

Model-based smoothing of noisy, intermittent biophysical signals

Quentin Huys¹, Peter Latham¹ and Liam Paninski²

¹ Gatsby Computational Neuroscience Unit, UCL; ² Dept. of Statistics, Columbia University
{qhuys,miller}@gatsby.ucl.ac.uk, liam@stat.columbia.edu

Noise is an omnipresent issue that is often handled suboptimally. For example, noise is an issue in voltage-sensitive imaging – even the best dyes achieve signal-to-noise ratios of no more than $\sim 1 - 6\%$. Averaging noise out is not always possible and sometimes not even desirable. Missing data are often burdensome too: in voltage dye experiments, the laser has to be moved between sites of interest and thus the data is not acquired simultaneously, leading to gaps in the data. Despite advances, this problem becomes more prominent the more sites one attempts to record from. More generally, we might even be interested in a variable that has not been observed directly at all, such as the voltage in a Ca^{++} imaging experiment. Principled methods to filter out noise, to interpolate between data points and to infer unobserved variables could substantially complement advances in data acquisition methodology.

Here we show how, when time series recordings of a dynamical system (*eg* the voltage of a cell) are made, knowledge of the dynamical system can be used to both filter and interpolate between the measurements, providing a principled alternative to heuristics such as temporal smoothing or low-pass filtering. Neural dynamics are usually specified as Markov chains. If these dynamics are hidden (due to noisy or indirect measurements), the task of recovering the distribution over the true underlying state evolution of the neuron over time is equivalent to inference in nonlinear state space models. These models, together with their discrete analogues such as intermittent Kalman filters and Hidden Markov Models have been analysed extensively and are very well understood. If the hidden variables do indeed evolve in a Markovian manner (as is often the case), a number of algorithms from the machine learning literature allow efficient sampling, despite the huge size of the state space.

We find that the combination of a nonlinear Gaussian state space model with Gaussian observation noise and a forward-backward formulation of a particle filter allows us to recover the true voltage of a Fitzhugh-Nagumo or a Hodgkin-Huxley spiking model very well at low signal-to-noise levels that qualitatively match those encountered in voltage imaging experiments. It is possible to formulate dynamical models of other, entirely unobserved variables (such as the voltage in a Ca^{++} imaging experiment) and apply the same techniques. The probabilistic form of this approach also naturally allows combination of measurements from different sources, such as voltage and Ca^{++} imaging, done simultaneously.

Finally, we relax the assumption that the underlying kinetic model of the cell is entirely known to the case in which the true channel densities are unknown. The above procedure then becomes the E-step of expectation-maximisation, while the M-step reduces to a simple constrained linear regression.