Fast low-SNR high-dimensional optimal filtering, applied to inference of dynamic receptive fields.

Liam Paninski, Kamiar Rahnama Rad, Jonathan Huggins, and Eftychios A. Pnevmatikakis

Manuscript:
http://www.stat.columbia.edu/ liam/research/pubs/fast-low-rank-kalman.pdf

Department of Statistics
Columbia University
February 24, 2011

Representation of the spatial environment in the brain

- **Place Field** $\Longleftrightarrow$ neurons in the rodent hippocampus **respond** selectively depending on the animal's **current location**.
- In many situations, e.g. learning, the place field is **time varying**.

$$
\begin{aligned}
n_t &\sim f(x_t, t) + \text{noise} \sim \langle B_t, q_t \rangle + \text{noise} \\
f(x, t) &\sim \text{time varying place field} \\
B_t &\sim N \times N\text{-pixel indicating the } \textbf{current location} \\
q_t &\sim N \times N\text{-pixel time varying } \textbf{place field}
\end{aligned}
$$

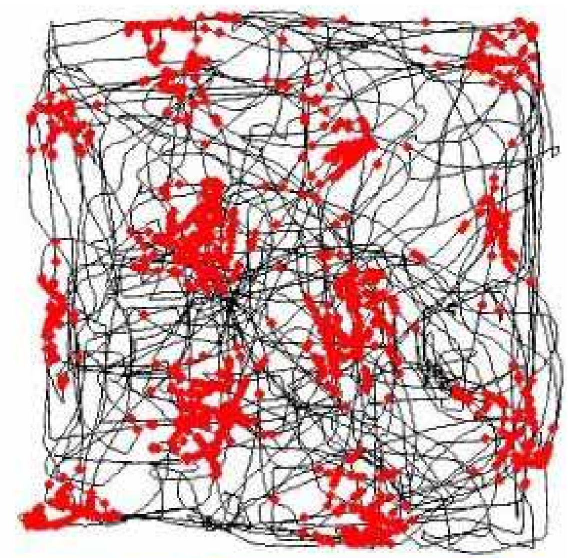


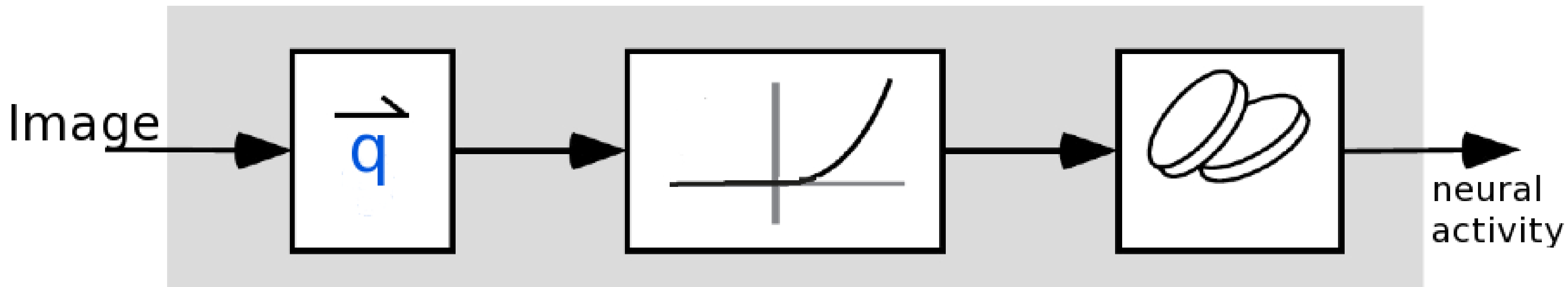
Figure 1: Trajectory of a rat through a square environment is shown in black. Red dots indicate locations at which the particular entorhinal cell being examined fired.

Dynamic receptive field estimation

The **activity** of a neuron in a sensory brain region depends on the **linear projection** of the **stimulus** into the **time varying receptive field**.

$$
\begin{aligned}
n_t &\sim \langle B_t, q_t \rangle + \text{noise} \\
B_t &\sim N \times N\text{-pixel time varying } \textbf{visual stimuli} \\
q_t &\sim N \times N\text{-pixel time varying } \textbf{receptive field}
\end{aligned}
$$

Main question: How to estimate the time varying receptive field?



One common problem:

- Understanding the dynamics of **large systems** for which **limited** and **noisy** observations are available.
- Classical solutions include **state space** models. See [1, 2, 3].
- Standard implementations of the Kalman filter require $O(\dim(q)^3)$ **time** and $O(\dim(q)^2)$ **memory** per time step, and are therefore impractical for applications involving very high-dimensional ($\dim(q) \sim 100 \times 100$) systems.

Fast low-SNR optimization

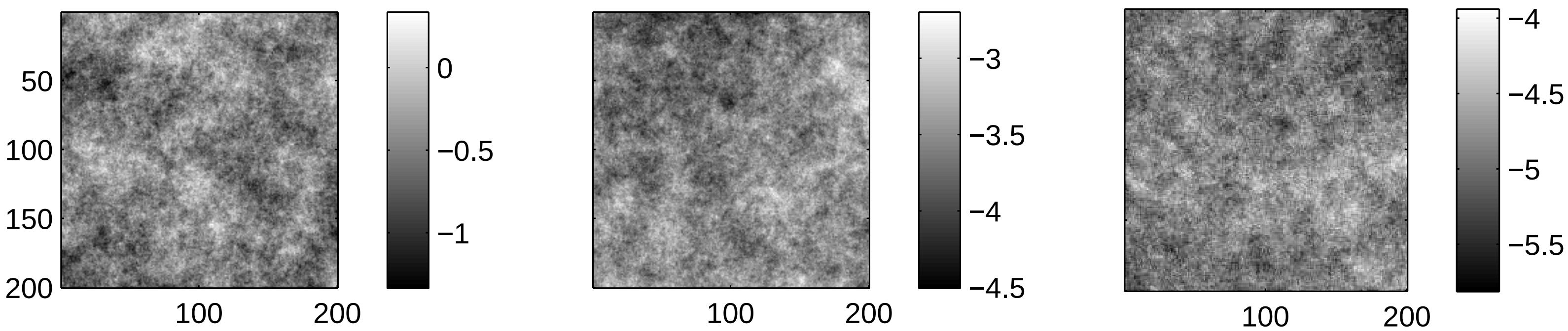
- When there are **no observations** the **uncertainty** reflects our **prior** belief such as smoothness and/or boundedness of the receptive/place fields.
- Observations **decrease** the uncertainty.
- The decrease in the uncertainty due to low snr observation is small in magnitude and only changes our uncertainty in one direction.
- The effect of previous observations decays exponentially fast.
- The difference between the uncertainty of no observation and low snr observation is *effectively* a low rank matrix, i.e. $C_t = C_0 + U_t D_t U_t^T$.
- All computations are fast: optimal smoother requires $O(n^3 + n \dim(q) \log \dim(q))$ time and $O(n \dim(q))$ space; $n = \mathrm{rank}(U_t)$.
- Can be used for fast experimental design. See [4, 5]

The model

- Smoothness along the **temporal** and **spatial** dimensions:

$$\begin{aligned} q_{t+1} &= Aq_t + \epsilon_t \quad q_t \sim \text{receptive/place field} \quad \epsilon \sim \mathcal{N}(0, V) \\ A &\sim \text{temporal correlation} \qquad V \sim \text{spatial correlation} \end{aligned}$$

Three independent samples ϵ_t drawn from the Gaussian prior with covariance matrix V.



- Noisy low dimensional **observations**:

$$y_{t+1} = B_t q_t + \eta_t \quad B_t \sim \text{visual/spatial stimuli} \quad \eta_t \sim \mathcal{N}(0, W_t)$$

Standard Kalman recursion

$$\mu_t = \mathrm{E}[q_t|y_{1:t}] \qquad C_t = \mathrm{cov}[q_t|y_{1:t}]$$

- no observation, equilibrium covariance: $AC_0A + V = C_0$ or $C_0 = V(I - AA^T)^{-1}$.

$$\begin{aligned}
\mu_t &= C_t\Big[(AC_{t-1}A^T + V)^{-1}A\mu_{t-1} + B^TW^{-1}y_t\Big] \\
C_t &= \Big[(AC_{t-1}A^T + V)^{-1} + B^TW^{-1}B\Big]^{-1}
\end{aligned}$$

- computational difficulty $\rightarrow C_t$ costs $O(\dim(q)^3)$ time ($O(\dim(q)^2)$ is B is low rank), and $O(\dim(q)^3)$ space

Low snr observation

- no observation: $C_t = C_0 = V(I - AA^T)^{-1}$
- single observation at $t = 1$ and no observation for $t > 1$:

$$\begin{aligned}
C_1 &= \left[C_0^{-1} + B_1^T W^{-1} B_1\right]^{-1} = C_0 - C_0 B_1^T (B_1 C_0^{-1} B_1^T + W^{-1})^{-1} B_1 C_0 \\
&= C_0 + U_1 D_1 U_1^T \qquad \operatorname{rank}(U_1) = \operatorname{rank}(B_1) \\
\text{similarly} \quad C_{t+1} &= C_0 + A^t U_1 D_1 (A^t U_1)^T.
\end{aligned}$$

Since A is stable, the perturbation to C_{t+1} around the equilibrium covariance C_0 caused by a lag t observation decays exponentially in t

Fast methods

- Approximating $C_t \sim C_0 + U_t D_t U_t^T$ where U_t is low rank, i.e.
$n := \text{rank}(U_t) << \dim(q)$ allows us to perform fast efficient recursion:
- Updating U_t and D_t costs $O(n^3 + nN \log N)$ time and $O(nN)$ space.

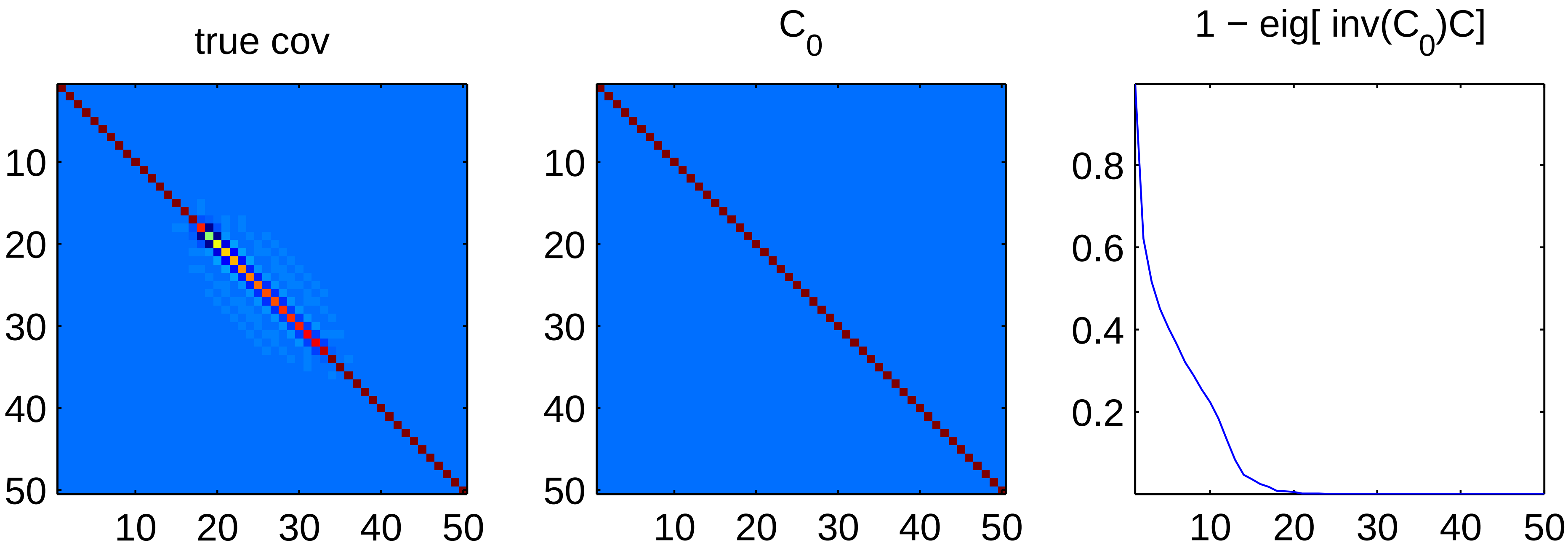
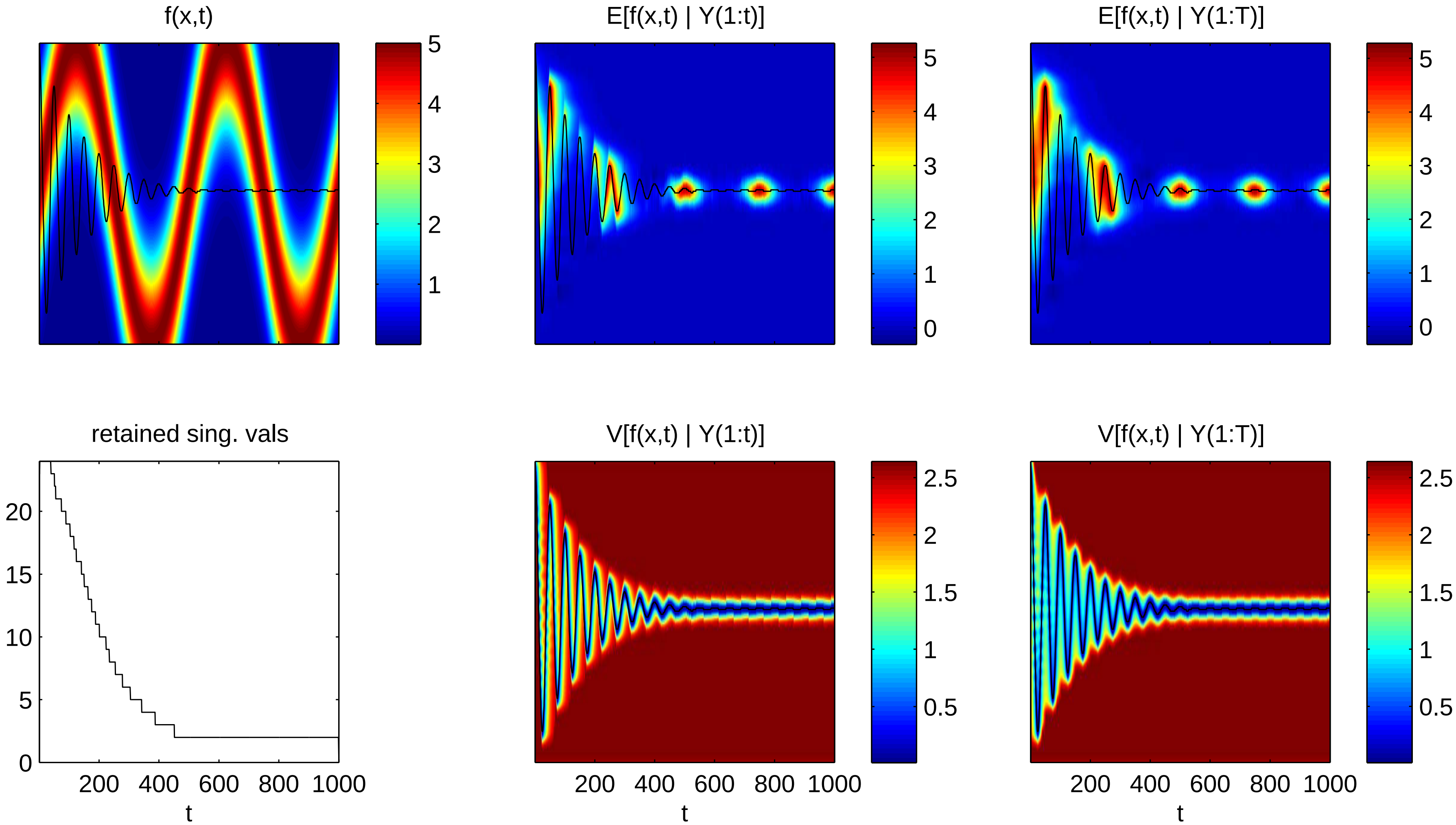
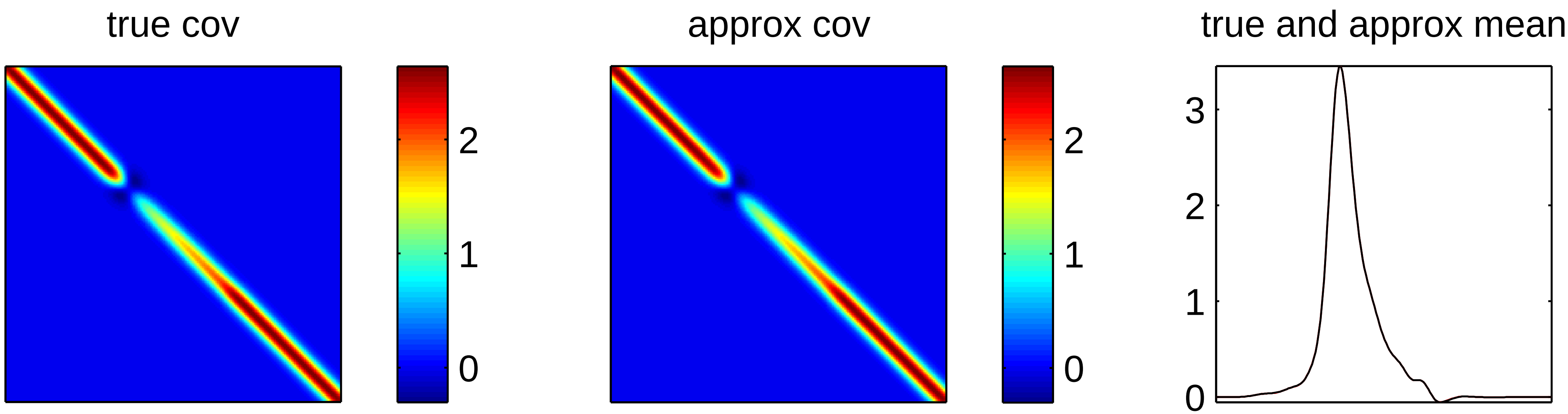


Figure 2: C_t is fairly close to C_0; in particular, $I - C_0^{-1} C_t$ has low effective rank. Left: true C_t. Middle: C_0.

One dimensional place field data



The superimposed black trace in all but the lower left panel indicates the simulated path x_t of the animal. Upper left: true simulated place field $q_t(x)$ is shown in color. Top middle and right panels: estimated place fields, forward $(\mathrm{E}(q_t|y_{1:t}))$ and forward-backward $(\mathrm{E}(q_t|y_{1:T}))$, respectively. Bottom middle and right panels: marginal variance of the estimated place fields, forward $(\mathrm{var}(q_t|y_{1:t}))$ and forward-backward $(\mathrm{var}(q_t|y_{1:T}))$, respectively. Lower left panel: effective rank of $C_0 - C_t^s$ as a function of t in the forward-backward smoother; the effective rank is largest when x_t samples many locations in a short time period.



Comparison of the true vs. approximate covariance. Left panel: true covariance. Middle panel: approximate covariance. The maximal pointwise error between these two matrices is about 1%. Right panel: true and approximate mean μ_t. The black trace indicates the true mean and the red trace (barely visible) the approximate mean.

Tracking a time-varying one-dimensional receptive field

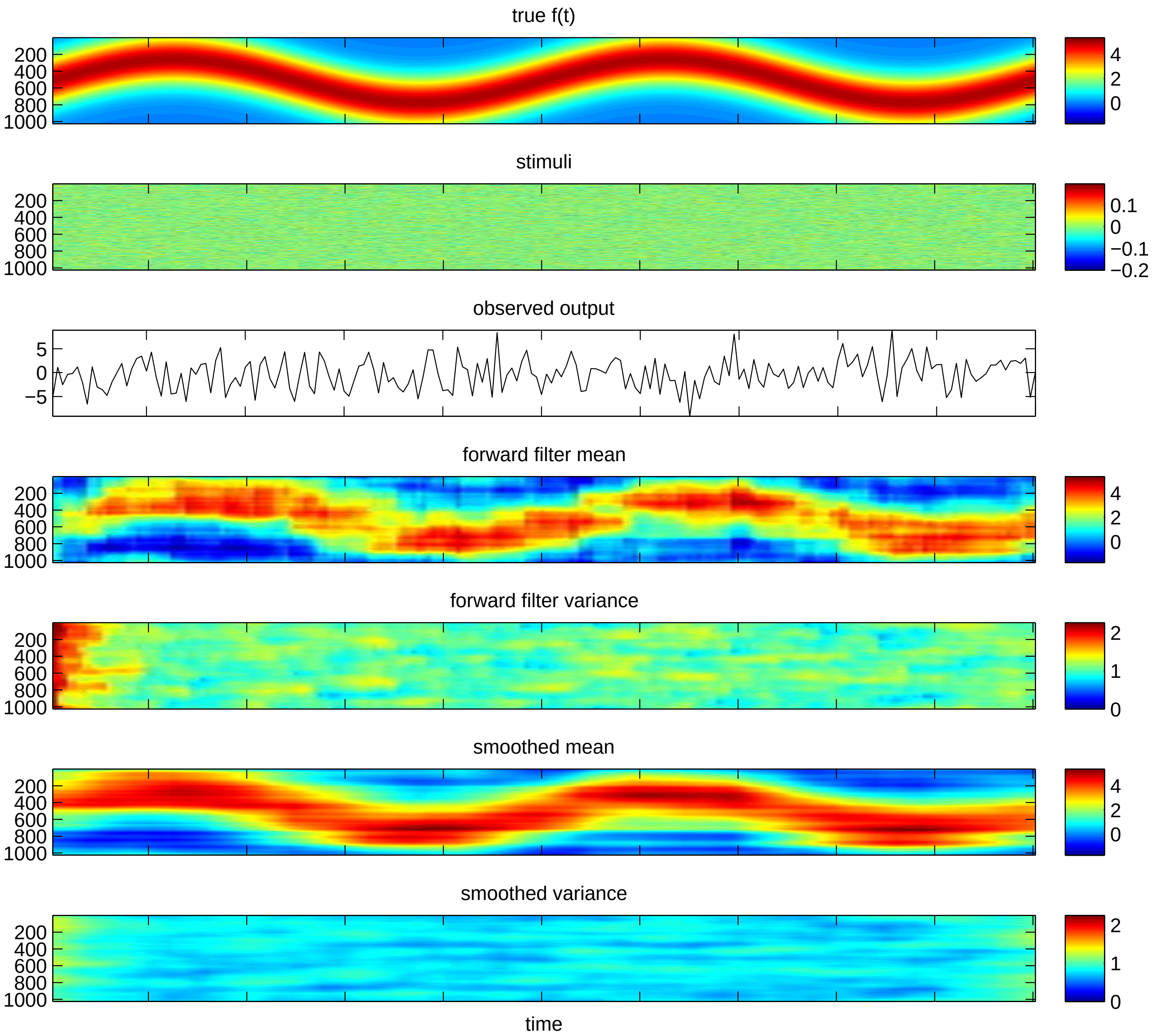


Figure 3: Second panel: the stimulus B_t was chosen to be spatiotemporal white Gaussian noise. Third panel: simulated output observed according to the Gaussian model $n_t = B_t q^t + \eta_t$.

Acknowledgment:LP is supported by a McKnight Scholar award and an NSF CAREER award. JHH is supported by the Columbia College Rabi Scholars Program. We thank P. Jercog for kindly sharing his hippocampal data with us.

References

[1] E. Brown, D. Nguyen, L. Frank, M. Wilson, and V. Solo. An analysis of neural receptive field plasticity by point process adaptive filtering. *PNAS*, 98:12261–12266, 2001.

[2] L. Frank, U. Eden, V. Solo, M. Wilson, and E. Brown. Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: An adaptive filtering approach. *J. Neurosci.*, 22(9):3817–3830, 2002.

[3] G. Czanner, U. Eden, S. Wirth, M. Yanike, W. Suzuki, and E. Brown. Analysis of between-trial and within-trial neural spiking dynamics. *Journal of Neurophysiology*, 99:2672–2693, 2008.

[4] J. Huggins and L. Paninski. Optimal experimental design for sampling voltage on dendritic trees. *Under review*, 2010.

[5] L. Paninski. Fast Kalman filtering on quasilinear dendritic trees. *Journal of Computational Neuroscience*, 28:211–28, 2010.