

# Coupled Markov Chain Model: Characterization of Membrane Channel Currents with Multiple Conductance Sublevels as Partially Coupled Elementary Pores

SHIN-HO CHUNG Department of Chemistry, Australian National University, Canberra, Australia

AND

RODNEY A. KENNEDY Research School of Information Science and Engineering, Australian National University, Canberra, Australia

Received 10 October 1994; revised 18 May 1995

## ABSTRACT

A parameterized Markov chain model is developed to represent the characteristics of channel currents that either are the superposition of many single channels or show multiple conductance sublevels. The simplified model takes the form of a set of binary chains that are interdependent according to a simple lumped coupling parameter. When varied, this parameter realizes a range of behaviors from tight coupling to complete independence. Other model parameters describe the intrinsic characteristics of the binary chains. An identification procedure for the model parameters is developed based on hidden Markov modeling ideas but incorporating a novel parameter estimation. The usefulness of the model in analyzing certain types of data is demonstrated with examples of real channel currents.

# 1. INTRODUCTION

Analysis of channel currents over the past decade has been dominated by the kinetic model proposed by Colquhoun and Hawkes [1]. It is based on a finite-state, continuous-time, homogeneous Markov process, where each state represents a hypothetical conformational state of the channel macromolecule. Thus, the channel molecule is believed to undergo a small set of discrete conformational changes, from closed states to open states. In this formulation, the observed channel currents are assumed to be binary; namely, the channel can be either open or closed. There are, however, instances in which channel currents cannot be represented as a two-state chain. First, a patch of membrane enclosed by the rim of the electrode tip frequently contains more than one active channel that open and close independently of each other, especially when the patch is maintained at physiological temperatures. Theoretical schemes for deriving properties of multiple channels have been proposed previously [2–4]. Second, numerous types of channels show multiple conductance levels. Among these are channels activated by acetylcholine [5, 6], neuroactive amino acids [7–9], or intracellular messenger systems [10–12]. The distinction between multiple channels and channels with subconductance levels becomes blurred when individual pores within a single channel become partially coupled. Indeed, when the action of a number of these pores couples rigidly, the properties of such multiple pores will be kinetically indistinguishable from those of a single two-state channel.

It is desirable to quantify whether multiple channels contributing to patch-clamp recordings are independent or have some dependency, as suggested by others [13, 14]. The ultimate goal is, naturally, to describe the collective behavior of channels with a set of dynamical equations. As a preliminary step toward such a physical model, a simplified mathematical model is developed in this paper. In the first part of the paper, we formulate a mathematical representation of a vector of partially coupled Markov chains. In our formulation, a finite number of stochastic processes, each characterized by an identical two-state, discrete-time, first-order, homogeneous Markov chain, is coupled with a dimensionless numerical quantity  $\kappa$ , which we call the coupling coefficient, in a way analogous to a system of linear differential equations linked with a static or dynamic coupling term. In the second part, we outline a scheme that enables the identification of the critical parameters of coupled Markov chains, when their realization is embedded in white Gaussian noise. The model identification problem is formulated in the framework of a hidden Markov model (HMM), and then a novel parameter estimation procedure is applied to obtain estimates, optimal in a sense to be defined, of the relevant parameters. When the coupling and other parameters of such a vector Markov process are altered, a rich variety of qualitative behavior is generated. In the final part of the paper, we illustrate how our model can be applied for the analysis of channel currents recorded from excised patches of the biological membranes.

# 2. PROBLEM FORMULATION

#### 2.1. OVERVIEW

Suppose there are two discrete open levels in patch-clamp recordings of ion channels, and the transition from the closed level to the higher open level or vice versa rarely occurs. We will say that the recordings represent the superposition of two independent or uncoupled "pores." We adopt this different terminology because the distinction between single-channel currents and multiple-channel currents, as commonly used in the literature, becomes ambiguous when several single channels contributing to a patch-clamp recording open and close in partial synchrony at times and in full synchrony at other times. When the pores are totally coupled, their collective behavior will appear as a single channel (composed of multiple pores) that can take on only two conductance levels. The model we propose here seeks to explain the discretetime (sampled) observed process by providing a plausible mathematical mechanism for the underlying process. The structure of the Markov chain model is guided by insight into the underlying physical process, the prominent aspect of which is the notion of coupling between *identical* pores which, as an ensemble, can also give an appearance of a single channel or a channel with multiple subconductance levels.

Whereas it is anticipated that the coupling mechanism may be quite involved and depend on the geometry of the channel, our model simplifies the description by characterizing the coupling through a single parameter. In this sense the parameter is *lumped* and the model approximate. This means that when we use the model as the basis for identifying a real channel, for example, we are able to determine in a coarse quantitative manner whether or not the pores exhibit any coupling. Moreover, with our formulation, channel currents representing an algebraic sum of many independent channels can be decomposed and the kinetics of single channels can be estimated.

# 2.2. NOTATION

We define in Table 1 the notation used throughout the paper. Note that matrices are represented with bold type and vectors with capital Roman letters.

### 2.3. SIGNAL MODEL

We make the following assumptions about the Markov process  $s_k^{(r)}$ .

Discrete-Time Representation. It is convenient to deal with discretetime Markov processes embedded in noise. In formulating the signal model, it makes little difference whether the process is envisaged as a continuous-time or discrete-time chain, since we can move between the two process descriptions by observing that the transition probability matrix **P** is the matrix exponential of the intensity matrix **Q** times the sampling period T. [We note here that a given **P** sometimes cannot be mapped back to a valid **Q** via  $\mathbf{Q} = (1/T)\log(\mathbf{P})$ , as  $\log(\mathbf{P})$  may be

TΑ	BL	Æ	1
----	----	---	---

Symbol	Description
k	Discrete time index
L	Number of pores in the channel
$q_1$	Closed-pore conductance level
$q_2$	Open-pore conductance level
$\mathbf{q} \triangleq \{q_1, q_2\}$	Set of possible pore conductance levels
$s_k^{(r)}$	State of $r$ th pore in channel (binary) at time $k$
$v_{ii}$	Transition probability for an isolated pore $s_k^{(r)} \rightarrow s_{k+1}^r, \forall r$
V	Transition probability matrix for a single pore
$\mathbf{P}^{(I)}$	Transition probability matrix for L uncoupled pores
$\mathbf{P}^{(C)}$	Transition probability matrix for L completely coupled pores
Р	Transition probability matrix for L partially coupled pores
$\{u_0,, u_L\}$	Set of $L + 1$ possible output channel conductance level values
$z_k$	Total channel conductance level before additive noise
$y_k$	Noisy measurement of total channel conductance level
$S_k$	State of channel (binary L-vector with components $s_k^{(r)}$ )
ζ	Pore closed-to-closed probability
ρ	Pore open-to-open probability
κ	Coupling coefficient (probability)

Notation and Description of Symbols

complex.] However, for any proposed model of channel dynamics, we need to obtain optimal estimates of the parameters featured in it and evaluate how well it describes the observation sequence. The measurement obtained in the laboratory invariably contains random noise in addition to the signal of interest. It is when dealing with these noisy observations that a continuous-time formulation of a Markov chain poses added technical difficulties. The mathematical tools for handling continuous-time processes, the realization of which is hidden in noise, are still being developed and involve advanced mathematics such as Ito calculus and Wiener processes. The discrete-time formulation, on the other hand, avoids these theoretical problems.

First-Order, Homogeneous Markovian. We assume that the probability of the elementary pore being open at time k + 1 depends solely on the state it was in at time k. The transition probabilities of passing from one state level (closed or open) at time k to another state level (closed or open) at time k + 1 form a 2×2 state transition probability matrix V. We further assume that the transition probabilities are invariant of time k. It is easy to extend the signal model to second- or higher order Markov chains or to semi-Markov processes, in which the transition probability is a function of the time the process spends in a particular

### COUPLED MARKOV CHAIN MODEL

state. A mathematical description of this extension is given by Krishnamurthy et al. [15].

Insensitivity to the First-Order Markovian Assumption. Theoretically, the open- or closed-time interval histogram tabulated from a first-order, homogeneous Markov chain is distributed according to a single exponential function. However, a finite-length data segment does not conform strictly to the first-order Markov statistics. Moreover, interval distributions, especially closed-time distributions, obtained from real channel currents can best be fitted with two or three exponential functions. In formulating our coupled Markov model, we have nevertheless represented the kinetics of an elementary pore as a binary first-order Markov process. This simplification makes the mathematics underlying the coupled chains tractable and reduces the number of parameters that have to be estimated from the observation sequence  $y_k$ . Having devised a parameter estimation scheme based on a simplifying assumption, it is desirable to demonstrate empirically (by deliberately violating the underlying assumption) that the processing scheme is not sensitive to deviation from such an assumption.

Elsewhere, we have demonstrated that the departure from the firstorder Markov assumption has little, if any, effect on the performance of the HMM processing scheme and fluctuation analysis [16-18]. Using these processing schemes formulated under the assumption that the underlying signal sequence can be construed as a first-order Markov process, we analyzed noisy records that contained periodic rectangular pulses or fictitious channel currents obeying a Colquhoun-Hawkes gating model. The accuracies of the estimated parameters (such as the transition probability matrix, signal amplitudes, and signal sequence) were relatively unaffected when the signal statistics deviated considerably from the underlying assumption. From an extensive series of simulations, we have also ascertained that the coupling parameters between the constituent pores can be deduced with a fair degree of confidence by using the algorithm presented here even when the modeling assumptions are invalid. Thus, our analytical method can effectively be applied to a collection of pores that do not strictly conform to first-order Markov processes.

*Remarks.* In the Colquhoun-Hawkes formulation [1], the channel molecule is assumed to undergo a series of conformational changes, from the closed states  $C_0, C_1, \ldots, C_n$  to the open states  $O_0, O_1, \ldots, O_m$ . The underlying transitions from one conformational state to the other, each represented as a continuous-time Markov process, are not directly observable, but some of their properties can be deduced from the behavior of single-channel currents. Because the n + m underlying

conformational states have been aggregated into just open- and closedchannel states, the signal sequence generated by this model is not a two-state Markov chain. When the number of underlying states is small. for example, n = 2 and m = 1, the process can be represented as a discrete-time, three-state, homogeneous hidden Markov model with aggregation to two observable states, and the maximum likelihood estimates of the relevant parameters featured in the model can be obtained [19]. If, on the other hand, a binary process embedded in noise is generated by a large number of underlying conformational states, or the observed data result from the superposition of several such processes, which are partially interdependent, the relevant parameters cannot be easily identified. It is not entirely clear how one can design a computationally efficient maximum likelihood estimator that is robust in the presence of white noise if the channel current is assumed to result from a large number of partially coupled pores, each of which can be modeled as an N-state, continuous-time Markov chain aggregated to a binary process.

We note here that the meaning of "state" in our representation differs from that adopted in the Colquhoun-Hawkes model [1] but is consistent with that used in mathematical literature [e.g., 20, 21]. The underlying conformational state, which is not directly observable from measurements, does not feature in our scheme.

# 2.4. BINARY VECTOR CHAIN PROCESS

Consider L identical discrete-time, binary, homogeneous Markov processes denoted  $s_k^{(1)}, \ldots, s_k^{(L)}$ , such that at each time k,  $s_k^{(r)}$ ,  $r = 1, \ldots, L$ , is a binary random variable taking on one of two states in the set  $\mathbf{q} \triangleq \{q_1, q_2\}$ . So one can for  $q_1$  read "closed" and for  $q_2$  read "open," and the state of the channel corresponds to some combination of L open or closed pores. Define the vector process at time k that describes the ensemble of the L scalar processes,

$$S_k \triangleq \left(s_k^{(1)}, s_k^{(2)}, \dots, s_k^{(L)}\right)$$

with state space

$$\mathbf{q}^{L} \triangleq \underbrace{\mathbf{q} \times \mathbf{q} \times \cdots \times \mathbf{q}}_{L \text{ times}},$$

that is,  $\mathbf{q}^L$  is the L times Cartesian product of  $\mathbf{q}$  (ordered binary L-tuple). Note that the set  $\mathbf{q}^L$  has  $N = 2^L$  elements. We index each of these states according to a binary ordering in Table 2.

#### TABLE 2

Index	State
0	$(q_1, q_1, \dots, q_1)$
1	$(q_2, q_1,, q_1)$
2	$(q_2, q_2,, q_1)$
3	$(q_2, q_2,, q_1)$
:	: -:
$N-1=2^{L}-1$	$(q_2, q_2,, q_2)$

Indices for the Corresponding Pore State Configurations

Each of the identical binary processes will be modeled as a Markov chain, with transition probabilities

$$v_{ij} \triangleq P(s_{k+1}^{(r)} = q_j | s_k^{(r)} = q_i), \quad i, j \in \{1, 2\}; \forall 1 \le r \le L.$$
(1)

So, for example,  $v_{11}$  is the probability that the pore remains closed at the next sampling measurement instant k + 1, given that the pore is closed at the current sampling instant k. These transition probabilities do not depend on values taken by the state at time instants before k; that is, the Markov property means that  $v_{ij}$ , as expressed in (1), does not depend on  $s_{k-1}^{(r)}$ ,  $s_{k-2}^{(r)}$ , or earlier (and further this probability is independent of time k by the homogeneous assumption). This mathematical property considerably simplifies the subsequent analysis and forms perhaps the simplest nontrivial example of a stochastic process.

The corresponding  $2 \times 2$  transition probability matrix that combines the above four transition probabilities will be explicitly represented as

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \triangleq \begin{bmatrix} \zeta & 1-\zeta \\ 1-\rho & \rho \end{bmatrix}$$
(2)

and is sufficiently parameterized, as indicated, by only two parameters,  $\zeta$  and  $\rho$ .

# 2.5. GENERAL MODEL FOR L PORES

When we consider the ensemble of L pores forming the channel, we have an analogous mathematical structure but this time of greater dimension. Let  $\mathbf{P}^{(G)} = (p_{mn}^{(G)})$  denote a generic transition probability matrix of the vector binary process  $S_k$ ; that is, for  $Q_m \triangleq (q_{i_1}, \dots, q_{i_L})$  and  $Q_n \triangleq (q_{j_1}, \dots, q_{j_L})$ ,

$$p_{mn}^{(G)} \triangleq P(S_{k+1} = Q_n | S_k = Q_m) = P(s_{k+1}^{(1)} = q_{j1}, \dots, s_{k+1}^{(L)} = q_{jL} | s_k^{(1)} = q_{i_1}, \dots, s_k^{(L)} = q_{i_L}).$$
(3)

Then  $\mathbf{P}^{(G)}$  is a square matrix with  $4^L$  elements.

### 2.6. UNCOUPLED CASE

We can determine the specific transition probability matrix when the constituent components of the vector process are *independent* as a special case of (3). This means that the pores forming the channel open and close with complete disregard for their neighbors. That is,  $\mathbf{P}^{(I)} = (p_{mn}^{(I)})$  is defined as the *L* times tensor (Kronecker) product of V (2):

$$\mathbf{P}^{(1)} \triangleq \underbrace{\mathbf{V} \otimes \mathbf{V} \otimes \cdots \otimes \mathbf{V}}_{l \text{ times}},\tag{4}$$

where  $\otimes$  denotes the tensor product. Alternatively, this means that the components of (4) can be factored as

$$p_{mn}^{(I)} = \prod_{r=1}^{L} P(s_{k+1}^{(r)} = q_{j_r} | s_k^{(r)} = q_{i_r}).$$

As an example, with L = 2 and V as in (2),

$$\mathbf{V} \otimes \mathbf{V} = \begin{bmatrix} \zeta^2 & \zeta(1-\zeta) & (1-\zeta)\zeta & (1-\zeta)^2 \\ \zeta(1-\rho) & \zeta\rho & (1-\zeta)(1-\rho) & (1-\zeta)\rho \\ (1-\rho)\zeta & (1-\rho)(1-\zeta) & \rho\zeta & \rho(1-\zeta) \\ (1-\rho)^2 & (1-\rho)\rho & \rho(1-\rho) & \rho^2 \end{bmatrix}.$$
(5)

To confirm that this description makes sense we can compute the probability that at time k the first pore is closed and transits to be open and the second pore is open and transits to be closed, say. Since the pores act independently, the probability of the first event is  $v_{12} = 1 - \zeta$ , that of the second is  $v_{21} = 1 - \rho$ , and that of the joint event, by independence, is the product  $(1 - \zeta)(1 - \rho)$ . This product can be identified in (5) as the entry in the second row and third column.

#### 2.7. COUPLED CASE

A second special case of (3), which we denote as  $\mathbf{P}^{(C)}$ , plays an important role. This vector chain represents the situation where the L binary chains are tightly coupled,

$$\mathbf{P}^{(C)} \triangleq \begin{bmatrix} \zeta & 0 & \cdots & 0 & 1-\zeta \\ 0.5 & 0 & \cdots & 0 & 0.5 \\ \vdots & \vdots & & \vdots & \vdots \\ 0.5 & 0 & \cdots & 0 & 0.5 \\ 1-\rho & 0 & \cdots & 0 & \rho \end{bmatrix}.$$
(6)

The selection of the parameters to be equal to those in  $\mathbf{P}^{(I)}$  reflects a modeling assumption that the behavior of an ensemble of pores that are tightly coupled is identical to that of a single isolated pore. The value of 0.5 in (6) affects only the one step transient and is not crucial to our development.

There are only two recurrent vector states:  $(q_1, q_1, ..., q_1)$ , corresponding to all pores being closed, and  $(q_2, q_2, ..., q_2)$ , corresponding to all pores being open. That is, after the short transient, all pores will either be simultaneously closed or simultaneously open and transit is in unison.

# 2.8. PARTIAL COUPLING

In the most useful model that we consider, the constituent chains need not be either independent (4) or fully coupled (6), leading to the transition probability matrix

$$\mathbf{P} \triangleq (1 - \kappa) \mathbf{P}^{(l)} + \kappa \mathbf{P}^{(C)}, \qquad 0 \leqslant \kappa \leqslant 1, \tag{7}$$

where  $\kappa$  (a probability) is the coupling factor. This means that (1) whenever  $\kappa$  is near 0, the pores open and close (essentially) independently; (2) whenever  $\kappa$  is near 1, the pores tend to open and close together as a group; and (3) for values of  $\kappa$  between 0 and 1 the pores have a tendency to open and close in groups, although at times they may seem to act independently.

Gathering together all the parameters, we see that  $\mathbf{P}$  can be parameterized by the set

$$\Theta \triangleq \{\zeta, \rho, \kappa\}. \tag{8}$$

The number of pores, L, also implicitly parameterizes the model. Despite the seemingly low number of parameters, we will see that a very rich range of qualitative behavior is possible. We argue that in a concise framework, this model can account, albeit approximately, for a wide range of seemingly disparate behavior observed in experiments.

# 2.9. AGGREGATED MARKOV CHAIN

Because of the nature of the observation process, it is possible to simplify, particularly in terms of level of computation required, the above Markov chain model to one of a significantly lower order via an aggregation procedure. This is achieved by grouping together the binary vector states into aggregated states in a natural way. We define a scalar finite-state stochastic process corresponding to a noiseless observation of the form

$$z_k \triangleq \mathbf{1}^T S_k = \sum_{r=1}^L s_k^{(r)},\tag{9}$$

that is, the inner product of the state vector  $S_k$  and the vector with all elements equal to 1. The state space of  $z_k$ , or the set of values that  $z_k$  can assume, then consists of L + 1 distinct values in the range

$$u_{0} \triangleq Lq_{1},$$

$$u_{1} \triangleq (L-1)q_{1} + q_{2},$$

$$u_{2} \triangleq (L-2)q_{1} + 2q_{2},$$

$$\vdots$$

$$u_{L} \triangleq Lq_{2}.$$
(10)

This means that, in terms of our observations, we will see L + 1 distinct levels uniformly distributed, although, once buried in noise, the levels may be not immediately evident from a laboratory measurement. However, even in noise a subset of these levels can usually be determined and, noting that the difference between levels is a constant equal to  $q_1 - q_2$ , one can infer the locations of the levels and even the number of levels.

This output process (9) adds the conductances of the L pores. Recall also that in our previous modeling, all the L pores were assumed to be identical. The marriage of these two facts makes it clear that on the basis of measurements of the (potentially noisy) observations of  $z_k$  it will be impossible to distinguish between channel configurations with the same number of open and closed pores. This leads to the notion of a Markov chain formed by aggregating states. That such an aggregation preserves the important Markov property is considered in the Appendix.

A separate motivation for considering the aggregation is that the dimensionality of the problem is reduced considerably. A naive application of (7) would have a transition probability matrix involving  $4^L$  elements, whereas if we define  $z_k$  (9) as the state at time k the transition probability matrix has only  $(L + 1)^2$  elements. For the range of L values that we wish to consider, the latter dimension is easily handled.

In light of the output equation (9), which defines the process  $z_k$ , we define aggregated states:

120

## AGGREGATED STATE

A state  $S_k$  belongs to aggregated state *i* if there are *i* pores that are open, that is, exactly *i* pores with conductance  $q_2$  and L-i pores with conductance  $q_1$ . As such, there are  $\binom{L}{i}$  (binary vector) states in aggregated state *i*.

In other words, states that are equivalent up to permutations of the binary components belong to the same aggregated state (equivalence class). The aggregated states are indexed by  $i \in \{0, 1, 2, ..., L\}$ . Aggregated state 0 corresponds to  $(q_1, q_1, ..., q_1)$ , and aggregated state L corresponds to  $(q_2, q_2, ..., q_2)$ . As a further explicit example, if L = 3 then we aggregate the  $2^L = 8 S_k$  states into the L + 1 = 4 aggregated states as shown in Table 3.

Finally, in terms of the output measurement process we have the simple observation

 $z_k = u_i \Leftrightarrow S_k \in aggregated state i \Leftrightarrow exactly i pores open.$  (11)

## 2.10. RELATIONSHIP BETWEEN ORIGINAL AND AGGREGATED SYSTEMS

It is possible to express the transition matrix of the aggregated system in terms of the transition matrix of the nonaggregated system (7). The  $(L+1)\times(L+1)$  aggregated probability transition matrix A corresponding to (7) takes the form

$$\mathbf{A} \triangleq \mathbf{LPR} = (1 - \kappa) \underbrace{\mathbf{LP}^{(I)}\mathbf{R}}_{\mathbf{A}^{(I)}} + \kappa \underbrace{\mathbf{LP}^{(C)}\mathbf{R}}_{\mathbf{A}^{(C)}}, \qquad (12)$$

where L is an  $(L+1) \times 2^{L}$  matrix with components

$l_{ii} \triangleq \left\{ \left( \frac{L}{i} \right)^{-1} \right\}$	if state $j \in aggregated$ state $i$ ,
0	otherwise

T	Ά	BL	Æ	3

Aggregated state index i	Set of States	$\begin{pmatrix} 3\\i \end{pmatrix}$
0	$\{(q_1, q_1, q_1)\}$	1
1	$\{(q_2, q_1, q_1), (q_1, q_2, q_1), (q_1, q_1, q_2)\}$	3
2	$\{(q_2, q_2, q_1), (q_2, q_1, q_2), (q_1, q_2, q_2)\}$	3
3	$\{(q_2, q_2, q_2)\}$	1

Indexing for the Aggregated States

and **R** is a  $2^L \times (L+1)$  matrix with components

$$r_{ji} \triangleq \begin{cases} 1 & \text{if state } j \in \text{aggregated state } i, \\ 0 & \text{otherwise.} \end{cases}$$

Here we note that under aggregation  $LP^{(C)}R$  maintains the form of (6) save that the dimensions are  $(L+1)\times(L+1)$  rather than  $2^L\times 2^L$ . However,  $LP^{(I)}R$  differs in form from (4); for example, if L = 2 then  $V \otimes V$  is given by (5), whereas

$$\mathbf{A}^{(I)} \triangleq \mathbf{L}(\mathbf{V} \otimes \mathbf{V}) \mathbf{R} = \begin{bmatrix} \zeta^2 & 2\zeta(1-\zeta) & (1-\zeta)^2 \\ \zeta(1-\rho) & \zeta\rho + (1-\zeta)(1-\rho) & (1-\zeta)\rho \\ (1-\rho)^2 & 2(1-\rho)\rho & \rho^2 \end{bmatrix}.$$
(13)

At this point we note that evaluation of the components of **A** according to (12) is somewhat impractical for large L because of the computation of the uncoupled component  $\mathbf{A}^{(I)}$ , which involves the Kronecker product expression.

#### 2.11. MEASUREMENT NOISE

Next we assume that the chain  $z_k$  is hidden, that is, indirectly observed by measurements  $y_k$  of the form

$$y_k = z_k + w_k, \tag{14}$$

where  $w_k$  is zero mean Gaussian noise of variance  $\sigma_w^2$ . Define the vector of conditional probability functions  $\mathbf{b}(\sigma_w^2, \mathbf{q}, y_k) = (b_i(\sigma_w^2, \mathbf{q}, y_k))$ , where

$$b_i(\sigma_w^2, \mathbf{q}, y_k) \triangleq P(y_k \mid z_k = u_i) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left[\frac{-(y_k - u_i)^2}{2\sigma_w^2}\right]$$
(15)

for  $0 \le i \le L$ . We denote the sequence of observations  $\{y_1, y_2, ..., y_k\}$  by  $Y_k$ . The problem takes the form of a hidden Markov model (HMM) because  $w_k$  is white, leading to the property

$$P(y_k | z_k = u_i, z_{k-1} = u_j, Y_{k-1}) = P(y_k | z_k = u_i)$$
(16)

(see also the Appendix). Also we assume that the initial state probability vector  $\pi = (\pi_m)$  is defined from  $\pi_i = P(z_1 = u_i)$ .

# 3. PARAMETER ESTIMATION

### 3.1. OVERVIEW

Now we address the problem of how to estimate the parameters of our model given real data measurements. Our emphasis lies in defining a procedure that obtains the estimates rather than focusing on the detailed development and analysis of the technique employed. We will therefore rely heavily on the literature and focus only on the novel aspects.

We summarize the estimation procedure as follows.

# Two-Step Estimation.

(1) Classical HMM identification of the transition probability matrix and associated relevant parameters (e.g., output levels  $q_1$ ,  $q_2$  and noise variance  $\sigma_w^2$ ).

(2) Novel optimal parameter fit according to the partially coupled model using gradient descent numerical techniques (recursions).

As motivation for the above, we note that the first step is independent of the "preferred" Markov model parameterization (7); that is, it imposes no bias toward the structure (12) we seek. Incorporating the first step also has the advantage of using existing theory (and software) with no modification. The second step is a well-defined procedure that can be employed on more general systems than those we consider; that is, it is not particularly restricted to the parameterization that we have developed. This feature enables the procedure to be easily modified in the event that, for example, the model were to be extended or changed.

## 3.2. HIDDEN MARKOV MODELING

We begin with a brief description of the standard HMM formulation that we use. The complete HMM parameter set is usually denoted

$$\lambda \triangleq \{\mathbf{A}, \mathbf{b}, \pi\}.$$

The matrix A is the transition probability matrix with entries

$$a_{ii} \triangleq P(z_{k+1} = u_i | z_k = u_i), \quad i, j \in \{0, 1, 2, \dots, L\}, \forall k$$

which gives complete information regarding the statistical properties of the transition between the states. The vector  $\mathbf{b}$  was defined in (15) and

depends on the parameters  $\{q_1, q_2, \sigma_w\}$ . This vector relates actual observations to the hidden state. The third element of  $\lambda$ ,  $\pi$ , defines the initial probability distribution across the states and is less crucial.

Some important things to note regarding standard HMM theory are

(1) Extensive work has been done developing computationally tractable recursions for evaluating maximum likelihood parameter estimates [22].

(2) These recursions crucially rely on estimating parameters corresponding to the "naive" parameterization; for example, the parameters  $a_{ii}$  are directly estimated.

(3) The estimation objective is to find the model estimate  $\hat{\lambda} = \{\hat{A}, \hat{b}, \hat{\pi}\}$  that maximizes the probability of the complete observation record  $Y_k$ , that is,

$$\hat{\boldsymbol{\lambda}} \triangleq \{ \hat{\mathbf{A}}, \hat{\mathbf{b}}, \hat{\boldsymbol{\pi}} \} = \arg \max_{\tilde{\boldsymbol{\lambda}}} P(Y_K \mid \tilde{\boldsymbol{\lambda}}), \tag{17}$$

where arg max means the argument (here  $\tilde{\lambda}$ ) that maximizes the indicated function.

In terms of the parameters that we seek, the parameters  $q_1$ ,  $q_2$  and  $\sigma_w^2$  can be directly estimated from the above standard HMM framework, whereas we need new methods to deal with the problem of estimating the parameters in  $\Theta$  that enter in a nonlinear fashion.

### 3.3. IDENTIFICATION OF THE NUMBER OF STATES

One disadvantage of the HMM technique—a common drawback for many identification schemes—is that the number of states in the underlying Markov process must be known *a priori*. The question of deducing the number of states in a Markov chain is an area of current research interest, and one proposed criterion for model order selection is the compensated likelihood approach [23]. Researchers in this area often focus on asymptotic estimators of the number of states. In practice the sample size is finite, and these estimators may exhibit bias in some cases. We will see later that in practice this problem can be overcome readily.

The maximum likelihood amplitude distribution of noisy channel currents, obtained under the assumption that there are a large number of evenly spaced current levels, frequently reveals prominent peaks corresponding to the closed state and several open states. Alternatively, we appeal to the principle of parsimony in comparing models with different numbers of free parameters. In other words, we measure the goodness of fit by evaluating the log likelihood of the model and weigh this against what is to be gained by increasing the number of parameters, which generally increases the likelihood. The error in fitting a model to a given set of data decreases with the number of free parameters in the model. Thus, it makes sense, in selecting a model from a set of models with different numbers of parameters, to penalize models that have too many parameters. Therefore, if a plot of log likelihood versus model order (i.e., the number of states) shows a "knee" for a certain model order, we would prefer this model to one of higher order. This approach has been used in determining the number of conductance substates in channel currents activated by  $\gamma$ -aminobutyric acid (GABA) [8].

#### 3.4. PARAMETER FITTING

In the above HMM parameter estimation problem, we assume that we have available optimal estimates of the transition probability matrix coefficients  $a_{ij}$ . Since our model developed in Section 2.2 impinges primarily on the transition probability matrix, estimating the parameters in the set  $\Theta$  (8) concerns only the  $a_{ij}$  parameters. Let the maximum likelihood matrix estimate generated from the data by the HMM processing have components given by  $\hat{a}_{ij}$  (17), and let the partially coupled Markov chain model have transition probability matrix denoted  $a_{ij}(\Theta)$ . We seek to estimate  $\Theta$ .

We pose the following parameter-fitting problem:

$$\Theta^* \triangleq \arg\min_{\Theta} \frac{1}{2} \|\mathbf{A}(\Theta) - \hat{\mathbf{A}}\|_F^2$$
  
=  $\arg\min_{\Theta} \underbrace{\frac{1}{2} \sum_{i=0}^{L} \sum_{j=0}^{L} (a_{ij}(\Theta) - \hat{a}_{ij})^2}_{\mathscr{I}(\Theta)}, \qquad (18)$ 

where  $\|\cdot\|_F$  is the Frobenius norm of a matrix, and argmin means the argument (in this case  $\Theta$ ) that minimizes the given index or cost  $\mathcal{I}(\Theta)$ .

We can visualize the cost  $\mathcal{I}(\Theta)$  to be minimized as a surface in the parameter space  $\Theta$ . The objective is to determine the parameter setting, here denoted  $\Theta^*$ , that realizes the minimum cost. This translates to trying to find the minimum of the surface and reading off the corresponding parameter setting (argument). Given that an initial parameter vector estimate is in error (an approximation or guess), a strategy that leads to the minimum, at least in a local sense, is to head downhill in a series of small steps, that is, attempt to locally decrease the value of the cost  $\mathcal{I}(\Theta)$ . Mathematically the direction to head is gleaned from

the gradient of the surface with respect to the parameter vector  $\Theta$ . The sequence of steps thus generated will be denoted  $\tilde{\Theta}(j)$ , where j is the step index. More explicitly this leads to the difference equation

$$\tilde{\Theta}(j+1) = \tilde{\Theta}(j) - \mu \frac{\partial \mathscr{I}}{\partial \Theta} \bigg|_{\Theta = \tilde{\Theta}(j)},$$
(19)

where

$$\frac{\partial \mathscr{I}}{\partial \Theta} \equiv \nabla_{\!\Theta} \mathscr{I} = \begin{bmatrix} \frac{\partial \mathscr{I}}{\partial \zeta} & \frac{\partial \mathscr{I}}{\partial \rho} & \frac{\partial \mathscr{I}}{\partial \kappa} \end{bmatrix}'$$

is the gradient with respect to  $\Theta$  (evaluated at  $\Theta$ ). Using (18), the partial derivatives in turn are written

$$\frac{\partial \mathscr{I}}{\partial \zeta} = \sum_{i=0}^{L} \sum_{j=0}^{L} \left[ a_{ij}(\Theta) - \hat{a}_{ij} \right] \frac{\partial a_{ij}(\Theta)}{\partial \zeta},$$

and similarly with respect to the remaining parameters in  $\Theta$ . Finally, in (19)  $\mu$  represents a small positive stepsize parameter, typically between 0.001 and 0.01.

At this point we highlight the need for efficient means to compute both the components  $a_{ij}(\Theta)$  and the partial derivatives, for example,  $\partial a_{ij}(\Theta)/\partial \zeta$ , above. When running the update equation (19) we initialize with

$$\tilde{\Theta}(0) \triangleq \begin{bmatrix} 0.5 & 0.5 & 0.5 \end{bmatrix}'$$

because the parameters to be estimated are known to be probabilities; that is, we take values between 0 and 1. The step size is typically taken as  $\mu = 0.01$  or smaller to ensure numerical stability. Also, in the algorithm, as a safety measure, we project the parameters of  $\hat{\Theta}$  to the range [0, 1] to ensure that they remain probabilities each time the estimates are updated.

# 4. APPLICATION TO ANALYSIS OF CHANNEL CURRENTS

#### 4.1. OVERVIEW

In this section we illustrate the application of the coupled Markov model to the analysis of channel currents recorded from biological membranes. These records were chosen solely to illustrate the method; the biological significance of the observations will not be dealt with in any detail here. Currents analyzed here were recorded from inside-out

#### COUPLED MARKOV CHAIN MODEL

patches of membranes from rat ventricular myocytes and neonatal hippocampal cells grown in culture for 7–10 days, using a current-to-voltage converter (Axopatch 200A amplifier, Axon Instrument).

# 4.2. DISCRIMINATION BETWEEN CONDUCTANCE SUBLEVELS AND MULTIPLE CHANNELS

When there are more than one open level in channel currents, either the lower levels are considered to be conductance substates of the fully open channel or the current record is believed to represent the superposition of two or more independent channels. The criterion used when inferring that one level is a conductance substate of another is that one can find examples of synchronous openings and asynchronous closings or vice versa. It is difficult to establish that channels show some form of linkage by visual inspection of current records, especially if the signalto-noise ratio is low. With the mathematical formulation given in the previous section, we can unambiguously determine whether the constituent channels are totally independent or partially coupled.

Current records generated by step changes in the membrane potential across an inside-out membrane patch revealed the presence of multiple open levels. In the example illustrated in Figure 1a, the three open levels, deduced accurately from the maximum likelihood amplitude distribution (see Figure 1b), were spaced nearly evenly, at 0.67, 1.24, and 1.91 pA. Are the two lower openings conductance sublevels of a channel whose open level is 1.91 pA, or does the patch contain three active channels? The transition probability matrix A constructed from an entire record failed to provide an unambiguous answer, which read

$$\mathbf{A} = \begin{pmatrix} 0.975 & 0.025 & 0.000 & 0.000 \\ 0.020 & 0.961 & 0.018 & 0.001 \\ 0.014 & 0.075 & 0.878 & 0.003 \\ 0.000 & 0.006 & 0.169 & 0.825 \end{pmatrix},$$

Although the predominant transitions were between the neighboring levels, there were finite probabilities that the current would jump between two discrete levels. For example, the probability that the current jumps to the 0.66 pA level at time t + 1 given that it was at the 1.91 pA level at time t was 0.006. Similarly, there was a finite probability (0.014) of closing to the baseline at time t + 1 given that the current was at the 1.24 pA level at time t. The maximum a posteriori estimate of the signal sequence, illustrated in Figure 1c, does not provide an unambiguous answer. By selecting a few examples of instantaneous transitions from the fully open level to the baseline, or from the baseline to the second open level, one could have argued that the





channel has one or two conductance sublevels at which it dwells preferentially. The coupling coefficient  $\kappa$  determined by our method from the matrix given above, however, is precisely zero, indicating that the three active channels contained in the patch opened and closed independently of each other. We thus conclude that the smallest open level of 0.66 pA represents the amplitude of a single voltage-gated sodium channel.

### 4.3. DECOMPOSITION OF MULTIPLE-CHANNEL CURRENTS

A membrane patch contains not just one but many active channels of the same type, especially if the recordings are made at physiological temperature. The signal then will be composed of an algebraic sum of Lidentical channels, and we will see L + 1 distinct levels (including the baseline level), although, once buried in noise, the levels may not be immediately evident from laboratory measurements. From such a realworld process, our task is then to (1) identify the number of active channels contributing to the summed current, (2) deduce whether or not the active channels have some form of linkage, and (3) determine the kinetics of single channels, such as the mean open duration and the open time distribution.

Figure 2a shows a segment of a current trace from a patch that contained many active voltage-gated sodium channels. From the histogram of the raw data, which merely shows a skewed distribution (not shown here), no information about the amplitude of single channels or the number of active channels contained in the patch could be gleaned. However, clearly defined levels can be discerned from the maximum likelihood amplitude distribution obtained with the HMM processing technique (Figure 2b). The 10 identifiable peaks in the amplitude distribution are separated by  $0.62 \pm 0.02$  pA. These current levels were then used to produce the maximum likelihood estimates of signal sequence (Figures 2c and 2d) and transition probability matrix A. The coupling coefficient deduced from the matrix was close to zero ( $\kappa = 0.012$ ), and  $\rho$  in the binary matrix given in Equation (2) was 0.946, thus indicating that the mean open duration of the constituent single channels was 1.85 ms.

### 4.4. PARTIALLY COUPLED PORES

The chloride currents flowing through a channel activated by GABA fluctuate over a wide range. Instead of switching between the fully open state and fully closed state instantaneously, currents dwell at intermediate levels at variable times. Three segments of such a record obtained with a pipette potential of 100 mV are shown in Figure 3a. The



FIG. 2. Decomposition of multiple-channel currents. (a) A 1000-point segment of channel currents generated by voltage-gated sodium channels. The patch was maintained at room temperature. In response to a step change of the membrane potential (from -150 to -50 mV), many single channels were activated and then inactivated. The trace contains responses to seven such step changes. (b) The maximum likelihood amplitude distribution reveals 10 or possibly 11 discrete peaks that are spaced evenly. The identifiable peaks are separated by  $0.62\pm0.02$  pA. These current levels were then used to produce the maximum a posteriori signal sequence shown in (c). (d) A response to one step change of the membrane potential shown on a larger time scale.



FIG. 3. Partially coupled pores. (a) Three 1000-point segments of GABA-activated chloride channel currents. The patch was held at a hyperpolarized potential ( $V_p = +100 \text{ mV}$ ). Downward deflection represents opening of the channel. (b) The maximum likelihood amplitude probability distributions were obtained using the HMM processing method, iterating the same observation sequence repeatedly. The sequential histograms estimated at successive iterations are presented as a three-dimensional graph. The presence of conductance sublevels becomes more and more distinct as the records are further iterated. Channel currents were recorded from an inside-out membrane patch of a cultured rat hippocampal neuron. The patch electrode contained 0.25  $\mu$ M GABA. The record was filtered at 2 kHz and sampled at 5 kHz. (c) A 1000-point segment of simulated channel currents. Five fictitious pores with  $\zeta = 0.933$  and  $\rho = 0.894$  were coupled with the coupling coefficient  $\kappa$  of 0.17. The amplitude of each pore and the standard deviation of noise were assumed to be -0.5 pA and 0.25 pA, respectively.

maximum likelihood amplitude distributions obtained with the HMM processing method, exhibited in Figure 3b, show several peaks, each separated by about 0.5 pA. These peaks, which become progressively more prominent with increasing number of iterations, can be construed as conductance sublevels at which currents dwell preferentially. Alternatively, we can envisage that the collective behavior of several partially coupled pores generates a complex pattern of gating kinetics that characterize many naturally occurring channels.

We estimated the transition probability matrix of constituent chains and their coupling coefficient. The channel currents exhibited in Figure 3a can be modeled as the superposition of five pores that are partially coupled with the coupling factor  $\kappa$  of 0.17. The estimated  $\zeta$  and  $\rho$  of the elementary pore are, respectively, 0.933 and 0.894. Using these parameters and taking the standard deviation of noise in the absence of channel activity to be 0.25 pA, we have simulated an observation sequence of fictitious channel currents generated by five partially coupled pores. A segment of such a record, exhibited in Figure 3c, shows that the gating behavior of the synthesized data sequence is broadly similar to that of the real channel currents. When the transmembrane potential was reduced stepwise from a hyperpolarized potential to zero, the value of  $\kappa$  decreased systematically as did the number of pores contributing to the summed currents. When the direction of the electric field was reversed, only two pores were active, each opening and closing independently of the other (i.e.,  $\kappa = 0$ ). The current-voltage relation of the summed currents, owing to the reduced number of pores and coupling strength at hyperpolarized potentials, showed a pronounced outward rectification, but for each elementary pore the current-voltage relationship obeyed Ohm's law.

#### 5. DISCUSSION

We developed a Markov chain model that incorporates partial coupling among a set of binary Markov chains and then formulated a scheme for estimating the parameters featured in a vector Markov chain hidden in noise. The model identification method we devised uses standard HMM techniques [17, 18, 22] and a gradient descent numerical iteration scheme in a novel way. When the parameters featured in our model are varied, the resultant vector Markov chains exhibit a rich range of characteristics that resemble channel currents recorded from patches of neuronal membranes. We have extensively tested the reliability of our estimation scheme in characterizing partially coupled or uncoupled Markov signals buried in noise. Our estimation scheme has been devised for characterizing summed channel currents obtained from patch-clamp recordings and thus allows up to 12 independent or partially independent pores contributing to the summed currents. For studying the records that are the sum of a larger number of channels, such as those recorded from a whole-cell configuration or with an intracellular electrode, a new analytical method such as the one described elsewhere [16] should be used.

Our analytical method rests on the premise that the signal sequence of each pore can be represented by a binary, first-order Markov chain. We emphasize that the processing scheme is relatively insensitive to deviation from what appears to be a severely restrictive assumption. One of the properties frequently observed in real channel currents is that channels remain closed for prolonged periods. Such gaps give two or sometimes three distinct exponential distributions in the closed-time interval histograms, and the observed binary channel can best be approximated as a switched Markov chain. The coupling coefficient we estimate, showing whether the constituent pores open and close independently or have some form of linkage, is unaffected by the presence or absence of long silent gaps. Also, it is relatively straightforward to introduce a further parameter, say  $\omega$ , that can take account of long closed intervals, which may be due to the diffusion of transmitter molecules near the receptor, binding of these molecules to the binding sites, and the phenomenon of desensitization. By introducing the parameter  $\omega$ , again a probability, we have been able to generate inter-burst interval distributions that are kinetically distinct from intra-burst interval distributions. We have made such an extension to the theory, but it is not reported here.

Using three examples of real channel currents, we have demonstrated the usefulness of our model in analyzing certain types of data. Instead of switching between the fully open and closed states instantaneously, observed channel currents appear to dwell at intermediate levels for variable times. These sublevels in some cases are clearly spaced in equal steps. We represented such an observation sequence with a coupled Markov model and noted how some of the model parameters, especially the coupling coefficient  $\kappa$ , changed systematically with the strength and direction of the applied electric field. Whether or not the mathematical representation we adopted for the analysis of GABA-activated chloride currents reflects the physical reality remains to be investigated. We have also illustrated how multiple channels contained in a recording can be analyzed using the mathematical framework presented here. The method enables us to deduce unambiguously whether or not the channels contributing to the summed currents are independent and provides a numerical measure of the coupling strength when they are not totally independent. Using our

model, one can readily characterize individual channel kinetics from a record containing multiple channels, coupled or otherwise. Such an analysis would have been difficult to carry out without the aid of a versatile mathematical model such as the one we present here.

The mathematical formulations for determining various statistics from the multichannel recordings have been proposed by others [3, 4]. Unlike the previous formulations, the model presented here links Markov chains with a coupling term. Such a coupling occurs in many physical systems, a prominent example of which is a series of damped harmonic oscillators that are interlinked with a coupling term. Although the dimensionless parameters featured in the model are introduced in a formal mathematical setting that describes the collective behavior of coupled Markov chains, they assume physical significance.

In our formulation, we have assumed for computational simplicity that the individual pores contributing to the total channel currents have identical amplitudes. In principle, our identification techniques can be readily modified for the cases where nonidentical pores are superimposed. However, since the pores are no longer describable by a common dynamical process, we would need to consider the nonaggregated Markov model, which consists of  $2^L$  distinct values, where L is the number of such pores. The details of the identification procedure represent a straightforward generalization of the gradient descent approach. However, when more than four or five such nonidentical pores are present, the identification of all the parameters will become computationally expensive. Fredkin and Rice [2] and Chung et al. [18] have considered the superposition of two nonidentical uncoupled channels.

This work was supported in part by grants from the Australian Research Council and National Health and Medical Research Council of Australia. We thank Dr. D. A. Saint for making the data illustrated in Figure 2 available to us.

# APPENDIX

# MARKOV PROPERTY OF THE AGGREGATED STATES

Given a finite state Markov process it is always possible to generate a lower dimensional process by grouping into aggregated states. This has two important effects: (1) information is thrown away, and (2) generally the Markov property is destroyed. However, in certain instances the information that was desired to be inferred about the original process is retained by the aggregated process because coarse versus fine information was desired. Also for certain processes the Markov property holds for the aggregated process. In other words, knowledge of the current state is all that is required for determining the transition probability. This is important in our theory as HMM identification techniques are employed.

In considering the relationship claimed by (16) we need to examine the definition of the aggregated states. Under the measurement (14) it is clear that the noise  $w_k$  is of secondary importance because it is independent of the state dynamics and is assumed to be an independent process; that is, given measurements up to time k-1, no information can be inferred about  $w_k$ . It therefore has no bearing on the Markov property with respect to  $y_k$  (i.e., does not introduce any additional difficulty). This means that  $\{y_k\}$  is a Markov process if and only if the noiseless random process  $\{z_k\}$  is one.

We focus on the partially coupled model of (7). Consider an arbitrary state, say  $Q_m$ , in aggregated state *i* at time *k* and the transition probability to aggregated state *j* at time k + 1 (i.e., the probability that we transit to some state in aggregated state *j*). This transition probability is given by

$$a_{ij} = \sum_{\mathbf{1}^T Q_n = j} P(S_{k+1} = Q_n | S_k = Q_m), \quad \mathbf{1}^T Q_m = i.$$
(20)

We now wish to show that all states in the aggregated state i share this same transition probability. This implies that we can form aggregated states with a sensible transition probability label.

By assumption, the pores are equivalent. Let  $\Pi$  be an arbitrary  $L \times L$  permutation matrix consisting of L(L-1) zeros and L ones where there is exactly one one appearing in every column and row. Such a matrix satisfies  $\Pi^T \Pi = \mathbf{I}$  (the identity matrix). Then the equivalence above can be expressed as

$$P(S_{k+1} = Q_n | S_k = Q_m) = P(S_{k+1} = \Pi Q_n | S_k = \Pi Q_m), \quad (21)$$

which can be thought of as a relabeling of the pores. Substituting (21) into (20) yields

$$a_{ij} = \sum_{\mathbf{1}^{T}Q_{n}=j} P(S_{k+1} = \Pi Q_{n} | S_{k} = \Pi Q_{m})$$
  
= 
$$\sum_{\mathbf{1}^{T}\Pi^{T}Q_{r}=j} P(S_{k+1} = Q_{r} | S_{k} = \Pi Q_{m})$$
  
= 
$$\sum_{\mathbf{1}^{T}Q_{r}=j} P(S_{k+1} = Q_{r} | S_{k} = \Pi Q_{m}).$$

Note that  $\mathbf{1}^T \Pi Q_m = i$  and that all elements of the aggregated state *i* can be expressed this way. We have used the property that  $\Pi^T$  is also a permutation matrix and therefore satisfies  $\mathbf{1}^T \Pi^T = \mathbf{I}$ .

The above argument is independent of i and j and reveals that the aggregated finite-state process possesses all the attributes of a finite-state Markov process.

#### REFERENCES

- 1 D. Colquhoun and A. G. Hawkes, On the stochastic properties of single ion channels, *Proc. Roy. Soc. Lond. B* 211:205-235 (1981).
- 2 D. R. Fredkin and J. A. Rice, On the superposition of currents from ion channels, *Phil. Trans. Roy. Soc. Lond. B* 334:347–356 (1991).
- 3 M. Jackson, Stochastic behavior of a many-channel membrane system, *Biophys. J.* 47:129-137 (1985).
- 4 C. F. Yeo, R. K. Milne, R. O. Edeson, and B. W. Madsen, Superposition properties of independent ion channels, *Proc. Roy. Soc. Lond. B* 238:155-170 (1989).
- 5 A. Auerbach and F. Sachs, Single-channel currents from acetylcholine receptors in embryonic chick muscle, *Biophys. J.* 45:187–198 (1984).
- 6 O. P. Hamill and B. Sakmann, Multiple conductance states of single acetylcholine receptor channels in embryonic muscle cells, *Nature (Lond.)* 294:462–464 (1981).
- 7 S. G. Cull-Candy and M. M. Usowicz, On the multiple-conductance single channels activated by excitatory amino acids in large cerebellar neurons of the rat, *J. Physiol. (Lond.)* 415:555–582 (1989).
- 8 P. W. Gage and S. H. Chung, Influence of membrane potential on conductance sublevels of chloride channels activated by GABA, *Proc. Roy. Soc. Lond. B* 255:167-172 (1994).
- 9 C. E. Jahr and C. F. Stevens, Glutamate activates multiple single channel conductances in hippocampal neurons, *Nature (Lond.)* 325:522-525 (1987).
- 10 L. S. Premkumar, S. H. Chung, and P. W. Gage, GABA-induced potassium channels in cultured neurons, *Proc. Roy. Soc. Lond. B* 241:153–158 (1990).
- 11 L. S. Premkumar, P. W. Gage, and S. H. Chung, Coupled potassium channels induced by arachidonic acid in cultured neurons, *Proc. Roy. Soc. Lond. B* 242:17–22 (1990).
- 12 A. L. Zimmerman and D. A. Baylor, Cyclic GMP-sensitive conductance of retinal rod consists of aqueous pores, *Nature (Lond.)* 321:70-72 (1986).
- 13 A. Queyroy and J. Verdetti, Cooperative gating of chloride channel subunits in endothelial cells, *Biochim. Biophys. Acta* 1108:159-168 (1992).
- 14 E. Yeramian, A. Trautmann, and P. Claverie, Acetylcholine receptors are not functionally independent, *Biophys. J.* 50:253-263 (1986).
- 15 V. Krishnamurthy, J. B. Moore, and S. H. Chung, On hidden fractal signal processing, *Signal Process.* 24:177–192 (1991).
- 16 S. H. Chung and G. Pulford, Fluctuation analysis of patch-clamp or whole-cell recordings containing many single channels, J. Neurosci. Methods 50:369-384 (1993).

- 17 S. H. Chung, J. B. Moore, L. Xia, L. S. Premkumar, and P. W. Gage, Characterization of single channel currents using digital signal processing techniques based on hidden Markov models, *Phil. Trans. Roy. Soc. Lond. B* 329:265–285 (1990).
- 18 S. H. Chung, V. Krishnamurthy, and J. B. Moore, Adaptive processing techniques based on hidden Markov models for characterizing very small channel currents buried in noise and deterministic interferences, *Phil. Trans. Roy. Soc. Lond. B* 334:357–384 (1991).
- 19 G. W. Pulford, J. C. Gallant, R. A. Kennedy, and S. H. Chung, Evaluation and estimation of various Markov models with applications to membrane channel kinetics, *Biomed. J.* 37:39–63 (1995).
- 20 P. Billingsley, *Statistical Inference for Markov Processes*, Univ. Chicago Press, Chicago, 1961.
- 21 J. G. Kemeny and J. L. Snell, *Finite-State Markov Chains*, Van Nostrand, Princeton, NJ, 1960.
- 22 L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE* 77:257–285 (1989).
- 23 L. Finesso, Consistent estimation of the order for Markov and hidden Markov chains, Ph.D. Dissertation, Univ. Maryland, College Park, MD, 1990.