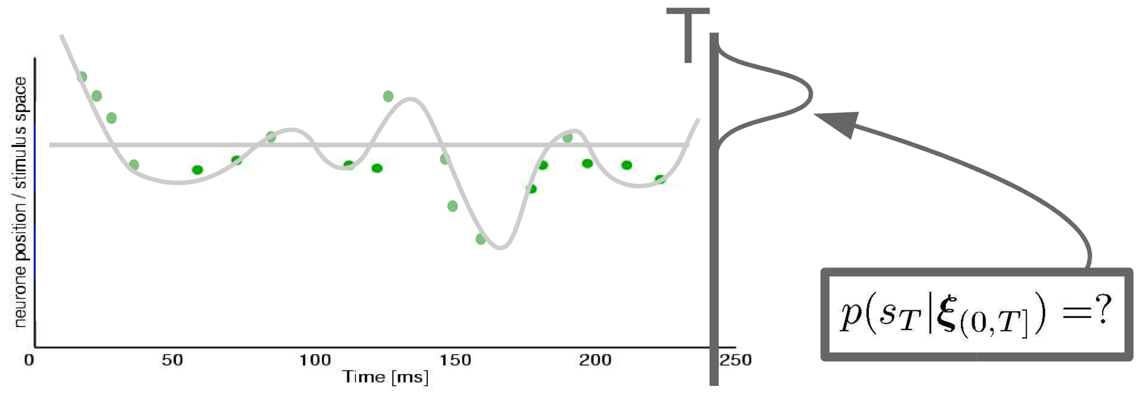
A simple population code in a fast-changing world

Quentin JM Huys[1], Richard S Zemel[2], Rama Natarajan[2], Peter Dayan[1]

[1]Gatsby Computational Neuroscience Unit, University College London, UK; [2] Dept. of Computer Science, University of Toronto, Canada
{qhuys, dayan}@gatsby.ucl.ac.uk, {zemel, rama}@cs.toronto.edu

Introduction

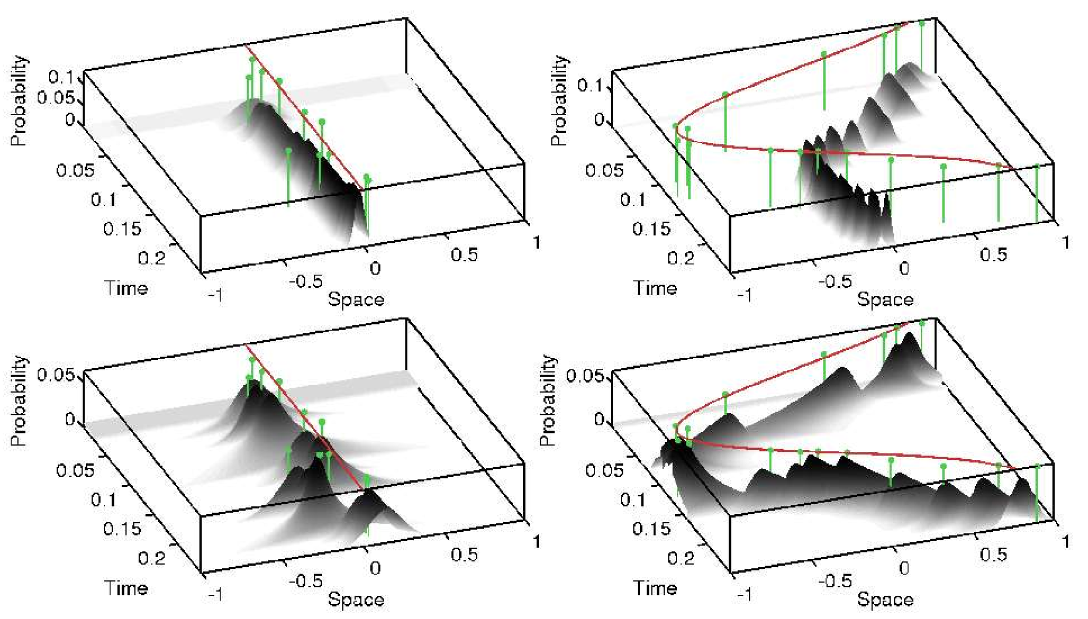


Stimulus inference on timescale of spike production is an underconstrained problem ⟶ need **prior**

We analyse a very simple case
- Gaussian process prior over stimulus trajectory
- Bell-shaped tuning functions
- Independent Poisson noise

Natural prior will make use of spikes for computations hard. A new set of spikes can be generated via recoding, which allows computations and take the prior into account.

Temporal priors matter

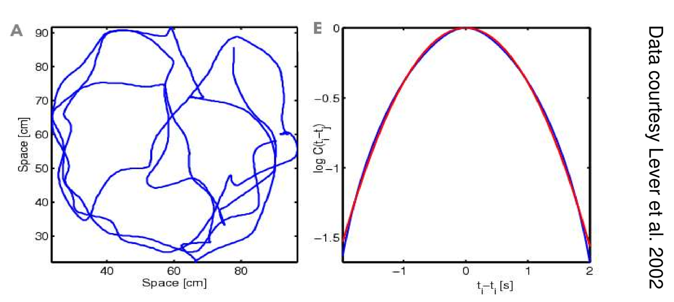


The problme is ill-defined. Need prior. Want to use the right, informative, natural prior

Natural temporal priors are smooth

Natural movements are smooth.

Quadratic exponential covariance function fits natural movements well.

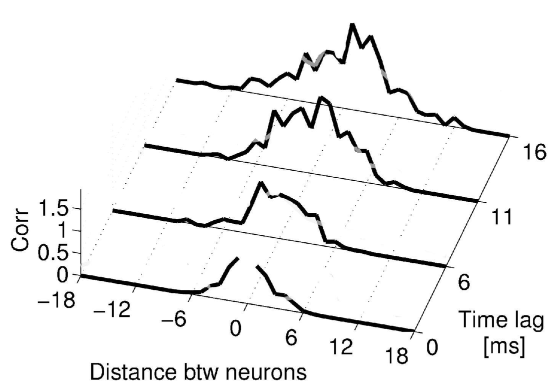


Stimulus-induced correlations

Neurons predict each other over time due to stimulus-induced correlations.

This is a statistical inefficiency.

Temporally efficient coding should flatten the crosscorrelations beyond time lag zero.



Exact inference

We find the posterior distribution over the stimulus at time T given all spikes observed so far

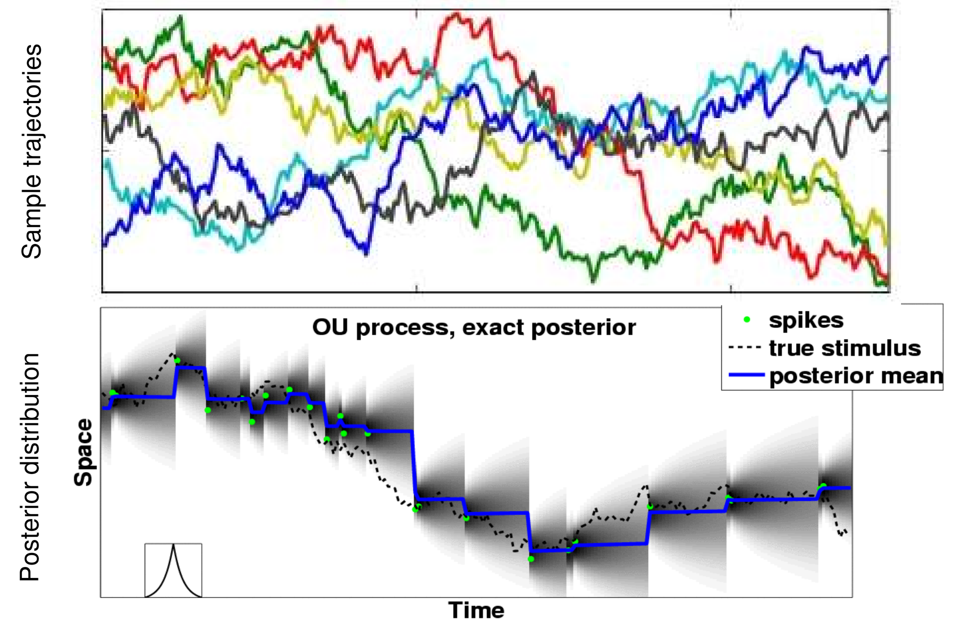
$$p(s_T|\xi_{(0,T]}) \overset{\text{Bayes}}{\propto} p(\xi_{(0,T]}|s_T)\,p(s_T)$$
$$= \int ds_{(0,T]}\, p(\xi_{(0,T]}|s_{(0,T]})\, p(s_{(0,T]}, s_T)$$

Assumption: independent, identical Poisson neurones

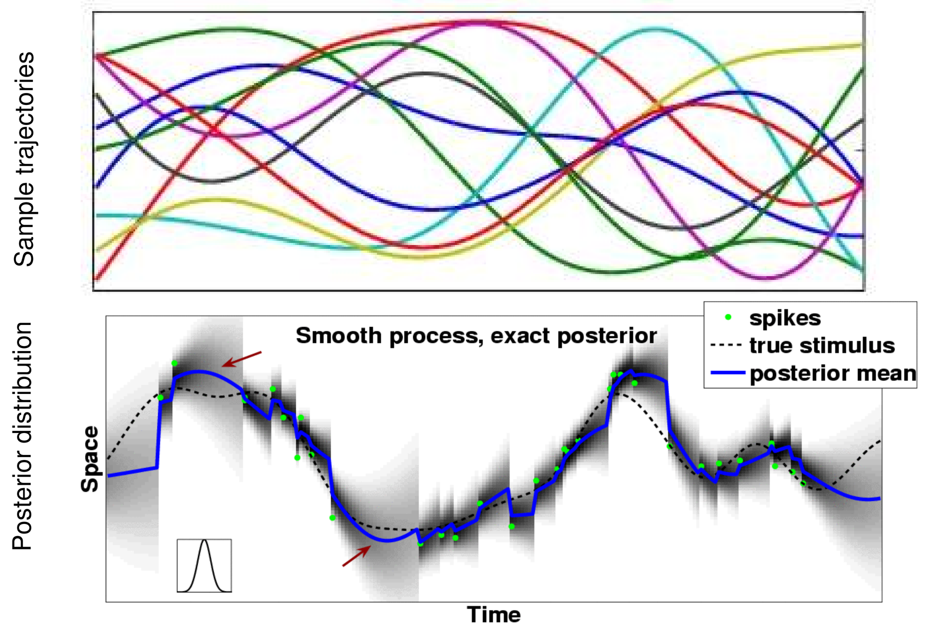
Assumption: Gaussian process prior over **entire** stimulus (trajectory)

Result: Posterior distribution is a simple Gaussian

OU prior



Smooth prior



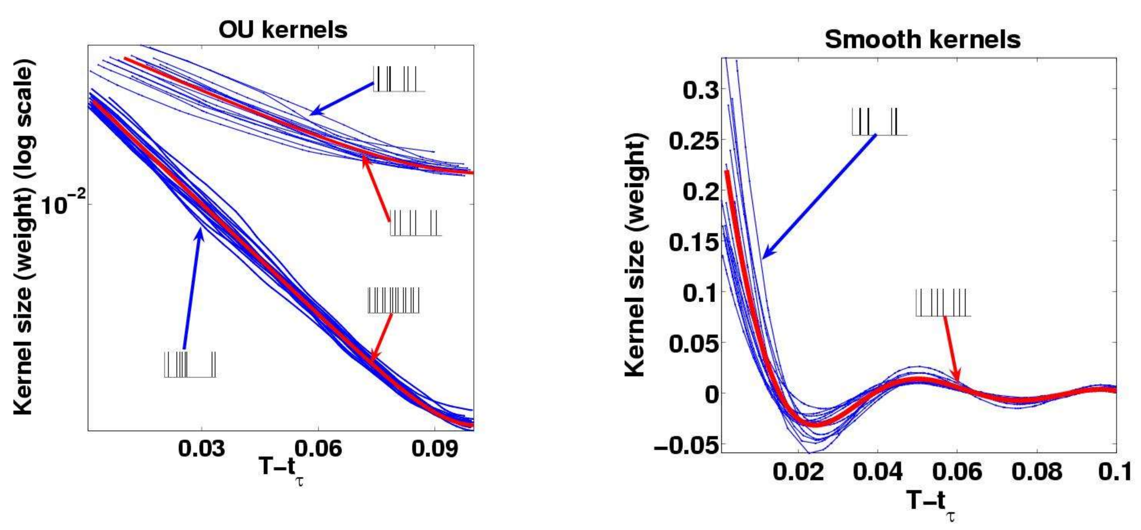
Exact kernels k

Static case: Simply count how many spikes $n_i(T)$ each neurone i has emitted up to time T.

$$\mu(T) = \frac{\sum_i \theta_i n_i(T)}{\sum_j n_j(T)}$$

Dynamic case: spikes are not just counted but are weighted by kernel **k**

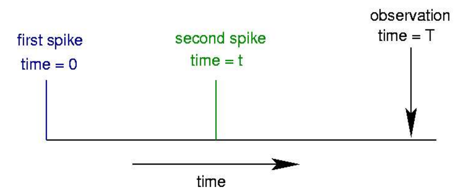
$$\mu(T) = \sum_i \theta_i \left(\sum_t \mathbf{k}(t)\xi_i(t) \right)$$



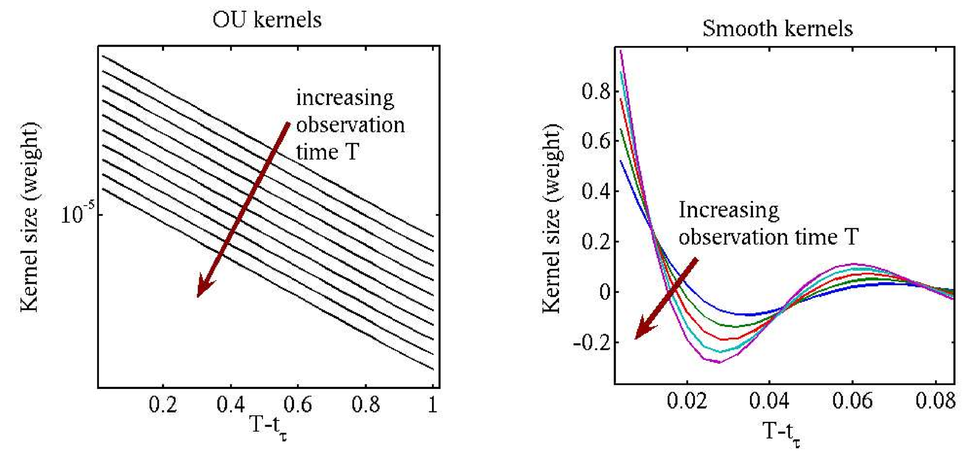
a) depends only on spike and observation times, not on spike locations
b) determines the weight of each spike
c) has a shape that is determined by the covariance of the Gaussian process prior

The structure of the code

Structure of the code apparent between spikes.



For OU prior, kernels change in simple manner. For smooth prior, kernels change in complex manner.



The prior determines the structure of decoding.
For the OU process, it can be decomposed into a product over spike duplets.
This decomposition allows a recursive formulation and thus the OU code generates a temporally compact decoder.

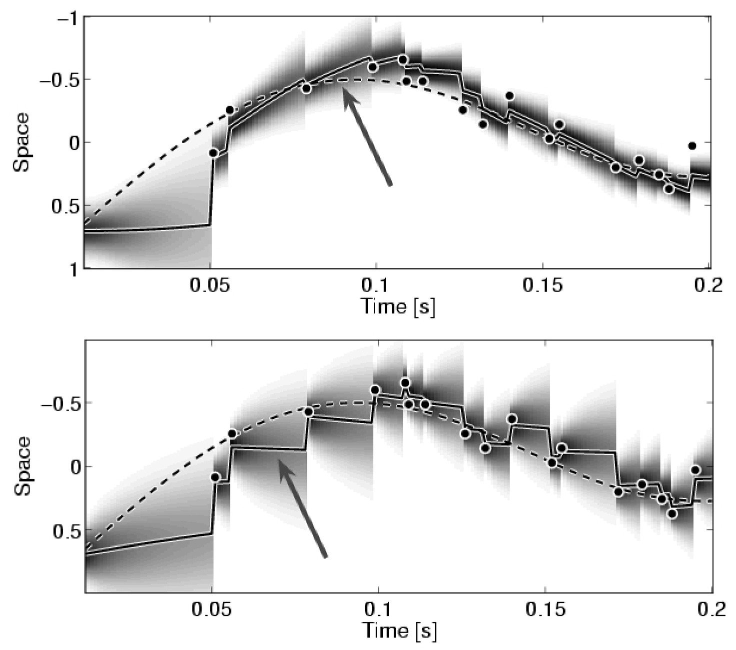
Smooth priors ⟷ complex decoders

There is NO such formulation for the smooth process.

With a smooth prior, decoding is NONLOCAL in TIME and across NEURONS

Neglecting the structure of the code is a bad approximation.

Natural temporal priors combined with a simple encoder lead to a code which is very hard to decode.



A powerful, simple code

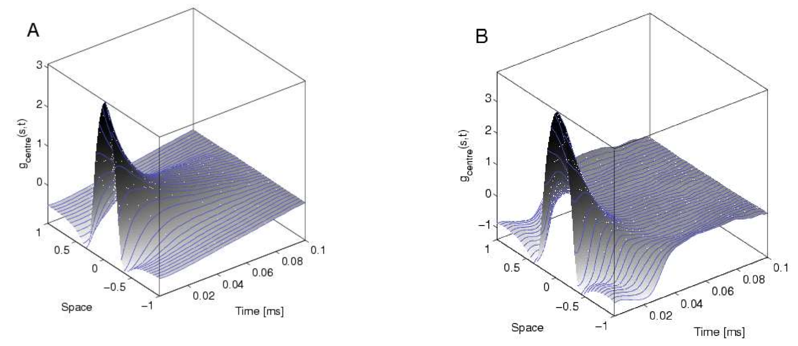
Decode each spike independently. Posterior is product over spikes.

$$\hat{p}(s_T|\xi_{(0,T]}) \propto \exp\left(\sum_{i,\tau} g(i,s,\tau)\xi(i,T-\tau) \right)$$

This is a computationally very powerful code that allows straightforward combination of information across modalities.

Neglecting correlation structure is bad, but maybe there is a good independent interpretation of spikes?

Best independent interpretation of spikes is similar for extremely different priors.



OU prior Smooth prior

Recoding

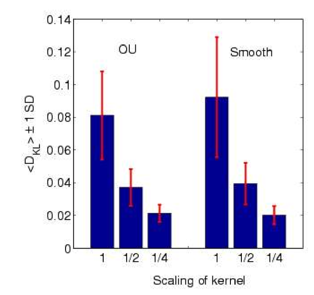
Fix decoding kernel g(i,s,t) and find a new set of spikes such that the original distribution

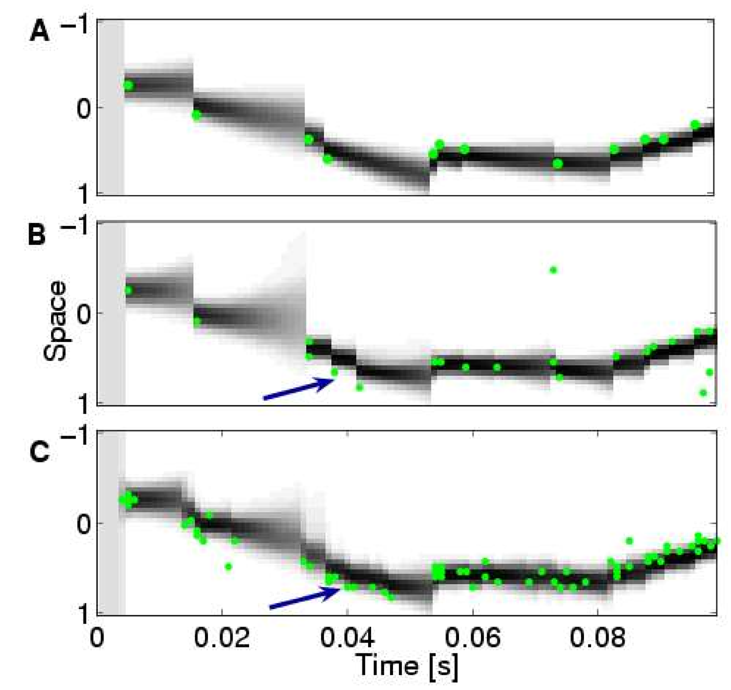
$$p(s_T|\xi_{[0,T]})$$

is matched by the new distribution

$$\hat{p}(s_T|\rho_{[0,T)})$$

with the new spikes ρ



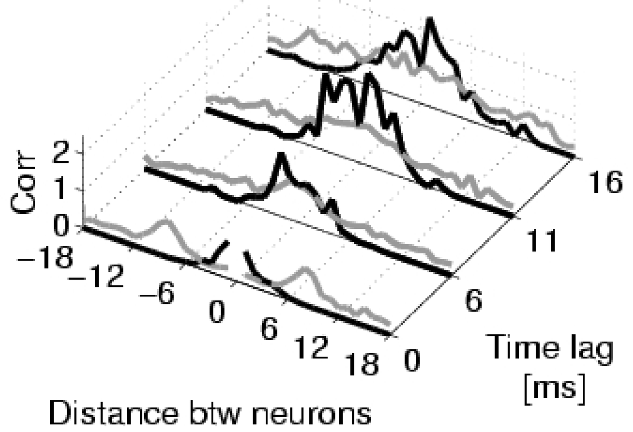


Thus the same information can be represented by **simply decodeable** spikes that are readily used in computations.

Statistically efficient coding in time

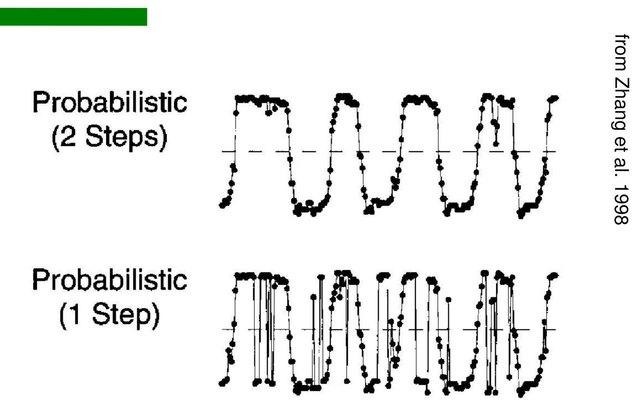
Recoding eliminates the temporal redundancies in the neural activities – the grey curve shows flat correlations across neurons after recoding.

This is the temporal analogue of adaptation to visual scene statistics (eg. Srinivasan et al. 1982)



Related work

Previous work has mainly used recursively defineable priors, ie priors within the OU and Kalman filter class (Brown et al. 1998, Gao et al. 2002, Barbieri et al. 2003, Twum-Danso and Brockett 2001), although Zhang et al. (1998) show that using a 2-step Bayesian decoder significantly increases the quality of decoding from the hippocampus.



The standard sliding temporal window corresponds to assuming that the stimulus is piecewise constant.

Kemere et al.(2004) have used informative priors to decode movement-related activity and found it to be strongly ameliorate performance.

Nirenberg et al. (2001) show that retinal Ganglion cells are independently decodeable.

Conclusions

Decoding in time necessitates an informative prior.

Natural priors combined with a simple encoder engender a computationally inflexible code. This is due to the stimulus-induced correlations which need to be taken into account.

The structure of a decoder tells us where the information is, in what format it is available. In our case, the information was not in an accessible format to downstream neurons.

We propose a recoding. Recoding engenders a computationally and representationally powerful and flexible code. The resulting spike trains seem to have “adapted” to the temporal statistics of the stimulus.