Eur Biophys J (1997) 26: 393–404 © Springer-Verlag 1997

ARTICLE

V. P. Pastushenko · H. Schindler
A. Moghaddamjoo Reza

Predictive optimality criterion for idealization of ion channel data and exact Akaike's criterion

Received: 23 July 1996 / Accepted: 9 May 1997

Abstract Different statistical or low-pass filters may be used for the idealization of ion channel data. We address the problem of predicting optimal filter parameters, represented by a threshold test-value for statistical filters and by a cut-off frequency for low-pass filters. Optimal idealization is understood in the sense of maximal similarity between recovered and real signals. Special procedures are suggested to quantitatively characterize the difference between the recovered and the real signals, the latter being known for simulated data. These procedures, called objective criteria, play the role of referees in estimating the performance of different predictive optimality criteria. We have tested standard Akaike's AIC and its modification by Rissanen, MDL. Both gave unsatisfactory results. We have shown analytically, that the Akaike-type criterion, based on the use of a certain penalty for the log likelihood function per transition, indicates the correct optimum point only if the penalty is set equal to half the optimal threshold. As the latter varies significantly for different data sets, this criterion is not particularly helpful. A new universal predictive optimality criterion, valid for real data and any idealization method, is suggested. It is formally similar to AIC, but instead of log likelihood it uses the doubled number of false transitions. The predictive power of the new criterion is demonstrated with different types of data for Hinkley and 50% amplitude methods.

Key words Threshold detection · Statistical filtering · Model ranking

List of notations A_q – Akaike-type criterion with the penalty per transition q · AIC – Asymptotic Information Criterion · B – boundary (midpoint) between the levels · D – expected dwell time (for two-state system the same for open and closed states) · D_k – dwells of real events · $\underline{D}_m$ – dwells of recovered events · E = error of the idealization in terms of mismatching events · F – number of false events in the record · f – likelihood of an ideal trace being recovered at a certain T-value · h – marking Hinkley-variable · H – state Hinkley-variable · i – index of data points; the position of the detector, i = 1, 2, …, N · i_s – the starting position of the detector · ITC – Information Theoretic Criteria · j – index of a current segment, j = 1, 2, …, J · J – total number of segments in a record · k – index of real events, k = 1, 2, …, K · K – total number of real events in a record · L – number of lost events · L_1, L_2 – level values (here 0 and 1 respectively) · m – index of recovered events, m = 1, 2, …, M · M – total number of recovered events in a record · M_x – number of recovered events declared by a corresponding criterion as optimal one (for a specific criterion, x = $\omega, \gamma, \alpha, \rho$) · MDL – minimum description length, same as R · n, $\underline{n}$ – noise and its estimate · N – total number of data points in the record r · p_j – a proportion of an event dwell, corresponding to segment j · P – number of free parameters in a model (here number of transitions) · q – penalty for log likelihood per transition · r – record · r_i – data point number i · R – the value of MDL · s_i, $\underline{s}_i$ – signal and its estimate (L_1 or L_2) in i-th point · S_x, $\underline{S}_x$ – signal and its estimate in x-th event or segment (x = k, m or j) · SNR – signal to noise ratio: SNR = Δ/σ · t – Hinkley test-variable · T – threshold value of t · T_x – values of T declared as optimal (for a specific criterion, x = $\omega, \gamma, \alpha, \rho$, 1) · $\alpha_q = A_q/N$-Akaike's potential, Eq. (8) · γ – matching potential (number of matching events per data point), Eq. (14) · Γ – predictive optimality criterion (predicted matching potential), Eq. (36) · δ – Kronecker delta-symbol · Δ – channel amplitude, $\Delta = L_2 - L_1$ · ε – mismatching potential, Eq. (13) · λ – potential of lost events · $\mu = M/N$ – potential of recovered events · $\mu_0 = K/N$ – potential of real events · μ_x – potential of recovered events declared by a corresponding criterion as optimal one (for a specific criterion, x = $\omega, \gamma, \alpha, \rho$) · Φ – average number of transitions found in a record of pure noise of length N · $\varphi = \Phi/N$

V. P. Pastushenko (✉) · H. Schindler
Institute for Biophysics, Johannes Kepler University of Linz, A-4040 Linz, Austria
(e-mail: Vassili.Pastushenko@jk.uni-linz.ac.at)

A. Moghaddamjoo Reza
Electrical Engineering, University of Wisconsin-Milwaukee, P.O. Box 784, Milwaukee, WI 53217-0784, USA

potential of false events (detected in pure noise) · ξ – normalized record r · ρ=R/N – Rissanen's potential · σ, $\underline{\sigma}$- noise amplitude (SD) and its estimate · θ – set of transition coordinates · η – Heavyside, or step function (one for positive and zero for negative arguments) · ω – potential of matching points, Eq. (11).

Introduction

In the last decade, the analysis of ion channel data has evolved remarkably (Colquhoun and Hawkes 1995 a; Colquhoun and Sigworth 1995; Heinemann 1995; Magleby 1992). Earlier approaches towards analysis of channel mechanisms involved first model identification and then parameter estimation. This is now changing to direct model fitting to data (Horn and Lange 1983; Bauer et al. 1987; Chung et al. 1990; Chung and Kennedy 1991; Fredkin and Rice 1992 a; Fredkin and Rice 1992 b; Becker et al. 1994; Albertsen and Hansen 1994; Qin et al. 1996). In these methods different models are first fitted to the given data, and then the best model, in terms of its likelihood, is selected. In this approach, Information Theoretic Criteria (ITC), such as Asymptotic Information Criterion, AIC (Akaike 1974), or Minimum Description Length, MDL (Rissanen 1983) are usually used to identify the optimum number of free parameters in the proposed model.

Markov models (Rabiner 1989) are superior to any competition, partly owing to their flexibility and convenience (Colquhoun and Hawkes 1995 a; Colquhoun and Hawkes 1995 b), and partly owing to their advantages in comparison with alternative models, demonstrated with the help of AIC (Horn 1987; Korn and Horn 1988; McManus et al. 1988). Thus ITC appear to have become the basis for the modern analysis of channel data. At the same time, many important results are based on idealization of the data (Horn and Lange 1983; Magleby and Weiss 1990 a; Magleby and Weiss 1990 b), as well as the test for the Markov property itself (Petracchi et al. 1991). Optimum idealization of channel data became especially important with the development of model fitting to idealized traces (Horn and Lange 1983; Qin et al. 1996). Again, AIC is suggested as a ranker for different channel mechanisms fitted to idealized data.

Classical 50% amplitude idealization (Colquhoun and Sigworth 1995) is based on the low-pass filtering of data, which may be done while recording the current, or afterwards. Low-pass filtering distorts the signal, and the event dwells need certain corrections. This decreases the elegance and the performance of the method. Different statistical filters, avoiding dwell aliasing, have been suggested: in the case of unknown levels and noise amplitude (noise amplitude is characterized by standard deviation, SD), F-test (Kirlin and Moghaddamjoo 1986; Moghaddamjoo 1989, 1991 a, b) or Student t-test (Pastushenko and Schindler 1992, 1997); in the case of the known noise SD, Gauss-test (Pastushenko and Schindler 1993); and in the case of no drift, when levels and noise SD are known and only transitions are unknown – Hinkley-test (Schultze and Draber 1993, 1994 a, b). In the context of transition finding, statistical tests differ from their classical interpretation. The corresponding modification for the Gaussian test is given by Pastushenko et al. (1996). The indicated choice of idealization methods reflects a simple rule: the less we have to guess, the fewer errors we make, and the more we know, the less we have to guess. Correspondingly, a smaller number of unknown parameters may be more reliably estimated from the same data (Hafner 1989).

A common feature for all idealization methods is the presence of a certain threshold, T, represented by the cutt-off frequency of a low-pass filter and by the threshold test values (or percentiles) of statistical filters. Selection of an optimum T-value, denoted by T_*, is an objectively difficult problem, owing to the fact that in general it depends on signal/noise ratio (SNR, defined as the ratio of channel amplitude to noise amplitude), event dwells and their correlations (cf. Results). The recommendations of standard T_*-values (Hinkley-test: Schultze and Draber 1993, 1994 a, b; cutt-off frequency of low-pass filter to reach a standard SNR: Colquhoun and Sigworth 1995) should be considered with due caution, because these recommendations are valid only for a specific class of data.

Within the problem of idealization, the situation is analogous to global searching for the channel mechanism. Thus, the ideal traces, obtained from the same data at different T-values, correspond to different statistical models of the signal. Within the limitations of each method, the likelihoods of these models are locally maximized with respect to variations of each transition. Therefore, searching for the T_*-value represents another way of maximizing the likelihood of the whole trace. For this reason, we have decided to test ITC-type criteria as model rankers within the idealization problem (recently, Moghaddamjoo (1991 b) suggested MDL as a tool for selecting the optimum number of events). To concentrate on the central question, i.e. of optimum idealization, we have assumed that the data have no drift and that the levels and the noise amplitude are known, which legitimizes the selection of Hinkley or 50% amplitude methods.

This study was initiated as an attempt to compare the predictive powers of MDL and AIC. The results we have obtained go beyond the question of the importance of optimal threshold selecting.

Methods

General

This section describes the different tools used to decide on the performance of different predictive criteria of optimality of idealization. Two known criteria, standard Akaike's AIC and Rissanen's MDL are uniformly described as two particular cases of a general Akaike-type criterion, corresponding to specific selections of the penalty per transition. The objective optimality criteria are introduced.

They allow one to find optimal filter parameters by comparison between recovered and real traces. These criteria play the role of referees in a competition between different predictive criteria. Four different types of data for a channel with two kinetic states were used. The demonstration is completed with data for a three-state system. Two different methods of idealization, the standard 50% amplitude method and the Hinkley detector are used. A short symmetric description of the Hinkley detector is given.

An optimal detection threshold T_* corresponds to the best ideal trace. The definition for the best ideal trace is implicitly given by Colquhoun and Sigworth (1995): “it is important that the idealized record be as complete and unbiased as possible …”, i.e. the difference between the recovered trace and the real one should be minimal. This difference can be understood in the sense of squared deviation, which leads to a criterion in terms of points (criterion ω), or in the sense of mismatching events, leading to a criterion in terms of events (criterion γ). These criteria serving as objective measures for optimality produce slightly different results. To be specific, by T_* we shall mean in most cases the optimal T-value, indicated by the γ-criterion, which is more appropriate for channel data analysis. To avoid confusion, the values of T claimed by different predictive criteria as optimal ones will be denoted separately.

Information theoretic criteria

A record of channel current, r, in digitized form represents a time series containing N “observations”:

$$r = \{r_1, r_2, \ldots, r_N\} . \tag{1}$$

Each observation is a sum of original signal s_i and noise n_i, or their estimations $\underline{s}_i$ and $\underline{n}_i$:

$$r_i = s_i + n_i = \underline{s}_i + \underline{n}_i , \quad i = 1, 2, \ldots, N \tag{2}$$

The second equality here is correct for noise estimation via the signal estimation, $\underline{n}_i = r_i - \underline{s}_i$. Both tested criteria, standard Akaike’s AIC and Rissanen’s MDL represent two specific cases of a general Akaike-type criterion A_q, corresponding to different values of the penalty per transition, $q > 0$:

$$A_q = -2 \ln f(r|\theta) + 2qP , \tag{3}$$

$f(r|\theta)$ is the likelihod function of the statistical model θ, represented in our case by P coordinates of transitions between events, considered as free parameters. Thus, the standard Akaike’s AIC is the same as A_q when $q = 1$, and Rissanen’s MDL is the same as A_q when $q = \ln(\sqrt{N})$. For this reason the standard AIC will be denoted A_1; MDL will be denoted R. The number of transitions, P, is one less than the number of recovered events, M:

$$P = M - 1 . \tag{4}$$

For white Gaussian noise with known variance σ^2, the expression for the likelihood function is (Moghaddamjoo 1991 b):

$$f(r|\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r_i - \underline{s}_i)^2}{2\sigma^2} \right). \tag{5}$$

As an optimality criterion, A_q may be replaced by any monotonic function of it (i.e. by a function with nonzero derivative), because this does not influence the position of the optimum. In particular, A_q may be subjected to an arbitrary linear transformation, i.e. it may be multiplied or increased by any nonzero real number, which should be independent of the model parameters. Therefore, after taking the logarithm of f and omitting the terms independent of M, Eq. (3) produces an equivalent expression for A_q:

$$A_q = N \underline{\sigma}^2/\sigma^2 + 2qM . \tag{6}$$

Here $\underline{\sigma}^2$ is the estimate of the noise variance for a given T-value:

$$\underline{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} (r_i - \underline{s}_i)^2 . \tag{7}$$

The quantity A_q represents, in a thermodynamic sense, an “extensive” property of the record, i.e. its value depends on data size. Therefore, as an asymptotic criterion, it is more consistent to consider an “intensive” property α_q, analogous to chemical potential and defined as:

$$\alpha_q = A_q/N = \underline{\sigma}^2/\sigma^2 + 2q\mu , \tag{8}$$

where μ stands for “event potential”,

$$\mu = M/N . \tag{9}$$

We shall also use the notations α_1 and α_0 for α_q at $q = 1$ and $q = 0$ respectively. Within the likelihood concept, N is considered as constant. Therefore, expression (8) is equivalent to expression (6). Correspondingly, Rissanen’s potential is

$$\rho = R/N = \underline{\sigma}^2/\sigma^2 + \ln(N)\mu . \tag{10}$$

Normalization per data point makes it immediately clear that ρ is not an asymptotic criterion, because it depends on data size.

Objective criteria of optimality

In order to decide, which of the predictive optimality criteria has higher performance, one needs an objective measure for the quality of the ideal trace, indicated by a corresponding criterion as an optimal one. The simplest criterion of this kind is based on the mean-square of the estimation error, found by comparison between the recovered and the real traces. The optimality is expressed in terms of maximum number of matching points. By definition, the matching takes place if the signal estimation coincides with real signal. This criterion is defined as the potential of matching points, ω:

$$\omega = \frac{1}{N} \sum_{i=1}^{N} \delta(s_i, \underline{s}_i) , \quad 0 \le \omega \le 1 . \tag{11}$$

Here $\delta(x, y)$ is the Kronecker symbol, equal to 1 if $x = y$, otherwise it is zero.

For kinetic analysis, correct recovery of events is more important than correct recovery of individual points. Therefore, it is more appropriate to minimize the sum of false and lost events, considering the errors of both types as equally harmful. Thus, the distance between the original and the recovered signals may be defined in terms of mismatching events. Each mismatching event increases this distance by one. A partial mismatching, e.g. in the case of transition shift, increases the distance by less than one. The counting of mismatching events proceeds as follows. We shall identify the real events by the index $k = 1, 2, \ldots, K$ and the recovered events by the index $m = 1, 2, \ldots, M$. The real and the recovered transitions are combined into a common set, which divides the data into J segments with dwells d_j, $j = 1, 2, \ldots, J$, signals S_j and signal estimates $\underline{S}_j$. Each segment j belongs simultaneously to a real event $k(j)$ (with dwell $D_{k(j)}$) and to a recovered event $m(j)$ (with dwell $\underline{D}_{m(j)}$). We want to consider all three types of error, i.e. false and lost events and transition shifts, in the same way. To this end we relate each segment to the shorter of the two dwells, $D_{k(j)}$ and $\underline{D}_{m(j)}$, so that each segment represents a proportion of an event, p_j:

$$p_j = d_j / \min(D_{k(j)}, \underline{D}_{m(j)}) . \quad (12)$$

Thus, the defined p_j is equal to one in the case of false or lost events and is less than one in the case of a transition shift. The total mismatching, E, is calculated as the sum of those p_j which correspond to mismatching between recovered and original signals:

$$E = \sum_{j=1}^{J} p_j [1 - \delta(S_j, \underline{S}_j)] . \quad (13)$$

We introduce further the mismatching potential, $\varepsilon = E/N$, and the matching potential γ:

$$\gamma = (K - E)/N = \mu_0 - \varepsilon . \quad (14)$$

The quantity μ_0 is defined as the potential of real events, $\mu_0 = K/N$. Because the quantities γ and ω are functions of T, we introduce the following notations for objectively optimal threshold values:

$$T_\gamma = \mathrm{argmax}(\gamma); \; T_\omega = \mathrm{argmax}(\omega); \quad (15)$$

The predicted optimal thresholds will be denoted as positions of minimum of the corresponding criterion:

$$T_\alpha = \mathrm{argmin}(\alpha_q); \; T_1 = \mathrm{argmin}(\alpha_1); \; T_\rho = \mathrm{argmin}(\rho) . \quad (16)$$

Denoting by x one of the letters γ, ω, α, ρ, we introduce the numbers of events (M_x) and the values of the event potential (μ_x) determined by the corresponding criteria as optimal ones:

$$M_x = M(T_x); \; \mu_x = M_x/N . \quad (17)$$

Different types of simulated data

We generated the data as jumps between two *known in advance* levels L_1 and L_2 (0 and 1), occurring at random time points, with white Gaussian noise superimposed (Box et al. 1994) of *known in advance* amplitude (SD) σ. Four different data types were used for a system with two kinetic states in combination with the Hinkley-detector, and one data type for a three-state system in combination with the Hinkley- and 50% methods. The simulation of the data for the three-state system will be described later.

We consider first the simplest case of a two-state Markov system with equal average lifetimes, i.e. expected dwells, D, in open and closed states, cf. Table 1. The time is measured in units of the sampling interval.

For the two-state system the record r is generated as a sequence of K events. The dwell times D_k, $k = 1, 2, \ldots, K$ are generated as the realizations of an exponentially distributed random variable with the expected value of D, so that N is a random number with the expected value of KD. The signal and its estimation may assume the values of either 0 or 1. The numbers n_i corresponding to noise values are independent realizations of a random variable distributed according to a Gaussian probability density with zero mean and variance σ^2. The values of σ, D and K are different for the four data types considered (see Table 1).

Table 1 Properties of different data types and numbers of repetitions, Runs

Type	σ	D	K	Runs
I	0.3	30	200	500
II	0.5	30	300	350
III	0.6	20	500	200
IV	1	20	1000	150

σ: noise amplitude; D: average dwell of open or closed event; K: number of events in a record, Runs: number of repetitions for averaging. Channel amplitude $\Delta = 1$, so that here $SNR = \Delta/\sigma = 1/\sigma$

These properties are organized in Table 1 in such a way that the difficulty of the data increases with the type number (more on that in the Discussion). One can see from Table 1 that the expected length of a record N increases with data type. This feature is selected to better illustrate the results for MDL, where the penalty per transition depends on the record length. The total number of data points for each data type, represented by the product of N with the number of repetitions, Runs, remains almost the same. In order to achieve the same statistical reliability of the results for all data types, the total number of data points should increase with the difficulty of the data. However, the lowest reliability is already sufficiently high, so that the difference in the reliabilities is of no importance.

Hinkley detector

We give here a more complete and symmetric description of the detector than is given in Schultze & Draber (1993). The boundary B between the levels L_1 and L_2, and the channel amplitude Δ are defined as:

$$B = (L_1 + L_2)/2 , \quad \Delta = L_2 - L_1 > 0 . \quad (18)$$

The data are normalized by the following linear transformation:

$$\xi(i) = (r_i - B)\,\Delta/\sigma^2 \,. \tag{19}$$

The detector is described by the following variables:

1) State variable H. For each data point, the variable H may assume the values of either –1 or 1, indicating the lower or the higher level respectively.
2) Position of the detector, i. This is the number of the current data point used for testing. The detector moves in steps of 1, i.e. the value of i increases only by 1.
3) Test and marking variables, t and h, calculated at each position of the detector.

At first, the detector searches for an initial position i_s, defined as the first point where $\mathrm{sign}(\xi) \neq 0$. Except of specially prepared data, with probability one this is the first data point, $i_s = 1$. The initial condition for h is zero, the initial condition for H is guessed and verified afterwards:

$$h(i_s) = 0; \quad H(i_s) = \mathrm{sign}(\xi(i_s)) \,. \tag{20}$$

The testing procedure begins by increasing i by one, so that at the beginning $i = i_s + 1$. The variables h and H are updated according to the equations:

$$H(i) = H(i-1) \,; \tag{21}$$

$$t = h(i-1) - H(i)\,\xi(i) \,; \tag{22}$$

$$h(i) = t\,\eta(t) \,. \tag{23}$$

Here $\eta(t)$ is the Heavyside or step function, equal to one for $t > 0$ and to zero for $t \leq 0$.

Subsequent operations depend on comparison of t with T. If $t < T$, no jump is recognized, and the testing begins from the next point (i is increased by one, Eqs. (21)–(23) are used again, etc.). However, if $t \geq T$, the jump is recognized. The position of the jump is just after the last point where $h = 0$. In all points above the jump, up to (and including) the detector position, the sign of H is changed, and h-values are replaced by zero. Then the testing is repeated in the next point.

The procedure described continues until i reaches the value of N. Consequent events are detected as sets of neighboring points with the same H-values. The initial guess for H is verified as follows: if $\left|\sum_{i=1}^{\underline{D}_1} \xi(i)\right| < T$, the first transition is deleted, i.e. the values of H(i), $1 \leq i \leq \underline{D}_1$, change their sign. The values of the recovered signal are calculated as

$$\underline{s}_i = B + H(i)\,\Delta/2 \,. \tag{24}$$

Analysis protocol

Each record was analyzed by the Hinkley detector with different T-values. The steps of T were selected to be smaller near the expected T_* (Fig. 1). For each T-value, an ideal trace with M(T) events is obtained. The corresponding dwell times $\underline{D}_m$, m = 1, 2, …, M and estimates of the signal $\underline{S}_m$ are stored. For each trace the values of the tested predictive criteria are calculated along with the values of the objective criteria. The whole procedure is repeated many times (Runs) and the results are averaged over repetitions for each value of T. The values of T are fixed for each data type.

For the three-state system, the analysis was made in a similar way; in addition, the 50% amplitude method of idealization was used.

Results

Comparison of AIC and MDL with objective criteria

Figure 1 shows the average values of $\mu/\mu_0 = M/K$ as a function of T for data I–IV, with μ_ω, μ_γ, and μ_ρ-values for each curve indicated. The optimum numbers of events, recovered per hundred real events along with their statistical errors are given in Table 2, and the corresponding average T_x-values are given in Table 3. The average values of the criteria ω, γ/μ_0 and ρ as a function of the average values of M/K are shown in Fig. 2. The criterion α_1 underestimates T_* so much that T_1 is outside the investigated intervals of T. Consequently, the corresponding curves are not shown in Fig. 2. By inspecting Figs. 1 and 2 and Tables 2 and 3, only a small difference between objective optimality criteria γ and ω is observed for data types III and IV.

In comparison with the α_1-criterion, the ρ-criterion gave much better results. However, they also deviate significantly from the objectively optimal value of $T_* = T_\gamma$, overestimating T_* and underestimating the corresponding value of M_γ. It follows from Table 3, that T_* taken as either T_ω or T_γ monotonically decreases with increasing

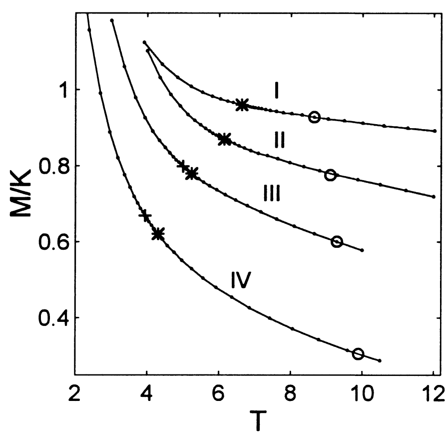

Fig. 1 The number of recovered events per real event as a function of T for different data types, described in Table 1. The values of μ_ω, μ_γ, and μ_ρ are shown by +, * and ○ respectively. Note that the value of T_ρ significantly deviates from the objectively optimal value of T_γ, showing even the opposite tendency with growing number of data type

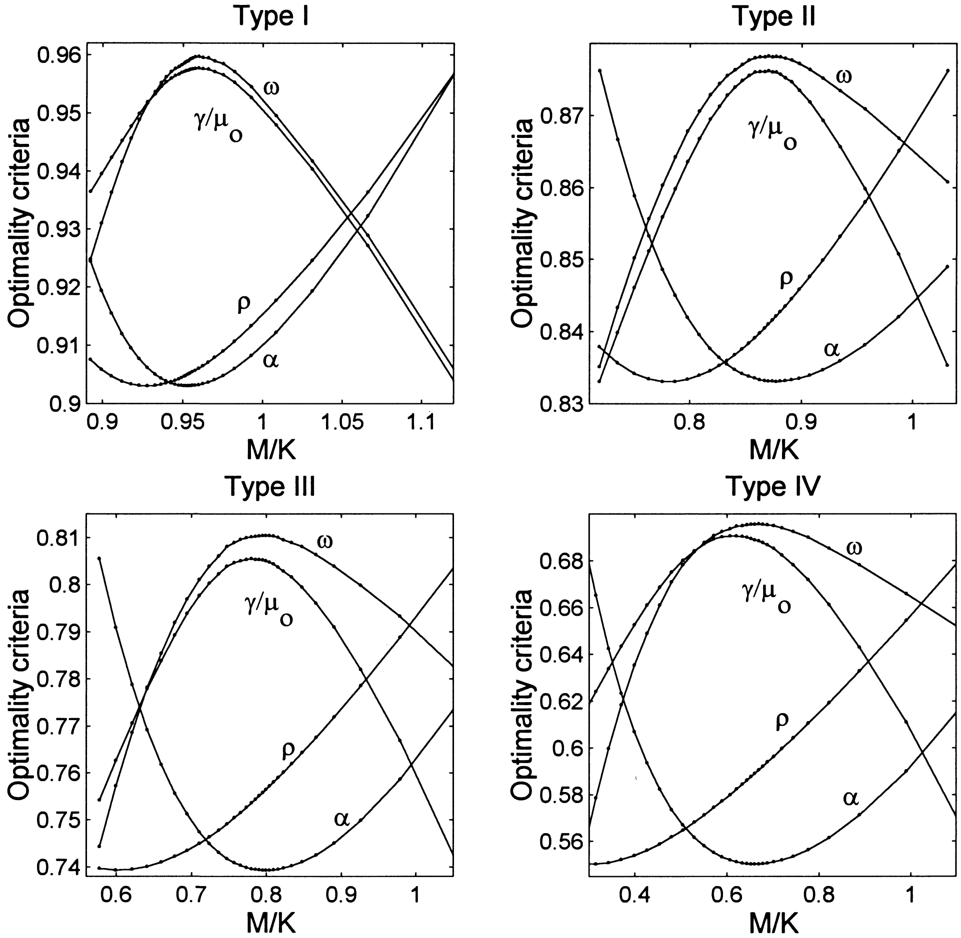

Fig. 2 Different criteria: ω, γ/μ_0, ρ and α_q (the latter for a specific choice of the q-value for each data type, cf. Tables 2 and 3) as a function of $\mu/\mu_0 = M/K$. The data types are indicated above each plot. The ordinate corresponds to the curves γ/μ_0, the other criteria are linearly transformed for better scaling. This does not affect the position of the optimal point

Table 2 Optimal numbers of recovered events per hundred real events, corresponding to different criteria. The values of q used with α_q are shown for each data type

Type	100 M_ω/K	100 M_γ/K	100 M_ρ/K	100 M_α/K	q
I	96.34 ± 0.11	96.31 ± 0.11	92.74 ± 0.12	95.26 ± 0.11	3.5
II	87.2 ± 0.2	86.89 ± 0.19	78.01 ± 0.17	87.76 ± 0.16	3
III	79.82 ± 0.33	78.81 ± 0.29	60.11 ± 0.2	79.94 ± 0.21	2.5
IV	66.33 ± 0.41	62.36 ± 0.32	31.21 ± 0.14	66.09 ± 0.2	2

Table 3 Optimal T-values for different criteria. The q-values used with α_q are shown

Type	T_ω	T_γ	T_ρ	T_α	q
I	6.71 ± 0.04	6.70 ± 0.04	8.63 ± 0.03	7.04 ± 0.01	3.5
II	6.21 ± 0.04	6.25 ± 0.03	9.08 ± 0.02	6.00 ± 0.006	3
III	5.08 ± 0.04	5.18 ± 0.03	9.26 ± 0.01	4.99 ± 0.003	2.5
IV	4.01 ± 0.03	4.30 ± 0.03	9.68 ± 0.02	3.99 ± 0.002	2

number of data type. At the same time, T_ρ monotonically increases (see also Fig. 1). This behavior correlates strongly with the size of the data, $N \cong KD$, indicating inconsistency of MDL in the context of the idealization problem.

As we have mentioned, the standard AIC, represented by α_1, strongly underestimates T_*, suggesting as optimum almost 3 times more events than real ones. Although the T_1-value is insensitive to data size, such a big deviation from reality can not be acceptable. Both criteria, α_1 and ρ, differ only in the value of the penalty per transition, q, one of them underestimating, the other overestimating the T_*-value. For this reason one may expect that with a proper choice of q-value the criterion α_q will give the exact T_*-value. It is of major importance, whether such a properly selected q-value is a universal one, i.e. independent of the data type. The universality of such a q-value would mean that the exact Akaike-type criterion can be used as a tool for predicting the T_*-value. In the hope that by an appropriate choice of the penalty per transition the situation might be improved, we have specially selected several q-values, one for each data type, in such a way that the T_α be not very different from T_γ. The average values of α_q are shown in Fig. 2 as curves α. The q-values selected, along with the corresponding M_α- and T_α-values are shown in Tables 2 and 3. Unfortunately, the q-values are significantly different for different data types. This means that a universal Akaike-type criterion, suitable for selecting the best ideal trace for different types of data, does not exist.

Exact Akaike-type criterion

One may see from Table 3 a striking proportionality between different q-values and corresponding T_α-values, expressed as:

$$T_\alpha = 2q\,. \tag{25}$$

In other words, in order to find the T_*-value with the help of "exact" Akaike-type criterion α_q, one has to select the penalty per transition $q = T_*/2$! Attempts to understand the discovered connection between the Hinkley method and the Akaike-type criterion have shown that it is not simply a coincidence. There is a simple proof that Eq. (25) is strictly correct in an asymptotic sence. It follows from Eq. (8), in order that Eq. (25) be true for any type of data, it is sufficient that the following equation holds:

$$\frac{d\alpha_0}{d\mu} = -T\,. \tag{26}$$

Here α_0 is the same as α_q when $q=0$. To prove Eq. (26), consider an ideal trace with the smallest value of h, equal to the threshold T, used for idealization. This means that the "weakest" event of this trace (say, with number m) should satisfy the equation

$$\left|\sum_{i=i_0}^{i_0+\underline{D}_m-1} \xi(i)\right| = T, \quad i_0 = \sum_{l=1}^{m-1} \underline{D}_l + 1\,. \tag{27}$$

To be specific, let us assume that the estimate of the signal in this event is L_1. If we slightly increase the T-value, the weakest event will disappear, and the estimation of the signal within this event will be changed to L_2. Correspondingly, the estimation of noise variance also changes, so that the change of $N\underline{\sigma}^2$ is:

$$\sum_{i=i_0}^{i_0+\underline{D}_m-1} (r_i - L_2)^2 - \sum_{i=i_0}^{i_0+\underline{D}_m-1} (r_i - L_1)^2$$
$$= -2\Delta \sum_{i=i_0}^{i_0+\underline{D}_m-1} (r_i - B) = 2T\sigma^2\,. \tag{28}$$

If the "weakest" event is the first or the last one, then the number of events, M, changes by –1, otherwise it changes by –2. Since we are interested in the asymptotic properties of the criterion, we should assume that the data are very long. Therefore, the chance that the "weakest" event is the first or the last one, is negligibly small. Thus, the number of events after a very small increase in T will be decreased by 2. Dividing the previous expression by $-2\sigma^2$ results in Eq. (26), which shows the correctness of this equation in an asymptotic sense. A more general proof, when several events may disappear as a result of an increase in T, is the same, because the increase in α_0 and the decrease in μ will both be proportional to the number of lost events. Thus, Eq. (26) demonstrates an internal connection between the Akaike-type criterion α_q and the Hinkley-threshold T. One may note also that T_ρ-values in Table 3 almost coincide with ln(N) for each data type, in agreement in Eq. (25). Thus, MDL produces correct results only for data of certain (optimal) length. For records longer than such an optimal length, the results of MDL become increasingly worse. This reflects the nonasymptotic character of MDL.

Summarizing our findings, we give two equivalent expressions for the "exact" Akaike-type criterion α_*, whose minimum indicates the point of the objective optimum for the Hinkley threshold, T_*. The two experessions are different according to the choice of the independent variable:

$$\alpha_*(\mu) = -\int_0^{\mu} T(\mu)\,d\mu + T_*\mu\,, \tag{29}$$

$$\alpha_*(T) = \int_0^{T} \mu(T)\,dT + (T_* - T)\,\mu(T)\,. \tag{30}$$

Here $T(\mu)$ is defined as the inversion of the function $\mu(T)$. This inversion is always uniquely possible owing to the monotonicitiy of $\mu(T)$, expressed as $d\mu/dT < 0$.

We have proved that the Akaike-type criterion produces the corret T_*-value only if this value is known in advance and is built into the expression for α_q. In practice, however, the T_*-value is what we want to find. We have developed a method for selecting the T_*-value based on statistical consistency of the regularized amplitude histograms (Pastushenko & Schindler 1995, Pastushenko and Schindler 1997). Here we present a simpler but closely related method, which is in a good agreement with the more detailed one and with objective criteria.

Predictive optimality criterion

We suggest an approximate expression for a predictive optimality criterion, closely related to the objective criterion γ and based on experimentally defined quantities. To this end we shall express the total error of the analysis in terms of mismatching events, E, as the sum of false and lost events, F and L respectively:

$$E = F + L\,. \tag{31}$$

This approach is based on a reasonable assumption that the contribution to the total error due to transition shifts caused by noise is only weakly dependent on T and may therefore by neglected. The quantity to be maximized within our approach is therefore

$$G = (K - E)/K\,. \tag{32}$$

This is the same as the minimization of E, because the quantity K is independent of filter parameters that are used for analysis. We shall consider quantity F as experimentally detectable by using the number of events Φ that are detected by the idealization of a record of the same length N, but consisting of noise only (i.e. containing no real events):

$$F = \Phi/2\,. \tag{33}$$

The number of lost events may be found from the following phenomenological equation:

$$M = K + 2F - 2L\,. \tag{34}$$

The meaning of this equation is transparent: the number of observed events is equal to the number of real events plus two times the number of false events (each false event increases the number of observed events by two) and minus two times the number of lost events. Substituting L from Eq. (34) into Eq. (32), and using Eq. (31, 33), we obtain G:

$$G = (K + M - 2\Phi)/2K\,. \tag{35}$$

Multiplying this equation by $2K/N$ and omitting the constant term, we can then write down the expression for the new predictive criterion Γ in terms of experimentally known potentials $\mu = M/N$ and $\varphi = \Phi/N$ as:

$$\Gamma = \mu - 2\varphi\,. \tag{36}$$

For illustrations with simulated data, where K is known, we shall use G instead of Γ, because the quantity G is analogous to the objective criterion $\gamma/\mu_0 = (K-E)/K$ where E is defined by Eq. (13). For this reason G and γ/μ_0 are directly comparable almost without additional rescaling.

Testing the new predictive criterion

(i) Two-state system

To show, how the Γ-criterion works, we have calculated the φ-potential by the Hinkley method from $6 \cdot 10^7$ normally distributed independent data points at $\sigma = 0.5$ and $\sigma = 1$ with $\Delta = 1$, i.e. for data types II and IV. Corresponding $\varphi(T)$ curves are shown in Fig. 3. Schultze and Draber 1993, suggested for white noise a universal dependency $\ln(\varphi) = -T$, Eq. (13) in this reference. As one may see, the curves in Fig. 3 are certainly nonlinear and different. This difference indicates that φ is dependent not only on T, but also on SNR. Although the nonlinearity is not very big, we have used polynomials of 5-th order for satisfactory interpolation. However, this nonlinearity can not account for big difference between the T_*-values found in this paper, and the "standard values" $T_* = 11$ (Schultze and Draber 1993) or $T_* = 16$, suggested as "the rule of thumb" in Draber and Schultze (1994 a). An obvious reason for such a big difference between T_*-values is connected with different data types. Another reason may lie in data preparation: we considered the data without additional filtering, whereas in Schultze and Draber (1993, 1994 b) the data were additionally low-pass filtered, which results in a shift in the σ-estimate, and therefore a shift in the T-estimate. In any case, "the rule of thumb" in the form of a standard T_*-value is not meaningful, because T_*-values are sensitive to data types.

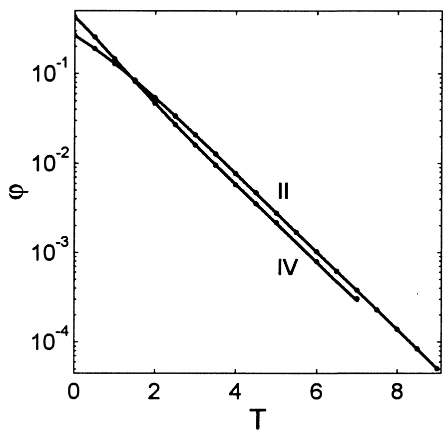

Fig. 3 The dependencies of the "false potential" φ of T for $\sigma = 0.5$ (*II*) and $\sigma = 1$ (*IV*) at $\Delta = 1$. The calculations were made using the "records" of pure Gaussian noise (no signal)

The dependencies of G and γ/μ_0 on M/K are shown in Fig. 4 for data types II (left) and IV (right). The curves γ/μ_0 are the same as in Figs. 2 II and 2 IV. The agreement between the positions of predicted and objective optima in both cases is perfect, although the data of type IV are very difficult. For instance, in order to reach the value of SNR = 10, recommended by the 50%-method (Colquhoun and Sigworth 1995), the filtering of data IV should be equivalent to averaging over a hundred points. After such filtering, in order to get independent data points, the record should be resampled with the sampling interval increased by a factor of a hundred. Thus, in terms of independent points, the average dwell time for type IV data becomes only 1/5 of the sampling interval. The same consideration for type II data results in the average dwell time of about 1.2 sampling interval, which is not very much better.

(ii) Three-state system

a) The method of 50% amplitude. We present now the results for the 50% amplitude method of analysis and for more typical data from ion channel practice. As an example we shall consider the 3-state channel model with two closed (states 1 and 2) and a third open state, as taken from Magleby and Weiss (1990 a). Channel amplitude is equal to one. The rate constants for transitions between states, measured in $\sec^{-1}$, are: $k_{12} = 100$, $k_{21} = 200$, $k_{23} = 1000$, $k_{32} = 500$ and $k_{13} = k_{31} = 0$. The amplitude (SD) of filtered noise was 0.14, i.e. SNR = 7.15. Using the known dead time (1.6 ms), SNR and sampling frequency (20 points per dead time), we have found that the efficiency of the digital unbiased exponential filter, which we used, should be 7.6 (efficiency = ratio of SD of white noise to SD of filtered noise). For Gaussian filters the efficiency would be 7.25. Thus we found the SD of nonfiltered noise, $\sigma = 7.6 \cdot 0.14 = 1.06$. We have reproduced such data (one example is shown in Fig. 5). The simulated record contained $1.5 \cdot 10^7$ data points and 500 000 transitions. Only 416 000 transitions were observable (transitions between two consecutive closed states are not observable even in the absence of noise). Thus, the average observable (in the absence of noise) dwell time was about 36 sampling intervals. This allows us to estimate the difficulty of these data as being between data types III and IV.

The simulated data were analyzed by the 50% method. The record was filtered at different efficiencies, the number of recovered events and objectively found error E were stored. Examples of differently filtered data are shown in

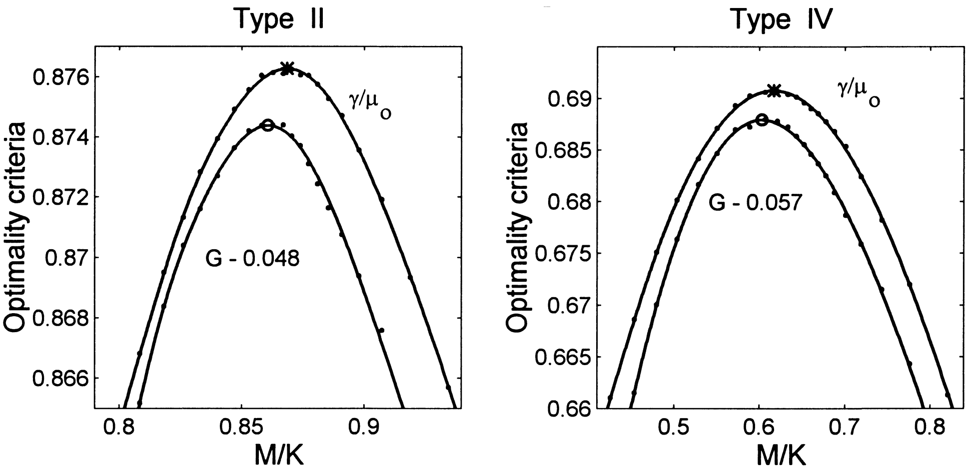

Fig. 4 Matching per real event, G and γ/μ_0. The ordinate corresponds to the values of γ/μ_0, the curves G are shifted downwards for better scaling. The interpolation was made with polynomials of the seventh order. The position of the optimum is practically independent of the order of the polynomial used. Data types: *left – II, right – IV.* The predicted (○) and the objective (∗) optimal μ-values practically coincide

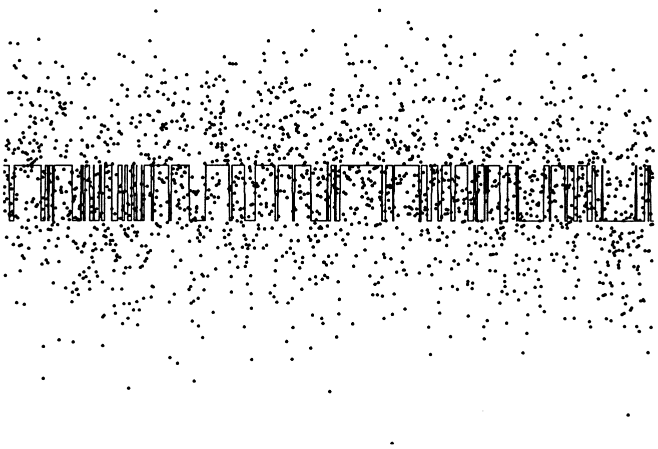

Fig. 5 Example of data for the system with three states. 2000 data points from nonfiltered record are shown (*points*) along with the real signal (*solid line*)

Fig. 6 B, C. Observable (in the absence of noise) transitions are shown in Fig. 6 A. Optimal filtering (Fig. 6 B, efficiency 5.7) allows one to find significantly more real events than the filtering with efficiency 7.6 (Fig. 6 C). To compensate for the transition shifts caused by the filtering, the transitions found were corrected by the Hinkley-test, applied to nonfiltered data.

To make the statistical uncertainty negligible, the false potential was calculated by the 50% method from the "record" of $1.75 \cdot 10^7$ normally distributed (with $\sigma = 1.06$) data points. Predicted (○) and objective (∗) optimal points are shown in Fig. 7 as maximum points of the criteria G

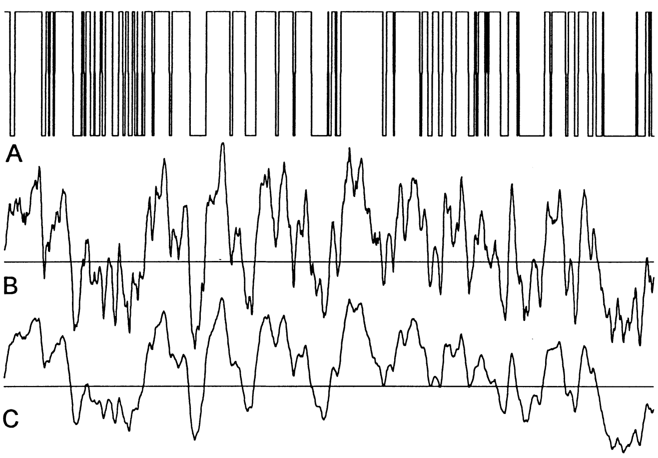

Fig. 6 Ideal signal **A**, and corresponding filtered data, **B, C** (the same data as in Fig. 5). **B**, optimal filtering: efficiency 5.7 (noise SD = 0.19), **C**, overfiltered: efficiency 7.6 (noise SD = 0.14). Optimal filtering allows one to recover more real events

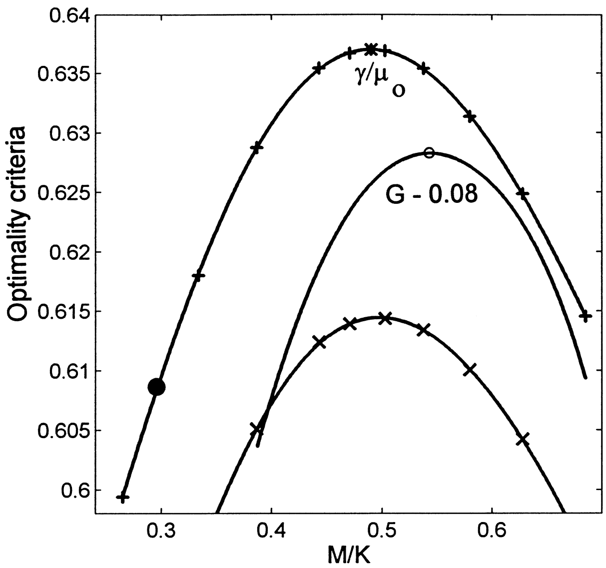

Fig. 7 Criteria G and γ/μ_0, calculated for three-state system using 50%-method. The ordinate corresponds to the values of γ/μ_0, the curves G are shifted downwards for better scaling. The calculated values of γ/μ_0 before and after transition correction are shown by marks × and + respectively, interpolated by cubic polynomials. *Filled circle* corresponds to the filtering efficiency of 7.6 ($M/K = 0.3$). Optimal filter efficiency is 5.7 ($M/K = 0.49$, shown by '∗'), the predicted efficiency is 5.4 ($M/K = 0.54$, shown by '○')

and γ/μ_0 respectively. The calculated values of G before and after eliminating transition shifts created by the low-pass filtering are indicated by marks × and + respectively. The calculated points were interpolated by cubic polynomials. As one may see, the transition correction increased the maximum value of γ/μ_0, but little influenced the position of the optimum.

The filled circle indicates the value of γ/μ_0 for the filtering with efficiency 7.6, which corresponds to the analysis in Magleby and Weiss (1990 a). For this filtering efficiency, M/K=0.3, whereas the objective optimum lies at M/K=0.49, and the predicted value is M/K=0.54. The average dwell times corresponding to these M/K-values are 121, 73 and 66 respectively. Almost 10% deviation of predicted value of M/K from objectively optimal value (found at efficiency 5.7) is partly the result of interference between real and false events, and partly is due to the method of analysis. However, this deviation is significantly smaller than the deviation for efficiency 7.6. These examples clearly indicate that the Γ-criterion works very well even in the case of data with a pronounced tendency to flickering. It works even better with the Hinkley-method, as described below.

b) Hinkley method. For comparison with the results of the 50% method, the relation between predicted and objective optima in the case of the Hinkley method of idealization is shown in Fig. 8 for the same data as in Fig. 7. The potential φ was calculated exactly. The predicted optimal value of M/K=0.475, and the objectively optimal value of M/K=0.495, so that the relative deviation is about 4%. Comparing Figs. 7 and 8, one may see that the objectively optimal value of γ is somewhat higher for the Hinkley-method than for the 50% method. This reflects the fact that the Hinkley method is a somewhat better idealizer than the 50%-method.

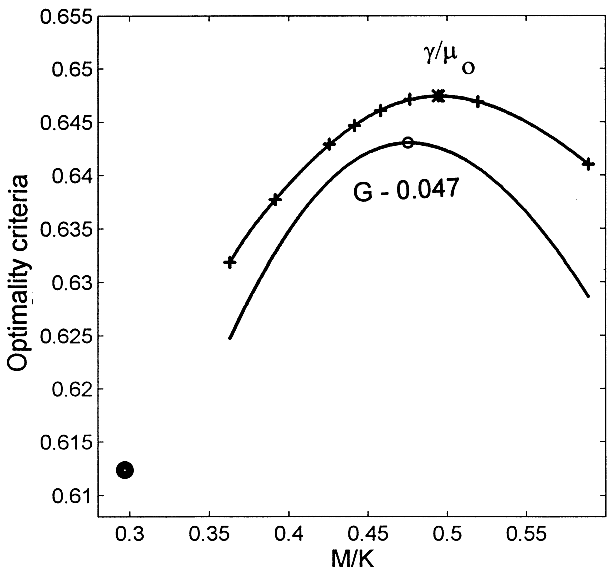

Fig. 8 Criteria G and γ/μ_0 for the same data as in Fig. 7, but analyzed with the Hinkley method. The objectively optimal value of $M_\gamma/K=0.495$ (shown by ‘*’), the predicted optimal value of $M/K=0.475$ (shown by ‘○’). The relative difference is about 4%

Discussion

The idealization of ion-channel data is generally considered as the first step dowards the recovery of the underlying mechanism of the channel activity. Sometimes, the first step strongly influences the subsequent results. In particular, the overfiltering of a record increases the number of missed events, rendering the recovery of the channel mechanism a more difficult problem. For this reason the optimal idealization is an especially important topic in the analysis of ion channel or similar data.

Following the suggestion of Colquhoun and Sigworth (1995) we have accepted the maximal similarity between the recovered and the real traces as the optimality principle. Two different procedures, expressing the degree of the similarity quantitatively, are suggested in this paper. Corresponding criteria in terms of matching points (ω) or events (γ) are considered as objective ones, because they allow one to decide objectively on the performance of different predictive optimality criteria. The difference between the two objective criteria is small, but we still prefer the criterion γ, because for kinetic analysis the correct recovery of events is more important.

The other results fall into three categories.

1) Information theoretic criteria, AIC and MDL, were tested as tools for predicting the optimal threshold value for the Hinkley-detector, the latter playing the role of a statistical filter for detecting transitions. Both criteria gave very unsatisfactory results.
2) We have first noticed numerically and then proved analytically a connection between the Akaike-type criterion and the Hinkley-method of idealization, expressed as $T_\alpha=2q$. This relation connects the penalty per transition q with the value of T_α, indicated by the Akaike-type criterion as an optimal threshold. Based on this finding, the expression for “exact” Akaike-type criterion α_* was derived.
3) A new predictive optimality criterion Γ is suggested, closely related to the objective criterion γ. In spite of neglecting the contribution of transition shifts, the criterion Γ works very well, as demonstrated with different types of data and idealization methods.

Although the results 1) seem to be negative, they surely contribute to the analysis of patch-clamp or related data, indicating the tools which *should not* be used and therefore saving the efforts of those who may occasionally come to the idea of using AIC or MDL. As we have shown, MDL is not an asymptotic criterion, and for this reason its results become increasingly worse with growing data size, if the data are sufficiently long. The “optimal” size of data for which MDL would work depends on the type of data, and this dependency is neglected by MDL.

Owing to the fact that our data were sufficiently long, MDL overestimated the T_*-value. The standard AIC, α_1 severely underestimated the T_*-value. As these criteria differ only in the value of the penalty per transition, we have undertaken an attempt to find a "correct" universal value for the penalty. The attempt was not successful, because such a "correct" penalty appeared to be sensitive to the data type. However, this has allowed us to find the expression for the "exact" Akaike-type criterion α_*. Corresponding Eqs. (29 – 30) confirm the existence of an Akaike-type criterion whose minimum is at the objectively optimal value of $T = T_*$. At the same time, the anatomy of these equations shows that α_* is not useful as a tool for seeking T_*, because it requires explicit knowledge of T_* in advance. In other words, the "exact" criterion α_* works according to principle WYPIWYG = "What You Put Is What You Get". This statement is engraved by the observation that the T_*-values vary significantly between different data types. Therefore, not only is the T_*-value unknown, it is also different for different data types. Consequently, a universal Akaike-type criterion as a model ranker does not exist. We do not see any significant difference between optimal idealization (i.e. searching for the optimal trace as the best statistical model) and searching for the most likely channel mechanism, which represents the ultimate aim of the analysis of patch-clamp data. The lack of any predictive power of the Akaike-type criterion in the idealization problem is an example which proves that in the general case the use of AIC as a model ranker is doubtful, and that the "best models" produced by AIC may be worse than those claimed by this criterion as unlikely ones.

The likelihood function for the analysis of ITC was written without using any specific information about the channel mechanism, except for two known levels and known noise amplitude. This reflects a typical situation where before the idealization no detailed information about the channel mechanism is available. As it should be, our likelihood function characterizes different statistical "models" of the hidden signal, corresponding to different ideal traces. These "models" are produced by the idealization method selected, which also takes into account the grouping of data points into events. For this reason, the likelihood function takes the grouping into account automatically. In fact one obtains the same log likelihood function, whether the summation in $\ln(f)$ is made point-wise, as Eq. (5) suggests, or event-wise, as in Moghaddamjoo (1991 b).

So far we have discussed ITC as likelihood maximizers within the idealization problem. Instead of these unsuitable criteria we have suggested a new predictive criterion Γ, which is based on the measurements of two potentials: the event potential μ and the "false" potential φ. The only simplifying assumption was in neglecting the contribution of transition shifts, caused by noise, to Γ. As the practical examples show, this assumption works reasonably well. No assumptions about noise properties or idealization method were made in the derivation of the new criterion Γ. Therefore, this criterion should be valid for any real data with two conductance levels and/or for any idealization method. We have considered the data with white Gaussian noise, which is uncorrelated. In the case of real data, the noise is usually correlated owing to filtering by registering devices or to additional filtering before digitization. To take this into account, the $\varphi(T)$ dependence should be found for site-specific noise, using the same statistical (or low-pass) filter, as for the idealization of the ion channel record.

We have used for demonstrations data of the Markov type. This was done because Markov models are most frequently used for interpretation of experimental data. Another reason is that it was the easiest way for us, because we have a program for simulating data according to an arbitrary kinetic scheme, described by rate constants. At the same time we have to stress that the method presented is in no way limited to Markov type data. This method is valid for any type of data, either Markov or non-Markow, which shows its additional advantage in comparison with methods based on specific assumptions about the mechanism of channel functioning.

Finally we should like to compare the standard AIC, $\alpha_1/2 = \mu - \ln(f)/N$, with the new criterion $\Gamma = \mu - 2\varphi = \mu - 2\Phi/N$. Formally, these expressions have some similarity, but we use the doubled number of transitions in pure noise 2Φ instead of the log likelihood $\ln(f)$. The principal difference between Γ and α_1 is: whereas α_1 is minimized, the criterion Γ is maximized. If we simply change the sign of Γ in the hope of reaching the version of minus log likelihood function, then the penalty per transition becomes negative. Thus, an essential difference with AIC still remains. Physically this difference is expressed by the fact that the derivative of μ over T defines the derivatives of Γ and α_1 on different sides of their optimal points.

We have demonstrated the predictive power of the new optimality criterion for different data types, using two different idealization methods. Equally good result should be expected for other idealization methods, based on different statistical tests.

Acknowledgements One of us (V. P.) expresses is gratitude to the City of Linz, High School Funds of Linz, and the government of Upper Austria for supporting scientific contacts, which enabled the initiation of this study, and to Dr Cleve Moler (Mathworks, Natick, USA) for the generous gift of the random number generators of the new generation. This work was partly supported by the Austrian Funds for Scientific Research, project S 066-10.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Control AC-19:716 – 723

Albertsen A, Hansen U-P (1994) Estimation of kinetic rate constants from multi-channel recordings by a direct fit of the time series. Biophys J 67: 1393 – 1403

Bauer RJ, Bowman BF, Kenyon JL (1987) Theory of the kinetic analysis of patch-clamp data. Biophys J 52: 961 – 978

Becker JD, Honerkamp J, Hirsch J, Fröbe U, Schlatter E, Greger R (1994) Analysing ion channels with hidden Markov models. Pflügers Arch 426: 328 – 332

Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis. Forecasting and control. Prentice-Hall, Inc Englewood Cliffs, N Jersey

Chung SH, Kennedy RA (1991) Forward-backward non-linear filtering technique for extracting small biological signals from noise. J Neurosci Methods 40: 71 – 86
Chung SH, Moore JB, Xia L, Premkumar LS, Gage PW (1990) Characterization of single channel currents using digital processing techniques based on hidden Markov Models. Phil Trans R Soc Lond B 329: 265 – 286
Colquhoun D, Hawkes AG (1995 a) The principles of the stochastic interpretation of ion channel mechanisms. In: Sakmann B, Neher E (eds) Single channel recording, 2nd edn. Plenum, New York London, pp 397 – 482
Colquhoun D, Hawkes AG (1995 b) A Q-Matrix Cookbook. In: Sakmann B, Neher E (eds) Single channel recording, 2nd edn. Plenum, New York London, pp 589 – 633
Colquhoun D, Sigworth FJ (1995) Fitting and statistical analysis of single channel records. In: Sakmann B, Neher E (eds) Single channel recording, 2nd edn. Plenum, New York London, pp 483 – 587
Draber S, Schultze R (1994 a) Correction for missed events based on a realistic model of a detector. Biophys J 66: 191 – 201
Draber S, Schultze R (1994 b) Detection of jumps in single-channel data containing subconductance levels. Biophys J 67: 1404 – 1413
Hafner R (1989) Wahrscheinlichkeitsrechnung und Statistik. Springer, Wien New York, pp 512
Fredkin DR, Rice JA (1992 a) Maximum likelihood estimation and identification directly from single-channel recordings. Proc R Soc Lond B 249: 125 – 132
Fredkin DR, Rice JA (1992 b) Bayesian restoration of single-channel patch clamp recordings. Biometrics 48: 427 – 448
Heinemann SH (1995) Guide to data acquisition and analysis. In: Sakmann B, Neher E (eds) Single channel recording, 2nd edn. Plenum, New York London, pp 53 – 91
Horn R (1987) Statistical methods for model discrimination: application to gating kinetics and permeation of acetylcholine receptor channel Biophys J 51: 255 – 263
Horn R, Lange K (1983) Estimating kinetic constants from single channel data. Biophys J 43: 207 – 223
Kirlin RL, Moghaddamjoo A (1986) A robust-window detector and estimator for step-signals in contaminated gaussian noise. IEEE Trans Acoustic, speech, signal processing. ASSP 34: 816 – 823
Korn SJ, Horn R (1988) Statistical discrimination of fractal and Markov models of single-channel gating. Biophys J 54: 871 – 877
Magleby KL (1992) Preventing artifacts and reducing errors in single-channel analysis. In: Rudy B, Iverson LE (eds) Methods in enzymology, vol 207. Ion channels. Academic Press, pp 763 – 791
Magleby KL, Song L (1992) Dependency plots suggest the kinetic structure of ion channels. Proc R Soc Lond B 249: 133 – 142
Magleby KL, Weiss DS (1990 a) Estimating kinetic parameters for single channels with simulation. A general method that resolves the missed event problem and accounts for noise. Biophys J 58: 1411 – 1426
Magleby KL, Weiss DS (1990 b) Identifying kinetic gating mechanisms by using two-dimensional distributions of simulated dwell times. Proc R Soc Lond B 241: 220 – 228
McManus OB, Weiss DS, Spivak CE, Blatz Al, Magleby KL (1988) Fractal models are inadequate for the kinetics of four different ion channels. Biophys J 54: 859 – 870
Moghaddamjoo A (1988) Step-like signal processing with distinct finite number of levels. IEEE Trans Indust Electron 35: 489 – 493
Moghaddamjoo A (1989) Constraint optimum well-log signal segmentation. IEEE Trans Geosci Remote Sensing 27: 633 – 641
Moghaddamjoo A (1991 a) Automatic segmentation and classification of ionic-channel signals. IEEE Trans Biomed Eng 38: 149 – 155
Moghaddamjoo A (1991 b) Estimation of the number of segments in a well-log signal via the information theoretic approach. IEEE Trans Geosci Rem Sensing 29: 177 – 178
Pastushenko VPh, Schindler H (1993) Statistical filtering of the single ion channel records. Acta Pharm 43: 7 – 13
Pastushenko VPh, Schindler H (1994) Idealization of ion channel records by Student T-test filtering with subsequent optimization of level distribution likelihood. Biophys J 66: A440
Pastushenko VPh, Schindler H (1995) Epoch oriented likelihood maximization of an idealized ion channel record. Biophys J 68: A148
Pastushenko VPh, Schindler H (1997) Level detection in ion channel records via idealization by statistical filtering and likelihood optimization. Phil Trans R Soc Lond B 352: 39 – 51
Pastushenko VPh, Schindler H, Waldl H, Hafner R (1996) A new statistical test for analysis of patch-clamp data. Biophys J 70: A200
Petracchi D, Barbi M, Pellegrini M, Pellegrino M, Simoni A (1991) Use of conditional distributions in the analysis of ion channel recordings. Eur Biophys J 20: 31 – 39
Qin F, Auerbach A, Sachs F (1996) Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. Biophys J 70: 264 – 280
Rabiner LR (1989) A tutorial on hidden Markov Models and selected applications in speech recognition. Proc IEEE 77: 257 – 285
Rissanen J (1983) A universal prior for the integers and estimation by minimum description length. Ann Statist 11: 417 – 431
Schindler H, Pastushenko VPh (1992) Statistical filtering of ion channel data. Biophys J 61: A248
Schultze R, Draber S (1993) A nonlinear filter algorithm for the detection of jumps in patch-clamp data. J Membrane Biol 132: 41 – 52