Idealisation of Patch Clamp Records Using an Any-level Local Modelling Scheme

J. Alex Stark*,
William J. Fitzgerald and Stephen B. Hladky

Departments of Pharmacology and Engineering,
University of Cambridge, CB2 1QJ, UK.
(*Now at: National Institute of Statistical Sciences,
Research Triangle Park, NC 27709-4006; stark@niss.org)

Abstract

Patch clamp records with small numbers of channels can be modelled effectively in a Bayesian framework. A statistical model was formulated in which the underlying signal is piecewise constant, any conductance level is permitted and there are no long-range constraints. (The levels used to fit two sections of data are independent of each other if the sections are sufficiently separated in time.) The posterior probability of a chosen set of conductance levels and transition times can be calculated from the values of these parameters and the experimental data. A Markov chain Monte Carlo procedure[a], in which the number of transitions is allowed to vary, is used to draw a sequence of fits from their posterior distribution.

The method is applied to filtered data contaminated with coloured noise and is illustrated with results from experimental data.

[a]: Stark, J.A., W.J. Fitzgerald, and S.B. Hladky. 1997. Bayesian Modelling of Patch-Clamp Records. *Biophysical Journal* **72**:A327.

Acknowledgements

This work was supported in part by grant #8/E03204 from the BBSRC and by Christ's College, Cambridge. Additional support was provided by the National Institute of Statistical Sciences, North Carolina.

Introduction

The idealisation of a patch-clamp record is a process of signal conversion which aims to eliminate noise and filtering effects. The complexity of the (signal) model that is imposed depends on the strategy employed. There are many situations in which a highly structured model is inappropriate, such as at early stages of analysis and when the details of the underlying process are unknown.

The aim of this research is to develop methods of patch-clamp analysis using relatively few assumptions and compensating for practical noise and filtering effects. We have employed emerging Bayesian statistical modelling methodologies. In particular, we have implemented a Markov chain Monte Carlo (MCMC) sampler. This type of computational procedure has revolutionised Bayesian statistics in recent years.

Signal, Model and Noise

The idealisation model used for this analysis assumes that the underlying signal is constant between transitions or 'changepoints' (CPs) which occur at any sample point. There are no restrictions on the conductance levels, and the model fit at one location is virtually independent of that at another sufficiently distant location.

The analysis is illustrated with an example 25-second segment of patch-clamp recording (Figs. 1a and b) from ATP-sensitive potassium channel(s) of a CRI-G1 cell sampled at 50kHz having been passed through two analogue low-pass Bessel filters with cut-off at 5kHz, the first of order 4 and the second of order 8.

The model required the noise to be white, and so a 8192-point section of baseline noise was characterised using an order-5 autoregressive (AR) model, and the signal was prewhitened by deconvolution. The noise was assumed to be approximately independent of the conductance level. The power spectra at different locations along the signal (Fig. 1c) were found to be similar. The zero-lag correlation was incremented by 3% so as to avoid excessive amplification at frequencies with low signal-to-noise. A Chebyshev low-pass filter (16kHz cut-off) was used to bandlimit the signal. Fig. 1d shows the original and whitened noise spectra.

The transition (step) response for the analogue filters was converted into digital form (Fig. 1e) and this was modified by the whitening filter in the same fashion as the noise (Fig. 1f). The noise variance used in the model was estimated from the baseline power spectrum. The record was analysed in 100 blocks.

The Sampler

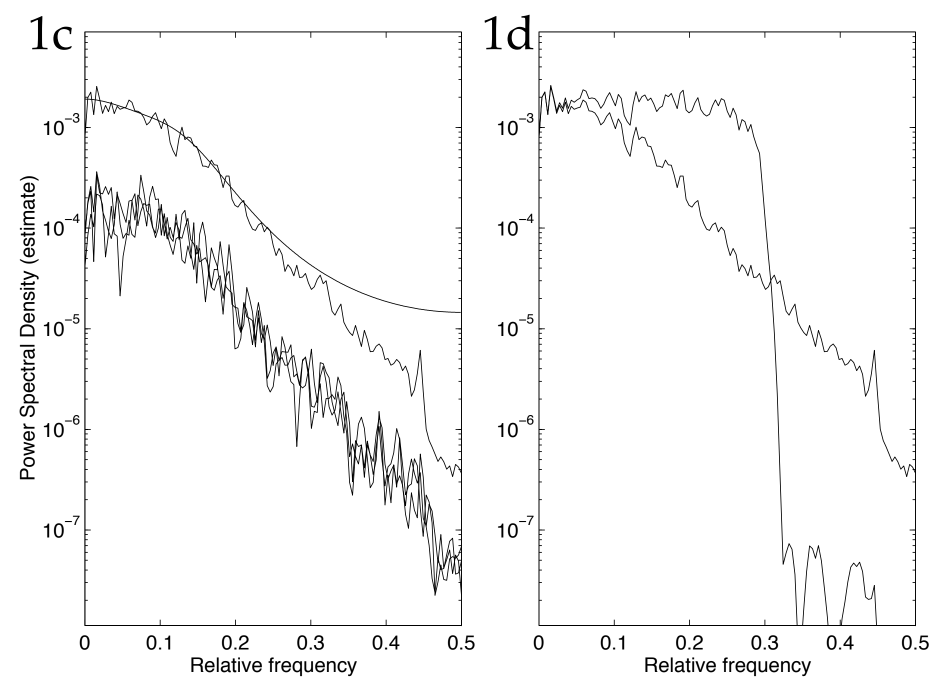
A Markov chain Monte Carlo (MCMC) sampler was implemented. This is rather complex, especially since the order of the model varies with the number of CPs (transitions). Metropolis-Hastings moves for CP insertion and removal were combined with Gibbs-type moves for relocating CPs and drawing samples of conductances. Three categories of moves were used:

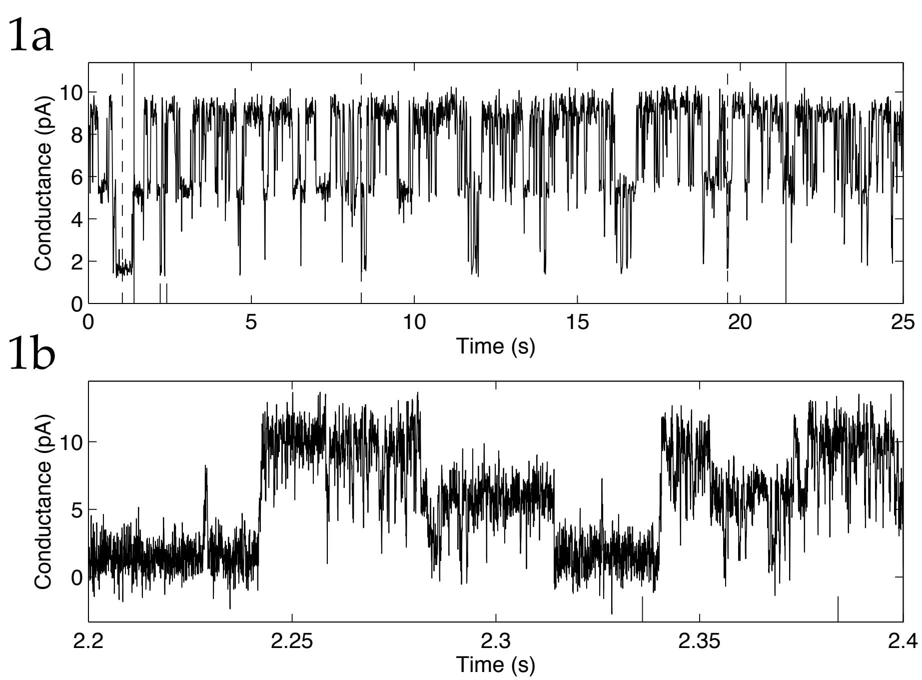
A. Relocating a CP between the preceeding and succeeding CPs, and reselecting the conductance level either side.

B. Inserting or removing one CP. When the CP was left out, the level for the combined segment was reselected, and when it was put in the levels either side were reselected.

FIGURE 1: THE SIGNAL, NOISE AND TRANSITION

Plot 1a is the patch-clamp record, smoothed slightly for clarity. Red/dashed bars locate three segments of noise. Green/continuous bars delimit the section which was analysed in 100 blocks. The fifth block is delimited by two shorter bars, and is displayed in plot 1b (raw form). Plot 1c shows the 1024-point estimate power spectral density (psd) for the noise locations. These have been shifted to separate them from the psd for the first location with 8192 points, to which a 5-th order AR model was fitted with an additive adjustment; its psd is overlayed. Plot 1d shows the noise psd before and after whitening and additional lowpass filtering. Plots 1e and 1f are the corresponding model transition shapes.





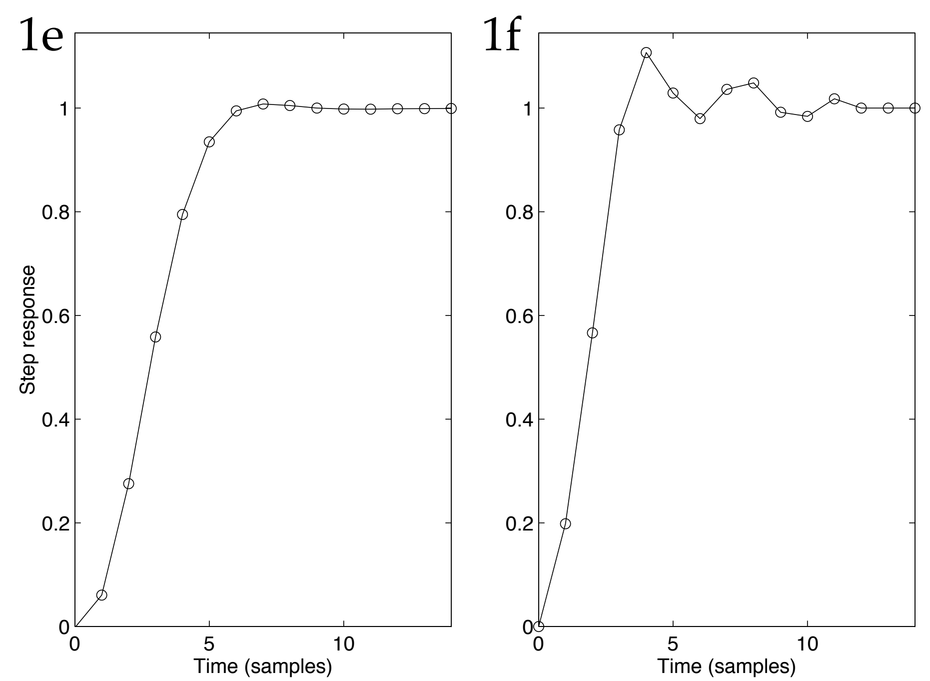
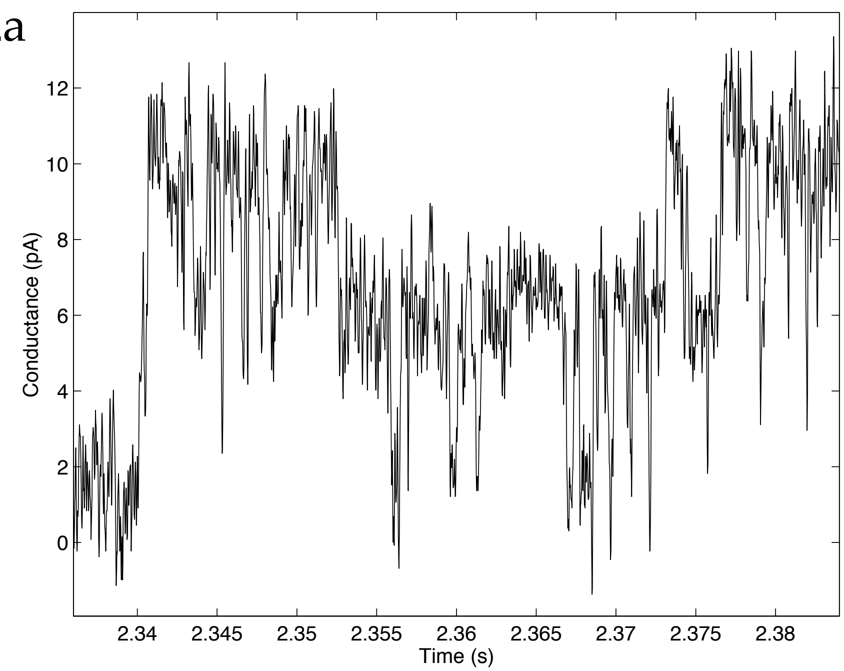


Figure 2: Sampler Output

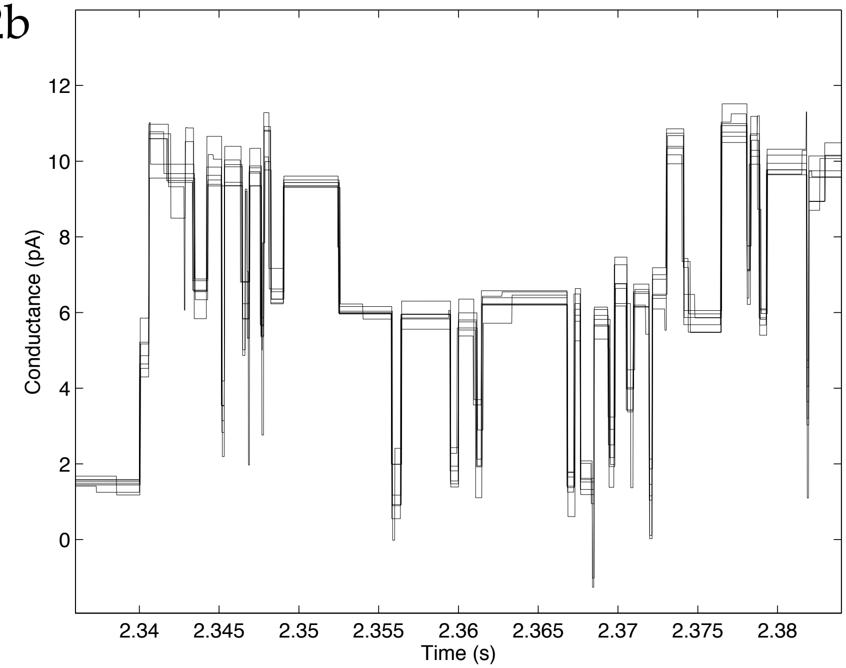
Plot 2a is a detail of the patch-clamp record, delimited by the short bars in plot 1b. The sampler was run, for this block (5), for 5346 burn-in iterations. Results were then collected for (a subset of) the next 8341 iterations. There are a variety of methods for presenting the sample fits. Plot 2b is a set of 7 examples taken evenly along the sampler run. This gives an indication of the way the modelling has handled the ambiguities in the measured signal, such as around time 2.347s. An alternative view is shown in plot 2c, which is the mean of the idealised signals.

Plot 2d shows the trace, as the sampling progressed, of the number of changepoints. The dashed line indicates the end of the burn-in period.

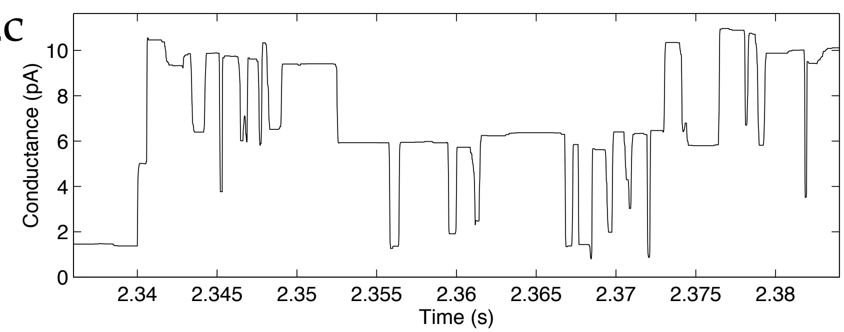
2a



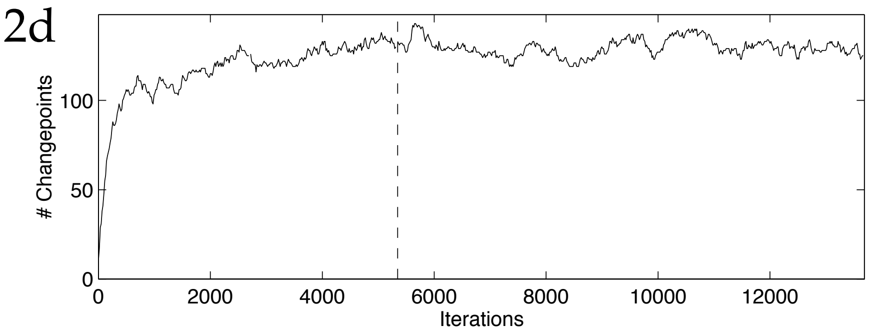
2b



2c



2d



C. Inserting or removing two CPs. The location of one of these was generated from the output of a detection filter. The details of this are very complex. Although the method of choosing the location influences the efficiency of the sampler, it does not affect the equilibrium distribution provided that detailed balance is maintained and convergence is achieved. The other details of these moves were similar to those of the one-CP moves.

The model also required prior estimates of the rate of events and size of transition. By inspection of the record these were set to 0.01 events per sample and 6.7pA. They were verified against the sampler output, and the results were not sensitive to their values. The (Gaussian) form of the prior permits transitions of any size and does not heavily penalise large ones.

The Fifth Block

The results for the fifth analysis block are shown in Fig. 2. There are many ambiguities in the measured signal. We were particularly interested in the short-duration events and whether or not there were partial closures.

Displaying a multidimensional distribution is difficult. The sampler provides a powerful method of exploring the distribution of fits, and insights can be gained by observing a sequence dynamically. A limited number of instances can be usefully displayed on paper (Fig. 2b), but the importance of any individual fit should be deemphasised. The mean conductance at each point (Fig. 2c) across the samples gives a different perspective. The progress of the sampler is illustrated by the trace of the number of CPs in Fig. 2d; the runs for each block were relatively short.

Histograms

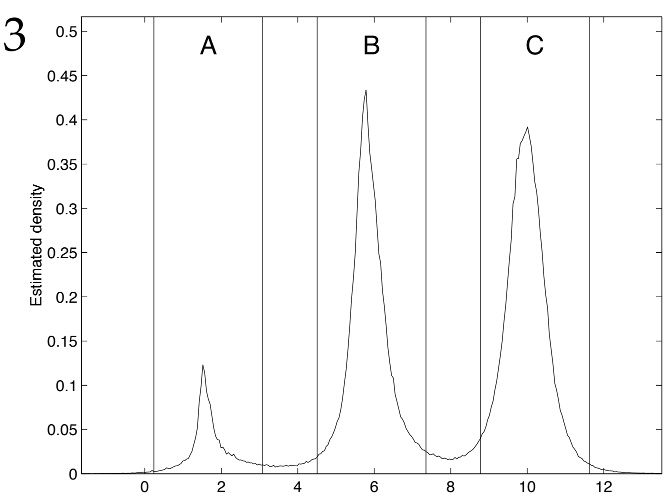
One of the best representations of samples from distributions is the histogram, which gives the marginal distribution of (a function of) one or more parameters. An all-points histogram of conductances can be generated by replacing the signal with sample idealisations. Such a histogram of samples for all signal blocks is shown in Fig. 3. There appear to be three separate levels.

A richer description (Fig. 4a) is a dual histogram, that is one across both conductances and segment sojourn time (duration). In this case there appear to be separate duration components for the lower two conductance bands. However, if more of the sequential dynamics of the channel is to be discerned, it is necessary to collect results based on the characteristics of successive segments.

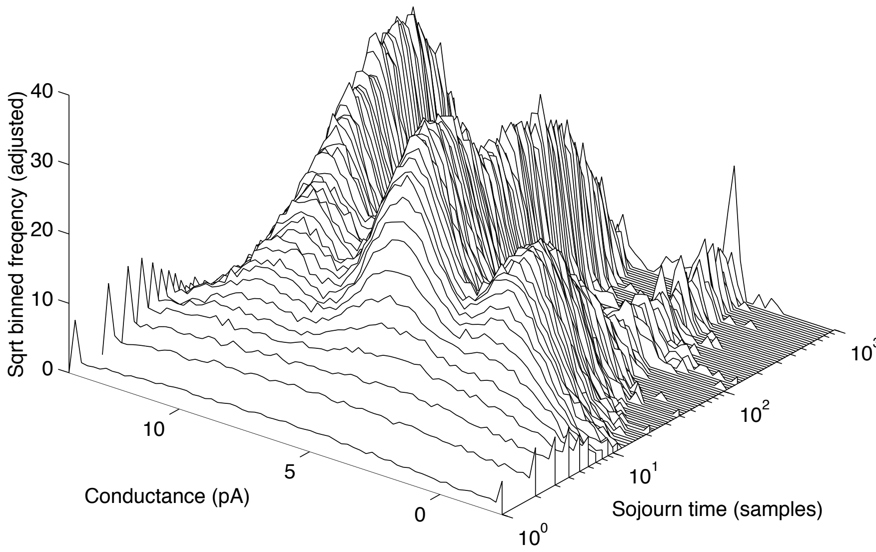
Figure 3: Idealised All-points Histogram

Plot 3 is the all-points histogram of the sample idealisations. This was formed from the conductance and duration of each segment between transitions in the sample idealised traces. For later analysis, signal segments were classed by conductance: A/B/C/other.

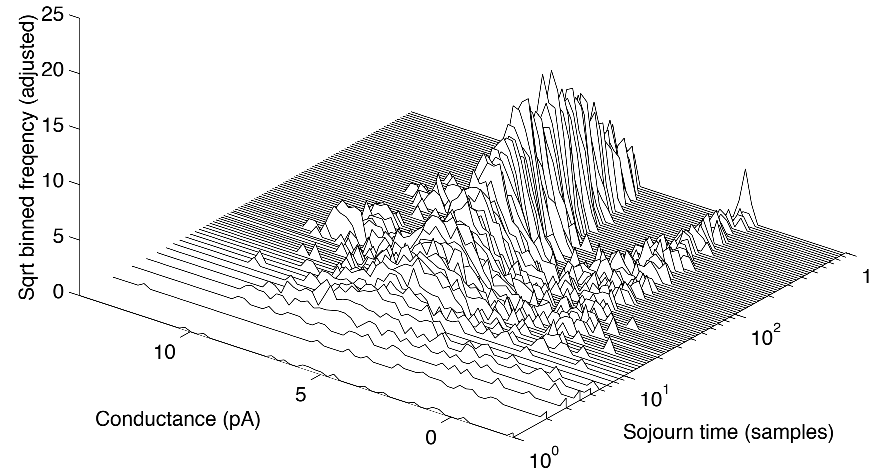
3



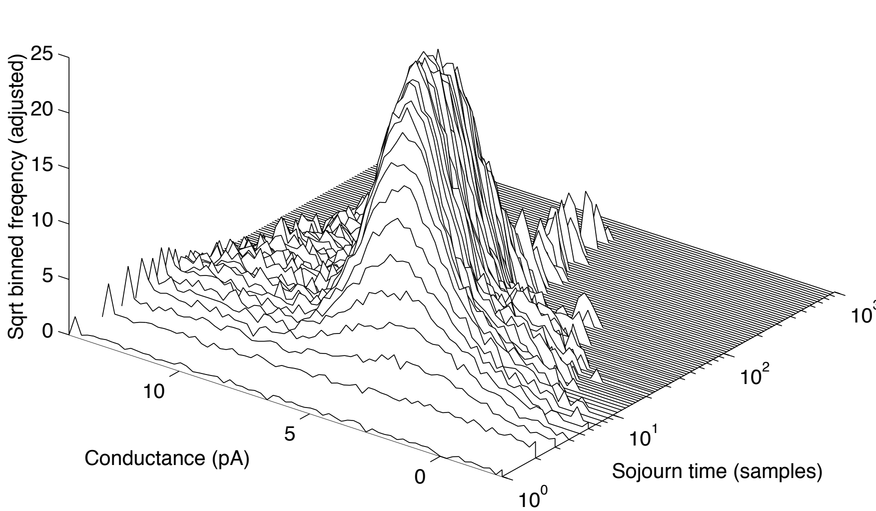
4a: All segments



4c: A∗A



4b: C∗C



4d: B∗B

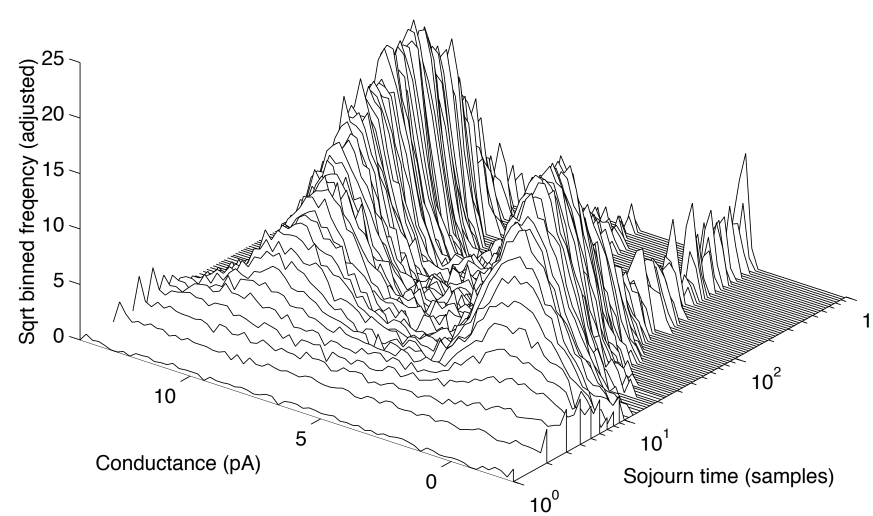


Figure 4: Dual Histograms

Plot 4a shows the (square-root) histogram of segments binned across conductances and log-sojourn times for the sample idealised traces. Classifying the segments by preceeding and succeeding conductances produces a rudimentary separation of the components. Plot 4b ('C∗C') is the histogram for segments with class C conductances (plot 3) before and after. Plots 4c and 4d are for those segments with A and B conductances either side.

A possible interpretation is that there are two independent channels, each with an open and closed state, and a short-duration closed state. The C∗C plot (4b) would then be the durations of short closures from both-open. The A∗A plot (4c) would be the durations *between* short closures from one-open, and the B∗B plot (4d) would be a combination of such effects.

Interpretation

Dealing with the histogram for two neighbouring segments would involve storage and visualisation problems, and there would be few instances in any histogram bin. Furthermore, an analysis of short excursions from one conductance to another requires treatment of three successive segments.

One approach is to use coarse binning of some of the parameters. In this case a rough separation of the constituents of the dual histogram was possible by generating a set of histograms for different preceeding and succeeding conductance ranges (A, B and C in Fig. 3). The histogram of segment levels and durations with segments of conductance type C before and after ('C∗C', Fig. 4b) suggests that there are short closures that return to the open state. There is a complementary component for the times *between* such events, as shown in the A∗A histogram (Fig. 4c). The B∗B histogram appears to have two consituent parts: one for excursions to A and one for intra-excursion segments in C.

These results appear consistent with the hypothesis that there are two independent channels, each with an open and closed state and a further short-duration closed state. Note that the chosen signal record has a high proportion of openings, and so the relative times spent in each state is unrepresentatve.

Conclusions

A variety of emerging statistical modelling techniques are being applied to the signals obtained from ion channels, and a number of these use MCMC methods. Model complexities vary, and in principle everything that is known about the signal should be included to maximise the quality of inferences that are drawn. However there are many situations, especially at early stages of data exploration, in which a simple model is desirable.

The model employed in this research was designed to minimise the number of assumptions whilst accounting for the noise spectrum and filtering. The conductances allowed for each segment were unconstrained, and the fit is localised. Despite the lack of model structure, it was possible to explore the signal quite extensively, and to propose a tentative hypothesis for the underlying process. It would be appropriate to take the signal on to analysis with a hidden Markov model. We would suggest that before doing so it might be analysed using an intermediate model, such as one in which the permitted conductances are constrained to a discrete global set.

FIGURE 5: TEMPLATE HISTOGRAM

Plot 5 shows a theoretical distribution of conductances and sojourns. This was derived for the case when the transition locations are correctly estimated. The sojourn times are exponentially distributed (before transformation to logarithmic axis). The sampled conductance levels for a given conductance are Gaussian.

The mean conductances and durations of the two components were selected by eye to approximate the histogram of plot 4d. The match is surprisingly close, given that there are missed and false events in the sampler output.

5