Alternate methods of representing single-channel data

S. V. Ramanan and P. R. Brink
Department of Anatomical Sciences, Health Sciences Center, State University of New York, Stony Brook, New York 11794

ABSTRACT Clustering of dwell times in data from single-channel recordings, which is in excess of the value predicted from the probability density function (pdf) alone, provides restrictions on modeling schemes. Two methods, (a) the probability density function of the running median for groups of any size of sequential dwell times, and (b) the distribution of cumulative probabilities associated with dwell times separated by any lag, or the second cumulative probability distribution, are proposed as alternative representations of single-channel data; these methods are suitable for the detection of such clusters or modes. Simulation of three models with and without modes is done to test the efficacy of these methods. It is found that they often yield a better estimate of moding parameters than the methods of running mean pdf and autocorrelation.

INTRODUCTION

A number of representations of single-channel data have been employed in an attempt to determine the mechanisms underlying channel behavior. The most frequently used methods include (*a*) binned histograms of current values to extract conductances, (*b*) probability density functions (pdf's) of open and closed dwell times (1), (*c*) autocorrelation functions (acf's) of sequences of dwell times (2, 3), and (*d*) moving means of dwell times (4).

In the modeling of the dynamics of channels by Markov chains, fitting the pdf of dwell times by exponential yields important information about the number of closed and open states. The number of schemes consistent with the pdf's alone, however, increases rapidly with the number of kinetic states. Thus any additional information which can be extracted from the data would be significant if it limited the large number of possible schemes. The last two methods are useful in this context as they provide information about the so-called "mode shifts" (4). Detection of these mode shifts is equivalent to knowledge about the connectivity (2) between different subsets of kinetic states, where gating is allowed in each subset.

To illustrate the effects of mode shifting, we consider the following example. Assume a channel characterized by the following behavior: Long closures interspersed with sporadic long openings (type I mode) are juxtaposed with extremely fast flickering between the closed and open levels (type II). This may be interpreted in the following fashion: The channel takes on two conformations corresponding to the behaviors labeled types I and II. Gating takes place in each of these conformations but at different rates. If the switching between these conformations is slow enough to permit several gating actions in each conformation, then the presence of modes is apparent visually in the raw data, with closed or open episodes of short (long) durations being much more likely to be preceded and followed by other short (long) episodes. The moving mean would change suddenly when it entered any of the modes and was restored when it left the mode (see reference 4 for illustrations of this). In addition, the correlation between the duration of the ith closed episode and that of the $(i + j)$th episode would be nonzero till the lag j is of the order of the average number of closed episodes in either of the modes.

It may be noted that in a Markov model with two closed states as above, there would exist a probability for closed states of short durations, say, to be found adjacent to other episodes of short duration even in the absence of modes. The signature of the presence of modes is a probability for abutment of short episodes, for instance, which exceeds the probability for such juxtapositions when the length of every dwell time is determined from the pdf alone. It is this property which we exploit for a quantitative description of modes.

If the rate at which the channel changes conformations becomes comparable to the rate at which gating occurs in any of the conformations, the ready visual distinction between modes which is observed in the case above is lost. It is in these circumstances that the utility of the autocorrelation and mean becomes apparent, as they are still able to provide a quantitative measure of presence of modes. Examples of such behavior are given in McManus and Magleby (4). In this paper, we present two methods which complement the enumerated list, and as we show by simulation, supersede them sometimes in detecting

0006-3495/90/04/893/09 $2.00

mode shifts. As stated above, sensitivity in detection of the existence of modes enables the reduction of the number of kinetic schemes consistent with the data.

THEORY

Median

The assumption of Markov processes in modeling channel kinetics and consequent use of exponentials in modeling dwell time pdf's is in some sense equivalent to the observation that these pdf's have long "tails" (5). Indeed the exaggerated tails often found in the data have led several authors (6) to prefer a power law falloff at large times as opposed to the sum of several exponentials, the number of which may sometimes exceed eight (4). Regardless of this, it is known that the presence of a tail with negligible area, but having a large first moment, makes the mean a poor estimator. The median is better than the mean for such distributions (reference 5, p. 459), as it fails only when the area of the tails is not negligible. This has consequences when the switching between conformations is relatively rapid in the sense described in the introduction. The moving mean cannot distinguish between the long tail of one distribution and the average of another, and tends to smear the two together. The median, being relatively immune to the appearance of infrequent tails, is better able to distinguish any occurrence of mode shifting from the occasional episode of long duration which may happen within any particular mode.

Distributions of cumulative probability

The data from single channels consists of a sequence of closed (and open) times. To each of these closed (or open) times we may assign a unique cumulative probability dictated by the distribution of the whole sequence. We may also look at the distribution of the sum of the probabilities associated with the ith closed event and the $(i + j)$th closed event. In the absence of modes, the distribution of this probability may be evaluated theoretically; this distribution turns out to be the same, irrespective of the number of closed states present. Any deviation of the distribution found from the actual sequence of events from this modeless distribution enables detection of moding present in the data.

Let t be a variable, for instance the closed dwell time, with pdf $F(t)$. Define a map $t \leftrightarrow y$ defined by $y = g(t) = \int_0^t F(t')\,dt'$. This is an invertible and monotonically increasing function of t mapping $[0, \infty) \leftrightarrow [0, 1)$. The pdf of y is then given by the first cumulative probability distribution (or *first cpd*)

$$G(y) = \left|\frac{dt}{dy}\right| F(t) = u(y) - u(y-1),$$

where $u(y)$ is the step function. $g(t)$ is merely the cumulative probability at t.

Given a sequence $t_0, t_1, \ldots, t_i, \ldots$ we can find a unique sequence $y_0, y_1, \ldots, y_i, \ldots$ by the mapping above. We look at the second cumulative probability distribution (or *second cpd*) of $\bar{y} = (y_i + y_{i+j})$ for fixed lag j. If the distribution of y_i is completely independent from that of y_{i+j} the sum $(y_i + y_{i+j})$ will have the second cpd

$$G_2^j(\bar{y}) = \int_0^{\bar{y}} dy\, G(y)G(\bar{y} - y),$$

this being the convolution $G \otimes G$. This can be evaluated to yield

$$\begin{aligned} G_2^j(\bar{y}) &= \bar{y} \qquad 0 \le \bar{y} \le 1. \\ G_2^j(\bar{y}) &= 2 - \bar{y} \qquad 1 \le \bar{y} < 2. \end{aligned} \tag{1}$$

This is the isosceles triangle over [0, 2) with unit area, which we notate by $\Delta(\bar{y})$. Deviations of $G_2^j(\bar{y})$ from this Δ function for any $\bar{y}$ and any lag j would indicate the presence of correlations or mode shifting in the data.

As an example, consider the situation when the pdf $F(t)$ can be decomposed into two pdf's as $F(t) = h_1F_1(t) + h_2F_2(t)$, where $F_1(t) = 0, t \ge \bar{t}$; $F_2(t) = 0, t \le \bar{t}$; and $h_1 + h_2 = 1$. This would then mimic the case when the dwell time pdf is the sum of two exponentials whose time constants are separated by several orders of magnitude, as in the example described in the introduction. Assume, further, that the sequence $t_0, t_1, \ldots, t_i, \ldots$ can be represented as below:

$$\underbrace{t_1, \ldots\ldots, t_{Mh_1}}_{\text{distribution 1}}, \underbrace{t_{Mh_1+1}, \ldots\ldots, t_M}_{\text{distribution 2}}, t_{M+1}, \ldots, t_{M+Mh_1}, \ldots$$

In the limit $M \to \infty$, we get the ideal separation of modes. The distribution of $\bar{y} = (y_i + y_{i+j})$ is then given by the sum of two nonoverlapping triangles:

$$G_2^j(\bar{y}) = h_1\Delta\left(\frac{\bar{y}}{h_1}\right) + h_2\Delta\left(\frac{\bar{y} - 2h_1}{h_2}\right), \text{ all } j.$$

The mode shifting would thus express itself as two distinct peaks in the distribution.

It is possible to write expressions for the second cpd G_2^j in the Markov formalism. In what follows we shall use the notation of Colquhoun and Hawkes (3). Thus the transition matrix $\mathbb{Q}$ between states is partitioned into $\mathbb{Q}_{AA}$, $\mathbb{Q}_{AF}$, etc., where subscripts A, F refer to open and closed states, respectively. Auxiliary matrices $\mathbb{G}_{AF} = -\mathbb{Q}_{AA}^{-1}\,\mathbb{Q}_{AF}$ and $\mathbb{X}_{AA} = \mathbb{G}_{AF}\,\mathbb{G}_{FA}$ may also be defined. The pdf of open dwell times is then $F(t) = \phi_0 e^{\mathbb{Q}_{AA}t}(-\mathbb{Q}_{AA})u_A$, u_A being a column

vector of ones and ϕ_0 the row vector of open state equilibrium probabilities. Defining the cumulative probability map $y = g(t)$ in the usual way, it follows directly from Eq. 2.2 of reference 3 that

$$G_2^j(\bar{y}) = \int_0^{\bar{y}} dy \frac{\phi_0 \mathbb{R}_{AA}(\bar{y} - y)\mathbb{X}_{AA}^j \mathbb{R}_{AA}(y)u_A}{(\phi_0 \mathbb{R}_{AA}(\bar{y} - y)u_A)(\phi_0 \mathbb{R}_{AA}(y)u_A)}, \quad j \geq 1 \quad (2)$$

where $\mathbb{R}_{AA}(y) = e^{\mathbb{Q}_{AA}g^{-1}(y)}(-\mathbb{Q}_{AA})$. Use of the spectral decomposition (Reference 3, p. 18),

$$\mathbb{X}_{AA}^j = u_A\phi_0 + \lambda_2^j \mathbb{A}_2 + \lambda_3^j \mathbb{A}_3 + \cdots, \quad (3)$$

enables us to rewrite the above equation in the form,

$$G_2^j(\bar{y}) = \Delta(\bar{y}) + (\lambda_2)^j H_2(\bar{y}) + (\lambda_3)^j H_3(\bar{y}) + \cdots, \quad (4)$$

where $H_r(\bar{y})$ is given by the RHS of Eq. 2 with $\mathbb{X}_{AA}^j$ replaced by $\mathbb{A}_r$. Inspection of this equation shows that deviations of G_2^j from the Δ function decrease exponentially with lag j but *uniformly* in $\bar{y}$. The time constants for this decay, given by $-1/\log(\lambda)$, are precisely those of the autocorrelation. The existence of nonzero λ's is an indication of moding behavior, because if all the "mode coefficient" λ's are zero, the second cpd reduces to the Δ function which is characteristic of no moding.

COMPUTATION

As identical procedures may be employed for analysis of sequences of closed and open dwell times, we restrict our attention in what follows to closed dwell times alone.

Pdf of running median

Given a sequence of N closed times and the number A of closed dwell times that we wish to lump in a group, the running median for each of the $N + 1 - A$ such groups may be computed. We distinguish here between the running and moving median: The moving median is for successive exclusive groups of A dwell times, there being N/A such groups; the running median is for all groups of A sequential times, which number $N + 1 - A$. We plot the pdf of the running medians of these A-groups according to the representation of Sigworth and Sine (7). Note that the pdf of A-groups with $A = 1$, i.e., 1-groups, is the dwell time pdf itself.

For determining the existence of modes, we need to find the pdf of the median of A-groups in a sequence where no modes exist, i.e., where the length of all dwell times is fixed by the dwell time pdf alone. This is done by randomly reshuffling the order of dwell times in the primary sequence, and computing the pdf of the running median for all of the A-groups in this reordered sequence. We repeat this process for a large numer of reshuffles and average to find the pdf of the A-group median for a modeless sequence; this sequence has the same dwell time pdf as the original sequence. Comparison of actual and modeless median pdf's then allows us to determine whether there exists a significant probability that differences between the two pdf's is attributable to some ordering present in the primary sequence due to existence of modes, or whether such differences can be explained by the random nature of the process alone.

For comparison, we may also calculate the pdf of A-groups of running mean similar to that for the running median. Note that the mean and median pdf's coincide for 1- and 2-groups, thus we need to consider at least 3-groups to find any contrast between the two.

Second cpd

The second cpd G_2 is computed by the following method: For the sequence $t_0, t_1, \ldots, t_N$ of closed times we find the associated sequence $y_0, y_1, \ldots, y_N$ of cumulative probability. Partitioning the interval [0, 2) into R disjoint parts $[z_r, z_{r+1})$, we define $G_2^j(r, N)$, $0 \leq r < R$ as the number of points y_i such that $z_r \leq (y_i + y_{i+j}) < z_{r+1}$. As there are $(N - j)$ points contributing to $G_2^j(r, N)$, we define the function $\Delta^j(r, N) = (N - j)\int_{z_r}^{z_{r+1}} d\bar{y}\, \Delta(\bar{y})$. We do not plot the second cpd itself; instead we graph the *normed deviation* of the second cpd from its base value, this being given by the expression,

$$\frac{G_2^j(r, N) - \Delta^j(r, N)}{\sqrt{\Delta^j(r, N)}} \quad (5)$$

The error bars are then fixed by assuming that the maximum deviation of this quantity from zero is 2 for 95% confidence. In plotting Eq. 5 we normally divide the interval [0, 2) into R equal parts.

For purposes of computing the mode coefficients λ, however, we make use of the fact that the decay of the quantity (Eq. 5) is uniform in $\bar{y}$. Thus for unions of any arbitrary collection of disjoint intervals $[z_r, z_{r+1})$ we may define the *normed absolute deviation:*

$$\frac{\Sigma' |G_2^j(r, N) - \Delta^j(r, N)|}{\sqrt{\Sigma' \Delta^j(r, N)}} \quad (6)$$

where the prime indicates that the sum is over only some of the intervals covering [0, 2). These intervals are chosen to maximize the quantity (Eq. 6). Illustrations of this procedure are given below.

The autocorrelation and its associated error bars are fixed by the methods described in Labarca et al. (8).

It is our claim that, with the choice of intervals which maximizes Eq. 6, the coefficients which multiply the λ's in the decay of the normed absolute deviation have greater

confidence (relative to the error bars) as compared to the corresponding coefficients for the autocorrelation. We have not been able to find any proof of this hypothesis, but it is verified in all our simulations.

METHODS

We use a random number generator and the method described in reference 5 (p. 200) to generate an exponential probability distribution. Simulation of $2N$ episodes equally distributed between closed and open states for any N is then done by the procedure described in Clay and DeFelice (9). Effects of finite deadtime are accounted for by appropriately pooling all episodes of duration shorter than this time.

An efficient search for the median is described in reference 5, p. 461. This takes 20–30 passes for 15,000 episodes and 5–10 passes for 100 episodes. Pdf's are plotted by the method of reference 7. The number of bins per decade is fixed at 25 in all the graphs.

To find the second cpd, we first need to order the sequence of N dwell times according to increasing duration. This may be done by constructing a rank table (reference 5, p. 234). The cumulative probability corresponding to the jth dwell time is then given by the formula (jth entry of rank table)/N.

When the sequence of dwell times is long, constructing a rank table becomes impossible due to constraints of memory. We then use the following procedure: The binned pdf of dwell times is first found by the methods described in reference 7. From this, the cumulative probability at the start and end of each bin is easily computed. One pass is then made through the sequence of dwell times and each dwell time is then assigned a cumulative probability given by the sum of two terms: (*a*) the cumulative probability at the beginning of the bin it is assigned to, and (*b*) a factor given by the formula, [(increment of cumulative probability over assigned bin) × (number of dwell times already encountered in this bin)]/(total number of dwell times in the bin).

The second cpd for a lag of 1, say, is then the distribution of the sum of the cumulative probabilities corresponding to adjacent dwell times. We plot this after binning, where the bins are equal divisions of the interval [0, 2). The corresponding weight of the modeless distribution in any bin is the area over this bin of the isosceles triangle over [0, 2) with total area of $N - 1$. The normed deviation in any bin is then the difference between the original cpd and the modeless cpd in that bin, scaled by the square root of the modeless value. Such scaling allows us to maintain a constant error estimate over all the bins.

We may now mark off any bins where the normed deviation differs sufficiently from the modeless zero value. The normed absolute deviation for a lag of 1 is then the sum of the absolute difference between the primary and modeless cpd for lag 1 over all marked bins, again scaled by the square root of the sum of modeless cpd over all such bins. We plot the decay of the normed absolute deviation for increasing lag for comparison with the autocorrelation.

All computations were done in QuickBasic on an IBM PC. The output of the pseudorandom generator was further randomized by the use of the algorithm of Bays and Durham as described in reference 5, p. 195.

RESULTS

The first model we use for comparison of the various discriminators of mode shifting is one of the simplest possible with two modes. The schema is represented by

$$C \underset{\alpha'}{\overset{\alpha}{\rightleftharpoons}} O \underset{\gamma'}{\overset{\gamma}{\rightleftharpoons}} O \underset{\beta'}{\overset{\beta}{\rightleftharpoons}} C \qquad (7)$$

and allows for two modes. The unique mode coefficient $\lambda = \lambda_2$, which is the same for closed and open states, may then be calculated from Eq. 3 to give $\lambda = \alpha\beta/\alpha\beta + \alpha\gamma' + \beta\gamma$. The kinetic parameters α(0.05/ms), α'(0.1/ms), β(0.02/ms), β'(0.0004/ms) are fixed in all the simulations. The parameters γ and γ' are varied to generate three different behaviors in Figs. 1–3. Values of these parameters are displayed in the legends.

Fig. 1 shows a case when both of the exponentials contributing to the dwell time pdf of the closed state have significant weight. The theoretical mode coefficient λ is 0.476. Fig. 1 *a* shows the normed deviation (Eq. 5) for a lag of 1 and a uniform partition with binwidth equal to 0.01. The presence of two modes is clearly seen, the positive deviation in the region [0, 0.6) showing that closed episodes of small durations are more likely to be found adjacent to each other than if there are no modes. The positive area over [1.3, 2) similiarly denotes that the probability of long dwell times being found abutting long dwell times is higher than it would be if mode shifting were absent. The intervals [0, 0.6), [0.6, 1.3), [1.3, 2) are therefore selected from Fig. 1 *a* to obtain the normed absolute deviation (Eq. 6) by the method detailed above. The resulting graph (Fig. 1 *b*) may be compared to the autocorrelation (Fig. 1 *c*) plotted on a scale such that the error bars appear the same. It is seen that the normed absolute deviation produces a value about twice the resolution of that from the autocorrelation when compared to the size of their respective error bars. The coefficient λ from the cpd method is 0.54 with a variance of 0.11, while λ from the autocorrelation is 0.51 with a variance of 0.18, the increased variance in the second case reflecting the larger error bars. Both values are seen to be compatible with the theoretical value within error limits.

The pdf's of running median and mean for 5-groups are displayed in Fig. 1 *d* and *e*, respectively, as the lighter of the two lines shown. The darker lines are the corresponding modeless pdf for 40 reshuffles. Note that the original median pdf differs systematically from the modeless median pdf for short and long times. Summing values of running median over all bins up to 100 ms yields 3,446 5-groups while the modeless median pdf predicts 2,526 5-groups for the same range. Assuming the distribution of 5-groups over sums of bins to be Poisson distributed, we find the difference to be $(3{,}446 - 2{,}526)/(\sqrt{3{,}446} + \sqrt{2{,}526/40})$, i.e., about 13 standard deviations (SD), with corresponding probability of 10^{-20} of such difference arising due to chance alone. For bins larger than 144 ms, we find a deviation of 11 SDs between actual and modeless median pdf's. The running mean pdf displays

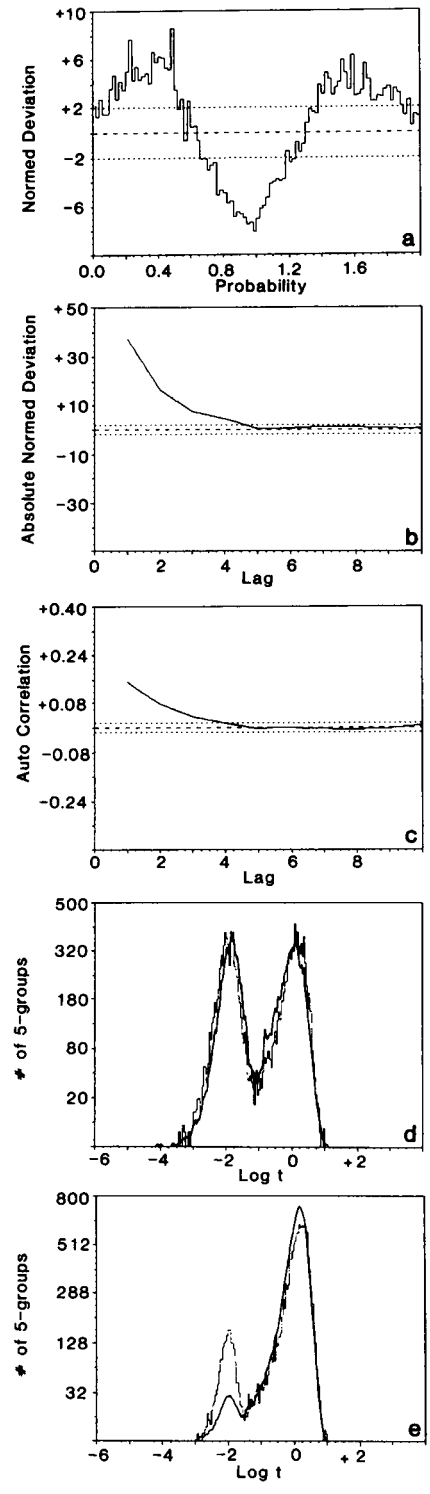

FIGURE 1 Plots of various estimators for a sequence of 15,000 episodes in the closed state of the model (Eq. 7) with $\gamma = 0.03$/ms and $\gamma' = 0.01$/ms. (*a*) Normed deviation (Eq. 5) as a function of the cumulative probability for a lag of 1. (*b*) Absolute normed deviation (Eq. 6) for lags from 1 to 10. (*c*) Autocorrelation for lags from 1 to 10. Note that the plots are made such that the error bars in *b* and *c* are the same size visually; this facilitates direct comparison. (*d*) Running median pdf of closed 5-groups as the lighter of the two lines. The darkened lines are the modeless median pdf calculated from 40 reshuffles. (*e*) Similar to *d*, but for the running mean pdf's of 5-groups. See discussion in text.

even better the effects of moding on short dwell times, differing by 30 SDs from modeless pdf's for 5-groups whose mean is <30 ms. For dwell times >30 ms we find that the running mean pdf does not show significant statistical difference from the modeless pdf.

Fig. 2 describes a situation where the system remains predominantly in one mode with brief excursions into another mode with short closed times. Fig. 2 *a* shows the normed deviation in the cumulative probability interval [0, 0.2) at a binwidth of 0.001 for adjacent episodes. The presence of the second mode can be seen in the positive area in the interval [0, 0.1), which indicates an excess of adjacent short closed times over the modeless case. We choose this interval to calculate the normed absolute deviation seen in Fig. 1 *b*. This is seen to decay rapidly, decaying to virtually zero at a lag of four. The autocorrelation (Fig. 1 *c*), however, fails to exhibit any correlation and remains well within the error bars. λ calculated from the cpd method is 0.30 with a variance of 0.14, which agrees with the theoretical value of 0.323. Examination of the median pdf for 5-groups in Fig. 2 *d* shows a persistent discrepancy between the original pdf and modeless pdf. Aggregation of the data from bins lesser than 48 ms where the difference is greatest allows an estimate of 10^{-9} for the probability of such deviations (11 SDs) occurring due to chance alone. The pdf of running medians thus confirms the prediction of the normed deviation that short episodes tend to cluster next to each other. The running mean for 5-groups shown in Fig. 1 *e*, also shows a difference between the original and modeless pdf's for short times; this corresponds to nine SDs for bins smaller than 48 ms, quite comparable to that found by the running median pdf.

The case studied in Fig. 3 may be considered the opposite of that in Fig. 2. Fig. 3 *a* is a plot of the normed deviation in the interval [1.9, 2) for a lag of one and binwidth of 0.001. The deviation is significant in the area over [1.985, 2]; thus long episodes tend to cluster more than would be expected if moding was absent. This area is therefore selected for computing the absolute normed deviation. Both this procedure (Fig. 3 *b*) and the autocorrelation (Fig. 3 *c*) show the existence of mode shifting. However, we are unable to calculate the mode coefficient λ from the autocorrelation method as it decays to within error bars at a lag of 2. The λ from the cpd method is found to be 0.17 with a variance of 0.11, which differs by 1 standard deviation from the theoretical value of 0.284.

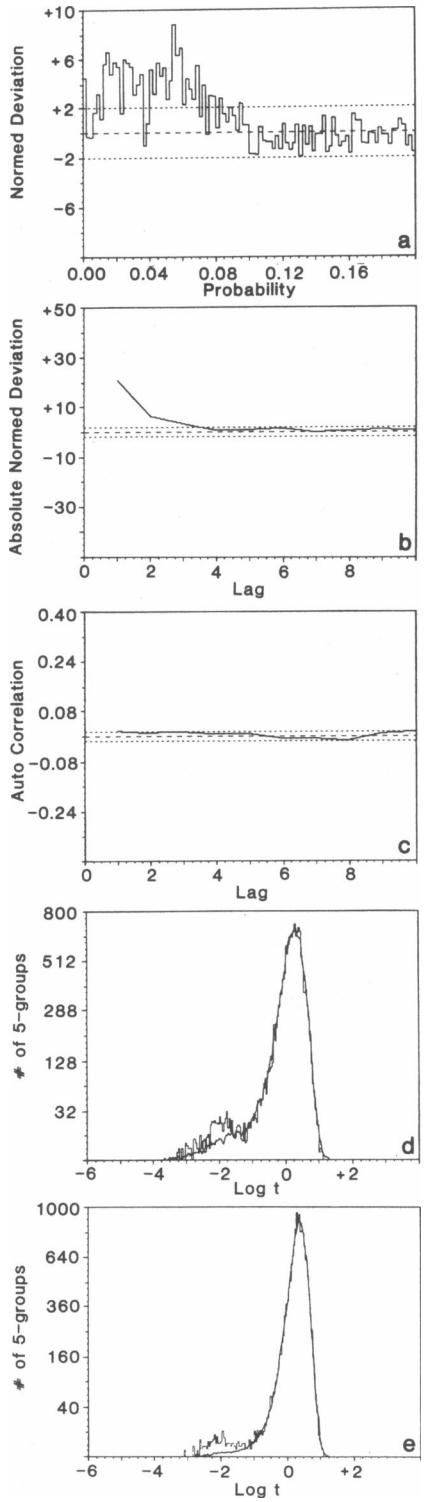

FIGURE 2 Same as Fig. 1, except that $\gamma = 0.1/\text{ms}$, $\gamma' = 0.002/\text{ms}$.

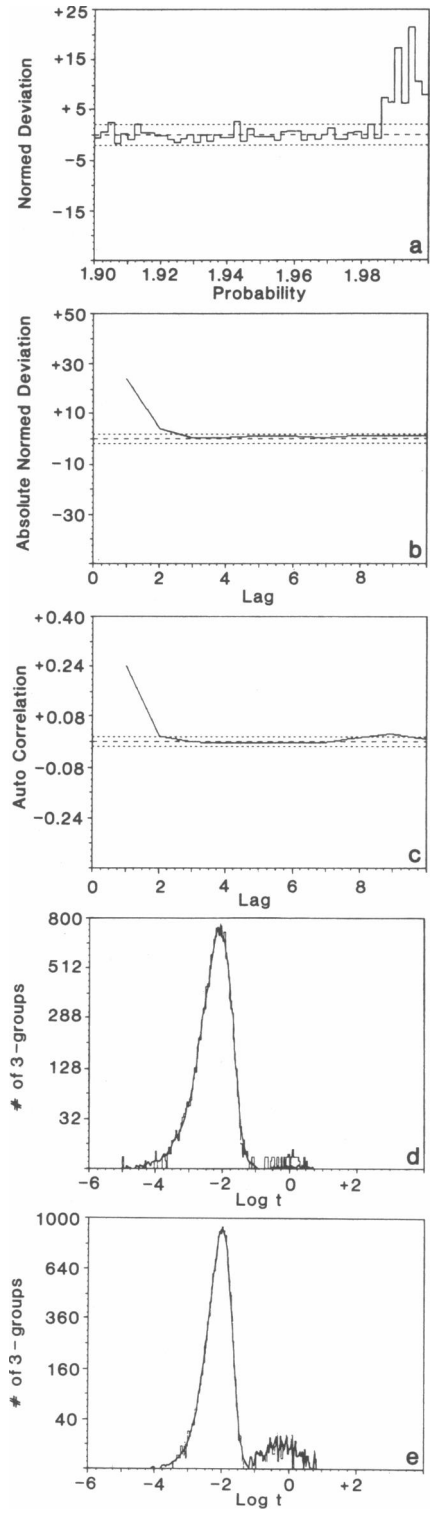

FIGURE 3 Same as Fig. 1 with $\gamma = 0.001/\text{ms}$, $\gamma' = 0.05/\text{ms}$, except that panels *d* and *e* refer to running median and mean pdf's, respectively, for 3-groups rather than 5-groups.

The 3-group median and mean pdf's are plotted in Fig. 3, *d* and *e*, respectively, along with the modeless 3-group pdf in darkened lines. The mean pdf shows no significant moding; the median pdf exhibits differences with the modeless pdf for long dwell times. Collecting all bins greater then 100 ms allows us to estimate a probability of $<10^{-4}$ for such differences (six SDs) to occur by random variations alone.

The second model we study has the kinetic scheme (10)

$$C \underset{10^4}{\overset{26}{\rightleftharpoons}} C \underset{1000}{\overset{1.9 \times 10^4}{\rightleftharpoons}} O. \qquad (8)$$

The parameters chosen (given in seconds^{-1}) are the same as in reference 10. This model allows for bursting but has only one mode and therefore can be used as a control model for testing of the methods presented. The normed deviation for 10,000 closed dwell times for this model is shown in Fig. 4. No significant deviations beyond those allowed by error bars are seen. Pdf's of both running mean and running median for A-groups with A ranging from 2 to 100, which we do not show here, do not reveal any significant deviations from the modeless pdf's.

The last model tested is one proposed (11) for Ca-activated K channels:

$$\begin{array}{ccccc} C & \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} & C & \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} & C \\ & {\scriptstyle \alpha_1' \downarrow \uparrow \beta_1'} & & {\scriptstyle \alpha_2' \downarrow \uparrow \beta_2'} & \\ & O & & O. & \end{array} \qquad (9)$$

The parameters chosen are $\alpha_1 = 34/s$, $\alpha_2 = 100/s$, $\alpha_1' = 120/s$, $\alpha_2' = 3950/s$, $\beta_1 = 180/s$, $\beta_2 = 333/s$, $\beta_1' = 2860/s$, $\beta_2' = 322/s$. All parameters are the same in reference 11, except that α_2 and β_2 have been changed to their current values from the original values of 285/s and 600/s, respectively. This is done to enhance the value of the sole closed mode coefficient λ which is given by the formula

$$\frac{\alpha_1'\alpha_2'}{\alpha_1'\alpha_2' + \alpha_2'\beta_1 + \alpha_1'\beta_2}.$$

For the parameters chosen above, this evaluates to 0.52.

The normed deviation for closed dwell times is displayed in Fig. 5 *a*. Significant deviation from zero is seen in the area over [1.75, 2), signifying clustering of long closed dwell times. This interval is chosen for evaluation of the absolute normed deviation displayed in Fig. 5 *b*. If compared to the autocorrelation in Fig. 5 *c* plotted on a

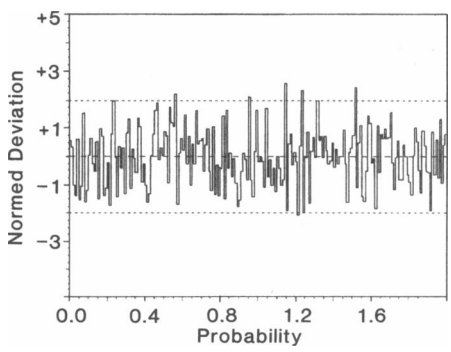

FIGURE 4 Plot of absolute normed deviation against cumulative probability for model (Eq. 8) which has bursts but no modes. Note the absence of any significant deviations outside the error bars.

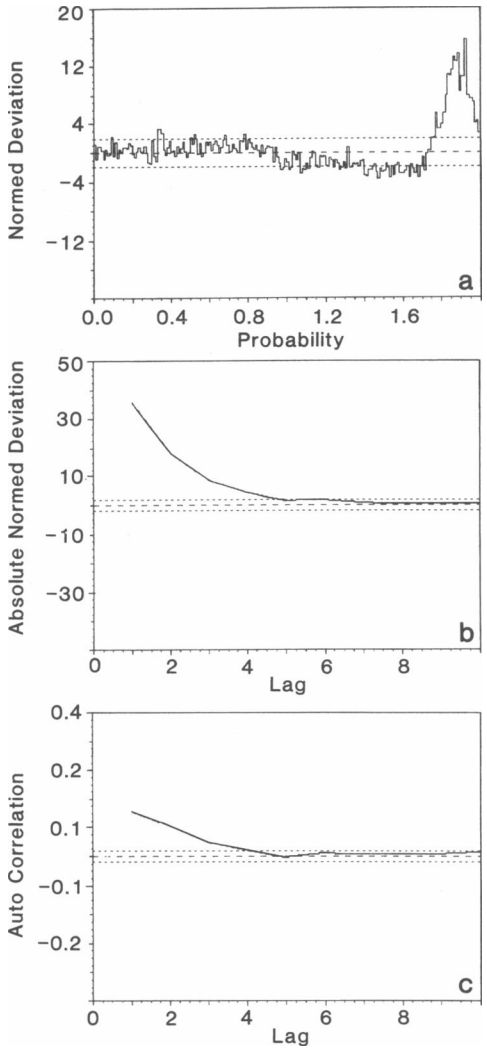

FIGURE 5 Plots for model (Eq. 9). (*a*) Normed deviation as a function of cumulative probability for a lag of 1. (*b* and *c*) Plots of absolute normed deviation and autocorrelation, respectively, for lags from 1 to 10. The error bars in *b* and *c* have been made the same size to permit direct comparison.

similar error scaling, it is seen that the deviation method produces a value about twice as significant as the autocorrelation. λ from the first method is 0.51 ± 0.09 and from the second is 0.42 ± 0.20, these quantities being both calculated from the values of the corresponding functions at lags 1 and 2.

We use this model for comparing the effects of changing sample size and finite dead time on the mode coefficients found from both the deviation and the correlation methods. Fig. 6 shows the influence of varying the sample size from 500 to 16,000 closed dwell times. On the *y*-axis we plot the absolute normed deviation and the autocorrelation for lags 1 and 2 scaled against their respective error bars, which are found as described in the methods section; such a plot provides a convenient way of estimating their relative statistical significance. The *x*-axis is the sample size. All scaled functions are zero to within or almost within error limits for small samples and seem to increase linearly with increasing sample size. However, the rate of increase for the scaled absolute deviation is seen to be at least double that for the scaled autocorrelation; note that the scaled deviation for a lag of two becomes as statistically significant as the scaled autocorrelation at a lag of one for sample sizes around 8,000. A similar comparison as in Fig. 6 but for varying dead times is shown in Fig. 7. All functions decay with increasing dead time. It may again be noted that the cpd method produces better estimates than the correlation for all dead times considered.

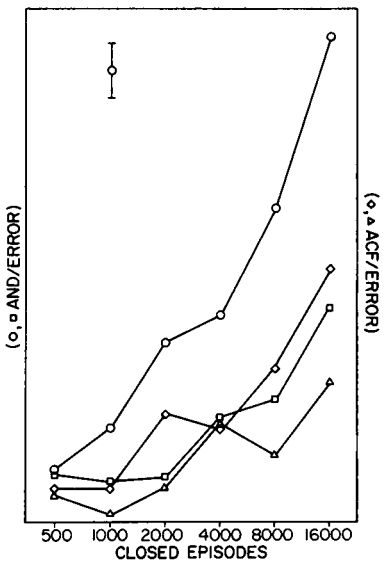

FIGURE 6 The *y*-axis on the left-hand side of the figure is the absolute normed deviation (AND) divided by the appropriate error for lags 1 (*circles*) and 2 (*squares*). The *y*-axis on the right-hand side represents the autocorrelation (ACF) for lags 1 (*diamonds*) and 2 (*triangles*) divided by its error. The *x*-axis is the number of closed episodes. As we plot the functions divided by their error bars, all of them have the same relative error, which is shown in the figure for a 95% confidence level (*upper left-hand corner*).

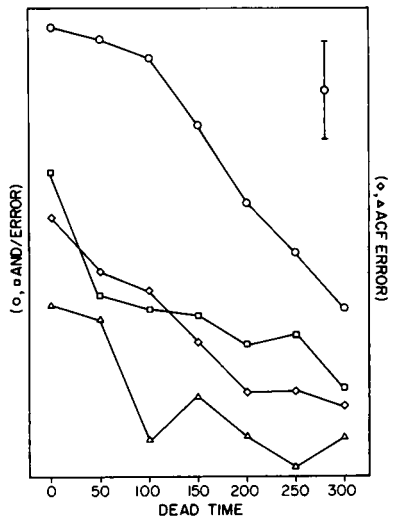

FIGURE 7 Similar to Fig. 6, except that the change in the various functions (again scaled to their respective error bars) is plotted as the deadtime is varied. Note the different physical size of the error bar (*upper right-hand corner*) for comparing with Fig. 6.

The dependence of both functions on dead time is a very model-dependent quantity. In the model under consideration the moding is produced by an excess of neighboring long dwell times; the functions would presumably not be as greatly affected by finite dead times as when moding produces an abundance of clustered short dwell times comparable to the dead time. Ball and Sansom (12) have analyzed theoretically the effects of finite dead time on the decay of the autocorrelation in a Markov framework. Their work indicated that the number of exponential components needed to fit the autocorrelation is bounded by the number of "gateway" (12) states in the Markov model, independent of the dead time. However, finite dead time does influence the weights and decay rate of these components. We have not performed a similar calculation for the absolute normed deviation; but the limited simulations we have done indicate that the dead time has a lesser effect on the weights of the components calculated from this function than it does in the case of the autocorrelation.

CONCLUSION

In summary, we note that the two discriminators presented here, the running median pdf and the second cpd, supplement the existing procedures using the mean and

correlation very well. Both the median pdf and second cpd pick up mode shifting in cases where the other two methods fail. The second cpd in particular does better in cases where the predictions of all methods agree qualitatively. The running median pdf is better than running mean pdf for detection of clustered long dwell times; for clustered short dwell times the two the comparable, with the running mean doing better in one of the simulations (Fig. 1 *e*) and the running median better in another (Fig. 2 *e*).

It may also be noted that calculation of the second cpd after the initial sort is less computationally intensive (N additions) than that for the autocorrelation (N multiplications). Although actual median and mean pdf's evaluate easily, calculation of modeless pdf's for running median and mean is demanding of computer time, with the running median for A-groups requiring ~log (A) passes through the data.

The authors thank Dr. K. Manivannan for many helpful discussions and a careful review of the manuscript, and the referees for their comments.

This work was supported by National Institutes of Health grant HL 31299.

Received for publication 7 August 1989 and in final form 18 December 1989.

REFERENCES

1. Sakmann, B., and E. Neher, editors. 1983. Single Channel Recording. Plenum Publishing Corp., New York.
2. Fredkin, D. R., M. Montal, and J. A. Rice. 1985. Identification of aggregated Markov models: application to the nicotinic acetylcholine receptor. *In* Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer. Vol. 1. Lucien M. Le Cam and Richard A. Olshen, editors. Wadsworth Press, Belmont, CA. 269–289.
3. Colquhoun, D., and A. G. Hawkes. 1987. A note on correlations in single ion channel records. *Proc. R. Soc. Lond. B Biol. Sci.* 230:15–52.
4. McManus, O. B., and K. L. Magleby. 1988. Kinetic states and modes of single large-conductance calcium-activated potassium channels in cultured rat skeletal muscle. *J. Physiol. (Lond.).* 402:79–120.
5. Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1988. Numerical Recipes. Cambridge University Press, Cambridge, MA. 818 pp.
6. Millhauser, G. L., E. E. Salpeter, and R. E. Oswald. 1988. Rate-amplitude correlation from single-channel records. *Biophys. J.* 54:1165–1168.
7. Sigworth, F. J., and S. M. Sine. 1987. Data transformations for improved display and fitting of single-channel dwell time histograms. *Biophys. J.* 52:1047–1054.
8. Labarca, P., J. A. Rice, D. R. Fredkin, and M. Montal. 1985. Kinetic analysis of channel gating. Application to the cholinergic receptor channel and the chloride channel from *Torpedo California. Biophys. J.* 47:469–478.
9. Clay, J. R., and L. J. DeFelice. 1983. Relationship between membrane excitability and single channel open-close recording. *Biophys. J.* 42:151–157.
10. Colquhoun, D., and A. G. Hawkes. 1981. On the stochastic properties of single ion channels. *Proc. R. Soc. Lond. B Biol. Sci.* B211:205–235.
11. Magleby, K. L., and B. S. Pallotta. 1983. Calcium dependence of open and shut interval distributions from calcium-activated potassium channels in cultured rat muscle. *J. Physiol. (Lond.).* 344:585–604.
12. Ball, F. G., and M. S. P. Sansom. 1988. Single-channel autocorrelation functions. Effects of time interval omission. *Biophys. J.* 53:819–832.