PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Persistence: A New Statistic for Characterizing Ion-Channel Activity

D. R. Fredkin, J. A. Rice, D. Colquhoun and A. J. Gibb

Phil. Trans. R. Soc. Lond. B 1995 **350**, 353-367
doi: 10.1098/rstb.1995.0170

Persistence: a new statistic for characterizing ion-channel activity

D. R. FREDKIN[1], J. A. RICE[2], D. COLQUHOUN[3] AND A. J. GIBB[3]

[1] *Department of Physics, University of California, San Diego, U.S.A.*
[2] *Department of Statistics, University of California, Berkeley, U.S.A.*
[3] *Department of Pharmacology, University College London, London, U.K.*

SUMMARY

We introduce and illustrate by examples a new statistical technique, the persistence function, for characterizing ion-channel activity in a single-channel patch-clamp recording. Persistence is a function of both current and time. It is the probability that the current is at a given level (conditional on it having been at that level at an earlier time). Viewed as a function of current it exhibits the prominent conductance levels present in the recording, and viewed as a function of time for a conductance level it portrays the kinetics at that level.

1. INTRODUCTION

The duration and amplitude of receptor-channel openings are of interest because they provide information about conformational changes occuring within the receptor protein, and about the gating of the ion channel. However, it is difficult to measure automatically the amplitude and duration of channel openings when there are many conductance levels.

A crude measure of ion-channel multiple conductance levels can be obtained from the histogram of all points in the data record, but it gives no information on the duration of channel openings. A more sensitive method of detecting channel conductance levels was described by Patlak (1988): the mean low-variance histogram.

In this paper we present a simple and effective method for the preliminary analysis of single-channel patch-clamp recordings. The analysis focuses on identifying conductance levels and qualitatively characterizing the kinetics associated with each of those levels. This is accomplished by the estimation and graphical presentation of an object we call the 'persistence function'.

For motivation, consider figure 1, which shows traces of three segments of a recording of NMDA receptor channel activity from a CA1 neurone in a rat hippocampal slice. The recording was filtered at 2 kHz and sampled at 20 kHz. Considerable variability in open current levels is displayed in this figure. In the top panel there are levels at about -2.2, -2.7, and -3.2 pA; in the middle panel there are levels at about -1.9 and -3.0 pA; the lower panel shows levels at about -1.5 and -2.5 pA. Although the reader might quarrel with these assignments (which were done graphically), the point is that there is such substantial variability in current levels that unambiguous assignment of the activations in the entire record to discrete conductance levels is a problematical undertaking. The reasons for this variability are not clear; from the figure we see that it cannot be explained by baseline drift. It is possible that the traces are from activations of different channels, which though identical in primary structure, are not identical in tertiary structure, due to cytoskeletal influences or post-translational modifications, for example. It is also possible that the recording is from a population of channels of varying molecular composition. However, multiple current levels are commonly seen during a single burst of activity where the channel open probability is high enough to suggest only a single channel.

A histogram of current levels is shown in figure 2. The sublevels apparent in figure 1 are not captured by the histogram. Figure 3 shows a mean low-variance plot (Patlak 1988) which reveals a number of sublevels.

Figure 4 displays the persistence function, estimated from the entire recording without any further attempt at restoration. The persistence, which will be defined precisely in the next section, is a function of two arguments (current and time) and is the probability that the recording is at a given current level conditional on it being at that current level t time units earlier. In addition to a level at 0 (no current), it shows the existence of current levels, or bands, centered around -0.8, -1.7, -2.5 and -3.2 pA, which are also apparent in the mean low-variance plot (see figure 3*a*). Levels at current more negative than -4.0 pA are due to the existence of simultaneous openings of more than one channel. The decay of the corresponding ridges with time reflects the persistence of openings at those levels and we see that the -3.2 pA level has the slowest kinetics of the non-zero current levels.

The remainder of the paper is organized as follows.

Phil. Trans. R. Soc. Lond. B (1996) **350**, 353–367
Printed in Great Britain

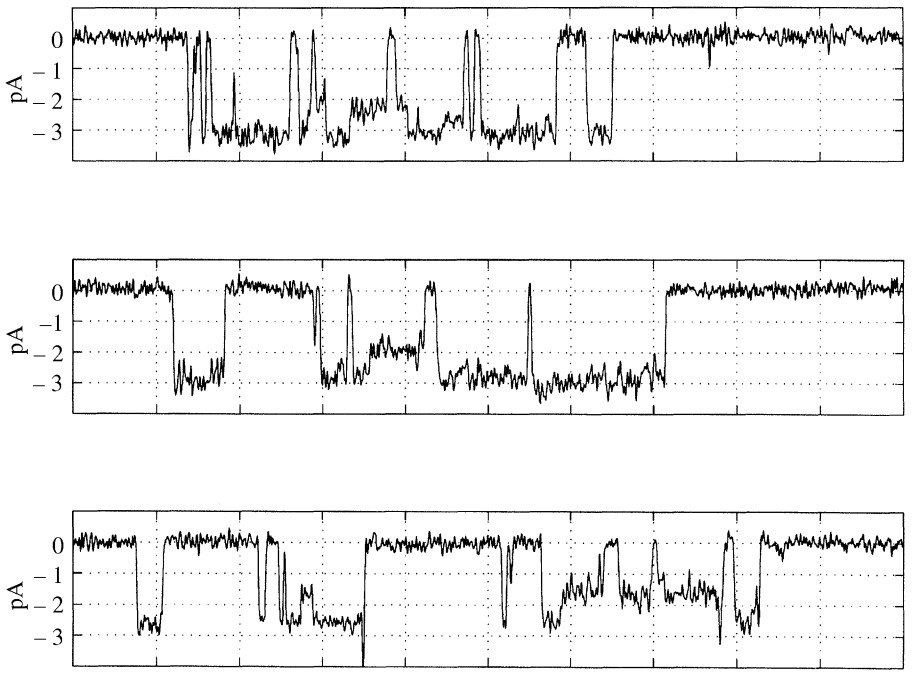

Figure 1. Three 100 ms segments of a recording from NMDA receptor activity in an outside-out patch isolated from a CA1 neurone in a rat hippocampal slice. The patch membrane potential was −60 mV and the channels were activated with 30 nM glutamate and 1 μM glycine. The recording was filtered at 2 kHz (−3 dB, Bessel characteristic) and sampled at 20 kHz.

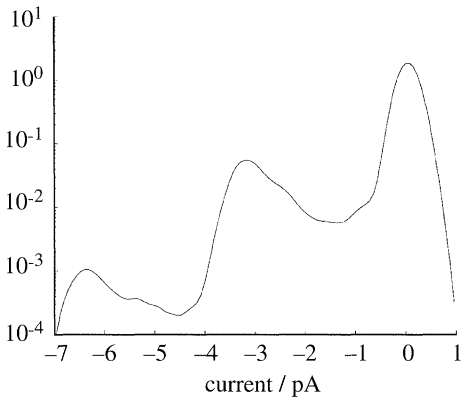

Figure 2. Smoothed (point amplitude) histogram of the data excerpted in figure 1.

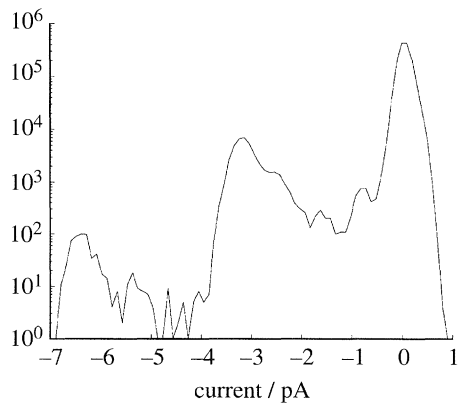

Figure 3. Mean low-variance (amplitude) histogram of the data excerpted in figure 1. We used ten data points in the running mean. The RMS noise was 0.161704 pA.

In §2 we define the persistence function of a random process and discuss its properties and interpretation. In §3 we present a method of estimating the persistence function. §4 is devoted to examples of the use of the persistence function, including a discussion of the effects of filtering. §5 consists of some concluding remarks.

2. DEFINITION AND PROPERTIES OF THE PERSISTENCE FUNCTION

We define the persistence function, illustrate its utility in the presence of noise, and give expressions corresponding to several common mathematical models. We also indicate that the persistence function continues to be useful for studying multichannel systems.

Let $X(t)$ be the current passing through an ion channel at time t. We will assume throughout that the channel is in equilibrium so that X is a stationary random process, and we assume initially that the current takes on a discrete set of values I_1, I_2, I_C, so that X is a step function. The time t may be continuous or discrete (sampled data). We define the persistence function as:

$$\rho(c,t) = P[X(u+t) = I_c | X(u) = I_c].$$

In other words, $\rho(c,t)$ is the conditional probability that the current is equal to I_c at time $t+u$ given that it was equal to I_c at time u. Note that, as in calculations of noise or macroscopic relaxations (but unlike most single-channel calculations) there is no requirement that the channel stays at the same current level throughout the time between u and $u+t$. Because the process is assumed to be stationary, ρ does not depend

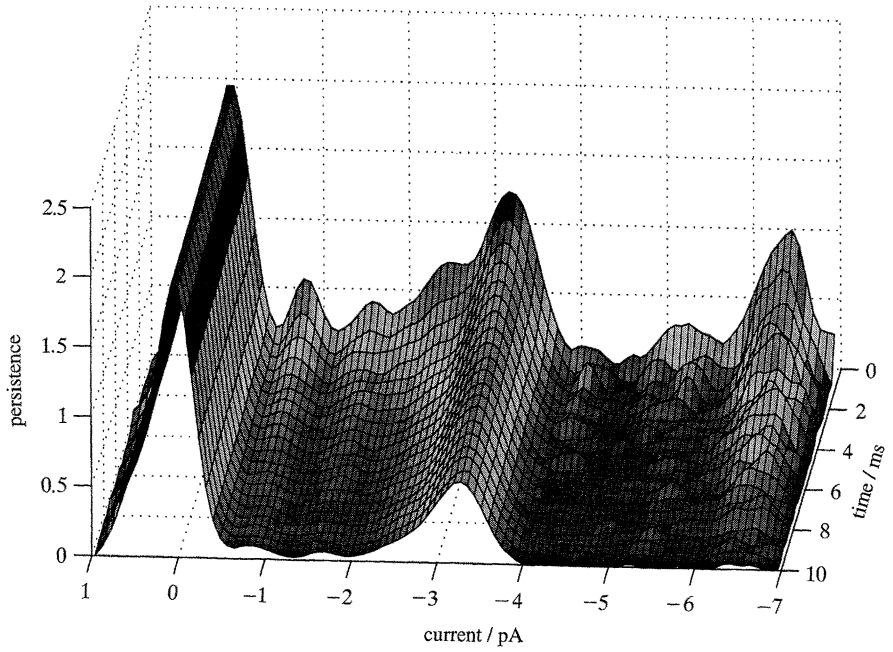

Figure 4. Persistence function of the data excerpted in figure 1.

on u and we may thus take $u = 0$ without loss of generality. We note that the persistence function is a function of two variables, c and t.

For fixed c, $\rho(c, t)$ as a function of t describes how the 'memory' of having been at current level c decays over time. As $t \to \infty$, $\rho(c, t) \to \pi_c$, the equilibrium probability of current level c, and as $t \to 0$, $\rho(c, t) \to 1$. Comparison of these curves for different current levels c reveals the relative strength of memory, or persistence, of the various current levels.

The persistence function may be viewed as a kind of correlation function and description of a record in terms of the persistence function as a kind of noise analysis. To see this, suppose that, for a given current level c, a new random process $Z_c(t)$ is defined to be $Z_c(t) = 1$ if $X(t) = I_c$ and 0 otherwise. The autocovariance function of Z_c is then:

$$\gamma(t) = \mathrm{Cov}\,[Z_c(u), Z_c(u+t)]$$
$$= E[Z_c(u)\, Z_c(u+t)] - E(Z_c(u))\, E[Z_c(u+t)].$$

Now, because Z_c only takes on values 0 and 1 and is stationary,

$$E(Z_c(u)) = 1 \times P(Z_c(u) = 1) + 0 \times P(Z_c(u) = 0)$$
$$= \pi_c.$$

Similarly,

$$E[Z_c(u)Z_c(u+t)] = P[Z_c(u) = Z_c(u+t) = 1]$$
$$= P[Z_c(u+t) = 1 | Z_c(u) = 1]$$
$$\times P(Z_c(u) = 1)$$
$$= \rho(c, t)\pi_c.$$

We thus have:

$$\gamma(t) = \pi_c \rho(c, t) - \pi_c^2.$$

As a specific example, consider the very simple two state Markov model C↔O. We take time to be discrete, and let a and b be the one time step probabilities for the transitions C→O and O→C respectively. By straightforward calculation, the persistence function for the state O is:

$$\rho_O(t) = [a + b(1-a-b)^t]/(a+b). \quad (1)$$

Note that $\rho_O(0) = 1$, that $\rho_O(t)$ has the correct limiting behavior for $t \to \infty$, $\pi_0 = a/(a+b)$, and that $\rho_O(t)$ has the same time constant as would be found for noise or macroscopic relaxation. We generated a sample of length $T = 100000$ using $a = 0.01$ and $b = 0.05$, and then we estimated $\rho_O(t)$ for $t = 1 \ldots 300$ using:

$$\hat{\rho}_O(t) = \frac{\Sigma_{u=1}^{T-t} Z_O(u)\, Z_O(u+t)}{\Sigma_{u=1}^{T-t} Z_O(u)}.$$

(We can imagine that we have 5 s of data, sampled at 20 kHz). The result is shown in figure 5 as the curve labeled 'no errors'. The expression (1) is also shown by

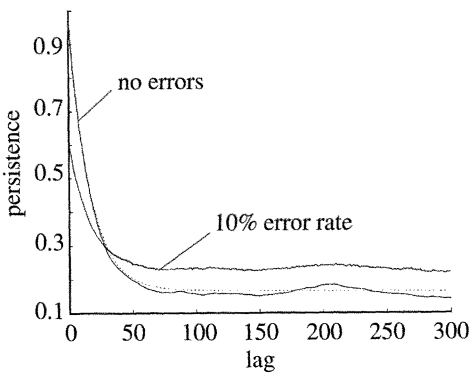

Figure 5. Persistence as a function of lag for a two state model. The dotted curve shows the theoretical persistence; the labelled solid curves show the persistence function estimated from a simulation, with and without state identification errors.

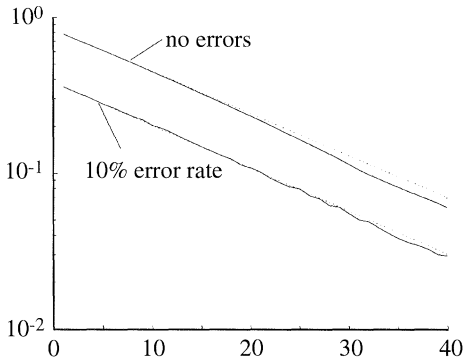

Figure 6. Semilogarithmic plots of $\hat{\rho}_O(t) - \bar{Z}_O$ (compare the two solid curves in figure 5), restricted to small lags. The dotted lines are straight lines fit to the solid curves. The corresponding estimates for the total one step transition probability $(a+b)$ are: theory 0.0600, without errors 0.0599, with errors 0.0611.

Table 1. *Results of dwell time analysis*

(τ_C denotes mean dwell time in closed state; τ_O denotes mean dwell time in open state.)

source	τ_C	τ_O	$a+b$
theory	100.0	20.0	0.06
no errors	98.31	20.06	0.0600
10% errors	7.94	2.46	0.532

the dotted curve. The estimated persistence is seen to agree with the theoretical value provide it has not yet approximately reached its asymptotic value. For long lags, the estimated persistence oscillates and the sampling fluctuations can be shown to have substantial autocorrelation. These properties are reminiscent of any autocovariance estimate.

Although the next section contains a detailed discussion of the effect of additive noise, we can get a preliminary idea of the persistence function's robustness by deliberately introducing errors with probability 0.1. At each time we reversed, with probability 0.1, the state obtained in our simulation of C↔O. (We can imagine that we have reconstructed the ideal signal from noisy data, and we are wrong 10% of the time.) The estimated persistence function for the corrupted data is shown in figure 5 as the curve labelled '10% error rate'. What useful information about the underlying Markov process is available in the estimate $\hat{\rho}_O(t)$ made from the corrupted data? Figure 6 shows a semilog plot of the estimated persistence functions for $t = 1 \ldots 40$. We see that the two curves are parallel straight lines for small t. We can estimate $a+b$ from the slope; the values we obtain in this way are given in the legend. We conclude that some kinetic information can be obtained reliably from the persistence function for noisy data. For comparison, we performed a conventional dwell time analysis on the simulated data, before and after introduction of errors, with the results shown in table 1. As expected, errors in state identification have a serious effect on the dwell times and makes estimation of kinetic parameters impossible.

The example demonstrates that the persistence function contains reliable information at short lags t, but the information at long lags is not reliable. Therefore, it is not useful to characterize the kinetic time scales by the integrated persistence $\int_0^\infty \mathrm{d}t\,(\rho_c(t) - \pi_c)$.

Our simple example also lets us see the advantage of our definition of the persistence function as a conditional probability, rather than as a joint probability. We shall see, when we apply persistence analysis to noisy data, that consideration of the persistence as a function of current level c at fixed lag t is a valuable method for detection of significant current levels. For our example, we have $\rho_O(1) = 0.95$ and $\rho_C(1) = 0.99$. By contrast, the joint probabilities $P[X(t) = O \;\&\; X(t+1) = O] = 0.158$ and $P[X(t) = C \;\&\; X(t+1) = C] = 0.285$. Although the persistence is merely a renormalized version of the joint probability, it emphasises low probability current levels in a useful way.

(a) *Various Models*

This section presents examples of theoretical persistence functions for various models in common use for the interpretation of single-channel patch-clamp recordings. None of the results are used in subsequent sections. This section, or any of its parts, can be omitted without loss of continuity.

(i) *Aggregated Markov Process*

The aggregated Markov process is the most popular model for channel kinetics. Here we have a finite state Markov process with states $s = 1, 2, \ldots, S$ and the current is J_s when the channel is in state s. In general, the current does not uniquely determine the state; for example, we might have $S = 3$, $C = 2$, and $J_1 = J_2 = I_1$ and $J_3 = I_2$. We again work with discrete time. Let the $S \times S$ matrix $\bar{P}$ be the one step transition matrix for the Markov process, and let the row vector $\bar{\pi}$ contain the equilibrium probabilities. Define the column vector u^c by:

$$u^c_s = \begin{cases} 1 & \text{if } J_s = I_c \\ 0 & \text{otherwise,} \end{cases}$$

and the row vector π^c by:

$$\pi^c_s = \begin{cases} \bar{\pi}_s & \text{if } J_s = I_c \\ 0 & \text{otherwise.} \end{cases}$$

Then the equilibrium probability for the current level c is $\bar{\pi}u^c$, the joint probability $P[X(u) = c \;\&\; X(t+u) = c]$ is $\pi^c \bar{P}^t u^c$, so the persistence function is:

$$\rho(c,t) = (\pi^c \bar{P}^t u^c)/(\bar{\pi}u^c). \qquad (2)$$

(ii) *Alternating Renewal Process*

An alternating renewal process (Cox 1962) is another model for channel kinetics. For the sake of variety, we work with continuous time.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

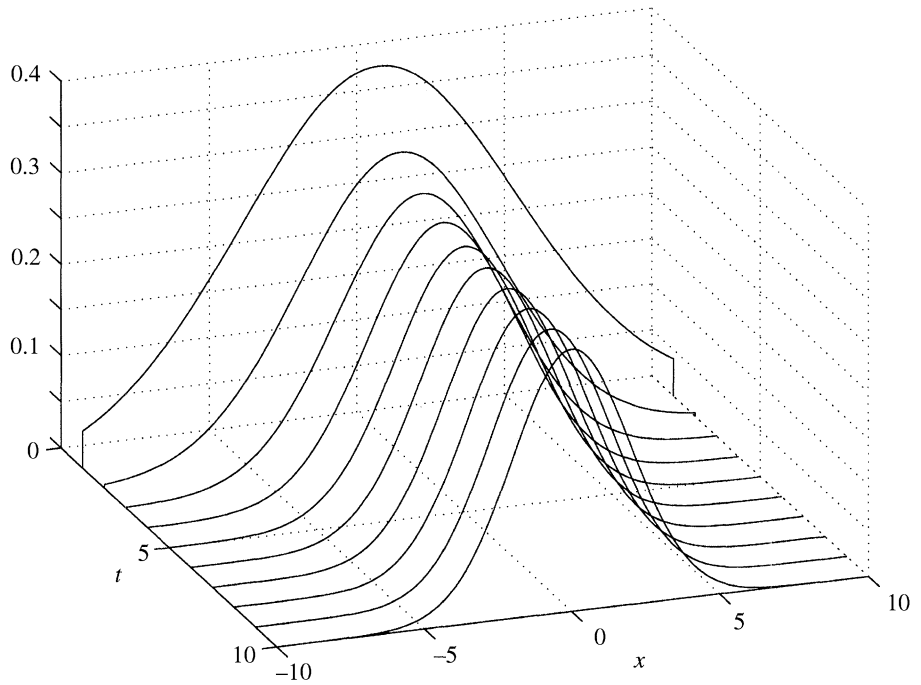

Figure 7. Persistence function for the autoregressive process $X(t) = 0.7X(t-1) + W(t)$, where $W(t)$ are independent Gaussian random variables with zero mean and unit variance.

Let $f_O(t)$, $f_C(t)$ be the probability densities for the open and closed state dwell times, respectively, and let τ_O and τ_C be the mean dwell times. It is convenient to introduce the survival functions:

$$\bar{F}_{O,C}(t) = \int_t^\infty \mathrm{d}u f_{O,C}(u).$$

We shall compute the persistence function for this model.

The probability that the channel is open at time u and has remained open until $u+t$ or later is:

$$P_O^{(0)}(t) = [1/(\tau_O+\tau_C)] \int_0^\infty \mathrm{d}t_0 \bar{F}_O(t_0+t), \tag{3}$$

where the integration variable t_0 may be thought of as the time since the last transition into the open state (Cox 1962). Similarly, the probability that the channel is open at u and at $u+t$, with exactly one sojourn is the closed state between u and $u+t$, is:

$$P_O^{(1)}(t) = [1/(\tau_O+\tau_C)] \int_0^\infty \mathrm{d}t_0 \int_0^t \int_0^{t-t_1} \mathrm{d}t_2 f_O(t_0+t_1) f_C(t_2) \bar{F}_O(t-t_1-t_2), \tag{4}$$

where the channel closed at $u+t_1$ and reopened at $u+t_1+t_2$. An expression for $P_O^{(n)}(t)$, the probability that the channel is open at u and at $u+t$, with exactly n sojourns in the closed state between u and $u+t$, can be written in the same way and like equation 4, takes the form of a multiple convolution. It is therefore convenient to use Laplace transforms which we denote, for example, by $f_O^*(s) = \int_0^\infty f_O(t) \exp(-st)\,\mathrm{d}t$. We have:

$$P_O^{(0)*}(s) = (\tau_O - \bar{F}_O^*(s))/[s(\tau_O+\tau_C)],$$

and, for $n \geqslant 1$,

$$P_O^{(n)*}(s) = (\bar{F}_O^*(s)^2 f_C^*(s)^n f_O^*(s)^{n-1})/(\tau_O+\tau_C).$$

The joint probability that the channel is open at times u and $u+t$ is then, given by:

$$P_O^*(s) = \frac{1}{\tau_O+\tau_C}\left[\frac{\tau_O - \bar{F}_O^*(s)}{s} + \frac{\bar{F}_O^*(s)^2 f_C^*(s)}{1-f_C^*(s)f_O^*(s)}\right],$$

and because $\pi_O = \tau_O/(\tau_O+\tau_C)$, the Laplace transform of the persistence function is:

$$\rho_O^*(s) = P_O^*(s)/\pi_O = \frac{1}{\tau_O}\left[\frac{\tau_O - \bar{F}_O^*(s)}{s} + \frac{\bar{F}_O^*(s)^2 f_C^*(s)}{1-f_C^*(s)f_O^*(s)}\right]. \tag{5}$$

It is helpful to remember, when applying equation 5, that:

$$\bar{F}_O^*(s) = (1-f_O^*(s))/s,$$

and

$$\bar{F}_O^*(0) = \int_0^\infty \bar{F}_O(t)\,\mathrm{d}t = \tau_O.$$

(iii) Continuous State Space: a First-Order Autoregressive Process

We usually think of a channel as having a discrete set of conductances, and the observed continuous range of current is regarded as a result of additive noise and the effect of filtering. Nevertheless, it is pedagogically useful to consider a simple example of a Markov process with a continuous state space. It requires exactly the same work to consider a diffusion process in

continuous time or a first order autoregressive process in discrete time. For simplicity, we discuss the latter.

Suppose, then, that the current $X(t)$ evolves according to:

$$X(t) = aX(t-1) + W(t),$$

where $|a| < 1$ and $W(t)$ are independent random variables, each with a Gaussian distribution with mean 0 and variance $EW(t)^2 = \sigma^2$. Then $X(t) = \Sigma_{n=0}^{\infty} a^n W(t-n)$, so $X(t)$ are jointly Gaussian with mean 0 and variance $\sigma^2/(1-a^2)$. Further, using the general properties of multidimensional Gaussian distributions, or from $X(t) = a^t X(0) + \Sigma_{n=0}^{t-1} a^n W(t-n)$, we see that, conditional on $X(0)$, $X(t)$ has a Gaussian distribution with mean $a^t X(0)$ and variance:

$$\mathrm{Var}\,[X(t)\,|\,X(0)] = \sigma^2\,(1-a^{2t})/(1-a^2).$$

When the current assumes values in a continuum, the persistence function is defined as a probability density. In our case, we have:

$$\rho(x,t) = \sqrt{\left[\frac{1-a^2}{2\pi\sigma^2(1-a^{2t})}\right]}\exp\left[-\frac{(1-a^2)\,(1-a^t)}{2\sigma^2(1+a^t)}x^2\right]. \tag{6}$$

Figure 7 is a plot of equation with $a = 0.7$ and $\sigma = 1$ for $t = 1 \ldots 10$. Note that, in contrast to the case of a discrete state space, the persistence is undefined at $t = 0$ because it is defined as a probability density, which does not exist for zero lag.

(*b*) *Multiple Channels*

The persistence function is useful for analysing data from multiple channels. A general discussion requires examination of many cases (see, for example, Fredkin & Rice 1991). Here we confine ourself to the illustrative case of two independent identical channels, each of which can carry current zero or one (in suitable units). Let the current carried by channel i be X_i, $i = 1, 2$, and let the observed current be $X = X_1 + X_2$, which assumes the values 0, 1, 2. Then:

$$\begin{aligned}
&P[X(u) = 0 \;\&\; X(u+t) = 0] \\
&\quad = P[X_1(u) = 0 \;\&\; X_2(u) = 0 \\
&\qquad \&\; X_1(u+t) = 0 \;\&\; X_2(u+t) = 0] \\
&\quad = P[X_1(u) = 0 \;\&\; X_1(u+t) = 0] \\
&\qquad \times P[X_2(u) = 0 \;\&\; X_2(u+t) = 0],
\end{aligned}$$

and

$$\begin{aligned}
P[X(u) = 0] &= P[X_1(u) = 0 \;\&\; X_2(u) = 0] \\
&= P[X_1(u) = 0]\,P[X_2(u) = 0],
\end{aligned}$$

so

$$\rho_0(t) = \rho_0^1(t)^2, \tag{7}$$

where $\rho^1(t)$ is the persistence function for a single channel. Similarly,

$$\rho_2(t) = \rho_1^1(t)^2. \tag{8}$$

The single channel persistence functions are, in principle, calculable from the two channel persistence functions.

The result, equation 8, is not very satisfactory because it requires use of the persistence function for double events, $\rho_2(t)$. These are usually rare, by design, so we cannot expect to estimate $\rho_2(t)$ very well. On the other hand, the more accessible persistence function $\rho_1(t)$ does not have any simple expression in terms of the persistence functions for a single channel. It is useful at this point to introduce the concept of cross persistence, analogous to cross correlation: For current levels c, c' we define, generally,

$$\rho_{cc'}(t) = P[X(u+t) = I_{c'}\,|\,X(u) = I_c].$$

Then we have:

$$\begin{aligned}
\rho_{01}(t) &= P[X(u+t) = 1\,|\,X(u) = 0] \\
&= P[X_1(t+u) = 1 \;\&\; X_2(u+t) \\
&\qquad = 0\,|\,X_1(u) = 0 \;\&\; X_2(u) = 0] \\
&\qquad + P[X_1(t+u) = 0 \;\&\; X_2(u+t) \\
&\qquad = 1\,|\,X_1(u) = 0 \;\&\; X_2(u) = 0] \\
&= P[X_1(t+u) = 1\,|\,X_1(u) = 0]\,P[X_2(u+t) \\
&\qquad = 0\,|\,X_2(u) = 0] \\
&\qquad + P[X_1(t+u) = 0\,|\,X_1(u) = 0]\,P[X_2(u+t) \\
&\qquad = 0\,|\,X_2(u) = 0] \\
&= 2\rho_{01}^1(t)\,\rho_0^1(t),
\end{aligned} \tag{9}$$

and, similarly, because

$$\begin{aligned}
&P[X(u+t) = 0 \;\&\; X(u) = 1] \\
&\quad = P[X_1(u+t) = 0 \;\&\; X_2(u+t) = 0 \;\&\; X_1(u) \\
&\qquad = 1 \;\&\; X_2(u) = 0] + P[X_1(u+t) = 0 \;\&\; X_2(u+t) \\
&\qquad = 0 \;\&\; X_1(u) = 0 \;\&\; X_2(u) = 1] \\
&\quad = P[X_1(u+t) = 0 \;\&\; X_1(u) = 1]\,P[X_2(u+t) \\
&\qquad = 0 \;\&\; X_2(u) = 0] + P[X_1(u+t) = 0 \;\&\; X_1(u) \\
&\qquad = 0]\,P[X_2(u+t) = 0 \;\&\; X_2(u) = 1] \\
&\quad = P[X_1(u) = 1]\,\rho_{10}^1(t)\,P[X_2 = 0]\,\rho_0^1(t) \\
&\qquad + P[X_1(u) = 0]\,\rho_0^1(t)\,P[X_2(u) = 1]\,\rho_{10}^1(t) \\
&\quad = 2\pi_0^1\pi_1^1\rho_0^1(t)\,\rho_{10}^1(t),
\end{aligned}$$

where π_c^1 are the probabilities that a single channel is found carrying current c,

$$\rho_{10}(t) = \rho_0^1(t)\,\rho_{10}^1(t). \tag{10}$$

Further, a similar argument leads to:

$$\rho_1(t) = \rho_1^1(t)\,\rho_0^1(t) + \rho_{01}^1(t)\,\rho_{10}^1(t). \tag{11}$$

Suppose we have determined $\rho_{ij}(t)$, where i and j take only the values 0 and 1, from a channel recording. Then, from equation 7 we know $\rho_0^1(t)$. Next, we can find the single-channel cross-persistence functions $\rho_{01}^1(t)$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

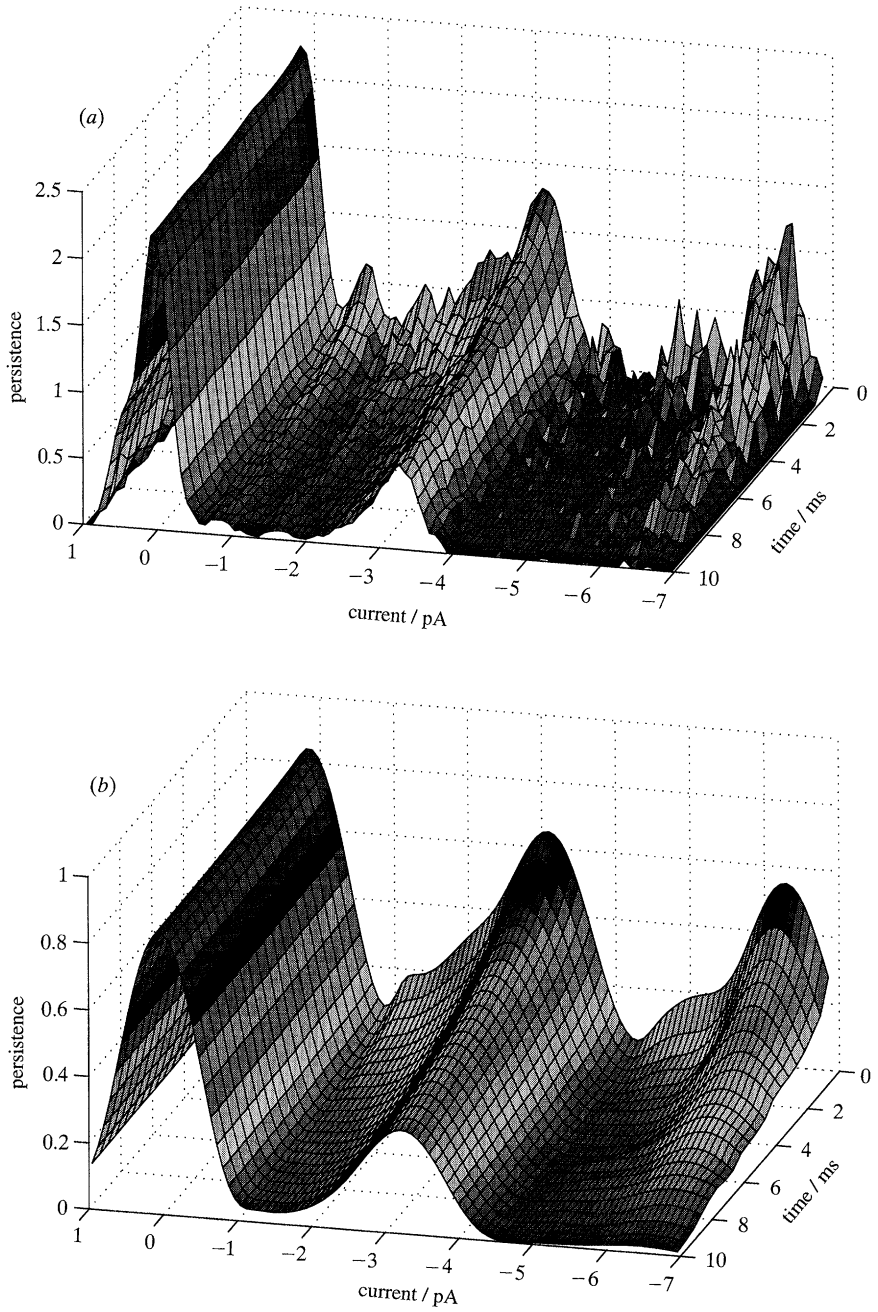

Figure 8. (*a*) Persistence function as in figure 4 with narrow window, $\tau = 0.1$ pA. (*b*) Persistence function as in figure 4 with wide window, $\tau = 1.0$ pA.

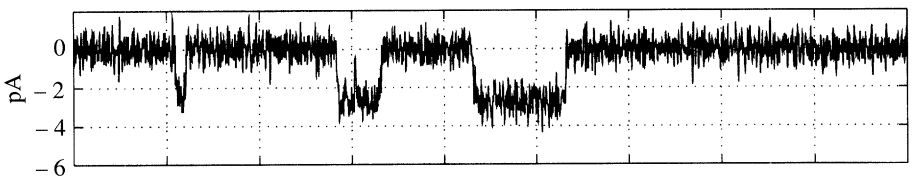

Figure 9. A 45 ms segment of lightly filtered data from an NMDA receptor. These data were filtered at 6.7 kHz and then sampled at 100 kHz, yielding 2×10^6 points.

and $\rho_{10}^{1}(t)$ from equations 9 and 10. Finally, we can use equation 11 to determine the single-channel open-state persistence function $\rho_{1}^{1}(t)$. We do not have to rely on observations of double events to determine the single-channel persistence functions.

3. ESTIMATION OF THE PERSISTENCE FUNCTION OF A NOISY RECORD

In this section we discuss the nature of the persistence function of a noisy record and how to estimate it.

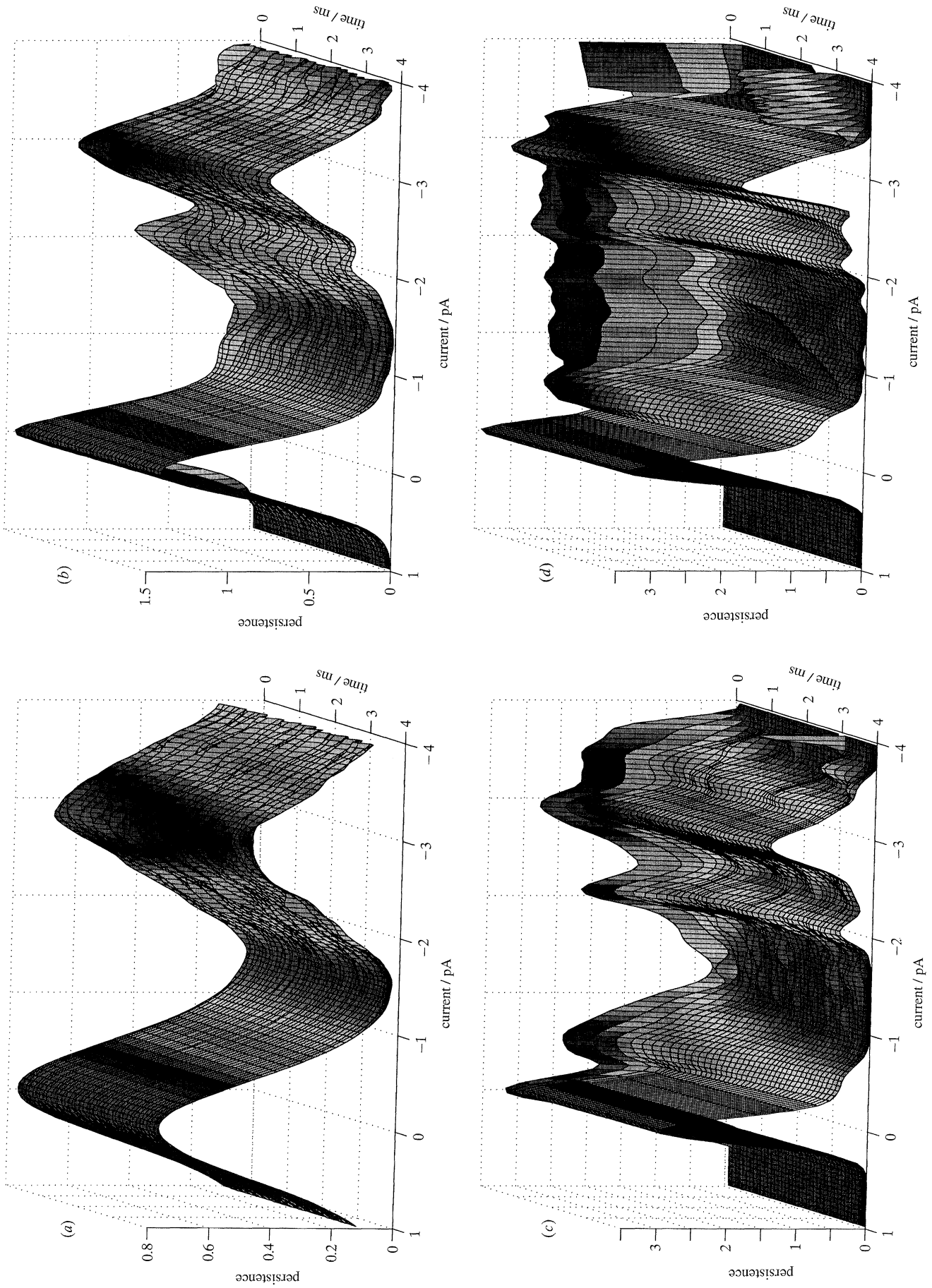

Figure 10. For description see opposite.

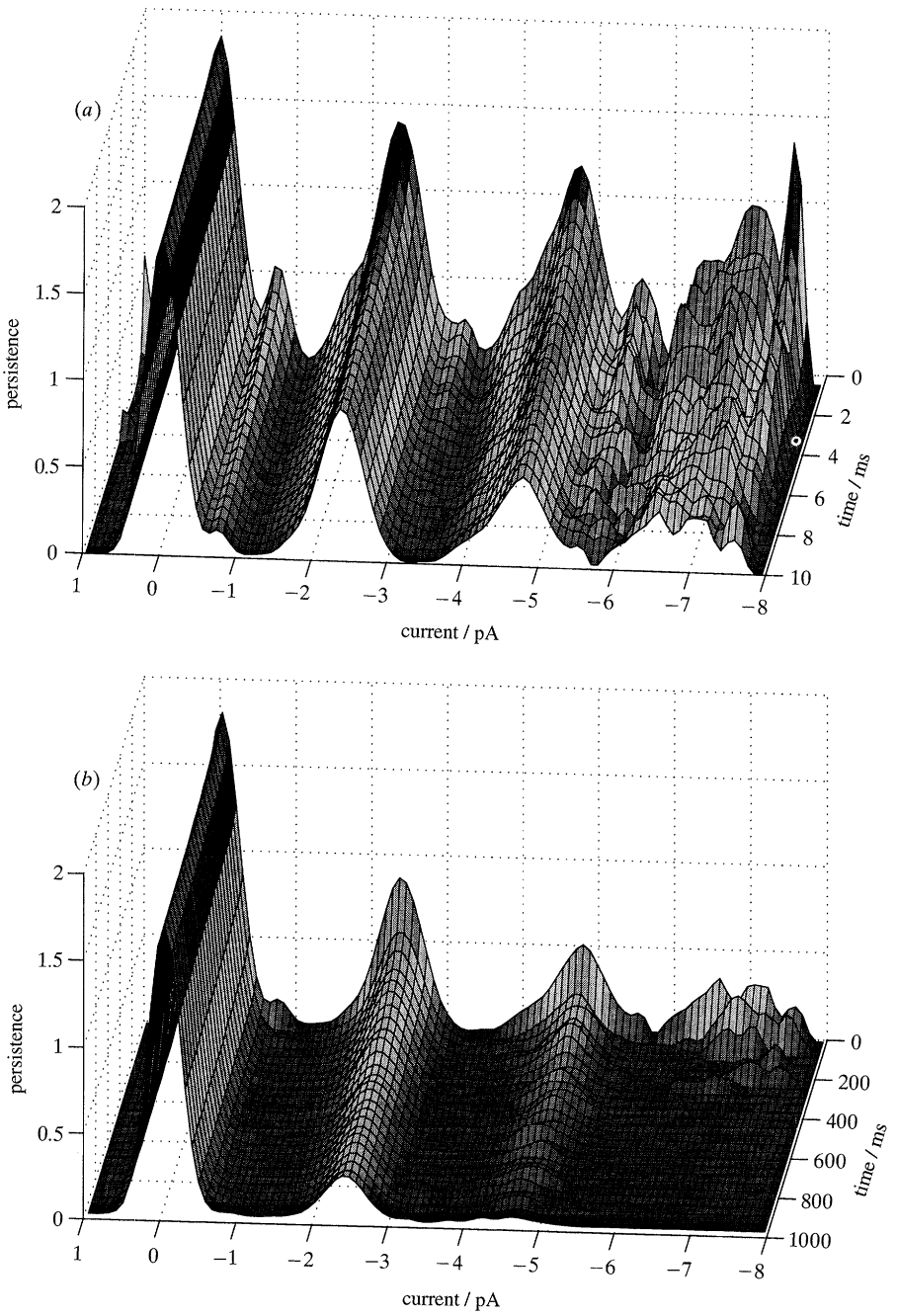

Figure 11. (*a*) Persistence function for a 140 s record of an NMDA receptor at −60 mV in 50 nM glutamate and 1 μM glycine. This record was filtered at 2 kHz and sampled at 20 kHz. Only the persistence for short lags is shown. (*b*) The same persistence function as in (*a*), but extended to long lags.

Suppose that X is the underlying quantal current process, which takes on values denoted by $I_1, I_2, \ldots, I_C$ with equilibrium probabilities π_c, and that one measures $Y(t) = X(t) + e(t)$ where $e(t)$ represents noise. Let the joint probability density of $Y(u)$, $Y(u+t)$ be denoted by $f_t(y_0, y_t)$ and the marginal probability density by $f_0(y)$. The persistence function of Y is then:

$$\rho^{(Y)}(y, t) = (f_t(y, y))/(f_0(y)).$$

We wish to examine the relation of this to the

Figure 10. (*a*) Persistence function for lightly filtered data from an NMDA receptor. Only a current peak near −3 pA is visible. (*b*) Persistence function for the same data as in (*a*), after application of a digital Bessel filter with bandwidth 5 kHz. A second current peak near −2 pA has appeared. (*c*) Persistence function for the same data as in (*a*), after application of a digital Bessel filter with bandwidth 1 kHz. Additional structure has appeared. (*d*) Persistence function for the same data as in (*a*), after application of a digital Bessel filter with bandwidth 100 Hz. Note how much detail has survived this extreme filtering.

persistence function of X. Denote the cross persistence function of X by $\rho^{(X)}_{cc'}(t)$, let the conditional probability density of $Y(t)$ given $X(t) = I_c$ be denoted by $\phi_c(y)$, and suppose that t is large enough so that $e(s)$ and $e(s+t)$ are independent. The denominator in the expression above is:

$$f_0(y) = \sum_c \pi_c \phi_c(y).$$

The numerator is:

$$\begin{aligned} f_t(y,y) &= \sum_c \sum_{c'} \phi_c(y)\, \phi_{c'}(y)\, P(X(0) \\ &\quad = I_c \,\&\, X(t) = I_{c'}) \\ &= \sum_c \sum_{c'} \phi_c(y)\, \phi_{c'}(y)\, \rho^{(X)}_{cc'}(t)\, \pi_c. \end{aligned}$$

We thus have:

$$\rho^{(Y)}(y,t) = \frac{\Sigma_c \Sigma_{c'} \phi_c(y)\, \phi_{c'}(y)\, \rho^{(X)}_{cc'}(t)\, \pi_c}{\Sigma_c \pi_c \phi_c(y)}.$$

If the noise level is small relative to the separations between the current levels, then $\phi_c(y)$ is small except when $y \approx I_c$. Thus if $y \approx I_c$,

$$\rho^{(Y)}(y,t) \approx \phi_c(y)\, \rho^{(X)}_c(t)$$

that is, the persistence function of Y is proportional to that of X. In summary, we have seen that if the noise level is small relative to the separations between the current levels and if t is larger than the correlation time of the noise, the persistence function of Y is proportional to that of X.

We now discuss a method for estimating the persistence function of Y from a recording. We do this by estimating the marginal density $f_0(y)$ and the joint densities $f_t(y,y)$ defined above. The simplest estimate of the marginal density is a histogram. The bin widths of the histogram would have to be chosen large enough so that the estimate was not too ragged and small enough so that the shape of the density was not too smeared out and obscured. The resulting estimate would be a step function rather than a continuous curve. There are a variety of ways to improve upon the histogram estimate (Silverman 1986; Scott 1992) and we choose a simple method for computational convenience and speed. The basic idea of the method is to first make a histogram with very small bin width and then smooth it by a weighted moving average. The same idea is used to estimate the joint density.

The marginal density is thus estimated in the following way. Let the observations be denoted by Y_1, $Y_2, \ldots, Y_T$. To form a histogram with n bins over the range $[a, b)$ and bin width $\Delta = (b-a)/n$, an array of counts $c(1), c(2), \ldots, c(n)$ is initialized to 0. Then for $t = 1, 2, \ldots, T$, let $j = \lfloor 1 + (Y_t - a)/\Delta \rfloor$ ($\lfloor u \rfloor$ denotes the greatest integer less than or equal to u) and increment $c(j)$ by one. This process is very fast. The histogram estimate of the probability density of Y for $a + (j-1)\Delta \leqslant y < a + j\Delta$ is $h(y) = c(j)/(T \times \Delta)$. As the bin width Δ is small, the histogram is very ragged and we next smooth it in the following way. Let $w(u)$ be a smooth non-negative weight function with support $[-1, 1]$; the precise choice does not make a great deal of difference – we have used $w(u) = (1-u^2)^2$. The weight function w is used to smooth h by first choosing a smoothing parameter or bandwidth τ and forming $w_\tau(u) = \tau^{-1} w(u/\tau)$. The weight function w_τ is thus a dilation of w with support $[-\tau, \tau]$. Let $y_j = a + j\Delta - \Delta/2$, the centre of the jth bin. Then for $a + \tau \leqslant y \leqslant b - \tau$, the smoothed estimate of $f_0(y)$ is:

$$\hat{f_0}(y) = [\Sigma_{j=1}^n w_\tau(y - y_j)\, h(y_j)]/[\Sigma_{j=1}^n w_\tau(y - y_j)].$$

In other words, to estimate $f_0(y)$, we form a weighted average of the histogram values in the range $[y-\tau, y+\tau]$. This provides a smoother estimate of $f_0(y)$ than that of the histogram, the smoothness being controlled by the choice of τ. If τ is too small, the estimate is too ragged and if τ is too large it is too blurred.

The joint density $f_t(y,y)$ is estimated analogously. For given t, a joint histogram h_2 of the observations $(Y_1, Y_{t+1}), (Y_2, Y_{t+2}), \ldots, (Y_{T-t}, Y_T)$ is formed over an $n \times n$ grid and the histogram is then smoothed yielding:

$$\hat{f_t}(y,y) = \frac{\Sigma_{j=1}^n \Sigma_{k=1}^n w_\tau(y - y_j)\, w_\tau(y - y_k)\, h_2(y_j, y_k)}{\Sigma_{j=1}^n \Sigma_{k=1}^n w_\tau(y - y_j)\, w_\tau(y - y_k)}$$

for $a + \tau \leqslant y \leqslant b - \tau$. Finally the persistence function of Y is estimated by:

$$\hat{\rho}^{(Y)}(y,t) = (\hat{f_t}(y,y))/(\hat{f_0}(y)).$$

We illustrate the effects of the choice of τ by returning to the example of §1. Figure 4, displayed there, was formed by smoothing histograms with 100 bins with a weight function with $\tau = 0.25$ pA. To illustrate the results of under and over smoothing, bandwidths of $\tau = 0.1$ pA and $\tau = 1$ pA were used to produce figures 8 (*a*) and (*b*), respectively. Figure 8 (*a*) is too ragged and details are blurred by oversmoothing in figure 8 (*b*). The computations producing the estimated persistence functions were quite fast: each took about 100 sec on a SUN SPARC 10. It is thus quite practical to choose a reasonable value of τ interactively, even for data sets of this size (1.84×10^6 points).

4. EXAMPLES

In this section we give some further examples of uses of the estimated persistence function, beginning with a discussion and illustration of the effects of filtering.

(*a*) *The Effects of Filtering*

The choice of a filter for patch-clamp recordings has been discussed extensively in the literature, see for example Colquhoun & Sigworth (1995). These discussions have generally focused on the effects of filter shape and bandwidth on the quality of the resulting approximation to the underlying quantal signal and the consequent accuracy of an idealized restoration. Here, in contrast, we wish to examine the consequences of filtering on the estimated persistence function. The effects of filter bandwidth on an estimated persistence function have been foreshadowed in the preceding section, where it was observed that if the current levels

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

of the channel were separated relative to the noise amplitude, the persistence function of the noisy record was approximately proportional to that of the noiseless current record for times greater than the correlation time of the noise. Now the effects of filtering are: (i) it decreases the noise amplitude; (ii) it increases the correlation time of the noise; and (iii) it rounds off transitions and depresses the amplitudes of short events. However, we note that it is not necessary to resolve all events in order for the estimated persistence function to resolve conductance levels. Thus increased filtering can result in sharper estimates of the persistence function at values of t that are not too small.

To illustrate these issues we examine a lightly filtered record and show how the estimated persistence function changes as the level of filtering is increased. A trace of the data is shown in figure 9. Figure 10 (*a*) shows the estimated persistence function which reveals only a current peak at about −3 pA.

The sampled data were then filtered with a digital implementation of an eight-pole Bessel filter with a 5 kHz bandwidth. Using the approximation that if two Gaussian or Bessel filters with bandwidths f_1 and f_2 are cascaded, then the bandwidth of the result, f_3, is given approximately by $f_3^{-2} = f_1^{-2} + f_2^{-2}$, the effective bandwidth of the cascaded operation was approximately 4.0 kHz. Figure 10 (*b*) shows the estimated persistence function. Filtering has resolved the broad peak around −3 pA into one at −2 pA and one at −3 pA. The figure also gives hints of shoulders at about −0.75 pA and −3.8 pA, but these are poorly resolved.

The sampled data were next filtered at 1 kHz (effective bandwidth equal to 0.99 kHz). The estimated persistence function is shown in figure 10 (*c*). Peaks at −0.5 pA, −2 pA, −3 pA are quite clear and there is also a peak at −3.5 pA which is not clearly split from the −3 pA peak. The persistence decays more slowly at −3 pA than at the other levels.

To illustrate the results of extreme filtering, the sampled data were digitally filtered at 100 Hz, resulting in the persistence function of figure 10 (*d*). There is clearly very poor resolution for times less than 1 ms, and the peak at −3.5 pA has been obscured, but it is rather surprising how much information has survived. The risetime (10–90 %) of the filter is about 3.4 ms. This is comparable with the slowest time constant of the NMDA receptor channel open time distribution which on average is about 4 ms (Gibb & Colquhoun 1992). The majority of the openings at −2 pA have a duration of less than 1 ms and thus many of the filtered openings will not reach a plateau before the channel closes. Despite this, the current amplitudes are reliably indicated.

These results illustrate that filtering can improve the detection of conductance levels via the persistence function. The persistence function contained a peak at almost exactly the same amplitude (−3 pA) in the very noisy data as was found when the data were more reasonably filtered. This suggests that the method will function reliably with smaller channels where the signal to noise ratio is not so good.

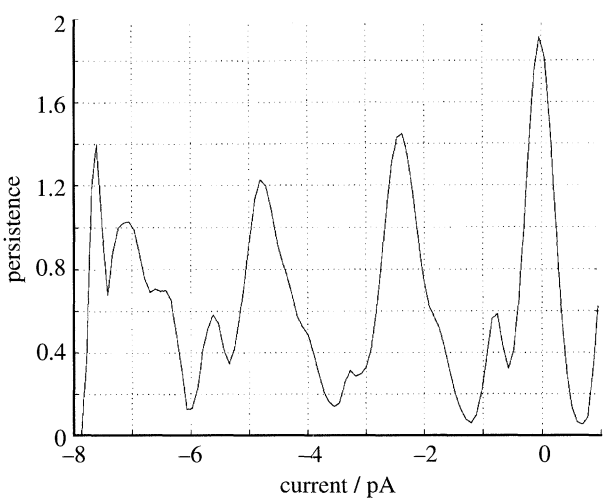

Figure 12. Persistence function from figure 11 (*a*) plotted against current for fixed lag 0.5 ms. This is a cross section of figure 11 (*a*). Plots like these are useful for the detection of current levels in a record.

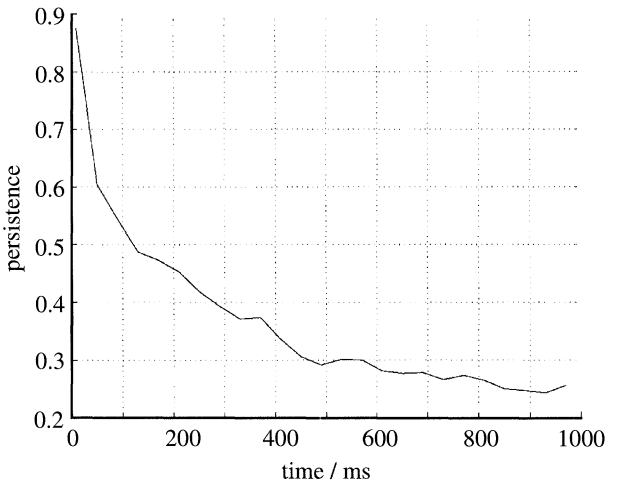

Figure 13. Persistence function from figure 11 (*b*) plotted against lag for fixed current 2.4 pA (the primary peak in figure 12). This is a cross section of figure 11 (*a*). The persistence almost extends to lag 1 s.

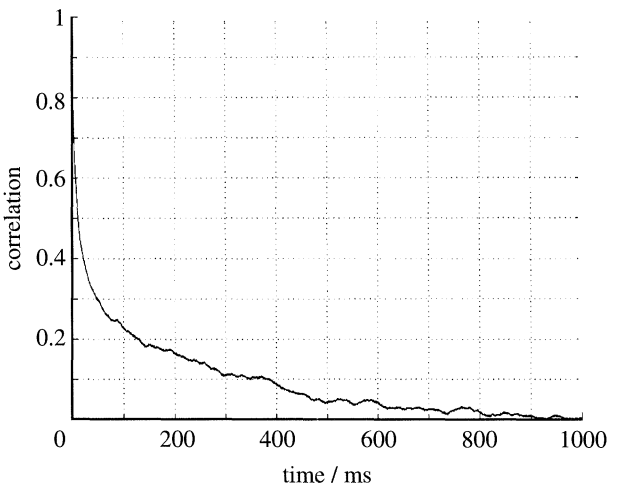

Figure 14. The autocorrelation function of the record for which the persistence function is shown in figure 11.

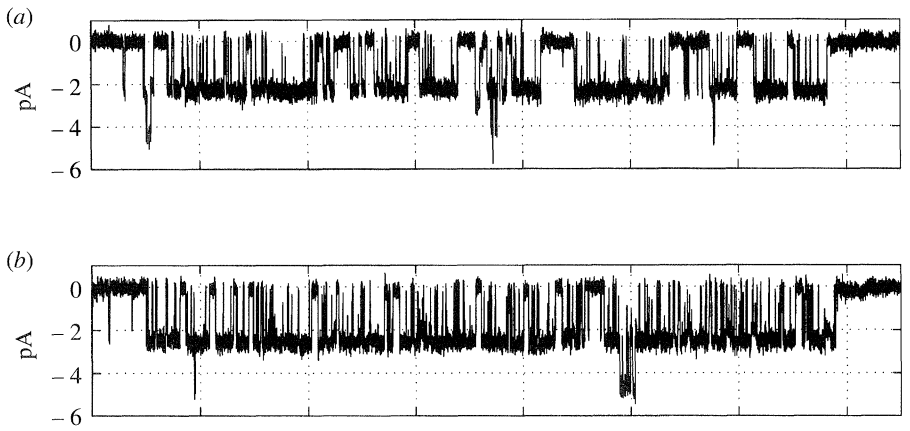

Figure 15. (*a*) A long burst at 2.4 pA which contributes to the large lag persistence shown in figure 13. Vertical grid lines are 0.2 s apart. (*b*) Another long burst at 2.4 pA which contributes to the large lag persistence shown in figure 13. Vertical grid lines are 0.2 s apart.

(*b*) *Time-Windowed Persistence Functions*

We consider a record consisting of 140 s of the activity of an NMDA receptor at −60 mV, 50 nM glutamate and 1 μM glycine. The recording was filtered at 2 kHz and sampled at 20 kHz, producing 2.8×10^6 points. Figures 11 (*a*) and (*b*) show the short and long term estimated persistence functions. The estimates were formed with 100 histogram bins smoothed with a bandwidth $\tau = 0.25$ pA. Figure 12 shows the persistence of 0.5 ms exhibiting clear peaks at −0.8 pA and −2.4 pA, current levels which have been previously noted in analyses of the NMDA receptor (Gibb & Colquhoun 1992). (The rise near 1 pA is due to several brief excursions to +1 pA and above; their brevity is consistent with figure 11 (*a*).) If there were two channels simultaneously active, additional levels at −1.6 pA, −3.2 pA and −4.8 pA would be expected. The latter two are clearly indicated in the figure. The absence of the level −1.6 pA may be due to the brevity of excursions at the level −0.8 pA; the shoulder on the right hand side of the −2.4 pA peak may reflect the −1.6 pA level. The superposition of three channels would give levels −2.4 pA, −4.0 pA, −5.6 pA and −7.2 pA; some of these are clearly present, but the resolution is poor as there were rarely three channels active simultaneously. From figure 11 (*b*) we see that the current level at 2.4 pA has the most slowly decaying persistence. Figure 13 shows that the persistence at this level only begins to level off at about 1 s, reflecting a very slow kinetic process. The decay of the autocorrelation function, figure 14, reflects this slow process but does not directly indicate the presence of faster processes shown in figure 11 (*a*).

This slow process corresponds to several long activations in the record. Figures 15 (*a*) and (*b*) show two such activations, each lasting slightly longer than one second. Figures 16 (*a*) and (*b*) show the persistence functions calculated from the recording during just those respective periods. The level at −0.8 pA is quite clear in figure 16 (*a*) but is absent in figure 16 (*b*). A level at −2 pA, just hinted at in the persistence functions of the entire record, is clear in figure 16 (*b*) but is absent in figure 16 (*a*). The current levels during these two long activations were thus different. Figure 17 shows the persistence function for a third activation, which lasted about 1.5 s, during which both of these intermediate levels were present.

The examples of the previous paragraph illustrate that the persistence function is useful not only as a summary of an entire record, but also as a description of activity during selected time segments. This method provides an alternative to the simple amplitude stability plot (Colquhoun & Sigworth 1995). A profile of dominant current levels through time can be constructed by partitioning the record into time blocks and estimating the persistence function during each block. Figure 18 was constructed from 1 s blocks: for each block the persistence function was estimated at a lag of 0.5 ms, yielding 140 curves like figure 12, which can be lined up to form a surface. The figure shows the contours of this surface and provides a visualization of the experimental record. The constancy of the peak at 0 pA reflects a stable baseline. The periods of time during which the level at −0.8 pA was active stand out quite clearly as do those time periods when more than one channel was active. The figure suggests some variability in current levels and directs our attention to periods of time in which the activity appears to be somewhat unusual. For example, the figure suggests that the period around 120 s may be interesting, and figure 19 shows a current trace during this time in which a current level at about −1.6 pA appears. Examination of the transitions to this level suggests that it is not due to two channels both with −0.8 pA of current. The figure also illustrates variability in the current levels around −2.4 pA, which may be due to baseline variability, but quite possibly is also due to variability in channel conductance.

(*c*) *Comparison of Persistence Functions*

To illustrate the comparative use of persistence functions we consider a sequence of recordings from the same patch with 30 nM glutamate and 1 μM glycine at

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

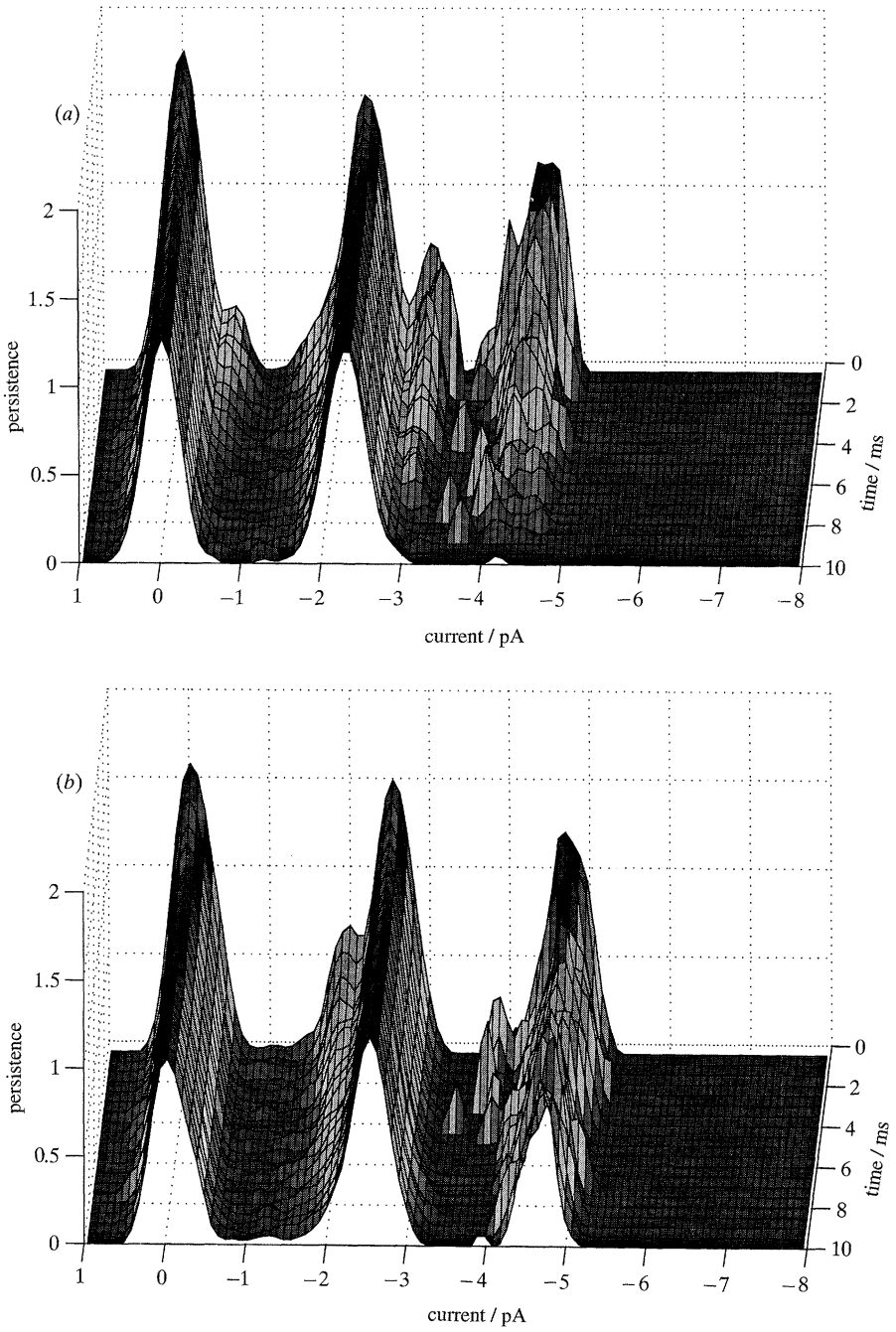

Figure 16. (*a*) Persistence function calculated from the fragment shown in figure 15 (*a*). The level at −0.8 pA is clearly seen, but the level at −2 pA is missing. (*b*) Persistence function calculated from the fragment shown in figure 15 (*b*). The level at −0.8 pA is missing, but the level at −2 pA is seen.

potentials of −40, −60, −80, and −100 mV. These recordings were made with extracellular divalent cation concentration buffered to very low levels. This results in a single-channel conductance of around 70 pS for the NMDA receptor (Gibb & Colquhoun 1992). The persistence functions (not shown) revealed major peaks at −2.3, −3.7, −5.2 and −6.9 pA. At greater hyperpolarization the signal to noise ratio increased and subconductance levels near the dominant levels became visible. The small subconductance levels shown in earlier figures in this paper did not appear. Figure 20 is a plot of the dominant current level versus voltage, indicating an ohmic relation. The slope of this line is 76.5 pS: almost exactly the same value as estimated from manually performed cursor measurements of individual single-channel openings at each membrane potential.

As well as comparing the current levels it is interesting to compare the decays of the persistence functions at varying voltage. The time decays of the ridges of the persistence functions at the current levels indicated above are compared in figure 21. The figure does not reveal any dependence of kinetics, as measured via the persistence function, on voltage. This is consistent with the known voltage independence of the NMDA channels in the absence of magnesium (Gibb & Colquhoun 1992).

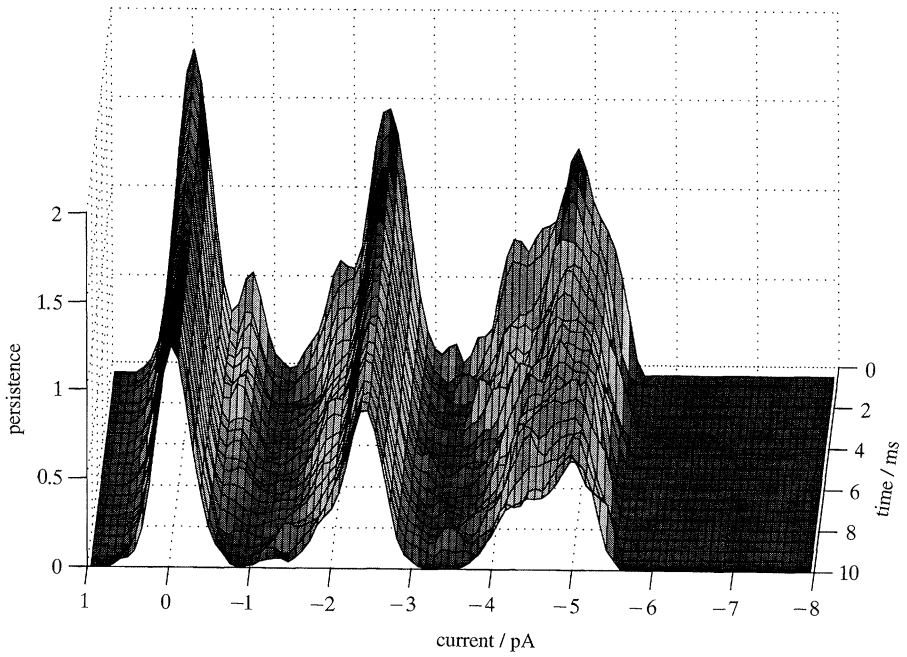

Figure 17. Persistence function calculated from a 1.5 s activation showing the minor levels at −0.8 pA and at −2 pA.

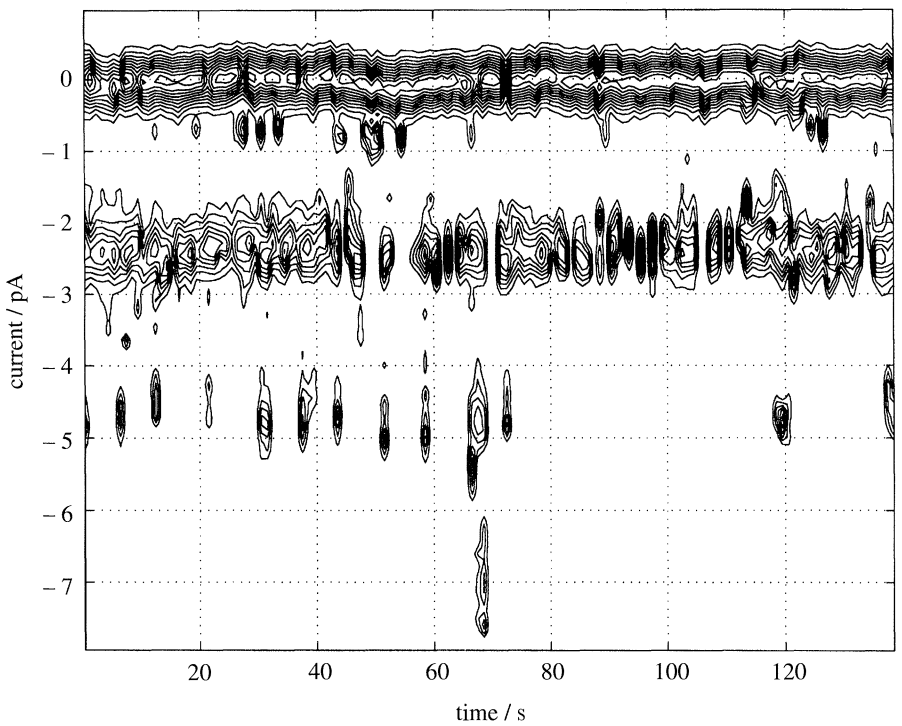

Figure 18. The persistence function was computed with a moving 1 s window and fixed lag 0.5 ms. A contour plot of the resulting function of window position in the record and of current is shown.

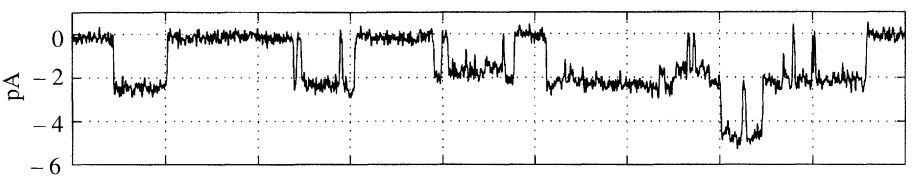

Figure 19. A record segment selected by reference to figure 18 to show activity at current −1.6 pA. Vertical grid lines are 20 ms apart.

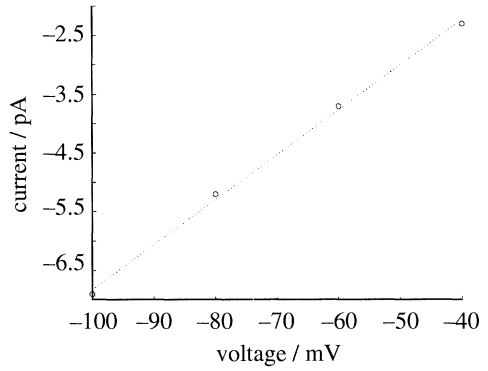

Figure 20. Current versus voltage, with current identified as the dominant peak in the persistence function at lag 5 ms. Ohm's law appears to hold.

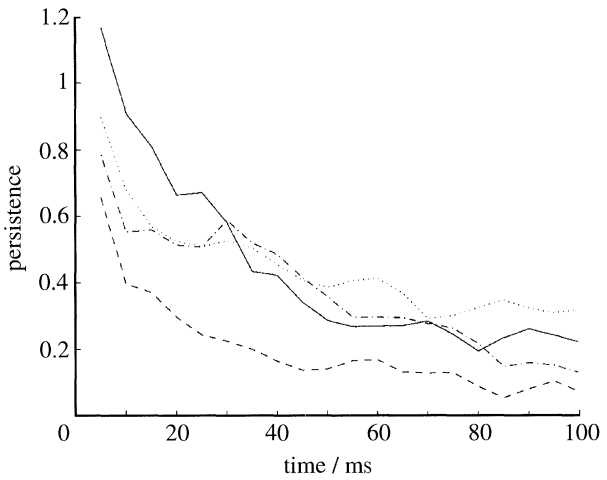

Figure 21. Time decay of the dominant peak in the persistence function for several voltages. Solid lines −40 mV, dashed line denotes −60 mV, dotted line denotes −80 mV, dot-dashed line denotes −100 mV.

5. CONCLUDING REMARKS

The vast amount of data produced by a patch-clamp recording creates a need for effective summary statistics. Such summaries are a necessary preliminary step in any analysis. The persistence function is straightforward to compute and conveys information about the conductance levels and the kinetics of a recording even when designation of discrete conductance levels is problematical. Like the all-points histogram, the persistence function does not require that amplitudes of openings be fitted individually (which takes time and risks operator bias), but as well as resolving amplitudes more sharply than does the all-points histogram, the persistence function simultaneously conveys kinetic information. It does not depend on a model, such as a Markov process, for the kinetics of the process although it can be interpreted within the context of a given model. We have shown that it is robust to filtering and hence to bandwidth limitation of the instrumentation. In this regard it can be a valuable supplement to histograms of dwell times, which can be quite sensitive to the quality of a restoration and to flickering. The persistence function is, however, sensitive to baseline drift, so trends should be removed from the recording before its estimation.

The necessary computations, although fairly extensive, are reasonably fast. We carried them out in an interactive mode using Matlab† as a computational and graphical platform, driving a specially written program in C.

D. R. F. and J. A. R. are supported by grants from the United States National Science Foundation and the Office of Naval Research. A. J. G. and D. C. are supported by the British Medical Research Council and the Wellcome Trust. We thank Bret Larget for programming the persistence function estimation.

REFERENCES

Colquhoun, D. & Sigworth, F. J. 1995 Fitting and statistical analysis of single-channel records. In *Single-channel recording* (ed. B. Sakmann & E. Neher), edn 2, pp. 483–587. Plenum.

Cox, D. R. 1962 *Renewal theory*, chapter 7. Methuen.

Fredkin, D. R. & Rice, J. A. 1991 On the superposition of currents from ion channels. *Phil. Trans. R. Soc. Lond.* B **334**, 347–356.

Gibb, A. J. & Colquhoun, D. 1992 Activation of N-methyl-D-aspartate receptors by L-glutamate in cells dissociated from adult rat hippocampus. *J. Physiol., Lond.* **456**, 143–179.

Patlak, J. 1988 Sodium channel subconductance levels measured with a new variance-mean analysis. *J. gen. Physiol.* **92**, 413–430.

Scott, D. W. 1992 *Multivariate density estimation: theory, practice, and visualization*. Wiley.

Silverman, B. W. 1986 *Density estimation for statistics and data analysis*. Chapman and Hall.

† The Math Works, Inc. The C program can also be used as a standalone routine. This program is publically available; email requests can be sent to drfredkin@ucsd.edu or rice@stat.berkeley.edn.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES