

Applying Hidden Markov Models to the Analysis of Single Ion Channel Activity

L. Venkataramanan* and F. J. Sigworth†

*Schlumberger-Doll Research, Ridgefield, Connecticut 06877 USA; and †Department of Cellular and Molecular Physiology, Yale University, New Haven, Connecticut 06520 USA

ABSTRACT Hidden Markov models have recently been used to model single ion channel currents as recorded with the patch clamp technique from cell membranes. The estimation of hidden Markov models parameters using the forward-backward and Baum-Welch algorithms can be performed at signal to noise ratios that are too low for conventional single channel kinetic analysis; however, the application of these algorithms relies on the assumptions that the background noise be white and that the underlying state transitions occur at discrete times. To address these issues, we present an “*H*-noise” algorithm that accounts for correlated background noise and the randomness of sampling relative to transitions. We also discuss three issues that arise in the practical application of the algorithm in analyzing single channel data. First, we describe a digital inverse filter that removes the effects of the analog antialiasing filter and yields a sharp frequency roll-off. This enhances the performance while reducing the computational intensity of the algorithm. Second, the data may be contaminated with baseline drifts or deterministic interferences such as 60-Hz pickup. We propose an extension of previous results to consider baseline drift. Finally, we describe the extension of the algorithm to multiple data sets.

INTRODUCTION

Recordings of single ion-channel currents provide a wealth of information about the activity of single allosteric protein molecules. The open-closed behavior of ion channels has generally been described in terms of continuous-time Markov models (Colquhoun and Hawkes, 1995) in which model states are taken to correspond to distinct states of protein conformation or ligand binding. Finding the best Markov model description of a channel's behavior is therefore taken to be equivalent to a complete elucidation of the kinetic behavior of the channel protein with the Markov transition probabilities corresponding directly to rate constants of ligand binding and unbinding and of conformational changes.

Finding the best Markov model involves two steps. First, the general topology of the model must be chosen, specifying the number of states and the connectivity that specify the allowable transitions among states. The second step is the optimization of Markov model parameters. Given a Markov model λ with N states, the parameters are the current levels μ_i corresponding to each state q_i , $i = 1, \dots, N$, the initial state probability π_i , and the transition rates contained in an $(N \times N)$ matrix \mathbf{Q} . Maximum-likelihood techniques are typically used to optimize these parameters, and the likelihood-ratio test is commonly used to identify the best model topologies.

Several methods have been used to compute likelihoods and estimate model parameters from single-channel data. Most commonly, threshold detection is used to identify channel-open and channel-closed intervals; the distributions of these dwell times are then fitted to the predictions of Markov models by maximum-likelihood techniques (Magleby and Weiss, 1990; Colquhoun and Sigworth, 1995). Alternatively, the likelihoods of models are computed on the basis of the entire sequence of open and closed dwell times (Horn and Lange, 1983; Ball and Sansom, 1989; Qin et al., 1997). An improved approach to the identification of open and closed intervals has been introduced through the use of the Viterbi algorithm (Fredkin and Rice, 1992a). Discussed in the present paper is an approach that does not require the identification of open and closed intervals at all but makes use of the raw single-channel recording in the form of sampled time course of membrane current. This application of signal processing based on hidden Markov models (HMMs) has already been demonstrated to be particularly useful in characterizing channel behavior when the signal-to-noise ratios are low and when multiple subconductance levels exist (Chung et al., 1990; Fredkin and Rice, 1992b; Chung and Gage, 1998). An excellent overview of the HMM approach to single channel analysis is given by Qin et al. (2000a,b). Various implementations of the HMM algorithms have been applied to experimental single channel data (Becker et al., 1994; Milburn et al., 1995; Michalek et al., 1999; Farokhi et al., 2000). The algorithms described here have been used in the recent studies by Sunderman and Zagotta (1999a,b), Wang et al. (2000), and Zheng et al. (2001).

In the HMM framework, the discrete-time data $Y(t)$ are assumed to be the sum of Gaussian noise $n(t)$ and the “noiseless” channel current as the channel makes transitions from one state to another. The Markov model λ underlying

Submitted August 22, 2001, and accepted for publication December 12, 2001.

Address reprint requests to F. J. Sigworth, Department of Cellular and Molecular Physiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520-8026. Tel.: 203-785-5773; Fax: 203-785-4951; E-mail: fred.sigworth@yale.edu.

© 2002 by the Biophysical Society

0006-3495/02/04/1930/13 \$2.00

the channel activity is considered to be “hidden” because the channel state at any given time is not directly observable. The channel state may be hidden due to the aggregated nature of the HMM (several states may have the same conductance level) and is also hidden by the noise. To mirror the discrete-time nature of the sampled recording, a discrete-time HMM replaces the usual continuous-time Markov model. The parameters of the discrete-time model are the transition probabilities, contained in a $(N \times N)$ matrix A ; the current levels μ_i corresponding to each state q_i , $i = 1, \dots, N$; and the initial state probability π_i .

The discrete-time transition probability matrix $A = \{a_{ij}, i, j = 1, \dots, N\}$ is related to the matrix of rate constants Q of the usual, continuous Markov model by

$$A = e^{Q\Delta t}, \quad (1)$$

in which Δt is the sampling interval; thus a_{ij} is the probability of making a transition from state q_i to state q_j during one sample interval. The objective of our analysis is to estimate the parameters of the discrete HMM $\lambda = \{A, \mu, \pi\}$ from the observed discrete-time data $Y(t)$ while realistically modeling the signal and noise.

Likelihood is defined as the probability of observing a particular data set given a model. The evaluation of the likelihood of HMMs has been made practical by an algorithm called the forward-backward procedure. Optimization of the parameters of the model is aided by the Baum-Welch procedure, which through iteration causes a maximum of the likelihood to be approached. These algorithms have been well described in the speech-recognition literature (Baum et al., 1970; Liporace, 1982; Rabiner, 1989).

The applicability of the HMM algorithms to patch-clamp recordings have been limited by several problems (Qin et al., 2000b). First, the traditional forward-backward procedure is highly efficient but is applicable only if the additive noise is white. In practice, the experimental background noise is correlated from sample to sample and has a spectral density that increases with frequency. Second, the data are usually low-pass filtered by an analog anti-aliasing filter such as a Bessel filter before digitization; this produces intermediate values for the sampled signal in the vicinity of state transitions.

Qin et al. (2000b) have addressed these issues by modeling the background noise by an autoregressive (AR) process, so that the data can be reduced to a higher-order (metastate) Markov process with white noise. Their model allows for different noise models for states of different conductances. Under the strong assumption that the noise in each metastate depends only on the current state, the analytical derivatives of the likelihood function with respect to all unknown model parameters are derived. This optimization approach has the advantage that it optimizes for the rate constants directly and thus easily allows constraints on rate constants such as detailed balance.

Our approach in analyzing the single channel data differs in several respects. First, we follow the traditional Baum-Welch approach to estimate the Markov model parameters. The algorithm is slow at low signal to noise ratios (SNR) but has the advantage that it allows a large number of parameters to be reestimated simultaneously, including for example vertices of a fit to a drifting baseline. Our focus is on the estimation of the transition probability matrix A , which can be considered to be equivalent to estimating the rate matrix Q when the sampling interval is small compared with any of the rate constants in Q . Second, we also model the noise as an AR process but make use of a more general description of state-dependent noise. Last, we model the randomness of sampling relative to transitions in the continuous time Markov process by approximating it by an additive noise source called the “ H -noise.”

In this paper we also address three issues that arise in the practical application of the algorithm in analyzing single channel data. First, the antialiasing filter used during data acquisition typically has a gradual roll-off in the transition region of the frequency response. However, for better performance and to reduce computational complexity of the algorithms, we prefer the use of an antialiasing filter with a sharp roll-off in the frequency domain. Therefore, we describe the implementation of an inverse filter to remove the effects of the gradual roll-off filter and in effect replace it with a digital filter with a sharp roll-off in the transition region. Second, the data may be contaminated with baseline drifts or deterministic interferences (such as 60-Hz pickup). We propose an extension of the results presented in Chung et al. (1990) to model and compensate for baseline drift. Finally, the traditional HMM algorithms focus on parameter estimation from a single data set. Often, in practice, multiple sets of data are recorded under the same experimental conditions. We describe the extension of the algorithm to multiple observation sequences.

This paper is organized as follows. The second section briefly describes the computation of likelihood and estimation of HMM parameters through use of the standard algorithms. The third section presents the noise model and highlights the main features of the H -noise algorithm that model the noise correlation in successive filtered, sampled data points. The fourth section describes the inverse filter and estimation of baseline from the measured single channel data. We conclude with a discussion of the approach in the fifth section.

COMPUTATION OF THE LIKELIHOOD AND ITS MAXIMIZATION

The likelihood of the hidden Markov model λ is defined as the probability of the observed data samples $Y(t) = \{y_t, t = 1, \dots, T\}$ given the model,

$$L = P(y_1, y_2, \dots, y_T | \lambda). \quad (2)$$

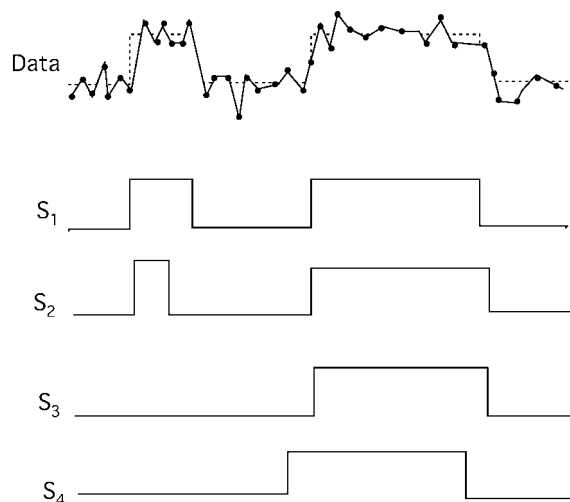


FIGURE 1 Illustration of the likelihood calculation. The top trace shows simulated data from a two-state Markov model. The underlying state sequence is shown in dashed lines. S_1 , S_2 , S_3 , and S_4 are four of the many possible state sequences that could have given rise to the observed data. The likelihood L of the HMM is obtained by summing the joint probability of the data and the state sequence over all possible state sequences.

The brute-force numerical computation of the likelihood is intensive as illustrated in the following example. The top trace in Fig. 1 shows noisy simulated data generated from a two-state Markov model and some of the several possible state sequences S_i that could have given rise to the observed data. The likelihood L of the HMM can be calculated by summing the probability of the data and the state sequence over all possible state sequences,

$$\begin{aligned} L &= P(\text{data}|\text{model}) \\ &= \sum_i P(\text{data}, S_i|\text{model}) \\ &= \sum_i P(\text{data}|S_i)P(S_i|\text{model}) \end{aligned} \quad (3)$$

In general, let s_t denote the state of the HMM at time t , and let $S = \{s_1, s_2, \dots, s_T\}$ denote a possible state sequence. The probability of this state sequence given the model λ can be computed in terms of the initial state probability π and the transition probabilities between states $a_{s_{t-1}s_t}$ as,

$$P(S|\lambda) = \pi_{s_1} \prod_{t=2}^T a_{s_{t-1}s_t} \quad (4)$$

Given a state sequence, the probability of the observed data is obtained as,

$$P(y_1, y_2, \dots, y_T|S) = \prod_{t=1}^T P(y_t|s_t) \quad (5)$$

in which $P(y_t|s_t)$ is the probability of obtaining the observation y_t given that the channel is in state s_t . This probability is considered to be independent of observations at previous

times; this is the uncorrelated or “white” noise assumption. Let us define a term called the emission probability $b_i(y)$ as the probability of observing datum y when channel is in state q_i ,

$$b_i(y)\delta y = P(y|q_i), \quad (6)$$

in which δy represents the resolution of measurement y . Assuming that δy is a constant, we can without loss of generality set $\delta y = 1$. In the case of Gaussian white noise with variance σ^2 the emission probability is a Gaussian function centered on the mean current μ_i ,

$$b_i(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(y - \mu_i)^2}{2\sigma^2}\right). \quad (7)$$

and the probability of the observations given the state sequence (Eq. 5) is seen to be the product of T Gaussians. Finally, the likelihood is obtained by summing the product of the probabilities specified in Eqs. 4 and 5 over all possible state sequences,

$$L = \sum_{\text{all } S} P(y_1, y_2, \dots, y_T|S)P(S|\lambda), \quad (8)$$

The evaluation of Eq. 8 is very computationally expensive, involving on the order of $(2TN^T)$ calculations.

Forward and backward variables

The forward procedure (Rabiner, 1989; Qin et al., 2000a) offers a vast improvement and can account for both white and correlated noise. The issue of correlated noise is addressed by introduction of the concept of a vector or metastate-HMM (Fredkin and Rice, 1992; Venkataramanan et al., 1998). If s_t is the state of the HMM at time t , then each k -tuple of successive states $(s_t, s_{t-1}, \dots, s_{t-k+1})$ forms a “metastate.” A metastate at time t is denoted by the notation $I = (q_{i_0}, q_{i_1}, \dots, q_{i_{k-1}})$ in which q_{i_j} is the state of the HMM at time $t - j$, $j = 0, \dots, k - 1$ and $i_j = 1, \dots, N$.

The forward variable at time t in metastate I is defined as the probability of the first t data measurements and being in metastate I at time t given that the model is λ , or

$$\alpha_t(i_0, i_1, \dots, i_{k-1}) = P(y_1, y_2, \dots, y_t, I|\lambda). \quad (9)$$

The forward variable at any time instant in any state can be computed efficiently in terms of the forward variable at the previous time instant and the emission probability. The forward variables at the final time instant can be used to compute the likelihood L of the HMM,

$$L = \sum_I \alpha_T(I) \quad (10)$$

In addition, the forward variable together with the backward variable defined below can be used in the estimation of the HMM parameters using the Baum-Welch procedure.

Let Y_t be the vector of k samples ending at time t . The backward variable at time instant t in metastate I is defined as the probability of the last $T - t$ data points given that the

model λ is in metastate I at time t and $(k - 1)$ previous data samples, or

$$\beta_t(i_0, i_1, \dots, i_{k-1}) = P(y_{t+1}, y_{t+2}, \dots, y_T | Y_t, I, \lambda). \quad (11)$$

Similar to the forward variable, it can be computed at each time step in terms of the backward variable at the next time step and the emission probability associated with the metastate at that time.

Reestimation

The Baum-Welch reestimation algorithm uses the forward and backward variables to update the model parameters. Let $\gamma_t(I)$ be the probability of being in metastate I at time t given all of the observed data and the model λ ,

$$\gamma_t(I) = P(I | y_1, y_2, \dots, y_T, \lambda). \quad (12)$$

It can be computed from the forward and backward variables as,

$$\gamma_t(I) = \frac{\alpha_t(I)\beta_t(I)}{L}, \quad t = 1, \dots, T. \quad (13)$$

Using the γ variables, new estimates can be made of the various model parameters. Let M_0 be the set of metastates where the state at time t is q_{i_0} and state at time $t - 1$ is q_{i_1} . Let M_1 be the set of metastates where the state at time $t - 1$ is q_{i_1} . The reestimated transition probability from state q_{i_1} to q_{i_0} can be understood as the ratio of the expected number of transitions from state q_{i_1} to q_{i_0} divided by the number of transitions from state q_{i_1} ,

$$\hat{a}_{i_1 i_0} = \frac{\sum_{t=1}^T \sum_{I \in M_0} \gamma_t(I)}{\sum_{t=1}^T \sum_{I \in M_1} \gamma_t(I)}. \quad (14)$$

The reestimated current levels can similarly be obtained by solving a linear system of equations (Venkataraman et al., 1998). The forward-backward procedure followed by the Baum-Welch reestimation formulae is a special case of the expectation maximization algorithm (Baum et al., 1970). Carried out iteratively, it is guaranteed to converge to a local maximum of the likelihood function. The algorithm converges slowly at low SNRs. In practical problems, the algorithm takes on the order of a hundred iterations to converge to the local maximum (Qin et al., 2000b; Zheng et al., 2001).

It is also possible to reconstruct the original, noiseless signal from the measured data in a number of ways. This reconstructed signal is model-dependent and is therefore not unbiased; however it gives the user useful feedback about the performance of the algorithm. In one approach, the Viterbi algorithm can be used to provide the most probable state sequence underlying the measured data (Rabiner, 1989). Alternatively, the cumulative mean or maximum a posteriori probability estimate of the signal can also be easily reconstructed from the a posteriori probabilities

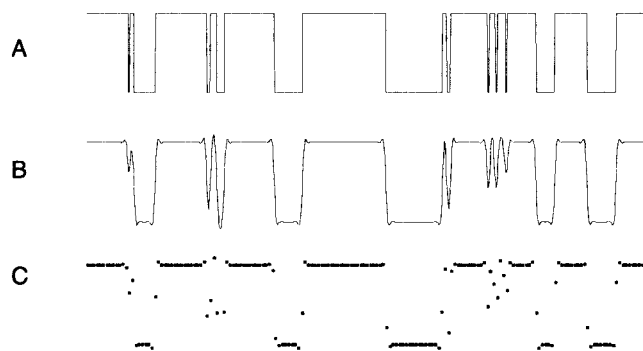


FIGURE 2 Effects of filtering and sampling continuous-time data. (A) Example of continuous-time data from a two-state HMM. (B) Data after passing through an antialiasing filter. (C) Filtered, sampled data. The impulse response of the filter that was used is given by Eq. 23 with the parameters $\sigma_f = 1$ and $f_c/f_s = 0.4$.

(Chung et al., 1990). The cumulative mean estimate gives the expectation value of the channel current based on the probabilities γ at each time point. This estimate is given by

$$\hat{x}_t = \sum_I \mu_{i_0} \gamma_t(I) \quad (15)$$

in which μ_{i_0} is the channel current in the final state of metastate I .

NOISE MODEL AND H-NOISE

The applicability of hidden Markov models to analyze single channel recordings is limited by two problems. First, the experimental background noise is correlated and has a spectral density that increases with frequency. Second, to remove high-frequency noise components, the data are necessarily filtered before sampling. The problem with filtering is illustrated with the following example. The top trace in Fig. 2 is the output of a continuous-time Markov model with two states—an open state and a closed state. The second trace is the signal after passing through an antialiasing filter. The last trace shows the filtered data after sampling. As seen in the last trace, when a channel closes from an open conductance level, there might be a sample taken on the falling edge of the filtered signal. Thus when the channel closes from an open state, the data will consist of a sequence of open current levels, one sample at a random intermediate level and a sequence of samples at the closed level. The problem is to distinguish the presence of the intermediate sample from a two-state channel from the case of a channel passing through a sublevel at say, one-half amplitude. We describe below the “H-noise” algorithm, which addresses the issues of filtering, correlated background noise, and the randomness of sampling relative to transitions.

Noise model for correlated background noise

The background current noise in patch-clamp recordings has a spectral density whose dependence on frequency f is

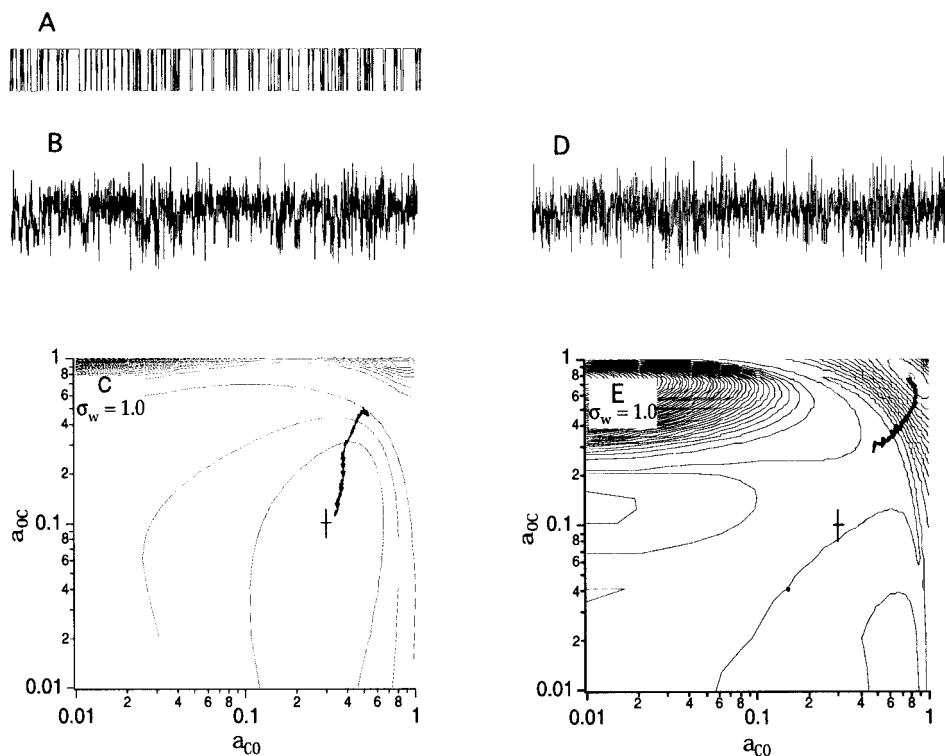


FIGURE 3 HMM likelihood estimation with white and colored noise. (A) Displayed are the first 1024 of the 20,000 data points of noiseless signal simulated from a two-state Markov model consisting of one closed state (C) and an open state (O). The values of the transition probabilities (a_{CO} , a_{OC}) were chosen to be (0.3, 0.1). The closed state had a current level of 0 and the open state a current level of 1. (B) White Gaussian noise with $\sigma_w = 1.0$ was added to the noiseless signal. Displayed are the first 1024 points of the data. (C) Log-likelihood contours obtained as the current levels, the initial state probability vector, and the noise variance were reestimated while fixing the two transition probabilities (a_{CO} , a_{OC}) at various values. The cross represents the true values of (a_{CO} , a_{OC}). The dotted contour represents points at two natural log units below the peak. The arrows show the sequence of reestimated values of the parameters obtained with iteration. The true values of (a_{CO} , a_{OC}) are seen to be close to this contour. (D) Colored noise was added to the signal by passing unity-variance, white noise $w(t)$ through a filter of the form,

$$c_t = m_0 w_t + m_1 w_{t-1}.$$

with $[m_0 \ m_1] = [0.8 \ -0.6]$. Displayed are 1024 points of the 20,000 points of simulated data. (E) Displayed are the log-likelihood contours obtained by traditional HMM analysis. The maximum value of (a_{CO} , a_{OC}) is near (0.9, 0.9) and is very far from the true values indicated by the cross.

ideally described by $S(f) = k_0 + k_1 f^2$ over the experimentally-accessible frequency range from DC to ~ 100 kHz (Levis and Rae, 1993; Sigworth, 1995). The white noise component k_0 is due to leakage conductances and shot noise associated with the patch-clamp amplifier. The f^2 -noise term is due to the amplifier voltage noise, which is imposed on the input capacitance; the result is the differential of a white noise process.

In one approach to handle the signal with correlated noise, the observed data $Y(t)$ are passed through a filter that removes the correlation in the noise. The correlated noise can thus be whitened by convolving it with a finite vector of filter coefficients comprising a so-called moving average filter. This is equivalent to modeling the original correlated noise as the output of the inverse of a moving average filter (an AR filter) driven by white Gaussian noise. Because the noise correlation may be unknown, the coefficients of the

AR filter are also considered to be unknown. The H -noise algorithm provides a framework for estimation of the AR coefficients from the observed data.

The simulation shown in Fig. 3 compares the performance of the traditional forward-backward and Baum-Welch algorithms on data with additive white and correlated noise. The traditional algorithms assume that the background noise is white. In this case, the number of elements k in a metastate can be reduced to 1, the metastate keeps track of only the current state, and Eqs. 9 and 11 reduce to the standard definition of the forward and backward variables (Rabiner, 1989). Noiseless data were generated from a two-state model consisting of one closed state and one open state,



with transition probabilities $a_{CO} = 0.3$ and $a_{OC} = 0.1$ and unit current amplitude. In the first simulation, the data $Y(t)$ were simulated by adding zero-mean, white, Gaussian noise $n(t)$ with variance $\sigma_w^2 = 1.0$ to the noiseless signal. In the second simulation, the data $Y(t)$ were simulated by adding correlated, Gaussian noise generated by passing white, Gaussian noise through a finite impulse response filter. These data were analyzed with a two-state HMM. The log-likelihood contour plot for data with white and correlated noise is shown in Fig. 3 (C and E), respectively. It can be seen in Fig. 3 C that the true values of $(a_{CO}, a_{OC}) = (0.3, 0.1)$ are close to the dotted two-unit log-likelihood contour. This indicates that the traditional forward-backward and Baum-Welch algorithm performs well when applied to data with additive white noise. On the other hand, for data with correlated noise, it is clear that the maximum value of (a_{CO}, a_{OC}) is near $(0.9, 0.9)$ and is very far from the true values of (a_{CO}, a_{OC}) . Thus, the traditional forward-backward and Baum-Welch algorithm perform poorly and yield biased estimates of the parameters when the noise is correlated.

H-noise algorithm

This algorithm considers the effects of the antialiasing filter before digitization of the data as well as correlation in the background noise, making use of the metastate framework. The randomness in the time of transition between states of the underlying continuous time Markov process is considered to be reflected as a randomness in the current level at each time instant at the output of the antialiasing filter. This randomness in the current level of a state at each time instant is modeled as a fictitious, nonstationary noise source called *H*-noise (Venkataramanan et al., 2000).

Shown in Fig. 4 A are three of the infinitely many possible realizations of making a transition from the closed state to the open state and back to the closed state. The response of the antialiasing filter to these three realizations of state transitions is shown in Fig. 4 B. Correspondingly, the current amplitude at each sampling time instant can be thought of as a random variable. As shown in Fig. 4 C, the random current amplitude at each time instant in the metastate can be regarded as the sum of two components—a mean current amplitude and an additional fictitious noise called the *H*-noise. The *H*-noise is nonstationary because its statistics are dependent on the underlying metastate; however, these statistics are easily computed from the known step response H of the antialiasing filter. Thus, the total noise in each sample has two components—the correlated background noise and the *H* noise. In the development of the algorithm, it is assumed that the data are sampled sufficiently quickly such that at most one transition takes place in a sampling interval. The *H*-noise is also assumed to be Gaussian, which is a reasonable approximation under poor signal to noise ratios.

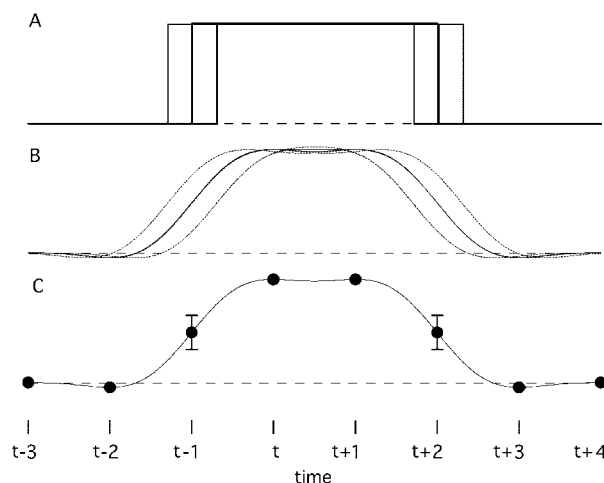


FIGURE 4 Origin of *H*-noise: randomness in time of transition between states in a sampling interval is reflected as randomness in current amplitude. (A) Three possible realizations of making a transition from a closed state to open state and back to a closed state. (B) Signal at the output of an anti-aliasing filter (same filter as in Fig. 2) from the three possible realizations in A. (C) Randomness in time of transition between states causes randomness in current amplitude at each sampling time, which is modeled as a fictitious noise called *H*-noise. Error bars indicate the standard deviation of the *H*-noise.

Correlated emission probability

The *H*-noise can be easily incorporated into the metastate HMM framework through defining the emission probability function to consider the noise correlation in successive data samples. This emission probability is then used in the computation of the forward and backward variables.

Recalling that Y_t is the vector of k successive data samples ending at time t , let the $(k-1)$ data samples ending at time $t-1$ be denoted by \bar{Y}_{t-1} . The correlated emission probability can be denoted by $P(y_t | \bar{Y}_{t-1}, I)$ and is the conditional probability of y_t given the metastate I at time t and the $(k-1)$ previous data samples. It is a Gaussian function of the present sample y_t and \bar{Y}_{t-1} and is given by,

$$P(y_t | \bar{Y}_{t-1}, I) = \frac{1}{\sqrt{2\pi}\sigma_I} \exp\left(-\frac{(h_I^T(Y_t - \Xi m_I))^2}{2\sigma_I^2}\right) \quad (16)$$

in which $m_I^T = [\mu_{i_0} \mu_{i_1} \dots \mu_{i_{k-1}}]$ is the vector of current levels in metastate I . In Eq. 16, Ξ is referred to as the filter matrix and can be easily computed from the known step response of the antialiasing filter (Venkataramanan et al., 2000). The vector Ξm_I denotes the modified current levels corresponding to the different states in metastate I after the temporal mixing of the samples due to the antialiasing filter. The emission probability is maximum when the measured data exactly equals the filtered current levels, i.e., $Y_t = \Xi m_I$. The different terms in the residual $(Y_t - \Xi m_I)$ are weighted and summed to cancel out their correlations according to the vector h_I .

In the special case when the effects of the antialiasing filter and therefore H -noise are ignored, the only contribution to the total noise in each metastate comes from the correlated background noise which is stationary. In this case, the filter matrix Ξ equals an identity matrix. The parameters h_1 and σ_1^2 then have a ready physical explanation. The vector h_1 is the same for all metastates and corresponds to the coefficients of the moving average prewhitening filter, which would whiten the correlated noise. The parameter σ_1^2 is the noise variance at the output of this prewhitening filter. These parameters h_1 and σ_1^2 can be estimated from the estimated correlation Σ_C of the colored background noise.

When the effects of the antialiasing filter are considered, due to the nonstationary nature of the H -noise, the total noise in each metastate is also nonstationary. Let C and O represent the closed and open states of a channel. The variance of the H -noise is highest in metastates that consist of different conductance states such as (C, O, C) or (O, C, O) and is zero for metastates that consist of the same conductance such as (C, C, C) or (O, O, O). Analogous to the previous case, the parameters h_1 and σ_1^2 are referred to as the metastate moving average (MMA) parameters corresponding to metastate I . They can be easily found from an estimate of the covariance matrix Σ_I of noise in the metastate.

Covariance of H -noise

Because the total noise in metastate I is the sum of two independent components (the correlated background noise and the fictitious H -noise) the second order moments of the total noise Σ_I are the sum of the corresponding second order moments of the correlated background noise and the H -noise in the metastate. Thus,

$$\Sigma_I = \Sigma_C + \Sigma_{H_I} \quad (17)$$

in which Σ_C is the autocorrelation of the colored background noise and Σ_{H_I} denotes the covariance matrix of H -noise corresponding to metastate I . It can be computed in terms of the known step response of the antialiasing filter and the estimated current levels of the states.

The time of transition from one state to the next is random relative to the sampling time. Let θ_0 be a random variable distributed between 0 and 1, $t - \theta_0$ indicating the precise time, relative to the sampling time, of a transition that occurs in the interval $(t - 1, t)$ (Fig. 5). Similarly, let $\theta_1, \theta_2, \dots, \theta_{k-2}$ be independent random variables, such that $t - j - \theta_j, j = 0, \dots, k - 2$ indicates the precise time of transition from the state at time $t - (j + 1)$ to the state at time $t - j$. Let $\vartheta = \{\theta_0, \theta_1, \dots, \theta_{k-2}\}$ be the set of these random variables. Our goal is to compute the mean and second moments of variations (the H -noise) in the filtered signal that arise from the random ϑ .

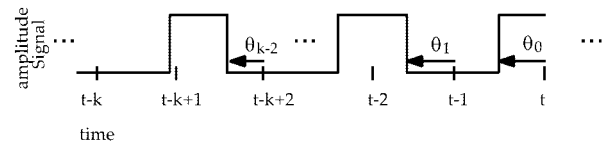


FIGURE 5 Definition of the variables θ_j , which specify the times of transition relative to the sampling times in metastate I . The precise time of transition from the state at time $t - (j + 1)$ to the state at time $t - j$ is given by the random variable $t - j - \theta_j, j = 0, \dots, k - 2$.

Let $m_I^T = [\mu_{i_0} \mu_{i_1} \dots \mu_{i_{k-1}}]$ be the vector of current levels associated with the metastate I . For a given value of ϑ , let $\tilde{m}_I(\vartheta)$ denote the current levels associated with the metastate modified due to antialiasing filter. Recalling that the covariance of two random variables x_1 and x_2 is defined as

$$\text{cov}(x_1, x_2) = E(x - \bar{x}_1)(x - \bar{x}_2) \quad (18)$$

in which the expectation operator $E(\cdot)$ is taken over all possible values of variables x_1 and x_2 and \bar{x}_1 and \bar{x}_2 denote the mean values of x_1 and x_2 , the covariance matrix of H -noise is similarly obtained as

$$\Sigma_{H_I} = E_{\vartheta}[(\tilde{m}_I(\vartheta) - \Xi m_I)^T (\tilde{m}_I(\vartheta) - \Xi m_I)] \quad (19)$$

in which $E_{\vartheta}(\cdot)$ denotes the expectation operator taken over all possible values of ϑ . The matrix Ξ is chosen such that the vector Ξm_I equals the mean value of $\tilde{m}_I(\vartheta)$, the average again being taken over possible values of ϑ . Because each element of ϑ uniformly takes values between 0 and 1, the above equation can be further reduced to,

$$\begin{aligned} \Sigma_{H_I} = & (\mu_{i_0} - \mu_{i_1})^2 V_0 + (\mu_{i_1} - \mu_{i_2})^2 V_1 + \dots \\ & + (\mu_{i_{k-2}} - \mu_{i_{k-1}})^2 V_{k-2}, \end{aligned} \quad (20)$$

in which the matrices $V_j, j = 0, \dots, k - 2$ are given by,

$$V_j = \text{cov}(H(\theta + j - m + 1), H(\theta + j - n + 1)). \quad (21)$$

These matrices can be computed from the known step response H of the antialiasing filter as θ uniformly takes values between 0 and 1.

Reestimation

Once the emission probability has been computed, the forward-backward and Baum-Welch algorithm can be used to calculate the likelihood and reestimate the HMM parameters (Venkataramanan et al., 2000). Briefly, our strategy for estimation of the noise model parameters from initial estimates is as follows. The matrix Ξ is fixed, being determined by the antialiasing filter. The reestimation of MMA parameters is equivalent to reestimating matrix Σ_C . The MMA reestimation is similar to reestimation of elements of the transition probability matrix A in Eq. 14. For example, in the

case that the noise statistics are the same for all metastates, the noise variance can be reestimated as,

$$\hat{\sigma}^2 = \frac{\sum_l \sum_{t=1}^T \gamma_t(l) ((h_l^T Y_t) - \Xi m_l)^2}{\sum_l \sum_{t=1}^T \gamma_t(l)}. \quad (22)$$

The reestimation formulas for other noise model parameters are given in Venkataramanan et al. (2000).

ISSUES IN PRACTICAL IMPLEMENTATION

Inverse filter

During acquisition of the data, the analog data are passed through an antialiasing filter to remove the high frequency noise and then are sampled at discrete times. Because the underlying signal can be thought of as a stochastic square wave, an analog filter with a sharp roll-off in the transition region in the frequency response produces overshoot and ringing phenomena. This is undesirable if the data are analyzed using traditional techniques such as threshold detection. Therefore, an analog filter with a gradual roll-off and linear phase response such as a Bessel filter has been conventionally used.

For analysis of the data using our algorithm, a filter with a sharp roll-off in the transition region is preferred for the following reasons. First, the noise spectral density is found to increase with frequency. Therefore, for the same cut-off frequency, a filter with a sharp roll-off has a smaller noise standard deviation than a filter with a gradual roll-off. Second, use of a filter with a gradual roll-off requires that the data be sampled at several times the Nyquist rate to prevent aliasing (Colquhoun and Sigworth, 1995). This would necessitate a large size for the metastate to characterize the long step response, thus increasing the computational complexity of the HMM algorithms. On the other hand, analog filters with a sharp roll-off characteristic do not have linear phase response, which can cause problems in analyzing the data.

Qin et al. (2000b) make use of a digital inverse filter to shorten the step response of the recording system. We describe here our implementation of a digital inverse filter that accomplishes a similar purpose. Its function is to remove the effects of the data acquisition system (including the effects of the patch clamp amplifier and the analog antialiasing filter) and produce a desired impulse response. We use an impulse response of the form (Venkataramanan, 1998)

$$h(n) = \exp\left(-\frac{f_x^2 n^2}{\sigma_f^2}\right) \frac{\sin(2\pi f_x n)}{(2\pi f_x n)}, \quad f_x = \frac{f_c}{f_s} \quad (23)$$

Here, f_x is the ratio of the bandwidth of the filter f_c to the data sampling rate f_s . The parameter σ_f controls the width and slope of the frequency response in the transition region. If $\sigma_f \ll 1$, the filter is essentially Gaussian with bandwidth

f_c/σ_f . If $\sigma_f \gg 1$, the filter has a sharp roll-off in the transition region with bandwidth f_c .

The inverse filter provides two main advantages. First, it provides the ability to increase or decrease the filter bandwidth and therefore allows the analysis of the data at a cut-off and sampling frequency that may be different from that used during data acquisition. Second, the resulting filter cut-off frequency can be placed close to one-half the sampling frequency. This allows the use of a lower-order noise model and a smaller size metastate, thus reducing the computational complexity of the HMM algorithms.

The digital inverse filter is built as follows. Let $t(n)$ and $e(n)$ denote the desired impulse response and measured impulse response of the acquisition system. Let $T(\omega)$ and $E(\omega)$ denote their corresponding Fourier transforms, here expressed as functions of the angular frequency $\omega = 2\pi f$. The impulse response $h_{\text{inv}}(n)$ of the inverse filter is then given by

$$h_{\text{inv}}(n) = F^{-1}\left(\frac{T(\omega)}{E(\omega)}\right) \quad (24)$$

in which F^{-1} denotes the inverse Fourier transform operation. Direct measurement of the impulse response $e(n)$ of the recording system is difficult to perform due to the limited dynamic range of the patch-clamp amplifier. Instead, the unit step response of the system denoted by $u(n)$ can easily be obtained by applying a triangle wave voltage through a capacitor into the head stage amplifier (Sigworth, 1995). In the Fourier domain, the impulse response of the acquisition system is related to its step response by

$$E(\omega) = i\omega U(\omega). \quad (25)$$

Let $\hat{e}(n)$ be an approximation to $e(n)$, obtained by passing the measured step response through a first-order differentiator of the form,

$$\hat{e}(n) = u(n) - u(n-1). \quad (26)$$

Taking the Fourier transform of Eq. 26, we obtain,

$$\begin{aligned} \hat{E}(\omega) &= U(\omega)[1 - e^{-i\omega}] \\ &= U(\omega)e^{-i\omega/2}[e^{i\omega/2} - e^{-i\omega/2}] \\ &= 2iU(\omega)e^{-i\omega/2} \sin(\omega/2) \end{aligned} \quad (27)$$

From Eqs. 25 and 27,

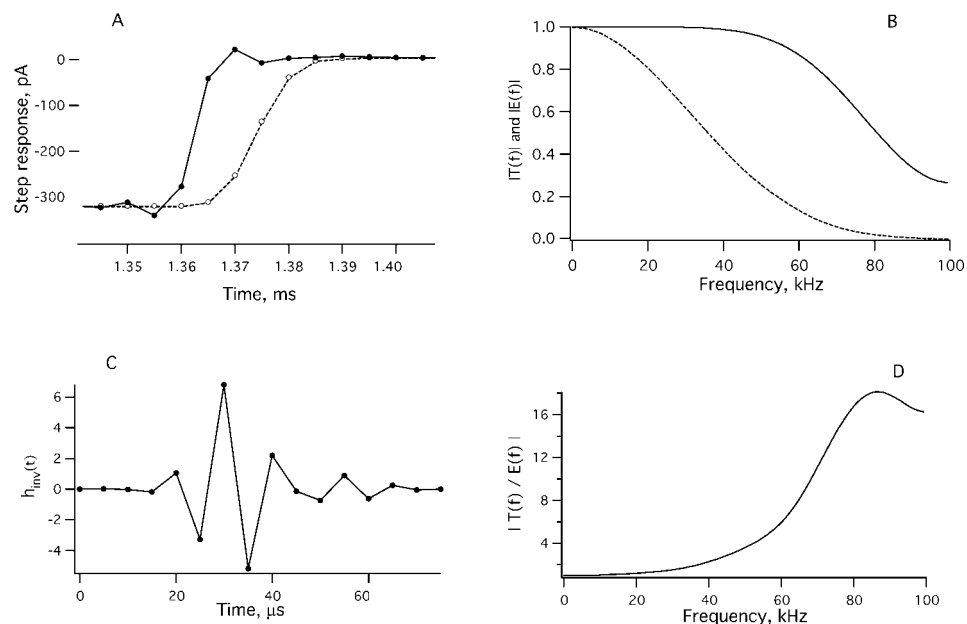
$$\frac{E(\omega)}{\omega} = \frac{\hat{E}(\omega)}{2 \sin(\omega/2) e^{-i\omega/2}} \quad (28)$$

Therefore, from Eqs. 24 and 28,

$$h_{\text{inv}}(n) = F^{-1}\left(\frac{2T(\omega)\sin(\omega/2)e^{-i\omega/2}}{\hat{E}(\omega)\omega}\right) \quad (29)$$

in which the factor $e^{-i\omega/2}$ represents a time delay of one-half sample interval and can be ignored.

FIGURE 6 Construction of the inverse filter used in Zheng et al. (2001). (A) Measured step response of the recording system (*dashed line*), which included an 8-pole, 30-kHz Bessel filter. The sampling frequency was 200 kHz. The final step response that was obtained by convolution with the inverse filter coefficients is shown by the solid line. (B) Frequency response of the recording system (*dashed curve*) and the desired frequency response after inverse filtering (*solid curve*). In building the inverse filter, the parameters were assigned $\sigma_f = 1/\sqrt{2}$ and $f_c/f_s = 0.4$, yielding an 80-kHz cutoff frequency. (C) The sixteen inverse filter coefficients $h_{inv}(n)$. (D) The frequency response of the inverse filter.



The impulse response of the inverse filter is given by Eq. 29 and is computed from the desired impulse response and the measured step response. Of course, the inverse filter coefficients are more accurately obtained when the sampling rate of the step response is larger than the data sampling rate. To avoid artifacts due to noise in the experimental $\hat{e}(n)$, we multiply it by a raised cosine window function that is centered on the impulse. To automate the construction of the window, we determine its location and width from the location and width of the impulse, as determined from the first and second moments of $|\hat{e}(n)|$.

In our implementations, values of f_c/f_s are in the range of 0.4 to 0.5. The cut-off frequency can further be reduced when the data are decimated after inverse filtering. Following are some guidelines in choosing the decimation factor. The greater the decimation factor, the smaller the bandwidth of the inverse filter. This results in good signal to noise ratios in the inverse filtered data as more noise is removed. However, a significant amount of signal information is lost due to the low cut-off frequency of the filter. This will lead to biased estimates of parameters of the Markov model if the transition rates approach the sampling rate. On the other hand, the smaller the decimation factor, the lower is the SNR of the inverse filtered data. Therefore, to obtain unbiased parameter estimates, the noise model must be more accurate leading to larger size of the metastate to capture the correlation in successive noise samples (Venkataramanan et al., 1998). The large size of the metastate leads to higher computational intensity of the H -noise algorithm. In practice, the decimation factor should be chosen such that the final sampling frequency of the data f_s should be larger than the expected highest rate in the Markov scheme. It should be noted that the inverse filter can be successfully constructed

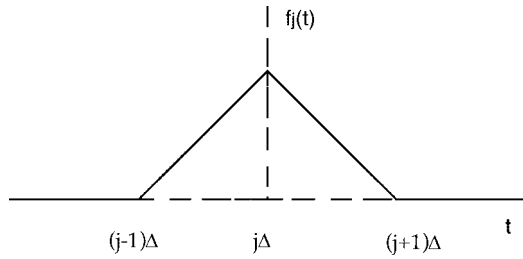
with an increase in bandwidth, if the frequency response of the acquisition system has a significant magnitude (at least roughly 1/10 of the maximum magnitude) at the desired bandwidth.

An example of the step response and the calculated frequency response of a patch-clamp recording system with a 30-kHz Bessel filter is shown in Fig. 6 (A and B) in the time and frequency domains (*dashed line*). The step response convolved with the inverse filter coefficients results in a sharp-cutoff filter with $f_c = 80$ kHz and is indicated by the solid lines. The 16-point inverse filter kernel is shown in Fig. 6 (C and D). There exists an overshoot in the final step response (visible also in Fig. 2, where a filter of the form of Eq. 23 was also used). However, data analysis using the algorithms described in this paper are insensitive to the overshoot because the filter matrix Ξ explicitly takes the shape of the step response into account.

Baseline estimation

In addition to the Gaussian noise, it is often seen that single-channel data are contaminated with nonrandom interferences such as a sinusoidal hum from the alternating current mains. In this subsection, we extend the results in Chung et al. (1990) to estimate the baseline iteratively from the single channel data in a metastate HMM framework.

The observed single channel data are modeled as the sum of three components—a noiseless signal that consists of current levels of the states obtained as the channel makes transitions from one state to another, Gaussian noise, and the drifting baseline $\eta(t) = \{\eta_t, t = 1, \dots, T\}$. We present below a method to characterize and extract the baseline from the single channel data.

FIGURE 7 The baseline kernel function $f_j(t)$ given by Eq. 32.

Generally, baseline drifts are slow variations in the background patch current. Therefore, the baseline can be modeled as a piecewise linear function. Let T denote the total number of points in the data set. Let the data set be divided into r segments of Δ points each, therefore $r = T/\Delta$. Let the baseline be approximated by a piecewise linear function of the form,

$$\eta_t = \sum_{j=0}^r \kappa_j f_j(t) \quad (30)$$

in which

$$f_j(t) = \begin{cases} 1 - \frac{|t - j\Delta|}{\Delta}, & |t - j\Delta| < \Delta, \\ 0 & \text{otherwise.} \end{cases} \quad t = 1, \dots, T \quad (31)$$

A graph of the function $f_j(t)$ is shown in Fig. 7.

The parameters κ_j ($j = 0, \dots, r$) in Eq. 30 are referred to as the vertices of the baseline. They are estimated iteratively in a metastate framework from the measured data $Y(t)$ as follows. *Step 1)* The baseline is computed from the initial or previous estimates of κ_j ($j = 0, \dots, r$). From Eqs. 30 and 31, the baseline at time t is computed from the vertices κ_j as

$$\eta_t = \kappa_j + (\kappa_{j+1} - \kappa_j) \left(\frac{t}{\Delta} - j \right), \quad j\Delta \leq t \leq (j+1)\Delta \quad (32)$$

Step 2) The baseline $\eta(t)$ is subtracted from the data $Y(t)$. The resulting data are then analyzed with the H -noise algorithm. *Step 3)* After obtaining the probability variables $\gamma_t(I)$ from the forward-backward algorithm, the vertices κ_j ($j = 0, \dots, r$) are updated. Let $\hat{\kappa}_n$ denote the reestimated values of the vertices and let $\hat{\eta}_t = \sum_{j=0}^r \hat{\kappa}_j f_j(t)$. On the lines of the Appendices in Venkataramanan (1998), a closed form expression for the reestimation of $\hat{\kappa}_j$ can be obtained by solving the system of equations,

$$\Omega \hat{\kappa} = \Gamma \quad (33)$$

in which $\hat{\kappa}^T = [\hat{\kappa}_0 \ \hat{\kappa}_1 \ \dots \ \hat{\kappa}_r]$ is the vector of reestimated vertices. Here,

$$\Omega_{nj} = \sum_t \sum_I \frac{\gamma_t(I)}{\sigma_1^2} (h_1^T F_n(t)) (h_1^T F_j(t)), \quad n, j = 0, \dots, r \quad (34)$$

and

$$\Gamma_n = \sum_t \sum_I \frac{\gamma_t(I)}{\sigma_1^2} (h_1^T (Y_t - \Xi m_1)) (h_1^T F_n(t)), \quad n = 0, \dots, r \quad (35)$$

in which

$$F_n(t) = [f_n(t) f_n(t-1) \dots f_n(t-k+1)]^T, \quad n = 0, \dots, r. \quad (36)$$

This iterative baseline estimation procedure is readily incorporated into the main body of the H -noise algorithm.

It should be noted that the average value of the baseline and the current levels of the states are not independent variables and, for this reason, it is necessary to impose a constraint in their reestimation. In one approach, it is assumed that the average value of the baseline is zero. This constraint is imposed in step 3 of the baseline estimation algorithm presented above by taking the average value of the vertices and subtracting it from each vertex. In this case, the final reestimated values of the current levels will be offset by the average value of the baseline. In an alternative approach, it is assumed that the closed states in the HMM always have a current level of zero. The current levels of all the other states and the average baseline offset are reestimated with respect to the current level of the closed states. The two approaches were found to be equivalent in our simulations.

We present in Table 1 simulation results on the two-state HMM specified in Scheme 1. These results are representative of the simulations performed on various HMMs at various signal to noise ratios. It can be clearly seen on comparing columns 4 and 5 that the parameters of the HMM and noise model have a smaller bias when the baseline is estimated. The log-likelihood of the model with baseline estimation is also much larger than the model without baseline estimation.

The original and estimated baseline and signal from this simulation are also displayed in Fig. 8. On comparison of the of the actual and estimated baselines, it is seen that the baseline reestimation procedure reasonably reconstructs the baseline and allows a good cumulative-mean reconstruction of the original signal.

Multiple sweeps

The traditional HMM algorithms focus on parameter estimation from a single data set. In practice, there may be a large number of sweeps of the data. Each sweep is a realization of the underlying Markov process. Because each sweep is independent, the modified forward and backward variables can be computed for each sweep independently. The likelihood of the data is the product

TABLE 1 Simulation results on the two-state HMM specified in Scheme 1

Parameters	Initial estimates	Data simulation model	Parameter estimates*	Parameter estimates†
a_{CO}	0.5	0.1	0.17	0.11
a_{OC}	0.5	0.3	0.42	0.30
μ_1	0.1	0.0	-0.02	0.01
μ_2	0.8	1.0	0.99	0.99
σ_w	0.5	0.5	0.59	0.49
$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$	$\begin{bmatrix} -0.3 \\ -0.1 \\ 0.0 \end{bmatrix}$	$\begin{bmatrix} m_0 \\ m_1 \end{bmatrix} = \begin{bmatrix} 0.8 \\ -0.6 \end{bmatrix}$	$\begin{bmatrix} -0.08 \\ 0.46 \\ 0.32 \end{bmatrix}$	$\begin{bmatrix} -0.61 \\ -0.33 \\ -0.10 \end{bmatrix}$
Maximal log-likelihood		MA Filter	-1320.48	1324.02

*Estimates on filtered, sampled data without baseline reestimation.

†Estimates on filtered, sampled data with baseline reestimation.

Comparison of parameter estimates without and with baseline estimation using the H -noise algorithm: The data were generated as from a two-state HMM as follows. The signal consisting of 160,000 data points was generated at fine time steps from the HMM specified by the transition probability matrix

$$B = \begin{bmatrix} 0.9845 & 0.0155 \\ 0.0464 & 0.9536 \end{bmatrix}.$$

The signal was passed through the antialiasing filter specified in (23) with $f_x = 0.5/8$, $\sigma_f = \sqrt{1/2}$ and then decimated by 8. Additive correlated noise was obtained by passing white noise $w(t)$ of standard deviation $\sigma_w = 0.5$ through a first-order moving average filter, $c_t = 0.8w_t - 0.6w_{t-1}$. In addition, a baseline drift of the form

$$\eta(t) = 0.5 \sin(4\pi t/T), \quad t = 1, \dots, T$$

was added to the signal at the output of the sampler. Because the data are decimated by 8, the transition probability matrix of the Markov process at the output of the sampler is

$$A = B^8 = \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix}.$$

The data thus obtained were analyzed by the H -noise algorithm with and without baseline estimation.

of the likelihood of each sweep. Similarly, the change in the reestimation formulae to consider multiple sweeps is readily apparent. Because the sweeps are independent of each other, the reestimation formulae described in Venkataramanan et al. (2000) are modified by summing the

left hand and right hand side (or numerator and denominator) of all the reestimation formulae over all sweeps (Venkataramanan, 1998). For example, let $\gamma_t^{(j)}(I)$ denote the a posteriori probability of being in metastate I at time t for the j -th sweep. The reestimation for the transition probabilities is obtained by modifying Eq. 14 by summing the numerator and denominator over all sweeps,

$$\hat{a}_{i_1 i_0} = \frac{\sum_j \sum_{t=1}^T \sum_{I \in M_0} \gamma_t^{(j)}(I)}{\sum_j \sum_{t=1}^T \sum_{I \in M_1} \gamma_t^{(j)}(I)} \quad (37)$$

SUMMARY

Patch clamp recordings typically contain correlated background noise. The data are also necessarily filtered before discretization and analysis. We have presented an algorithm for analysis of single channel current recordings at low signal-to-noise ratios when traditional algorithms such as threshold detection cannot be applied. The problem has been approached in a hidden Markov model framework through extension of traditional forward-backward and Baum-Welch algorithms.

The H -noise algorithm addresses the issues of correlated background noise and the effects of the antialiasing filter. The random time of transition between states of the underlying continuous-time Markov process produces a random-

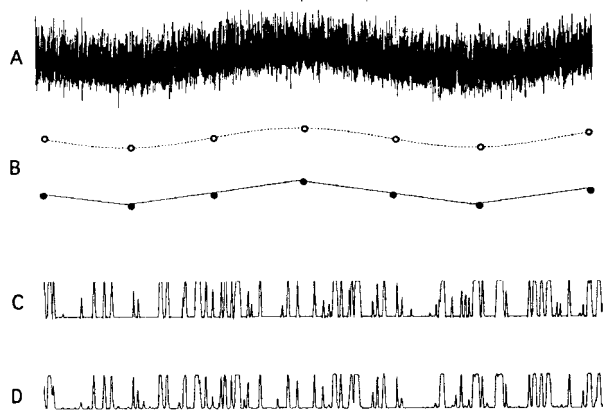


FIGURE 8 An example of baseline estimation. (A) First 1024 points of the filtered, sampled signal with additive correlated noise. The parameters of the HMM and noise model are given in Table 1. (B) True (dotted curve) and estimated (solid line) baseline functions. (C) The first 1024 points of the underlying filtered, noiseless signal from the HMM. (D) Cumulative-mean reconstruction of the signal, computed according to Eq. 15 from the a posteriori probabilities $\gamma_t(I)$.

ness in the observed current amplitude. This randomness in the amplitude is modeled as a fictitious noise source called *H*-noise. Although the parameters are not the maximum likelihood estimates and the *H*-noise method is not guaranteed to monotonically increase the likelihood with iteration, we have found that the *H*-noise method performs acceptably and the parameters are estimated with a tolerable bias (Venkataramanan et al., 2000).

The modified forward-backward algorithm requires on the order of $N^{k+1}kT$ operations, although in special cases and with improvements (Venkataramanan et al., 1998a; Qin et al., 2000b) the computational intensity can be reduced to as little as $r^k NkT$ with r being a constant between 2 and 3. Increasing the size k of the metastate by one introduces one more MMA coefficient for each metastate. Thus the computation of the modified forward-backward variables becomes intensive as the number of states constituting a metastate increases. In practice, for different values of k , the parameters of the HMM and AR noise model can be estimated and a minimum value of k chosen such that further increase in k does not produce an appreciable change in the HMM parameter estimates. As the signal-to-noise ratio of the data decreases, a larger k value is required to provide a more precise model of the noise. From our experience with simulations, we would estimate that a k value of 4 is typically appropriate for data whose SNR is too low by a factor of two to allow conventional threshold analysis.

We have also described the necessity and function of the inverse filter. At the output of the inverse filter, the data can further be decimated allowing a tradeoff to be made between time resolution and computational intensity.

In the case of baseline reestimation for each sweep, it is often found that the initial estimate of the baseline plays an important role in the convergence rate and the local maxima reached by the algorithm. In the case of multiple sweeps, the baseline of each sweep is considered to be independent and is considered to be a piecewise linear function whose vertices are estimated using Eq. 33. In our implementation, the vector of current levels and the average baseline of each sweep are reestimated with respect to the closed state current level. One important factor during baseline reestimation is that the baseline is assumed to vary much more slowly than the single channel data. Convergence to false maxima can occur if the number of points per vertex Δ is specified to be on the same order as the dwell times in one of the states in the model. In this case, it is possible to confuse the signal and the baseline.

We have also considered the extension of the algorithm to multiple sweeps. The modification of the reestimation formulae to consider multiple sweeps is seen to be quite simple; it essentially results only in the addition of a summation sign considered over all sweeps in the reestimation of the HMM, noise model, and baseline parameters.

An example of the practical application of the methods described here is given in Zheng et al. (2001) where the

H-noise algorithm was used to characterize brief subconductance levels in the deactivation time course of *Shaker* potassium channels. In that study a typical data set consisted of 1000 sweeps, each having a length of ~ 120 samples. With the metastate size $k = 4$, to perform 300 iterations of the reestimation algorithm required 2 h for a linear three-state model, or an overnight computation for a branched six to eight state model, on a 1998-vintage computer (Macintosh G3, 300 MHz).

We thank Daniel Brown and Lin Ci Brown (Bruxton Corporation) for integrating our algorithms into the TAC ion-channel analysis software. We also thank Profs. R. Kuc and P. Schultheiss (Yale University) for helpful discussions. This work was supported by NIH grants NS35282 and NS21501.

REFERENCES

- Ball, F. G., and M. S. P. Sansom. 1989. Ion-channel gating mechanisms: model identification and parameter estimation from single channel recording. *Proc. Royal Soc. Lond. B.* 236:385–416.
- Baum, L. E., T. Petrie, G. Soules, and N. Weiss. 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* 41:164–171.
- Becker, J. D., J. Honerkamp, J. Hirsch, U. Fröbe, E. Schlatter, and R. Greger. 1994. Analyzing ion channels with hidden Markov models. *Pflugers Arch.* 426:328–332.
- Chung, S.-H., and P. W. Gage. 1998. Signal processing techniques for channel current analysis based on hidden Markov models. In *Methods in Enzymology*, Vol. 293. Academic Press, New York.
- Chung, S. H., J. Moore, L. Xia, L. S. Premkumar, and P. W. Gage. 1990. Characterization of single channel currents using digital signal processing techniques based on hidden Markov models. *Phil. Trans. R. Soc. London B.* 329:265–285.
- Colquhoun, D., and A. G. Hawkes. 1995. The principles of stochastic interpretation of ion-channel mechanisms. In *Single-Channel Recording*. N. Sakmann, editor. Plenum Press, New York.
- Colquhoun, D., and F. J. Sigworth. 1995. Fitting and statistical analysis of single-channel records. In *Single-Channel Recording*. N. Sakmann, editor. Plenum Press, New York.
- Farokhi, A., M. Keunecke, and U.-P. Hansen. 2000. The anomalous mole fraction effect in Chara: gating at the edge of the temporal resolution. *Biophys. J.* 79:3072–3082.
- Fredkin, D. R., and J. A. Rice. 1992a. Bayesian restoration of single channel patch clamp recordings. *Biometrics.* 48:427–448.
- Fredkin, D. R., and J. A. Rice. 1992b. Maximum likelihood estimation and identification from single-channel recordings. *Proc. R. Soc. Lond. B.* 249:125–132.
- Horn, R., and K. Lange. 1983. Estimating kinetic constants from single channel data. *Biophys. J.* 43:207–223.
- Levis, R. A., and J. L. Rae. 1993. The use of quartz patch pipettes for low noise single channel recording. *Biophys. J.* 65:1666–1677.
- Liporace, L. A. 1982. Maximum likelihood estimation for multivariate observations of Markov sources. *IEEE Trans. Inform. Theory* 28: 729–734.
- Magleby, K. L., and D. S. Weiss. 1990. Estimating kinetic parameters of single channels with simulation: a general method that resolves the missed event problem and accounts for noise. *Biophys. J.* 58:1411–1426.
- Michalek, S., H. Lerche, M. Wagner, N. Mitrovic, M. Schiebe, F. Lehmann-Horn, and J. Timmer. 1999. On identification of Na(+) channel gating schemes using moving-average filtered hidden Markov models. *Eur. Biophys. J.* 28:605–609.

- Milburn, T., D. A. Saint, and S. H. Chung. 1995. The temperature dependence of conductance of the sodium channel: implications for mechanisms of ion permeation. *Recept. Chan.* 3:201–211.
- Qin, F., A. Auerbach, and F. Sachs. 1997. Maximum likelihood estimation of aggregated Markov processes. *Proc. Royal Soc. Lond. (Biol.)* 264: 375–383.
- Qin, F., A. Auerbach, and F. Sachs. 2000a. A direct optimization approach to hidden Markov modeling for single channel kinetics. *Biophys. J.* 79:1915–1927.
- Qin, F., A. Auerbach, and F. Sachs. 2000b. Hidden Markov modeling for single channel kinetics with filtering and correlated noise. *Biophys. J.* 79:1928–1944.
- Rabiner, L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*. 77:257–285.
- Sigworth, F. J. 1995. Electronic design of the patch clamp. In *Single-Channel Recording*. N. Sakmann, editor. Plenum Press, New York.
- Sunderman, E. R., and W. N. Zagotta. 1999a. Mechanism of allosteric modulation of rod cyclic nucleotide-gated channels. *J. Gen. Physiol.* 113:601–619.
- Sunderman, E. R., and W. N. Zagotta. 1999b. Sequence of events underlying the allosteric transition of rod cyclic nucleotide-gated channels. *J. Gen. Physiol.* 113:621–640.
- Venkataramanan, L. 1998. Hidden Markov modeling of ion channel currents. PhD Thesis. Yale University.
- Venkataramanan, L., R. Kuc, and F. J. Sigworth. 2000. Identification of hidden Markov models for ion channel currents. Part III: Band-limited, sampled data. *IEEE Trans. Signal Proc.* 48:376–385.
- Venkataramanan, L., J. L. Walsh, R. Kuc, and F. J. Sigworth. 1998. Identification of hidden Markov models for ion channel currents. Part I: Colored background noise. *IEEE Trans. Signal Proc.* 46:1901–1915.
- Wang, H. L., K. Ohno, M. Milone, J. M. Brengman, A. Evoli, A. P. Batocchi, L. T. Middleton, K. Christodoulou, A. G. Engel, and S. M. Sine. 2000. Fundamental gating mechanism of nicotine receptor channel revealed by mutation causing a congenital myasthenic syndrome. *J. Gen. Physiol.* 116:449–460.
- Zheng, J., L. Venkataramanan, and F. J. Sigworth. 2001. Hidden Markov model analysis of intermediate gating steps associated with pore gate of *Shaker* Potassium channels. *J. Gen. Physiol.* 118:547–562.