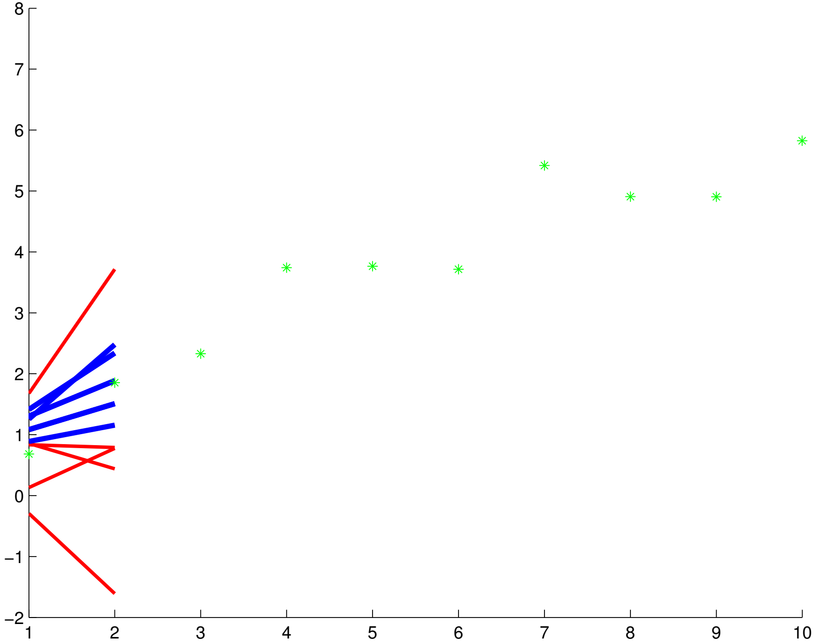
$$\begin{aligned}
w_n(x_{1:n}) &\propto \frac{p(x_{1:n}|y_{1:n})}{q_n(x_n|x_{1:n-1})q_{n-1}(x_{1:n-1})} \\
&= \frac{p(x_{1:n}|y_{1:n})}{q_n(x_n|x_{1:n-1})p(x_{1:n-1}|y_{1:n-1})} w_{n-1}(x_{1:n-1}) \\
&= \frac{f(x_n|x_{n-1})g(y_n|x_n)}{q_n(x_n|x_{n-1})p(y_n|y_{1:n-1})} w_{n-1}(x_{1:n-1})
\end{aligned}$$

# SIR Particle Filters [GSS93]

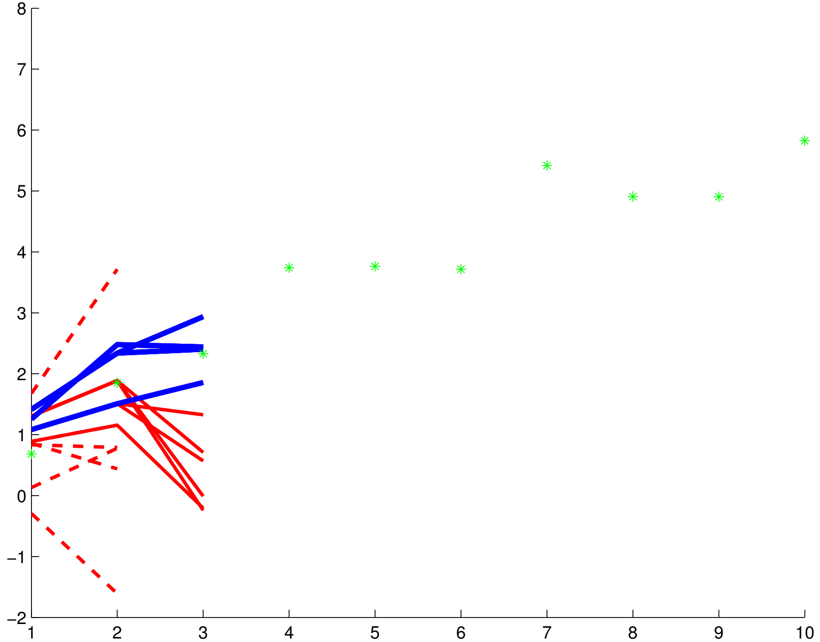
Can be viewed as an extension of importance sampling

- Algorithmically, at iteration $n$:
  - Given $\{W_{n-1}^i, X_{1:n-1}^i\}$ for $i = 1, \ldots, N$:
  - **Resample**, obtaining $\{1/N, \widetilde{X}_{1:n-1}^i\}$.
    - Sample $X_n^i \sim q_n(\cdot|\widetilde{X}_{n-1}^i)$
    - Weight $W_n^i \propto \frac{f(X_n^i|\widetilde{X}_{n-1}^i)g(y_n|X_n^i)}{q_n(X_n^i|\widetilde{X}_{n-1}^i)}$
- Actually:
  - Resample efficiently.
  - Only resample when necessary.

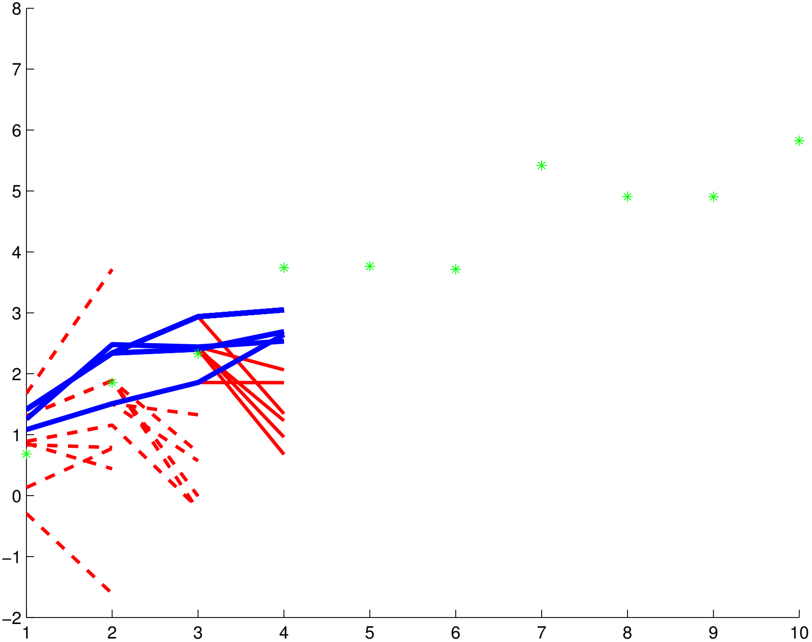
# Iteration 2



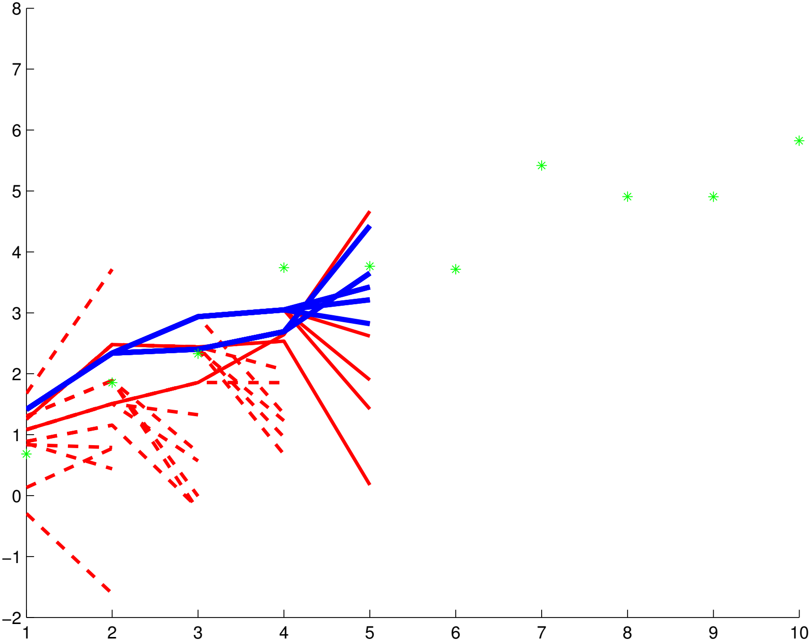
# Iteration 3



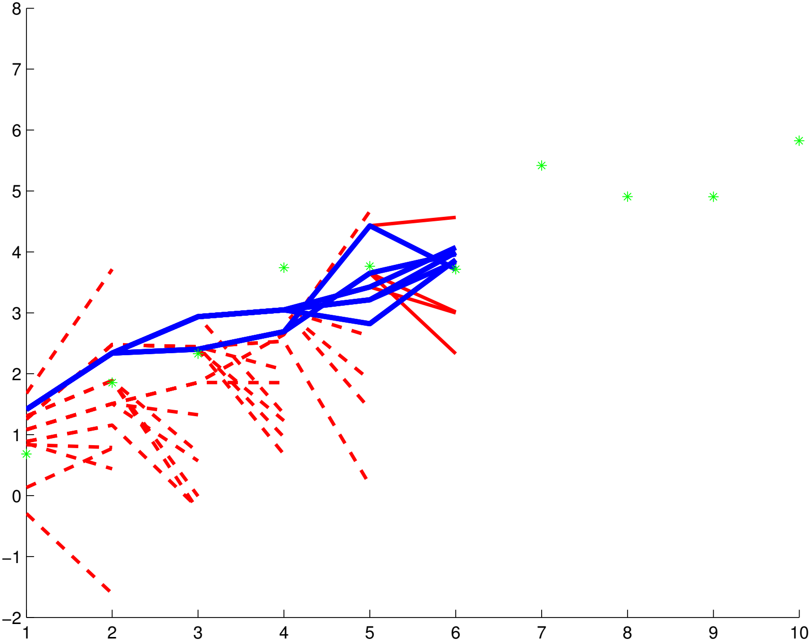
# Iteration 4



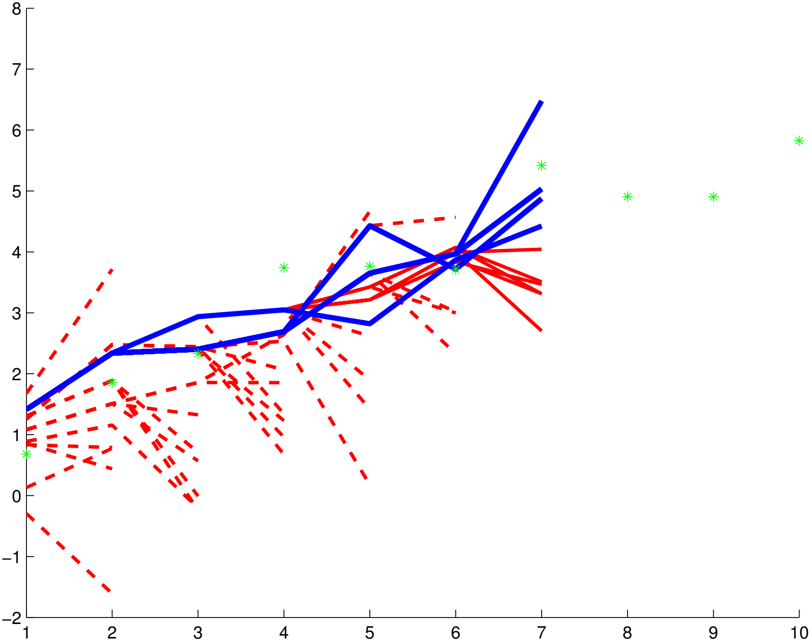
# Iteration 5



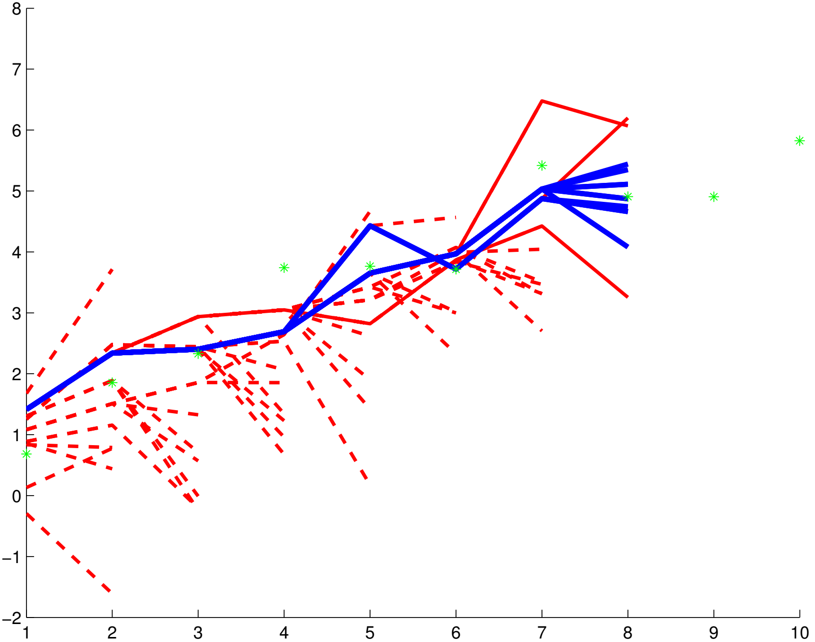
# Iteration 6



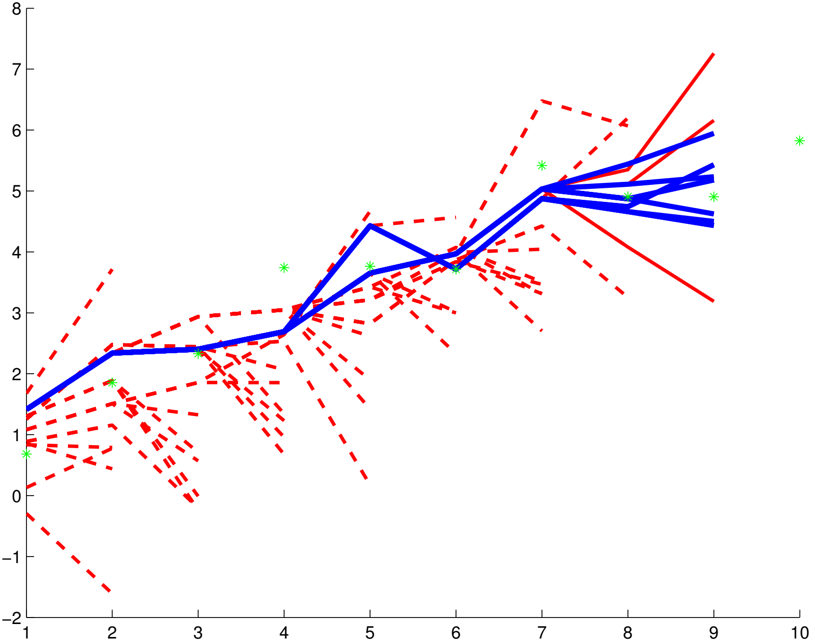
# Iteration 7



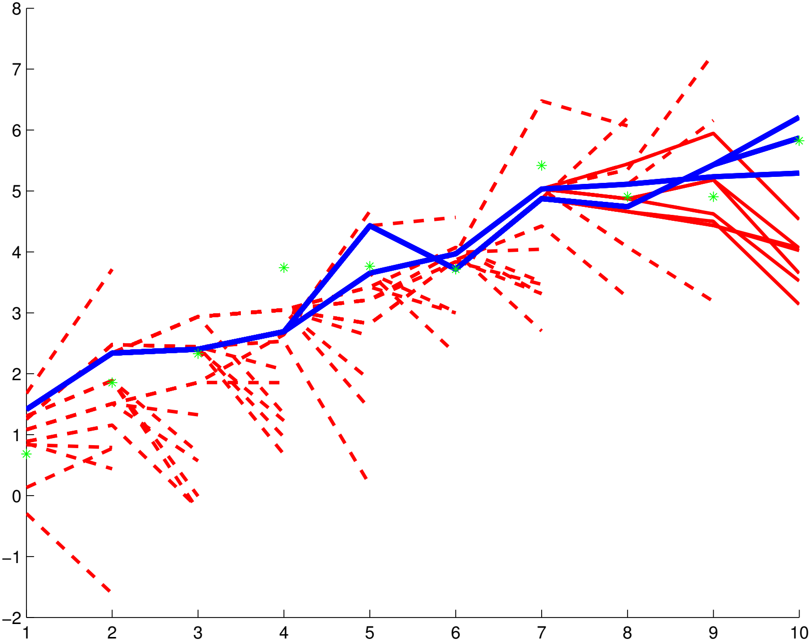
# Iteration 8



# Iteration 9



# Iteration 10



# Resample-Move Particle Filters [GB01]

- Originally, incorporate $p(x_t|y_{1:t})$-invariant MCMC kernels.
- Here we consider $p(x_{1:t}|y_{1:t})$-invariant kernels.
- Actually, go back only to the birth of the oldest surviving dipole.

# Measurement Model

- Work with discretized brain ($N_{\text{grid}} \approx 10000$ elements).
- Precompute ( $N_{\text{sensors}} \times 3N_{\text{grid}}$.) leadfield/gain matrix, $G$.

$$p(b_t|j_t) = \mathcal{N}\left(b_t; \sum_{i=1}^{N_t} G\left(r_t^{(i)}\right) q_t^{(i)}, \Sigma_{\text{noise}}\right)$$

where

$$j_t = \left(\left(r_t^{(1)}, q_t^{(1)}\right), \ldots, \left(r_t^{(N_t)}, q_t^{(N_t)}\right)\right)$$

and $G(r)$ is a $N_{\text{sensors}} \times 3$ matrix for each $r$.

# A Random Walk with Births and Deaths

Mixture of 3 components:

$$
\begin{aligned}
p(j_t|j_{t-1}) = & P_{\text{birth}}(j_{t-1})T_{\text{birth}}(j_t|j_{t-1}))+ \\
& P_{\text{death}}(j_{t-1})T_{\text{death}}(j_t|j_{t-1}))+ \\
& P_{\text{rw}}(j_{t-1})T_{\text{rw}}(j_t|j_{t-1}))
\end{aligned}
$$

Where:

$$
\begin{aligned}
T_{\text{birth}}(j_t|j_t) = & T_{\text{rw}}(j_t^{(1:N_{t-1})}|j_{t-1})U_{\text{grid}}(r_t^{(N_{t-1}+1)})\mathcal{N}(q_t^{(N_{t-1}+1)};0,\sigma_q I) \\
T_{\text{death}}(j_t|j_t) = & \frac{1}{N_{t-1}}\sum_{i=1}^{N_{t-1}} T_{\text{rw}}(j_t|j_{t-1} \setminus (r_{t-1}^{(i)}, q_{t-1}^{(i)})) \\
T_{\text{rw}}(j_t|j_t) = & \prod_{i=1}^{N_{t-1}} \mathcal{N}|_{\text{grid}}(r_t^{(i)}; r_{t-1}^{(i)})\mathcal{N}(q_t^{(i)}; q_{t-1}^{(i)}, \Delta)
\end{aligned}
$$

# A Static Dipole Model [SJA+13]

Mixture of 3 components:

$$\begin{aligned} p(j_t|j_{t-1}) = & P_{\text{birth}}(j_{t-1}) T_{\text{birth}}(j_t|j_{t-1})) + \\ & P_{\text{death}}(j_{t-1}) T_{\text{death}}(j_t|j_{t-1})) + \\ & P_{\text{sd}}(j_{t-1}) T_{\text{sd}}(j_t|j_{t-1})) \end{aligned}$$

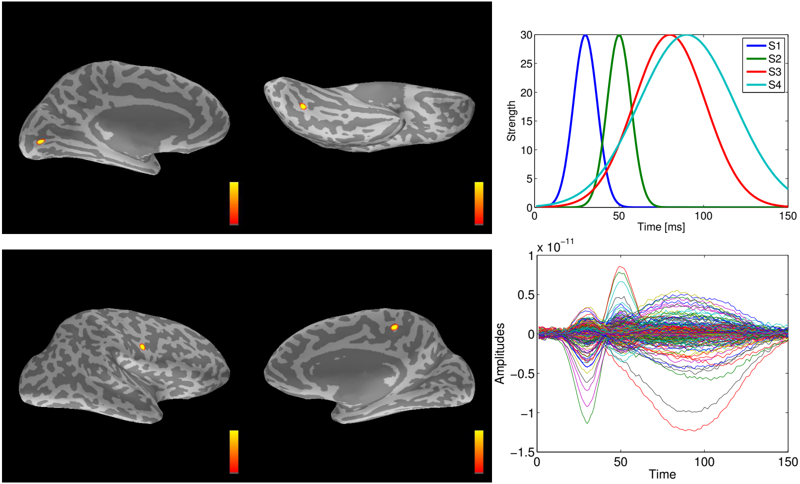
Where:

$$T_{\text{birth}}(j_t|j_t) = T_{\text{sd}}(j_t^{(1:N_{t-1})}|j_{t-1}) U_{\text{grid}}(r_t^{(N_{t-1}+1)}) \mathcal{N}(q_t^{(N_{t-1}+1)}; 0, \sigma_q I)$$

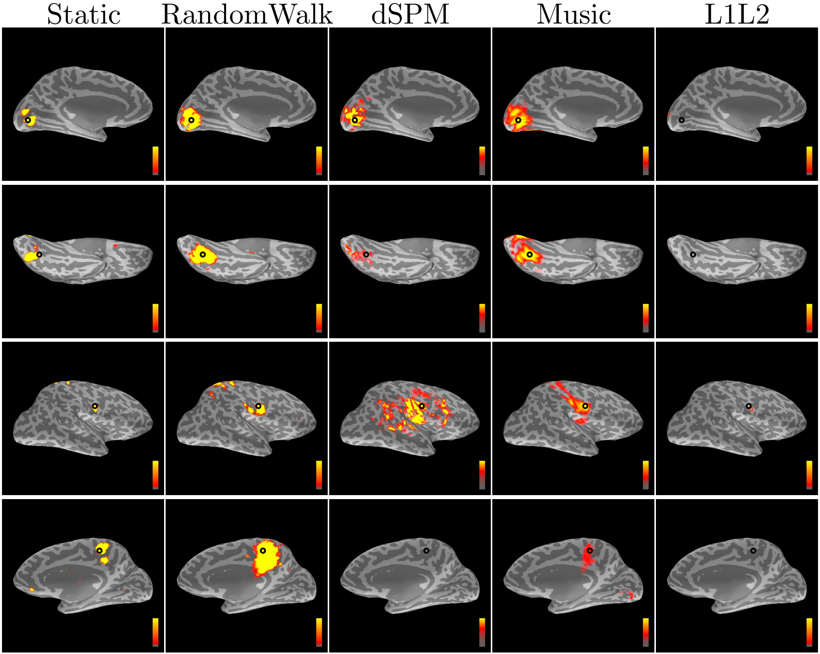
$$T_{\text{death}}(j_t|j_t) = \frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} T_{\text{sd}}(j_t|j_{t-1} \setminus (r_{t-1}^{(i)}, q_{t-1}^{(i)}))$$

$$T_{\text{sd}}(j_t|j_t) = \prod_{i=1}^{N_{t-1}} \delta_{r_{t-1}^{(i)}}(r_t^{(i)}) \mathcal{N}(q_t^{(i)}; q_{t-1}^{(i)}, \Delta)$$

# Simulation Experiment

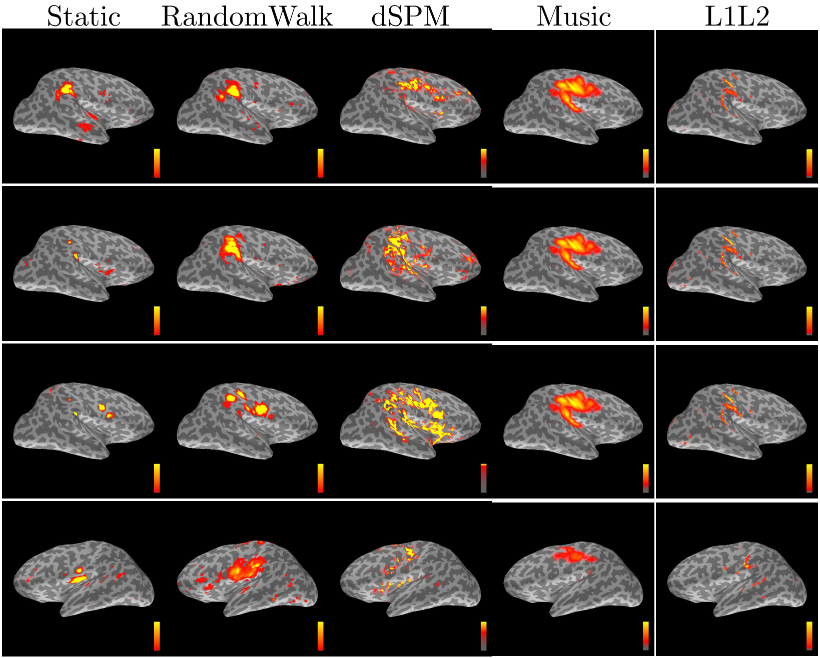



Source locations (top row S1 and S2, bottom row S3 and S4), source time courses and generated noisy field.

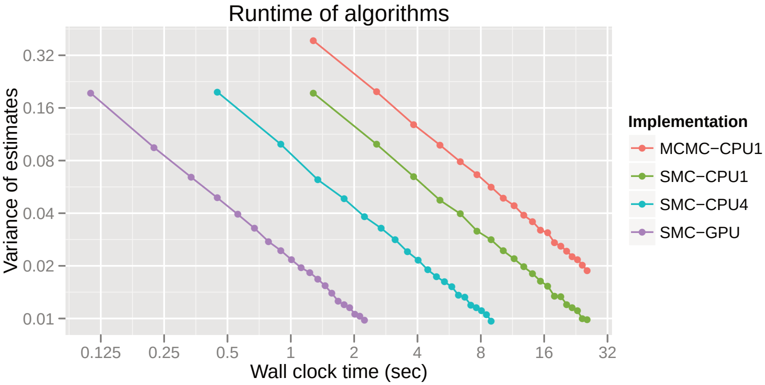

Static
RandomWalk
dSPM
Music
L1L2

# Application to Real Data

- Somatosensory Evoked Fields mapping experiment
- Acquired with 306-channel MEG device [204 planar gradiometers; 102 magnetometers]
- The left median nerve at wrist was electrically stimulated at the motor threshold (interstimulus interval randomly varying between 7.0 s and 9.0 s).
- Signals filtered to 0.1-200 Hz and sampled at 1000 Hz.
- Electrooculogram (EOG) used to monitor eye movements.
- Trials with EOG or MEG exceeding 150 mV or 3 pT/cm were excluded.
- 84 clean trials were averaged.
- To reduce external interference, signal space separation method was applied to the average.
- A 3D digitizer and four head position indicator coils were employed to determine the position of the subject's head.


Static
RandomWalk
dSPM
Music
L1L2

# Computational Considerations: A Related PET Algorithm [ZJA15]



# Conclusions

- Modelling and computational considerations interact.
- Static dipole models allow better interpretation.
- Online inference under such models is possible with care.
- Computational costs are manageable…
- and parallelisation is easy.

# References


W. R. Gilks and C. Berzuini.
Following a moving target – Monte Carlo inference for dynamic Bayesian models.
*Journal of the Royal Statistical Society B*, 63:127–146, 2001.

N. J. Gordon, S. J. Salmond, and A. F. M. Smith.
Novel approach to nonlinear/non-Gaussian Bayesian state estimation.
*IEE Proceedings-F*, 140(2):107–113, April 1993.

A. Sorrentino, A. M. Johansen, J. A. D. Aston, T. E. Nichols, and W. S. Kendall.
Filtering of dynamic and static dipoles in magnetoencephalography.
*Annals of Applied Statistics*, 7(2):955–988, 2013.

Y. Zhou, A. M. Johansen, and J. A. D. Aston.
Towards automatic model comparison: An adaptive sequential Monte Carlo approach.
*Journal of Computational and Graphical Statistics*, 2015.
In press.