Identification of Hidden Markov Models for Ion Channel Currents—Part III: Bandlimited, Sampled Data

Lalitha Venkataramanan, Roman Kuc, and Fred J. Sigworth

Abstract—Hidden Markov models (HMM's) have been used to model single channel currents as recorded with the patch clamp technique from living cells. Continuous time patch-clamp recordings are typically passed through an antialiasing filter and sampled before analysis. In this paper, an adaptation of the Baum-Welch weighted least squares (BW-WLS) algorithm called the H-noise algorithm is presented to estimate the HMM and noise model parameters from bandlimited, sampled data. The effects of the antialiasing filter and the correlated background noise are considered in a metastate or vector HMM framework. The “correlated emission probability,” which plays a central role in the algorithm, is redefined to consider the noise correlation in successive filtered, sampled data points. The performance of the H-noise algorithm is demonstrated with simulated data.

Index Terms—Filtered sampled data, hidden Markov model, ion channel.

I. INTRODUCTION

ION CHANNELS are protein molecules imbedded in the membranes of living cells that act as molecular transducers. They form pores that allow ionic currents on the order of 1 pA to pass across the membrane. Currents passing through individual channels can be recorded using the patch-clamp technique [1]. The open-close behavior of ion channels is commonly modeled as a continuous-time Markov process characterized by a rate matrix $Q$. A state $q_i, i = 1, \ldots, N$ in the Markov model corresponds to a physical configuration of the protein. Each state has a discrete current level $I_q$ associated with it. As the channel makes transitions from one state to another, a “noiseless” signal is obtained corresponding to the current levels of the underlying channel states. This signal is corrupted with additive Gaussian noise from the recording apparatus.

This paper is the third in a series that deals with the estimation of parameters of HMM’s for ion channel activity in situations when the signal-to-noise ratio in the recordings is poor and does not allow direct detection of “open” and “closed” dwells. The first paper in this series [2] addresses the issue of additive colored noise in the discrete-time data by considering “metastate” or vector hidden Markov processes. It introduces the discrete-time metastate algorithm to estimate the parameters of the underlying discrete-time HMM. The second paper [3] presents a noise model to address the problem of “open-channel” noise and introduces the Baum Welch weighted least squares (BW-WLS) algorithm to estimate the HMM and noise model parameters from the observed data. However, for simplicity, the approaches in [2], [3] ignore the effects of the anti-aliasing filter.

Fig. 1 illustrates the block diagram of the anti-aliasing filter and sampler. The output of a two-state Markov process with one closed state and one open state is shown in the top trace. The second trace displays the signal at the output of the antialiasing filter. The last trace displays the filtered, sampled signal. It is evident from the last trace that the bandlimited, sampled data can no longer be modeled as the output of the original two-state HMM. Due to filtering, the current amplitude at time is not only a function of the underlying state but the state sequence of the continuous time Markov process as well. Thus, the discrete-time HMM is “hidden” due to its aggregated nature, any additive noise, and due to the effects of filtering and sampling of the continuous-time data.

Manuscript received August 17, 1998; revised July 30, 1999. This work was supported by NIH Grants NS21501 and NS35282 to F. J. Sigworth. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lal C. Godara.

L. Venkataramanan is with Schlumberger–Doll Research, Ridgefield, CT 06877 USA.
R. Kuc is with the Department of Electrical Engineering, Yale University, New Haven, CT 06520 USA.
F. J. Sigworth is with the Department of Cellular and Molecular Physiology, Yale University, New Haven, CT 06520 USA.

Publisher Item Identifier S 1053-587X(00)00989-2.
In this paper, we present an adaptation of the BW-WLS algorithm called the H-noise algorithm, which incorporates the effects of the antialiasing filter into the metastate HMM framework. The key idea behind the algorithm is the redefinition and computation of the "correlated emission probability" [2]–[4] to consider the noise correlation in successive data samples.

II. THE NOISE MODEL, HMM AND METASTATE HMM

A. The Noise Model

The background additive correlated noise \( C(t) = \{c_t, t = 1, 2, \ldots, T\} \) at the output of the antialiasing filter in patch-clamp recordings is modeled as the output of a \((p-1)\)-order autoregressive (AR) filter with white Gaussian noise \( u(t) \) at the input. The parameters of the autoregressive noise model are the coefficients \( a^T = [d_1 \ d_2 \ \cdots \ d_{p-1}] \) and variance \( \sigma_u^2 \) of the process \( u(t) \) and are assumed to be unknown. Let the autocorrelation of \( C(t) \) estimated to \( k \) lags be \( R_k^C = [R(0) \ R(1) \ \cdots \ R(k-1)] \). The autocorrelation \( R_k^C \) of the correlated background noise is re-estimated from its initial estimates along with the parameters of the HMM in a procedure described in Section IV. After the algorithm converges, the AR parameters can be estimated from final estimates of using the Levinson-Durbin algorithm [5].

B. Hidden Markov Model and Metastate HMM

Let the underlying \( N \)-state continuous-time Markov process be characterized by rate matrix \( Q \) where current levels \( \mu_i \) correspond to each state \( q_i \) and initial state probability \( \pi_i, i = 1, \cdots, N \). If the data from the continuous-time Markov process are sampled at interval \( \Delta_t \), the resulting discrete-time HMM is characterized by transition probability matrix (TPM) \( A = \{a_{ij}, i, j = 1, \cdots, N\} \), where

\[
A = e^{Q \Delta t}.
\]

As in the first two papers in the series, the problem is to estimate the discrete-time HMM parameters from the data \( Y(t) \). From the estimate of the TPM, the rate matrix can be computed by making a suitable approximation to (1).

The correlated background noise is modeled as the output of a metastate Markov process wherein each \( k \)-tuple of successive states \( (s_t, s_{t-1}, s_{t-k+1}) \) forms a "metastate" [2], [6]–[9]. A metastate at time \( t \) is denoted by \( q_t = (q_{t_0}, q_{t_1}, \cdots, q_{t-k+1}) \), where \( q_{t_j} \) is the state of the HMM at time \( t-j \), \( j = 0, \cdots, k-1 \), and \( i_j = 1, \cdots, N \). Similar to the traditional forward-backward procedure described in [10], modified forward and backward variables can be computed in the metastate framework. They help compute the likelihood of the HMM and are useful in the re-estimation of the HMM and noise model parameters [2], [3], [8]. The modified forward and backward variables at each time instant in each metastate are computed iteratively using (3) and (7) in [3]. In the computation of the modified forward and backward variables, the expression \( b_j(y_t | \tilde{Y}_{t-1}) \) denotes the correlated emission probability and is the conditional probability of \( y_t \) given the metastate \( j \) at time \( t \) and the \((k-1)\) previous data samples. In the case of filtered, sampled data considered in this manuscript, the expression for the emission probability can be obtained in terms of the step response of the antialiasing filter, the current levels of the underlying metastate and the parameters of the noise model; it is derived in the next section.

The likelihood of the HMM can be computed from the modified forward variables at the last time instant. The modified forward and backward variables are useful in reestimation of the HMM parameters, which is described in Section IV. The computation of the modified forward-backward variables requires \( O(N^{k+1}T) \) computations; some methods to reduce the computational intensity of the forward backward variables are indicated in [2], [6].

III. CORRELATED EMISSION PROBABILITY

The correlated emission probability (which is also called the symbol probability) [2]–[4] is defined as the conditional probability of given the metastate at time and the previous data samples

\[
b_j(y_t | \tilde{Y}_{t-1}) = P(y_t | y_{t-1}, \cdots, y_{t-k+1}, s_t = q_{t_0}, q_{t_1}, \cdots, q_{t-k+1} = q_{t-k+1}),
\]

(2)

It is useful in computing the modified forward and backward variables that are, in turn, useful in the re-estimation of the HMM and noise model parameters described in Section IV.

In the context of filtered sampled data, the correlated emission probability in metastate \( I = (q_{t_0}, q_{t_1}, \cdots, q_{t-k+1}) \) at time instant \( t \) can be redefined as the conditional probability of \( y_t \), given the \((k-1)\) previous data samples and that transitions occur from state \( q_{t_j} \) to state \( q_{t_j+i} \) in \((t-j \ t-j+1)\), \( j = 1, \cdots, k-1 \), or

\[
b_j(y_t | \tilde{Y}_{t-1}) = P(y_t | \tilde{Y}_{t-1}, q_{t_0} \rightarrow q_{t_1} \rightarrow \cdots \rightarrow q_{t_k}, q_{t-k} \rightarrow q_{t_k-1} \rightarrow \cdots \rightarrow q_{t_0})
\]

(3)

To properly characterize the rates, the data must be sampled much more quickly than the rates in the rate matrix \( Q \). If the data are undersampled, it is possible that several transitions may occur among states of the Markov process in one sampling interval. However, oversampling the data requires that the size of the metastate be sufficiently long to model the strongly correlated noise. The computational intensity of the modified forward and backward variables become much more expensive if the size of the metastate increases. Therefore, we prefer not to over sample the data but assume that the data are sampled sufficiently fast to allow proper characterization of the underlying Markov process. We also make a first-order approximation to the multiple transitions in a sampling interval by assuming that a maximum of one transition occurs between states in a sampling interval. In practice, this assumption is not over-restrictive since it is experimentally straightforward to sample the data at high rates.

Let \( \theta \) be a random variable distributed between 0 and 1, and \( t - \theta \) indicating the precise time, relative to the sampling time, of transition from the state at time \( t - 1 \) to the state at time \( t \), as shown in Fig. 2. Similarly, let \( \theta_1, \theta_2, \cdots, \theta_{k-2} \) be independent random variables such that \( t - j - \theta \), \( j = 0, \cdots, k-2 \) indicates the precise time of transition from the state at time \( t - (j+1) \) to the state at time \( t - j \). Let \( \theta = \{\theta_0, \theta_1, \cdots, \theta_{k-2}\} \). As introduced in [2],
The parameters $\theta_0, \theta_1, \ldots, \theta_{k-2}$ are random variables distributed between 0 and 1 such that $t - j - \theta_{j+1} = 0, \ldots, k - 2$ indicates the precise time of transition from the state at time $t - j$ to the state at time $t - j$ in metastate $I$.

[3], the parameter $\mu_{ij}$ denotes the current level corresponding to state $q_{ij}$ in metastate $I = (q_{i0}, q_{i1}, \ldots, q_{ik-1})$. In the present case, the current amplitude at each time instant is a function of random variable $\theta$, the step response of the filter, and the current levels corresponding to the states of the underlying state sequence. Let $\mu_{I,j}(\theta)$ denote the observed current amplitude at lag $j$ in metastate $I$ for a given value of $\theta$. Fig. 3 illustrates the origin of $\mu_{I,j}(\theta)$ for a particular metastate. We provide below an expression for $\mu_{I,j}(\theta)$ in terms of $\theta$, the step response $H(t)$ of the antialiasing filter, and the current levels of the underlying metastate.

### A. Observed Current Amplitude Corresponding to States in a Metastate

In general, for a metastate comprising of $k$ states, let $\mathbf{m}_I = [\mu_{i0}, \mu_{i1}, \ldots, \mu_{ik-1}]$ be the vector of current levels associated with the metastate, where $\mu_{ij}$ is the current level of state $q_{ij}$, $i = 1, \ldots, N$, $j = 0, \ldots, k - 1$. Let $\mathbf{m}_I(\theta) = [\mu_{I,0}(\theta), \mu_{I,1}(\theta), \ldots, \mu_{I,k-1}(\theta)]^T$. Assuming that there is no change in the current level outside the metastate, i.e., $\mu_{ik}=\mu_{ik-1}$ and $\mu_{i0}=\mu_{i0-1} = \ldots$ and $\mu_{ik}=\mu_{ik+1} = \ldots$, the vector $\mathbf{m}_I(\theta)$ can be expressed in terms of $\theta$ and a matrix $\mathbf{J}(\theta)$ computed from the unit step response of the filter

$$
\mathbf{m}_I(\theta) = \mathbf{J}(\theta) \mathbf{m}_I
$$

(4)

where

$$
\mathbf{J}(\theta) = \begin{cases} 
H(\theta_i - i + 1), & j = 1, i = 1, \ldots, k; \\
1 - H(\theta_{i-2} + k - i - 1), & j = k, i = 1, \ldots, k; \\
H(\theta_{j-1} + j - i), & j = 1, i = 1, \ldots, k; \\
- H(\theta_{j-2} + j - i - 1), & \text{otherwise}.
\end{cases}
$$

(5)

In the following subsection, we describe an approach to compute the emission probability.

### B. The H-Noise Method

In this approach, the randomness in $\theta$ is considered to be reflected as a randomness in the current amplitude at each sampling time instant. For example, as shown in Fig. 4(A) are three of the 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(B), 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 some additional fictitious noise. This fictitious noise models the randomness in the current amplitudes at each time instant due to random time of transition between states in the underlying state sequence of the Markov process and is referred to as the H-noise.

Thus, the observed discrete time signal $Y(t)$ can be modeled as the sum of three components: the colored background noise, the H-noise, and the average current amplitude corresponding to each state. Let $Y(t) = [y_{t0}, y_{t1}, \ldots, y_{tk-1}]$. Let $\mathbf{m}_I$ denote the vector of average current amplitudes corresponding to the states of metastate $I$, Let $\Sigma_I$ denote the correlation matrix of the total noise in metastate $I$. It can be computed for each metastate $I$ in terms of the correlation matrix of the background noise and second-order moments of H-noise $\Sigma_H$ in that...
metastate. Thus, in metastate $I$, $Y_t$ is assumed to be Gaussian distributed according to $N(\bar{m}_I, \Sigma_I)$. As will be shown below, $\bar{m}_I$ and $\Sigma_I$ can be estimated from the known step response of the antialiasing filter and the current levels corresponding to the states of the metastate. These parameters are then used to evaluate the emission probability.

1) Average Current Amplitude $\bar{m}_I$ Corresponding to Metastate $I$: Let $E\theta(\cdot)$ denote the expectation operator over $\theta$. Assume that $\theta_j, j = 0, \cdots, k - 1$ are independent and identically distributed. Define

$$e_j = E\theta(H(\theta + j))$$

where the expectation is computed over $\theta$. Here, the subscript has been dropped from $\theta$ on the right-hand side of the equations since $\theta_j$ are assumed to be independent and identically distributed.

Let $\bar{m}_{I,j} = E\theta(\bar{m}_{I,j}(\theta))$ be the average current amplitude at lag $j = 0, \cdots, k - 1$ in metastate $I$. If $\bar{m}_I = [\bar{m}_{I,0}, \bar{m}_{I,1}, \cdots, \bar{m}_{I,k-1}]$ are the vector of average filtered current levels corresponding to metastate $I$, then from (4)

$$\bar{m}_I = E\theta(\bar{m}(\theta)) = E\theta(J(\theta))m_I = \Xi m_I$$

where the matrix $\Xi$ is of dimension $(k \times k)$. From (5) and (7)

$$\Xi_{ij} = \begin{cases} e^{-i-j}, & j = k, i = 1, \cdots, k \\ 0, & j = k, i = 1, \cdots, k. \end{cases}$$

$(8)$

In (8), $\Xi$ is referred to as the filter matrix. Assuming that $\theta$ varies uniformly between 0 and 1, the elements of the filter matrix can be evaluated from the known step response $H(t)$ of the antialiasing filter. Thus, the vector of average filtered current level $\bar{m}_I$ corresponding to metastate $I$ can be computed from (7) in terms of the unit step response of the antialiasing filter and the vector of current levels $m_I$ of the underlying metastate $I$.

2) Second-Order Moments $\Sigma_{H_I}$ of H-Noise Corresponding to Metastate $I$: The additive H-Noise is nonstationary since its statistics are dependent on the underlying metastate. It is also correlated; the correlation depends on the unit-step response of the antialiasing filter. It has zero-mean and is uncorrelated with the background noise. An expression for the correlation matrix of the H-Noise can be obtained, assuming that there are no changes in current levels outside the metastate. Let $\Sigma_{H_I}$ denote the correlation matrix of H-noise corresponding to metastate $I$, where

$$\Sigma_{H_I} = E\theta((\bar{m}_I(\theta) - \bar{m}_I)^T(\bar{m}_I(\theta) - \bar{m}_I)).$$

From (4) and (7)

$$\Sigma_{H_I} = (\mu_{i_0} - \mu_{i_1})^2V_0 + (\mu_{i_1} - \mu_{i_2})^2V_1 + \cdots + (\mu_{i_{k-1}} - \mu_{i_{k-2}})^2V_{k-2}$$

where

$$V_j = \{v_{j-m+1,j-n+1} \mid m, n = 1, \cdots, k, j = 0, \cdots, k - 2 \}$$

and $v_{m,n} = \text{cov}(H(\theta+m)H(\theta+n))$ and can be computed from the known step response of the antialiasing filter as $\theta$ uniformly takes values between 0 and 1.

3) Emission Probability Corresponding to Metastate $I$ in the H-Noise Framework: Since the total noise in metastate $I$ is the sum of two independent components, $\Sigma_I$ can be easily computed from second-order moments of H-noise and the estimated autocorrelation of the background noise

$$\Sigma_I = \Sigma_C + \Sigma_{H_I}$$

where $\Sigma_{H_I}$ is the H-noise covariance matrix in metastate $I$ and can be computed using (9). Let $\Sigma_I$ be partitioned into four submatrices

$$\Sigma_I = \begin{bmatrix} \Sigma_{H_{I1}} & \Sigma_{H_{I2}} \\ \Sigma_{I1} & \Sigma_{I2} \end{bmatrix}$$

where $\Sigma_{H_{I1}}$ is a scalar, $\Sigma_{H_{I2}} = \Sigma_{H_{I1}}^T$, $\Sigma_{I1}$ is a vector of dimension $(k-1)$, and $\Sigma_{I2}$ is a matrix of dimension $(k-1) \times (k-1)$. Using Theorem (2.5.1) in [11], the emission probability for each metastate at each sampling time is

$$b_I(y_t \mid V_{t-1}) = P(y_t \mid V_{t-1}, I) = \frac{1}{\sqrt{2\pi\sigma_I}} \exp \left( -\frac{(h_T^T(Y_t - \Xi m_I))^2}{2\sigma_I^2} \right)$$

where

$$h_T^T = [h_{0t}, h_{1t}, \cdots, h_{k-1t}] = [1, -\Sigma_{H_{I2}}(\Sigma_{I2})^{-1}]$$

and

$$\sigma_I^2 = \Sigma_{H_{I1}} - \Sigma_{H_{I2}}(\Sigma_{I2})^{-1}\Sigma_{I1}.$$
IV. THE H-NOISE ALGORITHM

A modification of the modified forward-backward and Baum-Welch algorithm, called the Baum-Welch weighted least squares (BW-WLS) procedure to characterize data with state dependent excess noise, has been described in [3]. In this section, we present an adaptation of the BW-WLS procedure called the H-noise algorithm to characterize bandlimited, sampled data to compute the likelihood of the HMM and estimate the parameters of the underlying HMM and noise model.

The re-estimation formulae for the parameters of the HMM and noise model are based on constructing an information theoretic Q-function [12]. The re-estimation formulae for the HMM in terms of the modified forward and backward variables are obtained by differentiating this function with respect to the parameters and setting it to zero. However, since the additive H-noise is nonstationary, we follow the strategy detailed in [3] to estimate the noise model parameters. The algorithm to calculate the likelihood of the HMM and estimation of the HMM and noise model parameters is called the H-noise algorithm. It is developed on the same lines as the BW-WLS with some differences. Since the BW-WLS algorithm has been described in detail in [3], we briefly present the H-noise algorithm, focusing only the differences between the two algorithms.

The first step involves computing the initial estimates of the autocorrelation \( R_C \) of noise process \( C(t) \) and is identical to the first step in the BW-WLS algorithm. The second step involves assigning the MMA coefficients and noise variance to each metastate. Due to the presence of H-noise, this step is different from the corresponding step in the BW-WLS procedure. The MMA parameters for each metastate are computed by solving (17), where correlation matrix \( \Sigma_f \) is computed from (12) using estimates of \( \Sigma_C \) and \( \Sigma_{h_f} \). Step 3 in the H-noise algorithm involves computing the modified forward and backward variables that are useful in the computation of the likelihood and re-estimation of the HMM and noise model parameters. The computation of the modified forward and backward variables is given in (3) and (7) in [3], with the expression for the emission probability being given in (14). Let \( \gamma(I) \) denote the a posteriori probability of being in metastate \( I \). It can be computed from the modified forward and backward variables [2], [3]. The initial state probability vector and the TPM are reestimated in terms of \( \gamma(I) \) according to (12) and (13) in [2].

The re-estimation formulae for the current levels \( \mu_i = 1, \ldots, N \) corresponding to each state of the HMM are given similarly to the procedure followed in [2] and [3]. Let \( z \) be the set of all metastates. Let the set of all states be partitioned into \( \rho \) disjoint sets \( Q^{(1)}, Q^{(2)}, \ldots, Q^{(\rho)} \) such that all states that belong to the same partition \( Q^{(i)} \) have the same current level \( \mu^{(i)} \) \( (i = 1, \ldots, \rho) \). Let \( \eta^{(i)} = [\mu^{(1)} \mu^{(2)} \ldots \mu^{(\rho)}] \) be the \( \rho \)-dimensional vector of unique current levels. Let \( \phi^{(i)} = [\phi_0, \phi_1, \ldots, \phi_{k-1}] = h_f(z)X \). The re-estimated current levels are obtained by setting the derivative of the information theoretic Q-function with respect to the current levels to zero, yielding the system of equations

\[
\hat{\eta} = G^{-1}V
\]

where

\[
G_{mn} = \sum_{l \in Z} k_l \sum_{t = 1}^{T} \sum_{j = 0}^{k-1} \gamma_t(I_0(I)) \phi_{j,t} \cdot \mu_{i}^{(m-n)} m, n = 1, \ldots, \rho
\]

and

\[
V_m = \sum_{l \in Z} k_l \sum_{t = 1}^{T} \sum_{j = 0}^{k-1} \gamma_t(I) (h_f(y)Y_t) \phi_{j,t} m = 1, \ldots, \rho
\]

and \( k_f = 1/\sigma_f^2 \). Therefore

\[
\hat{\mu}_i = \hat{\mu}^{(j)} \text{ if } q_i \in Q^{(j)}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, \rho
\]

Step 4 of the algorithm involves re-estimation of the MMA parameters \( h_f \) and \( \sigma_f^2 \) corresponding to each metastate \( I \). Due to the presence of H-noise, the expression for the re-estimation formulae for the MMA parameters \( h_f \) and \( \sigma_f^2 \) corresponding to metastate \( I \) are different from the expression in [3] and are obtained as follows. Let the set of all metastates \( Z \) be partitioned into \( U \) disjoint sets denoted by \( Z_1, Z_2, \ldots, Z_U \) such that all metastates that belong to the same partition have the same background noise and H-noise statistics. Therefore, all metastates that belong to the same set \( Z_J \) have the same set of MMA coefficients denoted by \( h_f \) and noise variance \( \sigma_f^2 \), \( 1 \leq J \leq U \).

Assuming that the current levels and the MMA parameters are independent parameters, they are re-estimated according to [3]

\[
\delta_{Z_j}^2 = \frac{\sum_{I \in Z_j} \sum_{t = 1}^{T} \gamma_t(I) (h_f(Y_t - \Xi m_t))^2}{\sum_{I \in Z_j} \sum_{t = 1}^{T} \gamma_t(I)}, Z_j = 1, \ldots, U
\]

and

\[
\hat{\mu}_{Z_j} = \delta^{-1} \Delta
\]

where \( \hat{H}_{Z_j} = [\hat{h}_{Z_j 1}, \hat{h}_{Z_j 2}, \ldots, \hat{h}_{Z_j k-1}] \) is the vector of re-estimated MMA coefficients, and

\[
\partial_{mn} = \sum_{l \in Z_j} k_l \sum_{t = 1}^{T} \gamma_t(I)(y_{n-m} - \overline{m}_{I,m})(y_{n-m} - \overline{m}_{I,m}) m, n = 1, \ldots, k-1
\]
and

$$\Delta_m = - \sum_{I \in \mathcal{M}} k_I \sum_{I \neq I_m} \gamma_{I,I_m}(y_{I,m} - \overline{y}_{I,m})$$

subject to

$$m = 1, \ldots, k - 1$$

(25)

where $k_I = 1/\sigma_I^2$. Note that $\partial$ and $\Delta$ depend on $Z_I$. If $\hat{\theta}_I$ and $\hat{\delta}_I^2$ are the re-estimated MMA coefficients associated with metastate $I$, then

$$\hat{\delta}_I^2 = \hat{\delta}_Z^2, \text{ if } I \in \mathcal{M}$$

(26)

and

$$\hat{h}_I^T = [\hat{h}_{0_I}, \hat{h}_{Z_I}^T] \text{ if } I \in \mathcal{M}, \text{ with } h_{0_I} = 1.$$  

(27)

The next step in the algorithm involves the estimation of the autocorrelation $\mathbf{R}_C$ of the background noise $C(t)$. Again, due to the presence of H-noise, this step differs slightly from Step 5 of the BW-WLS algorithm. For each metastate $I$, from (17), we have

$$(\Sigma_C + \Sigma_{H_I}) \hat{\theta}_I = \hat{G}_I.$$  

(28)

Thus

$$\Sigma_C \hat{h}_I = \hat{G}_I - (\Sigma_{H_I}) \hat{h}_I.$$  

(29)

Similar to the approach in [3], the left-hand side of (29) can be rewritten in terms of the variable $\mathbf{R}_C$ for each metastate. Stacking the equations from each metastate one below the other leads to an overdetermined system of equations of the form

$$\mathbf{F} \mathbf{R}_C = \mathbf{p} + \epsilon$$  

(30)

where $\mathbf{F}$ and $\mathbf{p}$ are computed by substituting the re-estimated values of the parameters for their expected values, and $\epsilon$ is the error vector that reflects errors in the estimation of the MMA parameters associated with each metastate. As before, the covariance matrix of errors $\mathbf{P} = \text{cov} (\epsilon)$ is assumed to be diagonal. A WLS estimate for $\mathbf{R}_C$ is obtained from (32) in [3].

Re-estimation of the noise model parameters along with the parameters of the HMM is performed by performing Steps 2 through 5 iteratively. After the H-noise algorithm has reached a local maximum, the parameters of the noise model are estimated from the estimated autocorrelation of the background noise $\hat{R}_C(\Delta), \Delta = 0, \ldots, k - 1$ by solving (2) in [2].

The H-noise algorithm shares the limitations of the BW-WLS procedure. The final estimates of the parameters are not guaranteed to be the maximum likelihood estimates. The log-likelihood is also not theoretically guaranteed to monotonically increase with iteration due to the various assumptions and approximations used in the development of the algorithm. However, as will be shown in the following section, the algorithm has been found to perform acceptably in simulations and does converge to estimates close to the maximum likelihood estimates of the HMM and the noise model.

In this section, we have ignored the effects of the state-dependent excess noise. Even though the H-noise algorithm provides a framework to incorporate the different types of noise sources, the algorithm did not work when the problem of state-dependent excess noise was considered in addition to the H-noise and correlated noise sources. The algorithm did not converge or was found to yield parameter estimates very far from the maximum likelihood estimates of the parameters. We suspect that the main reason behind its failure are the weights of the errors in the weighted least squares procedure. In this case, the errors in $\epsilon$ are clearly correlated, and it is no longer reasonable to assume that $\mathbf{P}$ is diagonal. We have been unable to theoretically find the correct form of the correlation matrix of errors and this remains a part of future research.

Summary of Approximations: In the development of the algorithm, we have made the following assumptions and approximations. It is assumed that the emission probability is Gaussian. Indeed, the probability density function (pdf) of $y_h$ is determined by the pdf of the H-noise $\eta_{H}$, which, in turn, depends on the underlying state sequence of the continuous time Markov process and the step response of the antialiasing filter. At signal to noise ratios, where the H-noise contributes a significant amount to the second moment, it has been verified through simulations via computations of moments that the resulting pdf of the $y_h$ is well approximated by a Gaussian.

It is assumed in the algorithm that $\theta_{k+1}$ are independent and identically distributed with a uniform distribution between 0 and 1 and that a maximum of one transition occurs between successive sampling time instants. If the sampling time of the data is smaller than the dwell times in the states or, equivalently, the off-diagonal elements of the TPM are small, then these assumptions are reasonable.

In computing the current level vector $\mathbf{m}_I$ and the second-order moments of H-noise in each metastate, it is assumed that no transition takes place outside the range of the metastate. Therefore, $\mu_{k-1} = \mu_k = \mu_{k+1} = \cdots$, and $\mu_{-1} = \mu_{-2} = \cdots$, Clearly, this is a very restricted assumption, and the average current amplitude $\overline{y}_{i,0}$ at time $t$ is strongly affected by transitions between state $s_i$ and state $s_{i+1}$. Similarly, the average current amplitude $\overline{y}_{i,k-1}$ at time $t - k + 1$ is affected by transitions between state $s_{i-k}$ and state $s_{i-k+1}$. In our application, it is found that the MMA parameters for each metastate are monotonically decreasing with $[\mu_{i,0}] > [\mu_{i,1}] > \cdots$ and converge to zero. Therefore, from (14), a bias in the estimate of $\overline{y}_{i,k-1}$ does not play a major role in the computation of emission probability. On the contrary, a bias in the estimate of $\overline{y}_{i,0}$ strongly influences the computation of the emission probability.

Two solutions to the problem suggest themselves. First, the effect of a transition one sampling time instant in the future can be considered by introducing a uniform random variable $\theta_{k-1}$, where $t-1-\theta_{k-1}$ indicates the time of transition from $s_i$ to $s_{i+1}$. The filtered current vector $\mathbf{m}_I(\theta)$ corresponding to a given $\theta$ and metastate $I$ can be modified to consider the transition from
For example, the filtered current level in metastate \( I = (q_0, q_1, \ldots, q_{N-1}) \) at time \( t \) can be modified as
\[
\hat{\mu}_{I,t}(\theta) = \mu_{i=1}^{N} (\mu_{i-1} - \mu_{i-2}) H(\theta_{k-2} + k - 2) + \cdots + (\mu_{i-2} - \mu_{i-3}) H(\theta_0) + \sum_{i=1}^{N} a_{i,i-1} (\mu_{i-1} - \mu_{i-2}) H(\theta_{k-1} - 1),
\] (31)

Similarly, random variable \( \theta_{k-1} \) can be introduced to consider the effect of transition between state \( s_{t-k} \) and state \( s_{t-k+1} \). However, from (31), the average current amplitude is dependent on the TPM of the HMM, thus complicating the problem of the TPM and current levels reestimation.

An alternate solution in the case where the impulse response of the antialiasing filter is symmetric in time about \( t = 0 \) is based on the fact that the transitions between \( s_t \) and \( s_{t+1} \) do not affect the average current amplitude \( \hat{\mu}_{I,t} \) two sampling time instants away to a great extent. Therefore, by delaying the response of the antialiasing filter to the signal by one sampling time instant, \( \hat{\mu}_{I,t} \) is not affected strongly by transitions between \( s_t \) and \( s_{t+1} \). Equivalently, the random variables \( \theta_j, j = 0, \ldots, k - 2 \) can be considered to be uniformly distributed between \(-1\) and \(0\). This is the solution that was used in our implementation.

V. SIMULATION RESULTS

In this section, we present simulation results of the performance of the H-noise algorithm on bandlimited, decimated data. The simulations presented in this section are representative of the simulations performed on various HMM’s at various signal-to-noise ratios. In Section V-A, we describe the method used for simulating the bandlimited, sampled signal. In Section V-B, we briefly describe the structure and specifications of the anti-aliasing filter. In Section V-C, we present the results of four simulations.

A. Simulation of the Signal

In the following simulations, the continuous-time signal was approximated by a finely sampled discrete-time Markov process characterized by TPM \( B \) and current levels \( \mu_i \) corresponding to state \( q_i, i = 1, \ldots, N \). Let \( f_{IN} \) denote the sampling frequency of this data. The signal was then passed through a FIR antialiasing filter with a half amplitude cut-off frequency of \( f_c \) and then decimated. Let the decimating process be specified by parameter \( c_x \); normally, \( c_x = 8 \) was used in our simulations. Let \( f_s = f_{IN}/c_x \) denote the sampling frequency of the filtered decimated data. The effective TPM of this signal is \( A = (B)^{f_x} \).

B. Antialiasing Filter

In our application, the antialiasing filter is implemented as a digital FIR filter with the impulse response
\[
h(n) = \exp(-m_0^2 f_c^2 sin(2\pi f_c n)/(2\pi f_c n))
\]
\[t_1 = \frac{f_s}{\sigma_f} \quad \text{and} \quad f_x = \frac{f_c}{f_{IN}}, \quad (32)\]

The parameter \( \sigma_f \) 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 \( f_c/\sigma_f \). If \( \sigma_f \gg 1 \), the filter essentially has a sharp rolloff in the transition region with bandwidth \( f_c \). The step response of the filter can be found computationally by integrating (32). Typically, \( f_c \) is 16–32 times lower than \( f_{IN} \); thus, \( f_x \) typically takes values between 1/16 and 1/32.

C. Simulation Results

We present below the results of three simulations. The first simulation compares the parameter estimates obtained from filtered, sampled data using the discrete-time metastate algorithm proposed in [2] with the H-noise algorithm. The H-noise algorithm parameter estimates are compared with the maximum likelihood estimates of the parameters in the second simulation. The last simulation indicates the parameter bias as the data are over and undersampled for a given cutoff frequency of the antialiasing filter.

In the simulations, the signal was obtained at fine time steps from one of the following TPM’s. The first Markov model (denoted M1) has three states:

1) closed state \( q_1 \) with a current level of 0;
2) open state \( q_2 \) with a current level of 1.0;
3) sub-level \( q_3 \) with a current level of 0.2.

The second and third Markov models (M2 and M3) have two states: an open state \( q_1 \) with a current level \( \mu_1 = 1 \) and a closed state \( q_2 \) with an associated current level of \( \mu_2 = 0 \). The TPM’s of the models are

\[
\text{M1: } B = \begin{bmatrix} 0.9699 & 0.0301 & 0.0 \\ 0.0245 & 0.9431 & 0.0324 \\ 0.0483 & 0.0 & 0.9517 \end{bmatrix}, \quad \text{and } \text{M2: } B = \begin{bmatrix} 0.985 & 0.015 \\ 0.046 & 0.954 \end{bmatrix} \quad \text{and } \text{M3: } B = \begin{bmatrix} 0.93 & 0.07 \\ 0.07 & 0.93 \end{bmatrix}.
\]

Model M1 has a short-lived dwell time in the sub-level, and model M2 has a short-lived closed state. Model M3 has a short dwell time in both the open and closed state. The signal generated from one of the above models was passed through the antialiasing filter with a sharp rolloff in the transition region, specified by choosing \( \sigma_f = 1 \sigma_f \). In the first three simulations, the cut-off frequency of the filter was specified by choosing \( f_c = 0.5/c_x \). After passing through the filter, the signal at the output of the filter was decimated by choosing \( c_x = 8 \). As before [3], we set the size of metastate \( k \) equal to the size of the AR noise model \( m_0 \). The additive correlated noise was obtained as the output of a first-order moving average (MA) filter driven by white, Gaussian noise \( u(t) \) with zero mean and variance \( \sigma_w^2 \)

\[
e_t = m_0 u_k - m_1 u_{k-1} \quad (33)
\]

In most cases, the coefficient values were chosen to be \([m_0, m_1] = [0.8, -0.6] \); the spectrum of the colored noise thus obtained is a good fit to that observed in a representative patch clamp recording sampled at 100 kHz during a “silent” period in an experiment in which acetylcholine channel receptor currents
were measured for kinetic analysis [13]. The colored noise thus obtained was added to the signal at the output of the decimator. The following simulation compares the performance of the previous discrete-time metastate algorithm and H-noise algorithm on bandlimited, sampled data. The signal was generated from M1. To the filtered, sampled signal was added correlated noise obtained from (33) with $\sigma_w = 0.3$. In columns 4 and 5 in Table I, we show the estimated parameters obtained with the two algorithms on 30 data sets of 30 000 points each. Comparing columns 3 and 4, it is seen that the parameter estimates obtained by analyzing data with the discrete-time metastate algorithm are biased and that the H-noise algorithm provides better estimates of the parameters. However, in other simulations (which are not shown in the paper), no significant statistical difference was found between the final estimates obtained with the two algorithms at poorer signal-to-noise ratios such as $\sigma_w = 1.0$. Hence, at very poor signal-to-noise ratios, the effects of the antialiasing filter are almost negligible. In the above simulations, the computational time for both algorithms is around 18.0 s/iteration on a 7300/200 Power Macintosh computer running a program written in C. On average, both algorithms converged in about 200 iterations.

The second simulation compares the HMM and noise model parameter estimates obtained from the H-noise algorithm with the maximum likelihood estimates. The difference between the two estimates can be attributed to the various approximations made in the H-noise algorithm. Similar to the BW-WLS procedure in [3], the weights in the weighted least squares procedure have been chosen by an educated guess; the relation of the MMA parameters to the current levels of the HMM has also been ignored during the re-estimation procedure. The simulation compares two sets of the parameter estimates obtained on data generated from M2. We added correlated noise obtained from (33) and $\sigma_w = 0.3$ to the filtered, sampled signal consisting of 25 000 points. The first 1024 points of this data set are shown in Fig. 5(a). First, the HMM and noise model parameters were estimated using the H-noise algorithm on the data set. Next, the maximum likelihood estimates of the parameters were found with the simplex method [14]. The estimated parameters obtained using the two methods are shown in Table II. The final log-likelihood reached by the H-noise algorithm was 0.89 log-units lower than the true value of the maximum. At a lower signal-to-noise ratio with $\sigma_w = 1.0$ (the results are not shown), the final log-likelihood reached by the H-noise algorithm was found to be only 0.04 log units lower that the maximum log-likelihood. Thus, we conclude that the H-noise algorithm parameter estimates are obtained within a tolerable bias.

The following simulation shows the effect of oversampling and undersampling the data. The cut-off frequency of the antialiasing filter was fixed at $f_{\text{c}} = 1/2\pi$ by setting $f_{\text{c}} = 0.5/8$ in (32). A signal consisting of 320 000 points was obtained from M3. To the signal was added correlated noise obtained by setting $[n_0, n_1] = [1/\sqrt{2}, -1/\sqrt{2}]$ and $\sigma_w = 10.0$ in (33). The data were then filtered and sampled. The decimating parameter $c_{\text{c}}$ was varied from 4 to 16. At values of $c_{\text{c}} > 8$, the data are undersampled, and the resulting noise at the output of the sampler is almost white. At values of $c_{\text{c}} < 8$, the data are oversampled, and the resulting noise at the output of the sampler is strongly correlated over several data samples. The data at the output of the decimator were analyzed with the H-noise algorithm and the discrete-time metastate algorithm. From the final estimates of the HMM parameters obtained from the analysis with an algorithm, the rms error for the data set at a particular $c_{\text{c}}$ was computed according to (34), shown at the bottom of the next page, where $\mu_2$, $\mu_1$, and $\mu_2$ are the true values of the off-diagonal elements of $A$ computed from the particular value of $c_{\text{c}}$ and the current levels of the two states, respectively. In

---

### Table I: Comparison of Discrete-Time and H-Noise Algorithms on M1

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Initial estimates</th>
<th>Data simulation model</th>
<th>Discrete -time parameter estimates</th>
<th>H-Noise parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{12}$</td>
<td>0.3</td>
<td>0.18</td>
<td>0.159 ± 0.0005</td>
<td>0.13 ± 0.004</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.1</td>
<td>0.18</td>
<td>0.128 ± 0.024</td>
<td>0.203 ± 0.019</td>
</tr>
<tr>
<td>$a_{23}$</td>
<td>0.4</td>
<td>0.18</td>
<td>0.207 ± 0.023</td>
<td>0.163 ± 0.016</td>
</tr>
<tr>
<td>$a_{31}$</td>
<td>0.2</td>
<td>0.3</td>
<td>0.400 ± 0.059</td>
<td>0.286 ± 0.018</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.1</td>
<td>0</td>
<td>0.042 ± 0.005</td>
<td>0.022 ± 0.003</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.8</td>
<td>1</td>
<td>0.960 ± 0.004</td>
<td>0.981 ± 0.004</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>0.4</td>
<td>0.2</td>
<td>0.218 ± 0.013</td>
<td>0.197 ± 0.007</td>
</tr>
</tbody>
</table>

---

### Table II: Comparison of H-Noise and Maximum Likelihood Estimates on M2

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Initial estimates</th>
<th>Data simulation model</th>
<th>H-Noise parameter estimates</th>
<th>Maximum likelihood parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{12}$</td>
<td>0.5</td>
<td>0.1</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td>$a_{21}$</td>
<td>0.5</td>
<td>0.3</td>
<td>0.315</td>
<td>0.315</td>
</tr>
<tr>
<td>$a_{23}$</td>
<td>0.8</td>
<td>1</td>
<td>0.992</td>
<td>0.992</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.2</td>
<td>0</td>
<td>0.016</td>
<td>0.015</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_w$</td>
<td>0.5</td>
<td>0.3</td>
<td>0.299</td>
<td>0.298</td>
</tr>
<tr>
<td>$\sigma_{\mu}$</td>
<td>0.0</td>
<td>MA filter</td>
<td>-0.614</td>
<td>-0.606</td>
</tr>
<tr>
<td>$d_1$</td>
<td>0.0</td>
<td>[m_1] = [0.8]</td>
<td>-0.306</td>
<td>-0.297</td>
</tr>
<tr>
<td>$d_2$</td>
<td>0.0</td>
<td>[m_1] = [-0.6]</td>
<td>-0.07</td>
<td>-0.068</td>
</tr>
</tbody>
</table>

---

Fig. 5. Data obtained from M2 through antialiasing filter specified in (32) with $\sigma_w = 1.0/\sqrt{2}$, $(f_{\text{c}}/f_{\text{c}_N}) = (1/16)$, and $c_{\text{c}} = 8$. (a) 1024 points of noiseless data after passing through an antialiasing filter. (b) Bandlimited data with additive colored noise obtained from (33) with $[n_0, n_1] = [0.8, -0.6]$ and $\sigma_w = 0.3$. 

---

Log-likelihood
Fig. 6. Mean and standard deviation of the root mean square error with respect to the sampling frequency. The impulse response of the antialiasing filter is given by (32), with $\sigma_f = 1.1/\sqrt{2}$ and the cut-off frequency $f_c$ of the antialiasing filter being fixed at $f_{1N}/16$. The noise was generated from a first-order differentiator with $\sigma_w = 10.0$. The decimating factor $e_x$ was varied between 4 and 16.

(34), $\hat{a}_{12}, \hat{a}_{21}, \hat{\mu}_1$, and $\hat{\mu}_2$ are the final estimates of the parameters obtained from analysis of a data set with either the H-noise or the discrete-time metastate HMM algorithm. The mean and the standard deviation of the rms error was computed from 30 independent data sets. Each data set was analyzed with

a) discrete-time metastate algorithm with $k = 4$;

b) discrete-time metastate algorithm with $k = 5$;

c) H-noise algorithm with $k = 4$;

d) H-noise algorithm with $k = 5$.

The plot of the rms error with respect to $1/e_x$ (or $f_s/f_{1N}$) is shown in Fig. 6. The following points can be made from the simulation. First, the parameter bias obtained with the H-noise algorithm is lower than when the data are analyzed with the discrete-time metastate algorithm. Second, when the data are undersampled, a significant amount of signal information is lost. Consequently, the parameter estimates have a larger bias. In addition, the additive noise at the output of the sampler is more strongly correlated. A large-size metastate needs to be considered for analysis to consider both the correlated noise and to ensure that the size of the metastate is at least comparable with the length of the impulse response of the antialiasing filter. The parameter estimates are biased due to sensitivity to the size of the metastate (or equivalently the size of the noise model).

VI. SUMMARY AND DISCUSSION

Continuous-time patch-clamp recordings are typically passed through an antialiasing filter and sampled before analysis. In [2], [3], and [6], a discrete-time metastate HMM approach has been proposed for considering correlation between noise in $k$ successive data samples. It has been shown that the effects of the antialiasing filter can also be incorporated into the same framework, thus handling both the issues of correlated background noise and effects of the antialiasing filter.

The key idea of the algorithm proposed in this paper is that the random time of transition between states of the underlying continuous-time Markov process results in a randomness in the observed current amplitude. This randomness is modeled as a fictitious noise source called H-noise. The second-order moments of H-noise along with the average current amplitudes in the computation of the emission probability.

The H-noise algorithm is similar to the BW-WLS method proposed in [3] to model state-dependent excess noise. Although the parameters obtained 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 that the parameters are estimated with a tolerable bias. It is seen that the H-noise method performs better than the discrete-time metastate algorithm at various signal-to-noise ratios. The effect of oversampling and undersampling the data when the filter cut-off frequency is specified is also demonstrated via simulations. The algorithms presented in this paper have been recently applied to data from a voltage-gated potassium channel. The results from this analysis will be described in a subsequent paper.

ACKNOWLEDGMENT

The authors are grateful to Professor P. Schultheiss for helpful discussions.

REFERENCES


---

Lalitha Venkataramanan was born in Bombay, India. She received the M.E. degree in electrical engineering from Indian Institute of Science, Bangalore, in 1992 and the Ph.D. degree from Yale University, New Haven, CT, in 1998.

She is currently a Senior Research Scientist with the Petrophysics and NMR Group, Schlumberger Doll Research Laboratories, Ridgefield, CT. Her current research includes time series, protein modeling, and digital signal processing.

Ms. Venkataramanan is a Member of Biophysical Society.

---

Roman Kuc received the B.S.E.E. degree from the Illinois Institute of Technology, Chicago, and the Ph.D. degree in electrical engineering from Columbia University, New York, NY.

He began his career at Bell Telephone Laboratories, where he investigated efficient speech coding techniques. As a Postdoctoral Research Associate with the Department of Electrical Engineering, Columbia University, he applied digital signal processing to diagnostic ultrasound signals to characterize liver disease. In 1979, he joined the Department of Electrical Engineering, Yale University, where, as the Director of the Intelligent Sensors Laboratory, he has been pursuing research in intelligent sensors that extract information from data for applications in robotics and bioengineering. He is the author of *Introduction to Digital Signal Processing* (New York: McGraw-Hill) and *The Digital Information Age* (New York: PWS). Current projects involve investigating biosonar systems, such as bats and dolphins, in order to mimic their behavior with a man-made system. In 1995, he began teaching EE101—The Digital Information Age, which quickly became the largest course at Yale.

Prof. Kuc is a past chairman of the Instrumentation Section of the New York Academy of Sciences. In 1998, he was elected Honorary Academician of the Higher Education Institute of the Academy of Sciences of the Ukraine.

---

Fred J. Sigworth was born in Berkeley, CA, in 1951. He received the B.S. degree in applied physics from the California Institute of Technology, Pasadena, in 1974 and the Ph.D. degree in physiology from Yale University, New Haven, CT, in 1979.

He was a Research Fellow at the Max Planck Institute for Biophysical Chemistry, Göttingen, Germany, from 1979 to 1984, where he worked with E. Neher on the development and applications of the patch-clamp technique for recording single ion-channel currents. Since 1984, he has been a Faculty Member with the Department of Cellular and Molecular Physiology, Yale University. His research interests include the structure and function of ion channel proteins and the development of techniques to study ion channels.

Dr. Sigworth is a Member of Biophysical Society, the Society of Neuroscience, and the American Scientific Affiliation.