

Estimating the Number of Channels in Patch-Clamp Recordings: Application to Kinetic Analysis of Multichannel Data from Voltage-Operated Channels

W. Baumgartner,* K. Hohenthanner,* G. F. Höfer,* K. Groschner,# and C. Romanin*

*Institute for Biophysics, University of Linz, A-4040 Linz, and #Institute for Pharmacology and Toxicology, University of Graz, A-8010 Graz, Austria

ABSTRACT Important kinetic information of voltage-operated ion channels can be obtained by estimating the open probability, the availability, and the first latency, and by applying run analysis. In the case of multichannel patches, estimation of the number of available channels is a prerequisite for the above analysis. Here we describe a method for calculation of the a posteriori probability of the number of available channels in each sweep by using the Bayes formula. This probability serves as a measure for the number of channels and allows for first latency determination and run analysis. The methods described were applied to simulated and experimental data obtained from L-type Ca^{2+} channel recordings.

INTRODUCTION

Ion channels play a major role in the transduction of biological processes. The membrane potential controls the function of voltage-operated channels, which open in a stochastic way, often described by a hidden Markov model (HMM) (Colquhoun and Hawkes, 1981; Chung et al., 1990). A depolarizing voltage step is usually applied to activate voltage-operated channels, which are then observed for a limited time interval, defined as a sweep. The parameters of interest used to describe the behavior of a single channel are availability (P_s), first latency, and open probability (p_o). P_s is defined as the chance to evoke channel activity after a voltage step, i.e., the ratio of sweeps showing channel activity to the total number of sweeps. First latency is the time to first channel opening in a sweep, and p_o represents the average time the channel is open. Furthermore, P_s has often been found to be clustered, which means that sweeps showing channel activity and those lacking activity are clustered in time. Such an activity pattern is examined by means of a run analysis (Horn et al., 1984; Sigworth and Zhou, 1992), which is a test based on the null hypothesis that “the probability to find channel activity in a sweep is independent of the channel’s availability in the previous sweep.”

Calculation of the above parameters in experiments with a single channel can easily be performed, whereas in multichannel experiments analysis is much more complicated. As channel clustering is often found in cells, giving rise to multichannel recordings, a sophisticated analysis is required to resolve single-channel characteristics from multichannel

data. One major task in the analysis of those data is the estimation of the number of channels. Estimation of the total number of channels in a patch employing several estimators has been carefully discussed (Horn, 1991), whereas determination of the number of available channels in each sweep is apparently more problematic (Sigworth and Zhou, 1992). Although the aforementioned estimators could be adapted for estimation of the number of channels in a sweep, the results are rather poor, as will be shown below.

Here we present a method to calculate the a posteriori probability for the number of available channels in a sweep and show its application to first latency determination and run analysis. Analysis of simulated as well as experimental data of L-type Ca^{2+} channels is demonstrated.

THEORY AND METHODS

General

The data are represented in digitized form, i.e., they are stored as individual sampling points. Without loss of generality we set the sampling time and the single-channel amplitude to 1. All calculations and simulations were carried out using MatLab Version 4.2c.1 (MATH WORKS Inc., Natick, MA) on a PC486-100. The electrophysiological experiments have been performed as described (Schmid et al., 1995; Höfer et al., manuscript submitted for publication). Single-channel data were idealized by using a t -test step detection algorithm (Pastushenko and Schindler, 1993, 1996).

Determination of the total number of channels, their availability, and their open probability

The total number of functional channels N in a membrane patch, their availability P_s , and their open probability p_o are essential parameters for calculation of the number of available channels per sweep, which will be shown in the next section. A maximum likelihood estimator has been recently

Received for publication 30 September 1996 and in final form 12 December 1996.

Address reprint requests to Dr. Christoph Romanin, Institute for Biophysics, University of Linz, Altenbergerstrasse 69, A-4040 Linz, Austria. Tel.: 732-2468-9272; Fax: 0732-2468-822; E-mail: christoph.romanin@jk.uni-linz.ac.at

© 1997 by the Biophysical Society

0006-3495/97/03/1143/10 \$2.00

established for the determination of N , P_s , and p_o (Schmid et al., 1995). Hence it will be referred to as p_o - P_s estimation and will be briefly described.

The following assumptions are made:

1. The channels are identical and behave independently.
2. The total number of channels remains constant.
3. The process regulating channels' opening and closing behavior is distribution ergodic with respect to the number of sampling points in each conductance level (no inactivation).

The variables are defined as: n_i , available channels in the i th sweep; t_{ik} , number of sampling points in conductance level k and sweep i ; \mathbf{T}_i , vector of t_{ik} for sweep i ; \mathbf{T} , matrix of all t_{ik} ; T , number of sampling points in a sweep; M , number of sweeps; A , single-channel amplitude.

Because of assumptions 1 and 2, the probability of finding an arbitrary sampling point of the i th sweep in the k th conductance level may be described by a binomial distribution with the parameters n_i and open probability p_o . Because of assumption 3, the elements of \mathbf{T}_i comprising one sweep obey a multinomial distribution with the probability for \mathbf{T}_i , given n_i and p_o ,

$$\Pr(\mathbf{T}_i | n_i, p_o) = \left\{ \prod_{k=0}^{n_i} \left[\binom{n_i}{k} \cdot p_o^k \cdot (1 - p_o)^{n_i - k} \right]^{t_{ik}} \right\} \cdot \frac{T!}{\prod_k t_{ik}!}, \quad (1)$$

if n_i is higher than or equal to the highest conductance level in sweep i , otherwise $\Pr(\mathbf{T}_i | n_i, p_o) = 0$. As the number of available channels is binomially distributed with parameter N and P_s , the probability for \mathbf{T}_i given N , P_s , and p_o is

$$\Pr(\mathbf{T}_i | N, P_s, p_o) = \sum_{n_i=0}^N \Pr(\mathbf{T}_i | n_i, p_o) \cdot \binom{N}{n_i} \cdot P_s^{n_i} \cdot (1 - P_s)^{N - n_i}, \quad (2)$$

leading to the probability of \mathbf{T} comprising all sweeps,

$$\Pr(\mathbf{T} | N, P_s, p_o) = \prod_{i=1}^M \Pr(\mathbf{T}_i | N, P_s, p_o). \quad (3)$$

Maximizing this likelihood leads to an estimator for N , P_s , and p_o for a given \mathbf{T} . This maximization is carried out using a nonlinear optimization algorithm including a penalty method, which requires start values for P_s and p_o . When analyzing simulated as well as original data, we never obtained a dependence of the results on the initial guesses. This has been extensively tested and has already been published with original data from L-type Ca^{2+} channels (Schmid et al., 1995; Groschner et al., 1996).

To evaluate the validity of the assumptions and the parameters calculated, a χ^2 test has been set up based on the maximum number of overlapping channels in each sweep. If assumptions 1–3 are fulfilled, the probability that n_{\max}

channels open simultaneously in the i th sweep given n_i is

$$\Pr(n_{\max} | n_i) = \left[\sum_{k=0}^{n_{\max}} \binom{n_i}{k} p_o^k (1 - p_o)^{n_i - k} \right]^T - \left[\sum_{k=0}^{n_{\max}-1} \binom{n_i}{k} p_o^k (1 - p_o)^{n_i - k} \right]^T. \quad (4)$$

Thus the probability for n_{\max} is

$$\Pr(n_{\max}) = \sum_{n_i=n_{\max}}^N \Pr(n_{\max} | n_i) \cdot \binom{N}{n_i} \cdot P_s^{n_i} \cdot (1 - P_s)^{N - n_i}. \quad (5)$$

The theoretical distribution $\Pr(n_{\max})$ can be compared with the empirical distribution using the χ^2 test (Papoulis, 1991) with $N - 2$ degrees of freedom (if $N > 2$).

Determination of the number of available channels in a sweep employing a Bayesian estimator

An idealized sweep is represented by $y(t)$, which may assume the values $0, A, 2A, \dots, nA$, with A and n corresponding to the single-channel amplitude and the number of available channels, respectively. Using the Bayes formula, the a posteriori probability for n available channels at given sweep $y(t)$ is

$$\Pr(n | y(t)) = \frac{\Pr(y(t) | n) \cdot \Pr(n)}{\sum_n (\Pr(y(t) | n) \cdot \Pr(n))}. \quad (6)$$

The a priori probability for n available channels in a sweep, given the availability P_s and the total number of channels N , as determined by p_o - P_s estimation described in the previous section, obeys a binomial distribution

$$n: B(N, P_s) \Rightarrow \Pr(n) = \binom{N}{n} \cdot P_s^n \cdot (1 - P_s)^{N - n}. \quad (7)$$

If the gating mechanism is known and corresponds to a Markov process, and if the matrix of transition probabilities is also known, the conditional probability $\Pr(y(t) | n)$ may be calculated by using Kolmogoroff's forward equation (Kolmogoroff, 1931). This method was found, however, to be extremely sensitive to wrong model assumptions (data not shown) and requires quite a lot of calculation time. To circumvent these constraints, we calculate the probability for an indicator function \bar{n} , which is the mean number of open channels

$$\bar{n} := \frac{1}{T \cdot A} \cdot \int_0^T y(t) dt. \quad (8)$$

In general, $\Pr(\bar{n} | n)$ is proportional to $\Pr(y(t) | n)$. The distribution of \bar{n} for one channel is approximately a β distribu-

tion, defined by

$$\beta_{a,b}(\bar{n}) := \frac{\Gamma(a+b)}{\Gamma(a) \cdot \Gamma(b)} \cdot (1-\bar{n})^{a-1} \cdot \bar{n}^{b-1} \quad (9)$$

for $0 < \bar{n} < 1$,

with parameters a and b , which have yet to be determined. These parameters are related to the expected value and variance by

$$E_{\bar{n}} := E(\bar{n}|n=1) = \frac{b}{a+b} \quad (10)$$

and

$$V_{\bar{n}} := \text{Var}(\bar{n}|n=1) = \frac{a \cdot b}{(a+b)^2 \cdot (a+b+1)}.$$

As the channels are independent and identical, the variance and the expected value are additive, leading to $E(\bar{n}|n) = n \cdot E_{\bar{n}}$ and $\text{Var}(\bar{n}|n) = n \cdot V_{\bar{n}}$. The resulting distributions of \bar{n} for general n are again β -shaped distributions but on the intervals $[0, n]$, i.e., \bar{n}/n is β -distributed with expected value $E_{\bar{n}}$ and variance $V_{\bar{n}}/n$, leading to

$$\text{Pr}(\bar{n}|n) = \frac{1}{n} \cdot \beta_{a(n),b(n)}\left(\frac{\bar{n}}{n}\right) \cdot d\bar{n}. \quad (11)$$

With n -dependent parameters,

$$a(n) = n \cdot \frac{[V_{\bar{n}}/n \cdot E_{\bar{n}}^2 - E_{\bar{n}}] \cdot [E_{\bar{n}} - 1]}{V_{\bar{n}}} \quad (12)$$

$$b(n) = n \cdot \frac{[V_{\bar{n}}/n \cdot E_{\bar{n}}^2 - E_{\bar{n}}] \cdot E_{\bar{n}}}{V_{\bar{n}}}.$$

Using $\text{Pr}(\bar{n}|n)$ instead of $\text{Pr}(y(t)|n)$, it is possible to calculate the a posteriori probability $\text{Pr}(n|y(t))$ (actually $\text{Pr}(n|\bar{n})$) as defined above in Eq. 6. The remaining problem is the determination of $E_{\bar{n}}$ and $V_{\bar{n}}$, which will be shown below.

Indirect method for calculation of the a posteriori probability of the number of channels per sweep

The variance and the expected value of \bar{n} are easy to find using the autocorrelation function $R_{yy}(\tau)$ of $y(t)$ (see Appendix). As we use only $R_{yy}(\tau)$, the method is rather insensitive to errors in model assumptions.

For approximation of the autocorrelation function of $y(t)$, it was found sufficient to use a two-state Markov process for most kinds of kinetic schemes:

$$C \xrightleftharpoons[\lambda_1]{\lambda_2} O.$$

It is important to note that the assumption of this model is used only to calculate the variance of \bar{n} , rendering the method largely insensitive to violations of the two-state scheme. The Markov model is employed here only as a tool

to approximate the decay length of the process underlying $y(t)$. Having the open probability p_o from the p_o - P_s estimation and an estimated value for the mean open time τ_o , i.e., the rate constant $\lambda_1 := 1/\tau_o$, we may approximate the rate constant λ_2 by

$$p_o = E(\bar{n}) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \Rightarrow \lambda_2 = \frac{p_o \cdot \lambda_1}{1 - p_o}. \quad (13)$$

Applying the formalism as described in the Appendix yields

$$E_{\bar{n}} = p_o \quad \text{and} \quad (14)$$

$$V_{\bar{n}} = \frac{\lambda_1 \lambda_2 \cdot [e^{-2T(\lambda_1 + \lambda_2)} + 2\lambda_1 T + 2\lambda_2 T - 1]}{2T^2 \cdot (\lambda_1 + \lambda_2)^4}.$$

Insertion of Eq. 14 in Eq. 12 leads finally to $\text{Pr}(\bar{n}|n)$.

Direct method for calculation of the a posteriori probability of the number of channels per sweep

For the case where assumptions 1 and 2 about ergodicity and stationarity are strongly violated, a second approach was developed that is applicable to experiments with >300 sweeps exhibiting channel activity. In this case, the β -distributions for $\text{Pr}(\bar{n}/n|n)$ could be fitted directly by using a maximum likelihood approach. This is possible because the β -distribution was found to be an excellent approximation for $\text{Pr}(\bar{n}/n|n)$ in the case of identical and independent channels lacking excessive moding. The log-likelihood function for one sweep is

$$L(\bar{n}|E(\bar{n}), P_s, \text{Var}(\bar{n})) = \sum_{n=0}^N \log \left[\frac{\text{Pr}(n)}{n} \cdot \beta_{a(n),b(n)}\left(\frac{\bar{n}}{n}\right) \right]. \quad (15)$$

Maximizing the sum of the $L(\bar{n}|p_o, P_s, V_{\bar{n}})$ values obtained from all sweeps leads to estimators for $E_{\bar{n}}$ ($= p_o$), P_s , and $V_{\bar{n}}$. These values are used to calculate $\text{Pr}(\bar{n}|n)$ as described above, leading to the a posteriori probability $\text{Pr}(n|y(t))$. This method is referred to as direct method, because the parameters p_o , P_s , and $V_{\bar{n}}$ were fitted directly to the experimental data, whereas the indirect method used the mean open time and required a kinetic assumption. To obtain good computational performance it was found convenient to use the values obtained from the p_o - P_s estimation and the indirect method as start values for the direct method.

Maximum likelihood estimator for the first latency

If standard kinetic assumptions are made, the first latency is multiexponentially distributed (Sigworth and Zhou, 1992). If the patch contains only one channel, the times to the first opening t_1 obey to the complementary distribution function

$$\text{Pr}(t_1|\vartheta) = \sum_j c_j \cdot e^{-(t_1/\tau_j)} \quad \text{with} \quad \sum_j c_j = 1, \quad (16)$$

for given parameter vector ϑ , which is

$$\vartheta = (c_1 \ c_2 \ \dots \ \tau_1 \ \tau_2 \ \dots). \tag{17}$$

In the case of n identical and independent channels, the complementary distribution function for t_1 follows as (Aldrich et al., 1983)

$$\Pr(t_1|\vartheta, n) = \left[\sum_j c_j \cdot e^{-(t_1 \cdot n/\tau_j)} \right], \text{ where } \sum_j c_j = 1. \tag{18}$$

As we do not exactly know n , we have to deal with the a posteriori probability for n , yielding

$$\Pr(t_1|\vartheta) = \sum_n \Pr(t_1|\vartheta, n) \cdot \Pr(n|y(t)). \tag{19}$$

Maximizing the probability for all observed t_1 values yields the maximum likelihood estimator for the first latency according to the method of Sigworth and Zhou (1992). To obtain for comparison the empirical distribution function of first latency, one must multiply the found first latencies in each sweep by the most likely value of n calculated for each sweep by

$$\langle n \rangle = \sum_{n=0}^N n \cdot \Pr(n|y(t)). \tag{20}$$

This is correct for one time constant and yields a good approximation for two time constants. The multiexponential time distribution calculated by the maximum likelihood estimator can then be verified by comparison with the empirical distribution function.

In the case of short sweeps and rather long first latencies, one has to deal with two sided censored data. Thus $\Pr(t_1|\vartheta, n)$ must be calculated as described (Colquhoun and Sigworth, 1995).

Run analysis (test for slow gating)

A convenient value for run analysis of multichannel data is the difference of the number of channels between neighbored sweeps denoted as Δn . As we do not know the exact value of n for each sweep, we have to use the a posteriori probability $\Pr(n|y(t))$. Denoting the i th sweep $y_i(t)$, the empirical distribution of Δn is given by

$$\widetilde{\Pr}(\Delta n) = \frac{1}{M-1} \sum_{i=1}^{M-1} \sum_n \Pr(n|y_i(t)) \cdot \Pr(n + \Delta n|y_{i+1}(t)). \tag{21}$$

If the number of channels in neighbored sweeps is not correlated, i.e., there is no slow gating, the theoretical distribution of Δn follows from the binomial distribution of n as

$$\Pr(\Delta n) = \sum_n \Pr(n)^* \cdot \Pr(n + \Delta n)^*. \tag{22}$$

With the probability for assignment,

$$\Pr(n)^* = \int_0^\infty \Pr(n|\bar{n}) \cdot \sum_{n=0}^N \Pr(\bar{n}|n) d\bar{n}. \tag{23}$$

The null hypothesis H_0 that there is no correlation in the number of channels of neighbored sweeps can be verified by using the χ^2 test. If the test rejects based on the significance level α , there are significant runs.

Three standard estimators used for comparison

Three different methods (Horn, 1991) of estimating the number of available channels in each sweep were used for comparison with our methods. The first was the so-called maximum estimator (MAX), which estimates n by the maximum number of simultaneously overlapping channels n_{MAX} in each sweep. Of course, this estimator is biased, as n_{MAX} is always smaller than or equal to the real n . The other two estimators were Bayesian estimators for the parameter of a binomial distribution, as introduced by Günel and Chilko (1989). The number of channels is estimated as

$$GC(\alpha, \beta) = \frac{\int_0^\infty G_1(y) \cdot e^{-y} dy}{\int_0^\infty G_2(y) \cdot e^{-y} dy}, \tag{24}$$

with

$$\begin{aligned} G_1(y) &= g(y) \cdot [t(y)]^{\alpha+\beta} \\ G_2(y) &= G_1(y)/t(y) \\ t(y) &= \frac{y \cdot \bar{x}}{\alpha} + n_{max} \end{aligned} \tag{25}$$

$$g = \frac{\prod_{i=1}^m \prod_{j=0}^{i-1} [t(y) - j]}{\alpha + \sum_i i \prod_{j=1}^i [u(y) - j]}$$

$$u(y) = (\alpha + \beta) + m \cdot t(y).$$

Here m refers to the number of sampling points used, which have to be independent, and α as well as β is a parameter of the β -distribution reflecting the a priori knowledge of p_o . If p_o is not known, both α and β should be set to 1, yielding the second estimator GC(1, 1) (Horn, 1991), which means that the open probability is uniformly distributed in the intervall [0, 1]. If p_o is known, one can adjust α and β appropriately, yielding GC(α , β) as a third estimator (Horn, 1991).

In contrast to our described method for calculating $\Pr(n|y(t))$, the GC estimator requires independent sampling points. Therefore, the GC estimators were calculated using sampling points with a distance more than the correlation length dependent on open and closed times (Liebovitch and

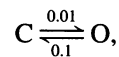
Fischbarg, 1985). This led to rather small sample sizes (5–20 for sweeps with 2000 sampling points), resulting in large errors in the estimation of available channels per sweep (n), as already described by Horn (1991). Nevertheless, it is worthy of note that the Bayesian estimators $GC(\alpha, \beta)$ are very useful for calculation of the total number of channels (N) in the patch.

RESULTS

The methods described in the previous section were put to a test employing simulated data as well as original recordings obtained from L-type Ca^{2+} channels. Simulations were carried out using different kinetic models for ion channels to evaluate the described methods.

Example 1

As a typical example that fulfills the assumptions made above, we simulated 140 sweeps with a total number of $N = 5$ channels. If the channel is available, it corresponds to the kinetic scheme



which means that the mean open time is $\tau_o = 1/\lambda_1 = 1/0.1 = 10$, and the mean closed time is $\tau_c = 1/\lambda_2 = 1/0.01 = 100$, leading to an open probability of 0.09 and a time constant for the first latency of $\tau = 100$. A slow gating with a time constant of 10 sweeps for the available as well as the unavailable state was programmed according to an availability of $P_s = 0.5$. Analysis of this example is shown in Fig. 1. Typical sweeps with different numbers of available channels can be seen in Fig. 1 A. Although $N = 5$, the maximum number of simultaneously overlapping channels in all 140 sweeps was 3. Application of the described p_o - P_s estimation (Schmid et al., 1995) yielded $N = 5$, $P_s = 0.58$, and $p_o = 0.08$ as the most likely result confirmed by the χ^2 test. If the number of available channels in each sweep is known, one could calculate the best estimator yielding $P_s = 0.52$ and $p_o = 0.09$, which is close to the results obtained by our p_o - P_s estimation.

Next we focused on determination of the available channels in each sweep. First, an open time analysis was performed using openings of nonoverlapping channels. Although the estimate of mean open time is biased, analysis of simulated data indicated that this bias will not change the results significantly. Then the indirect method was applied to calculate the posterior probability $\Pr(n|y(t))$ of the number of available channels, as exemplified in Fig. 1 B for each sweep in Fig. 1 A. It should be noted that the a posteriori probability is calculated for each possible channel number in a sweep. A diary plot of $\Pr(n|y(t))$ can be seen in Fig. 1 C, and a comparison of the simulated number of channels and the most likely number of channels calculated according to Eq. 20 is shown in Fig. 1 D. $\Pr(n|y(t))$ was then used for the first latency (τ) determination, yielding the empirical

and calculated distribution function as shown in comparison to the programmed distribution function (Fig. 1 E). Comparison of the optimal results for known n and the results obtained by the indirect method as well as the estimators MAX, $GC(1, 1)$, and $GC(1, 9)$ is given in Table 1. The results include the significance level α of the χ^2 test for run analysis. The α value calculated on the basis of a known number of channels (first column in Tables 1–3) represents the best estimate for evaluation of clustering of channel activity. If α is close to 1, the null hypothesis will be rejected and channel activity is apparently clustered, as clearly confirmed by all estimators except $GC(1, 9)$.

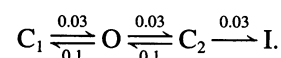
Example 2

In the previous example the kinetic scheme was known. Therefore, in this example we used experimental data obtained by standard patch-clamp recordings of cardiac L-type Ca^{2+} channels (Höfer et al., manuscript submitted for publication) expressed in Chinese hamster ovary cells (Welling et al., 1993). Our method was applied here to multichannel data artificially created from an experiment with a single channel comprising 596 sweeps. For the original single-channel data, P_s and p_o were calculated as 0.67 and 0.119, respectively. A first latency determination yielded two time constants with $c_1 = 0.37$, $\tau_1 = 7.21$ ms, $\tau_2 = 63.25$ ms. Then 149 artificial multichannel data were created by adding four independent sweeps (sweep 1 + sweep 150 + sweep 299 + ...). This produces data with a known number of channels per sweep. Analysis of example 2 is presented in Fig. 2. A representative set of artificial multichannel sweeps is depicted in Fig. 2 A, with the corresponding a posteriori probability $\Pr(n|y(t))$ in Fig. 2 B. A diary plot of $\Pr(n|y(t))$ is depicted in Fig. 2 C, and a comparison of the known numbers of available channels and the most likely number of channels is given in Fig. 2 D. Our results, together with those obtained by the other estimation methods, are listed in Table 2. Clustering of channel activity is unlikely in this example, as is evident from α calculated for known n and consistently found using the direct method or the MAX estimator. The remaining estimators yielded false results.

In the examples presented so far, all assumptions made for the indirect method were approximately fulfilled. The direct method as described below should not be used for these examples, as only 140 or 149 sweeps were available.

Example 3

As a third example we simulated 500 sweeps using the following kinetic scheme with two closed states and an inactivated state:



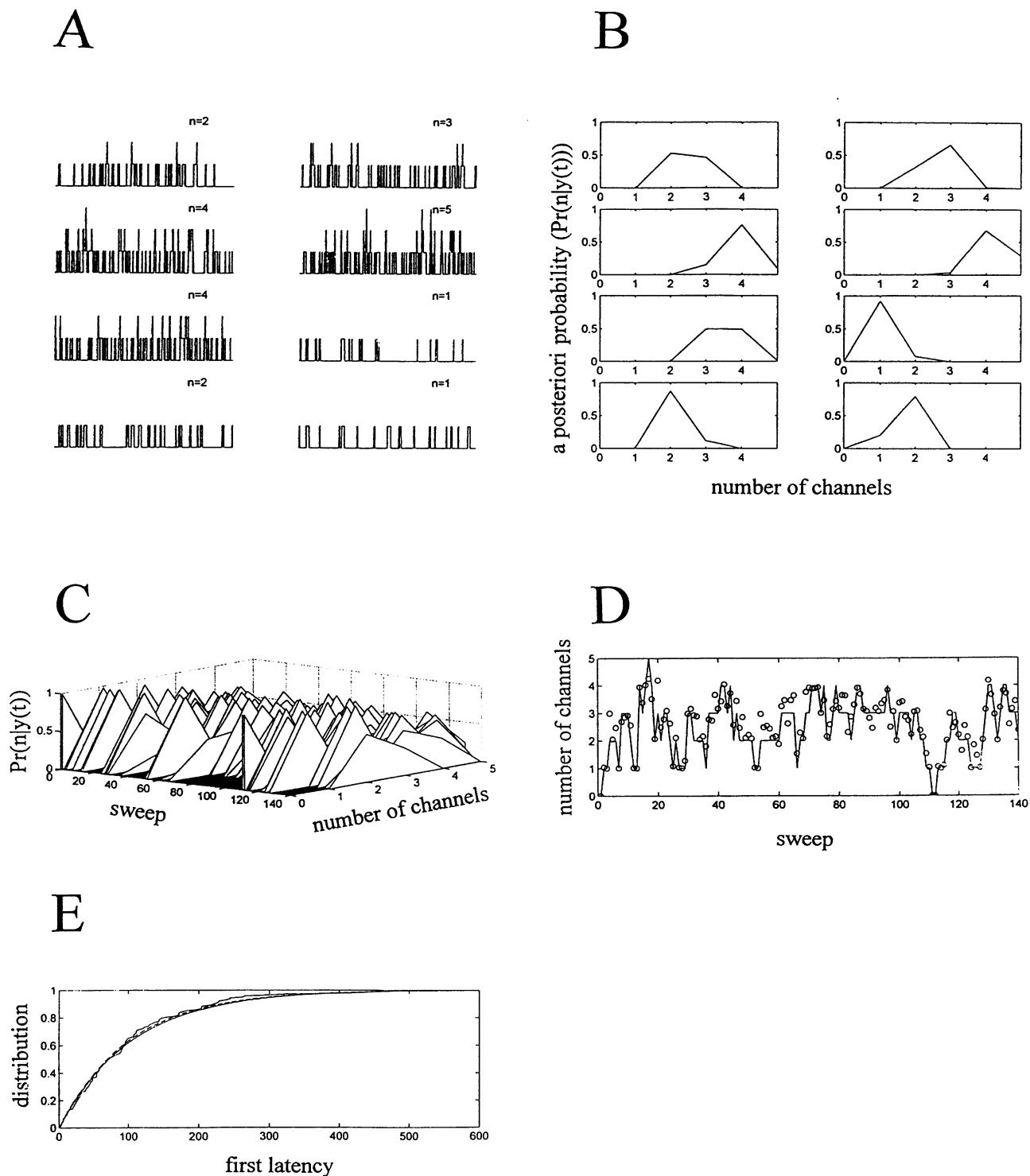


FIGURE 1 Analysis of 140 simulated sweeps corresponding to the scheme $C \rightleftharpoons O$. (A) Typical sweeps (idealized) with the programmed number of available channels as indicated. (B) A posteriori probability $\Pr(n|y(t))$ of the number of channels calculated for each sweep as shown in A. (C) $\Pr(n|y(t))$ of all simulated sweeps. (D) Comparison of the programmed number of available channels and the most likely number of channels (○) calculated from \Pr for each sweep. (E) A plot of the empirical first latency distribution (solid line) obtained by using the most likely number of available channels compared to the calculated (solid, smooth line) and programmed (dashed line) distribution function.

Channel availability was set to $P_s = 0.5$, although in contrast to the first example, no slow gating was programmed, i.e., the number of available channels is not correlated

between neighbored sweeps. If the channel is available, it starts in state C_1 . Analysis of example 3 is presented in Fig. 3. Representative sweeps are depicted in Fig. 3 A, yielding

TABLE 1 Comparison of the results in Fig. 1, calculated for a known number of available channels (n) in each sweep with those obtained by the indirect method, the MAX estimator, and the GC(1,1) and GC(1,9) estimators

	Known n	IM	MAX	GC(1,1)	GC(1,9)
N	5	5	3	3	6
p_o	0.091	0.084	0.114	0.183	0.152
P_s	0.520	0.576	0.650	0.423	0.293
τ	100.4	104.1	81.0	47.5	97.0
α	1.0000	0.9998	1.0000	1.0000	0.3989

Results are from data as shown in Fig. 1, corresponding to the kinetic scheme $C \rightleftharpoons O$.

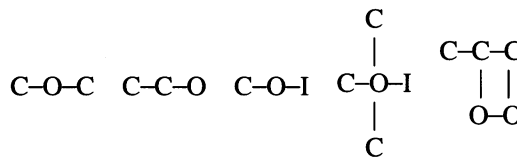
an ensemble average current with an inactivation of about 95%, as shown at the bottom. Such a strong inactivation would clearly violate assumption 3 made for the indirect method favoring application of the direct method. The $\Pr(n|y(t))$ for each sweep in Fig. 3 A is shown in Fig. 3 B, employing the direct and indirect (*dashed line*) method. A diary plot of the programmed number of channels per sweep is given in Fig. 3 C. Comparison of this programmed number of channels with that calculated by the different estimators is shown in Fig. 3 D. Specifically, the difference between the programmed and the calculated value is depicted. "Traces" with small deviations equally distributed around zero represent estimations of channel numbers close to the programmed ones, indicating that the direct method (DM) yields the best result. Table 3 shows a comparison of the results obtained by the different estimators. It is clearly evident that the direct and indirect methods yielded more accurate results than the other estimators.

DISCUSSION

The indirect and direct methods presented in this study allow for estimation of the number of available channels in individual depolarizations of a multichannel experiment. Moreover, we demonstrate that these methods can be used for a reasonable determination of first latency distribution as well as for run analysis. A comparison with three standard estimators shows that the described methods lead in most cases to more accurate results.

Multichannel data are often obtained because of substantial clustering of ion channels in most physiological systems. The described methods allow for analysis of ion channel kinetics from such recordings. Microheterogeneity in the channels leading to violation of assumption 1 (Dabrowski et al., 1990) and excessive moding (Hess et al., 1984) of channel activity might produce errors in the estimation of the number of channels in our methods as well as the other estimators (Horn, 1991). Nonetheless, multichannel data from L-type Ca^{2+} channel recordings have been successfully analyzed by the p_o - P_s estimation (Schmid et al., 1995) as well as the methods developed here (Höfer et al., manuscript submitted for publication). Furthermore, the described methods were carefully tested to examine their

applicability, by using simulated data based on a variety of different kinetic models. Analysis of kinetic schemes such as



together with variable transition probabilities suggests wide applicability.

Application guide

A diary plot depicting mean channel activity (Np) for each sweep should be initially used to decide how to proceed with the analysis methods. Specifically, experiments with more than 300 sweeps under each experimental condition should be analyzed by the direct method. The indirect method may be used for shorter experiments with data comprising >100 sweeps. Furthermore, channel recordings with strong inactivation as revealed by ensemble average currents should preferentially use the direct method.

To start with the indirect method, determination of mean open time and N , P_s , as well as p_o calculated by the p_o - P_s estimator (Schmid et al., 1995) is initially performed. In the case of strong inactivation, the total number of channels N should be determined by the GC(α , β) estimator. These parameters are then used by the indirect method to calculate the a posteriori probability of the number of channels in each sweep. The direct method can be applied in any case provided that more than about 300 sweeps are recorded and will yield more accurate results. The a posteriori probabilities are then used for determination of first latencies as well as for run analysis and may be employed for any further analysis.

Advantage and disadvantage of different estimators

The MAX estimator is biased, as n_{MAX} is always smaller than or equal to the real number of available channels in a sweep. This effect is dominant in the case of small open probability ($\leq 10\%$), which impairs further analysis. Use of the GC estimators is limited because of the requirement of independent sampling points. If the correlation length of channel gating is long compared to the sweep length, the number of sample points in each sweep used by the GC estimators is low. This results in large deviations in the calculated number of available channels, as evidenced in Fig. 3 D.

One further disadvantage of the standard methods in comparison to the indirect and direct methods is that they determine only one "optimal" value for the number of channels in a sweep, whereas our methods calculate a prob-

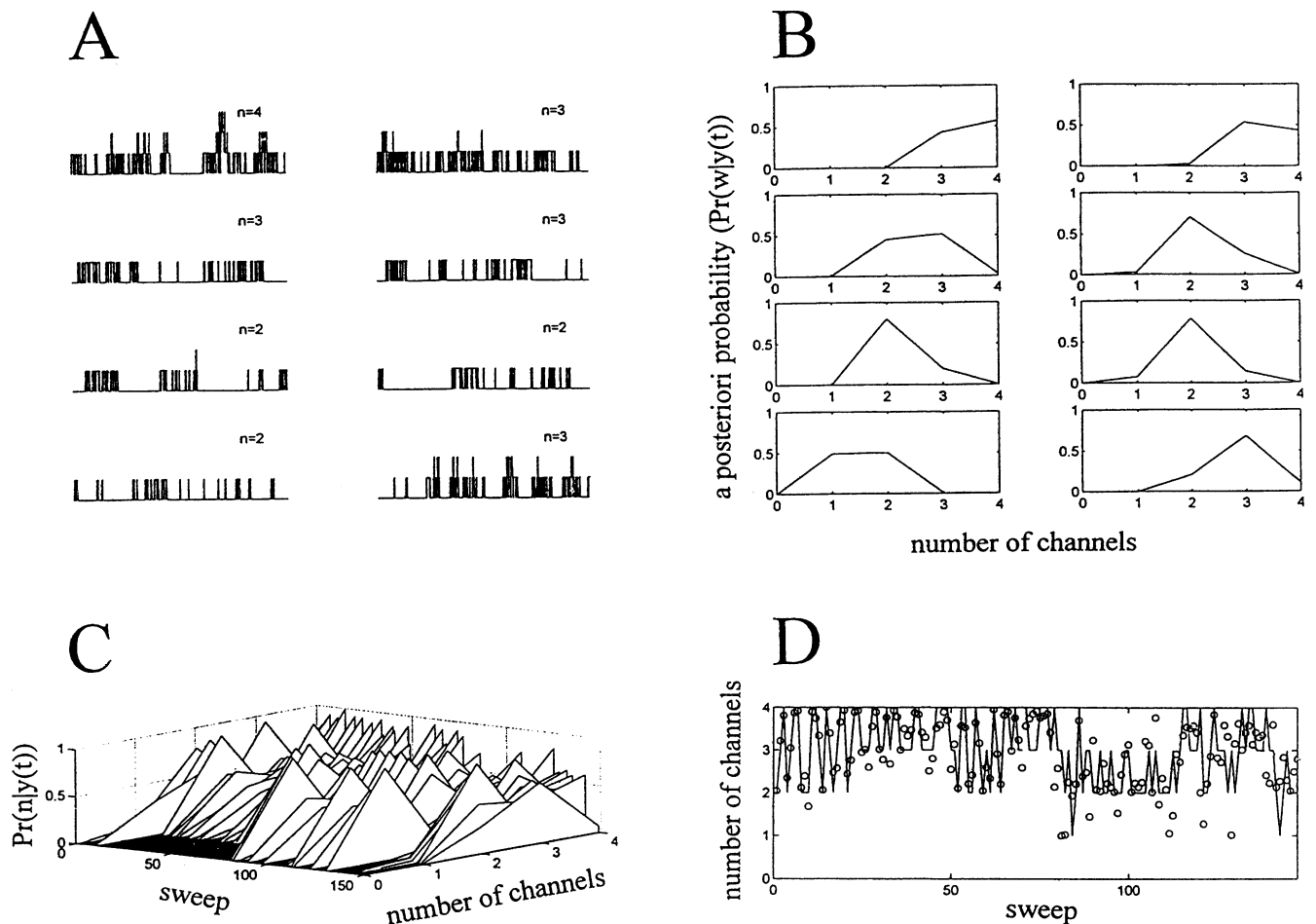


FIGURE 2 Analysis of 149 “artificial” multichannel sweeps from L-type Ca^{2+} channel recordings, each obtained by overlaying four single-channel sweeps. (A) Typical sweeps (idealized) with the known number of available channels as indicated. (B) A posteriori probability $\text{Pr}(n|y(t))$ of the number of channels calculated for each sweep as shown in A. (C) $\text{Pr}(n|y(t))$ of all sweeps. (D) Comparison of the known number of and the most likely number (○) of available channels in each sweep.

ability for each possible number of channels. Use of this a posteriori probability allows for most accurate results in further analysis of kinetic parameters, as demonstrated in the examples.

TABLE 2 Comparison of the results in Fig. 2, calculated for a known number of available channels (n) in each sweep with those obtained by the indirect method, the MAX estimator, and the GC(1,1) and GC(1,9) estimators

	Known n	IM	MAX	GC(1,1)	GC(1,9)
N	4	4	4	4	4
p_o	0.117	0.122	0.154	0.290	0.192
P_s	0.680	0.691	0.542	0.580	0.876
c	0.38	0.42	0.43	0.47	0.45
τ_1	7.01	7.46	4.98	2.05	4.56
τ_2	66.38	56.66	42.57	24.00	48.22
α	0.83	0.62	0.60	1.00	1.00

Typical data are shown in Fig. 2, obtained by overlaying four single-channel sweeps from L-type Ca^{2+} channel recordings.

It is anticipated that the described methods will help to reveal ion channel kinetics by enabling a detailed analysis of multichannel data.

APPENDIX

For convenience we assume the gating process to be wide sense stationary (WSS). This means that the mean of $y(t)$ is constant in time and the autocorrelation function $R_{yy}(t, \tau)$ is dependent only on τ . This was found to be approximately fulfilled in most types of experiments. As $y(t)$ is WSS, we find that

$$\text{Pr}(\bar{n}|n) = \text{Pr}(\tilde{n}|n), \quad \text{where} \quad \tilde{n} := \frac{1}{T \cdot A} \cdot \int_{-T/2}^{T/2} y(t) dt. \quad (26)$$

So \tilde{n} can be interpreted as the output of a system with input $y(t)/A$ and rectangular impulse response of height $1/T$ in the interval $[-T/2, T/2]$. The transfer function $G(\omega)$ of this system is the si-function:

$$G(\omega) = \text{si}(\omega T/2) := \frac{\sin(\omega T)}{\omega T}.$$

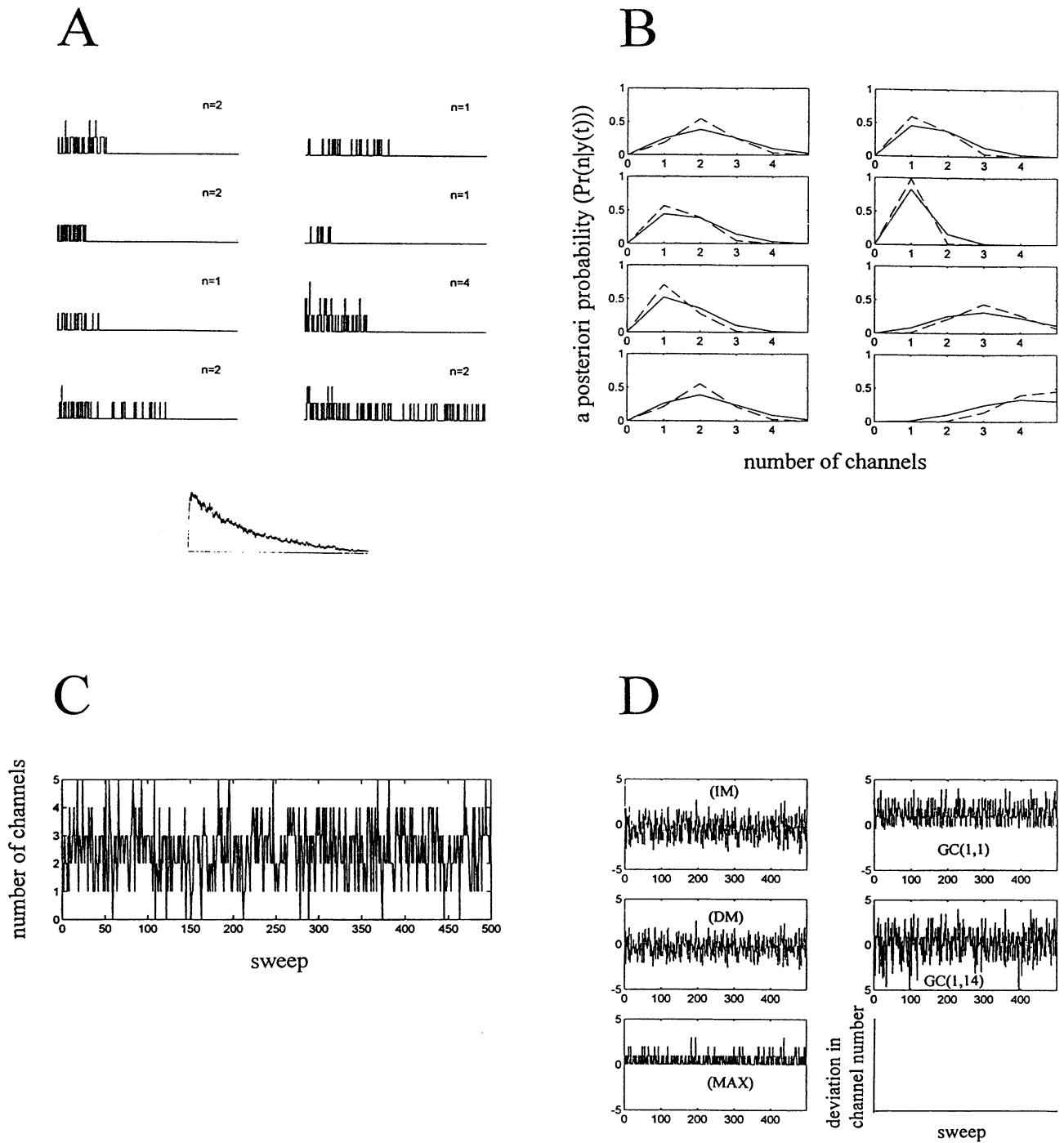


FIGURE 3 Analysis of 500 simulated sweeps corresponding to the scheme $C_1 \rightleftharpoons O \rightleftharpoons C_2 \rightarrow I$. (A) Typical sweeps (idealized) with the programmed number of available channels as indicated and the mean current at the bottom. (B) A posteriori probability $\Pr(n|y(t))$ of the sweeps in A calculated by the indirect (dashed line) and the direct (solid line) method. (C) Programmed number of available channels. (D) Programmed number of available channels minus number of channels found using the indirect method (IM), the direct method (DM), the MAX estimator, the GC(1, 1) estimator, and the GC(1, 14) estimator.

Knowing the autocorrelation function of the process underlying $y(t)$, one can calculate the autocorrelation function of \tilde{n} (Papoulis, 1991):

$$R_{\tilde{n}\tilde{n}}(\tau) = \frac{1}{T} \cdot \int_{-T}^T \left(1 - \frac{|t|}{T}\right) \cdot R_{yy}(t - \tau) dt, \quad (27)$$

leading to the expected value and the variance of \tilde{n} for $n = 1$,

$$E(\tilde{n}) = R_{\tilde{n}\tilde{n}}(\infty) \quad \text{and} \quad \text{Var}(\tilde{n}) = R_{\tilde{n}\tilde{n}}(0) - R_{\tilde{n}\tilde{n}}(\infty), \quad (28)$$

which are identical to the expected value and the variance of \tilde{n} . Defining

$$p_c := \frac{\lambda_1}{\lambda_1 + \lambda_2}, \quad (29)$$

TABLE 3 Comparison of the results in Fig. 3, calculated for a known number of available channels (n) in each sweep with those obtained by the indirect method, the direct method, the MAX estimator, and the GC(1,1) and GC(1,14) estimators

	Known n	IM	DM	MAX	GC(1,1)	GC(1,14)
N	5	5*	5	4	4	5
p_o	0.067	0.057	0.066	0.084	0.134	0.077
P_s	0.500	0.590	0.516	0.400	0.245	0.447
τ	37.2	40.2	36.2	28.1	20.9	36.3
α	0.88	0.84	0.80	1.00	1.00	1.00

Typical data are shown in Fig. 3, corresponding to the kinetic scheme $C_1 \rightleftharpoons O \rightleftharpoons C_2 \rightarrow I$.

* Applying the p_o - P_s estimation to this example yielded no optimum for the log-likelihood function at $N = 5$; it increased for higher N as expected, because of the violation of assumption 3. Not surprisingly, the χ^2 test rejected every possible N , rendering it indistinguishable, whether there were 4, 5, or 6 channels in the patch. Nevertheless, using the values obtained from the p_o - P_s estimation, one can set up a GC(1,27) estimator for the total number of channels, yielding $N = 5$.

we find the autocorrelation function of the stationary two-state Markov model to be

$$R_{yy}(\tau) = p_o^2 + p_o \cdot p_c \cdot e^{-(\lambda_1 + \lambda_2) \cdot \tau}, \quad (30)$$

leading to Eq. 14.

This work was supported by Austrian Research Funds S06605, S06606, and S06607.

REFERENCES

- Aldrich, R. W., D. P. Corey, and C. F. Stevens. 1983. A reinterpretation of mammalian sodium channel gating based on single channel recording. *Nature*. 306:436–449.
- Chung, S. H., J. B. Moore, L. Xia, L. S. Premkumar, and P. W. Gage. 1990. Characterization of single channel currents using digital signal processing techniques based on hidden Markov models. *Philos. Trans. R. Soc. Lond.* 329:265–285.
- 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 F. J. Sigworth. 1995. Fitting and statistical analysis of single-channel records. In *Single-Channel Recording*. Plenum Press, New York. 483–587.
- Dabrowski, A. R., D. McDonald, and U. Rösler. 1990. Renewal theory properties of ion channels. *Ann. Statist.* 18:1091–1115.
- Groschner, K., K. Schuhmann, G. Mieskes, W. Baumgartner, and C. Romanin. 1996. A type 2A phosphatase-sensitive phosphorylation site controls modal gating of L-type Ca^{2+} channels in human vascular smooth-muscle cells. *Biochem. J.* 318:513–517.
- Günel, E., and D. Chilko. 1989. Estimation of the parameter n of the binomial distribution. *Commun. Statistics. Simulation Computation*. 18: 537–551.
- Hess, P., J. B. Lansmann, and R. Tsien. 1984. Different modes of Ca channel gating behavior favored by dihydropyridine Ca agonists and antagonists. *Nature*. 311:538–544.
- Horn, R. 1991. Estimating the number of channels in patch recordings. *Biophys. J.* 60:433–439.
- Horn, R., C. Vandenberg, and K. Lange. 1984. Statistical analysis of single sodium channels. Effects of N -bromacetamine. *Biophys. J.* 45:323–335.
- Kolmogoroff, A. N. 1931. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* 104:415–458.
- Liebovitch, L. S., and J. Fischbarg. 1985. Determining the kinetics of membrane pores from patch clamp data without measuring the open and closed times. *Biochim. Biophys. Acta.* 813:132–136.
- Papoulis, A. 1991. Probability, Random Variables and Stochastic Processes. McGraw-Hill, New York.
- Pastushenko, V. Ph., and H. Schindler. 1993. Statistical filtering of single ion channel records. *Acta Pharmaceutica.* 43:7–13.
- Pastushenko, V. Ph., and H. Schindler. 1996. Level detection in ion channel records via idealization by statistical filtering and likelihood optimization. *Philos. Trans. R. Soc. B.* (in press).
- Romanin, C., P. Grösswagen, and H. Schindler. 1991. Calpastatin and nucleotides stabilize cardiac calcium channel activity in excised patches. *Pflugers Arch.* 418:86–92.
- Schmid, R., K. Seydl, W. Baumgartner, K. Groschner, and C. Romanin. 1995. Trypsin increases availability and open probability of cardiac L-type Ca^{2+} -channels without affecting inactivation induced by Ca^{2+} . *Biophys. J.* 69:1847–1857.
- Sigworth, F. J., and J. Zhou. 1992. Analysis of nonstationary single-channel currents. *Methods Enzymol.* 207:746–762.
- Welling, A., E. Bosse, A. Cavalié, R. Bottlender, A. Ludwig, W. Nastainczyk, V. Flockerzi, and F. Hoffmann. 1993. Stable co-expression of calcium channel α_1 , β and α_2/δ subunits in a somatic cell line. *J. Physiol. (Lond.)*. 471:749–765.