SPIE 1994 International Symposium on Optics, Imaging, and Instrumentation. 1994
DOI 10.1117/12.188812

Detecting Ionic Currents in Single Channels using Wavelet Analysis Part I: Zero Mean Gaussian Noise

A. Averbuch†, R.S. Eisenberg‡, M. Israeli§, Z. Schuss†
† School of Mathematical Sciences
Tel Aviv University, Tel Aviv 69978, Israel
‡ Department of Physiology
Rush Medical College
Chicago, Illinois 60612, USA
§ Faculty of Computer Science
Technion, Haifa 32000, Israel

Abstract

The patch clamp technique opened a new field in biological research and shed light on membrane permittivity for ionic currents.

The key element in patch clamp measurements is the detection of the ionic currents in a single biological channel. It is known that the channels open and close at random times, thus modulating the ionic currents. The measured current switches between two levels corresponding to the open and closed states of the channel.

Determining the statistics of the open and closed periods is of crucial importance to the experimenter, because it reflects the response of channel protein to drugs and other factors. The detected signal is strongly corrupted by instrumentation and other noises, rendering the detection of the opening and closing moments extremely difficult.

We describe the use the wavelet transform and its associated multiresolution (multiscale) analysis to detect the currents through single ionic channels corrupted with noise.

Keyword: Wavelets, biological research, ionic currents,channels,detection, initial estimation, training.

1 Overview

In this paper we present the investigation and implementation of new algorithms to detect the currents through ionic channels corrupted with noise.

The detection algorithm is based on the use of wavelet transform and its multiresolution decomposition analysis. The new algorithms for detecting the current levels of single ionic channel measurements are robust and link the wavelet transform to multiscale edge detection. Multiscale analysis generates internal information that can be used to optimize the wavelet algorithm and make it more robust, as insensitive as possible to noise.

Our results demonstrate the feasibility and utility of multiscale analysis and the algorithms are optimized for channels.

The wavelet algorithms work best if they are given some *a priori* estimates of channel amplitude. Often–indeed usually–such estimates are available from separate experimentation, already done, and often already

0-8194-1627-4/94/$6.00

published. In our case, a training algorithm (as a preprocessing step) is needed to generate a decent first guess from the raw data. This guess is then used by the detection algorithm. The training is performed once on initial raw data for the the complete duration of the detection process provided that just one ionic channel is present in the recording. If subconductance states or two channels or more are recorded then we can reapply the training process more often to have a new estimate for the channel amplitude.

The training algorithms work on raw data giving estimates of the two levels (namely of open and closed channels) and their spread. The algorithm is based on the (experimentally) well known stochastic properties of single channels and their detecting apparatus and amplifiers. The numerical suitability and stability of the algorithm is studied since it involves the solution of non-linear equations.

The algorithm initially smoothes the raw signal using the low-pass part of the wavelet transform, and then applies multiscale analysis to detect openings. A number of ways to do this are presented. We show the feasibility of Gaussian smoothing and edge detection previously developed for other purposes.

We were successful in detecting and analysis of short events containing as few as 7 sample points. (These events might be openings of 70 μsec in typical experimental situations.) Two possible improvements occur to us: one to oversample in general, hoping that the steep nonlinear dependence of wavelet resolution on number of samples will overcome the loss in independence of oversampled events. The other is to locally increase the resolution of analysis just at suspected locations of transitions.

The current work is implemented in a robust user friendly software package that performs training and detection with minimal user intervention. Indeed, it may be possible to implement it in an on-line, hands off, form.
The output is a cleaned record of channel openings, as they would appear in the absence of noise. From this output, an event list and all associated statistics can be generated. Our software can act as a front end to existing systems and so will be easy for workers to use.

One advantage of wavelet analysis is that it is surprisingly undemanding of computer resources and so can often be implemented in real time.

Our results reported here are performed on PC/486DX2 (33/66 MHz). The portable C program has been implementated successfully both on workstation and PC.

2 Wavelets - Background

Wavelet theory provides a unified framework for a number of techniques which had been developed independently for various signal processing applications. For example, multiresolution signal processing, used in computer vision; subband coding, developed for speech and image compression; and wavelet series expansions, developed in applied mathematics, have been recently recognized as different view of a single theory. In particular, the Wavelet Transform is of interest for the analysis of non-stationary signals because it provides an alternative to the classical Short-Time Fourier Transform.

For some applications it is desirable to see the wavelet transform as a signal decomposition onto a set of basis functions. In fact, basis functions are called *wavelets*, and they are obtained from a single prototype wavelet by dilations and scaling as well as shifts.

In a wavelet transform, the notion of *scale* is introduced as an alternative to frequency, leading to a so-called *time-scale representation.* This means that a signal is mapped into a time-scale plane (the equivalent of time-frequency plane used in short time Fourier transform).

Over the last few years we have seen the emergence of wavelet analysis as a new method for the study of transient signals, providing a good complement to traditional windowed Fourier transform analysis. The wavelet transform was introduced to remedy the inconvenience of a windowed Fourier transform. It is computed by expanding the signal on a family of functions which are the dilate and translate of a single function. Its discretization is completely known a priori. The wavelet transform allows us to decompose an arbitrary signal into its localized contributions labeled by a scale parameter. It decomposes an arbitrary function into a two-parameter family of elementary wavelets that are obtained by shifts in the time variable but also by dilations (or compression) that are both on the time and the frequency variables. This localization techniques substantially enrich our ability to deal with all types of signal analysis. The locality behavior of the transform is in contrast to the global nature of the classic Fourier transform. The complete representation is calculated by decomposing the signal on a wavelet orthonormal basis and this gives an intermediate representation between a Fourier and a spatial representation. Thus the new method is easier to implement; it permits a better convergence of the reconstruction formulas.

We believe that the wavelet transform theory has a good chance to succeed in many applications where the classical Fourier analysis has failed or is not adequate.

The wavelet transform [1, 2, 3] is a recent method of signal analysis and synthesis. It analyzes signals in terms of *wavelet*-functions limited both in the time and frequency domain.

3 Other methods

Single channel currents are detected in most labs by a combination of low pass filtering and threshold detection. These methods have been optimized over many years taking full advantage of biological knowledge of the signals. We present here other methods where some are based on analysis and others on traditional methods.

Denoising by Coifman[6, 18] - The adapted waveform analysis and denoising is used to denoise signals which are corrupted by noise. The algorithms utilize libraries of orthonormal waveforms (such as wavelet packets and local trigonometric libraries). The method extracts from a signal a coherent part which is well represented by the given waveforms and a noisy or incoherent part. The performance of the algorithm is independent from the statistical characterization of the signal. Therefore, the separation between noise and non-noise parts of the signal is a robust iterative method.

Denoising by thresholding the wavelet coefficients was proposed by Donoho[7].

Another use of thresholding on the measured data with smart filtering was proposed by [28, 29].

Identification of transients in noisy time series was proposed by [8, 9]. Reconstruction of a signal embedded in noise. The procedure relies on a bootstrap analysis of the statistical properties of the noise and a reconstruction algorithm from the multiscale zero-crossings of the wavelet transform of a signal. There are similarities with the generalized smoothing spline problem. The identification of the significant features of the wavelet transform is done by means of a bootstrap test on the size of a given extremum of the absolute value of the wavelet transform.

Wiener filter is another method in which the effect of observation noise is to produce a smoothed autoregressive (AR) spectral estimate. The practicality of the AR spectral estimation is very limited. The various methods for AR parameter estimation which were derived based on maximum likelihood principle are no longer MLEs when observation noise is present. Due to the difficulties of an MLE approach, the data is filtered with Wiener filter (which is a suboptimum estimator) to enhance the signal from the noise. It can be shown that the Wiener filter is an implicit part of the MLE.

4 Models of the noisy signal

The simplest experimentally measured signal switches between two states of the current, corresponding to open and closed states of the channel. The open and closed durations are typically random and the probability distributions of the open and closed periods may vary from one experiment to the other. It is often assumed that these periods are determined by a multi-state Markov process, but that only several of these states are observed [25]-[27]. For the purpose of testing our procedure, we adopt the simplest model of randomness, assuming a Markovian signal.

The measurements are typically very noisy, containing several components of noise. Depending on the instrumentation, different components of the noise may be dominant. While often white noise is the dominant component, in some cases other components, such as f noise, or even f^2 become dominant [28, 29]. We examine the performance of our algorithm on models with white noise and with filtered f^2 noise.

The signal. The random signal, x_t, is a telegraph process with exponential waiting times for switching. The signal switches between two levels, a and b, say. It stays at level a for an exponentially distributed random time with rate λ_a and then switches to b, where it stays with rate λ_b. The mathematical description of this signal is

$$\begin{aligned} x_{t+\Delta t} &= x_t && \text{with probability } 1-\lambda_{x_t}\Delta t \\ x_{t+\Delta t} &= a+b-x_t && \text{with probability } \lambda_{x_t}\Delta t. \end{aligned}$$

The noise. The noise intensity often varies between closed and open channels. We therefore add different noises to the two level signal. In the white noise model we add to each value of the telegraph signal a zero mean Gaussian variable, independently of one another, so that

$$\mathrm{Var}\{x_t\} = \frac{\sigma_{x_t}^2}{\Delta t},$$

with $\sigma_b^2 \approx 1.2\sigma_a^2$.

In the f^2 noise case the signal is sampled after low pass filtering through an 8-pole Bessel filter. The resulting noise is approximately Gaussian, as above, but the values of the noise are no longer independent, but rather correlated with correlation induced by the transfer function of the filter [29]. It should be noted that the filtering rounds the corners of the signal, spreading them on about 1.7 sample points.

5 Training for Level Estimates - Preprocessing Step

The current wavelet algorithm for detection, described in section 6, works best if they are given some *a priori* estimates of channel amplitude. A training algorithm (as a preprocessing step) is needed to generate a decent first guess from the raw data. This guess is then used by the detection algorithm. The training is performed once on initial raw data for the the complete duration of the detection process provided that we operate on single ionic channel.

The training algorithms work on a raw data giving estimates of the two levels (namely of open and closed channels) and their spread. The numerical suitability and stability of the algorithm is studied since it involves the solution of non-linear equations.

5.1 The model

The values are sampled from a population containing a mixture of two distributions, $A = (a_i)$ and $B = (b_i)$ given by $N(a, \sigma_A)$ and $N(b, \sigma_B)$ respectively. The distributions are independent of each other. The proportion of A is P_A and the proportion of B is P_B, such that $P_A + P_B = 1$. It follows that the pdf of a randomly drawn sample from the mixture is given by:

$$p(x=\alpha) = \frac{P_A}{\sqrt{2\pi}\sigma_A} e^{-(\alpha-a)^2/2\sigma_A^2} + \frac{P_B}{\sqrt{2\pi}\sigma_B} e^{-(\alpha-b)^2/2\sigma_B^2}. \tag{1}$$

Hence

$$\bar{x^0} = P_A + P_B$$

$$\bar{x^1} = P_A a + P_B b \tag{2}$$

$$\bar{x^2} = P_A(a^2 + \sigma_A^2) + P_B(b^2 + \sigma_B^2) \tag{3}$$

$$\bar{x^3} = P_A(a^3 + 3a\sigma_A^2) + P_B(b^3 + 3b\sigma_B^2) \tag{4}$$

$$\bar{x^4} = P_A(a^4 + 6a^2\sigma_A^2 + 3\sigma_A^4) + P_B(b^4 + 6b^2\sigma_B^2 + 3\sigma_B^4) \tag{5}$$

$$\bar{x^5} = P_A(a^5 + 10a^3\sigma_A^2 + 15a\sigma_A^4) + P_B(b^5 + 10b^3\sigma_B^2 + 15b\sigma_B^4) \tag{6}$$

(7)

and so on. To estimate the 6 parameters, $a, b, \sigma_A, \sigma_B, P_A, P_B$, one could in principle calculate the first 6 moments as explained below and solve the nonlinear system of 6 algebraic equations (2)-(7). The moment $\bar{x^j}$ is computed from

$$\bar{x^j} = \frac{\sum_{i=1}^{n} x_i^j}{n}.$$

5.2 Solution of the system by reduction to a single polynomial equation

The solution of systems of nonlinear equations is generally a non-trivial undertaking. The difficulty of such a solution by numerical means increases considerably with the number of equations. Even two non-linear equations could give considerable trouble. Therefore, it is advisable to devote much effort to reduction of the number of equations by elimination and simplification.

The first step is to eliminate the parameter $\bar{x}$ by shifting all values of measurements by that amount. After this step we work effectively in a zero mean system. Then we note that there is a symmetry in the form of the equations under an exchange of the variables corresponding to the two distributions.

We would like to preserve this symmetry by eliminating variables only in a symmetrical way. Since the variables P_A and P_B appear linearly in the first two equations and the σ variables do not appear there, we can eliminate P_A and P_B from all equations affecting only the powers of a and b.

Now we can eliminate σ_A^2 and σ_B^2 which also appear linearly in the third and fourth equation.

The two resulting equations are polynomial in a and b, we would like to eliminate one variable which cannot be done easily because of the high degree. We would also like to keep the symmetry, but this is impossible in terms of a or b. Pursuing the symmetry idea we look for the simplest symmetrical combinations of a and b which are $c = a + b$ and $e = ab$.

Eliminating a and b in terms of c and e we make the crucial observation that the first equation happens to be of second degree in c. This can be interpreted as a relation between c^2 and c so that we can reduce the remaining equation to a first degree in c. Now we can solve for c and substitute in the first equation to get one polynomial equation for e. By choosing $a > b$ we eliminate a trivial multiplicity of solutions. Since the mixture is now of zero mean we must have $a > 0$ and $b < 0$ so that only negative e's are acceptable. Positivity of the probabilities and standard deviations further restricts e to be $e > -3/2$.

Table 1 presents the results from applying the training algorithm on the raw signal described in figure 1. The signal has SNR=1.45 and its channel amplitude is 1.0. From 6000 samples we get a very good estimation. Even the estimated amplitudes from a small sample size such as 2000 and 4000 samples is more than sufficient for our initial estimate of the two levels (namely of open and closed channels) and their respective probability

No. of Samples	Amplitude A	Amplitude B	Estimated amplitude $A + \|B\|$	P_A	P_B
2000	0.64	-0.42	1.06	0.39	0.61
4000	0.66	-0.42	1.08	0.39	0.61
6000	0.63	-0.40	1.03	0.39	0.61
8000	0.61	-0.40	1.01	0.39	0.61
10000	0.61	-0.40	1.01	0.39	0.61

Table 1: Training results

6 The Detection Algorithm

The detection algorithm consists of two steps: smoothing the raw signal and then applying the wavelet transform and its multiscale analysis across the scales.

6.1 Smoothing

Before we apply the wavelet transform and its multiscale analysis to detect channel openings, the algorithm smoothes the raw signal. The smoothing is basically a denoising procedure.

We perform several cycles of Gaussian wavelet smoothing by applying the low-pass part of the wavelet transform. In other words, we apply the wavelet transform on the raw signal on several levels according to equation 8. The number of levels is predetermined in advance (it is an input parameter to the algorithm), and usually 3 levels is sufficient. Then, the low-pass part of the signal is the smoothed part of the signal that is being passed to the actual detection algorithm.

Figure 2 is the smoothed version of the signal given in figure 1. These were obtained after applying the wavelet transform across three levels. The low-passed part of this decomposition is the input to the detection algorithm.

The smoothing procedure (and other factors) misses short events containing 4 to 7 sample points. Perhaps improved procedure such as eliminating the smoothing procedure in the beginning will do better.

6.2 Detection

6.2.1 Overview

We decompose the signal by applying the wavelet transform on the smoothed data. Points of sharp variations are often among the most important features for analyzing the properties of transient signals. They are generally located at the boundaries of important signal structures. Sharp transients in a signal can be detected by multiscale edge detection. The scale defines the size of the neighborhood where the signal changes are computed. The wavelet transform is related to multiscale edge detection. The well kncwn Canny edge detector is equivalent to finding the local maxima of a wavelet transform modulus [12]-[14]. Our detection algorithm is based on detected edge that its amplitude survives multiscale decomposition.

There are many different types of sharp variations points in signals. To identify more precisely an edge that has been detected, it is necessary to analyze its local properties. Singularities of this type are usually characterized by their Lipschitz exponents. Wavelet theory proves that these Lipschitz exponents can be computed from the evolution across scales of the wavelet transform modulus maxima.

We derive a numerical procedure to measure these exponents across the scales. And if the edge is smooth, we can also estimate how smooth it is from the decay of the wavelet transform maxima across scales.

If we compute along a continuum of scales, the multiscale edges detected by the wavelet transform modulus maxima, defines curves in the scale-space. Within each curve, the points have similar gradient. A point propagates from scale 2^j to scale 2^{j+1}, if and only if, they are connected by a point curve in the 1D scale space. At each scale 2^j, we find which point propagates to a point at coarser scale 2^{j+1}.

This procedure also removes small signal fluctuations and we thus only detect the sharp variations of large structures. This can help in denoising the signal.

We compute across scales the multiscale edges detected by the wavelet transform modulus maxima. We detect the points that propagates from scale 2^j to scale 2^{j+1}.

If this propagation lasts over number of levels and has a certain maximal amplitude that is above some predetermined threshold (derived from training) then at this point we have an edge (singularity).

The detection of opening and closing of events is equivalent to locate edges in the above sense.

6.2.2 The algorithm

The edge detection is performed on the smoothed data. The signal is decomposed by a set of wavelet filters that were chosen in advance.

Assume that g_i and h_i, $i = 0, \ldots, R$ (length of the filter) are the low-pass and high-pass filters respectively. They consistute a complete subband coding scheme.

The regular application of the wavelet transform and its multiscale decomposition is performed in the following way:

$$\begin{aligned} L_j^l &= \sum_{i=0}^{R} g_i L_j^{l-1}(i+2j) \\ H_j^l &= \sum_{i=0}^{R} h_i L_j^{l-1}(i+2j) \end{aligned} \qquad (8)$$

where $l = 1, \ldots, L$ is the number of levels in its multiresolution decomposition, L_j^0 is the original signal in level 1 (before it decomposed), R is the range of the filters, L_j^l and H_j^l are the low-pass and high-pass results for point j, respectively, that are parts of the signal that were derived from the low-pass part of the signal in level $l-1$.

We do not actually perform "bimation" by factor of 2 (bimation of one sample out of two) when we progress from fine to coarse level as appeared in the definition of the wavelet $\psi_{jk}(x) = 2^{-\frac{j}{2}}\psi(2^{-j}x-k)$, $j,k \in Z$ (or in equation 8). Instead, in each level we insert zeros between the filter elements. The number of inserted zeros is proportional to the current level. We skip the decimation since the signal is analyzed by comparing the high-pass amplitudes (in H_j^l) across several scales of the wavelet decomposition. We compare features across scales that are in the same physical location in each scale. Therefore, decimation is eliminated in order that the comparison among the same locations across the high-pass decomposition in different levels of resolution can be done.

The filters that are being used in this application are symmetric and of short length. The symmetry of the filter is important but not crucial. It can help in chaining similar features across scales (explained later) without worrying about the shift of the maximum of the related wavelet coefficients across scales. The shift in the location of a feature across scale is the result of repeated application of the convolution in equation 8. Therefore, symmetric filer can reduce the shift problem across scales. We can use in equation 8 also non-symmetric filters such as Daubechies or Coiflets[4]. Then, the shift across scales should be dealt with extra care and can cause problems if the level of decomposition is increased.

Short filters are needed since when we go further in the multiscale decomposition in reaching coarser levels the length of the convolution as it is reflected in equation 8 is becoming wider and the influence of remote areas (when compare to the current convolved signal point) are considered. Then, the local computation of the feature across scales is influenced by irrelevant remote features. If there are two edges close to each other (opening and closing of channels), then, by using equation 8 we convolve them together while we progress in the depth of the decomposition and they can cancel or weaken each other. Then, the maximal amplitude of the wavelet coefficient will be significantly less than the predetermined threshold and the suspected feature will be rejected as a significant transition (as explained later).

6.2.3 The decision for the existence of channeling opening or closing

If computed along a continuum of scales, the extrema of the wavelet transform define curves in the scale-space. We say that an extrema propagates from a scale 2^j to a scale 2^{j+1}, if and only if, they are connected by an extrema curve in the scale space. At each scale 2^j we find which extrema propagates to an extrema at the coarser scale 2^{j+1}. From the decomposition across scales, at each scale 2^j, we detect the modulus maxima by finding the points where the discrete wavelet transform is larger than its two closest neighbor values and strictly larger than at least one of them. (Maximum is either a positive maximum or negative minimum of the wavelet transform.) This propagation algorithm computes the coordinates of the extrema of each high-passed detail of the wavelet transform in each level of the decomposition. We first locate the extremum points in each detail. Afterwards, each time an extremum point is found in a certain detail, a second order interpolation is performed to find the "exact" position of the extremum (and its "exact" value). The chaining propagates the extrema along the scales.

Figure 1 is a raw signal of 2,000 samples. Figure 2 is the smoothed version of the raw signal by applying on it three times the wavelet transform and taking its low-passed part.

The decomposition of the low-passed signal along seven levels is described in figure 3. We concentrate now on figure 3 to explain in detail its structure. There are groups of lines separated by a blank line. Each group is numbered where the number appears on the right handside in parenthesis. Each group represents an event that has been successfully chained across scales and so is a candidate to be classified as a transition. Each line in the group represents the complete information of the current level in the multiscale decomposition. The first column indicates the level. The second column indicates the exact location of the chained feature. Column three gives the amplitude of the measured feature in that location. Column four (A) is calculated by equation 11 and the ratio (in column five) is defined by equation 10. Column four (A), where its value is based upon column three (amp) and the ratio (column five), determines (as explained later) whether the group represents an edge that can be classified as an opening or closing of a channel.

For example, group (1) indicates that locations 6000,6001,6000,5999,5990,5977 are chained according to the algorithm described above. In group (3) there is only one level. In groups (4) and (5) there are only four levels. Probably we can not get the rest three levels in groups (4) and (5) because beyond level 4 the detected maximum disappeared or it was impossible to chain him to anything or they merged to other features. In figure 3 there are 17 groups in locations 6000-6381 which can be chained and therefore we classify them as "potential" or "suspected" edge that indicates a channel activity of opening or closing. More analysis is needed to determine whether each group represents an edge.

The list in figure 3 is incomplete in the sense that there are more groups that were chained and are not presented here in order to shorten the presentation.

So far, we decomposed the signal in seven levels and the related singularities in each high-passed part was chained. The suspected edges are the chained features for which the amplitudes survive to a substantial depth, i.e. number of levels. Points of maximum amplitude are generally located at the boundaries of important structures. The scale of the level defines the size of the neighborhood which is influenced by the signal changes. Thus the multiscale edge detection is related to the wavelet transform since it is equivalent to finding the local maxima of the wavelet transform modulus.

We analyze each one of the suspected group (as explained above) and decide whether or not it represents an edge. If the maximal amplitude of the wavelet coefficient in the chain will be significantly less than the predetermined threshold then the suspected feature will be rejected to be classified as a significant transition. In the following this idea is explained accurately.

In the following the superscript ref means reference which is the decomposition of an ideal step function (no noise, single level) with amplitudes 0.0 and 1.0. Tables 2 and 3 describe the results of this reference function along the multiscale decomposition. The entries in these tables are the reference number which determine for the smoothed raw data whether the inspected features is a "real" edge indicating channel opening or closing.

Level	Amplitude a_i^{ref}	Ratio between amplitudes r_i^{ref}
1	0.0469	0.047
2	0.0928	1.979
3	0.1777	1.916
4	0.3003	1.690
5	0.3989	1.328
6	0.4495	1.127
7	0.4747	1.056
8	0.4874	1.027
9	0.4937	1.013
10	0.4968	1.006

Table 2: The multiscale results from decomposition of a step ("clean synthetic") function with amplitude 1.0 along 10 levels after 3 levels of smoothing

Level	Amplitude a_i^{ref}	Ratio between amplitudes r_i^{ref}
1	0.0059	0.006
2	0.0117	2.000
3	0.0234	1.999
4	0.0467	1.995
5	0.0925	1.979
6	0.1771	1.915
7	0.2995	1.691
8	0.3984	1.330
9	0.4492	1.127
10	0.4746	1.057

Table 3: The multiscale results from decomposition of a step ("clean synthetic") function with amplitude 1.0 along 10 levels after 6 levels of smoothing

Let

$$r_i^{ref} = \frac{a_i^{ref}}{a_{i-1}^{ref}} \quad i = 0, \ldots, L \quad a_0^{ref} = 1.0 \tag{9}$$

Let a_i^m denote the measured amplitude along the levels of decomposition in each group.

$$r_i^m = \frac{a_i^m}{a_{i-1}^m} \quad i = 0, \ldots, L \tag{10}$$

where L is the number of levels in the decomposition.
For $i = 1, \ldots, L$ compute

$$A_i = \begin{cases} \frac{a_i^m}{a_i^{ref}} & |r_i^m - r_i^{ref}| < r_i^{ref} \cdot 0.2 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Let $A_{max} = \max_{4 \le i \le L} A_i$. The amplitude A_{max} is computed only from level 4 on because from there on we can see whether the amplitude is strong enough to survive even if far away samples are affecting the strength of the amplitude.

If $L \geq 5$ (there are at least five levels in the decomposition) and $|A_{max}| \geq 0.70 \times T$ then the current chained group is an opening or closing of a channel (depending on the sign of the measured amplitude). Otherwise we reject this feature.

We demonstrate the acceptance/rejection idea on figure 3. From the training process (see table 1) we have that $T = 1.0$. In group (1) $A_{max} = 0.1635$ in level=5 (we start from level 4). Since $|A_{max}| < 0.7$, then group (1) is rejected. We can also see that the amplitudes represented by A are decreasing and disappearing since they do not survive the decomposition. This group represents noise.
Group (2) is rejected because of the same reason. We can here that the $A = 0$ which mean that the amplitude did not survive through the decomposition. Group (3) does not have at least five levels.
Group (9) has strong A's along the decomposition. $A_{max} = 0.8270$ at level 5. This is an edge at location 6198 (determined by the location of the maximal A). The correct location is 6201. The sign of A indicates that this an opening.
Groups (10),(11),(12),(13) are rejected because their chaining do not survive at least five levels.
Group (15) has strong A's along the decomposition. $A_{max} = 0.9988$ at level 5. This is an edge at location 6361. This is the correct location. The sign of A indicates that this is a closing.
The correct location is 6422.

The smoothing procedure and the interference between close edges, that were discussed above, prevent the detection of short events containing 4 to 7 sample points. (These events might be openings of 40 to 70 μsec in typical experimental situations.)

6.2.4 Threshold verification

We use the threshold method as an additional tool to verify the validity of the detection results derived in the previous section (section 6.2.3) using wavelet and multiscale decomposition.

We take as an input the smoothed raw data (it is the same input as was used by the wavelet detection). Instead of checking the strength of the amplitudes across the the multiresolution decomposition we check whether each amplitude of the smoothed signal is greater or equal half of the amplitude computed by the training procedure. In such a case, we have an edge, otherwise it is ignored.

This procedure can work independently well in certain cases for single ionic channels [28, 29]. But it is limited to single channel, while the proposed algorithm is more general in the sense that it can be extended to two and more ionic channels (see section 8).

7 Software

A robust procedure and a stable user-friendly software was developed to implement the wavelet analysis with its multiresolution decomposition to detect and analyze currents through ionic channels.

Training, smoothing, multiresolution decomposition, chaining of similar features, edge detection and threshold verification were implemented in a software package that runs on either under DOS or UNIX.

8 Future Research

The key part of the wavelet algorithm smoothes the raw signal and then applies multiscale analysis to detect openings. A number of ways to do this are possible. Here we show the feasibility of Gaussian smoothing and edge detection previously developed for other purposes. One of the main areas of future work will be the choice of the correct basis and the best method to use that basis to detect and locate opening and closing events. Better smoothing and detection should be possible, for example, using a Haar (rectangular function) basis.

Our goal is to detect and analyze short events containing between 4 to 7 sample points. (These events might be openings of 40 to 70 μsec in typical experimental situations.) Some possible approaches are: one to oversample in general, hoping that the steep nonlinear dependence of wavelet resolution on number of samples will overcome the loss in independence of oversampled events. The other is to locally increase the resolution of analysis just at suspected locations of transitions.

Another possibility is to eliminate the smoothing procedure in the beginning and instead go further in the depth of the multiscale analysis. To compensate for the skipping of the smoothing procedure we will have to modify the rejection algorithms which is equivalent to the denoising step mentioned above. Our modified algorithm which will skip training and smoothing phases and will be reported in a subsequent paper. Thus, we plan to introduce in this step instead a more sophisticated "denoising" algorithms which is not based on predetermined parameters. Such procedures have been proven to work better than any linear filtering approach.

The multiscale edge detection method that was introduced in this paper can be extended to detect two and more channels. In a later stage, we plan to study the decay of the noise across the scales of a multiresolution analysis. The multichannel situation can be treated with modifications and extensions of the algorithms developed for single channels.

[1] Meyer, I., Wavelets and Operators, Cambridge, University Press, 1993.

[2] Meyer, I., Wavelets: Algorithms and Applications, SIAM, 1993.

[3] Daubechies, I., Orthonormal Bases of Compactly Supported Wavelets, Comm. on Pure and Applied Mathematics, Vol. XLI, pp. 909-996, 1988.

[4] Daubechies, I., Ten Lectures on Wavelets, SIAM, 1992.

[5] Coifman, R., Meyer, Y., Quake, S., Wickerhauser, V., Signal Processing Compression with Wavelet Packets, Progress in Wavelet Analysis and Application, Proceedings of the International Conference "Wavelet and Applications", Touluse, France,pp. 77-94, June 1992, edited by Meyer and Roques,Editions Frontiers.

[6] Coifman, R., Majid, F., Adapted waveform analysis and denoising, Progress in Wavelet Analysis and Application, Proceedings of the International Conference "Wavelet and Applications", Touluse, France,pp. 63-76, June 1992, edited by Meyer and Roques,Editions Frontiers.

[7] Donoho, D.L., Wavelet shrinkage and W.V.D.: a 10-minute tour, Progress in Wavelet Analysis and Application, Proceedings of the International Conference "Wavelet and Applications", Touluse, France,pp. 109-128, June 1992, edited by Meyer and Roques,Editions Frontiers.

[8] Carmona, R.A., Wavelet Detection of Transients in Noisy Time Series, to appear.

[9] Carmona, R.A., Wavelet Identification of Transients in Noisy Time Series, SPIE-Mathematical Imaging: Wavelet Applications in SIgnal and Image Processing, Vol. 2034, pp. 392-400,15-16 July, 1993, San Diego, USA.

[10] Mallat, S.G., A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, IEEE Trans. on PAMI, Vol. 11, No. 7, pp. 674-693, July 1989.

[11] Mallat, S.G., Review of Multifrquency Channel Decompositions of Images and Wavelet Models, Technical Report No. 412, Robotics Report No. 178, NYU, Courant Institute, November 1988.

[12] Mallat, S.G., Multiresolution Approximations and Wavelet Orthonormal Bases for $L^2(\Re)$, Trans. of AMS, Sept. 1989.

[13] Mallat, S.G., Zhong, S., Characterization of Signals from Multiscale Edges, IEEE Trans. on PAMI, Vol. 14, No. 7, pp. 710-732, July 1992.

[14] Mallat, S., Hwang, L., Singularity Detection and Processing with Wavelets, IEEE Trans. on Information Theory, Vol. 38, No. 2, pp. 617-643, March 1992.

[15] Froment, J., Mallat, S., Second Generation Compact Image Coding with Wavelets, Wavelets-A Tutorial in Theory and Applications, C.K. Chui (ed.), pp. 655-678, Academic Press, 1992.

[16] Wavelets: Time Frequency Methods and Phase Space, Proceedings of the International Conference, France, Dec. 14-18, 1987, Springer-Verlag 1989.

[17] Wavelets and Applications, Proceedings of the International Conference, Marseille,France, May, Springer-Verlag 1989, Edited by Y. Meyer.

[18] Progress in Wavelet Analysis and Application, Proceedings of the International Conference "Wavelet and Applications", Touluse, France, June 1992, edited by Meyer and Roques,Editions Frontiers.

[19] SPIE's 1993 International Symposium on Optics, Imaging and Instrumentation, San Diego, July, 1993.

[20] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM algorithm, J. Royal Statistical Society, B39, pp. 1-38, 1977.

[21] R.J.A. Little, D.B. Rubin, Statistical Analysis with Missing Data, John Wiley, 1987.

[22] A. Averbuch, R. Eisenberg, M. Israeli, Z. Schuss, Detecting ionic currents in single channels using wavelet analysis, Part I: zero mean Gaussian noise, Conference on "Mathematical Imaging: Wavelet applications in Signal and Image processing",SPIE coneference, July 27-29, San Diego, California, 1994.

[23] A. Averbuch, R. Eisenberg, M. Israeli, Z. Schuss, Detecting ionic currents in single channels using wavelet analysis, The 12th International Conference on Pattern Recognition, Jerusalem, Israel, Oct 9-13, 1994.

[24] B. Hille, Ionic Channels of Excitable Membranes, Sinauer, Sunderland MA, 1992.

[25] K.L. Magleby, B.S. Pallotta, Calcium dependence of open and shut interval distributions from calcuim-activated potassium channels in cultured rat muscle, J. Physiol Vol. 344, pp 585-604, 1983

[26] E. Mogzydlowski, R. Latorre, Gating kinetics of Ca+2 - activated K+ channels from rat muscle incorporated into planar lipid bilayers, J. Gen. Physiol. vol 82, pp 511-542, 1983.

[27] S.M. Sine, T. Claudio, F.J. Sigworth, Activation of Torpedo Acetylcholine receptors expressed in mouse fibroblasts single channle current kinetics reveal disticnt agonist binding affinities, J. Gen Physiol. vol 96, pp 395-437, 1990.

[28] D. Colquhoun, F.J. Sigworth, Fitting and Statistical Analysis of Single-Channel Records, in Single-Channel Recording, pp. 191-258, Chapter 11, edited by B. Sakmann and E. Neher, Plenum Press, 1983.

[29] R.A. Levis, Current Noise in Recording of Currents through Single Channels, private communication.