ESTIMATING KINETIC CONSTANTS FROM SINGLE CHANNEL DATA

RICHARD HORN AND KENNETH LANGE
Departments of Physiology and Biomathematics, University of California at Los Angeles School of Medicine, Los Angeles, California 90024

ABSTRACT The process underlying the opening and closing of ionic channels in biological or artificial lipid membranes can be modeled kinetically as a time-homogeneous Markov chain. The elements of the chain are kinetic states that can be either open or closed. A maximum likelihood procedure is described for estimating the transition rates between these states from single channel data. The method has been implemented for linear kinetic schemes of fewer than six states, and is suitable for nonstationary data in which one or more independent channels are functioning simultaneously. It also provides standard errors for all estimates of rate constants and permits testing of smoothly parameterized subhypotheses of a general model. We have illustrated our approach by analysis of single channel data simulated on a computer and have described a procedure for analysis of experimental data.

INTRODUCTION

In the last few years it has become possible to record current through individual ionic channels embedded in lipid bilayer membranes. This was accomplished by two techniques, patch recording from native membranes using glass pipettes to isolate a small membrane area (Neher and Sakmann, 1976), and recording from artificial planar lipid bilayers containing ionic channels (Bean et al., 1969). In most cases, the opening and closing of a channel, or the transitions between conductance states of a channel, are rapid compared with the dwell time of a channel in a given conductance level. In fact, these transitions are more rapid than the limits of present technology (Hamill et al., 1981). Dwell times within a given kinetic state also appear to be exponentially distributed (cf. Labarca et al.,1980; Fukushima, 1981). Thus, it is natural to consider Markov chain models when studying the kinetics underlying the opening and closing of ionic channels (Colquhoun and Hawkes, 1977, 1981). The transition rates, or rate constants, in such models are time invariant, although they may depend on voltage or the concentration of substances such as agonists or channel blockers. When the averaged behavior of many channels is studied, either as macroscopic current relaxations after a sudden perturbation of the transition rates, or as fluctuations in membrane currents, overall (or average) rates are obtained (Colquhoun and Hawkes, 1977). The theory that relates overall rates and individual transition rates in such Markov processes has already been developed (Conti and Wanke, 1975; Neher and Stevens, 1977; Colquhoun and Hawkes, 1977; DeFelice, 1981). Kinetic models with a small number of discrete Markov states are particularly attractive for analysis since they are often analytically or numerically tractable. In this case the overall rates are the decay rates (i.e. eigenvalues) in a finite sum of exponentially decaying components. The number of components is $N - 1$ for a Markov process with N states.

Single channel measurements greatly enhance the ability to relate physiological data to a given kinetic model for several reasons. First, some transition rates may be measured directly. For example, if a channel has two kinetic states, one open and the other closed, a single overall rate can be measured from macroscopic data. This rate is the sum of the rate constants for opening and closing. These two transition rates cannot be determined from the overall rate alone. For such a simple model, single channel recording can be used to estimate the transition rate for opening from the average closed time and the transition rate for closing from the average open time of the channel. Second, the number of exponentially decaying components in histograms of open time, or closed time, is usually less than that obtained from measurements of overall rates, and therefore easier to separate in general. For example, a Markov process with two closed states and two open states will have three overall rates in macroscopic currents; but histograms of open or closed time will only yield two time constants each. Last, some processes, especially rapid ones, may be difficult to detect in macroscopic measurements, but not in single channel measurements (Patlak, 1983). The principal technique for single channel data analysis is the lifetime histogram, in which dwell time of a channel in a given conductance level is plotted in bins on the abscissa, and the number of events observed in each bin is plotted on the ordinate (cf. Fig. 2). The histogram, when scaled properly, is an estimate of the probability density for the lifetime at that conductance level and therefore may be compared with a kinetic model. For Markov models of a single channel, histograms with their underlying probability densities are sums of exponential

components with as many rates as states at a given conductance level.

Several problems have been encountered in applying this type of analysis to data obtained from both biological membranes and artificial lipid bilayers. We list some of these problems below.

(*a*) *Multiple states.* In many, if not most, channels studied experimentally, more than one state has the same conductance level, making it difficult in some cases to determine transition rates between these states. For example, in the Hodgkin-Huxley (1952) model of the sodium channel of nerve there are eight kinetically distinct states, one of which is open. The other seven states are closed and have the same conductance. Therefore a histogram of closed times for a channel with a kinetic scheme like this will have seven exponential components. To make matters worse, several of these components in the Hodgkin-Huxley model differ from one another by a factor of 2 or less. This renders it extremely difficult to separate and identify the components.

(*b*) *Multiple channels.* In many experimental situations it is difficult or impossible to observe only one channel at a time. For example, when studying acetylcholine receptor channels with patch electrodes the number of channels in the membrane patch is usually not known, but is almost always >1. To observe only one channel open at a time, very low concentrations of agonist are applied to the membrane. When openings of two or more channels overlap, these data are usually discarded. In other cases the number of channels may be known (e.g., see Labarca et. al., 1980; Patlak and Horn, 1982), but again overlapping events are not used in histograms, unless the underlying kinetic model for each channel is very simple (Labarca et. al., 1980).

(*c*) *Nonstationarity.* For channels with more than two kinetic states, the dwell time in a given conductance level may have a time dependence. For example, fluctuation analysis of sodium current in the node of Ranvier suggests that these sodium channels have more than one kinetically distinct open state, each with the same conductance (Sigworth, 1981). During an activating depolarization, the dwell time in open states may change with time in some models (e.g., the model of Armstrong and Bezanilla, 1977). This is due to the fact that the probability of being in one (rather than another) open state is nonstationary. Therefore, the histograms of open time will depend on the interval between the onset of the depolarization and the time of data collection. This may lead to difficulties in obtaining sufficient data to yield meaningful histograms at all relevant times. Even in kinetic models with a single open state, if more than a single channel is present, then the dwell time for a given number of open channels is, in general, nonstationary. This is due to the fact that the probability of overlapping openings depends on time. One way to circumvent this problem is to wait until the overall process reaches a steady state level (Labarca et. al., 1980). This has two difficulties. The first is that kinetic information is lost on the way to steady state. The second is that in some cases the steady state probability of a channel being open is close to zero, as in the case of sodium channels, which inactivate.

(*d*) *Finite length of records.* Records of single channel events are always finite. In some cases channel behavior is studied during a finite voltage pulse. Perhaps more commonly, finite records are digitized and accumulated by a computer. If the histogram method is used to study transition rates, the last event before the end of the record has to be discarded. If, for example, a channel is open at the end of a record, the duration it remained open cannot be added to the open time histogram. Only open channels that have successfully closed can be used. However, useful information is being discarded, since this interval represents time during which the channel failed to close.

(*e*) *Combining collected data.* When all the data for a given experiment are gathered together, usually in the form of histograms fitted to theoretical functions, it is difficult or impossible to decide how to combine all the data to get a best model for the underlying kinetic process. Some of the histograms may be more reliable than others that have fewer events. Yet a strategy for weighting the histograms according to relative reliability is not available.

We have developed a method of analysis that is able to extract transition rates from single channel data under a wide variety of the conditions mentioned above. The method involves maximum likelihood estimates of these transition rates and uses combinatorial analysis to deal with the overlapping openings and closings of a small number of channels. In our opinion, the maximum likelihood approach is extremely powerful. It is conceptually appealing, but computationally demanding. However, for a variety of models the computations are feasible given current computer technology. To illustrate this method we will first develop the theory, then give examples of analyses of simulations of single channel data, and finally present a strategy for analysis of experimental data.

THEORY

Compositions

It is crucial to have an appropriate mathematical framework to handle the complications of multiple states and multiple channels. Suppose there are k states per channel and n channels. Label the states $1,2,\ldots,k$, and let n_i be the number of channels in state i at a given time. Because of the indistinguishability of channels, the system at any time can be characterized by a k-tuple of nonnegative integers $(n_1,\ldots,n_k)$ subject to the constraint

$$\sum_{i=1}^{k} n_i = n.$$

Such k-tuples are termed compositions of n into k parts (Nijenhuis and Wilf, 1978). For instance, if there are four states and three channels, then the composition (2,0,0,1) corresponds to two channels in state 1 and one channel in state 4. The number of compositions of n into k parts is given by the binomial coefficient $\binom{n+k-1}{n}$. (Nijenhuis and Wilf, 1978; Feller, 1968). This can be a large number, but it is always smaller and often much smaller than the number of ways k^n of distributing n distinguishable channels into k states. For instance, $\binom{3+4-1}{3} = 20$, while $4^3 = 64$. Table I lists all possible compositions of 3 into 4 parts.

The Composition Markov Process

As indicated above, we will be considering models in which each channel moves independently from state to state according to a continuous time Markov chain. Such motion is completely characterized for a single channel by the transition rates q_{ij} (Colquhoun and Hawkes, 1977). For $i \neq j$ and Δt the length of a small interval of time, $q_{ij}\,\Delta t$ is approximately the probability that a channel in state i will move to state j during the given time interval. If Δt is very small, the probability of more than one transition can be neglected.

In each of the models discussed in this paper, the states are divided into closed and open states. For simplicity, suppose the first ℓ states, $1 \leq \ell < k$, are closed and the remaining $k - \ell$ states are open. (Our labeling of states at this point is not meant to imply which states are neighbors in a kinetic scheme.) If the number of channels occupying closed states is c, then the number of channels occupying open states is clearly $n - c$. Each corresponding composition $(n_1, \ldots, n_k)$ satisfies the restrictions

$$\sum_{i=1}^{\ell} n_i = c$$

$$\sum_{i=\ell+1}^{k} n_i = n - c.$$

In other words, the collection of restricted compositions can be viewed as the Cartesian product of compositions of c into ℓ parts against compositions of $n - c$ into $k - \ell$ parts. It follows that there are $\binom{c+\ell-1}{c}\binom{n-c+k-\ell-1}{n-c}$ such restricted compositions. For example, if $k = 4, \ell = 2, c = 2$, and $n = 3$, there are $3 \times 2 = 6$ restricted compositions. One of these restricted compositions is (2,0,0,1).

In computing with the compositions of n into k parts it is convenient to be able to store certain probabilities connected with them in a single dimensional array. The alternative is to store these probabilities in a k-dimensional array with each array index ranging from 0 to n. This alternative storage arrangement can waste a lot of space. It has the further disadvantage of making all computer code depend rigidly on the number of parts k. To implement the single dimensional storage strategy it is necessary to have computer algorithms for two purposes. First, we need to generate all compositions in some appropriate sequence with a minimum of computational effort. Second, given a composition, we must be able to rank it according to the generating sequence. The value of a good ranking procedure will become evident when we discuss our methods for computing likelihoods. In computing these likelihoods it is necessary to consider transitions from a given composition to its neighbors. For instance, a typical neighbor of (2,0,0,1) is (1,1,0,1). The current probability connected with this neighbor must be accessed quickly. In other words, the neighboring composition must be ranked quickly.

Fortunately Nijenhuis and Wilf (1978) discuss both the generating and ranking tasks and give appropriate FORTRAN algorithms. By computing and storing certain binomial coefficients, we are able to improve on their ranking procedure. Because this technicality is not of central interest, we omit details. Table I illustrates the generated sequence for all compositions of 3 into 4 parts. Although the sequence of compositions appears nearly random, there is an underlying subtle pattern that Nijenhuis and Wilf (1978) fully explain.

The composition Markov process keeps track of the collective behavior of all n channels. The composition transition rate for the transition from

$$(n_1, \ldots, n_i, \ldots, n_j, \ldots, n_k),\ n_i > 0,$$

to

$$(n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_k)$$

is given by

$$n_i q_{ij}.$$

This expression simply reflects the fact that the n_i indistinguishable channels in state i act independently. Note that the composition process is Markovian because the minimum of a finite number of independent, exponentially distributed waiting times is still exponentially distributed (Feller, 1971, p.19).

The independent movement of different channels makes it easy to calculate some quantities of interest. If we start

TABLE I
ALL COMPOSITIONS OF 3 INTO 4 PARTS

Rank	Composition	Rank	Composition
1	(3, 0, 0, 0)	11	(2, 0, 0, 1)
2	(2, 1, 0, 0)	12	(1, 1, 0, 1)
3	(1, 2, 0, 0)	13	(0, 2, 0, 1)
4	(0, 3, 0, 0)	14	(1, 0, 1, 1)
5	(2, 0, 1, 0)	15	(0, 1, 1, 1)
6	(1, 1, 1, 0)	16	(0, 0, 2, 1)
7	(0, 2, 1, 0)	17	(1, 0, 0, 2)
8	(1, 0, 2, 0)	18	(0, 1, 0, 2)
9	(0, 1, 2, 0)	19	(0, 0, 1, 2)
10	(0, 0, 3, 0)	20	(0, 0, 0, 3)

with a single channel in state i and let it move randomly from state to state, then it is possible to calculate the probability $P_{ij}(t)$ that it is in some state j after t units of time. This calculation is reviewed in the next section. The point we wish to make here is that these elementary probabilities allow us to calculate the probabilities of all destination compositions after t units of time starting with the initial composition $(n_1, \ldots, n_k)$.

The possible destination compositions for the n_i channels starting in state i can be expressed as a random vector X_i. It is clear that X_i follows a multinomial distribution. In other words, if $m^i = (m_1^i, \ldots, m_k^i)$ is a possible value for X_i,

$$P[X_i = (m_1^i, \ldots, m_k^i)] = \binom{n_i}{m_1^i, \ldots, m_k^i} \prod_{j=1}^{k} P_{ij}(t)^{m_j^i},$$

where $\binom{n_i}{m_1^i \ldots m_k^i}$ is a multinomial coefficient. The random vector $X_1 + \ldots + X_k$ represents the random destination composition for all n channels. Its distribution is found by the vector convolution of the distributions of the X_i. If $n^* = (n_1^*, \ldots, n_k^*)$ is a possible value for $X_1 + \ldots + X_k$, then

$$P(X_1 + \cdots + X_k = n^*) = \sum_{n^* = m^1 + \cdots + m^k} \prod_{i=1}^{k} P(X_i = m^i).$$

Note that, for $1 \leq j \leq k$, $\sum_{i=1}^{k} m_j^i = n_j^*$ In practice, it is convenient to compute recursively the distribution of

$$X_1 + X_2$$
$$(X_1 + X_2) + X_3$$
$$\vdots$$
$$(X_1 + X_2 + \cdots + X_{k-1}) + X_k.$$

In doing so, note that $X_1 + \ldots + X_i$ ranges over all compositions of $n_1 + \ldots + n_i$ into k parts.

When each channel tends to an equilibrium distribution as $t \to \infty$, the equilibrium distribution for the composition process is multinomial. Thus, when

$$\lim_{t \to \infty} P_{ij}(t) = P_j,$$

the equilibrium probability of the composition $(n_1, \ldots, n_k)$ is

$$\binom{n}{n_1, \ldots, n_k} \prod_{j=1}^{k} P_j^{n_j}.$$

Review of Some Calculations for Continuous Time Markov Chains

One of the most important results in the theory of continuous time Markov chains is an explicit formula for the probabilities $P_{ij}(t)$ introduced in the last section (Cox and Miller, 1965; Karlin and Taylor, 1975; Colquhoun and Hawkes, 1977). This formula is general and does not depend on the present physiological interpretation. To state the formula let $\mathbf{Q}$ be the $k \times k$ matrix that, in our earlier notation, has off-diagonal elements q_{ij} and diagonal elements $q_{ii} = -\Sigma_{j \neq i} q_{ij}$. Then $P_{ij}(t)$ is the entry in row i and column j of the matrix exponential

$$\exp(t\mathbf{Q}) = \sum_{m=0}^{\infty} \frac{t^m}{m!} \mathbf{Q}^m.$$

For small $t > 0$, $\exp(t\mathbf{Q})$ can be computed by truncating the series expansion to

$$\sum_{m=0}^{r} \frac{t^m}{m!} \mathbf{Q}^m$$

for r of moderate size. For larger t such truncation can lead to serious errors. If the truncated expansion is sufficiently accurate for all $t \leq t_0$, where t_0 is a fixed constant, then for arbitrary t one can exploit the multiplicative property

$$\exp[(s + t)\mathbf{Q}] = \exp(s\mathbf{Q}) \exp(t\mathbf{Q}) \tag{1}$$

of the matrix exponential. Thus, if $t > t_0$, take some positive integer j so that $2^{-j}t \leq t_0$ and form $\exp(2^{-j}t\mathbf{Q})$. From Eq. 1 it follows that $\exp(t\mathbf{Q})$ can be computed by squaring $\exp(2^{-j}t\mathbf{Q})$, squaring the result $\exp(2^{-j+1}t\mathbf{Q})$, squaring the result of this and so forth, a total of j times. This is the simple procedure suggested by Moler and Van Loan (1978). It has the virtue of avoiding explicit calculation of the eigenvalues and eigenvectors of $\mathbf{Q}$.

For any matrix norm, $\| \ \|$, it is possible to derive the inequality

$$\left\| \exp(t\mathbf{Q}) - \sum_{m=0}^{r} \frac{t^m}{m!} \mathbf{Q}^m \right\| \leq \frac{|t|^{r+1} \|\mathbf{Q}\|^{r+1}}{(r+1)!} \cdot \frac{1}{1 - \frac{|t| \cdot \|\mathbf{Q}\|}{r+2}}. \tag{2}$$

See Isaacson and Keller (1966) for a discussion of matrix norms and Liou (1966) for a derivation of Eq. 2. Two easily computed norms are

$$\max_i \sum_{j=1}^{k} |q_{ij}|$$

and

$$\max_j \sum_{i=1}^{k} |q_{ij}|.$$

In practice, we have used the smaller of these norms and fixed r at 10. This allows us to solve for the maximum time t_0 that is consistent with a small value for the error estimate (Eq. 2).

Now suppose we have a collection of states like the closed states for a channel. As above we let these be the first ℓ out of the possible k states. The matrix $\mathbf{Q}$ can be

partitioned into four blocks

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{cc} & \mathbf{Q}_{c\infty} \\ \mathbf{Q}_{oc} & \mathbf{Q}_{o\infty} \end{pmatrix},$$

where $\mathbf{Q}_{cc}$ is the $\ell \times \ell$ matrix for transitions between closed states, $\mathbf{Q}_{oc}$ is the $(k - \ell) \times \ell$ matrix for transitions from open to closed states, and $\mathbf{Q}_{c\infty}$ and $\mathbf{Q}_{o\infty}$ are defined similarly. We can make the open states absorbing with respect to closed states by setting

$$\mathbf{Q}_{oc} = 0,$$

the $(k - \ell) \times \ell$ matrix with all entries 0. In other words, once the collection of open states is entered, it is never left. With this adjustment in $\mathbf{Q}$ it is possible to show that

$$\exp(t\mathbf{Q}) = \begin{pmatrix} \exp(t\mathbf{Q}_{cc}) & * \\ 0 & * \end{pmatrix},$$

where the * terms are unimportant.

Because the open states have been made absorbing with respect to the closed states, the entry $P_{ij}^{c}(t)$ of $\exp(t\mathbf{Q}_{cc})$ can be interpreted as the probability that a channel begins in state i at time 0, ends in state j at time t, and never leaves the closed states in the interim. Clearly the entries of $\exp(t\mathbf{Q}_{o\infty})$ bear a similar interpretation. It is important to note the obvious extension for the composition process, namely, we can calculate the probability that the initial composition $(n_1, \ldots, n_\ell)$ for the closed states changes to the final composition $(m_1, \ldots, m_\ell)$ for the closed states at time t with no channel passing through an open state during the interim. We simply replace $P_{ij}(t)$ by $P_{ij}^{c}(t)$ and argue as in the last section. The only difference is that we now concentrate exclusively on the closed states instead of all states.

Discrete Time Analogue

Any continuous time Markov chain with a finite number of states can be approximated by a discrete time analogue with the same set of states. Let Δt be a small time interval and consider observations of a single channel at times 0, $\Delta t, 2\Delta t, 3\Delta t, \ldots, (N-1)\Delta t$. If the channel is in state i at time $n\Delta t$, then it will move to state j at time $(n+1)\Delta t$ with approximate probability $q_{ij}\Delta t$. Alternatively, the channel will remain in the state i with the approximate probability

$$1 - \sum_{j \neq i} q_{ij}\,\Delta t.$$

This approximating discrete time chain best captures the behavior of the continuous time chain when all the sums $\Sigma_{j \neq i} q_{ij}\Delta t$ are significantly <1.

For the composition process, the one-step transition from

$$(n_1, \ldots, n_i, \ldots, n_j, \ldots, n_k), \quad n_i > 0,$$

to

$$(n_1, \ldots, n_i - 1, \ldots, n_j + 1, \ldots, n_k)$$

occurs with approximate probability

$$n_i\, q_{ij}\,\Delta t. \tag{3}$$

The process remains in its current composition with approximate probability

$$1 - \sum_i \sum_{j \neq i} n_i\, q_{ij}\,\Delta t. \tag{4}$$

In general, the larger the number of channels n, the smaller Δt should be. With Δt too large, the probability of multiple transitions during a time interval of length Δt cannot be neglected.

Likelihood for the Discrete Time Analogue

In the discrete time version, let E_i be the observed number of open channels at epoch $i\Delta t$. The E_i's can be viewed as probabilistic events. By definition, the likelihood of our experimental data is the probability

$$P(E_0 \cap E_1 \cap \cdots \cap E_{N-1}) \tag{5}$$

of the joint occurrence of all these N events. This probability depends on the transition rates as well as the events E_i. The basic idea of the maximum likelihood method is to maximize the likelihood with respect to the parameters, in this case the transition rates. The fundamental stumbling block to applying the method is that a closed form expression for Eq. 5 is apt to be incredibly complex. There can be an astronomical number of sample paths consistent with the observations E_0 through E_{N-1}. Our approach is to compute recursively probabilities akin to

$$P(E_0 \cap E_1 \cap \cdots \cap E_i),$$

starting with $i = 0$ and incrementing i by 1 until $N - 1$ is reached. While computationally demanding, this technique is conceptually simple and does permit exact calculation of the discrete time likelihood. In practice, we maximize the natural logarithm of the likelihood. This is equivalent to maximizing the likelihood itself (Rao, 1973).

What we actually compute recursively are the probabilities

$$P(F_{in} \cap E_0 \cap \cdots \cap E_i), \tag{6}$$

where F_{in} is the event that composition n occurs at epoch $i\Delta t$. In other words, we carry along the current destination of the process in the calculation. Note that only certain compositions are compatible with the observed number of open channels at epoch $i\Delta t$. Thus, either

$$F_{in} \subset E_i$$

or

$$F_{in} \cap E_i = \phi.$$

Also note that

$$P(E_0 \cap E_1 \cap \cdots \cap E_{N-1}) = \sum_{F_{N-1n} \subset E_{N-1}} P(F_{N-1n} \cap E_0 \cap \cdots \cap E_{N-1}),$$

so that the likelihood is recoverable at the final stage of the recursion.

The initial values

$$P(F_{0n} \cap E_0)$$

are taken as given. These can be computed as a multinomial distribution from the initial distribution for a single channel. For example, suppose that the probabilities of being in states S_1, S_2, and S_3 in a three-state kinetic model are p_1, p_2, and p_3, respectively at time zero. If four channels are present, the probability that the initial composition is (2,1,1) is given by

$$\binom{4}{2\,1\,1} p_1^2 p_2 p_3.$$

Now suppose the probabilities in Eq. 6 are known for epoch $i\Delta t$. Then for $F_{i+1n} \subset E_{i+1}$

$$P(F_{i+1n} \cap E_0 \cap \cdots \cap E_{i+1}) = \sum_{F_{im} \subset E_i} P(F_{i+1n} \cap F_{im} \cap E_0 \cap \cdots \cap E_i)$$

$$= \sum_{F_{im} \subset E_i} P(F_{i+1n} | F_{im} \cap E_0 \cap \cdots \cap E_i) \cdot P(F_{im} \cap E_0 \cap \cdots \cap E_i),$$

where $P(F_{i+1n} | F_{im} \cap E_0 \cap \ldots \cap E_i)$ denotes the conditional probability of F_{i+1n} given $F_{im} \cap E_0 \cap \ldots \cap E_i$. Since we are dealing with a Markov chain,

$$P(F_{i+1n} | F_{im} \cap E_0 \cap \cdots \cap E_i) = P(F_{i+1n} | F_{im}).$$

But the conditional probabilities $P(F_{i+1n} | F_{im})$ are given precisely by Eqs. 3 and 4. Hence, for $F_{i+1n} \subset E_{i+1}$, we can compute recursively via

$$P(F_{i+1n} \cap E_0 \cap \cdots \cap E_{i+1}) = \sum_{F_{im} \subset E_i} P(F_{i+1n} | F_{im}) P(F_{im} \cap E_0 \cap \cdots \cap E_i). \quad (7)$$

One further comment is in order. The individual terms in Eq. 7, and indeed the final likelihood itself, can be extremely small. We avoid problems of underflows in computing by periodically rescaling all the probabilities in Eq. 7. The rescaling factors can be accumulated and reincorporated in the final logarithm of the likelihood. Thus, if at some epoch $i\Delta t$

$$S = \max P(F_{im} \cap E_0 \cap \cdots \cap E_i)$$

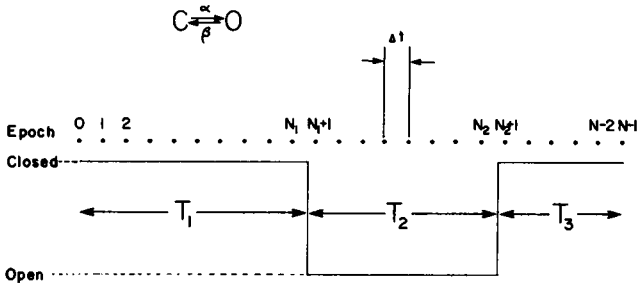

FIGURE 1 A hypothetical current trace is shown in which a single channel opens for a duration of T_2. In the discrete case the current is sampled at N epochs numbered $0,1,\ldots,N-1$, with a time interval of Δt between successive epochs. The kinetic model for this channel is shown above.

is too small, one tactic is to divide all terms $P(F_{im} \cap E_0 \cap \ldots \cap E_i)$ by S and add $\log(S)$ to the final logarithm of the likelihood.

Perhaps it would be useful to illustrate what is involved in a simple case stripped of all combinatorial complexities. Consider the simplest two-state model $C \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} O$ with only one channel. We are dealing, therefore, with two compositions, $a = (1,0)$ and $b = (0,1)$, corresponding to the single channel being closed or open, respectively. Fig. 1 gives a sample current record. In the discrete case we sample at times $0, \Delta t, \ldots, (N-1)\Delta t$. There is an initial quiet period from 0 to $N_1\Delta t$, a transition $C \rightarrow O$ during $N_1\Delta t$ to $(N_1+1)\Delta t$, another quiet period from $(N_1+1)\Delta t$ to $N_2\Delta t$, a transition $O \rightarrow C$ during $N_2\Delta t$ to $(N_2+1)\Delta t$, and a final quiet period from $(N_2+1)\Delta t$ to $(N-1)\Delta t$. The likelihood of this current record in the discrete time approximation is

$$L = (1-\alpha\Delta t)^{N_1} \alpha\Delta t (1-\beta\Delta t)^{N_2-N_1-1} \beta\Delta t (1-\alpha\Delta t)^{N-N_2-2}.$$

This formula follows directly from the recursion formula (Eq. 7).

To illustrate this in more detail, suppose in the above example that $N_1 = 2$, $N_2 = 4$, and $N = 6$. Then the events E_i, $i = 0$ to $N-1$, are 0, 0, 0, 1, 1, 0. If we assume that the channel is closed at the beginning of the observation period, then $P(F_{0a} | E_0) = 1$. The recursive computation, using Eq. 7, follows directly.

$$P(F_{1a} | F_{0a}) = 1 - \alpha\Delta t.$$

Therefore

$$P(F_{1a} \cap E_0 \cap E_1) = 1 - \alpha\Delta t.$$

This is the likelihood, L, for the first epoch, since all other compositions are inconsistent with the observed event at this epoch. For the next epoch,

$$P(F_{2a} | F_{1a}) = 1 - \alpha\Delta t.$$

Therefore,

$$P(F_{2a} \cap E_0 \cap E_1 \cap E_2) = (1 - \alpha\Delta t)^2,$$

which is the likelihood at this epoch. Then

$$P(F_{3b} | F_{2a}) = \alpha\Delta t,$$

and

$$L = P(F_{3b} \cap E_0 \cap E_1 \cap E_2 \cap E_3) = (1 - \alpha\Delta t)^2 \alpha\Delta t.$$

This method is applied recursively to give

$$L = P(F_{5a} \cap E_0 \cap E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5)$$
$$= (1 - \alpha\Delta t)^2 \alpha\Delta t(1 - \beta\Delta t)\beta\Delta t.$$

Likelihood for the Continuous Time Markov Model

The likelihood of the continuous time version of the above example is approximated by sending $\Delta t \to 0$ and N_1, N_2, $N \to \infty$ in such a way that

$$N_1\Delta t \to T_1$$
$$(N_2 - N_1)\Delta t \to T_2$$
$$(N - N_2)\Delta t \to T_3.$$

The adjusted discrete time likelihood $L/\Delta t^2$ approaches

$$\exp(-\alpha T_1)\,\alpha \exp(-\beta T_2)\,\beta \exp(-\alpha T_3). \qquad (8)$$

This quantity is by definition the likelihood for the continuous version of the process. It is no longer a probability, but it still summarizes the current record and can be maximized with respect to the parameters α and β. Observe that the omitted factor Δt^2 is irrelevant in determining the maximum of the likelihood with respect to α and β. Furthermore, if this factor were retained, the above procedure would give 0 in the limit for the likelihood.

For this simple model, the generalization to multiple channels is straightforward. Suppose we obtain the current record of Fig. 1 from a membrane containing three channels, of which one opened for a duration of T_2. Using our previous notation, an opening corresponds to a transition from composition (3,0) to composition (2,1). In the discrete case the likelihood for the current record is

$$(1 - 3\alpha\Delta t)^{N_1} 3\alpha\Delta t[(1 - \beta\Delta t)(1 - 2\alpha\Delta t)]^{N_2 - N_1 - 1} \cdot \beta\Delta t(1 - 3\alpha\Delta t)^{N - N_2 - 2},$$

and in the continuous case it is

$$\exp(-3\alpha T_1) 3\alpha [\exp(-\beta T_2) \cdot \exp(-2\alpha T_2)]\,\beta \exp(-3\alpha T_3).$$

Notice that if this current record were used for making a histogram, the interval T_3 would be discarded, thus losing information about the opening rate constant, α.

Let us now return to the full model with multiple states and multiple channels. Suppose the events $E_i, \ldots, E_{i+j}$ all exhibit the same number of open channels. In other words, the time between $i\Delta t$ and $(i + j)\Delta t$ is a quiet period. The key to summarizing this quiet period is to generalize the recurrence relation (Eq. 7). A straightforward extension of our earlier reasoning implies

$$P(F_{i+j,n} \cap E_1 \cap \cdots \cap E_{i+j}) = \sum_{F_{im} \subset E_i} P(F_{i+j,n} \cap E_{i+1} \cap \cdots \cap E_{i+j} | F_{im}) \cdot P(F_{im} \cap E_0 \cap \cdots \cap E_i).$$

To apply this formula it is important to observe that the conditional probability $P(F_{i+j,n} \cap E_{i+1} \cap \ldots \cap E_{i+j} | F_{im})$ can be represented as a product of two probabilities. The first probability is the probability that the subcomposition $(m_1, \ldots, m_\ell)$ for the ℓ closed states leads to the subcomposition $(n_1, \ldots, n_\ell)$ without any of the $m_1 + \ldots + m_\ell$ initially closed channels passing into an open state during the interim. The second probability is basically the same except that it pertains to the movement of the $m_{\ell+1} + \ldots + m_k$ initially open channels. The product of these two probabilities is appropriate because the closed and open channels behave independently during a quiet period. As mentioned above, these two probabilities are available exactly for a continuous time Markov chain. Thus, in the limit as $\Delta t \to 0$, the conditional probability

$$P(F_{i+j,n} \cap E_{i+1} \cap \cdots \cap E_{i+j} | F_{im})$$

can be calculated exactly for any quiet period. Clearly the update in the likelihood algorithm for a quiet period poses no problem. The infinitesimal times of occurrence for movements of channels from open states to closed states or vice versa are handled as in the simple example; namely, if a movement occurs between $i\Delta t$ and $(i + 1)\Delta t$, then the infinitesimal factor Δt in the one step transition probability $P(F_{i+1,n} | F_{im})$ must be dropped when using Eq. 7 (see simplest case of Fig. 1 and Eq. 8). With these two modifications, our basic likelihood algorithm carries over to the continuous case.

Estimation of the Closing Rate Constant

Although the likelihood function is generally complex, a simple closed form expression can be obtained for the maximum likelihood estimate of the closing rate constant in models where the last state in a linear sequence is the only open state. In fact, the only necessary condition for the model is that the one open state closes by the same path by which it opened. As models of this form are frequently discussed in the physiological literature, we present the derivation here.

The simplest case is that of a single two-state channel with one open and one closed state. Suppose we observe N samples of open time durations, $T_1, \ldots, T_N$. If the closing

rate constant is β, the probability density of open channel lifetime is $\beta \exp(-\beta T)$. The likelihood of observing exactly these N samples is given by

$$L(\beta) = \prod_{i=1}^{N} \beta \exp(-\beta T_i).$$

We will maximize the logarithm of $L(\beta)$ with respect to β as follows.

$$L^*(\beta) = \log [L(\beta)] = N \log(\beta) - \sum_{i=1}^{N} \beta T_i.$$

and

$$\frac{d}{d\beta} L^*(\beta) = N/\beta - \sum_{i=1}^{N} T_i.$$

Setting this derivative equal to zero, we find that the maximum likelihood estimate for $1/\beta$ is $\Sigma_{i=1}^{N} T_i/N$, i.e., the mean open time.

For the general case, with multiple channels and multiple closed states leading to a single open state, it can be seen that the likelihood function for such a model is a product of terms that include both simple exponential functions of quiet intervals, T_i, and also entries of matrix exponentials. Because of the limitations of this kinetic scheme, the closing rate constant, which we again denote as β, will only appear in two types of terms, either as $\exp(-n_i \beta T_i)$ or as $n_i \beta$, where n_i represents in the first case the number of open channels during T_i, and in the second case the number of channels open prior to a closing event. Taking the logarithm of the likelihood and differentiating with respect to β, all terms involving other rate constants drop out. By setting this derivative equal to zero, we find that the simple solution for the maximum likelihood estimate of β is

$$\hat{\beta} = \frac{n'}{\sum_i n_i T_i}, \tag{9}$$

where n_i is the number of open channels during the ith quiet period T_i, and n' is the total number of closings measured. Eq. 9 has been used by Fenwick et al. (1982) to estimate open time of sodium channels in chromaffin cells. The theoretical justification for and limitations of the use of this equation are presented here for the first time. The importance of this observation is its generality. Notice that the total number of channels in the membrane is irrelevant. It is not even required that the total number remains constant (as long as channels do not close by leaving the active population). Notice also that this estimation can be used rigorously, even in highly nonstationary conditions, where histograms of channel open time are themselves difficult to interpret. Stated in words, the maximum likelihood estimate of $1/\beta$ is the total time that all channels are open, divided by the total number of closing transitions observed.

Statistical Properties of Maximum Likelihood Estimates

We have presented a method for computing the likelihood of experimental data, given the rate constants of a particular kinetic model. The next step is to maximize the likelihood with respect to parameters of choice. The parameters are usually some or all of the rate constants in the model, but they can also be some function of the rate constants, as we will demonstrate below. In general, one assumes initial values for the rate constants in the model and then calculates the likelihood from the data. The parameters of interest can be varied to find the values which maximize the likelihood. Search of the likelihood surface can be achieved by a variety of methods. For the purposes of this paper we adapted the nonlinear regression program BMDPAR (Ralston, 1981) to perform approximate Newton-Raphson iterations. Among the advantages of this particular program are that it requires relatively little user-generated computer code and that it permits bounds and linear inequality constraints on the parameters. We are currently investigating more widely accepted and faster variable metric methods for function optimization (Powell, 1978). Both the nonlinear regression approach and the variable metric approach require numerically computed partial derivatives of the logarithm of the likelihood.

Maximum likelihood estimates can be used to test smoothly parameterized subhypotheses of a general model by the likelihood ratio test (Rao, 1973). Comparison of a submodel with the general model is achieved by forming the ratio of the maximum likelihoods in each case. Twice the natural logarithm of this ratio is distributed as chi-square. The number of degrees of freedom equals the difference between the number of independent parameters in the two models. The mechanics of the test procedure will become clear in our examples below. Asymptotic standard errors for all estimates can be obtained from the second partial derivatives of the logarithm of the likelihood function with respect to each of the parameters (Rao, 1973). In general, the standard errors are inversely proportional to the square root of the number of independent observations. In other words, doubling the number of independent but identical experimental trials will reduce all standard errors by $\sim 2^{-0.5}$.

RESULTS

In this section we present two examples of our method for obtaining individual transition rates for kinetic models. The examples are analyses of simulations of single channel activity of idealized AChR (acetylcholine receptor) channels and sodium channels. They are meant to be typical of situations encountered in patch recording. We have used the continuous time Markov chains for analysis, except where noted.

Simulation of AChR Channel

Initial studies of AChR channels, using voltage clamp techniques, were adequately characterized by a simple model having only two kinetically distinct states, one open and one closed (for review see Steinbach, 1980). Although it is generally agreed that two agonist molecules must bind before the channel can open, the binding and unbinding are usually considered to be very rapid compared with a rate-limiting conformational change, which is the transition to the open state (Steinbach, 1981). In this view the closed states can be considered to be in equilibrium with one another. Recently however, the dwell time in at least one of the closed states was shown to be longer than previously suspected (Colquhoun and Sakmann, 1981; Dionne and Leibowitz, 1982). The following kinetic scheme has been proposed to explain these data:

$$C_1 \underset{\ell}{\overset{k}{\rightleftarrows}} C_2 \underset{\alpha}{\overset{\beta}{\rightleftarrows}} 0.$$

Scheme I

Note that in this scheme α, by usual convention, is the closing rate constant, and β is the opening rate constant. At the low agonist concentrations usually used in physiological experiments the pseudo–rate constant k is much smaller than all the others and lumps together all of the agonist binding steps. In this scheme a channel must wait a long time before reaching state C_2. On the average it dwells in this state for $T = (\beta + \ell)^{-1}$ s, after which it opens with probability $p = \beta/(\beta + \ell)$. If T is long enough to be detectable physiologically, openings will appear to occur in bursts as the channel moves back and forth between states C_2 and O, until the channel makes the transition between C_2 and C_1. Such bursts have recently been observed in single channel recording (Nelson and Sachs, 1979; Colquhoun and Sakmann, 1981; Dionne and Leibowitz, 1982). The average burst duration is the "apparent open time" that was determined from earlier macroscopic measurements and single channel recordings with lower bandwidths (Colquhoun and Hawkes, 1981).

Simulations of such bursts are shown in Fig. 2 *A*, using the model in Scheme I. Transitions between states were generated on a computer using exponentially distributed random numbers. Although the number of channels is generally not known experimentally, the agonist concentration can be lowered, so that overlapping events occur only rarely. We have assumed for our simulation that openings occur so rarely that bursts cannot be mistaken for two different channels opening contiguously. For purpose of analysis this is equivalent to examining the behavior of a single channel with three kinetic states.

In our simulation, we begin in the open state, and wait for the burst to end. This corresponds to data obtained by triggering the sampling on the opening of a channel (Neher and Steinbach, 1978; Horn and Patlak, 1980). By summing together 300 bursts, all of which begin at the same time, we have simulated a miniature end-plate current in Fig. 2 *A* (Neher and Steinbach, 1978). Notice that it appears nearly exponential, although theoretically we should observe two exponential components. The observed relaxation has a time constant equal to the average burst duration.

For data such as these, the transition rates ℓ, α, and β can be estimated by examining open times, closed times, and burst durations (Colquhoun and Hawkes, 1981). (The rate constant k is, by assumption, unmeasurably small.) We have used our maximum likelihood method to do the same for 357 bursts. We determined the likelihood in the following manner. Suppose we sample for T s, during which our simulation produces a current record consisting of three consecutive time periods, T_1, T_2, and T_3, where T_1 and T_3 represent times during which the channel is open. For arbitrary values of k, ℓ, α, and β, the continuous time likelihood of observing this specific current record is

$$\exp(-\alpha T_1)\,\alpha P_{C_2}(T_2)\,\beta \exp(-\alpha T_3),$$

where $P_{C_2}(T_2)$ is defined as the probability of being in state C_2, T_2 s after entering it, without ever leaving the closed states. This probability can be calculated numerically from the matrix exponential, as discussed previously. In this simple model an analytic solution can be obtained by the method of spectral expansion (Colquhoun and Hawkes, 1977). For our model the matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} = \begin{pmatrix} -k & k & 0 \\ \ell & -(\ell + \beta) & \beta \\ 0 & \alpha & -\alpha \end{pmatrix}.$$

Note that $P_{C_2}(t)$ is an entry of the matrix exponential $\mathbf{exp}(t\mathbf{Q}_{cc})$. For this model $\mathbf{Q}_{cc}$ is given by

$$\mathbf{Q}_{cc} = \begin{pmatrix} -k & k \\ \ell & -(\ell + \beta) \end{pmatrix}$$

and

$$P_{C_2}(t) = (0\ \ 1)\,\mathbf{exp}(t\mathbf{Q}_{cc}) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In this example

$$P_{C_2}(T_2) = (R_1 - R_2)^{-1}\,[(R_1 - k)e^{-R_1T_2} - (R_2 - k)e^{-R_2T_2}],$$

where

$$R_1, R_2 = [k + \ell + \beta \pm \sqrt{(k + \ell + \beta)^2 - 4k\beta}]/2.$$

Note that for our example $k \to 0$, in which case $P_{C_2}(T_2) \to \exp[-(\ell + \beta)T_2]$. The likelihood for the entire series of 357 current records is obtained by the product of the likelihoods of each independent record, or the Log(likelihood) by the sum of the logarithms of the likelihoods. The

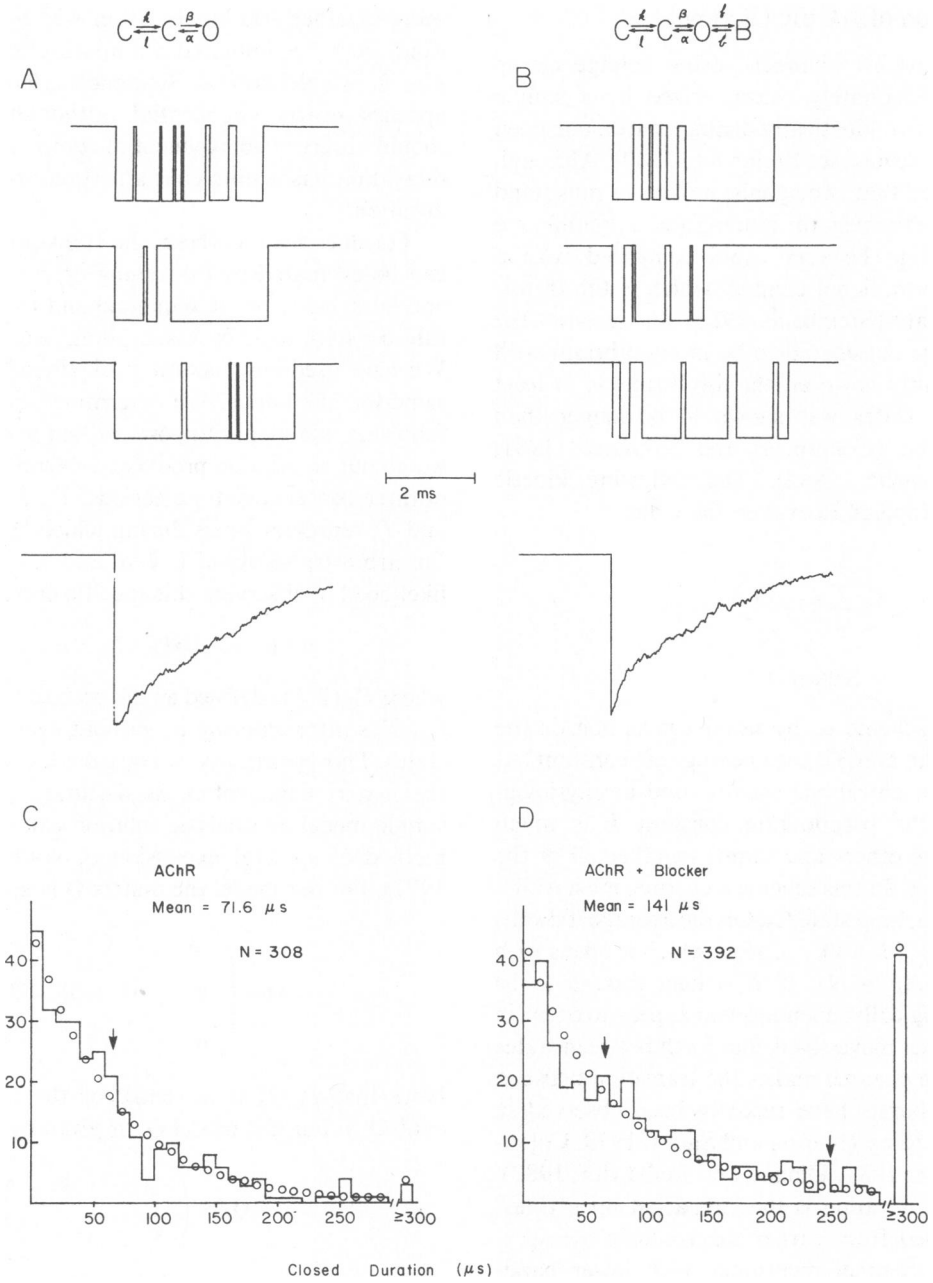

FIGURE 2 Simulations of acetylcholine receptor channels. (*A*) Three bursts of single channel activity are shown. As in other figures, the uppermost level is the zero current level, and downward deflections represent channel openings. These bursts were generated by the model shown above with values for the rate constants given in the text. The bursts were aligned so that the first openings for a sequence of 300 bursts began at the same time. Then these records were averaged to give the simulated miniature end-plate current below. The closed time histogram for the 300 bursts in *A* are plotted in *C*. The number of closed durations in these data were $N = 308$. The theoretical histogram for this model is a single exponential with a time constant of 67 μs, and is shown by open circles with the arrow depicting the theoretical time constant. By comparison, the mean closed duration of the simulated data was 71.6 μs. *B* and *D* are similar to *A* and *C*, except that an extra closed state has been included in the model. In this case, the theoretical histogram has two time constants of 67 and 250 μs. The theoretical mean closed time for this model is 128 μs, and the observed closed time was 141 μs.

procedure for maximizing the likelihood with respect to the parameters was discussed above.

When the simulated values of ℓ, α, and β were 5,000, 1,000, and 10,000 s^{-1}, respectively, our analysis produced estimated values of 5,172 ± 303, 1,063 ± 32, and 10,107 + 451 s^{-1}. The likelihood ratio test was applied to the likelihood values obtained from the estimated values and from the generating values of the above parameters. The chi-square value was 4.04 with 3 degrees of freedom ($P \simeq 0.25$).

Bursts of AChR channel openings and closings can also be caused by derivatives of lidocaine (Neher and Stein-

bach, 1978). These anesthetics block by a simple first-order reaction with the open state of the channel. The scheme for this reaction is as follows:

$$C_1 \underset{\ell}{\overset{k}{\rightleftarrows}} C_2 \underset{\alpha}{\overset{\beta}{\rightleftarrows}} O \underset{b}{\overset{f}{\rightleftarrows}} C_3 \text{ (blocked).}$$

Scheme II

The rate constant f is linearly dependent on the anesthetic concentration. The added complication in this scheme is that a closed interval in a burst can be due to a channel being either in C_2 or C_3. However, the closed states are indistinguishable in the current records. Simulations of Scheme II are shown in Fig. 2 *B*. If the average blocked time, i.e. $1/b$, is similar to the dwell time in state C_2, estimates of f and b become difficult using standard methods. For example one might compare the closed time histograms before and after adding anesthetic, as in Fig. 2, *C* and *D*. The success of this approach depends on the ability to separate the two exponential components in the presence of anesthetic. In our example we have deliberately chosen values of f and b such that the two components in the histogram are difficult to separate, which is obvious from the theoretical points in Fig. 2 *D*. We then estimated the transition rates f and b from these simulated data using our method.

The analysis of 300 bursts was of the discrete time Markov model discussed above. As an example of the method of calculation, suppose that the event E_i equals 1 when the channel is open, and 0 when it is closed. If we sample every Δt s, then the likelihood, L, of observing the sequence of events 1, 1, 0, 0, 1, 1 can be computed recursively using Eq. 7. For convenience let us arrange the states in the order shown in Scheme II, and we will denote compositions (0,1,0,0), (0,0,1,0), and (0,0,0,1) as a, b, and c, respectively. Note that a, b, and c correspond to the channel being in states C_2, O, or C_3, respectively. The initial composition, b, is taken as a certainty. Then,

$$P(F_{1b} \mid F_{0b}) = 1 - (\alpha + f)\,\Delta t,$$

and

$$L = P(F_{1b} \cap E_0 \cap E_1) = 1 - (\alpha + f)\,\Delta t.$$

Both a and c are consistent with the observed fact that the channel closes at epoch 2 Δt. Therefore we calculate both

$$P(F_{2a} \mid F_{1b}) = \alpha \Delta t,$$

and

$$P(F_{2c} \mid F_{1b}) = f \Delta t.$$

At this epoch the likelihood is the sum of all joint probabilities consistent with the observed events, or

$$L = P(F_{2a} \cap E_0 \cap E_1 \cap E_2) + P(F_{2c} \cap E_0 \cap E_1 \cap E_2),$$

where

$$P(F_{2a} \cap E_0 \cap E_1 \cap E_2) = [1 - (\alpha + f)\,\Delta t]\,\alpha \Delta t,$$

and

$$P(F_{2c} \cap E_0 \cap E_1 \cap E_2) = [1 - (\alpha + f)\,\Delta t]\,f \Delta t.$$

Further use of this recursion leads to the likelihood for the whole sequence of events, which is

$$[1 - (\alpha + f)\,\Delta t]\,\{f\Delta t\,(1 - b\,\Delta t) b\Delta t + \alpha\Delta t[1 - (\beta + \ell)\Delta t]\beta\Delta t\}[1 - (\alpha + f)\Delta t].$$

Notice that the ambiguity caused by multiple closed states introduces sums of probabilities into the likelihood calculation. This is due to the fact that the discrete time likelihood, in this case, is an estimate of a probability of mutually exclusive alternatives.

It was assumed that ℓ, α, and β were already known from other measurements and had the values 5,000, 1,000, and 10,000 s^{-1}. We sampled the data at intervals of 10^{-5} s. For simulated values of f and b of 500 and 4,000 s^{-1}, respectively, the maximum likelihood method gave estimates of 469 ± 44 and 3,708 ± 347 s^{-1}, converging after six iterations. The initial guesses of the parameters were the values used to generate the data. The natural logarithm of the likelihood ratio was 0.456, giving a chi-square value of 0.912 with 2 degrees of freedom ($P \simeq 0.6$).

The success of the maximum likelihood method depends greatly on the surface structure of the likelihood function. Local maxima could give spurious results, especially for poor initial guesses of parameters. In our experience this was usually not a problem. Even grossly inaccurate initial guesses tended to converge to the same estimates, although with a greater number of iterations. Using the previous example, eight iterations were required for convergence when values of 8,000 and 300 were used as initial values of f and b, respectively. However they converged to the same values. Fig. 3 plots the likelihood surface for this example as a function of the two parameters. It has a clear maximum for the best estimates found previously. While we have not examined in detail the likelihood surface for more complicated problems, we find in general that convergence occurs in relatively few iterations.

Simulation of Single Sodium Channel Currents

We have used the Hodgkin-Huxley (1952) model to simulate sodium channel currents in the absence of inactivation. The simplification of ignoring inactivation can be achieved experimentally by use of pronase or *N*-bromoacetamide (Armstrong et al., 1973; Oxford et al., 1978;

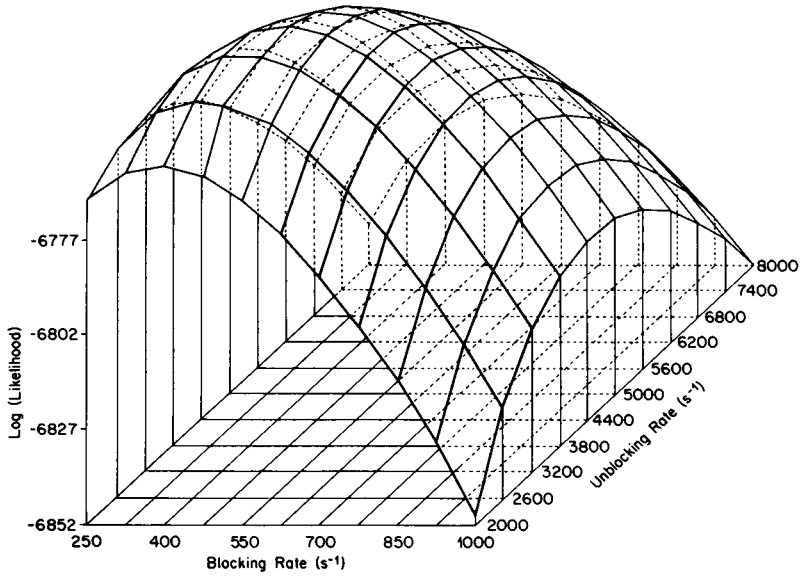

FIGURE 3 Likelihood surface for the four-state model of the acetylcholine receptor channel. The values of k, ℓ, α, and β were fixed at the values used to generate the simulated data, consisting of 357 bursts. The likelihood was calculated using the discrete Markov model, and is plotted as a function of f, the blocking rate constant, and b, the unblocking rate constant. The surface is smooth and shows a maximum near the values of f and b that were used to generate the data.

Oxford, 1981; Patlak and Horn, 1982). The model follows

$$C_1 \underset{\beta}{\overset{3\alpha}{\rightleftarrows}} C_2 \underset{2\beta}{\overset{2\alpha}{\rightleftarrows}} C_3 \underset{3\beta}{\overset{\alpha}{\rightleftarrows}} 0.$$

Scheme III

We simulated the current response to 165 voltage pulses of 30 ms from a patch with three identical and independent channels. The values of α and β for the simulation were 100 and 40 s^{-1}, respectively. For these values the steady state probability of a channel being open is 0.364, and the slowest overall time constant for activation is $(\alpha + \beta)^{-1} =$ 7.14 ms. Examples of simulated currents are shown in Fig. 4 *A*, and the averaged currents of 149 pulses is shown in Fig. 4 *B*. In this simulation, all three channels were simultaneously open in only 34 out of the 165 pulses. Two examples are shown in Fig. 4 *A*.

Let us assume that we know, from extensive macroscopic studies of sodium currents, that the kinetic scheme appropriate for these currents is as follows:

$$C_1 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C_2 \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} C_3 \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} 0$$

Scheme IV

We can use the simulated data and the maximum likelihood method to ask several questions about the kinetic model that generated these data. The questions we will ask are (*a*) What is the maximum likelihood estimate for the closing rate constant? (*b*) How well does a Hodgkin-

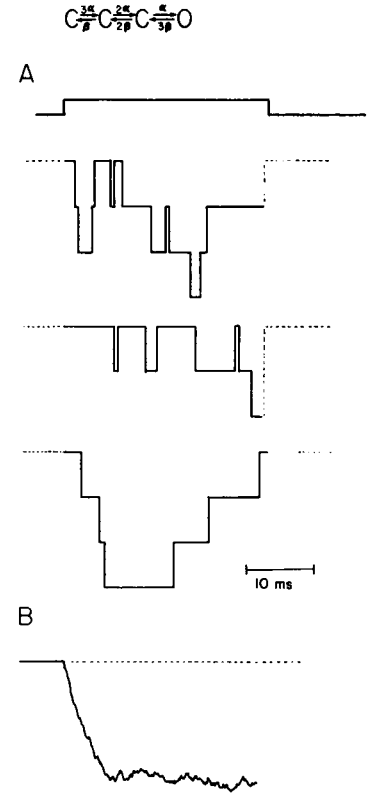

FIGURE 4 Simulations of sodium current. The model is shown above and is described in the text. *A* shows, at the top, a representative voltage step that activates the currents, shown below. Three channels were present in this simulation, and openings of the channels overlapped in some cases, as shown by three representative current records. The average response to 149 voltage steps is shown in *B*.

Huxley model, or variations on it, fit the data? and (*c*) What are the best estimates for all the transition rates?

Estimate of the Closing Rate Constant

The rate constant, k_{-3}, is simply estimated by use of Eq. 9. For the 165 records, this method produced an estimate of 114 s^{-1}, by comparison with the value of 120 s^{-1} used to generate the data. Doubling the number of records examined improved this estimate to 116 s^{-1}. In records such as these, where channel openings frequently overlap, k_{-3}cannot legitimately be estimated by averaging the open time for non-overlapping events. This procedure tends to preferentially select brief openings, because longer openings have a greater likelihood of overlapping with other openings. In our simulated data such a biased estimate of k_{-3} for 165 records was 273 s^{-1}, clearly larger than the generating value.

Variations on the Hodgkin-Huxley model

Using Scheme III we estimated α and β for these 165 records. Starting with values of 80 and 50 s^{-1}, respectively, six iterations were required for convergence at values of 94.5 ± 2.6 and 37.4 ± 2.6 s^{-1}. As might be expected, these two parameters were not significantly better at fitting the data than those used to generate the data. The chi-square value for the difference from the generating parameters was 4.20 with two degrees of freedom ($P \simeq 0.15$). (If we doubled the number of current records analyzed, the estimates of α and β improved to 96.6 ± 2.0 and 38.7 ± 1.5 s^{-1}, respectively).

Bezanilla and Armstrong (1975) noted that two variations of this model produce nearly identical macroscopic currents as those generated by the Hodgkin-Huxley model. In one variation, all the forward rate constants are equal and all the backward rate constants are equal. We examined the fit of our simulated data to this model using initial guesses of 120 s^{-1} for the forward rate constant and 110 s^{-1} for the backward rate constant. Best estimates of these two rate constants were 171 and 111 s^{-1}, respectively. This fit was significantly different from the model used to generate the data ($P \simeq 0.025$).

In the second variation, the rate constants have the opposite proportionality factors, i.e., $k_2 = 2k_1$, $k_3 = 3k_1$, $k_{-2} = 2k_{-3}$, and $k_{-1} = 3k_{-3}$. Using initial values of 120 and 110 s^{-1} for k_1 and k_{-3}, respectively, the best estimates for these rate constants were 121 and 102 s^{-1}, respectively. The chi-square value was 105.4 with 2 degrees of freedom, yielding a significant difference from the generating model ($P \simeq 0.001$). Thus, a statistical analysis allowed discrimination between these models, even though a deterministic analysis could not do this at one membrane potential.

We have also tried to fit these data with the m^2 model, which is the Hodgkin-Huxley model for two closed states leading to one open state. To do this without reducing the number of compositions, we used Scheme IV with very large k_1 and k_{-1}. This effectively lumps states C_1 and C_2 together kinetically. If k_1 also equals k_{-1}, then the probability of being in state C_2, conditional on being in either C_1 or C_2, is 0.5. Therefore, the rate of leaving these lumped states is reduced by 1/2. With this in mind we let $k_1 = k_{-1} = 5{,}000\ s^{-1}$, $k_2 = 4k_3$, and $k_{-3} = 2k_{-2}$. We then estimated k_3 and k_{-2} as α and β in the m^2 model. This model was also significantly different from the generating model ($P \simeq 0.025$). The results of this section are summarized in Table II.

Estimates for All Transition Rates

In our initial attempts to estimate all six rate constants simultaneously from our data, convergence often depended on the initial guesses for the parameters. Also, the estimates had very large standard errors. However, when the values used to generate the data were chosen as initial values, the convergence was rapid and the likelihood ratio test showed that the estimated values were statistically close to the generating values. This is hardly surprising and indicates that the likelihood surface for this model is somewhat complicated. To get better estimates from poor initial guesses we used more data. Instead of examining the current response to 165 voltage pulses with three channels, we examined the response of 483 voltage pulses with four channels. Note that this is not an inordinately large number of pulses for physiological experiments (e.g., see Sigworth, 1980).

TABLE II
MODELS OF SODIUM CHANNEL

	Estimates		2 × log (likelihood ratio)
	α	β	
$C \underset{\beta}{\overset{3\alpha}{\rightleftarrows}} C \underset{2\beta}{\overset{2\alpha}{\rightleftarrows}} C \underset{3\beta}{\overset{\alpha}{\rightleftarrows}} O$	94.5 ± 2.6	37.4 ± 2.6	4.20
$C \underset{\beta}{\overset{\alpha}{\rightleftarrows}} C \underset{\beta}{\overset{\alpha}{\rightleftarrows}} C \underset{\beta}{\overset{\alpha}{\rightleftarrows}} O$	171.4 ± 4.9	110.6 ± 4.0	12.20
$C \underset{3\beta}{\overset{\alpha}{\rightleftarrows}} C \underset{2\beta}{\overset{2\alpha}{\rightleftarrows}} C \underset{\beta}{\overset{3\alpha}{\rightleftarrows}} O$	121.3 ± 4.3	102.4 ± 5.4	105.40
$C \underset{\kappa}{\overset{\kappa}{\rightleftarrows}} C \underset{\beta}{\overset{4\alpha}{\rightleftarrows}} C \underset{2\beta}{\overset{\alpha}{\rightleftarrows}} O$ $\kappa >> \alpha,\beta$	75.9 ± 2.6	55.3 ± 2.9	11.56

Analysis of simulation of 165 pulses of 30 ms duration, using the Hodgkin-Huxley model with α = 100 s^{-1} and β = 40 s^{-1}. The total number of channels was 3. The estimates and standard errors are in units of s^{-1}.

We used, as initial guesses, the values obtained from the fit of one of the models tested above. Namely, we let $k_{-1} = k_{-2} = k_{-3} = 110\ s^{-1}$, and $k_1 = k_2 = k_3 = 170\ s^{-1}$. After 19 iterations convergence was reached. The estimated values for the transition rates are given in Table III. These estimates, judged by the likelihood ratio criterion, are excellent. By comparison with the values used to generate the data, it is immediately obvious that some estimates are better than others. The rate constants k_3 and k_{-3} are close to the generating values and have small standard errors. Estimates for the other four rate constants were statistically, but not absolutely, close to the generating values. Apparently the likelihood surface is relatively flat with respect to these parameters. This is to be expected, since the most information is obtained for rates of entering and leaving the open state.

We also tried constraining the parameters in order to improve the absolute values of the estimates for k_1, k_2, k_{-1}, and k_{-2}. On the basis of the above hypothesis testing, we might assume that the forward rate constants have decreasing magnitudes from left to right, and that the backward rate constants have the opposite trend. Although both arbitrary and post-hoc, these linear constraints did improve the absolute values of the estimates for the rate constants and reduced the standard errors, as shown in Table III, without having a large effect on the likelihood.

In spite of the large standard errors in some of these estimates, it is still possible to use them for hypothesis testing using the likelihood ratio test. For example, statistical tests of the above models against the model obtained from estimates of all the rate constants produce the same conclusion as reached when using the likelihood of the generating model. The conclusion is that, of the models examined, only the m^3 Hodgkin-Huxley model produces a satisfactory fit to the data. It is important to restate that the number of degrees of freedom for the chi-square test is different from our previous examples. In this case, if we estimate α and β for the Hodgkin Huxley model, and compare our model with the model obtained from estimating all rate constants simultaneously, the degrees of freedom are $6 - 2 = 4$, instead of 2 degrees of freedom used when testing our model against the generating values of the rate constants.

TABLE III
ESTIMATES OF ALL RATE CONSTANTS IN HODGKIN-HUXLEY MODEL

$$C \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} C \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} C \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} 0$$

	Generating	Constrained	Unconstrained
k_1	300	271.3 ± 86.3	218.5 ± 151.9
k_2	200	257.3 ± 79.8	333.7 ± 156.6
k_3	100	106.6 ± 9.98	105.4 ± 12.6
k_{-1}	40	102.6 ± 28.5	123.7 ± 121.6
k_{-2}	80	102.6 ± 28.5	109.2 ± 89.9
k_{-3}	120	118.9 ± 2.21	118.9 ± 0.79
Log(like):	6832.70	6834.20	6834.46

Analysis of simulation of 483 pulses of 30 ms duration in a population of four channels. The rate constants used to generate the simulated data are shown in the leftmost column. Estimates and standard errors are listed for the constrained and unconstrained cases (see text). The natural logarithm of the likelihood is shown for each case. All rate constants are in units of s^{-1}.

Strategy for Experimental Data

For experimental data, it is important to know as much as possible about the kinetic model underlying the single channel behavior before using our method of analysis. For this purpose it is desireable to construct all interpretable histograms under as many conditions as possible. From these data, and from macroscopic and noise measurements, one can then propose a kinetic model for the data that incorporates the number and nature of the various states and, if possible, initial estimates of rate constants for the model. If the data are nonstationary, one should try to determine the probabilistic starting condition for each channel. In general it is also necessary to know the number of channels in the data. Under these conditions one can try to estimate all the rate constants by maximizing the likelihood function with respect to these rates. In some cases, as above, it may be desirable to try subhypotheses of the model with fewer parameters. It is important, after obtaining estimates of all the rate constants, to use the model to generate theoretical histograms for the data (see Colquhoun and Hawkes, 1981). Marked discrepancies between the theoretical and observed histograms is an indication either that the estimates are not accurate or that the model is incorrect. When a good fit is obtained for the rate constants, this model and its associated likelihood function are treated as the model that generated the data. That is, subhypotheses of this model are tested against it by the likelihood ratio test. The general model has as many independent parameters as rate constants estimated, as opposed to zero when the generating model is known. The number of degrees of freedom for chi-square is the difference between the number of independent parameters in the two models being compared. One practical consideration is that more complicated models, i.e., models with more states, generally require more data for estimates of parameters. This becomes obvious from difficulties in convergence. The exact amount of data necessary depends on details of the model and on the desired precision of the parameter estimates.

DISCUSSION

We have described here a method for determining kinetic constants for Markov models from single channel data. The method is applicable in both patch recording and artificial bilayer studies. We have limited our analysis to linear kinetic schemes that fulfill the criteria for time

homogeneous Markov processes with fewer than six states. We have also limited our analysis to models with only two conductance states, open and closed. This is suitable for a wide variety of channel types (for some exceptions, see Hamill and Sakmann, 1981; Labarca and Miller,1981; Latorre and Alvarez, 1981). Our method with some modifications could be applied to branching or cyclic kinetic schemes, or to models with more than two conductance states. At this time it is not clear to us whether all kinetic parameters are identifiable in complex schemes. It may be useful to bear in mind the experience gained on similar identifiability questions in compartment models, econometrics, and systems engineering (Nguyen and Wood, 1982). The power of the technique is that it is able to estimate rate constants under conditions in which this would be very difficult with presently available methods. These conditions encompass many situations typically encountered during experiments and include nonstationarity, multiple kinetic states, multiple channels, and finite records. The method also has the advantage of using all the data at one time. This eliminates some of the strategy decisions, such as which histograms to construct and how to combine the information they yield.

This is not the place to enter into a detailed discussion of the classical technique of maximum likelihood (see Rao, 1973, for references). Suffice it to say that the maximum likelihood estimator is asymptotically unbiased and consistent. It is also efficient in that it has a variance at least as small as any other estimator. The technique generally produces good estimates in all but the most pathological cases, and Fig. 3 showed that the likelihood surface in one somewhat complicated model was quite regular. One of the main virtues of this technique is that standard errors can be computed for all estimates. Furthermore, with the estimates it is possible to test smoothly parameterized subhypotheses of a general model, as we have shown above. It is even possible to test hypotheses for the number of kinetic states by lumping states together, as we demonstrated above.

The power of our approach can perhaps be best appreciated in the Hodgkin-Huxley simulation we have used. In this model, especially in the case of multiple channels, the histogram method fails miserably for two reasons. First, histograms of dwell times at any conductance level are highly nonstationary. Perhaps the only histogram which is generally reliable and meaningfull is the histogram of latency between the onset of a voltage pulse and the first channel opening (Fukushima, 1981; Horn et al., 1981; Patlak and Horn, 1982). This is mainly due to the fact that the starting condition is the same for each channel. Even a histogram of single channel open time is distorted by the nonstationary probability of another channel opening. The second difficulty is that, even at steady state, histograms have many exponential components whose rates are very similar. The similarity of these rates explains why so many different kinetic schemes can account for the same macroscopic data. In spite of these complexities, our method can provide estimates of all six rate constants in the model, starting from relatively poor initial guesses of the parameters. The power of this approach answers some concerns about the ability of experiments, which examine conductance alone, to provide information about transitions between closed states (Armstrong, 1981). In the Hodgkin-Huxley simulation all of the rate constants could be estimated, not only the ones leading to or from the open state, although the latter rate constants were better determined.

In spite of the power of this type of analysis, several problems limit its usefulness. Perhaps the foremost problem is selecting a reasonable kinetic model to use. This generally requires a knowledge of the number of states, how they are connected, which are open or closed, and the initial condition for the channels (i.e., the initial probability of a given channel being in each of the possible kinetic states). Selecting the correct model is largely a matter of experimental intuition and the availability of other information, such as detailed macroscopic data, noise measurments, histograms of single channel currents, and gating currents.

Another problem with this approach is that we must know the number of channels present and the zero current level at which no channels are open. This is not always possible. In some cases a small population of channels fluctuate between open and closed without reaching the zero current level. It may not be possible, in the absence of model dependent assumptions, to determine the true zero current level in this case. If the zero level is known, it is possible to estimate the number of channels by use of the binomial distribution (Patlak and Horn, 1982). However, this is only feasible when the probability of overlapping events is reasonably high. Even without this information it may be possible to extract useful information for at least part of a kinetic scheme. For models that have a single open state, it is often possible to make an unambiguous estimate of the closing rate constant, without knowing either the number of channels or details of the kinetic scheme. Other rate constants can be estimated under conditions where only one of many channels is open at a time, as we have shown above for the case of the acetylcholine receptor channel.

One of the main practical difficulties with our method is the time required for computations. This time is mainly determined by the number of compositions in the model, which in turn depends on the number of channels and the number of kinetic states. Usually, the estimates of rate constants in the models we examined required 6–12 h of computing time. For the most difficult problem we have examined here, four states and four channels, calculation of the likelihood for 4,096 transitions, using the continuous Markov model, required 90 min on our laboratory PDP 11/34 computer(Digital Equipment Corp., Marlboro, MA). This calculation was reduced to ~12 min when a

floating-point processor was added to the system. Using the latter system the estimation of all six rate constants required nearly 5 d of continuous time on our computer. The computation time is reduced by $\sim 1/2$ if the number of channels is reduced to three. The inescapable conclusion is that, as experimentalists, we must attempt to reduce the number of channels under observation to only one or two, and also to manipulate the experimental conditions, if possible, to reduce the number of kinetic states to a minimum. We should point out that selecting experiments with fewer channels has a price. Although the computation time is reduced, the amount of information for the same number of transitions is significantly reduced, by comparison with an experiment with more channels. This will result in poorer estimates of parameters and larger standard errors.

Besides the time required for computation, our method has other limitations. We mention a few of them here. We have restricted our analysis to a kinetically homogeneous population of channels. This is not always the case (Clark and Adams, 1981; Hamill and Sakmann, 1981). In some experiments the homogeneity is not known (Jackson et al., 1982). Another limitation, which is common to all analytical methods, is the frequency response of the recording system. Obviously it is impossible to estimate rate constants faster than the sampling rate. In practice we have found that estimates are not affected when the data are sampled at intervals of Δt, where

$$\Delta t < (10nk_i)^{-1},$$

where n is the number of channels and k_i is the largest rate constant in the model.

Another experimental problem is drift, or "rundown." This could be seen as a gradual decrease in the number of functional channels over the period of observation. Our method of analysis is strictly dependent on a knowledge of the number of channels (with the exceptions noted above). Therefore it is essential to carefully monitor the preparation for rundown. One simple way to do this is to observe whether the overall reponse (e.g., in averaged records) decreases with time.

Many of the difficulties we have noted will be shared by other analytical techniques. We believe that our method of estimating rate constants will be a useful tool in the the arsenal of techniques available to physiologists for analyzing single channel data.

We thank Drs. Vincent Dionne, Mauri Krouse, and Elliot Landaw for helpful discussions, Dr. Haru Ohmori for carefully reading the manuscript, and Bradford A. Lubell for help with computer programming.

Supported by National Science Foundation grant PCM 76-20605, and National Institutes of Health grants NS 703-01, KO4 HD00307, and NS 186-08.

Received for publication 29 May 1982 and in final form 10 March 1983.

REFERENCES

Armstrong, C. M. 1981. Sodium channels and gating currents. *Physiol. Rev.* 61:644–683.

Armstrong, C. M., and F. Bezanilla. 1977. Inactivation of the sodium channel. II. Gating current experiments. *J. Gen. Physiol.* 70:567–590.

Armstrong, C. M., F. Bezanilla, and E. Rojas. 1973. Destruction of sodium conductance inactivation in squid axons perfused with pronase. *J. Gen. Physiol.* 62:375–391.

Bean, R. C., W. C. Shepherd, H. Chan, and J. Eichner. 1969. Discrete conductance fluctuations in lipid bilayer protein membranes. *J. Gen. Physiol.* 53:741–757.

Bezanilla, F., and C. M. Armstrong. 1975. Properties of the sodium channel gating current. *Cold Spring Harbor Symp. Quant. Biol.* 40:297–304.

Clark, R. B., and P. R. Adams. 1981. Two types of ACh receptor channels in *Xenopus* muscle membranes. *Int. Biophys. Congress.* 187. (Abstr.)

Colquhoun, D., and A. G. Hawkes. 1977. Relaxation and fluctuations of membrane currents that flow through drug-operated channels. *Proc. R. Soc. Lond. B. Biol. Sci. 199:231–262.*

Colquhoun, D., and A. G. Hawkes. 1981. On the stochastic properties of single ion channels. *Proc. R. Soc. Lond. B. Biol. Sci. 211:205–235.*

Colquhoun, D., and B. Sakmann. 1981. Fluctuations in the microsecond time range of the current through single acetylcholine receptor ion channels. *Nature (Lond.).* 294:464–466.

Conti, F., and E. Wanke. 1975. Channel noise in nerve membranes and lipid bilayers. *Q. Rev. Biophys.* 8:451–506.

Cox, D. R., and H. D. Miller. 1965. The Theory of Stochastic Processes. Chapman & Hall Ltd., London.

DeFelice, L. J. 1981. Introduction to Membrane Noise. Plenum Publishing Corp.,New York.

Dionne, V. E., and M. D. Leibowitz. 1982. Acetylcholine receptor kinetics. A description from single channel currents at snake neuromuscular junctions. *Biophys. J.* 39:253–261.

Feller, W. 1968. An Introduction to Probability Theory and Its Applications. Vol. I. John Wiley & Sons, Inc., New York. 38.

Feller, W. 1971. An Introduction to Probability Theory and Its Applications. Second ed. Vol. II. John Wiley & Sons, Inc., New York 19.

Fenwick, E. M., A. Marty, and E. Neher 1982. Sodium and calcium channels in bovine chromaffin cells. *J. Physiol. (Lond.).* In press.

Fukushima, Y. 1981. Identification and kinetic properties of the current through a single Na^+ channel. *Proc. Natl. Acad. Sci.* 78:1274–1277.

Hamill, O. P., A. Marty, E. Neher, B. Sakmann, and F. J. Sigworth. 1981. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv. Gesamte Physiol. Menschen Tiere.* 391:85–100.

Hamill, O. P., and B. Sakmann. 1981. Multiple conductance states of single acetylcholine receptor channels in embryonic muscle cells. *Nature (Lond.).* 294:462–464.

Hodgkin, A. L., and A. F. Huxley. 1952. A quantitative description of the membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.).* 117:500–544.

Horn, R., J. Patlak, and C. Stevens. 1981. Sodium channels need not open before they inactivate. *Nature (Lond.).* 291:426–427.

Horn, R., and J. Patlak. 1980. Single channel currents from excised patches of muscle membrane. *Proc. Natl. Acad. Sci.* 77:6930–6934.

Isaacson, E., and H. B. Keller. 1966. Analysis of Numerical Methods. John Wiley & Sons, Inc., New York.

Jackson, M. B., H. Lecar, V. Askanas, and W. K. Engel. 1982. Cholinergic single channel currents in cultured human muscle. *J. Neurosci.* 2:1465–1473.

Karlin, S., and H. M. Taylor. 1975. A First Course in Stochastic Processes. Second ed. Academic Press, Inc., New York.

Labarca, P., R. Coronado, and C. Miller. 1980. Thermodynamic and kinetic studies of the gating behavior of a K^+-selective channel from the sarcoplasmic reticulum membrane. *J. Gen. Physiol.* 76:397–424.

Labarca, P., and C. Miller. 1981. A K^+-selective, three-state channel from fragmented sarcoplasmic reticulum of frog leg muscle. *J. Membr. Biol.* 61:31–38.

Latorre, R., and O. Alvarez. 1981. Voltage-dependent channels in planar lipid bilayer membranes. *Physiol. Rev.* 61:78–150.

Liou, M.L. 1966. A novel method of evaluating transient response. *Proc IEEE.* 54:20–23.

Moler, C., and C. Van Loan. 1978. Nineteen dubious ways to compute the exponential of a matrix. *SIAM (Soc. Int. Appl. Math.) Rev.* 20:801–836.

Neher, E., and B. Sakmann. 1976. Single channel currents recorded from membrane of denervated frog muscle fibres. *Nature (Lond.).* 260:799–802.

Neher, E., and J. H. Steinbach. 1978. Local anaesthetics transiently block currents through single acetylcholine-receptor channels. *J. Physiol. (Lond.).* 277:153–176.

Neher, E., and C. F. Stevens. 1977. Conductance fluctuations and ionic pores in membranes. *Annu. Rev. Biophys. Bioeng.* 6:345–381.

Nelson, D.J., and F. Sachs. 1979. Single ionic channels oberved in tissue-cultured muscle. *Nature (Lond.).* 282:861–863.

Nguyen, V. V., and E. F. Wood. 1982. Review and unification of linear identifiability concepts. *SIAM (Soc. Int. Appl. Math.) Rev.* 24:34–51.

Nijenhuis, A., and H. S. Wilf. 1978. Combinatorial Algorithms. Second ed. Academic Press, Inc., New York. 46–51, 99–116.

Oxford, G. S. 1981. Some kinetic and steady-state properties of sodium channels after removal of inactivation. *J. Gen. Physiol.* 77:1–22.

Oxford, G. S., C. H. Wu, and T. Narahashi. 1978. Removal of sodium channel inactivation in squid giant axons by *N*-bromoacetamide. *J. Gen. Physiol.* 71:227–247.

Patlak, J. 1983. The information content of single channel data. *In* Membranes, Channels, and Noise. R. Eisenberg, M. Frank, and C.F. Stevens, editors. Plenum Publishing Corp., New York. In press.

Patlak, J., and R. Horn 1982. The effect of *N*-bromoacetamide on single sodium channel currents in excised membrane patches. *J. Gen Physiol.* 79:333–351.

Powell, M. J. D. 1978. A fast algorithm for nonlinearity constrained optimization calculations. *In* Numerical Analysis, Dundee 1977. A. Watson, editor. Springer-Verlag, Berlin.

Ralston, M. 1981. Derivative-free nonlinear regression. *In* BMDP Statistical Software 1981. W.J. Dixon, editor. University of California Press, Berkeley. 305–329.

Rao, C.R. 1973. Linear Statistical Inference and Its Applications. Second ed. John Wiley & Sons, Inc., New York.

Sigworth, F.J. 1980. The variance of sodium current fluctuations at the node of Ranvier. *J. Physiol. (Lond.).* 307:97–129.

Sigworth, F.J. 1981. Covariance of nonstationary sodium current fluctuations at Node of Ranvier. *Biophys. J.* 34:111–132.

Steinbach, J.H. 1981. Activation of nicotinic acetylcholine receptors. *In* The Cell Surface and Neuronal Function. C.W. Cotman and G.L. Nicolson, editors. Elsevier, Amsterdam. 119–156.