Rapid Kinetic Analysis of Multichannel Records by a Simultaneous Fit to All Dwell-Time Histograms

László Csanády
Laboratory of Cardiac/Membrane Physiology, The Rockefeller University, 1230 York Avenue, Box 297, New York, New York 10021, USA

ABSTRACT A method is presented for rapidly extracting single-channel transition rate constants from patch-clamp recordings containing signals from several channels. The procedure is based on a simultaneous fit of the observed dwell-time distributions for all conductance levels, using a maximum likelihood approach. This algorithm allows estimation of single-channel rate constants in cases where more advanced methods may be impractical because of their extremely long computational time. A correction is included for the limited time resolution of the recording system, according to theory developed by Roux and Sauvé (*Biophys. J.* 48:149–158, 1985), by accounting for the impact of undetected transitions on the dwell-time distributions, and by introducing an improved practical implementation of a fixed dead time for the case of more than one channel. This feature allows application of the method to noisy data, after filtering. A computer program implementing the method is tested successfully on a variety of simulated multichannel current traces.

INTRODUCTION

Gating of single ionic channels can be modeled by transitions among a finite number of conducting and nonconducting states, connected in some specified pattern. At any moment, a single channel will be in one of these states, and can undergo transitions into adjacent ones in a stochastic fashion. Such a system is conveniently described by a continuous-time Markov process (Colquhoun and Hawkes, 1977, 1981). Patch-clamp recording allows measurement of the conductance of individual ion channels. Because more than one state of a channel may be characterized by the same conductance, such states can be grouped into classes, transitions within which remain undetected. The observable signal is modeled by an aggregated Markov process. The challenge for the experimenter is to identify the model underlying the observed sequence of transitions, and to estimate the rate constants.

The classic approach to kinetic analysis of single-channel data consists of constructing dwell-time histograms for the open and closed times of a single channel, and then fitting them with exponential functions (Colquhoun and Sigworth, 1995). This approach is widely used because of its simplicity and relatively small computational demand.

In many cases, however, it may be impractical to restrict analysis to single-channel records. If the channels of interest tend to cluster, for example, most patches may contain multiple channels. Also, if the open probability is low, long records are required to amass enough events from a single channel. In some cases, it might be hard to maintain steady-state conditions for a sufficiently long time to collect enough events from a single channel.

Received for publication 25 May 1999 and in final form 2 November 1999.
Address reprint requests to László Csanády, The Rockefeller University, Lab of Cardiac/Membrane Physiology, Box 297, 1230 York Ave., NY, NY 10021. Tel.: 212-327-8617; Fax: 212-327-7589; E-mail: csanadl@rockvax.rockefeller.edu.

0006-3495/00/02/785/15 $2.00

As shown by Blunck et al. (1998), if the gating scheme is already known or assumed, data from patches containing multiple channels can be analyzed by a simultaneous fit of the dwell-time histograms for all the conductance levels. These authors used a least-squares fit to the observed set of histograms and showed that, in many cases, reasonable estimates of the parameters can be obtained after only short processing times. Unfortunately, the scatter of the estimates was considerable, and the approach failed for noisy data whenever the signal-to-noise ratio (SNR) was too low for correct event detection.

This paper presents an approach similar to the one introduced by Blunck et al. (1998), but with improvements that result in broader applicability and enhanced reliability. First, instead of least-squares fitting, which produces biased estimates depending on whether the fits are made to linear, semi-log or log–log histograms, a simultaneous maximum likelihood fit is made to the set of dwell-time histograms obtained from all conductance levels. Second, an effective correction for missed events is included, based on theory developed by Roux and Sauvé (1985) for a single channel, together with a practical implementation of a fixed dead time for multichannel records. This feature permits use of a range of filter settings, and hence, effective analysis of noisy data with fairly low SNR. Good results were obtained even with relatively small numbers of events (typically hundreds, for all but extreme ratios of rate constants).

A drawback of one-dimensional histogram fitting is that the information contained in cross correlations between neighboring events is ignored. As a result, the number of extractable parameters is limited, and the method cannot distinguish between alternative connectivities. More powerful methods exploit these cross correlations. Maximum likelihood fitting of the whole time series, first introduced by Horn and Lange (1983), is based on calculating the likelihood of an observed sequence of sample points (Horn and Lange, 1983) or dwell times (e.g., Qin et al., 1996, 1997), given a kinetic model and a set of rate constants. The

likelihood is then maximized with respect to the rate constants. However, the computational task involved is considerable, even if the likelihood is computed over dwell times only, and increases dramatically with the number of states in the model. Processing time and required memory space also increase with the length of the record. Alternatively, correlations are explored by two dimensional dwell-time histograms for pairs of adjacent events (Magleby and Song, 1992; Rothberg et al., 1997). The above maximum likelihood approaches can be applied directly to multiple channels by treating the system as a single large Markov scheme (Horn and Lange, 1983; Qin et al., 1996). However, even for simple single-channel models, the size of the compound scheme grows rapidly with channel number, making those computations extremely slow for more than 3 channels, especially for long records.

As for all histogram fitting methods, the approach presented here assumes that the gating scheme is specified at the outset, because correlations remain unused. However, for a given model, the reliability of the fit is ensured by the wealth of information exploited by considering the set of all dwell-time distributions. Attractively, computational time and required memory space do not depend on the length of the data while the quality of the fit improves with more events.

Suitable applications are records with many active channels obeying fairly simple, but nontrivial, gating schemes, where more advanced routines, like fitting the event sequence, become impractical. In reality, most ion channels have very complex gating schemes (e.g., Weiss and Magleby, 1989; Vandenberg and Bezanilla, 1991; Hoshi et al., 1994; Zagotta et al., 1994a,b; Schoppa and Sigworth, 1998a,b,c). Frequently, however, gating schemes can be simplified by pooling certain states while still retaining many essential features.

The general theory underlying the approach is developed first, then the method is extensively tested under a variety of conditions.

GENERAL THEORY

A many-channel system as a single Markov system

It is well known that a patch with multiple channels can be treated as a single Markov system (e.g., Horn and Lange, 1983; Qin et al., 1996). One state of this large system (macroscopic state) is a particular pattern of partitioning the individual channels among the single-channel states (microscopic states).

A simple important case, solved by Horn and Lange (1983), is when N channels are assumed to be identical and independent. If there are n microscopic states for a single channel, a macroscopic state of the system is represented by a vector of n entries, the ith component of which specifies the number of individual channels in the ith single-channel state. The entries of the macroscopic state vector sum to N. The number of macroscopic states is equal to $\binom{n+N-1}{N}$, the number of ways in which N can be divided into n parts. The transition rate from state $(N_1, \ldots, N_i, \ldots, N_j, \ldots, N_n)$ to state $(N_1, \ldots, N_i - 1, \ldots, N_j + 1, \ldots N_n)$ is given by

$$N_i \cdot r_{ij}, \tag{1}$$

where r_{ij} is the transition rate constant of a single channel from state i to j. Macroscopic states are numbered such that those belonging to the same conductance level are clustered, allowing partitioning of the transition matrix $\mathbf{Q}$ into submatrices relevant to each conductance level.

If the channels are distinguishable, the system can be described by a (usually much bigger) state space of n^N dimensions (see Qin et al., 1996). The concepts of this paper are presented for N identical and independent channels, but should be equally applicable to the more general case.

Dwell-time distributions of the various conductance levels

Given a Markov system described by the matrix $\mathbf{Q}$, distributions of the dwell times in any subset of the states can be calculated (Colquhoun and Hawkes, 1981). To calculate the dwell-time distribution of a conductance level k, a subset $\{k\}$ is defined as the set of all macroscopic states in which k channels are open, $\{\bar{k}\}$ as the set of all other macroscopic states. The survivor functions, defined as $\mathrm{surv}_{\{k\}}(t) = P(\text{no leaving from } \{k\} \text{ before time } t | \text{ entered } \{k\} \text{ at time } 0)$, are given by

$$\mathrm{surv}_{\{k\}}(t) = \mathbf{p}_k(0)^T e^{\mathbf{Q}_{kk}t}\mathbf{1}_k, \tag{2}$$

where $\mathbf{Q}_{kk}$ is the submatrix of $\mathbf{Q}$ describing transitions within $\{k\}$, $\mathbf{1}_k$ is the summation vector matching $\{k\}$ in dimensions. $\mathbf{p}_k(0)^T$ (matching $\{k\}$ in dimensions) is the initial row vector of probabilities of entering $\{k\}$ via any one of its component (macroscopic) states, given by

$$\mathbf{p}_k(0)^T = \frac{\mathbf{p}_{\bar{k}}(\infty)^T\mathbf{Q}_{\bar{k}k}}{\mathbf{p}_{\bar{k}}(\infty)^T\mathbf{Q}_{\bar{k}k}\,\mathbf{1}_k}, \tag{3}$$

where submatrix $\mathbf{Q}_{\bar{k}k}$ describes transitions from $\{\bar{k}\}$ to $\{k\}$, $\mathbf{p}_k(\infty)^T$ and $\mathbf{p}_{\bar{k}}(\infty)^T$ are row vectors of steady-state occupancy probabilities for macroscopic states in $\{k\}$ and $\{\bar{k}\}$ respectively. ($\mathbf{p}^T$ in general denotes the row vector corresponding to the column vector $\mathbf{p}$.) For binned maximum likelihood fitting (see Sigworth and Sine, 1987) it is convenient to use the survivor function, because the probability of a dwell time falling into a bin $t_i \le$ dwell time $< t_{i+1}$ is given by $\mathrm{surv}(t_i) - \mathrm{surv}(t_{i+1})$.

Simultaneous maximum likelihood fitting of a set of histograms

Binned maximum likelihood fitting (Sigworth and Sine, 1987) consists of a search for the set of kinetic parameters that maximizes the likelihood that a series of dwell times in one particular conductance level yields the experimentally observed dwell-time histogram. Here, this approach is extended to the set of dwell-time histograms constructed for all the conductance levels, which can be thought of as one generalized histogram, each bin of which represents a range of dwell times for a certain conductance level. Figure 1 illustrates this concept, showing the generalized dwell-time histogram for a simulated current trace (described in more detail later, in Fig. 7 *A*), from 3 channels. The fit (*solid black lines*) was found by the method described in this paper.

The likelihood function for an events list is defined as the likelihood of obtaining the observed generalized histogram (Fig. 1) given a model and a set of rate constants. The likelihood estimate is based on the calculation of the probability that an observed event belongs to level k and falls into the ith bin of the level k dwell-time histogram:

$$P(k, i) = P(\text{event is level } k \text{ and } t_{k,i} \le \text{dwell time} < t_{k,i+1}).$$

This is easily obtained using conditional probabilities,

$$P(k, i) = \text{pctlvl}(k) \cdot p_{ki}, \tag{4}$$

where pctlvl(k) specifies what fraction of all observed events is expected to be at level k,

$$\text{pctlvl}(k) = P(\text{observed event is level } k) = \frac{\mathbf{p}_{\bar{k}}(\infty)^{T}\mathbf{Q}_{\bar{k}\bar{k}}\mathbf{1}_{k}}{\sum_{j=0}^{N}\mathbf{p}_{\bar{j}}(\infty)^{T}\mathbf{Q}_{\bar{j}\bar{j}}\mathbf{1}_{j}}, \tag{5}$$

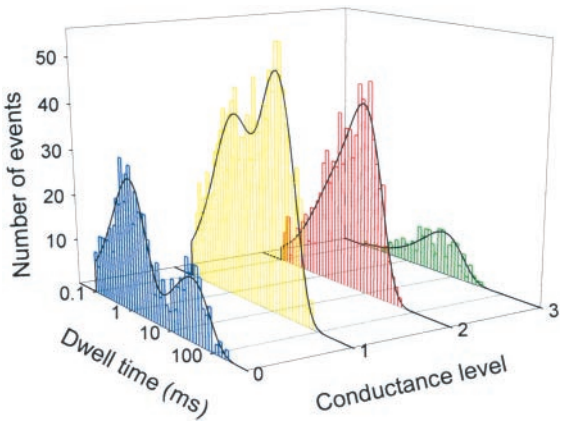

FIGURE 1 Set of dwell-time histograms (and fit, *solid black lines*) for all conductance levels of a simulated current record from 3 channels obeying the C–O–B scheme (see text for details). Rate constants (in s^{-1}) used for simulation: $r_{CO} = 10$, $r_{OC} = 20$, $r_{OB} = 40$, and $r_{BO} = 1000$. Estimates from fit: $r_{CO} = 10$, $r_{OC} = 20$, $r_{OB} = 38$, and $r_{BO} = 918$. Dead time: 0.1 ms (see text).

and p_{ki} is the fraction of all observed level k events predicted to fall into the ith bin,

$$p_{ki} = P(t_{k,i} \le \text{dwell time} < t_{k,i+1} | \text{event is level } k) = \text{surv}_{\{k\}}(t_{k,i}) - \text{surv}_{\{k\}}(t_{k,i+1}). \tag{6}$$

If the level k histogram contains r_k bins and $n_{k,i}$ is the occupancy of the ith bin, the likelihood function can be calculated as

$$L(\Theta) = \prod_{k=0}^{N}\prod_{i=1}^{r_k} P(k, i)^{n_{k,i}} = \prod_{k=0}^{N}\prod_{i=1}^{r_k} [\text{pctlvl}(k) \cdot p_{k,i}]^{n_{k,i}}, \tag{7}$$

where Θ represents the model and the set of rate constants. Both pctlvl(k) and $p_{k,i}$ are functions of Θ. In practice, the logarithm of the likelihood is used, given by

$$LL(\Theta) = \sum_{k=0}^{N}\sum_{i=1}^{r_k} n_{k,i} \cdot \ln p_{k,i} + \sum_{k=0}^{N} n_k \cdot \ln \text{pctlvl}(k), \tag{8}$$

where n_k is the total number of events in the level k histogram: $n_k = \sum_{i=1}^{r_k} n_{k,i}$.

Sometimes it is useful to exclude some events from binning by specifying lower and upper bin limits $t_{k,min}$ and $t_{k,max}$ for each conductance level k (see Sigworth and Sine, 1987). This will alter the scaling factors for the different levels, which can be compensated for by defining quantities analogous to those in Eqs. 4, 5, and 6, conditional on being binned. Following the derivation in Appendix A, the corrected log likelihood function then becomes

$$LL'(\Theta) = \sum_{k=0}^{N}\sum_{i=1}^{r_k} n_{k,i} \cdot \ln p_{k,i} + \sum_{k=0}^{N} n_k \cdot \ln \text{pctlvl}(k) - n_t \cdot \ln\left[\sum_{k=0}^{N} \text{pctlvl}(k) \cdot P_{k,\text{binned}}\right], \tag{9}$$

where

$$P_{k,\text{binned}} = P(\text{event is binned} | \text{event is level } k) = \text{surv}_{\{k\}}(t_{k,\min}) - \text{surv}_{\{k\}}(t_{k,\max}), \tag{10}$$

and

$$n_t = \sum_{k=0}^{N} n_k \quad \text{is the total number of binned events.}$$

Compared to the original expression for $LL(\Theta)$ in Eq. 8, the introduction of binning limits results in only one additional term, even though the binning limits may be different for each level. pctlvl(k), $p_{k,i}$, and $P_{k,\text{binned}}$ depend on the set of parameters Θ through Eqs. 5, 6, and 10, respectively.

The approach presented in this section discards the information contained in the correlations between neighboring events. This is the penalty accepted for the much smaller computational task of having to deal with only a number of terms equal to the total number of all histogram bins (a few tens typically, independent of the length of the events list) for the calculation of the likelihood, compared to the total number of events (some thousands typically, proportional to the length of record) involved in fitting the whole dwell-time series (see Qin et al., 1996, 1997). Even so, a lot of information is still exploited. If n_c and n_o are the numbers of closed and open states, respectively, of a single channel, then the dwell-time distribution of level k for a record with N channels contains

$$\binom{n_o + k - 1}{k} \cdot \binom{n_c + N - k - 1}{N - k}$$

exponential components. The time constants and relative amplitudes of each component, summed for all levels, plus, for $N > 1$, the relative scaling factor for each level provide a total number of

$$\left[2 \cdot \sum_{k=0}^{N} \binom{n_o + k - 1}{k} \cdot \binom{n_c + N - k - 1}{N - k}\right] - 1$$

constraints used for restraining the fit. For a linear three-state scheme with two closed states and one open state, there are four rate constants to be found. The number of parameters required to describe all dwell-time histograms are 4 for a single-channel record, 11 for a record with 2 channels, and 55 for a record with 6 channels. All of these parameters are functions of the four single-channel rate constants, providing strong constraints for accurate fitting. This does not necessarily mean that models that are unidentifiable from single-channel histogram fitting will become solvable from multichannel fitting. For instance, the simple loop model of CCO cannot be solved by single-channel histogram fitting, because the likelihood surface has no unique maximum, and it is easy to show that this property is inherited to the likelihood of the set of histograms of a multichannel patch as well.

Correction for missed events

Limited bandwidth of the patch-clamp recording system is the price paid for filtering to achieve an acceptable SNR. As a result of the filter delay, short events remain undetected, distorting the distribution of observed dwell times. A general theory for the treatment of the missed event problem was developed by Roux and Sauvé (1985), who assumed an absolute dead time (t_d): all events shorter than t_d remain undetected; all events longer than t_d are detected. Because of noise, such a fixed t_d cannot be calculated simply from the response characteristics of the filter, but it can be imposed retrospectively, by choosing t_d larger than the filter dead time, and by concatenating to the preceding sojourn any dwell time shorter than t_d (see Qin et al., 1996).

Following Roux and Sauvé (1985), given a fixed t_d, the survivor function for the observed dwell-time distribution of a conductance level k is defined by

$$\begin{aligned}\mathrm{surv}^{R\&S}_{\{k\}}(t, t_d) &= \mathrm{P}(\text{no observable leaving from } \{k\} \text{ before time } t \mid \\ &\quad \text{entered } \{k\} \text{ at time 0 after an observable stay in } \{\bar{k}\}) \qquad (11)\end{aligned}$$

where "observable" means longer than t_d. With the additional assumption that the total duration of missed event(s) within an observed event is negligible, the authors were able to write down the pdf, which, when integrated from t to ∞, yields the survivor function. In the present notation,

$$\mathrm{surv}^{R\&S}_{\{k\}}(t, t_d) = [\mathbf{initial}^{R\&S}]^T_k e^{\hat{\mathbf{Q}}_{kk}t}\mathbf{1}_k, \qquad (12)$$

where

$$[\mathbf{initial}^{R\&S}]^T_k = \frac{\mathbf{p}_k(\infty)^T\mathbf{Q}_{k\bar{k}}e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d}(-\mathbf{Q}^{-1}_{\bar{k}\bar{k}})\mathbf{Q}_{\bar{k}k}}{\mathbf{p}_k(\infty)^T\mathbf{Q}_{k\bar{k}}e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d}(-\mathbf{Q}^{-1}_{\bar{k}\bar{k}})\mathbf{Q}_{\bar{k}k}\mathbf{1}_k}, \qquad (12a)$$

$$\hat{\mathbf{Q}}_{kk} = \mathbf{Q}_{kk} - \mathbf{Q}_{k\bar{k}}(\mathbf{I}_{\bar{k}} - e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d})\mathbf{Q}^{-1}_{\bar{k}\bar{k}}\mathbf{Q}_{\bar{k}k}. \qquad (12b)$$

As noted by Qin et al. (1996), Eq. 11 does not take into account that no leaving can occur at all between time 0 and t_d after entering $\{k\}$, because that would produce a short $\{k\}$ event, which would be concatenated to the preceding event. The fact that a $\{k\}$ event is detected means that its initial section (between time 0 and t_d) is free of transitions. To account for this, an additional condition has to be introduced into the definition of the survivor function.

The correct definition of the survivor function will now be

$$\begin{aligned}\mathrm{surv}_{\{k\}}(t, t_d) &= P(\text{no observable leaving from } \{k\} \text{ before time } t \mid \\ &\quad \text{entered } \{k\} \text{ at time 0 after an observable stay in } \{\bar{k}\} \\ &\quad \text{and stayed in } \{k\} \text{ until time } t_d). \qquad (13)\end{aligned}$$

Using this definition, and following the steps described in Appendix B, the correct form of the survivor function can be written as

$$\mathrm{surv}_{\{k\}}(t, t_d) = \begin{cases} 1 & \text{for} \quad t < t_d \\ [\mathbf{initial}]^T_k e^{\hat{\mathbf{Q}}_{kk}(t - t_d)}\mathbf{1}_k & \text{for} \quad t \geq t_d, \end{cases} \qquad (14)$$

where

$$[\mathbf{initial}]^T_k = \frac{\mathbf{p}_k(\infty)^T\mathbf{Q}_{k\bar{k}}e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d}(-\mathbf{Q}^{-1}_{\bar{k}\bar{k}})\mathbf{Q}_{\bar{k}k}e^{\mathbf{Q}_{kk}t_d}}{\mathbf{p}_k(\infty)^T\mathbf{Q}_{k\bar{k}}e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d}(-\mathbf{Q}_{\bar{k}\bar{k}}{}^{-1})\mathbf{Q}_{\bar{k}k}e^{\mathbf{Q}_{kk}t_d}\mathbf{1}_k}, \qquad (14a)$$

$$\hat{\mathbf{Q}}_{kk} = \mathbf{Q}_{kk} - \mathbf{Q}_{k\bar{k}}(\mathbf{I}_{\bar{k}} - e^{\mathbf{Q}_{\bar{k}\bar{k}}t_d})\mathbf{Q}^{-1}_{\bar{k}\bar{k}}\mathbf{Q}_{\bar{k}k} \text{ (as in Eq. 12b).} \qquad (14b)$$

Thus, the additional condition results in a corrected initial vector and a translation of the survivor function along the time axis. The survivor function now takes on the value of 1 at time $t = t_d$, as expected. This correction is equivalent to the adjustment made by Qin et al. (1996); but merging the correction into the initial vector rather than the main operator seems more natural in the present application, and is required for the proper scaling of the survivor functions.

The relative scaling factor for each conductance level, required for simultaneous fitting, can be calculated by analogy to Eq. 5,

$$\mathrm{pctlvl}(k, t_d) = \frac{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d} (-\mathbf{Q}_{\bar{k}\bar{k}}^{-1}) \mathbf{Q}_{\bar{k}k} e^{\mathbf{Q}_{kk} t_d} \mathbf{1}_k}{\sum_{j=0}^{N} [\mathbf{p}_j(\infty)^T \mathbf{Q}_{j\bar{j}} e^{\mathbf{Q}_{\bar{j}\bar{j}} t_d} (-\mathbf{Q}_{\bar{j}\bar{j}}^{-1}) \mathbf{Q}_{\bar{j}j} e^{\mathbf{Q}_{jj} t_d} \mathbf{1}_j]}. \quad (15)$$

For $t_d = 0$, Eqs. 14 and 15 reduce to Eqs. 2 and 5, respectively.

Implementation of a fixed dead time

Some thought needs to be given to how the fixed dead time is implemented in practice. Qin et al. (1996) suggested imposing a fixed dead time (preferably longer than the filter dead time) retrospectively by concatenating each event shorter than the chosen t_d to the preceding sojourn. As pointed out by the authors, if a channel undergoes many short transitions in a row (buzz mode) the fixed dead time cannot be enforced in the above way. However, even in the absence of a buzz mode, when many channels are present, other types of event sequences will appear for which the above method violates the definition in Eq. 13. These cases can be dealt with in the way discussed below, even though buzz mode events will still remain unidealizable.

For instance, as illustrated in Fig. 2 *A* (*left*), a leaving from $\{k\}$ to $\{\bar{k}\}$ much longer than t_d would be ignored that way, if it consisted of a series of dwells—each shorter than t_d—at different conductance levels in $\{\bar{k}\}$, before returning to $\{k\}$. This is contrary to Eq. 13, and would result in concatenation of the two bracketing $\{k\}$ events plus the intervening sojourn in $\{\bar{k}\}$ into one long $\{k\}$ event, leading to underestimation of the rates. A different, but smaller, problem is that, according to Eq. 13 a $\{k\}$ event should be considered terminated once there is a leaving into $\{\bar{k}\}$ longer than t_d. However, as illustrated in Fig. 2 *A* (*right*), if the new sojourn in $\{\bar{k}\}$ begins with an event shorter than t_d in a particular conductance level of $\{\bar{k}\}$ (or a series of such events) that event (series) would be concatenated to the end of the last $\{k\}$ event. Note that both of these problems arise only if there are more than two conductance levels.

To circumvent these problems, the following strategy was used (see flow chart in Fig. 2 *B*) for imposing a fixed t_d in a manner that more closely follows the assumptions of Eq. 13. (Because, in this study, half-amplitude threshold crossing was used for idealization of raw current traces, the algorithm could be applied on line, yielding an events list with the required properties. For other methods of idealization, this can be done retrospectively.) Scanning along the record, events are written to the output events list until a short ($\leq t_d$) event is reached. Incoming short events are temporarily stored in a buffer, until the next long ($> t_d$) event is reached or the conductance returns to the level of the last long event. At this point, there are three options.

If the new (long) event belongs to a different conductance level than the last long event, the contents of the buffer are discarded, and a flag is written to the events list stating the total length of discarded time, followed by the new long event.

If the new event belongs to the same conductance level as the last long event, and the total time stored in the buffer is

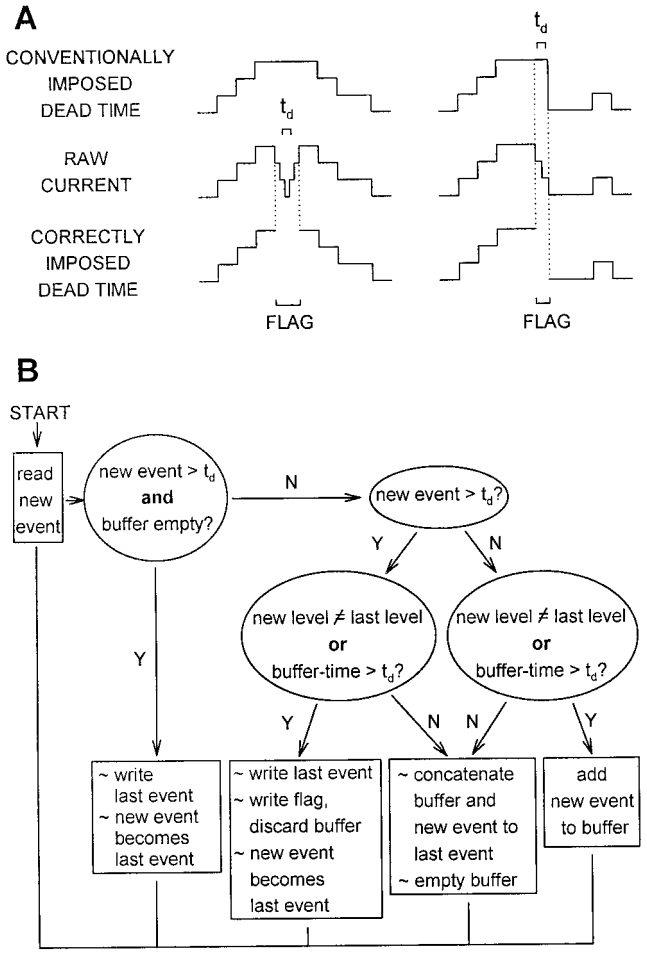

FIGURE 2 Application of a fixed dead time. (*A*) Two examples of raw event sequences where conventional concatenation of brief events (*top*) is inappropriate: *left*, a leaving from level 3 longer than t_d is incorrectly ignored because it is composed of only sub-dead-time components; *right*, a level 3 event is prolonged by incorporation of a series of brief sojourns (staircase), although the latter are part of a long dwell in $\{\bar{3}\}$. For comparison, idealized event sequences generated by the algorithm described in the text and in (*B*) are shown below. (*B*) Flow chart of the algorithm used to avoid inappropriate concatenations.

shorter than t_d, the contents of the buffer and the bracketing events are concatenated into a single event of conductance level given by the bracketing events and of duration given by the sum of the three sources.

If the new event belongs to the same conductance level as the last long event, but the total time stored in the buffer exceeds t_d, the last long event is written to the events list. If the new event is short, it is added to the buffer, otherwise the buffer is discarded, a flag is written to the events list, and the new event is considered separately.

The flow chart in Fig. 2 *B* summarizes the strategy described above. Fig. 2 *A* provides a visual comparison of the resulting idealized event sequences obtained in the conventional way or correctly, according to the above algorithm.

As a result, the only concatenated events are of the type $\{k\} \rightarrow \{\bar{k}\} \rightarrow \{k\}$, where the total time spent in $\{\bar{k}\}$ is shorter than t_d, as required by Eq. 13. The first sojourn in every detected event still has to be longer than t_d (see Eq. 13), otherwise it would be discarded from the buffer. The flags can be thought of as dividing the record into small sections—considered separately—within which the fixed dead time can be imposed consistently, whereas the regions where this is impossible are discarded. The final histograms can be viewed as the sums of the histograms for the good sections. This approach improved the results of subsequent histogram fitting substantially, especially in cases where, because of frequent rapid gating events and many channels, sequences of the type shown in Fig. 2 *A* occurred often (see Fig. 6).

Program cycle

To implement the procedures just described, a computer program (see Fig. 3) was written in C language. The inputs for this analysis program are 1) the model gating scheme, 2) an events list with sub-dead-time events suppressed, 3) the value of t_d, 4) the number of channels, and 5) an initial guess at the single-channel rate constants. Figure 3 illustrates the program cycle in the form of a flow chart.

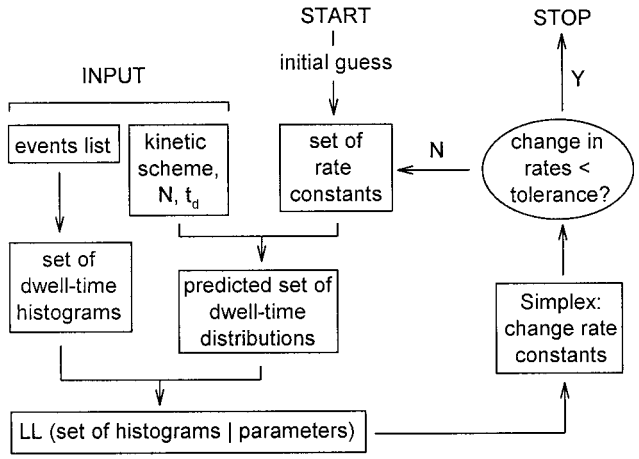

FIGURE 3 Flow chart illustrating the program cycle of the fitting software.

First, the program scans through the events list and constructs the dwell-time histograms, with logarithmic time axes (Sigworth and Sine, 1987), for each conductance level. Unless $t_d = 0$, the lower binning limit is t_d for all levels (i.e., no short events are excluded from binning). Bin density is 6 per *e*-fold. The maximum number of bins is 60 for each level, which usually accomodates all events. If the range of dwell durations exceeds e^{10} ($2 \cdot 10^4$), some long events may remain unbinned, in which case the correction explained in Eqs. 9 and 10 is implemented. Event binning is the only step whose processing time depends on the actual length of the events list. Because this step represents only about one percent of the computational time, the latter is largely independent of the length of the record, in contrast to routines that fit the event sequence (see Horn and Lange, 1983; Qin et al., 1996, 1997).

Next, the program enters a loop of iterations aimed to maximize the likelihood of the set of histograms. The loop starts with the construction of the transition rate matrix **Q** for the model macrosystem, as given by Eq. 1. The survivor functions of the conductance levels are calculated according to Eq. 14, with scaling factors given by Eq. 15. Finally, the log likelihood function is calculated according to Eq. 9. The likelihood is maximized with respect to the rate constants using the simplex algorithm (Caceci and Cacheris, 1984), chosen because of its simple programming code and proven robustness.

The key step is evaluation of the survivor function, of the form $\mathbf{p}^T e^{\mathbf{Q}t}\mathbf{1}$. The row vector $\mathbf{p}^T e^{\mathbf{Q}t}$ is calculated as a power-series, truncated when a term drops below a predefined error,

$$\mathbf{p}^T e^{\mathbf{Q}t} = \mathbf{p}^T + \mathbf{p}^T\mathbf{Q}t + (\mathbf{p}^T\mathbf{Q}t)\left(\frac{t}{2}\mathbf{Q}\right) + \left[(\mathbf{p}^T\mathbf{Q}t)\left(\frac{t}{2}\mathbf{Q}\right)\right]\left(\frac{t}{3}\mathbf{Q}\right) + \cdots$$

Each term is obtained from the previous one as a vector-matrix product, involving n_Q^2 add-multiply operations, n_Q being the dimension of **Q**. Thus, processing time is proportional to the square of the number of macroscopic states.

RESULTS AND DISCUSSION

The analysis program was tested on a range of simulated multichannel current traces to establish how the efficiency depends on the gating model, relative sizes of the rate constants, channel number, dead time, noise, and length of the record. For reasonable schemes, i.e., if all microscopic states were visited at some non-negligible frequency by the individual channels, the iterations converged, starting from a broad range of initial values.

The C_1–O–C_2 model was chosen to test the efficiency of the routine given a variety of parameter combinations. Because this is the scheme used to describe open-channel block, the following nomenclature will be adopted:

$$C \underset{r_{OC}}{\overset{r_{CO}}{\rightleftharpoons}} O \underset{r_{BO}}{\overset{r_{OB}}{\rightleftharpoons}} B.$$

This single-channel scheme has one conducting state O (open), and two nonconducting states C (closed) and B (blocked). Because of the symmetry of the scheme (exchanging r_{OC} with r_{OB} and r_{CO} with r_{BO} yields a scheme indistinguishable from the original one), the likelihood surface is also symmetrical, with two peaks mirroring each other. Depending on the seed parameters, the program ends up at one of the two peaks.

Simulation of multichannel currents and noise

Channel currents were constructed using the macroscopic Markov scheme described earlier and a simulator similar to that described by Blunck et al. (1998). Channel current traces were overlaid by noise synthesized with specified Gaussian amplitude distribution and Lorentzian power spectrum, to mimic the signal that arises after low-pass filtering broad-bandwidth noise, whose initial power spectrum is flat in the pass-range of the filter. Variance σ^2 and corner frequency $f_{c,n}$ were adjustable. A typical setting for $f_{c,n}$ was 5 kHz, a bandwidth at which patch-clamp data are commonly acquired. To test the validity of the dead-time correction, noisy data traces were Gaussian-filtered at a corner frequency $f_{c,G}$, chosen to achieve a SNR sufficient for idealization by half-amplitude threshold crossing (see Fig. 7 *A*).

Distribution of the parameter estimates

To establish the nature of the scatter of the estimates, 100 simulations were carried out using the C–O–B scheme, with a fixed set of parameters (in s^{-1}, $r_{CO} = 50$, $r_{OC} = 10$, $r_{OB} = 50$, $r_{BO} = 1000$; $N = 3$, 900 events), but different random seed values. Figure 4 shows the fit results for all parameters, normalized to the simulated values and collected into bins of width 0.05. The histograms were fitted reasonably well by Gaussian functions (*solid lines* in Fig. 4). In subsequent sections, results will be depicted as the mean of the estimates normalized to the simulated values, with error bars representing the normalized standard deviations.

Sensitivity to relative values of rate constants, channel number, and dead time

Next, fitting efficiency was tested on the C–O–B model with various combinations of rate constants. The difficulty of the fit was found to depend on two factors: the difference

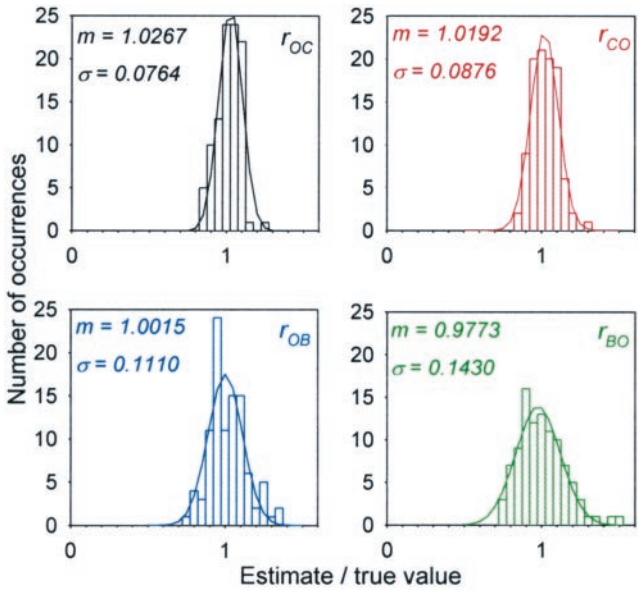

FIGURE 4 Histograms describing the scatter of the estimates. 100 simulations with different random seed values (900 events each) were made for 3 channels obeying the C–O–B scheme, $r_{CO} = 50$, $r_{OC} = 10$, $r_{OB} = 50$, $r_{BO} = 1000$ (in s^{-1}). Dead time was 0.4 ms. Parameter estimates, normalized to the respective true values, were collected into bins of width 0.05. Distributions of the estimates were fitted by Gaussian functions (*solid lines*) of the form $A \cdot \exp[-(x - m)^2/2\sigma^2]$.

in the mean dwell times of B and C, and the relative frequencies of entering C or B, respectively. These factors are expressed as ratios of rate constants: the ratio of the mean dwell times in C and B is r_{BO}/r_{CO}, the mean number of blocked (B) events between two closures (C events) is r_{OB}/r_{OC}. In more general terms, fitting is easy if all microscopic states are visited at some non-negligible frequency, and dwell times of microscopic states in the same conductance class are sufficiently distinct—criteria that already arise for single-channel histogram fitting.

Ratio r_{BO}/r_{CO} was varied from 100 to 4, and the fitting procedure performed well over the entire range, even when the difference between mean blocked and closed times was as small as four-fold. As an example, after simulating 2500 events with 4 channels, with input parameters (in s^{-1}) $r_{CO} = 50$, $r_{OC} = 10$, $r_{BO} = 50$, and $r_{BO} = 200$, followed by idealization with a dead time of 1 ms, the fitting yielded estimates $r_{CO} = 60$, $r_{OC} = 13$, $r_{OB} = 39$, and $r_{BO} = 185$, respectively.

Next, ratio r_{OB}/r_{OC} was varied in a total of 1800 experiments. While r_{BO}/r_{CO} was set to either 100 or 20, r_{OB}/r_{OC} values of 0.1, 0.2, 1, 5, and 10 were tested. Each combination of rate constants was tried with one to six channels, and dead times of 0, 0.1, 0.2, 0.3, 0.4, and 0.5 times the mean blocked time. (Because, for N channels, the fastest macroscopic rate constant is $N \cdot r_{BO}$, a dead time of $0.5/r_{BO}$ is three times longer than the mean dwell time of the shortest lived state, when $N = 6$.) Each of these 360 parameter

combinations was simulated five times, starting with different random seed numbers. For a single channel, 600 to 2200 transitions were simulated, traces with multiple channels were simulated for the same length of time as the corresponding single-channel ones.

Figure 5 is a summary of 300 experiments. Results were similar for all channel numbers ranging from 1 to 6, those for 2 (*Panel A*) and 5 channels (*Panel B*) are shown. With $r_{BO}/r_{CO} = 20$, r_{OB}/r_{OC} was varied (x axis) between 0.1 and 10, with t_d ranging (y axis) from 0 to 0.5 times the mean blocked time ($1/r_{BO}$). Means and standard deviations of the fit results for the individual rate constants are shown (z axis) normalized to their simulated values. A symbol at height 1 means accurate estimation, values higher or lower than 1 represent over- and underestimates, respectively. The solution surfaces tend to be close to 1; i.e., good estimates were obtained over a broad range of input parameters.

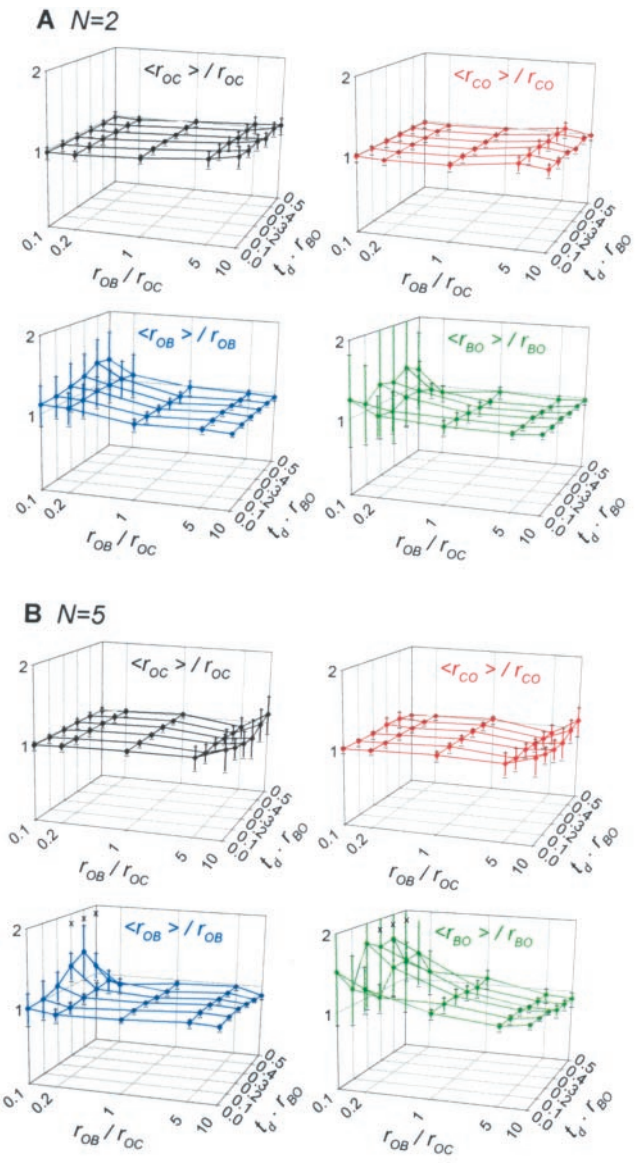

FIGURE 5 Performance with different rate constants. The means of the parameter estimates from five simulations, normalized to the true values, are plotted against r_{OB}/r_{OC} and $t_d \cdot r_{BO}$. Error bars are normalized standard deviations. A symbol at height 1 means accurate estimation. Error bars of r_{OB} and r_{BO} are large when r_{OB}/r_{OC} is small, whereas those of r_{CO} and r_{OC} increase somewhat for large r_{OB}/r_{OC}. (*A*) $N = 2$. (*B*) $N = 5$; in three cases, marked with x, estimation of r_{OB} and r_{BO} was impossible: for those combinations, symbols and error bars represent only four simulations.

For r_{OB}/r_{OC} small, transitions between O and B are rare, and many of these are lost if t_d is long compared to the mean blocked time (i.e., $r_{BO} \cdot t_d$ is comparable to 1). Under these conditions the likelihood surface is relatively insensitive to r_{OB} and r_{BO}, the estimates of which become uncertain, whereas r_{CO} and r_{OC} are still estimated precisely. This is verified in Fig. 5, where the error bars of r_{OB} and r_{BO} increase for $r_{OB}/r_{OC} < 0.2$. When r_{OB}/r_{OC} is large, transitions between states O and B of the individual channels are prevalent. Whereas estimation of r_{OB} and r_{BO} becomes easy even for long t_d, the error on r_{CO} and r_{OC} increases (but remains around 20% even for $r_{OB}/r_{OC} = 10$). Overall errors become smaller with larger numbers of fitted events (see later).

To demonstrate the impact of the implementation of t_d on subsequent fitting, the same simulated current traces were idealized by concatenating brief events either in the conventional or the improved way (cf. Fig. 2). Figure 6 shows parameter estimates for 6 channels, after correct (depicted as in Fig. 5), or conventional idealization (cyan symbols, error bars omitted for clarity). In cases where rapid blocked

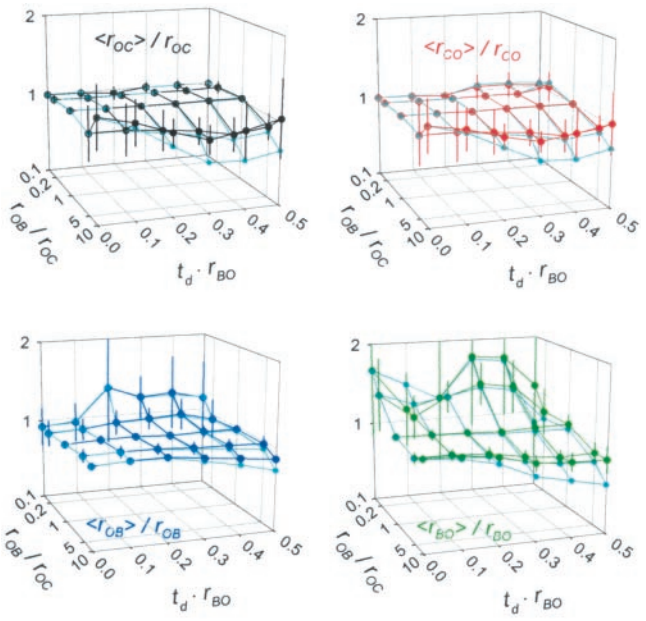

FIGURE 6 Comparison of fit results for 6 channels after conventional (*cyan symbols*) or correct (depicted as in Fig. 5.) implementation of a fixed t_d. For $r_{OB}/r_{OC} \geq 5$ and $t_d \cdot r_{BO} \geq 0.3$, conventional idealization resulted in considerable underestimation of the rate constants.

events were frequent (r_{OB}/r_{OC} high), especially with longer t_d, rate constants were seriously underestimated after conventional concatenation, whereas accurate estimates were obtained after improved idealization. This effect is expected to be less significant for methods that also exploit information from adjacent interval correlations (e.g., Horn and Lange, 1983; Qin et al., 1996). Also, the types of events described in Fig. 2 only arise for three or more conductance levels, and their impact was found to become significant only with 5–6 channels and many rapid transitions (see Fig. 6).

The correction for missed events as given in Eq. 14 is approximate, it assumes that the total duration of missed events is short compared to the observed event (Roux and Sauvé, 1985). This assumption becomes less accurate for more channels, especially if all rate constants are fast. As an example, for the C–O–B scheme, 1000 events simulated with 1 channel and rate constants (in s^{-1}) $r_{CO} = 50$, $r_{OC} = 100$, $r_{OB} = 200$, and $r_{BO} = 1000$, were well fitted with $t_d = 1$ ms ($r_{BO} \cdot t_d = 1$), yielding estimates of $r_{CO} = 44$, $r_{OC} = 91$, $r_{OB} = 163$, and $r_{BO} = 838$, although more than 50% of the events were omitted. For the same scheme with 6 channels and 6000 events, the estimates were good with $t_d = 0.5$ ms ($r_{CO} = 57$, $r_{OC} = 118$, $r_{OB} = 150$, $r_{BO} = 1208$), but started to deviate at $t_d = 1$ ms ($r_{CO} = 46$, $r_{OC} = 98$, $r_{OB} = 83$, and $r_{BO} = 855$). When three of the rates were slowed down, but r_{BO} left unchanged, the fit worked well even for 6 channels and $t_d = 1$ ms, although 60% of the events was lost, yielding, for simulated parameters $r_{CO} = 10$, $r_{OC} = 20$, $r_{OB} = 100$, and $r_{BO} = 1000$, estimates of $r_{CO} = 10$, $r_{OC} = 19$, $r_{OB} = 83$, and $r_{BO} = 834$.

Performance of the fit procedure on noisy data

The advantage of a correction for missed events is applicability to noisy data. If the SNR (i.e., the ratio of the single-channel current amplitude to the standard deviation of the noise) is small, idealization by any method is difficult. If the SNR is smaller than ~5, idealization by half-amplitude threshold crossing is impossible, because too many false events will appear. Using a routine similar to the one described here, but without missed-event correction, Blunck et al. (1998) reported a failure of their procedure on noisy data, attributed to their event-detection algorithm. This problem can be circumvented by filtering the raw data to increase the SNR, and subsequently correcting for the loss of brief events due to the filter dead time.

To explore the theoretical limits of this type of missed-event correction, the following facts need to be considered. (The numerical values below are derived for Lorentzian noise, a Gaussian filter, and for the C–O–B scheme, but the same reasoning applies with only minor variations to other types of noise, filters, and schemes.) First, t_d cannot be very long compared to the shortest lived state in the kinetic scheme. For the C–O–B scheme, the analysis worked well with all parameter combinations tested, for up to 6 channels, with $t_d \leq 0.5/r_{BO}$. (Longer dead times also worked as long as r_{CO} and r_{OC} were not very fast.) This imposes a lower limit to the corner frequency used for filtering. For a Gaussian filter, with corner frequency $f_{c,G}$, $t_d \approx 0.18/f_{c,G}$. Hence, the constraint $f_{c,G} \gtrsim 0.36 \cdot r_{BO}$ is obtained. Second, reliable idealization by half-amplitude threshold crossing requires a SNR $\gtrsim 8$. In the common case where the corner frequency of the noise ($f_{c,n}$) is much greater than $f_{c,G}$, the relationship between the SNRs before (SNR_n) and after (SNR_G) filtering is given by

$$SNR_G^2 = SNR_n^2 \cdot (\pi \cdot \ln 2)^{1/2} \cdot f_{c,n}/f_{c,G}. \quad (16)$$

Combining this result with the above inequalities yields

$$SNR_n^2 \cdot f_{c,n} \gtrsim 8^2 \cdot 0.36 \cdot (\pi \cdot \ln 2)^{-1/2} \cdot r_{BO} \approx 16 \cdot r_{BO}. \quad (17)$$

This is an approximate guideline for the quality of the raw data, required for efficient use of this approach. As an example, if $r_{BO} = 1000\ s^{-1}$ (i.e., the mean blocked time is 1 ms), raw data acquired at an initial bandwidth of 5 kHz will have to satisfy $SNR_n \gtrsim 1.8$. In contrast, SNR_n at 5 kHz can be as low as 1 if the mean blocked time is longer than 3 ms. (For other schemes, r_{BO} is replaced by the fastest rate constant, and, for different noise characteristics and filters, the numerical values will be somewhat modified.)

To test the missed-event correction algorithm on more realistic records than the noise-free simulations used for Figs. 4, 5, and 6, the fit procedure was tested on noisy current traces. The testing was structured as illustrated in Fig. 7 *A*. Channel currents were simulated and overlaid by noise. SNRs as low as 3, 2 and 1 were tested, in which cases idealization by threshold crossing was impossible. Noisy traces were Gaussian-filtered digitally to increase the SNR to 8–10. Filtered traces were idealized by half-amplitude threshold crossing combined with the imposition of a fixed t_d, about twice the dead time of the filter, to minimize uncertainties caused by noise (see representative noisy, filtered, and idealized traces in Fig. 7 *A*). Finally, events lists were fitted to yield estimates of the rate constants.

The program performed well on noisy data after filtering. Figure 7 *B* provides a summary of one such series of 150 experiments structured like the one shown in Fig. 7 *A*. With a fixed ratio of $r_{BO}/r_{CO} = 20$, r_{OB}/r_{OC} was varied between 0.1 and 10 (x axis), with channel numbers ranging from 1 to 6 (y axis). Each parameter combination was simulated 5 times. Initial noise was characterized by $SNR_n = 2$ and $f_{c,n} = 5.0 \cdot r_{BO}$. Noisy traces were filtered at $f_{c,G} = 1.0 \cdot r_{BO}$, and idealized with $t_d = 0.4/r_{BO}$. A sampling rate of 20 times $f_{c,G}$ was used. From Fig. 7 *B*, the performance of the analysis under these conditions is comparable to that seen for noise-free simulations (cf. Fig. 5), confirming the applicability of the approach in realistic situations.

It is of interest to visualize the impact of missed-event correction. The same set of noisy traces described in the previous paragraph were also filtered at $f_{c,G} = 0.45 \cdot r_{BO}$,

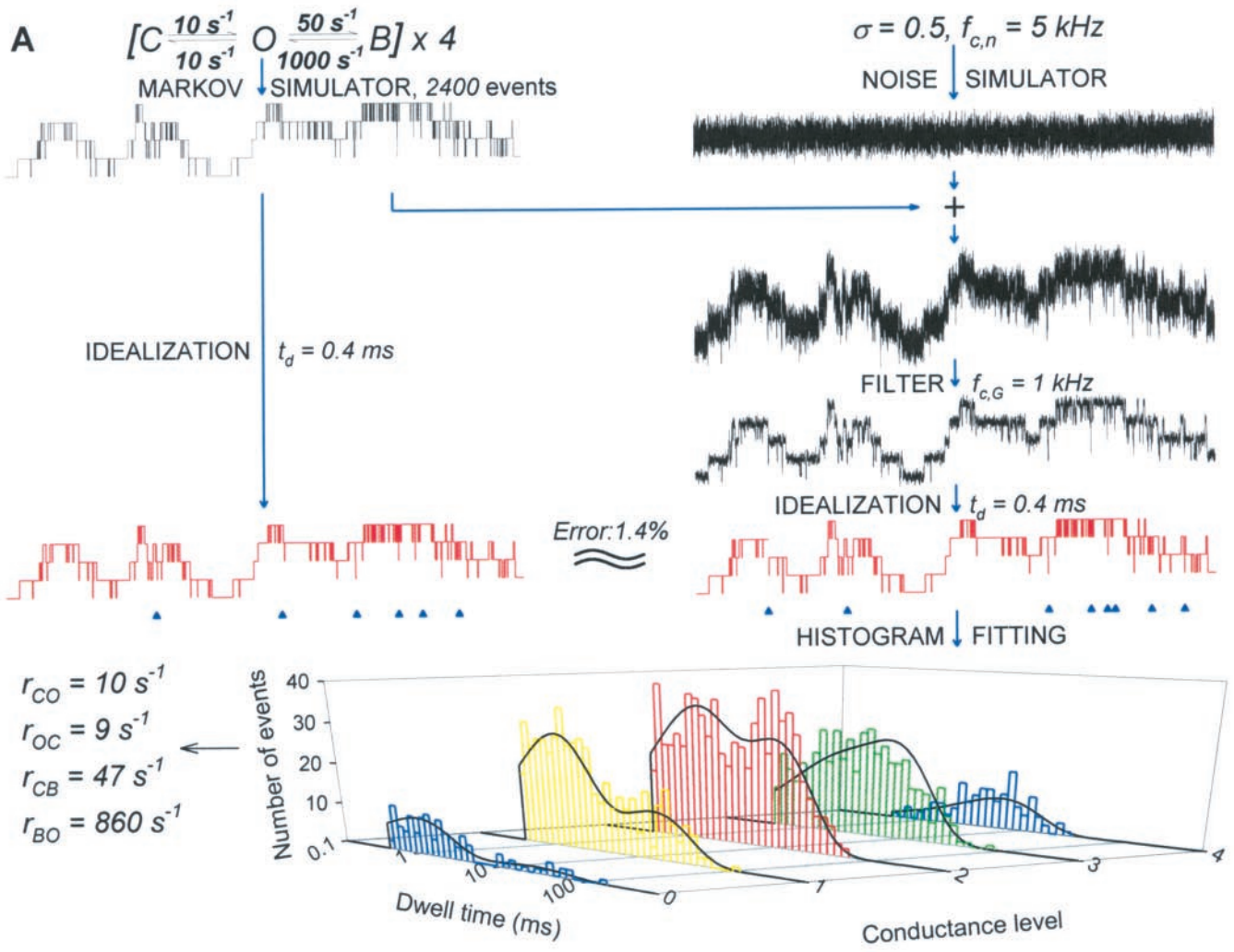

FIGURE 7 Testing of the fit procedure on noisy current traces. Simulated noise was added to channel currents. Noisy traces were Gaussian-filtered digitally. Filtered traces were idealized by half amplitude threshold crossing. Events lists were fitted to extract rate constants. (*A*) Structure of the testing procedure. In the example shown, using the same t_d, idealized currents obtained from noise-free channel currents and noisy traces after filtering were 98.6% identical. Small triangles mark the sites of flags (see text). (*B*) Summary of 150 tests of the type shown in *A*. The means of the parameter estimates from five simulations, normalized, are plotted against r_{OB}/r_{OC} and channel number. Error bars are SD. Simulation parameters: C–O–B scheme, $r_{BO}/r_{CO} = 20$, $SNR_n = 2$, $f_{c,n} = 5.0 \cdot r_{BO}$. Filter: $f_{c,G} = 1.0 \cdot r_{BO}$. Idealization: $t_d = 0.4/r_{BO}$. Cyan symbols are estimates obtained without missed-event correction. Without correction r_{CO} and r_{OC} are fairly well estimated for $N = 1$, but less so for higher numbers of channels.

resulting in a filter dead time of $0.4/r_{BO}$, equal to t_d above. These traces were then idealized by simple half-amplitude threshold crossing, without imposing any artificial dead time. The events lists were then analyzed without missed-event correction, i.e., Eqs. 2 and 5 were used in the program cycle instead of Eqs. 14 and 15. The resulting estimates for the rate constants are plotted in Fig. 7 *B* in cyan. (Error bars are omitted for clarity.) Roux and Sauvé (1985) showed that, for a single channel with a C–O–B scheme, as long as t_d is short compared to the mean duration of sojourns in C, the estimates of r_{CO} and r_{OC} remain undistorted, even if most blocked events are filtered out. This is verified in Fig. 7 *B*, where, for 1 channel, lack of missed-event correction results in underestimation of r_{OB} and r_{BO} only. However, with many channels, lack of correction results in underestimation of r_{CO} and r_{OC} as well. This is because, even if r_{CO} and r_{OC} are slow compared to the dead time, transitions between C and O are lost when there are several channels, if mirroring transitions of two different channels occur close to each other in time.

Dependence of the scatter of the estimates on the length of the record

It is useful to establish, for a given scheme, how long a record is required for reliable fitting. The C–O–B model was chosen with fixed parameters $r_{CO} = 50$, $r_{OC} = 10$, $r_{OB} = 2$, $r_{BO} = 1000$, $t_d = 0.2/r_{BO}$, $N = 4$, and simulations were done with event numbers ranging from 40 to 5000. Each simulation was repeated 10 times with different random seed values. Means and standard deviations of the normalized parameter estimates are plotted against total event number in Fig. 8 *A*. As expected from the ratio $r_{OB}/r_{OC} = 0.2$, the errors of the estimates of r_{OB} and r_{BO} exceed those of r_{CO} and r_{OC}. For all rate constants, however, the scatter becomes smaller with increasing numbers of events. Figure 8 *B* shows the sizes of the error bars from Fig. 8 *A* plotted against the number of events. The dependence of the errors of r_{OB} and r_{BO} on total event number is shifted to the right compared to the errors of r_{CO} and r_{OC}. From the total number of events, the expected number of individual transitions along any of the four pathways is easily calculated. These numbers are plotted in red as an alternative abscissa in Fig. 8 *B*. The shift of r_{OB} and r_{BO} compared to r_{CO} and r_{OC} is less dramatic from the perspective of this new abscissa. The standard deviations of r_{CO} and r_{OC} decrease to ~20% with ~200 corresponding transitions, and a similar number of respective transitions brings the errors of r_{OB} and r_{BO} also down below 30%.

Dependence of the processing time on channel number and length of record

To illustrate the speed of the fitting, the processing time was measured on events lists of various lengths and different

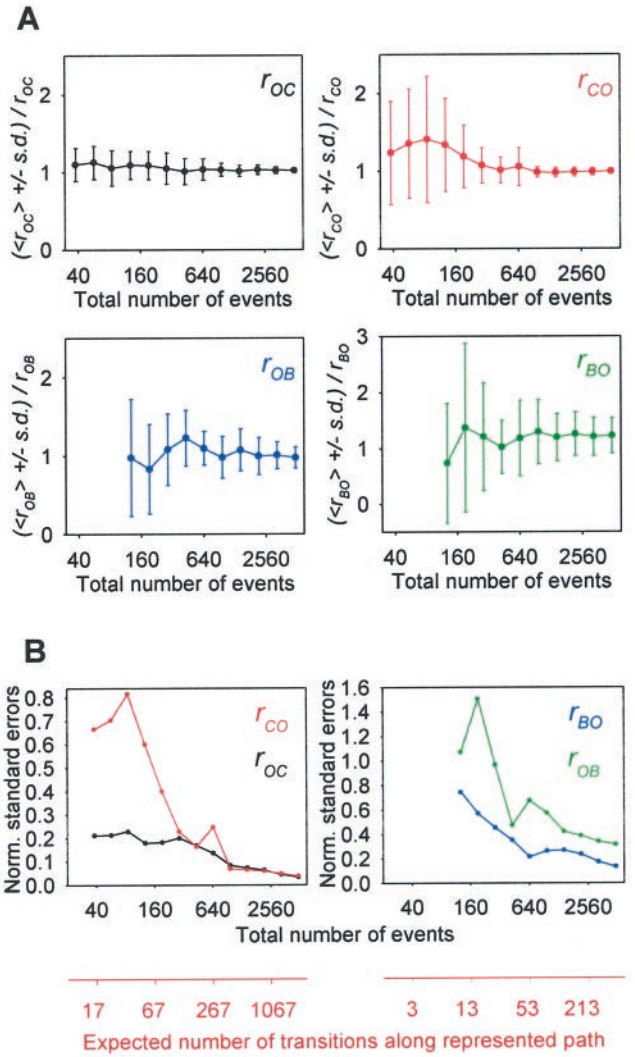

FIGURE 8 Dependence of the scatter of the estimates on the length of the record. Channel currents of different lengths were simulated with the same fixed set of parameters (C–O–B scheme, $r_{CO} = 50$, $r_{OC} = 10$, $r_{OB} = 2$, $r_{BO} = 1000$, $N = 4$), idealized with $t_d = 0.2/r_{BO}$, and fitted. (*A*) Normalized parameter estimates plotted against total event number. Symbols and error bars are means and SD from 10 simulations. Because $r_{OB}/r_{OC} = 0.2$, errors of r_{OB} and r_{BO} are larger than those of r_{CO} and r_{OC} for any given number of total events. All errors decrease with larger numbers of events. (*B*) Lengths of the error bars from *A* plotted against total event number. Alternative abscissa (*red*) describes expected numbers of transitions along represented pathways. Error sizes for various rate constants are more uniform with respect to the alternative abscissa: ~200 corresponding transitions keep the errors below 20% for r_{CO} and r_{OC}, and below 30% for r_{OB} and r_{BO}.

numbers of channels. The processing time depends on many parameters, including kinetic scheme, rate constants, and initial parameter guesses. In the following experiments, the C–O–B scheme was used with $r_{CO} = 50$, $r_{OC} = 10$, $r_{OB} = 50$, $r_{BO} = 1000$. Seed parameters were chosen based on a quick estimate of the cycle-time, obtained by counting the number of transitions per unit time. Fitting was performed on a Pentium 266 MHz machine.

Results are summarized in Fig. 9. Channel number was varied from 1 to 6, with event numbers of 3300, 10,000, or 30,000. Red bars represent the mean of three measurements. As expected, the processing time (note logarithmic ordinate) increases with channel number (as the square of the number of macroscopic states), but remains independent of the length of the data. Typical processing times were ~1 s for 1 channel, ~6 s for 4 channels, and ~33 s for 6 channels.

As a comparison, the same events lists were analyzed with MIL, part of the QuB single-channel analysis package, which maximizes the likelihood of the joint probability density for the whole dwell time series (Qin et al. 1996, 1997). Results are shown as blue bars in Fig. 9. Large differences in processing time are apparent for $N = 4$ and 5, where MIL typically converged after ~500 s and ~5000 s, respectively, compared to ~6 s and ~14 s for the present histogram method. With 6 channels, MIL failed to initialize, but the extrapolated processing time is ~50,000 s (14 hours, cyan bars in Fig. 9) as opposed to ~30 s with the method given here. The resulting estimates were essentially identical for the two approaches.

Applicability to different gating schemes

Two three-state schemes are encountered frequently in ion-channel physiology. In previous sections, the C–O–B scheme (open-channel block model) was tested extensively. A second common scheme is C_1–C_2–O, referred to as the ligand-gated channel. Both schemes result in a bursting gating pattern for a single channel, where openings occur in clusters interrupted by short (flickery) closures, and flanked by long interburst closures.

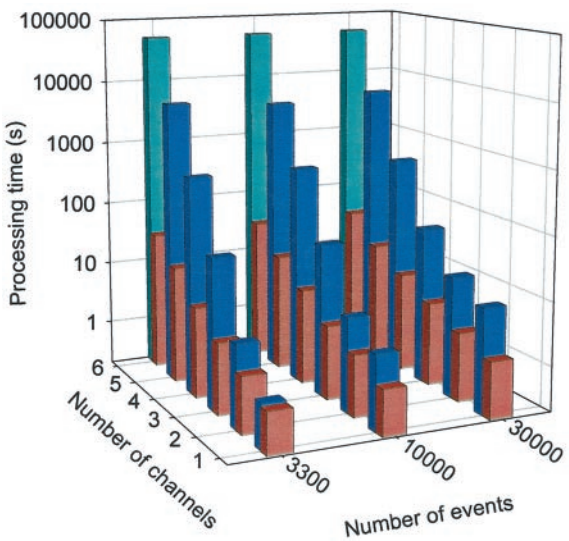

FIGURE 9 Comparison of processing times for simultaneous histogram fitting (*red bars*) and complete time series fitting (MIL, *blue bars*). Bars are means of three measurements. Processing time for simultaneous histogram fitting increases proportionally with the square of the number of macroscopic states, but is independent of the length of the record. Typical processing times were 1, 2, 3, 6, 14, and 33 s for $N = 1$, 2, 3, 4, 5, and 6, respectively. For comparison, MIL produced identical estimates after processing times of 3, 6, 50, 500, and 5000 s for $N = 1$, 2, 3, 4, and 5, respectively, and 10,000 events. Extrapolated processing time for $N = 6$ (*cyan bars*) is ~50,000 s (14 hours) with MIL.

In single-channel records, burst analysis can be applied, where closings shorter than some cutoff t_c are ignored. Various criteria exist for choosing t_c (Jackson et al., 1983; Magleby and Pallota, 1983; Colquhoun and Sakmann, 1985), all of which give satisfactory results if the flickery closures are short compared to the interburst time, and the relative frequencies of the two types of closure are not very different. This is equivalent to filtering, which, as shown by Roux and Sauvé (1985), will not distort subsequent estimation of the slow transitions if the latter are slow compared to the filter. This approach becomes problematic, however, when many channels are present, because short events can arise from mirroring transitions of two different channels occurring close to each other in time. As shown earlier in this paper, even short dead times can seriously distort the estimation of slow rate constants in multichannel patches, unless appropriate care is taken (see, e.g., Fig. 7 *B*, *cyan symbols*).

In records with multiple channels, if the open probability of a single channel is very low, such that the average number of open channels at any given time is small, Jackson (1985) showed that the rate constants can be estimated from the distribution of closed times (all channels closed) and open times (one channel open, no superimposed openings). If the activity of the channels is high, however, such events become rare, limiting the applicability of that approach.

The method described in this paper was also tried on the C_1–C_2–O model, with channel numbers ranging from 1 to 6, and, as with the C–O–B scheme, was found to perform well, regardless of the open probability ($0.01 \leq P_o \leq 0.99$) of the channels. Table 1 summarizes the results of a more limited range of fitting experiments for the C_1–C_2–O scheme. Essentially identical results were obtained, although after much longer processing times (as in Fig. 9), with the more sophisticated program MIL (Qin et al. 1996, 1997).

CONCLUSION

A procedure has been developed for quickly and reliably extracting rate constants of channel gating from patch recordings containing multiple channels. The algorithm is based on a simultaneous maximum likelihood fit to the dwell-time histograms for the various conductance levels. It does not exploit correlations between adjacent events, but results in a greatly reduced computational task and, hence, processing time. This feature makes it attractive for the analysis of multichannel patches in cases where the gating scheme itself is already known and relatively simple. Of course, patches with multiple channels, and analyses using

TABLE 1 Parameter estimates for the C_1–C_2–O_3 scheme

	r_{12}	r_{21}	r_{23}	r_{32}	N	t_d
True value	50	800	200	100	2	0.2
Estimate	56 ± 11	750 ± 215	171 ± 55	97 ± 5		
True value	50	500	500	100	2	0.2
Estimate	50 ± 4	480 ± 103	437 ± 67	97 ± 4		
True value	50	200	800	100	2	0.2
Estimate	49 ± 9	189 ± 26	721 ± 25	95 ± 2		
True value	50	800	200	100	4	0.2
Estimate	47 ± 3	616 ± 14	166 ± 17	99 ± 2		
True value	50	500	500	100	4	0.2
Estimate	50 ± 8	446 ± 90	420 ± 30	95 ± 3		
True value	50	200	800	100	4	0.2
Estimate	48 ± 4	179 ± 16	701 ± 94	94 ± 5		
True value	50	800	200	100	6	0.2
Estimate	51 ± 3	616 ± 50	152 ± 11	98 ± 2		
True value	50	500	500	100	6	0.2
Estimate	52 ± 6	480 ± 99	439 ± 43	96 ± 2		
True value	50	200	800	100	6	0.2
Estimate	40 ± 5	138 ± 26	644 ± 66	94 ± 1		
True value	50	200	800	100	4	0.0
Estimate	52 ± 7	244 ± 29	894 ± 60	100 ± 4		
True value	50	200	800	100	4	0.2
Estimate	48 ± 4	179 ± 16	701 ± 94	94 ± 5		
True value	50	200	800	100	4	0.4
Estimate	49 ± 5	231 ± 45	960 ± 192	108 ± 10		

Rates are s^{-1}, t_d is ms, errors are SD ($n = 5$).

this algorithm, are not the tools of choice for solving complicated schemes, nor for distinguishing between gating models. For those purposes, single-channel records are desirable, along with correlation analysis using event sequence fitting or two-dimensional histograms.

The theory is illustrated for N identical and independent channels, but can be applied straightforwardly to more general cases. The algorithm is tested on two common three-state schemes, C–O–B and C_1–C_2–O, both of which yield a burst-type gating pattern, but it is expected to work on other simple (e.g., three-state) schemes as well. It can be recommended for such schemes when 3 or more channels contribute to the recorded current, because a considerable decrease in processing time is gained in comparison to more advanced routines (see Fig. 9 for a comparison with MIL, the fitting routine in the QuB package), without loss of accuracy.

Accurate parameter estimates are obtained from as little as a few hundred events, if there are no large discrepancies between the frequencies of occurrence of the single-channel gating steps. In contrast, the accuracy increases with longer recordings, without any cost in processing time or memory space requirement.

A robust correction for missed events due to filtering allows the algorithm to be applied to noisy data. Limitations for such cases are summarized in Eq. 17, which provides acceptable limits for the relationship between speed of channel gating and magnitude and bandwidth of noise.

APPENDIX A: LIKELIHOOD FUNCTION IN THE CASE OF BINNING LIMITS

Suppose that lower and upper bin limits $t_{k,min}$ and $t_{k,max}$ are imposed. Quantities analogous to those in Eqs. 5, 6, and 4, conditional on being binned, are defined as

$$\text{pctlvl}'(k) = P(\text{event is level } k \mid \text{event is binned}), \quad \text{(A1)}$$

$$p'_{ki} = P(t_{k,i} \le \text{dwell time} < t_{k,i+1} \mid \text{event is level } k \text{ and event is binned}), \quad \text{(A2)}$$

$$P'(k, i) = P(\text{event is level } k \text{ and } t_{k,i} \le \text{dwell time} < t_{k,i+1} \mid \text{event is binned}). \quad \text{(A3)}$$

Next, $P_{k,binned}$ is defined as the fraction of all observed level k events predicted to be binned given bin limits $t_{k,min}$ and $t_{k,max}$,

$$P_{k,binned} = P(\text{event is binned} \mid \text{event is level } k) = \text{surv}_{\{k\}}(t_{k,min}) - \text{surv}_{\{k\}}(t_{k,max}). \quad \text{(A4)}$$

It follows from the theory of probabilities that

$$\text{pctlvl}'(k) = \frac{\text{pctlvl}(k) \cdot P_{k,binned}}{\sum_{j=1}^{N} \text{pctlvl}(j) \cdot P_{j,binned}}, \quad \text{(A5)}$$

and

$$p'_{ki} = \frac{p_{ki}}{P_{k,binned}}. \quad \text{(A6)}$$

The probability of an event falling into the ith bin of the level k histogram, conditional on being binned, is given by

$$P'(k, i) = \text{pctlvl}'(k) \cdot p'_{ki} = \frac{\text{pctlvl}(k) \cdot p_{k,i}}{\sum_{j=1}^{N} \text{pctlvl}(j) \cdot P_{j,binned}}. \quad \text{(A7)}$$

Proceeding analogously to steps 7 and 8 in the main text, the likelihood function becomes

$$L'(\Theta) = \prod_{k=0}^{N} \prod_{i=1}^{n_k} P'(k, i)^{n_{k,i}} = \prod_{k=0}^{N} \prod_{i=1}^{n_k} \left[\frac{\text{pctlvl}(k) \cdot p_{k,i}}{\sum_{j=1}^{N} \text{pctlvl}(j) \cdot P_{j,binned}} \right]^{n_{k,i}}, \quad \text{(A8)}$$

while $LL'(\Theta)$, defined as $LL'(\Theta) = \ln(L'(\Theta))$, is given by

$$LL'(\Theta) = \sum_{k=0}^{N} \sum_{i=1}^{r_k} n_{k,i} \cdot \ln p_{k,i} + \sum_{k=0}^{N} n_k \cdot \ln \mathrm{pctlvl}(k) - n_t \cdot \ln\left[\sum_{k=0}^{N} \mathrm{pctlvl}(k) \cdot P_{k,\mathrm{binned}}\right], \quad \text{(A9)}$$

where $n_t = \Sigma_{k=0}^{N} n_k$ is the total number of binned events.

APPENDIX B: INTRODUCING ADDITIONAL CONSTRAINT OF FIRST DWELL > t_d

Because of the properties of a stationary Markov system (see e.g., Colquhoun and Hawkes, 1977), Eq. 13 can be rewritten in the form,

$$\begin{aligned}&\mathrm{surv}_{\{k\}}(t, t_d)\\ &= P(\text{no observable leaving from } \{k\} \text{ before time}\\ &\quad t - t_d| \text{ entered } \{k\} \text{ at time } -t_d \text{ after an}\\ &\quad \text{observable stay in } \{\bar{k}\} \text{ and stayed in } \{k\} \text{ until time } 0)\end{aligned} \quad \text{(B1)}$$

The expression in Eq. B 1 has the form $P(C|A \text{ and } B) = P(C \text{ and } A \text{ and } B)/P(A \text{ and } B)$, where C = no observable leaving from $\{k\}$ before time $t - t_d$, A = entered $\{k\}$ at time $-t_d$ after an observable stay in $\{\bar{k}\}$, B = stayed in $\{k\}$ between time $-t_d$ and 0. Following the reasoning of Roux and Sauvé (1985),

$$P(C \text{ and } A \text{ and } B) = \frac{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d} (-\mathbf{Q}_{\bar{k}\bar{k}}^{-1}) \mathbf{Q}_{\bar{k}k} e^{\mathbf{Q}_{kk} t_d}}{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} \mathbf{1}_{\bar{k}}} e^{\hat{\mathbf{Q}}_{kk}(t-t_d)} \mathbf{1}_k, \quad \text{(B2)}$$

$$P(A \text{ and } B) = \frac{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d} (-\mathbf{Q}_{\bar{k}\bar{k}}^{-1}) \mathbf{Q}_{\bar{k}k} e^{\mathbf{Q}_{kk} t_d} \mathbf{1}_k}{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} \mathbf{1}_{\bar{k}}}, \quad \text{(B3)}$$

from which the survivor function can be assembled:

$$\mathrm{surv}_{\{k\}}(t, t_d) = \begin{cases} 1 & \text{for} \quad t < t_d \\ [\mathrm{initial}]_k^T e^{\hat{\mathbf{Q}}_{kk}(t-t_d)} \mathbf{1}_k & \text{for} \quad t \geq t_d \end{cases}, \quad \text{(B4)}$$

where

$$[\mathrm{initial}]_k^T = \frac{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d} (-\mathbf{Q}_{\bar{k}\bar{k}}^{-1}) \mathbf{Q}_{\bar{k}k} e^{\mathbf{Q}_{kk} t_d}}{\mathbf{p}_k(\infty)^T \mathbf{Q}_{k\bar{k}} e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d} (-\mathbf{Q}_{\bar{k}\bar{k}}^{-1}) \mathbf{Q}_{\bar{k}k} e^{\mathbf{Q}_{kk} t_d} \mathbf{1}_k}, \quad \text{(B4a)}$$

$$\hat{\mathbf{Q}}_{kk} = \mathbf{Q}_{kk} - \mathbf{Q}_{k\bar{k}} (\mathbf{I}_{\bar{k}} - e^{\mathbf{Q}_{\bar{k}\bar{k}} t_d}) \mathbf{Q}_{\bar{k}\bar{k}}^{-1} \mathbf{Q}_{\bar{k}k} \quad \text{(as in Eq. 12b).} \quad \text{(B4b)}$$

We thank Dr. David Gadsby for constant support and supervision, and Drs. Benoît Roux and Olaf Andersen for careful revision of the manuscript and insightful comments.

This work was supported by the The Rockefeller University, a William O'Baker Graduate Fellowship, and National Institutes of Health Grant DK-51767 (to D.G.).

REFERENCES

Blunck, R., U. Kirst, T. Riessner, and U.-P. Hansen. 1998. How powerful is the dwell-time analysis of multichannel records? *J. Memb. Biol.* 165:19–35.

Caceci M. S., and W. P. Cacheris. 1984. Fitting curves to data. The simplex algorithm is the answer. *Byte.* May:340–348.

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

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

Colquhoun, D., and B. Sakmann. 1985. Fast events in single channel currents activated by acetylcholine and its analogues at the frog muscle end-plate. *J. Physiol.* 369:501–557.

Colquhoun, D., and F. J. Sigworth. 1995. Fitting and statistical analysis of single channel records. *In* Single Channel Recording. 2nd edition. B. Sakmann and E. Neher, editors. Plenum Press, New York. 483–587.

Horn, R., and K. Lange. 1983. Estimating kinetic constants from single channel data. *Biophys. J.* 43:207–223.

Hoshi, T., W. N. Zagotta, and R. W. Aldrich. 1994. *Shaker* potassium channel gating. I. Transitions near the open state. *J. Gen. Physiol.* 103:249–278.

Jackson, M. B., B. S. Wong, C. E. Morris, H. Lecar, C. N. Christian. 1983. Successive openings of the same acethylcholine receptor channel are correlated in open time. *Biophys. J.* 42:109–114.

Jackson, M. B. 1985. Stochastic behaviour of a many channel membrane system. *Biophys. J.* 47:129–137.

Magleby, K. L., and B. S. Pallotta. 1983. Burst kinetics of single calcium-activated potassium channels in cultured rat muscle. *J. Physiol.* 344: 605–623.

Magleby, K. L., and L. Song. 1992. Dependency plots suggest the kinetic structure of ion channels. *Proc. R. Soc. Lond. B.* 249:133–142.

Qin, F., A. Auerbach, and F. Sachs. 1996. Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys. J.* 70:264–280.

Qin, F., A. Auerbach, and F. Sachs. 1997. Maximum likelihood estimation of aggregated Markov processes. *Proc. R. Soc. Lond. B.* 264:375–383.

Rothberg, B. S., R. A. Bello, and K. L. Magleby. 1997. Two-dimensional components and hidden dependencies provide insight into ion channel gating mechanisms. *Biophys. J.* 72:2524–2544.

Roux, B., and R. Sauvé. 1985. A general solution for the time interval omission problem applied to single channel analysis. *Biophys. J.* 48: 149–158.

Schoppa, N. E., and F. J. Sigworth. 1998a. Activation of *Shaker* potassium channels. I Characterization of voltage-dependent transitions. *J. Gen. Physiol.* 111:271–294.

Schoppa, N. E., and F. J. Sigworth. 1998b. Activation of *Shaker* potassium channels. II. Kinetics of the V2 mutant channel. *J. Gen. Physiol.* 111: 295–311.

Schoppa, N. E., and F. J. Sigworth. 1998c. Activation of *Shaker* potassium channels. III. An activation gating model for wild-type and V2 mutant channels. *J. Gen. Physiol.* 111:313–342.

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.

Vandenberg, C. A., and F. Bezanilla. 1991. A sodium channel gating modelbased on single channel, macroscopic ionic, and gating currents in the squid giant axon. *Biophys. J.* 60:1511–1533.

Weiss, D. S., and K. L. Magleby. 1989. Gating scheme for single GABA-activated chloride channels determined from stability plots, dwell-time distributions, and adjacent-interval durations. *J. Neurosci.* 9:1314–1324.

Zagotta, W. N., T. Hoshi, J. Dittman, and R. W. Aldrich. 1994a. *Shaker* potassium channel gating. II. Transitions in the activation pathway. *J. Gen. Physiol.* 103:279–319.

Zagotta, W. N., T. Hoshi, and R. W. Aldrich. 1994b. *Shaker* potassium channel gating. III. Evaluation of kinetic models for activation. *J. Gen. Physiol.* 103:321–362.