PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Level detection in ion channel records via idealization by statistical filtering and likelihood optimization

Vassili Ph. Pastushenko and Hansgeorg Schindler

Phil. Trans. R. Soc. Lond. B 1997 **352**, 39-51
doi: 10.1098/rstb.1997.0004

Level detection in ion channel records via idealization by statistical filtering and likelihood optimization

VASSILI PH. PASTUSHENKO AND HANSGEORG SCHINDLER

Institute for Biophysics, Johannes Kepler University of Linz, Altenbergerstr. 69, A-4040 Linz-Auhof, Austria

SUMMARY

A parameter-free method is presented for the level detection in ion channel records via recovery of step wise current changes. No assumptions about ion channel mechanism are made. The primary detection of the transitions is made by statistical filtering the data using the Student's *t*-test. The event currents are calculated as the average value of the current between two adjacent transitions. An optimal ideal trace is found by maximization of a likelihood function. The distribution of event currents recovered from the raw data is then analysed, again by using the Student's *t*-test, for their grouping into separate statistical ensembles, defining current levels. The method is subjected to rigorous test using simulated data, and is compared with several other methods. It produces the levels of channel current, their noise amplitudes and distributions of dwell times, the desired information for constructing the channel mechanism.

1. INTRODUCTION

Statistical analysis of ion channel records has been described in Colquhoun & Hawkes (1983) and Colquhoun & Sigworth (1983) (both papers are now updated in the second edition of the book), Jackson (1992), Magleby (1992) and French & Wonderlin (1992). Several new approaches have been published recently (Kirlin & Moghaddamjoo 1986; Moghaddamjoo 1988, 1989, 1991; Patlak 1988; Schultze & Draber 1993; Qin *et al*. 1996; VanDongen 1996). For a discussion of acquisition and analysis of ion-channel data see French & Wonderlin (1992) and Heinemann (1995). The final aim of such an analysis is a proposal on the mechanisms underlying the channel activity in the form of kinetic schemes. Some methods assume a certain hidden Markov model (HMM) for the mechanism (Rabiner 1989). This enables the parameters of HMM to be directly fitted to experimental data (Dempster *et al*. 1977; Horn 1987; Korn & Horn 1988; Chung *et al*. 1990; Chung & Kennedy 1991; Heinemann & Sigworth 1991; Terrien 1992; Fredkin & Rice 1992*a*; Becker *et al*. 1993). However, even if the Markov assumption is not unreasonable (for testing the Markov property, see Petracchi *et al*. 1991), it is rather difficult to select or identify the most appropriate type of HMM. In particular, even the number of levels may be unclear. The popularity of HMM as the main model for ion channel may have two reasons: (i) the properties of Markov systems are well studied and (ii) accepting a certain model represented by several parameters simplifies the model identification (Hafner 1989) which is then reduced to a parameter estimation (cf. Horn & Lange 1983; Magleby & Weiss 1990*a*, *b*).

The main information for constructing the channel mechanism is contained in the event currents and distributions of dwell times. Therefore, the first step in the analysis is the detection of transitions in the experimental record, sometimes referred to as 'detecting clusters' (Kirlin & Moghaddamjoo 1986), 'segmentation' (Moghaddamjoo 1989), 'restoration' problem (Fredkin & Rice 1992*b*), 'estimating the state sequence' (Terrien 1992), or 'record idealization' (Magleby 1992). The purpose of the idealization is to produce a possibly complete list of *dwells* for consecutive *events* including all data points. Recently, an approach for model identification directly from the idealized trace was suggested (Qin *et al*. 1996).

One standard method for record idealization, the 50% amplitude threshold method (Colquhoun & Sigworth 1983), frequently requires deep low-pass filtering. This method uses preliminary knowledge of levels, which allows avoidance of some natural errors in the case of unknown levels. At the same time, if the initial information about levels is incorrect, this method gives less chances of finding real levels. In addition, the statistical filtering described here avoids transition shifts caused by low-pass filtering.

The idealization is based on testing the hypothesis of 'a real transition within a testing window' with asymmetric Student's *t*-test (Heinhold & Gaede 1968; Oppenheim & Schaefer 1975) and on maximization of a likelihood function. It has no separate limitations on SNR and distance between levels, as in VanDongen (1996). For preliminary reports see Pastushenko & Schindler (1992, 1993, 1994).

2. STATEMENT OF THE PROBLEM

A record of an ion-channel current, r, in digitized form, represents a time series, $r = \{r_1, r_2, \ldots, r_N\}$, where N is the total number of data points. The time is measured in units of a sampling interval t_s. Thus, each dwell is equal to the amount of data points in a corresponding event. As usual (Jackson 1992), we make the following assumptions.

1. The record is a superposition of mutually independent background noise (noise) and step wise changing current (signal):

$$r = \text{signal} + \text{noise}. \tag{1}$$

Signal and noise are two sets of numbers of the same length N.

2. The signal assumes one of K (unknown) values, corresponding to different current levels (Box *et al.* 1994) of a single channel or several ion channels; K is unknown. For examples of sub-onductance states see Fox (1987).

3. Transitions between levels occur at random instants; the distribution of dwells is unknown.

4. The noise is a realization of a white noise process (Box *et al.* 1994), i.e. it represents N independent samples of a normally distributed random variable with zero mean and variance σ^2 which is not known.

To consider the noise in adjacent points as independent, the sampling interval t_s should be sufficiently large. A practically acceptable rule is $t_s = 1/(2\pi f_c)$, where f_c is the cut-off or -3 dB frequency (Sigworth 1983). The factor 2π corresponds to an analog RC filter, in which case t_s coincides with the characteristic time of the correlation function of filtered white noise. Close recipes for the sampling rate were given in Sachs (1983), Sigworth (1983) and Colquhoun (1987).

3. CONSTRUCTION OF A STUDENT'S T FILTER

We first introduce the Student's t-test in the context of finding transitions within some window by testing a null hypothesis. It is then used for constructing a two-step algorithm for finding transitions in the record (T filter).

(*a*) *Student's* t*-test as a transition finder*

Consider an arbitrary window of w data points from the record ($w \geqslant 3$), starting from $i_0 \geqslant 1$ and numbered as r_i, where $i = i_0, i_0+1, \ldots, i_0+w-1$. For convenience, we renumber these points as r_j, where $j = 1, 2, \ldots, w$. It is assumed that not more than one transition may occur within the window. Let the transition be expected immediately after the point $j = l$, with $w > l \geqslant 1$. The test is made using the average values of the current to the left and to the right of the suspected transition, A_l and A_r respectively

$$A_l = \frac{1}{l}\sum_{j=1}^{l} r_j; \quad A_r = \frac{1}{w-l}\sum_{j=l+1}^{w} r_j. \tag{2}$$

Random quantities A_l, A_r and $A_r - A_l$ have expected variances V_l, V_r and $V_l + V_r$ respectively:

$$V_l = \frac{\sigma^2}{l}; \quad V_r = \frac{\sigma^2}{w-l}; \quad V_l + V_r = \frac{\sigma^2}{l(1-l/w)}. \tag{3}$$

If the null hypothesis , H_0 ('the transition is false, i.e. generated by noise') is true, the quantity G, defined as

$$G = (A_r - A_l)\sqrt{(l(1-l/w)}/\sigma, \tag{4}$$

will be normally distributed with zero average and variance 1. This test function is known as the 'Gaussian test' (Heinhold & Gaede 1968). Usually the value of σ^2 is not known in advance. The Student's t-test overcomes this difficulty, replacing σ^2 by a local estimate of the noise variance, σ^2_{lr}, constructed from the σ^2 estimates on both sides of the expected transition, σ^2_l and σ^2_r

$$\sigma^2_{lr} = \{(l-1)\,\sigma^2_l + (w-l-1)\,\sigma^2_r\}/n, \tag{5}$$

$$n = w-2. \tag{6}$$

Here n is the degree of freedom for σ^2_{lr} estimation. The estimates for the noise variances are

$$\sigma^2_l = \frac{1}{l-1+\epsilon}\sum_{j=1}^{l}(r_j - A_l)^2;$$

$$\sigma^2_r = \frac{1}{w-l-1+\epsilon}\sum_{j=l+1}^{w}(r_j - A_r)^2. \tag{7}$$

Here $\epsilon > 0$ is a negligibly small number, $\epsilon \lll 1$, really meaningful in case $l = 1$ or $l = w-1$. If H_0 is true, the quantity $n\sigma^2_{lr}/\sigma^2$ obeys the χ^2 distribution with n degrees of freedom, and is independent of G. Therefore, for true H_0 the quantity T defined as

$$T = G\sigma/\sigma_{lr}, \tag{8}$$

obeys the Student's t-distribution with $n \geqslant 1$ degrees of freedom. To find the most likely transition, T values are calculated for all possible values of l. At a particular l value, say l_0, the value of T^2 will be maximum. Therefore, l_0 will be the most likely position of a possible transition (the transition to be tested is between l_0 and l_0+1). One can show that the Student's t-test (like Gaussian test) minimizes the noise σ^2_{lr} for $l = l_0$. In other words, the candidates for transitions suggested by these tests are found in the sense of maximum likelihood. To reject H_0, one has to compare $|T(l_0)|$ with a threshold value of the test, $X_n(P)$, which depends on a certain transition reliability P, satisfying $0 \leqslant P < 1$, and on $n = w-2$. The criterion for accepting the most likely transition as a real one is

$$T^2(l_0) > X^2_n(P). \tag{9}$$

The quantity P corresponds to the proportion of correct decisions only in the absence of the signal, i.e. only in the case of pure noise, and only if the position of the tested transition is selected at random. Nevertheless, in the presence of a signal, the higher the P value is, the less likely the super-threshold transition is false one. By this reason we may consider P as a threshold transition reliability. The existence of an optimal P value is obvious: for too small P values one obtains too many false transitions, and if P approaches sufficiently close to one, one loses more real transitions than false ones. We discuss the selection of the P value later; here it is assumed to be known. The Student's distribution function $F(T|n)$ is defined as (Abramowitz & Stegun 1965)

$$F(T|n) = \int_{-|T|}^{|T|} g(t|n)\,\mathrm{d}t, \tag{10}$$

where $g(t|n)$ is the probability density of the Student's t-distribution. To find $X_n(P)$, one has to solve for X_n the equation

$$F(X_n|n) = P. \tag{11}$$

We introduce now the main format of the final and intermediate results of the record analysis, corresponding to $M \geqslant 1$ events. The term 'event' means the data points delimited from either side by a transition. The properties of the mth event ($m = 1, 2, \ldots, M$) will be characterized by its dwell D_m, mean current A_m, and by a local estimate of event variance, V_m

$$V_m = \sigma_m^2 / D_m. \tag{12}$$

The local estimates of noise variance, σ_m^2, are obtained as in equation (7), but the summation is made over all the points in mth event. The global estimates are considered in practice as estimates of expected values. By this reason we shall denote the global estimates of the noise variance in the same way as its expectation, σ^2. As long as the levels are unknown, we use a straightforward generalization of equations (5) and (6)

$$\sigma^2 = \frac{\sum_{m=1}^{M} (D_m - 1)\sigma_m^2}{N - M}; \quad N = \sum_{m=1}^{M} D_m. \tag{13}$$

Later on, having estimates for levels, we may improve the σ^2 estimate (see the discussion). The quantities D_m, A_m and V_m are organized into an *event matrix* with M rows. Different event matrices will be denoted by letter Z with an additional letter, reflecting the matrix specificity. For instance, ZP will be a *P*rogrammed event matrix (i.e. the matrix created by a simulation program), ZE: *E*xpected event matrix (it differs from ZP, for instance, due to transition shifts created by noise; exact definition is given later), ZO: *O*ptimum event matrix, etc. In the same format we shall introduce a *level matrix* ZM (Z *M*acroscopic) with K rows, corresponding to the number of current levels identified ($K \ll M$). Due to the analogy with usual filters, which retain only sufficiently large entities and neglect small ones, we call the algorithm for finding super-threshold transitions a Transition filter (or T filter).

(*b*) *T filter as asymmetric* t*-testing with dynamic self-adjusting window*

In the first step we find candidates for transitions by t-testing. This produces a ZT event matrix. In the second step, the t-testing is improved by increasing testing windows, when possible. The final product of the T filter is an event matrix ZC ('Z *C*orrected'). These steps will be illustrated by the test record r with $K = 3$ (figure 1*a*, *b*). For longer records, the analysis gives more reliable results. A relatively short selected record allows us to show the performance of the method for short records and to give complete illustrations.

(i) *Step 1(ZT): Finding transition candidates*

According to the construction of the Student's t-test, not more than one transition is expected within the testing window. Since we have no *a priori* information about possible transition points, the transition may occur at any point. Therefore, the testing begins in the first three data points (i.e. the initial testing window $w = 3$). If no transition is found, w is increased by 1, i.e. one more point is added to the set of tested points, and the test is repeated. In this way the testing window may increase up to a certain maximum $w = W$. The specification of $W \ll N$ while finding the candidates for transitions serves mainly to save time in those cases when we are not interested in maximum amplitude resolution. If the transition is not found in the testing window $w = W$, then the window is shifted one point to the right. At each position of the testing window, all possible transitions are considered. If at some position and width of the testing window a transition is found, then the procedure described is repeated, starting from the first three points to the right of the found transition. This continues until the last point of the record is tested. The outcome of the preliminary T filtering will be a matrix ZT with M(ZT) events. Figure 1*c* shows the ideal trace for ZT with 273 events obtained at $X = 2$ and $W = 20$, where X is defined by P via

$$\mathrm{erf}(X/\sqrt{2}) = P. \tag{14}$$

The quantity X represents the threshold value of the Gaussian test, or Gaussian P percentile. This equation shows that the finding of an optimum P value is equivalent to the finding of an optimum X value.

(ii) *Step 2(ZC): Transition correction/ghost deletion*

Now we improve the estimates of positions and reliabilities of the found candidates for transitions, using maximum possible number of data points from two events on both sides of the tested transition. For instance, if we consider the transition between events m and $m+1$, then $w = \min(D_m + D_{m+1}, W)$. If $D_m + D_{m+1} > W$, the centre of the window is positioned as close to the transition tested as possible.

Due to the increased (in average) testing windows, the T values of real transitions are expected to increase; simultaneously, the transition coordinates become more accurate. For 'ghosts', i.e. false transitions, the value of T has a good chance of decreasing. Those transitions for which the T values drop below the threshold are then deleted. In this way, the assignment of the transitions is improved both for real and false transitions. Transition correction/ghost deletion is finished if no ghosts and no corrections are found. Sometimes, after all ghosts are deleted, the position(s) of one (or several) transition(s) undergo cyclic changes. For $W = N$ we did not observe such oscillations. To avoid cycling, the number of consecutive ghost-free loops is limited to two. After this procedure, the T values are estimated using all data points in adjacent events. Therefore, some $|T|$ values may become smaller than X_n. Such transitions are also deleted. The resulting ZC matrix has all super-threshold transitions.

The ZC matrix obtained from our test record at $X = 2$ and $W = 20$ has $M(ZC) = 168$ events (figure 1*d*) – fewer than ZT, but still much more than the programmed 50 events. One way of improving the results is increasing the X value. Figure 3*a* shows M(ZC) in dependence on X in the range $2 \leqslant X \leqslant 4$. By comparison with the original matrix ZP one can see that the closest to ZP is ZC matrix with 48 events, found at $X = 3.3$ or 3.4. When dealing with real ion-channel data, one needs a certain criterion in order to

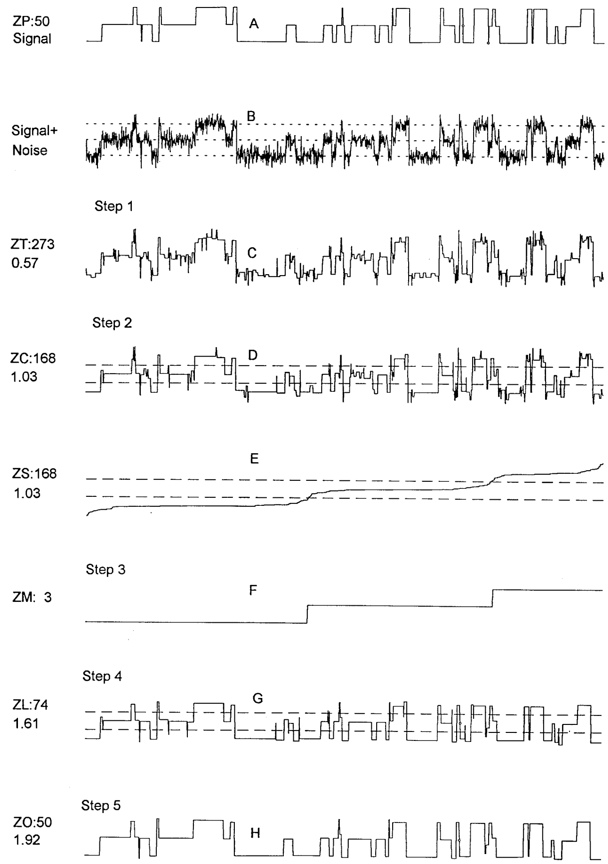


Figure 1. The signal is of HMM type with 3 levels at 0, 1, 2 (*a*), 50 real events, exponentially distributed dwells with average dwell 30, equal transition probabilities. Two one-point events are shown by circles. Addition of the white noise with $s = 0.3$ gives the simulated record (*b*). Different idealized traces obtained at $X = 2$ and $W = 20$: (*c*) ZT, (*d*) ZC, (*e*) ZS, (*f*) ZM, (*g*) ZL and (*h*) ZO = ZE. Horizontal dashed lines in (*d*), (*e*) and (*g*) show the boundaries. Numbers of events and sharpnesses are shown near each trace.

select an optimum event matrix. This criterion will be described in the next section. Another question is, whether the best ZC matrix is the absolutely best result which may be found from this record. Using the same criterion, we have developed a procedure for obtaining an optimum event matrix (ZO), which may be found for given X value (as a rule, the higher X value, the less events in ZO will be). This has allowed us to recover the expected matrix ZE. For our test record three expected transitions are different from original ones.

4. OPTIMAL IDEALIZATION

We consider the optimization of an event matrix as a method for increasing its statistical consistency. In section (*a*) we introduce the optimization principle and describe the detection of levels from events. In section (*b*) we present the global optimization procedures. Section (*c*) shows the influence of deviations from white noise, simulated by additional low-pass filtering. For our test record, a 100% recovery of expected events is shown both from raw and additionally filtered data. In section (*d*) a choice of the parameters X and W is discussed.

(*a*) *Detection of levels from event currents*

The number of levels, K, and their values are commonly estimated from pronounced peaks in point amplitude histograms. Low-pass filtering of the record is usually used to sharpen the peaks. However, very deep filtering shifts the positions of the peaks, influencing the estimates of the levels. The idealization without additional low-pass filtering allows better estimates of the level currents. One alternative to the point amplitude histogram is a histogram of event currents. Such a histogram is irregular for relatively short data (of the order of a hundred of events). A regularized histogram (p.d.f.) has been presented (Colquhoun & Sigworth 1983, formula 26) as a superposition of p.d.f. of event currents. In this approach, however, short and long events contribute equally. This differs essentially from point histograms. Another alternative is the superposition of the Gaussian p.d.fs, corresponding to event currents and weighted with coefficients proportional to dwells. This is a straightforward generalization of point amplitude histograms. It has a meaning of probability density to find a certain signal estimate in a point selected at random. The sensitivity of this function to the set of recovered transitions, the latter defining a statistical model of the signal, is of main interest for us. For this reason we shall call it a likelihood density function (l.d.f.), which is a standard term in statistical literature for probability densities considered as functions of model parameters. For the very same reason, in this paper, it would be more logical to use the term 'likelihood' with respect to p.d.f. of event currents. However, we shall retain the term 'p.d.f.' used by Colquhoun & Sigworth (1983), because both versions will be compared (cf. figure 7*c*) and we need different names. We have to note, that the term 'p.d.f.' is also correct, if the record is sufficiently long and if the transitions are not considered as variable model parameters.

(i) *Likelihood density function (l.d.f.)*

Let us denote by $\boldsymbol{V}_m$ a global estimate of the mth event current variance. Assuming that all the data points in mth event have the same value of the signal (i.e. neglecting the possibility of missed events), we may write

$$\boldsymbol{V}_m = \sigma^2 / D_m. \qquad (15)$$

A possible scattering of A_m is characterized by a p.d.f. $\psi_m(A)$ given by a normal distribution density function $\boldsymbol{N}(\alpha, \beta) = \exp(-\alpha^2/2\beta)/(2\pi\beta)^{1/2}$

$$\psi_m(A) = \boldsymbol{N}(A - A_m, \boldsymbol{V}_m). \qquad (16)$$

Denoting the l.d.f. by $\psi(A)$, we define it as a superposition of $\psi_m(A)$, with statistical weights proportional to dwells

$$\psi(A) = \sum_{m=1}^{M} p_m \psi_m(A); \quad p_m = D_m / N. \qquad (17)$$

The function $\psi(A)$ is an analogue of the point amplitude histogram made for an optimally filtered record (the optimality is understood in the sense that the averaging is made only within each event with maximum possible averaging window, equal to event dwell). Considered as such an 'optimal point histogram', the function $\psi(A)$ displays a remarkably higher sharpness and separation of peaks in comparison with usual point amplitude histograms even after low-pass filtering. It is also sharper in comparison with the p.d.f. of event currents defined in Colquhoun & Sigworth (1983) analogously to equation (17), with $p_m = 1/M$ (for comparison of p.d.f. and l.d.f. see figure 7*c*). Examples of ψ-functions for different event matrices are shown in figure 3*b*.

Usually the term 'likelihood' is used for a probability density of a given set of measurements (represented in our case by the record r), considered as a function of the model parameters to be estimated (in our case transition coordinates). We do not use any complete model which would allow us to write down such a function. In other words, our model is only partly defined by assumptions of Gaussian noise and the mere existence of levels. Therefore, the function $\psi(A)$, being dependent on transition coordinates, should be considered as a likelihood function for an incomplete model. No methods have been suggested previously for maximization of likelihood functions of this kind. The next paragraph contains an approach to the solution of this problem. Some simpler approaches, e.g. maximizing the sum of peak values, or the sums of ψ-values calculated at A values corresponding to the levels, gave much weaker results (at least for relatively short records).

(ii) *Optimality criterion: sharpness of l.d.f.*

Comparing the amplitude histograms for different degrees of filtering, one recognizes higher and narrower (i.e. sharper) peaks as better indicators of levels. Therefore, the statistical consistency of an event matrix may be characterized by the sharpness of the corresponding $\psi(A)$ function: the closer events are to levels and the lesser event variances are, the sharper the $\psi(A)$ function will be. Any $\psi(A)$ function is normalized to unity

$$\int_{-\infty}^{\infty} \psi(A)\,\mathrm{d}A = 1. \qquad (18)$$

Due to the fixed area below the curve $\psi(A)$, an increase in sharpness is accompanied by an increase in the average value of ψ. Therefore, the sharpness of a ψ function, S, will be defined as its average value represented by the scalar product of the $\psi(A)$ function with itself

$$S = \int_{-\infty}^{\infty} \psi(A)^2 \mathrm{d}A = \sum_{m=1}^{M} p_m \sum_{m'=1}^{M} p_{m'} \boldsymbol{N}(A_m - A_{m'}, \boldsymbol{V}_m + \boldsymbol{V}_{m'}). \qquad (19)$$

The quantity S may be considered as a measure for the likelihood of an event matrix. Therefore, our approach is to maximize S.

The scalar product of p.d.f.s for two arbitrarily chosen events is proportional to $\exp(-G^2/2)$, as is clear from equations (4) and (19). For this reason it decreases rapidly with increasing G value. Therefore, if due to deletion of false transitions, the event currents are more compactly distributed around corresponding levels, the sharpness will increase. The same property shows that the cross products, i.e. the products of the events, belonging to different levels, practically do not contribute to the sharpness (for more details see the discussion).

The values of S for ZC matrices at different X are shown in figure 3*a*. The sharpest ZC matrix is found at $X = 3.3$ or 3.4. It is the same matrix as found by direct comparison with the original matrix ZP and has 48 events. All the transitions in the sharpest ZC coincide with expected ones, although the expected matrix has 50 events.

(iii) *Step 3 (ZM): Detecting levels as a secondary quantization*

Level detection may be considered as a sequence of quantization steps. By quantization we mean a procedure of finding a smaller set of numbers, representing system properties. A trivial quantization is represented by the digitization of the analogue data, replacing a continuum by N data points. The idealization of the record may be considered as a primary quantization, replacing N data points by a much smaller set of numbers contained in a corresponding event matrix with $M \ll N$ rows. In full analogy, the detection of levels from the events found corresponds to a secondary quantization, resulting in a level matrix with K rows, $K \ll M$ (for our test record $K = 3$). Such a secondary quantization may be done in a matrix form, producing the level matrix immediately from the event matrix. Two further alternatives are presented.

To obtain the level matrix ZM immediately from the event matrix ZC, the lines in ZC are first sorted in the ascending order of event currents. This gives a sorted event matrix, ZS (the ideal trace for ZS is shown in figure 1*e*). The third column in ZS contains the local estimates of the event variances. This allows t-testing in ZS matrix, which corresponds to classical use of the Student's t-test, because the positions of the transitions tested are already predefined. This testing does not change the transition coordinate, but deletes subthreshold transitions. Due to the sorting, the differences of event currents in consecutive rows of ZS (and therefore the T values) are significantly decreased in comparison with those in ZC. In order to obtain a guess for the level matrix ZM, we delete all subthreshold transitions in ZS at some X value, which is accepted independently of initial X value, used with T filter. The number of rows in the remaining matrix, recognized as a guess for ZM for a given X value, is considered as a guess for K. A certain specific feature of the dependence $K(X)$ allows one to find the real level matrix.

Figure 2 shows the dependencies $K(X)$ for different ZS matrices, each of them being obtained at $W = 20$ and different initial X values (o, $X = 2$; +, $X = 2.5$; ×, $X = 3$). Further increase of the initial X value up

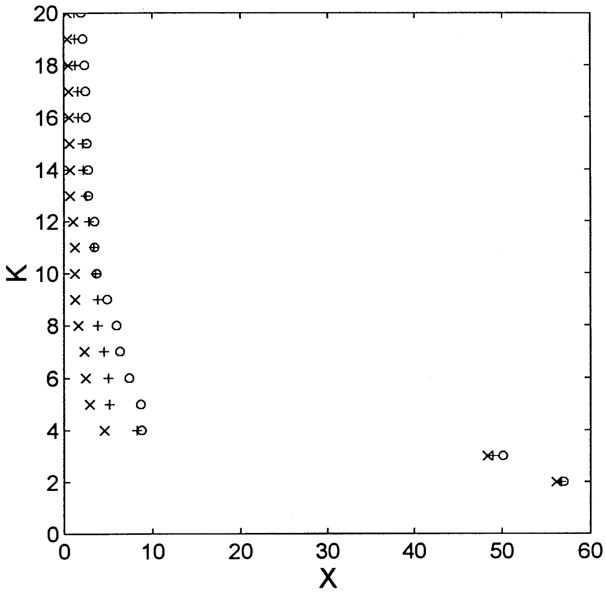


Figure 2. Guesses for K, corresponding to different ZS matrices. Starting ZS matrices are obtained at initial X values: (o, 2, +, 2.5; × 3). The big jump of X value at transition from $K = 4$ to $K = 3$ is typical for appearance of real levels.

to four gives results close to those for $X = 3$. The big jump in X value (figure 2) between $K = 4$ and $K = 3$ indicates that the real value of K is three. The ideal trace corresponding to ZM constructed in this way is shown in figure 1*f*. The wide gap in X values, separating the real level matrix from different guesses, is the consequence of the macroscopic nature of ZM matrices, in which rows are constructed from many events of ZS matrix. This is quantitatively reflected by the asymptotic (i.e. for sufficiently high N) proportionality of the gap width to $N^{1/2}$.

We have described the procedure of detecting levels from the sorted event matrix ZS. Two further alternatives are based on a 'record' r_s, defined as a record, correspondent to the event matrix ZS. It is obtained from the record r in the same way as the matrix ZS was obtained from ZC. The first point(s) of r_s is (are) the data point(s) belonging to the first event in ZS and taken in the same sequence. The data point(s) belonging to the second event in ZS is (are) taken as next point(s) of r_s, etc.

The first and the simplest alternative is standard one, e.g. the levels are detected as pronounced peaks in the point amplitude histogram of the low-pass filtered record r_s, which may be low-pass filtered much deeper than r due to essentially decreased number of real transitions in r_s ($K \ll M$).

The second, slower but somewhat more accurate alternative uses statistical filtering instead of the low-pass filtering. The level matrix ZM is obtained from r_s as an event matrix produced by T filter at $W = N$ and sufficiently high X, taken as any value from the wide gap shown in figure 2. The justification of using T filter for this version of the secondary quantization is the fact that the noise values in the record r_s are uncorrelated. An insignificant amount of adjacent one-point events in ZS may be omitted while constructing r_s. This may be done if one-point events do not form their own level, which is almost always the case.

As we have seen from figure 2, the level matrices can be obtained from ZC matrices, produced by the T filter in a wide range of initial X values (from two to four). One can expect an absolutely best result for ZM in the case when the starting matrix is as close to the original matrix ZP as possible. The question is, whether there is a better starting matrix than the sharpest ZC matrix. If the levels are well expressed, as in our example, the practical necessity of looking for a better starting matrix is negligible. For instance, the levels produced from the sharpest ZC matrix are very close to those produced from the programmed event matrix. Moreover, the estimates of the levels for our test record, obtained from ZC corresponding to $X = 2$ and $X = 4$, differ only by about 0.03. For worse data, looking for a better starting matrix may be more significant. The subsequent steps are described to find an optimum event matrix, called ZO.

(*b*) *Optimization of event matrices at constant threshold*

The Student's *t*-test works without preliminary knowledge of levels. As a consequence, sometimes adjacent events most likely belong to the same level (the closest one to both event currents). This is an unavoidable price for increase of amplitude resolution with growing testing window. As we have seen, the deletion of all such false transitions by increasing X value may cost several extra real transitions. The knowledge of the levels changes the problem principally: it allows us to get rid of this kind of false transitions by switching to a different test with built-in information about levels (such as Hinkley test, Schultze & Draber 1993, see also discussion). However, even if we would switch, we would still have to find the optimum threshold value for the new test (cf. Pastushenko *et al.* 1997). Moreover, such a switching assumes that the levels are already known, whereas we are still trying to improve the level estimates. By these reasons, we describe here a simpler alternative, based on explicit deletion of the false transitions. This allows us to demonstrate the performance of the new optimality criterion, which may be used also with different statistical tests, such as Gaussian or Hinkley tests.

(i) *Step 4 (ZL): Transition correction/ghost deletion using levels*

Let us denote the levels (the second column in ZM) by A_k, $k = 1, 2, \ldots K$. The boundaries between levels, B_k, are defined as midpoints between neighbouring A_k values

$$B_{\underline{k}} = (A_{\underline{k}} + A_{\underline{k}+1})/2, \quad \underline{k} = 1, 2, \ldots, K-1. \tag{20}$$

If two adjacent event currents are not separated by any boundary, then the events are most likely belonging to the same level. The corresponding transition is then deleted. Such a process of merging dwells produces a linked matrix ZL. Merging dwells gives a better

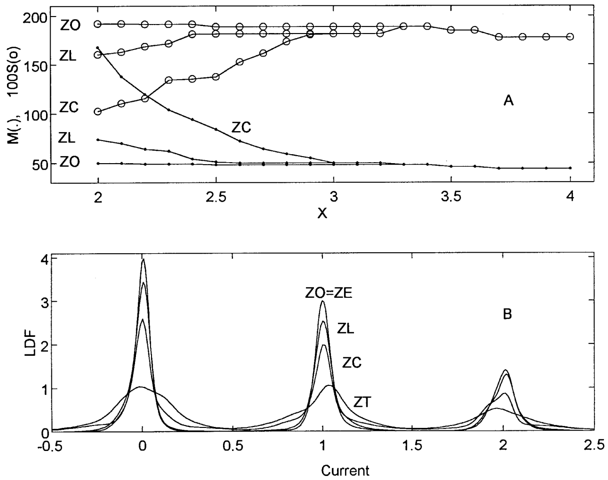


Figure 3. Stages of T filtering/optimization, performed at different values of X at $W = 20$. (a) Sharpness is shown as 100 S (circles), numbers of events in different matrices are shown by points. (b) different l.d.f.s at $X = 2$, $W = 20$. Note the growth of the sharpness in sequence ZT < ZC < ZL < ZO.

estimate for the ideal trace and increases the sharpness (cf. figure 1 d, g) but does not affect the level estimates. The sharpnesses (circles) and the numbers of events (points) in ZL matrices are shown in figure 3 a in relation to the X value. The sharpest ZL matrix has 48 events. ZL matrix with 50 events is also different from ZE. *By definition*, the expected matrix ZE is obtained from the original matrix ZP by correcting the transitions and subsequent merging dwells. The transition correction is made by step 2 of the T filter at $X = 0$ and $W = N$. Due to noise, in our test record two transitions are shifted by one point and one transition is shifted by two points; merging dwells after correcting ZP does not change the number of events.

The ψ-functions for ZT, ZC and ZL produced at $X = 2$ and $W = 20$, are shown in figure 3 b. Almost the same curves would be obtained for higher values of W. As expected, the peaks become higher and sharper in the sequence ZT → ZC → ZL. What could be the next step to improve further the estimates for transitions? We have constructed the following procedure to obtain an optimum matrix ZO from ZL, using the sharpness as a guiding line.

(ii) *Step 5 (ZO): maximization of sharpness at constant threshold*

Suppose that a matrix ZL was obtained at some X value. In this subsection, we describe the procedure of obtaining ZO from ZL at fixed X value. The first guess for matrix ZO is given by ZO = ZL. The transitions in this ZO are numbered in the ascending order of T^2 (it would be more logical to make this sorting in the ascending order of G^2, but this gives very small improvement at the cost of increased time). We make then an attempt to increase the sharpness of ZO by deleting the first transition (in spite of the fact that it is super-threshold one) and applying to the remaining transitions the second step of the T filter at the initial X value, with subsequent merging dwells. If the attempt is successful, i.e. if the new matrix is sharper, it is accepted as the next guess for ZO. Then the whole procedure repeats. If the new matrix has smaller sharpness, it is not accepted as the next guess for ZO. In this case the next attempt starts from the next transition. In this way the attempts may be continued until all the transitions have been verified. The result is recognized as the optimum matrix, ZO.

Figure 1 h shows an ideal trace corresponding to a certain ZO for our test record. The same matrix ZO is obtained at higher values of W (30 and 40). The matrices ZO coincide exactly with the expected matrix ZE for $X < 2.2$. The sharpnesses and the numbers of events in ZO matrices for different X values are shown in figure 3 a. An example of l.d.f. for ZO is shown in figure 3 b. For our test record, all the matrices ZO contain only expected transitions (for different test records of the same type, this was true in most cases).

(c) *Effects of correlated noise/additional filtering*

According to its derivation, the Student's t-test (like Gaussian test) may be applied to records with uncorrelated Gaussian noise. However, due to proportionality between cut-off frequency f_c and sampling frequency, the noise is usually correlated. The main effect of the correlated noise is expressed in an increase of the T values due to underestimated noise variances. For dwells longer than several correlation times the variances will be underestimated by approximately the same factor. Thus, the correlated noise causes mainly a shift of the relevant T or X values. If the sampling rate is not much higher than the standard one, this shift has little influence on the optimal recovery of transitions.

To illustrate this with the same test record, we have filtered it by an unbiased digital exponential finite impulse response (FIR) filter with an efficiency of $\sqrt{2}$, which corresponds approximately to averaging over two points, figure 4 a. By 'efficiency' we mean the factor by which the r.m.s. of white noise (in the absence of signal) decreases after filtering. It can be shown analytically that the exponential filter is the best of all other filter types (compared at the same efficiency) in the sense of minimum distortion of the signal at a single transition (i.e. between two long real events). The exponential filtering with efficiency of $\sqrt{2}$ produces almost the same correlation in noise as found with white noise after filtering with analogue RC filter and sampling with the standard interval $t_s = 1/(2\pi f_c)$.

The results of all five steps of the analysis are shown in figures 4 and 5. One difference with the previous results for non-filtered data (figure 1) is that ZL is now much closer to ZO and even coincides with ZO at $X > 2.1$.

Another difference to the previous results is that the sharpest ZO matrix has now 52 events. Two additional events, each consisting of one point, are shown by circles in figure 4 f. Working only with additionally filtered data, we would have to recognize 52 events as the best result. However, we can recalculate the event currents and variances in ZO matrix, using the dwells from ZO, but the data points from the non-filtered record r. After such recalculation, the matrix with 52 events will not be the sharpest one any more. Moreover, the matrices ZL and ZO with 50 events, obtained in the range $2.2 \leqslant X \leqslant 2.6$ after recalculation of event currents and variances, are less sharp than ZE for non-filtered data. This is due to the shifts of two transitions by the low-pass filtering. The effect of transition shifts can be removed by correcting transitions, using the non-filtered record r (the transition correction is made by the step 2 of the T filter at $X = 0$, with subsequent merging dwells). After this operation ZL and ZO matrices with 50 events coincide with ZE for non-filtered data. This illustrates the decrease of the time resolution by the low-pass filtering due to shifts of transitions (Colquhoun & Sigworth 1983). It indicates simultaneously the possible way of correction, which represents an alternative to the time-course fitting. For practical recordings, this suggests that higher cut-off frequencies for the recording devices are preferential in spite of the increased noise amplitude, unless the noise power spectrum is steeply growing at f_c.

The shift of the highest X value, at which the expected matrix can still be recovered, from 2.1 to 2.6, is caused by the filtering, and, due to the fact that the filtering was not very deep, is of no importance.

Figure 5 shows the sharpnesses and the $\psi(A)$ functions for different event matrices obtained from the filtered record at $X = 2$ and $W = 20$. The l.d.f. of ZE for filtered data is not shown: it optically coincides with l.d.f. of ZO. The sharpnesses of ZC and ZT now differ considerably less than in figure 3. Thus, the low-pass filtering has already used up a good part of reserves for increasing the statistical consistency. Therefore, additional low-pass filtering and T filtering are competing for essentially the same reserves. However, the low-pass filtering is equivalent to some averaging of the data points, ignoring the transitions, whereas T filter avoids the averaging across transitions. This is the main reason for better resolution of statistical filters, represented in given case by T filter.

The matrices ZL and ZE for filtered data coincide in the range of X values from 2.2 to 2.6, whereas for non-filtered data none of ZL matrices coincided with ZE. This does not mean that the low-pass filtering has rendered the idealization an easy problem. For instance, simply attributing the data points to the closest level, we would still obtain 91 events – much more than programmed 50 events (and before the filtering we would get in this way 224 events). If we low-pass filter the data so that SNR $= \Delta/\sigma$ increases up to six (the 50% method recommends to increase this ratio up to 10), we still find 63 events by the 50% method. At the same time, several real transitions are already lost (over-filtered). This situation is typical for records where transitions between remote levels are possible (i.e. if transitions are allowed not only between adjacent levels). For records with only two levels this difficulty does not exist.

(*d*) *Choice of the initial parameters X and W*

For maximum amplitude resolution, the maximum window is unlimited, $W = N$. Then the problem is reduced to selection of X, based on the search for the sharpest event matrix. The price for maximum amplitude resolution can be some extra transitions (deleted later by merging dwells), and somewhat higher calculation time. This technical detail is the main reason to accept $W \ll N$.

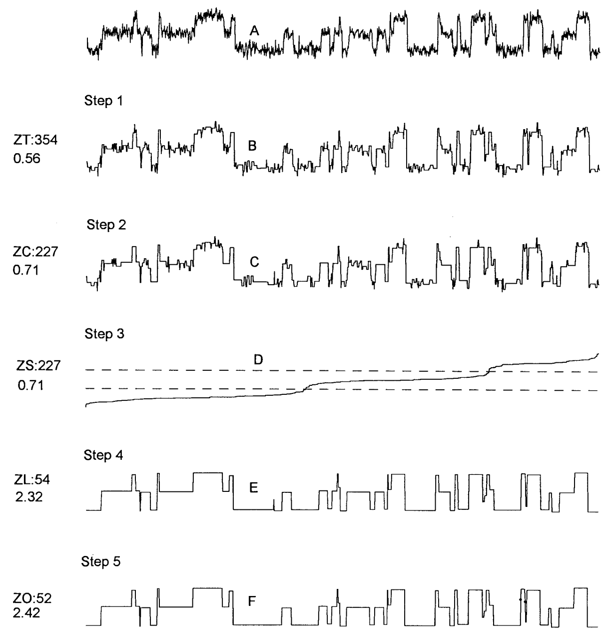


Figure 4. Some stages of the analysis of the test-record r filtered with efficiency 1.41: (*a*) filtered record; (*b*) to (*f*) different traces (matrix name, number of events and sharpness are shown near each trace).

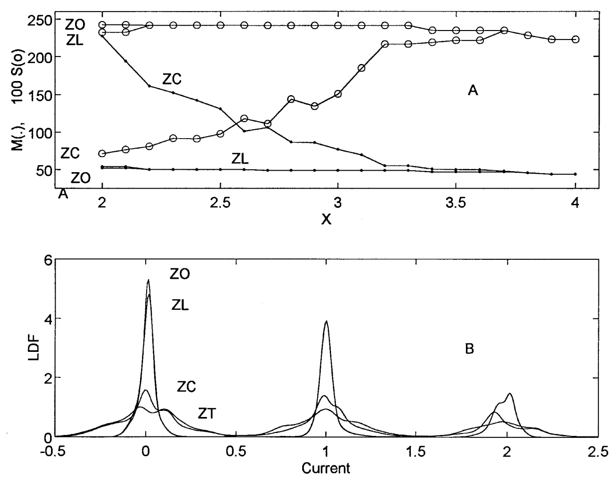


Figure 5. As figure 3, except record r was filtered with efficiency 1.41.

A logically possible strategy for selecting X and W parameters as those which lead to sharpest ZO matrices is time consuming. A much simpler and faster practical version is based on relative robustness of levels, which are almost equally well estimated from the sharpest of ZL matrices. By this reason the term 'optimum X value' will be understood in the sense of the sharpest ZL matrices. The fluctuations of optimum X values from record to record decrease with increasing N. To illustrate the selection of optimum X and W values, a longer record was generated. The t-testing was carried out in the range of X values from 1.9 to 3.9. Maximum testing windows assumed one of the values $W = 10$, 16, 22, 28, 34, 40, 46, 52, 70, 100, 200. For matrices ZC and ZL, calculated for $W > 10$, numbers of events and sharpnesses are shown in figures 6*d*, *e*. The maximum sharpness of ZL at fixed W is shown in figure 6*f* in relation to $\ln(W)$. The sharpest ZL is found at $X = 3.3$ and $W = 28$, which are optimum for this case.

For practical purposes, one may further simplify the choice of X and W, using the fact that the results are mostly sensitive to X values and much less to W values, if the latter are sufficiently high. This allows some standard choice of W to be made for a given X, screening subsequently only different X values, i.e. in one dimension. A practically acceptable rule is $W = 20\lambda$, where λ is the dead time, defined as the shortest detectable dwell time for transitions between two closest levels. The distance between the closest levels is denoted by Δ. Strictly speaking, due to fluctuations the shortest time for transitions to the first or to the last levels is always equal to one. Thus by λ we mean the average dwell of such events, which in the neighbourhood of very long events produce the threshold value of the test. We shall estimate λ for noisy records, in which case $\lambda \gg 1$. Then the T value can be replaced by the G value, and X_n by X. Assuming that one of the two tested dwells is equal to $\lambda \ll W$, one may write

$$T \approx G = \sqrt{\lambda(1-\lambda/W)}\Delta/\sigma \approx \lambda^{1/2}\Delta/\sigma \geqslant X, \quad \text{and} \quad \lambda = (\sigma X/\Delta)^2. \qquad (21)$$

For our test record (optimum value of $X = 3.3$ and $\sigma/\Delta = 0.3$) this gives $\lambda \approx 1$. In other words, one point events are partly resolvable at $X = 3.3$, if they have sufficiently long neighbours.

5. DISCUSSION

Several scientific and practical aspects of the new method should be discussed and compared with existing methods. One of the first aspects is the quality of the idealizer. The T filter could be compared with F testing by Kirlin & Moghaddamjoo (1986) or with published recently method of maximum slopes (VanDongen 1996). The asymmetric testing in T filter in combination with the dynamic self-adjusting testing window provides for its applicability to a wider class of data, or for a higher resolution, if applied to the same data.

Presently, the standard method of level detection based on analysis of the point amplitude histograms of additionally low-pass filtered data with low SNR works satisfactorily only in the case of sufficiently long average dwells. An attempt to increase the resolution using p.d.f. or l.d.f. in combination with idealization methods based on preliminary knowledge of levels (such as 50% and Hinkley methods), is not always helpful. This is illustrated in figure 7, where our results are compared with those obtained by the 50% amplitude method applied under the assumption of two existing levels, which are easily revealed from the point amplitude histograms.

There are no indications of three levels in the point amplitude histograms even for deep filtering (up to efficiency 6, actually remarkably over-filtered). An attempt to see the third level with the help of p.d.f. of the event matrix produced by 50% amplitude method, fails (cf. figure 7*b*), in spite of the fact that we used the optimal low-pass filter (exponential). Efficiency 1 corresponds to non-filtered data. With efficiency 1.41 one should reach approximately $\Delta/\sigma = 4.5$, which is somewhat less than the value of 5–6 recommended for this method (Sigworth 1983). For efficiency 3 the SNR value is 9.5, well above the recommended ratio. No indication of three levels is seen from any of the three curves. Practically the same results would be found with Hinkley method, which is a somewhat better idealizer than the 50% method, if the data require additional low-pass filtering.

The level detection by our method is much better. Figure 7*c* shows the p.d.f. and l.d.f. for ZL matrix. The l.d.f. for levels, denoted by l.d.f._m(three narrow sharp peaks), is reduced 40 times for better scaling.

This example should not be understood in the sense that the idealization by 50% amplitude or by Hinkley method is *always* worse than the idealization by T filter. For instance, if we compare 50%, Hinkley, Gauss and T filters as idealizers for records with *two known* levels and *known* noise variance, then the best ideal traces will be ordered as Hinkley > 50% > Gauss > T filter (for comparison of Student and Gauss see Pastushenko *et al.* (1996)). This sequence is quite natural, because the first two methods use more additional information, and therefore they solve an easier problem than the last two. Correspondingly, the Gaussian test is better than Student, because 'it knows more than Student'. The operation of merging dwells in combination with Gaussian or Student testing decreases their difference with other methods, and may even change the hierarchy for multilevel records with sufficiently low SNR. The example in figure 7 shows that using information about levels, frequently considered as the advantage of the first two methods, may become their disadvantage, if the level detection is not a trivial problem. The other way around, the 'disadvantage' of the Student's *t*-test may become its advantage if levels and noise variances are unknown. In fact one should not speak about disadvantage of Student or Gaussian tests – these tests are appropriate in corresponding conditions, which may lead to their superiority, as demonstrated in figure 7. We speak about both these tests because all the procedures described in this paper (with exception of estimating the noise variance) are applicable after replacing the Student's *t*-test by the Gaussian test, which gives the Gaussian equivalent of the T filter. The switch to Gaussian filter may be even recommended after finding the noise variance, which is very well estimated at the optimum X value. One can ask, whether the Gaussian filter does not occupy the whole niche of the T filter. If the noise amplitude σ is known in advance, and if σ is surely the same for different levels, then the Gaussian filter has a preference. As we have indicated, this is possible not in all experiments. Both Student's *t*-test and Gaussian test are derived in assumption of the same σ for different levels. Nevertheless, the Student's *t*-test may appear preferential in comparison with Gaussian test for multilevel records, if σ significantly varies from level to level.

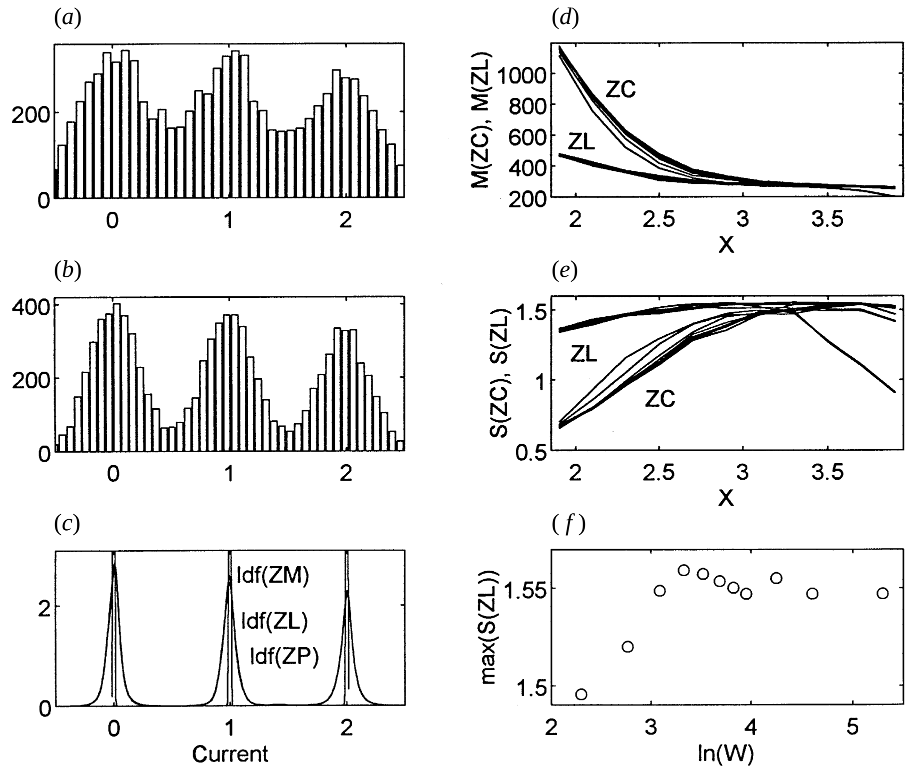


Figure 6. Illustration for choice of the parameters X and W. The test record contains 300 real events, expected average dwell 30, $\sigma/\Delta = 0.3$. Point amplitude histograms: (*a*) for white-noise data; (*b*) for additionally filtered data with efficiency 1.41. (*c*) different l.d.f.s: ψ(ZP) and ψ(ZL) for the sharpest ZL practically coincide. ψ(ZM) is very sharp and high (its main part is out of the picture); the width of ψ(ZM) reflects uncertainty in detected levels. (*d*) M(ZC) and M(ZL) versus X for $W > 10$. Note that the two bundles of curves fuse in the region of optimum X. (*e*) Sharpnesses of ZC and ZL versus X. (*f*) Maximal at given W sharpness versus ln(W). The sharpest ZL is found at $W = 28$, $X = 3.3$.

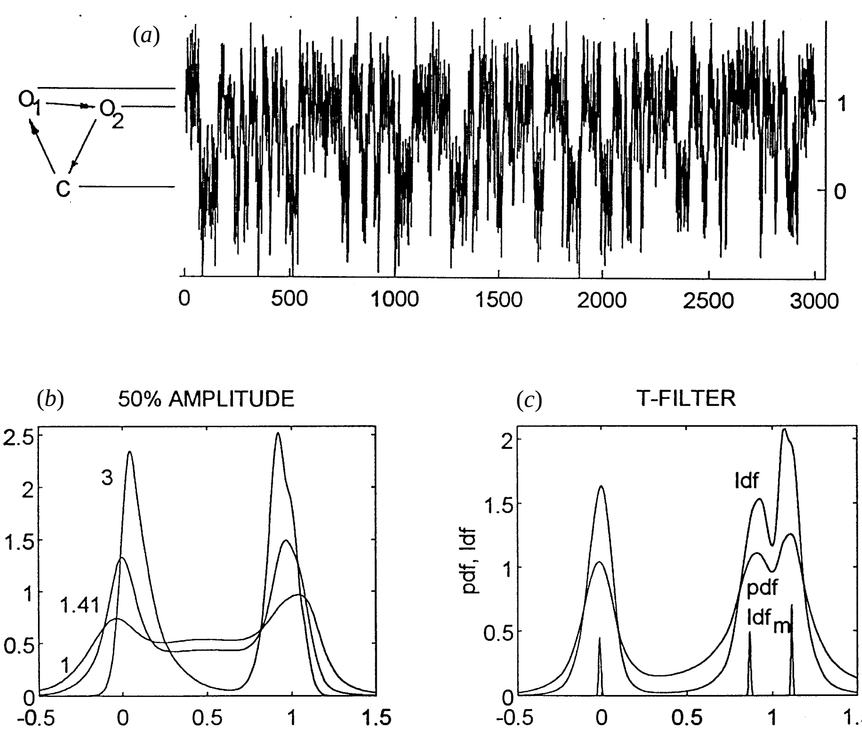
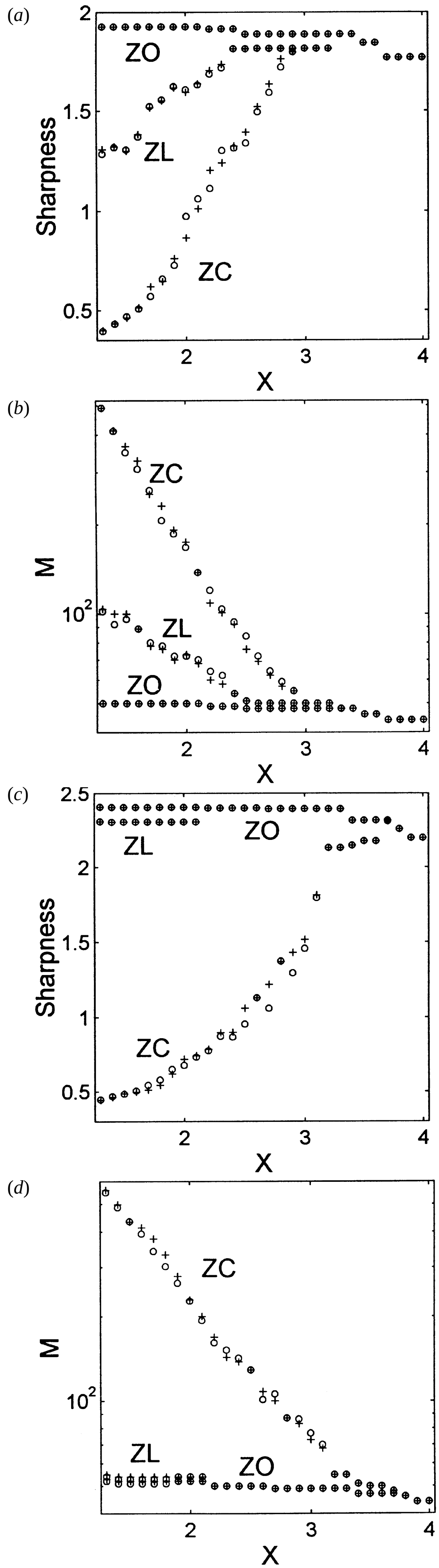


Figure 7. Comparison of different methods of level detection. The record with 300 real events was simulated, with three levels at 0, 1.1 and 0.9, shown as closed state C and open states O_1 and O_2 respectively. (*a*) First 3000 simulated data points and kinetic scheme for channel activity. Levels are shown by horizontal lines. (*b*) p.d.f.s for matrices obtained by 50% method from differently filtered data (filtering efficiency is shown near the curves).(*c*) Results of T filter: p.d.f. and l.d.f. for the same ZL matrix obtained at $X = 2.5$ and $W = 50$. Note that the l.d.f. is sharper than p.d.f. The l.d.f. for levels (l.d.f.$_m$) is plotted with scaling factor 1/40.



There are several technical points that should be mentioned. One of them is related to the global estimation of the noise variance. As long as the levels are unknown, or the drift is significant, the definition (equation (13)) gives the only possibility. However, as soon as the level estimates in a stable record are known, one can improve the estimate for σ^2, using the level matrix obtained as event matrix for the sorted record, r_s. The difference between the two σ^2 estimates grows with decreasing X.

In full analogy with the estimate of the noise variance, one can somewhat improve the definition of the sharpness using the knowledge of the levels. As one may see from equation (19), a certain contribution to the sharpness is made by cross products due to events belonging to different levels. Such events belong to different classes and it is more logical to omit the corresponding cross products. As long as we did not know the levels, we could only rely on small contribution of the cross products into the sharpness. There are several reasons for this. (i) The cross products quickly decrease with increasing difference of event currents. (ii) Short events belonging to different levels may occasionally appear close to each other. However, their contribution to the total sharpness is small. This is partly due to their small statistical weights, partly due to their high variances (level variance is inversely proportional to dwell), and partly due to their deviations from levels. (iii) Relatively long events, which give the main contribution to the sharpness, make extremely small cross products due to small values of event variances. With growing noise variance, or with decreasing X values, when the average dwell decreases, the omitting of cross products becomes more important, because in both cases the event variances grow. Formally omitting cross products corresponds to the view on ψ-function as a vector with K components. Within this concept, the cross products are automatically not included into the sharpness, defined in the same way, i.e. as the squared norm of the (vector) ψ-function (Pastushenko & Schindler 1995).

The third point is the question, whether the results are the same when the data are analysed in reversed sequence. All the three points are addressed in figure 8 for the non-filtered (a, b) and filtered (c, d) test record. The calculations were made at $W = 20$ with improved estimates of the noise variance. Correspondingly, cross products in the sharpness were omitted. The sharpnesses (a, c) and numbers of events (b, d) are presented for matrices ZC, ZL and ZO in a wider range of X values (beginning from $X = 1.3$ in steps of 0.1). We did not try even smaller values of X because already at $X = 1.3$ we have ZC with 560 events for non-filtered data and with almost 600 events for filtered data.

For X values from 2 to 4, the results for forward analysis (circles) practically coincide with those pre-

Figure 8. Results of the analysis in forward (○) and backward (+) directions. The same record as in figures 1 to 5. (a, b) Non-filtered: the difference between both directions exists only for ZC and ZL matrices, and for not very high X values. (c, d) Filtered: ZO matrices in backward direction have one event more than in forward direction only for $X < 1.9$.

sented earlier. In this region the dwells and the event currents are exactly the same for all matrices. The sharpnesses are changed only in fourth digit. This justifies the use of the definition (19).

For non-filtered data, ZO matrices calculated in both directions are the same in the whole range $1.3 \leqslant X \leqslant 4$. The difference between forward and backward directions exists only for ZC and ZL matrices, as long as X value is not very high. The sharpest ZC and ZL matrices for both directions are the same. For smaller X values, both ZC and ZL matrices represent intermediate results, therefore the difference between both directions is not important. It is possible to construct the time symmetric version of the T filter. However, due to the fact that the time symmetry is not the main problem, we have selected for presentation the simplest, solid and logically consistent version, aiming to achieve the highest possible resolution. One of the advantages of this version is that ZT matrices, which may be produced practically simultaneously with the recording, are relatively close to ZC matrices at sufficiently high X values. This enables a preliminary level detection without even saving the record. Therefore, the first step of the T filter in combination with matrix level detection may be realized in a form of a hardware with relatively short memory.

It is interesting that for non-filtered data the range of X values where ZO coincides with ZE is now increased (1.3–2.1). For filtered data, analogous interval remains the same (2.2–2.6). For $X < 2.2$ we find ZO matrices for filtered data with 51, 52 and even with 53 events (the latter in the reversed direction at $X = 1.3$). Taking into account that these 53 events are found from initially 600 events, it should be recognized as a good result.

The main scientific novelty of this paper is the new optimality principle, maximum of an incomplete likelihood. As we have demonstrated, for data of moderate quality it works very well even in its simplest form. At the same time, one can ask for a statistically reliable assessment of the performance of the method for a wider class of data. Recently, we have developed special procedures for such assessment via direct comparison with ZP (Pastushenko *et al.* 1997). These procedures are especially efficient for comparison of the qualities of idealization by different methods. Additionally, they allow one to find an objectively optimum test threshold value (in our case X value). Using these procedures, we have made a rigorous test for S criterion, having combined it with Hinkley test. The only problem for very difficult data was to take into account the probability of missed events while estimating event variances, equation (15). We were able to show that S criterion excellently recovers the objectively optimum X values, even for data with very high noise and relatively short events, including non-HMM-type data (preliminary report: Pastushenko & Schindler 1995). Together with previous demonstrations, these findings strongly suggest the high predictive power of S criterion.

Our method may be used as a prefilter or a control in combination with other methods of analysis, including HMM methods. The estimates for levels and noise variances, used as the starting values, save considerable computation time. On the other hand, for data of reasonable quality, the optimum event matrix may be immediately used for the recovery of the channel mechanism.

Another advantage of our method is its possible use for detrending data. The local character of the t-test allows finding transitions from a limited number of adjacent points where the role of the drift is negligible. The analysis of the event currents allows to detect the drift and to eliminate it from the record for subsequent optimal analysis.

A version of the method may also be applied to unstable records without drift elimination. We did this for simulated channel data and for jump wise growing capacitive current. In this case one should optimize the l.d.f. for jump amplitudes instead of the l.d.f. for event currents. The jump amplitudes are given by the differences of adjacent event currents, and their expected variances are estimated exactly as in derivations of the statistical tests.

These additional possibilities illustrate the attractivity of the new method as a tool for analysis of ion-channel records or similar data.

The authors express their gratitude to Dr F .J. Sigworth (Yale University), Dr K. L. Magleby (University of Miami), Dr J. B. Cooper, Dr R. Hafner, C. Romanin, Dr P. Weiß (University of Linz) and Dr G. Zamponi (University of Calgary) for useful discussions and remarks on the subject of this article. Our deep thanks are to Dr C. Moler (Mathworks, Natick) for generous gifting random number generators of new generation. Supported by the Austrian Research Funds, grant S066–10 MED.

REFERENCES

Abramowitz, M. & Stegun, I. A. 1965 *Handbook of mathematical functions with formulas, graphs and mathematical tables*. New York: Dover.

Becker, J. D., Honerkamp, J., Hirsch, J., Fröbe, U., Schlatter, E. & Greger, R. 1994 Analysing ion channels with hidden Markov models. *Pflügers Arch. GES. Physiol.* **426**, 328–332.

Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. 1994 *Time series analysis. Forecasting and control*. Englewood Cliffs, New Jersey: Prentice-Hall.

Chung, S. H. & Kennedy, R. A. 1991 Forward-backward non-linear filtering technique for extracting small biological signals from noise. *J. neurosci. Methods* **40**, 71–86.

Chung, S. H, Moore J. B, Xia, L., Premkumar, L. S., & Gage, P. W. 1990 Characterization of single channel currents using digital processing techniques based on hidden Markov Models. *Phil. Trans. R. Soc. Lond.* B. **329**, 265–285.

Colquhoun, D. 1987 Practical analysis of single channel records. In *Microelectrode Techniques. The Plymouth Workshop Handbook* (ed. N. B. Standen, P. T. Gra & M. J. Whitaker), pp. 83–104. Cambridge.

Colquhoun, D. & Hawkes, A. G. 1983 The principles of the stochastic interpretation of ion channel mechanism. In *Single channel recording* (ed. B. Sakmann & E. Neher), pp. 135–175. New York: Plenum Press.

Colquhoun, D. & Sigworth, F. J. 1983 Fitting and statistical analysis of single channel records. In *Single channel recording* (ed. B. Sakmann & E. Neher), pp. 191–263. New York: Plenum Press.

Dempster, A. P., Laird, N. M. & Rubin, D. B. 1977 Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.* B. **39**, 1–38.

Fox, J. A. 1987 Ion channel subconductance states. *J. Membrane Biol.* **97**, 1–8.

Fredkin, D. R. & Rice, J. A. 1992*a* Maximum likelihood estimation and identification directly from single-channel recordings. *Proc. R. Soc. Lond.* B. **249**, 125–132.

Fredkin, D. R. & Rice, J. A. 1992*b* Bayesian restoration of single-channel patch clamp recordings. *Biometrics* **48**, 427–448.

French, R. J. & Wonderlin, W. F. 1992 Software for acquisition and analysis of ion channel data: choices, tasks, and strategies. In *Methods in enzymology. Ion channels* vol. 207 (ed. B. Rudy & L. E. Iverson), pp. 711–728. Academic Press, Inc.

Hafner, R. 1989 *Wahrscheinlichkeitsrechnung und Statistik, Wien.* New York: Springer-Verlag.

Heinemann, S. H. 1995 Guide to data acquisition and analysis. In *Single channel recording*, 2nd edn (eds. B. Sakmann & E. Neher), pp. 53–91. New York, London: Plenum Press.

Heinemann, S. H. & Sigworth, F. J. 1991 Open channel noise. Analysis of amplitude histograms to determine rapid kinetic parameters. *Biophys. J.* **60**, 577–587.

Heinhold, J. & Gaede, H. 1968 *Ingeneur-statistik*. München, Wien: R. Oldenbourg.

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.

Jackson, M. B. 1992 Stationary single-channel analysis. In *Methods in enzymology. Ion channels* vol. 207 (ed. B. Rudy & L. E. Iverson), pp. 729–746. Academic Press.

Kirlin, R. L. & Moghaddamjoo, A. 1986 A robust-window detector and estimator for step-signals in contaminated gaussian noise. *IEEE Trans. Acoustic, Speech, Signal Processing*. **34**, 816–823.

Korn, S. J. & Horn, R. 1988 Statistical discrimination of fractal and markov models of single-channel gating. *Biophys. J.* **54**, 871–877.

Magleby, K. L. 1992 Preventing artifacts and reducing errors in single-channel analysis. In *Methods in Enzymology. Ion channels*, vol. 207 (ed. B. Rudy & L. E. Iverson), pp. 763–791. Academic Press

Magleby, K. L. & Weiss, D. S. 1990*a* Estimating kinetic parameters for single channels with simulation. *Biophys. J.* **58**, 1411–1426.

Magleby, K. L. & Weiss, D. S. 1990*b* Identifying kinetic gating mechanisms by using two-dimensional distributions of simulated dwell times. *Proc. R. Soc. Lond.* B. **241**, 220–228.

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 Automatic segmentation and classification of ionic-channel signals. *IEEE Trans. Biomed. Eng.* **38**, 149–155.

Oppenheim, A. V. & Schaefer, R. W. 1975 *Digital signal processing*. Englewood Cliffs, New Jersey: Prentice Hall.

Pastushenko, V. Ph. & Schindler, H. 1993 Statistical filtering of the single ion channel records. *Acta Pharm.* **43**, 7–13.

Pastushenko, V. Ph. & Schindler, H. 1994 Idealization of ion channel records by Student's *t*-test filtering with subsequent optimization of level distribution likelihood. *Biophys. J.* **66**, A440.

Pastushenko, V. Ph. & Schindler, H. 1995 Epoch oriented likelihood maximization of an idealized ion channel record. *Biophys. J.* **68**, A148.

Pastushenko, V. Ph., Schindler, H. & Reza, A. M. 1997 Predictive optimality criterion for idealization of ion channel data and exact akaike's criterion. *Eur. Biophys. J.* (Submitted.)

Pastushenko, V. Ph., Schindler, H., Waldl, H. & Hafner, R. 1996 A new statistical test for analysis of patch-clamp data. *Biophys. J.* **70**, A200.

Patlak, J. B. 1988 Sodium channel subconductance events measured with a new Variance-Mean Analysis. *Biophys. J.* **54**, 413–430.

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, L. R. 1989 A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE.* **77**, 257–285.

Sachs, F. 1983 Automated analysis of single-channel records. In *Single channel recording* (ed. B. Sakmann & E. Neher), pp. 265–285. New York: Plenum Press.

Schindler, H. & Pastushenko V. Ph. 1992 Statistical filtering of ion channel data. *Biophys J.* **61**, 1–248

Schultze, R. & Draber, S. 1993 A nonlinear filter algorithm for the detection of jumps in patch-clamp data. *J. Membrane Biol.* **132**, 41–52.

Sigworth, F. J. 1983 An example of analysis. In *Single channel recording* (ed. B. Sakmann & E. Neher), pp. 301–321. New York: Plenum Press.

Terrien, C. W. 1992 *Discrete random signals and statistical signal processing*. Englewood Cliffs: Prentice Hall.

Tyerman, S. D., Terry, B. R. & Findlay, G. P. 1992 Multiple conductances in the large K channel from *Chara corallina* shown by a transient analysis method. *Biophys. J.* **61**, 736–749.

VanDongen, A. M. J. 1996 A new algorithm for idealizing single ion channel data containing multiple unknown conductance levels. *Biophys. J.* **70**, 1303–1315.

Received 16 November 1995; accepted 13 August 1996